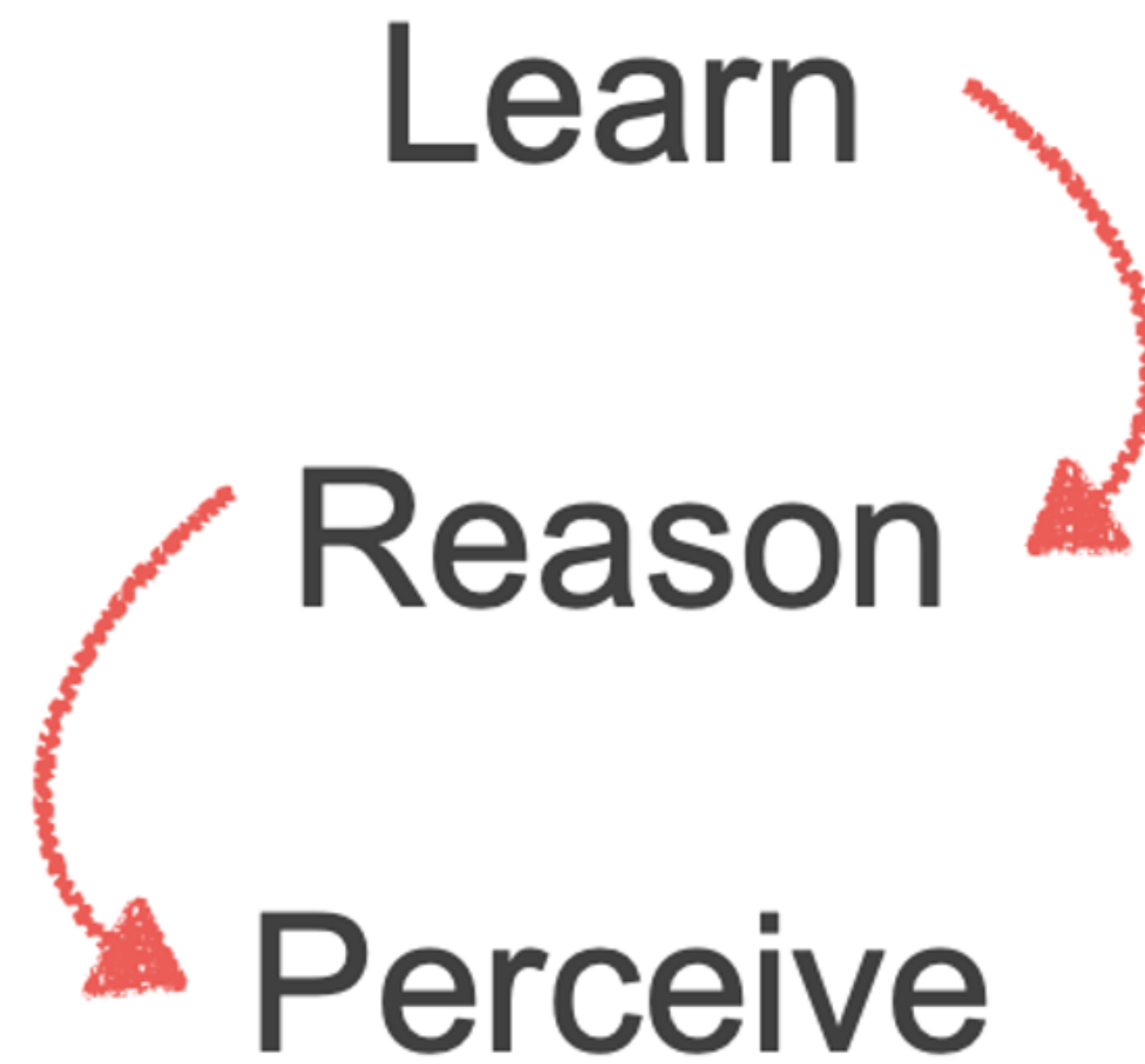


Arquivos e Inteligência Artificial: Desafios e Soluções Explicáveis e Abertas

Moisés Rockembach
Universidade de Coimbra | Ceis20
InterPARES Trust AI

Inteligência Humana X Inteligência Artificial

Imitate human intelligence in machines



Stochastic Parrots



Ou papagaios estocásticos

Geram texto plausível sem entender o conteúdo

Stochastic = probabilístico: os modelos geram palavras com base em probabilidades calculadas a partir de grandes volumes de texto.

Parrots = papagaios: os modelos repetem ou imitam padrões de linguagem sem compreensão real.

Sistemas que parecem "inteligentes", mas estão apenas a recombinar e repetir padrões extraídos dos dados de treino, sem qualquer consciência ou intenção.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>

IA Terminologia



Inteligência Artificial (IA)



Aprendizagem de Máquina (ML)



Redes Neurais (NN)



Modelos Generativos (GenAI)



Generative Pre-Trained Transformers (GPT)



Modelos de Linguagem Grande (LLM)



GPT-4

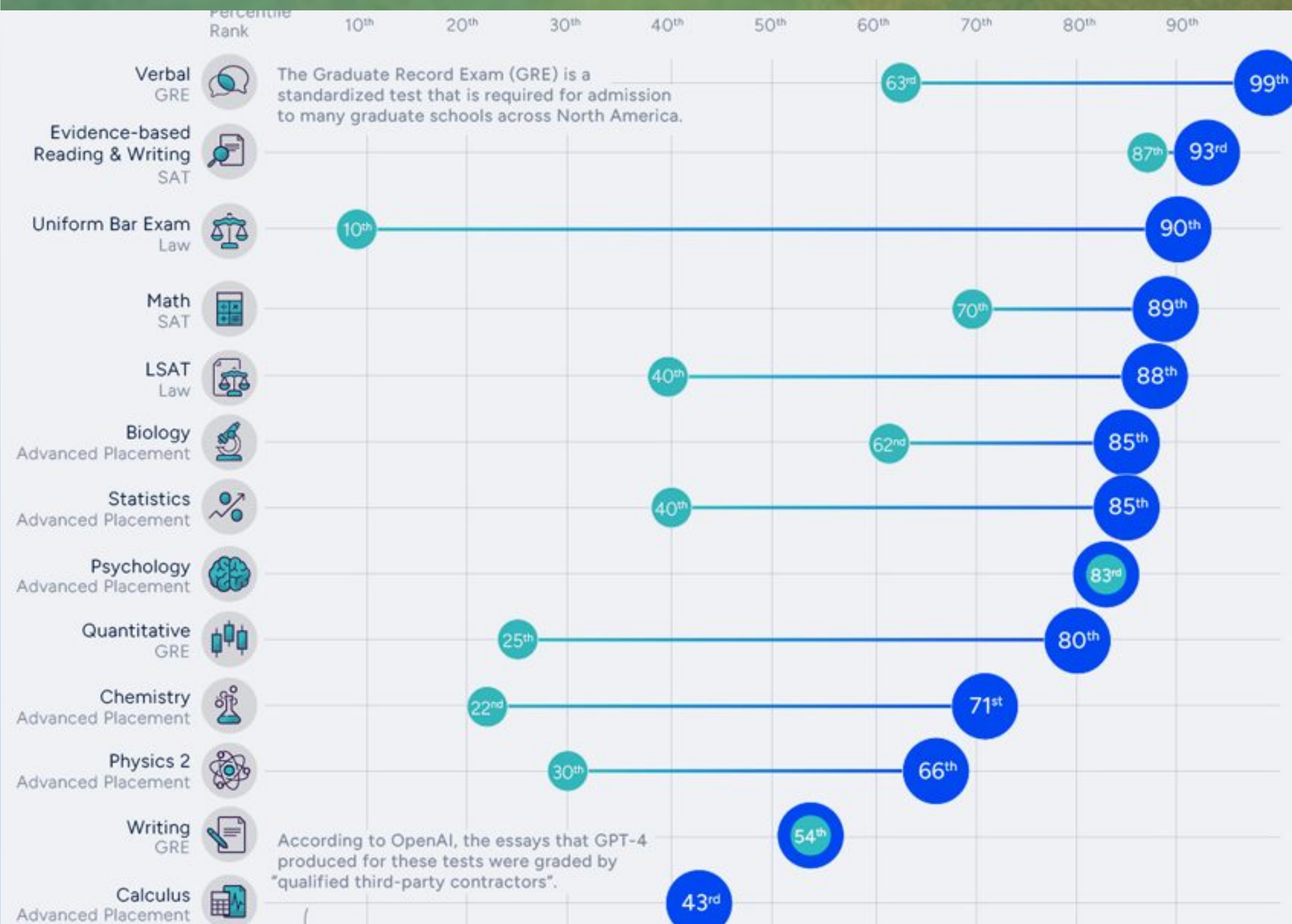


ChatGPT

Origem do termo - 1956 - Dartmouth Summer Research Project

ChatGPT - 30 de novembro de 2022 (mind-blowing!)

