IBM Open-Source Software Contribution Request Form

## Research Template

Request Issue Id:

Current Date: 09/24/2024

# A. Summary Section

1. **Project Name:** Please provide the name of the Project.  
   EMNLP 2024 LLM robustness
2. **Project Description:** Provide the description of the Open Source project, including the planned IBM contributions, what significant functionality will you provide, the programming languages used, how many lines of code etc.  
   The project will contain results (csv files) and visual analysis (in Juypter notebooks), as well as some minimal Python code, relating to a paper we had accepted at the EMNLP conference: https://arxiv.org/html/2408.01963v1. The materials will not replicate the results in the paper in entirety, just showcase the final results.
3. **Contribution Type:** Does the community you are contributing to already exist? Please indicate one of the following:

* a.) Existing community
* b.) New community under the IBM org
* c.) New community outside the IBM org
* d.) Small contribution, prototype, Proof of Concept, Paper/Conference, external collaboration with University etc.

1. **Ownership/Author:** Does the code you are contributing already exist? Please indicate one of the following:

* a.) Existing code written only by IBM
* b.) Existing code written by IBM and third parties
* c.) Existing code written by third party(s) (including open source)
* d.) New Code. If this is New Code, is this IBM created code?

1. **Publicity:** Will there be any widespread publicity? IBM or partner press release, conference announcement, new community, major Blog, or X (formerly Twitter) posts etc.  
   No
2. **Legal Review:** Does this Open Source contribution request require legal review? For example: any unique, non-preapproved Contributor License Agreement (CLA/DCO), IP issues, major contribution, patents, or patent disclosures pending require legal review. (Most minor contribution proposals and pre-approved CLAs do not require legal review. For more information, see: [https://w3.ibm.com/developer/docs/open-source/oss-contributing-non-special/#snippet](https://w3.ibm.com/developer/docs/open-source/oss-contributing-non-special/#snippe) (doesn’t apply to IBM Research)).  
   No
3. **Have you completed the mandatory open source training?** <https://yourlearning.ibm.com/activity/EL01-00001160>

# B. Business Section

1. Will you be acting as a committer/maintainer in the project?
2. What resources are in place to participate in the project?   
   Describe the resources, state how many developers, and for how long.   
   *What are the ramifications if your team doesn’t move this project into Open Source?*  
   The code already exists in entirety. There are two contributors. The move to open source is meant simply to make the results for the paper transparent and available in the future to readers.
3. Will you provide community building support? If yes, by who? If no, who will build and manage the community?  
   No
4. Who are the IBM technical leads and key contributors?  
   * Development manager (Name and IBM e-mail)
   * Submitters (Name(s) and IBM e-mail(s)) Samuel Ackerman (samuel.ackerman@ibm.com), Ella Rabinovich (ella.rabinovich1@ibm.com)
   * Director's Email (Name and IBM e-mail) Wesam Ibraheem (wesam@il.ibm.com)

All the Open-Source requests required Director level approvals. Please append to your email a screenshot or a copy of the email/Slack with Director's approval for this project.

If you have no Director to approve, then please ask your VP, or have your VP nominate a delegate to approve.

