

# Atribuição autoral de textos digitais

José Eleandro Custódio

Programa de Mestrado em Sistemas de Informação - PPgSI  
Universidade de São Paulo - USP

Novembro de 2018

# Informações gerais

- **Orientador:** Prof. Dr. Ivandré Paraboni
- **Semestre no curso:** 4o.
- **Qualificação:** 29/10/2018
- **Defesa:** realização planejada para 30/06/2019
- **Linha de pesquisa:** Inteligência de sistemas
- **Área de pesquisa:** Inteligência artificial
- **Área de aplicação:** Linguística computacional / Língua natural

# Agenda

- 1 Introdução
- 2 Conceitos
- 3 Trabalhos relacionados
- 4 Experimentos
- 5 Projeto de pesquisa
- 6 Referências

## Contexto

- A atribuição autoral de textos digitais (AA) (do inglês, *Authorship Attribution*) visa identificar quem é o autor de um determinado texto a partir de um conjunto de autores possíveis (POTTHAST et al., 2017).
- A premissa principal da AA é que o autor deixa rastros de seu estilo, sendo que esses rastros podem ser a preferência por certas palavras, o tamanho do vocabulário, a utilização de pontuação e a repetição de certos elementos gramaticais.
- A quantificação do estilo de escrita, ou estilometria, compreende um vasto conjunto de medidas e técnicas que buscam extrair uma “biometria” textual (NEAL et al., 2017).

# Aplicações da Atribuição Autoral

## Aplicações da AA

Sua aplicação pode ajudar:

- em casos de escândalos de corrupção, como no caso Enron (KLIMT; YANG, 2004; CHEN et al., 2011).
- na identificação de abusos na utilização da internet (VARTAPETIANCE; GILLAM, 2012).
- na detecção de notícias falsas (PENG; CHOO; ASHMAN, 2016).
- na detecção de casos onde uma pessoa tenta se passar por outra (KOPPEL; SEIDMAN, 2018).
- na atribuição autoral de código-fonte (ALSULAMI et al., 2017)
- na detecção de pseudônimos (JUOLA, 2015)

## Áreas de interesse

- Humanidades Digitais
- Análise Forense
- Linguística computacional

# Métodos para atribuição autoral

Os métodos computacionais para atribuição autoral utilizam:

- Análise estatística multivariada (SAVOY, 2016; EVERT et al., 2017).
- Métodos baseados em vizinho mais próximo (KOCHER; SAVOY, 2017; KOPPEL; SEIDMAN, 2018; VARELA et al., 2016).
- Modelos de compressão (HALVANI; GRANER, 2018).
- Aprendizado de máquina com SVM (SCHWARTZ et al., 2013; STAMATATOS, 2017).
- Redes neurais recorrentes (BAGNALL, 2016).
- Redes neurais de convolução (SHRESTHA et al., 2017; SARI; STEVENSON, 2016).

# Conceitos - Fatores que influenciam a AA

## Canal

- E-mail, jornais, livros, SMS
- Textos mais ou menos formais.

## Idioma

- Complexidade morfológica e lexical diferentes.

## Tópico

- Economia, celebridades, dia-a-dia
- Influencia o vocabulário.

## Domínio ou Gênero do texto

- Contos, artigos, avaliações de produtos
- Influencia no rigor formal e no vocabulário.

## Tamanho do texto

- Métodos probabilísticos são afetados pelo número de observações.

## Número de autores

- O aumento do número de classes requer o aumento do número de classes.

# Conceitos - Subtarefas da análise autoral

## AA de conjunto fechado

Os textos do conjunto de teste pertencem a um dos autores candidatos presentes no córpus de treinamento.

## AA de conjunto aberto

Os textos do conjunto de teste não necessariamente foram escritos por um dos autores do córpus de treinamento

## K-Atribuição ou ordenação

As saídas do classificador são ordenadas pela probabilidade e são retornados os K autores mais prováveis.

## Caracterização

São extraídas informações demográficas do autor do texto podem reduzir a lista de candidatos.

## Verificação

Verifica-se se dois documentos foram escritos pelo mesmo autor, não sendo necessário saber quem são os autores.

## Demais

Agrupamento, Ligação e Quebra de estilo.



# Conceitos - Tipos de conhecimentos usados

A abordagem estilométrica tradicional utiliza as seguintes fontes de conhecimento:

- **Categoria lexical:** tamanho médio das palavras, número de letras maiúsculas, tamanho das sentenças, etc.
- **Categoria sintática:** frequência da pontuação, palavras de função, etc.
- **Categoria semântica:** contagem das palavras, analisadores semânticos, *word embeddings*, etc.
- **Categoria específica de domínio:** palavras-chave, *tags* HTML, *emoticons*, etc.

