

# O digital nas humanidades: desafios na formação profissional

Uma contribuição de Letras

---

CLÁUDIA FREITAS

PUC-RIO/PPGEL

LINGUATECA

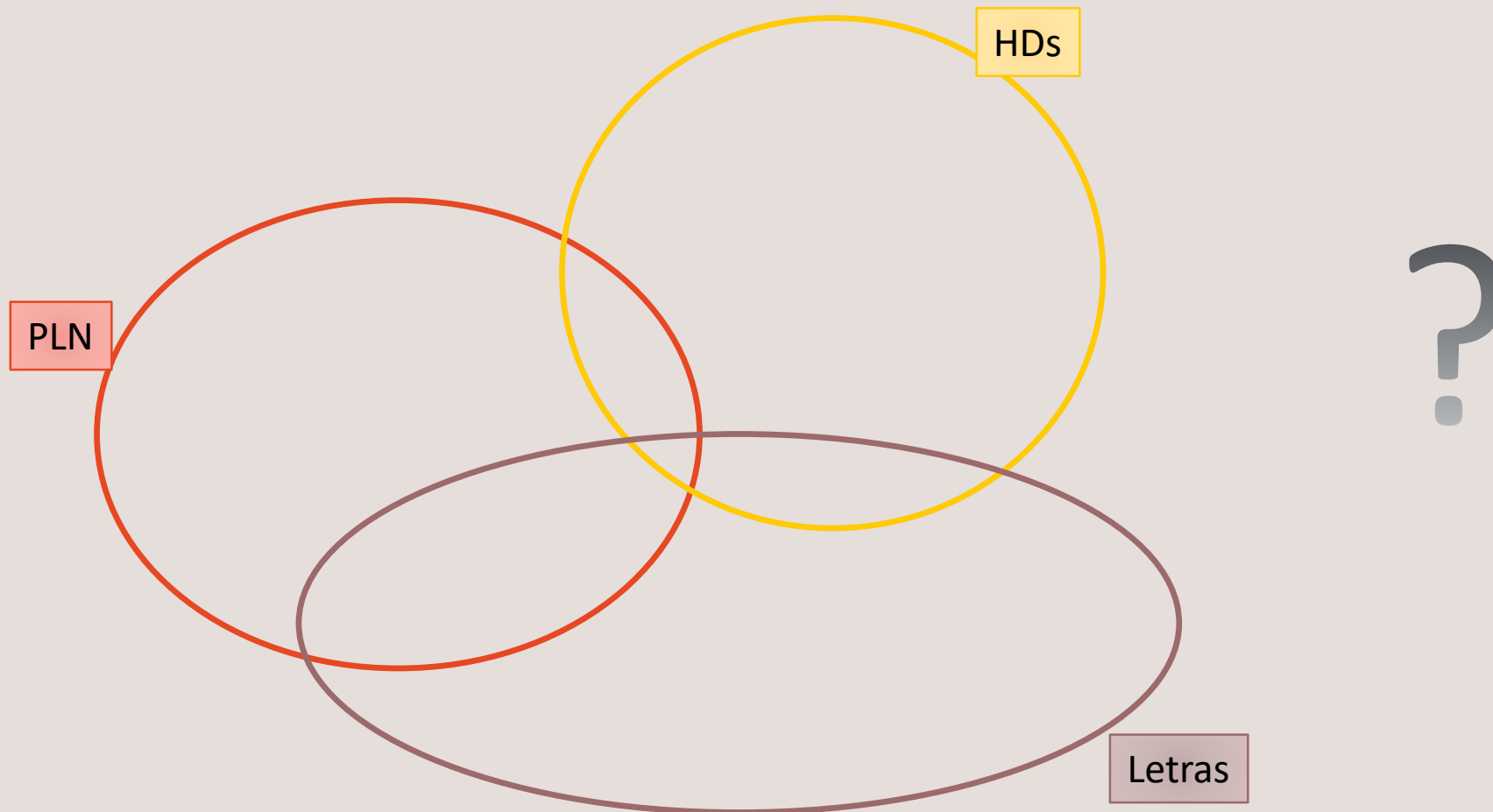
# Linguateca

<http://www.linguateca.pt>

- Centro distribuído para a criação de recursos para o PLN em português
  - Uma infraestrutura para investigar “língua\_cultura” com computadores
- Criada em 1998!!!!
- Hoje: acervo de mais de 1 bilhão de palavras
  - Anotadas morfossintaticamente
  - Anotadas semanticamente
- Repositório de outros serviços e ou atividades:
  - Tradução; Ensino; Datasets; Avaliações conjuntas (de sistemas)
- Todo o material é PÚBLICO e ABERTO

# InterInterdisciplinaridade

---



# Um exemplo

---

- PLN: Tarefa de Extração de Citações (“Quotation Extraction”)
  - alguém DIZER algo
- Linguística: O que é “dizer”? Quais as maneiras de “dizer” em português?
  - Quando estamos diante de um discurso relatado?

