

# Extração e Validação de Ontologias a partir de Recursos Digitais

**Objetivo:** Apresentar uma metodologia para extração e validação semiautomática de ontologia.

**Estudo de caso:** Ontologia derivada do PAPEL (Palavras Associadas Porto Editora Linguatca, [Oliveira et al. 2008]).

**Materiais:**

PAPEL: conjunto de relações entre palavras em Português

Nº de entradas	199.672
Nº de relações	8 (hiperonímia, sinonímia ...)
Nº de relações de hiperonímia	61.263
Nº de relações de sinonímia	36.504

**Foco:** Extração de relações de hiperonímia e de sinonímia.

**Etapas de Hiperônimos:** Conversão automática em duas etapas: (1) criação da ontologia com todas as classes sem relações; (2) inserção gradativa das relações.

**Etapas de Sinônimos:** Devido à polissemia das palavras e à transitividade da relação de sinonímia, a conversão é semiautomática: (1) É criada uma lista (inicialmente vazia) de conjuntos (inicialmente vazios); (2) Para cada relação de sinonímia, se as palavras já se encontram em algum conjunto, o linguista edita manualmente (com base em dicionários e contextos de ocorrência) os conjuntos onde a palavra deve ser inserida, se não, é criado um novo conjunto com ambas as palavras.

**Resultado:** A ontologia resultante contém 46.904 entradas, com 40.614 relações de hiperonímia e 20.096 de sinonímia.

**Conclusão:** Framework para extração e validação de ontologias possibilitou geração de recurso que pode ser utilizado em sistemas de Processamento de Linguagem Natural, tais como tradução automática, sistemas de busca e extração de informação, sistemas conversacionais, sistemas de inferência e extração automática de ontologias.