# Conceitos - Tipos de conhecimentos usados

## Palavras

- As palavras mais frequentes são independentes de domínio e utilizadas de forma inconsciente (KESTEMONT, 2014).
- **Palavras de função** (do inglês, *function words*) compreende artigos, preposições, locuções adverbiais, e outros.
- Capturam semântica e elementos de conexão entre sentenças.

## Caracteres

- As sequências de caracteres são considerados os modelos mais efetivos para AA (KJELL; WOODS; FRIEDER, 1994; NEAL et al., 2017).
- Os **caracteres mais frequentes (CNG)** (do inglês, *common n-grams*) (KEŠELJ et al., 2003; SAPKOTA et al., 2014).
- Os n-gramas de caracteres conseguem capturar pontuação, utilização de espaços, preferências temporais, palavras de função de tamanho curto.

## Modelo tradicional de representação textual

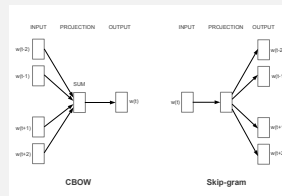
- Hipótese distribucional aplicada a documentos (TURNEY; PANTEL, 2010)
- N-gramas
- Conhecimentos: Caracteres, Palavras, POS
- TF-IDF
- TF-IDF equações alternativas utilizadas nos experimentos:
  - $TF_{sublinear} = 1 + \log TF_{t,d} \rightarrow$  Definição I do SMART.
  - $IDF_{Suavizado}(t, D) = \log \left( \frac{D}{DF(t)} \right) + 1$

# Conceitos - Modelos computacionais

## Modelo de representação distribuída

- Hipótese distribucional aplicada símbolos
- Modelos neurais de língua natural (BENGIO et al., 2003)
- *Word embeddings*
  - Representa cada palavra, ou símbolo, em  $k$  dimensões, sendo  $k$  menor que o tamanho do vocabulário.
  - Word2Vec (MIKOLOV et al., 2013)
  - Doc2Vec (LE; MIKOLOV, 2014)
  - FastText (BOJANOWSKI et al., 2017)

## Modelo Word2Vec



Fonte: Mikolov et al. (2013)

Table 1: Trabalhos selecionados

Estudo	Idioma	Tarefa	Conhecimento	Método
Sapkota et al. (2015)	EN	A	<i>C</i>	SVM
Stamatatos (2017)	EN	A,V	<i>C, W</i>	SVM
Schwartz et al. (2013)	EN	A	<i>C, W</i>	SVM
Rocha et al. (2017)	EN	A,V	<i>C, W, P</i>	SVM, RF e SCAP
Evert et al. (2017)	EN	C	<i>W</i>	Clusterização
Varela et al. (2016)	PT-BR	A,V	<i>P</i>	SVM
Posadas-Durán et al. (2017)	EN	A	D2V de <i>W</i>	Softmax e SVM
Rhodes (2015)	EN	A	W2V	CNN-Softmax
Shrestha et al. (2017)	EN	A	<i>C One-hot</i>	Softmax
Bagnall (2016)	PAN2015	C	<i>C One-hot</i>	RNN-Softmax

José Eleandro Custódio, 2018

# Experimentos

## Experimentos

# Experimento 1: Verificação autoral

## Publicação 1

CUSTÓDIO, J. E.; PARABONI, I. Similaridade de Textos aplicada à Verificação Autoral. In: 1st International Congress on Digital Humanities in Rio de Janeiro. [S.l.]: Fundação Getúlio Vargas, 2018.

## Verificação autoral ou atribuição por similaridade

- Deseja-se saber se pares de documentos foram escritos pelo mesmo autor. (KOPPEL et al., 2012)
- Aplicável quando não se sabe quem são os autores.
- Modelo supervisionado por vizinho mais próximo.
  - O documento é atribuído ao vizinho mais próximo.
  - A distância pode ser usada no agrupamento autoral.
- Modelo transformado
  - Documentos são uma representação única.

# Experimento 1: Verificação autoral

## Extração de características

Modelo de espaço de vetores (BOW) com n-gramas de caracteres normalizados com norma L1 (TF).

Foram selecionadas os n-gramas presentes em 90% do cópulus (*Common n-grams* (KEŠELJ et al., 2003)).

