Enabling Personal Communication for Voice-Based Health Assistants in Geriatric Care

Murilo Bellatini

08.08.2023, Master Thesis Kick-off

Chair of Software Engineering for Business Information Systems (sebis)
Department of Computer Science
School of Computation, Information and Technology (CIT)
Technical University of Munich (TUM)
wwwmatthes.in.tum.de

AI-Supported Care: The Current Solution of a Digital Health Assistant

“Oma” Hilda

- 73 years old
- Rollator user
- Requires medical care

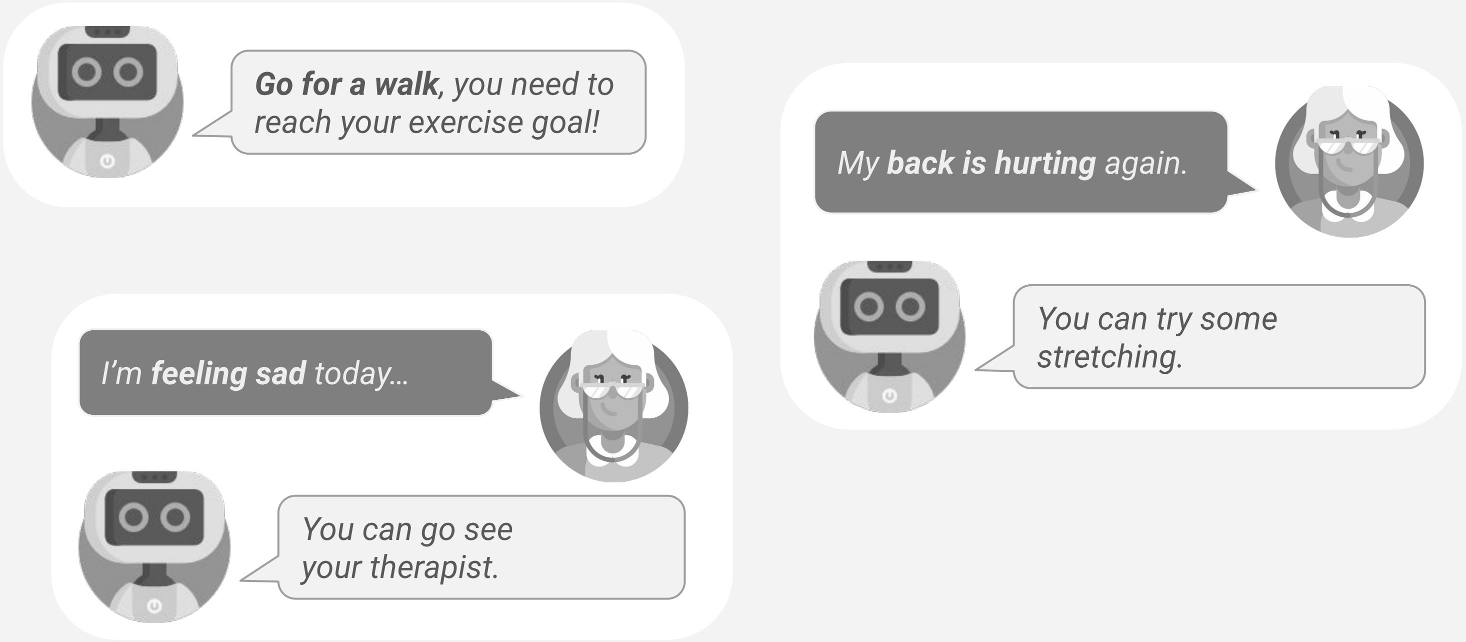
Digital Health Assistant

- Voice-based
- Task-based

Task-Based Assistants: Lack of Personal Communication in Healthcare

Most Health Assistants

1. Lack personal touch
2. Feel robotic



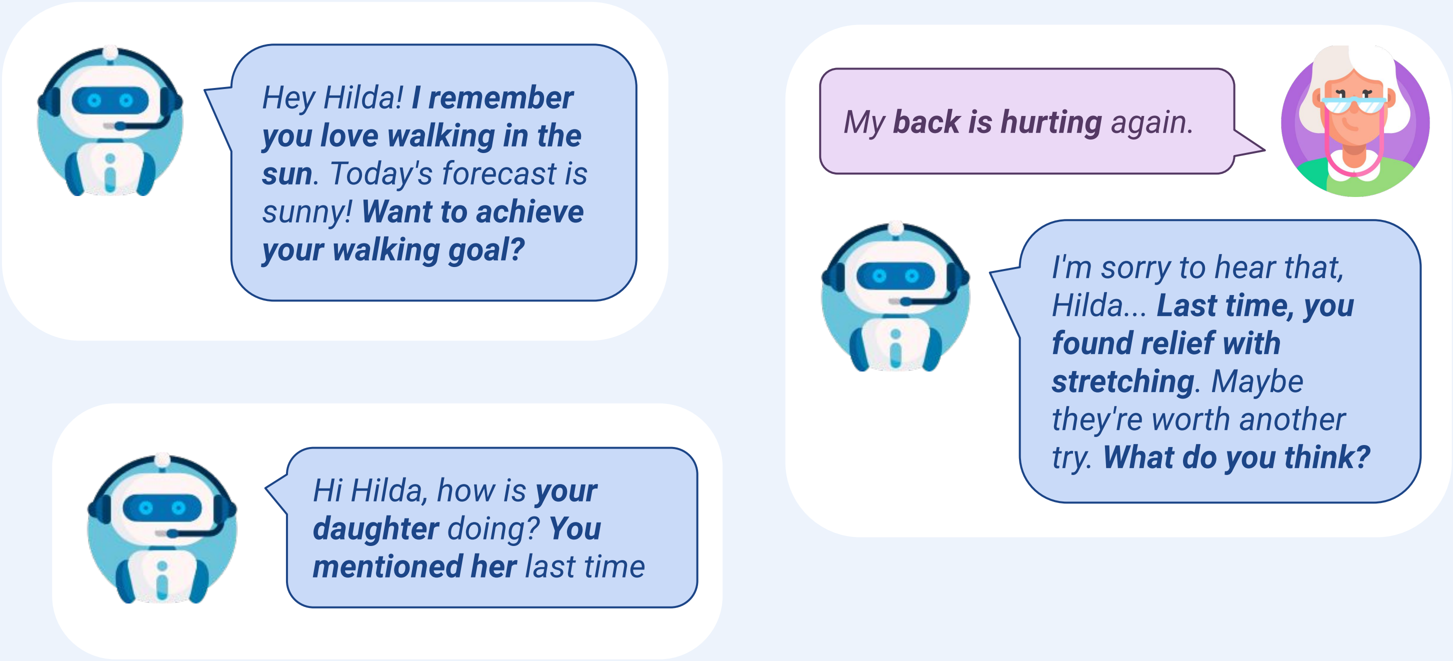
Outcome

- Suboptimal user engagement

Enhancing User Experience: Personalized and Engaging Health Assistants

Ideal Agent: Personalized and Engaging

1. Remembers user specifics
2. Tailors responses
3. Stimulates user engagement



Expected Outcome

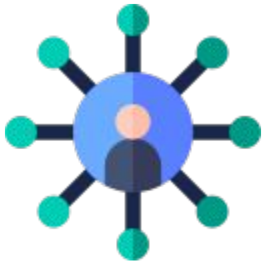
- Increased user satisfaction
- Enhanced user engagement

Outline

Our Approach: Harnessing User-centric Knowledge for Engaging Dialogue

High-level concept:

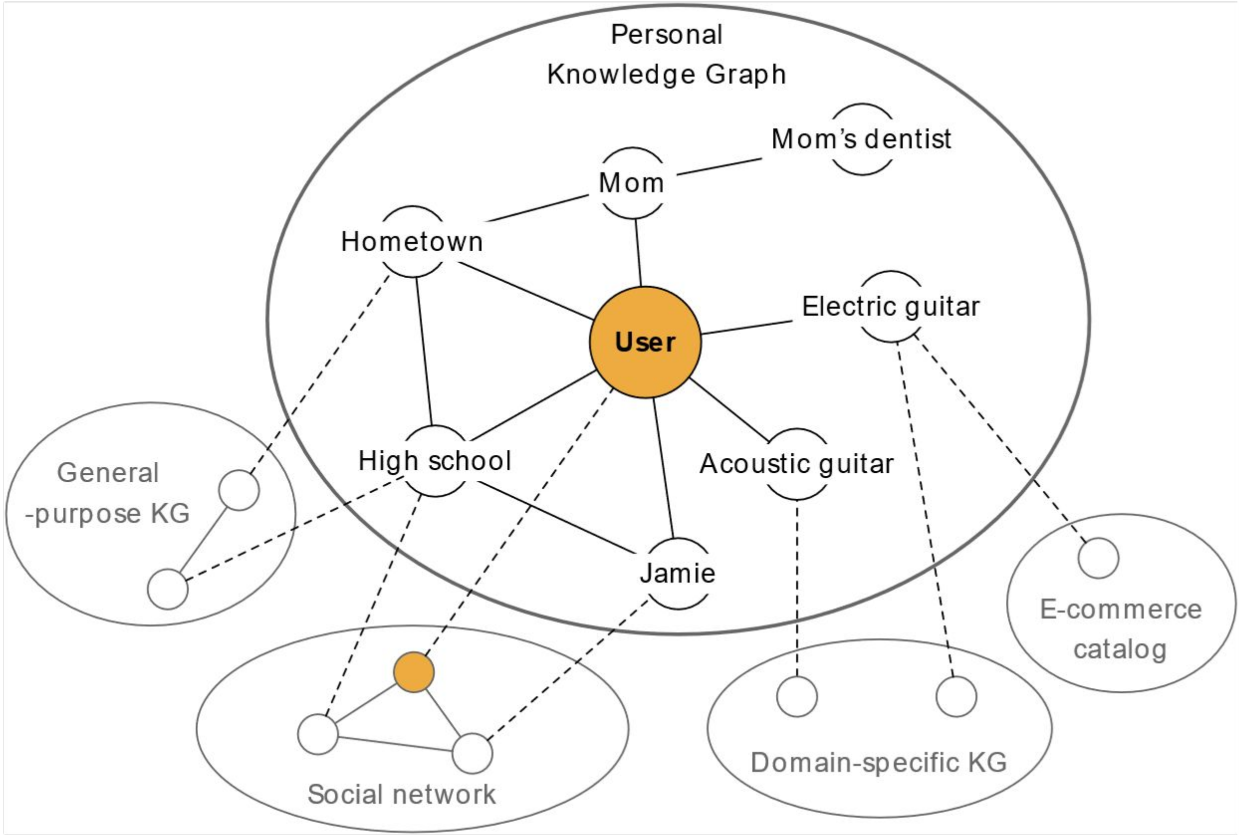
Stimulate Dialogue
Promote continuous sharing of personal information by users



Curate Knowledge
Convert user chat history into related entities

Personalize Interactions
Use knowledge to personalize interactions

Personal Knowledge Graphs: Enabling Personalized Recommendations



Definition *Individual knowledge representation and relationships (spiderweb).*

Why? *Tap into chat history in a structured manner for recommendations.*

Google Research

According to **Balog and Kenter** [1], PKGs:

1	*Enable personalized search and recommendations*
2	*Support personal information management*
3	*Facilitate personal digital assistant services*

[1] Balog, K., and Kenter, T. 2019. Personal Knowledge Graphs: A Research Agenda. In Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval (pp. 217–220). Association for Computing Machinery.

Main Pillars: Personal Data Mapping & Contextual Recommendations

Research focus is two-fold:

		Pillar	References
1		**Dynamic Personal Knowledge Graph Construction**	GraphWOZ [2], PKG [3], PKG Recommendations [4]
2		**Tailored Conversation Topic Recommender**	PKG Recommendations [4], RAG [5], Prompt Template [6]

Further References: [Click Here]

[2] Nicholas Thomas Walker, Stefan Ultes, and Pierre Lison. (2022). GraphWOZ: Dialogue Management with Conversational Knowledge Graphs.
[3] Krisztian Balog and Tom Kenter. Personal knowledge graphs: A research agenda. In Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval, ICTIR '19, page 217–220, New York, NY, USA, 2019. Association for Computing Machinery.
[4] Yang, Y., Lin, J., Zhang, X., and Wang, M. 2022. PKG: A Personal Knowledge Graph for Recommendation. In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 3334–3338). Association for Computing Machinery.
[5] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich K¨uttler, Mike Lewis, Wen tau Yih, Tim Rockt¨aschel, Se- bastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive nlp tasks, 2021
[6] https://python.langchain.com/en/latest/modules/prompts/prompt_templates.html

Research Questions

RQ1 What should our **data model** include for better **personalization**?

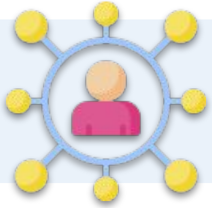
RQ2 Which **techniques** can **populate the Personal Knowledge Graph** effectively?

RQ3 How can we **integrate knowledge** for **personalized responses**?

RQ4 How can we **evaluate** system performance and effectiveness?

Outline

Literature-Based Data Model for Geriatric Communication



Q1 What should our **data model** include for better **personalization**?

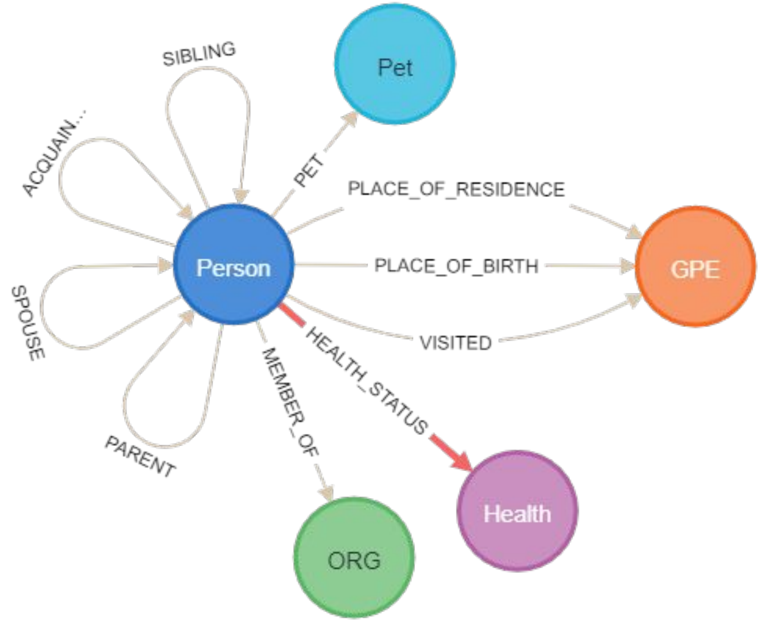
Literature Review

1. Independent Research [7,8,9,10]
2. Input from Domain Expert - Richard Paluch, Uni Siegen [11,12]

<u>**Key Insight:**</u> Tom Kitwood's Person-Centered Framework

Available Datasets

DialogRE *36 classes ⇒ mapped to 5 needs ⇒ 9 selected*



✅ Successfully reproduced paper results [13]

<u>Challenges</u>: a) Class imbalance, b) underperformance in many classes, c) absence of "no relation" class.

[7] Fukaya et al., 2009. "Education for Caregivers in Geriatric Care Facilities." Japan J. of Nursing Science, 6(2).; [8] Office et al., 2020. "Reducing Seniors' Social Isolation during COVID-19." J. of the American Medical Directors Assoc., 21(7).; [9] D'Onofrio et al., 2019. "Assistive Robots for Elderly Socialization." Aging Clinical and Experimental Research, 31.; [10] Sgorbissa et al., 2018. "Culturally Competent Robot for Elderly Care." IEEE/RSJ International Conference on Intelligent Robots and Systems.; [11] Kitwood, 2013. "Person-centered Approach in Dementia." Huber.; [12] Kitwood & Brooker, 2019. "Dementia Reconsidered Revisited." Open University Press. [13] Yu, D., Sun, K., Cardie, C., Yu, D., 2020. "Dialogue-Based Relation Extraction." arXiv preprint arXiv:2004.08056.

Knowledge Graph Construction from Chat Histories

Q2 Which **techniques** can **populate the Personal Knowledge Graph** effectively?

