

GENIAC-PRIZE（NEDO 懸賞金活用型プログラム）

～生成 A I の安全性確保に向けたリスク探索及びリスク低減技術の開発～

1. 事業趣旨

1. 1. 背景及び目的

- 生成 A I（文章、画像、プログラム等を生成できる A I モデルにもとづく A I の総称）は、生産性や付加価値の向上等を通じて大きなビジネス課題を引き出すとともに様々な社会課題の解決に資することが期待されています。
- 国際的な生成 A I の開発競争は激しさを増し、また、生成 A I のリスクへの対処の必要性が議論されている中、経済産業省や NEDO では、様々な取組を行っています。
 - 経済産業省と NEDO は、国内の生成 A I 開発力の底上げを目的とし、2024 年 2 月から「GENIAC (Generative AI Accelerator Challenge)」プロジェクトを立ち上げ、基盤モデルの開発に必要な計算資源の提供支援、データの利活用やナレッジシェアに向けたコミュニティの運営等を行っています。
 - 経済産業省は総務省と共同で、A I（生成 A I を含む）の安全性にかかる A I ガバナンスの統一的な指針として、2024 年 4 月に「A I 事業者ガイドライン」を策定・公表しています。
 - 関係府省庁との連携により 2024 年 2 月に新たに独立行政法人情報処理推進機構（I P A）内に設置された A I セーフティ・インスティテュート（A I S I）は、米英をはじめとする国際的なパートナーとも連携しながら、相互運用可能な生成 A I の安全性評価手法の確立を目指しています。
- 一方で、生成 A I のリスクの観点においては、生成 A I が新たなコンテンツ（テキスト・画像・音声など）を出力する創造的なタスクを処理することから、活用される様々な場面において具体的にどのようなリスクが生じうるか、国際的にも共通の理解が醸成されている訳ではありません。現在も、国際 A I 安全会議の場などで継続的に議論されている状況です。以下に示すようなリスクは顕在化し社会的に認識されているものの、具体的なリスクを網羅的に示すことは非常に難しい課題です。
 - 生成 A I で特に顕在化したリスクの例
（「A I 事業者ガイドライン（第 1.1 版）」別添より抜粋）
 - ☆ ハルシネーション：生成 A I が事実と異なることをもっともらしく回答する。
 - ☆ 偽・誤情報等の流通・拡散：生成 A I が生み出す誤情報を鵜呑みにすることがリスクとなりうる。ディープフェイクは、各国で悪用例が相次いでいる。
 - ☆ 知的財産権等の侵害：知的財産権の取扱いへの議論が提起されている。
- 生成 A I により生じ得るリスクの低減にあたっては、リスクを特定することだけではなく、対策となる技術の開発も同時に求められます。

- 本事業を通じて、社会全体として生成A Iにより生じ得るリスクを認識するとともに、リスクを低減する技術開発を促進することで、わが国における生成A Iの受容性を高め、生成A Iの利活用・社会実装を後押しすることを目的とします。

1. 2. 事業概要

- 本事業では、生成A Iの安全性に関わるリスクを特定すると同時に、当該リスク低減を目的とする技術開発を広く募集します。
- 例示として、以下のような事象（「A I事業者ガイドライン（第1.1版）」別添より抜粋）への対策となる技術開発が考えられます。
 - ▶ データ汚染攻撃等のA Iシステムへの攻撃
 - ▶ バイアスのある出力、差別的出力、一貫性のない出力等
 - ▶ ハルシネーション等による誤った出力
 - ▶ ブラックボックス化、判断に関する説明の不足

そのほか、A Iが生成した結果であるかを区別する技術の開発や、生成A Iの安全性そのものを評価するベンチマークの開発なども考えられますが、これに限るものではありません。

- 本事業においては、アイデアのみの応募に対して懸賞金は支払われません。提案技術によるリスク低減の可能性や実効性を示す成果物の提示や評価の実施、（デモ実演を含む）審査を受けることなどが求められます（詳細は項目3の応募内容・審査内容を参照ください）。
- 生成A Iの安全性に資するものとして活用するため、応募いただいた内容は基本的に全て公開されますが、事業性を保つ目的など非公開であることが必要な範囲については考慮しますので、公開を限定する範囲及びその理由についても併せて応募書類に記載してください。ただし、成果の公開度は審査項目にもなっておりますのでその点は留意ください。
- なお、本事業では、国内で開発された基盤モデルだけでなく、国外で開発された基盤モデルを活用してもよいものとしますが、情報漏洩などに配慮し、安全性が確保された信頼性のある環境にて開発を実施いただきますよう留意ください。

