Reviewer Instructions and Checklist

Your job as a reviewer is to check the tasks completed by other attemptors in order to detect errors that make the task low quality. It is important to assign an appropriate score and feedback to help the person improve.

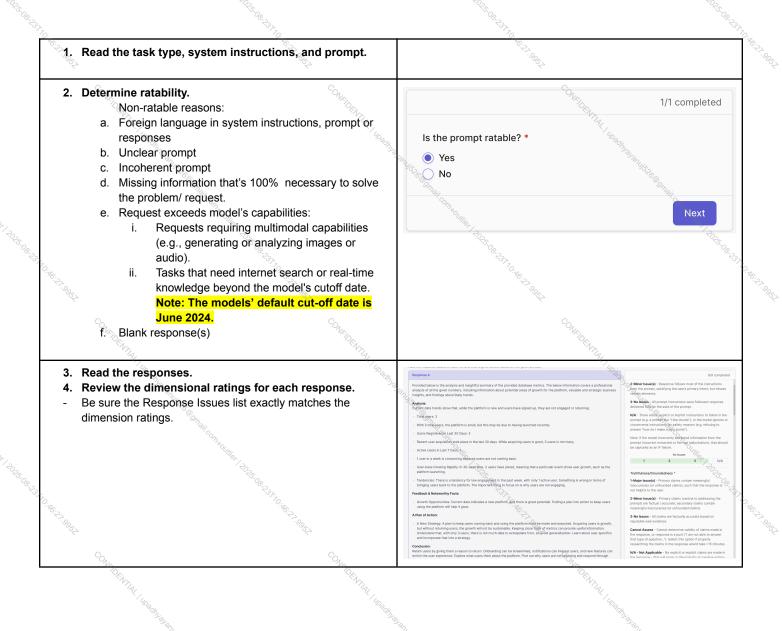
It is crucial **not** to focus on either approving as many tasks as possible or disapproving as many tasks as possible. **Your job is to stick to the rubric provided in this document as closely as possible.** To do this, it is essential to first understand the **attempter instructions** in detail.

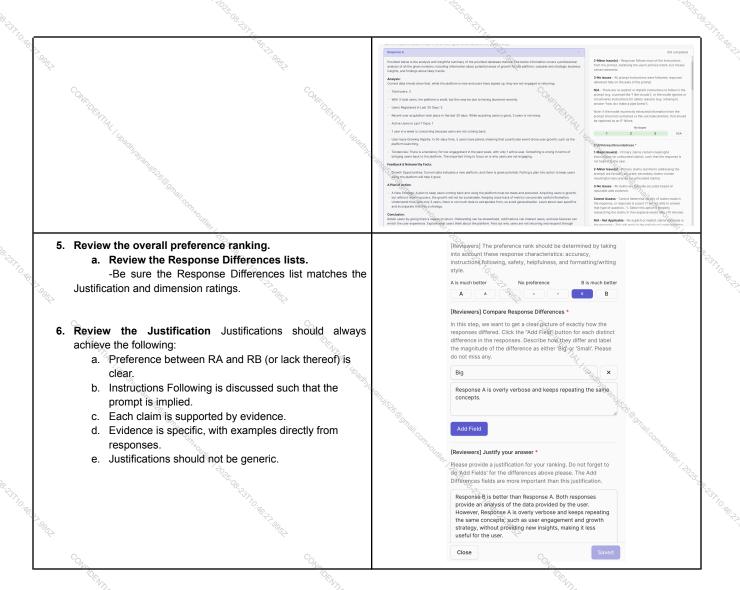
Attempter Instructions

♣ For this project, we have some specific steps to focus on:

- 1. The prompt must be correctly categorized as **ratable** or **not ratable**.
- 2. All dimensions must be assigned the correct score. Pay particular attention to the 3 strict dimensions: Instruction Following, Content Conciseness & Relevance, and Truthfulness.
- The issues list must be correctly labeled and reflect the problems from the dimensions above.
- 4. The **Likert** rating and its **justification** must match and align with the scores previously assigned.

Review Workflow





Score	Description O	Action	Criteria (Criteria)	14020th
4-5	The task is perfect! It was well-written, creative, appropriately	Approve	The prompt is correctly classified as ratable or non-to-	ratable according to the instructions.