Methods from Literature

OpenIE [14], GraphWOZ [15], DialogRE [13]

Challenges

Differing domain specifics;
only a part of the pipeline

Solution

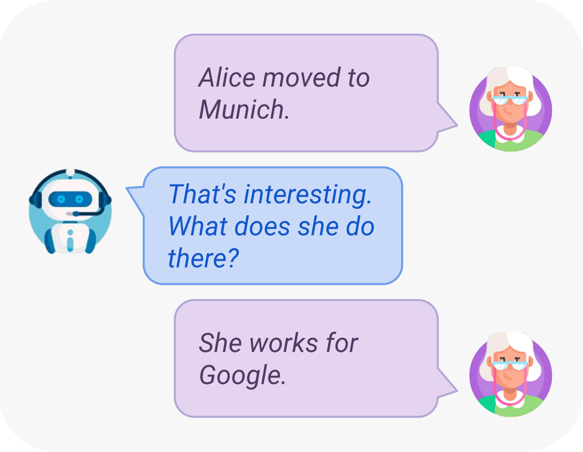
Developed custom pipeline
for relation extraction

[13] Yu, D., Sun, K., Cardie, C., Yu, D., 2020. "Dialogue-Based Relation Extraction." arXiv preprint arXiv:2004.08056.
[14] Temperoni, A., Biryukov, M., Theobald, M., 2022. "Enriching Relation Extraction with OpenIE." arXiv preprint arXiv:2212.09376.
[15] Melnyk, I., Dognin, P., Das, P., 2022. "Knowledge Graph Generation From Text." arXiv preprint arXiv:2211.10511.

Our Custom Pipeline for Relation Extraction: Performance Evaluation

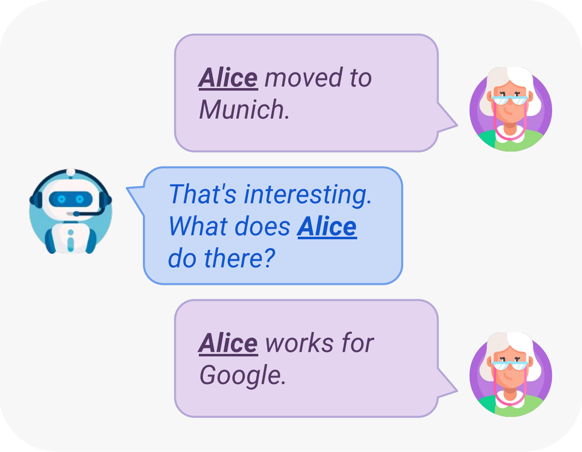
Input Data

Sample Dialogue



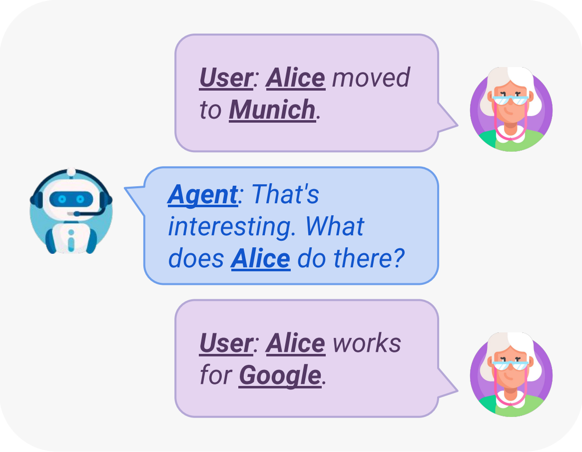
Coreference Resolution

F-COREF (F1: 78% [a]) [16]



NER

SpaCy + Rule-based NER (F1: 20% [b])



Relation Identification

XGBoost on DialogRE (F1: 49% [c])

Alice - Munich	**yes**	User - Munich	**no**
Alice - Google	**yes**	User - Google	**no**
Alice - User	**no**	Agent - Munich	**no**
Alice - Agent	**no**	Agent - Google	**no**
User - Agent	**no**	Munich - Google	**no**

Relation Classification

BERT on DialogRE (F1: 60% [d]) [13]

Alice - Munich	**lives_at**
Alice - Google	**works_at**

Cumulative errors ⇒ unviable results

[13] Yu, D., Sun, K., Cardie, C., Yu, D., 2020. "Dialogue-Based Relation Extraction." arXiv preprint arXiv:2004.08056.
[16] Otmazgin, S., Cattan, A., Goldberg, Y., 2022. "F-coref: Fast, Accurate and Easy to Use Coreference Resolution." arXiv preprint arXiv:2209.04280.

a) According to paper data distribution, not DialogRE! F-COREF Paper.
b) F1 Score for DialogRE entities, but potentially useful according to qualitative analysis.
c) Potential improvement with sentence based input.
d) Potential improvement via HiDialog.

Lesser Problem

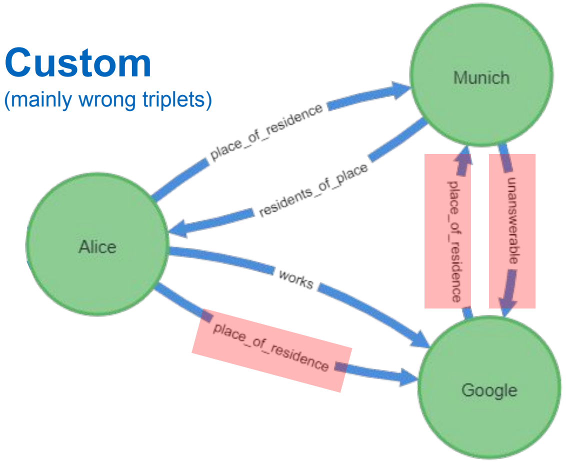
False predictions

Comparative Analysis of Relation Extraction: Custom Pipeline vs. LLMs

Input Dialogue

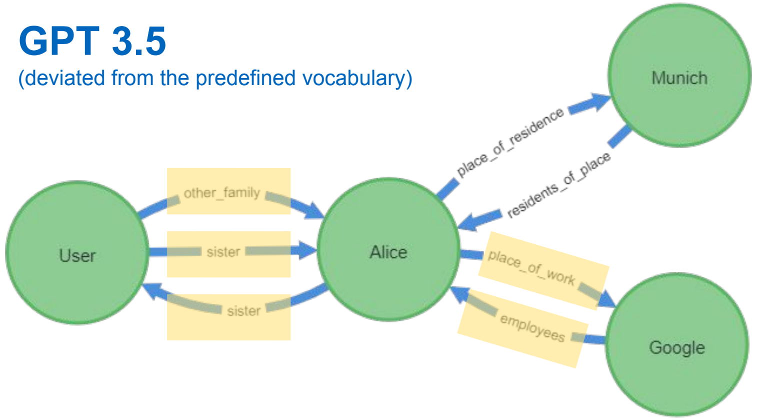
Custom

(mainly wrong triplets)



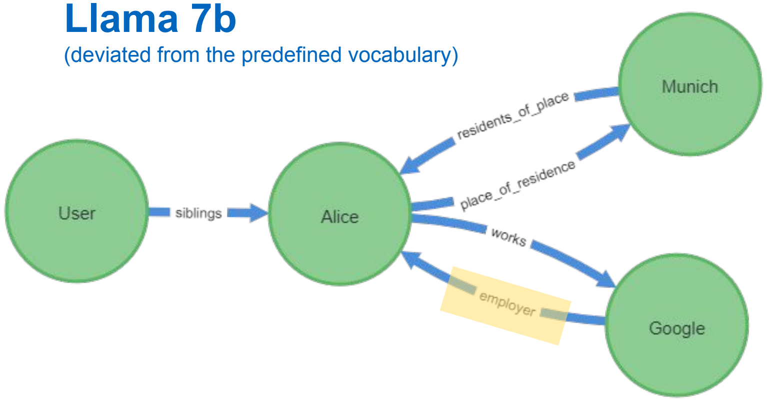
GPT 3.5

(deviated from the predefined vocabulary)



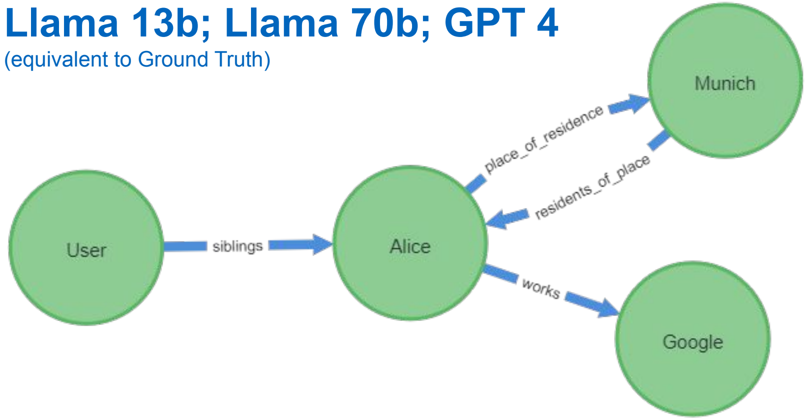
Llama 7b

(deviated from the predefined vocabulary)



Llama 13b; Llama 70b; GPT 4

(equivalent to Ground Truth)



Personalized Chatbot: End-to-End Proof of Concept Using LLMs

Q3

How can we **integrate knowledge** for **personalized responses**?

