



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Supervisor:

Dr. David Lewis

Dr. Subrahmanyam Murala

Ontology-Based Modelling of AI Fundamental Rights Impacts(FRIA): Integrating Impact Assessment based on VAIR, AIRO, CIDS and LLMs

Kaiyu Chen (23330889)

MSc in Computer Science - Intelligent Systems

- [1] Mantelero, Alessandro and Esposito, Samantha, An Evidence-Based Methodology for Human Rights Impact Assessment (HRIA) in the Development of AI Data-Intensive Systems (March 22, 2021). Computer Law & Security Review, 2021, Available at SSRN: <https://ssrn.com/abstract=3829759>
- [2] Delaram Golpayegani, Harshvardhan J. Pandit, and Dave Lewis. 2023. To Be High-Risk, or Not To Be—Semantic Specifications and Implications of the AI Act’s High-Risk AI Applications and Harmonised Standards. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23). Association for Computing Machinery, New York, NY, USA, 905–915. <https://doi.org/10.1145/3593013.3594050>
- [3] <https://www.aiaaic.org/aiaaic-repository>

Despite the early stage of **AI impact prediction** [1], the EU’s **proposed AI Act** may require **impact assessments** for new AI projects, focusing on **fundamental rights** and other key areas. This project aims to explore how to make semantic models become a help for AI risk assessment [2] and how it can be applied to publicly available AI incident reports from AIAAIC[3] to support compliance and oversight under the AI Act.

Ontologies provide a formal representation of knowledge and the relationships between concepts of a domain. They are used in the requirements specification to guide formal and unambiguous specification of the requirements, particularly in expressing concepts, relations and business rules of domain model with varying degrees of formalization and precision^{[4][5]}

[4] Emebo, Onyeka & Varde, Aparna & Daramola, Olawande. (2021). Common Sense Knowledge, Ontology and Text Mining for Implicit Requirements.

[5] Sugumaran, V., & Storey, V. C. (2002). Ontologies for conceptual modeling: their creation, use, and management. *Data & knowledge engineering*, 42(3), 251-271.

Fundamental Rights Impact Assessment (FRIA)^[6]

Artificial intelligence can incredibly enhance law enforcement agencies' capabilities to prevent, investigate, detect and prosecute crimes, as well as to predict and anticipate them. However, despite the numerous promised benefits, the use of AI systems in the law enforcement domain raises numerous ethical and legal concerns.

The ALIGNER Fundamental Rights Impact Assessment (AFRIA) is a tool addressed to LEAs who aim to deploy AI systems for law enforcement purposes within the EU. The AFRIA is a reflective exercise, seeking to further enhance LEAs' already existing legal and ethical governance systems, by assisting them in building and demonstrating compliance with ethical principles and fundamental rights while deploying AI systems.

Fundamental Rights Impact Assessment Template				
Name				
Organisation/Position				
Date				
Contributors				
AI system assessed				
Detailed description of the technology and input data				
Detailed description of the purposes and context of use				
1. Presumption of innocence and right to an effective remedy and to a fair trial				
Everyone charged with a criminal offence must be presumed innocent until proved guilty according to law. Everyone whose rights and freedoms are violated has the right to an effective remedy before a tribunal. Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law, including rights: ❖ to be informed promptly of the nature and cause of the accusation; ❖ to bring their arguments and evidence as well as scrutinise and counteract the evidence presented against them; and to obtain an adequately reasoned and accessible decision.				
Challenge	Evaluation	Estimated impact level		
1.1 The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision		-		
1.2 The AI system does not provide percentages or other indication on the degree of likelihood that the outcome is correct/incorrect, prejudicing the user that there is no possibility of error and therefore that the outcome is undoubtedly incriminating		-		
1.3 The AI system produces an outcome that forces a reversal of burden of proof upon the suspect, by presenting itself as an absolute truth, practically depriving the defence of any chance to counter it		-		
1.4 There is no explanation of reasons and criteria behind a certain output of the AI system that the user can understand		-		
1.5 There is no indication of the extent to which the AI system influences the overall decision-making process		-		
1.6 There is no set of measures that allow for redress in case of the occurrence of any harm or adverse impact		-		

AI Risk Ontology (AIRO)^[7]

AIRO is an ontology for expressing risk of harm associated with AI systems based on the EU AI Act and ISO/IEC 23894 on AI risk management.

Vocabulary of AI Risks (VAIR)^[8]

VAIR is an open vocabulary for AI risks. VAIR is intended to assist with identification and documentation of risks by providing a common vocabulary that facilitates knowledge sharing and interoperability between actors in the AI value chain. VAIR provides semantic specifications for cataloguing AI risks in a FAIR (Findable, Accessible, Interoperable, Reusable) manner.

Common Impact Data Standard(CIDS)^[9]

The Common Impact Data Standard is a standardized way to represent a social purpose organization's (SPO) impact model (i.e. their theory of change, logic model, outcome chain, etc). It is a way to represent impact as defined by the Impact Management Project Norms (now housed at Impact Frontiers). It enables the exchange of impact information between organizations regardless of the impact models being used.

^[7] <https://w3id.org/airo>

^[8] <https://w3id.org/vair>

^[9] <https://www.commonapproach.org/common-impact-data-standard/> 5

Research Questions

1. How can we integrate **Fundamental Rights Impact Assessment (FRIA)** into existing ontological frameworks such as **AI Risk Ontology (AIRO)**, **Vocabulary of AI Risks (VAIR)**, and **Common Impact Data Standard (CIDS)** to create a more comprehensive ontological structure for impact assessment?
2. To what extent can **Large Language Models (LLMs)** be effectively utilized to populate Fundamental Rights Impact Assessment (FRIA) reports and related ontologies, thereby assisting in the completion of AI impact assessments?

Motivation



New technologies have profoundly changed how we organise and live our lives. In particular, new data-driven technologies have spurred the development of **artificial intelligence (AI)**, including increased automation of tasks usually carried out by humans. The COVID-19 health crisis has boosted AI adoption and data sharing – creating new opportunities, but also challenges and threats to human and fundamental rights.

Artificial Intelligence (AI) impact assessments **are crucial** for evaluating the potential impact of AI systems on **fundamental rights** such as privacy, non-discrimination, and freedom of expression.

LLMs can extract and categorize information from vast amounts of text into structured data that can be integrated into ontologies, contributing to the creation of comprehensive and accurate knowledge graphs.

Main Goal

- Develop FIRA ontology
- Find the relationship between FRIA and CIDS, AIRO, VAIR
- Develop the relationship
- Have some useful prompt for LLMs to generate the instances based on the designed ontology and given incidents
- Evaluation
- GraphDB show cases (competency questions)

Main Goal

Research Question 1

- Develop FIRA ontology
- Find the relationship between FRIA and CIDS, AIRO, VAIR
- Develop the relationship

Research Question 2

- Have some useful prompt for LLMs to generate the instances based on the designed ontology and given incidents
- Evaluation
- GraphDB show cases (competency questions)

Why I did this?

- Respond to the **EU AI Act**
- Aligning with **legal** and regulatory requirements, promoting compliance and responsible AI development.
- Sharing and reusing knowledge across domains, contributing to develop **a basic farmwork for AI impact Assessment.**
- Advancing AI governance and ensuring AI systems respect **fundamental rights** and ethical principles.
- Ensuring consistency and compatibility across various AI impact assessment frameworks. Make sure the information sharing and collaboration.
- Developing a comprehensive and holistic approach to assessing AI system risks and impacts.

Design of the ontology

```
1 @prefix fria: <http://www.example.org/fria-report#> .
2 @prefix airo: <https://w3id.org/airo#> .
3 @prefix vair: <https://w3id.org/vair#> .
4 @prefix cids: <http://www.example.org/cids#> .
5 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
6 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
7 @prefix owl: <http://www.w3.org/2002/07/owl#> .
8 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
9
10 # Basic Things
11 fria:FRIA-report a rdfs:Class ;
12     rdfs:comment "A class representing the FRIA report." .
13
14 fria:FRIA-reportName a rdfs:Class ;
15     rdfs:comment "A class representing the name of the FRIA report." ;
16     owl:equivalentClass cids: hasName .
17
18 fria:hasReportName a rdf:Property ;
19     rdfs:domain fria:FRIA-report ;
20     rdfs:range xsd:string ;
21     rdfs:comment "A property to hold the name of the FRIA report." .
22
23 fria:FRIA-reportorganisationPosition a rdfs:Class ;
24     rdfs:comment "A class representing the organisation position in the FRIA report." .
25
26 fria:hasOrganisationPositionDescription a rdf:Property ;
27     rdfs:domain fria:FRIA-reportorganisationPosition ;
28     rdfs:range xsd:string ;
29     rdfs:comment "A property to hold the description of the organisation position in the FRIA report." .
30
31 fria:FRIA-reportcontributors a rdfs:Class ;
32     rdfs:comment "A class representing the contributors to the FRIA report." .
33
34 fria:hasContributorDetails a rdf:Property ;
35     rdfs:domain fria:FRIA-reportcontributors ;
36     rdfs:range xsd:string ;
37     rdfs:comment "A property to hold the details of the contributors to the FRIA report." .
38
39 fria:FRIA-reportaiSystemAssessed a rdfs:Class ;
40     rdfs:comment "A class representing the AI system assessed in the FRIA report." ;
41     owl:equivalentClass cids: hasConsequence, cids: forOutcome, airo: producesOutput, airo: hasConsequence, vair: Assessment, vair: AssessingPeopleRelatedRisk .
```

FRIA with CIDS, AIRO, VAIR

Basic Things

FRIA	CIDS	AIRO	VAIR
hasName	hasName		
organisationPosition			
date			
contributors			
aiSystemAssessed	hasConsequence	producesOutput	Assessment
	forOutcome	hasConsequence	AssessingPeopleRelatedRisk
technologyAndData		usesTechnique	
purposesAndContext	hasDescription	haspurpose	Purposes

Some of the Challenges Related Things

Challenge11

The AI system does not communicate that a decision/advice or outcome **is the result of an algorithmic decision**

VAIR

Transparency

Property of a system that appropriate information about the system **is made available to relevant stakeholders**

DecisionMaking

Generation of decisions

FRIA with CIDS, AIRO, VAIR

Some of the Challenges Related Things

Challenge16

There is no set of measures that allow for redress in case of the occurrence of any harm or adverse impact

AIRO

hasRisk

Indicates risks associated with an AI system, an AI component, etc.

hasSeverity

Indicates severity of a consequence or an impact

CIDS

hasImportance

Specifies the nature of the importance. One of {"high importance", "moderate important", "neutral", "unimportant"}.

intendedimpact

Identifies the intended direction of the change. Note that ImpactReport captures the actual direction. This helps to inform the interpretation of the ImpactReport. helps to inform the interpretation of the ImpactReport.

Research Questions

1. How can we integrate **Fundamental Rights Impact Assessment (FRIA)** into existing ontological frameworks such as **AI Risk Ontology (AIRO)**, **Vocabulary of AI Risks (VAIR)**, and **Common Impact Data Standard (CIDS)** to create a more comprehensive ontological structure for impact assessment?
2. To what extent can **Large Language Models (LLMs)** be effectively utilized to populate Fundamental Rights Impact Assessment (FRIA) reports and related ontologies, thereby assisting in the completion of AI impact assessments?

Prompt

Now, when given an incident report, **find the relevant information** for each part of the RDF definition and **fill in the Turtle format** using the information found. So you need to return Turtle representation for the incident report. Always include **all of the properties** in the RDF output. If there is no relevant information for a given property in the report, return a blank value in the RDF. However, if you can understand and fill in the blank value based on the available information, start the value with "LLM understand: ". For Basic Things like fria:hasAssessmentContent, if the provided information is too long, you then need to **summarize appropriately** based on the content. For the date, try to access the specific AIAAIC Link and retrieve the time information, if you can't access the provided link, try to find it based on provided information. If there are some details you cannot find or understand, leave those properties blank.

For the Challenges, Evaluation, and Impact Level sections, follow the guidelines provided:

Review the FRIA template's pre-listed challenges, adding any system-specific ones. Evaluate each challenge's relevance to the AI system, explaining its embedding and impact **within the law enforcement context**. Assess the severity of prejudice and affected population to determine the overall impact level using the provided matrix. Consider the predetermined context of use, including target group, geographical area, deployment period, and trigger conditions. Provide detailed explanations for each evaluation and impact estimate. Regularly update the FRIA to reflect changes in the AI system's functioning or deployment circumstances.

When evaluating the impact level, **follow the guidelines provided in the prompt**:

- 1.Determine the severity of prejudice: Negligible, Critical, or Catastrophic.
- 2.Evaluate the number of affected individuals: Low, Medium, or High.
- 3.Use the impact matrix to determine the overall impact level: Low, Medium, High, or Very High.

Based on the instructions provided, you can only have 4 impact level! Decide inside 4 of them(Low, Medium, High, or Very High)!!

You need to return complete RDF/Turtle representation for the incident report, including all basic things, challenges, evaluations and impact levels for each challenge.

Do you understand? if so reply yes with no further follow up and await an incident report.

