APPENDIX I: INTERVIEWS

PART I: AGENCY USE CASES

A. HHS: Reg Explorer and the Regulatory Cleanup Initiative

Agency Officials

- HHS Former Official A: [Zoom Interview, Feb. 11, 2022]
- HHS AI Official B:[Zoom Interview, Feb. 28, 2022]
- HHS AI Official C: [Zoom Interview, March 22, 2022]
- HHS ONC (Office of the National Coordinator) Official D: [Zoom Interview, April 06, 2022]
- HHS ASPE (Assistant Secretary for Planning and Evaluation) Official E: [Zoom interview, October 24, 2022]
- FDA Official A: [Zoom interview, April 13, 2022]

Deloitte Representatives

- Deloitte Product Manager: [Zoom Interview, June 29, 2021]
- Deloitte Managing Director: [Zoom Interview Feb. 9, 2022]

B. DOT: RegData Dashboard

Agency Officials

- DOT Official A**:** [Zoom interview, March 28, 2022]
- DOT Officials B, C, ad D: [Zoom interview, March 11, 2022]

C. DoD: GAMECHANGER

Agency Officials

- DoD Former Official A: [Zoom Interview 04/25/22]
- DoD Former Official B:[Zoom Interview 04/13/22]
- DoD Former Official C: [Zoom Interview 05/03/22]
- DoD Officials D & E: [Zoom Interview March 10, 2022]
- DoD Official F: [Email Interview 12/05/22]

BAH Representatives

- BAH Data Scientist: [Zoom Interview April 19, 2022]
- BAH Lead Associate: [Zoom Interview April 7, 2022]

D. GSA/CMS: Regulatory Analytics Proof of Concept

Agency Officials

- GSA Official A: [Zoom Interview Jan. 12, 2022]
- CMS Official A: [Zoom Interview, April 1, 2022]

- CMS Official B: [Zoom Interview, June 27, 2022]
- CMS Official C: [Zoom Interview 04/01/2022]

BeInformed Representatives

- BeInformed Team Members A & B: [Zoom Interview, March 21, 2022]

PART II: RETROSPECTIVE REVIEW

A. Independent Agencies

- FTC Official A:[Zoom Interview, March 8, 2022]
- STB Officials A & B: [Zoom Interview, March 29, 2022]
- NCUA Official A: [Zoom Interview, April 12, 2022]
- OCC Officials A, B & C: [Zoom Interview, April 18, 2022]

B. Executive Branch Departments & Agencies

- DoEd Official A: [Zoom Interview, March 25, 2022]
- BSEE Official A: [Zoom interview, Mar. 25, 2022]
- Coast Guard Official A: [Zoom Interview, April 11, 2022]
- DOC Official A: [Zoom Interview, April 20, 2022]

C. Stakeholders

- NFIB Representative A: [Zoom Interview, 03/28/22]
- Public Citizen Representative A: [Zoom Interview 04/11/22]
- Unidos US Representative A: [Zoom Interview 04/11/2022]
- CDT Representatives A & B: [Zoom Interview, April 13, 2022]
- OTI Representatives A & B: [Zoom Interview 04/21/22]
- About ML Representative A: [Zoom Interview, April 29, 2022]
- NAACP LDF Representative A : [Zoom Interview April 29, 2022]

ON BACKGROUND: ADDITIONAL INTERVIEWS

Agency Officials

- USCIS (US Citizen and Immigration Services) Official A: [Zoom Interview Feb. 29, 2022]
- VA (Veterans Administration) Official A: [Zoom Interview, March 4, 2022]

Academics

- George Washington Law Professor (January 12, 2022, via Zoom)
- Duke Law Professor (January 28, 2022, via Zoom)

Nonprofits

- National Academy of Public Administration (NAPA) Representatives A and B (January 20, 2022, via Zoom)
- Mercatus Center Representatives A and B (Feb. 7, 2022, via Zoom)
- Ford Foundation Representative A (April 14, 2022, via Zoom)

Firms

- IBM Representative A (January 7, 2022, via Zoom)
- IBM Representative B (January 21, 2022, via Zoom)
- IBM demonstration of Federated Learning (February 18, 2022, via Zoom)
- Regulatory Group Representative A (January 20, 2022, via Zoom).

