

# Preservação Digital e Inteligência Artificial

Moisés Rockembach  
Universidade de Coimbra, CEIS20

**"PRESERVAÇÃO DA  
INFORMAÇÃO DIGITAL:  
FERRAMENTAS E  
PRÁTICAS"**

# Inteligência Artificial e Ciência da Informação

- Automatiza a descrição e indexação de documentos e arquivos históricos
- Automatiza a catalogação e classificação de livros e documentos nas bibliotecas
- Melhora a precisão dos metadados através de etiquetagem e indexação orientadas por IA.
- Chatbots para atendimento automatizado, oferecendo respostas rápidas e personalizadas aos utilizadores
- Automatiza a transcrição e tradução de documentos
- Permite a criação de gémeos digitais para a exploração virtual

# Inteligência Artificial e Ciência da Informação

- Automatiza a descrição e indexação de documentos e arquivos históricos
- Automatiza a catalogação e classificação de livros e documentos nas bibliotecas
- Melhora a precisão dos metadados através de etiquetagem e indexação orientadas por IA.
- Chatbots para atendimento automatizado, oferecendo respostas rápidas e personalizadas aos utilizadores
- Automatiza a transcrição e tradução de documentos
- Permite a criação de gémeos digitais para a exploração virtual

Gotcha!



# Inteligência Artificial e Ciência da Informação

ChatGPT ▾

Qual a relação entre ciência da informação e inteligência artificial?

A relação entre a ciência da informação e a inteligência artificial (IA) é cada vez mais estreita, com a IA a transformar profundamente as práticas de gestão e acesso à informação. Eis algumas formas concretas dessa integração:

- **Automatiza a descrição e indexação de documentos e arquivos históricos**, acelerando o processamento e facilitando a recuperação da informação.
- **Automatiza a catalogação e classificação de livros e documentos nas bibliotecas**, garantindo maior consistência e eficiência nos sistemas bibliográficos.
- **Melhora a precisão dos metadados** através de técnicas de etiquetagem e indexação orientadas por IA, como o reconhecimento de entidades e temas em grandes volumes de texto.
- **Chatbots para atendimento automatizado**, oferecendo respostas rápidas, personalizadas e acessíveis aos utilizadores de serviços informacionais.
- **Automatiza a transcrição e tradução de documentos**, permitindo o acesso mais inclusivo e rápido a conteúdos em diversos idiomas ou suportes (por exemplo, áudio ou manuscritos).
- **Permite a criação de gémeos digitais** de bibliotecas, arquivos ou museus, possibilitando experiências de exploração virtual de acervos e ambientes informacionais.



