日本学術会議　第26期
記者会見（令和7年2月27日）
資料1-3

# 生成AIを受容・活用する社会の実現に向けて

黒橋　禎夫
日本学術会議 情報学委員会 幹事

2025年2月27日（木）

# 日本学術会議情報学委員会メンバー

| | | | |
|---|---|---|---|
| 委員長 | 下條　真司 | （第三部会員） | 青森大学ソフトウェア情報学部教授／大阪大学名誉教授 |
| 副委員長 | 高田　広章 | （第三部会員） | 名古屋大学未来社会創造機構教授 |
| 幹　事 | 黒橋　禎夫 | （第三部会員） | 大学共同利用機関法人情報・システム研究機構国立情報学研究所所長／京都大学大学院情報学研究科特定教授 |
| 幹　事 | 佐古　和恵 | （第三部会員） | 早稲田大学理工学術院教授 |
| | 浅川智恵子 | （第三部会員） | IBM Fellow／日本科学未来館館長／ Carnegie Mellon University IBM特別功労教授 |
| | 有村　博紀 | （第三部会員） | 北海道大学大学院情報科学研究院教授 |
| | 内田　誠一 | （第三部会員） | 九州大学理事／副学長 |
| | 大場みち子 | （第三部会員） | 京都橘大学工学部情報工学科教授 |
| | 田浦健次朗 | （第三部会員） | 東京大学執行役／副学長 |
| | 永井由佳里 | （第三部会員） | 北陸先端科学技術大学院大学理事／副学長 |

- 情報学委員会メンバーが中心となり、提言を審議
- 生成AIの現状と動向、脅威と課題、活用による波及効果について、学術の立場から深く洞察
- 提言のまとめに際し、生成AI研究開発に係る多分野（情報学、法学、教育学、科学技術政策など）の専門家や産業界の有識者が参加

# 生成AIの特徴

**包括性**
あらゆる学術分野、産業分野、社会全体に対する大きな影響

**革新性**
将来的には人間と共存する知的レベルとなる可能性

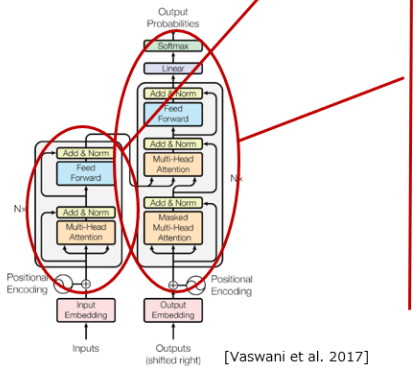
**加速性**
技術の進展や社会への影響が加速度的に進展

**予見の難しさ**
人類とAIの共存社会の姿を正確に予見することは極めて困難

- 大規模言語モデル（Large Language Model：LLM）を基盤とするChatGPTは、2022年11月に公開後２ヶ月でアクティブユーザー数が１億人に到達
- モデルサイズ（パラメータ数）と学習データ量に対して、対数スケールで性能が向上
- 国際的な研究開発競争が激化

2014 Attention
機械翻訳において目的言語の次の語を生成する際に原言語の文のどこに着目するか

2017 Transformer
attentionの精緻化、原言語文内、目的言語内でのattention



[Vaswani et al. 2017]

2018 BERT
Transformerのencoder側を単言語の分類問題等に

2018 GPT (117Mパラメータ)
Transformerのdecoder側を言語モデルに

2019 GPT-2 (1.5Bパラメータ)

2020 GPT-3 (175Bパラメータ)

2022 GPT-3.5 / InstructGPT

2022 ChatGPT

2023 GPT-4 (2Tパラメータ?)
画像も扱える、多言語能力も大幅向上
- 米司法試験で人間受験者の上位10%の成績
- 米大学入試テストSATで1600点中1410点
- 米医師試験USMLEでも合格レベルの点数

# 提言の背景（１）

## 生成AI技術の急速な発展・普及→教育・研究を含め社会全体に大きな影響

- 教育機関における生成AI利用ガイドラインの整備など

## 生成AI活用のリスク

- ハルシネーション（事実と異なる情報を出力すること）
- 高品質な生成メディアによる詐欺や世論操作、非社会的な回答の生成、機密情報漏洩
- 著作権侵害、名誉棄損
- 芸術的活動への脅威、社会の価値観や文化への影響

