

## XXV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - XXV ENANCIB

### GT 2 – Organização e Representação do Conhecimento

#### INDEXAÇÃO TEMÁTICA DE IMAGENS COM A UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL

##### *THEMATIC INDEXING OF IMAGES THROUGH ARTIFICIAL INTELLIGENCE*

**Adriana Aparecida Lemos Torres** – Universidade Federal de Minas Gerais (UFMG)

**Benildes Coura Moreira dos Santos Maculan** – Universidade Federal de Minas Gerais (UFMG)

**Alexandre Alves da Rocha** – Universidade Federal de Minas Gerais (UFMG)

**Felipe Moreira de Assunção** – Universidade Federal de Minas Gerais (UFMG)

**Francis Bento Marques** – Universidade Federal de Minas Gerais (UFMG)

#### **Modalidade: Trabalho Completo**

**Resumo:** este estudo investiga o potencial da inteligência artificial na indexação temática de imagens, tomando como base o método iconográfico proposto por Erwin Panofsky. O objetivo é avaliar em que medida modelos computacionais são capazes de interpretar e categorizar imagens a partir dos três níveis analíticos estabelecidos pelo autor, que abrangem a descrição formal dos elementos visuais, o reconhecimento de significados simbólicos e a contextualização histórica e cultural. A pesquisa adota uma abordagem qualitativa de natureza comparativa, utilizando um corpus composto por duas imagens de caráter histórico e artístico. As imagens foram analisadas por dois especialistas humanos e por dois modelos de linguagem com capacidade multimodal. As interpretações produzidas foram comparadas por meio de critérios estruturados, organizados em uma escala ordinal simplificada com três níveis de correspondência. Os resultados indicam que os modelos computacionais demonstram bom desempenho na identificação de formas e objetos, mas apresentam limitações nos níveis mais profundos de interpretação, especialmente na leitura de símbolos e na articulação com o contexto sociocultural das imagens. Considera-se que a inteligência artificial pode atuar como ferramenta de apoio à organização da informação imagética, mas que sua aplicação plena depende da mediação crítica de especialistas humanos, especialmente em domínios que envolvem repertórios simbólicos complexos.

**Palavras-chave:** organização da informação; indexação de imagens; inteligência artificial; método iconográfico; Erwin Panofsky.

**Abstract:** this study investigates the potential of artificial intelligence in the thematic indexing of images, based on the iconographic method proposed by Erwin Panofsky. The objective is to assess the extent to which computational models are capable of interpreting and categorizing images using the three analytical levels defined by the author, which involve the formal description of visual elements, the recognition of symbolic meanings, and the historical and cultural contextualization. The research adopts a qualitative and comparative approach, using a corpus composed of two historically and artistically significant images. These images were analyzed by two human experts and two multimodal language models. The interpretations generated were compared using structured criteria, organized into a simplified three-point ordinal scale. The results indicate that computational models perform well in identifying forms and objects, but show limitations in the deeper levels of interpretation, particularly in the recognition of symbols and in establishing connections with the sociocultural context of the

images. It is considered that artificial intelligence can serve as a support tool for the organization of visual information, but its full application depends on the critical mediation of human experts, especially in domains involving complex symbolic repertoires.

**Keywords:** information organization; image indexing; artificial intelligence; iconographic method; Erwin Panofsky.

## **1 INTRODUÇÃO**

A indexação de imagens difere-se substancialmente da indexação de textos, considerando sua natureza e linguagem específicas (Smit, 1996). Para uma eficiente representação de imagens, são necessários a elaboração de políticas, a adoção de metodologias próprias, a utilização de vocabulário controlado, a qualificação de profissionais, dentre outros aspectos.

No contexto de desenvolvimento e utilização de inteligência artificial (IA) para diversas finalidades, a Ciência da Informação coloca-se diante de oportunidades e desafios. Nesse sentido, esse estudo propôs verificar a capacidade de ferramentas de inteligências artificiais em identificar, analisar e interpretar imagens a partir da utilização do método iconográfico de Erwin Panofsky (1979). Tal método, embora oriundo do contexto da História da Arte, é muito utilizado na Ciência da Informação para a indexação de diferentes tipologias de imagens e em contextos variados, além de constituir-se base para a elaboração de outras metodologias de indexação de imagens (Layne, 1986; Manini, 2002; Smit, 1996; Torres, 2019).

Os resultados obtidos na utilização das ferramentas de IA, devidamente treinadas e ajustadas para a realização da indexação temática de imagens com o método iconográfico, foram analisados considerando os princípios estabelecidos na literatura e comparados à indexação feita por profissionais da informação, com vistas a oferecer subsídios teóricos e práticos para o aprimoramento dessa atividade relevante.

Para tanto, depois desta introdução, este artigo se organiza da seguinte forma: a seção 2 apresenta o referencial teórico sobre representação temática de imagens com o uso do método de Panofsky e sobre o uso de inteligência artificial na indexação de imagens; a seção 3 apresenta a metodologia; a seção 4 expõe os resultados e as análises; e na seção 5 finaliza-se com as considerações finais.

## **2 REFERENCIAL TEÓRICO**

Esta seção apresenta os fundamentos teóricos que embasam este estudo, discorrendo sobre a representação temática de imagens a partir do método iconográfico proposto por Panofsky (1979) e a utilização da inteligência artificial nos processos de indexação temática de imagens. A articulação desses referenciais possibilita uma compreensão aprofundada dos desafios inerentes à interpretação imagética, ao mesmo tempo em que permite avaliar criticamente o potencial das tecnologias baseadas em IA como suporte a essa atividade de elevada complexidade.