2. 応募内容及び審査内容

2. 1. 応募資格

- 本邦の企業、大学等の法人であることを必須とします。
- 1つの法人から応募いただく数に制限は設けませんが、トライアル審査においては、提案技術の開発に必要な体制が整っているか確認を行う場合があります。

2. 2. 応募内容

本事業において応募者に応募いただく内容は、以下とします。

- 生成A I のリスクの特定とその低減に資する技術（以下の2つの要件をいずれも満たすものとします）

① 生成A I の安全性に関わるリスクを特定すること

② ①のリスクを低減する技術を提案すること

※①・②について、定量的な評価手法を併せて提案することが望ましい

2. 3. 審査内容

2. 3. 1. 審査方式

（トライアル審査）

- トライアル審査は、書面審査及び応募者によるプロトタイプの実演を通じた審査とします。
 - 「プロトタイプ」とは、以下を満たすものとします（当該技術を用いた製品の販売有無は問いません）。
 - ☆ 応募者が提案する技術が、応募者が特定したリスクを低減するという目的において有効である可能性を示すことができる状態であるもの
- プロトタイプの提示と実演は必須の要件とします（プロトタイプは既存のものを活用いただいても問題ありません）。
- 2. 3. 2. に記載した審査基準に基づき、生成A I に関連する知見を有する者が審査を行います。
- トライアル審査において懸賞金を獲得した者は、原則本審査への応募を必須とし、応募されなかった場合には原則懸賞金を全額返金いただきます。

（本審査）

- 本審査は、書面審査、プロダクトのテスト結果及び実演の結果を加味した審査とします。
 - 「プロダクト」とは、以下を全て満たすものとします。（当該技術を用いた製品・サービスの販売有無は問いません）。
 - ☆ 応募者が開発した技術が、応募者が特定したリスクを低減するという目的において有効であることを示すことができる
 - ☆ 技術的な専門性を有する者であれば応募者側・事務局・審査員を問わず、審査期間中にテストを実施することができる状態である
- 本審査においては、応募者が開発した技術の実効性について、応募者側の提案する評価方法にて評価した結果の提示を必須とします。（審査ではその評価方法の妥当性も問われます）
- 本審査においては、応募者が開発した技術の実効性を、主催者側の審査員が実際に操作することで、もしくは、審査員の指示に従い応募者に操作していただき、確認ができる環境を応募者側が用意することを必須とします。（参考欄もご参照ください）

- 本審査においては、応募者が開発した技術についての課題（特定したリスク低減における課題や派生して発生するリスク）についても説明を求めます。
- トライアル審査の結果に関わらず、また、トライアル審査に参加しない応募者であっても、本審査に参加することは可能です。
- 2. 3. 2. に記載した審査基準に基づき、生成A Iに関連する知見を有する者が審査を行います。

2. 3. 2. 審査基準

(トライアル審査)

審査項目（大項目）	審査項目（中項目）
懸賞広告との合致性 (必須)	<ul style="list-style-type: none"> ・生成A Iの安全性に関わるリスクの特定 ・当該リスクを低減する技術を提案 ・プロトタイプの提示
特定したリスクの評価	<ul style="list-style-type: none"> ・特定したリスクの妥当性や影響度
対策技術（プロトタイプ）の評価	<ul style="list-style-type: none"> ・提案する対策技術（特定したリスクに対する低減手法）の妥当性 ・提案する対策技術（特定したリスクに対する低減手法）において目標（可能な限り定量的な目標）とする技術レベル・性能の妥当性 ・提案する対策技術の評価手法の妥当性
新規性の評価	<ul style="list-style-type: none"> ・新規性（応募者が提示する根拠の妥当性を含む）
公共性の評価	<ul style="list-style-type: none"> ・成果の公開度

・ 補足事項

トライアル審査では、特に提案の妥当性の観点を重視します。

(本審査)

審査項目	審査の内容
懸賞広告との合致性 (必須)	<ul style="list-style-type: none"> ・生成A Iの安全性に関わるリスクの特定 ・当該リスクを低減する技術を提案 ・プロダクトの提示
特定したリスクの評価	<ul style="list-style-type: none"> ・特定したリスクの妥当性や影響度
対策技術（プロダクト）の評価	<ul style="list-style-type: none"> ・提案する対策技術（特定したリスクに対する低減手法）の妥当性 ・提案する対策技術の評価手法の妥当性 ・開発した対策技術（特定したリスクに対する低減手法）の性能評価の結果
新規性及び将来性の評	<ul style="list-style-type: none"> ・新規性（応募者が提示する根拠の妥当性を含む）