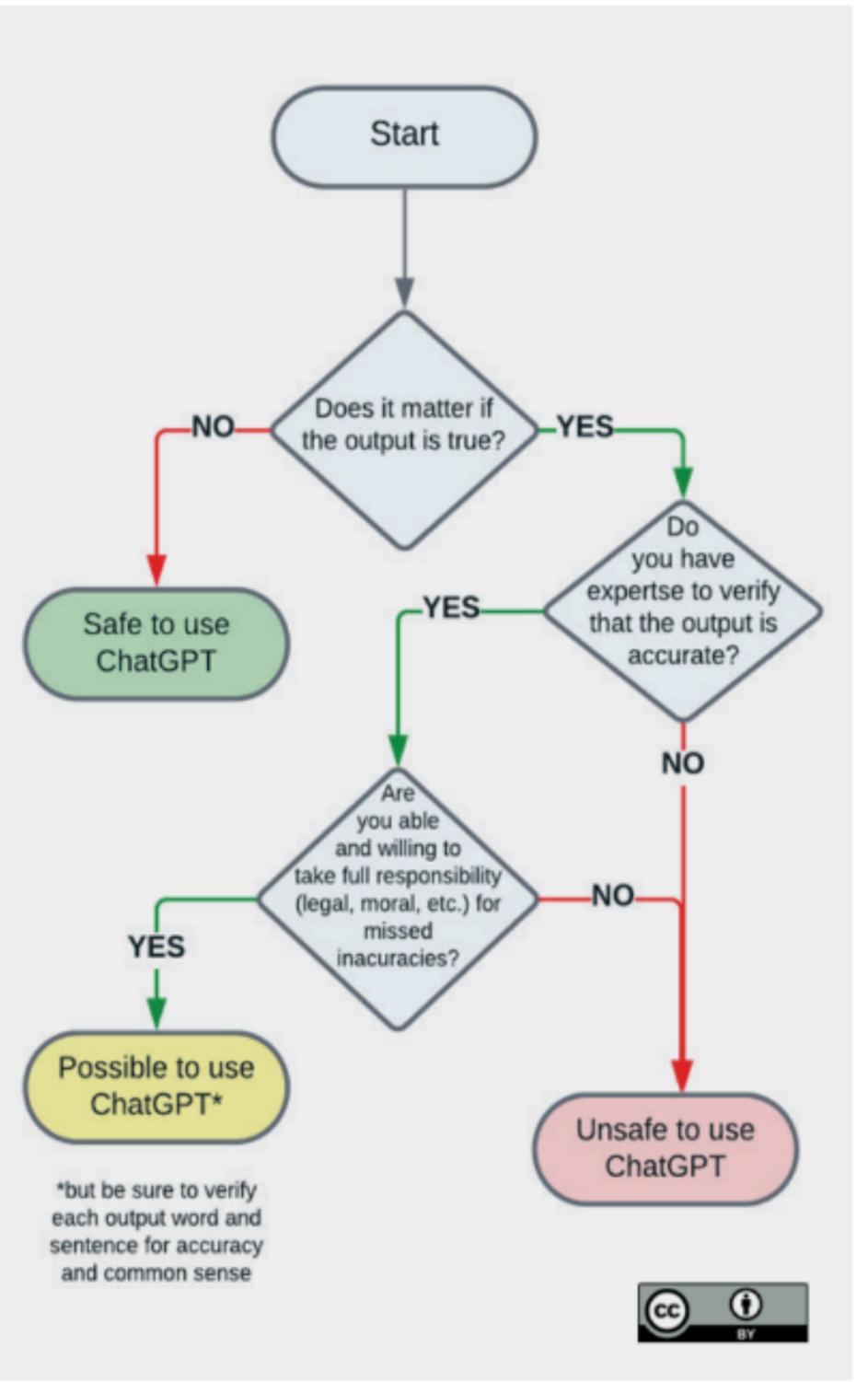
**ChatGPT**

# Inteligência Artificial e Ciência da Informação

ChatGPT ▾

A relação entre a IA e a ciência da informação está se tornando cada vez mais estreita, com a IA assumindo algumas formas concretas de aplicação:

- Automatiza a produção e o processamento de conteúdo
- Automatiza a organização e a manutenção de maior consistência
- Melhora a precisão da busca em IA, como o reconhecimento de padrões
- Chatbots para atendimento ao usuário, acessíveis aos usuários
- Automatiza a produção de conteúdos em massa
- Permite a criação de experiências de aprendizagem



...ção e inteligência artificial?

...ez mais estreita, com a IA a assumir algumas formas concretas

...cos, acelerando o

...bibliotecas, garantindo

...e indexação orientadas por texto.

