



UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS SOCIAIS APLICADAS  
FACULDADE DE BIBLIOTECONOMIA

FRANCISCO MIKAEL CARVALHO DE LIMA

**INTELIGÊNCIA ARTIFICIAL NA CATALOGAÇÃO: comparação de registros MARC  
21 gerados por IA generativas**

Belém – PA

2025

FRANCISCO MIKAEL CARVALHO DE LIMA

**INTELIGÊNCIA ARTIFICIAL NA CATALOGAÇÃO: comparação de registros MARC  
21 gerados por IA generativas**

Trabalho de Curso, apresentado como requisito parcial  
para obtenção de grau de Bacharel em Biblioteconomia,  
pela Universidade Federal do Pará.

Orientadora: Dra. Jacquelin Teresa Camperos Reyes

Belém – PA  
2025

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD  
Sistema de Bibliotecas da Universidade Federal do Pará  
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

---

L732i Lima, Francisco Mikael Carvalho de.  
Inteligência artificial na catalogação : comparação de registros  
MARC 21 gerados por IA generativas / Francisco Mikael Carvalho  
de Lima. — 2025.  
29 f. : il.

Orientador(a): Prof<sup>ª</sup>. Dra. Jacquelin Teresa Camperos Reyes  
Trabalho de Conclusão (Graduação) - Universidade Federal do  
Pará, Instituto de Ciências Sociais Aplicadas, Faculdade de  
Biblioteconomia, Belém, 2025.

1. Catalogação. 2. Inteligência Artificial - Generativa. 3.  
MARC 21. 4. Representação descritiva. I. Título.

CDD 025.32

---

FRANCISCO MIKAEL CARVALHO DE LIMA

**INTELIGÊNCIA ARTIFICIAL NA CATALOGAÇÃO: comparação de registros MARC  
21 gerados por IA generativas**

Trabalho de Curso, apresentado como requisito parcial para obtenção do grau de Bacharela em Biblioteconomia, pela Faculdade de Biblioteconomia da Universidade Federal do Pará.

Data da avaliação: 03/09/2025

Conceito: \_\_\_\_\_

**Banca Examinadora:**

---

Profª. Dra. Jacquelin Teresa Camperos Reyes (Orientadora)  
Universidade Federal do Pará – UFPA

---

Dra. Maurila Bentes de Melo e Silva  
Universidade Federal do Pará – UFPA

---

Me. Doris Campos Mendonça  
Universidade Federal do Pará – UFPA

## INTELIGÊNCIA ARTIFICIAL NA CATALOGAÇÃO: comparação de registros MARC 21 gerados por IA generativas

Francisco Mikael Carvalho de Lima<sup>1</sup>  
Jacquelin Teresa Camperos Reyes<sup>2</sup>

**Resumo:** Este trabalho investiga o desempenho de três modelos de inteligência artificial generativa – ChatGPT, Gemini e DeepSeek – na catalogação de e-books no formato MARC 21, com base no Código de Catalogação Anglo-Americano. Trata-se de uma pesquisa básica, de caráter exploratória, de abordagem qualitativa e método comparativo, que analisou os registros bibliográficos produzidos por cada inteligência artificial, com base em critérios de exatidão descritiva, conformidade normativa e consistência na estruturação dos campos MARC. Os resultados mostraram que, embora capazes de gerar registros parcialmente adequados, os modelos apresentaram falhas recorrentes, incluindo uso incorreto de campos, inconsistência normativa e variabilidade de respostas. Entre os sistemas avaliados, o ChatGPT demonstrou maior aderência ao Código de Catalogação Anglo-Americano e adequação aos campos MARC 21; o Gemini apresentou dificuldades na organização lógica da descrição; e o DeepSeek mostrou-se inconsistente ao alternar entre Código de Catalogação Anglo-Americano e o *Resources Description and Access*. Conclui-se que inteligência artificial generativas possuem potencial para atuar como ferramentas de apoio à catalogação, gerando rascunhos de registros bibliográficos, mas ainda não substituem o trabalho especializado do bibliotecário. As limitações observadas reforçam a necessidade de supervisão humana e apontam para a importância de realizar novas testagens à medida que a tecnologia evolui, abrangendo diferentes tipologias documentais, códigos de catalogação e métricas de avaliação, de modo a reavaliar sua precisão, consistência e aplicabilidade na Biblioteconomia.

**Palavras-chave:** catalogação; Código de Catalogação Anglo-Americano; inteligência artificial generativa; MARC21; representação descritiva.

## ARTIFICIAL INTELLIGENCE IN CATALOGUING: comparison of MARC 21 records generated by generative AI

**Abstract:** This paper investigates the performance of three generative artificial intelligence models - ChatGPT, Gemini and DeepSeek - in cataloging e-books in the MARC 21 format, based on Anglo-American Cataloguing Rules. This is a basic, exploratory study with a qualitative approach and a comparative method, which analyzed the bibliographic records produced by each artificial intelligence based on criteria of descriptive accuracy, normative compliance and consistency in the structuring of MARC fields. The results showed that, although they were able to generate partially adequate records, the models had recurring flaws, including incorrect use of fields, normative inconsistency and response variability. Among the systems evaluated, ChatGPT showed greater adherence to Anglo-American Cataloguing Rules and suitability for MARC 21 fields; Gemini showed difficulties in the logical organization of the description; and DeepSeek proved to be inconsistent when

---

<sup>1</sup> Graduando de Biblioteconomia na Universidade Federal do Pará – UFPA. Email: f.mikaelcarvalho@gmail.com

<sup>2</sup> Professora adjunta do Instituto de Ciências Sociais Aplicadas (ICSA) na Universidade Federal do Pará – UFPA. Email: jacquelin@ufpa.br

switching between Anglo-American Cataloguing Rules and Resources Description and Access. It can be concluded that generative artificial intelligence has the potential to act as cataloguing support tools, generating draft bibliographic records, but they do not yet replace the specialized work of the librarian. The limitations observed reinforce the need for human supervision and point to the importance of carrying out new tests as the technology evolves, covering different document typologies, cataloging codes and evaluation metrics, in order to re-evaluate its accuracy, consistency and applicability in Librarianship.

**Keywords:** Anglo-American Cataloguing Rules; artificial intelligence generative; cataloging; MARC21; descriptive representation.

## 1 INTRODUÇÃO

A Inteligência artificial (IA) é frequentemente associada à otimização, ou até mesmo substituição, de tarefas realizadas por humanos. Esse debate pode se estender para o campo da Biblioteconomia com operações realizadas na área, especialmente ao que se refere ao processo da representação descritiva em *Machine Readable Cataloging – MARC* (Catalogação Legível por Máquina), denominado de MARC 21. O estudo tem como tema correlacionar a IA com a catalogação em MARC21, visando comparar os registros bibliográficos gerados por diferentes modelos de IA generativa.

O tema pode ser considerado atual, pois como afirmado por Rodriguez e Prudencio(2024, p. 5-6)

Qualquer estudo que se debruçar a avaliar quantitativamente o número de publicações relacionadas ao tema da IA, verificará um grande aumento de publicações nos últimos três anos. [...] IA deixou de ser apenas um campo de estudo restrito a nichos especializados como o da Computação e vem tornando-se uma área de interesse interdisciplinar.

Logo, no campo da Biblioteconomia é visto esse interesse em trabalhos abordando aspectos como: tendências, ética, serviço de referências, indexação automática e entre outros. Além disso, se engloba como atual, visto que ainda há necessidade de investigar outras questões relacionadas ao campo científico.

As bibliotecas e outras unidades de informação utilizam catálogos online para facilitar a consulta ao acervo pelos usuários. Esses catálogos operam por meio de sistemas de gerenciamento de bibliotecas e acervos, os quais devem ser capazes de armazenar e estruturar metadados para a descrição de recursos bibliográficos, além de possibilitar o intercâmbio dessas informações. O padrão de metadado MARC 21 é amplamente adotado por bibliotecas em todo o mundo para descrição bibliográfica e o intercâmbio de informação, apesar de ter sofrido modificações ao longo dos anos, ainda se destaca por sua significativa adesão e uso da comunidade bibliotecária (Dumer; Albuquerque, 2020).

O formato MARC 21 é utilizado para a representação descritiva de recursos informacionais. Para que esses recursos sejam localizados em um catálogo online é fundamental que as descrições bibliográficas sigam um padrão de qualidade garantindo uma recuperação da informação mais eficiente. As autoras Dumer e Albuquerque (2020) ressaltam que, um grande volume de informação, sem um tratamento adequado e uma descrição de qualidade, não proporciona praticidade para o usuário.

Dumer e Albuquerque (2020) também destacam que a realização da descrição bibliográfica demanda tempo e dedicação por parte do bibliotecário. Sob o mesmo ponto de vista, Breending (2022) estabelece que para criar um registro MARC pode variar entre 1 hora à 20 minutos dependendo da complexidade da descrição.

O uso do MARC 21 permite a importação de registros bibliográficos já existentes em outras unidades informacionais, cabendo ao catalogador apenas “revisar os dados da importação e finalizar o cadastro conforme as necessidades de sua instituição” (Dumer; Albuquerque, 2020, p. 33). Esta prática pretende otimizar o trabalho do bibliotecário, entretanto, quando não há registros disponíveis a catalogação deve ser feita do zero.

