



25°enancib

Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação
Informação, Decolonialidade e Direitos Difusos para o Desenvolvimento Sustentável

3 a 7 de novembro de 2025 – Rio de Janeiro



XXV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – XXV ENANCIB

GT 2 – Organização e Representação do Conhecimento

AMPLIAÇÃO DE ESTRUTURAS SEMÂNTICAS EM DOMÍNIOS ESPECIALIZADOS: UM ESTUDO DE CASO NA INTENSIFICAÇÃO AGROPECUÁRIA

EXPANSION OF SEMANTIC STRUCTURES IN SPECIALIZED DOMAINS: A CASE STUDY ON AGRICULTURAL INTENSIFICATION

Vinícius Corrêa Braga – Universidade Federal de Minas Gerais (UFMG)
Gercina Ângela de Lima – Universidade Federal de Minas Gerais (UFMG)
Lucinéia Souza Maia – Universidade Federal de Ouro Preto (UFOP)

Modalidade: Trabalho completo

Resumo: a intensificação agropecuária demanda estruturas conceituais capazes de representar, com precisão, as relações entre os termos do domínio. Este trabalho teve como objetivo aplicar a ferramenta *Semantizar* para identificar e propor relações semânticas nesse contexto, utilizando como corpus o livro *Evolução Agrária e Pressão Demográfica*. Adotou-se a metodologia de estudo de caso, fundamentada em estratégia mista que combinou análise qualitativa das relações semânticas com a quantificação dos resultados, por meio da comparação entre a saída da ferramenta e uma estrutura de referência previamente validada por especialistas. O corpus textual foi convertido para formato compatível com ferramentas de Processamento de Linguagem Natural (PNL) e analisado a partir de uma lista de conceitos selecionados. O protocolo metodológico contemplou as etapas de análise textual, extração de dados, elaboração de regras, aplicação da ferramenta e avaliação das relações geradas. Como resultado, foram identificadas vinte e oito novas relações semânticas, classificadas conforme diretrizes internacionais. Conclui-se que o *Semantizar* contribui para o enriquecimento de estruturas conceituais especializadas, favorecendo a organização do conhecimento com clareza, coerência e relevância semântica.

Palavras-chave: extração de relações; relações semânticas; estruturas conceituais; *Semantizar*.

Abstract: the intensification of agricultural production requires conceptual structures capable of accurately representing the relationships between terms in the domain. This study aimed to apply the *Semantizar* tool to identify and propose semantic relations in this context, using the book *Evolução Agrária e Pressão Demográfica* as the corpus. A case study methodology was adopted, based on a mixed strategy that combined qualitative analysis of semantic relations with the quantification of results, by comparing the tool's output with a reference structure previously validated by experts. The textual corpus was converted into a format compatible with Natural Language Processing (NLP) tools and analyzed based on a selected list of concepts. The methodological protocol encompassed the stages of textual analysis, data extraction, rule development, tool application, and evaluation of the generated relations. As

a result, twenty-eight new semantic relations were identified, classified according to international guidelines. It is concluded that *Semantizar* contributes to the enrichment of specialized conceptual structures, fostering knowledge organization with clarity, coherence, and semantic relevance.

Keywords: relation extraction; semantic relations; conceptual structures; *Semantizar*.

1 INTRODUÇÃO

A crescente complexidade e o volume de dados em ambientes digitais impõem desafios significativos aos profissionais da Ciência da Informação, exigindo estratégias cada vez mais eficientes para a organização e a recuperação da informação. Nesse contexto, as Estruturas Conceituais (EC) desempenham um papel central ao sistematizar o conhecimento e permitir conexões semânticas consistentes entre conceitos. A automatização da Extração de Relações Semânticas (ERS) surge como alternativa promissora frente às limitações da construção manual dessas estruturas, contribuindo para maior escalabilidade, consistência e atualização contínua. Esta pesquisa investiga o potencial da ferramenta *Semantizar* na identificação e proposição de Relações Semânticas (RS) em textos em linguagem natural, com ênfase no domínio da Intensificação Agropecuária (IA).