Interview Questionnaire Templates

Federal Agencies

For All Agencies

1) How does your agency currently attempt to identify rules that are:
    a. Outdated
    b. Redundant
    c. Contain inaccurate cross references
    d. Contain typographical errors
    e. Are in need of elaboration or clarification?
2) What other types of retrospective review (beyond identifying such rules) does your agency conduct?
3) What principal "pain points" and sources of inefficiencies does your agency face when performing retrospective review?
4) Does your agency involve the public, if at all, in identifying rules to be reviewed?
5) Does your agency engage in retrospective review, for all or a portion of your rules, on a set interval (e.g., every five years)?
6) Do you conduct the same sort of retrospective review process for things that aren't "legislative rules"? (For example, statements of policy or guidance documents.)
7) Has your agency used AI tools in retrospective review?
    ➔ *Branch out to line of questions A or B*

A. For Agencies Currently Using AI for Retrospective Review

AI for retrospective review:

8) What form of AI does your agency use to assist with retrospective review of rules (e.g., natural language processing-based software)?
9) Did your agency develop these tools in house or did it procure them from an outside vendor?
    a. If your agency developed these tools in house, which kind of employees (e.g., data scientists, engineers, etc.) were involved in developing them? Did your agency face any challenges with internal capacity building?
    b. If your agency procured them from an outside vendor, what decided your agency to not develop them in house? Who were the decisionmakers for the procurement process? What considerations went into selecting the vendor? How does your agency oversee the vendor? What types of training does your staff get on the vendor's tools?
10) Has your agency trained staff involved in the rulemaking process in how to use AI-based tools?
    a. If yes, how effective has this training been? What challenges have your agency encountered in this training?
    b. If not, are you considering this training?
11) Did your agency consider potential issues with litigation and/or violating statutory or APA requirements? If yes, what type of claims was your agency concerned about?
12) Developing, procuring, deploying, and overseeing AI-based tools to assist with retrospective review cost money and staff time. Among the many competing priorities

your agency faces, what caused it to decide to allocate its limited resources to these endeavors?

13) How does your agency avoid overreliance on these tools to conduct retrospective review? Put another way, how does your agency ensure that these tools aren't making final decisions?
14) Does your agency inform the public of its use of AI for retrospective review?
15) Does your agency seek public input on its use of AI for retrospective review?

AI for rulemaking more broadly:

16) Has your agency used AI in other aspects of rulemaking (e.g., comment analysis)?
17) For what additional areas of rulemaking would your agency consider using AI? If you could automate any step in the rulemaking process, what would it be?
18) Would your agency be concerned about potential issues with litigation and/or violating statutory or APA requirements if it were to implement such use cases of AI rulemaking?

B. For Agencies Not Currently Using AI for Retrospective Review

8) If not AI, has your agency used any type of computer tools / statistical techniques (including CBA type) to conduct retrospective review?
9) Is your agency open to the idea of using AI-based tools to assist with retrospective review?
10) Would your agency likely develop these tools in house or procure them?
    a. What considerations does your agency take into account for deciding whether to develop in house or procure a technology or tool?
    b. If your agency would likely develop these tools in house, which kind of employees (e.g., data scientists, engineers, etc.) would be involved in developing them?
    c. If your agency would likely procure them from an outside vendor, what considerations would go into selecting the vendor? Who would be the decisionmakers for the procurement process? How would your agency oversee the vendor?
11) Would your agency be concerned about potential issues with litigation and/or violating statutory or APA requirements? If yes, what type of claims would your agency be concerned about?
12) If you could accurately automate one step in the retrospective review process, what would it be?
13) Developing, procuring, deploying, and overseeing AI-based tools to assist with retrospective review cost money and staff time. Among the many competing priorities your agency faces, what would cause it to decide to allocate its limited resources to these endeavors?
14) How, if at all, would your agency avoid overreliance on these tools? Put another way, how would your agency ensure that these tools aren't making final decisions?
15) How, if at all, would your agency inform the public of its use of AI for retrospective review?
16) How, if at all, would your agency seek public input on its use of AI for retrospective review?

AI for rulemaking more broadly:

17) Has your agency used AI in other aspects of rulemaking (e.g., comment analysis)?
18) For what additional areas of rulemaking would your agency consider using AI? If you could automate any step in the rulemaking process, what would it be?
19) Would your agency be concerned about potential issues with litigation and/or violating statutory or APA requirements if it were to implement such use cases of AI rulemaking?