Logo, a importação apesar de facilitar a operação, ainda há catalogações que necessitam do conhecimento, tempo e desempenho do catalogador. Além disso, a importação não equivale ao uso de uma IA generativa, que, segundo Almeida (2023), é um sistema capaz de criar diferentes tipos de conteúdo, como textos, imagens, músicas, vídeos e voz, a partir de instruções fornecidas pelo usuário e com treinamento gerar novos tipos de conteúdo. Nesse caso, podemos considerar a criação de novos registros em MARC 21.

Dado esse contexto, a IA tem se mostrado uma ferramenta promissora para otimizar processos e gerenciamento em atividades da Biblioteconomia. Segundo Silva (2024), atualmente, a Biblioteconomia vivencia mudanças expressivas com a integração da IA, que redefine a forma como os ambientes de informação funcionam e prestam serviços.

Observado que a catalogação em MARC21 é um serviço operacional essencial na Biblioteconomia, exige tempo e precisão do profissional para garantir uma recuperação ágil e eficiente da informação. A IA surge como uma solução possível de otimização, reduzindo a carga de trabalho manual e minimizando erros, permitindo que os bibliotecários direcionem seus esforços para outras atividades. A IA generativa pode ser instruída para otimizar o processo de catalogação. Dessa forma, investigar essa abordagem é essencial para avaliar seu impacto e viabilidade.

Diante desse cenário, emerge a seguinte questão de pesquisa: quais são as possibilidades da utilização da IA generativa, para criar registros bibliográficos em MARC 21 com qualidade e eficiência?

Cientificamente, a pesquisa se justifica pela necessidade de investigar a integração da IA às práticas biblioteconômicas, semelhante ao que Godinho (2019, p. 15) discute com “[...] a utilização de novos recursos dentro da biblioteca, trazendo as novas tendências, definidas por organizações relevantes do universo biblioteconômico, para a área e revelando como a IA tem sido aplicada em diferentes bibliotecas do mundo”.

Apesar da familiaridade dos bibliotecários com bibliotecas automatizadas, as tendências para a aplicação da IA em unidades informacionais ainda é um tema recente e pouco debatido na área, principalmente em território nacional. Logo, a relevância da pesquisa decorre do fato de abordar um tema ainda pouco explorado no Brasil, mas em expansão em outros países (Silva; Nathanhson, 2018).

Profissionalmente, esta pesquisa se justifica por buscar compreender como essas tecnologias podem vir a fazer parte da atuação do bibliotecário, ampliando suas possibilidades e otimizando processos. Conforme ressaltado por Rodriguez e Prudencio (2024, p. 19) a “aplicação dessas tecnologias nas bibliotecas já é uma realidade, ainda que não ampla”. Silva (2024, p. 31), por sua vez

o profissional bibliotecário, no início do século XXI, continua se deparando com os avanços tecnológicos e algumas mudanças na sua atuação [...]. Olhando para o futuro, esse profissional precisa continuar a prospectar novas e distintas frentes de trabalho. Os que assumirão o mercado do futuro deverão possuir uma postura flexiva, mente aberta para o novo, atualizar-se constantemente, ter os olhos nos mercados futuro e, mais do que nunca, ser inovadores.

Nesse sentido, é fundamental que o profissional esteja atento às inovações tecnológicas. Logo, se a catalogação em MARC21 gerada por IA demonstrar resultados positivos, a adoção da IA pode contribuir para eficiência na organização da informação, fortalecendo a prática profissional e oferecendo novas ferramentas para a organização da informação.

O estudo tem como objetivo geral: analisar a catalogação em MARC21 realizada pelas IA generativa, considerando a qualidade, eficiência e conformidade com as regras de catalogação. Ademais, como objetivos específicos: a) explorar o desempenho da IA na catalogação em MARC21; b) verificar a conformidade com as normas catalográficas e qualidade da descrição; c) identificar vantagens e limitações de cada abordagem na geração de registros de descrição bibliográfica.

## 2 CATALOGAÇÃO BIBLIOGRÁFICA

O processo de catalogação, na Biblioteconomia, é considerado um dos mais relevantes para a efetiva utilização do acervo. Segundo Mey e Silveira (2009, p. 7), a catalogação é

O estudo, preparação e organização de mensagens codificadas, com base em registros do conhecimento, reais ou ciberespaciais, existentes ou passíveis de inclusão em um ou vários acervos, de forma a permitir interseção entre as mensagens contidas nestes registros do conhecimento e as mensagens internas dos usuários.

Essa definição destaca a catalogação como uma atividade de mediação entre o conhecimento registrado e as necessidades informacionais dos usuários, contribuindo para a recuperação eficiente da informação. Complementando essa perspectiva, Cunha e Cavalcanti (2008, p. 7) definem a catalogação como um “processo técnico para registro e descrição de itens tendo em vista a organização de catálogos”. Apesar das diferentes abordagens, os conceitos apresentados pelos autores convergem ao enfatizar a catalogação como um processo de representação descritiva dos materiais, essencial para sua organização e recuperação.

No ambiente das bibliotecas, a localização dos materiais é viabilizada por meio dos catálogos, que funcionam como instrumentos para a recuperação dos documentos disponíveis. Dentro desse contexto, o registro bibliográfico deve estar inserido em um catálogo, o qual, segundo Mey e Silveira (2009, p. 12), é entendido como

[...] meio de comunicação, que veicula mensagens sobre os registros do conhecimento, de um ou vários acervos, reais ou ciberespaciais, apresentando-as com sintaxe e semântica próprias e reunindo os registros do conhecimento por semelhanças, para os usuários desses acervos.

Com base nessa compreensão, a Declaração dos Princípios Internacionais de Catalogação da *International Federation of Library Associations and Institutions*<sup>3</sup> – IFLA (2016) estabelece objetivos e funções que o catálogo deve constituir um instrumento eficiente e funcional permitindo ao usuário: encontrar recursos bibliográficos específicos ou conjuntos relacionados; identificar corretamente a entidade procurada; selecionar o material mais adequado às suas necessidades; adquirir ou acessar o item desejado; e navegar de forma lógica entre os dados disponíveis.

Para que essas funções sejam cumpridas adequadamente, Mey e Silveira (2009) destacam que a catalogação deve apresentar características como: integridade, que garante a fidelidade e a honestidade das informações; clareza, ao tornar os dados compreensíveis ao usuário; precisão, assegurando informações sem ambiguidades; lógica, na organização coerente dos dados; e consistência, com a aplicação uniforme das regras. Essas qualidades são

---

<sup>3</sup> Traduzido para o português como Federação Internacional de Associações e Instituições Bibliotecárias

fundamentais para garantir que os usuários encontrem com facilidade a informação de que necessitam, reafirmando a catalogação como uma atividade técnica e comunicacional essencial na organização e no acesso ao conhecimento.

Para garantir consistência e interoperabilidade, diferentes códigos de catalogação foram desenvolvidos. Entre eles, destaca-se o AACR2 (*Anglo-American Cataloguing Rules*, 2ª edição), ainda amplamente utilizado no Brasil. Segundo Teixeira (2019), cerca de 95% das bibliotecas universitárias brasileiras pesquisadas adotam o AACR2 em seus processos catalográficos, o que evidencia sua força institucional, mesmo frente a propostas mais recentes como o *Resources Description and Access* (RDA).

Além disso, segundo Dumer e Albuquerque (2020), a representação descritiva precisa considerar as variações entre os diferentes tipos de materiais de um acervo, exigindo que o catalogador utilize adequadamente os instrumentos normativos para garantir a qualidade dos registros e facilitar sua recuperação.

## **2.1 Formato MARC21**

O formato MARC 21 é um padrão de metadados desenvolvido para representar dados bibliográficos de forma legível por máquinas. Ele facilita o intercâmbio de registros e constitui um padrão essencial para a automação dos processos catalográficos permitindo que os registros bibliográficos sejam lidos, interpretados e utilizados por diferentes softwares de gestão de acervos. Alves (2010, p. 47-48) descreve um padrão de metadados como:

[...] estruturas de descrição constituídas por um conjunto predeterminado de metadados (atributos codificados ou identificadores de uma entidade) metodologicamente construídos e padronizados. O objetivo do padrão de metadados é descrever uma entidade gerando uma representação unívoca e padronizada que possa ser utilizada para recuperação da mesma.

No âmbito da Biblioteconomia, o formato MARC 21 se destaca como padrão de metadados por sua estrutura sistematizada, que organiza os dados em campos numéricos, subcampos e indicadores, permitindo a representação estruturada de elementos descritivos de um item informacional, sejam eles físicos ou digitais.

De acordo com a Library of Congress (1996), o formato MARC 21 é composto por cinco tipos principais de dados: bibliográficos, de acervo, de autoridade, de classificação e para informações da comunidade. Cada um desses formatos atende a uma função específica no contexto da organização de informações em bibliotecas. No entanto, esta pesquisa se concentra no formato de dados bibliográficos, e conforme a Library of Congress (1996), utilizado para descrever recursos informacionais, como livros, periódicos, documentos

digitais, entre outros. Esse formato contém especificações para a codificação de elementos de dados necessários à descrição, recuperação e controle de diversas formas de material bibliográfico.

