



Inteligência Artificial e memória profunda transformando o acesso e a interpretação dos arquivos

Luis Borges Gouveia

Imbg@ufp.edu.pt, Universidade Fernando Pessoa

Inteligência Artificial e memória profunda

transformando o acesso e a interpretação dos arquivos

BAD, 11 de abril de 2025

Os arquivos são repositórios de memória, mas o acesso tradicional, muitas vezes limitado por descrições e indexações manuais, apenas “toca” a superfície do seu potencial informativo. Propõe-se a exploração de como a Inteligência Artificial pode desbloquear uma “memória profunda”, permitindo novas formas de acesso e interpretação dos acervos documentais. Técnicas como o Processamento de Linguagem Natural (NLP) e a Visão Computacional podem auxiliar na extração de significado, identificar relações e visualizar padrões em grandes massas documentais, transcendendo estratégias de pesquisa tradicionais como por palavra-chave. A IA pode facilitar a análise de sentimento, a identificação de temas emergentes e a reconstrução de contextos históricos de formas anteriormente difíceis, senão impossíveis de conceber. Contudo, esta “memória profunda” mediada por algoritmos exige cautela, levando à análise crítica do potencial de má interpretação algorítmica e à redescoberta da importância do papel do arquivista e do investigador na validação e contextualização das percepções geradas pela IA, fomentando uma hermenêutica digital (na sua procura de sentido inter e transdisciplinar) que combine o poder computacional com a sensibilidade humana.

