ufjf | PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

PROVA DE CONHECIMENTOS EM LINGUÍSTICA – QUESTÃO ESPECÍFICA
LINGUÍSTICA COGNITIVA E COMPUTACIONAL
MESTRADO EM LINGUÍSTICA 2027.1

CANDIDATA/O No: _____

ATENÇÃO!

1. As respostas deverão ser escritas à tinta, na folha de almaço fornecida juntamente a esta prova.
2. Não é permitido qualquer tipo de consulta para a realização desta prova.
3. É vedada a identificação da/o candidata/o em qualquer das folhas da prova.
4. O tempo de duração total da prova (considerando as duas questões) é de 4 (quatro) horas.

ÁREA TEMÁTICA LINGUÍSTICA COGNITIVA E COMPUTACIONAL

No artigo "Multimodalidade: abordagens cognitivas e representações computacionais", Torrent e Coneglian (2025) contextualizam a atuação da rede ReINVenTA no âmbito da Linguística Cognitva-Computacional. Para tanto destacam que a discussão sobre a construção de datasets padrão ouro (*gold standard*) é um ponto fundamental para a implementação computacional da teoria linguística:

> A grande hipótese que se busca desenvolver é que teorias linguísticas, particularmente a Linguística Cognitiva, representada pelo modelo da Semântica de Frames (Fillmore, 1982) e pela metodologia da FrameNet (Fillmore; Baker, 2009), desempenham um papel fundamental na construção de datasets gold standard e no treinamento de modelos de Inteligência Artificial. Por aí se pode ver muito explicitamente o diálogo entre linguística e computação. (Torrent; Coneglian, 2025, p.16).

Lilian Ferrari (2011) apresenta a Semântica de Frames na obra *Introdução à Linguística Cognitiva*. Ela define o conceito de frame como “um sistema estruturado de conhecimento, armazenado na memória de longo prazo e organizado a partir da esquematização da experiência" (Ferrari, 2011, p. 50). A autora destaca a importância do ancoramento cultural no processo de atribuição de significado às palavras, o que acontece, segundo a Semântica de Frames, pelo padrão de evocação de frames estabelecido por uma unidade lexical em uso na língua.

Tomando essas considerações como ponto de partida, responda de forma objetiva:

a) Como a aplicação da Semântica de Frames pode beneficiar a construção de datasets padrão ouro? (20 pontos)
b) Como o uso de Inteligência Artificial pode auxiliar esse processo? (15 pontos)
c) Qual o papel dos linguistas humanos na construção destes datasets? (15 pontos)

Referência de correção:

a. A semântica de frames, ao prover o ancoramento cultural no processo de atribuição de significado às palavras em contexto real de uso, pode oferecer as associações necessárias para que um dataset carregue as informações específicas de cena de fundo providas pelos frames, aproximando-o ao máximo da percepção humana, o que levaria ao padrão ouro.

b. O uso de inteligência artificial se aplica na ampliação da capacidade de processamento de dados e identificação de padrões, o que impacta diretamente no volume de dados possível de ser processado, o que por sua vez amplia a base de dados disponível para treinamentos futuros, avaliação, etc.

c. Os linguistas humanos têm o papel fundamental de validar qualquer auxílio prestado por IA, por exemplo corrigindo as limitações culturais, mas, frequentemente também, provendo os inputs iniciais a partir da singular experiência humana.