

AI PICASSO

AI時代の 知的財産権検討会

2023年10月18日

AI Picasso株式会社 代表取締役 富平準喜

自己紹介

富平 準喜

AI Picasso株式会社 代表取締役

1997年生まれ、国立東京工業高等専門学校情報工学科卒業後、筑波大学情報学群知識情報図書館学類に編入学、その後同大学院に進学。シンガポール国立大学の研究室に短期留学。深層学習の研究に従事。

全国高専プログラミングコンテストにて最優秀賞、文部科学大臣賞を受賞、ものづくり大賞内閣総理大臣賞を受賞。

2020年8月CommentScreen株式会社 代表取締役就任。

2021年1月 株式会社AldeaLab 代表取締役就任。

2022年10月 AI Picasso株式会社 代表取締役就任。



会社紹介

会社名

AI Picasso株式会社

設立年月日

2022年10月



AI領域に特化した筑波大学発ベンチャー 内閣総理大臣賞を受賞メンバーで創業

プログラミングコンテストで優勝し、文部科学省賞、内閣総理大臣賞を受賞しました。受賞メンバー同士で、AIの研究室のメンバーで創業したため、技術力のあるメンバーが中心です。



文部科学省

地域イノベーション・エコシステム形成プログラム
基盤構築プロジェクト採択



TSUKUBA QUALITY

2022年度つくばクオリティ認定

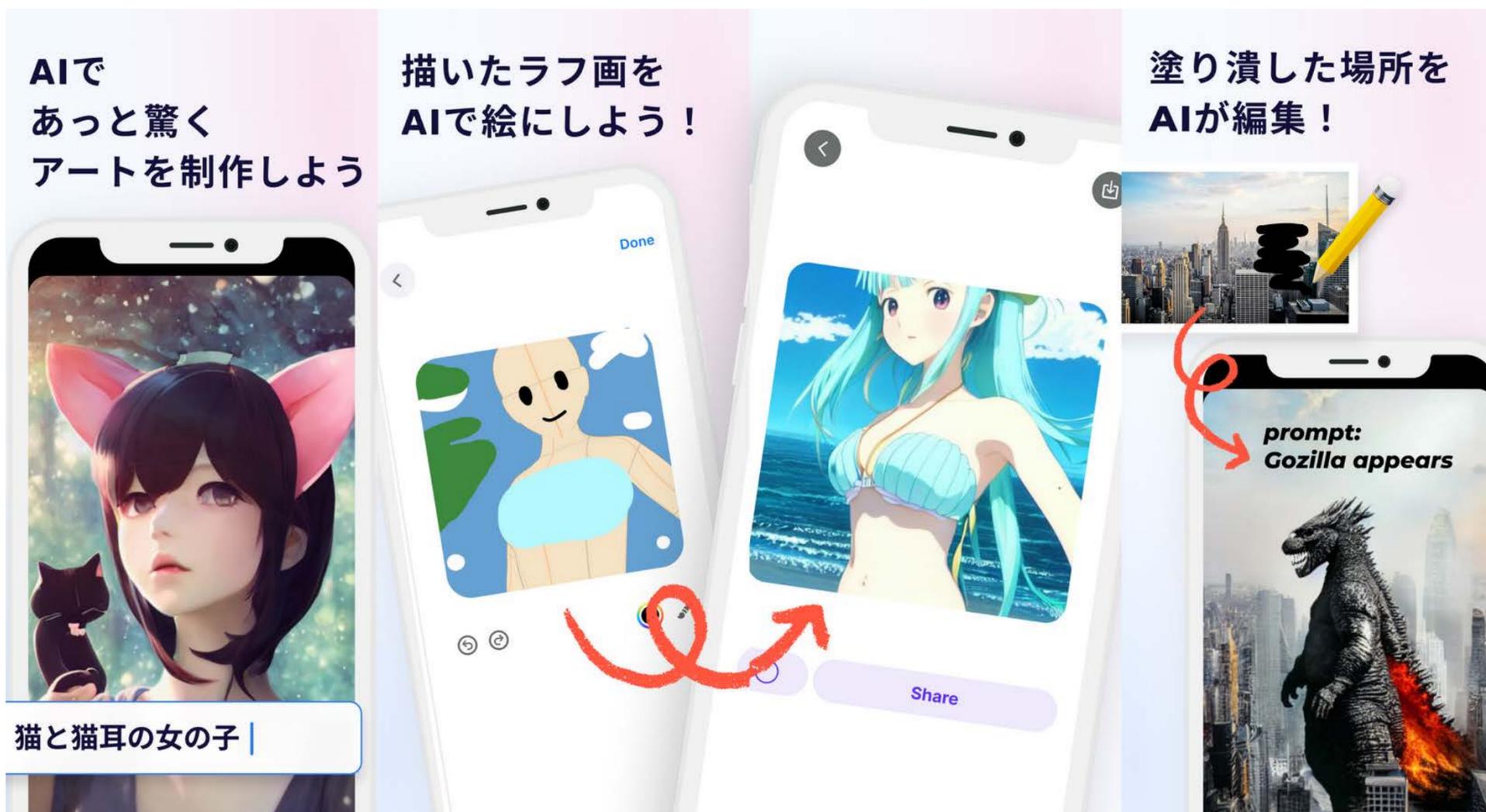
AI PICASSO

事業について

Our Business



AI Picasso



Download
85万DL

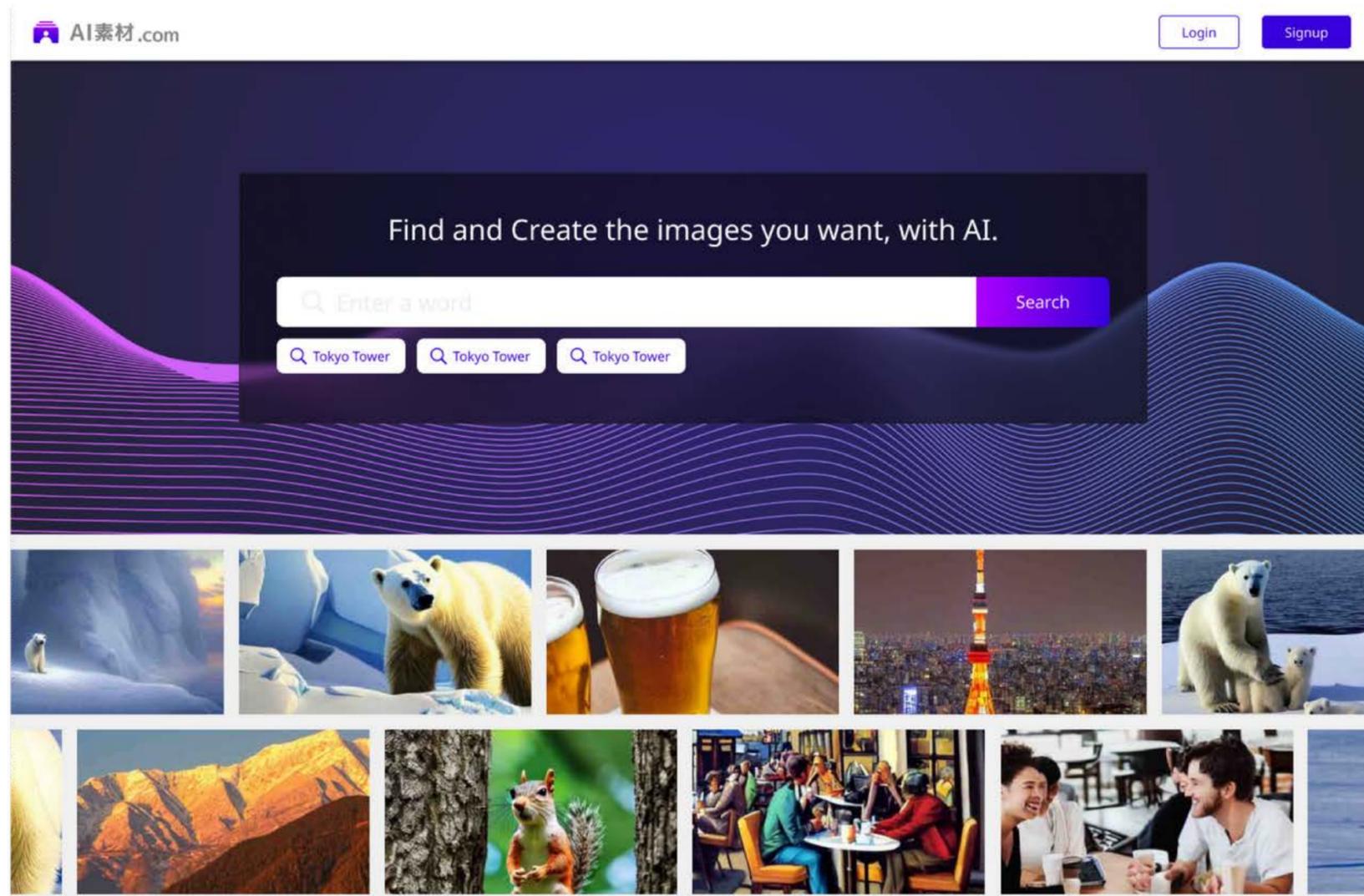
たくさんのメディアで紹介される

07L Abema^{TV}

NHK TBS
TELEVISION

AIで文章から画像を生成できる日本初の画像生成アプリ、画像生成AIモバイルアプリにてユーザ数国内No.1

AI素材.com



登録ユーザ
5万人

生成画像数
53万件

AIで生成した画像に特化した素材サイト

AIいらすとや

いらすとやは、かわいいイラストが沢山見つかるフリー素材サイトで、企業・学校・自治体など様々な場面で使われています。

いらすとやの作者のみふねたかし様と連携して、いらすとや風の画像を生成したり、ダウンロードすることができるWebプラットフォームが、「AIいらすとや」です。

AIによって生成されたいらすとや風画像



歩く人工知能



馬に乗る宇宙飛行士



ケーキを食べる女の子

The screenshot shows the AI Irasutoya website interface. At the top, there is a logo for 'AIいらすとや' with the URL 'いらすとや x AI素材.com'. Below the logo, there are two steps: 'STEP1 テキストで入力' (Input text) and 'STEP2 AIで生成' (Generate with AI). To the right, there is an illustration of a person painting a rabbit on an easel, with a thought bubble containing a girl's face and another containing a rabbit. Below this, there is a section titled 'AI画像生成' (AI Image Generation) with a text input field containing the prompt '着物を着たうさぎ、正面を見ている。' (A rabbit wearing a kimono, looking forward). A button labeled '生成する 980円/月~' (Generate 980 yen/month~) is positioned below the input field. At the bottom, there is a section titled 'イラストを検索 無料' (Search for illustrations Free) and a note: 'すでに生成されている画像は、無料でお使いいただけます。' (Images already generated are available for free use).