Em um registro MARC, as informações são distribuídas em campos específicos, identificados por etiquetas numéricas, ou chamadas também de *tags*, compostas por três caracteres. Os campos variáveis são organizados em blocos, conforme o primeiro dígito da *tag*, o qual indica a função geral dos dados dentro do registro, como entrada principal, entrada secundária ou entrada de assunto. Já os dois dígitos restantes da *tag* especificam o tipo de informação representada no campo, como nome pessoal, entidade ou evento (Library of Congress, 1996). As *tags* para o formato de dados bibliográficos estão estruturados como apresentado no Quadro 1:

Quadro 1 – Campos MARC para o formato de dados bibliográficos.

<i>Tag</i>	Conteúdo do campo
0XX	Informações de controle, números, códigos
1XX	Entrada principal
2XX	Títulos, edição, impressão
3XX	Descrição física, etc.
4XX	Declarações de série
5XX	Notas
6XX	Campos de acesso ao assunto
7XX	Entrada secundária
8XX	Entradas adicionadas à série; acervos e localizações
9XX	Reservado para implementação local

Fonte: Adaptado de Library of Congress (1996).

Picco e Ortiz Repiso (2012) afirmam que para registrar as informações no formato MARC 21 precisa seguir o código de catalogação, combinando a aplicação dos dois padrões, um para descrever e outro para armazenar. Assim, a padronização promovida pelo uso combinado do AACR2 e do MARC21 assegura a integridade dos registros, possibilita a interoperabilidade entre sistemas e favorece a construção de catálogos mais precisos e úteis para os usuários e bibliotecários.

### 3 INTELIGÊNCIA ARTIFICIAL (IA)

A IA é um campo da ciência da computação que visa desenvolver sistemas capazes de simular a inteligência humana. Entre suas abordagens mais comuns destacam-se os modelos

generativos, como os *Large Language Model* (LLM), que utilizam vastas quantidades de dados e técnicas de Processamento de Linguagem Natural (PLN)<sup>4</sup> para gerar respostas em linguagem natural (Brady, 2024). Os modelos de IA generativa selecionados – ChatGPT, Gemini e DeepSeek – para a pesquisa podem ser classificados como chatbots, conceituados como

[...] um programa de computador que tenta simular um ser humano em conversa com pessoas. Seu objetivo é responder a perguntas de forma que as pessoas tenham a impressão de que estão conversando com outro ser humano e não com uma máquina. Depois de receber perguntas em linguagem natural, o programa consulta uma base de conhecimento e, em seguida, fornece uma resposta que tenta imitar o comportamento humano (Maciel, 2023, p. 110, tradução própria).

De acordo com Stryker e Scapicchio (2024), a IA generativa proporciona uma série de benefícios relevantes, tanto para indivíduos quanto para organizações. Entre os benefícios, os autores destacam: o aumento da produtividade, uma vez que a tecnologia pode automatizar tarefas repetitivas; o estímulo da criatividade ao auxiliar na geração de ideias e versões de conteúdo, facilitando o processo criativo; a melhoria na tomada de decisões, já que a IA é capaz de analisar grandes volumes de dados, identificar padrões e sugerir hipóteses úteis para decisões estratégicas; e, por fim, a personalização dinâmica também é um diferencial permitindo gerar conteúdos adaptados ao perfil do usuário em tempo real.

Apesar dos benefícios, a IA generativa também apresenta riscos e limitações importantes que não podem ser ignorados. Stryker e Scapicchio (2024) ressaltam a ocorrência das chamadas “alucinações”, situações que o modelo gera informações falsas ou imprecisas que, embora pareçam plausíveis, não têm respaldo na realidade. Sendo assim, esse problema pode comprometer a confiabilidade das respostas.

Ademais, a variabilidade das respostas é outro ponto crítico. Por conta da natureza probabilística desses modelos, uma mesma entrada pode resultar em saídas diferentes, o que pode ser indesejável em contextos que exigem consistência (Stryker; Scapicchio, 2024). Diante disso, é fundamental compreender que todo conteúdo gerado por uma IA generativa é passível de revisão.

### **3.1 IA na Biblioteconomia**

A IA, no contexto da Biblioteconomia, já pode ser considerada uma realidade concreta e irreversível. Embora a adoção dessas tecnologias ainda seja desigual entre as instituições, os impactos da IA sobre os serviços, práticas e processos biblioteconômicos vêm se tornando

---

<sup>4</sup> Área da Computação e da IA que usa aprendizado de máquina para que computadores entendam a linguagem humana, tanto escrita quanto falada, para resposta e comando (Stryker; Holdsworth, 2024).

cada vez mais evidentes. Conforme defendido por Wheatley e Hervieux (2019), ainda são poucas as bibliotecas que incluíram a IA em seus planejamentos estratégicos. No entanto, as autoras enfatizam que: “A revolução da IA não está no horizonte, ela já está aqui, e as bibliotecas precisam aceitar esse fato e iniciar o processo de coexistência.” (Wheatley; Hervieux, 2019, p. 354, tradução própria). Essa coexistência exige não apenas a adoção tecnológica, mas também uma revisão crítica dos papéis das bibliotecas e dos profissionais da informação diante de novos modelos de organização e recuperação da informação mediados por algoritmos inteligentes.

A IA já se apresenta nas bibliotecas principalmente como uma solução para melhorar a eficiência no atendimento aos usuários. Godinho (2019) analisa, por exemplo, o caso da Bibliotecária Informativa Automatizada (BIA), um chatbot desenvolvido pela Divisão de Bibliotecas e Documentação da PUC-Rio, cujo objetivo é prestar atendimento virtual por meio do Facebook Messenger. A experiência da PUC-Rio demonstra como a IA, ainda que em soluções simples como assistentes virtuais, pode contribuir significativamente para a mediação da informação e a aproximação entre usuários e serviços bibliotecários, especialmente em ambientes digitais cada vez mais demandados.

Algumas outras aplicações da IA em bibliotecas pelo mundo foi sintetizado por Godinho (2019):

Quadro 2 – Iniciativas e projetos de IA em bibliotecas.

<b>Nome</b>	<b>Local</b>	<b>Usado Para</b>	<b>Características</b>
Vincent e Nancy (Robôs com IA)	Estados Unidos	Aulas e workshops	<ul style="list-style-type: none"> <li>- Programáveis;</li> <li>- Aprendizado de máquina (correção de comportamento);</li> <li>- Reconhecimento de fala;</li> <li>- Processamento da linguagem natural ;</li> <li>- Reconhecimento facial e processamento de imagens;</li> <li>- Interação com pessoas e objetos</li> </ul>
Scout (Assistente de IA)	Canadá	Atendimento ao usuário	<ul style="list-style-type: none"> <li>- Chatbot;</li> <li>- Independente;</li> <li>- Aprendizado profundo;</li> <li>- Redes neurais;</li> <li>- Processamento da linguagem natural;</li> <li>- Interação com pessoas.</li> </ul>
Laboratório de IA	Estados Unidos	Ensino	<ul style="list-style-type: none"> <li>-Disponibilização de materiais;</li> <li>-Estrutura para estudo;</li> <li>-Assistência de bibliotecários e professores.</li> </ul>
Xiaotu (Chatbot para serviço de referência)	China	Serviço de referência	<ul style="list-style-type: none"> <li>- Chatbot;</li> <li>- Independente;</li> <li>- Aprendizado profundo;</li> <li>- Processamento da linguagem natural;</li> <li>- Interação com pessoas</li> </ul>

Fonte: Adaptado de Godinho (2019).

Outras aplicações são explicitadas no Livro da IFLA (Balnaves, 2025) “*New Horizons in Artificial Intelligence in Libraries*”<sup>5</sup>

- Malmsten *et al.* (2025) relatam a experiência da Biblioteca Nacional da Suécia com a criação automatizada de um sistema de assuntos vinculados, utilizando técnicas de aprendizado de máquina para análise de documentos e indexação automática, o projeto empregou redes neurais para estabelecer conexões entre termos relacionados. O objetivo era substituir os títulos de assunto não estruturados por uma estrutura mais coerente e navegável.
- Kasprzik (2025) apresenta a experiência da ZBW – Biblioteca Nacional da Alemanha para Economia, que aplicou o uso de aprendizado de máquina à prática de indexação temática. O projeto combinou classificações humanas e automatizadas com o objetivo de aprimorar a precisão dos descritores e reduzir inconsistências.
- Saccucci e Potter (2025) descrevem a experiência da Library of Congress dos Estados Unidos com o uso de aprendizado de máquina para criar ou aprimorar registros bibliográficos de conteúdos nascidos digitais. O projeto Explorando a Descrição Computacional consistiu em testar diferentes modelos de aprendizado de máquina para gerar metadados no formato MARC a partir do conteúdo de e-books, com o objetivo de avaliar a viabilidade da automação parcial da catalogação e a qualidade dos registros gerados em comparação aos produzidos por catalogadores humanos.

