Sim-X-HCOME+ PD DOMAIN

Results from Sim-X-HCOME+ methodology in PD domain.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| LLM-Method | Evaluation | | | | | Metrics | | | | |
| **OOPS**  **Results** | **Ontology Reusability** | **Consistent**  **(Pellet Reasoner)** | **Syntactical errors** | **Opened**  **by Protege** | Axiom | Classes | Object Properties | Data Properties | Individuals |
| *ChatGPT-4* | P36 | - | Yes | No | Yes | 71 | 17 | 11 | 5 | 0 |
| *ChatGPT-3.5* | P36 | - | Yes | No | Yes | 205 | 21 | 18 | 6 | 15 |
| *Gemini* | P36 | - | Yes | Yes | Yes | 36 | 22 | 13 | 1 | 0 |
| *Claude* | P36 | - | Yes | No | Yes | 96 | 29 | 11 | 7 | 0 |

**Table .** Evaluation metrics on Sim-X-HCOME+ generated ontologies in PD domain (classes).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Method | Number of  Classes | True  Positives | False  Positives | False  Negatives | Precision | Recall | F-1 Score |
| *Gold ontology* | 41 |  |  |  |  |  |  |
| *ChatGPT-4* | 17 | 9 | 8 | 32 | 52% | 21% | 0,3103 |
| *ChatGPT-3.5* | 21 | 14 | 7 | 27 | 66% | 34% | 0,4516 |
| *Gemini* | 22 | 15 | 7 | 26 | 68% | 36% | 0,4761 |
| *Claude* | 24 | 12 | 12 | 29 | 50% | 29% | 0,3692 |

**Table .** Evaluation metrics on Sim-X-HCOME+ generated ontologies in PD domain (object properties).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Method | Number of  Classes | True  Positives | False  Positives | False  Negatives | Precision | Recall | F-1 Score |
| *Gold ontology* | 29 |  |  |  |  |  |  |
| *ChatGPT-4* | 13 | 1 | 12 | 28 | 7,6% | 3,4% | 0,0475 |
| *ChatGPT-3.5* | 18 | 1 | 17 | 28 | 5,5% | 3,4% | 0,0425 |
| *Gemini* | 11 | 1 | 10 | 28 | 9% | 3,4% | 0,05 |
| *Claude* | 11 | 1 | 10 | 28 | 9% | 3,4% | 0,05 |

**Table .** Evaluation metrics on Sim-X-HCOME+ generated ontologies in PD domain (NL2SWRL).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Method | Number of  Atoms | True  Positives | | False  Positives | | False  Negatives | | Precision (%) | | Recall (%) | | F-1 Score | |
| *Gold ontology* | 8 | SC | LC | SC | LC | SC | LC | SC | LC | SC | LC | SC | LC |
| *ChatGPT-4* | 13 | 0 | 3 | 13 | 10 | 8 | 5 | 0 | 23 | 0 | 27 | 0 | 0,125 |
| *ChatGPT-3.5* | 17 | 1 | 3 | 16 | 14 | 7 | 5 | 5 | 17 | 12,5 | 3 | 0,08 | 0,111 |
| *Gemini* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Claude* | 12 | 0 | 5 | 12 | 7 | 8 | 3 | 0 | 41,6 | 0 | 28,4 | 0 | 0,2 |