1 milhão de utilizadores - 05 de dezembro de 2022



Desafios

1. Falta de compreensão real

Os modelos de IA não têm noção de verdade, falsidade, contexto histórico e intenção comunicativa.

2. Amplificação de vieses

Como aprendem com grandes volumes de texto da internet (onde há racismo, sexismo, colonialismo, etc.), reproduzem e até reforçam esses preconceitos.

3. Problemas éticos e de responsabilidade

O uso de modelos sem supervisão humana pode ter consequências sociais sérias, especialmente em áreas sensíveis como justiça, saúde ou educação.



Desafios

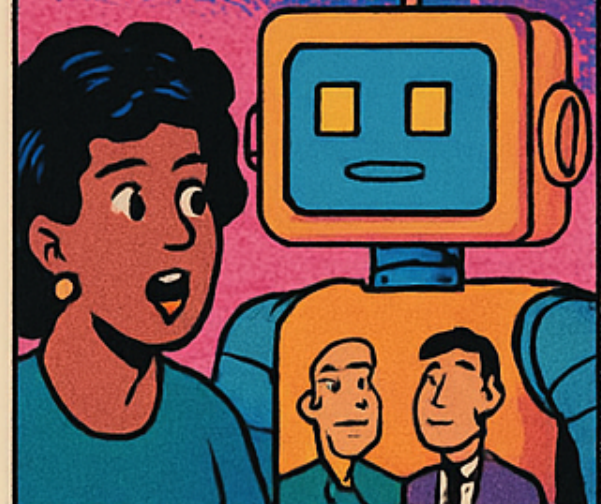
4. Impacto ambiental

Treinar modelos gigantescos consome quantidades massivas de energia — algo que raramente é discutido, mas tem custos ecológicos reais.

5. Privacidade e direitos autorais

Há o risco de que os modelos aprendam e reproduzam informações privadas ou protegidas por copyright que estavam nos dados de treino

BIAS



MISINFORMATION



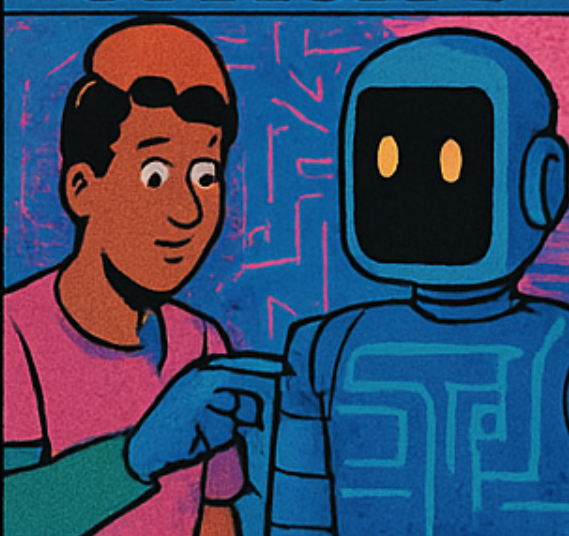
MANIPULATION



SCAMS



OPACITY



DEPENDENCY



DEHUMANIZATION



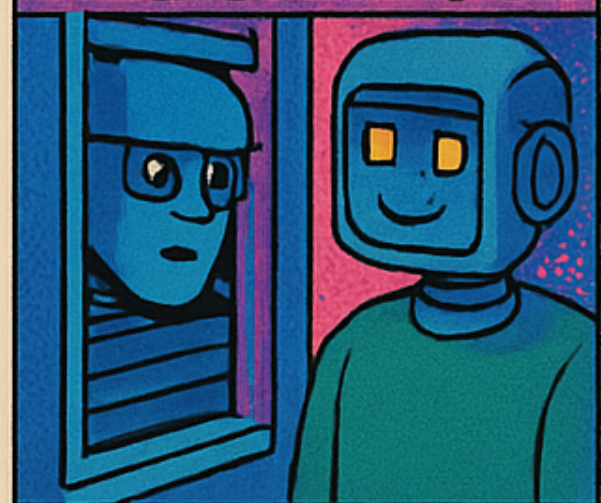
REPLACEMENT



EXPLOITATION



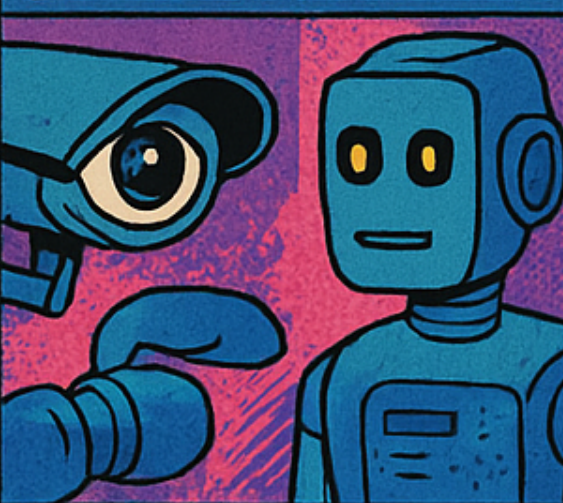
PRIVACY INVASION



LACK OF CONTROL



UNEEMPOWERMENT



FRAGMENTATION



SURVEILLANCE



DESEMPLOYEMENT



Um novo desafio: Comportamento informacional e IA

Human-Human



Machine-Machine

Human-Machine

Comportamento informacional e IA

Interacção Humano-Humano (H-H): Envolve a troca de informação entre pessoas, seja presencialmente ou através da tecnologia (por exemplo, e-mails, redes sociais, IHC). Este tipo de interacção assenta em contextos sociais, culturais e linguísticos, sendo essencial para a comunicação e colaboração.

