1: Secure Sufficient Access and Resources			
Item		**Fulfilled?**	**Notes/Evidence**
1.1	**<u>Requirement:</u> The evaluator secured sufficient technical access to assess the specific system characteristics being evaluated.** *Note: for the below sub-items of item 1.1, only 1.1.1 is required. The rest may be important to a valid evaluation, but will depend on the nature of the evaluation methods.*		
	1. The evaluator had query access.	Yes	
	2. The evaluator had access to the system's scaffolding.	No	
	3. The evaluator had exemptions from system safeguards.	No	
	4. The evaluator had access to intermediate system states.	No	
	5. The evaluator had finetuning access.	No	Yes for open weight models
	6. The evaluator had access to model weights.	No	Yes for open weight models
	7. The evaluator was granted access to other tools for elicitation or otherwise supporting external validity.	No	

	8. The evaluator had access to relevant user data.	No	
1.2	**<u>Recommendation:</u> The system provider shared with the evaluator information relevant to carrying out a trustworthy and useful evaluation.**	No	
1.3	**<u>Requirement:</u> The evaluator had access to sufficient computational resources to complete a thorough evaluation.**	Yes	Maybe not sufficient if we want to fine-tune a model for evaluation.
1.4	**<u>Requirement:</u> The evaluator had adequate time to carry out a thorough evaluation.**	Yes	
1.5	**<u>Recommendation:</u> The system provider provided legal safe harbor for actions by the evaluator that are within the agreed upon scope of the evaluation.**	No	

2: Minimized Conflicts of Interest

Item		Fulfilled?	Notes/Evidence
2.1	**<u>Requirement:</u> The evaluator did not receive compensation contingent on the results of the evaluation.**	Yes	
2.2	**<u>Requirement:</u> The system provider did not exercise organizational or financial control over the evaluator.**	Yes	
2.3	**<u>Requirement:</u> The evaluator has published a conflict of interest policy, and it was applied to the evaluation.**	Yes	https://airiskmonitor.net/doc/en/coi

2.4	**<u>Requirement:</u> The evaluator clearly disclosed conflicts of interest relevant to the evaluation.** *Note: For each of the below sub-items of 2.4, indicate whether the answer is "yes" or "no" in the "Notes/Evidence" column, and add accompanying information if the answer is "yes".*		
	1. Was the evaluator paid by the system provider or its direct competitor to conduct the evaluation?		No
	2. Does a meaningful fraction of the evaluator's funding come from the system provider, its employees, or its direct competitors?		No
	3. Do the system provider, its employees, or its direct competitors own equity in the evaluator organization?		No
	4. Do evaluator staff who carried out the evaluation own equity in the system provider or its direct competitors?		No
	5. Do any evaluation staff working on the evaluation simultaneously work for the system provider or its direct competitors?		No
	6. Did the evaluator have any other conflicts of interest relevant to the evaluation?		No
2.5	**<u>Requirement:</u> The evaluator recused any individuals with a significant financial interest in the system provider from carrying out the evaluation.**	Yes	
2.6	**<u>Recommendation:</u> The evaluator disclosed any separate agreements with the system provider that significantly impact the independence and trustworthiness of the evaluation.**	No	

3: Analytic Autonomy			
Item		Fulfilled?	Notes/Evidence
3.1	**<u>Recommendation:</u> The evaluator retained flexibility to define which specific properties of the system to evaluate.**	Yes	
3.2	**<u>Requirement:</u> The evaluator had autonomy in deciding the methods of the evaluation.**	Yes	
3.3	**<u>Recommendation:</u> The evaluator ran the evaluations themselves via direct system access.**	Yes	
3.4	**<u>Requirement:</u> The evaluator retained editorial control over how they present the results of their evaluation.**	Yes	
4: Transparent Methods and Results			
Item		Fulfilled?	Notes/Evidence
4.1	**<u>Requirement:</u> The evaluator shared sufficient methodological details to allow for independent review of the results.**	Yes	https://airiskmonitor.net/doc/zh/about
4.2	**<u>Recommendation:</u> The system provider granted any necessary rights to disclose results to the evaluation's intended audiences upfront.**	/	We are doing independent third-party evaluation of publicly available systems; no prior authorization from the system provider was required, nor did they have grounds to restrict publication of the findings.
4.3	**<u>Recommendation:</u> The intended audiences for an evaluation were not narrowed based on the results.**	Yes	

4.4	**<u>Recommendation:</u> The system provider did not misrepresent the evaluation's findings.**	Yes	
4.5	**<u>Recommendation:</u> The system provider or other external parties did not unduly delay the evaluation from being disclosed based on the content of the results.**	Yes	
4.6	**<u>Requirement:</u> The system provider did not have authority to redact results to conceal concerning findings.**	Yes	
4.7	**<u>Recommendation:</u> The evaluator clearly disclosed the redaction authorities granted to the system provider or other external entities.**	/	We don't have **redaction authorities granted to the system provider or other external entities** now**.**
5: Protection of Sensitive Information			
Item		Fulfilled?	Notes/Evidence
5.1	**<u>Requirement:</u> The evaluator had the system provider's permission to release any results based on non-public information or systems.**	/	We don't have **results based on non-public information or systems** now.
5.2	**<u>Recommendation:</u> The evaluation methods were not gamed or leaked.**	No	Our evaluation methods are public, which may be gamed.
5.3	**<u>Requirement:</u> The evaluator implemented measures to protect any confidential information they received from the system provider.**	/	We don't have **confidential information received from the system provider** now.
5.4	**<u>Requirement:</u> The evaluator established and followed a responsible disclosure policy.**	Yes	https://airiskmonitor.net/doc/en/report/2025-Q4#information-hazard