Claude 3.5 Sonnet

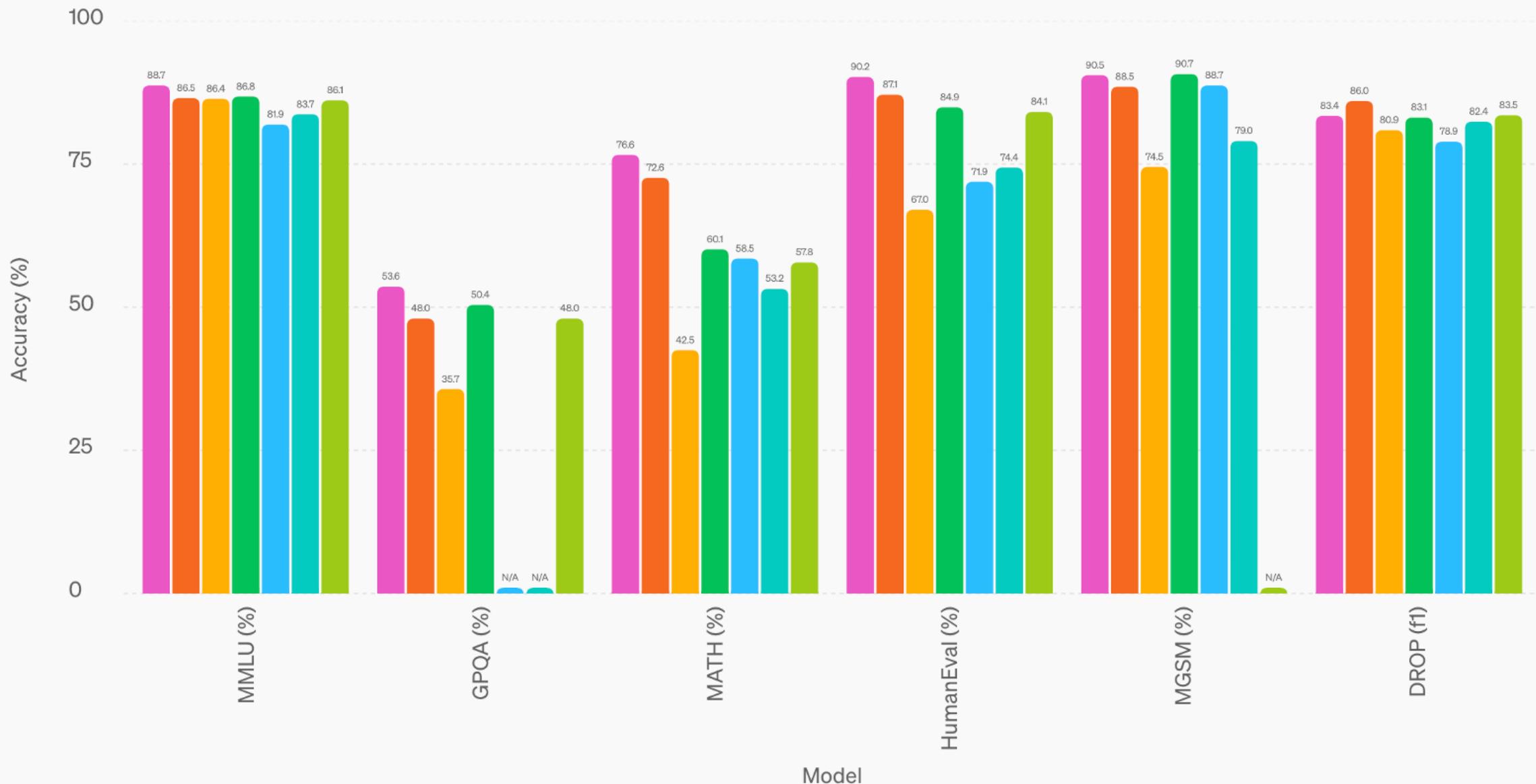
Claude 3.5 Sonnet raises the industry bar for intelligence, outperforming competitor models and Claude 3 Opus on a wide range of evaluations, with the speed and cost of our mid-tier model, Claude 3 Sonnet.

GPT-4o

GPT-4o (“o” for “omni”) is a step towards much more natural human-computer interaction—it accepts as input any combination of text, audio, image, and video and generates any combination of text, audio, and image outputs. It can respond to audio inputs in as little as 232 milliseconds, with an average of 320 milliseconds, which is similar to human response time([opens in a new window](#)) in a conversation.

Text Evaluation

GPT-4o GPT-4T GPT-4 (Initial release 23-03-14) Claude3 Opus Gemini Pro 1.5 Gemini Ultra 1.0 Llama3 400b



	Claude 3.5 Sonnet	Claude 3 Opus	GPT-4o	Gemini 1.5 Pro	Llama-400b (early snapshot)
Graduate level reasoning <i>GPQA, Diamond</i>	59.4%* 0-shot CoT	50.4% 0-shot CoT	53.6% 0-shot CoT	—	—
Undergraduate level knowledge <i>MMLU</i>	88.7%** 5-shot	86.8% 5-shot	—	85.9% 5-shot	86.1% 5-shot
	88.3% 0-shot CoT	85.7% 0-shot CoT	88.7% 0-shot CoT	—	—
Code <i>HumanEval</i>	92.0% 0-shot	84.9% 0-shot	90.2% 0-shot	84.1% 0-shot	84.1% 0-shot
Multilingual math <i>MGSM</i>	91.6% 0-shot CoT	90.7% 0-shot CoT	90.5% 0-shot CoT	87.5% 8-shot	—
Reasoning over text <i>DROP, F1 score</i>	87.1 3-shot	83.1 3-shot	83.4 3-shot	74.9 Variable shots	83.5 3-shot Pre-trained model
Mixed evaluations <i>BIG-Bench-Hard</i>	93.1% 3-shot CoT	86.8% 3-shot CoT	—	89.2% 3-shot CoT	85.3% 3-shot CoT Pre-trained model
Math problem-solving <i>MATH</i>	71.1% 0-shot CoT	60.1% 0-shot CoT	76.6% 0-shot CoT	67.7% 4-shot	57.8% 4-shot CoT
Grade school math <i>GSM8K</i>	96.4% 0-shot CoT	95.0% 0-shot CoT	—	90.8% 11-shot	94.1% 8-shot CoT

* Claude 3.5 Sonnet scores 67.2% on 5-shot CoT GPQA with maj@32

** Claude 3.5 Sonnet scores 90.4% on MMLU with 5-shot CoT prompting

Generated Instances Example

```
@prefix fria: <http://www.example.org/fria-report#> .
@prefix airo: <https://w3id.org/airo#> .
@prefix vair: <https://w3id.org/vair#> .
@prefix cids: <http://www.example.org/cids#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

fria:FRIA-report-3D-masks-fool-facial-recognition a fria:FRIA-report ;
    fria:hasReportName "3D masks fool payment, airport facial recognition systems" ;
    fria:hasOrganisationPositionDescription "Kneron, an artificial intelligence company, conducted the research" ;
    fria:hasContributorDetails "Researchers from Kneron, including CEO Albert Liu" ;
    fria:hasAssessmentContent "Researchers found that facial recognition technology can be fooled by using 3D-printed masks depicting
    fria:hasTechnologyAndDataDescription "Facial recognition technology used in payment systems, airport security, and transportation
    fria:hasPurposesAndContextDescription "To test the security and reliability of facial recognition systems in public spaces and pa
    fria:reportDate "2023-10-10"^^xsd:date ;
    fria:hasAIAAICLink "https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/3d-masks-fool-payment-airpor

fria:FRIA-reportChallenge11 a fria:FRIA-reportChallenge ;
    fria:FRIA-reportHasEvaluation fria:FRIA-reportEvaluation11 ;
    fria:FRIA-reportHasImpactLevel fria:FRIA-reportImpactLevel11 ;
    rdfs:comment "The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision" ;
    rdfs:subClassOf fria:FRIA-reportChallenge1 ;
    owl:equivalentClass airo:Transparency, vair:OperatingCriticalDigitalInfrastructure .

fria:FRIA-reportEvaluation11 a fria:FRIA-reportEvaluation ;
    fria:hasEvaluationContent "The facial recognition systems tested did not indicate that they were using AI for identification." .

fria:FRIA-reportImpactLevel11 a fria:FRIA-reportImpactLevel ;
    fria:hasImpactLevelContent "High" .
```

Generated Instances

I've generate 50 instances using both the most advanced LLMs right now:
Claude 3.5 Sonnet and GPT-4o

Evaluation

file_name	report_name	organisation_description	contributor_details	assessment_content	technology_description	purposes_description	report_date	aiaaic_link	challenge_11	evaluation_11	impact_level_11	challenge_12
fria-instance-gpt-1.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-2.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-3.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-4.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-5.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-6.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-7.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-8.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-9.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	0.5	1.0	1.0000000000000000
fria-instance-gpt-10.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	0.5	1.0	1.0000000000000000
fria-instance-gpt-11.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	0.5	1.0	1.0000000000000000
fria-instance-gpt-12.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-13.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-14.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	0.0	0.0	0.0
fria-instance-gpt-15.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-16.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-17.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-18.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-19.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-20.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-21.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-22.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-23.ttl	1	0.5	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-24.ttl	1	0.5	0.5	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-25.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-26.ttl	1	0.5	0.5	1	0	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-27.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-28.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-29.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-30.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-31.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	1.0	1.0	0.815275442423370
fria-instance-gpt-32.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-33.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-34.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-35.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-36.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-37.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-38.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-39.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-40.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-41.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-42.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-43.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-44.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	0.0	1.0	1.0
fria-instance-gpt-45.ttl	1	1.0	1.0	0	1	1	1	1	1.0000000000000000	0.5	0.5	1.0000000000000000
fria-instance-gpt-46.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	1.0000000000000000
fria-instance-gpt-47.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-48.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-49.ttl	1	1.0	1.0	1	1	1	1	1	1.0000000000000000	1.0	1.0	0.0
fria-instance-gpt-50.ttl	1	1.0	1.0	1	1	1	1	1	1.04968883751210100	1.0	1.0	0.1292814965223100

Evaluation

1	1.0000000000000000	1.0	1.0	1.0000000000000000
1	1.0000000000000000	1.0	1.0	0.0
1	1.0000000000000000	1.0	1.0	0.0
1	1.0000000000000000	1.0	1.0	0.0
1	0.14968883751210100	1.0	1.0	0.12928149652243100

Similarity Scores Calculated for evaluate the challenge in Frequency Analysis

The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision.

The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision.

The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision.

Personal data may be captured from people who are not even aware that the device is there, or that it records and processes audio and personal data.

For Challenge, they should look the same since it's designed in the ontology.

So if it's different, it means that the LLMs didn't follow the prompt well and may have wrong understanding of the ontology.

In this example, the challenge is wrong. The Similarity Scores are used to see the correctness of the ontology instance.

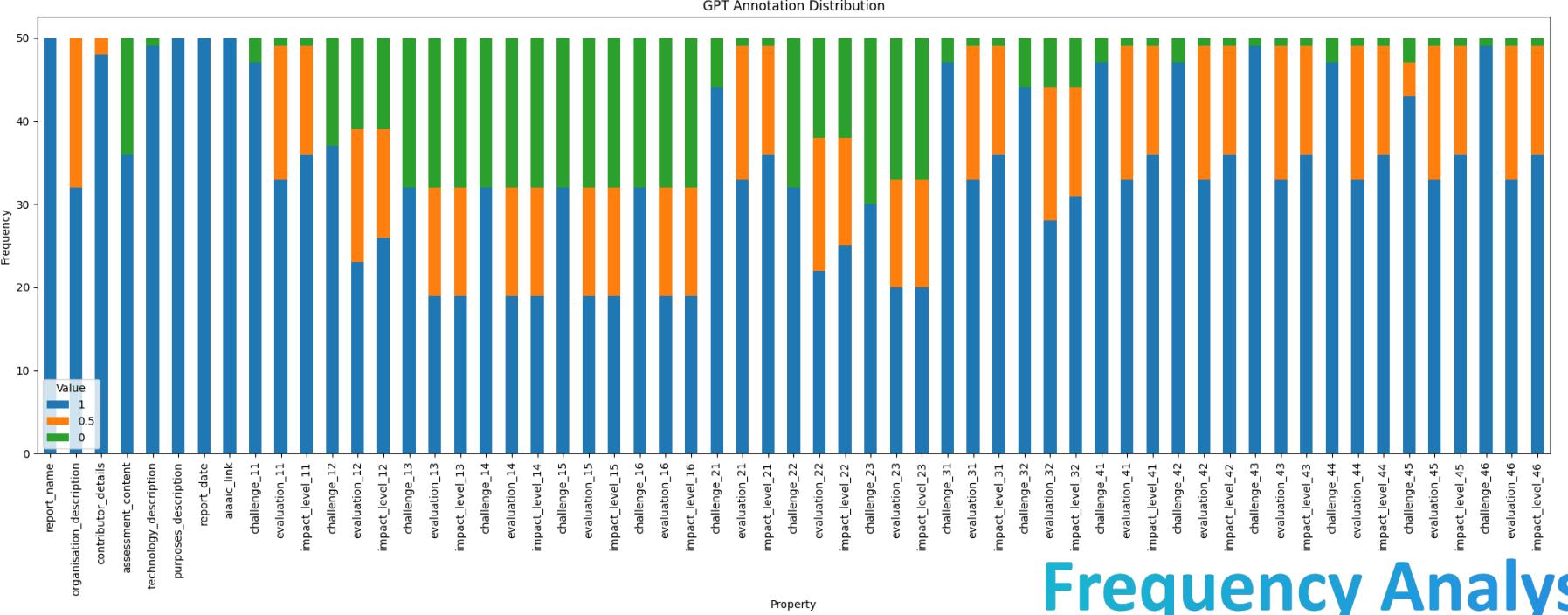
Higher the scores, more correct the instance is.