4

Dimensões

*Para além da automação,
rumo à augmentação*

*Ética e transparência no
centro*

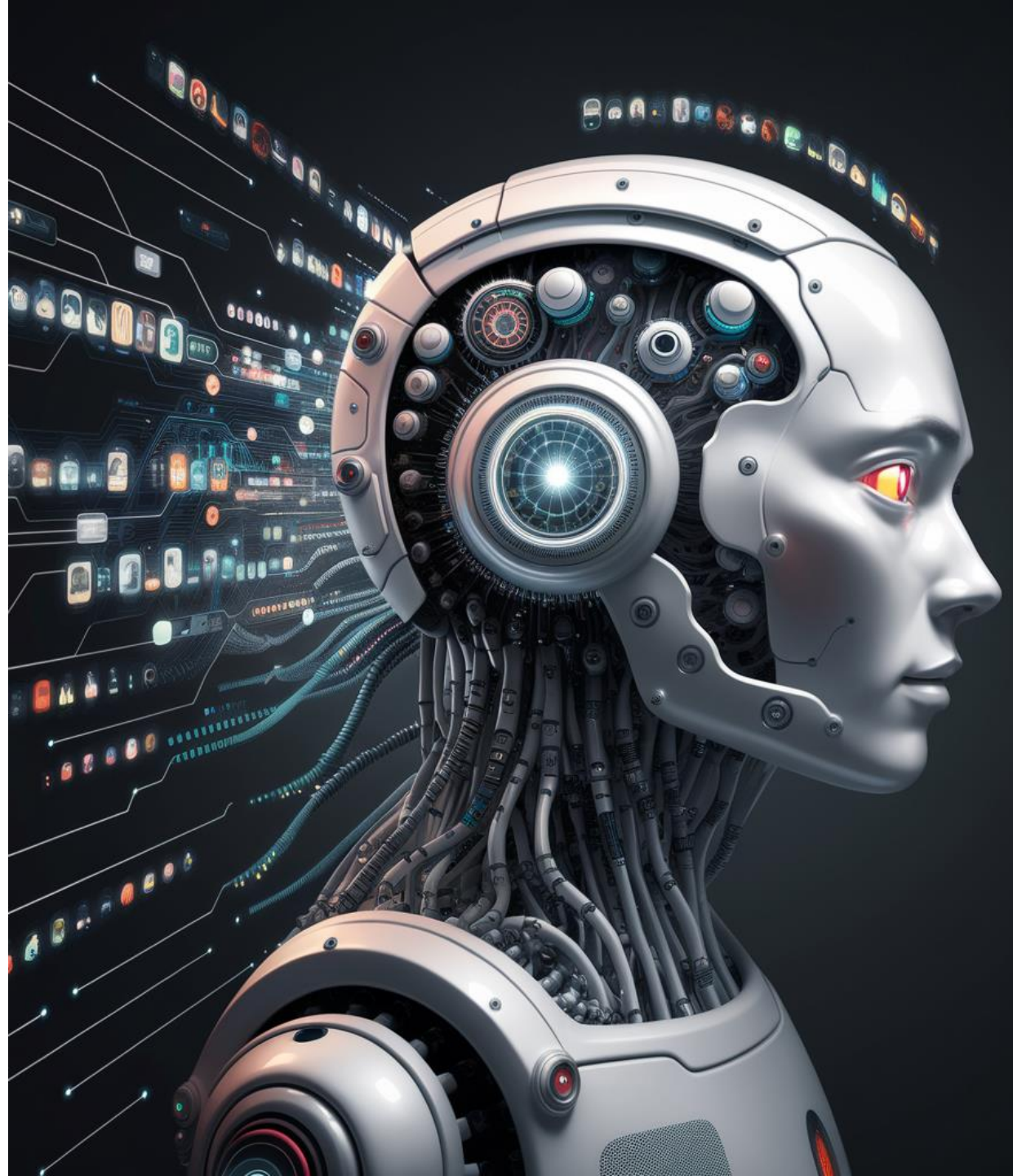
*Redefinindo acesso e
descoberta*

*O arquivo do futuro e o
futuro do arquivo*



Para além da automação, rumo à augmentação

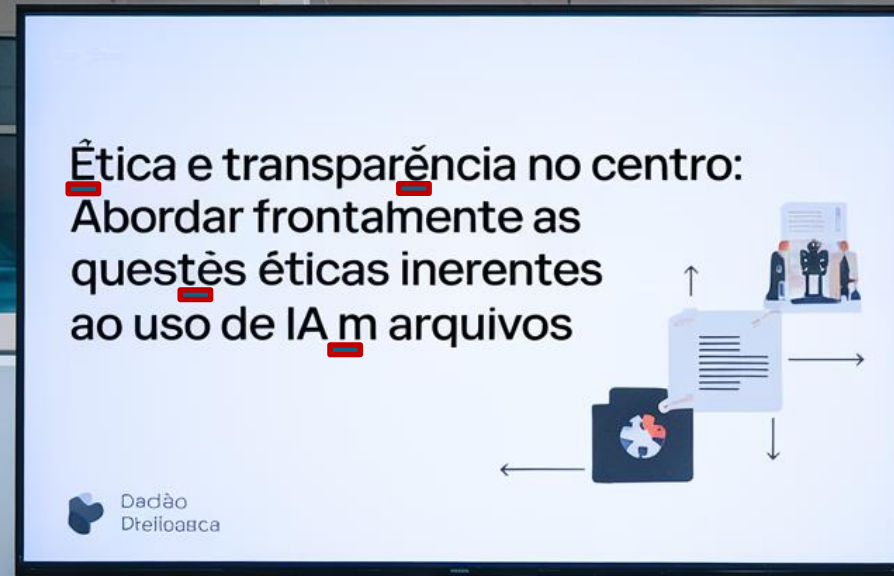
A discussão deve transcender a visão da IA como mera ferramenta de automatização de tarefas (catalogação, indexação, ...) e explorar como ela pode *aumentar* as capacidades intelectuais e de curadoria dos arquivistas, permitindo novas formas de análise, interpretação e contextualização



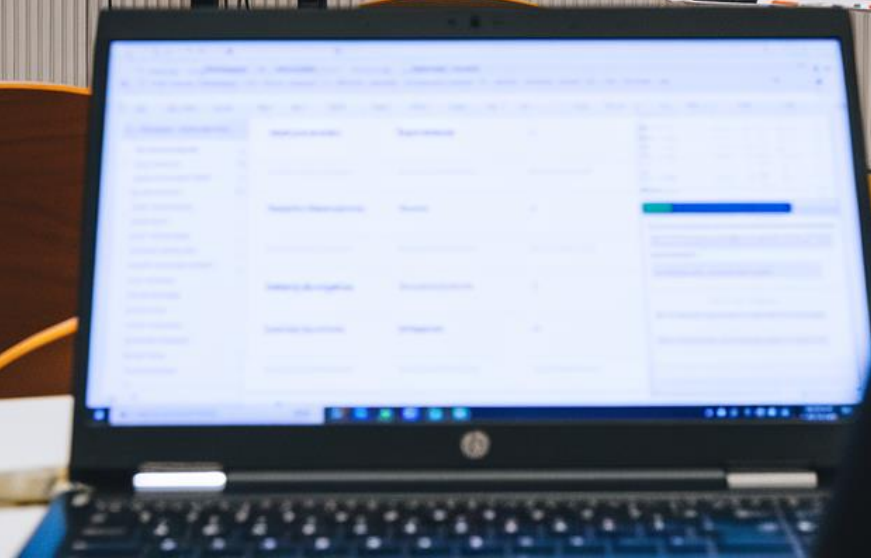
Ética e transparência no Centro

Abordar frontalmente as questões éticas inerentes ao uso de IA em arquivos – distorção algorítmica, privacidade, opacidade dos sistemas (“caixa preta”), representatividade e o potencial de perpetuar desigualdades históricas.

A inovação reside em propor modelos de IA explicável (XAI) e eticamente conscientes para o contexto arquivístico.



- Caixa Preta
- Viés Algorítmico
- Privacidade
- Representatividade
- Desigualdades históricas



Redefinindo acesso e descoberta

Explorar como a IA (NLP, visão computacional, grafos de conhecimento, ...) pode(m) revolucionar a descoberta, das pesquisas baseadas em palavras-chave para buscas semânticas, identificação de relações complexas entre documentos e personalização da experiência do utilizador, tornando os arquivos mais acessíveis a públicos diversificados.



O arquivo do futuro e o futuro do arquivo

Considerar o duplo desafio: como usar a IA para gerir e preservar os arquivos tradicionais e digitais, e como arquivar os próprios sistemas de IA, os seus dados de treino e os conteúdos por eles gerados, que constituirão o património digital futuro.



AI Meets Archives: The Future of Machine Learning in Cultural Heritage

Posted October 21, 2024 by Stacey Patton in Featured, News



<https://www.clir.org/2024/10/ai-meets-archives-the-future-of-machine-learning-in-cultural-heritage/>

Jane Winters, professor of Digital Humanities at the School of Advanced Study, University of London

*“The message for GLAMS professionals is clear: **AI is not going away**, and the stakes are high if institutions do not embrace its potential.*

*[GLAMS institutions] are responsible for **high-quality data** that could be used to develop large language models with transparency and fairness in mind. Without their input, decisions about AI’s application in cultural heritage will be made without their valuable perspectives”*

Archives and AI: An Overview of Current Debates and Future Perspectives

Authors:


Giovanni Colavizza,

Tobias Blanke,

Charles Jeurgens,

Julia Noordegraaf

[Authors Info & Claims](#)