### **2.1 A indexação de imagens com o método iconográfico de Panofsky**

Possibilitar o acesso ao conhecimento contido na informação é o objetivo do processo de organização da informação, que engloba a descrição física e temática dos documentos, sendo a representação da informação um produto desse processo descritivo (Bräscher; Café, 2008). A representação da informação visa à atribuição de descritores ou rótulos temáticos a distintas tipologias documentais, tais como livros, artigos, vídeos e, de modo particular, imagens, estas últimas representando um desafio específico devido à sua natureza eminentemente subjetiva.

Metodologias diversas para a representação temática e descritiva de imagens são propostas e tratadas na literatura. Dentre elas, tem destaque o método iconográfico proposto por Erwin Panofsky, historiador de arte, que, em sua obra de 1979, sistematizou um modelo para a leitura e análise de obras de arte do Renascimento, mas vem sendo utilizado na representação de diversas tipologias de imagens no âmbito da Ciência da Informação. Tal método de representação temática de imagens estrutura-se em três níveis inter-relacionados e hierarquicamente organizados: pré-iconográfico ou primário, iconográfico ou secundário (de identificação) e iconológico (de compreensão dos significados intrínsecos) (DIRECCIÓN..., [202-]; Torres, 2019), os quais possibilitam uma compreensão aprofundada das imagens analisadas.

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB  
Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

**Quadro 1 – Níveis de análise de Panofsky**

<b>NÍVEIS</b>	<b>ABORDAGEM</b>	<b>APLICAÇÃO À ICONOGRAFIA</b>	<b>APLICAÇÃO À FOTOGRAFIA</b>
<b>PRÉ-ICONOGRÁFICO</b>	<b>TEMA</b> Tema natural ou primário (factual e expressivo).	<b>DESCRIÇÃO</b> Possibilita a enumeração dos motivos artísticos do mundo das formas puras, portadoras de significados.	Descrição dos elementos constitutivos da imagem, o REFERENTE da imagem.
<b>ICONOGRÁFICO</b>	<b>MENSAGEM</b> Tema secundário ou convencional (o motivo artístico).	<b>ANÁLISE</b> Faz-se a ligação entre os motivos e as composições com assuntos e conceitos e descrevem-se os assuntos específicos ou conceitos manifestados em imagens, histórias e alegorias.	Reconhecimento do significado atribuível ao referente (ou o MOTIVO fotografado).
<b>ICONOLÓGICO</b>	<b>SIGNIFICADO</b> Valores simbólicos, significados intrínsecos ou conteúdos somente detectáveis e / ou observáveis cultural, social, filosófica ou ideologicamente.	<b>INTERPRETAÇÃO</b> Requer algo mais que a familiaridade com conceitos ou temas específicos transmitidos através de fontes literárias.	<b>INTERPRETAÇÃO</b> da imagem.

**Fonte:** Torres e Silva (2017, p. 7).

O nível pré-iconográfico ou primário é a categoria mais básica de descrição e corresponde a uma leitura inicial, de caráter descritivo e objetivo, centrada nos elementos visuais básicos da composição, sem recorrer a interpretações culturais ou simbólicas. Nesse nível, identificam-se formas visuais elementares, como linhas, cores, volumes e massas, além dos objetos representados (como figuras humanas, animais, plantas, elementos naturais e artefatos), suas inter-relações (fatos) e qualidades expressivas (como gestos, posturas, atmosferas e espaço). Neste nível, descreve-se de forma genérica o que se vê na imagem (objetos e ações), sem contextualização de sua criação e/ou interpretação. Por exemplo, em uma análise da representação da "Última Ceia", pode-se descrever a cena como composta por treze figuras masculinas dispostas ao redor de uma mesa.

O segundo nível, denominado iconográfico ou secundário, busca identificar os conteúdos temáticos da imagem, exigindo o reconhecimento de conhecimentos prévios, advindos de repertórios culturais, históricos, literários e artísticos. Nessa fase, os elementos visuais são analisados quanto aos seus motivos iconográficos (temas ou conceitos), cuja combinação permite determinar o tema central da imagem, que pode assumir a forma de

narrativas, alegorias ou representação simbólica. Esse nível depende do nível anterior, uma vez que pressupõe a identificação prévia do tema representado na imagem, e demanda um grau mais elevado de familiaridade com conceitos específicos e com o repertório de fontes literárias, históricas e artísticas pertinentes. No exemplo anteriormente citado, os elementos identificados permitem reconhecer a cena como a "Última Ceia", episódio amplamente representado na tradição iconográfica cristã.

Por fim, o nível iconológico refere-se à interpretação mais profunda e contextual da imagem, com ênfase nos significados intrínsecos e nos valores simbólicos subjacentes à obra. Essa etapa demanda um conhecimento abrangente do contexto histórico, social, filosófico e religioso no qual a imagem foi produzida. A análise iconológica visa desvendar pressupostos culturais e ideológicos incorporados à representação visual, permitindo uma leitura crítica e contextualizada. Assim, no caso da "Última Ceia" de Leonardo da Vinci, é possível interpretá-la à luz do pensamento renascentista, das técnicas de representação espacial empregadas e dos valores sociais e religiosos prevalentes à época de sua produção.

