Request for Feedback

AI Grand Challenge for Resilience:
Impact of U.S. Government Policy on COVID-19 using Natural Language Processing & Text Analytics

Agency

General Service Administration's Office of Technology Transformation Services (TTS)

Action

Request for Feedback (RFF)

Summary

GSA TTS is planning for an AI Challenge for Resilience to help to understand the impact of U.S. Government policy on the COVID-19 pandemic within the U.S., to be posted on Challenge.gov in the June - July timeframe.

The AI Challenge for Resilience is TTS' first AI Challenge and the first step to expanding the federal AI community with public, private, academic, and other partners. The primary objective of this challenge is to bring together AI practitioners, with particular emphasis on those who work in the natural language processing and text analytics fields, and policy experts, who understand and navigate the policy domain. The secondary objective of this challenge is to glean insight from the collective impact of regulations across government on our national resilience during the COVID-19 pandemic.

In essence, the policy response is a reflection of where stressors and shocks on the nation begin to show. Rather than waiting until the pandemic is over, the Challenge

Team hopes to begin the retrospective now so the United States can be better prepared for unpredictable crises in the future.

Through this RFF, we seek input on approaches to constructing the challenge to better achieve the above goals. The public input provided in response to this RFF will inform the TTS team who will conduct the challenge and will be beneficial for the development of future challenges. We are particularly interested in alternative or complementary data sources, analytic approaches, and relevant policy domains as well as more general feedback on the tentative challenge process defined below.

Dates
Response Deadline: June 4th, 2021

To Request Further Information and Submit Feedback

For further information and to submit feedback, contact the AI Challenge team at the team email: tts_ai_challenge@gsa.gov. We have provided 8 questions for your feedback at the bottom of this RFF.

Questions, comments, or RFF submissions via email should include (1) "AI Grand Challenge for Resilience" in the subject line of the message and (2) how you came across this RFF in the body of the email. Responses to questions will be posted publicly in github (https://github.com/GSA/AI_Grand_Challenge_For_Resiliency/) and accessible to all. Please refer to the proposed challenge document below in your RFF submissions.

Proposed Challenge

TTS is considering launching the AI Challenge for Resilience: Natural Language Processing & Text Analytics to understand the impacts of the pandemic, assess the U.S. Government response, and strengthen the resilience of communities across the United States.

Since the start of the global pandemic, demand for federal government services has dramatically increased. In 2020, a wide range of policy changes were made to respond to COVID-19, exposing limitations of the existing system. Are executive orders, new regulations, or changes to existing regulations resulting in their intended outcomes?

With the AI Grand Challenge for Resilience series, GSA's primary objective is to bring AI practitioners and policy experts together for one or more challenges to assess resilience in the face of COVID-19. The first proposed AI Challenge for Resilience would seek to understand the landscape and impact of COVID-19 on federal policy-making and the impact of policies enacted during the pandemic. The objective of this challenge would be to use natural language processing (NLP), text analytics, and Artificial Intelligence/Machine Learning (AI/ML) methods to improve the analysis of existing unstructured policy information, demonstrate ways that structure can be applied for improved analysis, and lead to recommendations for improving data-driven regulatory review.

We know that the data created and managed at the federal level is not exhaustive, but a representation of data collected across the country. We invite Challenge participants to leverage additional data in order to fill gaps that would impede the ability to analyze policy outcomes.

The top-level questions we are interested in better understanding through this challenge are:

- How can we characterize the federal response to COVID-19 through policy actions like new regulations or changes to existing regulations?
- What are the relationships between the Federal policy responses and outcomes?
  - Are there data gaps that would impede the ability to analyze the outcomes of policies?
  - How can practices for regulatory review be improved and better supported with data?
- In what ways has COVID-19 related regulatory activity impacted Americans? What can we add or change to the methods for representing, or processing regulatory information that will improve our ability to answer the above questions?

Regulations are primarily expressed in unstructured natural language text, and regulatory analysis involves human resource-intensive knowledge work. Legal policy analysts must discern general characteristics of policies as well as interpret nuance and context within a single policy domain. They often must also navigate complex semantics and interrelationships across policy domains. Overall changes to policies and rules span dozens, if not hundreds, of organizations, are fragmented across silos, and have downstream effects that may not be immediately visible to either rulemakers or the public.

Data

The following datasets associated with this challenge are key to understanding the government response to the pandemic.

Candidate datasets associated with this challenge include:
- Documents published on Regulations.gov https://www.regulations.gov/, reflecting federal level rule changes, additions, and exemptions;
    - Information on the Regulations.gov API can be found here.
    - The Challenge team pulled ~4GB of starter data here. This data contains documents pulled from Regulations.gov under the keyword search terms "COVID", "COVID-19", and "coronavirus" from Nov 2020-Mar 2021. The documents are a mix of JSON, HTML, and PDF formats.
- The Code of Federal Regulations (CFR) from both 2019 and 2020;
- Search logs from coronavirus.gov, which provide some insight into the questions the federal government received from the public at different stages of the pandemic;
- An archive of Executive Orders post-pandemic 2020;

We anticipate that solvers will bring in additional data sets such as census, public health-related, and economic data sets depending on their approach to a solution.