<u>Stakeholders</u>

1) How are you engaged with rulemaking and/or retrospective review?
2) Is it appropriate for agencies to use ML/AI in the rulemaking/retrospective review process?
3) What parts of the retrospective review process do you think are most amenable to being supported with an AI-based tool?
4) More broadly, what parts of the entire rulemaking process do you think are most amenable to being supported with an AI-based tool?
5) Conversely, which parts of the rulemaking process (and the retrospective review process more specifically) do you think are out of reach for current or currently developing technology?
6) What priorities should agencies keep in mind as they consider whether and how to use AI in the rulemaking/retrospective process?
7) How sanguine are you about agencies' commitments to "Trustworthy AI," namely "the design, development, acquisition, and use of AI in a manner that fosters public trust and confidence while protecting privacy, civil rights, civil liberties, and American values, consistent with applicable laws?"
8) How about agencies' commitments to "Explainable AI," including attention to:
   (a) legal and regulatory risk (defined as "unfair practices, compliance violations, or legal action due to biased data or a lack of explainability"); and
   (b) enhancing public trust
9) When adopting an AI tool, how should an agency ensure that it remains faithful to important principles of administrative law such as:
   (a) transparency
   (b) reason-giving
   (c) public participation, and
   (d) accountability
10) Agencies typically assert that any AI tool is designed not to replace but rather to "augment" human judgment, interpretation, and decision making. What safeguards (if any) should the agency put in place to avoid potential overreliance on an AI tool?
11) What factors should agencies consider when deciding whether to develop a tool in house or procure one from a vendor? Should the use of "open-source" standards and/or the public disclosure of training datasets be a pre-requisite? How much disclosure to the public is warranted in either situation?
12) Are you aware of any legal concerns or risks that may be associated with using an AI-based tool in the rulemaking/retrospective review process? (e.g., violating APA/statutory requirements)

APPENDIX II: Technical Details of AI-Enabled Tools

A. HHS/Deloitte RegExplorer[1]

- *Keyword Technology*: HHS has described "keyword technology" as "a structured and iterative approach to process, analyze, and return keyword search results."[2] This includes keyword extraction (or keyword detection or keyword analysis), which is "the automatic identification of a set of the terms that best describe the subject of a document."[3] Methods of keyword extraction vary from simpler, statistical approaches, to more complex linguistic, machine learning, or graph-based approaches. For example, a common statistical method is TF-IDF (term frequency-inverse document frequency), which identifies the "importance" of a word by calculating the normalized number of times it appears in a document (the term frequency) and multiplying it by a logarithmically scaled inverse fraction of the documents containing that word (the inverse document frequency).[4] More complex ML methods can be either supervised (trained on a set of keywords) or unsupervised, and graph-based text representation methods vary widely.[5]

- *Clustering Algorithms*: HHS has defined a "cluster" as "a machine-generated group of regulatory documents that have been algorithmically gathered together based on a set of similar characteristics, such as the relevant sub-agency, placement of text within the regulatory dataset, similarity of text content, and text format and structure."[6] More specifically, RegExplorer uses neural networks (a subset of ML) to create these clusters, which have been validated by statistical tests and regulatory specialists.[7] Among other things, these neural networks allow computers to "understand how concepts in a given piece of text relate to each other—for example, that boat and ship are similar"[8]—enabling a deeper comparison of and a more faithful clustering of regulations. Neural networks

---

[1] In response to a comment inquiring about the underlying algorithms used by Deloitte in a 2019 analysis, HHS answered:

> While RegExplorer is proprietary technology, some of the models deployed within RegExplorer include keyword technology (a structured and iterative approach to process, analyze, and return keyword search results); a clustering algorithm (a cluster is a machine-generated group of regulatory documents that have been algorithmically gathered together based on a set of similar characteristics, such as the relevant sub-agency, placement of text within the regulatory dataset, similarity of text content, and text format and structure); citation extraction and mapping; and similar section analysis.

Securing Updated and Necessary Statutory Evaluations Timely, 86 Fed. Reg. 5694, 5710 (Jan. 19, 2021). Additional categories of algorithms were mentioned in an interview with a Deloitte product manager. See Interview with Deloitte Product Manager.