## Distâncias

Medidas de similaridade textual entre os documentos A e B do cópulus C:

$$\text{Cossenos}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} \quad (1)$$

$$\text{Jaccard}(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

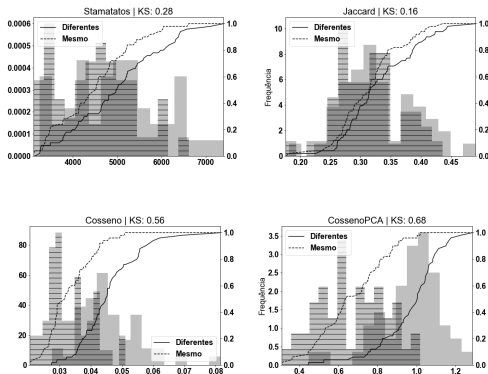
$$\text{Stamatatos}(A, B, N) = \sum_i \left( \frac{2 * (A - B)}{A + B} \right)^2 * \left( \frac{2 * (A - C)}{A + C} \right)^2 \quad (3)$$



# Experimento 1: Verificação autoral

Análise da capacidade de separação das medidas de similaridade aplicadas a corpus PAN-CLEF 2014 (STAMATATOS et al., 2014)

Figure 1: Diagnósticos



- Histograma para mesma autoria (tracejado).
- Histograma para autorias diferentes (liso).
- Distribuição acumuladas (linhas).
- Separação pela métrica Kolmogorov-Smirnov.
- Métricas AUC e acurácia.

# Experimento 1: Verificação autoral

## Modelo proposto 1 - MP1

- As distâncias foram utilizadas como variáveis para o modelo.
- Aplicado a normalização minmax.
- Aplicado a regressão logística.

## Modelo proposto 2 - MP2

- Os documentos conhecidos  $C$  e de autoria desconhecidas  $D$  foram unificados em um único BoW através da equação:

$$MP2(C_{ij}, D_{ij}) = \log \left( 1 + \frac{(C_{ij} - D_{ij})^2}{C_{ij} + 1} \right) \quad (4)$$

- Aplicado a normalização minmax.
- Aplicado a regressão logística.

# Experimento 1: Verificação autoral

**Table 2:** Verificação autoral - Resultados médios das métricas AUC e acurácia em 5-partições.

Modelo	PAN2014 (EE e EM)		PAN2014-SP	
	ROC	Acurácia	ROC	Acurácia
Jaccard	0,60	0,56	0,57	0,52
Cossenos	0,63	0,50	0,88	0,77
Cossenos_PCA	0,63	0,55	<b>0,92</b>	<b>0,83</b>
Keselj	0,61	0,54	0,71	0,60
Stamatatos	0,60	0,55	0,59	0,54
MP1 – Mix	<b>0,75</b>	<b>0,67</b>	0,72	0,62
MP2 – BOW	0,62	0,53	<b>0,93</b>	<b>0,85</b>

PAN2014 (EE e EM) corpus com textos em língua inglesa, PAN2014-SP textos em língua espanhola.

# Experimento 2: Atribuição Autoral

## Publicação 2

CUSTÓDIO, J. E.; PARABONI, I. EACH-USP Ensemble Cross-domain Authorship Attribution: Notebook for PAN at CLEF 2018. In: CAPPELLATO, L. et al. (Ed.). Working Notes Papers of the CLEF 2018 Evaluation Labs. [S.l.]: CLEF and CEUR-WS.org, 2018. (CEUR Workshop Proceedings). ISSN 1613-0073.

## Atribuição por aprendizado de máquina supervisionado

- Tem-se um conjunto de documentos para os quais se sabe quem são os autores e um documento do qual deseja-se atribuir.
- O classificador extrai a “assinatura do estilo”.
- Aspectos inconscientes, como a sintaxe, são mais importantes que a semântica.
- O trabalho apresentado foi parte da participação da tarefa de AA da competição PAN-CLEF2018.

# Experimento 2: Atribuição Autoral

## Baseline *Bas.PAN*

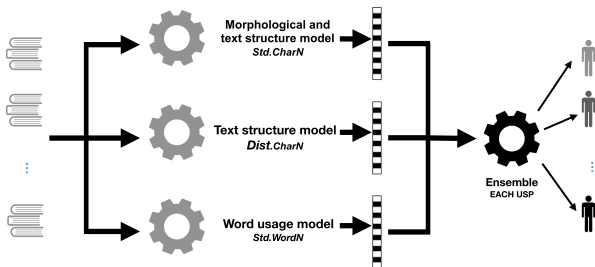
Os organizadores forneceram um sistema *baseline* pelos com as seguintes características:

- N-gramas de caracteres de tamanho fixo.
- Normalização no documento por TF.
- Sem normalização no cópús.
- Frequência mínima de 4 ocorrências.
- Classificador SVM encapsulado nas estratégias um-contra-um e um-contra-todos.
- Foi otimizado por *grid search* com validação cruzada com 5 partições, e os melhores parâmetros foram:
  - n-gramas de tamanho 4.
  - Frequência mínima de 5 documentos.
  - SVM com estratégia um-contra-todos.

