Cartografia da IA generativa

2024

Pesquisa e design de mapas
@tallerestampa

Tradução
Priscila Gonsales

Discursos sobre IA

- Ecologias de IA
- Crítica por meio das artes
- Conhecimento crítico
- Ética da IA
- Pesquisa acadêmica
- Ciências de dados
- Inovação de Big Tech
- Mistificação
- Determinismo tecnológico
- Imaginários do Vale do Silício
- Promoção exagerada (hype) constante
- Pânico da IA
- IA e extinção humana

Computação em escala planetária

Consumo de energia

1 a 1,3% da eletricidade mundial
As plataformas de Big Data estão se tornando cada vez mais parceiras na exploração de recursos de petróleo e gás

Combustíveis fósseis (+80%)
Principalmente movido a carvão, gás natural ou petróleo

Mais de 1% das emissões globais de gases com efeito de estufa

Aquecimento global

Estratégias Greenwashing

Mercado de compensação de emissão de carbono
As empresas compensam suas emissões de gases de efeito estufa pagando a terceiros para evitar emissões ou remover dióxido de carbono do ar. A compensação é feita por meio de projetos florestais e eólicos questionáveis, mais simbólicos do que reais

Grandes quantidades de resíduos radioativos

Fissão nuclear
Oklo (Sam Altman)

Energia renovável

Colonialismo verde:
Conflitos de uso da terra com as comunidades

Preços locais de eletricidade

Poder de cálculo

TSMC
(Companhia de Fabricação de Semicondutores de Taiwan)

Produz mais de 60% dos chips do mundo e mais de 90% dos chips mais avançados do mundo

Importância estratégica:
Medos crescentes de uma invasão chinesa a Taiwan (batalha tecnológica entre EUA e China)

Nvidia: Sistema proprietário CUDA para IA
- Intel
- AMD
- Google TPU
- AWS Trainium

Unidades de processamento gráfico (GPU)

Data centers

8.000 data centers em todo o mundo. Previsão de crescimento contínuo de 10% ao ano (para manter a garantia de treinamento de dados barato)

Consumo de energia
Consumo de energia
Um único data center pode consumir o equivalente a 50.000 residências

Sistemas de resfriamento (CRACs/CRASs)
Mais de 40 por cento do consumo de eletricidade em data centers

Resfriamento a água
Como o consumo extra de energia gera mais calor, os data centers de IA também precisam de métodos alternativos de resfriamento,como sistemas de resfriamento líquido

Torres de água

Calor

Servidores

Trocador de calor a placas

Estação de bombeamento

Centros de dados de IA
precisam de infraestrutura adicional que possa fornecer muito mais energia

Data centers atuais
5 a 10 quilowatts por rack

Requisitos de IA
60 ou mais quilowatts por rack

Azure Cloud AWS

Plataformas de Big Data

Microsoft Google Amazon

Software como serviço (SaaS) e plataforma como serviço (PaaS)
Interfaces de programação de aplicativos (APIs)

Dados de treinamento

Comunidades indígenas e locais:
300 bilhões de toneladas de carbono são armazenadas em terras administradas por comunidades indígenas em todo o mundo (devido a menores taxas de desmatamento). Essas comunidades são frequentemente excluídas das decisões sobre o uso da terra. No Quênia, milhares de pessoas foram despejadas de suas terras ancestrais como parte de um programa para reduzir o desmatamento

Colonialismo Digital

Afetações do lençol freático local

Secas e conflitos locais
No ano de 2023, moradores de Montevidéu, sofrendo com uma grave escassez de água, protestam contra os planos de construção de um data center do Google

Água

4-9 litros por kWh de potência do servidor.
O Google afirma que em 2022 seus data centers ao redor do mundo teriam usado 5,2 bilhões de galões de água

Resíduos
Cabos, baterias, fontes de alimentação ininterruptas (UPS), condicionadores de ar (CRACs e CRAHs), unidades de distribuição de energia (PDUs) e transformadores também são desativados e descartados periodicamente, quando as garantias expiram

Poluição sonora
O zumbido dos data center causa efeitos fisiológicos agudos e longitudinais

Geradores a diesel
Poluição do ar e impactos climáticos

Mais de 50.000 toneladas de lixo eletrônico, que leva milênios para se decompor, são enviadas para a África todos os anos. Elas acabam em lixões informais como os de Agbogbloshie, Gana, onde materiais perigosos são queimados, expondo as pessoas a vapores tóxicos e elementos radioativos

Mercúrio, cobre, chumbo e arsênico penetram no solo e nas vias navegáveis, causando acúmulo de substâncias químicas nocivas no ecossistema e em suas cadeias alimentares

Resíduos eletrônicos globais equivalentes a um ano: 7,3 kg per capita
82,4% depositados em aterros sanitários ou reciclados informalmente