Evaluation

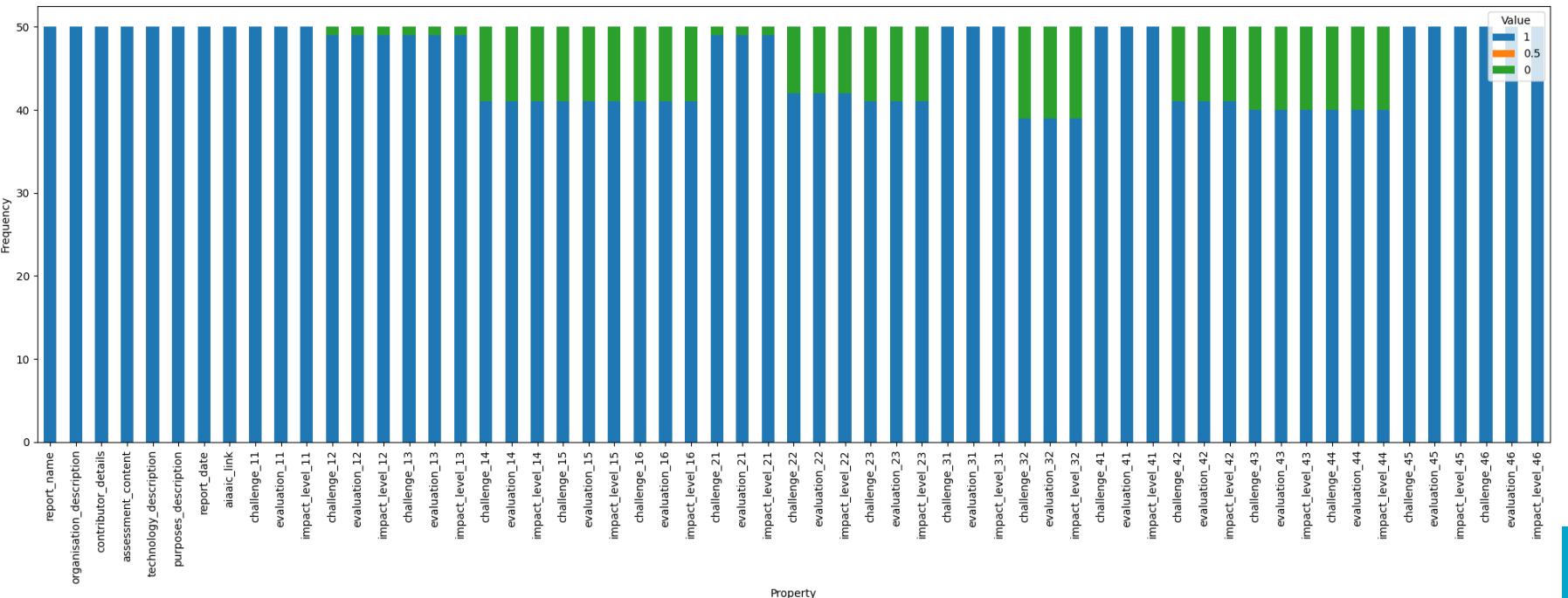
file_name	report_name	organisation_description	contributor_details	assessment_content	technology_description	purposes_description	report_date	aiaaic_link	challenge_11	evaluation_11	impact_level_11
fria-instance-gpt-1.ttl	1	0.5	1.0	1	1	1	1	1	1.000000000000000	1.0	1.0
fria-instance-gpt-2.ttl	1	0.5	1.0	1	1	1	1	1	1.000000000000000	1.0	1.0
fria-instance-gpt-3.ttl	1	0.5	1.0	1	1	1	1	1	1.000000000000000	1.0	1.0
fria-instance-gpt-4.ttl	1	1.0	1.0	1	1	1	1	1	1.000000000000000	1.0	1.0
fria-instance-gpt-5.ttl	1	1.0	1.0	1	1	1	1	1	1.000000000000000	1.0	1.0
fria-instance-gpt-6.ttl	1	1.0	1.0	1	1	1	1	1	1.000000000000000	1.0	1.0
fria-instance-gpt-7.ttl	1	1.0	1.0	1	1	1	1	1	1.000000000000000	1.0	1.0
fria-instance-gpt-8.ttl	1	1.0	1.0	1	1	1	1	1	1.000000000000000	1.0	1.0

Since this is used for frequency analysis,
if the defined properties exist and are filled by the LLMs, then the value is 1,
If the if the defined properties exist but aren't filled by the LLMs, then the value is 0.5,
If the defined properties don't exist, then the value is 0.

Evaluation



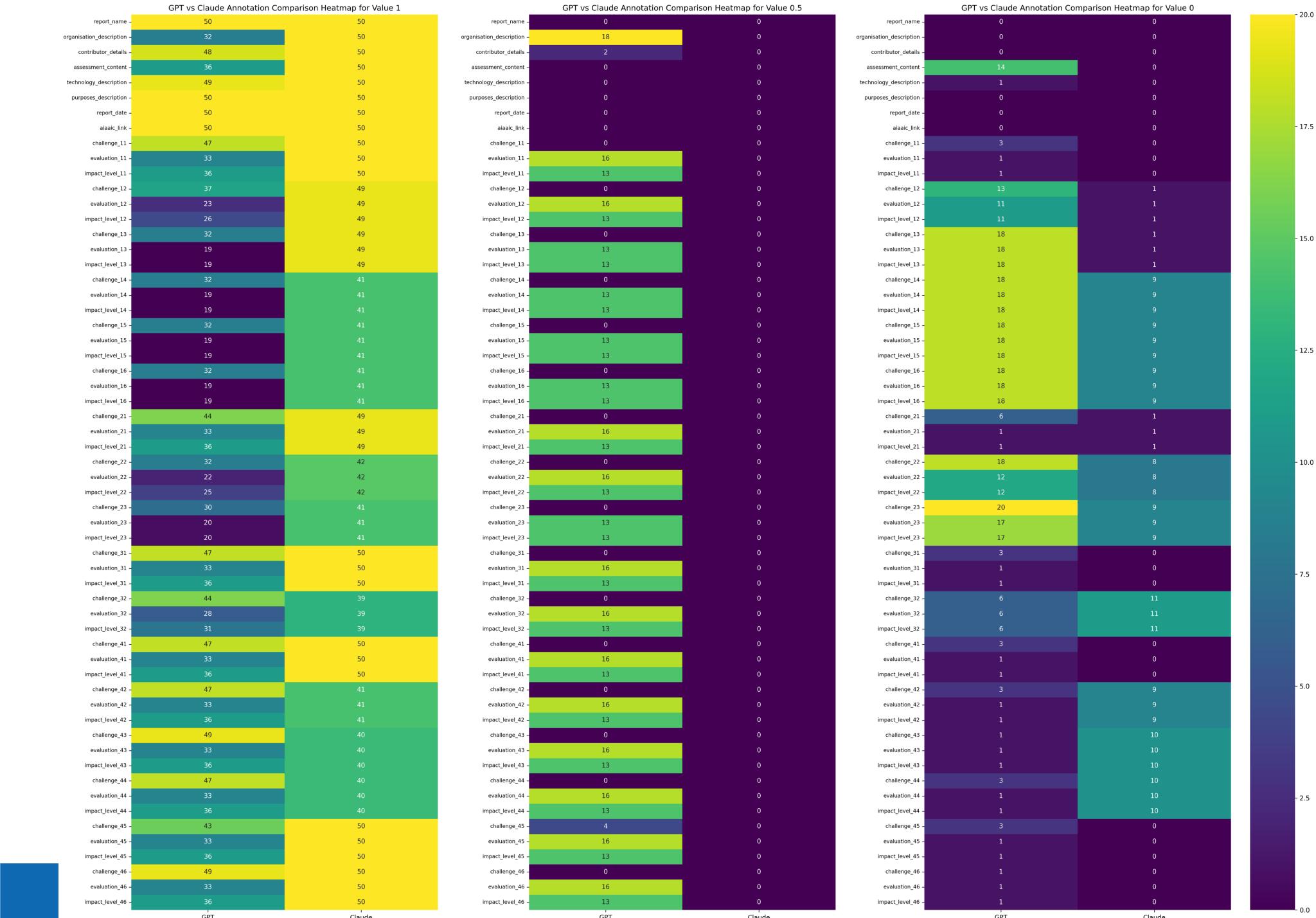
Frequency Analysis



Claude's
performance is
better than ChatGPT
in this analysis.

Frequency Analysis

Claude's performance is better than ChatGPT in this analysis.



Evaluation

file_name	report_name	organisation_description	contributor_details
fria-instance-gpt-1.ttl	3D masks fool payment, airport facial recognition systems		AI development company Kneron
fria-instance-gpt-2.ttl	4 Little Trees (4LT) student emotion recognition		Find Solution AI
fria-instance-gpt-3.ttl	7-Eleven customer survey facial recognition		7-Eleven
fria-instance-gpt-4.ttl	Aadhaar COVID-19 facial recognition marginalisation	Plans by the Indian government to use facial recognition integrated within the Aadhaar biometric ID system have raised fears that millions of vulnerable people without mobile phones or internet access will lose out on receiving benefits.	Reuters, Internet Freedom Foundation
fria-instance-gpt-5.ttl	Aadhaar glitches result in villagers' starvation	India's Aadhaar biometric ID system faced technical problems leading to villagers being unable to access food rations or subsidized grain, resulting in starvation and deaths.	Campaigners, The Guardian
fria-instance-gpt-6.ttl	AccessiBe automated web accessibility	The use of AccessiBe's automated web accessibility solution by Eyebobs led to a lawsuit due to failure to provide users of its website with equal access, especially for screen readers.	Accessibility advocates, software deve
fria-instance-gpt-7.ttl	Adobe Creative Cloud content analysis	Adobe's automated content analysis of Creative Cloud files to train AI algorithms raised privacy and employment concerns among users.	Adobe, Adobe Creative Cloud users, B
fria-instance-gpt-8.ttl	Adobe Firefly AI art generator training	Adobe's use of Adobe Stock content to train the Firefly AI art generator without explicit consent from contributors raises significant ethical, copyright, and employment concerns.	Adobe, Adobe Stock contributors, Ven
fria-instance-gpt-9.ttl	Adobe Sensei Project Morpheus		Adobe
fria-instance-gpt-10.ttl	Aespa virtual K-pop anthropomorphism, sexualisation		SM Entertainment
fria-instance-gpt-11.ttl	Agricultural Bank of China facial recognition age bias		Agricultural Bank of China
fria-instance-gpt-12.ttl	AI confuses bus ad for jaywalker		Ningbo Police
fria-instance-gpt-13.ttl	AI converts Asian-American student into Caucasian		Playground AI
fria-instance-gpt-14.ttl	AI Dungeon GPT-3 offensive speech filter	Content moderation system NLP/text analysis	AI Dungeon, Latitude, OpenAI, Suchin
fria-instance-gpt-15.ttl	AI-generated article calls fake tanning 'racist'		Irish Times
fria-instance-gpt-16.ttl	AI impersonation scams Canadian couple of USD 21,000		Canadian Authorities, Washington Post
fria-instance-gpt-17.ttl	AI invents 40,000 biochemical warfare agents		Researchers from USA, UK, Switzerlan
fria-instance-gpt-18.ttl	AI meal planner app suggests chlorine gas recipe		Pak 'n Save, New Zealand political cor
fria-instance-gpt-19.ttl	AI Portrait Art racial bias	Generative adversarial network (GAN) Neural network Machine learning	Mashable journalist Morgan Sung disc
fria-instance-gpt-20.ttl	AI satellite location spoofing		University of Washington
fria-instance-gpt-21.ttl	AI Stefanie Sun (AI孙燕姿)		Bilibili
fria-instance-gpt-22.ttl	AI text detector language bias		Stanford University
fria-instance-gpt-23.ttl	Airbnb Smart Pricing algorithm racial bias		Carnegie Mellon University
fria-instance-gpt-24.ttl	Ajin USA worker crushed to death by robot		
fria-instance-gpt-25.ttl	Alexei Navalny Smart Voting Bot Blocking Report	This report was conducted to evaluate the blocking of Alexei Navalny's Smart Voting bot by Apple, Google, and Telegram.	Research Team at Example.org
fria-instance-gpt-26.ttl	Alfi personalised, real-time advertising		
fria-instance-gpt-27.ttl	Algorithm misses gambling addict red flags	Machine learning	Luke Ashton, from Leicester, UK, was c
fria-instance-gpt-28.ttl	Allocation algorithm wrongly places thousands of Italian teachers	Resource allocation algorithm	An algorithm used by the Italian govern
fria-instance-gpt-29.ttl	Alistate car insurance 'suckers list' overcharging	Price adjustment algorithm	A joint investigation by The Mark Up ar
fria-instance-gpt-30.ttl	Alonzo Sawyer facial recognition wrongful arrest, jailing	CCTV Facial recognition	54-year old Alonzo Sawyer was arreste
fria-instance-gpt-31.ttl	Amazon, Waterstones algorithms promote vaccine misinformation	Researchers at the University of Washington and Sky News conducted studies revealing the promotion of vaccine misinformation by Amazon, Waterstones, and Foyles algorithms.	Researchers at the University of Washi
fria-instance-gpt-32.ttl	Amazon Astro home robot	Robotics; Computer vision; Facial recognition	USA
fria-instance-gpt-33.ttl	Amazon UK automated pricing glitch	Pricing automation	UK
fria-instance-gpt-34.ttl	Amazon AWS Panorama automated workplace surveillance	CCTV, Computer vision	USA
fria-instance-gpt-35.ttl	Amazon botches delivery drone commercial launch	Drone	USA
fria-instance-gpt-36.ttl	Amazon chemical food preservative suicides	Recommendation algorithm	USA; India
fria-instance-gpt-37.ttl	Amazon Driveri delivery driver safety monitoring	CCTV Computer vision	USA
fria-instance-gpt-38.ttl	Amazon delivery drone malfunctions, sparks 25-acre fire	Drone	USA
fria-instance-gpt-39.ttl	Amazon Echo Dot Kids remembers kids' conversations	Speech recognition Natural language understanding (NLU)	USA
fria-instance-gpt-40.ttl	Amazon employees use Ring to spy on customers	CCTV Computer vision	USA
fria-instance-gpt-41.ttl	Amazon Flex algorithm fires delivery drivers	Automated management system Image recognition	USA
fria-instance-gpt-42.ttl	Amazon Flex delivery drivers forced to take unsafe routes	Routing algorithm	USA; EU; UK; Australia
fria-instance-gpt-43.ttl	Amazon Go fails to inform NYC customers about facial recognition	Facial recognition; Computer vision; Deep learning	USA
fria-instance-gpt-44.ttl	Amazon India own brand search engine rigging	Search engine algorithm	Amazon, Diego Piacentini, Russell Gra
fria-instance-gpt-45.ttl	Amazon Mentor delivery driver scoring	Performance scoring algorithm	USA
fria-instance-gpt-46.ttl	Amazon One Palmprint Biometric Opacity	Palm print scanning	TechCrunch, Albert Fox Cahn, Surveill
fria-instance-gpt-47.ttl	Amazon Ring Always Home Cam	Drone; Computer vision	Amazon, TechCrunch, Evan Greer, Fin