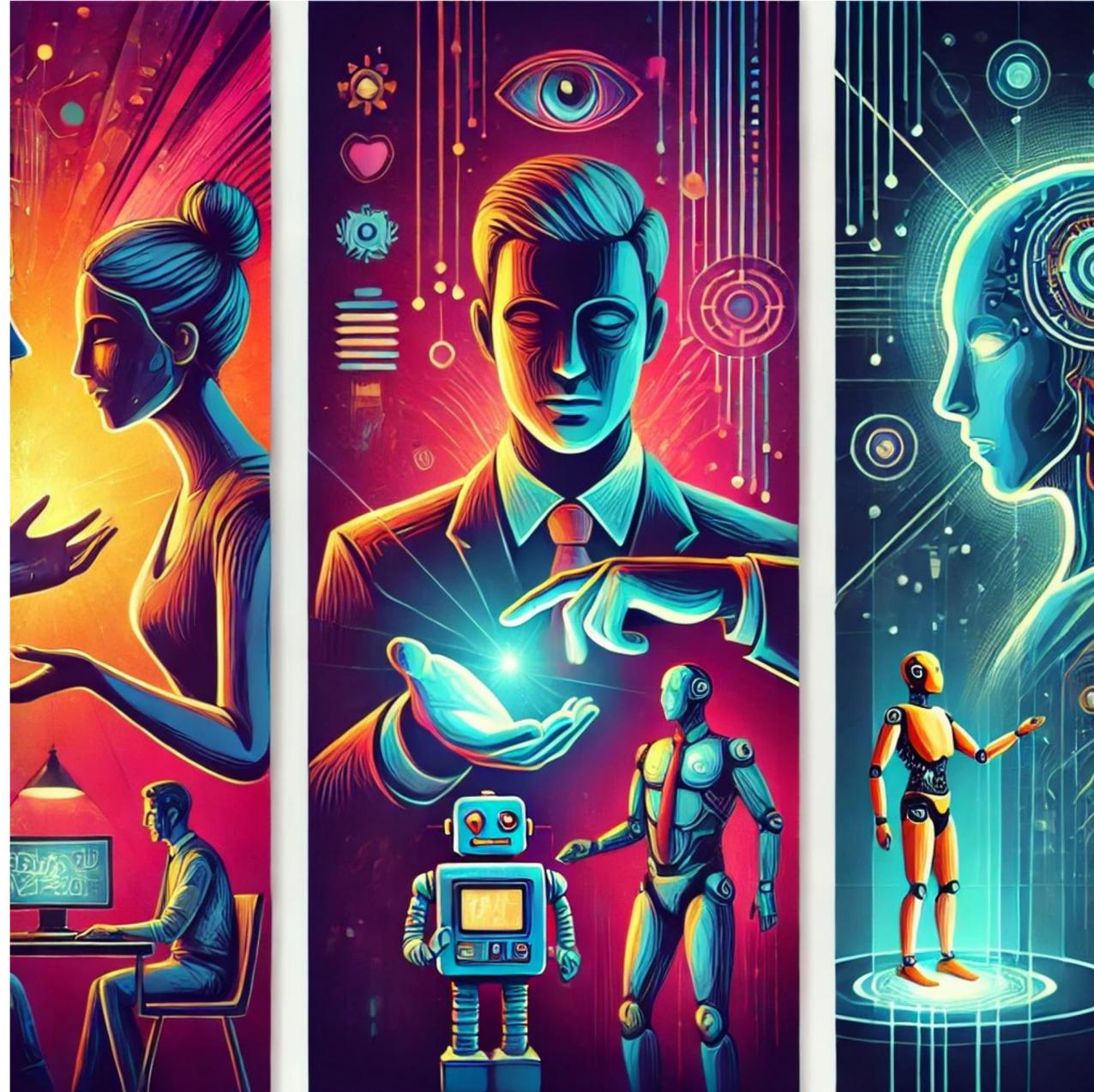
...as, personalizadas e

...sso mais inclusivo e rápido a (usuários).

...s, possibilitando

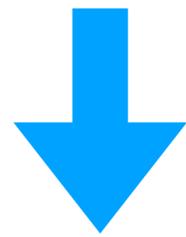
...

# Então, qual caminho seguir?

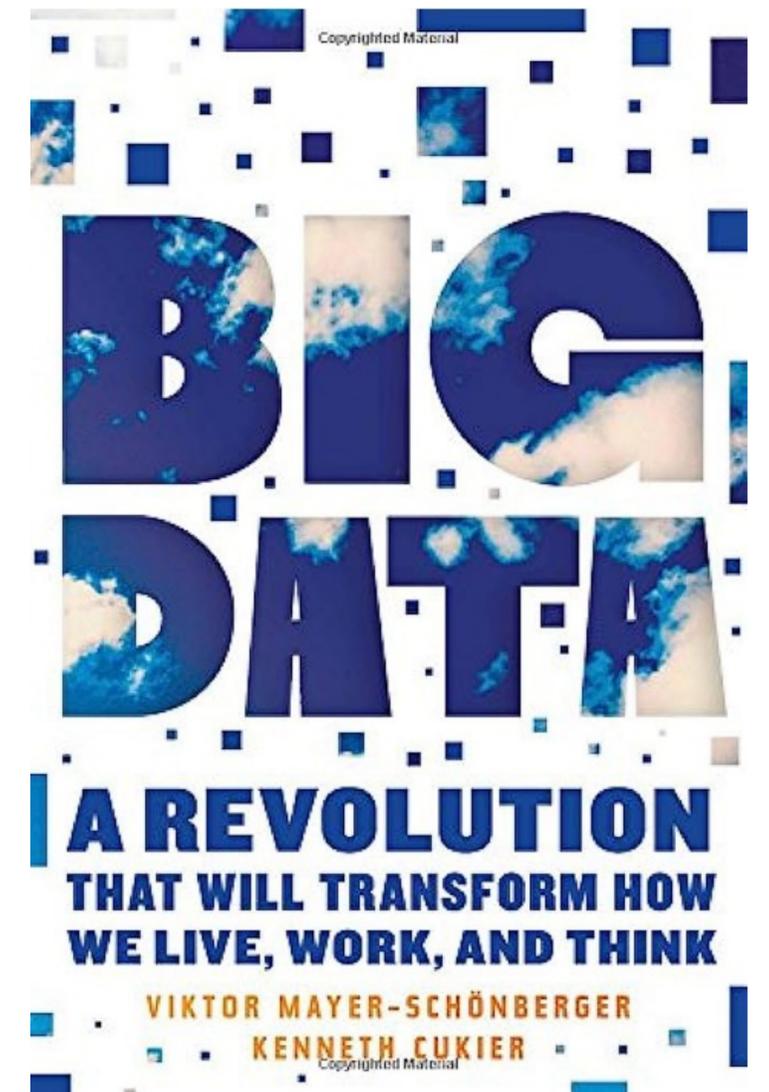


# O “paradigma” dos dados

Documentos, informação



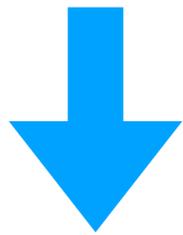
Dados —> Dataficação



# O Papel da Inteligência Artificial

Arquivos  
Bibliotecas  
Museus

Web | Redes Sociais



Dados  
(Dataficação)

- A IA engloba métodos como aprendizagem automática, processamento de linguagem natural e visão computacional.
- Possibilita a preservação digital através da automatização de tarefas como a geração de metadados, categorização /classificação e calendarização (cron jobs).
- Análise de grandes volumes de dados, permitindo estratégias de preservação mais eficientes.

