

# A avaliação no DIP e seus problemas

Diana Santos

d.s.m.santos@ilos.uio.no



21 de novembro de 2022



## A tarefa e como avaliá-la

- Identificar personagens com nome próprio, quais os nomes próprios que se referem à mesma personagem, qual o género e a profissão dessa personagem, e quais as relações familiares entre essa personagem e as outras.
- Idealmente, essa avaliação seria feita por personagem. Quantas informações sobre essa personagem conseguiram os sistemas recolher?
- Na prática, não conseguimos obter um algoritmo que permitisse alinhar sempre a resposta do sistema com a coleção dourada, e decidimos dividir para reinar.
- Ou seja, cada uma das subtarefas é avaliada independentemente.

- Avaliação de identificação (AI)
- Avaliação da co-identificação (ACI) – ou co-referência
- Avaliação do género (AG)
- Avaliação da profissão/ocupação/estatuto social (APOES)
- Avaliação das relações familiares (AR)

De notar que estas medidas não estão relacionadas entre si, e sobretudo não estão relacionadas ao número de personagens de uma dada obra.

## Avaliação de identificação (AI)

Dadas as duas listas de identificadores em relação a uma dada obra, a da coleção dourada e a do sistema participante, usar precisão e abrangência para classificar o sistema.

**Precisão:** o número de elementos que estão nas duas listas/o número de elementos na lista do sistema

**Abrangência:** o número de elementos que estão nas duas listas/o número de elementos na lista dourada

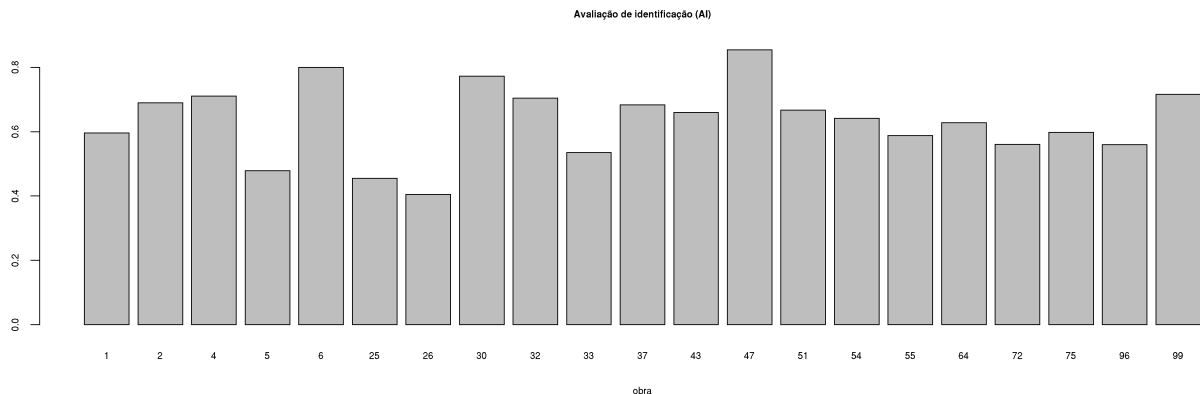
Problemas: Quando há personagens com o mesmo nome... Para isso adicionamos o identificador num passo preliminar

203,9, João da Esquina|Sr. João da Esquina|Sr. João|João\_9,M,boticá

203,15, João\_15,M,

É necessário fazer isso 19 vezes (004) + 2 vezes (006) + 2 (026) + 8 (030)+ 4 (037) + 2 (043) + 2 (054) + 2 (064) + 4 (072) + 6 (099) + 4 (201) + 2 (203) : 57 vezes em 12 obras diferentes!

# Resultados do PALAVRAS na avaliação de identificação



Melhor: 47 - *A escrava Isaura*, de Bernardo Guimarães: 0,854

Pior: 26 - *Simá*, de Lourenço Amazonas: 0,405

## Avaliação da co-identificação (ACI) – ou co-referência

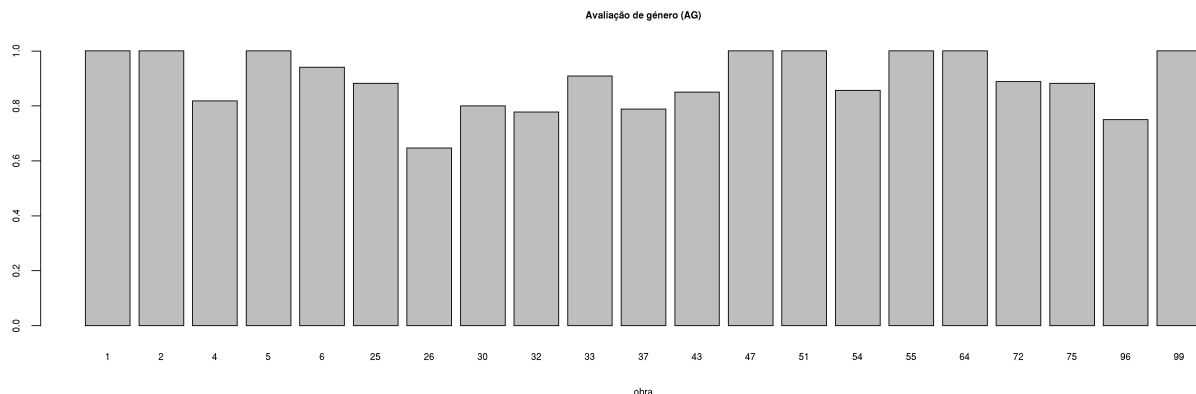
Podemos separar a noção de co-identificação em casos binários, identificando os casos em que só há um nome para uma personagem pela relação X – ZERO, e usar outra vez precisão e abrangência para essas relações binárias.

É preciso portanto expandir todas as relações, depois de retirar da coleção dourada os nomes não identificados (e cuja falta já foi penalizada pela medida anterior).

3,Margarida Guida,F 4,Guidinha dos Meadas,M	1,Margarida Guida Guidinha dos Meadas D. Guida,F  Margarida Guida Guidinha dos Meadas
Margarida—Guida Guidinha dos Meadas—ZERO	Margarida—Guida Margarida—Guidinha dos Meadas Guida—Guidinha dos Meadas
Precisão: 1/2 Abrangência: 1/3	



# Resultados do PALAVRAS na avaliação de género



Piores: 26 - *Simá*, de Lourenço Amazonas: 0,647

96 - *Amar, verbo intransitivo*, de Mário de Andrade: 0,750

## Avaliação da profissão/ocupação/estatuto social (APOES)

Esta medida apenas se refere às personagens a que o sistema ou a CD atribuiu uma profissão. Não pontuamos nem prejudicamos os casos sem profissão. Se o nome da personagem não existir na CD, não conta para a avaliação. (Já foi prejudicado na medida da identificação). Idem para se os nomes não corresponderem à mesma personagem (tratado pela ACI). Usamos também precisão e abrangência, sobre pares nome – profissão.

139,1, Sr. Bentinho, M, escravo|advogado  
139,2, Capitu, M,  
139,3, Dom Casmurro, M, advogado  
139,4, Sancha|D. Sancha|Sanchinha, F,  
139,5, Padre Cabral, M, padre|agregado  
139,6, Dona Fortunata|Tartaruga, F, funcionário de repartição pública  
139,7, Tijuca, F, lavradora  
139,8, Prima Glória|Mana Glória, F,  
139,9, Thomaz, M,

139,3, D. Glória|Dona Glória|D. Maria da Glória Fernandes Santiago|Prima Glória|Mana Glória, F,  
139,5, Sancha|Sanchinha|D. Sancha|Sinhazinha Sancha, F,  
139,7, Padre Cabral|Cabral, M, padre  
139,8, Pádua|João|Sr. Pádua|Joãozinho|Tartaruga, M, funcionário de repartição pública  
139,9, Dona Fortunata, F,  
139,10, Bento|Padre Bentinho|Sr. Bentinho|Doutor Santiago|Dom Casmurro, M, advogado  
139,13, Capitu|Capitolina, F,  
139,18, Thomaz, M, escravo

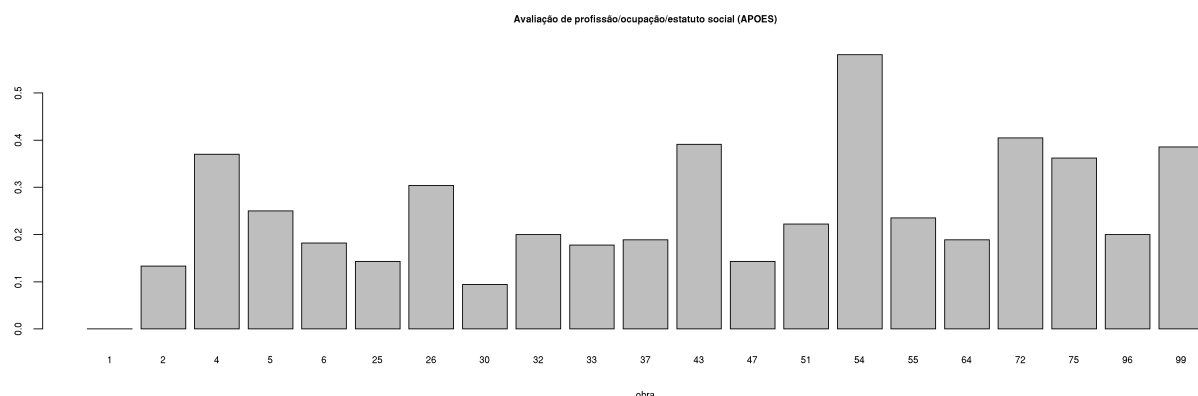
- Sr. Bentinho, escravo
- + Sr. Bentinho, advogado
- + Dom Casmurro, advogado
- + Padre Cabral, padre
- Padre Cabral, agregado
- Dona Fortunata, funcionário de repartição pública
- + Tartaruga, funcionário de repartição pública
- 0 Tijuca, lavradora
- Thomaz,

- Sr. Bentinho, advogado
- Dom Casmurro, advogado
- Padre Cabral, padre
- Tartaruga, funcionário de repartição pública
- Thomaz, escravo

Precisão: 4/8

Abrangência: 4/5

# Resultados do PALAVRAS na avaliação de profissão, ocupação e estatuto social



Melhor: 54 - *O Barão de Lavos*, de Abel Botelho: 0,581

Pior: 1 - *Miss Kate*, de Cosme Velho: 0

## Avaliação das relações familiares (AR)

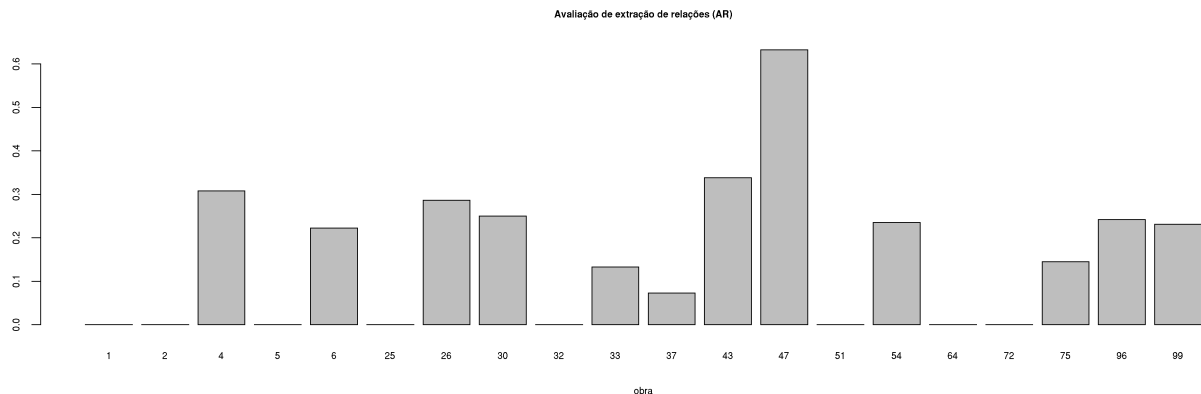
É a parte mais complicada de toda a avaliação no DIP, porque é preciso “alinhar” os identificadores, e esse alinhamento não é 100% fiável.

Além disso, é preciso expandir as relações tanto na CD como no sistema, para poder pontuar corretamente.

Como nos casos anteriores, apenas se leva em conta as “personagens” que existem na CD e na resposta do sistema.

Depois disso, é apenas precisão e abrangência sobre triplos (que têm, além disso, um vocabulário fixo).

# Resultados do PALAVRAS na avaliação das relações



Melhor: 47 - *A escrava Isaura*, de Bernardo Guimarães: 0,632

Piores: 8 obras sem nenhuma identificação certa...

## Avaliação final

Por agora e por falta de alternativa, temos feito uma média das cinco medidas.

Mas não é interpretável, e funcionaria melhor como métrica comparativa.

Em 21 obras,

AI = 0.634,

ACI = 0.678,

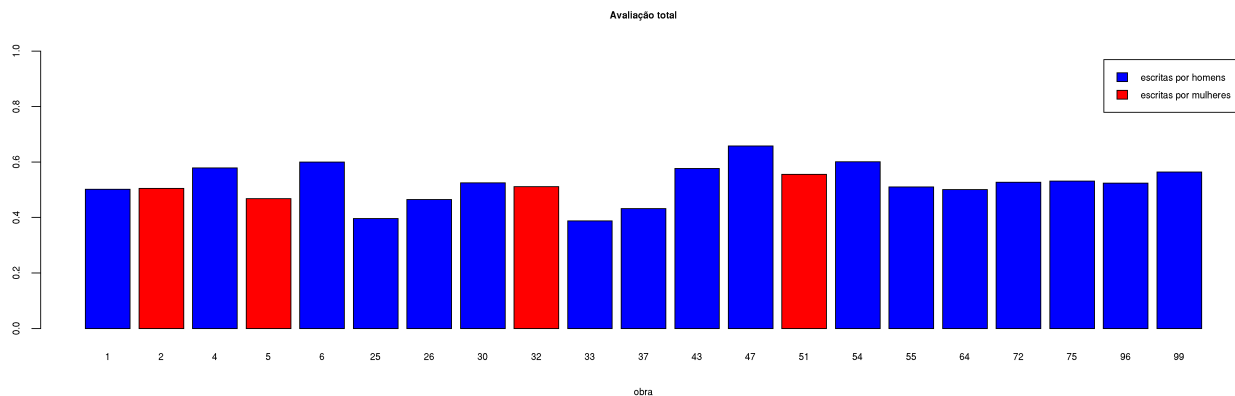
AG = 0.895,

APOES = 0.246,

AR = 0.155

Total = 0.5154

# Resultados do PALAVRAS numa avaliação total



Melhor: 47 - *A escrava Isaura*, de Bernardo Guimarães: 0,6578

Pior: 33 - *A ermida de Castromino*, de Teixeira de Vasconcellos: 0,3874