Evaluation

challenge_45	evaluation_45	impact_level_45	challenge_46	evaluation_46
1.0000000000000000	0.3040753256113660		1.0	1.0000000000000000
1.0000000000000000	0.31258809797206800		1.0	1.0000000000000000
1.0000000000000000	0.22430556321584300		1.0	1.0000000000000000
0.49304922133780900	0.3250015465650920		1.0	1.0000000000000000
0.7897746636343570	0.27824562220363600		1.0	1.0000000000000000
0.49304922133780900	0.1238273358432390	0.6666666666666670	1.0000000000000000	0.12034344390398000
0.49304922133780900	0.25817837206566500		1.0	1.0000000000000000
0.49304922133780900	0.18383925004159200	0.6666666666666670	1.0000000000000000	0.17022838043477500
1.0000000000000000	0.07701564005870900		1.0	1.0000000000000000
1.0000000000000000	0.13045377565199000		1.0	1.0000000000000000
1.0000000000000000	0.15923229953482200	0.6666666666666670	1.0000000000000000	0.14471270582616800
1.0000000000000000	0.5024805573663780	0.6666666666666670	1.0000000000000000	0.2796724903912870
1.0000000000000000	0.36775666078015000	0.6666666666666670	1.0000000000000000	0.25534187135075000

Similarity Scores Calculated for evaluation the different LLMs' Accuracy of Judgments

For Impact Level, I took the impact level out if it's existed. (It should be in "low", "medium", "high" or "very high")

If there is no available value from any of the LLMs' output, then it is 0.

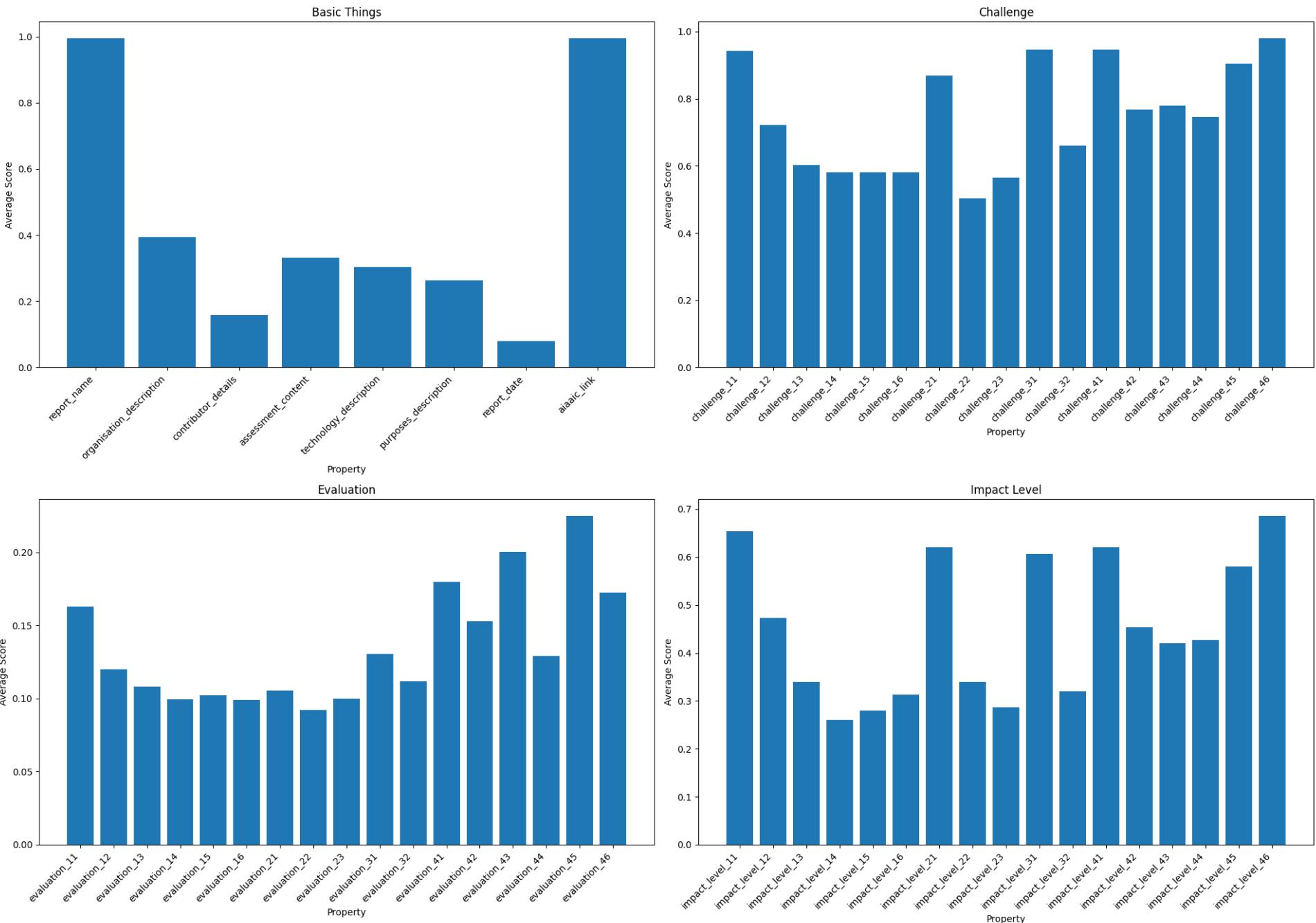
If it's same, then it is 1.

If they are different, it calculates the distance and returns score based on distance.

Evaluation Performance Analysis

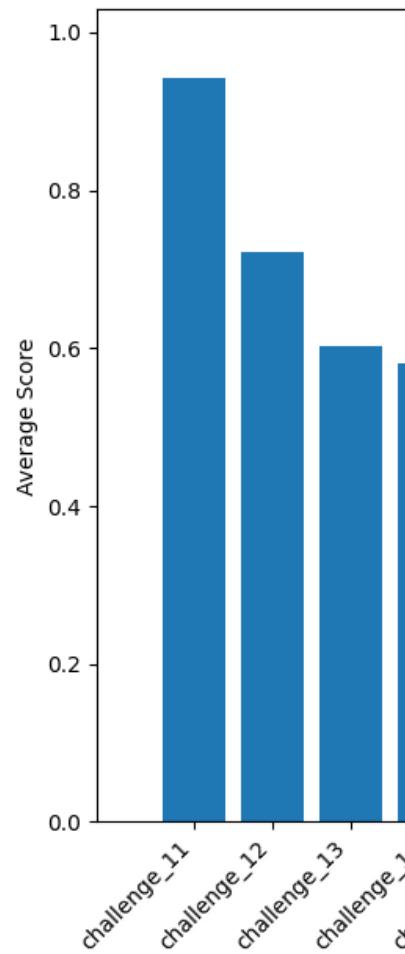
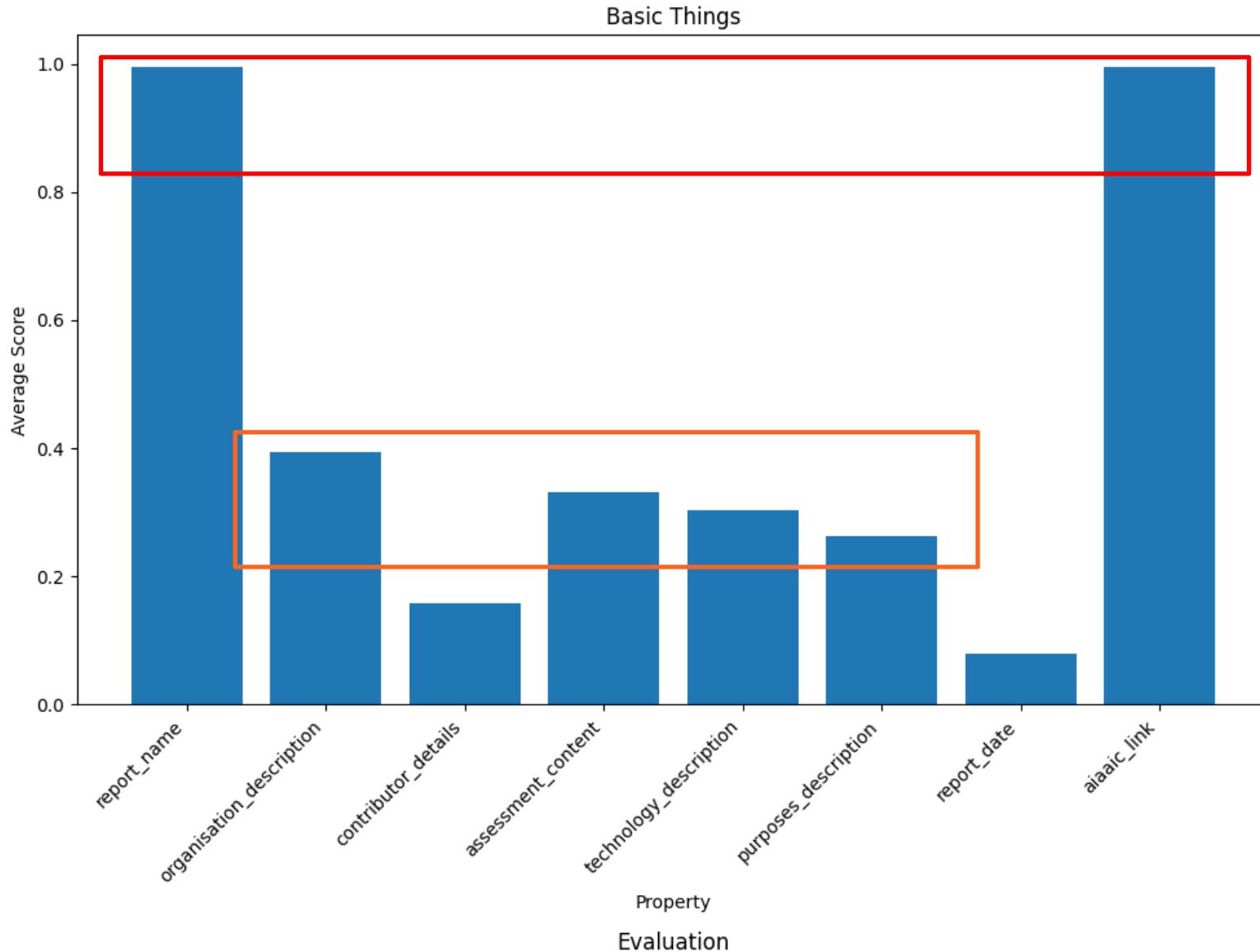
The overall similarity of the Impact Level is the best. But for incidents, they have different thoughts.