1. What is the reason behind contributing this code to Open Source?   
   Describe your reasoning. For example: Drive developers to IBM's platforms, provide support for additional architectures etc.  
   To make the results of the paper transparent and available to readers.
2. Will this be  a one-time contribution or  an ongoing contribution?
3. Is this a contribution for open hardware?
4. Will this be a contribution related to foundation models?   
   If yes, please follow one or more processes as appropriate.
   * a.) **Software** – model architecture, Inference and Training code, scripts for processing the model, e.g., Python code, or publications, no models, no data. (*You can continue processing this form*).
   * b.) **Weights and Measures** – actual model weights, numerical values and variables, used for transforming the input data   
     (You cannot continue with this submission and must obtain the clearance from the DMG (Data and Model Governance) team <https://one-ibm-tool.dal1a.cirrus.ibm.com/>).
   * c.) **Data Sets** used for training and inference.   
     (You cannot continue with this submission and must obtain the clearance from the DMG (Data and Model Governance) team <https://one-ibm-tool.dal1a.cirrus.ibm.com/>).
   * d.) **Something else**, please explain.  
     The datasets were already publicly released. This simply contains the outputted results of LLM generation in csv form, and the code to visualize the outputs.
5. Will you open source dataset or data content (including foundation models and weights and measures)?   
   If yes, please obtain legal approval as specified under the following document: <https://w3.ibm.com/ibm/privacy/knowledge-base/using-data-ibm>.
6. Does the project require integrating other existing external open-source code?   
   List open source projects (if any) that this project requires to be integrated with, including any dependencies.  
   None
7. What non-open technical dependencies does the target Open-Source project have?   
   For example, Db2.  
   None
8. Does open sourcing this code disrupt a competitor (either Open-Source Software or proprietary)?   
   Please explain and list your project competitors (if any) in Open Source providing similar features.   
   No
9. Are any IBM and/or non-IBM products or solutions affected if your team does not move this project to Open Source?   
   If yes, in what way?  
   No
10. List the IBM product or customer engagement(s) that owns the code.   
    Include the owning executive's approval email to contribute.  
    The project was part of Research challenge 4898 "Trust in large language models". Since this is a one-time contribution for a paper repo, there will be no other contributors other than those listed already. The owning executive's email is wesam@il.ibm.com.

# C. Legal Section:

1. If the contribution is made to new open-source projects:
   * What Open-Source license(s) will the new project use?
   * Will it accept 3rd party contributions and if so, what CLA will be utilized?
2. Does the contribution contain any valuable IBM IP?   
   Valuable IP must be reviewed with the owning business units and legal. Please explain. No
3. Does the contribution have any cross-business unit business or legal concerns? Does your contribution affect the business interests of another group, or have broad legal implications for IBM (for example, is this code being used by other teams at IBM/Red Hat or outside IBM) ?   
   Note: Answer "Yes" here if you seek to open-source code owned by another team, or the open source code contribution will be consumed in products and services in another business unit.  
   No
4. Provide the Open-Source project's steps you must follow to contribute material to the project.   
   If the Open-Source project has a unique Contributor License Agreement (CLA), Developer Certificate of Origin (DCO), or any other agreement you must accept to contribute material to the project, provide the link or attach the text for review.   
   None

# D. Security Section

*TBD*

# E. Export Section

*TBD*

# F. Technical Details

1. Business Unit: **Research**
2. Research Pillar:
3. Research Challenge Link *e.g.:* [*https://challenges.apps.res.ibm.com*](https://challenges.apps.res.ibm.com)   
   https://challenges.apps.res.ibm.com/challenges/4898

Note: we review the challenge to ensure that a project overview together with the rationale for open sourcing appears in the challenge. If there is no adequate description, we will ask you to create a challenge Progress Report that contains the following items:

1. quick overview of what you are open-sourcing  
2. quick overview of why you are open-sourcing particular code  
3. link to the current, internal GitHub repository  
4. link to the intended, target, public repository  
5. then please capture this part of the challenge and include the screenshot with this submission.

Our motivation is to alert the challenge members and relevant executives, giving them the opportunity to show their support and address any concerns about the open sourcing request.

1. Internal repository URL:   
   Please provide the current internal repo for this project.  
   e.g.: <https://github.ibm.com/my-project>  
   https://github.ibm.com/Samuel-Ackerman/EMNLP\_2024\_LLM\_robustness
2. Proposed public repository URL:   
   Please provide the proposed, public repo for this project. (Note: code cannot be published prior to the approval of this request):   
   e.g: <https://github.com/IBM/my-project>  
   https://github.com/IBM/EMNLP\_2024\_LLM\_robustness

# G. Comments and Discussion:

# H. Approvals Section:

This section is intended for OSS Managers only.

* Legal Review Required:
* Manager Review Required: Name:       Date:
* Executive Review Required: Name:       Date:
* Foundation Models DMG Review Required: Name:       Date:
* OSS Current Approval Status:
* OSS Approval Record:   Date: