

Tipificação Manual do Alinhamento do Corpus CSTNews

Renata Tironi de Camargo
Verônica Agostini

Grupo de Sumarização Automática





ALINHAMENTO

- Exemplo **com** sobreposição lexical

Sumário: Antes de chegar à Jamaica, Dean matou ao menos nove pessoas nas ilhas de Santa Lúcia, Dominica, República Dominicana e Haiti, no Caribe.

Documento: O furacão matou ao menos nove pessoas em sua passagem pelas ilhas do Caribe.

ALINHAMENTO

- Exemplo **sem** sobreposição lexical

Sumário: Vários moradores e turistas nas regiões, inclusive brasileiros, foram retirados dos locais, enquanto outros estão se preparando para a passagem do furacão.

Documento: Na Jamaica, muitos estocaram alimentos, água, lanternas e velas.



TIPIFICAÇÃO DOS ALINHAMENTOS

Definição da Tarefa



Análise dos processos de refacção textual entre os textos-fonte e os sumários manuais do *corpus* CSTNews

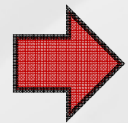


Atribuição de **tipos** a cada par de sentenças alinhadas



TIPIFICAÇÃO DOS ALINHAMENTOS

Motivação



Verificar quais alinhamentos um alinhador automático acerta mais e quais erra mais



Descobrir a natureza das transformações que um humano pode realizar quando faz a sumarização multidocumento

TIPOLOGIA

Quanto à forma

Igual

Parcial

Diferente

Quanto ao conteúdo

Contradição

Neutro

Generalização

Especificação

Inferência

Outro

QUANTO À FORMA

Tipo	Definição
Igual	As duas sentenças são idênticas, incluindo pontuação
Parcial	As duas sentenças possuem algumas palavras em comum
Diferente	As duas sentenças possuem poucas palavras em comum ou nenhuma



QUANTO AO CONTEÚDO

Tipo	Definição
Contradição	Quando há informação contraditória
Inferência	Quando uma informação foi inferida
Generalização	Quando a informação do sumário é mais genérica que a do texto-fonte
Especificação	Quando a informação do sumário é mais específica que a do texto-fonte
Neutro	Quando não há modificação
Outro	Quando não se concorda com o alinhamento

ALÉM DISSO...

Onomástica

Toponímia

Antroponímia

Estudo de nomes de
lugares


Estudo dos nomes
próprios das pessoas



EXEMPLO: parcial/neutro, especificação

Sumário: O Brasil conseguiu um gol logo nos **primeiros 4 minutos do jogo**, fazendo os argentinos apertarem o ataque no jogo, restando ao Brasil os contragolpes, chegando ao segundo gol, que foi um gol contra.

Documento: É verdade que o Brasil deu sorte de conseguir um gol logo no **início da partida**.





EXEMPLO: diferente/neutro

Sumário: A seleção brasileira, sob direção de Dunga, conquistou o oitavo título da Copa América, goleando a Argentina por 3 a 0.

Documento: O Brasil é a melhor seleção das américas pela oitava vez.

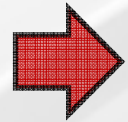


EXEMPLO: diferente/inferência - antroponímia

Sumário: O Brasil, mesmo sem duas de suas estrelas, bateu a Argentina, que tinha a melhor campanha do campeonato e **contava com seus principais jogadores.**

Documento: Após acompanhar o belo futebol apresentado pelos “hermanos” durante toda a Copa América, que foi **conduzida pelos habilidosos pés de Riquelme e Messi**, o Brasil foi a campo sem o tradicional status de favorito que o acompanha há muito.

DÚVIDAS



Regra:

- Alinhamentos do tipo forma
 - Sentença inteira
- Alinhamentos do tipo conteúdo
 - N-gramas

DÚVIDAS - EXEMPLO

Sumário: Dentre esses está Juan Carlos Ramirez Abadia, um dos traficantes mais procurados do mundo.

Documento: Entre elas, o maior traficante da Colômbia, Juan Carlos Ramirez-Abadia, 44 anos, um dos mais procurados pelo Departamento Antidrogas dos Estados Unidos (DEA).

Generalização? Extra?

DÚVIDAS - EXEMPLO

Sumário: Dentre esses está **Juan Carlos Ramirez Abadia**, um dos traficantes mais procurados do mundo.

