

GT1 – Catalogação Descritiva Contemporânea: Metadados, Ética, IA e Softwares de Biblioteca e Repositórios Digitais e Ciência Aberta.

INTELIGÊNCIA ARTIFICIAL E METADADOS SEMÂNTICOS

Desafios para a Catalogação Descritiva Contemporânea

ARTIFICIAL INTELLIGENCE AND SEMANTIC METADATA

Challenges for Contemporary Descriptive Cataloging

Brenda de Melo Silva¹

Carlos Henrique Almeida²

Rosangela Zanforlin de Almeida³

RESUMO

O presente trabalho descreve o desenvolvimento de um modelo de serviço de Disseminação Seletiva da Informação (DSI) suportado por Inteligência Artificial (IA) no Instituto de Pesquisas Tecnológicas (IPT). Diante da sobrecarga informacional e do crescimento do volume de patentes e normas técnicas, a catalogação descritiva tradicional apresenta limitações na recuperação proativa. A metodologia utiliza Processamento de Linguagem Natural (PLN) para a geração de metadados semânticos e criação de perfis informacionais dinâmicos. O ciclo de validação baseia-se em metodologias ágeis de quatro semanas, focando na curadoria humana como elemento indispensável para mitigar riscos como "bolhas de filtro" e perda de serendipidade. Conclui-se que o modelo DSI-IA otimiza o monitoramento de inovações e redefine o papel do especialista da informação como curador de sistemas inteligentes.

Palavras-chave: Inteligência Artificial; Metadados Semânticos; DSI; Catalogação Descritiva; IPT.

¹Bacharel em Biblioteconomia pela Escola de Comunicações e Artes da Universidade de São Paulo (ECA-USP). Foi estagiária no Instituto de Pesquisas Tecnológicas (IPT), na GITEB, até dezembro de 2025. E-mail: brendamelo@usp.br.

² Bacharel em Biblioteconomia pela Escola de Comunicações e Artes da Universidade de São Paulo (ECA-USP). Foi estagiário no Instituto de Pesquisas Tecnológicas (IPT), na GITEB, até dezembro de 2025. E-mail: carlos.anis.silva@usp.br.

³ Pesquisadora assistente e Chefe da Seção de Planejamento e Gestão da Informação Tecnológica e Bibliográfica (GITEB) do Instituto de Pesquisas Tecnológicas (IPT). E-mail: zanfor@ipt.br.

ABSTRACT

This paper describes the development of a Selective Dissemination of Information (SDI) service model supported by Artificial Intelligence (AI) at the Institute for Technological Research (IPT). Given the information overload and the increasing volume of patents and technical standards, traditional descriptive cataloging presents limitations in proactive recovery. The methodology utilizes Natural Language Processing (NLP) for the generation of semantic metadata and the creation of dynamic informational profiles. The validation cycle is based on four-week agile methodologies, focusing on human curation as an essential element to mitigate risks such as "filter bubbles" and the loss of serendipity. It is concluded that the AI-SDI model optimizes the monitoring of innovations and redefines the role of the information specialist as a curator of intelligent systems.

Keywords: Artificial Intelligence; Semantic Metadata; SDI; Descriptive Cataloging; IPT.

Data de envio do artigo: 23/01/2026.

1 INTRODUÇÃO

O cenário contemporâneo da pesquisa de vanguarda enfrenta o obstáculo da dispersão de publicações científicas, patentes e normas técnicas. Softwares de biblioteca tradicionais, dependentes de metadados estáticos, mostram-se insuficientes para lidar com a complexidade e o volume dos repositórios digitais atuais. Parte-se da premissa de que a integração de inteligência artificial à curadoria humana pode converter a recuperação de dados em um sistema de antecipação de demandas. Diante disso, a Gerência de Gestão da Informação Tecnológica e Bibliográfica (GITEB)⁴ do IPT⁵ propõe evoluir da recuperação reativa para um modelo de mediação proativo. O objetivo deste artigo é descrever o desenvolvimento deste modelo de Disseminação Seletiva da Informação (DSI) suportado por IA, analisando sua eficácia no monitoramento de inovações tecnológicas.

