

## DESENVOLVIMENTO DE UM GLOSSÁRIO A PARTIR DO DISCURSO DA INOVAÇÃO TECNOLÓGICA

Ketlin Marcelly P. da Silva (IC), Prof<sup>a</sup>. Dr<sup>a</sup>. Márcia de S. Luz-Freitas (PQ)

*Universidade Federal de Itajubá*

**Palavras-chave:** Software. Glossário. Inovação.

### Introdução

Considerando que a Terminologia se firmou como campo de estudo e que tem atuado para auxiliar os especialistas no tratamento dos termos das suas rotinas de trabalho, criou-se um projeto de pesquisa para estudar o discurso especializado da inovação tecnológica. Este relatório busca detalhar as etapas da elaboração de dois glossários terminológicos resultantes de uma pesquisa sobre o discurso da inovação tecnológica, a partir de ferramentas de auxílio para pesquisa e agrupamento de dados.

O uso crescente dos termos da área da inovação tecnológica na mídia digital, na área acadêmica, nas empresas e em outros espaços importantes da sociedade reforça a relevância desta pesquisa e, portanto, estabelecer os termos que compõem e caracterizam esse discurso é um passo necessário para reforçá-lo. É necessário não apenas estabelecer esses termos, mas também dar a conhecer seus significados à comunidade acadêmica que os utiliza, principalmente na construção da base conceitual e no desenvolvimento de pesquisas e a outros profissionais que precisam disso em sua prática diária (LUZ-FREITAS, 2019).

A razão da terminologia ter-se tornado objeto de interesse dos linguistas, deve-se ao fato de deixar de ser vista apenas como um instrumento de normalização de termos, para tornar-se mais um instrumento de comunicação (CABRÉ, 1993). Mas, Sakar (2007) alerta para a possibilidade de os termos utilizados corriqueiramente tornarem-se clichês, quando mal-empregados, deixando de gerar o impacto e a credibilidade esperados.

A elaboração de um glossário com definições é um grande aliado na tarefa de registrar os usos dos termos e definições de uma área especializada, no caso deste projeto o discurso da inovação tecnológica, e assim fornecer instrumentos de consulta para que o mal emprego dos termos seja evitado.

Desta forma, a intenção do projeto como um todo é elaborar um glossário em duas linguagens, sendo uma mais técnica e outra voltada ao público leigo por meio do

auxílio de softwares de pesquisa e organização, como o BootCat e AntConc. O glossário, de forma geral, tem como objetivo apresentar definições para alguns termos da área da inovação terminológica que o leitor ou a leitora podem não conhecer. Entretanto, podem também contemplar termos que apenas determinadas regiões geográficas utilizam, já que se trata de uma área em desenvolvimento e que muitos termos ainda são instáveis.

Para realizar a etapa de elaboração dos glossários do projeto de pesquisa, foi selecionada uma bolsista de iniciação científica cujas atividades principais foram definir qual a plataforma mais adequada para criação dos glossários propostos, realizar a migração dos dados já previamente coletados e preparar as páginas para publicação dos glossários em uma plataforma de acesso livre para consulta.

### Metodologia

Para executar as atividades e atingir os objetivos propostos do projeto como um todo, este estudo foi dividido em três fases. Na primeira fase, foram coletados e analisados dados sobre produtos textuais publicados nas últimas quatro décadas completas, ou seja, nos seguintes intervalos de tempo: 1981-1990; 1991-2000; 2001-2010; 2011-2020. Para agrupar textos disponíveis na mídia digital relacionados à inovação tecnológica, foram coletadas publicações de instituições oficiais e órgãos de mídia regional com o auxílio do programa BootCat. O software automatiza o processo de encontrar textos de referência na web e reuni-los em um único corpus.

Para isso, deve-se fornecer uma lista de palavra-chave a serem usados para a coleta automática de textos na web. Esses mesmos termos são combinados entre si formando tuples (cruzamentos) e enviados como consulta a um mecanismo de pesquisa, que retorna uma lista de URLs potencialmente relevantes. Foram utilizadas como palavras-chave, a princípio, as palavras/expressões Itajubá, inovação e inovação tecnológica. A organização e a sistematização do material coletado foram realizadas de acordo com os princípios metodológicos da lexicologia e terminologia (CABRÉ, 1999; BARROS,

2004; KRIGER e FINATTO, 2004).