AI PICASSO

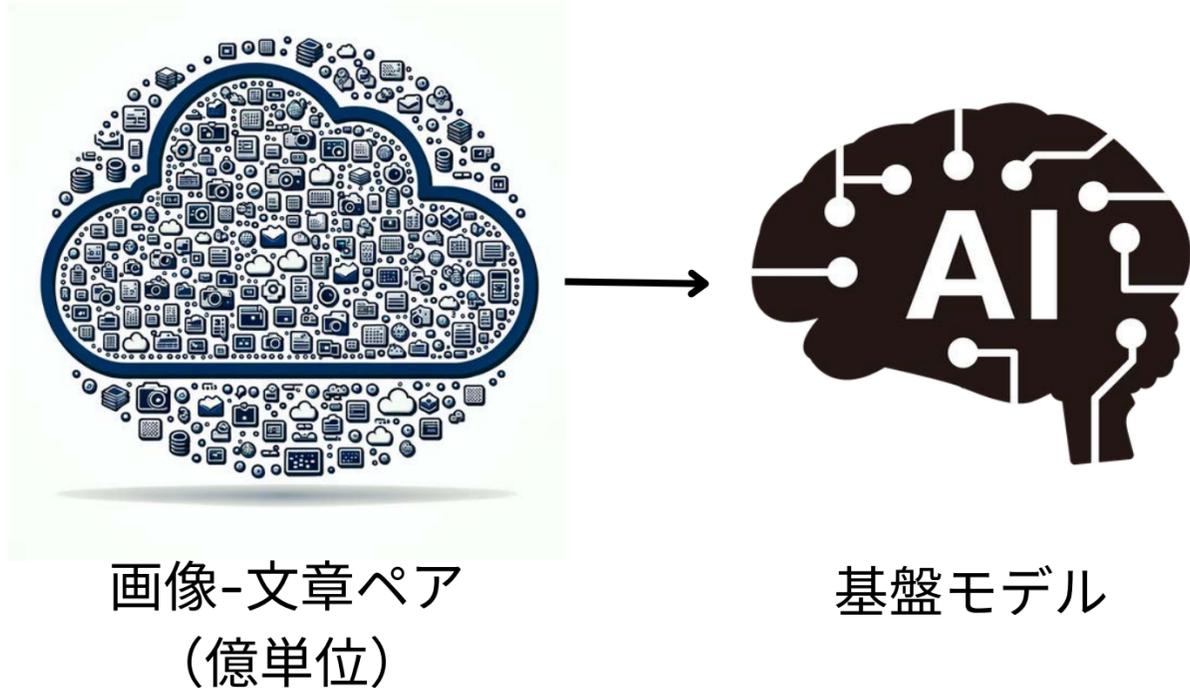
AIいらすとやの
学習と生成
について



参考: 画像生成AIの学習方法

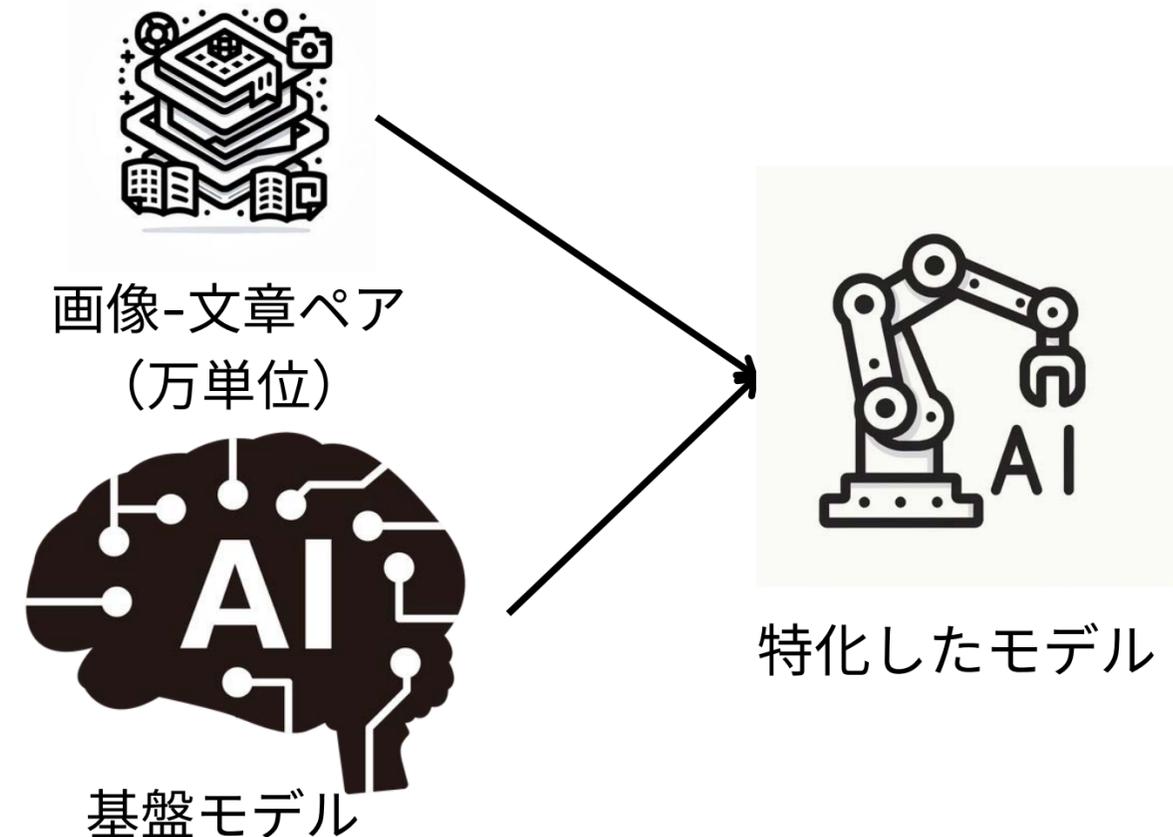
画像生成AIの学習には基盤モデルを作る事前学習と、基盤モデルを特化させる追加学習の二段階があります。