# Experimento 2: Atribuição Autoral

Dado nossas premissas, o sistema final PAN2018 para AA consistiu de um comitê que concatenou as fontes de informações em uma saída única.

Figure 2: Método proposto final



O sistema foi otimizado por *grid search* com validação cruzada de 5 partições.

# Experimento 2: Atribuição Autoral

Premissa: O estilo de escrita de autor pode ser capturado através de diversas fontes de informação, como sintática, léxica e semântica.

## Método proposto *Std.word*

Consistiu de um modelo BOW de n-gramas de palavras tradicional.

## Método proposto *Std.char*

Consistiu de um modelo BOW de n-gramas de caracteres tradicional.

## Método proposto *Dist.char*

Consistiu de um modelo BOW de n-gramas de caracteres onde são letras maiúscula e minúsculas sem acento são distorcidas, deixando a pontuação, espaços e letras com diacríticos.

# Experimento - Modelo de distorção textual em caracteres

O modelo *Dist.charN* utiliza distorção textual e foi inspirado em Stamatatos (2017). O objetivo da distorção é mascarar conteúdos que estariam se comportando como ruído para o modelo. Stamatatos (2017) utiliza distorções para mascarar palavras menos frequentes, que estariam relacionadas ao tópico. O modelo *Dist.charN* atua sobre carácter, filtrando a utilização de espaços, pontuação e caracteres com diacríticos, de forma a mascarar caracteres comuns, ou seja, letras minúsculas e maiúsculas. A tabela 3 ilustra a aplicação desse método.

**Table 3:** Exemplo de distorção de texto aplicado o 1o. documento do 9o. problema da base de treinamento.

Texto original	Texto transformado
-¿Y cómo sabes que no lo ama?	-¿* *ó** ***** *** ** ** ***?
-Inglaterra se preguntó a su	-***** ** *****ó * **
vez si habría un muñeco del	*** ** *****í* ** **ñ*** **
esposo también.	***** *****é*.



# Experimento 2: Atribuição Autoral

Table 4: Valores ótimos encontrados para PAN2018

Módulo	Parâmetros	Valores ótimos
Extração de características	Faixa n-gram	Std.charN - Início=2 Fim=5 Dist.charN - Início=2 Fim=5 Std.WordN - Início=1 Fim=3
	Freq. min. doc.	0,05
	Freq. max. doc.	1,0
	TF	Sublinear
	IDF	Suavizado
	Normalização no documento	L2
Transformação	PCA	0,99

# Experimento 2: Resultados obtidos em treinamento

Table 5: F1 para PAN-CLEF 2018 AA no corpus de desenvolvimento

Problema	Língua	Autores	Bas.PAN	Std.charN	Dist.charN	Std.wordN	Comitê
01	EN	20	0,514	0,609	0,479	0,444	<b>0,625</b>
02	EN	5	0,626	0,535	0,333	0,577	<b>0,673</b>
03	FR	20	0,631	0,681	0,568	0,418	<b>0,776</b>
04	FR	5	0,747	0,719	0,586	0,572	<b>0,820</b>
05	IT	20	0,529	0,597	0,491	0,497	<b>0,578</b>
06	IT	5	0,614	0,623	0,595	0,520	<b>0,663</b>
07	PL	20	0,455	0,470	0,496	0,475	0,554
08	PL	5	0,703	<b>0,948</b>	0,570	<b>0,922</b>	0,922
09	ES	20	0,709	<b>0,774</b>	0,589	0,616	0,701
10	ES	5	0,593	0,778	<b>0,802</b>	0,588	<b>0,830</b>
Média			0,612	0,673	0,551	0,563	<b>0,714</b>

# Experimento 2: Resultados obtidos em produção

Resultado geral apresentados pelos organizadores do PAN2018 (KESTEMONT et al., 2018).