Em um segundo momento, foram incluídas novas palavras-chave que foram agrupadas em três categorias para a coleta de dados, conforme a figura 1. Para cada categoria foi indicado um conjunto de palavras-chaves para realizar o cruzamento de dados que o Corpus de referência necessitaria para realizar as consultas de URLs e gerar o agrupamento para ser utilizado o AntConc.

Figura 1: Categorias e sementes para cruzamento

Categoria	Sementes
Núcleo conceptual	Ciência
	Desenvolvimento
	Inovação
	Inovação Tecnológica
	Pesquisa
Aplicação	Tecnologia
	Método
	Processo
	Produto
	Serviço
Implementação	Sistema
	Base Tecnológica
	Parque Tecnológico
	Setor Tecnológico

Fonte: Equipe executora do projeto

Os textos que compõem o *corpus* documental, foram processados pelo programa AntConc para organização de dados e quantificação de resultados. O software AntConc permite realizar as principais técnicas de linguística de corpus como: contador de frequência de palavras, colocações, concordância, n-grams e comparação de corpora a qualquer tipo de texto.

Após a finalização das etapas anteriores, iniciou-se a etapa para definição de qual plataforma seria a mais adequada para a criação do glossário. Contou-se com auxílio de planilhas para a migração dos dados obtidos nos softwares. Essa fase está em desenvolvimento e será concluída na segunda etapa da pesquisa.

## Resultados e discussão

Foi realizada a coleta de dados em diversos sites, como Congresso Nacional (Leis sobre o tema), Ministério da Ciência, Tecnologia e Inovações (MCTI), Empresa Brasileira de Pesquisa e Inovação Industrial (EMBRAPPI), Instituto Nacional da Propriedade Industrial (INPI) e sites de todas as Universidades Públicas do Brasil (Agências de Inovação, Núcleos de Inovação, etc.).

Para iniciar o teste no programa BootCat foram realizados noventa cruzamentos, ou seja, escolheu-se uma palavra de cada categoria e adicionou-se à plataforma para gerar o agrupamento. Como resultado, foram coletadas 137 URLs ao final do processamento.

Após esse processo, foi realizada uma averiguação minuciosa nas páginas selecionadas e foram descartados 55 URLs. Dentre esses 55, 24 arquivos estavam em “.txt,” o arquivo 10 e 11 eram iguais entre si e os arquivos 34, 35, 39, 42, 43, 45 não traziam o texto todo. Além disso, a delimitação de sites com a terminação “.gov.br” nesta etapa de coleta não funcionou como o esperado, já que houve uma eliminação significativa de URLs com essa terminação. Para solucionar o problema, foi criada uma lista de instituições de ensino superior e um Corpus Universidades e a partir deles foram coletados materiais que proporcionaram a obtenção de *corpora* documental adequado, composto por um *corpus* oficial, com 81 arquivos, e um *corpus* midiático e acadêmico, composto por 237 arquivos. Para as universidades públicas, foi realizado um levantamento de dados, totalizando 81 sites para consulta que foram organizados em uma planilha. Foi adicionado o nome e sigla de cada universidade, além do link para o redirecionamento aos sites. Para a coleta de dados em fontes não oficiais utilizaram-se textos acadêmicos, jornais regionais, notícias e entrevistas.

Para a escolha da plataforma adequada, foi feito um levantamento de quais softwares se encaixariam no perfil da pesquisa. Algumas plataformas de gestão de glossários foram analisadas, entre elas a Wordsmith Tools e a plataforma Terminus. A plataforma WordSmith Tools, mesmo contando inúmeras ferramentas como Wordlist, Concord e Key Words, não se adequou ao perfil da pesquisa já que é uma plataforma paga e não pode ser disponibilizada de forma gratuita para a população. Já a plataforma (Terminus) é uma aplicação web para gerenciamento de corpus e terminologia que permite criar e gerenciar projetos terminológicos e selecionar categorias de dados, mas que também tem como ponto negativo ser um recurso pago.