事前学習



基盤モデルの例: Stable Diffusion、Adobe Firefly

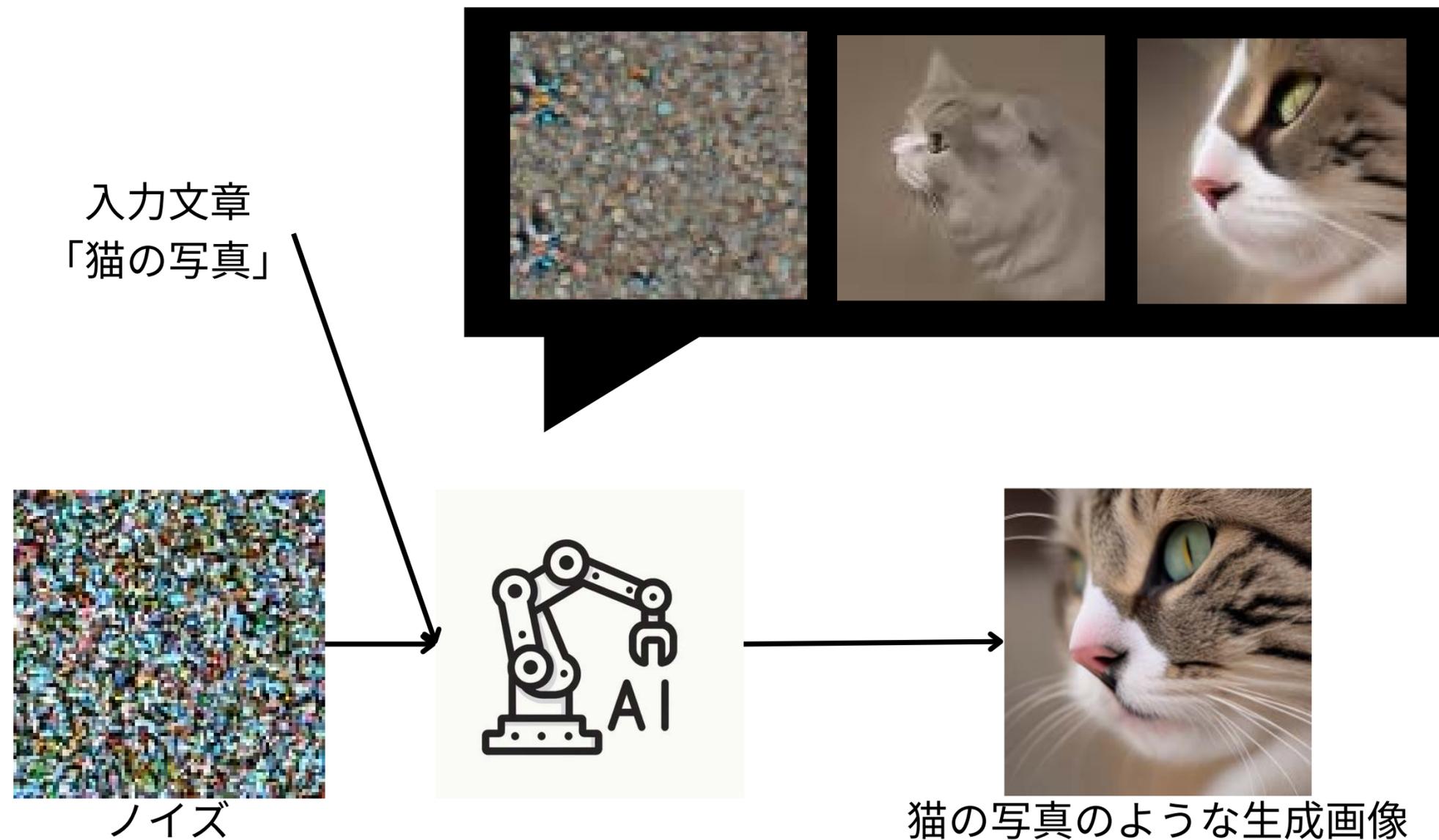
追加学習



追加学習した例: Emi (Stable Diffusionベース)

参考: 画像生成AIの生成方法

- Stable Diffusionなどの拡散モデルの場合、ノイズからノイズを取り除くことで画像を生成しています。
- 画像を切り貼りされているとよく勘違いされますが、そうではありません

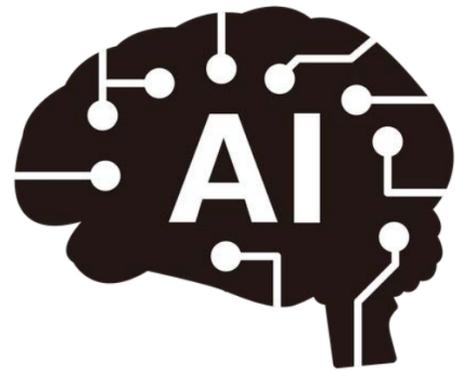


AIイラストの学習と生成

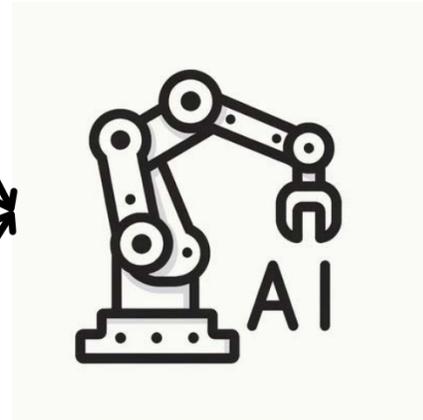
独自の追加学習



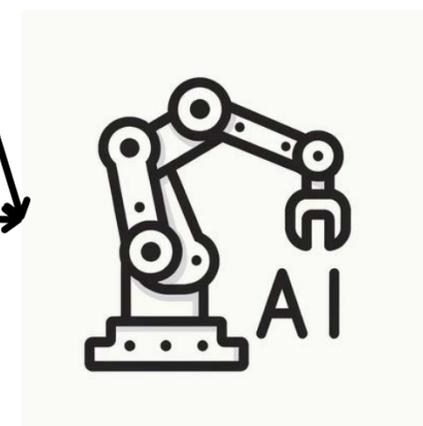
いらすとやの
画像-文章ペア
(万単位)