Table 6: PAN-CLEF 2018 - 3 melhores equipes - por língua

Equipe	F1 Geral	EN	FR	IT	PL	ES
Custódio e Paraboni (2018)	<b>0,685</b>	0,744	<b>0,668</b>	0,676	0,482	<b>0,856</b>
Murauer, Tschuggnall e Specht (2018)	0,643	<b>0,762</b>	0,607	0,663	0,450	0,734
Halvani e Graner (2018)	0,629	0,679	0,536	<b>0,752</b>	0,426	0,751
PAN18-BASELINE	0,584	0,697	0,585	0,605	0,419	0,615

Table 7: PAN-CLEF 2018 - 3 melhores equipes - por língua

Equipe	Quantidade de autores			
	20	15	10	5
Custódio e Paraboni (2018)	<b>0,648</b>	<b>0,676</b>	<b>0,739</b>	<b>0,677</b>
Murauer, Tschuggnall e Specht (2018)	0,609	0,642	0,680	0,642
Halvani e Graner (2018)	0,609	0,605	0,665	0,636
PAN18-BASELINE	0,546	0,532	0,595	0,663

# Experimento - Características mais importantes

Table 8: Características textuais mais relevantes para *Std.charN*

Candidatos				
01	02	03	04	05
_as_l	_Sti	_sub	_joi	_day,
_'	_" Can	_suc	_gh	_dev
_prec	_" Ca	_l_fi	_er	_dete
_l'd	_" Be	_succ	_glow	_plac
_" Are	_K	_subs	_ls	_mut
_Re	_but_	_l_f	_sta	_must
_smel	_Ofte	_" T	_gor	_Dro
_leak	_posi	_a_t	_sorr	_day_
_is_s	_For	_" St	_eat_	_she_
_spu	_Ri	_a_sw	_lf_t	_chi

Extraído do subconjunto 02 com textos em inglês e com 5 autores.

# Experimento - Anexo: Características mais relevantes

**Table 9:** Características textuais mais relevantes para *Dist.charN*

Candidatos				
01	02	03	04	05
*_**	##_	"*'	*_~	*_~*
##_	##_ (	"*_**	*_~	'*_*
*'	##_*	!)*_*	*_~*	"_
**).	*!	*!!	'***	*_~
**),_	##_'	*'*_*	'****	*_~
*_~*	*!_*	**_*'	"_**'	'*.
*_~	*_""**	**_*	_É***	_""*
'**	_~	**_*'	_""*	_~
!),	_~_*	_**!	_**..	_~

Extraído do subconjunto 02 com textos em inglês e com 5 autores.

# Experimento - Características mais relevantes

Table 10: Características textuais mais relevantes em *Std.wordN*

Candidatos				
01	02	03	04	05
about_what and_practically any_of any_more and_nearly and_pulled agree all_tony ah and_wet_and	against_his and_it_was and_so and_already and_steve and_say accent and_wet apparently after	an_odd and_then_he acknowledged and_he_had are_your again_to and_tell and_forth are_just and_grabbing	although an_eye and_said and_takes and_just ancient amount_of always and_grinned about_the	and_pulled_him and_pulling across_his across_the and_all against_her among about_what_to acting about_their

Extraído do subconjunto 02 com textos em inglês e com 5 autores.

# Projeto de pesquisa

Projeto de pesquisa

# Projeto - Lacunas e motivação

## Lacunas gerais:

- A AA é um problema de pesquisa não totalmente resolvido (POTTHAST et al., 2017).
- É o tema da série de competições PAN-CLEF (KESTEMONT et al., 2018) (KOCHER; SAVOY, 2017; KOPPEL; SEIDMAN, 2018; VARELA et al., 2016).
- Estudos desta área exploram técnicas independentes de idioma e de domínio, subutilizando recursos linguístico-computacionais.

O trabalho em Custódio e Paraboni (2018) apresentou o melhor desempenho global na edição de 2018 da competição PAN-CLEF, no entanto, deixa as seguintes lacunas:

- não tirou proveito de conhecimentos dependentes de idioma como *part-of-speech* (POS).
- e modelos de representação distribuída (*word embeddings*).
- foi restrito ao domínio *Fanfic*.
- não considerou dados em português brasileiro.



# Projeto - Hipóteses

Este trabalho considera as seguintes hipóteses:

H1:

O uso de modelos independentes de idioma do tipo de distorção textual permite filtrar aspectos específicos do texto, e a combinação de diversos tipos de distorção pode aumentar o desempenho de sistemas de AA.

H2:

O uso de modelos dependentes de idioma do tipo *part-of-speech* extraídos por anotadores baseados em aprendizado profundo pode aumentar o desempenho de sistemas de AA.

H3:

O uso de modelos dependentes de idioma do tipo representação distribuída (*embeddings*) pode aumentar o desempenho de sistemas de AA.