Conversation Triggers

Event-based Logic

Build opener to explore user interests.

Chat Mapping

Prompt Template

1. Instruct LLM to extract relations..
2. Merge relations into Neo4j
3. Stimulate & expand dialogue via **Active Listening**

Memory Retrieval

Search Strategy

Find minimal paths, between:

1. User node
2. Specific entities.

Current heuristic (simple):

1. Select a random node.
2. Find paths to the user.
3. Randomly select one path.

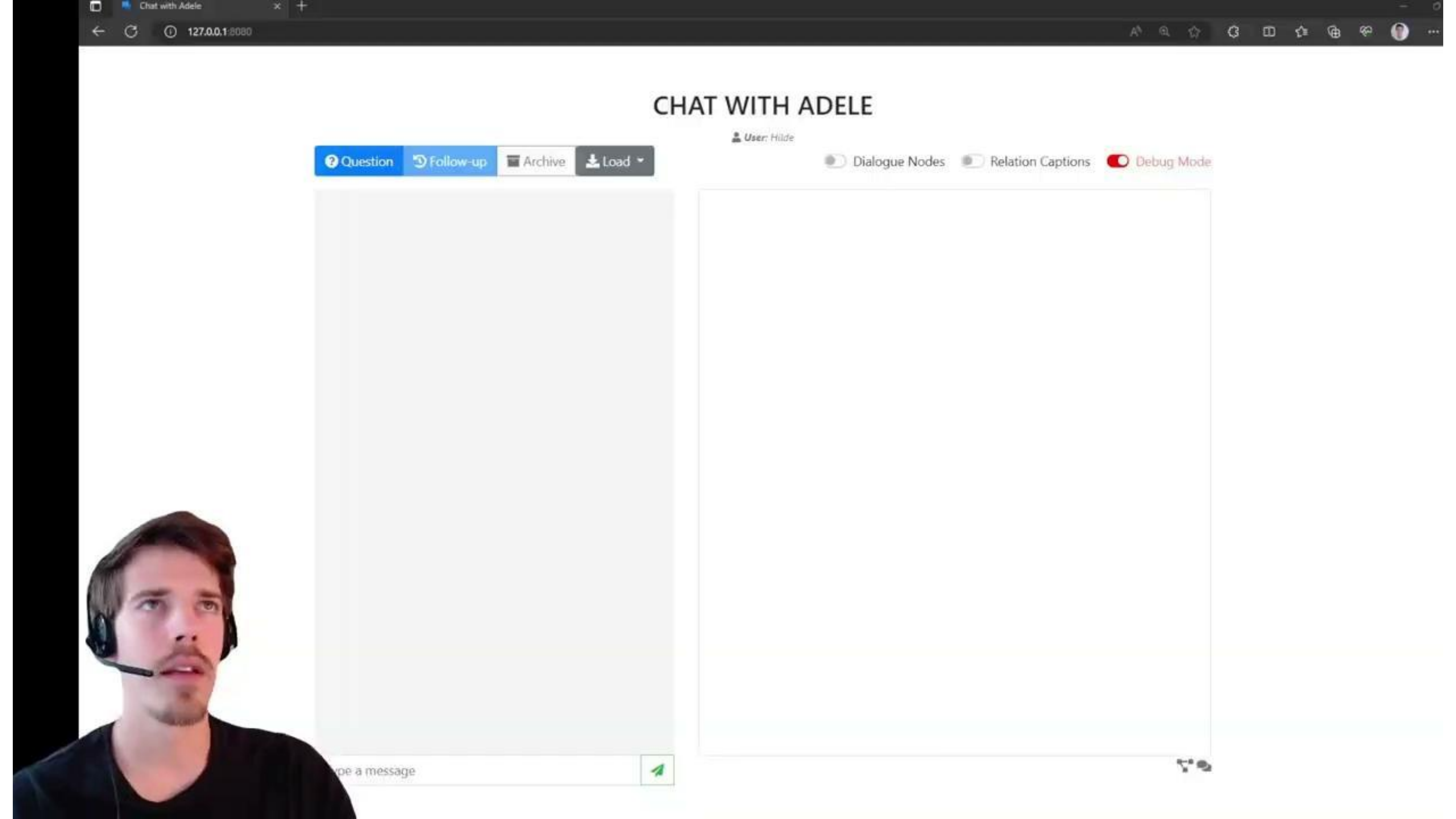
Personalized Triggers

Prompt Template

Instruct LLM to restart conversation using memory

Continue until user opts to stop chatting

Outline


Chat with Adele
127.0.0.1:8080
CHAT WITH ADELE
User: Hilde
Question
Follow-up
Archive
Load
Dialogue Nodes
Relation Captions
Debug Mode
ype a message

Outline

Motivation & Goal

Approach & Research

Initial Results

Demo

Next Steps & Timeline

Next Steps

Independent Relation Extraction

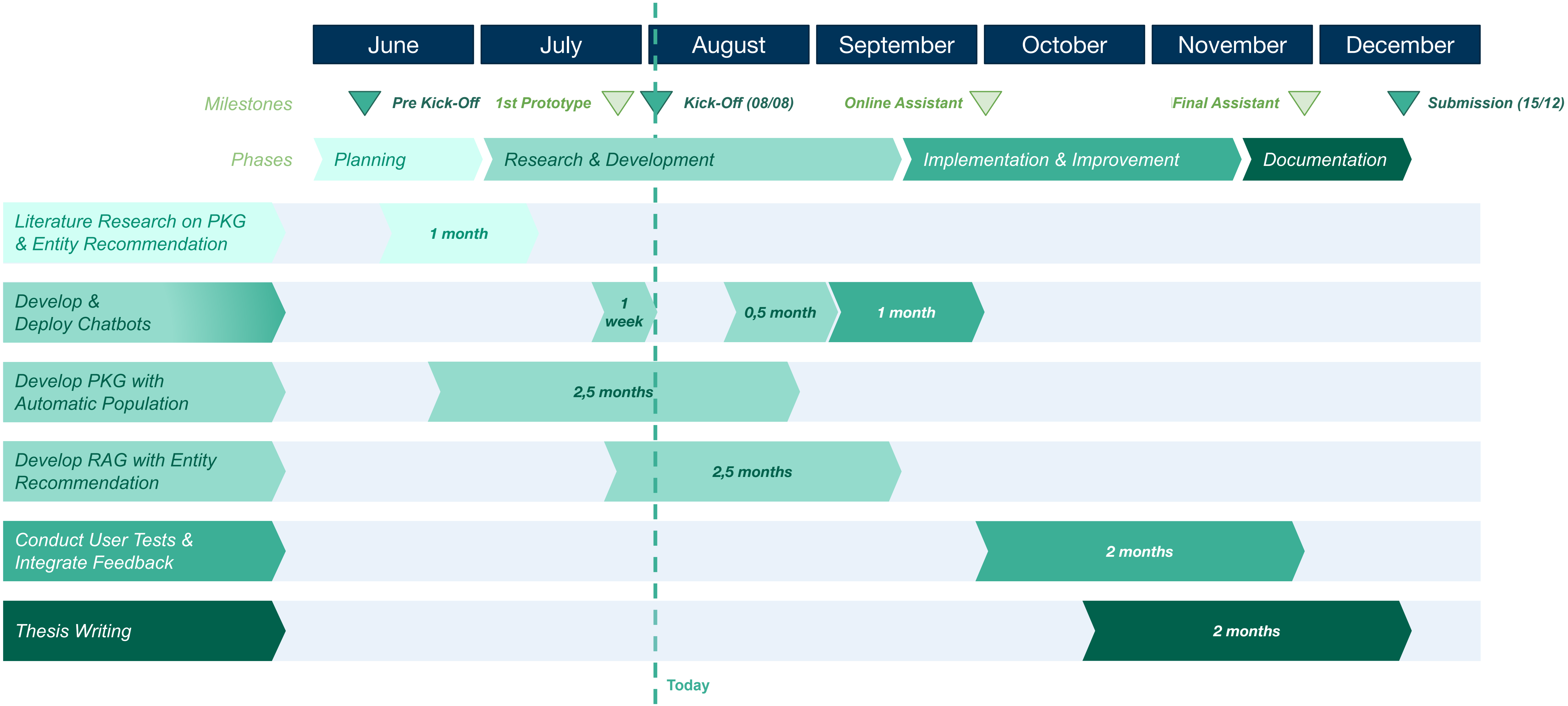
1. Fine-tune Llama-2; assess outcomes
2. Develop dataset
   a. Perform German translation
   b. Address class imbalance, if necessary
3. Utilize relation extraction metrics for results

