



จดหมายข่าวความปลอดภัยของข้อมูล



Vol2 Generative AI

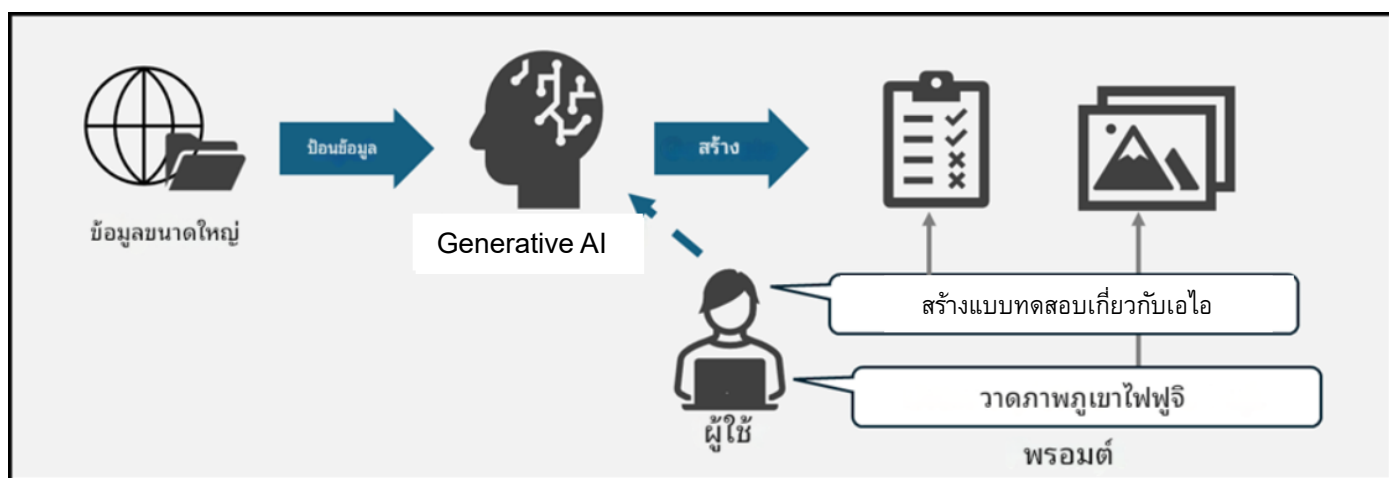
ปัญญาประดิษฐ์ที่ใช้การเรียนรู้เชิงลึก (Generative AI) คืออะไร?

Generative AI หมายถึง เทคโนโลยีที่สามารถสร้างเนื้อหาโดยอัตโนมัติ เช่น ข้อความ รูปภาพ และเสียง

กลไกการทำงานของ Generative AI อาศัยข้อมูลจำนวนมากมหาศาลที่ได้จากการเรียนรู้ของเครื่อง (Machine Learning) ในอดีต เพื่อนำมาสร้างเนื้อหาตามข้อมูลที่ผู้ใช้ป้อนเข้าไป

ประเภทหลักของ Generative AI ได้แก่ ปัญญาประดิษฐ์สำหรับสร้างข้อความ ปัญญาประดิษฐ์สำหรับสร้างภาพ และปัญญาประดิษฐ์สำหรับสร้างเสียง

กลไกการทำงานของ Generative AI



ความเสี่ยงจากการใช้ Generative AI

แม้ว่า Generative AI จะมีความสะดวกสบายในการใช้งาน แต่ก็มีความเสี่ยงบางประการเช่นกัน ที่เราขอแนะนำความเสี่ยงที่พบได้บ่อย

การละเมิดข้อมูล

มีความเสี่ยงที่ข้อมูลที่ป้อนเข้าระบบ generative AI อาจรั่วไหลออก
สู่ภายนอก ข้อมูลอินพุตอาจถูกนำไปใช้ในการฝึกอบรมระบบ generative AI และปรากฏในการตอบกลับผู้ใช้รายอื่น ข้อมูลอินพุตอาจถูกเก็บไว้ในบันทึกของผู้ให้บริการเพื่อป้องกันการใช้งานในทางที่ผิด อย่างไรก็ตาม หากเกิดการกระทำที่มุ่งร้ายภายในองค์กรของผู้ให้บริการ หรือหากการโจมตีจากภายนอกประสบความสำเร็จ ข้อมูลดังกล่าวอาจรั่วไหลได้