## **2.2 Aplicação de Inteligência Artificial na Indexação Temática de Imagens**

A indexação automática de imagens tem sido consolidada como alternativa estratégica frente às limitações dos processos manuais tradicionais, os quais são marcados por alto custo operacional, subjetividade e baixa escalabilidade. Nesse contexto, a indexação tem sido realizada por meio de sistemas computacionais capazes de atribuir descritores temáticos de forma mais padronizada e eficiente (Nascimento; Martins; Albuquerque, 2023). Com o avanço das técnicas de inteligência artificial, especialmente aquelas baseadas em aprendizado de máquina e processamento de linguagem natural (PLN), tem-se observado um aprofundamento dessa abordagem, ampliando-se o alcance e a precisão dos resultados obtidos (Deng; Lin, 2022).

Por meio dessas tecnologias, a extração de informações visuais tem sido automatizada com maior grau de sofisticação, permitindo que descritores temáticos sejam atribuídos com base em padrões reconhecidos em grande volume de dados. A subjetividade inerente à análise humana tem sido, em parte, mitigada pela padronização promovida por tais modelos, embora ainda persistam desafios relacionados à interpretação de nuances semânticas, simbolismos culturais e variações contextuais (Devlin *et al.*, 2018; Kocón *et al.*, 2023).

Tem-se verificado que a qualidade dos dados de treinamento, a clareza dos critérios semânticos empregados e a robustez dos modelos utilizados exercem influência direta sobre a efetividade dos processos de indexação automática. Embora resultados promissores tenham sido apresentados, a atuação humana continua a ser requerida para a curadoria, supervisão e validação das inferências produzidas, especialmente em domínios nos quais a polissemia visual e os conteúdos simbólicos ocupam papel central.

Ferramentas baseadas em modelos generativos multimodais, como ChatGPT-4o<sup>1</sup>, Claude 2 e Gemini 1.5 Flash<sup>2</sup>, têm sido exploradas na atribuição automatizada de descrições estruturadas a imagens. Nesses casos, a geração de metadados tem sido realizada com base em instruções linguísticas específicas, possibilitando aplicações em contextos como a representação iconográfica de obras de arte, o tratamento de arquivos fotográficos e a organização de acervos visuais. Ainda assim, limitações têm sido observadas, sobretudo no que se refere à compreensão de representações culturalmente complexas e à interpretação iconológica de imagens (Dendrites.io, 2024; Devlin et al., 2018).

### **3 METODOLOGIA**

Esta pesquisa adota uma abordagem qualitativa e comparativa, centrada na análise da capacidade de modelos contemporâneos de inteligência artificial em realizar a indexação temática de imagens com base no método iconográfico proposto por Erwin Panofsky. Busca-se compreender em que medida esses modelos são capazes de reproduzir ou se aproximar da interpretação humana ao lidar com três níveis distintos de análise: o pré-iconográfico, o iconográfico e o iconológico. A investigação parte do pressuposto de que a interpretação de imagens requer não apenas o reconhecimento de formas visuais, mas também a atribuição de significados contextuais e simbólicos, o que impõe um desafio substancial às tecnologias automatizadas de processamento de linguagem e imagem.

A estrutura metodológica foi concebida para possibilitar uma comparação sistemática entre a indexação realizada por especialistas humanos e por modelos multimodais de inteligência artificial. A escolha desses modelos se fundamenta em sua capacidade de

---

<sup>1</sup> OPENAI. ChatGPT [*inteligência artificial*]. Disponível em: <https://openai.com/>. Acesso em: 28 mar. 2025.

<sup>2</sup> GOOGLE. Gemini [*inteligência artificial*]. Disponível em: <https://deepmind.google/technologies/gemini/>. Acesso em: 28 mar. 2025.

interpretar simultaneamente entradas visuais e textuais, característica essencial para a aplicação da metodologia panofskyana em ambiente automatizado.

Em lugar de métricas estatísticas tradicionais, como o coeficiente de concordância de Cohen (1960) ou medidas vetoriais de similaridade, foi adotada uma escala ordinal interpretativa de três pontos, desenvolvida especificamente para esta pesquisa. Essa escala tem por finalidade capturar o grau de correspondência entre as interpretações produzidas pelos modelos de IA e aquelas fornecidas pelos especialistas, de forma sensível às sutilezas semânticas envolvidas em cada nível de análise. Os valores atribuídos expressam, respectivamente, uma correspondência satisfatória (+1), uma correspondência parcial ou ambígua (0) e uma ausência de correspondência ou erro interpretativo claro (-1). A avaliação foi conduzida de forma criteriosa, considerando os aspectos específicos de cada nível da abordagem panofskyana, desde a identificação formal dos elementos visuais (nível pré-iconográfico), passando pelo reconhecimento de símbolos e convenções culturais (nível iconográfico), até a atribuição de significados profundos inseridos em contextos históricos, sociais ou filosóficos (nível iconológico).

As análises foram organizadas por meio de uma ficha descritiva padronizada, conforme descrito no Quadro 2, elaborada para registrar as respostas produzidas por humanos e máquinas de maneira comparável.

