Coaching Diabetes Patients by Conversational AI through Event-centric Knowledge Graphs.

Piek Vossen, June, 26th 206

Hybrid Intelligence

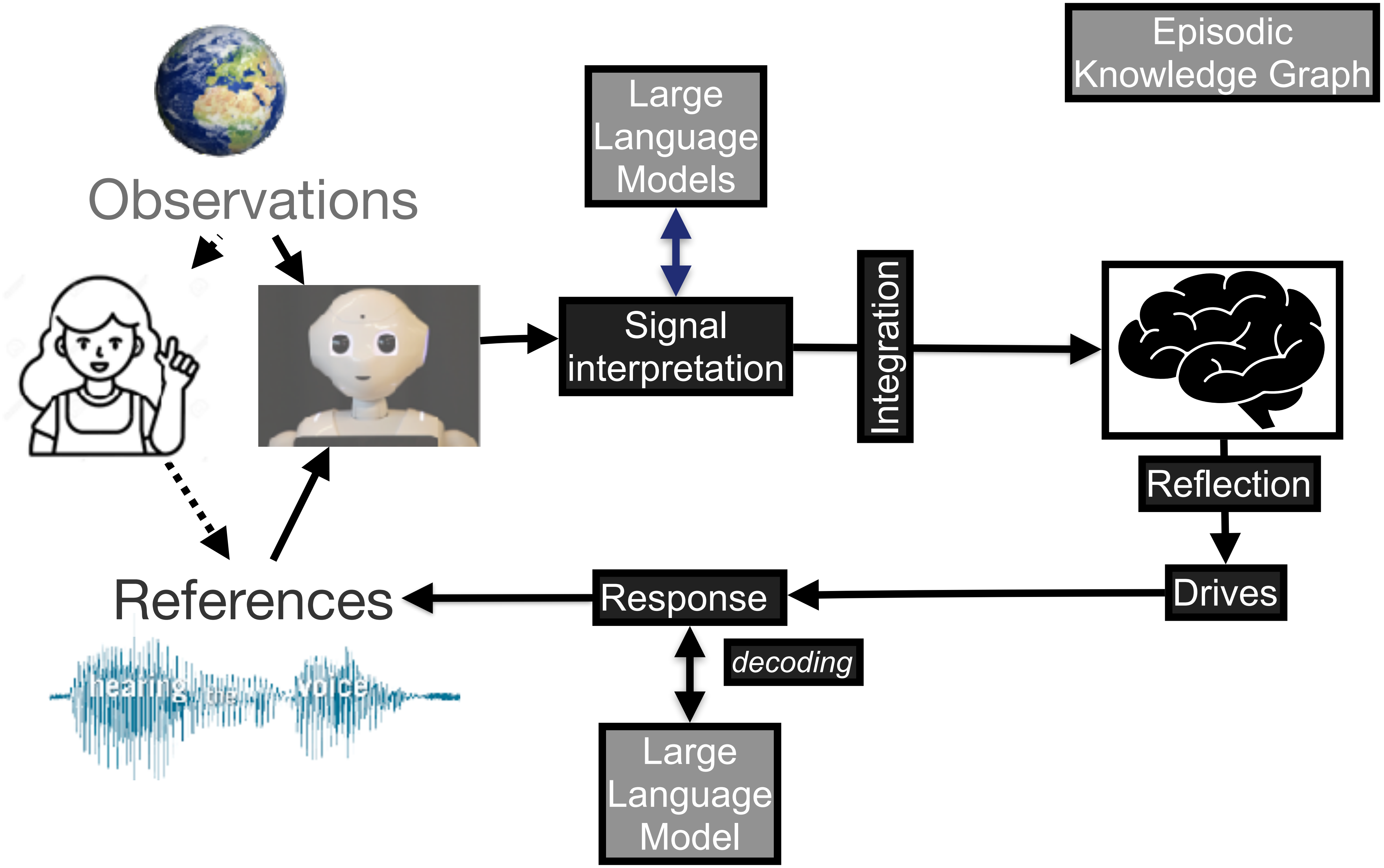
CLaDA BG

Overview

- Leolani: agent platform for controlled Knowledge Graph driven interaction
- Event-centric Knowledge Graphs for capturing interactions and activities of daily life
- The Hybrid Intelligence diabetes use case
- Bootstrapping the semantics of diabetes lifestyle coaching: from open to closed world model

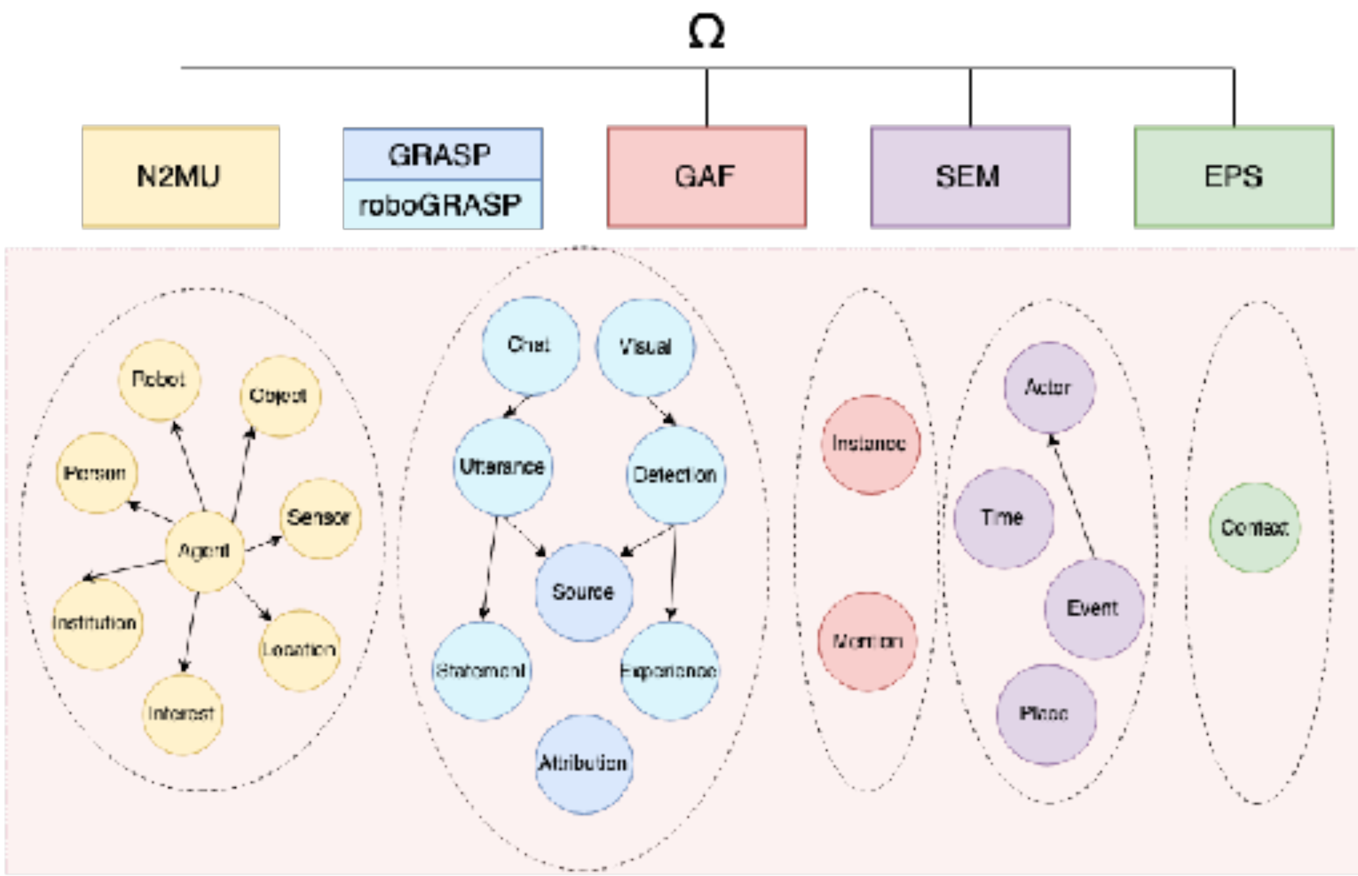
Leolani agent platform

https://leolani.github.io/

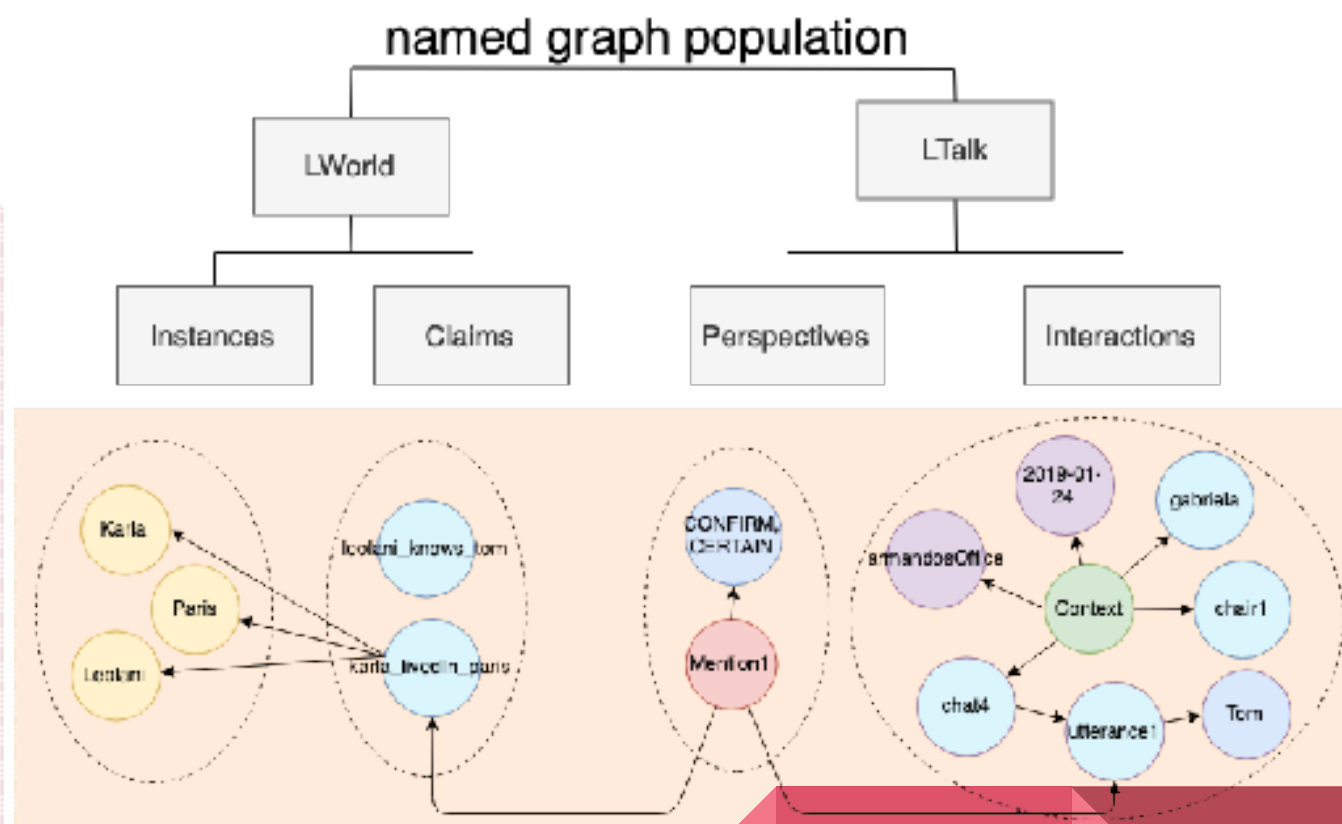


Knowledge representation

- TBox = Ontology
  - Model of the world



- ABox = Vocabulary
  - Instances existing in the world



→ Vossen, Piek, et al. "Leolani: A robot that communicates and learns about the shared world." 2019 ISWC Satellite Tracks (Posters and Demonstrations, Industry, and Outrageous Ideas), ISWC 2019-Satellites. CEUR-WS, 2019.

Identities

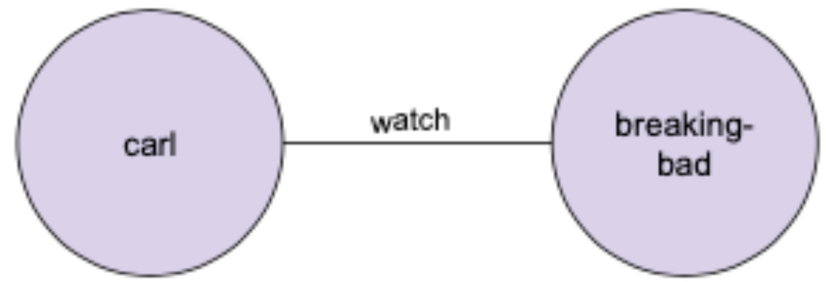


Figure 2.3: Visualization of the RDF triples representing that Carl watches a TV series called 'Breaking Bad'". Nodes in purple signal instances with an *identity*.

Named graph: lTalk:Instances

lWorld:carl	a	n2mu:person, gaf:Instance ;
	rdfs:label	"Carl" .
lWorld:breaking-bad	a	n2mu:tv-series, gaf:Instance ;
	rdfs:label	"Breaking Bad" .
lWorld:watch	a	sem:Event;
	rdfs:label	"watch" .

Named graph: lWorld:Claims

lWorld:carl_watch_breaking-bad	a	gaf:Statement ;
	owl:sameAs	lWorld:claim1 .

Named graph: lWorld:carl_watch_breaking-bad

lWorld:carl	lWorld:watch	lWorld:breaking-bad.

Table 2.1: RDF triples representing that Carl watches a TV series called 'Breaking Bad'". Columns represent the subject, predicate, and object values in the triples.