Potential Addition: End-to-end Evaluation

If timely feasible, two possible ways

1. Simulate user chats; monitor engagement metrics
   a. Record Conversation Turn Count
   b. Assess Sentiment Analysis
2. Seek user feedback from ALMA specialists.

Research Journey: Our Current Plan



MSc. Student

Murilo Bellatini

Technical University of Munich (TUM)
TUM School of CIT
Department of Computer Science (CS)
Chair of Software Engineering for Business
Information Systems (sebis)

Boltzmannstraße 3
85748 Garching bei München

+49.89.289. 17132
matthes@in.tum.de
wwwmatthes.in.tum.de

Back-up Slide

Knowledge Graphs in Literature: A Review of Sources

[1] Balog, K., and Kenter, T. 2019. Personal Knowledge Graphs: A Research Agenda. In Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval (pp. 217–220). Association for Computing Machinery.

[2] Nicholas Thomas Walker, Stefan Ultes, and Pierre Lison. (2022). GraphWOZ: Dialogue Management with Conversational Knowledge Graphs.

[3] Krisztian Balog and Tom Kenter. Personal knowledge graphs: A research agenda. In Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval, ICTIR ’19, page 217–220, New York, NY, USA, 2019. Association for Computing Machinery.

[4] Yang, Y., Lin, J., Zhang, X., and Wang, M. 2022. PKG: A Personal Knowledge Graph for Recommendation. In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 3334–3338). Association for Computing Machinery.

[5] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich K ̈uttler, Mike Lewis, Wen tau Yih, Tim Rockt ̈aschel, Se- bastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive nlp tasks, 2021

[6] https://python.langchain.com/en/latest/modules/prompts/prompt_templates.html

[7] Fukaya, Y., Koyama, S., Kimura, Y. and Kitamura, T., 2009. Education to promote verbal communication by caregivers in geriatric care facilities. Japan Journal of Nursing Science, 6(2), pp.91-103.

[8] Office, E.E., Rodenstein, M.S., Merchant, T.S., Pendergrast, T.R. and Lindquist, L.A., 2020. Reducing social isolation of seniors during COVID-19 through medical student telephone contact. Journal of the American Medical Directors Association, 21(7), pp.948-950.;

[9] D’Onofrio, G., Fiorini, L., Hoshino, H., Matsumori, A., Okabe, Y., Tsukamoto, M., Limosani, R., Vitanza, A., Greco, F., Greco, A. and Giuliani, F., 2019. Assistive robots for socialization in elderly people: results pertaining to the needs of the users. Aging clinical and experimental research, 31, pp.1313-1329.;

[10] Sgorbissa, A., Papadopoulos, I., Bruno, B., Koulouglioti, C. and Recchiuto, C., 2018, October. Encoding guidelines for a culturally competent robot for elderly care. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 1988-1995). IEEE.;

[11] Kitwood, T.M., 2013. Demenz: Der person-zentrierte Ansatz im Umgang mit verwirrten Menschen. Huber.

[12] Kitwood, T., Brooker, D., 2019. Dementia reconsidered revisited: The person still comes first. Open University Press, McGraw-Hill Education, London, UK.

[13] Yu, D., Sun, K., Cardie, C., Yu, D., 2020. "Dialogue-Based Relation Extraction." arXiv preprint arXiv:2004.08056.

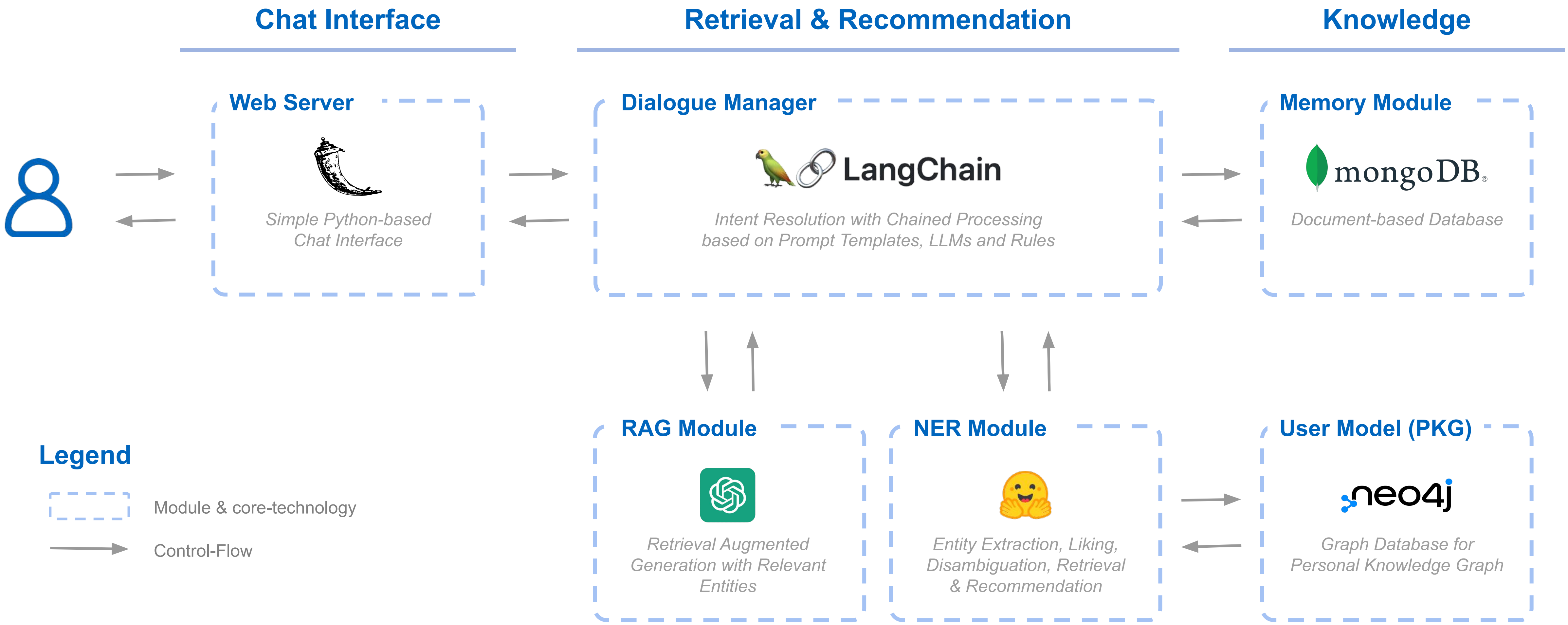
[14] Temperoni, A., Biryukov, M., Theobald, M., 2022. "Enriching Relation Extraction with OpenIE." arXiv preprint arXiv:2212.09376.

[15] Melnyk, I., Dognin, P., Das, P., 2022. "Knowledge Graph Generation From Text." arXiv preprint arXiv:2211.10511.

