



**XXXIII SIC** SALÃO INICIAÇÃO CIENTÍFICA

<b>Evento</b>	Salão UFRGS 2021: SIC - XXXIII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
<b>Ano</b>	2021
<b>Local</b>	Virtual
<b>Título</b>	METODOLOGIAS PARA A COMPILAÇÃO DE UM CORPUS E PARA A EXTRAÇÃO DE LEXICAL FRAMES
<b>Autor</b>	KAROL TRASLATTI SOARES
<b>Orientador</b>	ANA ELIZA PEREIRA BOCORNY

## METODOLOGIAS PARA A COMPILAÇÃO DE UM CORPUS E PARA A EXTRAÇÃO DE LEXICAL FRAMES

Karol Traslatti Soares

Resumos (*Abstracts*) são a porta de entrada para artigos de pesquisa. É a partir deles que pesquisadores e estudantes decidem se irão ou não ler o texto do artigo completo. Dentro do contexto acadêmico, artigos de pesquisas científicas exercem um papel importante durante a formação universitária. Levando em consideração estes aspectos, é importante que os pesquisadores e estudantes sejam capazes de escrever resumos informativos e estruturados de acordo com as necessidades e a linguagem considerada autêntica pela comunidade acadêmica da área de pesquisa da qual fazem parte. Buscamos, com a presente pesquisa, contribuir para a melhoria da qualidade da produção científica da comunidade acadêmica da UFRGS. Mais especificamente, tratamos neste trabalho da identificação, extração e classificação de *Lexical Bundles (LBs)*, entendidos neste estudo como sequências contínuas e recorrentes de palavras características de uma coleção de textos de um domínio e ou gênero específico, e a posterior construção de *Lexical Frames (LFs)*, entendidos como sequências descontínuas e recorrentes, onde as palavras formam uma 'estrutura' em torno de um *slot* variável (ex. I don't \* to, it is \* to)" (Gray and Biber 2013:109). A razão pela qual tal objeto foi escolhido para este estudo está no fato dos *LFs* representarem importantes blocos textuais do discurso acadêmico escrito. Neste estudo, descrevemos a metodologia utilizada para a compilação do corpus utilizado neste estudo e o processo de construção manual de *LFs* a partir dos *LBs*. Os resultados indicam a validade da utilização desta metodologia para a extração dos referidos objetos de corpora de outras áreas de estudo.