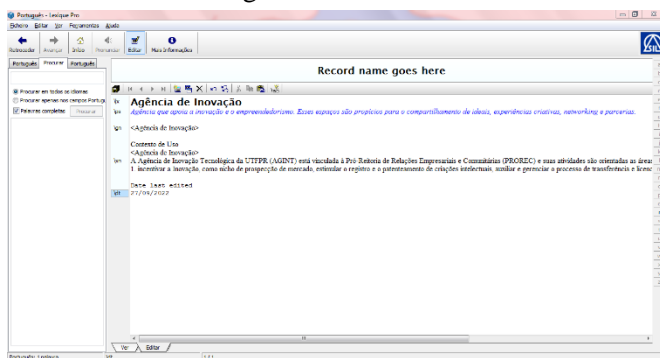
Sendo assim, a plataforma escolhida foi a Sil Language Technology. Ela desenvolve e oferece suporte a uma ampla gama de softwares para auxiliar nas atividades de desenvolvimento linguístico como pesquisa linguística e cultural, alfabetização e educação na língua materna, desenvolvimento de dicionários, tradução comunitária e assistida por computador e produção de mídia vernácula. No site da instituição, existem inúmeras opções de softwares e o escolhido foi o Lexique Pro. A Lexique Pro é uma ferramenta gratuita para elaboração de glossários e é de fácil entendimento.

Para a preparação das páginas iniciais do glossário, os dados obtidos nas etapas anteriores foram organizados em duas planilhas. A primeira planilha, é direcionada a dar embasamento para a criação do glossário com definições em linguagem técnica e a segunda, com linguagem acessível ao público leigo. As planilhas são formadas pelos seguintes itens: termos, definições, fontes

de consulta, contextos de uso, fontes dos contextos de uso, notas e remissivas, caso necessário.

Na plataforma, foram inseridos os termos e suas definições para teste, conforme a figura 2. A plataforma permite que seja indicado um exemplo de contexto de uso que aquele termo foi encontrado e é necessário também indicar qual a data da última edição. A plataforma tem o inglês como língua oficial, mas é possível criar um glossário em português, que é a língua em que se está realizando esta pesquisa. Também pode ser realizada a consulta em algumas línguas na aba de procura. A plataforma também disponibiliza a opção de pronúncia para cada termo, sendo uma ótima ferramenta para auxiliar o público leigo com os termos mais técnicos da pesquisa.

Figura 2: Glossário



Fonte: Autor (2022)

## Conclusões

Diante disso, foi possível concluir que o uso da terminologia e de seus procedimentos metodológicos para organizar os termos de uma área e criar um glossário para auxiliar a consulta é um ótimo caminho, se combinado com as ferramentas adequadas. Pode-se criar um artefato potente que facilita o dia a dia, tanto de um profissional especializado que depende de um material confiável para consulta, do público mais geral, que tem interesse e curiosidade sobre o tema e deseja realizar uma consulta em um material produzido por uma instituição de pesquisa.

Além disso, vale salientar que a pesquisa ainda não foi concluída já que ainda restam 10 meses para a conclusão do projeto e a disponibilização dos glossários para a consulta geral da população.

## Agradecimento

Agradece-se à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e à Diretoria de Pesquisa da Universidade Federal de Itajubá (UNIFEI), pela bolsa concedida para a aluna para realização desta

pesquisa. Também se agradece à orientadora, prof<sup>a</sup>. Dr<sup>a</sup>. Márcia de Souza Luz Freitas pelo apoio e orientação.

## Referências

ANTHONY, L. **AntConc** (Version 3.5.8) [Windows]. Tokyo, Japan: Waseda University. Disponível em: <<http://www.laurenceanthony.net/software/antconc/>>. Acesso em: 12 fev. 2021.

BARONI, M. et al. **BootCat**. (Version 1.41) [Computer Software]. Universities of Bologna (Forlì Campus), 2020. Disponível em: <<https://bootcat.dipintra.it/>>.

BARROS, L. A. **Curso básico de terminologia**. São Paulo: EDUSP, 2004.

CABRÉ, Maria Teresa. **La Terminología: teoría, metodología, aplicaciones**. Barcelona: Antártida/Empúries, 1993.

CABRÉ, M. T. **Hacia una teoría comunicativa de la terminología: aspectos metodológicos**. La terminología: representación y comunicación. Barcelona: Institut Universitari de Lingüística Aplicada, 1999, p. 129-150.

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à Terminologia: teoria e prática**. São Paulo: Contexto, 2004.

LUZ-FREITAS, M. S. **A neologia no entrecruzar das ciências médicas e biológicas e da engenharia: estudo terminológico do léxico pertinente à engenharia biomédica**. Tese de Doutorado. Universidade de São Paulo, 2019.

SAKAR, S. Inovação: metamorfoses, empreendedorismo e resultados. In: TERRA, J.C.C **Inovação: quebrando paradigmas para vencer**. São Paulo: Saraiva, 2007. Cap. 2, p. 27-31.