Apesar dos avanços proporcionados por ferramentas de ERS como o *Semantizar*, persistem incertezas quanto à sua eficácia em diversos aspectos fundamentais. Questiona-se se a ferramenta é capaz de propor relações além das estruturas previamente estabelecidas, se tais relações estão de acordo com as recomendações normativas, como a ANSI/NISO Z39.19-2005 (R2010). Além disso, busca-se, ainda, verificar se a estrutura do vocabulário gerada apresenta uma organização lógica e navegabilidade adequada, se a ferramenta atua de forma eficaz no domínio específico da IA e se adota critérios adequados na seleção de termos, priorizando relevância, precisão e consistência. Diante disso, formula-se a seguinte questão de pesquisa: até que ponto a ferramenta *Semantizar* é adequada para apoiar a construção de EC por meio da ERS em textos redigidos em linguagem natural?

O objetivo desta pesquisa consiste em avaliar a ferramenta *Semantizar* quanto à sua capacidade de propor relações semânticas além da estrutura conceitual previamente definida; verificar seu alinhamento com as recomendações da Norma ANSI/NISO Z39.19-2005 (R2010); analisar sua aptidão para categorizar as relações em hierárquicas, associativas e de equivalência; examinar a coerência da estrutura do vocabulário gerada; averiguar sua conformidade com as diretrizes para identificação e organização de relações semânticas no

domínio específico da IA; e avaliar os princípios adotados para a seleção de termos, com foco na relevância, precisão e consistência das conexões estabelecidas.

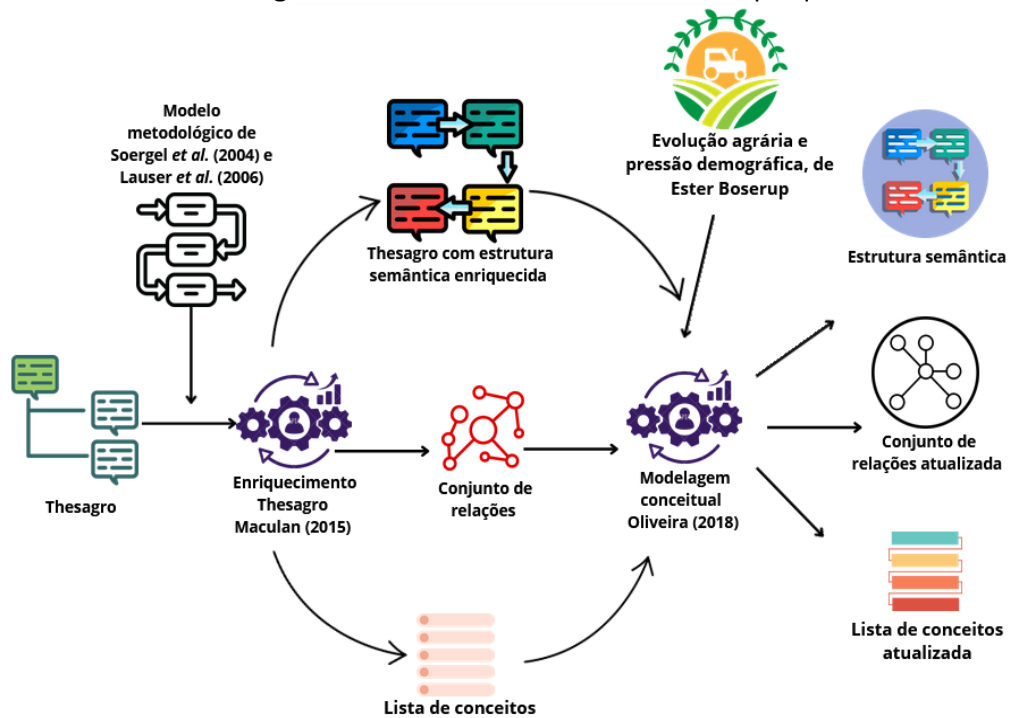
A construção manual de EC configura-se como uma tarefa complexa, morosa e suscetível a inconsistências. A automação da ER, especialmente em língua portuguesa, representa um avanço expressivo para a Ciência da Informação, ao possibilitar o tratamento de grandes volumes de dados textuais. A escolha da ferramenta *Semantizar* justifica-se por sua especificidade para o português e por sua aderência a normas internacionais, como a ANSI/NISO Z39.19-2005, além de seu potencial para sugerir relações semânticas inéditas, contribuindo para o enriquecimento das EC e para ampliação das possibilidades de representação do conhecimento em domínios especializados.