当社の基盤モデル



入力文章
「立っている灰色の猫」



生成

AI PICASSO

クリエイターへの 収益還元方法に ついて



収益還元方法

AI PICASSO



いらすとやとはレベニューシェアのやり方で、ユーザからの課金金額を分配しています。

収益還元方法

AI PICASSO



今後はいらすとやだけではなく、他のイラストレーターともコラボレーションして、複数人のイラストレーターに対して、分配していく予定です。

AI PICASSO

知的財産権に 関わる技術等につ いて



画像生成AIに対する学習への同意

学習前のオプトアウトが多く、厳密なオプトインである画像生成AIはほぼありません。
学習後にもオプトアウトできることが研究によって示唆されています。

原則オプトイン方式

Adobe Fireflyは原則として著作権者から許諾を得た画像を使用して学習しています。

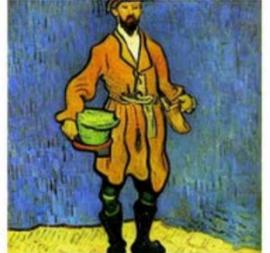
オプトアウト方式

Stable Diffusionは第3版から学習前に学習対象の画像を著作権者の申請で除くことにしています。
Stable Diffusionは学習後に概念を消せるということが示されています。[Gandikota et al. 2023]



Erasing Artistic Style

Original Model



Edited Model



Erased from model:
"Van Gogh"

学習を拒否する方法

現在はソフトロー (robot.txt) により自動収集プログラム (クローラ) を拒否する方法が一般的です。自動収集されてしまっても学習できないように妨害する技術も研究されています。

導入されている方法

- robot.txt に自動収集していい範囲をコンテンツ提供側が明示する。
 - OpenAIにはGPTBotというクローラに対して認可情報をrobot.txtで通知します。
 - GoogleにはGoogle-Extendedというクローラに対して認可情報をrobot.txtで通知します。

検討されている方法

- 収集されそうな画像に特殊な画像処理を施すことで学習を妨害することができます。
 - Glazeは画像生成AIを利用して学習を妨害するノイズを画像に付与します。[Shan et al., 2023]
 - 他にもMist [Liang et al., 2023] などがあります。

生成された画像を識別する仕組み

生成された画像を識別する技術は民間が主体となって標準化を進めているように見られます。

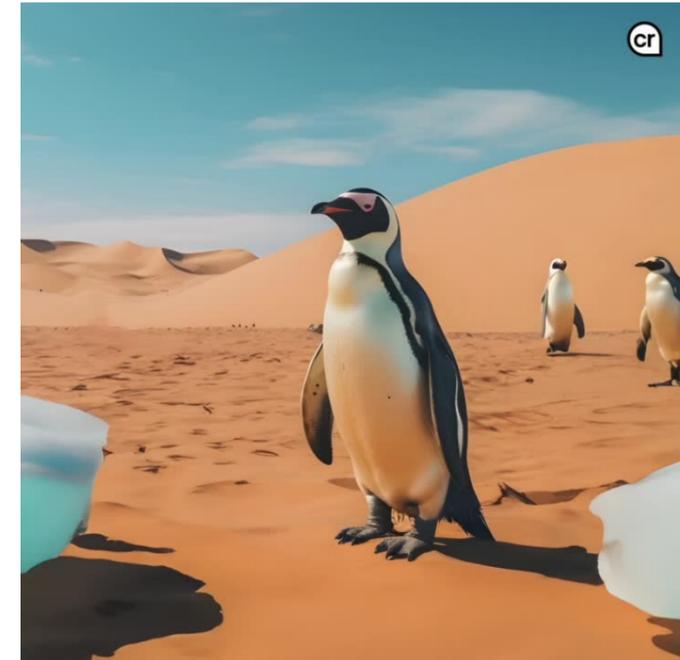
提案されている方法

- Stable Diffusionの開発元は電子透かしによる識別方法を使っています。
- AdobeやMicrosoftなどが参加するC2PA*1はContent Credentialsを提案しています。
 - Content Credentialsでは電子透かしに加え、コンテンツの由来まで考慮します。
 - Content CredentialsはAdobe Fireflyを使うサービスにて使用されています。
- NTTやNHKなどが参加するOP CIP*2はOriginator Profileを提案しています。
 - Originator Profile (OP) ではコンテンツの由来を検証可能な形で付与します。
 - OPの仕様策定と試験実装を進め、Web標準化と社会実装を目指す取り組みを進めています。

*1 Coalition for Content Provenance and Authenticity

*2 Originator Profile 技術研究組合

<https://contentcredentials.org/>
より引用



フィルタリングの方法

画像間の類似度を評価する方法と画像を検索する方法は数多く提案されています。
知的財産権を侵害する画像を出力しない方法は生成前と生成後で提案されています。

画像間の類似度

- 画素同士が似ているMSE、深層距離学習による類似度などがあります。

画像検索方法

- 厳密な手法として最近傍探索があり、高速な手法としてLSHなどがあります。

知財権を侵害しない方法

- 生成前
 - 知財権を侵害しうる入力文章に対する生成を拒否します
- 生成後
 - 知財権を侵害しうる可能性を画像認識で推定し、ユーザへの提供を拒否します。

類似度が高い画像を生成する確率

ランダムに生成したとき、類似度が0.5以上の画像をStable Diffusionが生成する確率は1.88%です。[Somepalli et al., 2023]
しかし、Somepalliらが定義した類似度であって、著作権侵害の要件である類似性を満たすとは限りません。
この現象が拡散モデルで発生しないようにする方法が現在研究されています。[Somepalli et al., 2023]



Figure 1. *Stable Diffusion* is capable of reproducing training data, creating images by piecing together foreground and background objects that it has memorized. Furthermore, the system sometimes exhibits *reconstructive* memory, in which recalled objects are semantically equivalent to their source object without being pixel-wise identical. Here, we show this behavior occurring with a range of prompts sampled from LAION, and with a hand-crafted prompt (rightmost pair). The presence of such images raises questions about the nature of data memorization and the ownership of diffusion images. Top row: generated images. Bottom row: closest matches in the LAION-Aesthetics v2 6+ set. Sometimes source and match prompts are quite similar, and sometimes they are quite different. See Fig. 6 for more examples with prompts, or the Appendix for prompts from this figure.