# Preservação Digital

## Quais são os objetivos e fundamentos

- Processo de manutenção e garantia da longevidade da informação digital ao longo do tempo.
- Envolve estratégias como migração de formatos, emulação e redundância para combater a obsolescência tecnológica e a degradação dos dados.
- Essencial para salvaguardar património cultural, dados científicos e registos governamentais

# Preservação Digital

**Quais são os objetivos e fundamentos**

**Mas também garantir:**  
**Controle sobre as fontes**  
**Confiabilidade**  
**Recuperação eficiente**  
**Interoperabilidade...**

- Processo de manutenção e garantia da longevidade da informação digital ao longo do tempo.
- Envolve estratégias como migração de formatos, emulação e redundância para combater a obsolescência tecnológica e a degradação dos dados.
- Essencial para salvaguardar património cultural, dados científicos e registos governamentais

# IA e Preservação Digital

Um gêmeo digital é um ecossistema de subsistemas multidimensionais e interoperáveis compostos por coisas físicas no mundo real, versões digitais dessas coisas reais, conexões de dados sincronizadas entre elas e as pessoas, organizações e instituições envolvidas na criação, gestão e utilização dessas coisas.



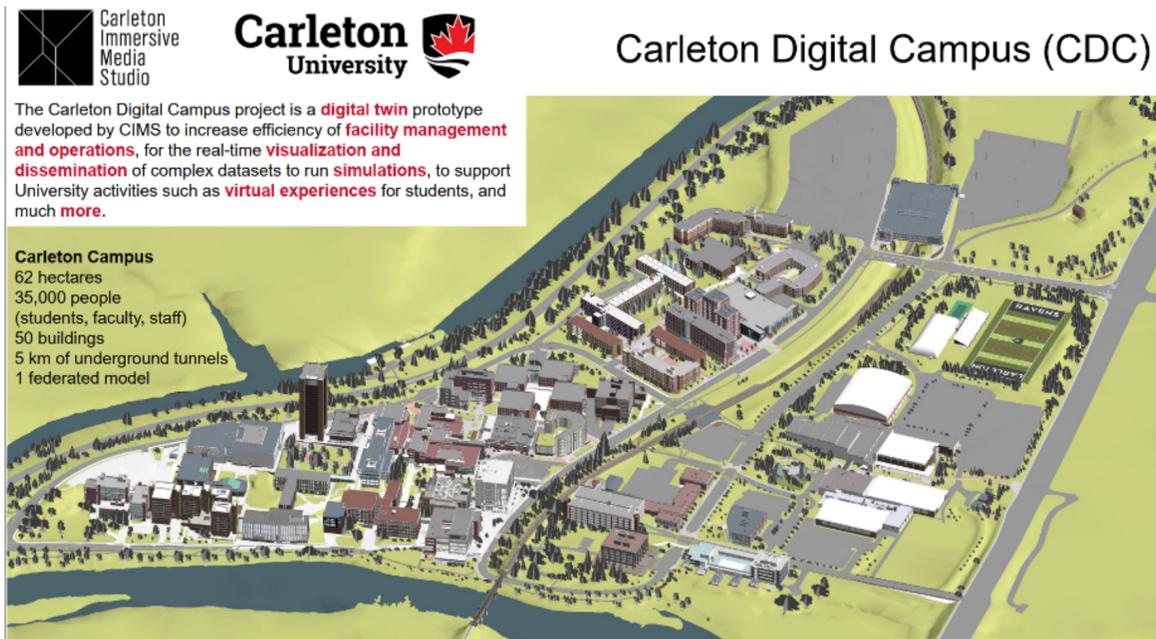
<https://canadasdigitaltwin.ca>

Lauriant, Tracey (2023)

# IA e Preservação Digital

Desenvolver uma DT (Digital Twin) que permita aos produtores e usuários de dados — independentemente de seu porte, capacidade ou contexto geográfico — gerenciar conjuntos de dados confiáveis e autoritativos, processos automatizados e sistemas técnicos em benefício de suas próprias comunidades e de acordo com suas próprias responsabilidades jurisdicionais e cultura + Interoperabilidade

Lauriant, Tracey (2023)



# IA e Preservação Digital

O processo de busca e captura da informação descrito no texto, certamente poderá sofrer aportes futuros, especialmente ao se considerar o aporte de ferramentas novas, como as de Inteligência Artificial (Apresentação, Prof. Demi Getschko)



SCAN ME

# IA e Preservação Digital



Situa-se entre a Web 3.0 (web semântica) e a web 4.0 (...em uma enorme teia de interações altamente inteligentes, parecidas com o cérebro humano - simulações)

pg.49-51



SCAN ME

# IA e Preservação Digital

Conforme Schafer e Winters (2021), a Inteligência Artificial (IA) certamente desempenhará um papel fundamental no avanço do campo, mas é importante utilizá-la de forma crítica, fortalecer a “boa governança” dos arquivos da web pode ser uma solução para lidar com esses desafios.



# IA e Preservação Digital



NotebookKIm (IA Google)

# IA e Preservação Digital



## Archives, Records and AI Literacy Framework



Arias Hernández, R., & Rockembach, M. (2025). Building trustworthy AI solutions: integrating artificial intelligence literacy into records management and archival systems. *AI & SOCIETY*, 1-18.