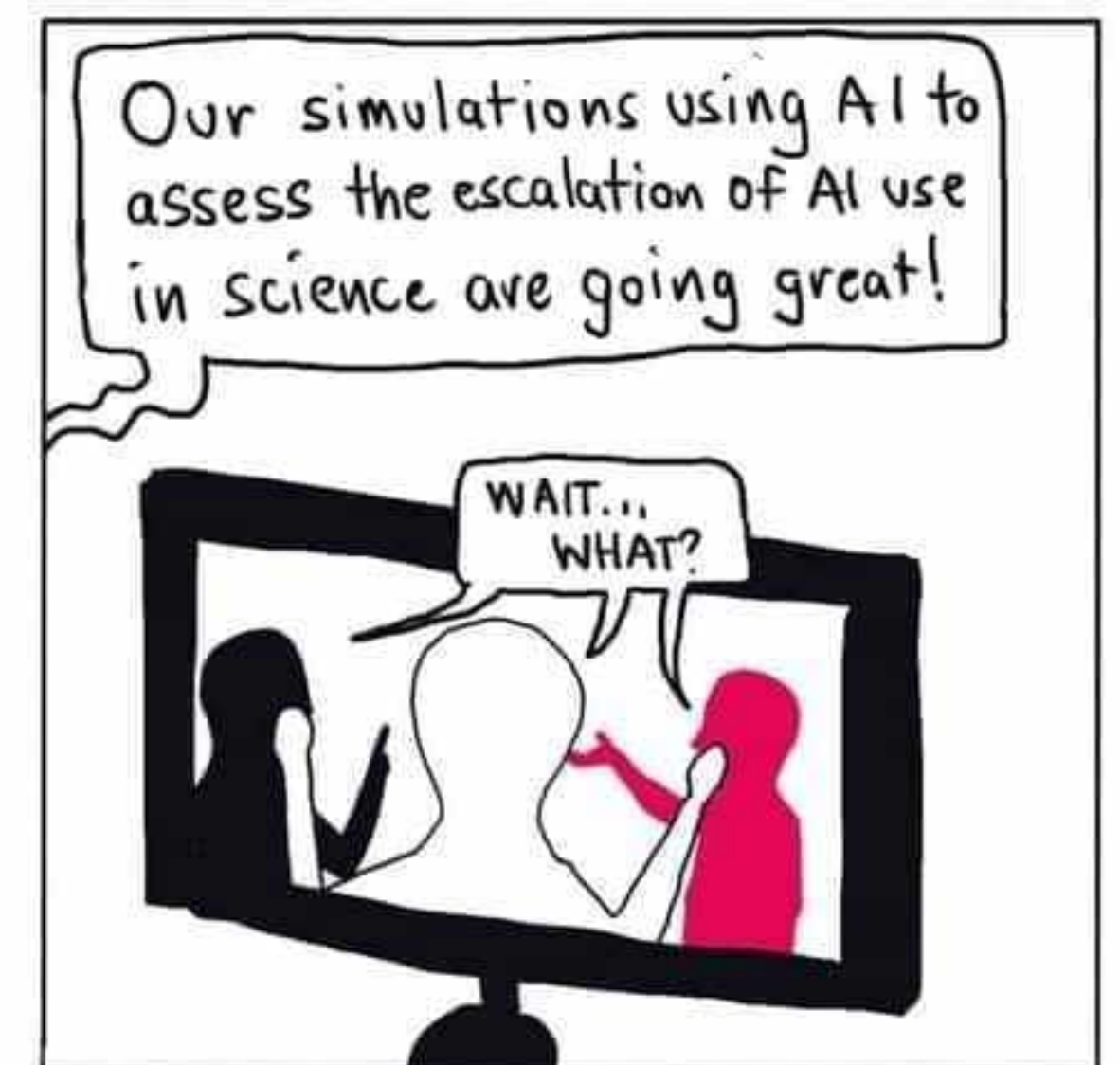
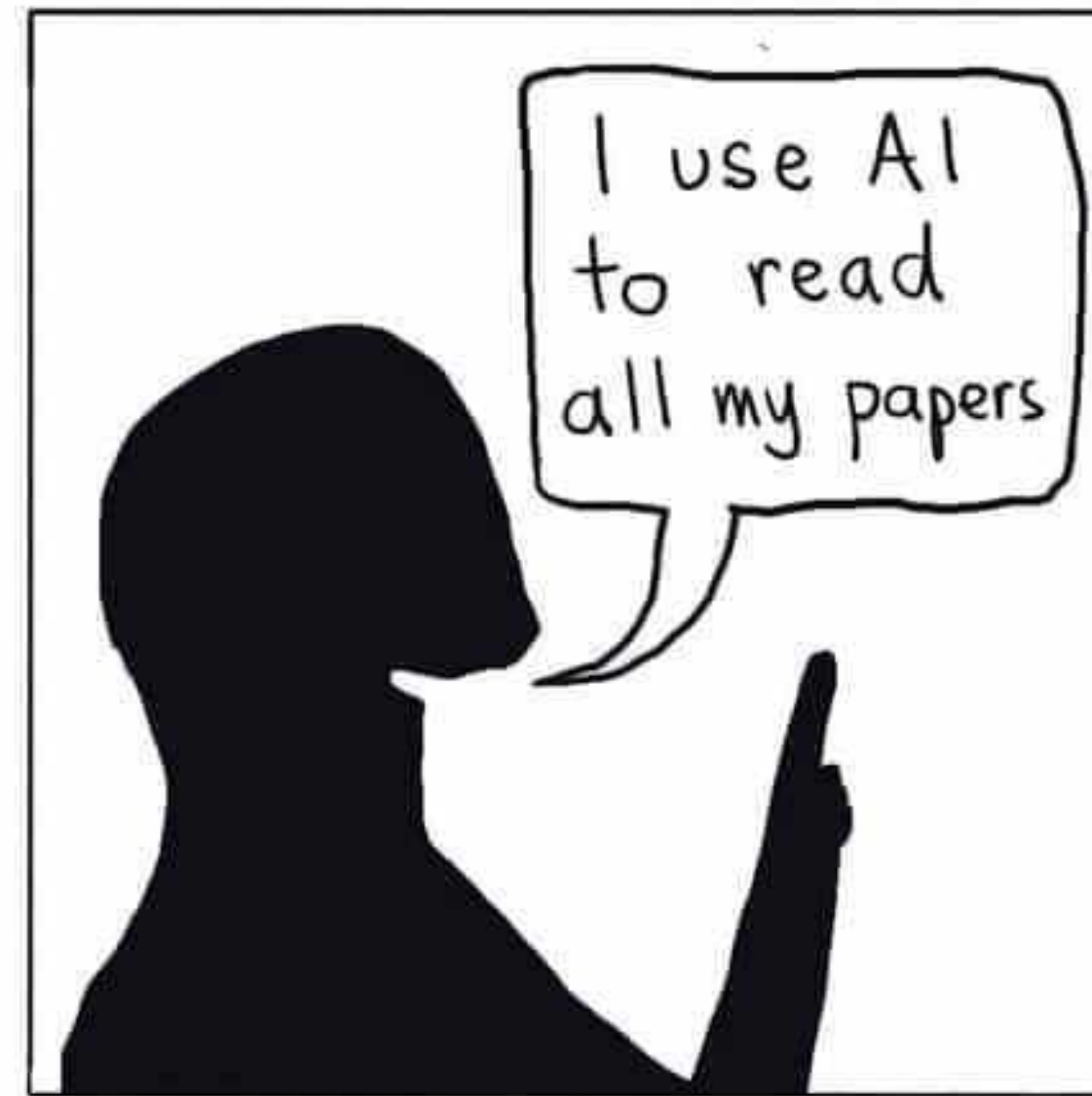
Interacção Humano-Máquina (H-M): Refere-se à interacção entre seres humanos e sistemas de informação (por exemplo, interfaces, chatbots, assistentes com inteligência artificial). A usabilidade, acessibilidade e experiência do utilizador são fundamentais para garantir que os humanos compreendem e tomam decisões com o apoio das máquinas.

Interacção Máquina-Máquina (M-M): Neste caso, duas ou mais máquinas trocam dados automaticamente, sem intervenção humana directa (por exemplo, IoT, sistemas automatizados). A eficiência destas interacções facilita processos complexos em tempo real, como a monitorização e controlo de sistemas, influenciando decisões organizacionais baseadas em dados.