生成画像と学習画像との関係

生成した画像から学習した画像を推定する研究が行われています。
生成過程を見ることで生成に寄与した学習画像を推定する研究があります。[Georgiev et al., 2023]
還元方法を考える一つの見方になるかもしれません。

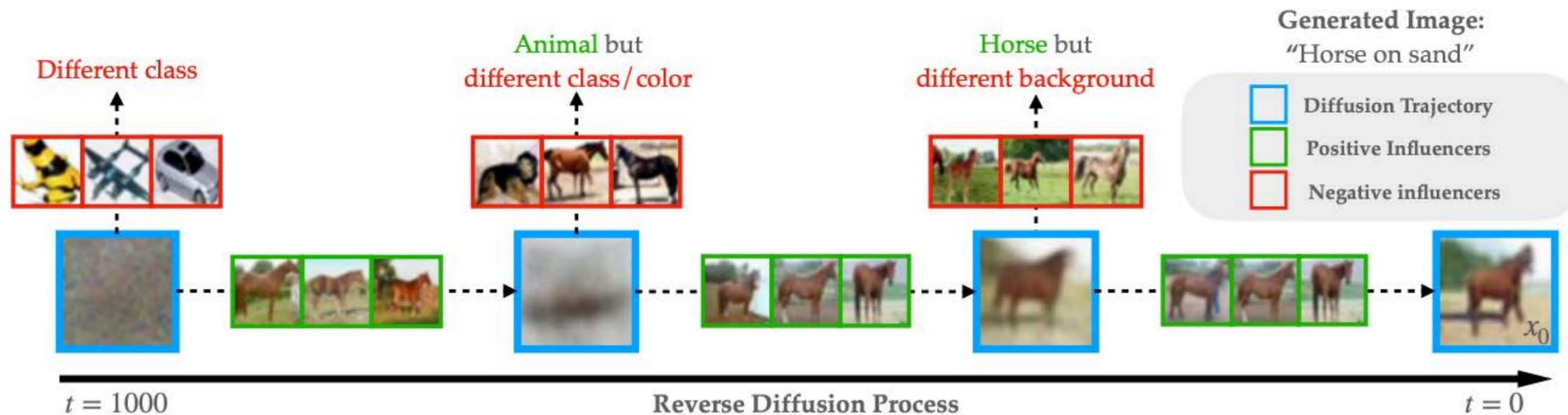


Figure 1. Overview of our attribution method. For a given synthesized image x_0 , we apply our attribution method at individual steps (denoted by t) along the sampling trajectory. At each step, our method pinpoints the training examples with the highest influence (positive in green, negative in red) on the generative process. In particular, training examples with positive attribution scores guide the trajectory towards x_0 ; in contrast, ones with negative scores guide the trajectory away from x_0 . For more examples, see Appendix G.

技術監修

尾崎 安範

AI Picasso株式会社 AIエンジニア

1989年生まれ、国立豊田工業高等専門学校情報工学科卒業後、名古屋大学現情報学部に編入学、その後東京大学大学院情報理工学系研究科修士課程を修了。AIいらすとやなどの画像生成AIの開発に注力。また、ロボットビジョンとRLHFを用いたロボットの研究に従事。

2014年4月に日本電信電話株式会社 (NTT研) 入社。

2019年4月より民間企業研究部門にて研究員。

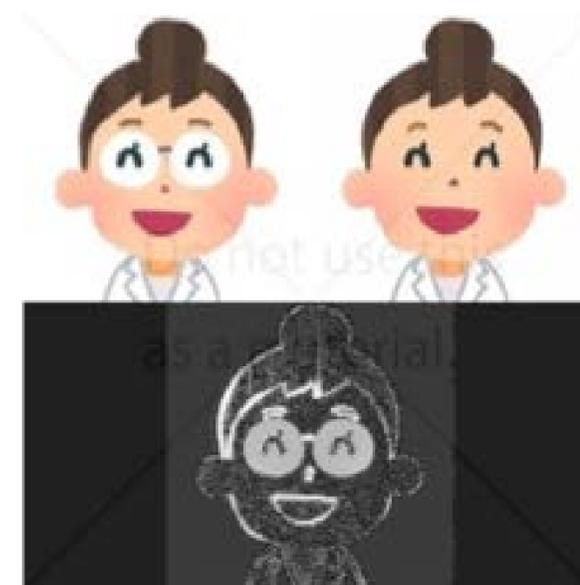
2019年4月より国立大学にて客員研究員。

2023年10月よりAI Picasso株式会社にてAIエンジニア。



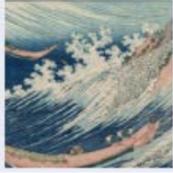
AIいらすとやができた背景

1. Googleが画像生成AI、Imagen [Saharia et al., 2022] を発表し、当時話題になった
2. 尾崎が主体となって行動したところが次の通り
 - a. 画像生成AIを個人で学習したら話題になるのでは？と思いつく
 - b. いらすとやの顔アイコンを試しに学習してみた
 - c. その結果、数時間あれば画風のコピーができることを発見してしまった (Stable Diffusionができる前)
 - d. いらすとやの人に報告して、緊急で技術報告書をプレ公開
 - e. 国会議員や弁護士、東大の先生と対策を相談
 - f. よくよく考えてみると、画風のコピーを許可して使う分には活用できるのではとのアイデアが出てくる
3. いらすとやの中の人と当社が共同でいらすとやの画風をコピーしたAI「AIいらすとや」を開発

