# 仕　様　書

## １．件名

MLflow の環境構築および大規模言語モデルの評価手法の調査作業

## ２．研究の概要・目的

### 2-1.概要・目的

国立研究開発法人産業技術総合研究所（以下「産総研」という）人工知能研究センターでは、機械学習の品質評価研究開発と AI 国際標準化の推進において、機械学習システムの開発に必要なデータセットと機械学習モデルパイプラインを管理しながら、品質を検査・評価するオープン型テストベッドの構築を目指している。本作業は、AWS クラウドサービスを利用して MLflow の環境を構築し、MLflow 上に大規模言語モデル（以下「LLM」という）を評価する手法を調査する作業である。

### 2-2.用語の定義

本仕様書で使用される用語とその意味について、以下に記す。

| カテゴリ | 用語 | 説明 |
|---|---|---|
| 組織及び人物 | 産総研担当者 | 本システムの企画及び運用等を担当する者及び所管部署の業務運用担当者。 |
| | 調達担当者 | 本調達の契約手続き等を担当するもの。 |
| | 受注者 | 本調達の対象となる業務に従事する事業者。 |
| その他 | 情報セキュリティインシデント | 産総研が望まない単独若しくは一連の情報セキュリティ事象、又は予期しない単独若しくは一連の情報セキュリティ事象であって、事業運営を危うくする確率及び情報セキュリティを脅かす確率が高いもの。 |
| | 情報セキュリティポリシー | 産総研の情報セキュリティ基本方針、情報セキュリティ規程、情報セキュリティ実施要領及び情報セキュリティ実施ガイドの総称。 |
| | MLflow | 機械学習のライフサイクルを管理するためのオープンソースソフトウェア<br>https://mlflow.org/ |
| | 機械学習品質マネジメントガイドライン | 機械学習を利用した AI システムのライフサイクル全体にわたる品質マネジメントのために産総研が作成した品質要求の仕様ガイドライン<br>https://www.digiarc.aist.go.jp/publication/aiqm/ |
| | AWS クラウドサービス | Amazon 社が提供しているインターネット経由でコンピューティング、データベース、ストレージなど、さまざまな IT リソースをオンデマンドで利用することができるサービス<br>https://aws.amazon.com/jp/cloud/ |

３．作業項目

（１）MLflow の環境構築作業

（２）LLM 評価手法の調査作業

４．作業項目別仕様

（１）MLflow の環境構築作業

① 環境構築のため、AWS クラウドサービスを利用し（②～⑤の要件を満たす）、プライベートネットワーク上でシステムアーキテクチャを設計すること。

② Amazon RDS（AWS マネージドリレーショナルデータベースサービス）を利用してモデルのメタデータとログを保存すること。保存項目を Github に記入すること。

③ Amazon S3（AWS オブジェクトストレージサービス）もしくは Hugging Face に共有されている LLM を読み込むための設定を行うこと。MLflow の設定情報は Github に記入すること。

④ Amazon EC2（AWS 仮想コンピューティングサービス）インスタンスを利用して Linux OS 上 MLflow をホスティングすること。ホスティングのインスタンスの仕様を調査し、Github に記入すること。

⑤ Amazon VPC（AWS 仮想ネットワークサービス）を利用してプライベートネットワークを構成すること。

（２）LLM 評価手法の調査作業

① mlflow.evaluate にパッケージ化された MLflow の LLM 評価機能を調査し、Github に記入すること。

② MLflow の評価メトリックを調査し、Github に記入すること。

③ LLM 評価のためカスタム pyfunc モデル（Python プログラム）を作成すること。

④ 産総研が貸与するテストデータを利用して、４（２）③で作成したカスタム pyfunc モデルを利用して LLM を評価すること。評価結果を Github にまとめること。

⑤ LLM を評価するメトリックと機械学習品質マネジメントガイドラインで示している評価指標と比較し、Github に記入すること。

⑥ 機械学習品質マネジメントガイドラインで示している評価指標と MLflow のログを利用して、AITHub 用の機械学習システムの品質評価手法（以下、AIT）を開発する方法を調査し、AIT 開発ガイドラインを作成すること。

⑦ 以下の SAFE 手法を調査し、４（２）③のように MLflow に取り込むこと。

https://github.com/google-deepmind/long-form-factuality

５．貸与品

① 作業内容や進捗報告用の Github のリポジトリー（電子　メール貸与）
② AWS にアクセスする作業用のノート PC（Windows 11 Pro）
③ AWS サービスを利用するアカウント（電子　メール貸与）
④ Docker Desktop、Office365のライセンス(電子　メール貸与)
⑤ テストデータセット（電子　ファイル転送アプリにて貸与）

６．特記事項

① １回/週の頻度で Online 会議を実施し、作業状況、中間結果などを共有すること。
② 貸与した情報及びシステム内容について第三者への提示を行わないこと（ただし、産総研担当者の了解を得た場合は除く）。
③ 作業者は下記の能力、要件を満たすものとする。
- Github についての十分な知識を有し、Github を利用した開発経験があること。
- Python のプログラミング経験があること。
- Linux 上での開発経験があること。
- AWS サービスを利用してシステム開発・運営の経験があること。
- 本プロジェクトに関連する機械学習品質マネジメントガイドラインについての十分な知識を有し、AIT を開発した経験があること。

７．納入物品

以下、メールや産総研指定のファイル転送アプリなどにより納入すること。

① Python ソースコード　一式
② 作業報告書　１部
・ 環境構築の手順書、メトリックの調査書、AIT 開発ガイドライン

８．納入の完了

作業完了の後、「７．納入物品」に記載された納入物品が過不足なく納入されたことを確認して、納入の完了とする。

９．納入期限及び納入場所

納入期限：2024年9月30日
納入場所：東京都江東区青海2-4-7
国立研究開発法人産業技術総合研究所人工知能研究センター
臨海副都心センター別館9F 92020室