References (GAF)

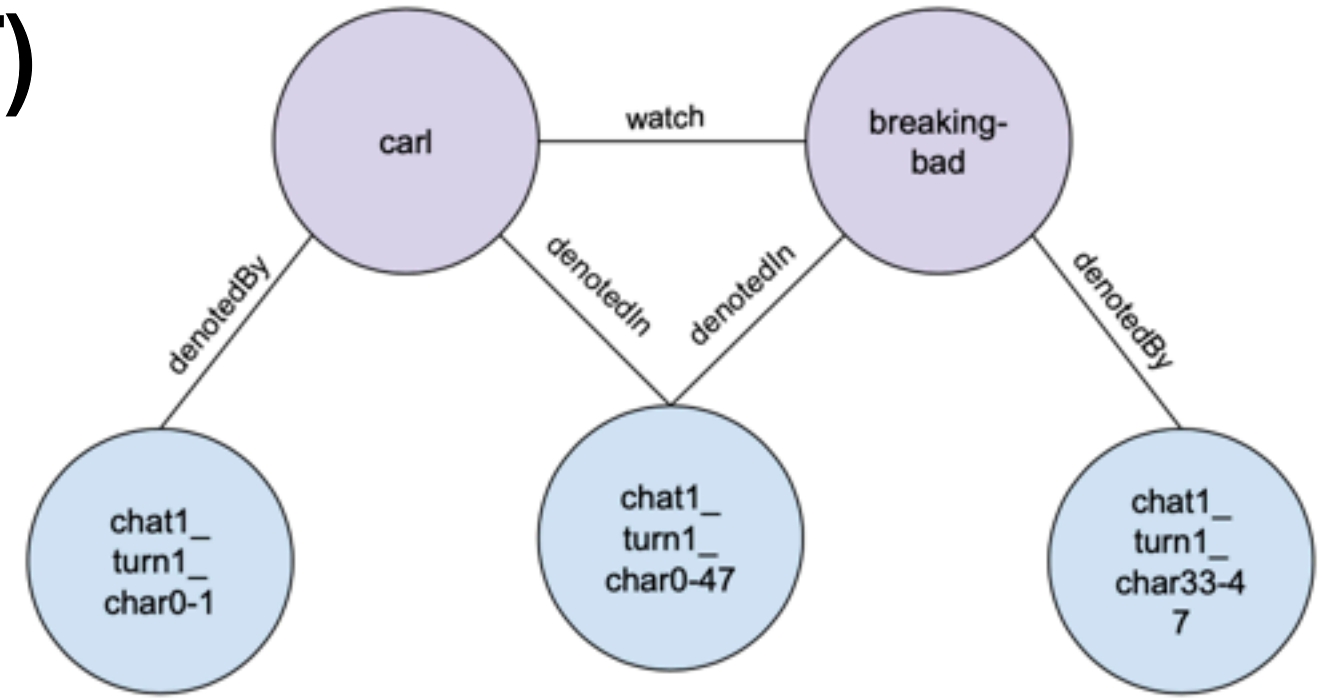


Figure 2.4: Visualization of the RDF triples representing that Carl said: "I am watching a TV series called 'Breaking Bad'". Nodes in purple signal instances with an *identity* while nodes in blue represent *mentions* referring to these instances.

Named graph: lTalk:Instances

lWorld:carl	a	n2mu:person, gaf:Instance ;
	rdfs:label	"Carl" ;
	gaf:denotedBy	lTalk:chat1_turn1_char0-1 ;
	gaf:denotedIn	lTalk:chat1_turn1_char0-47.
lWorld:breaking-bad	a	n2mu:tv-series, gaf:Instance ;
	rdfs:label	"Breaking Bad";
	gaf:denotedBy	lTalk:chat1_turn1_char33-47 ;
	gaf:denotedIn	lTalk:chat1_turn1_char0-47.

Table 2.2: RDF triples representing that Carl said: "I am watching a TV series called 'Breaking Bad'". Columns represent the subject, predicate, and object values in the triples.

Interactions as events

Named graph: lTalk:Interactions		
lTalk:chat1_turn1_char5-13	a	gaf:Mention ;
	rdf:value	"I am watching a TV series called 'Breaking Bad'" ;
	prov:wasDerivedFrom	lTalk:chat1_utterance1 .
lTalk:chat4_utterance1	a	gaf:Utterance .
lTalk:chat4	a	gaf:Chat;
	sem:hasSubevent	lTalk:chat4_utterance1 ;
	sem:hasActor	lFriends:carl ;
	sem:hasTimeStamp	lContext:20180512 ;
	sem:hasPlace	lContext:armandosOffice .
lFriends:carl	a	sem:Actor, gaf:Source .
lContext:20180512	a	sem:Time .
lContext:armandosOffice	a	sem:Place .

Table 2.3: RDF triples representing the interaction with Carl as situated events.

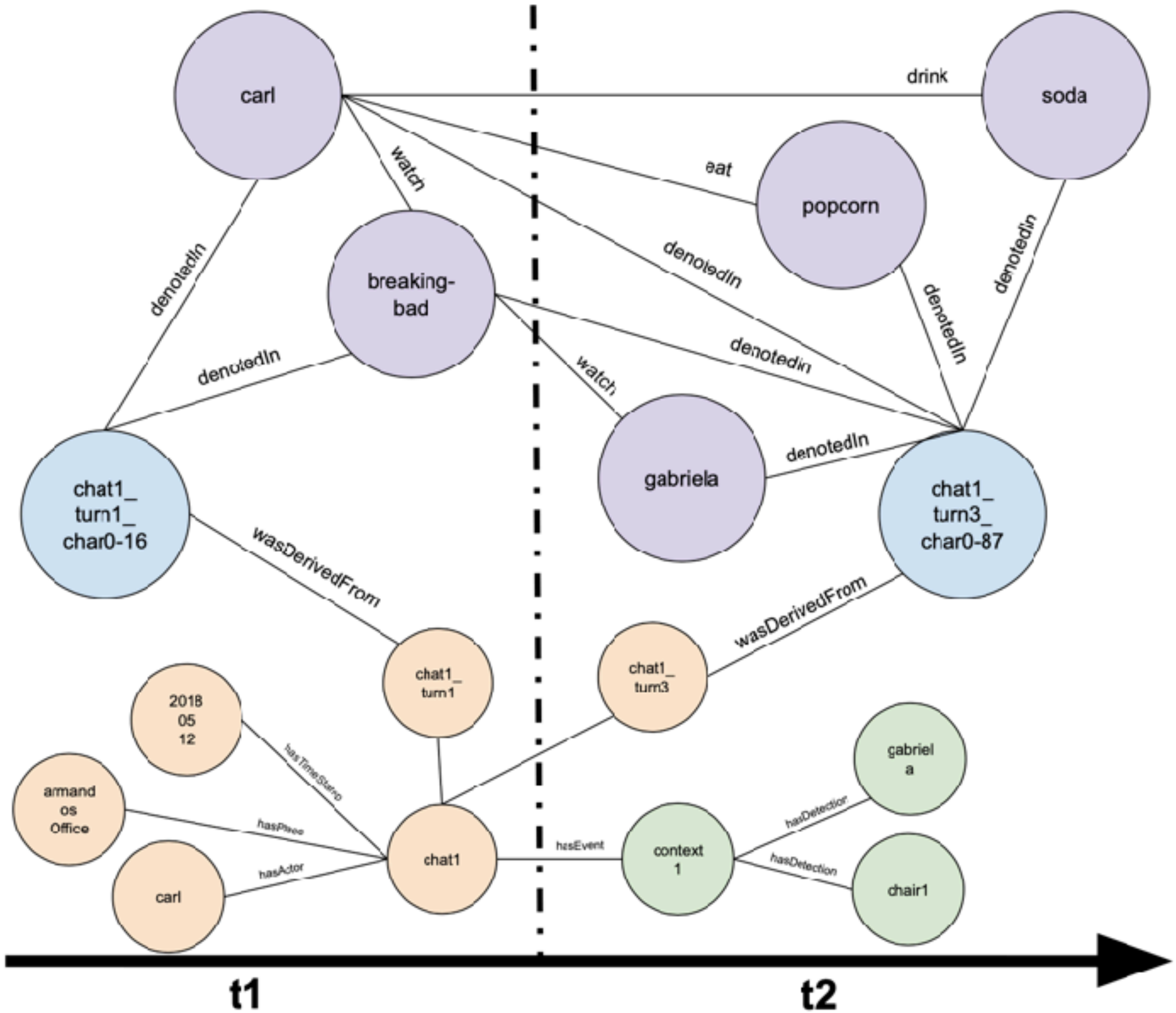
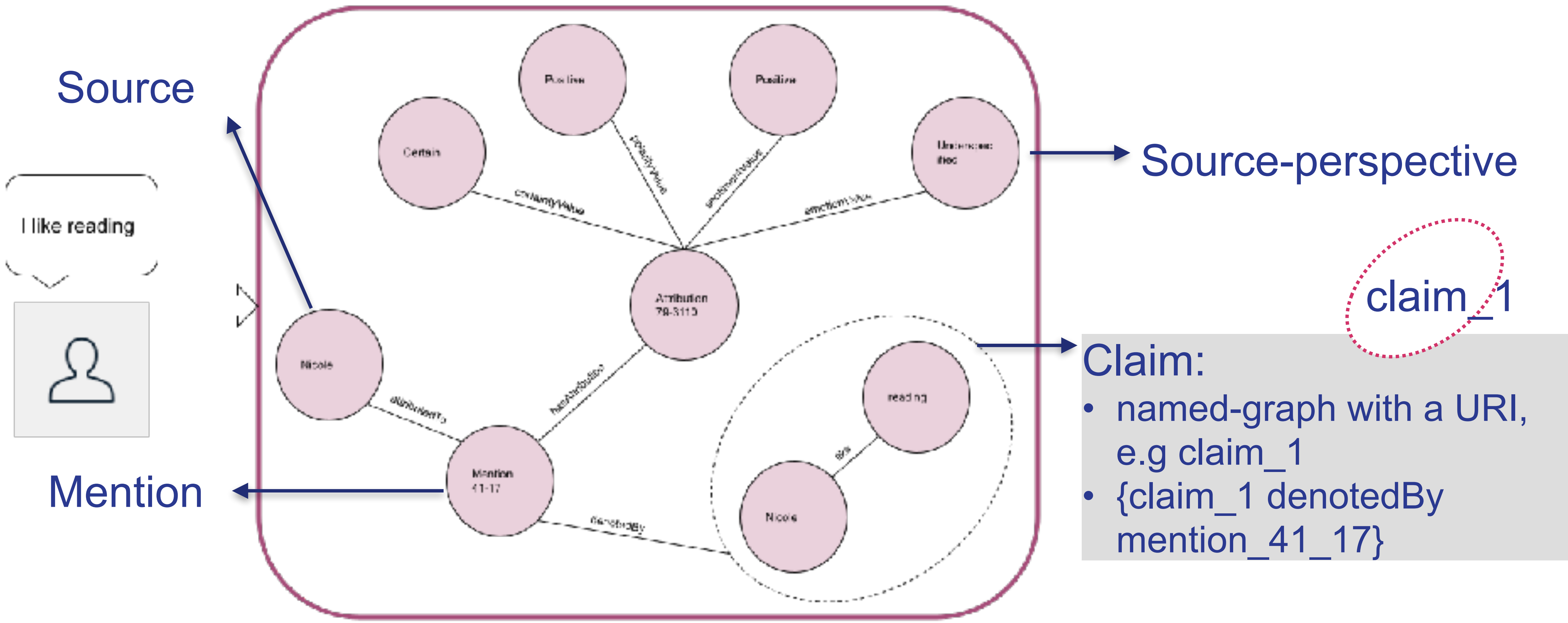


Figure 2.7: Visualization of the accumulation of information as RDF triples over time. Nodes in orange represent event provenance information, while nodes in green represent contextual information. As before, nodes in purple signal instances with an *identity* while nodes in blue represent *mentions* of these instances.

Episodic knowledge graphs consisting of claims



Claim:

- named-graph with a URI, e.g claim_1
- {claim_1 denotedBy mention_41_17}

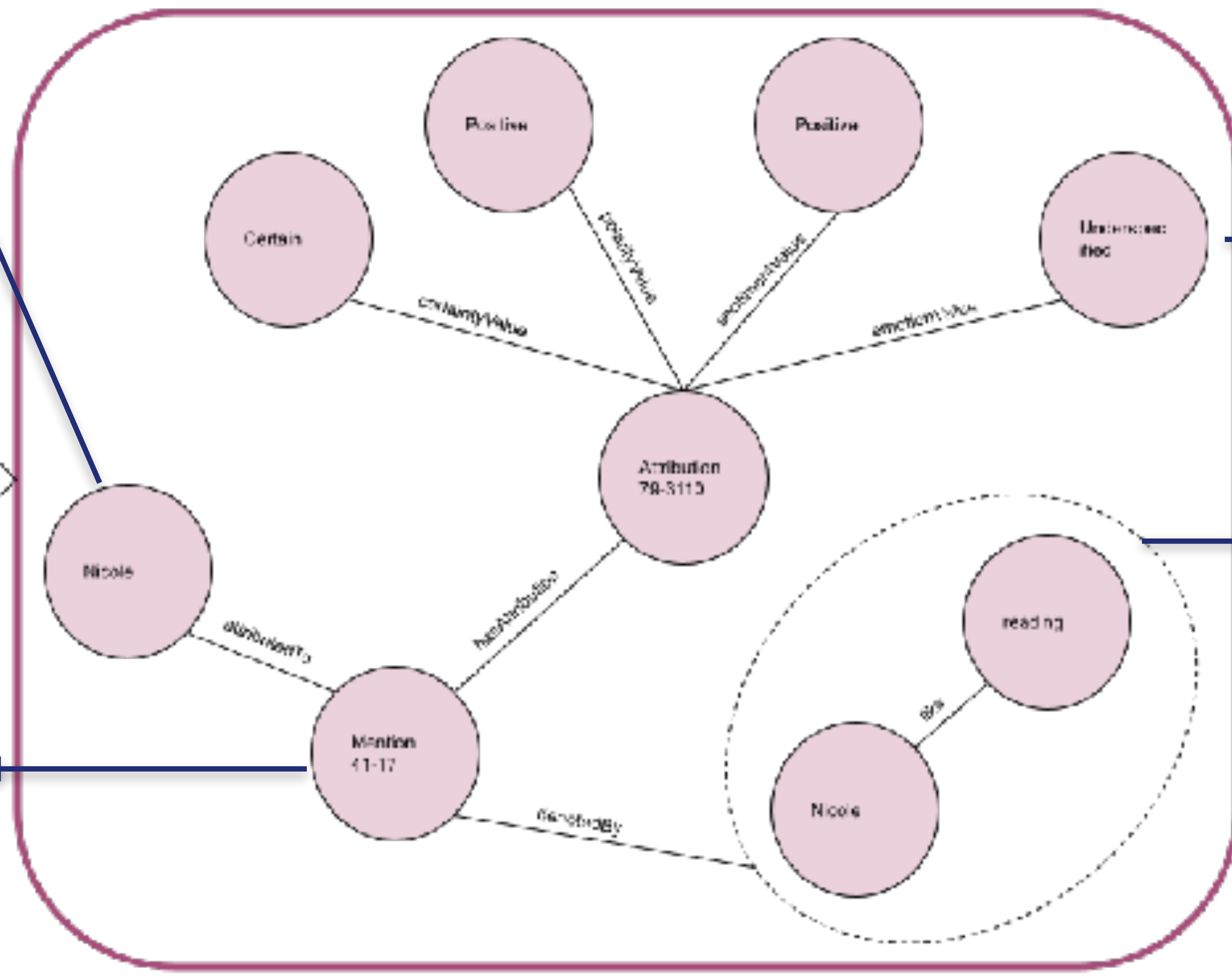
➔ Fokkens, Antske, et al. "GRaSP: Grounded Representation and Source Perspective." Proceedings of the Workshop Knowledge Resources for the Socio-Economic Sciences and Humanities associated with RANLP 2017. 2017.