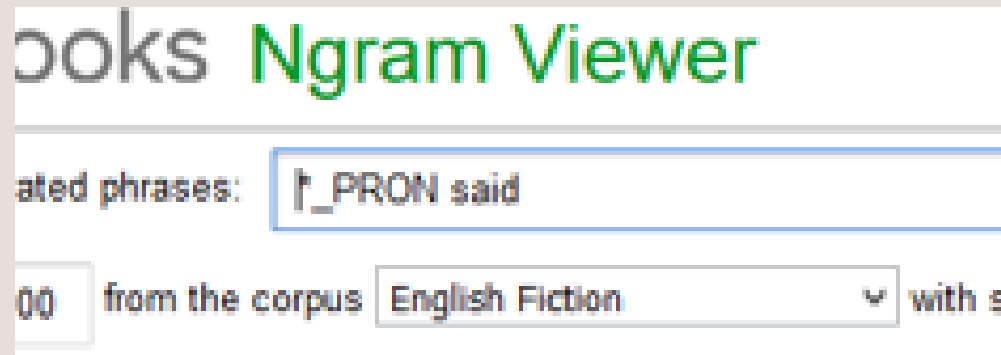
*“Hoje **diria** que ha um movimento que se gera a partir do Me e da movimentação de base que existe no Tecnico!”*

- HDs: Detecção de viés em notícias de jornal
  - Quem fala (e, portanto, quem cala?) → materialidade para a ideia de “viés”?



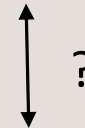
# As palavras se distribuem em classes conforme as suas propriedades fundamentais (?)

---



Empresas usam liminares para manter nas ruas painéis proibidos.

Empresas usam liminares para manter painéis (*que foram*) **proibidos** (VERBO)



Empresas usam liminares para manter painéis **proibidos** (=ilegais) (ADJETIVO)

# O significado da palavra *está* na palavra (?)

---

“São muitas as palavras que não têm um significado preciso. Mas isto não é um defeito. Pensar o contrário seria como afirmar que a luz do meu candeeiro não é uma luz verdadeira porque não tem um limite bem definido”

(Wittgenstein, L. Livro Azul, p.61).

# O significado do texto *está* no texto (?)

*Quando os deuses coroaram de conquista as armas, quando os soberbos muros de Troia fumegaram por terra, a Grécia, para recompensar as galhardas fadigas de seu soldado, cumulou sua armada de incontáveis despojos. Assim, grande em glória, voltou seguro do estrondo marcial, sem uma cicatriz hostil, e embora as lanças se fechassem à sua volta em tormentas de ferro, seu jogo inútil foi inocente de ferimentos.*

2 traduções **consagradas** dos mesmos  
versos da *Odisseia*

*Uma vez ocupada a cidade, ele pôde apanhar e embarcar sua parte de benefícios havidos, que era uma forte soma. Saiu sem um arranhão de toda essa perigosa campanha. Já se sabe: tudo *está* em ter sorte.*

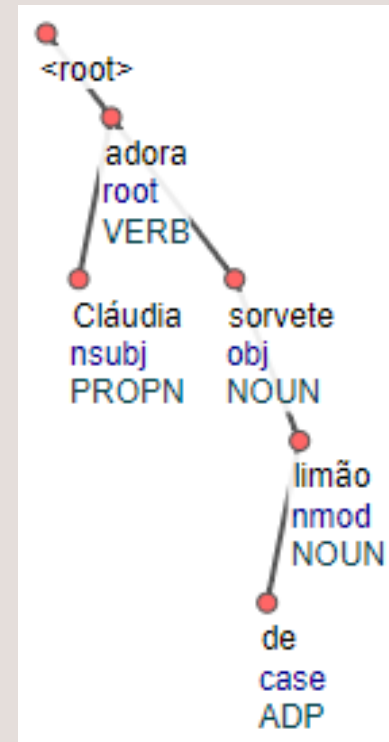


# A análise linguística é linear (?)



Cláudia adora sorvete de limão

<b>Cláudia</b>	Cláudia	PROPN	2 nsubj	agente	pessoa
<b>adora</b>	adorar	VERB	0 root	predicado	Amar/sentimento
<b>sorvete</b>	sorvete	NOUN	2 obj		alimento
<b>de</b>	de	ADP	5 case		---
<b>limão</b>	limão	NOUN	3 nmod	tema	fruta



Com estas palavras, André Coruja, além de **quebrar o gelo que havia esfriado o clima**, devolveu ao recinto a eloquência necessária para que a sessão continuasse.