Tools

Regulatory analytics is a narrow domain, but to the extent possible we encourage solvers to leverage existing open source tools such as Quantgov, GameChanger, eRegulations, and Policynet in their solutions to this challenge.

Also, we are interested in solutions that leverage ontologies, XML, and other formalisms, and contribute to knowledge generation in the regulatory ontology space. We encourage solvers to present solutions that leverage prior work in this area. Research presented at JURIX and ReMep, and packages such as LKIF, Lexicog, CEN Metalex, Akomantoso, USLM, are good places to start when considering technical approaches to this challenge.

Stakeholders

Personnel from the GSA Office of Regulation Management are key stakeholders for this challenge, have helped to frame it, and will be part of the cohort of challenge evaluators. Part of the Office of Regulation Management's mission is to: "Support Federal agencies as they develop regulations, including management of agencies' recordkeeping systems, known as dockets, as well as review of public comments." During the course of the challenge, personnel from the Office of Regulation Management will be made available during public meetings, such as webinars, to ensure that solvers have the information they need to ensure their solutions are relevant to the regulatory process.

Additionally, organizations like the Undersecretary of Defense - Comptroller's Digital Transformation Office are sponsoring technology modernization efforts around policy discovery, evaluation, and analytics. The Gamechanger initiative is an example of one such activity, which has open sourced their data science and engineering tools to the public.

Challenge Categories

Data-driven policy analysis has the potential to improve and better coordinate policymaking, leading to improved outcomes across a range of issues such as food supply, education, health care, economic recovery, and resilience.

The TTS Challenge Team is considering accepting submissions across the following three categories once the challenge competition begins:

1. **Text Summarization, Classification, & Topic Modeling:** Using natural language processing techniques, describe the types of changes and updates made to federal regulations, policies, and rules.
2. **Named Entity Recognition, Data Annotation, & Document Linking:** Combining natural language processing (NLP) and knowledge representation techniques to create policy-domain specific knowledge graphs, which show relationships between organizations, topics, and policy issuances. Alternatively, leveraging labeling and annotation methods to create high quality, enriched datasets to be used for training NLP or AI/ML models for policy domain-specific tasks.
3. **Impact Modeling:** Quantitatively understanding the impact of policy on resilience. The solver will leverage analytic techniques from disciplines like operations research, statistical analysis, and mathematical modeling to enable inferences about the social and economic impact of COVID-19 related

policy changes. The provided solution must include an NLP and/or other text analytics, but can be augmented with other data, analytic techniques, and data visualization.

Judging Criteria

The Challenge team is considering using the following evaluation criteria for the submissions:

1. **Data Preparation and Analytic Output :** Creative use of at least one of the required data sets and when necessary, supplement with additional datasets to achieve a meaningful analytic output. Additionally, this score focuses on the performance, output, and interpretation of NLP, text analytics, and AI/ML techniques employed in the project.
2. **Technique Selection and Analytic Methodology:** This score focuses on the technical approach to employing NLP, text analytics, and other AI/ML techniques in the project.
3. **Project Overview and Analytic Storytelling:** This score focuses on the overall communication of the project, central thesis, and the "so what."
4. **Explainability, Transparency, Responsibility, and Trust:** This score also includes communication about strengths, weaknesses, and limitations behind the analytic techniques used, which could include things like bias in the data, limited data, technical limitations of AI/ML techniques.
5. **Future Practical Application:** This score focuses on how the solution provided could be applied in combination with data practices, processes and services to support ongoing policy analysis.

Provide Your Feedback to the Following Questions

1. Are there specific data sets (E.g. Federal policy corpus, geospatial, COVID-19 related) that you think would be critical to include as part of the Challenge?
2. While evaluating policy impact and outcomes is difficult, is there sufficient data and possible research areas to test and analyze hypotheses related to policy impact?
3. What are your thoughts on the level of guidance/access to policy domain expertise required to make this challenge a success?
4. Are there additional submission categories that you believe further the overall goals of the Challenge that should be considered?
5. Are there additional evaluation criteria or modifications to these criteria that

you think should be included/made?

6. What are your thoughts on our plan to use a publicly-viewable repository, version control and collaboration platform as the method for challenge dissemination and submission collection?

7. Currently, the Challenge is structured to provide monetary awards to winner(s). Do you think the Challenge would be most effective as a prize competition including monetary and/or honorary awards, or a non-competitive crowdsourcing activity?

8. What have we missed, what other thoughts do you have in relation to the proposed Challenge?