อาการประสาทหลอน

บางครั้งการสังเคราะห์แสง (Generative AI) อาจสร้างข้อมูลที่แตกต่างจากข้อเท็จจริง ปรากฏการณ์นี้เรียกว่า "ภาพหลอน" หากผลลัพธ์มีข้อมูลที่ผิดพลาดหรืออคติ การเชื่อถือข้อมูลดังกล่าวอาจนำไปสู่การตัดสินใจหรือการกระทำที่ไม่ถูกต้อง



การละเมิดทรัพย์สินทางปัญญา

มีความเสี่ยงที่เนื้อหาที่สร้างโดย AI อาจละเมิดลิขสิทธิ์หรือเครื่องหมายการค้าของผู้อื่น ตัวอย่างเช่น หากรูปภาพหรือข้อความที่สร้างขึ้นมีความคล้ายคลึงกับงานที่มีอยู่เดิมมาก อาจนำไปสู่ปัญหาทางกฎหมายได้ แม้ว่าบริการสร้างงานของ AI จะถูกระบุว่า "อนุญาตให้ใช้ในเชิงพาณิชย์" ก็ตาม แต่ก็ไม่ได้รับประกันว่าผลลัพธ์จะปราศจากการละเมิดลิขสิทธิ์ของผู้อื่น



ปลอมลึก

อัลทำให้สามารถสร้างเสียงและวิดีโอปลอมที่มีความแม่นยำสูงจนแทบแยกไม่ออกว่าเป็นของจริง ส่งผลให้เกิดปัญหาต่างๆ เช่น การแพร่กระจายข้อมูลเท็จและการฉ้อโกงที่ทำลายชื่อเสียงของบุคคลหรือบริษัท



ความไม่แน่นอนด้านกฎระเบียบ

ปัจจุบัน กฎระเบียบของรัฐบาลและองค์กรต่างๆ ยังตามไม่ทันการพัฒนาและการใช้ AI เชิงกำเริบ หากมีความไม่แน่นอนใดๆ โปรดพิจารณาว่าสิ่งเหล่านั้นอาจส่งผลกระทบต่อความไว้วางใจของลูกค้าหรือนำไปสู่ปัญหาทางกฎหมายอย่างไร และต้องแน่ใจว่าได้รับการอนุมัติอย่างถูกต้องก่อนใช้งาน



มาตรการสำหรับผู้ใช้งานรายบุคคล

เพื่อให้สามารถใช้ Generative AI ได้อย่างปลอดภัย ผู้ใช้งานแต่ละคนควรคำนึงถึงประเด็นต่อไปนี้



อย่าใส่ข้อมูลส่วนตัวหรือข้อมูลลับอื่นๆ ลงในระบบ Generative AI



ตรวจสอบให้แน่ใจว่าเนื้อหาที่สร้างขึ้นจะไม่ละเมิดลิขสิทธิ์ เครื่องหมายการค้า สิทธิการออกแบบ หรือสิทธิในทรัพย์สินทางปัญญาอื่น ๆ ของผู้อื่น



ควรตรวจสอบข้อมูลที่ได้รับจาก generative AI กับแหล่งข้อมูลที่เชื่อถือได้อื่นๆ เสมอ



ตรวจสอบให้แน่ใจว่าเนื้อหาที่สร้างขึ้นไม่มีข้อมูลส่วนบุคคลที่เป็นเท็จ คำพูดที่หมิ่นประมาท หรือการแสดงออกที่เลือกปฏิบัติ

【ข้อมูลติดต่อสอบถาม】

ฝ่ายเลขานุการคณะกรรมการความปลอดภัยสารสนเทศ (กองเทคโนโลยีสารสนเทศ แผนกระบบสารสนเทศ ส่วนที่ 2)