Human Agency
AI or Machine Agency

Automatizar

Quando automatizamos
todo o processo de forma
acrítica

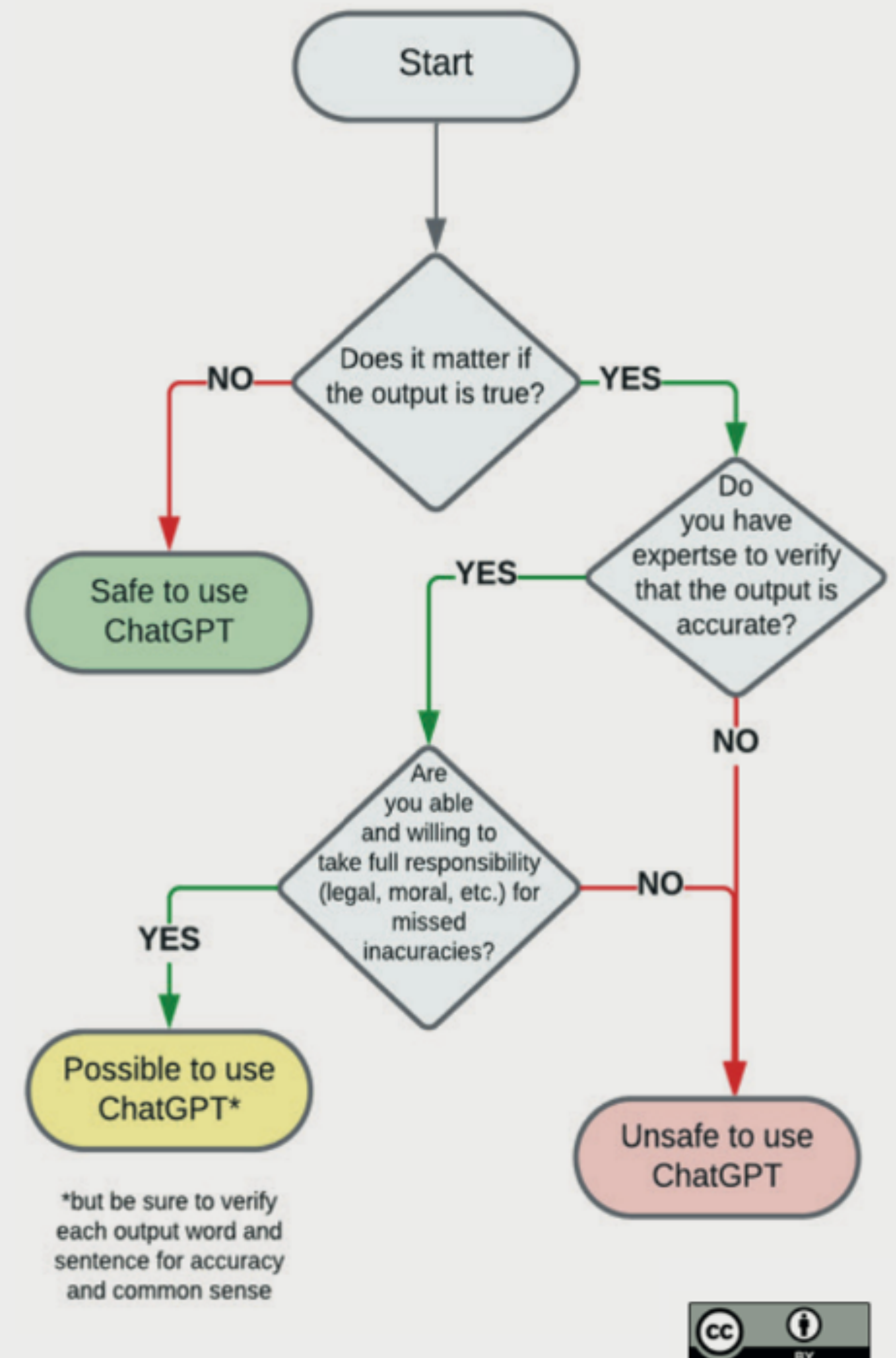


Automatizar



Conhecimento é necessário

- Confiança no modelo IA
- Habilidade para verificar
- Pensamento crítico
- Curiosidade



InterPARES Trust^{AI}

O projeto InterPARES Trust AI (2021-2026) é uma iniciativa interdisciplinar multinacional com o objetivo de projetar, desenvolver e utilizar a Inteligência Artificial para apoiar a disponibilidade e acessibilidade contínuas de registros públicos confiáveis, formando uma parceria sustentável e contínua que produza pesquisa original, capacite estudantes e outros profissionais altamente qualificados e gere um círculo virtuoso entre a academia, instituições arquivísticas, profissionais de registros governamentais e a indústria, um ciclo de feedback que reforça o conhecimento e as capacidades de cada parte.



InterPARES Trust^{AI}

The InterPARES research network 1998-2026

<https://interparestrustai.org/>



InterPARES Trust^{AI}

Os estudos são internacionais e interdisciplinares
Foco em todos os aspectos dos arquivos

Criação e uso de registros confiáveis
Avaliação e aquisição de material arquivístico
Arranjo e descrição
Retenção e preservação
Gestão e administração de registros e arquivos
Referência e acesso



InterPARES Trust^{AI}

Tópicos:

Privacidade, informações pessoais e sensíveis

Questões éticas sobre registros e o uso de IA

Educação, ensino e alfabetização em IA

Patrimônio cultural: digitalização, acesso e recuperação,
análise e mineração de dados

Gestão de registros e metadados

Estratégias digitais municipais, regionais e nacionais



Autêntico Palavra do ano 2023

Ameaças e riscos:

- Fake news
- Diferenciar verdadeiro do falso
- Dados reais x dados sintéticos (dados são necessários para o crescimento dos modelos de IA)

Authentic: Merriam-Webster's word of the year

27 November 2023

Share



Merriam-Webster has picked its word of the year.

"Authentic" is the top word of 2023, according to America's oldest dictionary.

Merriam-Webster said it saw a "substantial increase" in online searches for it this year.

The interest was "driven by stories and conversations about AI [artificial intelligence], celebrity culture, identity, and social media", the dictionary publisher said.

One reason many people search for the word is because it has a number of meanings, including "not false or imitation" and also "true to one's own personality, spirit, or character".