Dando destaque para os estudos que estão em desenvolvimento pela Library of Congress, a instituição tem explorado o uso de inteligência artificial na catalogação, aplicando modelos baseados em transformadores para gerar metadados a partir do conteúdo completo das obras. Em fases recentes testou-se a extração automática de informações como título, autor e assunto. A iniciativa também levanta reflexões sobre os desafios éticos e técnicos da automação, reforçando o papel essencial dos profissionais na curadoria dos dados (Library of Congress, [202-]).

Constata-se que a aplicação da IA em bibliotecas já é uma realidade em diferentes partes do mundo, acompanhada por pesquisas em andamento sobre seu uso em unidades de informação.

---

<sup>5</sup> Em tradução livre: Novos horizontes da Inteligência Artificial em bibliotecas.

## 4 METODOLOGIA

Trata-se de uma pesquisa básica, pois busca ampliar o conhecimento sobre a aplicação da IA na área da Biblioteconomia, com ênfase no processo de catalogação em MARC21. Entretanto, não há a intenção de desenvolver um produto ou solução prática imediata, mas sim de compreender melhor as potencialidades e limitações da IA nesse contexto. Além disso, a pesquisa adotou um caráter exploratório, com abordagem qualitativa.

O método empregado será o comparativo que, segundo Gil (2008, p. 16) busca examinar “indivíduos, classes, fenômenos ou fatos, com vistas a ressaltar as diferenças e similaridades entre eles”. A comparação foi realizada com base em três variáveis principais: os registros de catalogação gerados automaticamente por três modelos de IA generativa. Para realizar a análise comparativa, os modelos escolhidos foram: ChatGPT (versão IA GPT-4o), da OpenAI; Gemini (versão 2.5 Flash), do Google; e DeepSeek (versão V3), da empresa DeepSeek.

A coleta de dados conceituais e teóricos foi realizada por meio de buscas em bases científicas e repositórios de acesso aberto, incluindo as plataformas SciELO, Brapci (Base de Dados em Ciência da Informação) e Google Acadêmico. Para aprofundar os resultados e obter publicações mais específicas, foi utilizada a pesquisa avançada com o emprego de operadores booleanos. Foram aplicadas combinações como: Biblioteconomia AND “inteligência artificial”; “inteligência artificial” AND “catalogação bibliográfica”; “formato MARC21” AND IA. Os critérios de seleção para o escopo da IA priorizaram publicações recentes, textos completos em português ou inglês, com foco na aplicação prática de IA em unidades de informação. Além disso, para fundamentar os conceitos de catalogação e do formato MARC 21, foi consultado código de catalogação - AACR2, fontes institucionais como o site oficial da Library of Congress e bibliografias especializadas de AACR2 em MARC 21.

Como procedimento metodológico, primeiramente, foi selecionado aleatoriamente um conjunto de 10 materiais em formato de *e-book*, de duas fontes de acesso aberto: a plataforma Scielo Livros e da Biblioteca Digital da Câmara dos Deputados (BDGD). Optou-se pelo código de catalogação AACR2, por ainda ser amplamente utilizado no contexto brasileiro. Antes de solicitar a catalogação a cada modelo, foi fornecido o seguinte *prompt*<sup>6</sup> detalhando todas as informações necessárias para orientar a forma como o modelo de IA deveria atuar:

Você é um bibliotecário catalogador experiente, especializado na descrição bibliográfica de materiais digitais, como e-books, PDFs, documentos eletrônicos e recursos online. Seu trabalho é seguir as normas do AACR2 (Anglo-American

---

<sup>6</sup> “Os prompts são instruções ou perguntas específicas fornecidas ao modelo de linguagem para que ele execute uma tarefa.” (Ferreira, 2023, p. 5)

Cataloguing Rules, 2ª edição) e aplicá-las corretamente ao formato de metadados MARC 21.

Você deve preencher todos os campos bibliográficos cabíveis ao material, atentando-se principalmente aos seguintes campos, quando aplicáveis:

1XX (autoria)

245 (título e responsabilidade)

250 (edição)

300 (descrição do recurso digital)

5XX (notas)

650 (assuntos)\*

7XX (entradas secundárias)

\*O campo 650 deve conter três termos de assunto (tópicos) Seu objetivo é criar registros bibliográficos de alta qualidade.

(Comunicação pessoal com os modelos de IA – ChatGPT, Gemini e DeepSeek, 2025)

Após a configuração inicial, os documentos integrais foram anexados aos modelos de IA, seguido de *prompts* mais simples “faça a catalogação em MARC21” (Comunicação pessoal com os modelos de IA – ChatGPT, Gemini e DeepSeek, 2025) para a execução da catalogação. Os registros MARC21 gerados foram submetidos a uma análise comparativa entre si, com base em critérios como: exatidão da descrição bibliográfica, conformidade com as normas do AACR2, e adequação na utilização dos campos MARC.

Para a avaliação dos resultados, foi adotada uma escala qualitativa composta por quatro níveis: excelente, quando há total conformidade com as normas, com execução precisa, completa e padronizada; positivo, para casos com poucas falhas ou omissões, mas com desempenho satisfatório; parcialmente positivo, quando o resultado é funcional, porém com falhas claras que comprometem a qualidade bibliográfica; e insatisfatório, quando há múltiplos erros, omissões significativas e uso inadequado das normas ou dos campos MARC, indicando desempenho incompatível com os critérios mínimos de catalogação.

A partir dessa análise, buscou-se identificar semelhanças, divergências, pontos fortes e limitações de cada modelo de IA no processo de catalogação de recursos bibliográficos digitais.

## 5 RESULTADOS E DISCUSSÃO

Os modelos de IA analisados apresentaram registros que incluíam os campos de controle e os campos numéricos/codificados do formato MARC21. No entanto, esses campos não serão considerados nesta análise. A ênfase da análise recairá sobre os seguintes campos: entrada principal (1XX), área do título e responsabilidade (245), edição (250), área da publicação, distribuição, etc. (260), descrição física (300), notas (5XX), assuntos (650) e entradas secundárias (7XX).

Na análise da entrada principal (campo 1XX) houve um padrão de erro comum entre os três modelos: quando a obra possui organizador(es) ou mais de três autores, todos os modelos atribuíram incorretamente essas responsabilidades como entrada principal. No entanto, conforme estabelece o Código de Catalogação Anglo-Americano (2004), ambos os casos devem adotar o título como entrada principal.

Ainda assim, os modelos demonstraram desempenho adequado ao lidar com obras de até três autores, aplicando corretamente a entrada principal do campo 1XX. O Quadro 3 apresenta os erros cometidos pelos modelos no campo 1XX.

Quadro 3 – Descrição do campo 1XX

Modelo de IA	Descrição gerada
ChatGPT	=100 1\\\$aCaldas, Rosângela Formentini, \$organ.
Gemini	100 1# \$a Caldas, Rosângela Formentini, \$e organizador.
DeepSeek	100 1 ‡a Caldas, Rosângela Formentini, ‡e organizadora.

Fonte: Comunicação pessoal com os modelos de IA – ChatGPT, Gemini e DeepSeek (2025).

Todos os modelos – ChatGPT, Gemini e DeepSeek – estruturam adequadamente o campo 245, identificando corretamente os subcampos: \$a (título), \$b (subtítulo) e \$c (indicação de responsabilidade). Também demonstraram competência ao incluir, no subcampo \$c, outros tipos de responsabilidade além da autoria principal, como prefaciadores.

Entretanto, um erro recorrente no campo 245 foi a má aplicação dos indicadores, especialmente o segundo indicador, que serve para ignorar caracteres iniciais no título, como artigos definidos e indefinidos (Library of Congress, 2022). Em duas obras analisadas, cujos títulos iniciavam por artigo definido, nenhuma das IA ajustou corretamente o valor do segundo indicador, deixando de contabilizar o número de caracteres a serem ignorados. Exposto no Quadro 4 a falha dos modelos:

Quadro 4 – Descrição do campo 245

Modelo de IA	Descrição gerada	Observação
ChatGPT	=245 10\$aOs verbos de padrão especial no português do século XVI /\$cZenaide de Oliveira Novais Carneiro.	Todos os modelos não ajustam o segundo indicador, que nesse exemplo deveria inserir 3, para desconsiderar os 3 primeiros caracteres – o artigo e o espaço.
Gemini	245 10 \$a Os verbos de padrão especial no português do século XVI / \$c Zenaide de Oliveira Novais Carneiro.	
DeepSeek	245 10 a Os verbos de padrão especial no português do século XVI /  c Zenaide de Oliveira Novais Carneiro	

Fonte: Comunicação pessoal com os modelos de IA – ChatGPT, Gemini e DeepSeek (2025).

Um aspecto observado na descrição gerada pelo DeepSeek no campo 245 foi o uso ocasional do subcampo \$h – Designação Geral do Material (DGM) – em dois registros. Esse subcampo, embora opcional, serve para indicar a classe geral do material e é recomendado, especialmente, para recursos não convencionais, como *e-books*.

No campo da edição (250), o ChatGPT seguiu corretamente o padrão de abreviação estabelecido pelo AACR2, utilizando numerais arábicos acompanhados da forma abreviada "ed.", conforme orientações da regra 1.2B1 e do Apêndice B e C (Código, 2004). Assim, em todos os exemplares analisados, o modelo ChatGPT adiciona o campo 250, mesmo quando a obra não informava explicitamente ser uma primeira edição.