**Quadro 2** - Ficha de análise das indexações realizadas

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

IMAGEM X		
Comparação por nível panofskyano		
<b>1. Nível Pré-Iconográfico (descrição formal dos elementos visuais)</b>		
Critério	Avaliação (-1, 0, +1)	Comentário
A IA identificou corretamente os elementos visuais?		
Houve correspondência com a descrição feita pelos especialistas?		
Termos atribuídos foram pertinentes e objetivos?		
<b>2. Nível Iconográfico (identificação de significados convencionais e simbólicos)</b>		
Critério	Avaliação (-1, 0, +1)	Comentário
A IA reconheceu símbolos ou elementos com significado cultural específico?		
Atribuiu significados alinhados ao contexto da imagem?		
Houve convergência com os significados apontados pelos especialistas?		
<b>3. Nível Iconológico (interpretação profunda, contexto histórico-social)</b>		
Critério	Avaliação (-1, 0, +1)	Comentário
A IA reconheceu símbolos ou elementos com significado cultural específico?		
Atribuiu significados alinhados ao contexto da imagem?		
Houve convergência com os significados apontados pelos especialistas?		

**Fonte:** Elaborado pelos autores (2025).

Essa estrutura permitiu não apenas o registro dos resultados numéricos atribuídos segundo a escala proposta, mas também a inclusão de comentários qualitativos para enriquecer a interpretação dos dados, apontando nuances sobre os acertos e limitações observados nas respostas.

Foram selecionadas duas imagens como base para as análises, a partir de repositórios públicos, e tratam de representações simbólicas fortemente enraizadas no imaginário religioso e cultural brasileiro, como esculturas de cunho barroco e imagens de forte apelo devocional. A escolha se orientou pelo objetivo de simular um cenário real de indexação em repositórios digitais de patrimônio cultural, nos quais a mediação semântica é especialmente relevante, cuja seleção está descrita no Quadro 3.

**Quadro 3 - Amostra de imagens usadas para a indexação**

AMOSTRA DE IMAGENS	
Imagem 1	Imagem 2



Fonte: elaborado pelos autores (2025).

A indexação das imagens foi realizada a partir de orientações explícitas sobre a utilização de palavras-chave para cada um dos três níveis de análise do método iconográfico de Panofsky: pré-iconográfico, iconográfico e iconológico. O objetivo foi garantir a consistência na indexação, fornecendo uma estrutura clara para a seleção de palavras-chave que refletissem as características e interpretações de cada nível. Tanto as ferramentas de IA quanto os especialistas humanos foram orientados a selecionar entre cinco e dez palavras-chave por nível de análise, com o intuito de garantir uma indexação detalhada e robusta. Essa faixa de palavras-chave foi definida para evitar a superficialidade nas descrições e ao mesmo tempo oferecer flexibilidade para que as nuances das imagens fossem capturadas adequadamente.

O processo de indexação humana foi conduzido por dois especialistas com formação em Biblioteconomia e Ciência da Informação, que atuaram de forma independente. Cada especialista realizou a indexação inicial das imagens, utilizando as fichas descritivas padronizadas, as quais foram projetadas para registrar as respostas de maneira clara, objetiva e comparável. Após a conclusão da indexação inicial, os especialistas se reuniram para discutir os resultados e alinhar as interpretações, com o objetivo de assegurar a coerência metodológica e consolidar uma abordagem comum para a análise das imagens.

Simultaneamente, a indexação automatizada foi conduzida utilizando os modelos de IA multimodais GPT-4o e Gemini 1.5 Flash, com a utilização de técnicas de prompt engineering. Para garantir que as IAs seguissem os níveis de Panofsky de maneira consistente, os prompts foram cuidadosamente estruturados. Cada *prompt* foi formulado de maneira a orientar os modelos a realizar a interpretação das imagens de acordo com os três níveis de

análise, solicitando a seleção de palavras-chave e a atribuição de significados conforme cada nível específico.

Exemplo de *prompt* utilizado:

Nível Pré-Iconográfico: "Descreva os elementos visuais formais presentes nesta imagem, identificando as formas, cores e composições, e forneça entre cinco e dez palavras-chave que melhor representam essas características."

Nível Iconográfico: "Identifique os símbolos ou convenções culturais presentes nesta imagem e forneça entre cinco e dez palavras-chave que descrevam o significado convencional ou simbólico desses elementos."

Nível Iconológico: "Forneça uma interpretação mais profunda desta imagem, considerando o contexto histórico, social ou filosófico, e atribua entre cinco e dez palavras-chave que capturam os significados mais profundos da imagem."

O processo de *prompt* engineering foi desenvolvido para garantir isonomia entre os modelos e comparabilidade com a análise humana, evitando enviesamentos que poderiam surgir da formulação das perguntas. Cada comando foi desenhado de forma a permitir que as IAs aplicassem a metodologia de forma consistente, buscando as palavras-chave apropriadas para cada nível de análise.

Após a geração das respostas pelos modelos de IA, as respostas foram transcritas integralmente nas fichas descritivas (quadro 2) e avaliadas com base na escala interpretativa adotada na pesquisa, assegurando que os dados fossem comparáveis entre as respostas dos especialistas humanos e das IAs. A comparação entre os resultados de humanos e máquinas foi conduzida de forma criteriosa, levando em consideração não apenas os valores atribuídos, mas também os aspectos qualitativos envolvidos nas escolhas interpretativas feitas por ambos os grupos. Essa abordagem permitiu identificar diferenças, limitações recorrentes e acertos surpreendentes por parte das IAs, que foram registradas nas fichas de análise, sem que qualquer inferência interpretativa fosse realizada na seção de metodologia.

O conjunto dos dados coletados foi posteriormente analisado de forma integrada na seção de resultados, onde as implicações mais amplas sobre o uso das tecnologias de IA em sistemas de recuperação de informação visual foram discutidas. Esse processo também refletiu sobre as potencialidades e limitações das IAs no contexto da mediação simbólica e cultural das imagens, identificando as oportunidades e desafios da aplicação dessas ferramentas em ambientes automatizados de indexação.