_		\o'\	7 ₂
	complex, and completely followed the instruction guide.		If the prompt is non-ratable , the user correctly selected the reason why the prompt is non-ratable.
	C _{O4}		Dimension Ratings:
	WEIDENTIAL.		You agree with all the scores assigned by the attempter for each of the dimensions. OR
	Logith alanging Commencer of the Commenc		You agree with all the scores assigned by the attempter for the strict dimensions (Instruction Following, Truthfulness, and Content Concisness and Relevance) and have minor disagreements (Minor vs Major or No Issues vs Minor Issues) for the lenient dimensions (Style, Completeness, and Harmlessness).
	97,	VI. Co.	Overall Scores
		Th _X OLH:	You agree with the overall score assigned to both responses,
		**************************************	You disagree with the overall score for one of the responses by only 1 point (a subtle difference, e.g., 6 vs. 7, 1 vs. 2, 3 vs. 2, etc.).
		7)	Preference Rankings: The preference ranking(s) are accurate. The Likert scale score(s) match the justification.
			The Likert scale score(s) match the justification. The justification(s) are concise, insightful, and detailed.
	No.		Prompt
	The task is good but it needs some minor fixes		The prompt was correctly marked as non-ratable, but the category explaining why it is non-ratable is incorrect.
	or slight adjustments to		Dimension Ratings:
	some of the ratings. If you rate a task a 3 you		Difficultion (Manage).
	must fix the identified	\$ //	 You agree with all the scores assigned for the strict dimensions, except for one minor disagreement in one strict dimension (e.g., you think it is Minor Issue vs. Major Issue or No
√ 3	issues before submitting it.	Fix + Approve	Issues vs. Minor Issue, and vice versa).
		**************************************	Note: If more than two strict dimensions have minor disagreements, the score must be 2, not 3.
	Note: only one of the listed criteria is	2,08,5	Preference Rankings:
	enough for the task to be rated as 3.	`3>	You disagree with the Likert rating by a maximum difference of 2 points (e.g., 1 vs. 3).
	ne rateu as 3.		The Likert scale score(s) match the justification.
57			 The justification(s) are verbose, not very insightful, and/or lack details, but the justification matches the Likert score.
	Th. 162 - 1. : -		\$\langle \cdot \c
	The task is spam, it does not follow any of	Fix but solost s	Prompt(s): • The prompt was marked as ratable when it is actually non-ratable, and vice versa.
	the instructions, it needs to be completely redone,	Fix but select a 1 or a 2 and	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1-2	and/or it is not helpful	provide detailed	Dimension Ratings:
1-2	data at all for the customer.	feedback to	You disagree due to a major disagreement in at least one strict dimension — for example, the
		the attempter	user marked No Issues when it should have been Major/Significant or Critical Issues.
	Note: only one of the	W.can	181 ₁₈
		"XOUTIE	Trough
		1/2024	*/ _{***********************************}
<u> </u>		,06,53°	, which is the second of the s
70. 46			70. (70
.66.			The state of the s
7			7
	So		9

	*/ *d25.00	*/************************************	**/***********************************
listed criteria is enough for the task to be rated	AND/OR		
as 2 or 1.	The prompt require	ed a word/character constraint and the response did not ed No Issues instead of Critical Issues for Content Con	
TOENTRAL 400		user selected <i>Critical Issues</i> when <i>No Issues</i> was correct	
*Unaparillista	The user correctly	penalized a strict dimension with a <i>Major/Significant/C</i> the response is higher than 2. Or vice versa.	Critical Issue, but the
~	AND/ OR	The control of the co	Str _{iki COPr} ro,
da d	The user penalized violations table. (i)	d the wrong dimension with a major issue , contrary to the More info here)	he overlapping
To the state of th		nking(s) are completely inaccurate.	
% <u>></u>		core(s) do not match the justification. are vague and inaccurate/ 1 or more of the claims contr	radict the dimensional
William William	900 100 100 100 100 100 100 100 100 100	- William - Will	
\\(\(\text{U}_{\text{2}(f)_{\text{0}}}\)	1 quality	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	
THE STATE OF THE S	in _{ell}		Tigge Canali
	Conradillation of the contradition of the cont	COTACOLINE / FEE	COTRICULTION PROPERTY.
FOR	Sarge Sarro	**************************************	**************************************
1867 1867	1867 - 18	. 1877. 1877. 1877. 1877. 1879. 1879. 1879. 1879. 1879. 1879. 1879. 1879. 1879. 1879. 1879. 1879. 1879. 1879.	
CONFIDE	Canada	CONFIDE	
NTAI/UDA	`NI _{I4} ₁₄₂₂₀₁	**************************************	
19 April 19 19 19 19 19 19 19 19 19 19 19 19 19		Na _{RILI} STO	idh _{lific} e
<i>۵</i>	Projection of the control of the con	Ordi Corridge	Oneil County Office
	16. 1 202. 108. 2	Mer Right Ca.	Mer/ FOR DR.
R. C.	TREEL COPPROVINGE ARRA OR SAN TO SECOND SE	OENTRAL LUCARINAN. REPRESENTAL LUCARINAN. REPRESENTAL LUCARINAN.	Ongiconrollice design
COM.	Con.	CON.	

137,00gs

Appendix:

Overlapping Violations Table

Here are some common overlaps:

	How to rate when these dimensions overlap		Dimension 1			
			Style & Format/Structure	Groundedness	Relevance & Conciseness	Completeness & Detail
	Dimension 2	Instruction Following	Mark down Style & Format/Structure without marking down Instruction Following if the Response fails to follow a Style & Format/Structure instruction.	Mark down Groundedness without marking down Instruction Following if the Response fails to follow a Groundedness instruction.	Mark down Relevance & Conciseness without marking down Instruction Following if the Response is too verbose. IMPORTANT UPDATE REGARDING WORD	Mark down Completeness & Detail without marking down Instruction Following if the Response fails to meet instruction related to Completeness & Detail.
	ROENTAL LEGISTATURE	Official Companies of the Companies of t	in the state of th	#EXCEPTION!! For tasks that explicitly involve information retrieval or extraction from a prompt, a failure to extract information correctly would be penalized for Instruction Following only.	COUNTS/Character Counts: If a response exceeds a word count/character count limit explicitly specified in prompt, even by a single unit, you should mark this as a Critical Issue under Response Relevance & Conciseness.	CO _{NFIDENTRAL} LOGARDISTANIA
		Style & Format/Struct ure	N/A Zo.		Mark down both Style & Tone and Relevance & Conciseness if repetition is found.	N/A

How to Rate JSON Errors:

ERROR	DIMENSION
Information is present within JSON, but incorrect	Truthfulness/Groundedness
JSON includes fields (or are different) that are not requested by the user or the defined schema	Instruction Following
JSON is missing or incomplete	Content Completeness
JSON is invalid	Writing Style