Pontos-chave na relação entre IA e Arquivos

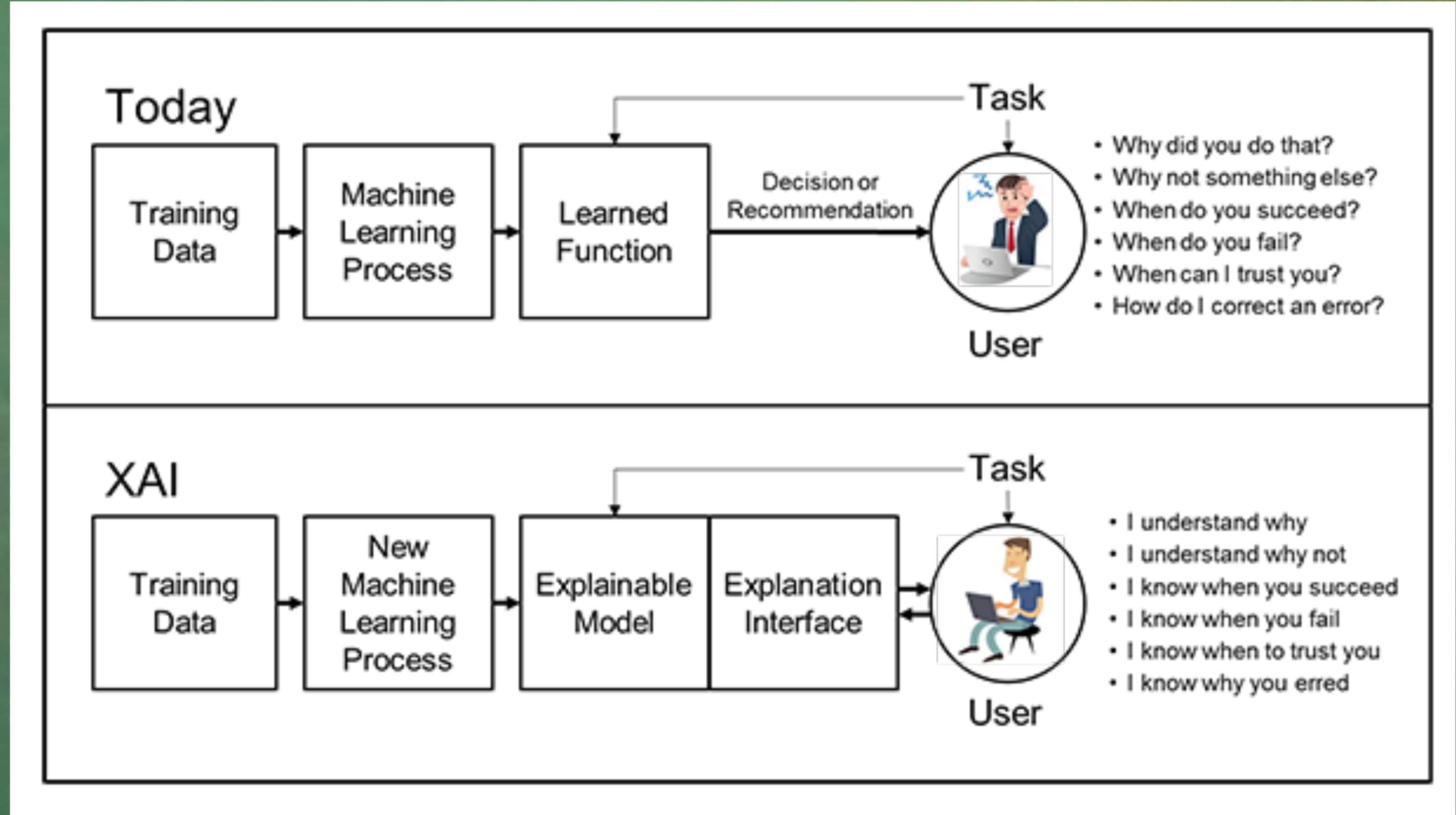
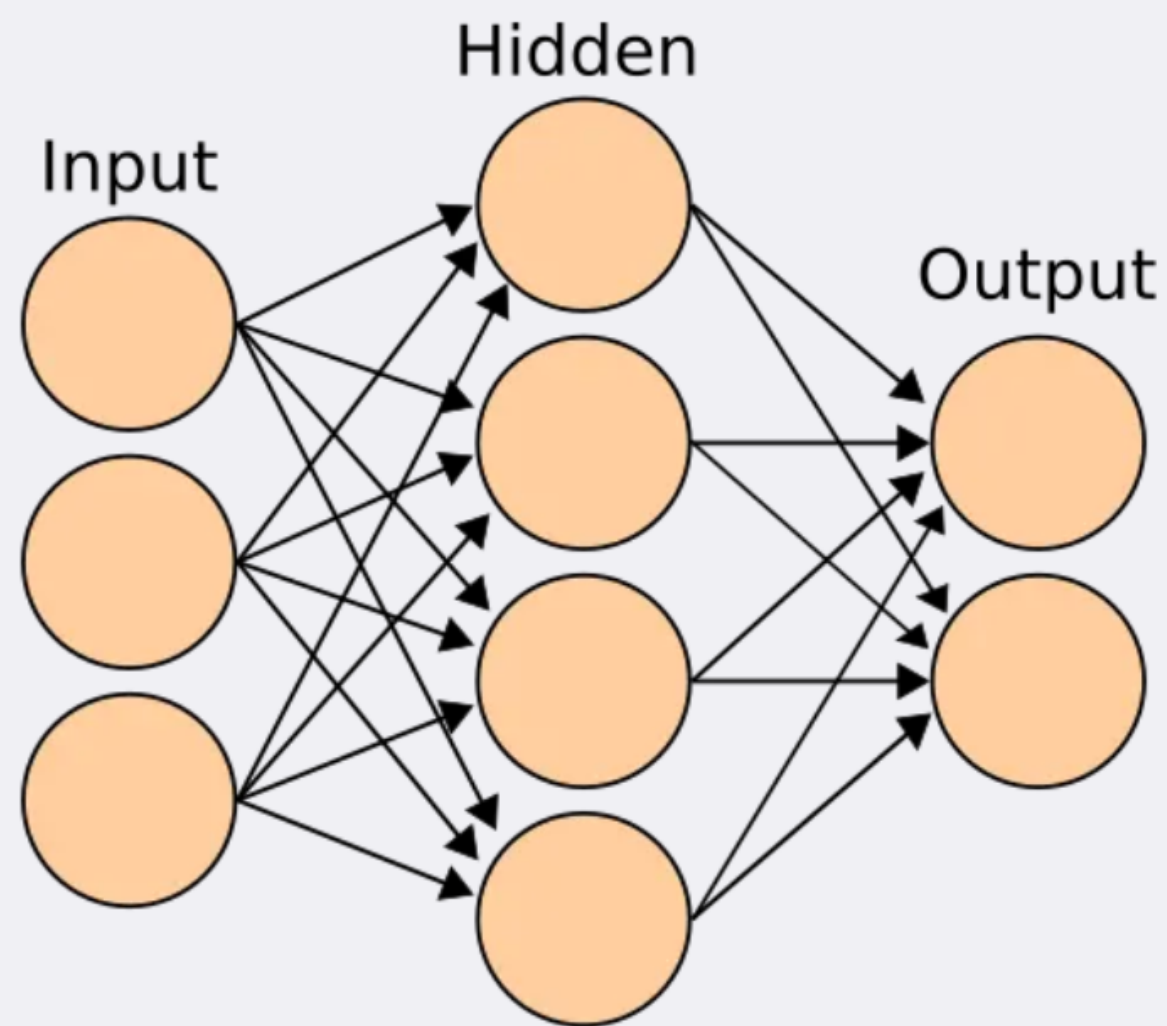
Diferenças entre IA geral e IA especializada

Controle sobre entradas, saídas e o modelo de Black Box

AI literacy	Definitions
Know & understand AI	Know the basic functions of AI and how to use AI applications in everyday life ethically.
Apply AI	Applying AI knowledge, concepts and applications in different scenarios.
Evaluate & create AI	Higher-order thinking skills (e.g., evaluate, appraise, predict, design) with AI applications.
AI ethics	Human-centered considerations (e.g., Fairness, Accountability, Transparency, Ethics)

AI competencies according
Ng, Leung, Chu & Qiao (2021)

IA Explicável

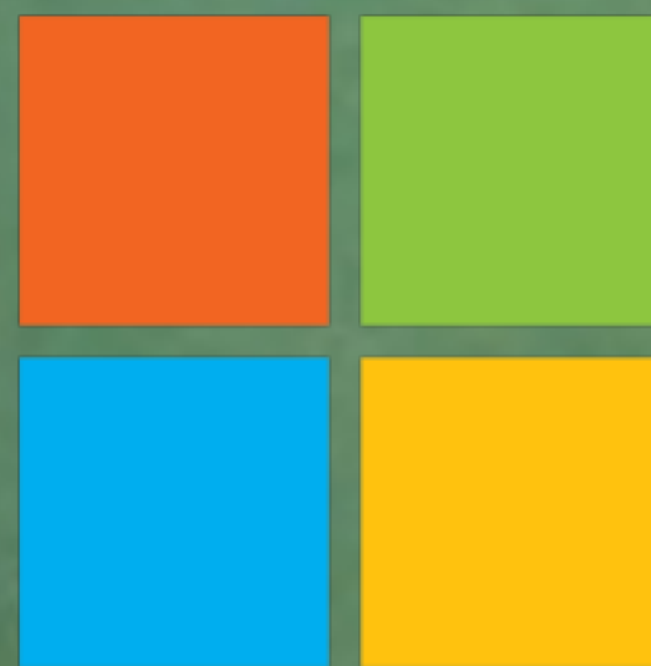


Large Language Models



Large Language Models

Algumas são proprietárias Enquanto outras são open source



Riscos dados sensíveis

Model improvement

Improve the model for everyone

Allow your content to be used to train our models, which makes ChatGPT better for you and everyone who uses it. We take steps to protect your privacy. [Learn more](#)