**InterPARES Trust<sup>AI</sup>**

Luciana Duranti and Muhammad Abdul-Mageed (coord.) - UBC

# IA e Preservação Digital



A relevância da  
proveniência e  
contexto

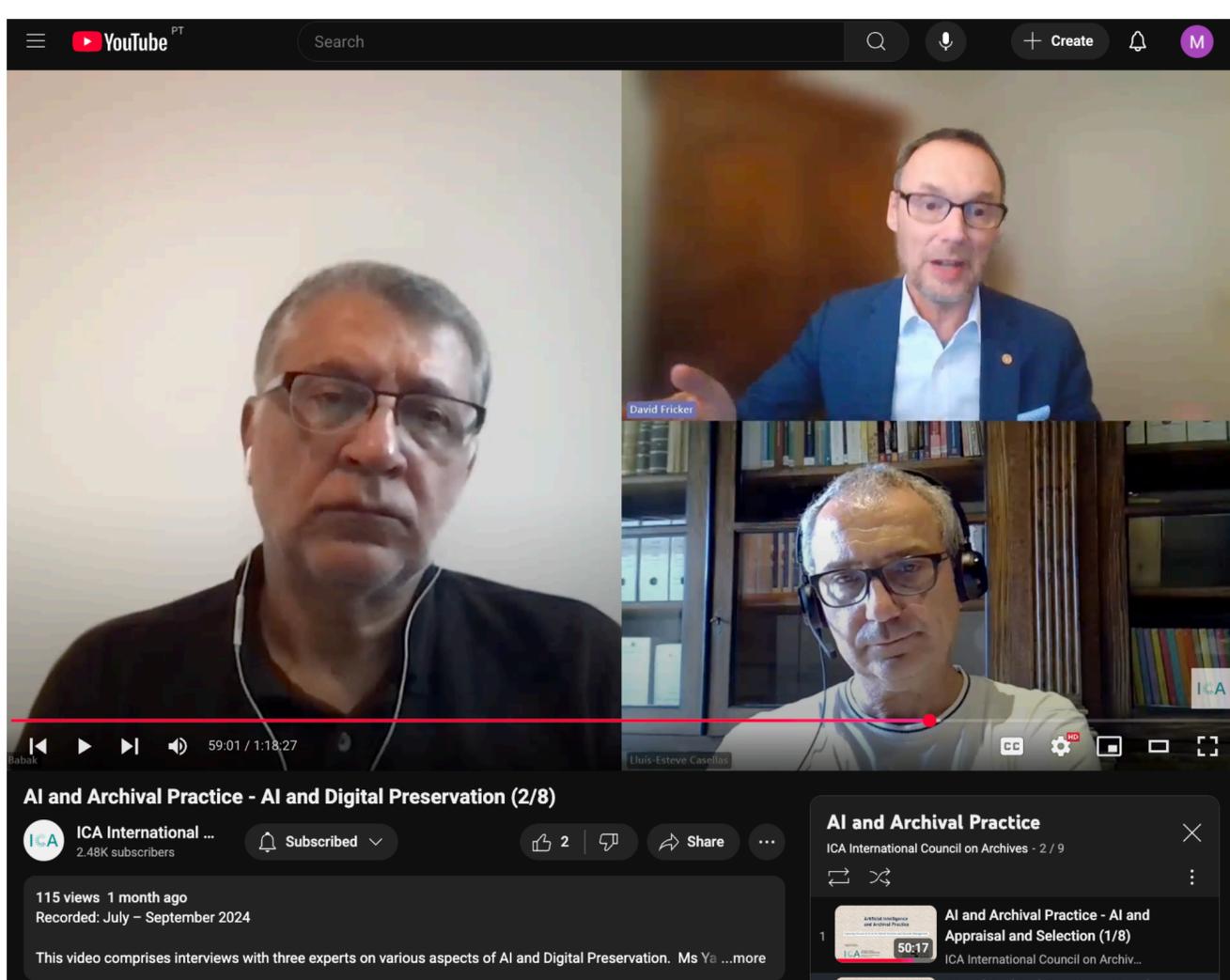
Arias Hernández, R., & Rockembach, M. (2025). Building trustworthy AI solutions: integrating artificial intelligence literacy into records management and archival systems. *AI & SOCIETY*, 1-18.