[2] Securing Updated and Necessary Statutory Evaluations Timely, 86 Fed. Reg. at 5710.

[3] Slobodan Belinga et al., *An Overview of Graph-Based Keyword Extraction Methods and Approaches*, 39 J. INFO. & ORG. SCIS. 1, 1 (2015).

[4] *See* Anand Rajaraman & Jeffrey D. Ullman, MINING OF MASSIVE DATASETS 8 (Cambridge Univ. Press 2012).

[5] *See* Slobodan Belinga et al., *supra* note 3, at 2–4. In a graph-based model, a "document is modelled as a graph where terms (words) are represented by vertices (nodes) and their relations are represented by edges (links)." *Id.*

[6] Securing Updated and Necessary Statutory Evaluations Timely, 86 Fed. Reg. at 5710.

[7] *See* Daniel Byler, Beth Flores & Jason Lewis, *Using Advanced Analytics to Drive Regulatory Reform* 8, DELOITTE, https://www2.deloitte.com/content/dam/Deloitte/us/Documents/public-sector/us-ps-using-advanced-analytics-to-drive-regulatory-reform.pdf.

[8] *Id.*

themselves do not produce clusters, but they can provide meaningful information from which clusters can be created.

- *Citation Extraction and Mapping*: Deloitte does not provide information on what techniques or methodology it uses for citation extraction and mapping. But, generally, the purpose of such algorithms is to faithfully identify, extract, and map any citation in a given text.[9] RegExplorer focuses on extracting citations—mainly citations to other regulations—in any given regulation.[10]

- *Guided LDA*: Latent Dirichlet Allocation (LDA) is a method of topic modeling—discovering topics in a collection of documents and then automatically classifying an individual document within a discovered "topic."[11] Generally, topic modeling is an unsupervised class of ML algorithms, but the adjective "guided" implies at least a semi-supervised algorithm. In fact, "GuidedLDA" is a publicly available Python package that implements traditional LDA in a semi-supervised process.[12]

B. <u>DoD/Booz Allen Hamilton GAMECHANGER</u>

- *Rule Data Crawler*: Data crawling is a method of data extraction, typically performed over large quantities of data to speed up the data collection process, which automatically collects data from the Internet or from any document or file.[13] GAMECHANGER provides a Data Crawler API (Application Programming Interface) that enables organizations to automatically upload regulatory documents from their web sources (e.g., website listing all their policies) into GAMECHANGER.[14]

- *Knowledge Graph Generation, Search, and Pattern Recognition*: GAMECHANGER uses knowledge graph technology to structure data and organize it into a policy knowledge graph. Such graph technology creates and updates the nodes and edges within the knowledge graph, mainly relying on automation and complemented by review and

---

[9] For information on citation extraction algorithms, see generally BRETT POWLEY & ROBERT DALE, HIGH ACCURACY CITATION EXTRACTION AND NAMED ENTITY RECOGNITION FOR A HETEROGENEOUS CORPUS OF ACADEMIC PAPERS (2017), https://web.science.mq.edu.au/~rdale/publications/papers/2007/49.pdf

[10] For an example of a RegExplorer citation map, see REGULATING FOR NSW'S FUTURE, NSW TREASURY 8 fig. 3 (July 2020), https://www.treasury.nsw.gov.au/sites/default/files/2020-07/FINAL%20Treasury%20report%20210720.pdf.

[11] *See* David M. Blei et al., *Latent Dirichlet Allocation*, 3 J. MACH. LEARNING RSCH. 993 (2003).

[12] *Welcome to GuidedLDA's Documentation!*, READ THE DOCS, https://guidedlda.readthedocs.io/en/latest/ (last visited March 5, 2022).

[13] *See Web Scraping vs Web Crawling: The Differences*, Oxylabs (May 4, 2021), https://oxylabs.io/blog/crawling-vs-scraping.