O que contraria as regras do AACR2, segundo as quais, quando não há menção de edição na fonte principal de informação, considera-se tratar-se da primeira edição – situação em que o registro do campo 250 é desnecessário; apenas quando a primeira edição estiver expressamente indicada na fonte principal, ela deve ser descrita (Ribeiro, 2008).

O Gemini, por sua vez, realizou apenas a abreviação do numeral, mantendo a palavra "edição" por extenso, em desacordo com a padronização prevista. O DeepSeek alternou entre descrições padronizadas (2. ed.), e não padronizadas (3. edição), demonstrando oscilação na aplicação da descrição.

Quadro 5 – Descrição do campo 250

Modelo de IA	Descrição gerada	Observação
ChatGPT	=250 \a2. ed.	Utilização do campo adequadamente seguindo as indicações de abreviações.
	=250 \a1. ed.	Uso incorreto do campo, a obra não registrava na fonte principal de informação a edição.
Gemini	250 ## \$a 2. edição.	Sem o uso de abreviação em “edição”
DeepSeek	250 ## \$a3ª edição.	Padrão de abreviação incorreta e sem uso de abreviação em “edição”
	250 ‡a2. ed.	Utilização do campo adequadamente seguindo as indicações de abreviações.

Fonte: Comunicação pessoal com os modelos de IA – ChatGPT, Gemini e DeepSeek (2025).

Após a análise da edição, observou-se que DeepSeek em três registros MARC optou fazer a catalogação pelas regras do RDA, mesmo o *prompt* da pesquisa deixando claro o AACR2 como código de catalogação, essa tomada de escolha influenciou em como o modelo desenvolveu o restante da representação descritiva.

Conforme Fernandes (2014), o RDA modifica diretamente o uso de campos MARC, introduzindo os campos específicos 336 (tipo de conteúdo), 337 (tipo de mídia) e 338 (tipo de suporte) em substituição ao DGM do campo 245; implementa o campo 264 em lugar do 260 para detalhar local e data de publicação e elimina várias abreviaturas presentes no AACR2, como “p.”, “ed.”, adotando termos por extenso, como “páginas” e “edição”. O Quadro 6 é um registro completo do MARC 21 gerado pelo DeepSeek adotando o RDA como regra:

Quadro 6 – Registro MARC 21 gerado pelo DeepSeek

020	‡a 9786558461593 ‡q (eISBN)
040	‡a BRUeB ‡b por ‡e rda ‡c BRUeB
041	0 ‡a por
100	1 ‡a Abers, Rebecca Neera, ‡e organizador.
245	10 ‡a Ativismo institucional : ‡b criatividade e luta na burocracia brasileira / ‡c Rebecca Neera Abers (org.).
250	‡a 1a edição.
264	1 ‡a Brasília : ‡b Editora UnB, ‡c 2021.
300	‡a 1 recurso online (356 páginas) : ‡b ilustrações.
336	‡a texto ‡b txt ‡2 rdacontent
337	‡a computador ‡b c ‡2 rdamedia
338	‡a recurso online ‡b cr ‡2 rdacarrier 500 ‡a Inclui referências bibliográficas.
505	0 ‡a Ação criativa, ativismo e lutas no interior do Estado / Rebecca Neera Abers -- Parte 1. Ativismo situado na história -- Parte 2. Ativismo situado em contextos -- Parte 3. Ativismo em ação.
520	‡a O livro examina o ativismo institucional no Brasil, explorando como servidores públicos contestam hierarquias e o status quo para promover políticas públicas alternativas ou resistir a tentativas de desmonte de conquistas anteriores. A obra reúne estudos de caso que analisam as causas, estratégias e efeitos do ativismo institucional em diferentes políticas públicas brasileiras.
546	‡a Texto em português.
650	7 ‡a Burocracia ‡z Brasil ‡2 larpcal
650	7 ‡a Movimentos sociais ‡z Brasil ‡2 larpcal
650	7 ‡a Políticas públicas ‡z Brasil ‡2 larpcal
700	1 ‡a Amaral, Lucas Alves, ‡e autor.
700	1 ‡a Brandão, Igor, ‡e autor.
700	1 ‡a Viana, Rafael, ‡e autor.
700	1 ‡a Oliveira, Marília Silva de, ‡e autor.
700	1 ‡a Vilaga, Luiz Henrique, ‡e autor.
700	1 ‡a Pereira, Ana Karine, ‡e autor.
700	1 ‡a Cartaxo, Mariana Gomes, ‡e autor.
700	1 ‡a Lima, Ana Marusia Pinheiro, ‡e autor.
700	1 ‡a Lima, Paula Pompeu Fiuza, ‡e autor.
710	2 ‡a Universidade de Brasília, ‡e editora.
856	40 ‡u [URL do recurso digital] ‡z Acesso online

Fonte: Comunicação pessoal com o modelo de IA DeepSeek, 2025, grifo próprio.

No campo 260 (área da publicação, distribuição), o modelo ChatGPT demonstrou bom desempenho neste campo, preenchendo de forma consistente os subcampos \$a (local), \$b (editora) e \$c (data), mesmo em situações mais complexas com múltiplas cidades e editoras. As informações foram representadas de forma adequada e padronizada, conforme previsto pelo AACR2.

Já os modelos Gemini e DeepSeek, no campo 260, embora tenham apresentado bons resultados de forma geral, deixaram de registrar a segunda cidade e editora em um dos casos analisados. Ademais, conforme mencionado anteriormente, o modelo DeepSeek adotou as normas do RDA, motivo pelo qual utilizou o campo 264 para descrever a área de publicação e distribuição, em vez do 260.

Além disso, o Gemini ao descrever os *ebooks* retirados da BDCD utilizou colchetes e pontos de interrogação para indicar ausência ou incerteza dos dados, recurso comum, mas que sugere dificuldade em localizar ou interpretar as informações presentes na fonte principal. Logo, o Quadro 7 apresenta as observações do campo 260:

Quadro 7 – Descrição do campo 260

Modelo de IA	Descrição gerada	Observação
ChatGPT	=260 \\\$aMarília :\$bOficina Universitária ;\$aSão Paulo :\$bCultura Acadêmica,\$c2021.	Seguindo corretamente todas as informações constantes na fonte principal e inserindo todas.
Gemini	260 ## \$a [Brasília] : \$b Edições Câmara, \$c [2019?].	Utiliza de colchetes e pontos de interrogação, demonstrando que a obra não tinha dados completos e incerteza sobre dados inseridos.
DeepSeek	260 ‡a São Paulo : ‡b Cultura Acadêmica, ‡c 2021.	Esse item apresenta mais uma cidade e editora, informação omitida.

Fonte: Comunicação pessoal com os modelos de IA – ChatGPT, Gemini e DeepSeek (2025)

Na análise do campo 300 (descrição física), o ChatGPT e Gemini apresentou descrições adequadas na maioria dos casos, mencionando corretamente o suporte digital. Portanto, o modelo Gemini omitiu informações de ilustrações em dois dos registros. O DeepSeek oscila nos registros, descreve dimensões da versão impressa dos *e-books*, e nos registros que assume o RDA, o modelo não faz as abreviações. Destaca-se que os modelos de IA tiveram dificuldade de executar esse campo com os *e-books* da BDCD, inserindo a contagem de página errada.

De maneira breve, sobre o campo 490 de série, ChatGPT e o DeepSeek registraram adequadamente o campo 490 sempre que necessário. O Gemini, por outro lado, utilizou inadequadamente o campo de nota para representar a série, não representando corretamente a série.

Passando para a análise do campo de notas, as notas são os campos que devem “contém informações úteis que não puderam ser incluídas nas outras áreas da descrição. Podem ser relativas a qualquer aspecto da apresentação do item ou do seu conteúdo. Devem ser claras e concisas e na língua da agência catalogadora” (Ribeiro, 2008).

Os modelos de IA apresentaram tanto acertos quanto inconsistências no preenchimento do campo de notas. O ChatGPT foi o modelo de IA que apresentou melhor desempenho nestes campos, apesar de ter sido o melhor entre os três, também cometeu erros, utilizando as notas para indicar responsabilidades (como prefaciadores), que deveriam ter sido inseridas no campo 7XX. Também, no ChatGPT e DeepSeek, foi identificado o uso do campo 500 para registrar conteúdos que deveriam estar no campo 504(Referências).

O Gemini preencheu corretamente poucos campos de notas, destacando-se no uso do campo 506 (restrições de acesso), que evidenciava que os *e-books* eram de “acesso livre”. Em contrapartida, utilizou o campo de notas para inserir links de acesso - informação mais adequada ao campo 856 - e aplicou de forma equivocada o campo 546(idioma), mesmo quando os itens possuíam apenas um idioma.

Já DeepSeek foi o único modelo a aplicar o campo 505 (conteúdo). Portanto, possui erros similares aos do Gemini, utilizando as notas para inserir campos de acesso. Como acerto, o DeepSeek e ChatGPT apresentaram informação correspondente ao item no campo 520 (Resumo), com resumos sucintos.