# Paradata

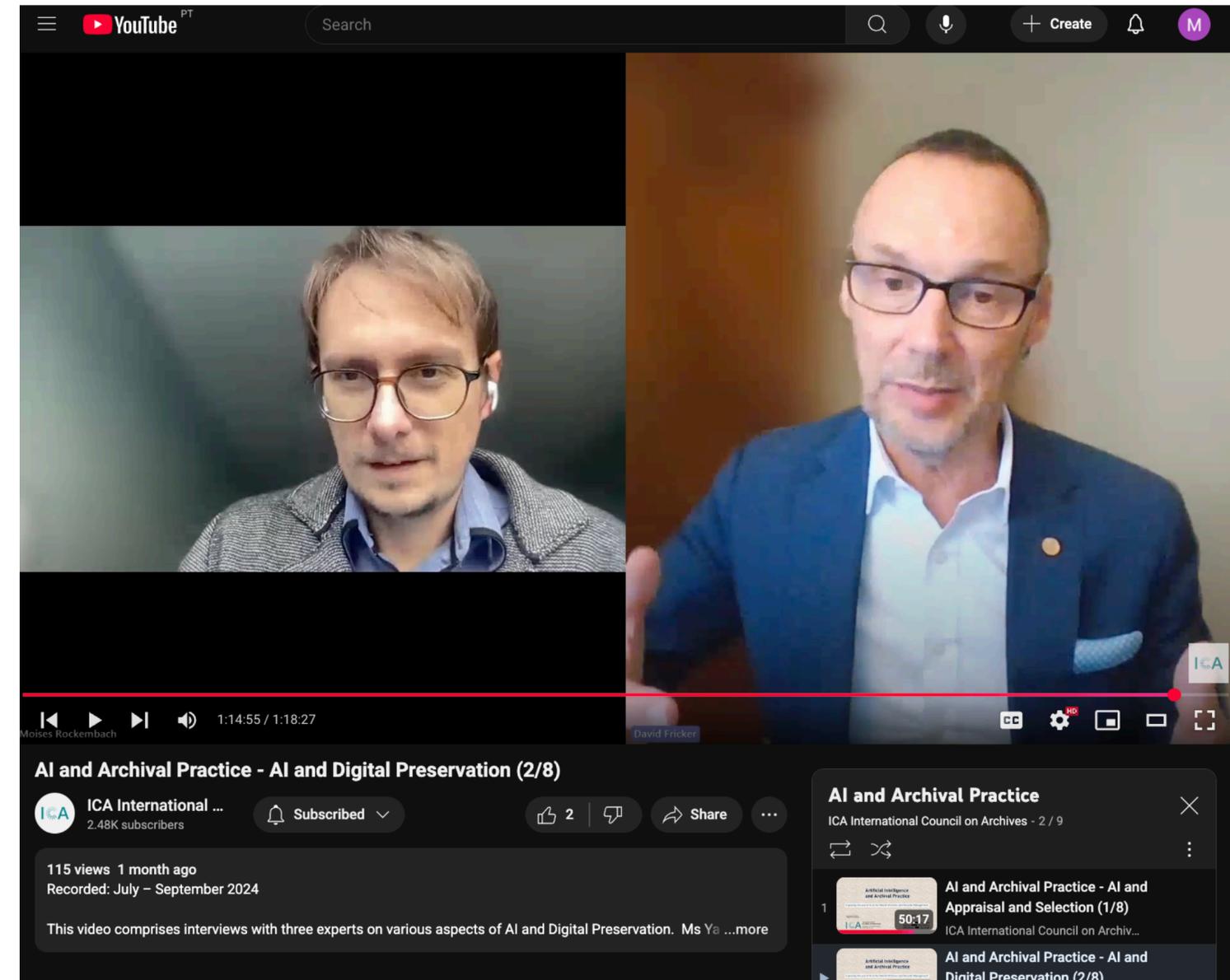
“informações sobre os procedimentos e ferramentas utilizados para criar e processar recursos de informação, juntamente com informações sobre as pessoas que os executam.” (Cameron et al., 2023; Davet et al., 2022). Embora a operacionalização deste conceito em fluxos de trabalho e procedimentos de gestão de documentos ainda esteja em andamento, o conceito em si é uma nova contribuição para o corpo de conhecimento relevante de literacia em IA para profissionais de arquivo e pode agora ser incluído como parte das competências “princípios e práticas arquivísticas” e “IA/ML de dados críticos e ética para gestão de documentos e arquivos”.

Arias Hernández, R., & Rockembach, M. (2025). Building trustworthy AI solutions: integrating artificial intelligence literacy into records management and archival systems. *AI & SOCIETY*, 1-18.