➔ Santamaria, Selene Baez, et al. "EMISSOR: A platform for capturing multimodal interactions as episodic memories and interpretations with situated scenario-based ontological references." Proceedings of the First workshop Beyond Language: Multimodal Semantic Representations in conjunction with IWCS 2022. 2021.

Episodic knowledge graphs consisting of claims

Source

Mention



Source-perspective

claim_1

Claim:
- named-graph with a URI, e.g claim_1
- {claim_1 denotedBy mention_41_17,

mention_52_4,

mention_86_2,

mention_82_8,

mention_90_3,

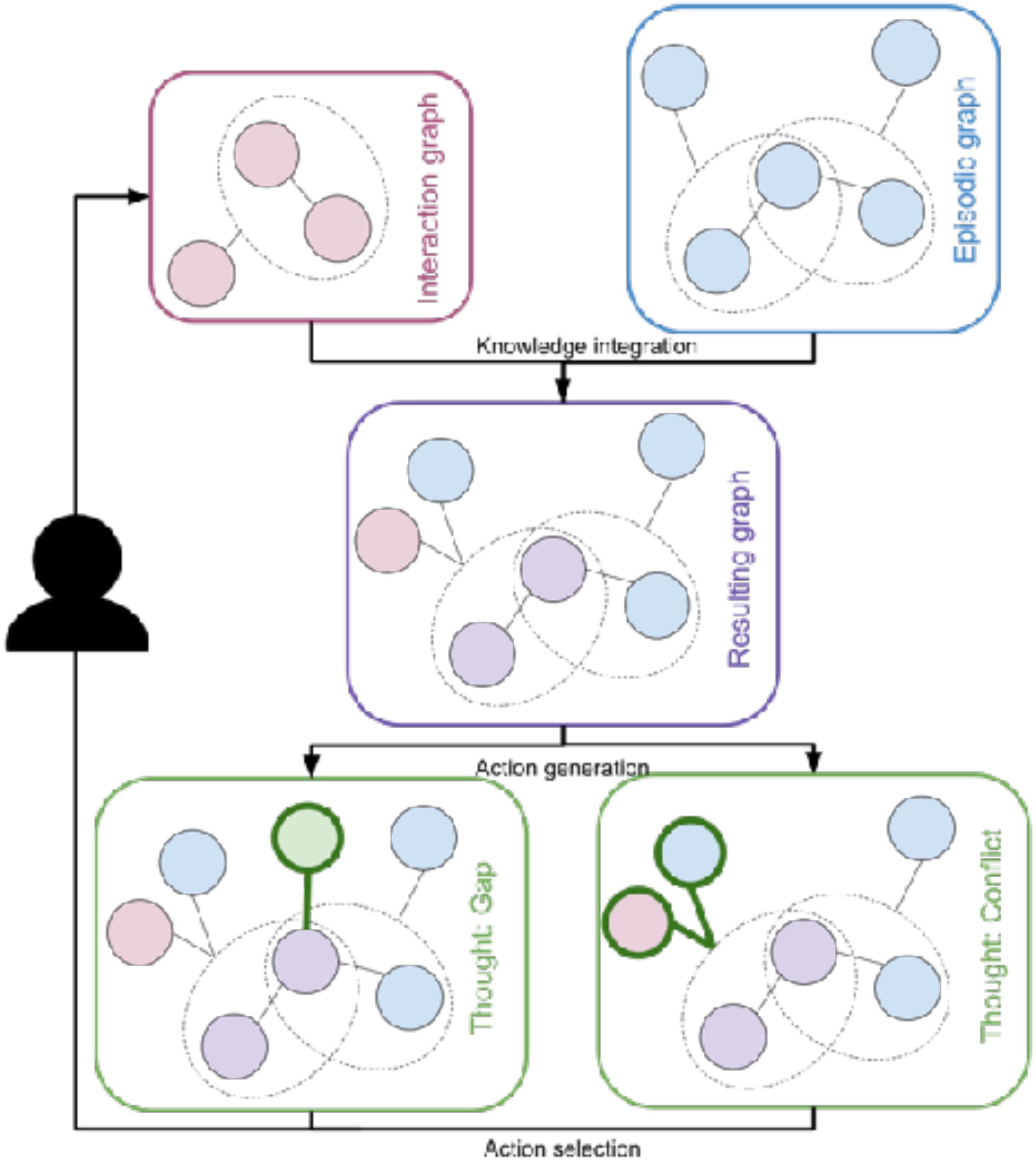
mention_91_5,

mention_92_2

➔ Fokkens, Antske, et al. "GRaSP: Grounded Representation and Source Perspective." Proceedings of the Workshop Knowledge Resources for the Socio-Economic Sciences and Humanities associated with RANLP 2017. 2017.
➔ Santamaria, Selene Baez, et al. "EMISSOR: A platform for capturing multimodal interactions as episodic memories and interpretations with situated scenario-based ontological references." Proceedings of the First workshop Beyond Language: Multimodal Semantic Representations in conjunction with IWCS 2022. 2021.

A pro-active agent with a drive to improve the quality of the knowledge graph

Information integration cycle

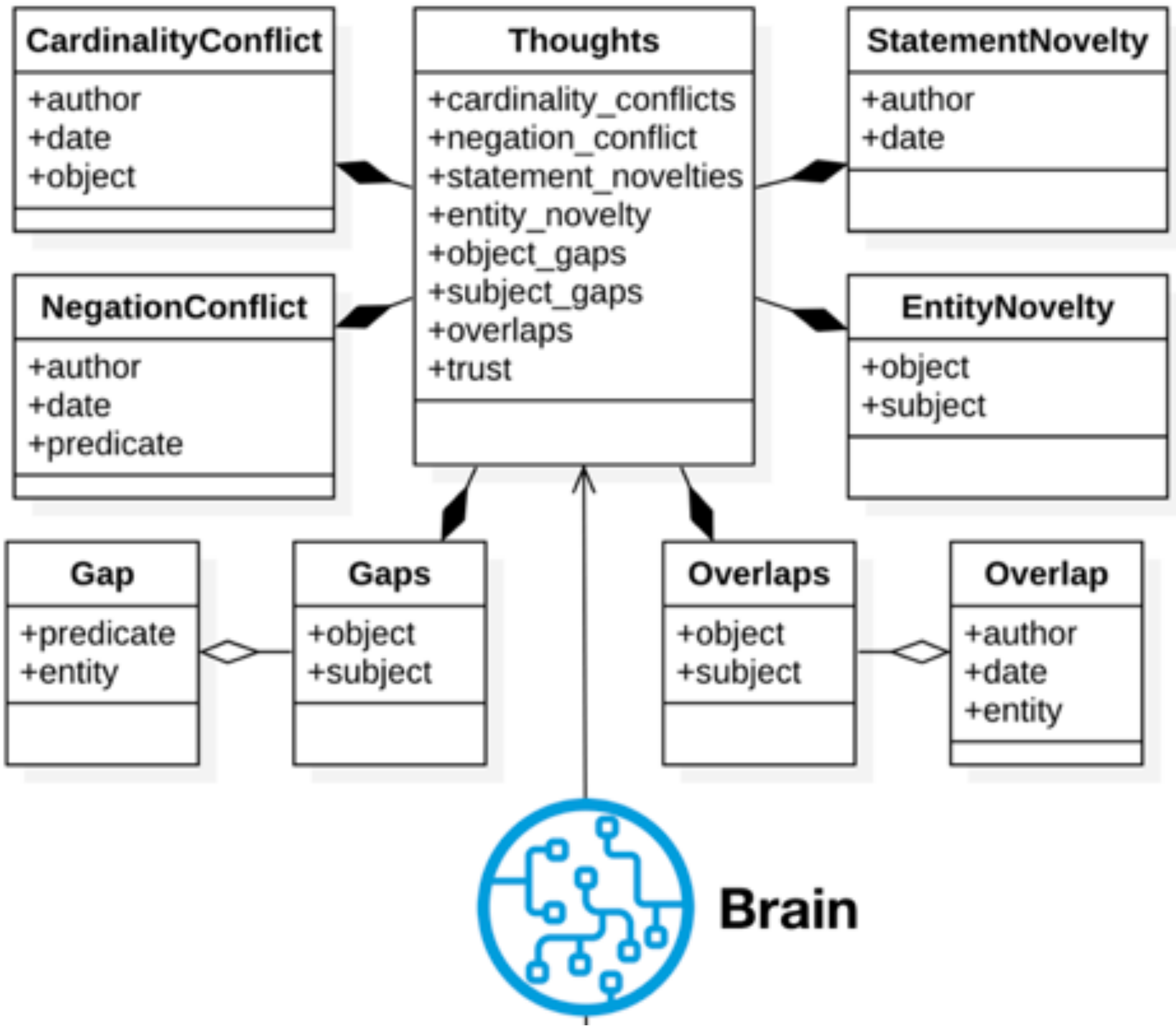


- Information conveyed at every turn is represented as an interaction graph.
- This is integrated into the existing episodic graph of the artificial agent.
- Focus on specific predefined semantic graph patterns arising in the integrated neighbourhoods of the resulting graph —> ***Thoughts***
- One ***thought***-pattern gets selected to respond to the user in the hope to acquire more knowledge.

Selene Baez, 2025, Knowledge-centered conversational agents with a drive to learn, PhD Thesis, Vrije Universiteit Amsterdam