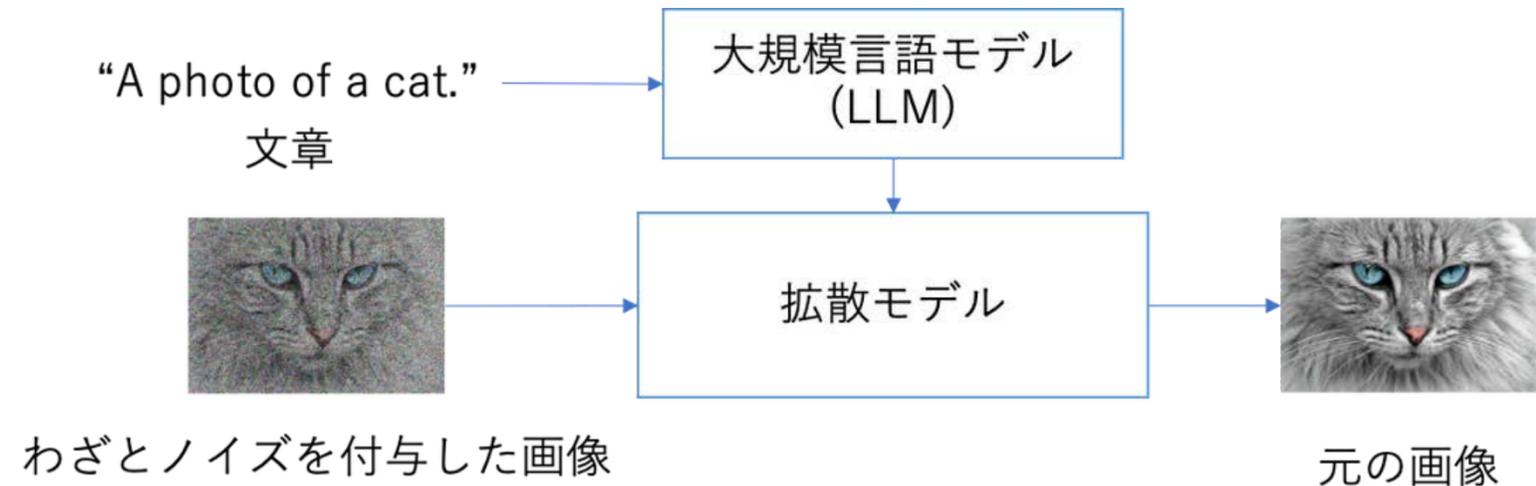


事前学習のやり方

数億以上の画像-文章ペアを必要とします。
事前学習には数ヶ月単位の莫大な計算量を必要とします。

文章	画像
a painting of a bridge with boats on it	
a japanese woodblock print depicting waves crashing on the shore	
a japanese woodblock print depicting people fishing	
a painting of a mountain with clouds in the sky	

画像-文章ペアの具体例



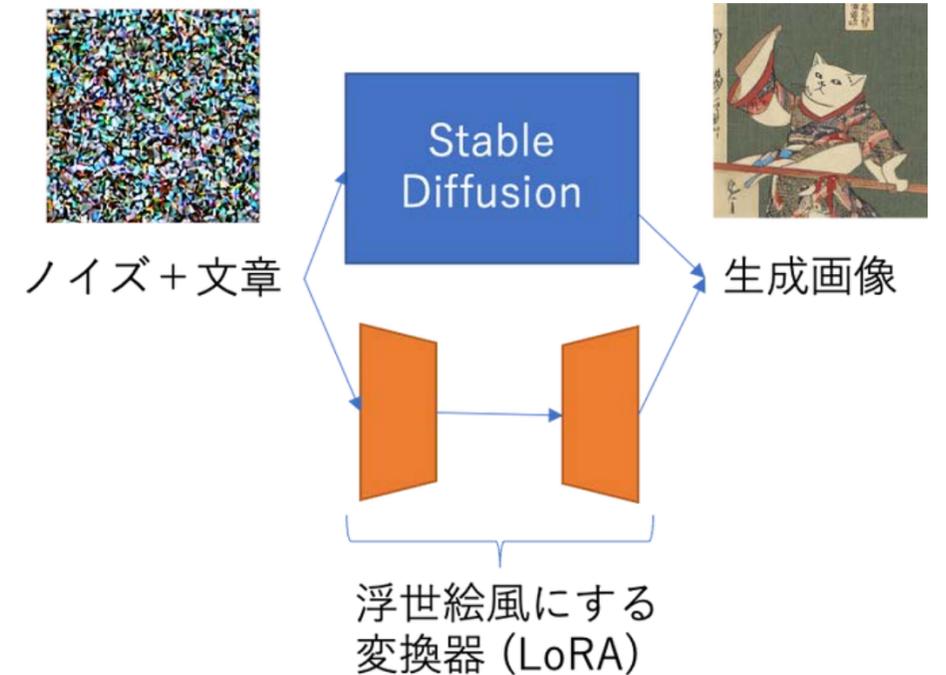
事前学習の具体例
(拡散モデルでの一例。厳密には他の学習法もあります)