[16] Otmazgin, S., Cattan, A., Goldberg, Y., 2022. "F-coref: Fast, Accurate and Easy to Use Coreference Resolution." arXiv preprint arXiv:2209.04280.

Components Overview: Integrated Modules for Contextualized Interaction

Preliminary diagram:



Automated Graph Construction - LLM One Shot - Prompt Design

Extract personal relevant entities, and their relations. Return only the jsonl format list .

Ontology:

- relations: {"acquaintance", "children", "other_family", "parents", "siblings", "spouse", "place_of_residence", "visited_place", "pet", "residents_of_place", "visitors_of_place"}
- types: {"ORG", "GPE", "PERSON", "DATE", "EVENT", “ANIMAL”}

Input:

[
"User: My daughter, Emma, recently moved to London.",
"Agent: That's exciting! Does she like it there?",
"User: Yes, she loves it! She even adopted a cat named Whiskers.",
]

Output:

[
{'x': 'User', 'x_type': 'PERSON', 'y': 'Emma', 'y_type': 'PERSON', 'r': 'children'},
{'x': 'Emma', 'x_type': 'PERSON', 'y': 'London', 'y_type': 'GPE', 'r': 'place_of_residence'},
{'x': 'London', 'x_type': 'GPE', 'y': 'Emma', 'y_type': 'PERSON', 'r': 'residents_of_place'},
{'x': 'Emma', 'x_type': 'PERSON', 'y': 'Whiskers', 'y_type': 'ANIMAL', 'r': 'pet'},
{'x': 'Whiskers', 'x_type': 'ANIMAL', 'y': 'Emma', 'y_type': 'PERSON', 'r': 'pet'},
]

Input:

[
"User: My son, John, went to visit Tokyo last month.",
"Agent: That sounds like a fun trip. Did he go alone?",
"User: No, he went with his wife, Mary. They even brought their dog, Rover."
]

Output:

Expected output:

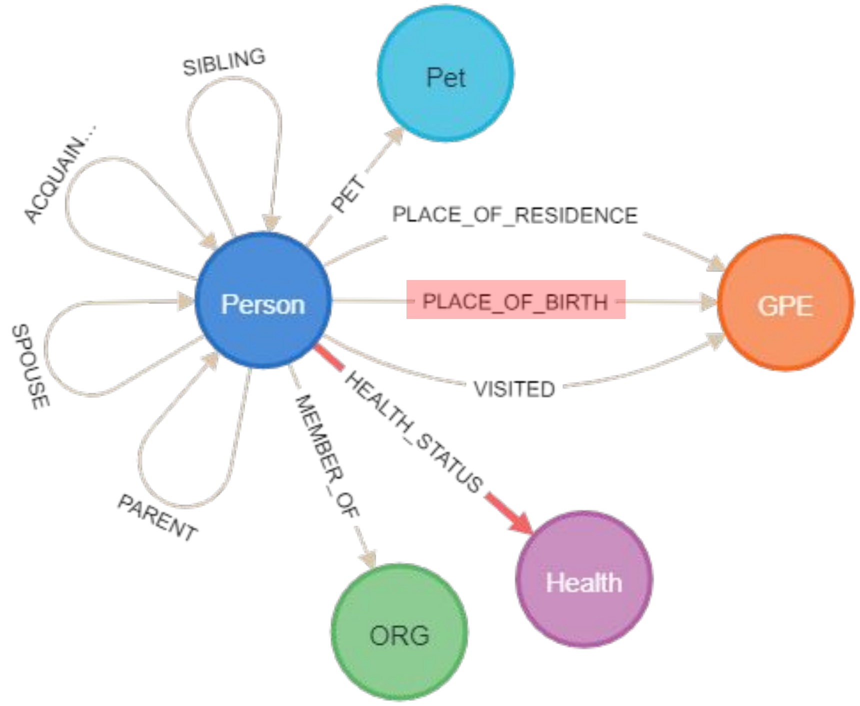
[
{'x': 'User', 'x_type': 'PERSON', 'y': 'John', 'y_type': 'PERSON', 'r': 'children'},
{'x': 'John', 'x_type': 'PERSON', 'y': 'Tokyo', 'y_type': 'GPE', 'r': 'visited_place'},
{'x': 'Tokyo', 'x_type': 'GPE', 'y': 'John', 'y_type': 'PERSON', 'r': 'visitors_of_place'},
{'x': 'John', 'x_type': 'PERSON', 'y': 'Mary', 'y_type': 'PERSON', 'r': 'spouse'},
{'x': 'Mary', 'x_type': 'PERSON', 'y': 'John', 'y_type': 'PERSON', 'r': 'spouse'},
{'x': 'John', 'x_type': 'PERSON', 'y': 'Rover', 'y_type': 'ANIMAL', 'r': 'pet'},
{'x': 'Mary', 'x_type': 'PERSON', 'y': 'Rover', 'y_type': 'ANIMAL', 'r': 'pet'},
{'x': 'Rover', 'x_type': 'ANIMAL', 'y': 'John', 'y_type': 'PERSON', 'r': 'pet'},
{'x': 'Rover', 'x_type': 'ANIMAL', 'y': 'Mary', 'y_type': 'PERSON', 'r': 'pet'},
]

Nationality (no city granularity)!

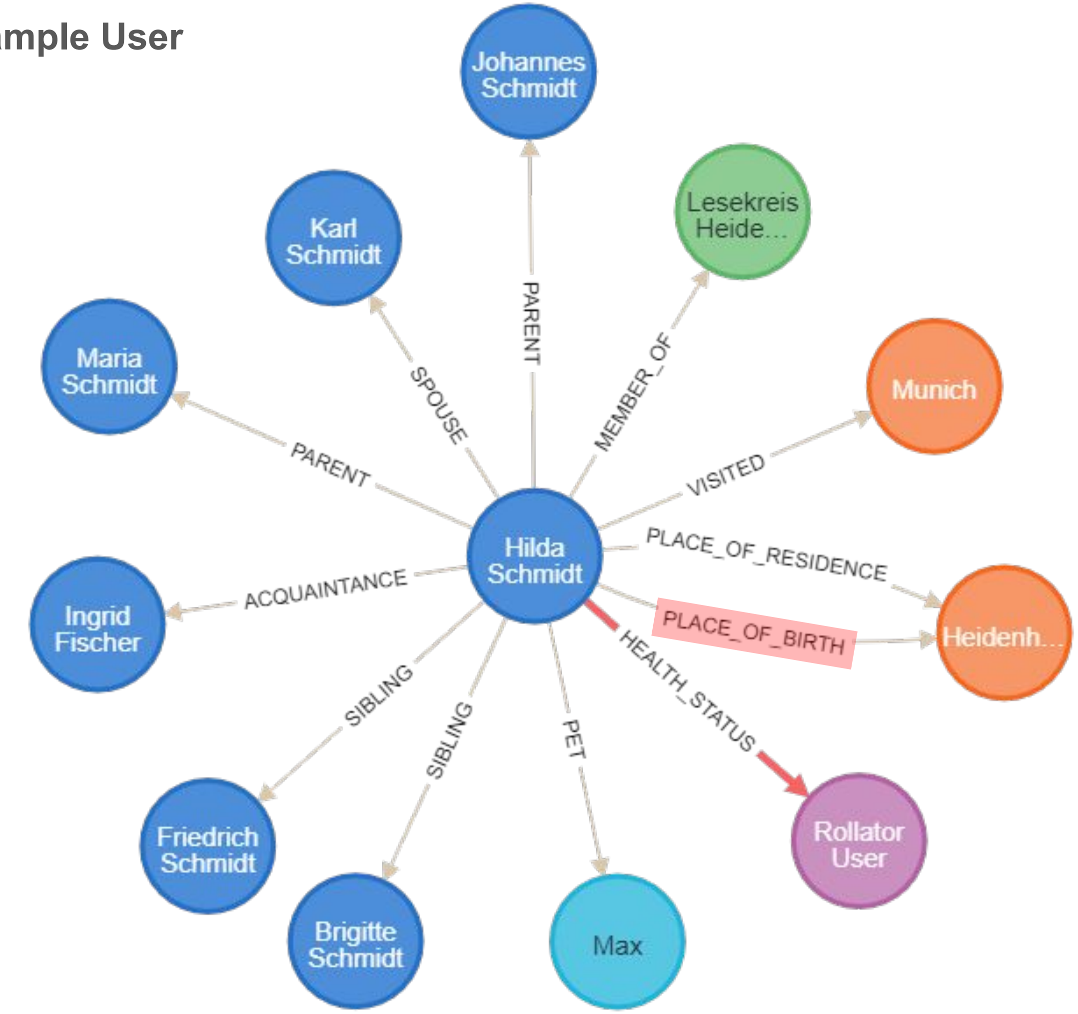
Proposed Data Model using DialogRE (Friends Dataset)