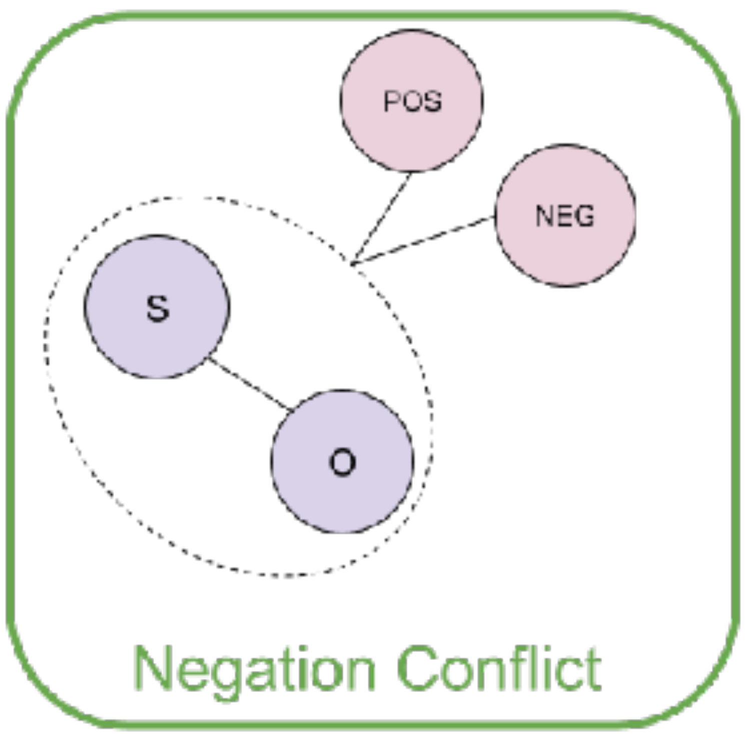
Vossen, 2025

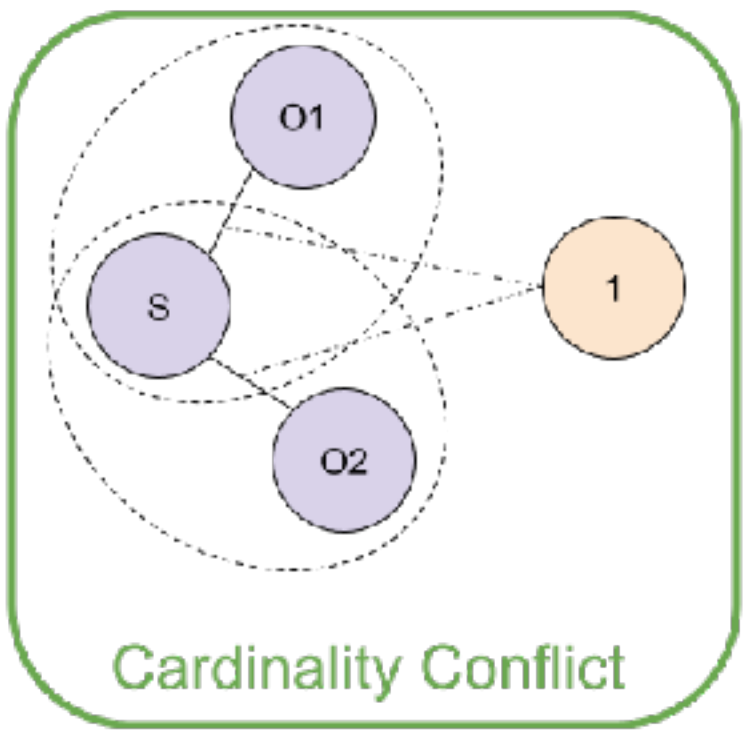
Reasoning over the state of knowledge



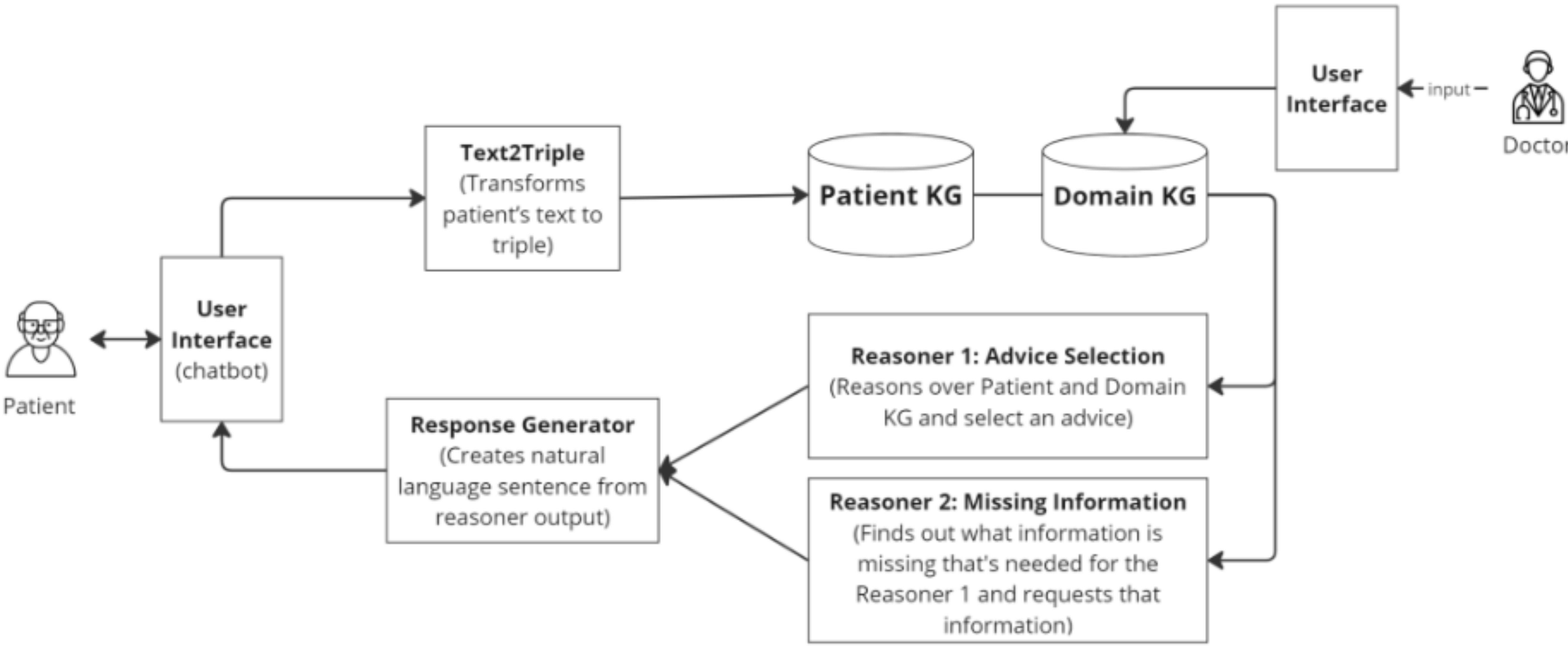
Graph patterns as ‘thoughts’

Goal	Thought type	Graph pattern	Example response
Correctness	Negation Conflict	lWorld:<CLAIM> grasp:PolarityValue grasp:Positive . lWorld:<CLAIM> grasp:PolarityValue grasp:Negative .	*”You say that Karla lives in Paris, but I have heard she does not”*
	Cardinality Conflict	n2mu:<PREDICATE> owl:cardinality "1"xsd:int . lWorld:<SUBJECT> n2mu:<PREDICATE> lWorld:<OBJECT1> . lWorld:<SUBJECT> n2mu:<PREDICATE> lWorld:<OBJECT2> .	*”I heard Karla lives in Amsterdam, not in Paris”*





Hybrid Intelligence use case



Hybrid Intelligence

Ontology Recommendations

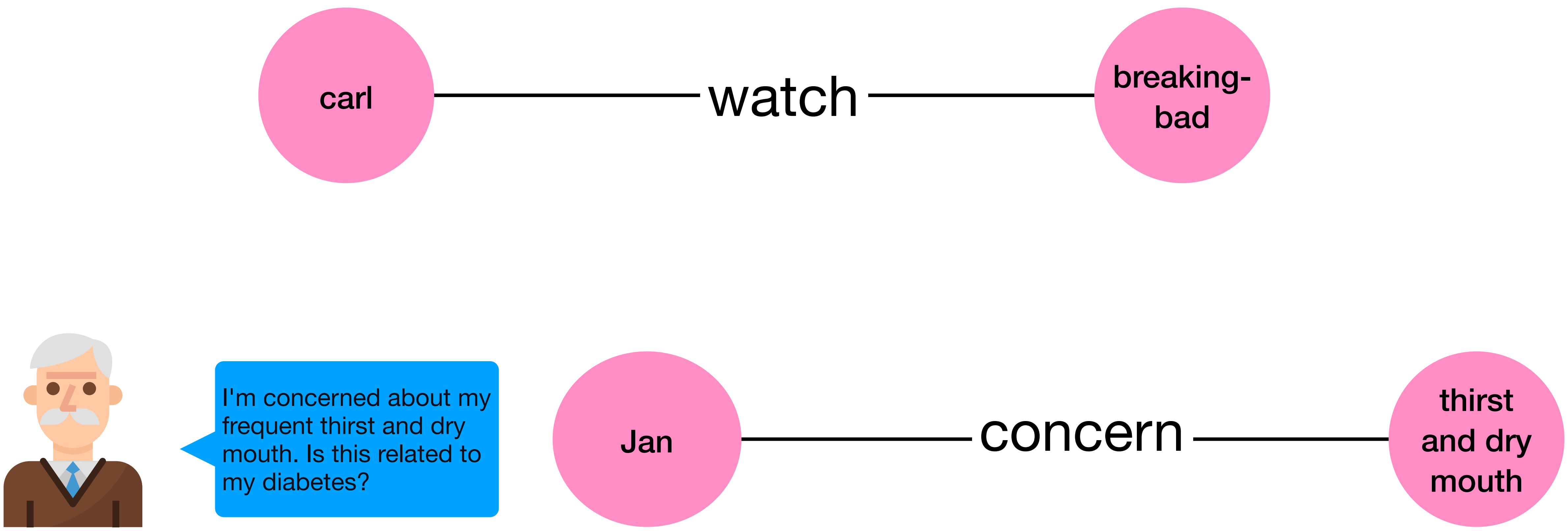
- symptoms that the patient reports having (e.g. blurry vision)
- regular self-reported blood sugar measurements (currently only measurements that I think are usually taken by doctors, e.g. FastingGlucose, HbA1c, TwoHourGlucose)
- the patient's current routine ("I usually take short walks in the park") to register changes in the routine and bring them up in conversation.
- commitments that the patient makes in conversation ("I will look into these exercises then"), could be helpful to follow-up and potentially remind the patient

Ontology Recommendations

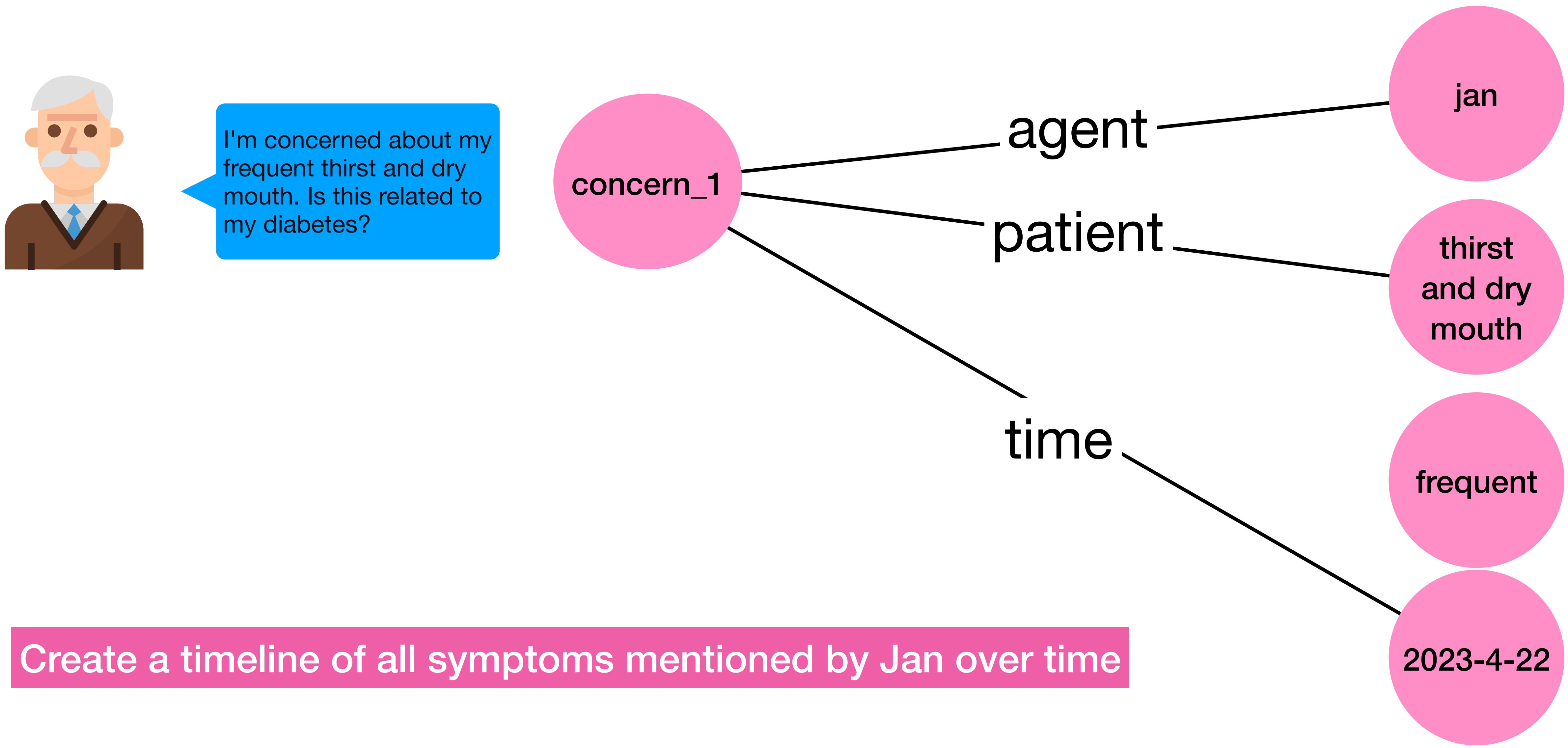
- reasons for why the patient deviates from the recommended lifestyle (e.g. “I know [Moroccan sweets] are high in sugar, but it’s a tradition that brings me joy”). These could potentially be subclasses of the “Challenge” class.

- Constraints are represented (e.g. TimeConstraint), consequences of the constraints are not directly modeled (e.g. “[The weather has] been quite unpredictable recently, so I haven’t been out as often as I’d like”). It could be nice to do that, to be able to query the graph for these connections.