2 METODOLOGIA

A pesquisa caracteriza-se como de natureza aplicada, pois visa a utilização prática de conhecimentos para solucionar problemas concretos, promovendo, conforme Boaventura (2004), práticas mais eficazes. Quanto aos fins, classifica-se como pesquisa exploratória, por buscar proporcionar maior familiaridade com o problema investigado (Gil, 2002, p. 43). A aplicação da ferramenta *Semantizar*, comparada a um padrão ouro, teve como finalidade explorar sua eficácia, sem a formulação de hipóteses rígidas. Em relação à abordagem do problema, adotou-se uma estratégia mista, combinando análise qualitativa das Relações Semânticas (RS) com a quantificação dos resultados obtidos, em termos de precisão e desempenho. Do ponto de vista técnico, trata-se de um estudo de caso, modalidade que segundo Prodanov e Freitas (2013) e Gil (2010, p. 37), permite analisar fenômenos contemporâneos dentro de seu contexto real, examinando uma ou mais unidades de análise.

O universo da pesquisa corresponde ao domínio da agropecuária brasileira, com recorte no *Thesaurus Agrícola Nacional* (THESAGRO), único tesouro nacional da área. A escolha fundamenta-se na afinidade com o experimento de Soergel *et al.* (2004) com o AGROVOC, tesouro agrícola americano. A Figura 1 apresenta a linha histórica e elementos das pesquisas referenciadas.

Figura 1 - Linha histórica e elementos das pesquisas



Fonte: elaborada pelos autores (2024).

Maculan (2015) aplicou ao THESAGRO um modelo de reengenharia de tesouros, baseado em Soergel *et al.* (2004) e Lauser *et al.* (2006), com o objetivo de ampliar a estrutura semântica do domínio da Intensificação Agropecuária (IA). Como resultado, foram gerados três artefatos principais: uma estrutura semântica, um conjunto de relações e uma lista de conceitos.

Com base nessa representação conceitual, Oliveira (2018) propôs um modelo de conversão de texto linear em hipertexto semântica, utilizando como corpus o livro “Evolução agrária e pressão demográfica”, versão em português traduzida do original em inglês “The Conditions of Agricultural Growth: The Economics of Agrarian Change Under Population Pressure”, de Ester Boserup (1985), como base para a modelagem conceitual. Para a seleção dos termos, Oliveira (2018) utilizou três fontes: (1) conceitos da amostra de Maculan (2015); (2) extração manual com base em diagrama do especialista; e (3) extração automática por minerador de texto. Como resultado, obteve-se uma amostra de 26 termos validados por um especialista, com definições construídas a partir da Teoria da Terminologia e nos critérios de Maculan (2015).

2.1 Procedimentos metodológicos

O corpus da pesquisa é o livro *Evolução agrária e pressão demográfica*, de Ester Boserup, obra clássica e amplamente referenciada nas Ciências Agrárias por tratar da relação entre crescimento populacional e práticas agrícolas. Sua escolha se justifica pela relevância teórica e metodológica para estudos sobre intensificação agropecuária, e por seu uso consolidado em pesquisas prévias na área de Organização do Conhecimento. Mesmo tendo sido originalmente publicado em 1985, o conteúdo permanece atual por abordar fundamentos conceituais e dinâmicas estruturais que ainda influenciam políticas agrárias e estratégias de uso do solo em contextos contemporâneos.

A estrutura semântica enriquecida de Maculan (2015), foi utilizada como padrão ouro, previamente validado por especialistas, servindo como base comparativa para avaliar a eficácia do *Semantizar* na identificação e categorização de RS. A lista de conceitos atualizada do domínio da Intensificação Agropecuária, contendo termos como *pousio*, *fertilidade do solo* e *pressão populacional*, foi inserida na ferramenta para guiar a identificação de pares conceituais semanticamente relevantes e por fim, o conjunto de relações utilizada foi utilizada para determinar e classificar as conexões entre os conceitos.

O texto do corpus foi convertido do formato *Portable Document Format* (PDF) para .txt, preservando a estrutura textual para compatibilidade com o *Semantizar*, que opera em ambiente Windows com suporte do *Uniform Server* e banco de dados MySQL. Desenvolvido por Maia (2018), o *Semantizar*¹ realiza extração semiautomática de relações semânticas (RS) em português, com arquitetura baseada em regras linguísticas e padrões de notação semântica, sem uso de algoritmos de aprendizado de máquina ou redes neurais. O sistema reconhece pares conceituais em sentenças e atribui relações com base em padrões predefinidos.