## 生成AI活用の利点

- 科学技術：仮説生成、仮説検証、論文による知識流通等の科学技術での利用、分野横断的な新たな知の創造
- 産業界：業務効率化・業務量削減、労働力不足や長時間労働などの解消
- 教育：効率化や高度化、
- その他：文章の推敲・要約・翻訳、挨拶文やメールの作成支援、旅行計画の提案、投資のアドバイス、音楽・アート・デザインの生成

# 提言の背景（２）

**国際動向の活発化**

- 2023.12　G7 広島AIプロセス：「広島AIプロセス包括的政策枠組み」が承認
- 2024.2　AIセーフティ・インスティテュート（AISI）が設立
- 2024.4　Gサイエンス学術会議2024：「人工知能と社会」についての共同声明が公表
- 2024.5　欧州連合：AI法（The Artificial Intelligence Act：AI Act）が採択
- 2024.7　サイエンス20：人工知能の倫理や社会的影響の共同声明が公表

**我が国の取るべき施策（提言のキーポイント）**

- リスク対策に対する十分な工夫
- 生成AIの研究開発や社会での積極的な活用
- 人類とAIの共存社会のデザインにおいて世界をリード
- リスクと活用の二つの側面の調和

# 提言の概要


提言１
生成AI研究開発の望ましい体制
①生成AIの技術開発を国家戦略として位置づける
予算の確保　研究体制の整備
学術間や産学官の連携　オープンな研究開発の支援
②生成AIの研究基盤の強化と国際的研究連携の推進
研究資源の拡充　データインフラの構築
データ共有　国際的な知見共有や人材交流の加速化
③生成AI開発における透明性の確保とAIガバナンスへの包括的な取り組み
ガイドラインの作成　問題を回避するための技術開発
国際的ルールメイキングへの関与
提言２
生成AIモデルの適切な運用
①生成AIに対する攻撃を検知・回避する頑健なシステム構築
ベンチマーク作成・評価　改竄されていないことを証明する仕組み
②AI利用のリスク最小化と迅速に問題に対処する体制の整備
技術的特性や倫理的影響　AI技術の標準化
③人間中心の原則に基づく持続可能な社会の実現に向けたAI利活用の継続的議論
人間中心の原則　インセンティブ設計や規制設計
【提言】
生成AIを
受容・活用する社会の
実現に向けて
SCIENCE COUNCIL OF JAPAN
日本学術会議
提言４
生成AIモデル以降の教育とリテラシー
①AIとの共存・共生のための社会変革に対応する人材育成
リスキング　教育プログラムの推進　地域格差の解消
②AIとの共存を目指した新たな教育への転換
AIの批判的利用　課題解決　創造
ELSI　先進事例の共有
③AIの学際性を活用するための学術分野間および産学間の対話・連携の促進
広いステークホルダー間の議論
提言３
責任ある生成AI実装に向けた制度設計
①アジャイルかつマルチステークホルダー型のガバナンスの志向
最適な価値のバランス　官民の役割の変化
②政府の役割：オープンなルール形成・情報共有の促進、制裁に関する新たな制度設計
枠組みやツールの提供　法令の解釈に関するガイダンス　ベストプラクティスの共有
③民間主体の役割：主体的なリスク評価とAIベネフィットの最大化、ガバナンスの恒常的な改善
十分な質と量の情報開示　アカウンタビリティ　ステークホルダーからのフィードバック

# 提言１：生成AI研究開発の望ましい体制

## ①生成AIの技術開発を国家戦略として位置づける

日本の技術競争力を強化するため、国家戦略として生成AIの研究開発を推進すべきである。特に、**オープンな研究開発の取り組みへの支援**を重視・強化することが必要である。