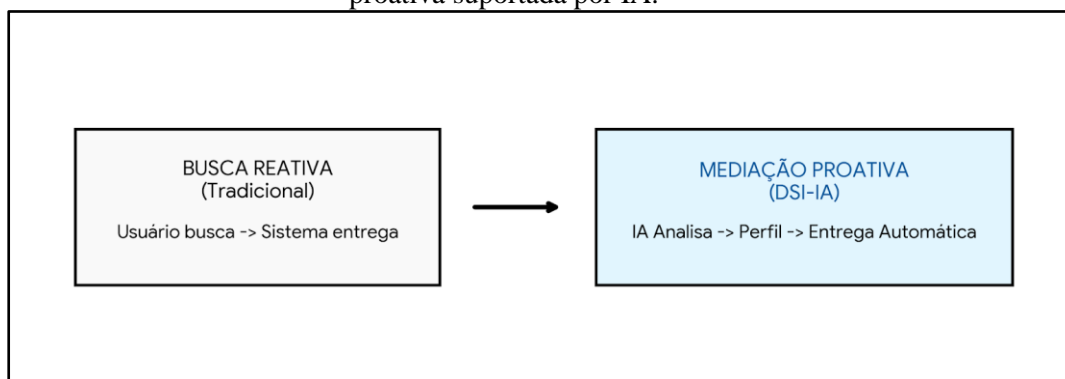
2 O PROBLEMA DA BUSCA REATIVA

⁴ A GITEB é a unidade do IPT responsável pela curadoria, preservação e disseminação do acervo técnico-científico da instituição, atuando na gestão de normas técnicas, patentes e publicações especializadas para suporte à inovação.

⁵ O Instituto de Pesquisas Tecnológicas (IPT) é um dos maiores institutos de pesquisa do Brasil, vinculado à Secretaria de Ciência, Tecnologia e Inovação do Estado de São Paulo. Atua há mais de 120 anos no suporte tecnológico à indústria e na gestão de informações técnicas de vanguarda.

A recuperação tradicional depende da consulta explícita do usuário, o que gera riscos estratégicos, como o não monitoramento de inovações críticas. Para o pesquisador, a otimização do tempo gasto na busca é um fator decisivo. Na Ciência da Informação, o desafio reside em incorporar camadas de análise semântica à catalogação para agregar valor informacional e garantir que a informação relevante chegue ao interessado de forma automática. Essa mudança de paradigma é fundamental para instituições de tecnologia. Enquanto o modelo tradicional sobrecarrega o pesquisador com buscas manuais exaustivas, a integração de camadas inteligentes permite que o sistema antecipe necessidades. A eficácia dessa transição, que move o eixo da busca do usuário para o sistema, é detalhada graficamente na Figura 1.

Figura 1 – Evolução do modelo de recuperação de informação: da busca reativa à mediação proativa suportada por IA.



Fonte: Elaboração própria (2026)

3 REFERENCIAL TEÓRICO: A TRANSIÇÃO PARA A CATALOGAÇÃO SEMÂNTICA

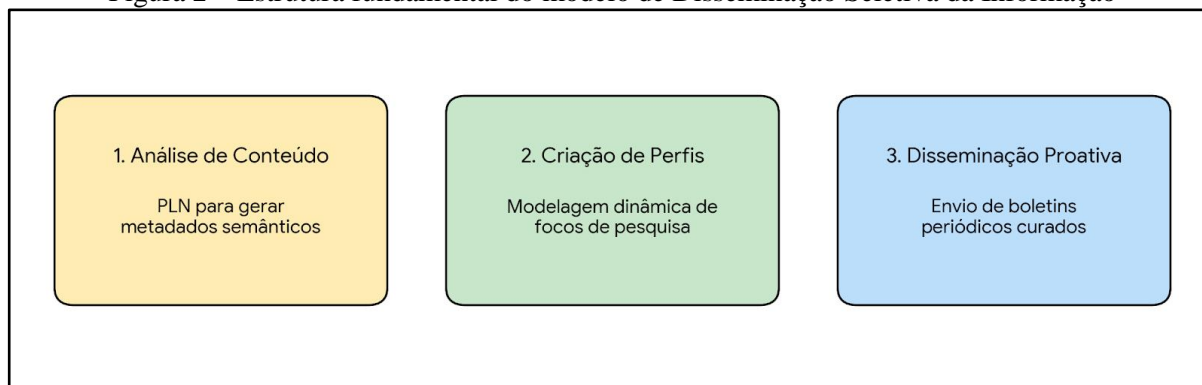
A catalogação descritiva contemporânea enfrenta o desafio de transpor a barreira da representação física para a representação do conhecimento. Enquanto os modelos tradicionais dependem de metadados estáticos, a evolução para a Disseminação Seletiva da Informação (DSI) permite que a informação "encontre" o usuário (Lancaster, 2004). O uso de Processamento de Linguagem Natural (PLN) visa gerar metadados semânticos, como identificação de entidades e tópicos emergentes. Conforme apontam Manning e Schütze (1999), o PLN é essencial para que sistemas capturem nuances que registros legados ignoram sem intervenção manual exaustiva.