A escolha da ferramenta, fundamentada em revisão de literatura, justifica-se por seu alinhamento aos objetivos do estudo e por ter sido concebida especificamente para atuar sobre textos acadêmicos em português. Diferentemente de ferramentas como *ReVerb* e

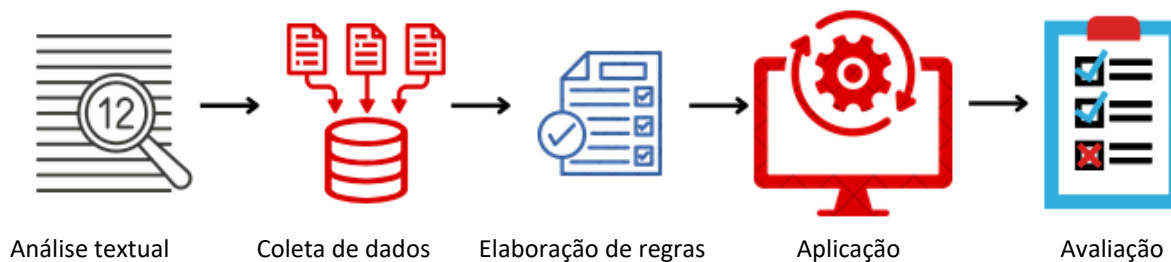
¹ Durante a pesquisa, elaborou-se um plano para aprimorar o código-fonte da ferramenta *Semantizar*, com o objetivo de expandir suas funcionalidades e melhorar sua eficiência. Além disso, foi prevista a criação de um site dedicado, que permitirá a disponibilização de informações mais detalhadas sobre a ferramenta, incluindo sua aplicação prática, manuais de uso e resultados de estudos anteriores. Essa iniciativa visa não apenas aumentar a acessibilidade e a transparência do *Semantizar*, mas também promover sua utilização por outros pesquisadores e profissionais interessados na extração de relações semânticas, contribuindo assim para o avanço do campo e para o fortalecimento da colaboração acadêmica.

OpenNLP, Cruz e Weitzel (2018), originalmente desenvolvidas para o inglês e que demandam extensas adaptações para o português brasileiro.

Após a instalação, o banco de dados foi configurado com credenciais específicas (nome, usuário, senha e host) para integração com o *Semantizar*. Essa configuração garantiu o ambiente para a realização das análises, assegurando a integridade do processo e a confiabilidade dos resultados.

Devido à sua flexibilidade metodológica e à ausência de estruturas rígidas, que caracterizam outros métodos de pesquisa, o estudo de caso, conforme Gil (2002) requer atenção às fases da pesquisa e à construção de um protocolo estruturado, com uma visão global do projeto. O protocolo metodológico compreendeu as cinco etapas, apresentadas na Figura 2, estruturadas de forma a assegurar a replicabilidade do procedimento.

Figura 2 – Visão geral do protocolo do estudo de caso



Fonte: elaborada pelos autores (2024).

A análise textual, etapa inicial do protocolo, teve como objetivo identificar os temas centrais de cada capítulo e sistematizar as informações relacionadas aos conceitos-chave e às implicações apresentadas no texto.

Na sequência ao protocolo, a coleta de dados foi estruturada para registrar os conceitos-chave e as relações explícitas ou implícitas entre eles, garantindo que as informações essenciais fossem sistematizadas e alinhadas ao objetivo da pesquisa, que compreende a construção de uma estrutura semântica e a validação das RS identificadas pela ferramenta. A Figura 13 apresenta as duas etapas explicadas até o momento:

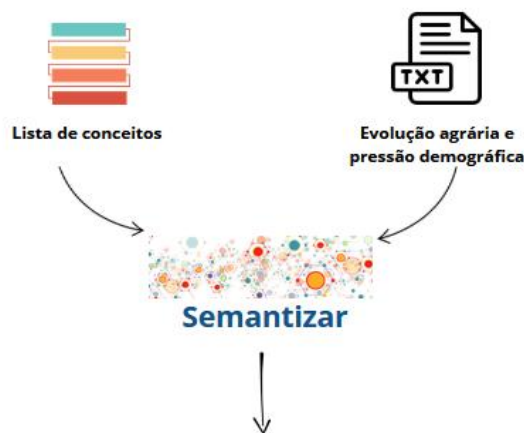
A terceira etapa constituiu na construção de regras, contemplando a definição de forma de apresentação, notação e tipo de RS, em conformidade com a Norma ANSI/NISO Z39.19-2005. Ressalta-se que não há um formato único considerado ideal para todas as situações; podendo estes ser adaptados conforme o contexto e as necessidades específicas.