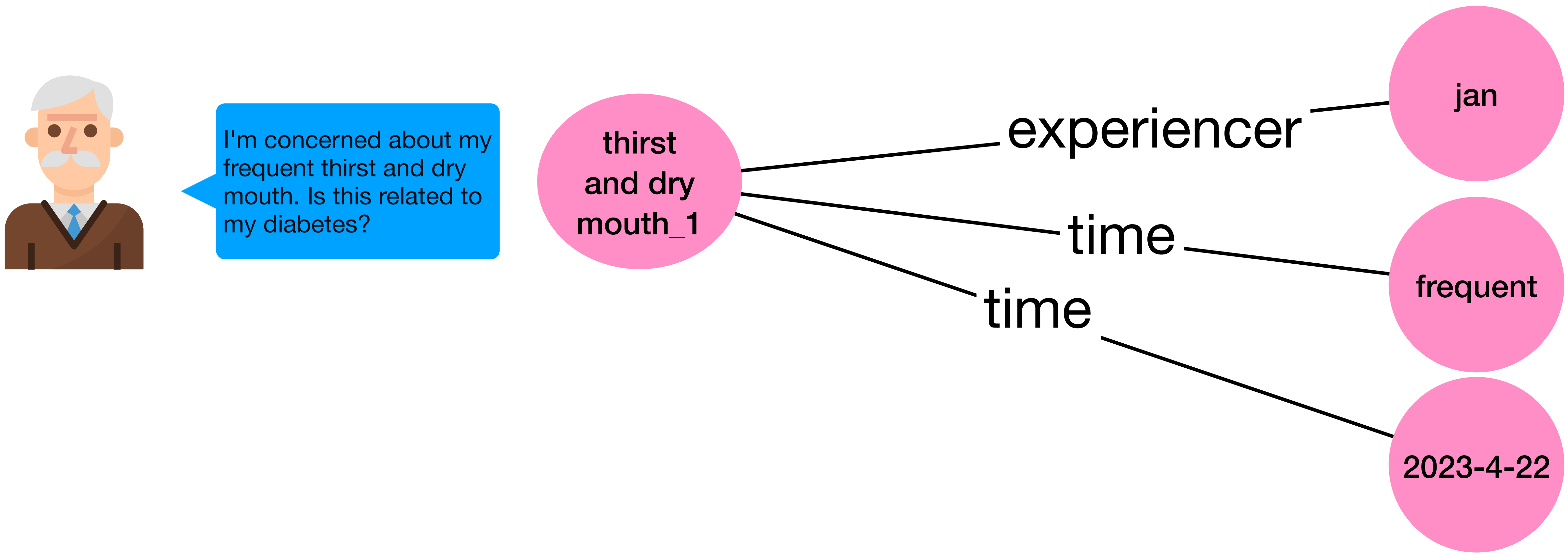
Entity-centric knowledge graphs



Event-centric knowledge graph for diabetes

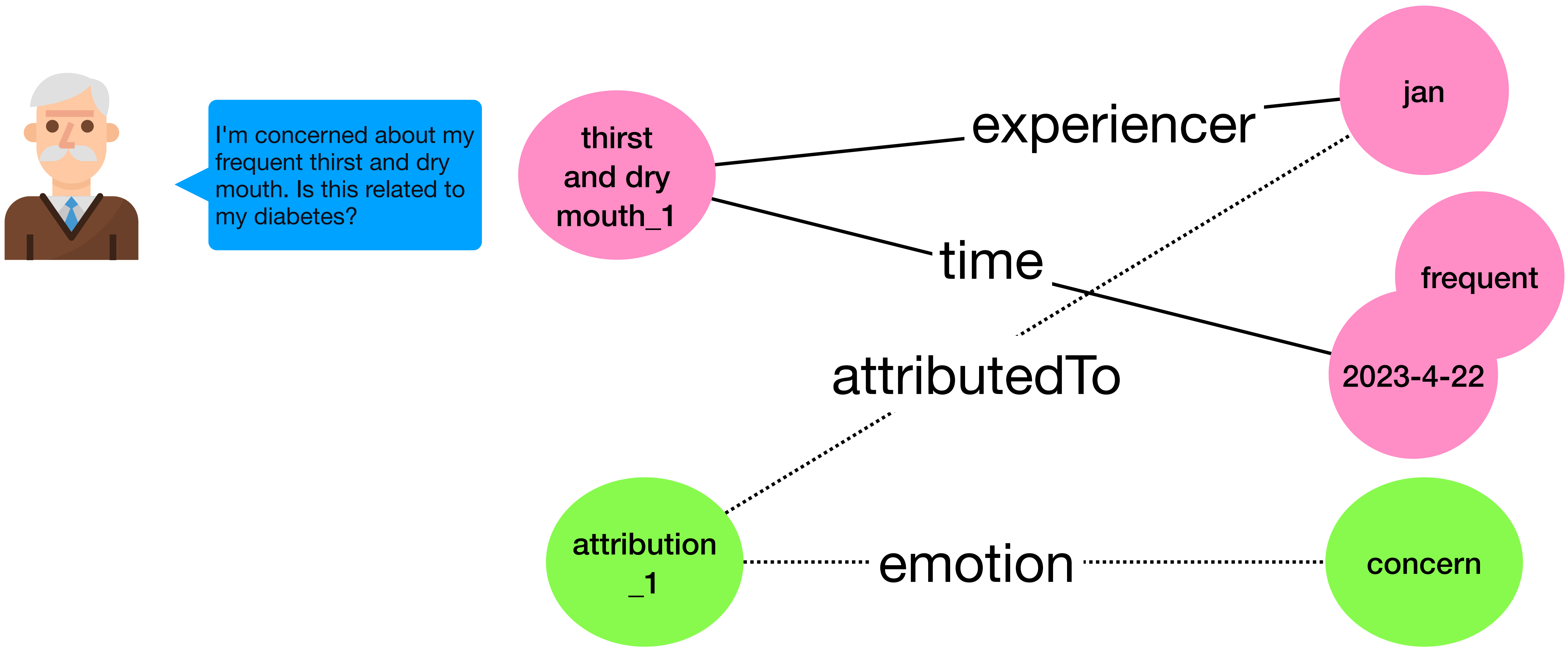


Event-centric knowledge graph for diabetes



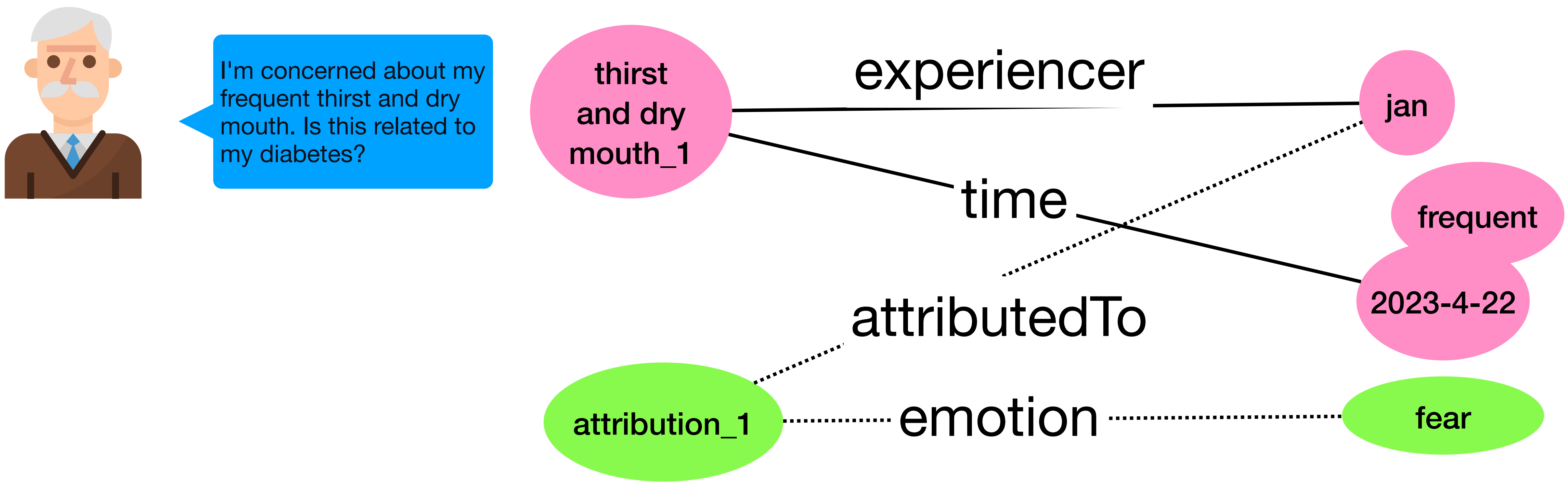
Create a timeline of all symptoms mentioned by Jan over time

Event-centric knowledge graph for diabetes



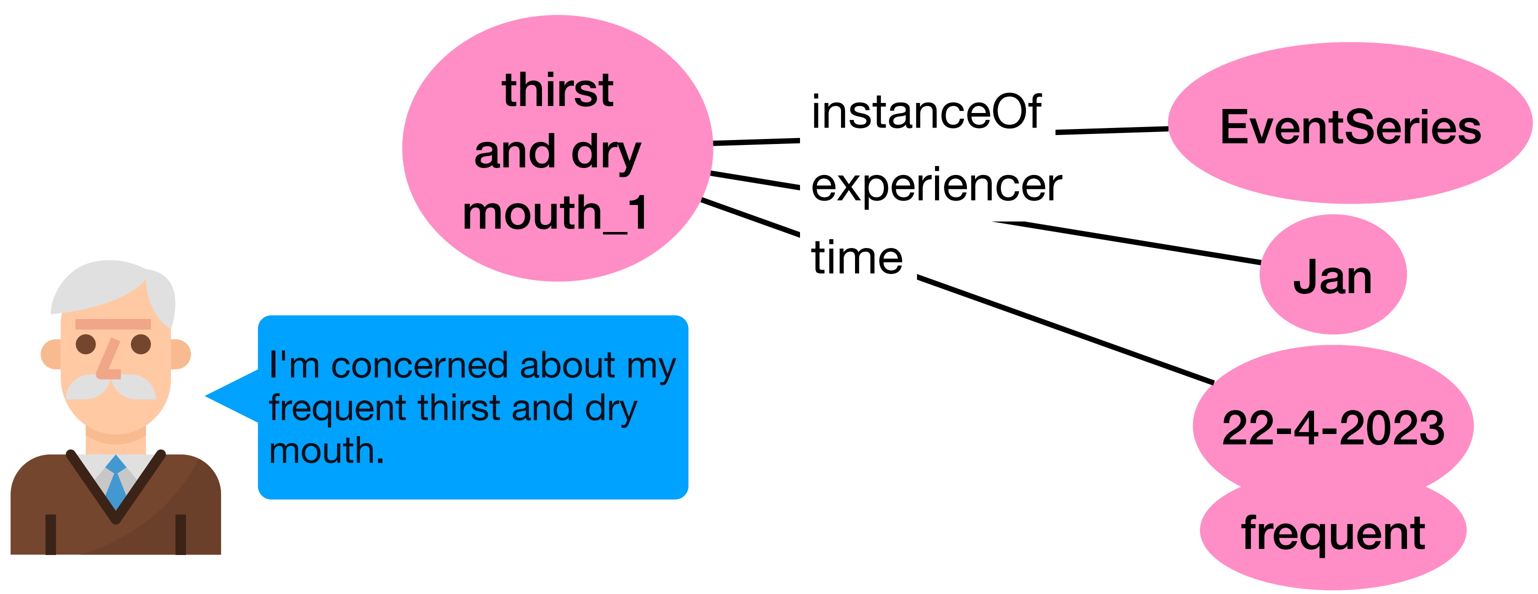
Create a timeline of all symptoms mentioned by Jan over time

Event-centric knowledge graph for diabetes



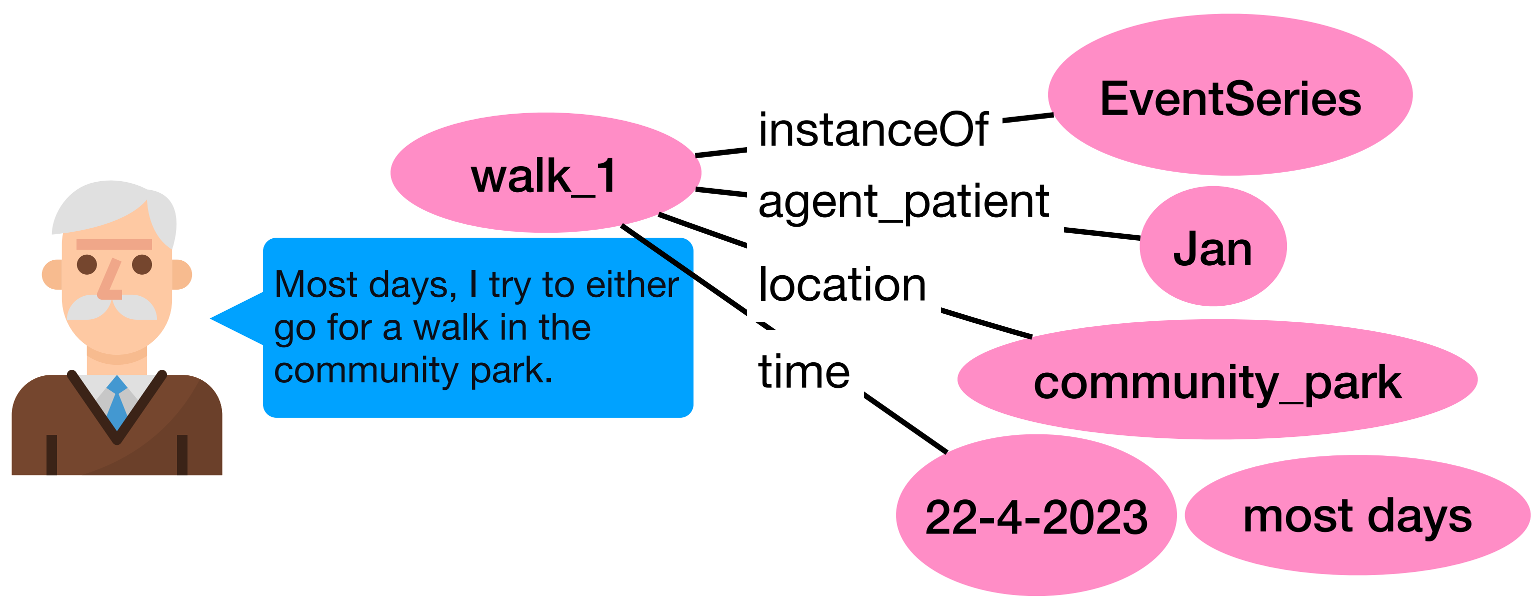
Create a timeline of all symptoms mentioned by Jan over time

Event-centric knowledge graph for diabetes



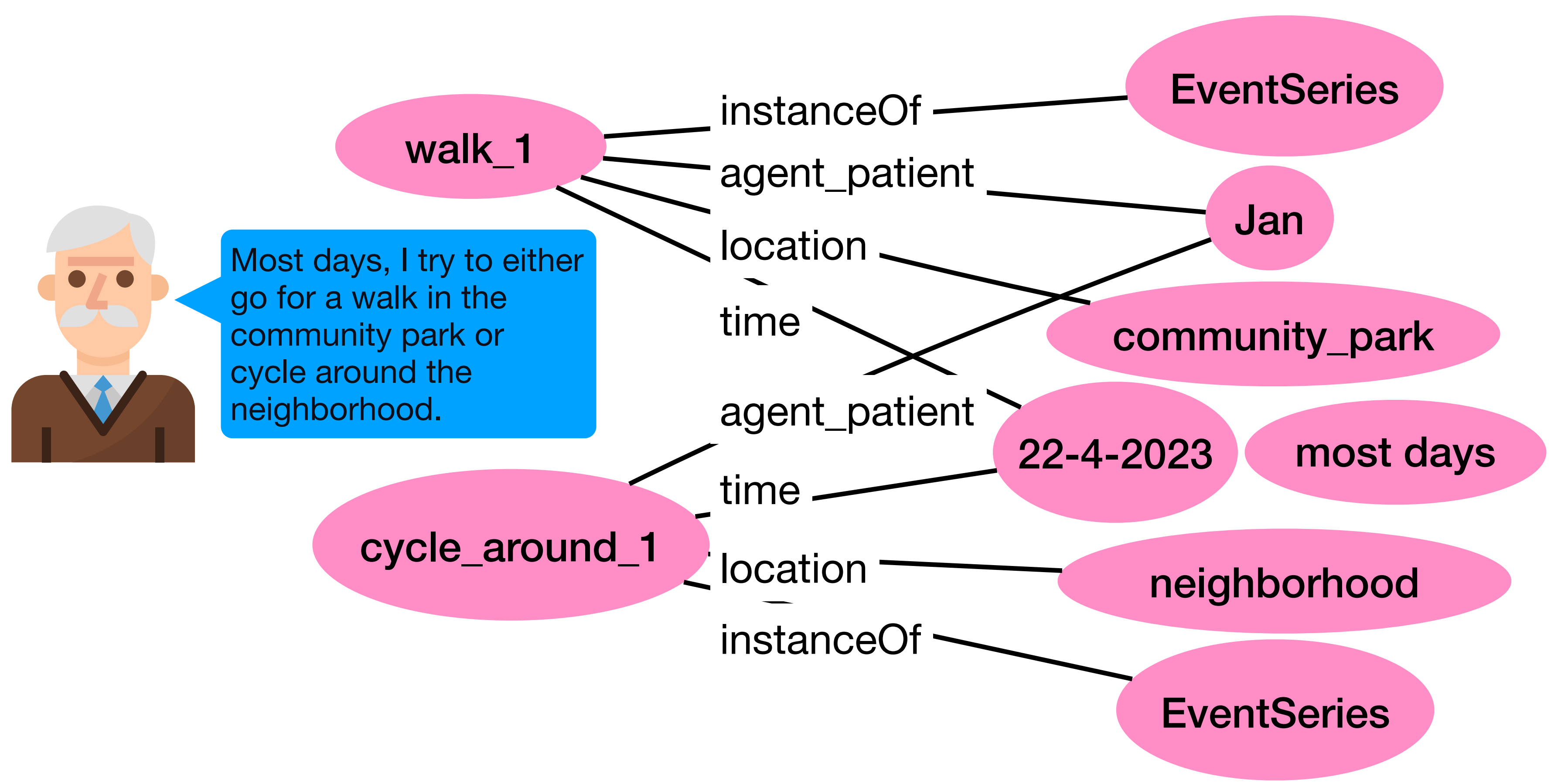
Create a timeline of all symptoms mentioned by Jan over time