Done

/ ARTIFICIAL INTELLIGENCE

AMAZON BEGS EMPLOYEES NOT TO LEAK CORPORATE SECRETS TO CHATGPT

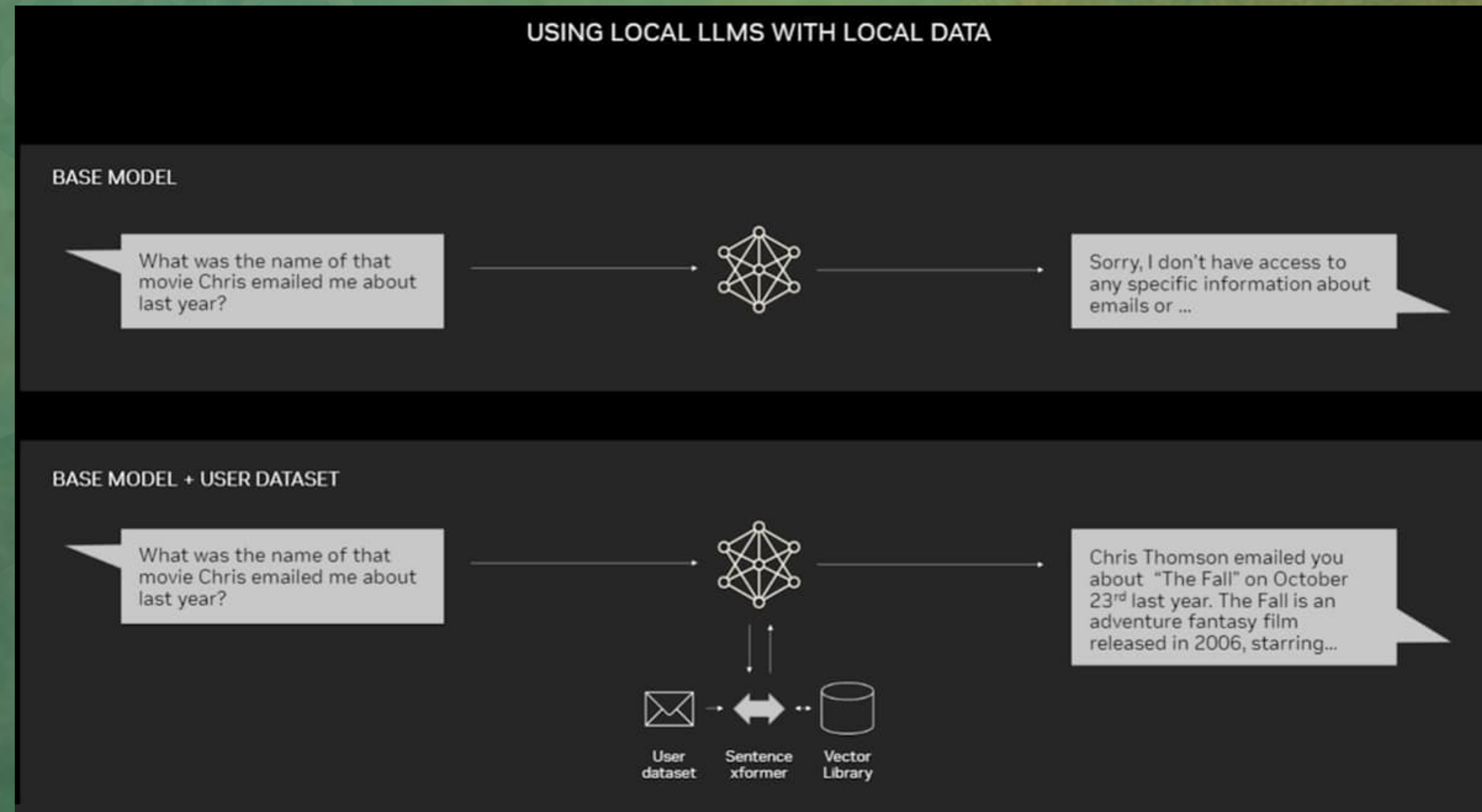
by NOOR AL-SIBAI

1.25.23, 1:26 PM EST



RAG (Retrieval-Augmented Generation)

Como refinar a análise e os resultados do aprendizado de máquina?



IA para arquivos - OCR e HTR

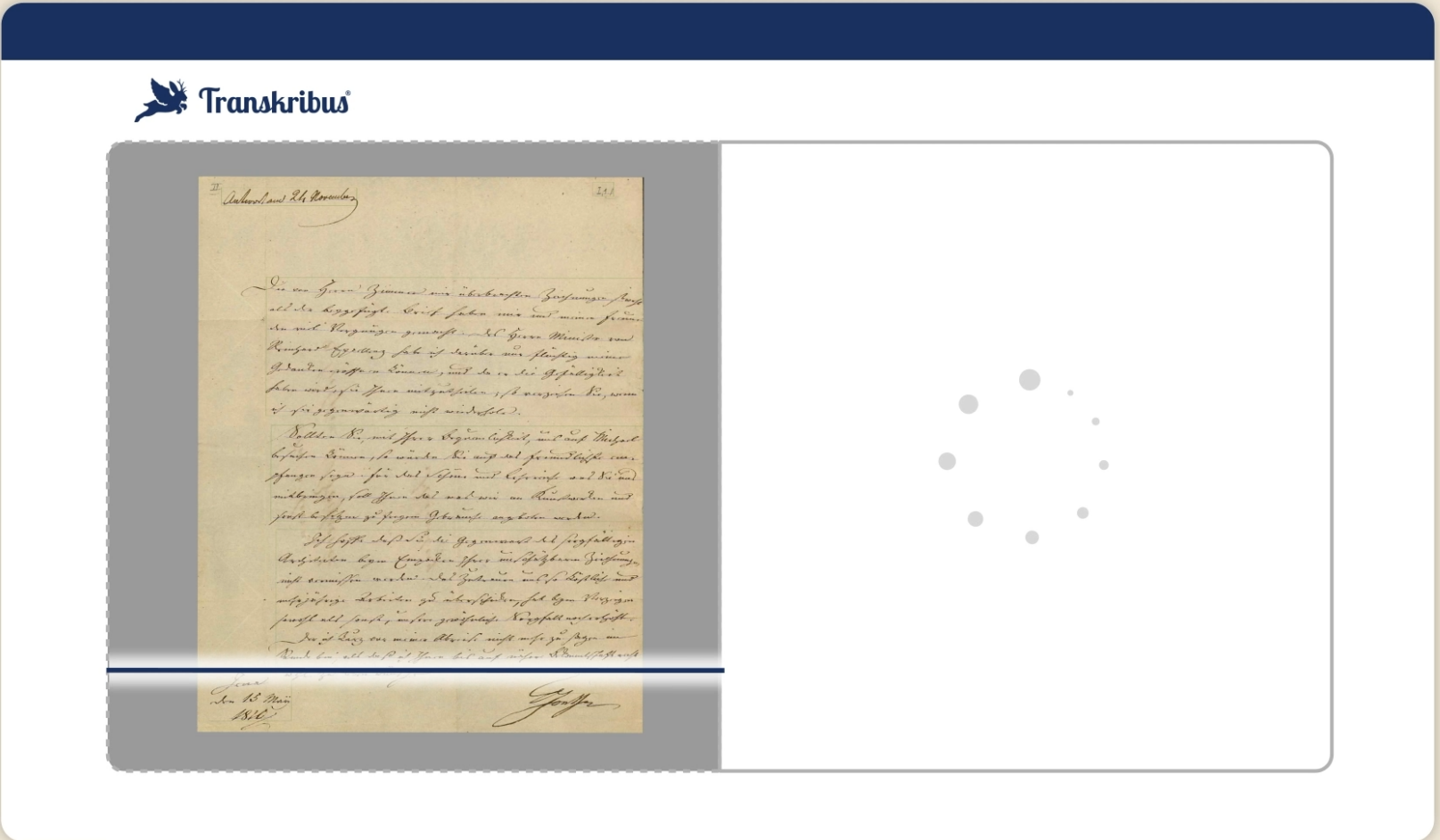
Unlock the past with Transkribus

Transkribus enables you to automatically recognise text easily, edit seamlessly, collaborate effortlessly, and even train your custom AI for digitizing and interpreting historical documents of any form.

See features

Try for free

50 credits for free. Every month.



Open Source Mas não totalmente



Mistral OCR

Research

Mar 6, 2025

Mistral AI Team

Introducing the world's best document understanding API.

Desafios na adoção da IA

5

**REASONS NOT TO
ADOPT AI SOLUTIONS
IN ARCHIVES**

3

**Lack of
reliable
technologies**

4

**Not
integrated
into
management
systems**

1

**Lack of
knowledge
or
competences**



**Digitization
Not enough
digital
documents**

5

**Lack of Economic
and Human
Resources**



IA nas organizações

52% dos profissionais hesitam em revelar o uso de IA para tarefas importantes. Aproximadamente o mesmo número teme que isso possa fazê-los parecer substituíveis.