Nesta pesquisa, optou-se por apresentar as relações em formato de tabela, de modo a atender às demandas do estudo.

Para a notação de RS, adotou-se o conjunto de relações apresentado nos insumos, para determinar e classificar as conexões entre os conceitos, seguindo as diretrizes da norma ANSI/NISO Z39.19-2005 (R2010), que orienta a estruturação de vocabulários controlados. As relações foram organizadas em três categorias principais: equivalência (ex.: *temSinônimo*), hierárquica (ex.: *éParteDe*, *possuiTipoDe*) e associativa (ex.: *afeta*, *influencia*, *dependeDe*).

Com base nos insumos definidos, procedeu-se à configuração do ambiente computacional necessário para a execução do protocolo. A entrada da ferramenta foi composta pela lista de conceitos previamente selecionados e pelo texto do corpus convertido para o formato .txt, garantindo compatibilidade com os requisitos do *Semantizar*. Em seguida, iniciou-se a aplicação da ferramenta, que processou os dados e identificou indícios de relações semânticas entre os conceitos analisados, conforme a Figura 4.

Figura 4 – Aplicação da ferramenta e apresentação dos indícios de relação semântica



Se o sistema feudal de posse e uso da terra existe em regiões onde a crescente população rural é incapaz de encontrar emprego suficiente por causa da sobrevivência de formas extensivas de uso do **solo**, uma explicação que com frequência se utiliza é a de que o comportamento dos senhores feudais é antieconômico, revelando mais interesse no prestígio que a propriedade da terra lhes confere do que nos lucros que uma **agricultura intensiva** poderia proporcionar.

Fonte: elaborada pelos autores (2024).

Na etapa avaliação, a estrutura semântica gerada após a aplicação da ferramenta foi analisada quantitativamente, por meio das métricas de precisão, revocação e F1-Score. Na avaliação qualitativa, considerando-se a estrutura semântica como padrão de referência e seguindo as recomendações da Norma ANSI/NISO Z39.19-2005, examinaram-se os tipos de

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

relações explicitadas (hierárquicas, associativas e de equivalência) e as RS sugeridas além da estrutura original.

O Quadro 1 mostra um exemplo de RS explicitada após aplicação do protocolo. As colunas “Conceito 1” e “Conceito 2” (“Agricultura intensiva” e “Solo”, respectivamente) denotam os conceitos em análise para a ERS. A coluna “Nome da Relação” descreve o vínculo semântico entre os conceitos, cujo exemplo é "afeta". A coluna “RS explicitada” estabelece relação entre os conceitos: “Agricultura intensiva afeta o solo”. Já o “Tipo de RS” denota a tipologia das RS, que, no exemplo, é Associativa.

Quadro 1 – Exemplo da forma de apresentação

Conceito 1	Conceito 2	Notação da RS	RS explicitada	Tipo de RS
Agricultura intensiva	Solo	afeta	Agricultura intensiva afeta o solo.	Associativa

Fonte: elaborado pelos autores (2024).

3 ESTUDO DE CASO

A aplicação do protocolo de extração corresponde à execução do estudo de caso. A análise textual e coleta de dados, detalhados na Figura 3, tiveram como finalidade interpretar as informações contidas nos 12 (doze) capítulos do livro de Ester Boserup e estruturando-as em: Capítulo, Tema Central, Termos mais Relevantes e Implicações.

Dando seguimento ao estudo de caso, procedeu-se à aplicação da ferramenta *Semantizar*, conforme descrito na introdução. Durante o processo, foram sugeridas 144 (cento e quarenta e quatro) Relações Semânticas (RS) a partir do corpus analisado. Essas relações foram geradas por meio da identificação automática de pares conceituais nas sentenças, seguida da atribuição de vínculos com base em padrões linguísticos previamente definidos. O Quadro 2, composto pelo identificador, frase extraída do livro e termos encontrados na frase, tem o objetivo de ilustrar um exemplo dos indícios gerados, permitindo compreender como o sistema interpreta o texto e propõe as possíveis conexões semânticas.