Average Similarity Scores for AIAAIC Properties by Category



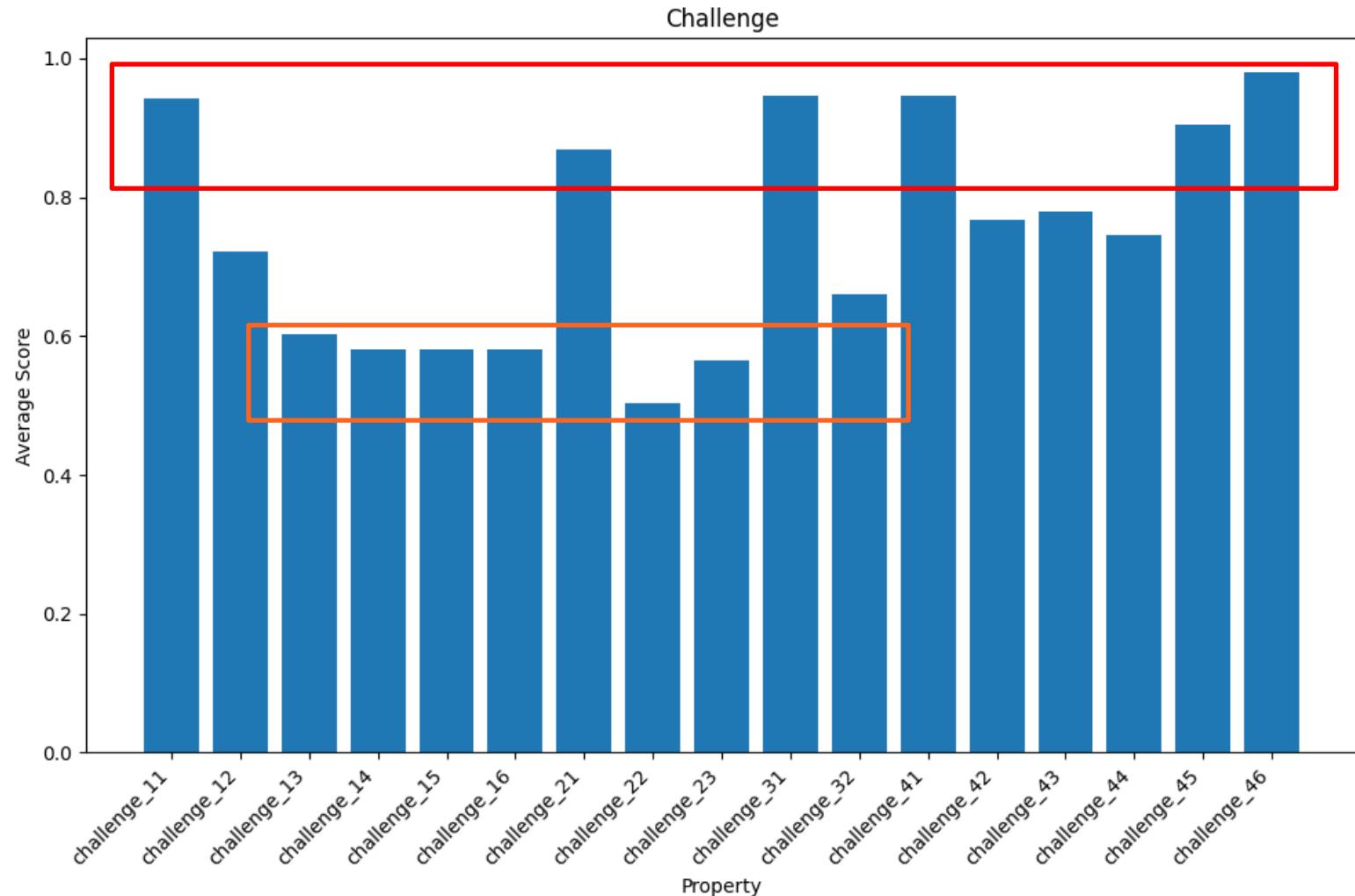
Evaluation Performance Analysis

In Basic Things,
LLM extracted all
the basic thing
successfully.
Except the
Report Date.
This may
because they
don't have
access to the
webpage



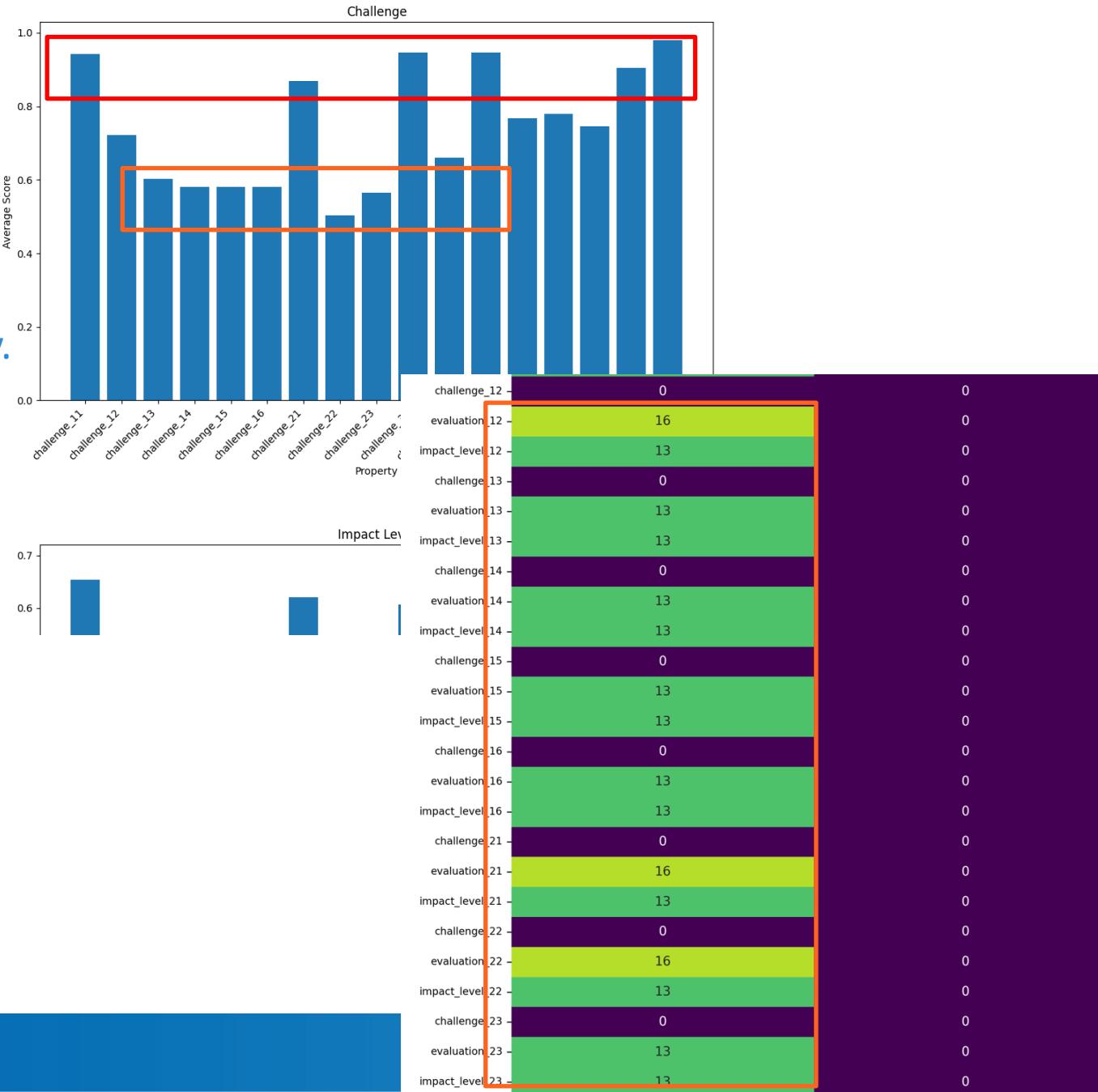
Evaluation Performance Analysis

Most of the challenge are extracted successfully. The low similarity scores may because of the low performance of GPT4o. It missed these thing.



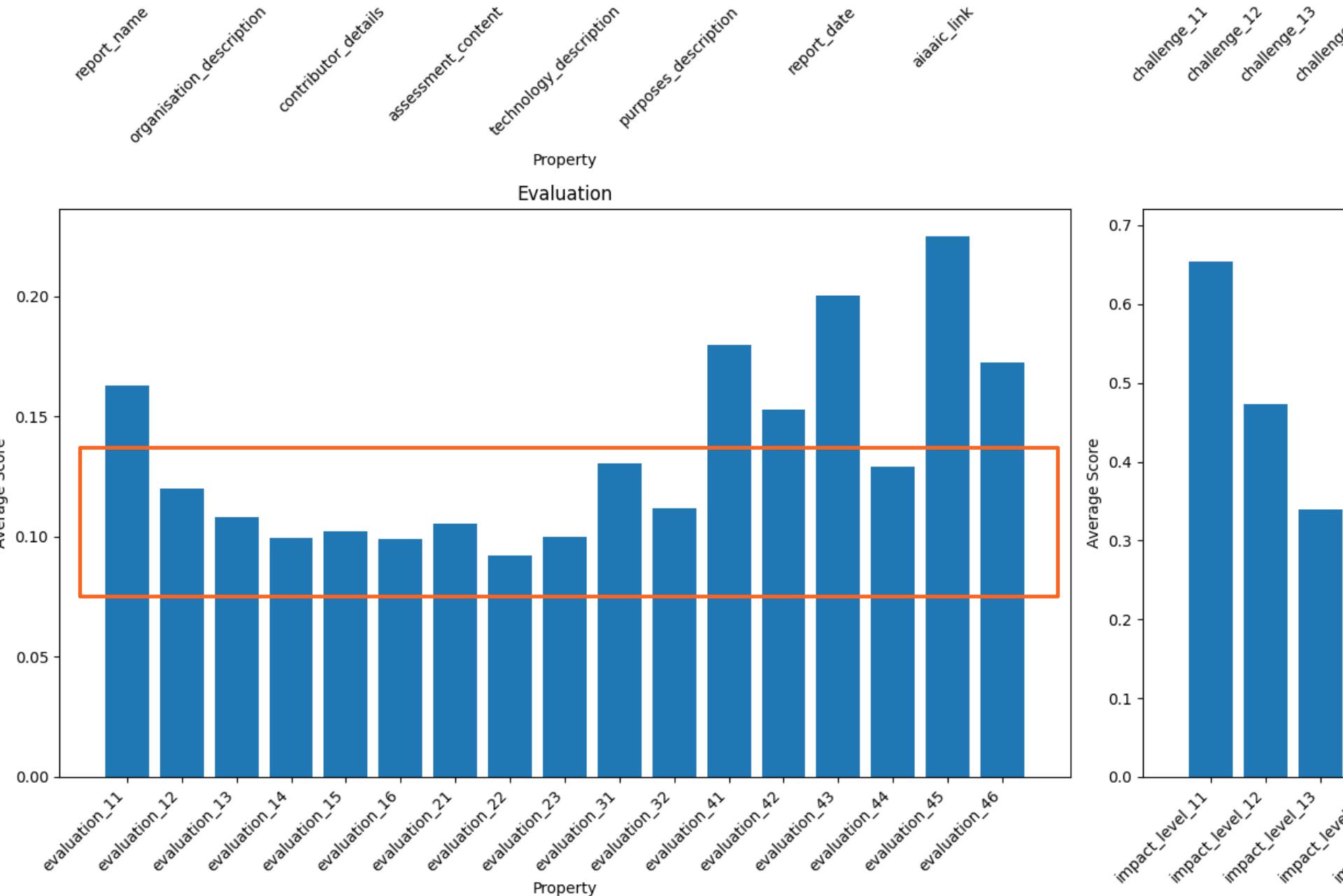
Evaluation Performance Analysis

Most of the challenge are extracted successfully. The low similarity scores may because of the low performance of GPT4o. The low performance of GPT4o may be result of the low impact level in some challenges' impact level. The instances from GPT don't have the specific impact levels like 13.



Evaluation Performance Analysis

For the evaluation, the scores are low and this is expected.



