■ 公募情報

公募No.	2027T-87
職種	有期研究技術員
部署	サイバーセキュリティ研究所AIセキュリティ研究センター
業務名	大規模言語モデル・AIエージェントのサイバーセキュリティ能力に関するベンチマーク評価
業務内容	AIセキュリティ研究センターでは、大規模言語モデル（LLM）やAIエージェントが有するサイバーセキュリティに関する能力を客観的かつ継続的に把握するため、各種サイバーセキュリティタスクを用いたベンチマーク評価を実施している。 本業務では、商用及びオープンソースのLLM・AIエージェントを対象として、脆弱性の発見・解析、コード生成・解析、CTF等のセキュリティ課題、その他のサイバーセキュリティ関連タスクに対する能力評価を実施する。また、評価を継続的かつ再現可能に実施するためのベンチマーク、評価データセット、評価環境及び評価ツールの整備・改良を行う。 具体的には、経験・適性等を踏まえ、以下に示す業務のうち、いずれか又は複数の業務を担当する。 ・商用及びオープンソースのLLM・AIエージェントに対するサイバーセキュリティ能力評価の実施 ・脆弱性解析、コード解析・生成、CTFその他のサイバーセキュリティタスクを用いたベンチマークの実施 ・LLM及びAIエージェントの実行環境、API接続環境、サンドボックス等の評価環境の構築・運用 ・ベンチマーク評価システム、評価スクリプト及び評価ツールの実装・改良・自動化 ・評価データセット、タスク及び評価シナリオの作成・整備 ・複数のAIモデル・AIエージェントに対する反復的な評価の実施及び結果の収集・管理 ・評価結果の集計、統計処理、可視化、比較分析及び技術文書・報告書の作成 ・新たに公開されたAIモデル、AIエージェント及び評価手法への継続的な対応
自発的な研究活動等の実施に関して	職務への専念義務があります。
科学技術・イノベーション創出の活性化に関する法律第15条の2の対象業務該当の有無	【有】
応募要件	【求める経験・スキル】 ・Python等のプログラミング言語を用いて、評価スクリプトやツールの作成・修正ができること ・Linux環境におけるソフトウェアの実行、設定及び実験環境の構築に関する知識又は経験を有すること ・大規模言語モデル（LLM）、生成AI又はAIエージェントに関する基礎的な知識を有すること ・サイバーセキュリティに関する基礎的な知識を有すること ・多数の実験を一定の条件の下で実施し、その結果を正確に収集・整理・分析できること ・研究員その他の関係者と連携し、主体的に評価業務を遂行できること ・英文の技術文書、論文、モデル・ソフトウェアのドキュメント等を読むための基礎的な英語力を有すること 【歓迎する経験・スキル】 ・LLM又はAIエージェントを用いたシステムの開発・評価経験 ・LLM API、オープンソースLLM、推論フレームワーク等の利用経験 ・CTF、脆弱性解析、ペネトレーションテスト、セキュリティツール等に関する知識又は経験 ・AI又はソフトウェアのベンチマーク評価、データセット構築又は大規模実験の経験 ・Docker等を用いたサンドボックス・実験環境の構築経験 ・実験結果の統計処理、データ分析又は可視化の経験 ・AIエージェントによるツール利用や自律的なタスク実行に関する知識又は経験
募集人員	1 人
本年度契約期間	採用日　～　令和10年3月31日（更新の可能性：有り）
更新した場合の雇用期間（又は期日）	一定の条件を満たした場合に、採用日より最長5年
給与(基本給)	485,000円　～　541,000円/月 本給は学歴や職務経験等を考慮し決定します。ただし、本給については、国家公務員の給与に準拠していることから国家公務員の給与に改正があり、当機構労働組合等の合意後に本給の改定が生じた場合は変更する。
勤務地名称	イノベーションセンター (東京都中央区)
勤務頻度	週5日（週37時間30分勤務）　※時間外労働有

※従事する業務及び勤務地の変更の範囲：原則として変更無し
※部署の名称、勤務地の名称、及び業務名や業務内容内の表現に関しては、組織改編等により変更となる場合があります。

NICT有期雇用職員公募情報：https://www.nict.go.jp/employment/apply.html