Quadro 2 – Exemplo de indício de relação semântica

Identificador	Frase	Termos
48	Mesmo onde o sistema de cultivo é do tipo pousio arbustivo combinado com o cultivo permanente de arroz irrigado, a quantidade de trabalho pode permanecer pequena comparada àquela utilizada na moderna	Pousio, Sistema de

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

	agricultura dos países ocidentais.	cultivo
--	------------------------------------	---------

Fonte: elaborado pelos autores (2024).

Do total de indícios apresentados, 28 (vinte e oito) não estavam presentes na estrutura conceitual de referência construída manualmente, evidenciando a capacidade da ferramenta de propor conexões adicionais entre os conceitos analisados. Embora a ferramenta tenha gerado falsos positivos, relações genéricas excessivamente amplas ou associações contextualmente deslocadas, optou-se, por questões de concisão e devido às restrições de espaço, por não apresentar no presente artigo a análise pormenorizada dessas ocorrências. Essa avaliação detalhada, contendo exemplos e categorização dos casos, encontra-se registrada integralmente na dissertação.²

As relações adicionais, consideradas semanticamente válidas, reforçam o potencial do *Semantizar* para ampliar estruturas conceituais e revelar associações relevantes que não haviam sido identificadas na construção manual. O Quadro 3 apresenta as RS sugeridas além da estrutura conceitual de referência.

² BRAGA, V. C. **A extração de relações semânticas como suporte para a construção de estruturas conceituais:** um estudo de caso na Intensificação Agropecuária. 180 f. 2025. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2025. (Disponível em: <https://repositorio.ufmg.br/handle/1843/79949>. Acesso em: 12 ago. 2025.



25°enancib

Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação
Informação, Decolonialidade e Direitos Difusos para o Desenvolvimento Sustentável

3 a 7 de novembro de 2025 - Rio de Janeiro



Quadro 3 – Relações semânticas sugeridas além da estrutura conceitual

N.º	Conceito 1	Conceito 2	Notação da RS	RS explicitada	Tipo de RS
1	Agricultura intensiva	Solo	afeta	Agricultura intensiva afeta o solo.	Associativa
2	Agricultura intensiva	Uso da terra	influencia	A agricultura intensiva influencia o uso da terra.	Associativa
3	Crescimento populacional	Período de pousio	influencia	Crescimento populacional influencia o período de pousio.	Associativa
4	Crescimento populacional	Sistema de cultivo	influencia	A taxa de crescimento populacional influencia o sistema de cultivo.	Associativa
5	Crescimento populacional	Uso da terra	influencia	Crescimento populacional influencia o uso da terra.	Associativa
6	Crescimento populacional	Solo	afeta	Crescimento populacional afeta o solo.	Associativa
7	Crescimento populacional	Terra agrícola	influencia	Crescimento populacional influencia a terra agrícola.	Associativa
8	Crescimento populacional	Uso da terra	afeta	Crescimento populacional afeta o uso da terra.	Associativa
9	Fertilidade do solo	Pousio	dependeDe	Fertilidade do solo depende do pousio.	Associativa
10	Fertilidade do solo	Pousio	afeta	Pousio afeta a fertilidade do solo.	Associativa

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

(Continua...)

N.º	Conceito 1	Conceito 2	Notação da RS	RS explicitada	Tipo de RS
11	Período de cultivo	Pousio curto	dependeDe	Período de cultivo depende de pousio curto.	Associativa
12	Período de cultivo	Pousio	influencia	O pousio influencia o período de cultivo.	Associativa
13	Período de pousio	Pousio	éParteDe	O período de pousio é parte do sistema de pousio.	Associativa
14	Período de pousio	Uso da terra	éParteDe	Período de pousio é parte do uso da terra.	Associativa
15	Período de pousio	Uso da terra	temPrática	Uso da terra tem prática de período de pousio.	Associativa
16	Período de pousio	Uso da terra	afeta	Período de pousio afeta o uso da terra.	Associativa
17	Pousio	Pousio curto	influencia	O pousio influencia o pousio curto.	Associativa
18	Pousio	Pousio curto	éTipoRelacionadoA	O pousio curto é uma forma evoluída do pousio.	Associativa
19	Pousio	Sistema de cultivo	éParteDe	O pousio é parte do sistema de cultivo.	Associativa
20	Pousio	Solo	afeta	Pousio afeta o solo.	Associativa

(Continua...)