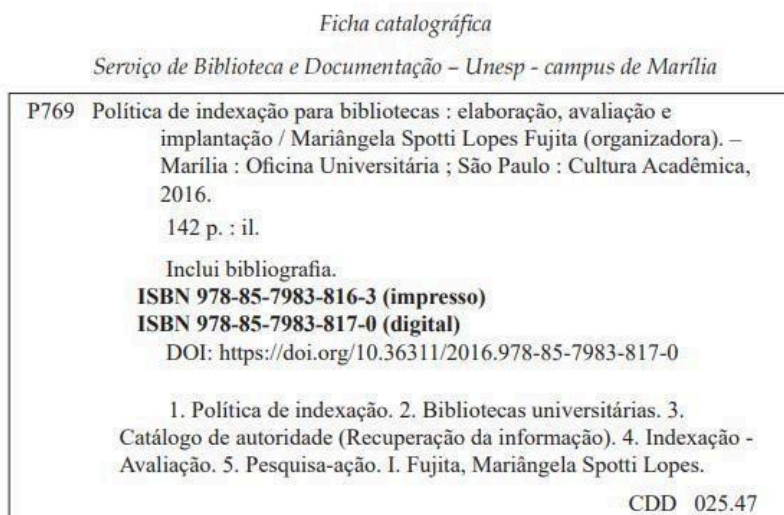
Percebe-se que os modelos de IA utilizaram os campos de notas para inserir informações irrelevantes ou que deveriam ter sido alocadas em outros campos, comprometendo a qualidade e coerência dos registros.

Para a análise do campo de assuntos tópicos (650), considerou-se os termos que eram coerentes com o conteúdo dos itens. Todos os modelos demonstraram bom desempenho neste critério. Gemini e DeepSeek utilizaram corretamente os subcampos disponíveis, ao passo que o ChatGPT se restringiu ao subcampo \$a, apresentando informações nesse subcampo que seriam mais adequadas aos outros subcampos do campo.

Analisou-se, também, se as IA retiraram os termos diretamente da ficha catalográfica presente nas obras ou se também se basearam no conteúdo do livro. Observou-se que tanto o ChatGPT quanto o DeepSeek apresentaram pelo menos dois assuntos tópicos, em conformidade com os pontos de acesso de assunto indicados na ficha catalográfica do livro, sendo que a omissão ocorreu apenas nos casos de livros editados pela BDCD, cujas fichas catalográficas traziam apenas um ponto de acesso de assunto.

Em um dos casos, o modelo de IA – ChatGPT – retirou todos os três assuntos tópicos exatamente conforme apresentados na ficha catalográfica da Figura 1:

Figura 1 – Ficha catalográfica do livro “Política de indexação para bibliotecas”



Fonte: Fujita (2016)

O Quadro 8 apresenta os campos 650 (assuntos tópicos) gerado pelas IA na descrição bibliográfica deste item da Figura 1:

Quadro 8 – Descrição do campo 650.

<b>Modelo de IA</b>	<b>Descrição gerada</b>
ChatGPT	=650 \$\$\$aPolítica de indexação. =650 \$\$\$aBibliotecas universitárias. =650 \$\$\$aIndexação – Avaliação.
Gemini	650 #4 \$a Indexação (Documentação). 650 #4 \$a Biblioteconomia. 650 #4 \$a Políticas de informação.
DeepSeek	650 7 _a Política de indexação. 650 7 _a Bibliotecas universitárias. 650 7 _a Catalogação de assuntos.

Fonte: Comunicação pessoal com os modelos de IA – ChatGPT, Gemini e DeepSeek (2025)

Os erros na entrada principal (1XX) refletiram diretamente nas aplicações das entradas secundárias, que obras com mais de três autores responsáveis, sem responsabilidade principal indicada, faz a entrada secundária para o primeiro autor citado na fonte principal de informação e organizadores entram por secundária.

O ChatGPT deixou de inserir prefaciadores no campo 7XX. O Gemini inseriu as editoras como entrada secundária em todos os registros analisados. O DeepSeek também incluiu editoras no campo 7XX, mas não o fez de forma consistente em todos os itens. Nesses casos de editoras mencionadas as IA agiram de forma errônea, pois não cabia aos itens ter essas informações como destaque de entrada secundária.

A seguir, apresenta-se o Quadro 9, que reúne os registros bibliográficos completos gerados por cada modelo de IA para um mesmo e-book da amostra analisada. O objetivo do quadro é destacar, de forma visual e objetiva, as diferenças no preenchimento dos campos, evidenciando aspectos já mencionados na análise dos dados.

Quadro 9 – Registros bibliográficos completos gerados por ChatGPT, Gemini e DeepSeek

ChatGPT	Gemini	DeepSeek
<p>=020 \\\$a9786586546903 (e-book)</p> <p>=100 1\\\$aJorente, Maria José Vicentini.</p> <p>=245 14\$aA emergência do design da informação na contemporaneidade da ciência da informação /\$cMaria José Vicentini Jorente, Natalia Nakano, Mariana Cantisani Padua.</p> <p>=250 \\\$a2. ed.</p> <p>=260 \\\$aMarília :\$bOficina Universitária ;\$aSão Paulo :\$bCultura Acadêmica,\$c2020.</p> <p>=300 \\\$a1 recurso online (202 p.) :\$bil.</p> <p>=500 \\\$aObra publicada com apoio da CAPES e do Programa de Pós-graduação em Ciência da Informação - UNESP.</p> <p>=504 \\\$aInclui bibliografia.</p> <p>=520 \\\$aDiscussão sobre os fundamentos e aplicações do design da informação na ciência da informação contemporânea, incluindo aspectos visuais, cognitivos e tecnológicos.</p> <p>=650 \\\$aDesign da informação.</p> <p>=650 \\\$aCiência da informação.</p> <p>=650 \\\$aComunicação visual.</p> <p>=700 1\\\$aNakano, Natalia,\$eautor.</p> <p>=700 1\\\$aPadua, Mariana Cantisani,\$eautor. =856 40\$uhttps://books.scielo.org/id/xjr8k\$yTexto completo</p> <p>=856 42\$uhttps://doi.org/10.36311/2020.978-6586546-90-3\$Ydoi</p>	<p>020 ## \$a 978-65-86546-90-3 (e-book)</p> <p>100 1# \$a Jorente, Maria José Vicentini, \$e autor. 245 10 \$a A emergência do Design da Informação na contemporaneidade da Ciência da Informação / \$c Maria José Vicentini Jorente, Natalia Nakano, Mariana Cantisani Padua.</p> <p>250 ## \$a 2. edição.</p> <p>260 ## \$a Marília : \$b Oficina Universitária ; \$a São Paulo : \$b Cultura Acadêmica, \$c 2020. 300 ## \$a 1 recurso online (202 p.) 500 ## \$a Título da tela.</p> <p>500 ## \$a "SciELO Books / SciELO Livros / SciELO Libros."</p> <p>500 ## \$a Disponível em: <a href="https://books.scielo.org/id/xjr8k">https://books.scielo.org/id/xjr8k</a>. Acesso em: 17 jun. 2025.</p> <p>506 ## \$a Acesso livre.</p> <p>546 ## \$a Em português.</p> <p>650 #4 \$a Design da informação.</p> <p>650 #4 \$a Ciência da informação.</p> <p>650 #4 \$a Sociedade da informação.</p> <p>700 1# \$a Nakano, Natalia, \$e autor.</p> <p>700 1# \$a Padua, Mariana Cantisani, \$e autor.</p> <p>710 2# \$a Oficina Universitária (Editora).</p> <p>710 2# \$a Cultura Acadêmica (Editora).</p> <p>710 2# \$a SciELO Books.</p> <p>856 40 \$u <a href="https://books.scielo.org/id/xjr8k">https://books.scielo.org/id/xjr8k</a> \$z Acesso livre.</p>	<p>020 _a978-65-86546-90-3</p> <p>1001 _aJorente, Maria José Vicentini</p> <p>24514 _aA emergência do Design da Informação na contemporaneidade da Ciência da Informação _h[recurso eletrônico] / _cMaria José Vicentini Jorente, Natalia Nakano, Mariana Cantisani Padua.</p> <p>250 _a2ª ed.</p> <p>260 _aMarília : _bOficina Universitária ; _aSão Paulo : _bCultura Acadêmica, _c2020.</p> <p>300 _a1 arquivo digital (202 p.) _bil.</p> <p>500 _aPublicado também em formato impresso.</p> <p>500 _aInclui bibliografia.</p> <p>520 _aDiscute a emergência do Design da Informação como área interdisciplinar com a Ciência da Informação, abordando aspectos teóricos, metodológicos e aplicados na organização e representação da informação em ambientes digitais.</p> <p>538 _aRequer navegador de internet.</p> <p>546 _aTexto em português, com resumos em inglês e espanhol.</p> <p>650 0 _aDesign da informação.</p> <p>650 0 _aCiência da informação.</p> <p>650 0 _aTecnologia da informação.</p> <p>7001 _aNakano, Natalia</p> <p>7001 _aPadua, Mariana Cantisani</p> <p>7102 _aUniversidade Estadual Paulista "Júlio de Mesquita Filho". _bFaculdade de Filosofia e Ciências.</p> <p>85640 _uhttps://books.scielo.org/id/xjr8k _zAcesso online via SciELO Books</p>

Fonte: Comunicação pessoal com os modelos de IA – ChatGPT, Gemini e DeepSeek (2025).