# Projeto - Objetivo

## Objetivo Geral

O objetivo geral deste trabalho é enriquecer modelos de atribuição autoral de texto digitais com conjunto fechado de autores utilizando conhecimentos dependentes e independentes de idioma combinados com técnicas de aprendizados de máquina, de modo a obter resultados superiores ao estabelecido em trabalhos anteriores.

# Projeto - Conjunto de dados e Avaliação

## Avaliação das hipóteses

- Comparação com *baselines* pertinentes, como o modelo apresentado em Custódio e Paraboni (2018).
- Serão utilizadas as medidas tradicionais de AM, como *medida F*, acurácia, auROC e outros.
- Espera-se que o resultado médio seja superior ao dos modelos de *baseline*.

Table 11: Córpus para avaliação dos métodos de AA

Córpus	No. Autores	Idioma	Domínio/Gênero
PAN-CLEF2014	-	EN, ES, DU, GR	NV, AR, RV, ES
PAN-CLEF2018	20	EN, ES, FR, IT, PL	NV
RCV1	50	EN	AR
Nus-SMS	116	EN	SMS
b5-post	1.019	PT-Br	Facebook
BlogSet-BR	4.331	PT-Br	AR

# Projeto - Escopo e limitações

Este projeto de pesquisa se limita

- ao estudo das técnicas de distorção textual
- ao estudo das técnicas de anotações linguísticas
- ao estudo das técnicas de representação distribuída
- e utilizará métodos de aprendizado de máquina.
- aos idiomas considerados primordialmente são inglês e português brasileiro.

Não serão considerados

- modelos computacionais baseados em grafos, redes complexas e modelos de compressão.

# Projeto - Atividades

- 1 **Revisão bibliográfica** Concluído.
- 2 **Participação na PAN-CLEF 2018** Concluído.
- 3 **Preparação dos dados** Concluído.
- 4 **Modelos independentes de idioma** Estudo dos tipo de distorção textual, construção dos modelos computacionais e refinamentos específicos.
- 5 **Modelos baseados em anotações** Estudo dos pacotes para anotações POS, como NLTK<sup>1</sup> e Spacy, construção dos modelos computacionais e refinamentos específicos.
- 6 **Modelos baseados em *embeddings*** Preparação de bases de dados de *embeddings*, estudo de *embeddings* específicos para AA, construção de modelos computacionais e refinamentos.
- 7 **Refinamentos**
- 8 **Avaliação**
- 9 **Redação da dissertação**
- 10 **Divulgação**

Table 12: Cronograma

	2018							2019						
Atividades	1-6	7	8	9	10	11	12	1	2	3	4	5	6	
01. Revisão bibliográfica	x	x	x	x	x									
02. Participação PAN-CLEF2018	x													
03. Preparação dos dados	x	x				x								
04. Modelos independentes de idioma						x	x	x						
05. Modelos baseados em anotações		x				x	x	x						
06. Modelos baseados em <i>embeddings</i>		x						x	x					
07. Refinamentos								x	x	x				
08. Avaliação final										x				
09. Redação da dissertação										x	x	x		
10. Divulgação												x	x	



 ALSULAMI, B., DAUBER, E., HARANG, R. E., MANCORIDIS, S., AND GREENSTADT, R.

Source code authorship attribution using long short-term memory based networks.

In *Computer Security - ESORICS 2017 - 22nd European Symposium on Research in Computer Security, Oslo, Norway, September 11-15, 2017, Proceedings, Part I* (2017), pp. 65–82.




BAGNALL, D.

Authorship clustering using multi-headed recurrent neural networks.

In *CEUR Workshop Proceedings* (2016), M. C. B. K. Cappellato L. Ferro N., Ed., vol. 1609, CEUR-WS, pp. 791–804.

# Referências II

 BENGIO, Y., DUCHARME, R., VINCENT, P., AND JANVIN, C.

A Neural Probabilistic Language Model.

*Journal of Machine Learning Research* 3 (2003), 1137–1155.

 BOJANOWSKI, P., GRAVE, E., JOULIN, A., AND MIKOLOV, T.

Enriching word vectors with subword information.

*Transactions of the Association for Computational Linguistics* 5 (2017), 135–146.



# Referências III

 CHEN, X., HAO, P., CHANDRAMOULI, R., AND SUBBALAKSHMI, K. P.

Authorship similarity detection from email messages.


In *Machine Learning and Data Mining in Pattern Recognition - 7th International Conference, MLDM 2011, New York, NY, USA, August 30 - September 3, 2011. Proceedings* (2011), pp. 375–386.