4 METODOLOGIA: O SERVIÇO DSI-IA E O FLUXO GITEB/IPT

O serviço opera sustentado por três pilares fundamentais que compõem a base da arquitetura informacional da GITEB. A eficácia do modelo não depende apenas do

processamento isolado, mas da coexistência dessas frentes de atuação para transformar registros técnicos em conhecimento acionável. Esses componentes — Análise de Conteúdo, Criação de Perfis e Disseminação Proativa — estão estruturados em três eixos principais, apresentados na Figura 2.

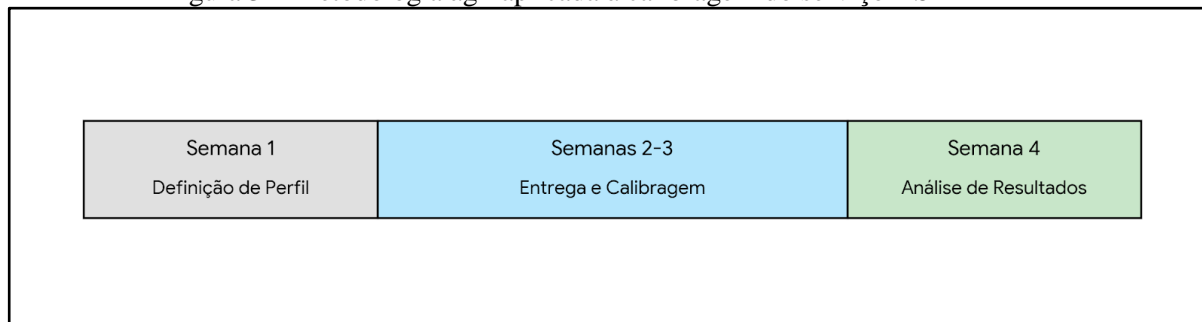
Figura 2 – Estrutura fundamental do modelo de Disseminação Seletiva da Informação



Fonte: Elaboração própria (2026)

A estrutura de trabalho foi baseada em Metodologias Ágeis, com ciclos de quatro semanas. O fluxo operacional dessas etapas, desde o mapeamento inicial até a entrega final, é sistematizado na Figura 3, detalhada a seguir:

Figura 3 – Metodologia ágil aplicada à calibragem do serviço DSI-IA



Fonte: Elaboração própria (2026)

- **Semana 1 (Definição):** Mapeamento do perfil e seleção de fontes.
- **Semanas 2 e 3 (Execução):** Ciclos de entrega de boletins. O sistema sugere documentos via algoritmos, que passam pela validação do bibliotecário.
- **Semana 4 (Consolidação):** Feedback e refinamento do algoritmo.

5 DIMENSÕES ÉTICAS E O PAPEL DA CURADORIA HUMANA

A implementação de IA na catalogação exige transparência. O sistema não substitui o profissional; a curadoria humana é o que combate o "viés de confirmação". Pariser (2011) alerta

para o risco das "Bolhas de Filtro", onde a alta personalização diminui a serendipidade (descobertas casuais). No IPT, o bibliotecário atua como um auditor da IA, garantindo a diversidade informacional.

6 RESULTADOS E IMPACTO INSTITUCIONAL

A implementação do modelo DSI-IA no IPT aponta para uma redução no tempo de busca e um aumento na precisão do monitoramento de patentes. Ao funcionar como uma camada de inteligência sobre os sistemas de biblioteca existentes, o serviço garante que inovações normativas não passem despercebidas, protegendo a propriedade intelectual e fomentando a pesquisa científica.

7 CONSIDERAÇÕES FINAIS

O presente trabalho buscou enfrentar o desafio da dispersão de publicações e normas técnicas por meio de um modelo de Disseminação Seletiva da Informação suportado por IA. Os resultados da aplicação no IPT indicam que os objetivos propostos foram atingidos, uma vez que a integração entre Processamento de Linguagem Natural e curadoria humana permitiu converter a recuperação reativa em um sistema de antecipação de demanda. Confirmou-se a premissa de que a tecnologia, quando validada por especialistas, reduz o tempo de busca e amplia a precisão do monitoramento estratégico.

Conclui-se que o modelo DSI-IA é promissor para otimizar o acesso à informação, posicionando o bibliotecário como um validador essencial de sistemas inteligentes. O desafio futuro reside em aperfeiçoar algoritmos para aumentar a diversidade das recomendações e expandir o serviço para outras áreas do conhecimento.

REFERÊNCIAS

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Briquet de Lemos, 2004.

MANNING, C. D.; SCHÜTZE, H. *Foundations of statistical natural language processing*. MIT Press, 1999.

PARISER, Eli. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press, 2011.