#### **4 RESULTADOS E ANÁLISES**

A análise comparativa da indexação realizada por especialistas humanos e pelos modelos de inteligência artificial teve como objetivo avaliar a capacidade das IAs em replicar a interpretação iconográfica do método proposto por Panofsky. A pesquisa focou em identificar as convergências e divergências nas escolhas interpretativas entre humanos e máquinas, com ênfase nas limitações e eficácia das IAs em replicar as abordagens metodológicas tradicionais. A análise também visou avaliar a aptidão dos modelos de IA em lidar com as três dimensões da metodologia de Panofsky: pré-iconográfica, iconográfica e iconológica, destacando o grau de correspondência com as interpretações feitas por especialistas da área.

##### **4.1 Resultados da indexação**

Após a realização das indexações por profissionais da informação e pelas ferramentas de inteligência artificial, procedeu-se à comparação dos resultados com o objetivo de avaliar a capacidade da IA em representar imagens com base no método iconográfico de Panofsky. Os dados resultantes dessa comparação estão apresentados no Quadro 4.

**Quadro 4 - Resultado da indexação das fotografias**

Profissional ou Tipo de IA	Imagem 1			Imagem 2		
	Pré-iconográfico	Iconográfico	Iconológico	Pré-iconográfico	Iconográfico	Iconológico
GPT 4o	Igreja barroca; Esculturas de pedra-sabão; Portões de ferro trabalhado; Escadarias; Fachada ornamentada; Torres sineiras simétricas; Céu azul; Cruz no topo.	Santuário do Bom Jesus de Matosinhos (Congonhas, MG); Profetas de Aleijadinho; Arquitetura barroca e rococó; Conjunto histórico e religioso; Escultura religiosa colonial brasileira; Símbolo da fé católica.	Expressão da religiosidade do período colonial; Representação do barroco mineiro como apogeu artístico; Relação entre arte, poder e fé no Brasil colonial; Patrimônio cultural da humanidade (UNESCO); Reflexo da espiritualidade e da cultura local.	Esculturas de madeira policromada; Mesa redonda coberta com toalha branca; Doze figuras humanas ao redor; Figura central em destaque; Interior com estantes e janela ao fundo; Expressões faciais detalhadas; Cenário simétrico.	Representação da Última Ceia de Jesus Cristo; Apóstolos em interação dinâmica; Arte sacra do período colonial; Produção típica da Escola Mineira; Influência do cristianismo europeu; Cena bíblica central na tradição católica.	Expressão da espiritualidade no Brasil colonial; Reflexão da importância da Eucaristia na fé católica; Transmissão de valores religiosos por meio da arte; Representação da coletividade e do sacrifício cristão; Elemento educativo e catequético para a época.

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

	Imagem 1			Imagem 2		
Gemini 1.5 Flash	Arquitetura; materiais; composição.	Religião; arte barroca; contexto histórico	Ideias; função social; significado universal.	Cena bíblica; elementos visuais; composição.	Última Ceia; contexto histórico; significado religioso.	Ideias, função social; significado universal.
Profissional 1	Igreja com 6 esculturas de homens em pé; Igreja com 2 torres.	Aleijadinho; Profetas; Santuário do Bom Jesus de Matosinhos Congonhas; doze estátuas de profetas.	Arquitetura colonial brasileira.	Mesa redonda com toalha branca e 13 homens; Homens sentados à mesa.	A Última Ceia; Jesus Cristo e os discípulos.	Momento de partilha de Jesus e os discípulos.
Profissional 2	Igreja; estátuas; turismo; religião; céu azul; esculturas de pedra; estilo barroco.	Aleijadinho; Congonhas; Minas Gerais; turismo religioso; tempo colonial; arte sacra.	Colonialismo; resistência pela arte; espiritualidade; turismo nacional.	Estátua; mesa circular; jantar; construção histórica; símbolo religioso.	Jesus e os apóstolos; representação simbólica; religiosidade; arte sacra.	Cristianismo; catolicismo; espiritualidade; Expressão da fé; religiosidade; devoção e fé; arte religiosa.

**Fonte:** elaborado pelos autores (2025).

Tendo em vista a complexidade da indexação temática de imagens, conforme tratado na literatura, devido à sua linguagem específica (Smit, 1996), à sua natureza polissêmica (Manini, 2002), à subjetividade dos indexadores (Torres *et al.*, 2018), as metodologias contribuem significativamente para este trabalho intelectual (Manini, 2002; Ferreira; Santos, 2013; Torres, 2019).

Seja feita por profissionais humanos ou por ferramentas de IA, os resultados das indexações são possibilidades diversas de se representar imagens, afinal, a indexação de imagens nem sempre é um trabalho completo ou terminado (Smit, 1987). Mas os resultados das indexações são proposições plausíveis, considerando sempre o contexto, o público-alvo e a finalidade do acervo ou repositório de imagens.