# IA e Preservação Digital



Dr Babak Hamidzadeh,  
University of Maryland  
Libraries

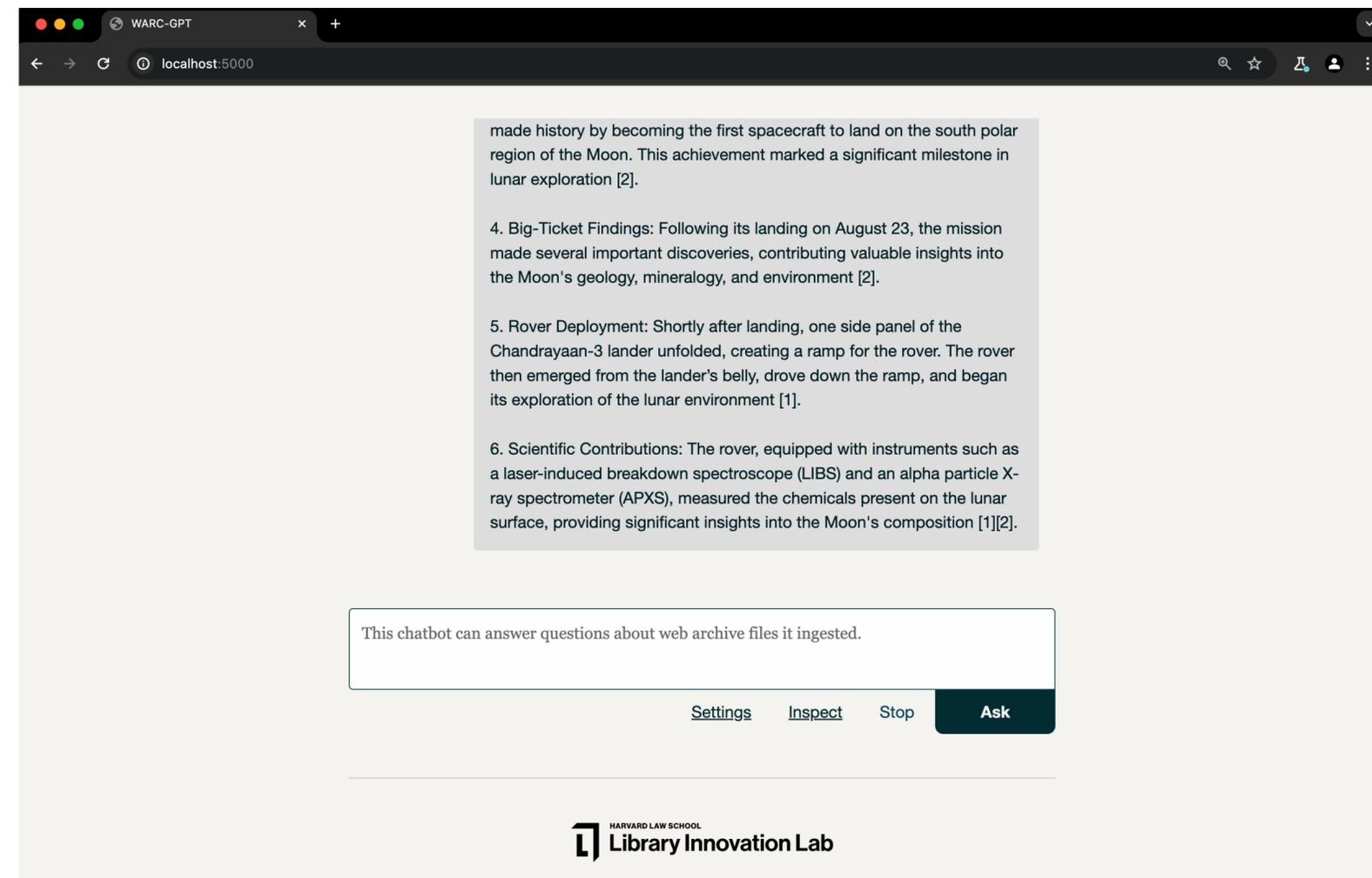
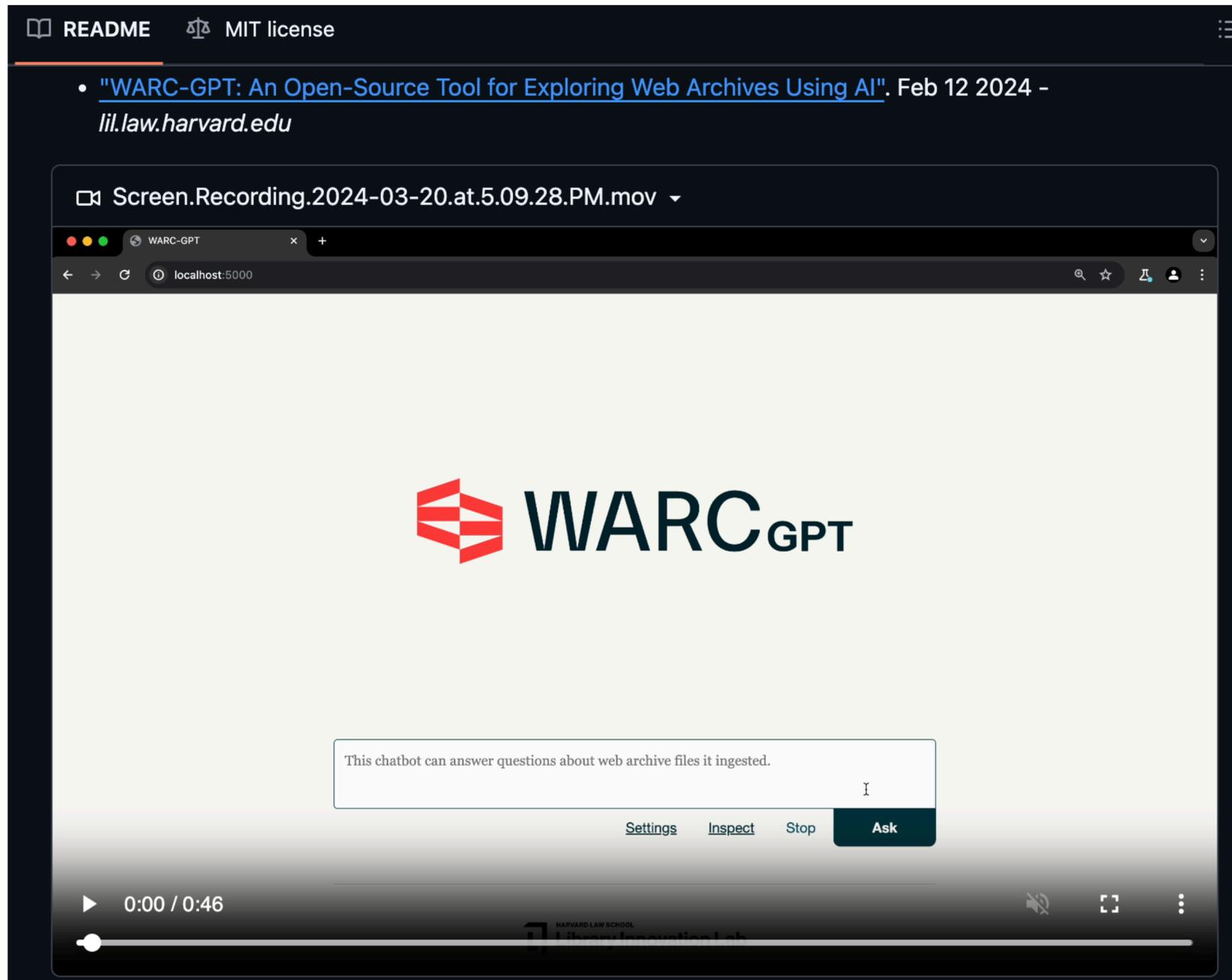