ACM Journal on Computing and Cultural Heritage (JOCCH), Volume 15, Issue 1
•
Article No.: 4, Pages 1 - 15
•
<https://doi.org/10.1145/3479010>

Published: 14 December 2021
[Publication History](#)



*“[...] archivists increasingly becoming **custodians of larger and larger datasets and with no established approaches yet to address these new challenges**, we have analysed the growing explorations of AI technologies for archives.*

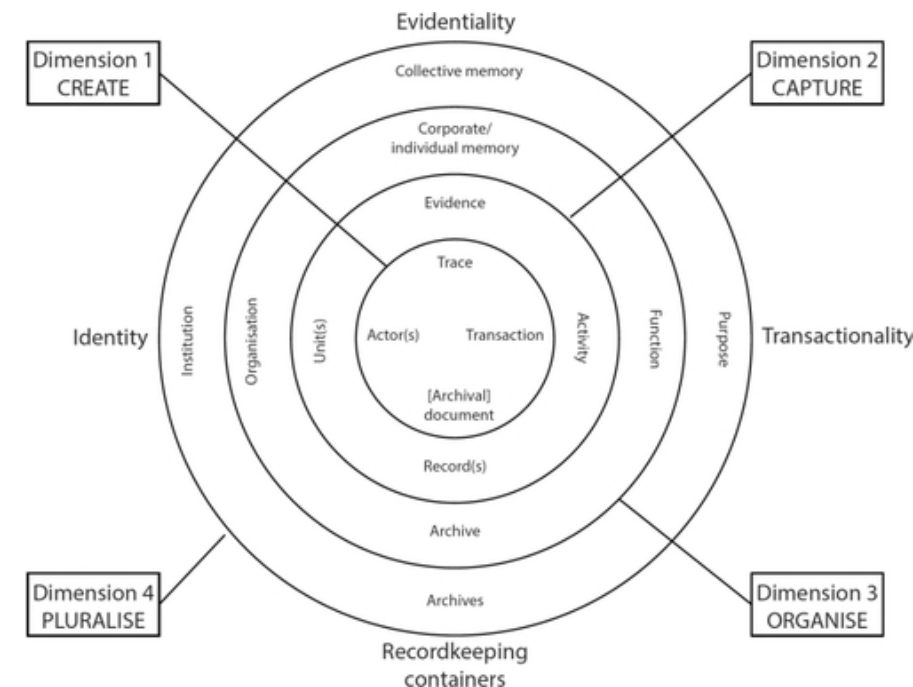
*Using the **Records Continuum model**, we have identified four distinct areas [...]. We started with the question how the integration of AI changes the theoretical foundations of archival practice. Here, we found **several areas of overlap with more general concerns in archives about digital transformations**. Beyond theoretical concerns, AI methodologies and technologies are already actively used to **automate parts of the archival workflows**, especially around the **“Capture”** and **“Organise”** dimensions of the Records Continuum model.*

*[...] **novel AI techniques** used with regards to the **“Organise”** and **“Pluralise”** Records Continuum dimensions. There is much excitement within the archival communities about the potential of AI technologies here.*

*In terms of organisation, **AI helps archives break out of the limitations of traditional archival units while also maintaining their strengths**, e.g., in terms of **provenance**.*

*AI techniques are also widely used to **gain new insights from the now-digital collections**, **pluralise** them through novel forms of retrieval and distant reading that **allow archives to connect directly to new communities** such as digital humanities.*

*[...] the role of AI technologies to open up new types of archives such as **social media collections or diaries**, which **extend our current institution-focused ideas of archives**. Because of their scale, these pose distinct challenges in terms of organisation and access to these collections that can only be solved with further AI technology development [...]*”



Record Continuum Model (Upward, 2018)



Applying AI to digital archives: trust, collaboration and shared professional ethics

Lise Jaillant , Arran Rees

Digital Scholarship in the Humanities, Volume 38, Issue 2, June 2023, Pages 571–585,

<https://doi.org/10.1093/lhc/fqac073>

<https://academic.oup.com/dsh/article/38/2/571/6832097>

Volume 38, Issue 2
June 2023

*This article has shown that the **lack of communication between civil servants, GLAM professionals and academics often leads to mistrust of other stakeholders.***

*Interviewees anticipated situations where **data would leak, reputations would be damaged, and in the worst cases, national security would be at risk.** [...] it is precisely the lack of trust in other stakeholders, and in new technologies, that was highlighted in our interviews.*

***Mistrust of technology** makes it difficult to implement AI tools. The concerns highlighted by interviewees—in terms of accountability, control, and bias—make it difficult to accelerate the development of AI applied to archives.*

*The article recommends **more cross-sector and cross-disciplinary dialogue** to surface the shared professional ethics between stakeholders. **Deeper understanding between humans is an essential step to apply AI to born-digital records.** [...] the issues of innovative technologies applied to digital archives have too often been explored within single disciplines such as archival studies or computer science. Yet, applying AI to archives requires a variety of skillsets: not only computational skills, but also domain knowledge, i.e. knowledge about the entities that are being treated by the algorithms.*

*Thinking within disciplinary boxes **hinders our collective ability to apply technology to digital records.** Digital Humanities—which is in essence an **interdisciplinary field**—has an important role to play in bridging the gap between record creators, archivists, researchers, and other users.*

Artificial intelligence and the future(s) of archival theory and practice

Archeion, 2024, 125, pp. 10-32, <https://doi.org/10.4467/26581264ARC.24.008.20572>

Geoffrey Yeo

*“Some recent commentators have suggested that the growth of artificial intelligence may require us to **reconsider the principles of provenance and original order** [...], it may lead us to rethink our ideas of what records and archives might be.*