**Quadro 5** - Número de palavras usados na indexação das imagens por níveis, profissionais e IA

Profissional ou Tipo de IA	Imagem 1			Imagem 2		
	Pré- iconográfico	Iconográfico	Iconológico	Pré- iconográfico	Iconográfico	Iconológico
GPT 4.0	8	6	5	7	6	5
Gemini 1.5 Flash	3	3	3	3	3	3

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

Profissional 1	2	3	1	2	2	1
Profissional 2	7	6	4	5	4	7

**Fonte:** elaborado pelos autores (2025).

Tendo recebido as mesmas orientações para a realização da representação das imagens, os profissionais da informação e as ferramentas de IA deveriam considerar a utilização de cinco a dez palavras-chave para cada nível de análise do modelo de Panofsky. Ainda assim, percebe-se uma variação significativa entre as IAs e os profissionais, conforme apresentado no quadro 5. Quanto às IA, somente o Gemini 1.5 Flash manteve um padrão quantitativo de números de palavras-chave: três por nível. O GPT 4.0 não teve padrão, mas considerou o limite mínimo e máximo da orientação e foi mais elaborada e informativa, o que pode ser considerado um diferencial positivo em contextos que exigem maior precisão descritiva. Em relação aos profissionais, não houve padrão quantitativo, inclusive desconsideraram o limite mínimo de cinco palavras-chave definidas para a realização da indexação.

Para além dos aspectos quantitativos, observam-se diferenças significativas nas descrições analisadas. Em relação à imagem 1, por exemplo, no nível iconográfico o GPT 4.0 é bem minucioso na descrição dos elementos presentes na imagem, mencionando inclusive os “portões de ferro trabalhado” que não constituem elementos principais na imagem, enquanto o Gemini 1.5 Flash foi bem sucinto e genérico em suas descrições: “arquitetura, materiais, composição”. Os profissionais, por sua vez, descrevem a igreja e as esculturas, o que pode sugerir esses elementos como centrais. Contudo, apenas o profissional 2 menciona o “céu azul”, assim como o GPT 4.0, sendo grande o espaço ocupado pelo céu na imagem.

Descrições mais neutras ou mais específicas são notáveis, embora esse aspecto não comprometa necessariamente a exatidão ou informações corretas sobre a imagem, como é o caso do seguinte exemplo: o Profissional 1 descreve a Imagem 2, no nível pré-iconográfico, como “13 homens”, recorrendo a uma linguagem específica em relação ao gênero, enquanto o modelo GPT 4.0 opta por uma formulação mais neutra: “doze figuras humanas”. Tal escolha pode indicar uma intencionalidade interpretativa, possivelmente orientada por diretrizes voltadas à neutralidade descritiva do GPT 4.0.

As imprecisões na identificação de elementos nas imagens, no que concerne aos sistemas algorítmicos, podem ser atribuídas à presença de características visuais semelhantes

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

na imagem, como contornos, cores, formas e texturas, bem como à ocorrência de sobreposições, variações dimensionais ou perspectivas específicas de captura da imagem, fatores que comprometem a acurácia na detecção de padrões. Já sob a ótica da percepção humana, a subjetividade desempenha papel central na identificação de elementos nas imagens. O profissional 1, por exemplo, infere a existência de treze figuras humanas, presumivelmente Jesus Cristo e os doze apóstolos, desconsiderando a décima quarta figura. O profissional 2, por sua vez, não faz qualquer menção ao número de personagens representados. É plausível supor que o profissional 1 tenha sido influenciado pela centralidade narrativa dos personagens bíblicos, o que pode ter conduzido à omissão da contagem exata das figuras escultóricas efetivamente presentes na composição visual.

Apesar das inconsistências nas descrições acima, do ponto de vista da compreensão das imagens, tais variações não comprometem substancialmente a compreensão da cena representada, desde que os elementos centrais sejam preservados. Isso reforça a ideia de que a fidelidade descritiva não depende apenas da terminologia empregada, mas também da coerência contextual e da capacidade de evocar uma representação do contexto de patrimônio histórico-cultural subjacente à imagem.

É possível identificar ainda variações de ordem léxica e estilística tanto nas descrições produzidas tanto pela inteligência artificial quanto naquelas elaboradas por profissionais humanos, sendo estas últimas, em geral, mais diversificadas em termos de vocabulário, especialmente quando há ausência do uso de vocabulário controlado.

Ao comparar os resultados, constatou-se - conforme os dados apresentados no quadro 6 - que a indexação realizada pelas ferramentas de IA apresenta correspondência satisfatória ou parcial com aquelas geradas pelos profissionais da informação no que se refere à análise da imagem 1.

**Quadro 6 - Análise das indexações da Imagem 1 feitas por IA e por profissionais da informação**

<b>IMAGEM 1</b>		
<b>Comparação por nível panofskyano</b>		
<b>1. Nível Pré-Iconográfico (descrição formal dos elementos visuais)</b>		
<b>Critério</b>	<b>Avaliação (-1, 0, +1)</b>	<b>Comentário</b>
A IA identificou corretamente os elementos visuais?	+1	Todas as descrições das IA correspondem aos objetos retratados na imagem. O GPT 4.0 foi o mais exaustivo, inclusive mais do que um dos profissionais.
Houve correspondência com a descrição feita pelos especialistas?	+1	Houve correspondências da indexação das IA com a indexação dos profissionais.