- Como decisões algorítmicas podem influenciar tomadas de decisão?
- Data provenance” e “model lineage”
- Arquivamento da IA”: pensar em como arquivar não só os produtos da IA, mas a própria IA (modelos, versões, documentação, decisões)

- Dados de treino usados (origem, qualidade);
- Algoritmos e técnicas aplicadas;
- Parâmetros e decisões de design;
- Quem criou ou treinou o modelo, e quando;
- Modificações e versões subsequentes;
- Contexto de uso e impacto previsto.

International Council on Archives  
<https://youtu.be/sSauycemlqc?si=ykNdL4mPWHvQ2LHt>

# Explorar formatos de preservação com IA

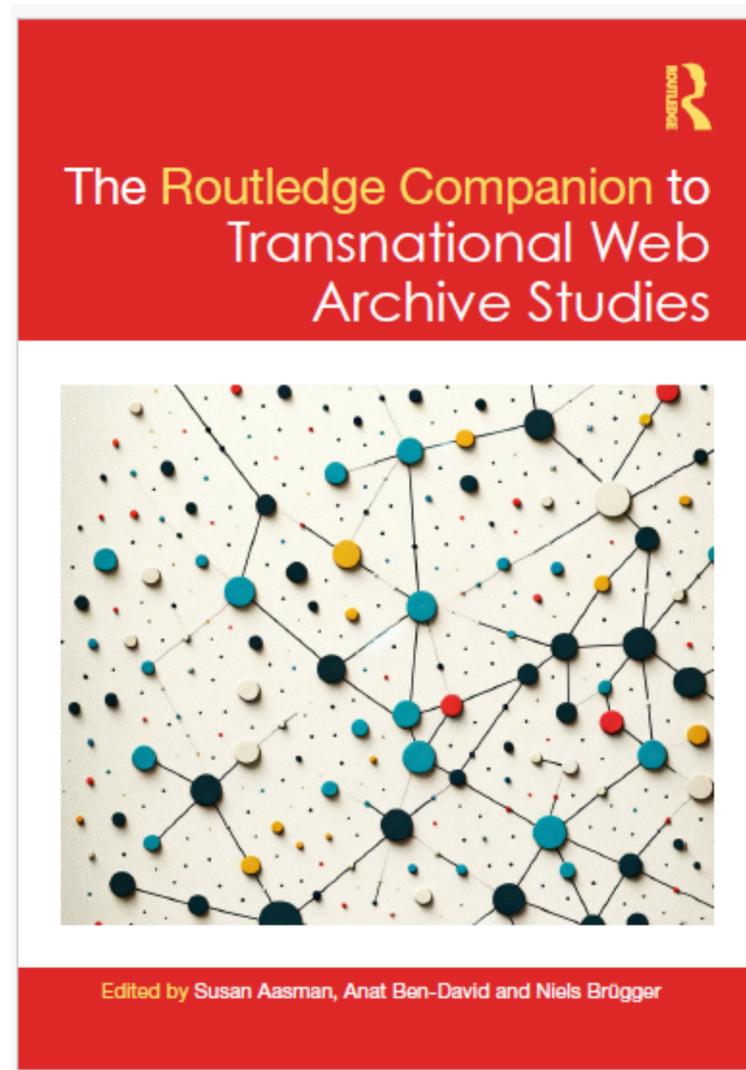


<https://github.com/harvard-lil/warc-gpt>

# Desafios e Considerações Éticas

- Garantir a autenticidade e integridade do conteúdo gerado por IA.
- Abordar os vieses nos algoritmos de IA que podem afetar os resultados da preservação.
- Equilibrar a automação com a expertise dos profissionais da informação.
- Implementar práticas transparentes e responsáveis de IA no contexto da preservação digital. A colaboração entre investigadores de IA e profissionais da informação é essencial.

# The Routledge Companion to Transnational Web Archive Studies



**Obrigado!**

[moises.rockembach@uc.pt](mailto:moises.rockembach@uc.pt)