Para finalizar a análise, apresenta-se a seguir uma síntese comparativa entre os três modelos de IA generativa avaliados – ChatGPT, Gemini e DeepSeek. O Quadro 10 resume os principais pontos observados quanto à precisão, consistência e aderência às normas catalográficas, permitindo visualizar de forma consolidada os padrões de acerto e erro identificados em cada sistema, categorizados com os critérios pré-estabelecidos.

Quadro 10 - Comparação entre os modelos de IA generativa (ChatGPT, Gemini e DeepSeek)

<b>Critério de Avaliação</b>	<b>ChatGPT</b>	<b>Gemini</b>	<b>DeepSeek</b>
Aderência ao AACR2	Positivo – significativa conformidade com o código	Parcialmente positivo – não seguiu abreviações adequadas do AACR.	Insatisfatório – Oscilação entre padrões AACR2/RDA.
Entrada principal (1XX)	Parcialmente positivo – erro na entrada de organizadores e obras com mais de três autores.	Parcialmente positivo – erro na entrada de organizadores e obras com mais de três autores.	Parcialmente positivo – erro na entrada de organizadores e obras com mais de três autores.
Título e indicação de responsabilidade (245)	Positivo – erros no segundo indicador.	Positivo – erros no segundo indicador.	Parcialmente positivo – erros no segundo indicador e inconsistência na descrição.
Edição (250)	Parcialmente positivo – inserção indevida de "1. ed.".	Insatisfatório – não aplicou abreviações conforme o AACR2.	Parcialmente positivo – oscilação entre certo e errado.
Área da publicação, distribuição (260)	Excelente – descrição correta e completa.	Parcialmente positivo – omitiu segunda cidade e editora em um registro. Utilização de colchetes em <i>ebooks</i> da BDCD.	Insatisfatório – omitiu segunda cidade e editora em um registro. Oscilação entre o campo 264, comum do RDA.
Descrição física (300)	Positivo – poucas falhas na contagem de páginas em títulos da BDCD.	Parcialmente positivo – contagem incorreta de páginas em obras da BDCD e ausência de ilustração em dois registros.	Insatisfatório – erro na paginação de obras da BDCD, uso incorreto do padrão para recursos online e dimensões físicas.
Série (490)	Excelente – descrição correta e completa.	Insatisfatório – inseriu série no campo de nota.	Excelente – descrição correta e completa.
Notas (5XX)	Parcialmente positivo – uso inadequado para indicação de responsabilidade, mas acertos com o campo 520.	Insatisfatório – informações irrelevantes, indicação de série e uso inadequado para links e acessos.	Parcialmente positivo – uso inadequado para links e acesso, mas acertos com o campo 505 e 520.
Assunto tópico (650)	Positivo – termos coerentes com a obra, mas uso limitado de subcampos.	Excelente – termos coerentes, com bom uso dos subcampos.	Excelente – termos coerentes, com bom uso dos subcampos.

Entradas secundárias (7XX)	Insatisfatório – deixou de inserir organizadores e prefaciadores.	Insatisfatório - deixou de inserir organizadores e incluiu de maneira errônea editoras.	Insatisfatório - deixou de inserir organizadores e inseriu editoras de forma inconsistente.
----------------------------	---	---	---

Fonte: Elaboração própria (2025)

Com base na sistematização apresentada, observa-se que os três modelos de IA generativa analisados demonstram níveis distintos de desempenho técnico e aderência normativa. O ChatGPT apresentou maior conformidade com o AACR2 na maior parte dos registros.

O Gemini, apesar de utilizar alguns campos corretamente, teve maior dificuldade em organizar os dados de forma coesa e cometeu erros conceituais, como o uso inadequado do campo de notas. Já o DeepSeek, embora tenha gerado descrições, mostrou uma tendência clara ao uso do padrão RDA, contrariando as instruções dadas. Isso revela uma variabilidade significativa nos comportamentos dos modelos, refletindo diferenças em suas arquiteturas e bases de treinamento.

Os resultados obtidos demonstram que, embora as IA generativas consigam simular registros bibliográficos em formato MARC 21 com considerável precisão, ainda enfrentam obstáculos para alcançar a mesma qualidade dos registros elaborados por catalogadores humanos.

Aspectos abordados por Dumer e Albuquerque (2020), como a necessidade de considerar as variações entre os diferentes tipos de materiais presentes no acervo, evidenciam uma limitação recorrente nos modelos testados. As IA, por exemplo, não conseguiram aplicar adequadamente os critérios normativos para obras com organizadores ou com mais de três autores – situações que requerem decisões específicas conforme as regras do AACR2.

Complementar a essa limitação, os modelos apresentaram desempenho inferior ao lidar com publicações da BDCD nos campos 260 e 300, nas quais houve falhas frequentes, como descrição imprecisa. Esses resultados reforçam a importância da atuação humana no processo de representação descritiva, uma vez que os sistemas de IA ainda não são capazes de interpretar com a devida profundidade as particularidades de diferentes tipos de publicação e aplicar com consistência os instrumentos normativos.

Outro ponto importante a ser considerado diz respeito às características fundamentais da catalogação apontadas por Mey e Silveira (2009), como integridade, clareza, precisão, lógica e consistência. A análise dos registros gerados revelou que os modelos de IA

generativa ainda não atendem plenamente a esses princípios, especialmente no que se refere à consistência das informações e a lógica dos dados.

Em diversos casos observou-se que dados essenciais foram apresentados de forma incompleta ou ambígua, comprometendo a fidelidade das informações descritivas. O modelo DeepSeek, em particular, apresentou problemas recorrentes de inconsistência: ora aplicava corretamente regras do RDA, ora misturava elementos do AACR2, resultando em registros com padrões normativos híbridos e contraditórios. Tal comportamento compromete diretamente a característica de consistência.

Conforme apontado por Stryker e Scapicchio (2024), essa variabilidade decorre da natureza probabilística dos modelos de IA generativa, os quais podem produzir saídas distintas mesmo a partir de uma mesma entrada. No caso do DeepSeek, isso resultou em registros inconsistentes para obras com características similares, com diferenças na escolha de campos, na forma de entrada de autores e na aplicação das regras catalográficas.

Além disso, o Gemini apresentou dificuldades quanto à lógica na organização dos dados, especialmente no uso do campo de notas (500). O Gemini, por exemplo, incluiu anotações genéricas ou pouco úteis em registros nos campos de nota.

Portanto, a discussão aponta para a complementaridade entre IA e profissionais da informação, pois os modelos de IA são capazes de gerar apenas rascunhos de registros, mas a avaliação crítica, a aplicação normativa e a mediação contextual continuam sendo imprescindíveis do catalogador, com domínio nos campos MARC 21 e nos códigos de catalogação.

Stryker e Scapicchio (2024), propõem o uso seguro e eficaz da IA generativa pois exige que humanos estejam no centro da tomada de decisão, atuando como supervisores e validadores dos resultados produzidos pela máquina. Ainda assim, sobre as novas tecnologias de IA, Rodriguez e Prudencio (2024, p. 19) destacam “a necessidade de ampliar os processos de aprendizagem sobre essas tecnologias e sobre possíveis impactos de sua operação no mundo do trabalho informacional.”

## **6 CONSIDERAÇÕES FINAIS**

A pesquisa analisou o desempenho de três modelos de inteligência artificial generativa – ChatGPT, Gemini e DeepSeek – na catalogação de e-books em formato MARC 21, à luz das normas do AACR2. A comparação dos registros gerados permitiu avaliar a qualidade

técnica, a precisão descritiva e a aderência às regras normativas de cada modelo, identificando seus pontos fortes e limitações.

Esses achados indicam que a viabilidade da IA na catalogação ainda deve ser compreendida como uma atividade de apoio, ao gerar rascunhos incompletos dos registros. As limitações encontradas, como a variabilidade de respostas e a aplicação imprecisa de normas, confirmam a importância da atuação humana, sobretudo no que diz respeito à tomada de decisão técnica, à coerência descritiva e à curadoria das informações.

Como limitações, destaca-se o tamanho reduzido da amostra, restrita a e-books, e a avaliação voltada apenas para o AACR2, sem aplicação comparativa direta com outros códigos de catalogação. Além disso, os testes foram realizados em um recorte temporal específico, o que significa que eventuais atualizações dos modelos de IA podem alterar seu desempenho, é fundamental que novas testagens sejam realizadas periodicamente, a fim de acompanhar melhorias técnicas e reavaliar sua eficácia diante de atualizações tecnológicas.

Para estudos futuros, sugere-se ampliar a análise para outros tipos de materiais, incluir diferentes normas catalográficas, como o RDA, e explorar métricas para aferir o desempenho. Também se recomenda que novas testagens sejam realizadas periodicamente, acompanhando a evolução tecnológica e verificando se as melhorias implementadas pelos desenvolvedores se refletem em maior precisão e consistência na catalogação.

Espera-se que esta pesquisa contribua para o debate sobre a integração responsável dessas tecnologias na Biblioteconomia e estimule novos estudos voltados ao aprimoramento da catalogação automatizada.