価	・将来性（対策技術の課題点を把握できているかという観点を含む）
公共性の評価	・成果の公開度 ・国民生活や社会への波及効果

・ 補足事項

本審査では、特に開発された技術の有効性や、公共性の観点を重視します。

2. 3. 3. 審査員

政府機関（AISI、国立研究機関）・アカデミア・民間などの幅広い領域の有識者で構成します。

審査員名（敬称略）	所属情報
大塚 玲	人工知能学会 安全性とセキュリティ研究会(SIG-SEC) 主査 情報セキュリティ大学院大学 教授
小田 悠介	国立情報学研究所（NII）特任准教授 シェルパ・アンド・カンパニー株式会社 AI 事業部テクニカル・ディレクター
加藤 敏洋	独立行政法人情報処理推進機構（IPA） AI セーフティ・インスティテュート（AISI）研究員
関根 聡	国立情報学研究所（NII） 科学主幹 特任教授
津田 麻紀子	西村あさひ法律事務所 弁護士
浜谷 千波	アドソル日進株式会社 執行役員 AI 研究所部長
村川 正宏	産業技術総合研究所 インテリジェントプラットフォーム研究部門 研究部門長
森 正弥	博報堂 DY ホールディングス 執行役員兼 CAIO

・ 補足事項

なお、今後、審査員を変更・追加する可能性があります。

3. 懸賞金等

● トライアル審査

➤ 懸賞金

審査にて高評価を得た応募者に対し、最低 500 万円を分配

✧ 審査の結果、評価点が必要な基準に到達していない場合、応募要件に即していないと判断された場合には、懸賞金は支払われません。

- ◇ トライアル審査において懸賞金を獲得した応募者は、原則本審査への応募を必須とします（本審査には、トライアル審査において応募した内容で応募することを原則必須とします）。注意いただく点として、トライアル審査において懸賞金を獲得した応募者が本審査に応募しなかった場合、トライアル審査において獲得した懸賞金は原則全額返金いただきます。

- 本審査

- 懸賞金

1 位：7,000 万円 2 位：5,000 万円 3 位：3,000 万円

※このほか、新規性や将来性において、高い評価を受けた応募者には、特別賞を付します。

審査の結果、評価点が必要な基準に到達していない場合、応募要件に即していないと判断された場合には、懸賞金は支払われません。

- 本審査におけるその他のインセンティブ

- ◇ 総合点 1～3 位、特別賞受賞者：表彰式への招待、 GENIAC のイベントでの PR 支援

- ◇ 全ての応募者※ 1：

- ・ 応募概要（提案書の一部）の特設サイト等での公表※ 2
- ・ 成果普及イベントへのご招待
- ・ 技術段階によって、応募者を国立研究機関・政府機関などと連携

※ 1）著しく応募内容が目的に合致しない場合は、公表/招待しないこともあり得ます。

※ 2）原則、応募いただいた内容の全てを、特設サイト等において公表します。特設サイト URL は NEDO の公募ページを参照ください。なお、応募者の事前の申し出により応募内容における公表の範囲を限定することは可能であるため、限定する範囲及びその理由についても併せて提案書に記載してください。

4. スケジュール

- 懸賞広告：2025 年 5 月 9 日
- 応募説明会：2025 年 5 月 26 日
 - 応募説明会の詳細は別添を参照してください。
- トライアル審査応募〆切：2025 年 7 月下旬
- トライアル審査結果発表：2025 年 9 月下旬
- 本審査応募〆切：2025 年 12 月下旬
- 表彰式（結果発表）：2026 年 3 月下旬
- 応募概要等の公表、成果普及イベント：2026 年 3 月下旬以降

参考

「本審査においては、応募者が開発した技術の実効性を、主催者側の審査員が実際に操作することで、もしくは、審査員の指示に従い応募者に操作していただき、確認ができる環境を応募者側が用意することを必須とします」という記載について、事務局の動作環境を参考のために記載しておきます。

- 動作環境の詳細（パターン１）
 - OS：Windows 11 Enterprise バージョン 22H2（x64）
 - プロセッサ：AMD Ryzen 5 5600U with Radeon Graphics（x64）
 - メモリ：16GB
 - ストレージ：25GB
 - 利用ブラウザ：Microsoft Edge for Business（公式ビルド）（64 ビット）

- 動作環境の詳細（パターン２）
 - OS：macOS Sequoia 15.3.2（x64）
 - プロセッサ：Apple M3 Max（x64）
 - メモリ：36GB
 - ストレージ：Machintosh HD 1TB
 - システムの種類：64 ビット オペレーティング システム、x64 ベース プロセッサ
 - 利用ブラウザ：Microsoft Edge for Business（公式ビルド）（arm64）