## ②生成AIの研究基盤の強化と国際的研究連携の推進

日本国内の**生成AI研究者コミュニティの強化**と**国際的研究連携の推進**が必要である。プライバシーやセキュリティに配慮した**データインフラの構築を支援**するとともに、**公共データの開放や産業界とのデータ共有プロジェクト**を奨励すべきである。

## ③生成AI開発における透明性の確保とAIガバナンスへの包括的な取り組み

生成AIによる判断や行動が、人間の価値観や倫理観に合うことが極めて重要であり、学習データや学習手法を含む**開発プロセスの透明性を確保**すること、**AIの設計・開発・評価においてガイドラインを作成**してリスクを最小化すること、**AIガバナンスの国際的ルールメイキングに日本の考え方を反映させる体制を構築**することが必要である。

# 提言２：生成AIモデルの適切な運用

## ①生成AIに対する攻撃を検知・回避する頑健なシステム構築

生成AIモデルが**サイバー攻撃や物理的攻撃から適切に保護**される必要があり、これらの**攻撃を検知・回避する頑健なシステムが構築**されるべきである。

## ②AI利用のリスク最小化と迅速に問題に対処する体制の整備

AI技術に起因する問題が発生した場合に、**迅速かつ適切に対処できる体制**を整えることが必要である。また、**国際的な協力を通じて、AI技術の標準化やベストプラクティスを共有**し、グローバルな視点でのAIの発展と運用を推進することが重要である。

## ③人間中心の原則に基づく持続可能な社会の実現に向けたAI利活用の継続的議論

人間中心の原則に基づく持続可能な社会の実現に向けて、**市場原理や競争原理に任せるのではなく、地球規模の課題や社会・経済にとって最重要な問題**に対してAIの利活用・運用を優先すべきである。

# 提言３：責任ある生成AI実装に向けた制度設計

**①アジャイルかつマルチステークホルダー型のガバナンスの志向**

従来型の規制モデルでは、複雑で変化の速いAIがもたらす様々なリスクに適切に対処することができない。**アジャイル（迅速かつ反復的）でマルチステークホルダー型のガバナンスを志向**すべきである。

**②政府の役割：オープンなルール形成・情報共有の促進、制裁に関する新たな制度設計**

政府は、**オープンな場でのルール形成の促進、事故調査への関係者の積極的な協力を促す制度設計、事故被害者に対する迅速な救済制度の設計**などを行う必要がある。

**③民間主体の役割：主体的なリスク評価とAIベネフィットの最大化、ガバナンスの恒常的な改善**

民間主体は、**政府を含むステークホルダーに十分な質と量の情報開示**を行うとともに、ステークホルダーからのフィードバックを得て、**常にガバナンスのあり方を改善**する必要がある。

# 提言４：生成AIモデル以降の教育とリテラシー

## ①AIとの共存・共生のための社会変革に対応する人材育成

社会全体で**生成AIの教育やリスキリング**に取り組んでいくことが必要であり、それを推進するための**リテラシーを持つ人材の養成と教育プログラムの推進、リスキリング支援**があまねく必要である。この際、地域格差に配慮し、むしろ**地域格差を解消**することを目指すべきである。

## ②AIとの共存を目指した新たな教育への転換

慎重な議論を行った上で、AIの活用を前提として**AIとの共存を目指した新たな教育への転換**を図るべきである。従来の知識の伝達に偏重するのではなく、**AIを批判的に利用し、課題を解決し、創造する能力を高める教育・カリキュラム**が必要である。また、**新たな教育について情報共有・議論する場**を国として支援することも重要である。

## ③AIの学際性を活用するための学術分野間および産学間の対話・連携の促進

AIの活用は、学術を学際的に深め、複合的な社会課題の解決につながるが、そのためには、**科学者が高いAIリテラシー**を身につけることが必要であり、**学術分野間および産学間の対話・連携の促進**が必要である。