*Artificial intelligence **makes it easy to put these new understandings into practice**. It enhances our ability to **re-organise** and **re-aggregate** the components of an archive to meet the requirements of different stakeholders; instead of a single arrangement (a so-called original order), we can have multiple arrangements simultaneously. **Provenance and context remain crucial**, and artificial intelligence can be expected to facilitate the **modelling or documentation of archival contexts in powerful ways that do not depend on physical ordering**.*

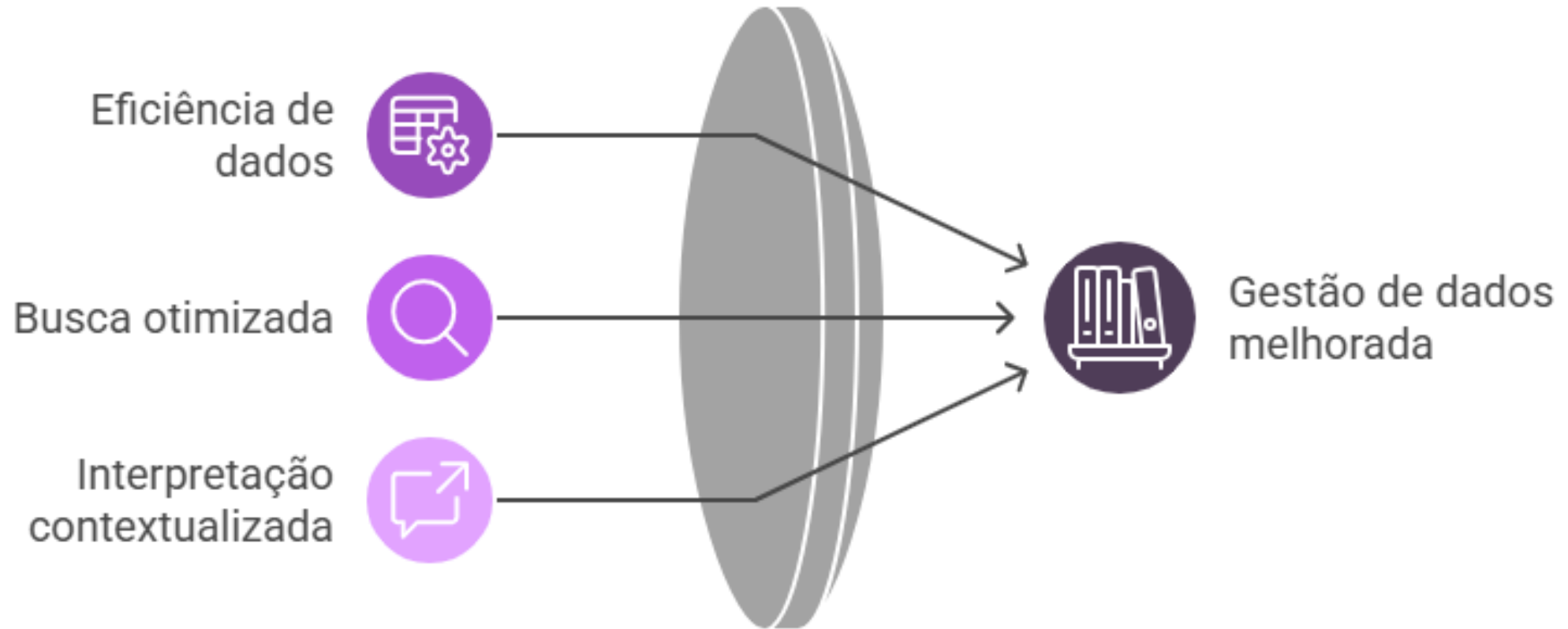
*Even in the era of artificial intelligence, **archives will still represent, and emerge from, activities and events that take place in the world**. Their **provenance and their contextual relationships will still need to be safeguarded, alongside their content and structure**. They will continue to fulfil important roles in society, because there will still be demands for the **evidence and information they can provide, the accountability they can support, and the aids to memory they can supply**.*

*We must explore ways of managing what we **capture**, perhaps using advanced artificial intelligence tools to **manage the outputs of other artificial intelligence tools**. And we must remember that we do not simply require search tools; we need **tools to support digital preservation and tools that can provide overviews of complex aggregations and uncover their interrelationships**. As yet, we do not fully understand the cultural and emotional responses that such tools will evoke.”*