Event-centric knowledge graph for diabetes



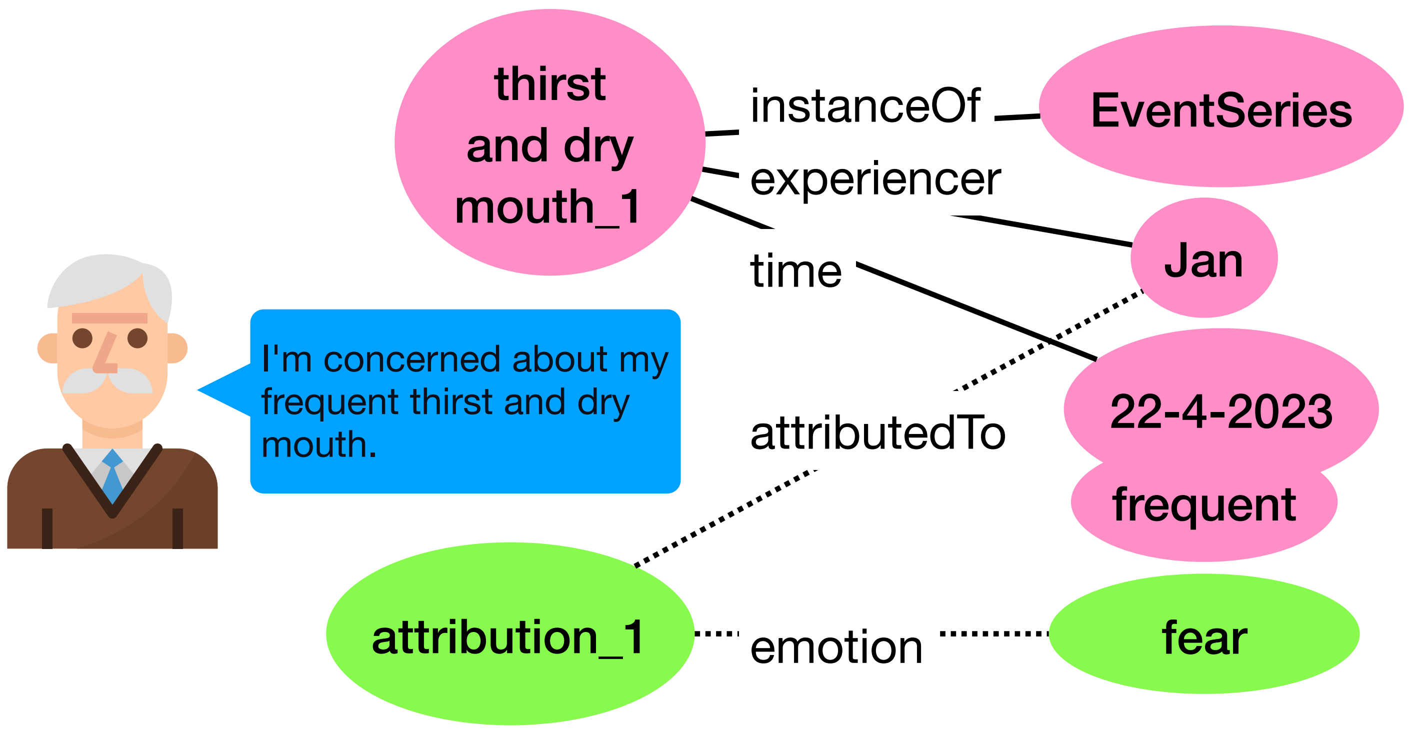
Create a timeline of all symptoms mentioned by Jan over time

Event-centric knowledge graph for diabetes



Create a timeline of all symptoms mentioned by Jan over time

Event-centric knowledge graph for diabetes

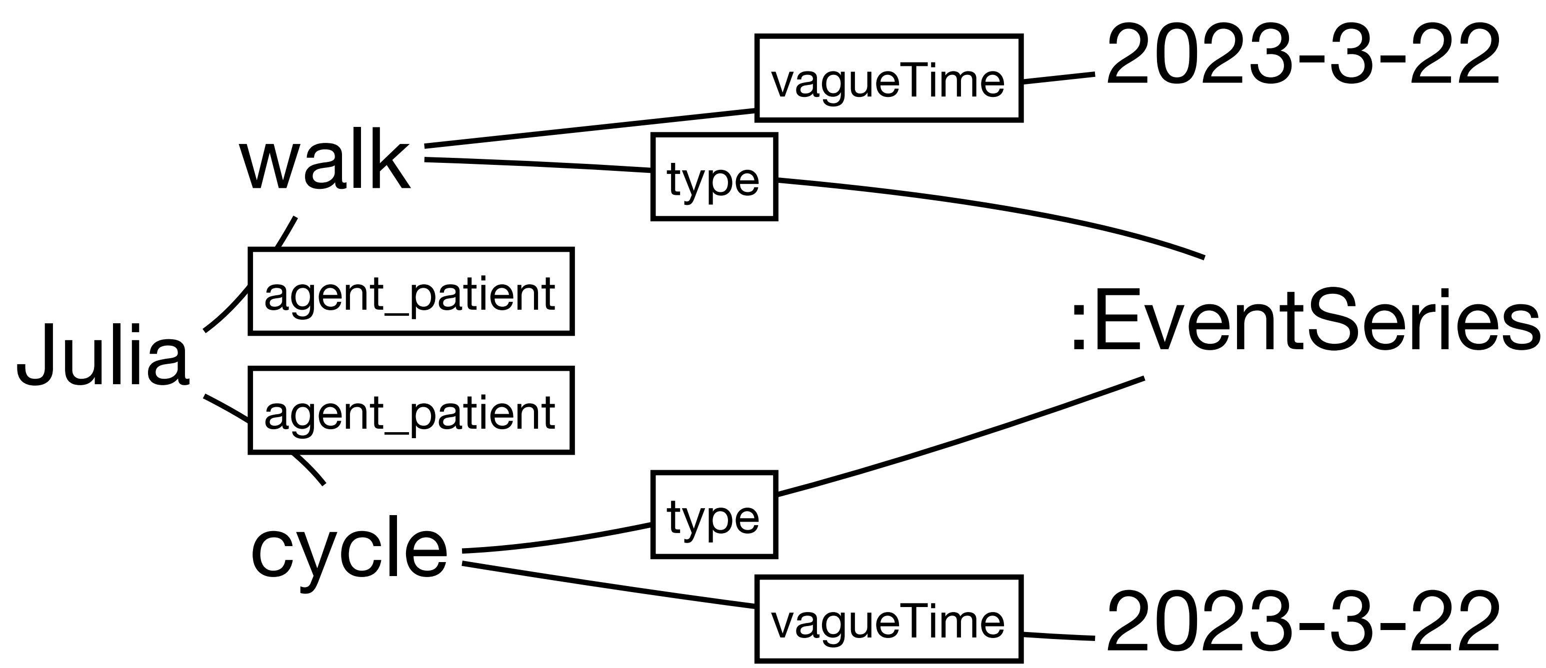


Create a timeline of all symptoms mentioned by Jan over time

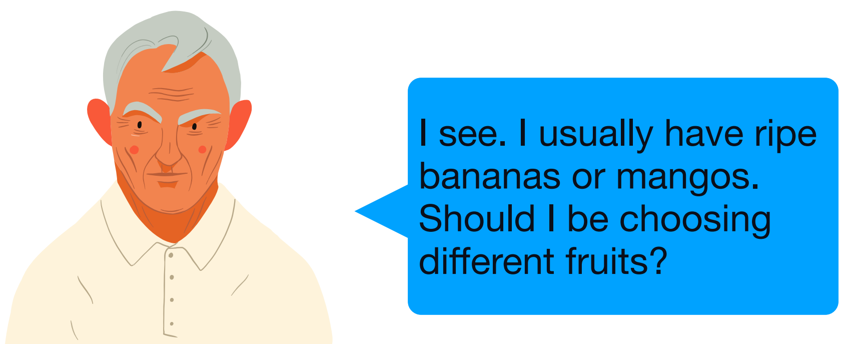
Grounding series in time

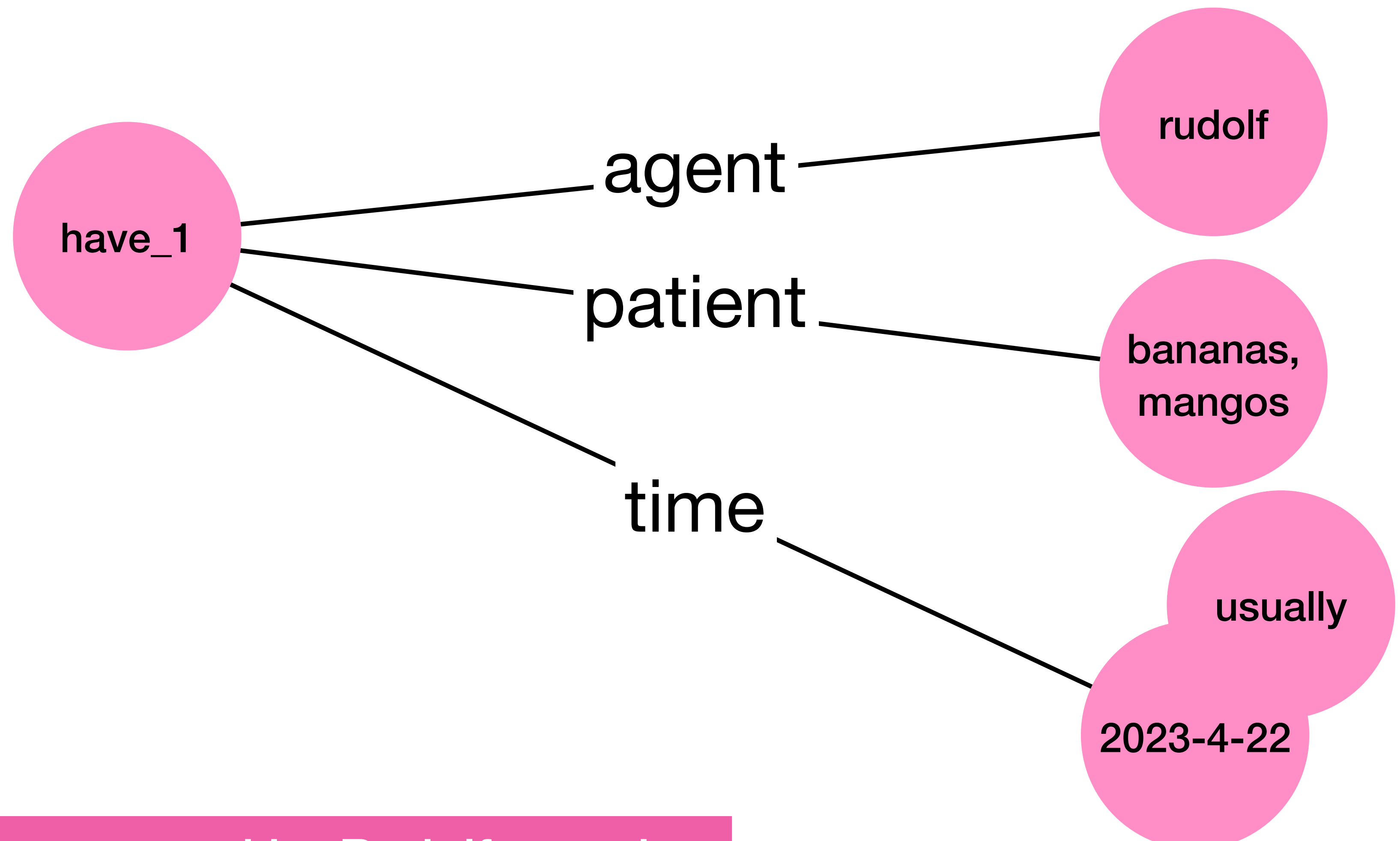
2023-4-22

Most days, I try to either go for a walk in the community park or cycle around the neighborhood.



Event-centric knowledge graph for diabetes





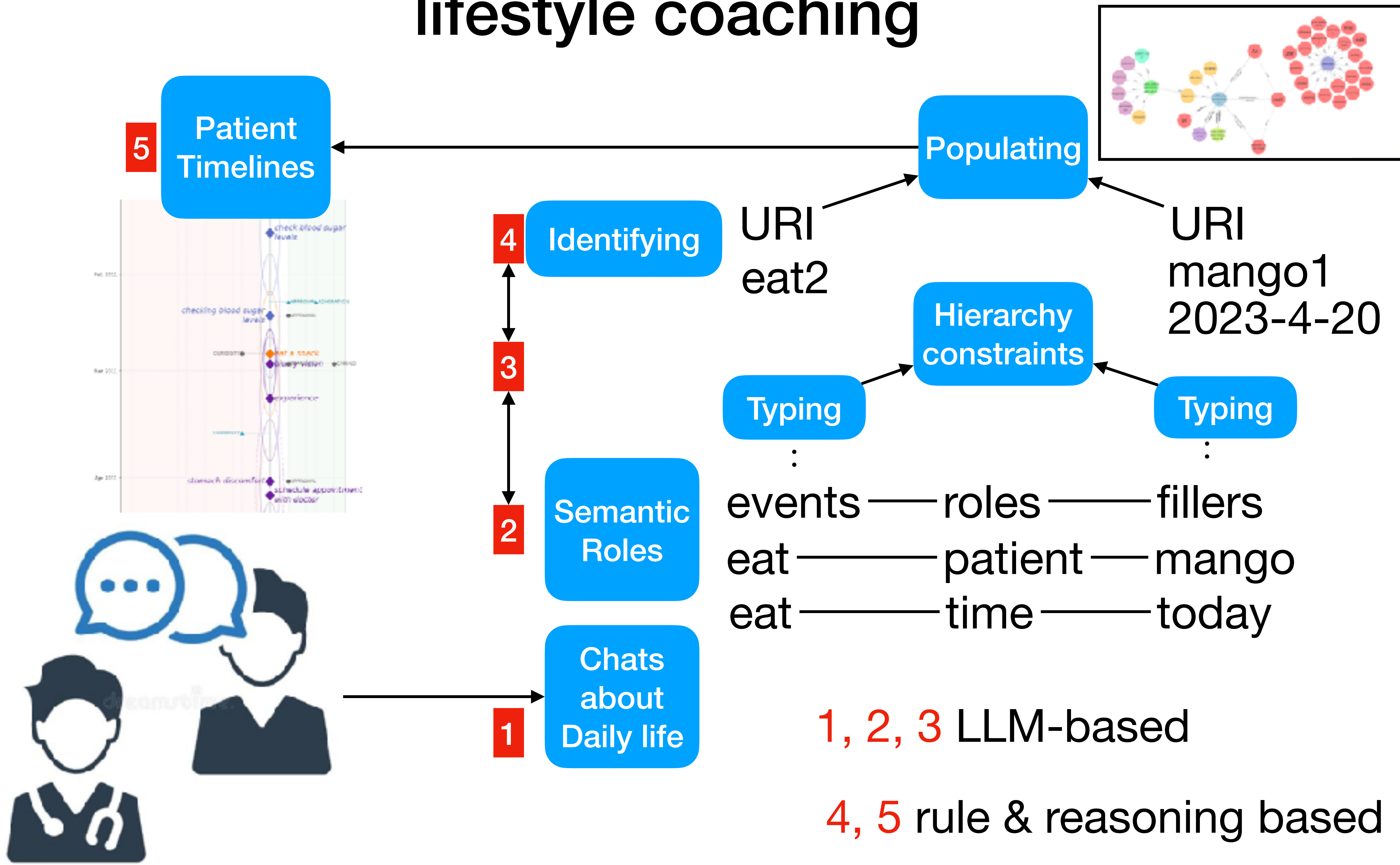
Create a timeline of all fruits consumed by Rudolf over time