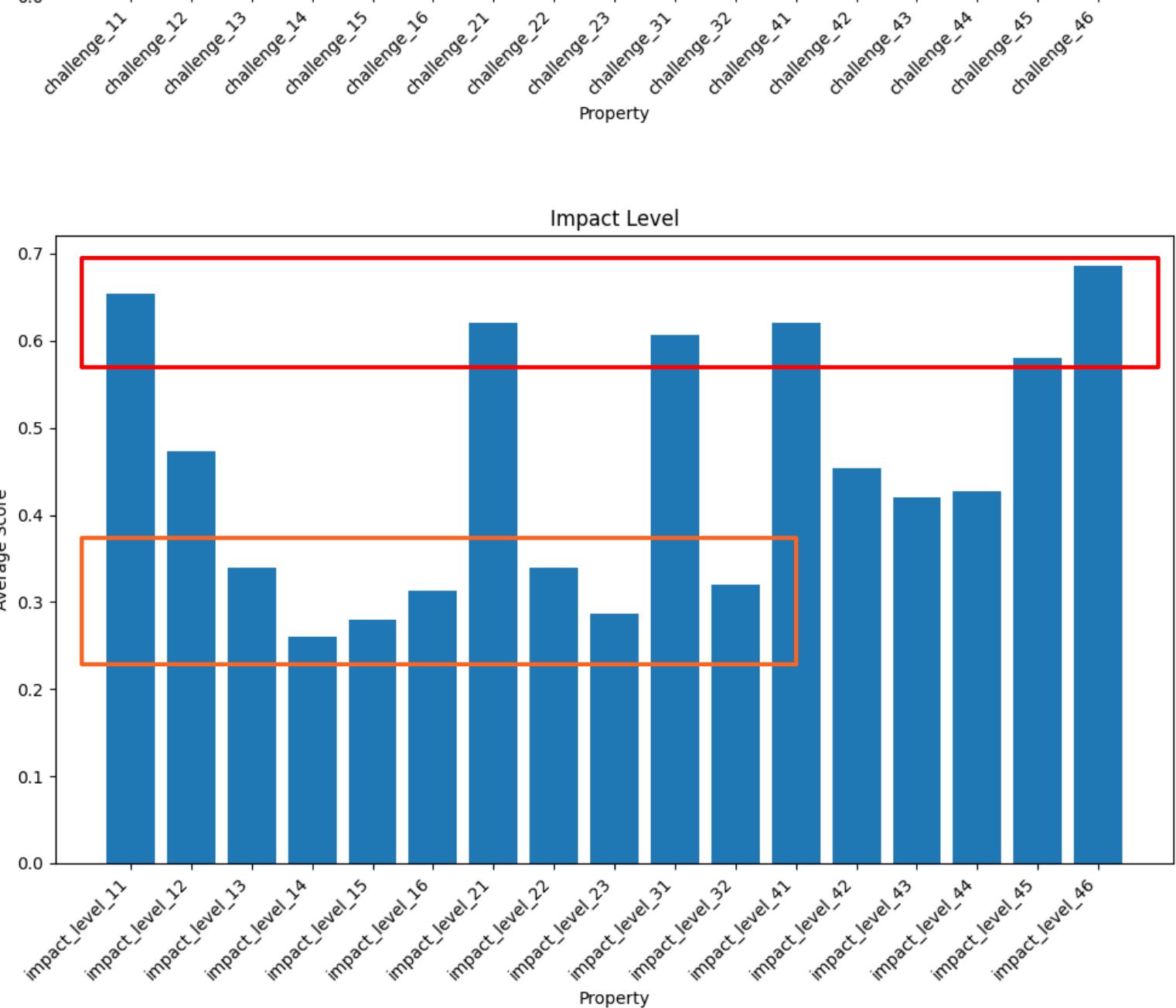
Evaluation

Performance

Analysis

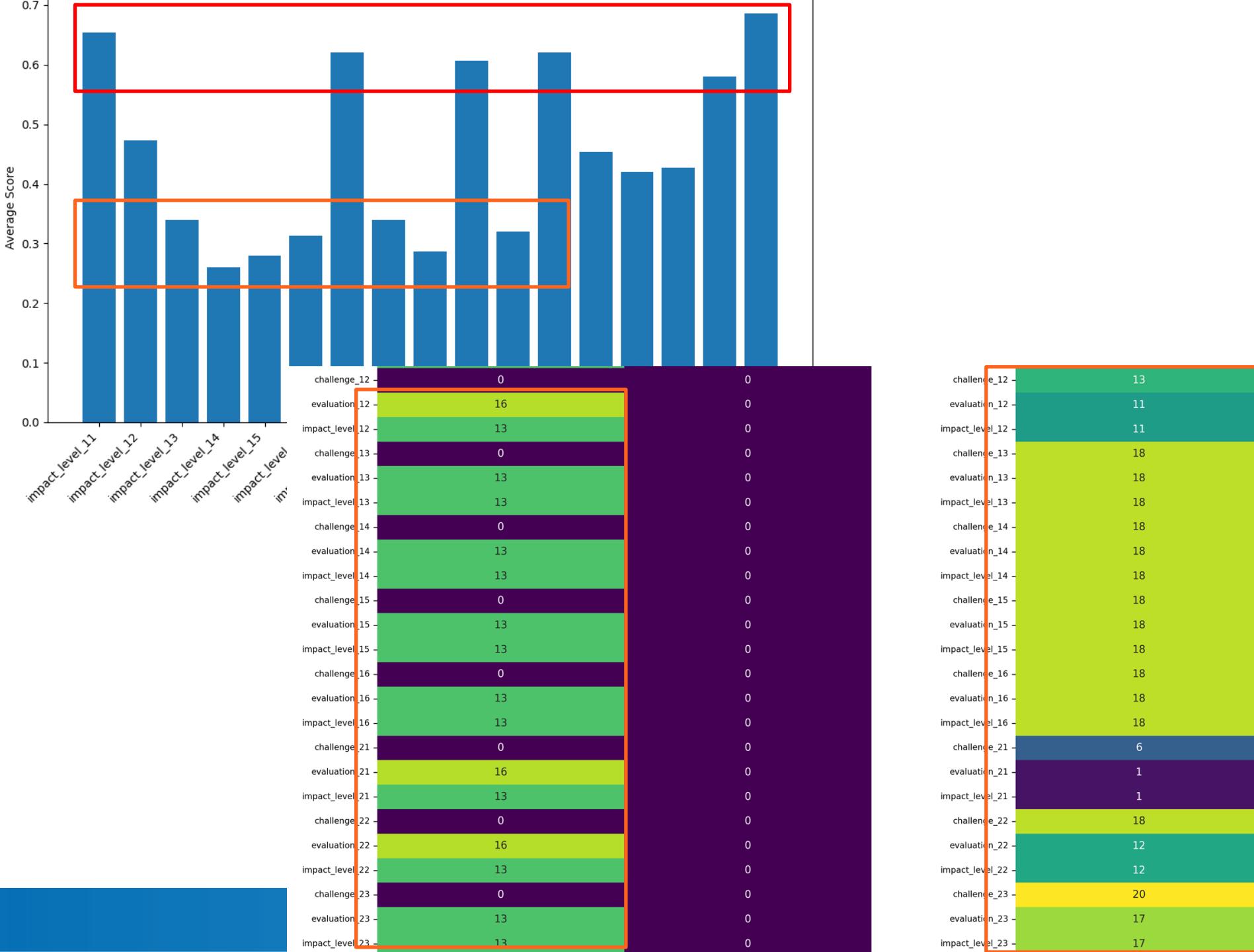
The overall similarity of the Impact Level is the best. Which means that in these context, different LLMs can still have similar result.

This make usage of LLMs in AI Impact Assessment Domain and to face EU AI Act possible.



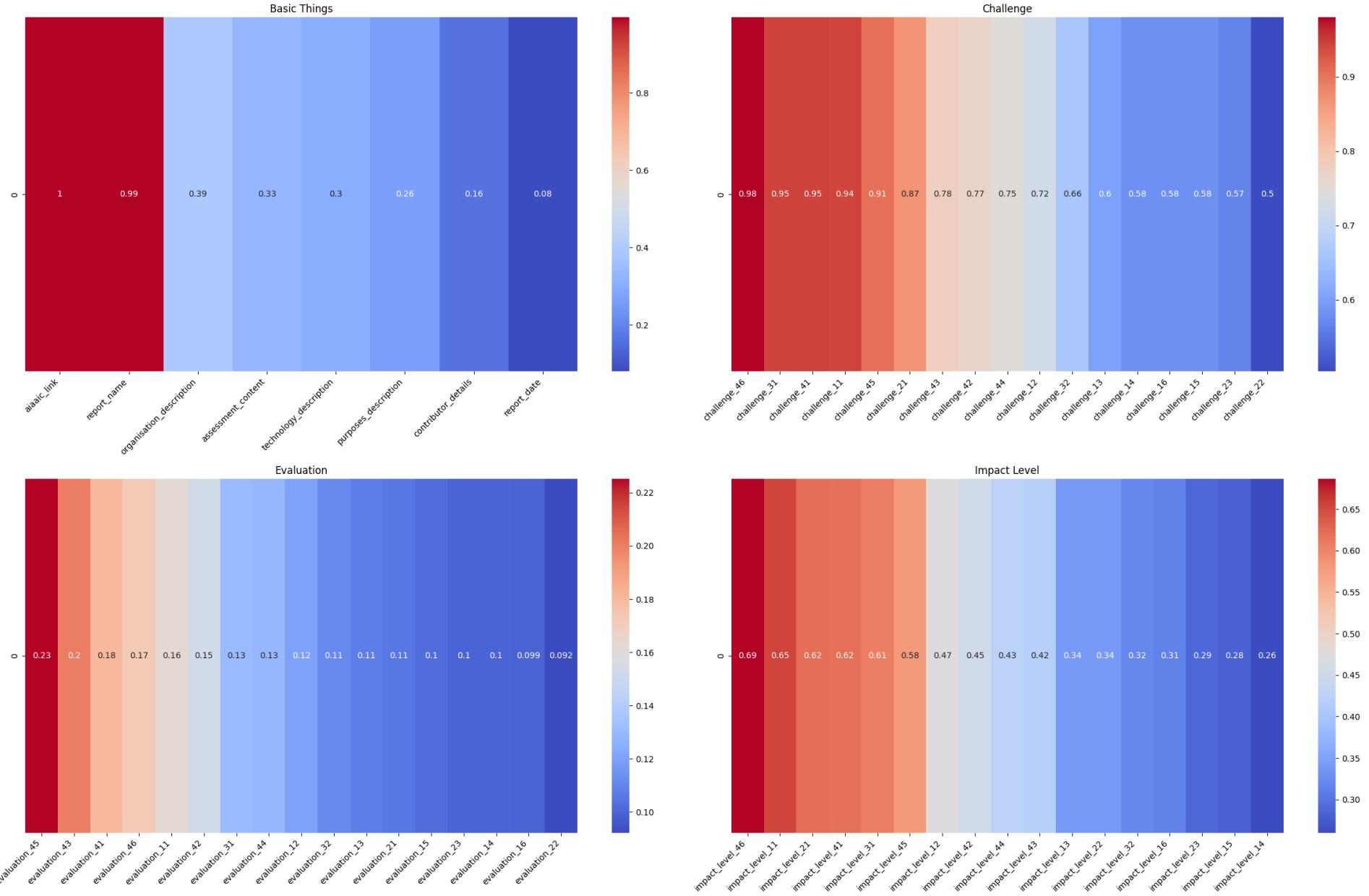
Evaluation Performance Analysis

The low performance of GPT4o may be result of the low impact level in some challenges' impact level. The instances from GPT don't have the specific impact levels like 13.



Evaluation Performance Analysis

Average Similarity Scores Heatmaps for AIAAIC Properties by Category



The overall similarity of the Impact Level is the best.

Evaluation

Performance

Analysis

We can see that some incident have very high scores. In these case, the LLM is not lazy.

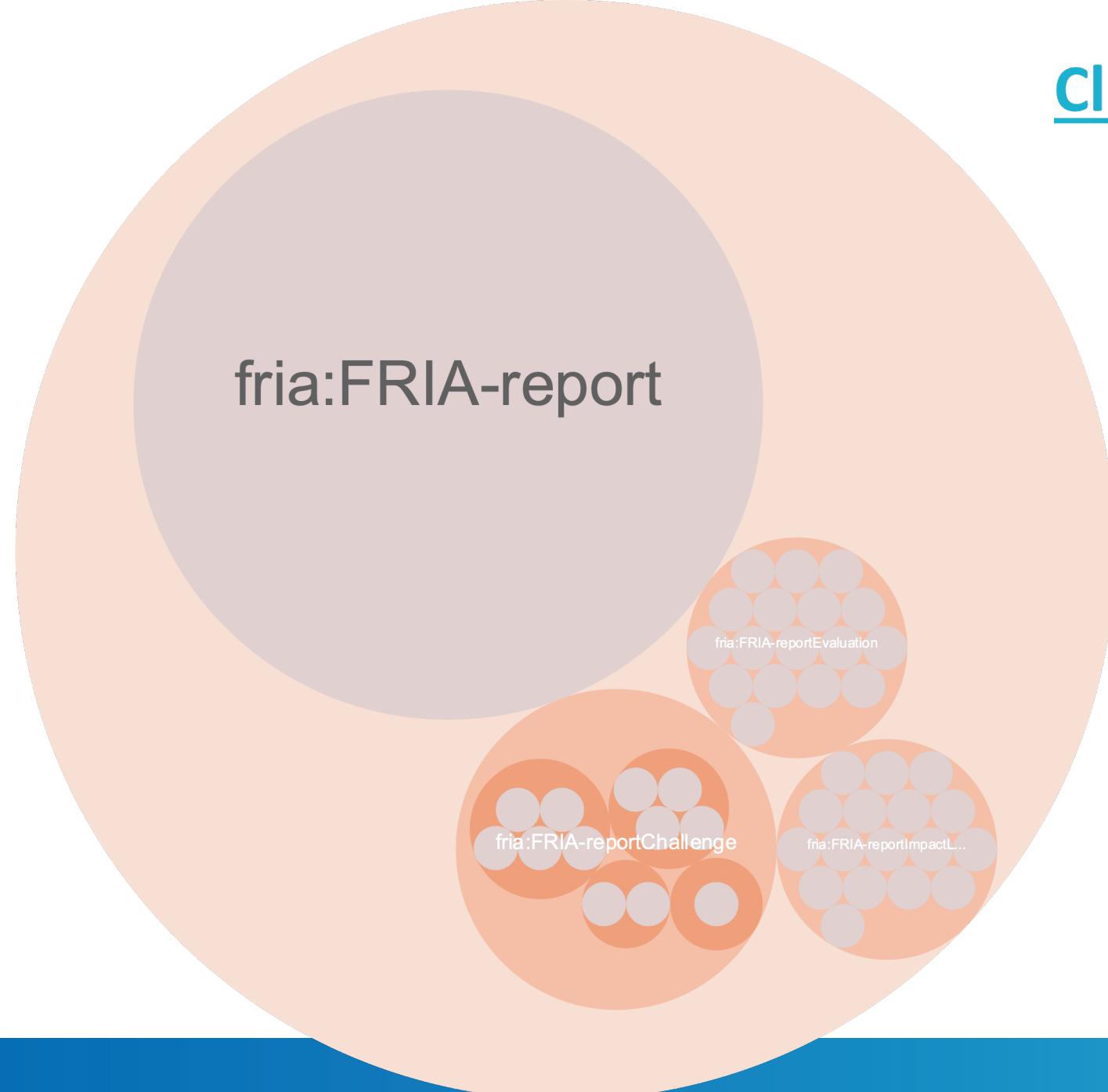


Research Questions

1. How can we integrate **Fundamental Rights Impact Assessment (FRIA)** into existing ontological frameworks such as **AI Risk Ontology (AIRO)**, **Vocabulary of AI Risks (VAIR)**, and **Common Impact Data Standard (CIDS)** to create a more comprehensive ontological structure for impact assessment?
2. To what extent can **Large Language Models (LLMs)** be effectively utilized to populate Fundamental Rights Impact Assessment (FRIA) reports and related ontologies, thereby assisting in the completion of AI impact assessments?