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

Termos atribuídos foram pertinentes e objetivos?	+1	As palavras e expressões atribuídas pelas ferramentas de IA foram pertinentes e objetivas, mas houve variação de quantidade atribuídas por cada uma: o GPT 4.0 é mais exaustivo que o Gemini 1.5 Flash.
<b>2. Nível Iconográfico (identificação de significados convencionais e simbólicos)</b>		
<b>Critério</b>	<b>Avaliação (-1, 0, +1)</b>	<b>Comentário</b>
A IA reconheceu símbolos ou elementos com significado cultural específico?	0	As palavras e expressões atribuídas pelo GPT 4.0 foram específicas, correspondendo à proposta deste nível representação, enquanto que as atribuídas Gemini 1.5 Flash foram genéricas e atendendo parcialmente à proposta deste nível.
Atribuiu significados alinhados ao contexto da imagem?	0	O Gemini 1.5 Flash utilizou palavras e expressões mais genéricas e atendendo parcialmente à proposta deste nível.
Houve convergência com os significados apontados pelos especialistas?	+1	Houve correspondência forte entre as palavras-chave ou expressões utilizadas pelas IA e pelos profissionais.
<b>3. Nível Iconológico (interpretação profunda, contexto histórico-social)</b>		
<b>Critério</b>	<b>Avaliação (-1, 0, +1)</b>	<b>Comentário</b>
A IA reconheceu símbolos ou elementos com significado cultural específico?	0	As palavras e expressões atribuídas pelo GPT 4.0 foram específicas, correspondendo à proposta deste nível representação, enquanto que as atribuídas Gemini 1.5 Flash foram genéricas e atendendo parcialmente à proposta deste nível.
Atribuiu significados alinhados ao contexto da imagem?	0	As palavras e expressões atribuídas pelo GPT 4.0 foram específicas, correspondendo à proposta deste nível representação, enquanto que as atribuídas Gemini 1.5 Flash foram genéricas e atendendo parcialmente à proposta deste nível.
Houve convergência com os significados apontados pelos especialistas?	0	Houve correspondência parcial entre as palavras-chave ou expressões utilizadas pelas IA e pelos profissionais.

**Fonte:** elaborado pelos autores (2025).

De maneira semelhante, a análise das indexações relativas à imagem 2 indica uma correspondência satisfatória ou parcial entre as categorizações realizadas pelas ferramentas de inteligência artificial e aquelas efetuadas pelos profissionais da informação, conforme exposto no quadro 7.

**Quadro 7 - Análise das indexações da Imagem 2 feitas por IA e por profissionais da informação**

<b>IMAGEM 2</b>		
<b>Comparação por nível panofskyano</b>		
<b>1. Nível Pré-Iconográfico (descrição formal dos elementos visuais)</b>		
<b>Critério</b>	<b>Avaliação (-1, 0, +1)</b>	<b>Comentário</b>
A IA identificou corretamente os elementos visuais?	0	As palavras e expressões atribuídas pelo GPT 4.0 identificaram melhor a temática da imagem do que as atribuídas pelo Gemini 1.5 Flash, que foram mais genéricas e pouco elucidativas.
Houve correspondência com a descrição feita pelos especialistas?	+1	Houve boa correspondência entre as palavras e expressões utilizadas pelas ferramentas de IA e pelos profissionais.
Termos atribuídos foram pertinentes e objetivos?	0	As palavras e expressões atribuídas pelo Gemini 1.5 Flash foram mais genéricas e pouco elucidativas.
<b>2. Nível Iconográfico (identificação de significados convencionais e simbólicos)</b>		

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

<b>Critério</b>	<b>Avaliação (-1, 0, +1)</b>	<b>Comentário</b>
A IA reconheceu símbolos ou elementos com significado cultural específico?	+1	Em geral, as palavras e expressões e expressões apontaram para símbolos ou elementos com significado cultural específico.
Atribuiu significados alinhados ao contexto da imagem?	+1	Em geral, as palavras e expressões e expressões apontaram para símbolos ou elementos com significado cultural específico alinhado ao contexto.
Houve convergência com os significados apontados pelos especialistas?	0	Houve convergência parcial entre as palavras e expressões utilizadas pelas ferramentas de IA e pelos profissionais.
<b>3. Nível Iconológico (interpretação profunda, contexto histórico-social)</b>		
<b>Critério</b>	<b>Avaliação (-1, 0, +1)</b>	<b>Comentário</b>
A IA reconheceu símbolos ou elementos com significado cultural específico?	0	As palavras e expressões atribuídas pelo GPT 4.0 atenderam à proposta deste nível de representação, enquanto que as atribuídas pelo Gemini 1.5 Flash não atenderam.
Atribuiu significados alinhados ao contexto da imagem?	0	As palavras e expressões atribuídas pelo GPT 4.0 atenderam à proposta deste nível de representação, enquanto que as atribuídas pelo Gemini 1.5 Flash não atenderam.
Houve convergência com os significados apontados pelos especialistas?	0	Houve correspondência parcial entre as palavras-chave ou expressões utilizadas pelas IA e pelos profissionais.

**Fonte:** elaborado pelos autores (2025).

Os resultados evidenciam que a IA possui um potencial considerável para atuar como suporte na representação temática de imagens, notadamente na identificação de elementos visuais no nível pré-iconográfico. Isso é significativo como recurso potencial para facilitar a atuação de profissionais, sobretudo em contextos em que o número reduzido de profissionais para as tarefas de indexação compromete a produtividade.

Contudo, a eficácia da IA mostra-se ainda limitada nos níveis interpretativos mais complexos, iconográfico e iconológico, especialmente no que se refere à decodificação de símbolos e à contextualização histórica. Tal limitação reforça a relevância da mediação exercida pelos profissionais da informação, cuja atuação é essencial para a análise crítica, a inserção contextual e os ajustes interpretativos necessários à representação adequada dos conteúdos imagéticos.