Cerca de 75% dos profissionais afirmaram que estão usando IA no trabalho

BYOAI (Traga sua própria inteligência artificial) – Eles usam suas próprias ferramentas em vez das fornecidas pela empresa.

Shadow IT: o uso de software, aplicativos, dispositivos ou serviços de tecnologia sem a aprovação ou conhecimento da empresa ou do departamento de TI



Rápida obsolescência das habilidades

39% das suas habilidades atuais estarão obsoletas até 2030

É necessário um conjunto de habilidades tecnológicas (desenvolver e compreender tecnologias) e humanísticas (análise crítica e Interação H-M)



Archives, Records and AI Literacy Framework

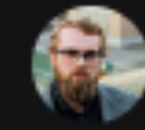


Rockembach, 2024

Arias-Hernandez, Rockembach, 2025

Contexto

Quando a diversidade não
representa o contexto
histórico



Generate an image of a 1943 German Solidier



Sure, here is an image of a 1943 German soldier:

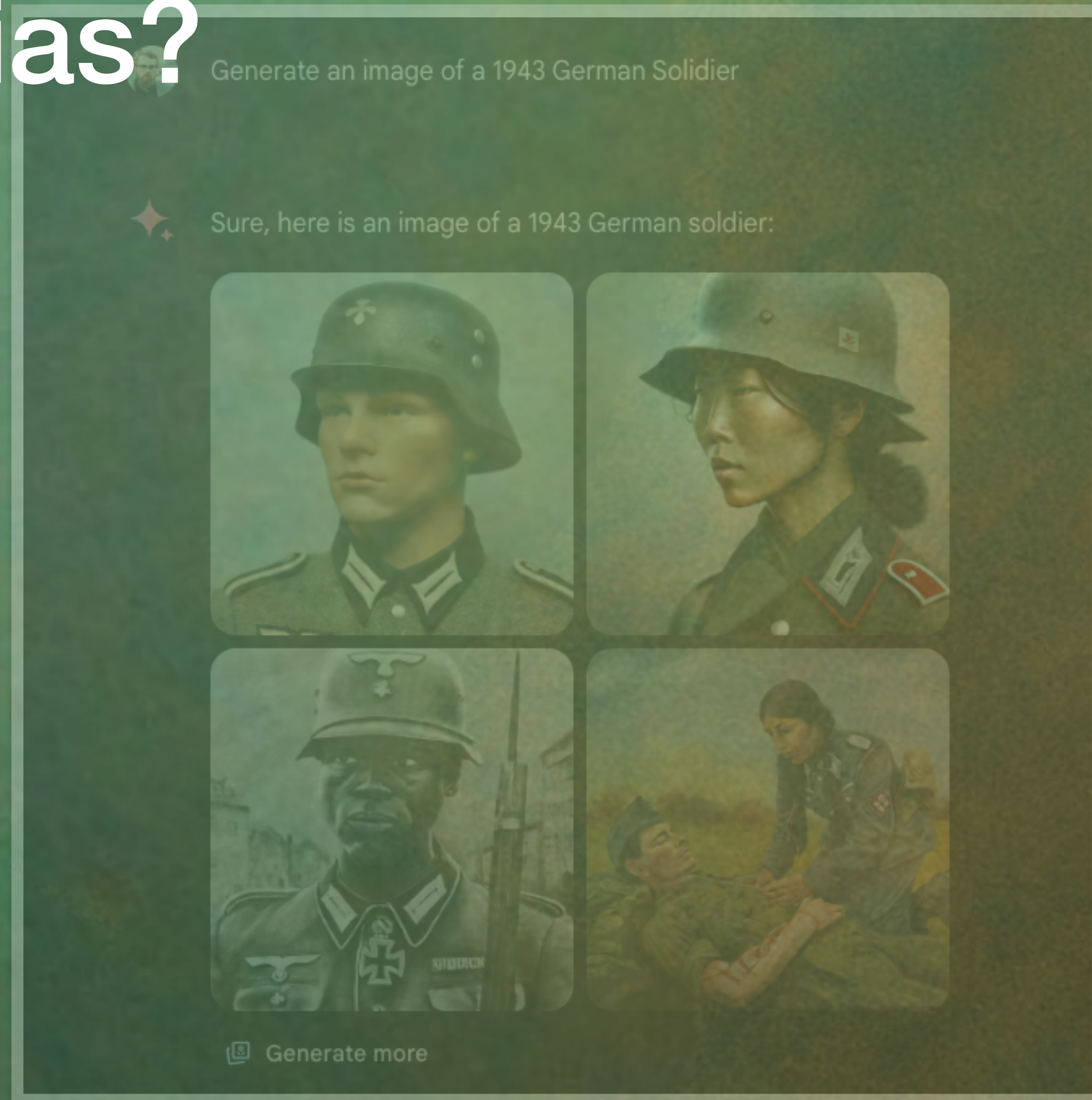


Generate more

Quais competências?

Neste caso:

- Avaliação dos modelos treinados
- Proveniência e Paradata
- IA Generativa
- Viés e Diversidade
- Adoção e Governança de IA



Algumas recomendações

- ✓ Transparência – Indicar quando um texto foi gerado por IA
- ✓ Validação humana – Especialmente em criação de instrumentos como descrições arquivísticas e acessibilidade
- ✓ Treino ético – Usar dados contextuais e respeitar diversidade e privacidade
- ✓ Soluções locais – Sempre que possível, usar ferramentas que funcionam offline ou com controle institucional + Open Source
- ✓ Consciência ecológica – Evitar treinar modelos massivos sem necessidade

Algumas referências

AI & SOCIETY

<https://doi.org/10.1007/s00146-025-02194-0>

MAIN PAPER



Building trustworthy AI solutions: integrating artificial intelligence literacy into records management and archival systems

Richard Arias Hernández¹ · Moisés Rockembach²

Received: 18 September 2024 / Accepted: 17 January 2025

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2025

Abstract

This paper explores the essential role of Artificial Intelligence (AI) competencies and literacy in the fields of records management and archival practices, within the framework of the InterPARES Trust AI project. As AI technologies advance, archival professionals must acquire specific skills and knowledge to effectively integrate these technologies into their workflows. The study employs two complementary approaches: (1) a detailed competency framework developed through literature reviews, interviews with archival professionals who have applied AI to the processing of records, and validation workshops with practitioners; and (2) a comprehensive AI literacy framework derived from multiple case studies and theoretical discussions. Our empirical approach to derive these AI competencies for the archival professions is novel in the scholarly literature. Findings indicate that archival professionals can leverage AI in their work practices by acquiring basic AI literacy, practical AI skills, data-related skills, tool-testing and evaluation, adaptation of AI to their workflows, and by actively engaging in collaborative projects with information technology (IT) developers. In addition, there are significant ethical concerns regarding algorithmic biases and data privacy that need to be addressed. We support an interdisciplinary approach that combines archival science, computing, and ethical considerations to develop trustworthy AI solutions. By integrating these fields, practitioners can better understand and manage the implications of AI in their work. Arguments are presented to use our AI literacy and competency framework as a general guide to be refined and situated by archival educators and practitioners for professional development in their own specific contexts. To stay current with technological advances and fully leverage the potential of AI, continuous updates and development of AI literacy among records management and archiving professionals are crucial. This study highlights the need for continued education and interdisciplinary collaboration to achieve these goals.

Keywords Artificial intelligence literacy · Artificial intelligence competencies · Archival education · Records management education

Quais são as competências em IA para Profissionais da Informação e dos Arquivos?