[14] dod-advana/gamechanger-crawlers, GITHUB, https://github.com/dod-advana/gamechanger-crawlers. This crawler accesses the relevant web sources, generates PDF documents from the webpages or downloads PDF files already hosted on the webpages, and uploads to the GAMECHANGER database both these PDFs and a JSON representation of these policy documents. *See* Interview with BAH Data Scientist. Alternatively, the organization can manually upload their policy files in one file location referenced in their GAMECHANGER configuration and run the GAMECHANGER Data API to generate JSON documents from that data. dod-advana/gamechanger-data, GitHub, https://github.com/dod-advana/gamechanger-data/tree/dev/dataPipelines (API to perform the "data engineering work" of GAMECHANGER, including "turning raw publication data into processed JSON format").

additions by human subject matter experts. On top of the graph, GAMECHANGER leverages search and pattern recognition technologies to let users find policy documents and access document metadata.

GAMECHANGER uses on an open-source knowledge graph algorithm called neo4j.[15] Most of the graph can be created from policy documents automatically. AI extracts patterns of words within documents to understand relationships between terms in the policy domain's vocabulary and to extract the relevant "entities" within the policy domain. Based on such pattern recognition, AI also can infer relationships between documents and between documents and entities.[16] On top of this automated process, human subject matter experts can review the graph for inaccuracies or manually code important entities or known relationships.[17]

GAMECHANGER automatically augments the metadata available in the knowledge graph with topic generation and topic modeling technologies. Topic generation identifies a list of most important words within a document based on their relative frequencies in the document and in the rest of the policy corpus. GAMECHANGER uses the tf-idf (term frequency–inverse document frequency) statistic to identify such "topic" terms.[18] Topic modeling groups documents into "topic" clusters (instead of identifying top terms for individual documents).[19] For topic modeling, GAMECHANGER relies on non-negative matrix factorization (NMF), a type of linear algebra algorithms using matrix properties to identify clusters of similar documents.[20] GAMECHANGER's NMF topic modeling features are still in development/testing.

---

[15] *See Knowledge Graph*, NEO4J, https://neo4j.com/use-cases/knowledge-graph (mentioning that neo4j also is used by technology companies such as Lyft, Airbnb, Cisco, and eBay). Neo4J offers paid licenses but provides a fully open source "Community Edition." *See Licensing*, NEO4J, https://neo4j.com/licensing.

[16] Possible relationships include "child of," "mentions," "is similar to," or "is related to." *See* Interview with BAH Data Scientist.

[17] *Id.* Tuning graph generation models requires a fine balance. If the models are made too broad, then the human reviewers need to discard a lot of inaccurate entities and relationships. But if the models are too narrow, many entities and relationships may be missing from the graph, thereby forcing human experts to add them manually—assuming they identify their absence.

As of April 2022, graph generation was not yet leveraging the advanced NLP models used in other parts of GAMECHANGER, but porting it to such advanced models was on the team's short-term roadmap. *See* Interview with BAH Data Scientist (mentioning that the cross-references in the graph are not generated by ML, but that the team plans on building transformer models for these operations).

[18] tf-idf is a standard method to identify important terms within a document. *See* tf-idf, WIKIPEDIA, https://en.wikipedia.org/wiki/Tf%E2%80%93idf. GAMECHANGER uses the open-source implementation of tf-idf provided by Gensim. *See TF-IDF Model*, GENSIM, https://radimrehurek.com/gensim/models/tfidfmodel.html.
It computes the tf-idf metric for all terms within a document and returns the top five terms as "topic" metadata.

[19] Topic modeling is an unsupervised learning method (meaning that it can find patterns in "input" data without being shown "input-output" pairs). *See Unsupervised Learning*, WIKIPEDIA, https://en.wikipedia.org/wiki/Unsupervised_learning. As a result, topic modeling would not provide a name or category for the document clusters it identifies. A human subject matter expert would have to review the cluster and determine what it stands for.

[20] *See* Interview with BAH Data Scientist. NMF is an alternative to Latent Dirichlet Allocation (LDA), described *supra* at text accompanying notes 11-12. *See also* CDO, IMPLEMENTING FEDERAL-WIDE COMMENT ANALYSIS, CDO COUNCIL SPECIAL PROJECTS FINAL RECOMMENDATIONS (June 2021) (using LDA in the CDO pilot for comment analysis to identify clusters of similar comments).