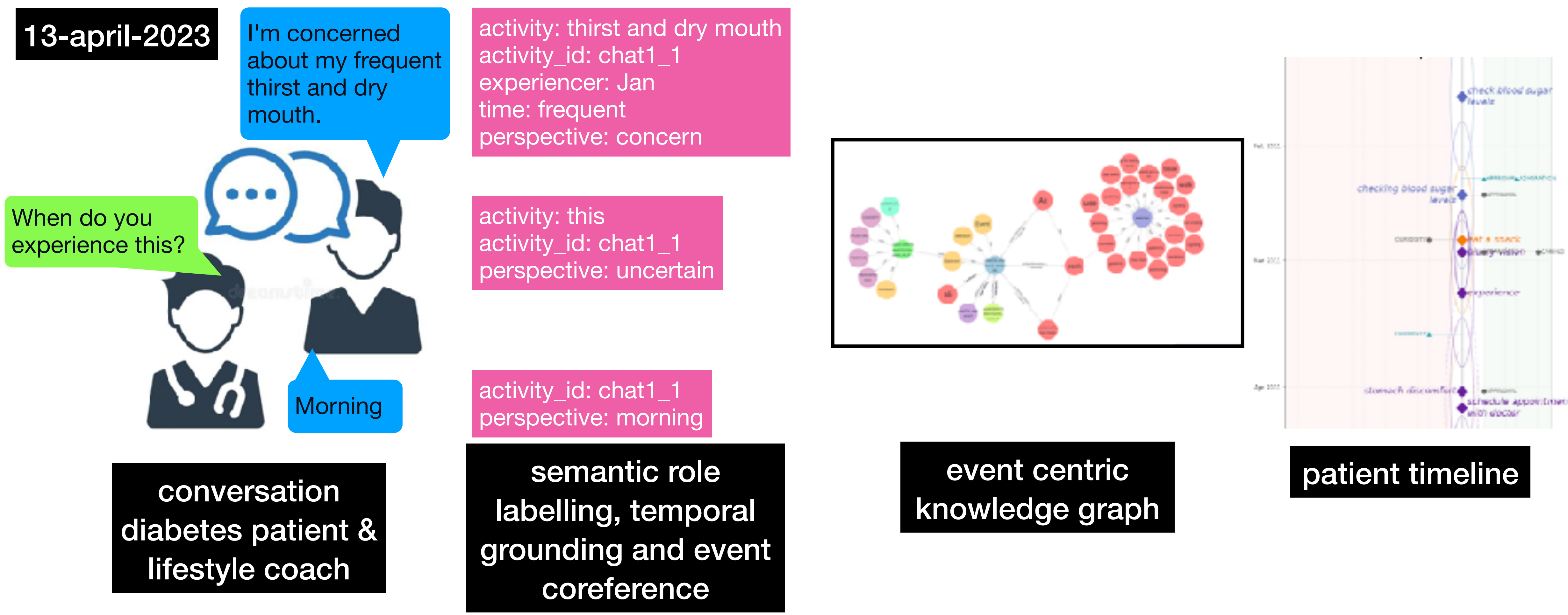
Event-centric

- Advantages:
  - Define specific properties of instances of events
  - Open world semantics depends on the meaning of predicates, closer to semantic-role labelling in NLP
  - Anything that is said, can be represented
  - Small number of predicates
- Disadvantages
  - Semantics remains vague, especially how to identify different events
  - More complex to query and reason
  - (Very) large number of events populated in the Knowledge Graph

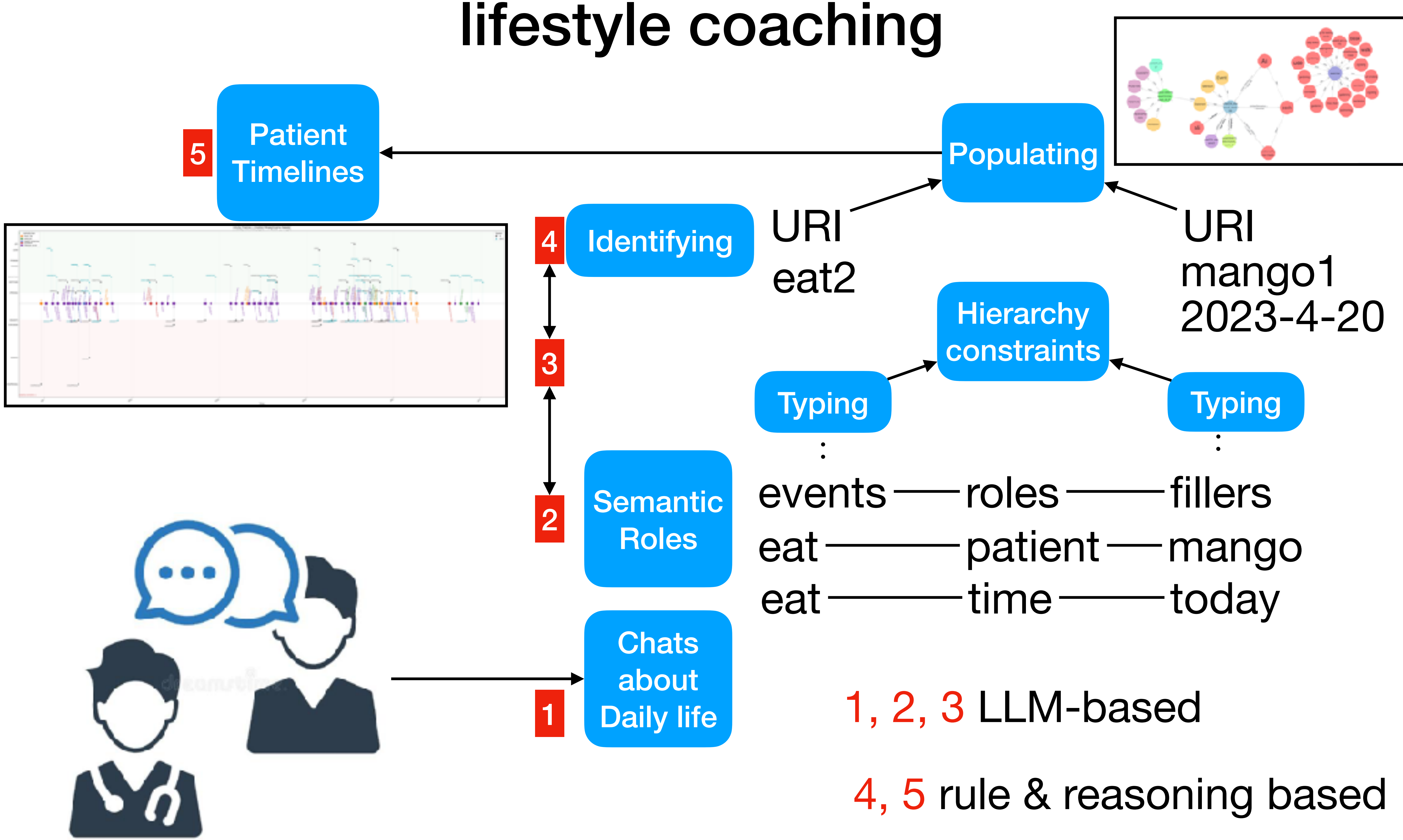
From open to closed world: bootstrapping the semantics of diabetes lifestyle coaching

From open to closed world: bootstrapping the semantics of diabetes lifestyle coaching




13-april-2023
I'm concerned about my frequent thirst and dry mouth.
When do you experience this?
Morning
activity: thirst and dry mouth
activity_id: chat1_1
experiencer: Jan
time: frequent
perspective: concern
activity: this
activity_id: chat1_1
perspective: uncertain
activity_id: chat1_1
perspective: morning
conversation diabetes patient & lifestyle coach
semantic role labelling, temporal grounding and event coreference
event centric knowledge graph
patient timeline

From open to closed world: bootstrapping the semantics of diabetes lifestyle coaching



1. Generating conversations

GPT-4-0613, June 2023, temp = 0.9, few-shot

Persona profiles → Biography → Conversation

Name	Gender	Ethnicity	Age
Jan	Male	Dutch	70
Maria	Female	Dutch	60
Abdullah	Male	Moroccan	70
Mohammed	Male	Moroccan	75
Aicha	Female	Moroccan	75
Julia	Female	Surinamese	55
Ingrid	Female	Surinamese	75
Johan	Male	Surinamese	60
Fatima	Female	Moroccan	65
Rudolf	Male	Surinamese	55
Ayşe	Female	Turkish	70
Pieter	Male	Dutch	75
Johanna	Female	Dutch	75
Fatma	Female	Turkish	65
Ali	Male	Turkish	60
Mehmet	Male	Turkish	65

Table 1: Selected Profiles

Figure 1: Example of conversation generation from a persona profile. The biography of "Pieter" (75-year-old Dutch male with Type 2 Diabetes) is used to generate a sample dialogue with the caretaker agent.

Stergios Ntanavaras, Maaike H. T. de Boer, Piek Vossen, 2026, A Synthetic Conversational Dataset for Type 2 Diabetes Management, CL4Health-worksop, LREC-2026, Mallorca

Github: https://github.com/stergios97/synthetic-conversational-data-diabetes-management

1. Generating conversations

Overview		Dialogue Acts	
Number of dialogues	256	Statements	77.8%
Total utterances	2,027	Comments	7.5%
Total tokens	52,019	Opinions	7.3%
Unique word types	5,337	Positive answers	2.8%
Avg. utt./dialogue	7.9	Commands	2.8%
		Back-channeling	0.7%
Emotions			
Neutral	785	Curiosity	432
Gratitude	211	Approval	142
Fear	33	Disappointment	11
Sadness	19		

Table 2: Dataset statistics, dialogue acts, and emotions.

(1 = Very Poor, 2 = Poor, 3 = Acceptable, 4 = Good, 5 = Excellent)

Dimension	Avg. Human Score ($\pm$SD)	α
Fluency	4.65 ± 0.54	-0.035
Naturalness	3.95 ± 0.89	0.020
Clarity	4.48 ± 0.77	0.021
Structure	4.54 ± 0.66	0.050

Table 5: Human evaluation over 20 dialogues and 10 expert raters. For each dimension, we report the mean and standard deviation (SD) across all ratings, alongside inter-rater agreement using Krippendorff's α (ordinal).

Stergios Ntanavaras, Maaike H. T. de Boer, Piek Vossen, 2026, A Synthetic Conversational Dataset for Type 2 Diabetes Management, CL4Health-worksop, LREC-2026, Mallorca

Conversation input


```
{"chat": 230,
  "human": "Ali",
  "date": "2013,Nov,29",
  "turns": [
 {"turn": 1,
     "speaker": "Ali",
     "utterance": "I've been needing more insulin lately, and I'm worried my diet might be the reason."},
  { "turn": 2,
     "speaker": "agent",
     "utterance": "It's possible, Ali. Have you been tracking your food intake to see any patterns with your blood sugar levels?"},
  { "turn": 3,
     "speaker": "Ali",
     "utterance": "Yes, I try to log everything my wife cooks, but I love Turkish food and it's hard to control portions."},
  {"turn": 4,
     "speaker": "agent",
     "utterance": "That's understandable. Traditional meals can be high in carbs and sugars. Consider substituting some ingredients with
healthier options to manage carb intake better."},
  {"turn": 5,
     "speaker": "Ali",
     "utterance": "Do you have any specific suggestions for Turkish dishes?"},
  {"turn": 6,
     "speaker": "agent",
     "utterance": "Certainly. For example, you can use whole grain bulgur instead of white rice and avoid adding too much sugar to
desserts like baklava. Balancing meals with plenty of vegetables and lean proteins can help, too."},
```

2. Activities and semantic roles

- Conversations are spread in time over 5 years.
- Prompted **gpt-3.5-turbo-0125** to apply semantic role labelling to each turn in a conversation:
  - Activity or Condition,
  - Activity type: exercise, take_food, take_drink, advise, social, diet, treatment, medicine, physical_condition, mental_condition, symptom, disease
  - Agent, Patient, Instrument, Manner
  - Place, Time
- Each conversation processed independently
- Turn-by-turn, keeping the history as context
- Prompted twice to derive the activity types relevant for the detected activity phrases

Prompt for activities

You will receive a conversation in JSON format between two speakers: a diabetes patient and a lifestyle coach.

The conversation contains the name of the diabetes patient and the date on which the conversation took place.

You need to extract activities and conditions of the diabetes patient from the last turn in the conversation. You can use the preceding turns as the context.

Only extract Activities of Daily Life in which the diabetes patient participates or physical, social or mental conditions of the patient in relation to the patient's lifestyle.

Try to represent as much details about the Activity or the Condition from the conversation.

For each activity or condition, extract the WHAT as activity, the TYPE of activity or condition as activity_type, the WHO as agent, the WHOM as patient, the HOW as manner, the WITH as instrument, WHERE as location and WHEN as time.