The diagram below illustrates the potential capabilities achievable with DialogRE. However, it is important to measure the quality of the output. ***Next step:*** *Compute the F1-score for each relation and entity label.*

Data Model: Entity and Relation Types



Example User



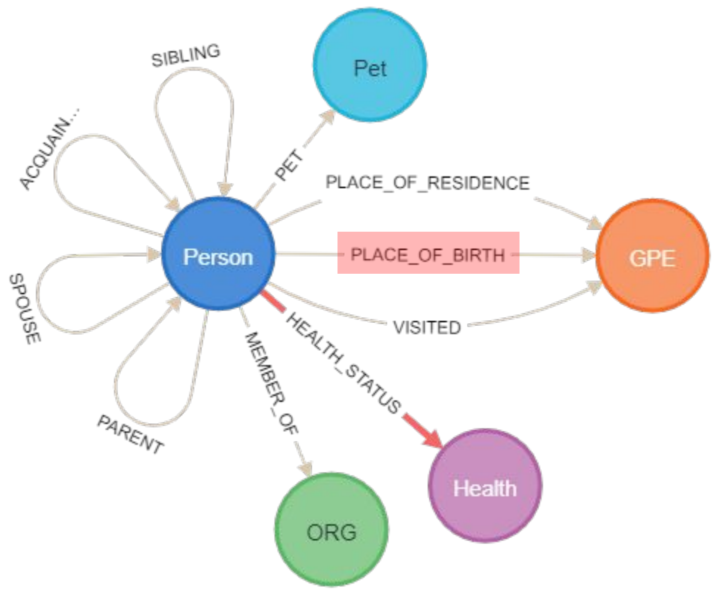
Legend

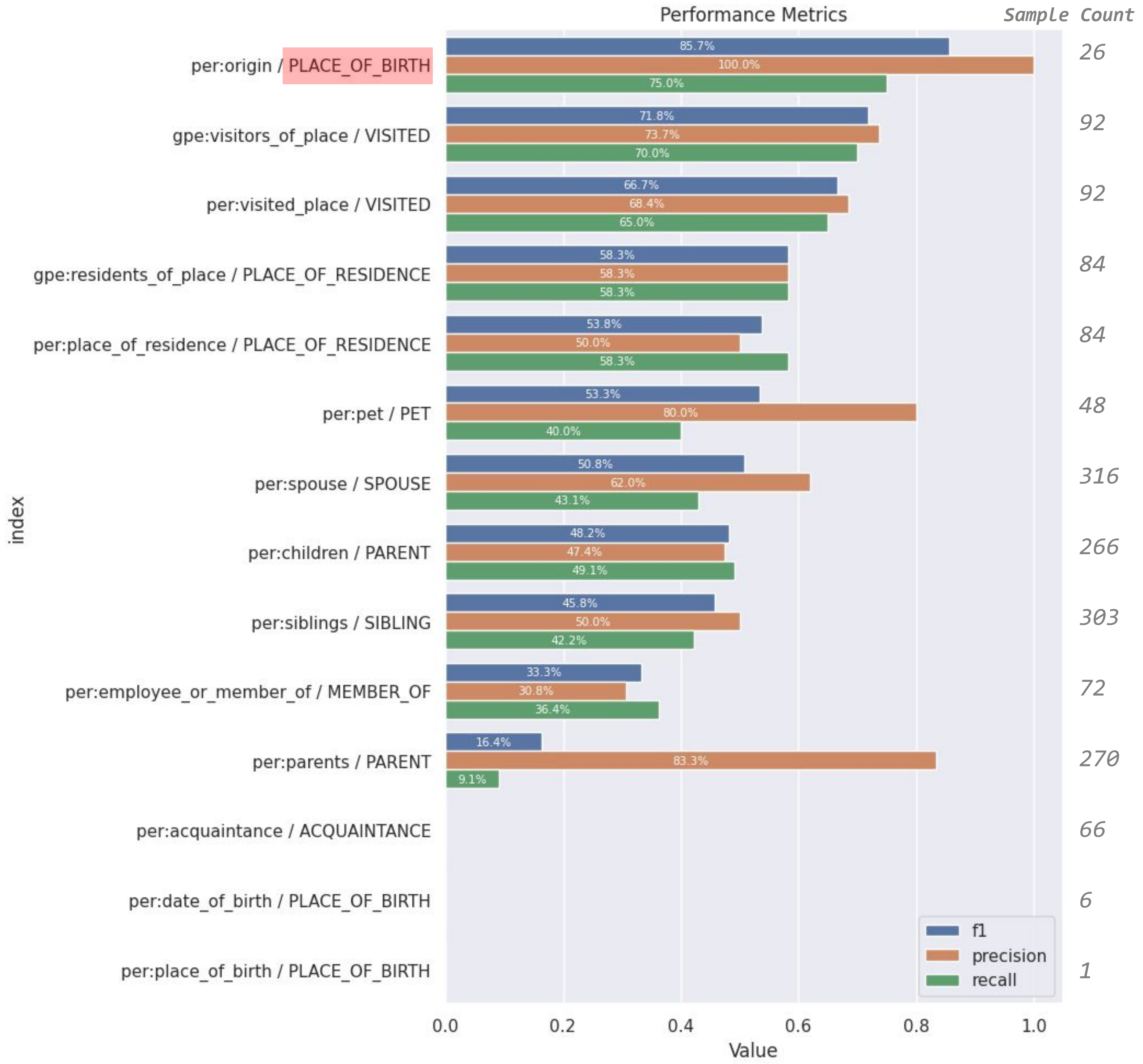
- ORG: Organization
- GPE: Geopolitical Entity (city, state, region, country)
- Relation in Red: Still not modelled (not present in DialogRE) → *Potential fix*: *Keep Me Updated!*

Nationality (no city granularity)!