 CUSTÓDIO, J. E., AND PARABONI, I.

EACH-USP Ensemble Cross-domain Authorship Attribution: Notebook for PAN at CLEF 2018.

In *Working Notes Papers of the CLEF 2018 Evaluation Labs* (Sept. 2018), L. Cappellato, N. Ferro, J.-Y. Nie, and L. Soulier, Eds., CEUR Workshop Proceedings, CLEF and CEUR-WS.org.

# Referências IV

 EVERT, S., PROISL, T., JANNIDIS, F., REGER, I., PIELSTRÖM, S., SCHÖCH, C., AND VITT, T.  
Understanding and explaining delta measures for authorship attribution.  
*Digital Scholarship in the Humanities* 32, suppl\_2 (2017), ii4–ii16.

 HALVANI, O., AND GRANER, L.  
Cross-Domain Authorship Attribution Based on Compression:  
Notebook for PAN at CLEF 2018.  
In *Working Notes Papers of the CLEF 2018 Evaluation Labs*  
(Sept. 2018), L. Cappellato, N. Ferro, J.-Y. Nie, and L. Soulier,  
Eds., CEUR Workshop Proceedings, CLEF and CEUR-WS.org.

# Referências V

 JUOLA, P.

The rowling case: A proposed standard analytic protocol for authorship questions.

*Digital Scholarship in the Humanitie* 30, Suppl-1 (2015), i100–i113.

 KEŠELJ, V., PENG, F., CERCONE, N., AND THOMAS, C.

N-Gram-Based Author Profiles for Authorship Attribution.

In *Proceedings of the conference pacific association for computational linguistics (PACLING)* (2003), vol. 3, pp. 255–264.

# Referências VI

 KESTEMONT, M.

Function Words in Authorship Attribution From Black Magic to Theory ?


*3rd Workshop on Computational Linguistics for Literature (CLfL 2014)*, January 2014 (2014), 59–66.

 KESTEMONT, M., TSCHUGNALL, M., STAMATATOS, E., DAELEMANS, W., SPECHT, G., STEIN, B., AND POTTHAST, M.

Overview of the Author Identification Task at PAN-2018:  
Cross-domain Authorship Attribution and Style Change Detection.

In *Working Notes Papers of the CLEF 2018 Evaluation Labs* (Sept. 2018), L. Cappellato, N. Ferro, J.-Y. Nie, and L. Soulier, Eds., CEUR Workshop Proceedings, CLEF and CEUR-WS.org.

# Referências VII

 KJELL, B., WOODS, W. A., AND FRIEDER, O.  
Discrimination of authorship using visualization.  
*Inf. Process. Manage.* 30, 1 (1994), 141–150.

 KLIMT, B., AND YANG, Y.  
The enron corpus: A new dataset for email classification research.

In *Machine Learning: ECML 2004, 15th European Conference on Machine Learning, Pisa, Italy, September 20-24, 2004, Proceedings* (2004), J. Boulicaut, F. Esposito, F. Giannotti, and D. Pedreschi, Eds., vol. 3201 of *Lecture Notes in Computer Science*, Springer, pp. 217–226.

# Referências VIII

 KOCHER, M., AND SAVOY, J.

A simple and efficient algorithm for authorship verification.

*Journal of the Association for Information Science and Technology* 68, 1 (2017), 259–269.

 KOPPEL, M., SCHLER, J., ARGAMON, S., AND WINTER, Y.

The "Fundamental Problem" of Authorship Attribution.

*English Studies* 93, 3 (2012), 284–291.

 KOPPEL, M., AND SEIDMAN, S.

Detecting pseudepigraphic texts using novel similarity measures.

*Digital Scholarship in the Humanities* 33, 1 (2018), 72–81.

# Referências IX

 LE, Q. V., AND MIKOLOV, T.

Distributed representations of sentences and documents.

In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014* (2014), pp. 1188–1196.

 MIKOLOV, T., CHEN, K., CORRADO, G., AND DEAN, J.

Efficient estimation of word representations in vector space.

*CoRR abs/1301.3781* (2013).

 MURAUER, B., TSCHUGGNALL, M., AND SPECHT, G.

Dynamic Parameter Search for Cross-Domain Authorship Attribution: Notebook for PAN at CLEF 2018.

In *Working Notes Papers of the CLEF 2018 Evaluation Labs* (Sept. 2018), L. Cappellato, N. Ferro, J.-Y. Nie, and L. Soulier, Eds., CEUR Workshop Proceedings, CLEF and CEUR-WS.org.