# Linguagem

---

Línguas são sistemas complexos, **abertos** e **dinâmicos**

Sentido das palavras, e dos textos, é **instável**

- Faz pouco sentido a ideia de *uma* representação *única e verdadeira* do **sentido**

A análise linguística **não é** necessariamente **linear**

A metalinguagem linguística é um **produto humano**, situado historicamente e motivado por problema específico (que nunca foi o processamento automático de uma língua, aliás).

*“(...) Quando aprendi sobre essas coisas [diferentes abordagens para explorar pensamento e linguagem, como retórica, e teoria literária], percebi o quão limitada era minha visão tecnológica do pensamento e da linguagem. Eu aprendi como a visão quantificável, individualista, a-histórica – isto é, computacional – que eu tinha de cognição não conseguia explicar a vastidão da experiência cognitiva (incluindo, por exemplo, muito de Shakespeare). Eu aprendi como perspectivas pragmáticas e contextualistas refletem melhor a diversidade e flexibilidade de nossas práticas linguísticas do que os modelos de linguagem formais. (...)”*

*Mais impressionante: eu aprendi que havia precedentes históricos para exatamente o tipo de simplificações lógicas que caracterizaram o meu trabalho em IA. Na verdade, havia inclusive precedentes para a minha motivação em embarcar, em um primeiro momento, neste tipo de perspectiva. Encontrei tais precedentes em episódios que vão desde os tempos antigos – com o fascínio de Platão pelas formas matemáticas como fonte de verdades atemporais – até o século 20 – o Positivismo Lógico e sua busca pela criação de uma linguagem não ambígua para expressar as bases seguras para todo o conhecimento. Eles também tinham uma noção acrítica do progresso; e eles, também, lutaram em suas tentativas de quantificar formalmente conceitos humanos que eu agora vejo como inextricavelmente ligados às preocupações e práticas humanas.”*

Damon Horowitz (2011) - “Why you should quit your technology job and get a Ph.D. in the humanities”  
<http://chronicle.com/article/From-Technologist-to/128231/>

# Desafios na formação do ponto de vista de Letras...

---

Áreas com metalinguagens MUITO diferentes

Áreas com conhecimentos específicos com valorização social muito diferente

-----

A legitimidade de certas atividades enquanto “atividades de pesquisa”, sobretudo nas humanidades, é uma questão: o volume colossal de trabalho necessário à preparação do material -- limpeza de ruídos, inclusão de metadados, conversões e adaptações de formato -- não costuma ser considerado atividade de pesquisa, ainda que só possa ser feito por especialistas (WILLIFORD & HENRY (2012))

-----

FISH (2012): Promoção de uma “interdisciplinaridade profunda” ou transdisciplinaridade, que **não se assenta nas tradicionais fronteiras de disciplinas.**

Política da interdisciplinaridade: áreas diferentes, avaliações diferentes?

# Desafios na formação...



André Dahmer



# Referências

---

Bianca Freitas. “O DIZER em português: diálogos entre tradução, descrição e linguística computacional”. PPGEL / Mestrado. PUC-Rio. 2016.

Cláudia Freitas, Bianca Freitas & Diana Santos. "QUEMDISSE?: Reported speech in Portuguese". In: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)

Stanley FISH. “The Digital Humanities and the Transcending of Mortality”. The Opinion Pages/New York Times, 9 de janeiro de 2012.

SILVA, Rafael dos Reis. “Direct and Indirect Quotation Extraction for Portuguese”. Dissertação de Mestrado. Departamento de Informática. PUC-Rio. 2017.

WILLIFORD, C. & HENRY, C. “One Culture - Computationally Intensive Research in the Humanities and Social Sciences - A Report on the Experiences of First Respondents to the Digging into Data Challenge”. Council on Library and Information Resources Washington, D.C. 2012.

WITTGENSTEIN, Ludwig. “O Livro Azul” (1958). Lisboa:Edições 70 (1992)