64% dos resíduos eletrônicos enviados para centros de reciclagem na Europa são então enviados para a África

Cadeia de suprimentos de matérias-primas

O cobalto é usado em baterias recarregáveis de íons de lítio. Há relatos de uso de trabalho infantil e abusos de direitos humanos nas minas de cobalto congolesas

Mineração em águas profundas:
Impactos: destruição de ecossistemas, compactação do fundo do mar e criação de plumas de sedimentos que perturbam a vida aquática

Indústrias de mineração

Extração de metais e minerais exportados para o Sul Global

Depressão central — Salar — Deserto do Atacama — Cordilheira dos Andes

Argentina — Bolívia — Chile — Peru — Amazônas

"Corredor de mineração" de cobre (sul do Peru)

Li — O Salar de Atacama, uma planície de sal com quase quatro vezes o tamanho da cidade de Nova York, é uma das maiores minas de lítio do mundo, um componente essencial na produção de baterias. A crescente demanda afeta cada vez mais as comunidades locais, ameaçando seu acesso à água e empobrecendo a biodiversidade

Cu — Os fabricantes utilizam cobre nos chips de computador mais potentes por sua superior condutividade elétrica. O "corredor de mineração" do sul do Peru (usado pela chinesa MMG Ltd, pela suíça Glencore e pela canadense Hubbay) é um dos pilares da economia, mas também uma das principais fontes de conflito devido à distribuição desigual da renda da mineração e aos problemas de saúde causados pela poluição da água

Au — A legislação oficial brasileira não permite a mineração em terras indígenas. Mas a mineração ilegal de ouro na Amazônia brasileira disparou desde 2019. Pesquisadores documentaram dezenas de milhares de garimpeiros de pequena escala e mais de 320 garimpos ilegais, sendo que o número real provavelmente é muito maior. A mineração de ouro em pequena escala levou ao desmatamento generalizado, à poluição por mercúrio e à sedimentação de rios vitais

Solda	Sn	China, Indonésia, Birmânia, Peru
Bateria	Li	Triângulo do Lítio (Bolívia, Chile, Argentina), Austrália, China
Circuitos	Ta W	Ruanda, Brasil, Congo, China
Condutores	Au Cu Ag	Congo, Zâmbia, Chile, Peru, México, China, Brasil
Processador	Si	China
Tela	Al K	Canadá, Rússia, Bielorrússia
LCD	Sn In Ga	China, Indonésia, Birmânia, Peru
	Elementos de terras raras	China
Vibração	Dy Nd Pr Tb	
Tela	Dy Eu Gd	
Microfone	Dy Nd Pr Tb	

Desde 2007, mais de sete bilhões de smartphones foram fabricados

A fabricação de racks de servidores envolve cobre, silício e lítio. Cabeamento, barramentos, aparelhagem de comutação e outros componentes utilizam grandes quantidades de cobre

Falta de transparência

Capital de risco do Vale do Silício

Corrida para implementar IA inovadora

Onda de especulação tecnológica

Líder "visionário" celebrado pela mídia

Startups de IA

OpenAI — DeepMind — Anthropic

Agenda política

Institutos para o futuro

Lobbies

IA e extinção humana
Demandas por autorregulação

Regulações Públicas
- Lei de IA da UE
  Requisitos de transparência em bases de dados
- Acordo voluntário de empresas dos EUA

Conteúdos sintéticos
Regimes de verdade: discursos públicos e afetações

Populismos autoritários
Polarização política

Entretenimento personalizado

Desinformação e propaganda
- Sistemas de Informação Notícias
- Recomendação Categorização de conteúdo Governança algorítmica

Redes neurais profundas

Modelo de IA

Ajuste fino
- Pontuação de dados
- Anotação de dados
- Rotulagem

(RLHF)

Filtragem
- Moderação de conteúdo
- Prompts proibidos
- Detecção de "toxicidade"
- Alinhamento

Alinhamento

Fotografia com IA
Mecanismos de busca
Assistentes de IA
Processadores de texto
Atendimento ao cliente

LAION-5B
5,8 bilhões de pares de imagem e texto

CommonPool
12,8 bilhões de pares de imagem-texto

Conjuntos de dados

As informações contidas em conjuntos de dados estão sempre inseridas nos contextos que as criam e utilizam. Permanecem vinculadas às comunidades, indivíduos, organismos e ambientes em que foram criadas

Maior conjunto de dados possível

Deepen IA treina refugiados sírios para anotar dados para empresas como Google e Amazon. Sama treina refugiados em Uganda, Quênia e Índia para realizar tarefas curtas de dados

Campos de refugiados (Shatila do Quênia e do Líbano)