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

(Conclusão)

N.º	Conceito 1	Conceito 2	Notação da RS	RS explicitada	Tipo de RS
21	Pousio	Solo	influencia	Pousio influencia o solo.	Associativa
22	Pousio	Uso da terra	influencia	O pousio influencia o uso da terra.	Associativa
23	Pousio curto	Sistema de cultivo	afeta	O pousio curto afeta o sistema de cultivo.	Associativa
24	Produção de alimentos	Solo	afeta	A produção de alimentos afeta o solo.	Associativa
25	Produção de alimentos	Solo	influencia	A produção de alimentos influencia o solo.	Associativa
26	Produção de alimentos	Solo	afeta	A produção de alimentos afeta o solo.	Associativa
27	Sistema de cultivo	Uso da terra	influencia	O sistema de cultivo influencia o uso da terra.	Associativa
28	Sistema de cultivo	Uso da terra	dependeDe	Sistema de cultivo depende do uso da terra.	Associativa

Fonte: elaborado pelos autores (2024).

4 AVALIAÇÃO

Ferramentas como o *Semantizar* são projetadas para otimizar o desempenho em seus domínios-alvo, por meio da aplicação de regras cuidadosamente definidas, visando garantir consistência e precisão. Em avaliações quantitativas, métricas como precisão, revocação e F1-score são amplamente utilizadas para mensurar o desempenho de sistemas de extração de relações, permitindo uma análise objetiva sobre sua eficácia.

A precisão foi de 0,21, indicando que pouco mais de 20% das relações propostas pela ferramenta estavam corretas segundo o padrão de referência validado. A revocação, também conhecida como cobertura, foi de 0,37, refletindo a capacidade parcial da ferramenta em recuperar as relações esperadas. O F1-score, média harmônica entre precisão e revocação, resultou em 0,1339. Tais resultados podem ser atribuídos a limitações da arquitetura da ferramenta, baseada exclusivamente em regras fixas, sem incorporar algoritmos de Aprendizado de Máquina (AM). Essa abordagem. Embora mais simples, tende a comprometer a capacidade do sistema em capturar relações semânticas mais complexas. Estudos da literatura indicam que ferramentas que combinam AM e uso de ontologias especializadas apresentam desempenho superior, sugerindo caminhos para os aprimoramentos futuros do *Semantizar*.

Apesar dessas limitações, a análise qualitativa revelou que a ferramenta foi capaz de propor 28 (vinte e oito) novas relações semânticas não previstas na estrutura conceitual de referência, conforme listadas no Quadro 3, evidenciando seu potencial para expandir estruturas conceituais previamente estabelecidas.

As RS sugeridas mostraram conformidade com as recomendações da Norma ANSI/NISO Z39.19-2005 (R2010), atendendo aos critérios relativos aos tipos de RS, à estrutura do vocabulário, ao escopo do temático e aos princípios de seleção de termos. Esse alinhamento evidencia a adequação da ferramenta como suporte na de EC no domínio da IA.

Quanto à tipologia de RS, a ferramenta categorizou corretamente as conexões entre conceitos em relações hierárquicas, associativas e de equivalência, recomendando que elas sejam claras e semanticamente consistentes. As relações identificadas pelo *Semantizar*, como "Pousio afeta Fertilidade do Solo", "Crescimento populacional influencia Uso da Terra" e "Período de pousio é parte de Pousio", exemplificam a capacidade do sistema de capturar adequadamente relações.

A estrutura do vocabulário gerada apresenta organização lógica e favorece a navegabilidade, conforme recomendado pela norma. Quanto ao propósito e escopo do

vocabulário está bem definido, concentrando-se no domínio específico da IA garantindo que as conexões sejam relevantes e úteis para esse contexto. A delimitação temática promovida pela ferramenta assegura que o vocabulário gerado seja consistente com as necessidades do domínio e com os objetivos da pesquisa.