追加学習のやり方

追加学習には簡易なものと同格的なものがあります。
数分で終わる簡易なもの代表がLoRAで、数日かかる本格的なもの代表がFull Finetuningです。

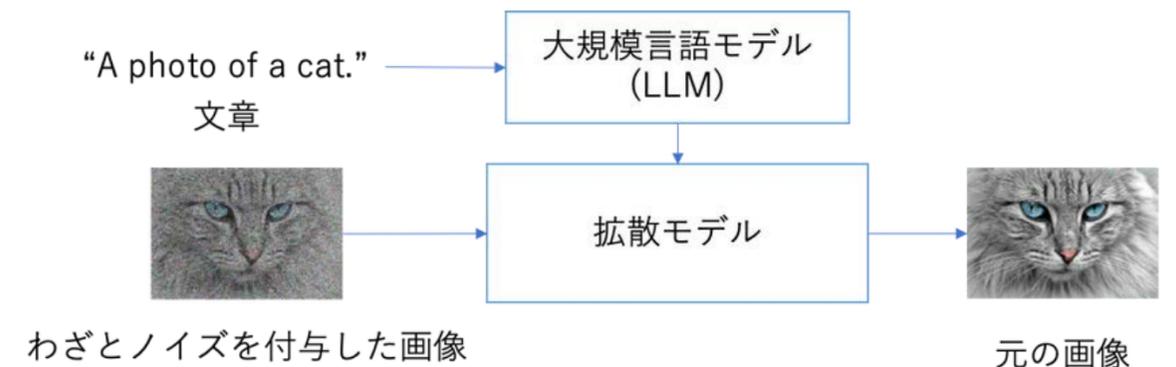
LoRA

LoRA (Low-Rank Adaptation) はサイズの小さい変換器を画像生成AIにくっつけることで画風などを変化させる技術です。



Full Finetuning

事前学習と同じ方法で学習するやり方です。
学習する具合を表す学習率が事前学習よりも小さく設定されます。



AI PICASSO

独自技術について

Our Technology

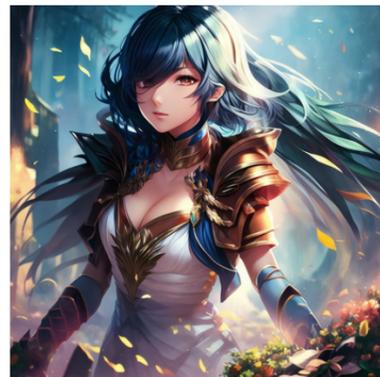


画像生成AI技術

独自の3つの画像生成AI技術を保有し、国内最高峰の画像生成AIの技術を誇ります。
会社が保有する画像データの画風やスタイルを再現するAIの共同研究も行っています。

アニメ/マンガ生成AI

- 世界最大クオリティのアニメ・漫画に特化した画像生成モデルを独自開発し、オープンソースとして公開しています。
- 著作権問題を解決した初の生成モデルです。



実写生成AI

- 通常のオープンソースの画像生成AIよりも、高クオリティかつ高解像度の画像を生成できる技術を独自で開発しています。

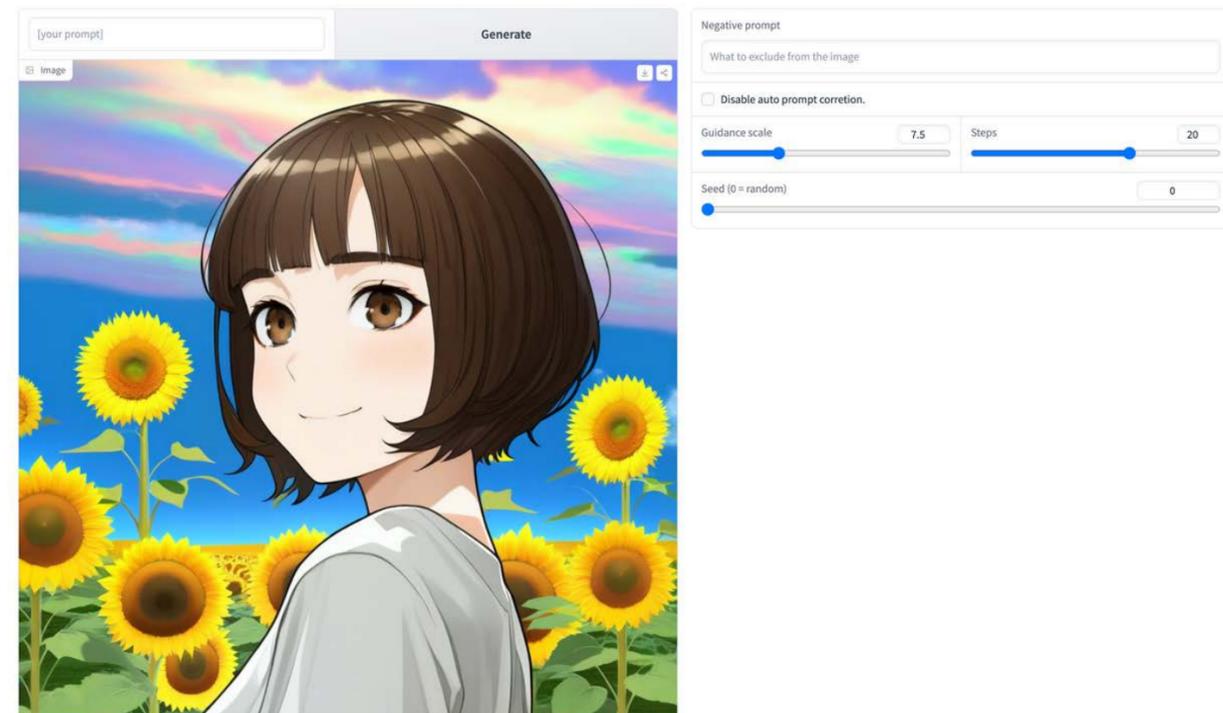


AIアバター

- 数枚の写真から、顔や画風を学習することができるAIです。
- 弊社独自の技術により、少数でも高クオリティな画像が生成可能です。



イラストAI



[ここから](#)上記のようなEmiによる画像生成を試せます

私たちはクリエイターを尊重した画像生成AIの開発を行っています。3段階の画像生成AIモデルを公開しています。

Clean Diffusion

著作権が切れた画像のみで学習したAI

Emi

無断転載画像を追加学習していないAI

Cool Japan Diffusion

クールジャパンを表現するために学習したAI