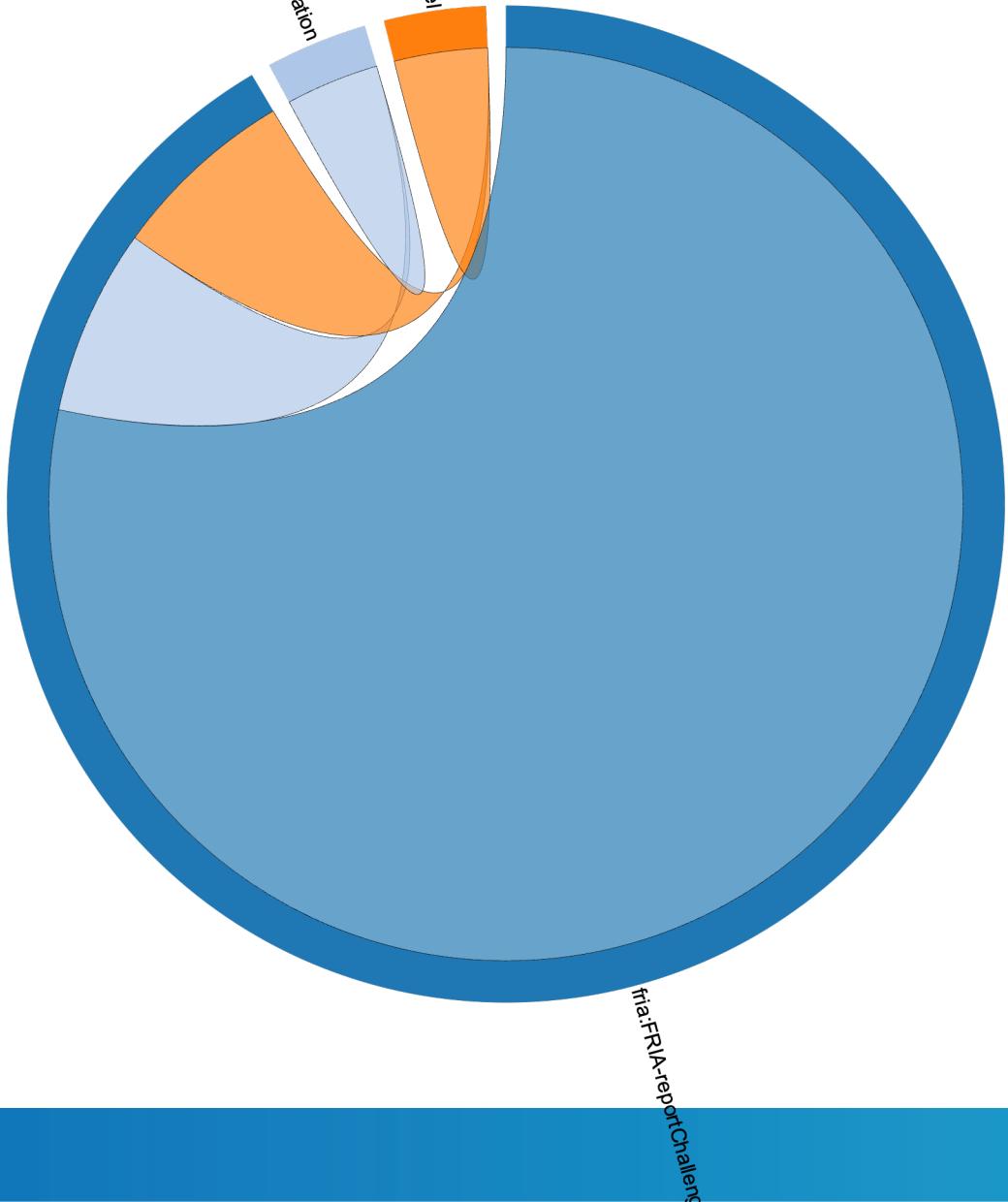
GraphDB Visualizations

Class Hierarchy

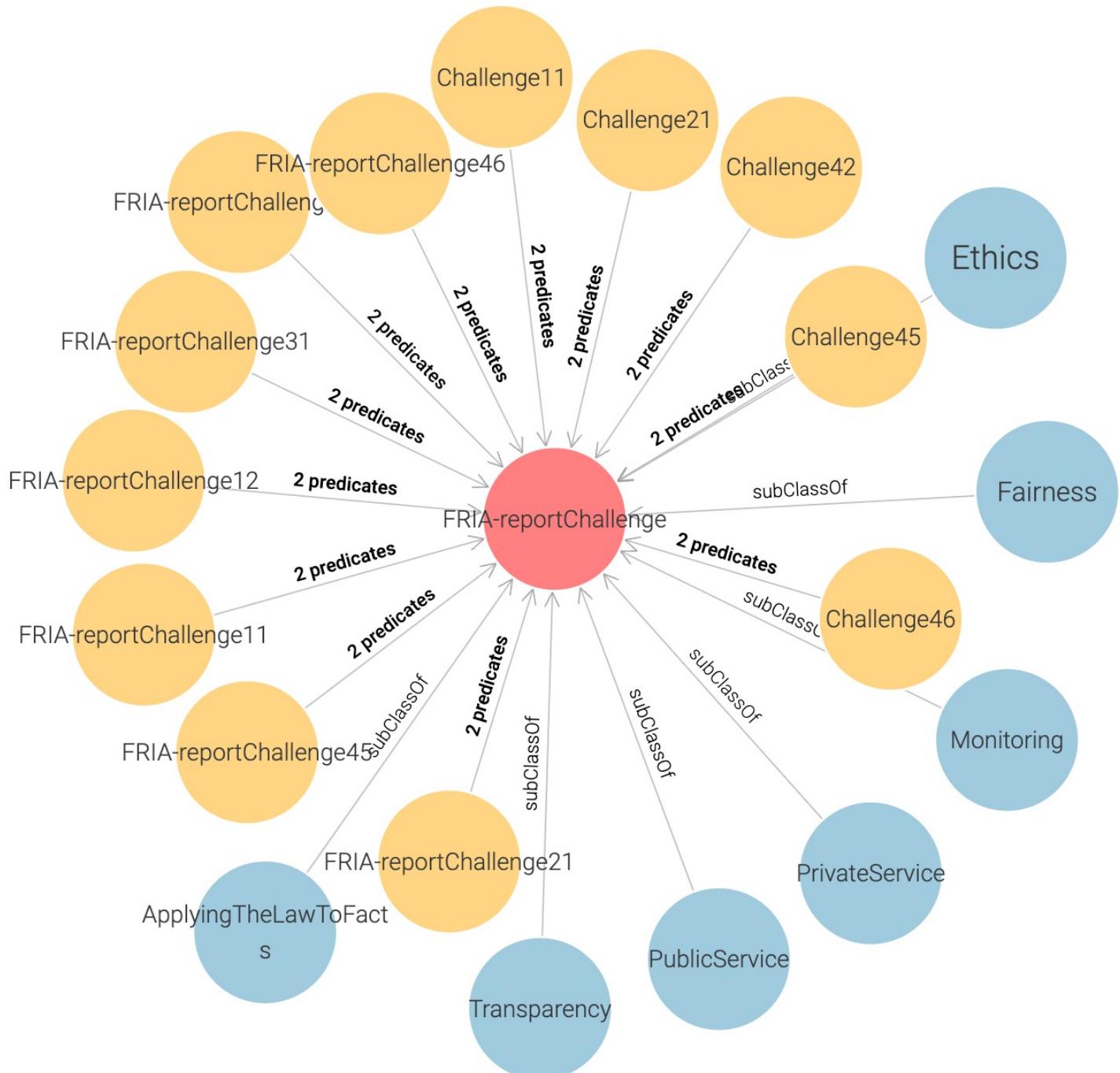
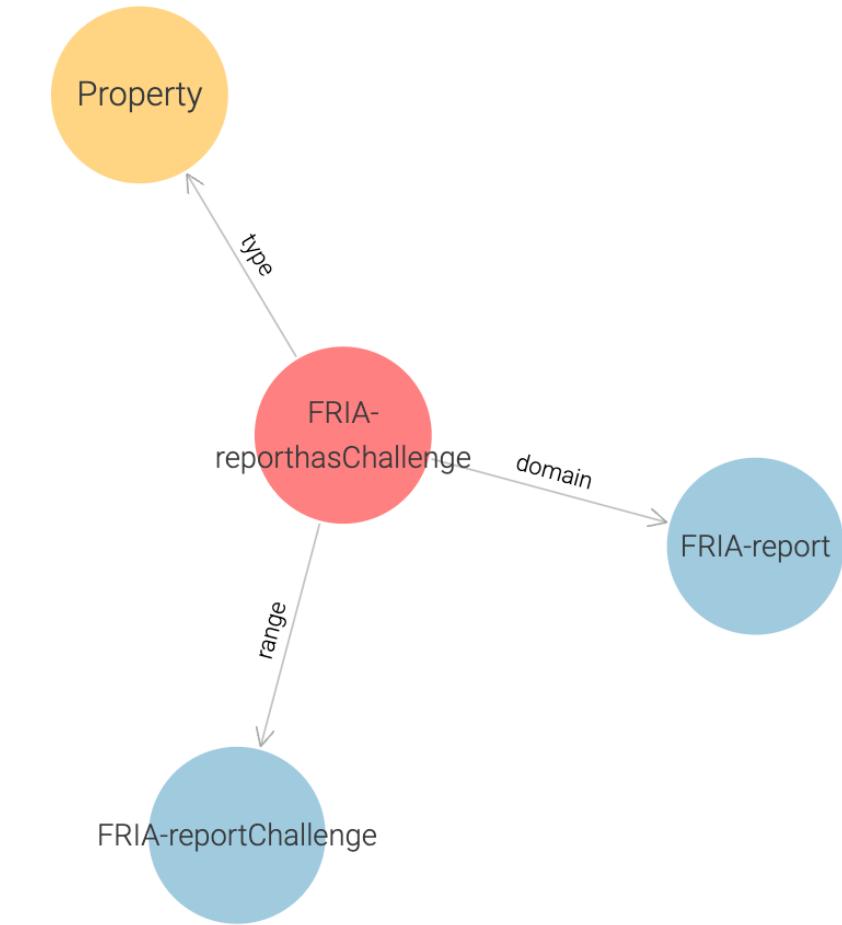


GraphDB Visualizations

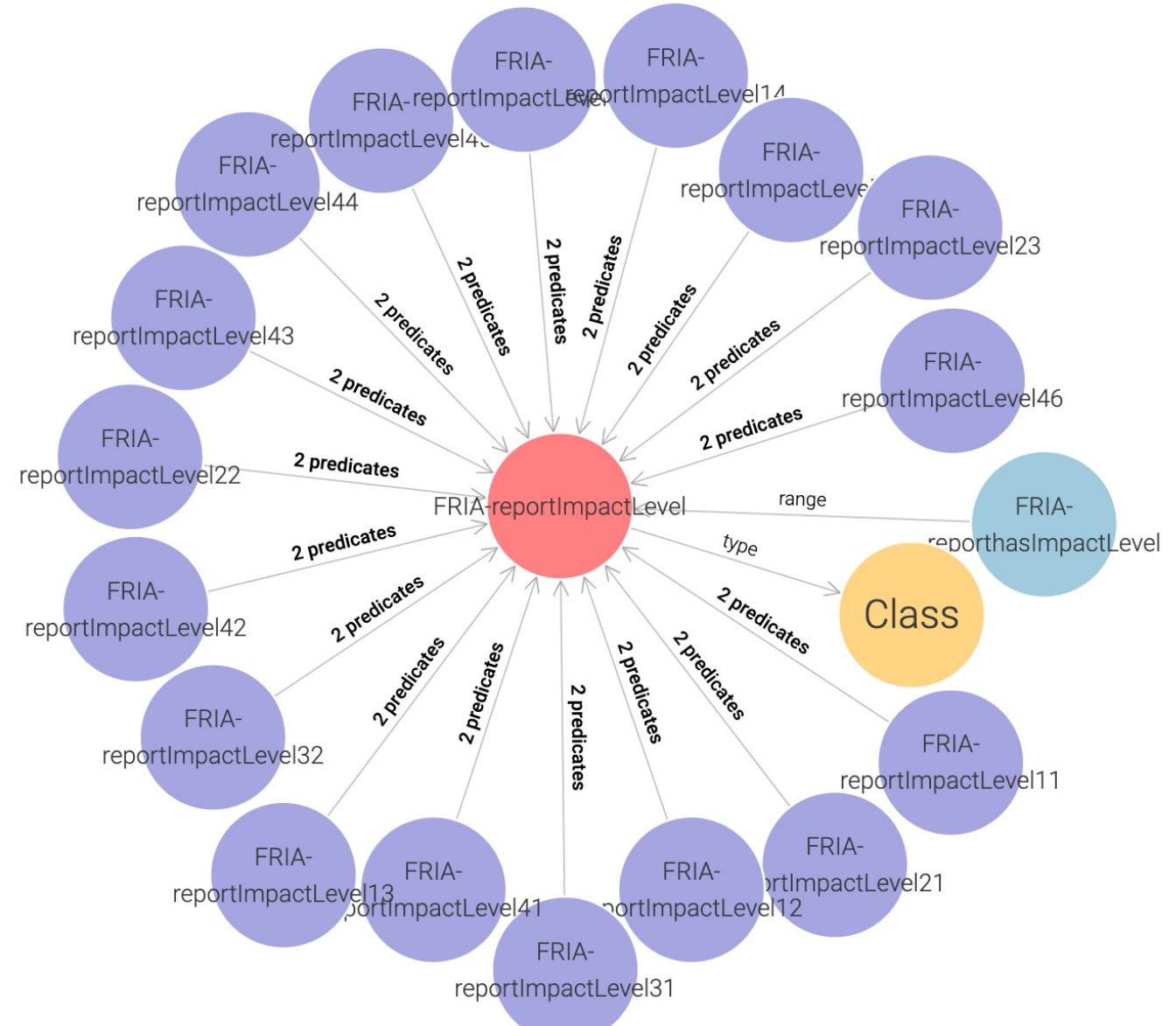
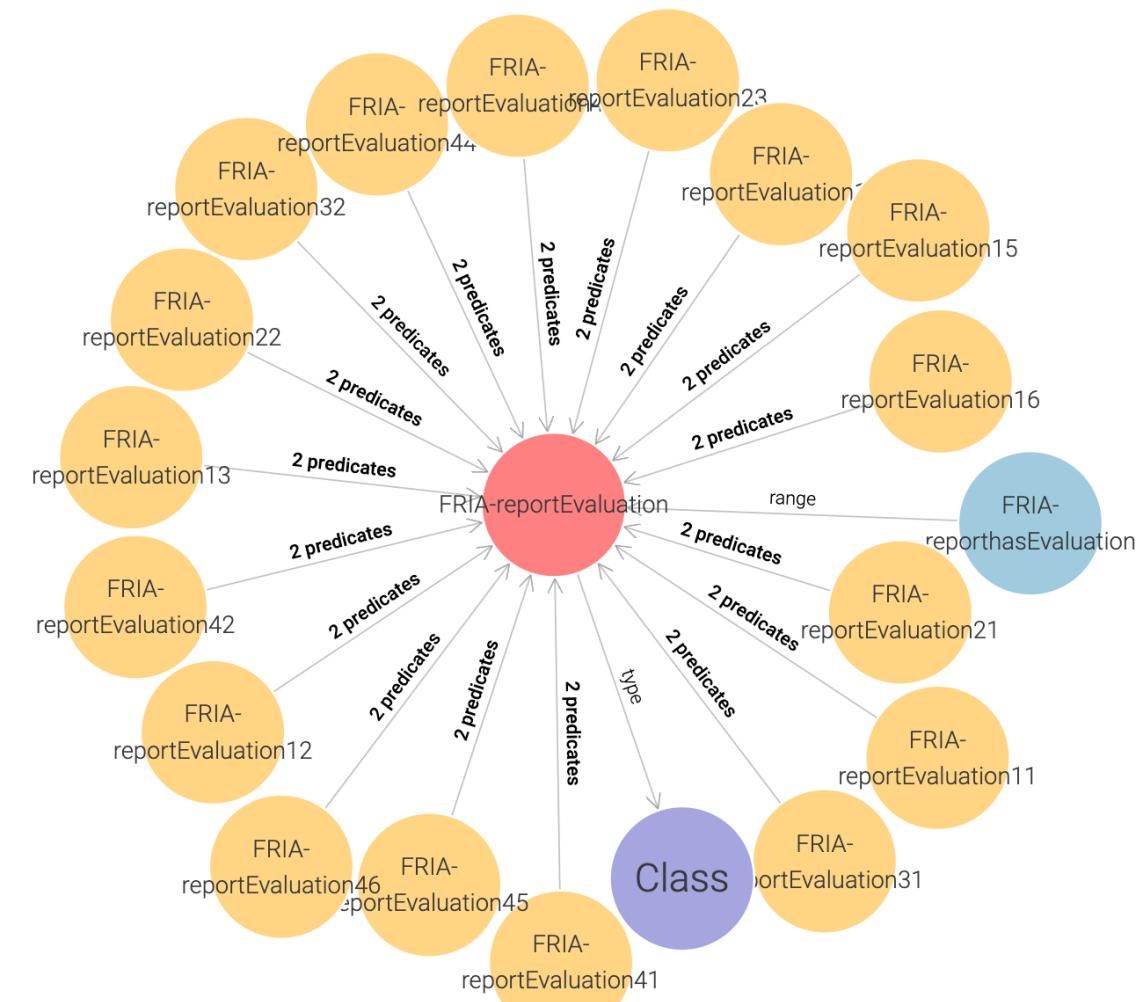
Dependencies