Violação das leis de propriedade intelectual

Extração e privatização
"O direito de ler é o direito de minerar": apropriação do discurso "A informação quer ser livre" versão de meados dos anos 90 da Liberdade na Internet

Codificação de visões do mundo
- Estereótipos, ideologias e pontos de vista raciais e de gênero extremistas
- Arquétipos narrativos
- Cultura dominante
- Estilos como padrões estatísticos que podem ser replicados automaticamente

Espaços latentes massivos e privados

Texto estatístico — Imagens estatísticas — Som estatístico

Motores de IA generativa

ChatGPT Bard Meta Llama Google Palm Falcon Midjourney Stable Diffusion Stable Video Runway Gen-2 Google Lumiere Sora Adobe Firefly

Dependência de serviços
Em agosto de 2023: 34 milhões de imagens/dia

Web scraping

Técnicas para download e processamento automático de conteúdo web. Common Crawl e OpenWebText2: instituições de pesquisa científica para mineração de dados em contextos não comerciais

- Wikipedia
- Projeto Gutenberg
- Bancos de imagens

- livros
- blogs
- comentários
- postagens
- avaliações
- artigos
- shares
- likes
- etiquetas
- pontuações

- imagens
- vídeos
- ilustrações
- pinturas
- marcas
- estilos

- fotografia
- trabalho
- engenharia
- equipamento
- conhecimento técnico

- música
- sons
- vozes

420 cabos submarinos distribuídos seguindo as rotas dos telégrafos e das navegações coloniais

Provedores de serviços de Internet

Conteúdo digital da internet imagem, texto e som

Social media — Web2

Reconhecimento de imagem (rotulagem, segmentação, pose, profundidade)

Criação de perfil e filtros de conteúdo

Mídias sociais

Trabalho humano

Mercados de trabalho: Quênia (36% abaixo da linha internacional de pobreza), Índia (68%), Filipinas (18%), Venezuela, Colômbia (39%), Uganda (28%)

Empresas: Sama, Appen, Clickworker, Scale AI

Extração de dados

Termos e condições

Empregos para freelancers — Trabalhadores terceirizados — Instrutores de IA não remunerados

Designer de produto — Roteirista — Editor/a de filme — Escritor/a — Arquiteto/a — Redator/a — Desenvolvedores web — Cenógrafo/a — Web designer — Renderizador/a — Ator/Atriz — Estilista — Rotoscopistas — Usuário/a

Músicos — Compositor/a — Artista de voz — Artista de textura — Designer de UX — Artista de storyboard — Moda — Gaming — Designer gráfico — Ilustradores — Motion graphics — Fotógrafos

Chips avançados

ASML
Fabricante holandês de equipamentos para a produção de circuitos integrados. É o maior fornecedor mundial de sistemas de fotolitografia para a indústria de semicondutores, com mais de 5/6 do mercado mundial

TSMC
A Taiwan Semiconductor Manufacturing Company controla uma grande fatia do mercado global de chips avançados e tem uma posição forte na fabricação de semicondutores de ponta

Montagem
Os dispositivos são montados em fábricas como a Foxconn, localizadas no Sul Global, onde a mão de obra pode ser adquirida a baixo custo e as proteções legais para os trabalhadores são escassas

Gadgets digitais

Suas vidas úteis são, em média, inferiores a dois anos, uma consequência da obsolescência programada

Ascensão da economia da atenção e suas tecnologias

Descritores de imagens e textos < Virada Estética

IA
Modelos em larga escala executados em infraestruturas privadas

Computação global

- 2021 OpenAI apresenta a rede neural chamada CLIP, que classifica imagens usando descrições textuais
- 2021 Latent Diffusion Processo de difusão dentro de um espaço latente
- 2019 GPT-2 (Generative pre-trained transformer)
- 2016 Geração de imagens (32 x 32 pixels) a partir de legendas
  Elman Mansimov et al.

Controle da comunicação por meio de imagens > Imagem para descritores

Pesquisa →

AI
Código aberto ou disponível publicamente em papel

Computação local

- 2003 Primeiro modelo de linguagem de rede neural feed-forward Yoshua Bengio et al.
  Conjunto de dados Caltech 101 10.000 imagens
  Fei-Fei Li et al.
- 2009 Conjunto de dados ImageNet
  14 milhões de imagens e mais de 20.000 categorias
  Fei-Fei Li et al.
- 2012 Alexnet. Primeira implementação rápida de uma CNN com GPU a vencer um concurso de reconhecimento de imagem
  Alex Krizhevsky, Ilya Sutskever e Geoffrey Hinton
- 2014 Rede Adversarial Generativa (GAN)
  Ian Goodfellow et al.
- 2015 Modelo de difusão
  Usado para geração de imagens, começando com uma imagem composta de ruído aleatório para que a rede elimine ruído iterativamente