

Diretrizes de Avaliação de Perguntas e Apresentação do Sistema.

Neste documento apresentamos a interface do *QuestionEvaluator* e as classes de erros que as perguntas geradas podem pertencer. Nosso objetivo é avaliar perguntas geradas automaticamente através de recursos computacionais que servirão de base para a criação de sistemas de perguntas e respostas. O trabalho de avaliação é definir se a formulação da pergunta está correta ou incorreta. Isso deve ser feito analisando o conteúdo da pergunta, o texto base e a resposta. As perguntas são sempre respondidas por trechos do texto base. Chamamos esses trechos de *spans*.

Taxonomia de Erros

A categorização de erros baseia-se em Laban et al 2022¹ que avaliou perguntas geradas no contexto educacional e na língua inglesa. Neste trabalho, adaptamos essa categorização para a língua portuguesa. Os erros de perguntas podem pertencer à três categorias macro: fluência, alvo e contexto. Caso uma pergunta pertença a mais de uma classe de erros, usar a precedência: Fluência > Alvo > Contexto.

Categoria Macro	Categoria Micro	Definição
Fluência	Conjugação errada	Erros relacionados à estrutura e ao português da pergunta.
	Fraseamento Estranho	
	Não questão	
	Repetições	
Alvo	Não respondível	Erros relacionados à conexão pergunta-resposta.
	Outro span	
Contexto	Muito específica	Erros relacionados à semântica e à racionalidade da pergunta.
	Revela resposta	
	Inconsistente	
	Não específica o bastante	

¹ Philippe Laban, Chien-Sheng Wu, Lidiya Murakhov's'ka, Wenhao Liu, and Caiming Xiong. 2022. [Quiz Design Task: Helping Teachers Create Quizzes with Automated Question Generation](#). In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 102–111, Seattle, United States. Association for Computational Linguistics.

Fluência

Conjugação Errada

Exemplo: O que *foram* historicamente usado para privar de direitos as minorias raciais.

- Deveria ser "O que foi.."

Fraseamento Estranho

Exemplo: Quando o mamute-lanoso *morreu*?

- A forma correta é: *foi extinto*.

Não questão

Exemplo: Em janeiro de 2020, cientistas reportaram que a modelagem climática dos eventos de extinção favorecem o impacto por asteroide e não vulcanismo?

Exemplo: O SBT foi criado a partir de concessões cassadas da extinta Rede Tupi?

- Trata-se de uma sentença no formato declarativo com um ponto de interrogação.

Repetições

Exemplo: Quem *liderou* o time que *liderou* o nível K-Pg?

- Liderou é repetido duas vezes.

Exemplo: Como são chamados os termos de vacinação e "vacinação"?

Alvo

Não Respondível

Exemplo: Por que os estudos de DNA são tão importantes?

- A resposta não está presente no texto base.

Outro span

Texto base: ... A estátua de cobre, projetada pelo escultor francês Frédéric Auguste Bartholdi, que se baseou no Colosso de Rodes para edificá-la, foi construída por Gustave Eiffel e inaugurada em 28 de outubro de 1886 ...

Exemplo: Quem projetou a Estátua da Liberdade?

- O trecho marcado (vermelho) está como Eiffel, porém a verdadeira resposta é Bartholdi.

Texto base: ... o Grupo Silvio Santos também é formado por **outras empresas** de grande sucesso como SBT Filmes (empresa cinematográfica), SBT Music (gravadora musical), SBT Vídeos (vídeos na internet), SBT Licensing (**licenciadora de marcas**), Liderança Capitalização (responsável pela emissão do título de capitalização conhecido como Tele Sena), ...

Exemplo: O SBT licensing é responsável por quê?

- O trecho marcado (vermelho) está como outras empresas, porém a verdadeira resposta é licenciadora de marcas.

Contexto

Muito Específica

Exemplo: Quando a estátua de cobre, projetada pelo escultor francês Frédéric Auguste Bartholdi, que se baseou no Colosso de Rodes para edificá-la e construída por Gustave Eiffel, foi inaugurada?

- Muito específica e pouco natural. Melhor seria: Quando a Estátua da Liberdade foi inaugurada?

Revela a resposta

Exemplo: Qual é o nome do Campo de Namorado?

- A resposta é Campo de Namorado e foi revelada na própria pergunta.

Texto base: ...Liderança Capitalização (responsável pela emissão do título de capitalização conhecido como Tele Sena), Jequití Cosméticos (empresa de cosméticos), Hotel Jequitimar (**hotelaria**), entre outras ...

Exemplo: Qual o nome do complexo Jequitimar?

- O trecho está marcado de maneira incorreta, porém a resposta é revelada mesmo assim.

Inconsistente

A pergunta parte de uma premissa errada.

Exemplo: Quais são os únicos *dois animais* que sobreviveram à extinção do Cretáceo-Paleogeno?

- O texto base menciona espécies e não animais.

Exemplo: Qual é a profissão do SBT?

- SBT não tem profissão

Não específica o bastante

Texto base: ... Os termos " vacina " e " vacinação " são derivados de Variolae vaccinae (varíola da vaca), ...

Exemplo: De que é derivada a vacina?

- Faltou mencionar que estamos falando do termo vacina.

Exemplo: Quais são algumas espécies ectotérmicas?

- Muitas espécies são mencionadas no artigo.

Sistema QuestionEvaluator

QuestionEvaluator

geoRegis admin (sair)

Tarefa 1	0/4	contornitos e as correntes de Fundo Contornitos, por definição, s...
Tarefa 2	0/4	A Bacia do Recôncavo representa o marco da exploração de petróle...
Tarefa 3	0/4	Posamentier e James ressaltam que os conceitos da Estratigrafia ...
Tarefa 4	0/3	A maioria dos modelos de preenchimento de riftes descreve, nas fa...
Tarefa 5	0/4	Seguindo a linha de Posamentier e Allen , após o reconhecimento d...
Tarefa 6	0/1	AMPLITUDE MODERADA A ALTA MODERADA A ALTA BAIXA A MODERADA MODERA...
Tarefa 7	0/3	soerguimento e erosão, corroborando a propagação lateral do rifte...

A Formação Karapotó é constituída por arenitos mineralogicamente maduros, quartzo-sos, brancos a creme-claros, também róseos, de granulação fina a muito grossa, angulosos a su-barredondados, contendo grânulos ou níveis con-glomeráticos, gradando a conglomerados . Os conglomerados são formados por seixos bem selecionados e de tamanho relativamente uniforme, constituídos por **quartzo leitoso, com esfericidade alta a média, e, eventualmente, fragmentos líticos oriundos do embasamento proximal** . Dispõem-se em corpos de geometria tabular a sigmoidal, exibindo estratificação tabular e acanalada. Em subsuperfície, os arenitos são róseos a cinza-claros, também avermelhados, quartzosos, com fragmentos líticos, finos a grossos, subangu-lares, friáveis, ocorrendo intercalados a folhelhos vermelho-acastanhados a cinza-esverdeados, localmente negros, micáceos, piritosos, em parte betuminosos. É comum a presença de gradação normal e inversa e feições de erosão e preenchimento .

Tarefa 29/61
0% concluída

PRÓXIMO

Os fragmentos líticos são formados por quartzo leitoso ou quartzo leitoso?

☐ Correta

☒ Não avaliada

Fluência

☐ Conjugação Errada

☐ Fraseamento Estranho

☐ Não é pergunta

☐ Repetições

Alvo

☐ Não respondível

☐ Outro span

Contexto

☐ Muito específica

☐ Revela resposta

☐ Inconsistente

☐ Não específica o bastante