## **5 CONSIDERAÇÕES FINAIS**

As análises conduzidas neste estudo permitiram verificar em que medida modelos de inteligência artificial são capazes de replicar os três níveis de interpretação propostos por Erwin Panofsky. Os resultados demonstram que os objetivos inicialmente propostos foram alcançados, ao evidenciar que os modelos avaliados desempenham bem a identificação de

elementos visuais no nível pré-iconográfico, com desempenho parcialmente satisfatório no nível iconográfico, mas encontram limitações significativas no nível iconológico. Entre os principais desafios observados estão o viés algorítmico, a dificuldade em compreender simbolismos culturais e a dependência da qualidade dos dados de treinamento.

A IA mostra-se útil como ferramenta de apoio à indexação temática de imagens, principalmente em fases iniciais de análise, mas ainda carece de intervenção humana para lidar com significados mais profundos e contextos históricos complexos.

Os achados reforçam a necessidade de metodologias híbridas que combinam processamento automatizado com interpretação especializada, especialmente em contextos nos quais a representação da informação exige sensibilidade simbólica e responsabilidade ética. Assim, conclui-se que este estudo contribui tanto para a compreensão das potencialidades e limitações da IA na análise iconográfica quanto para o aprimoramento de sistemas de organização da informação visual, indicando caminhos promissores para futuras pesquisas e aplicações em ambientes digitais multimodais.

## **REFERÊNCIAS**

BRÄSCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento? *In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO*, 9., 2008, São Paulo, **Anais** [...]. São Paulo: ANCIB, 2008.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, [s. l.], v. 20, n. 1, p. 37–46, 1960. Disponível em: DOI: 10.1177/001316446002000104. Acesso em: 11 ago. 2025.

DENDRITES.IO. **Glossário de Inteligência Artificial de A a Z**. 2024. Disponível em: <https://www.i2ai.org/content/glossary/>. Acesso em: 12 jul. 2024.

DENG, J.; LIN, Y. The benefits and challenges of ChatGPT: an overview. **Frontiers in Computing and Intelligent Systems**, [s. l.], v. 2, n. 2, p. 81–83, 2022.

DEVLIN, J. et al. BERT: pre-training of deep bidirectional transformers for language understanding. **arXiv preprint**, [s. l.], 2018. Disponível em: <https://arxiv.org/abs/1810.04805>. Acesso em: 12 jul. 2024.

DIRECCIÓN GENERAL DE BIBLIOTECAS Y SERVICIOS DIGITALES DE INFORMACIÓN. El método iconográfico e iconológico de Erwin Panofsky. **LIEC – Lee, Investiga, Escribe y Comunica**, Coyoacán, [202-]. Disponível em: <https://liec.dgb.unam.mx/index.php/lee/la-lectura-de-imagenes/el-metodo-iconografico-e-iconologico-de-erwin-panofsky>. Acesso em: 11 ago. 2025.

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

FERREIRA, S. L.; SANTOS, M. Elementos da descrição de imagens de arte em ambiente eletrônico: considerações sobre o padrão VRA Core 4.0. *In: ENCONTRO INTERNACIONAL DE CATALOGADORES; ENCONTRO NACIONAL DE CATALOGADORES*, 9., 2., 2013, Rio de Janeiro. **Anais** [...]. Rio de Janeiro: Biblioteca Nacional, 2013.

KOCÓN, J. *et al.* ChatGPT: jack of all trades, master of none. **Information Fusion**, [S. l.], v. 99, e.C, 2023. Disponível em: DOI: 10.1016/j.inffus.2023.101861. Acesso em: 11 ago. 2025.

LAYNE, S. S. Analyzing the subject of a picture: a theoretical approach. **Cataloging & Classification Quarterly**, New York, v. 6, n. 3, p. 39-62, 1986.

MANINI, M. P. **Análise documental de fotografias**: um referencial de leitura de imagens fotográficas para fins documentários. 2002. 231 f. Tese (Doutorado em Ciências da Comunicação) - Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2002.

NASCIMENTO, G. D.; MARTINS, G. K.; ALBUQUERQUE, M. E. B. C. de. Automação da indexação: evidências e tendências da produção científica indexada na Brapci. **Encontros Bibli**: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 28, p. 1–20, 2023.

PANOFSKY, E. **Significado nas artes visuais**. 2. ed. São Paulo: Perspectiva, 1979.

SMIT, J. A análise da imagem: um primeiro plano. *In: SMIT, Johanna. (coord.). Análise documental: a análise da síntese*. Brasília, DF: IBICT, 1987. p. 99–111.

SMIT, J. A representação da imagem. **Informare: Caderno do Programa de Pós-Graduação em Ciência da Informação**, Rio de Janeiro, v. 2, n. 2, p. 28- 36, jul./dez. 1996.

TORRES, A. A. L.; MACULAN, B. C.; DIAS, C. C.; SILVA, G. R. Subjective aspects in the thematic representation of photographic. **Collection and Curation**, México, v. 37, n. 4, p. 151-157, 2018. Disponível em: <https://www.emeraldinsight.com/doi/abs/10.1108/CC-03-2018-0005>. Acesso em: 27 out. 2024.

TORRES, A. A. L. **Metodologia para a representação de registro fotográfico de esculturas de arte sacra**. 2019. 206 f. Dissertação (Mestrado em Gestão e Organização do Conhecimento) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2019.

TORRES, A. A. L.; SILVA, G. R. **Metodologias de representação de imagens**. Belo Horizonte, set. 2017. 40 slides.

## **AGRADECIMENTOS**

Agradecimento ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio à pesquisa (Processo 307765/2023-7)

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**