cases	doc_1		doc_2		decision	id
	authors	• Myeongjun Jang • Thomas Lukasiewicz	**authors**	• Jang, Myeongjun • Lukasiewicz, Thomas	DUPLICATES	101
	title	Consistency Analysis of ChatGPT	**title**	Consistency Analysis of ChatGPT		
	publication_date	2023-03-11 00:00:00	**publication_date**	2023-03-10 00:00:00		
	source	SupportedSources.INTERNET_ARCHIVE	**source**	SupportedSources.CORE		
	journal		**journal**			
	volume		**volume**			
	doi		**doi**	None		
	urls	• https://web.archive.org/web/20230314050900/https://arxiv.org/pdf/2303.06273v1.pdf	**urls**	• http://arxiv.org/abs/2303.06273		
	id	id-2548784724965518265	**id**	id6760267510230648194		
	abstract	ChatGPT, a question-and-answer dialogue system based on a large language model, has gained huge popularity since its introduction. Its positive aspects have been reported through many media platforms, and some analyses even showed that ChatGPT achieved a decent grade in professional exams, including the law, medical, and finance domains, adding extra support to the claim that AI now can assist and, even, replace humans in industrial fields. Others, however, doubt its reliability and trustworthiness. In this paper, we investigate ChatGPT's trustworthiness regarding logically consistent behaviours. Our findings suggest that, although ChatGPT seems to achieve an improved language understanding ability, it still fails to generate logically correct predictions frequently. Hence, while it is true that ChatGPT is an impressive and promising new technique, we conclude that its usage in real-world applications without thorough human inspection requires further consideration, especially for risk-sensitive areas.	**abstract**	ChatGPT, a question-and-answer dialogue system based on a large language model, has gained huge popularity since its introduction. Its positive aspects have been reported through many media platforms, and some analyses even showed that ChatGPT achieved a decent grade in professional exams, including the law, medical, and finance domains, adding extra support to the claim that AI now can assist and, even, replace humans in industrial fields. Others, however, doubt its reliability and trustworthiness. In this paper, we investigate ChatGPT's trustworthiness regarding logically consistent behaviours. Our findings suggest that, although ChatGPT seems to achieve an improved language understanding ability, it still fails to generate logically correct predictions frequently. Hence, while it is true that ChatGPT is an impressive and promising new technique, we conclude that its usage in real-world applications without thorough human inspection requires further consideration, especially for risk-sensitive areas.Comment: 11 page		
	versions		**versions**			