Acesso e interpretação de Arquivos com IA e memória profunda



A gestão de dados com IA



Made with  Napkin

A abordagem para maximizar o potencial dos arquivos



Gestão Tradicional de Dados

Limitações no potencial
informativo

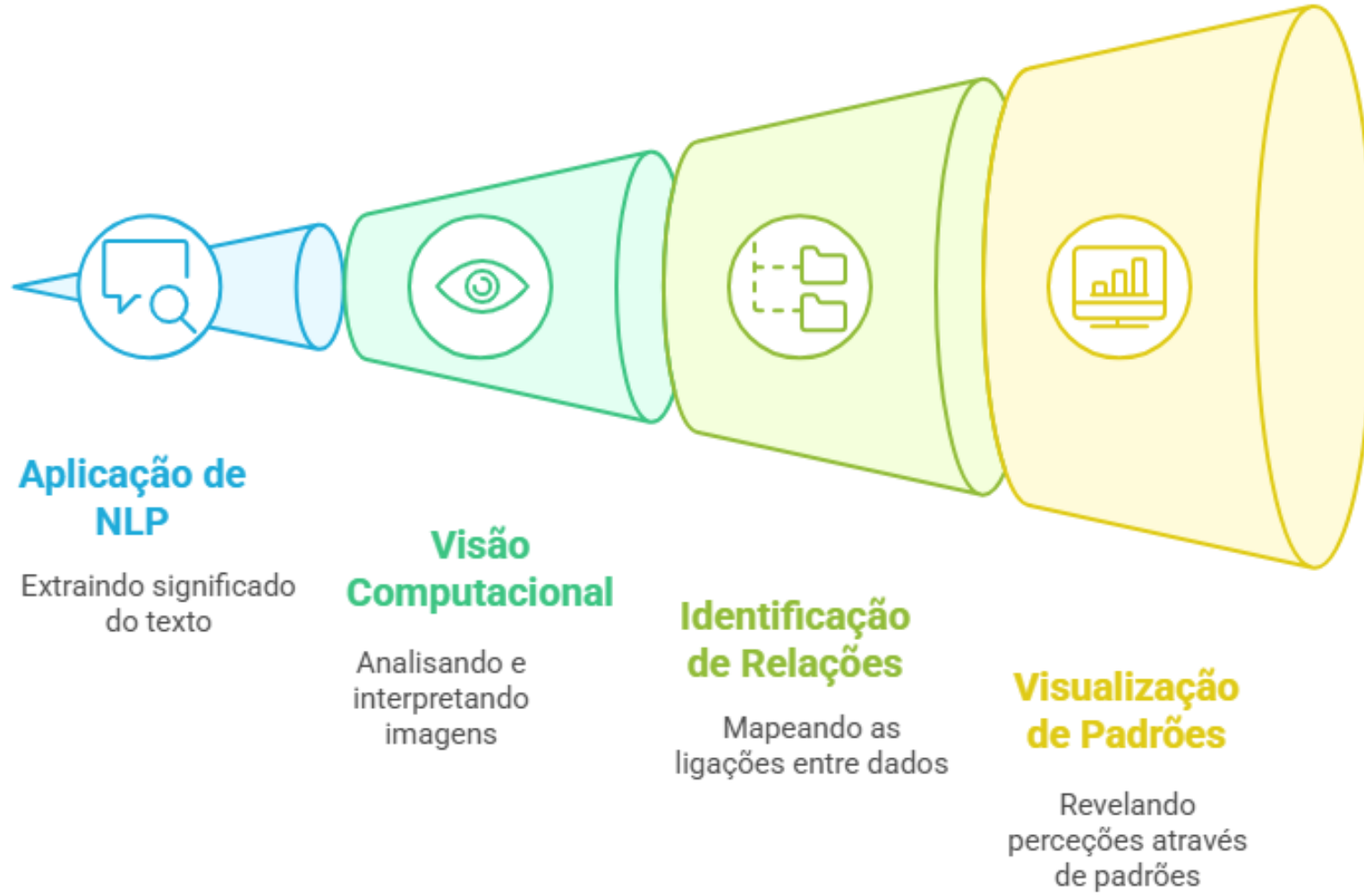


Gestão Avançada de Dados

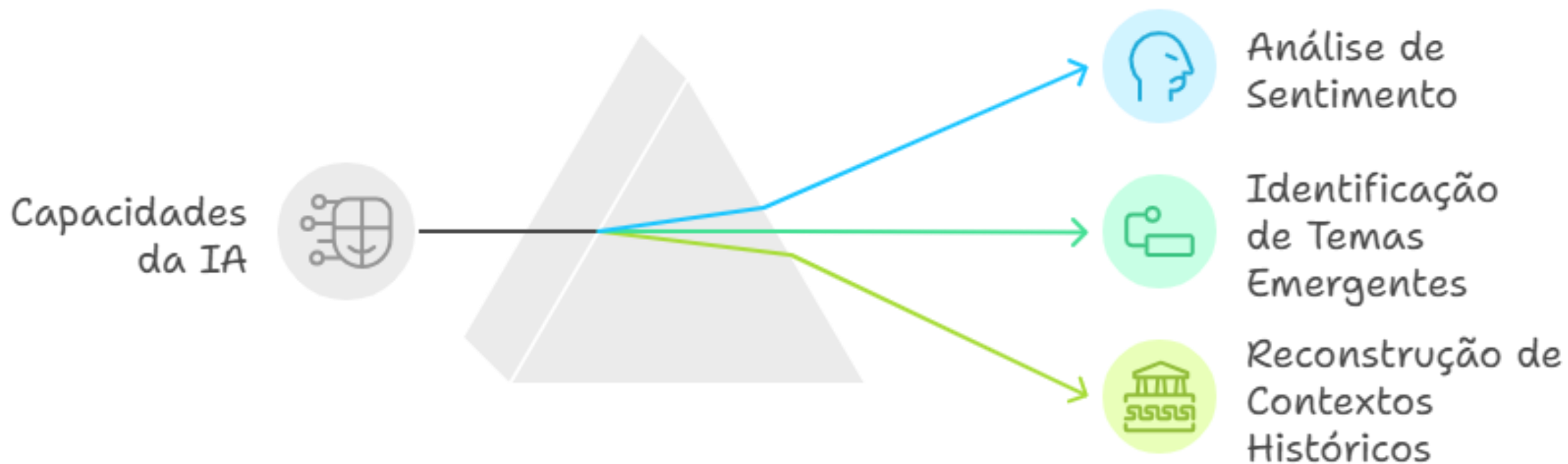
Maximiza o potencial
informativo

Made with  Napkin

Desbloqueando o Potencial dos Arquivos com IA



Capacidades da IA na Análise de Dados



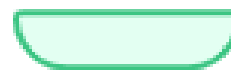
Escala na
Processamento
de Dados
Eficiência na
Análise de
Dados