GAMECHANGER applies search technology on top of its knowledge graph of policy documents. It deploys ElasticSearch, a widely used open-source search and analytics engine.[21]

- *Transformer "Large Language Models" for Natural Language Processing (NLP):* GAMECHANGER relies on Natural Language Processing (NLP) to perform most of its policy language analytics functions, such as "Responsibility Explorer" or "Document Comparison." Its NLP models use "transformers," a deep learning technique adapted to training "large language models" relying on millions (if not billions) of parameters.[22] Transformer models have become the benchmark of high-performance NLP.[23] GAMECHANGER uses the open-source versions of these language models wherever possible, to increase transparency and applicability across use cases and organizations.[24]

  Transformer NLP models within GAMECHANGER implement the SBERT (or Sentence-BERT) open-source transformer framework, which is available in the Python programming language.[25] SBERT provides semantic comparisons and semantic search functionalities,[26] which make it especially suited to GAMECHANGER's main policy use cases. SBERT converts each analyzed paragraph into a "paragraph embedding," a representation of the paragraph into a vector space in which such embeddings can be spatially compared.[27] SBERT "embeds" all paragraphs in the policy document corpus into the vector space.[28] To conduct a semantic search, SBERT would convert the query's text into the same vector space and identify the closest embeddings from the corpus. These closest embeddings are estimated to have a high semantic overlap with the query and are returned at the top of search results.

  GAMECHANGER has finetuned its SBERT transformer models for the analysis of policy documents. It used a pre-trained model called "distilroberta-base"[29] as its baseline and trained it on the corpus of forty thousand policy documents within GAMECHANGER to

---

[21] ElasticSearch provides a search engine that centrally stores and organizes data to make searches faster (even as database size increases) and more relevant. *See ElasticSearch*, ELASTIC.CO, https://www.elastic.co/elasticsearch.
[22] *See* Julien Simon, *Large Language Models: A New Moore's Law?*, HUGGINGFACE BLOG (Oct. 26, 2021), https://huggingface.co/blog/large-language-models (mentioning that the BERT-Large model, run by AI company HuggingFace, has 340 million parameters).
[23] *See* Britney Muller, *BERT 101: State of The Art NLP Model Explained*, HUGGINGFACE BLOG (Mar. 2, 2022), https://huggingface.co/blog/bert-101 ("Since their introduction in 2017, Transformers have rapidly become the state-of-the-art approach to tackle tasks in many domains such as natural language processing, speech recognition, and computer vision. In short, if you're doing deep learning, then you need Transformers!").
[24] *See* Interview with BAH Data Scientist.
[25] *See SentenceTransformers Documentation*, SBERT, https://sbert.net/index.html.
[26] *Id.* (mentioning that SBERT is "useful for semantic textual similar, semantic search, or paraphrase mining").
[27] This means that the model can calculate a metric of the "distance" that separates two embeddings in this vector space.
[28] *See Semantic Search*, SBERT, https://sbert.net/examples/applications/semantic-search/README.html (providing an overview of embedding based semantic search and a graphical illustration of how to conceptualize the vector space distance between two embeddings).
[29] distilroberta-base is a variation of the BERT model that has been "distilled" to reduce the size of the model and increase its speed while retaining high levels of accuracy. *See* distilroberta-base, HUGGINGFACE, https://huggingface.co/distilroberta-base.

teach it to operate on policy language.[30] GAMECHANGER also trains different models to refine them for specific policy analysis tasks, such as identifying cross-references within documents.[31] GAMECHANGER implements "symmetric" semantic models as it always runs comparisons between two paragraphs—units of text of roughly the same length.[32]

Transformer models operate under the hood of multiple GAMECHANGER features. With regard to the "Document Comparison Tool," an "embedding" model translates each document's paragraph into an embedding in the vector space,[33] and then a "similarity" model ranks the top results by semantic similarity to the query.[34] Although (as mentioned above) most search functionalities use ElasticSearch, GAMECHANGER's search results display at the top the document most semantically similar to the query using a transformer model.[35] Transformer models also power the "Query Expansion" feature. The model used there embeds the query text into a vector space and finds similar search queries in this space.

C. GSA/CMS Beinformed: KRR POC

- *Machine-Readable Representation of Rules*: KRR leverages machine-readable versions of regulatory information instead of applying predictive models to unstructured regulatory text (like NLP tools do). To translate information into a machine-readable format, KRR relies on ontology models, which create a graphical architecture of the regulatory domain. The ontology's knowledge graph constitutes an abstract representation of the actors, actions, and duties referenced in the regulatory rules and guidance.