Documento: Entre elas, o maior traficante da Colômbia, **Juan Carlos Ramirez-Abadia**, 44 anos, um dos mais procurados pelo Departamento Antidrogas dos Estados Unidos (DEA).

44 anos → Generalização? Extra?

Nenhum



EXEMPLO – DISCORDÂNCIAS

Sumário: O Brasil conseguiu um gol logo nos primeiros 4 minutos do jogo, fazendo os argentinos apertarem o ataque no jogo, restando ao Brasil os contragolpes, chegando ao segundo gol, que foi um gol contra.

Documento: Ainda aos quatro minutos, Júlio Baptista dominou dentro da área, driblou Ayala e chutou forte, marcando um golaço.

Parcial? Diferente?

Diferente



EXEMPLO – DISCORDÂNCIAS

Sumário: Amanhã será decidido, pela Mesa Diretora do Senado, se a quarta representação será encaminhada ao Conselho de Ética.

Documento: Na quinta-feira, a Mesa Diretora do Senado se reúne às 14 horas para decidir se aceita a quarta representação contra o presidente da Casa.

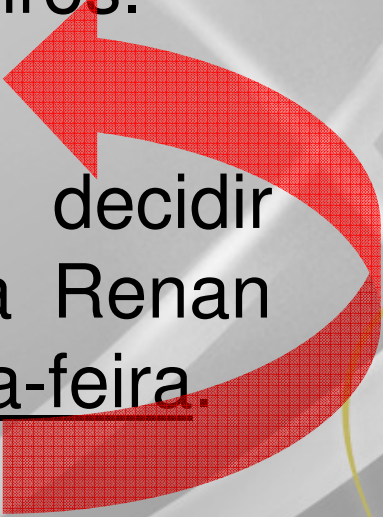
Generalização? Especificação? Neutro?
parcial/neutro



EXEMPLO – DISCORDÂNCIAS

Sumário: Na próxima semana, o Conselho de Ética vai decidir sobre a unificação dos processos contra o presidente da Casa, Renan Calheiros.

Documento: O Conselho de Ética vai decidir sobre a unificação dos processos contra Renan em reunião marcada para a próxima quarta-feira.



Generalização? Especificação? Neutro?
parcial/generalização



EXEMPLO - EXERCÍCIO

Sumário: Um porta-voz da ONG disse que a organização nunca sofreu uma perda tão grande em 25 anos.

Documento: O diretor da ACF no Sri Lanka, Benoit Miribel, confirmou a morte de seus funcionários e afirmou, comovido, que a ONG "não sofreu uma perda similar em seus mais de 25 anos de existência".

**parcial/neutro, inferência
generalização - antroponímia**

EXEMPLO - EXERCÍCIO

Sumário: Mais de 300 policiais federais de vários estados participaram das buscas e prisões durante a operação.

Documento: Ao menos 300 policiais de Amapá, Distrito Federal, Mato Grosso, Acre e Rondônia trabalharam na Operação Dominó.

**parcial/neutro, generalização -
toponímia**



EXEMPLO - EXERCÍCIO

Sumário: A pressão argentina continuou no segundo tempo, mas o Brasil fechou a goleada com um gol marcado pela sua dupla de volantes.

Documento: O atacante, que na Copa América marcou apenas uma vez, deu um passe preciso para Daniel Alves chutar cruzado de dentro da área e vazar mais uma vez o goleiro Abbondanzieri.

**diferente/neutro, contradição,
generalização - antroponímia**



EXEMPLO - EXERCÍCIO

Sumário: A expectativa de lideranças da Câmara e do Conselho de Ética é que pouco mais de 10% dos 69 deputados denunciados no relatório parcial da CPI abrirão mão de seus mandatos.

Documento: A Folha Online apurou que o Congresso espera por um pedido em massa de renúncia até o final do prazo.

parcial/neutro, contradição



CONSIDERAÇÕES SOBRE A TAREFA

- ➔ Tarefa mais complexa do que imaginávamos
- ➔ Necessidade de revisão da anotação
- ➔ Aparentemente, a maioria dos alinhamentos são do tipo “parcial/neutro”
- ➔ Foram encontrados alguns erros de digitação na tarefa de alinhamento, os quais foram corrigidos

Tipificação Manual do Alinhamento do Corpus CSTNews

Renata Tironi de Camargo
Verônica Agostini

Grupo de Sumarização Automática