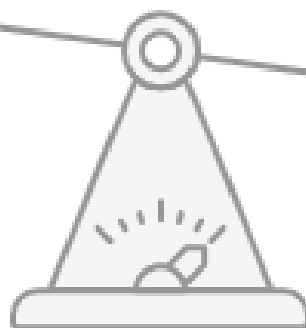
Contexto
Cultural



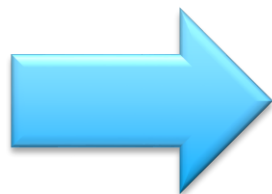
Precisão na
Validação



Algoritmos de IA

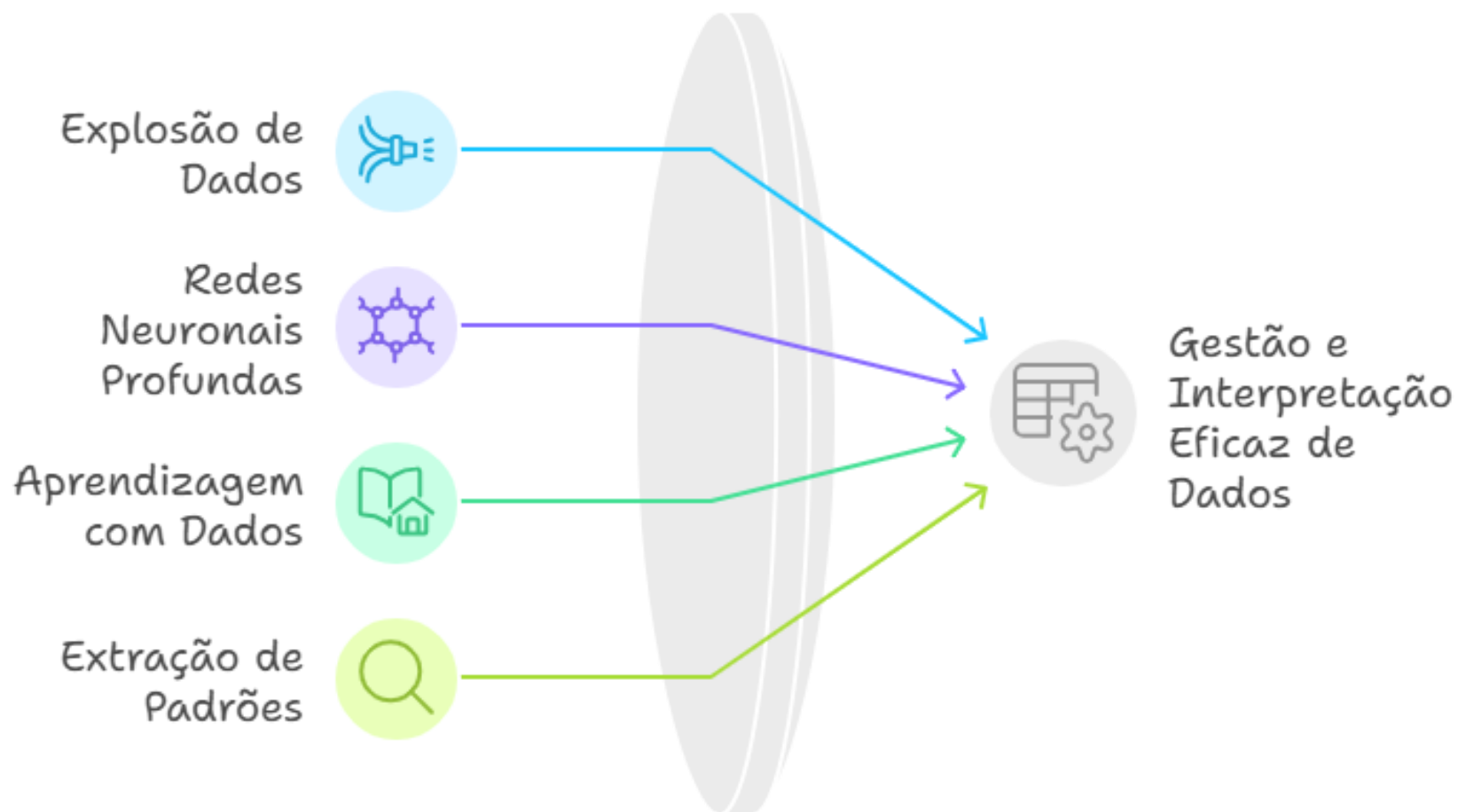


Arquivistas Humanos



Equilibrando Eficiência da IA com Precisão Humana

A IA e a compreensão de dados



IA na Gestão de Arquivos

Automação Eficiente

Agiliza a
organização de
documentos



S

Dependência de Dados

Requer entrada de
dados de qualidade



W

Experiência do Utilizador Melhorada

Melhora a
recuperação de
informação



O

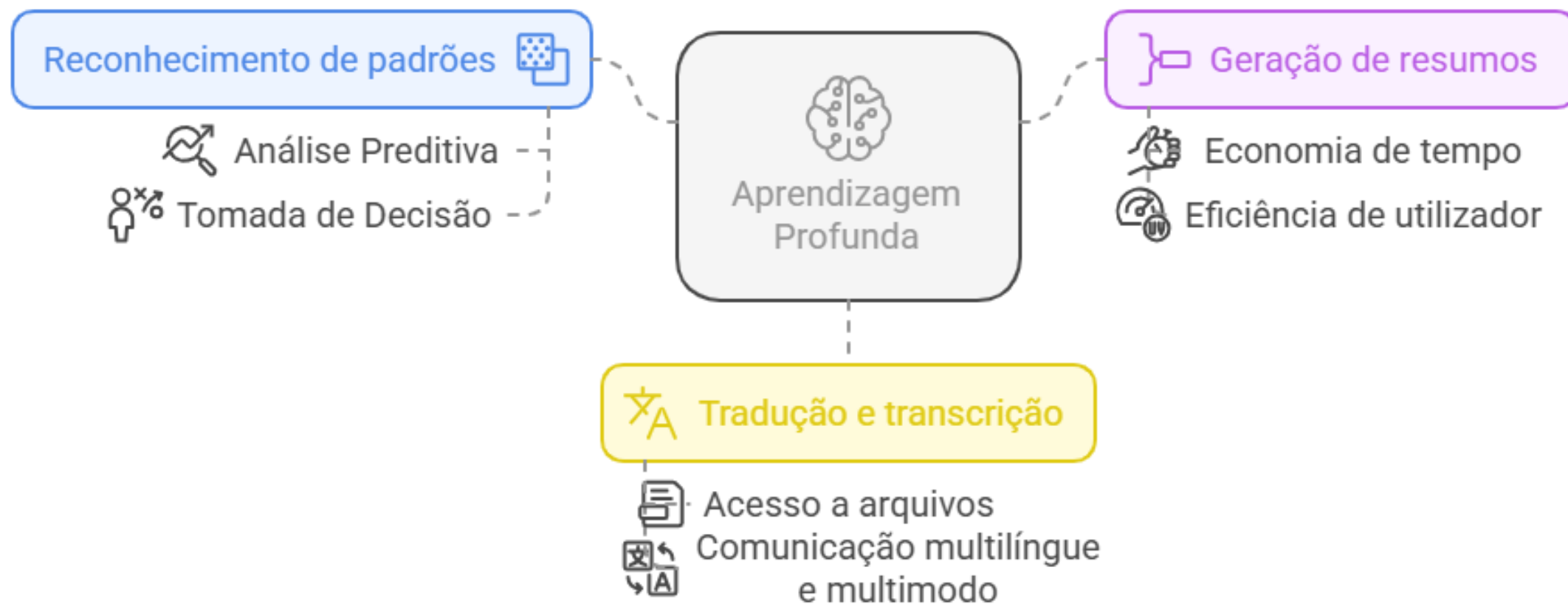
Preocupações com a Privacidade

Riscos de perda e
fuga de dados

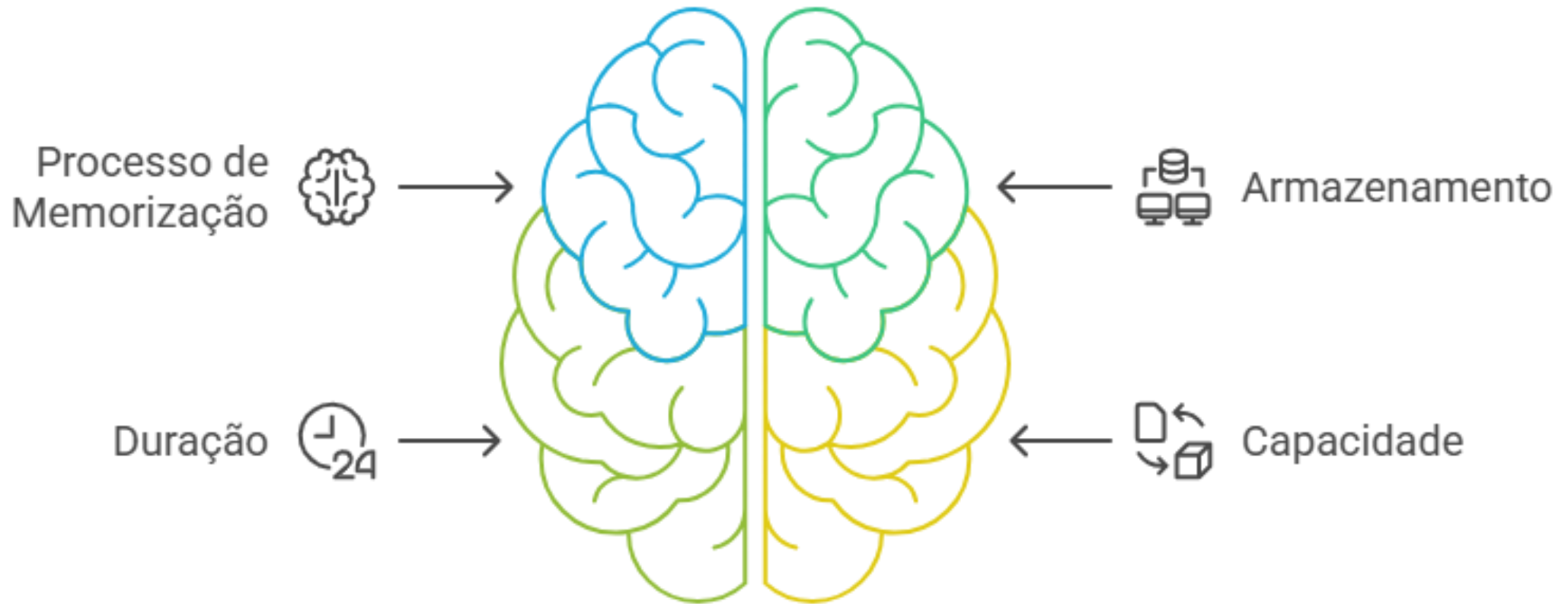


T

Aplicações da Aprendizagem Profunda

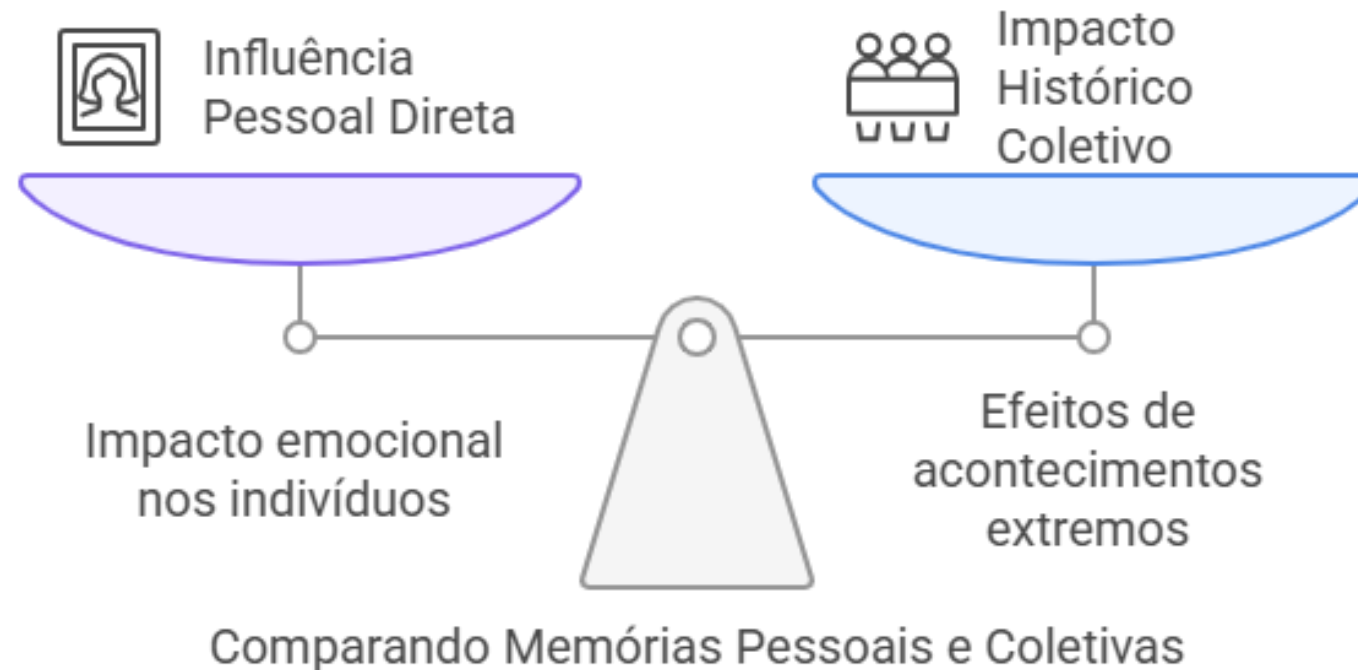


Compreendendo a Memória de Longo Prazo



A memória profunda no **contexto da história humana** (elementos que ilustram como a memória influencia experiências pessoais e coletivas ao longo da história):