- *"Flint frames"*: To code an ontology of the CMS regulations, the pilot leveraged the open source eFLINT standard. eFLINT represents a legal knowledge graph by storing metadata for a regulation that categorizes actors, actions, and duties related to this regulation. Metadata are organized in computer objects called "Flint frames," which include database

---

[30] *See* Interview with BAH Data Scientist.
[31] *Id.*
[32] *See Semantic Search*, *supra* note 28 ("For symmetric semantic search your query and the entries in your corpus are of about the same length and have the same amount of content. An example would be searching for similar questions: Your query could for example be 'How to learn Python online?' and you want to find an entry like 'How to learn Python on the web?'."). The GAMECHANGER team is considering giving users the option to select different levels of comparisons beyond paragraph, including page-level or document-level comparisons. *See* Interview with BAH Data Scientist. It also is considering offering the option to compare between different levels, for example finding all paragraphs that are semantically similar to a queried sentence, which would require the implementation of "asymmetric" semantic models. *See Semantic Search*, *supra* note 28 ("For asymmetric semantic search, you usually have a short query (like a question or some keywords) and you want to find a longer paragraph answering the query. An example would be a query like "What is Python" and you want to find the paragraph 'Python is an interpreted, high-level and general-purpose programming language. Python's design philosophy …'."). The BAH data scientist whom we interviewed mentioned that his team would have to find a new model to implement asymmetric models, but that once they found an appropriate model, making the change would take "two line[s]" of code. Interview with BAH Data Scientist.
[33] *See* Annex Figure 14.
[34] *Id.*
[35] Users can provide feedback on the results (thumbs up / thumbs down), which gets incorporated into further model training.

elements such as the name of an action (e.g., "granting an immigration visa"), the pre-conditions for the action (e.g., "visa applicant must have filled all necessary application forms"), the interested party for the action (e.g., "visa applicant"), and the results from the action (e.g., "grant immigration visa").[36]

Designing a regulatory ontology and filling in the corresponding Flint frames remains a manual process. Subject matter experts can use the Flint Editor, a computer code editor program, to translate a regulation's text into eFLINT's domain-specific programming language.[37] Subject matter experts also would need to continuously monitor updates to regulations previously added to eFLINT and manually reflect these updates into the eFLINT code.[38]

BeInformed and ontology researchers have been evaluating the use of NLP to pre-fill Flint frames.[39] While such NLP tools could not fill the interpretation elements of a Flint frame and therefore will never fully automate the process of creating and maintaining the KRR knowledge graph,[40] they could significantly speed up the process by identifying action, actors, and other frame elements within the regulatory text and automatically build links between Flint frames within the ontology knowledge graph.[41]

- *"Calculemus"*: The domain-specific knowledge stored in eFLINT frames enables other AI programs to search, interpret, and process the underlying regulatory information. BeInformed has developed a protocol called "Calculemus" to program such operations on regulatory rules.

---

[36] The BeInformed team demonstrated examples of eFLINT Frames representing an "act" and a "duty." *See* Interview with BeInformed Team Members.

[37] The BeInformed team demonstrated the use of the Flint Editor, which subject-matter experts use to translate regulations into Flint Frames. *See* Interview with BeInformed Team Members.

[38] The BeInformed team mentioned that agency staff could leverage a notification service alerting them to rule changes, but that updating the rule interpretation within eFLINT (if necessary) would be a manual task. However, a subset of the operations required to update Flint frames could be automated using NLP techniques. *See* Interview with BeInformed Team Members.

[39] *See* BeInformed, *W3.1*, at 13 ("An obvious direction is to leverage NLP tools where possible to support the manual annotation of semantic metadata."); *See Proof-of Concept Briefing*.

[40] *See Proof-of Concept Briefing* (BeInformed Team Member B stating that "making the interpretation" is definitely something that has to be done by a human).

[41] The BeInformed team estimates that NLP can pre-fill up to approximately 80% of the Flint Frames content. *See* Interview with BeInformed Team Members. Although this would never become a fully automated process, such NLP support in building Flint Frames could substantially drive down implementation costs, assuming that it can guarantee sufficient accuracy.