

Modelagem de tópicos com o algoritmo LDA aplicado a obras representativas das ondas do feminismo

Priscila Cristina de S. Mizukami Rocha

José Artur Quilici Gonzalez

Organização e
Apoio:



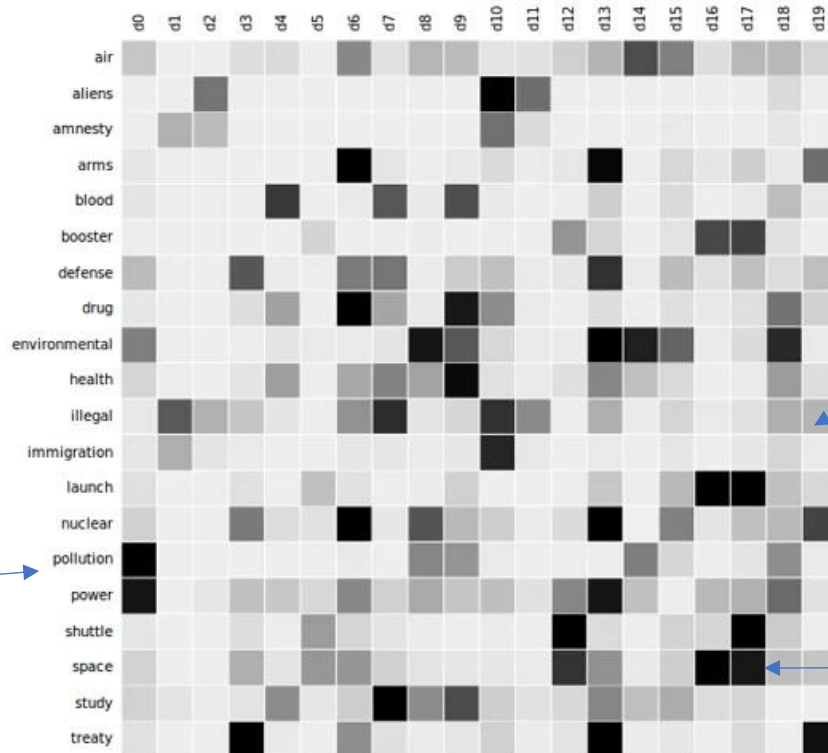
Objetivo e Metodologia

- **Base de Dados:** Corpus criado a partir de obras feministas representativas das quatro ondas de ativistas
- Aplicação do **algoritmo LDA** nos textos para descobrir computacionalmente os **principais temas tratados**
- Tecnicamente, trata-se de extrair dos documentos **tópicos representativos** das reivindicações de cada onda feminista
- Arquivos nos formato **Json**, script em **Python** para **PLN**

O algoritmo Latent Dirichlet Allocation (LDA)

**Matriz
Documento-
Palavra**

Cada linha
corresponde
a uma palavra



Cada coluna
corresponde a
um documento

Cada célula
especifica a
frequência de
uma palavra num
documento

Células mais
escuras
representam alta
frequência de
uma palavra num
documento

https://en.wikipedia.org/wiki/Topic_model#

**Organização e
Apoio:**



Resultados

- **Primeira Onda (1830-1920):** Luta por Igualdade Política como o direito ao voto para as mulheres
- **Tópicos encontrados pelo LDA**

(9,

$0.120 \cdot \text{"woman"} + 0.064 \cdot \text{"work"} + 0.029 \cdot \text{"fact"} + 0.020 \cdot \text{"suffrage"} +$
 $0.020 \cdot \text{"superior"} + 0.020 \cdot \text{"year"} + 0.019 \cdot \text{"political"} + 0.018 \cdot \text{"pay"} +$
 $0.017 \cdot \text{"equal"} + 0.017 \cdot \text{"class"}),$

(10,

$0.171 \cdot \text{"woman"} + 0.050 \cdot \text{"great"} + 0.047 \cdot \text{"true"} + 0.030 \cdot \text{"learn"} +$
 $0.026 \cdot \text{"good"} + 0.025 \cdot \text{"prove"} + 0.024 \cdot \text{"force"} + 0.017 \cdot \text{"emancipation"} +$
 $0.014 \cdot \text{"freedom"} + 0.012 \cdot \text{"demand"}),$

