

Explaining Scientific Hypotheses in Drug Development with Knowledge Graphs

Knowledge Graphs (KGs) are powerful tools for representing and reasoning with complex scientific information, particularly in biomedical research. Link prediction models, a critical application of KGs, can generate valuable scientific hypotheses, such as identifying potential drug repurposing opportunities and drug-target interactions. However, these predictions often lack scientifically grounded explanations, limiting their practical adoption and trustworthiness among biomedical researchers.

This user study investigates the scientific validity of explanations generated by different methods. Participation in this study is strictly voluntary, and will take approximately 60 minutes, but there is no mandatory minimum time. The goal of this study is to evaluate the explanations, not to evaluate you. Any recommendations or suggestions for improvement are welcome.

The study will be conducted fully online with this questionnaire. All your responses are confidential and will be used for research and academic purposes only and will only be available to the researcher team in charge. Members of the research team will only have access to data on a need to know basis. If your data is used for publications or presentations, it will be aggregated and/or anonymised.

No questions that would reveal your racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, health condition or sexual orientation will be asked in this study.

You will be presented with 10 predictions for **drug repurposing (Task 1)** and 10 predictions for **drug-target indication (Task 2)**. For each prediction there are 4 explanations generated with different systems.

Your objective:

- Fill the **Profile** form below so we can know your background better and then move on to Task 1 and Task 2.
- For each explanation, provide the following ratings:
 - **Scientific Validity:** Rate from 1 (least valid) to 5 (most valid). *Scientific Validity refers to how scientifically correct, plausible, and coherent the explanation is based on existing biomedical knowledge.*
 - **Completeness:** Rate from 1 (too simple) to 5 (too complex). *Completeness refers to the extent to which the explanation provides sufficient detail to make the prediction understandable. It should be informative without being unnecessarily overwhelming.*
 - **Relevance:** Rate from 1 (least relevant) to 5 (most relevant). *Relevance relates to the usefulness and informativeness of the explanation for understanding the prediction. An explanation can be scientifically valid but still irrelevant if it does not help clarify why the prediction matters or how it connects to the task at hand.*
- Provide brief comments on the **strengths** and **weaknesses** of the explanations like: the length of the path, preferred properties, or the complexity of the explanation (more than one path or if the ontological expansion are helpful).
- Fill the **Final Evaluation** of each system.

There is no foreseeable risk involved in participating in this study. In our experience, participants enjoy collaborating in this type of research studies.

This study is restricted to participants 18 years and above.

Acknowledgement of the Information provided

<input type="checkbox"/>	<i>YES, I have read this form and fully understand the contents of it.</i>
<input type="checkbox"/>	<i>YES, I understand that my participation in the study is voluntary and that I am free to give up at any time, without having to give any explanations and without any consequences.</i>
<input type="checkbox"/>	<i>YES, I understand that the data collected during the study may be available to the members of the research team, whenever necessary for the study. I authorize team members to have access to this data.</i>
<input type="checkbox"/>	<i>YES, I declare that I make my decision entirely freely and I agree to participate in this study.</i>

User Profiling

Please select your age range

Not answered

What is your background? Choose all that apply.						
	No knowledge	Some knowledge	BSc level or 3+ years	MSc level or 5+ years	PhD level or 10+ years	
Computer Science and Engineering	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Life Sciences	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Information Sciences	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Health Sciences	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

What is your background? Choose all that apply.					
	No knowledge	Novice	Competent	Expert	
Knowledge Graphs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
AI systems	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Molecular Biology	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

How likely are you to trust a decision made by an AI system in the area of healthcare?	
Not answered	