In-Depth Analysis of DialogRE Paper Results - Data Model Impact

Current metrics broken down by class





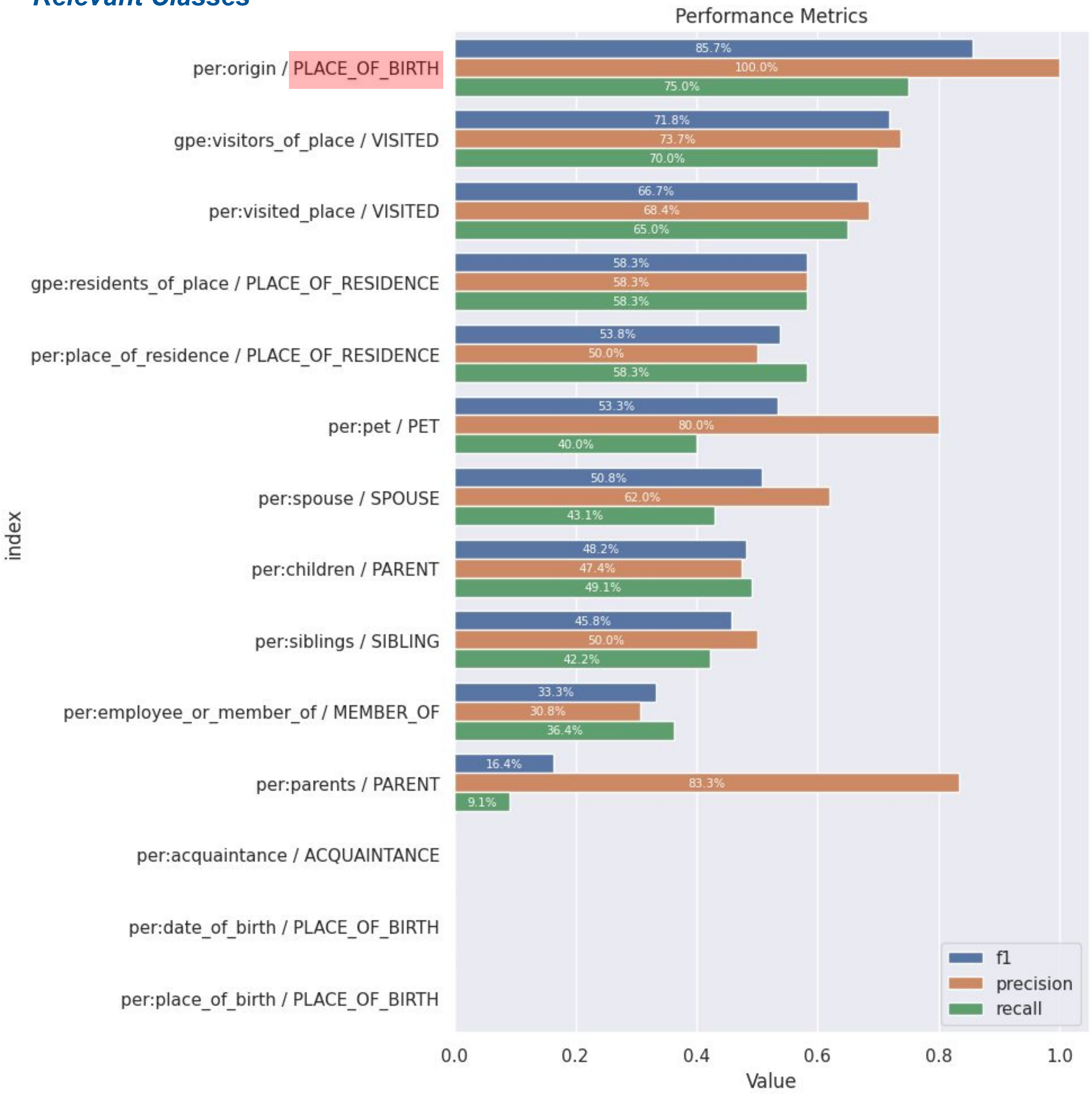
Macro-Averaging

	focus relations (all)
f1	0.385888 (0.3645)
precision	0.459342 (0.4438)
recall	0.369568 (0.3490)

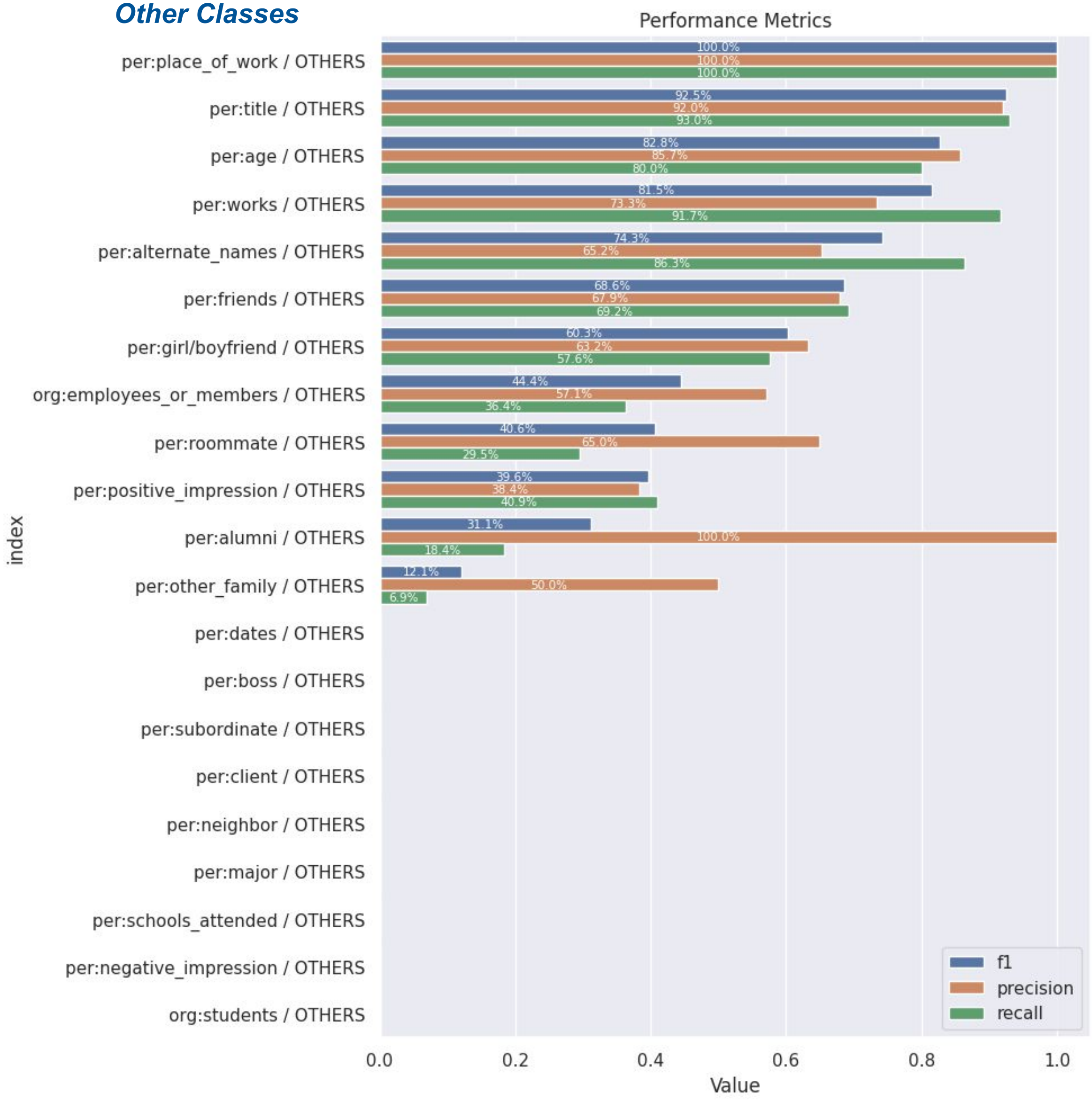
Nationality (no city granularity)!

In-Depth Analysis of DialogRE Paper Results - Data Model Impact

Relevant Classes



Other Classes



Kitwood's Categories in DialogRE Data

While there may be some overlapping classes, the current proposed allocation is as follows:

Group: Attachment | Sample Count: 3,088 (40.4%)

group	relation_type	Counts	%
Attachment	per:roommate	193	2.5
	per:pet	48	0.6
	per:client	87	1.1
	per:dates	35	0.5
	per:other_family	120	1.6
	per:children	266	3.5
	per:parents	270	3.5
	per:acquaintance	66	0.9
	per:spouse	316	4.1
	per:friends	648	8.5
	per:girl/boyfriend	736	9.6
	per:siblings	303	4.0

Group: Identity | Sample Count: 2,667 (34.8%)

group	relation_type	Counts	%
Identity	per:date_of_birth	6	0.1
	per:title	414	5.4
	per:major	6	0.1
	per:origin	26	0.3
	per:place_of_birth	1	0.0
	per:age	78	1.0
	per:alternate_names	2136	27.9

Group: Comfort | Sample Count: 879 (11.5%)

group	relation_type	Counts	%
Comfort	per:negative_impression	222	2.9
	per:positive_impression	657	8.6

Comfort Group: Potential extension with the MELD dataset for sentiment classification.

Group: Occupation | Sample Count: 607 (7.8%)

group	relation_type	Counts	%
Occupation	per:place_of_work	71	0.9
	org:employees_or_members	72	0.9
	per:subordinate	63	0.8
	per:boss	72	0.9
	per:works	89	1.2
	org:students	8	0.1
	per:schools_attended	8	0.1
	per:alumni	152	2.0
	per:employee_or_member_of	72	0.9

Group: Inclusion | Sample Count: 408 (5.3%)

group	relation_type	Counts	%
Inclusion	per:neighbor	56	0.7
	per:place_of_residence	84	1.1
	gpe:residents_of_place	84	1.1
	gpe:visitors_of_place	92	1.2
	per:visited_place	92	1.2

Group: Others | Sample Count: 1 (0.0%)

group	relation_type	Counts	%
Others	gpe:births_in_place	1	0.0