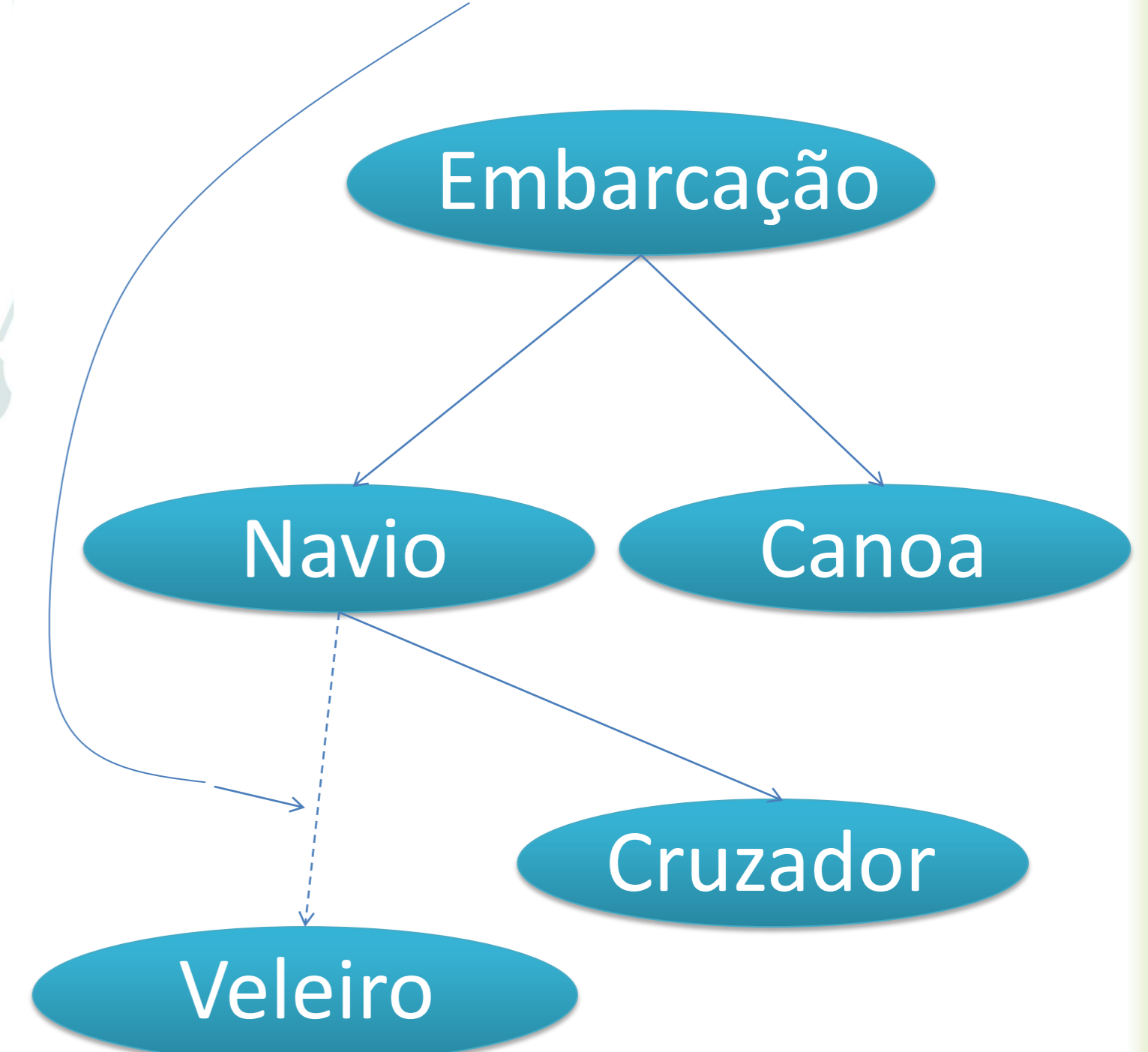
**Referências:**

Oliveira, H. G., Santos, D., and Gomes, P. (2009). Extração de relações semânticas entre palavras a partir de um dicionário: o papel e sua avaliação. STIL 2009, Linguamática, páginas 77–93.

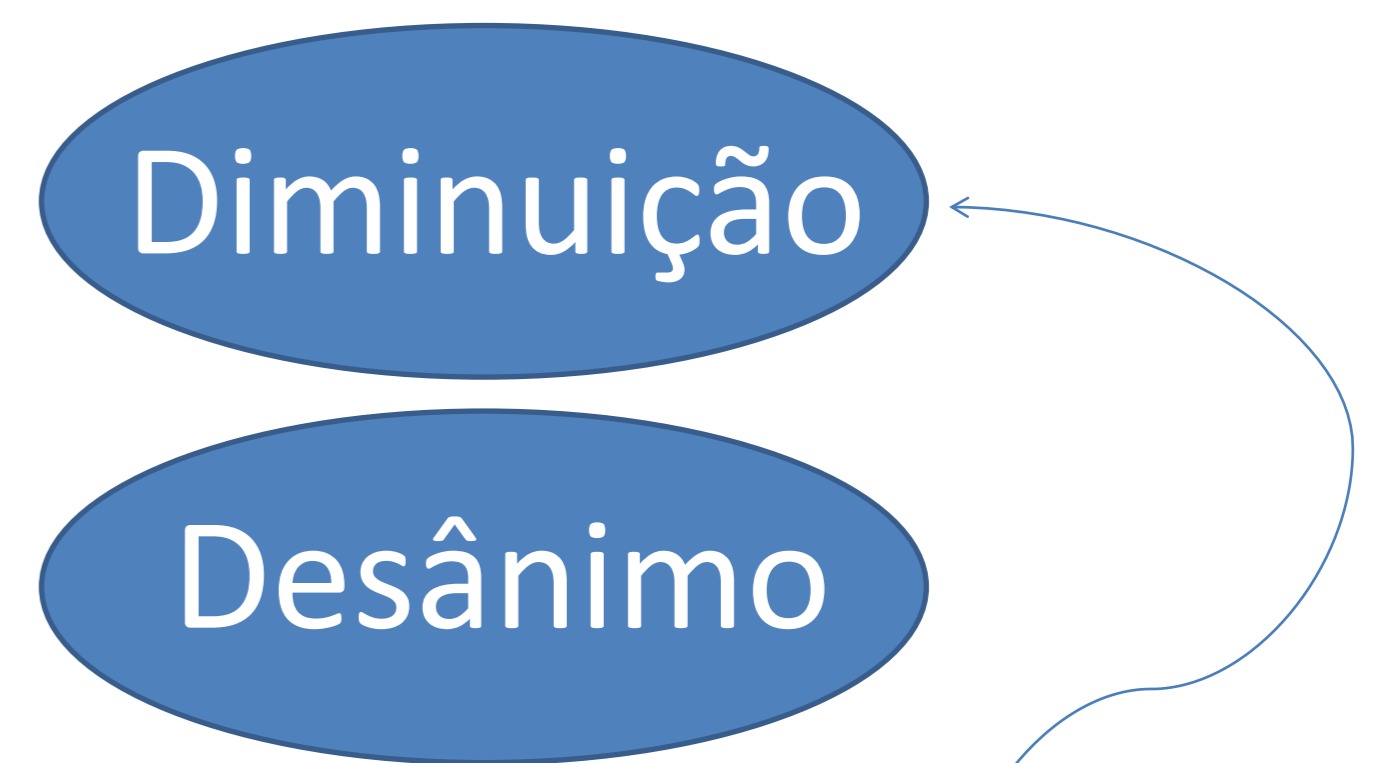
## ONTOBRAS-MOST 2011

**Etapas de hiperônimos:** Inclusão da relação entre navio e veleiro

navio HIPERONIMO DE veleiro



**Etapas de sinônimos:** Decisão de incluir sinônimo pelo número de ocorrências: de abatimento de preços(diminuição) ocorre mais vezes do que abatimento emocional(desânimo)



**Abatimento**

- de preços(maior ocorrência)
- emocional
- de animais

Agradecimentos: Esta pesquisa tem apoio dos projetos COMUNICA (FINEP/SEBRAE 1194/07), CAPES-COFECUB (707/11) e CNPq (479824/2009-6 e 309569/2009-5).