１０．付帯事項

（１）本仕様書の技術的内容及び知り得た情報については、守秘義務を負うものとする。

（２）本仕様書の技術的内容に関する質問等については、産総研担当者と協議すること。また、本仕様書に定めのない事項及び疑義が生じた場合は、調達担当者と協議のうえ決定する。

（３）本作業完了後１年以内の故障・不具合については、その修理・調整作業等を無償で実施すること。

（４）受注者の責において及ぼした損害は、受注者が賠償すること。

（５）サプライチェーン・リスクに対応するため、「IT 調達に係る国等の物品等又は役務の調達方針及び調達手続きに関する申合せ」（平成30年12月10日関係省庁申合せ）に基づき対応を求めることがあるので応じること。

１１．セキュリティ要件

11.1.情報セキュリティポリシーに関する要件

①本業務の履行に当たっては、産総研の情報セキュリティポリシー(別途定める読み替え条項に従うものとする。以下同じ。)を遵守するとともに、情報セキュリティポリシーにおいて産総研に求められる水準の情報セキュリティ対策を講じること。なお、産総研の情報セキュリティ規程については、下記 URL を参照のこと。その他の情報セキュリティポリシーの詳細については受注者決定後に提示する。

【国立研究開発法人産業技術総合研究所情報セキュリティ規程】

https://www.aist.go.jp/Portals/0/resource_images/aist_j/outline/comp-legal/pdf/securitykitei.pdf

②産総研の情報セキュリティポリシーの見直しが行われた場合は、見直しの内容に応じた情報セキュリティ対策を講じること。なお、対応内容については産総研担当者に事前に報告し承認を得ること。

11.2.その他セキュリティに関する要件

① 受注者は、本業務の履行に際して、秘密である旨を示されて貸与を受けた秘密情報を秘密として適切に保持することとし、第三者に開示又は漏洩してはならない。

② 受注者は、本業務の履行によって知った一切の情報を本業務の履行以外の目的に利用してはならない。契約終了後も同様とする。

③ 産総研の所外へ持ち出した資料については一覧を作成し、産総研担当者に提出すること。なお、契約終了後、速やかに返却または廃棄し、産総研担当者に報告すること。

④ 本業務にかかる情報に関する情報セキュリティインシデントが生じた

場合、速やかに報告の上、原因の分析を実施し、請求担当者と対処内容及び再発防止策を検討すること。当該インシデントへの対処を実施するにあたっては、事前に請求担当者の確認を得ること。

⑤ 情報セキュリティインシデントが生じたことで、受注者の作業環境等の確認が必要となった場合には、産総研の調査に協力を行うこと。

⑥ 本業務の遂行における情報セキュリティ対策の履行状況を確認するために、産総研が情報セキュリティ監査の実施を必要と判断した場合は、産総研が定めた実施内容（監査内容、対象範囲、実施者等）に基づく情報セキュリティ監査を受注者は受け入れること。

⑦ 産総研の許可なく、作業の一部又は全部を第三者（再委託先）に請け負わせてはならない。ただし、受注者に求めている情報セキュリティ対策を、再委託先が実施することを再委託先に担保させるとともに、再委託先の情報セキュリティ対策の実施状況を確認するために必要な情報を産総研に提供し、産総研の許可を受けた場合はこの限りではない。

## １２．成果の取扱い

⑴ 産総研は、受注者がプログラム作成により得られた技術上の成果のうち産総研が指示するもの（以下「成果」という。）についての利用及び処分に関する権利を専有するものとする。

⑵ 受注者は、成果に係るソフトウェアの著作権（著作権法第27条及び第28条に規定する権利を含む。）及び意匠登録を受ける権利を産総研に譲渡するものとし、著作者人格権を行使しないものとする。ただし、パッケージ製品に係るものは除く。

⑶ 受注者は、契約条項に定める検査に合格後、直ちに別紙様式による著作者財産権譲渡証書及び著作者人格権不行使証書を産総研に提出しなければならない。

⑷ 受注者は、産総研に対し、納品した成果品が第三者の知的財産権を侵害しないことを保証するものとする。なお、納品した成果品について、第三者の権利侵害の問題が生じ、その結果、産総研又は第三者に費用や損害が生じた場合は、受注者は、その責任と負担においてこれを処理するものとする。

別紙様式

〇〇〇〇年〇〇月〇〇日

# 著作者財産権譲渡証書

国立研究開発法人産業技術総合研究所　殿

受注者
住　所
会社名
代　表　者　氏　名
印

作業請負契約　　（〇〇〇〇年〇〇月〇〇日　契約）
件　名

上記契約により作成したソフトウェア等の成果物の所有権及び著作権（著作権法第27条及び第28条に規定する権利を含む）は、国立研究開発法人産業技術総合研究所に譲渡したことに相違ありません。ただし、本契約前に自己所有していた権利は除くものとします。ただし、上記契約締結前に自己所有していた権利は除くものとします。

別紙様式

〇〇〇〇年〇〇月〇〇日

# 著作者人格権不行使証書

国立研究開発法人産業技術総合研究所　殿

受注者
住　所
会社名
代　表　者　氏　名
印

作業請負契約　　（〇〇〇〇年〇〇月〇〇日　契約）
件　名

上記契約により作成したソフトウェア等の成果物の著作権（著作権法第27条及び第28条に規定する権利を含む）に係わる著作者人格権を行使しないことを約束します。

なお、著作者人格権を行使しようとする場合は、国立研究開発法人産業技術総合研究所の承認を得るものとします。