อีเมล securitylearning@fanuc.co.jp



Vol2 Generative AI

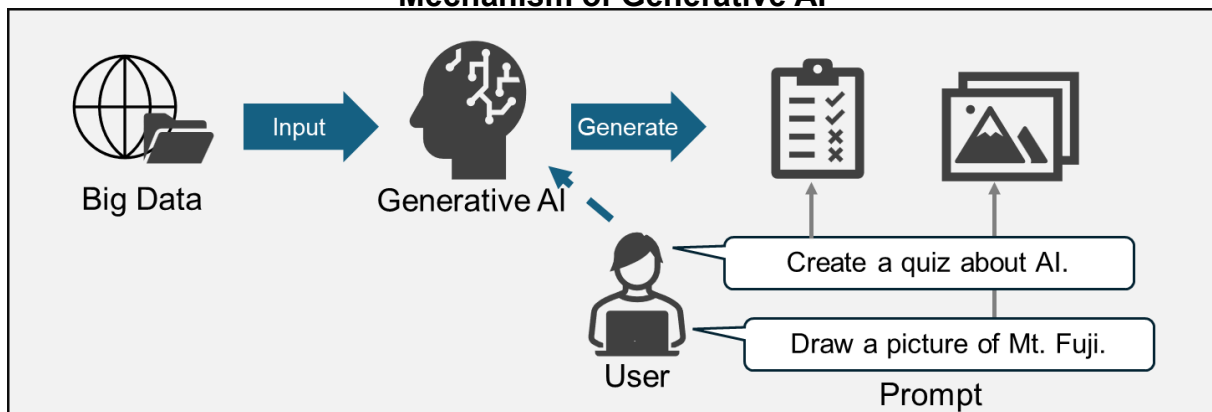
What is Generative AI?

Generative AI refers to technology that automatically creates content such as text, images, and audio.

Its mechanism involves using vast amounts of data acquired through previous machine learning to generate content based on user input.

Major types of generative AI include text generation AI, image generation AI, and audio generation AI.

Mechanism of Generative AI

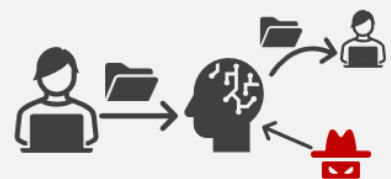


Risks of Using Generative AI

Generative AI is convenient, but it also comes with certain risks. Here, we will introduce some of the most common ones.

Data breach

There is a risk that **data entered generative AI may be leaked externally**. Input data may be used for AI training and appear in responses to other users. Input data may also be stored in service providers' logs to prevent misuse. However, **if a malicious act occurs within the provider's organization or if an external attack is successful**, the information could be leaked.



Hallucination

Generative AI may sometimes **produce information that differs from facts**. This phenomenon is known as "hallucination." If the output contains misinformation or bias, trusting it could lead to incorrect decisions or actions.



Intellectual Property Infringement

There is a risk that content generated by AI may infringe on others' copyrights or trademark rights. For example, if a generated image or text closely resembles an existing work, it could lead to legal issues. Even if a generative AI service is labeled as "commercial use allowed," **it does not guarantee that the output will be free from infringement of others' rights.**



Deep fake

AI has made it possible to create fake audio and video with such high precision that they are indistinguishable from real ones. As a result, issues such as the spread of misinformation and fraud that damage the reputation of individuals or companies have emerged.



Regulatory uncertainty

Currently, **regulations by governments and organizations have not yet caught up** with the development and use of generative AI. "If there are any uncertainties, **consider how they may affect client trust or lead to legal issues**, and be sure to obtain proper approval before use.



Individual Measures

To use generative AI safely, each person should keep the following points in mind.



Do not enter personal or other highly confidential information into generative AI.



Ensure that the generated content does not infringe on others' copyrights, trademarks, design rights, or other intellectual property rights.



Always verify information provided by generative AI with other reliable sources.



Ensure that the generated content does not contain false personal information, defamatory statements, or discriminatory expressions.

【Inquiry】

Information Security Committee Secretariat
(IT Division information System Department Section 2)
securitylearning@fanuc.co.jp