 NEAL, T. J., SUNDARARAJAN, K., FATIMA, A., YAN, Y., XIANG, Y., AND WOODARD, D. L.

Surveying stylometry techniques and applications.

*ACM Comput. Surv.* 50, 6 (2017), 86:1–86:36.

 PENG, J., KWANG RAYMOND CHOO, K., AND ASHMAN, H.

Astroturfing detection in social media: Using binary n-gram analysis for authorship attribution.

In *2016 IEEE Trustcom/BigDataSE/ISPA* (2016), pp. 121–128.



# Referências XI

 POSADAS-DURÁN, J.-P., GÓMEZ-ADORNO, H., SIDOROV, G., BATYRSHIN, I., PINTO, D., AND CHANONA-HERNÁNDEZ, L.

Applications of the distributed document representation in the authorship attribution task for small corpora.

*Soft Computing* 21, 3 (feb 2017), 627–639.

 POTTHAST, M., RANGEL, F., TSCHUGGNALL, M., STAMATATOS, E., ROSSO, P., AND STEIN, B.

Overview of PAN'17: Author identification, author profiling, and author obfuscation.


*Lecture Notes in Computer Science 10456 LNCS* (2017), 275–290.

# Referências XII

 RHODES, D.

Author Attribution with CNN's.

*Stanford University - CS224D Projects* (2015), 1–8.

 ROCHA, A., SCHEIRER, W. J., FORSTALL, C. W.,  
CAVALCANTE, T., THEOPHILO, A., SHEN, B., CARVALHO,  
A., AND STAMATATOS, E.

Authorship Attribution for Social Media Forensics.


*IEEE Transactions on Information Forensics and Security* 12, 1  
(2017), 5–33.

# Referências XIII

 SAPKOTA, U., BETHARD, S., MONTES-Y-GÓMEZ, M., AND SOLORIO, T.

Not all character n-grams are created equal: A study in authorship attribution.

In *NAACL HLT 2015, The 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Denver, Colorado, USA* (2015), pp. 93–102.

 SAPKOTA, U., SOLORIO, T., MONTES, M., BETHARD, S., AND ROSSO, P.

Cross-Topic Authorship Attribution: Will Out-Of-Topic Data Help?

# Referências XIV

In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers* (2014), pp. 1228–1237.



SARI, Y., AND STEVENSON, M.


Exploring Word Embeddings and Character N -Grams for Author Clustering Notebook for PAN at CLEF 2016.  
*CEUR Workshop Proceedings* (2016).



SAVOY, J.

Estimating the probability of an authorship attribution.  
*Journal of the Association for Information Science and Technology* 67, 6 (2016), 1462–1472.

# Referências XV

 SCHWARTZ, R., TSUR, O., RAPPOPORT, A., AND KOPPEL, M.

Authorship Attribution of Micro-Messages.

In *Empirical Methods in Natural Language Processing* (2013), pp. 1880–1891.

 SHRESTHA, P., SIERRA, S., GONZÁLEZ, F., ROSSO, P., MONTES-Y-GÓMEZ, M., AND SOLORIO, T.

Convolutional Neural Networks for Authorship Attribution of Short Texts.


In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics* (2017), vol. 2, Association for Computational Linguistics (ACL), pp. 669–674.

# Referências XVI

 STAMATATOS, E.

Authorship attribution using text distortion.

*Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017 - Proceedings of Conference 1* (2017).

 STAMATATOS, E., DAELEMANS, W., VERHOEVEN, B., STEIN, B., POTTHAST, M., JUOLA, P., SÁNCHEZ-PÉREZ, M. A., AND BARRÓN-CEDENO, A.

Overview of the author identification task at PAN 2014.

In *Working Notes for CLEF 2014 Conference, Sheffield, UK, September 15-18, 2014*. (2014), pp. 877–897.

 TURNEY, P. D., AND PANTEL, P.

From frequency to meaning: Vector space models of semantics.  
*Journal of artificial intelligence research abs/1003.1141* (2010).

# Referências XVII

 VARELA, P. J., JUSTINO, E. J., BORTOLOZZI, F., AND OLIVEIRA, L. E.

A computational approach based on syntactic levels of language in authorship attribution.

*IEEE Latin America Transactions* 14, 1 (2016), 259–266.

 VARTAPETIANCE, A., AND GILLAM, L.

Quite simple approaches for authorship attribution, intrinsic plagiarism detection and sexual predator identification.

In *CLEF 2012 Evaluation Labs and Workshop, Online Working Notes, Rome, Italy, September 17-20, 2012* (2012).