## REFERÊNCIAS

ALMEIDA, Virgílio. **IA Generativa: oportunidades, riscos e governança**. [S. l.], 18 maio 2023. Apresentação de Power Point. Disponível em: <https://iagenerativa.ceweb.br/files/apresentacao/arquivo/1616/IA%20Generativa%20-%20Oportunidades,%20Riscos%20e%20Governan%C3%A7a.pdf>. Acesso em: 26 fev. 2025.

ALVES, Rachel Cristina Vesu. **Metadados como elementos do processo de catalogação**. 2010. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010. Disponível em: <http://hdl.handle.net/11449/103361>. Acesso em: 5 jul. 2025.

BALNAVES, Edmund *et al.* (ed.). **New Horizons in Artificial Intelligence in Libraries**. Berlin/Boston: De Gruyter Saur, 2025. (IFLA Publications; 185). Disponível em: <https://repository.ifla.org/items/0695f7ce-d945-468b-8768-737a8321df11>. Acesso em: 4 ago. 2025.

BRADY, Damian. **What developers need to know about generative AI**. GitHub Blog, 7 fev. 2024. Disponível em: <https://github.blog/ai-and-ml/generative-ai/what-developers-needto-know-about-generative-ai/>. Acesso em: 11 jul. 2025.

BREENDING, Marshall. Disruption in the library bibliographic services arena. **Library Technology Newsletter**, Nashville, v. 1, n.06, jun. 2022. Disponível em: <https://librarytechnology.org/document/27429>. Acesso em: 4 ago. 2025.

CÓDIGO de catalogação anglo-americano. 2. ed. rev. 2002. Preparado sob a direção do Joint Steering Committee for Revision of AACR. Tradução para língua portuguesa sob a responsabilidade da FEBAB. São Paulo: FEBAB / Imprensa Oficial do Estado, 2004.

CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robalinho de Oliveira. **Dicionário de Biblioteconomia e Arquivologia**. Brasília: Briquet de Lemos, 2008.

DUMER, Luciana; ALBUQUERQUE, Maria Elizabeth Baltar Carneiro de. **O MARC 21 e os outros formatos de intercâmbio bibliográfico**. João Pessoa: Editora UFPB, 2020. Disponível em: <https://www.editora.ufpb.br/sistema/press5/index.php/UFPB/catalog/download/179/911/7524>. Acesso em: 18 fev. 2025.

FERREIRA, Giuliano. O impacto da inteligência artificial na gestão do conhecimento em bibliotecas universitárias: desafios e soluções. *In*: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 22., 2023, Florianópolis, SC. **Anais** [...]. Florianópolis, SC, 2023. Disponível em: <https://portal.febab.org.br/snbu2023/article/view/2934>. Acesso em: 4 ago. 2025.

FUJITA, Mariângela Spotti Lopes (org.). **Política de indexação para bibliotecas: elaboração, avaliação e implantação**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2016.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GODINHO, Keila Ingrid dos Santos. **Inteligência Artificial em bibliotecas: Bibliotecária Informativa Automatizada (BIA) da Divisão de Bibliotecas e Documentação da PUC-Rio**. 2019. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) – Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2019. Disponível em: [https://bdm.unb.br/bitstream/10483/25870/1/2019\\_KeilaIngriddosSantosGodinho\\_tcc.pdf](https://bdm.unb.br/bitstream/10483/25870/1/2019_KeilaIngriddosSantosGodinho_tcc.pdf). Acesso em: 18 fev. 2025.

IFLA. **Declaração dos Princípios Internacionais de Catalogação (PIC)**. 2016. Disponível em: [https://cdn.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/icp/icp\\_2016-pt.pdf](https://cdn.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/icp/icp_2016-pt.pdf). Acesso em: 02 jul. 2025

KASPRZIK, Anna. Transferring Applied Machine Learning Research into Subject Indexing Practice. *In*: BALNAVES, Edmund *et al.* (ed.). **New Horizons in Artificial Intelligence in Libraries**. Berlin/Boston: De Gruyter Saur, 2025. p. 199–212. DOI: <https://doi.org/10.1515/9783111336435-015>. Disponível em:

<https://repository.ifla.org/items/0695f7ce-d945-468b-8768-737a8321df11>. Acesso em: 4 ago. 2025.

LIBRARY OF CONGRESS. Exploring Computational Description. **Library of Congress Labs**, [202-]. Disponível em: <https://labs.loc.gov/work/experiments/ECD/>. Acesso em: 4 jul. 2025.

LIBRARY OF CONGRESS. **The MARC 21 Formats**: background and principles. Washington, D.C., 1996. Disponível em: <https://www.loc.gov/marc/96principi.html>. Acesso em: 10 jul. 2025.

MACIEL, Leandro. Editorial: ChatGPT and the ethical aspects of artificial intelligence. **REGE - Revista de Gestão**, São Paulo, v. 30, n. 2, p. 110–112, 2023. DOI: 10.1108/REGE04-2023-207. Disponível em: <https://revistas.usp.br/rege/article/view/211585>. Acesso em: 10 jul. 2025.

MALMSTEN, Martin et al. Without Heading? Automatic Creation of a Linked Subject System. In: BALNAVES, Edmund et al. (ed.). **New Horizons in Artificial Intelligence in Libraries**. Berlin/Boston: De Gruyter Saur, 2025. p. 179–198. DOI: <https://doi.org/10.1515/9783111336435-014>. Disponível em: <https://repository.ifla.org/items/0695f7ce-d945-468b-8768-737a8321df11>. Acesso em: 4 ago. 2025.

MEY, Eliane Serrão Alves.; SILVEIRA, Naira Christofolleti. **Catálogo no plural**. Brasília: Briquet de Lemos, 2009.

PICCO, Paola; ORTIZ REPISO, Virginia. RDA, el nuevo código de catalogación: cambios y desafíos para su aplicación. **Revista Española de Documentación Científica**, [S. l.], v. 35, n. 1, p. 145–173, 2012. DOI: 10.3989/redc.2012.1.848. Disponível em: <https://redc.revistas.csic.es/index.php/redc/article/view/727>. Acesso em: 10 jul. 2025.

RIBEIRO, Antônia Motta de Castro Memória. **Catálogo de recursos bibliográficos: AACR2R em MARC 21**. 3 ed. Brasília: Ed. do Autor: 2008.

RODRIGUEZ, Gabriela Mendes; PRUDENCIO, Dayanne da Silva. O que se discute sobre a aplicação de inteligência artificial em bibliotecas? análise da produção do campo informacional. **P2P & Inovação**, Rio de Janeiro, v. 11, n. 1, p. 1-24, jul./dez. 2024. Disponível em: <https://brapci.inf.br/v/309170>. Acesso: 9 set. 2025.

SACCUCCI, Caroline; POTTER, Abigail. Assessing Machine Learning for Cataloging at the Library of Congress. In: BALNAVES, Edmund et al. (ed.). **New Horizons in Artificial Intelligence in Libraries**. Berlin/Boston: De Gruyter Saur, 2025. p. 227–238. DOI: <https://doi.org/10.1515/9783111336435-017>. Disponível em: <https://repository.ifla.org/items/0695f7ce-d945-468b-8768-737a8321df11>. Acesso em: 4 ago. 2025.

SILVA, Janete Fernandes. A integração da inteligência artificial na biblioteconomia: um caminho em construção. **Código 31**, v. 2, n. 1, p. 26-46, jan./jun. 2024. Disponível em: <https://revista.fumec.br/index.php/codigo31/article/view/9842>. Acesso em: 19 fev. 2025

SILVA, Narjara; NATHANSON, Bruno Macedo. Análise da produção científica em inteligência artificial na área da Ciência da Informação no Brasil. *In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO*, 19., 2018, Londrina, PR. **Anais [...]**, 2018.

STRYKER, Cole; HOLDSWORTH, Jim. O que é PNL(processamento de linguagem natural)? **IBM**, 11 ago. 2024. Disponível em: <https://www.ibm.com/brpt/think/topics/natural-language-processing>. Acesso em: 4 ago. 2025.

STRYKER, Cole; SCAPICCHIO, Mark. O que é IA generativa?. **IBM**, 22 mar. 2024. Disponível em: <https://www.ibm.com/br-pt/think/topics/generative-ai>. Acesso em: 4 ago. 2025.

TEIXEIRA, Heytor Diniz. **Disseminação do RDA no Brasil**: estudo acerca da produção científica brasileira sobre o tema e a implementação do novo código de catalogação nas bibliotecas universitárias do país. 2019. Monografia (Bacharelado em Biblioteconomia) – Instituto de Ciências Humanas e da Informação, Universidade Federal do Rio Grande – FURG, Rio Grande, 2019. Disponível em: <https://repositorio.furg.br/handle/1/10065>. Acesso em: 25 jun. 2025.

WHEATLEY, Amanda; HERVIEUX, Sandy. Artificial intelligence in academic libraries: an environmental scan. **Information Services and Use**, Montreal, v. 39, n. 4 p. 347-356, nov. 2019. Disponível em: <https://journals.sagepub.com/doi/epub/10.3233/ISU-190065>. Acesso em: 10 jul. 2025.