GraphDB Visualizations



GraphDB Visualizations



GraphDB SPARQL

SPARQL Query & Update ?

Editor only Editor and results Results only □

```
Unnamed × fria:FRIA-reportChallenge × fria:FRIA-reportImpactLevel × fria:FRIA-reportChallenge × fria:FRIA-reportEvaluation × Unnamed × +  
1 PREFIX onto: <http://www.ontotext.com/>  
2 SELECT ?s ?p ?o  
3 WHERE {  
4     ?s a <http://www.example.org/fria-report#FRIA-reportEvaluation> .  
5     ?s ?p ?o .  
6 }  
7
```

Save Folder Link More Run keyboard shortcuts

Table Raw Response Pivot Table Google Chart

Download as ▼ 1 2 >

Filter query results			
Showing results from 1 to 1,000 of 1,059. Query took 0.1s, moments ago.			
	s	p	o
1	fria:FRIA-reportEvaluation11	rdf:type	rdfs:Class
2	fria:FRIA-reportEvaluation11	rdf:type	fria:FRIA-reportEvaluation
3	fria:FRIA-reportEvaluation11	rdfs:subClassOf	fria:FRIA-reportEvaluation
4	fria:FRIA-reportEvaluation11	rdfs:subClassOf	fria:FRIA-reportEvaluation11
5	fria:FRIA-reportEvaluation11	fria:hasEvaluationContent	"This is the evaluation content for FRIA-reportEvaluation11."
6	fria:FRIA-reportEvaluation11	fria:hasEvaluationContent	"The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision."
7	fria:FRIA-reportEvaluation11	fria:hasEvaluationContent	"The AI-generated songs do not explicitly communicate that they are artificial creations, potentially misleading listeners into believing they are authentic works by Stefanie Sun."

GraphDB

SPARQL

Unnamed × fria:FRIA-reportChallenge × fria:FRIA-reportImpactLevel × fria:FRIA-reportChallenge × fria:FRIA-reportEvaluation × Unnamed ×

fria:FRIA-reportImpactLevel × Unnamed × 

```

1 PREFIX onto: <http://www.ontotext.com/>
2 PREFIX fria: <http://www.example.org/fria-report#>
3 SELECT ?s ?p ?o
4 WHERE {
5   ?s a fria:FRIA-reportImpactLevel .
6   ?s ?p ?o .
7   FILTER(?s IN (fria:FRIA-reportImpactLevel46, fria:ImpactLevel46))
8 }
9

```

 Table  Raw Response  Pivot Table  Google Chart 

Filter query results Showing results from 1 to 14 of 14. Query took 0.2s, moments ago.

	s	p	o
1	fria:FRIA-reportImpactLevel46	rdf:type	rdfs:Class
2	fria:FRIA-reportImpactLevel46	rdf:type	fria:FRIA-reportImpactLevel
3	fria:FRIA-reportImpactLevel46	rdfs:subClassOf	fria:FRIA-reportImpactLevel
4	fria:FRIA-reportImpactLevel46	rdfs:subClassOf	fria:FRIA-reportImpactLevel46
5	fria:FRIA-reportImpactLevel46	fria:hasImpactLevelContent	"This is the impact level content for FRIA-reportImpactLevel46."
6	fria:FRIA-reportImpactLevel46	fria:hasImpactLevelContent	"High"
7	fria:FRIA-reportImpactLevel46	fria:hasImpactLevelContent	"Very High"
8	fria:FRIA-reportImpactLevel46	fria:hasImpactLevelContent	"Medium"
9	fria:FRIA-reportImpactLevel46	fria:hasImpactLevelContent	"High impact due to lack of impact assessment procedures."
10	fria:FRIA-reportImpactLevel46	fria:hasImpactLevelContent	"The impact level content for FRIA-reportImpactLevel46."
11	fria:FRIA-reportImpactLevel46	fria:hasImpactLevelContent	"Without impact assessments, potential risks to data protection are not proactively addressed."
12	fria:FRIA-reportImpactLevel46	fria:hasImpactLevelContent	"High impact as it risks non-compliance with data protection regulations."
13	fria:ImpactLevel46	rdf:type	fria:FRIA-reportImpactLevel
14	fria:ImpactLevel46	fria:hasImpactLevelContent	"This is the impact level content for FRIA-reportImpactLevel46."

GraphDB SPARQL

SPARQL Query & Update

[Editor only](#)[Editor and results](#)[Results only](#)

Unnamed × fria:FRIA-reportChallenge × fria:FRIA-reportImpactLevel × fria:FRIA-reportChallenge × fria:FRIA-reportEvaluation × Unnamed ×

```
fria:FRIA-reportImpactLevel × Unnamed × fria:FRIA-report × Unnamed × +  
1 PREFIX onto: <http://www.ontotext.com/>  
2 PREFIX fria: <http://www.example.org/fria-report#>  
3 SELECT ?s ?p ?o  
4 WHERE {  
5   ?s a fria:FRIA-report .  
6   ?s ?p ?o .  
7   FILTER(?s IN (fria:Report21, fria:FRIA-report-021))  
8 }  
9
```

[Run](#)

keyboard shortcuts

[Table](#) [Raw Response](#) [Pivot Table](#) [Google Chart](#)[Download as](#) ▾

Showing results from 1 to 18 of 18. Query took 0.1s, moments ago.

	s	p	o
1	fria:Report21	rdf:type	fria:FRIA-report
2	fria:Report21	fria:hasReportName	"Al Stefanie Sun (AI孙燕姿)"
3	fria:Report21	fria:hasOrganisationPositionDescription	" "
4	fria:Report21	fria:hasContributorDetails	"Bilibili"
5	fria:Report21	fria:hasAssessmentContent	"Al-cloned songs in the name and voice of Singapore-based Mandopop singer Stefanie Sun have gone viral on China's most popular video platform Bilibili, raising questions about copyright and jobs in the music industry. The videos were generated by so-vits-svc fork, an open source software that enables anyone to train their own AI model to speak in any

GraphDB SPARQL

FRIA-reportEvaluation11

Source: <http://www.example.org/fria-report#FRIA-reportEvaluation11>

subject	predicate	object	context	all	Explicit only	Show Blank Nodes	Download as	Visual graph
	subject		predicate		object		context	
1	fria:FRIA-reportEvaluation11		fria:hasEvaluationContent		"4 Little Trees does not explicitly communicate to students that their emotions and performance are being analyzed by an AI system."		http://www.ontotext.com/explicit	
2	fria:FRIA-reportEvaluation11		fria:hasEvaluationContent		"7-Eleven did not adequately inform customers that their facial images were being captured and processed by an AI system for survey validation and demographic profiling."		http://www.ontotext.com/explicit	
3	fria:FRIA-reportEvaluation11		fria:hasEvaluationContent		"7-Eleven did not provide adequate notice to customers that their facial images were being used to validate survey responses. This lack of transparency means that customers were unaware of the AI system's role in the process."		http://www.ontotext.com/explicit	
4	fria:FRIA-reportEvaluation11		fria:hasEvaluationContent		"AWS Panorama does not explicitly communicate to employees that their actions are being monitored and analyzed by an AI system."		http://www.ontotext.com/explicit	
5	fria:FRIA-reportEvaluation11		fria:hasEvaluationContent		"AccessiBe does not clearly communicate to users that the website's accessibility features are being provided by an AI-powered overlay, which may lead to confusion when issues arise."		http://www.ontotext.com/explicit	
6	fria:FRIA-reportEvaluation11		fria:hasEvaluationContent		"Adobe did not clearly communicate to Adobe Stock contributors that their content was being used to train the Firefly AI model, leading to surprise and concern when the practice was discovered."		http://www.ontotext.com/explicit	
7	fria:FRIA-reportEvaluation11		fria:hasEvaluationContent		"Adobe did not clearly communicate to		http://www.ontotext.com/explicit	

FRIA-reportEvaluation11

Source: <http://www.example.org/fria-report#FRIA-reportEvaluation11>

	subject	predicate	object	context	all	Explicit only	Show Blank Nodes	Download as	Visual graph
	subject			predicate			object		
1	fria:FRIA-reportEvaluation11			fria:hasEvaluationContent			"4 Little Trees does not explicitly communicate to students that their emotions and performance are being analyzed by an AI system."	http://www.ontotext.com/explicit	
2	fria:FRIA-reportEvaluation11			fria:hasEvaluationContent			"7-Eleven did not adequately inform customers that their facial images were being captured and processed by an AI system for survey validation and demographic profiling."	http://www.ontotext.com/explicit	
3	fria:FRIA-reportEvaluation11			fria:hasEvaluationContent			"7-Eleven did not provide adequate notice to customers that their facial images were being used to validate survey responses. This lack of transparency means that customers were unaware of the AI system's role in the process."	http://www.ontotext.com/explicit	
4	fria:FRIA-reportEvaluation11			fria:hasEvaluationContent			"AWS Panorama does not explicitly communicate to employees that their actions are being monitored and analyzed by an AI system."	http://www.ontotext.com/explicit	
5	fria:FRIA-reportEvaluation11			fria:hasEvaluationContent			"AccessiBe does not clearly communicate to users that the website's accessibility features are being provided by an AI-powered overlay, which may lead to confusion when issues arise."	http://www.ontotext.com/explicit	
6	fria:FRIA-reportEvaluation11			fria:hasEvaluationContent			"Adobe did not clearly communicate to Adobe Stock contributors that their content was being used to train the Firefly AI model, leading to surprise and concern when the practice was discovered."	http://www.ontotext.com/explicit	
7	fria:FRIA-reportEvaluation11			fria:hasEvaluationContent			"Adobe did not clearly communicate to	http://www.ontotext.com/explicit	

Conclusion

- Successfully find a possible way to face the need of EU AI Act.
- Developed the FRIA ontology based on the FRIA report template.
- Established the relationship with FIRA and CIDS, AIRO, VAIR.
- Explored a new way to do the Impact Assessment by using the prompt engineering with the latest LLMs.
- Discovered a way to evaluate the performance of different LLMs when facing the same context.
- **Claude 3.5 Sonnet** has better performance than the ChatGPT 4o.
- Successfully answer both the Research Question 1 and 2.

Future Works

- Develop more relationships than now.
- Update the prompt to prevent the “lazy LLM”.
- Have some general way to automatically grab the data and update the ontology instances in server.
- Make everyone can see the result from internet.
- Make the evaluation and impact level more accurate.



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Supervisor:

Dr. David Lewis

Dr. Subrahmanyam Murala

Thank You!

Kaiyu Chen (23330889)
MSc in Computer Science - Intelligent Systems