No que se refere à seleção dos termos, o *Semantizar* demonstrou priorizar conceitos relevantes, precisos e consistentes no que tange a escolha dos conceitos e das relações. A ferramenta identifica termos e estabelece conexões claras entre eles, minimizando ambiguidades e garantindo que os resultados sejam semanticamente significativos. Além disso, a inclusão de termos específicos do domínio, como "Pousio", "Solo" e "Uso da Terra", demonstra a adequação dos conceitos selecionados à realidade temática da IA.

5 CONSIDERAÇÕES FINAIS

A avaliação da ferramenta *Semantizar* evidenciou seu potencial como ferramenta de apoio à construção de Estruturas Conceituais (EC) no domínio da Intensificação Agropecuária (IA), especialmente em contextos de análise textual em linguagem natural. A ferramenta demonstrou competência em propor Relações Semânticas (RS), além daquelas previstas na estrutura conceitual de referência, contribuindo para o enriquecimento informacional.

Os resultados apontaram alinhamento significativo com os critérios estabelecidos pela Norma ANSI/NISO Z39.19-2005 (R2010), sobretudo no que se refere à categorização adequada das relações em hierárquicas, associativas e de equivalência, bem como à organização lógica da estrutura vocabular gerada. A ferramenta também mostrou coerência ao manter o foco no domínio da IA respeitando o propósito e escopo do vocabulário analisado.

Destaca-se, ainda, o cumprimento dos princípios de seleção de termos, com destaque para a escolha de conceitos pertinentes e na formação de conexões semanticamente relevantes e consistentes. No entanto, observou-se a necessidade de aprimoramentos técnicos voltados ao aumento da precisão e à identificação de relações mais sutis ou contextualmente complexas.

Conclui-se, portanto, que o *Semantizar* se mostra uma alternativa viável e promissora para apoiar processos de organização do conhecimento, oferecendo vantagens em termos de escalabilidade, sistematização e atualização semântica, elementos indispensáveis em contextos marcados por alta densidade informacional e crescente complexidade temática.

REFERÊNCIAS

- AMERICAN NATIONAL INFORMATION STANDARDS ORGANIZATION. **NISO Z39.19**: Guidelines for the construction, format and management of monolingual controlled vocabularies. Bethesda: NISO, 2005.
- BOAVENTURA, E. M. **Metodologia da pesquisa**: monografia, dissertação, tese. São Paulo: Atlas, 2004.
- BOSERUP, E. **Evolução agrária e pressão demográfica**. São Paulo: Hucitec; Polis, 1987.
- CRUZ, C. G. A.; WEITZEL, L. Evaluation of relation extraction systems for Portuguese language pt-br. *In*: IBERIAN CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGIES (CISTI), 13., 2018, Lisboa. **Proceedings** [...]. Lisboa: IEEE, 2018. p. 1-6.
- GIL, A. C. **Como elaborar um projeto de pesquisa**. 6. ed. São Paulo: Atlas, 2010.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. 4. ed. São Paulo: Atlas, 2002.
- LAUSER, B. *et al.* From AGROVOC to the Agricultural Ontology Service: Concept Server na OWL model for creating ontologies in the agricultural domain. *In*: INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS, 2006, Colima. **Proceedings** [...] Colima: DCMI, 2006.
- MACULAN, B. C. M. S. **Estudo e aplicação de metodologia para reengenharia de tesouro**: remodelagem do THESAGRO. 2015. 339 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2015.
- MAIA, L. S. **Extração e explicitação de relações semânticas para representação de documentos acadêmicos**: um estudo de caso a partir de uma estrutura classificatória. 2018. 248 f. Tese (Doutorado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em: <https://repositorio.ufmg.br/handle/1843/VAFA-BB5JHW>. Acesso em: 15 jun. 2023.
- OLIVEIRA, E. D. **Proposta de um modelo de hipertexto com abordagem semântica para a representação do conhecimento no domínio temático da intensificação agropecuária**. 2018. 233 f. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018.
- PRODANOV, C. C.; FREITAS, E. C. **Metodologia do trabalho científico**: métodos e técnicas da pesquisa e do trabalho acadêmico. 2. ed. Novo Hamburgo: Feevale, 2013.
- SOERGEL, D. *et al.* Reengineering thesauri for new applications: the AGROVOC example. **Journal of Digital Information**, United Kingdom, v. 4, n. 4, 2004.