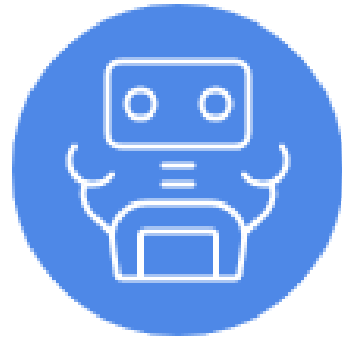
- significa o **impacto emocional duradouro de eventos**, sobre aqueles ao seu redor, destacando a sua importância na vida (*rituais*)
- **efeitos duradouros de eventos coletivos extremos**, refletindo como eventos históricos moldam **a memória e a identidade coletivas** ao longo do tempo



Desafios Éticos na Gestão de Arquivos com IA e Memória Profunda



Como equilibrar a inovação tecnológica com a responsabilidade ética na gestão de arquivos?



Inteligência Artificial

Melhora o acesso e a interpretação dos arquivos



Memória Profunda

Preserva a compreensão humana e o contexto

MANTER



STATUS QUO

ARRISCAR



A LIBERDADE

Texto e estrutura:

- Humano: Luis Borges Gouveia
- Imagens IA (geradas por Luis Borges Gouveia)*
- Ideogram (slides 4 e 5)
 - Image Bing Creator (slides 3, 6 e 7)
 - Napkin (slides: 12 a 24)
 - chatGPT 4º (slide 25)



Luis Borges Gouveia

Dip (UPT), MSc (FEUP), PhD (ULANCS), PD (FLUP) <http://homepage.ufp.pt/lmbg>

Os seus interesses estão relacionados com o digital e como o seu uso e exploração pode beneficiar indivíduos e organizações, como nas questões associadas com o uso e exploração da Inteligência Artificial. Consultor para a transformação digital.

Professor Catedrático da Universidade Fernando Pessoa (**UFP**)

<https://www.ufp.pt/>

Membro Integrado do grupo Informação, Comunicação e Cultura Digital do **CITCEM**, FLUP

<https://citcem.org/>

Colaborador do **LIACC**, Laboratório de Inteligência Artificial e Ciência de Computadores, FEUP

<https://liacc.fe.up.pt/>

Sócio e Membro da Direção da Delegação Norte da **APDSI** (ONG que promove a discussão do digital e de como promover uma sociedade mais capaz de lidar com o digital)

<https://apdsi.pt/>

