



## **BIG DATA MEDIA ANALYSIS AND THE REPRESENTATION OF URBAN VIOLENCE IN BRAZIL - RESULTADOS PARCIAIS**

Tainá Copetti Bernardi (PIBIC-CNPq), Heloisa Pedroso De Moraes Feltes (Orientador(a))

Esta comunicação tem o objetivo geral apresentar os principais resultados parciais do projeto de pesquisa DATA-UV. Tomando um *corpus* de quatro jornais de dois estados brasileiros: no RS, Zero Hora e Pioneiro; em SP, Estado de São Paulo e Folha de São Paulo, em edições diárias do ano de 2014, quantificaram-se e analisaram-se dados referentes à veiculação de casos de violência urbana na mídia impressa por meio de métodos e técnicas computacionais da Linguística de Corpus. Com a utilização do *software* AntConc processou-se o *corpus* de 1.778.282 milhões de palavras e gerou-se uma lista de *keywords* (palavras-chave) para o *corpus* geral. Verificou-se, com essa lista, que os crimes com maior frequência - crimes entre as 100 primeiras *keywords* -, durante o ano de 2014, no *corpus* geral, foram: roubo (1962 ocorrências), homicídio (1827), assalto (1143), assassinato (931), furto (843) e estupro (432). Essa análise também foi realizada para cada jornal separadamente, a fim de se compararem os resultados obtidos. Na etapa seguinte, analisaram-se os *collocates* (colocados) desses crimes no *corpus* geral e em cada jornal separadamente. Observou-se, tomando como exemplo os três crimes mais frequentes no *corpus* geral, que a maior coocorrência para 'roubo' foi 'veículo' e 'veículos'. Para 'assassinato' foram termos referentes ao caso de Bernardo Boldrini, ocorrido em 2014, de modo a evidenciar o destaque que o crime obteve na mídia. Com 'homicídio', a maior coocorrência foi 'ocultação' e 'cadáver', o que aponta a existência frequente de crimes qualificados. A coocorrência de 'tráfico' e 'drogas' com quase todos os crimes analisados torna possível pontuar, no *corpus* geral, a influência do narcotráfico nos casos relatados de violência urbana no país. Além disso, os dados relativos ao ranking dos crimes que reportados na mídia, convergem com os resultados encontrados no Atlas da Violência 2016, documento publicado pelo Instituto de Pesquisa Econômica Aplicada (IPEA).

Palavras-chave: Linguística de Corpus, Violência urbana, Mídia impressa, Resultados

Apoio: CNPq