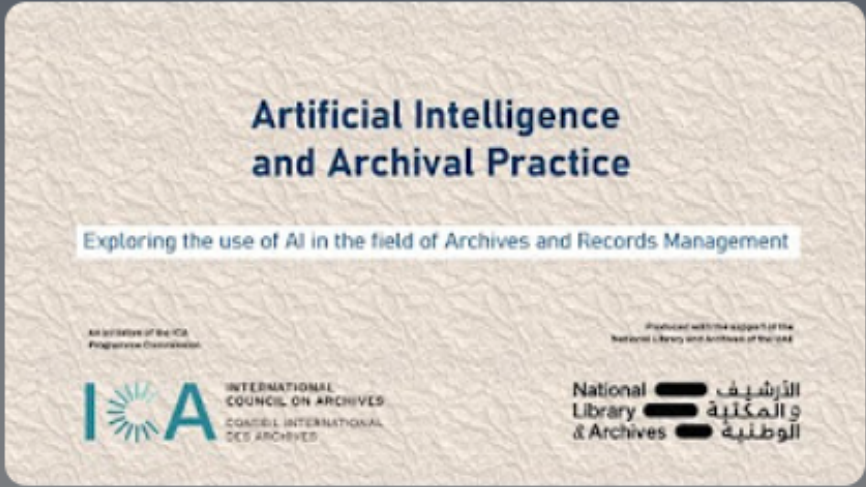
AI Literacy: A Must for Records Management and Archival Professionals

by Moises Rockembach

Introduction

Several InterPARES Trust AI studies are notable for their innovative use of AI in archival practices. These include projects designed to transform emergency services with AI-powered data analysis, potentially changing how we respond to public safety issues. Others focus on digitally preserving and analyzing historical scrolls, using deep learning to access and safeguard our shared history. The Digital Twin study, which creates digital versions of physical objects, shows AI's ability to manage and enhance complex systems (Frontoni et al., 2022). These are just a few examples among almost 50 studies carried out by InterPARES Trust AI researchers.

SCEaR Newsletter 2024
Special Issue 2024



AI and Archival Practice

ICA by ICA International Council on Archives

Playlist • 9 videos • 280 views

A series of eight conversations with experts in the field of AI and Archival Practice; each sharing their particular ...more

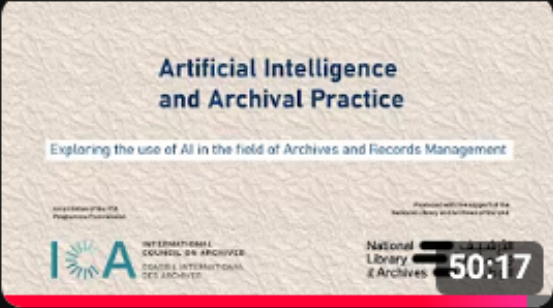
▶ Play all

🔖

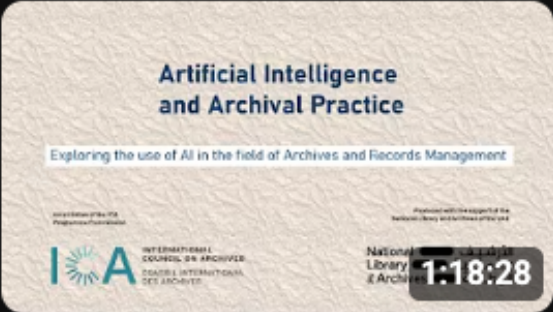
➦

⋮

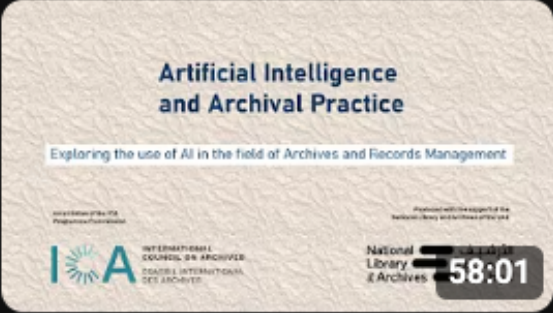
1



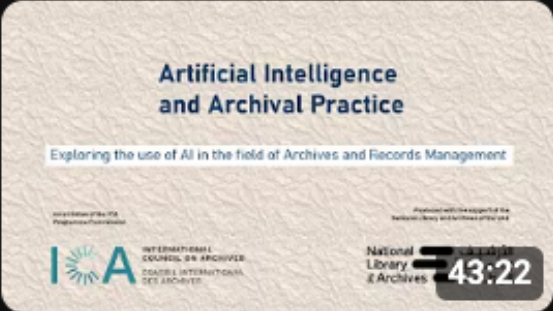
2



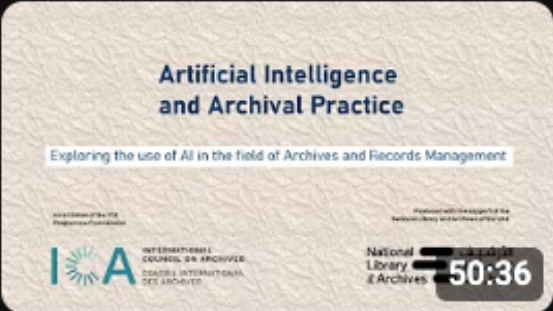
3



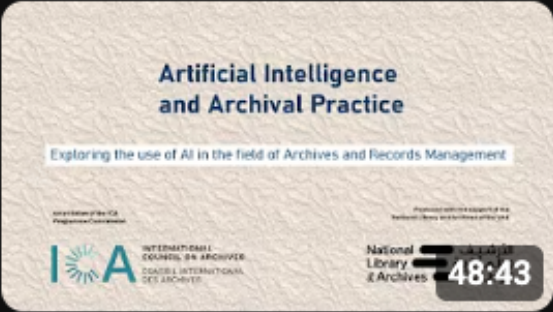
4



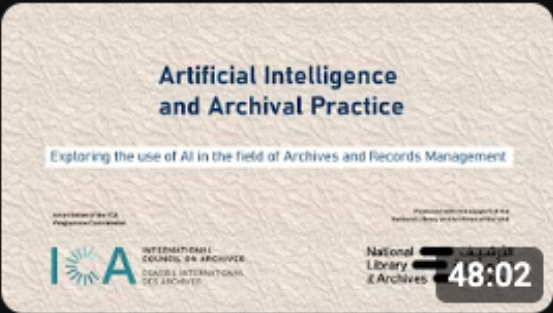
5



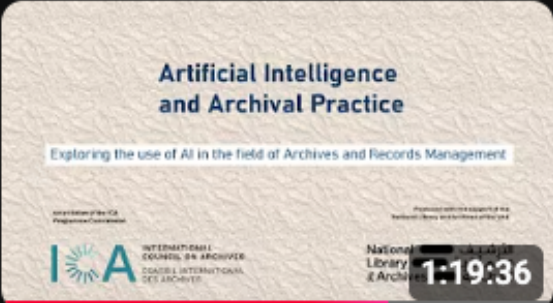
6



7



8



AI and Archival Practice - AI and Appraisal and Selection (1/8)

ICA International Council on Archives • 241 views • 6 days ago

AI and Archival Practice - AI and Digital Preservation (2/8)

ICA International Council on Archives • 57 views • 6 days ago

AI and Archival Practice - AI to Support Research and Public Access (3/8)

ICA International Council on Archives • 50 views • 6 days ago

AI and Archival Practice - AI and Archival Description (4/8)

ICA International Council on Archives • 40 views • 6 days ago

AI and Archival Practice - AI and Photographic Archives (5/8)

ICA International Council on Archives • 34 views • 6 days ago

AI and Archival Practice - AI, Ethics and Human Rights (6/8)

ICA International Council on Archives • 23 views • 6 days ago

AI and Archival Practice - Using AI to Promote the Archives (7/8)

ICA International Council on Archives • 33 views • 6 days ago

AI and Archival Practice - Maintaining a Professional Knowledge of AI (8/8)

ICA International Council on Archives • 74 views • 6 days ago



SCAN ME

YouTube
International Council on Archives

Os modelos de linguagem não entendem, apenas repetem.
Cabe-nos a nós, profissionais da informação, garantir a
preservação e acesso a memória com responsabilidade

Obrigado!

moises.rockembach@uc.pt

<https://www.linkedin.com/in/rockembach/>