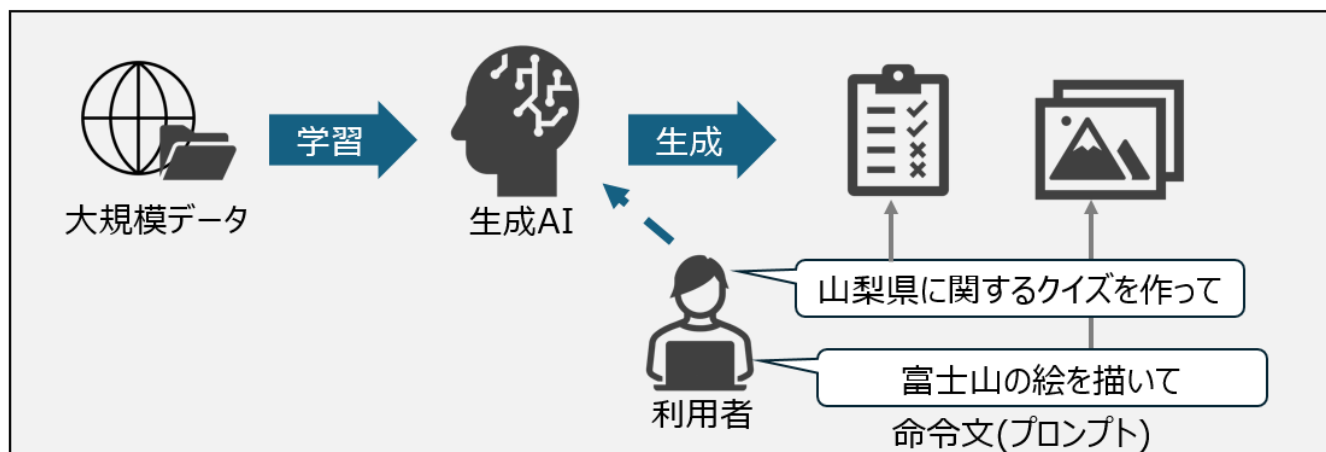
第2回 生成 AI

生成 AI とは？

生成 AI とは、テキスト、画像、音声などのコンテンツを自動的に生成する技術です。

仕組みとしては、AI が過去の機械学習で取り入れた膨大なデータを基に、利用者が入力した情報に応じてコンテンツを生成する技術・サービスを指します。

代表的な生成 AI の種類には、テキスト生成 AI をはじめ、画像生成 AI、音声生成 AI などがあります。



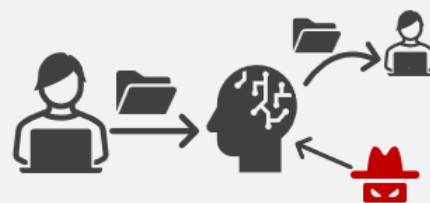
生成 AI のしくみ

生成 AI の利用に伴うリスク

生成 AI は便利ですが、使用する上でのリスクもあります。ここではその代表的なものをご紹介します。

情報漏えい

生成AIに入力したデータが外部に漏れるリスクがあります。入力した情報がAIに学習され、その情報は他の人がAIに対して質問をする際に回答として使用される可能性があります。また、生成AIサービス事業者のログに情報が残ることもリスクとなりえます。事業者は生成AIの悪用を防ぐために、入力内容を保存しています。しかし、**事業者内で不正行為が行われた場合**や、**外部からの攻撃**によって情報が漏れる可能性があるのです。



誤情報（ハルシネーション）

生成AIは時として、**実際の事実とは異なる情報を生成する**ことがあります。この現象を「ハルシネーション」と呼びます。誤情報や偏見が含まれると、それを信じてしまうことで誤った判断や行動を引き起こす可能性があります。



権利侵害

生成AIが出力した他者の著作権や商標権を侵害するリスクがあります。例えば、生成された画像やテキストが既存の著作物に似ている場合、法的な問題が発生する可能性があります。生成AIサービスが「商用利用可」とされていても、**出力される内容が他者の権利を侵害しないことを保証しているわけではありません。**



ディープフェイク

AIにより、本物と区別がつかないほどの精度で偽の音声や映像を作成することが可能となりました。これにより、個人や企業の信用を損なうような偽情報や詐欺の拡散などの問題が生じています。



規制の不確実性

現在、生成AIの開発や利用に関して、**政府や組織による規制が追いついていません。**不明確な点がある場合は、**取引先の信頼や法的問題**にどのように関与するか検討して、必ず許可を得た上で利用してください。



個人の注意点

生成 AI を安全に利用するため、個人個人が以下のことに注意しましょう。



個人情報など機密性の高い情報を生成AIに入力しない



生成されたコンテンツが他者の著作権、商標権、意匠権などを侵害していないか確認する



生成AIが提供する情報は、必ず他の信頼できるソースで確認する



生成物に虚偽の個人情報や名誉棄損表現、差別的表現が含まれていないか確認する

【お問い合わせ先】

IT 本部 情報システム部 二課

情報セキュリティ委員会事務局

[情報セキュリティ委員会事務局へのお問合せ](#)