Organização e
Apoio:



Resultados

- Segunda Onda (1960-1980): Luta por Igualdade Social e de Gênero, por oportunidades para entrar no mercado de trabalho, por salários iguais, por liberdade sexual
- Tópicos encontrados pelo LDA

(12,

$0.049 \cdot \text{"human"} + 0.047 \cdot \text{"experience"} + 0.039 \cdot \text{"case"} + 0.034 \cdot \text{"play"} +$
 $0.033 \cdot \text{"equal"} + 0.024 \cdot \text{"society"} + 0.024 \cdot \text{"begin"} + 0.023 \cdot \text{"different"} +$
 $0.022 \cdot \text{"pay"} + 0.019 \cdot \text{"help"}),$

(19,

$0.069 \cdot \text{"also"} + 0.053 \cdot \text{"come"} + 0.049 \cdot \text{"self"} + 0.025 \cdot \text{"flesh"} +$
 $0.022 \cdot \text{"feminine"} + 0.022 \cdot \text{"nature"} + 0.020 \cdot \text{"succeed"} + 0.019 \cdot \text{"slave"} +$
 $0.018 \cdot \text{"partner"} + 0.018 \cdot \text{"worth"})$

Resultados

- **Terceira Onda (1990-2005):** Conscientização da Interseccionalidade de Diferentes Formas de Opressão (de Gênero, de Raça, de Classe, ...)
- **Tópicos encontrados pelo LDA**

(12,

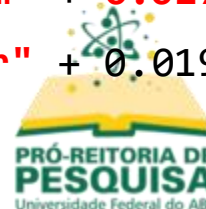
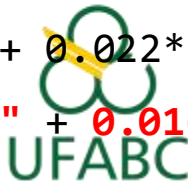
$0.286 * \text{"woman"} + 0.148 * \text{"man"} + 0.077 * \text{"black"} + 0.019 * \text{"oppression"} +$
 $0.015 * \text{"feel"} + 0.010 * \text{"struggle"} + 0.010 * \text{"suffrage"} + 0.009 * \text{"begin"} +$
 $0.008 * \text{"enjoy"} + 0.008 * \text{"end"}),$

(13,

$0.045 * \text{"movement"} + 0.042 * \text{"rape"} + 0.031 * \text{"racism"} + 0.027 * \text{"lynch"} +$
 $0.026 * \text{"racist"} + 0.022 * \text{"support"} + 0.021 * \text{"color"} + 0.019 * \text{"turn"} +$
 $0.019 * \text{"violence"} + 0.016 * \text{"anti"}),$

Organização e

Apoio:



Resultados

- **Quarta Onda (2008-????): Campanhas nas redes sociais, #MeToo, Time's Up, Women's March, Queer-inclusiva, Trans-inclusiva**
- **Tópicos encontrados pelo LDA**

(4,

$0.216 * \text{"man"} + 0.075 * \text{"woman"} + 0.043 * \text{"class"} + 0.036 * \text{"power"} +$
 $0.025 * \text{"hold"} + 0.024 * \text{"race"} + 0.022 * \text{"food"} + 0.020 * \text{"religion"} +$
 $0.019 * \text{"patriarchal"} + 0.014 * \text{"fact"}),$

(17,

$0.094 * \text{"people"} + 0.045 * \text{"thing"} + 0.031 * \text{"sexuality"} + 0.029 * \text{"white"} +$
 $0.026 * \text{"understand"} + 0.024 * \text{"problem"} + 0.024 * \text{"happen"} +$
 $0.022 * \text{"privilege"} + 0.022 * \text{"public"} + 0.021 * \text{"abortion"}),$

Conclusão e Trabalhos Futuros

- Escolha de autoras para o corpus
- Escolha mais abrangente de autoras
- Movimentos feministas x Obras publicadas
- Refinamento no número de tópicos
- Sobreposição de tópicos em diferentes ondas