Always include the activity_type in the output, which can be either "exercise", "take_food", "take_drink", "advise", "social", "treatment", "physical condition", "social condition" or "mental condition".

Required Fields:

- activity: the activity or condition that the patient is doing or experiencing.
- activity_type: the specific type or category of the activity.
- agent: The entity performing the activity.

Optional Fields:

- patient: The entity affected by the activity.
- instrument: The tool or means used.
- location: Where the activity takes place.
- time: When the activity occurs.

Output the results in the specified JSON format. Do not output any other text than the JSON. The output JSON MUST include fields for the activity, activity_type and agent, and can optionally include patient, instrument, location and time if that information is present in the text.

<3 few shot examples>

SRL output

{ "chat": 230,

"date": "2013,Nov,29",

"human": "Ali",

"Input": {"turn": 1,

"speaker": "Ali",

"utterance": "I've been needing more insulin lately, and I'm worried my diet might be the reason."},

"Output": [

{"activity": "needing more insulin",

"activity_type": "treatment",

"agent": ["I"],

"patient": [“insulin"],

"time": ["lately"]},

{"activity": "worried",

"activity_type": "mental condition",

"agent": ["I"],

"patient": ["my diet"], }

]}

"Input": {

"turn": 2,

"speaker": "agent",

"utterance": "It's possible, Ali. Have you been tracking your food intake to see any patterns with your blood sugar levels?"},

"Output": [

{"activity": "needing more insulin",

"activity_type": "treatment",

"agent": [“Ali"],

"time": ["lately"]},

{"activity": "tracking food intake",

"activity_type": "diet",

"agent": [“Ali”],

"time": ["currently"]}]},

SRL statistics


```
{
    "name": "activities_1",
    "Total phrases": 1419,
    "Phrase counts": {
        "gardening": 31,
        "walking": 25,
        "exercise": 20,
        "managing diabetes": 19,
        "manage diabetes": 19,
        "managing blood sugar levels": 14,
        "quit smoking": 14,
        "quitting smoking": 14,
        "discuss": 13,
        "inquire": 13,
        "manage blood sugar levels": 12,
        "consult": 12,
        "monitoring blood sugar levels": 11,
        "experience": 11,
        "monitor": 11,
        "experiencing": 10,
        "manage": 10,
        "inform": 10,
        "walk": 9,
        "advise": 9,
        "conversation": 8,
        "notice": 8,
        "explain": 8,
        "ask": 8,
        "feeling dizzy": 8,
        "feeling thirsty": 7,
```

```
{
    "name": "agents_1",
    "Total phrases": 130,
    "Phrase counts": {
        "agent": 411,
        "Mehmet": 111,
        "Aicha": 109,
        "Ali": 109,
        "Johan": 108,
        "Rudolf": 106,
        "Pieter": 105,
        "Ingrid": 104,
        "Ay\u015fe": 104,
        "Fatma": 104,
        "Maria": 99,
        "Abdullah": 99,
        "Jan": 97,
        "Mohammed": 96,
        "Fatima": 96,
        "Julia": 92,
        "Johanna": 91,
        "Agent": 37,
        "doctor": 13,
        "healthcare provider": 8,
        "healthcare agent": 5,
        "Rudolf's sister": 4,
        "children": 4,
        "grandchildren": 4,
        "patient": 3,
        "church family": 3,
```

```
{
    "name": "patients_1",
    "Total phrases": 837,
    "Phrase counts": {
        "blood sugar levels": 215,
        "diabetes": 54,
        "diet": 44,
        "Maria": 35,
        "feet": 32,
        "blood sugar": 30,
        "meals": 28,
        "Jan": 27,
        "anxiety": 27,
        "Abdullah": 27,
        "Fatima": 26,
        "Pieter": 25,
        "Julia": 24,
        "Johanna": 24,
        "Mohammed": 23,
        "Ali": 23,
        "doctor": 22,
        "exercise routine": 21,
        "dizziness": 21,
        "sweets": 19,
        "Ay\u015fe": 19,
        "Rudolf": 17,
        "Mehmet": 17,
        "diabetes management": 16,
        "fruits": 16,
```

SRL statistics


```
{
    "name": "instruments",
    "Total phrases": 1160,
    "Phrase counts": {
        "conversation": 34,
        "diet": 32,
        "vegetables": 32,
        "exercise": 31,
        "whole grains": 31,
        "portion control": 29,
        "balanced diet": 22,
        "fruits": 21,
        "medication": 20,
        "walking": 19,
        "reminders": 15,
        "blood sugar monitoring": 13,
        "meditation": 13,
        "chair yoga": 11,
        "physical activity": 10,
        "nicotine patches": 10,
        "lean proteins": 10,
        "discussion": 10,
        "regular exercise": 9,
        "monitoring blood sugar levels": 9,
        "protein": 9,
        "light stretching": 9,
        "medications": 8,
        "nicotine replacement therapy": 8,
        "legumes": 8,
```

```
{
    "name": "manners",
    "Total phrases": 1040,
    "Phrase counts": {
        "informative": 50,
        "regularly": 47,
        "supportive": 41,
        "consistently": 22,
        "challenging": 19,
        "helpful": 17,
        "relaxing": 17,
        "gently": 15,
        "daily": 14,
        "carefully": 13,
        "moderation": 12,
        "lately": 11,
        "concerned": 11,
        "effectively": 11,
        "light": 11,
        "in moderation": 11,
        "balanced": 10,
        "supportively": 10,
        "slowly": 10,
        "encouraging": 9,
        "gentle": 9,
        "mindfully": 8,
        "recently": 8,
        "proactively": 8,
        "difficult": 7,
```

```
{
    "name": "locations",
    "Total phrases": 266,
    "Phrase counts": {
        "home": 50,
        "family gatherings": 37,
        "indoors": 32,
        "garden": 25,
        "at home": 20,
        "park": 14,
        "community park": 13,
        "feet": 11,
        "Rotterdam": 11,
        "outside": 10,
        "indoor": 7,
        "community gatherings": 7,
        "special occasions": 7,
        "": 6,
        "local community center": 6,
        "Amsterdam": 6,
        "neighborhood park": 6,
        "outdoors": 5,
        "Home": 5,
        "legs": 5,
        "social events": 5,
        "neighborhood": 5,
        "conversation": 5,
        "not specified": 5,
        "at night": 5,
```

```
{
    "name": "times",
    "Total phrases": 609,
    "Phrase counts": {
        "lately": 183,
        "recently": 137,
        "ongoing": 56,
        "daily": 34,
        "future": 27,
        "present": 27,
        "regularly": 21,
        "after meals": 15,
        "currently": 14,
        "usually": 13,
        "during Ramadan": 13,
        "during the conversation":
        "current": 12,
        "morning": 12,
        "Lately": 11,
        "throughout the day": 11,
        "sometimes": 10,
        "during family gatherings":
        "often": 8,
        "Ramadan": 8,
        "this week": 8,
        "evenings": 8,
        "turn 3": 7,
        "today": 7,
```

2. Sources & Perspectives

GO-Emotion classifier: “bhadresh-savani/bert-base-go-emotion”,
https://github.com/google-research/google-research/tree/master/goemotions

emotion	Abdullah	Aicha	Ali	Ayşe	Fatima	Fatma	Ingrid	Jan	Johan	Johanna	Julia	Maria	Mehmet	Mohammed	Pieter	Rudolf	agent
admiration	2	8	3	3	2	3	5	1	1	3	1	2	3	1	7	2	143
amusement	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	2
anger	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
annoyance	0	0	0	0	0	0	0	0	1	1	0	0	0	2	0	0	0
approval	11	14	8	10	11	6	13	10	15	12	10	16	9	14	15	8	256
caring	5	5	6	12	7	11	4	11	8	2	3	4	6	3	7	4	236
confusion	2	2	3	1	0	4	7	8	1	1	5	2	7	2	4	8	7
curiosity	29	19	29	22	29	24	27	20	22	25	27	21	18	27	33	26	266
desire	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
disappointm	0	1	0	0	0	1	0	1	0	2	1	1	0	1	1	2	0
disapproval	1	1	2	2	0	0	0	2	0	0	0	1	0	2	2	1	0
disgust	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
fear	1	3	2	5	3	3	2	0	2	1	1	1	3	3	3	1	0
gratitude	11	6	9	11	16	11	12	11	10	12	14	9	11	8	9	14	74
joy	2	3	2	0	0	4	2	0	3	1	1	3	2	2	2	1	31
love	0	2	2	2	0	3	3	0	0	0	0	3	0	1	2	2	9
nervousness	1	1	1	3	1	2	0	0	1	0	0	0	2	1	3	0	0
optimism	0	1	1	1	0	1	1	0	1	0	1	2	1	3	0	1	13
realization	2	1	0	0	1	1	1	2	0	2	2	3	2	0	1	1	0
remorse	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
sadness	3	5	1	0	1	3	1	2	4	1	2	3	3	4	2	3	1
surprise	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	1

3. Activity types & roles

Prompted **gpt-4o-mini** to assign types to the role fillers

Activity types	
advise	629
diet	426
disease	2
exercise	1050
measurement	168
medicine	4
mental condition	259
physical condition	676
social condition	10
social	84
symptom	88
take_drink	35
take_food	421
treatment	196

Filler types and roles				
type	**agent**	**patient**	**instrument**	**location**
activity	9	97	97	41
body_part	25	304	24	184
concept	26	246	65	50
device	0	19	77	19
food	14	628	177	61
location	1	5	7	320
meal	0	2	0	0
medication	2	44	26	0
other	372	78	87	55
person	3624	266	11	7
quantity	0	41	9	9
time	0	2	0	6


Event-Centric Knowledge Graph — Concept Hierarchy with Frequency Counts and Examples
owl:Thing
ActivityType
EntityType
TemporalType
Exercise n=1,050
Condition n=945
Advise n=629
Diet n=882
Social n=84
Medical n=458
Person n=3,908
Food n=880
Other n=592
Body Part n=537
Concept n=387
Location n=333
Vague n=1,019
Recurring n=716
Range n=622
Point n=19
jan, Alcha
baklava, dates
blood sugar
monitoring
indoor, Dutch
recently, often
daily, weekly
this week
today, now
Physical Cond. n=676
Mental Cond. n=259
Social Cond. n=10
Diet (general) n=426
Take Food n=421
Take Drink n=35
Treatment n=196
Measurement n=168
Symptom n=88
Medicine n=4
Disease n=2
Activity n=244
Device n=115
Medication n=72
Quantity n=59
Time n=8
Meal n=2
feel, tired
worry, stress
smoking
diet, eat
eat, have
drink, tea
take, quit
check, monitor
fatigue, thirst
medication
nerve damage
brisk walking
glucometer
insulin
15g carbs
morning
dinner
Activity group (intermediate)
Activity type (leaf)
Entity role type
Temporal type

4. Activity instances

- How to distinguish one event from the other?
- Baseline:
  - One mention = One instance / conversation
  - Grounded to the date of the utterance
  - 1642 instances of activities
- Challenges:
  - Differentiate past, now and future, relative to the time of the utterance
  - Paraphrases and pronominal references, “cycle”, “cycling”, “it”
  - Coordination (cycled Sunday and Monday) and quantification (cycles twice)
  - Vague expressions: lately, regularly, sometimes, last week, most days

4. Grounding activities in time

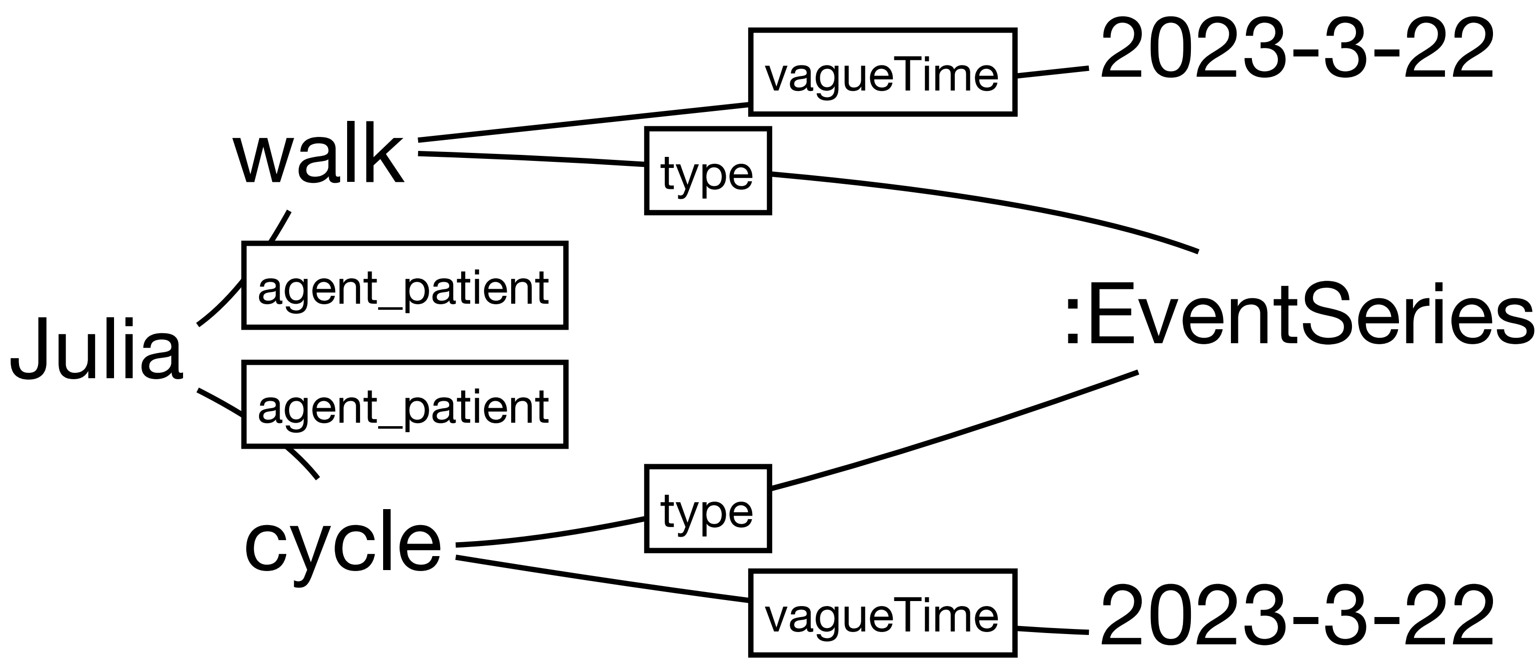
- 2376 time expression annotations across 641 distinct expressions
- 19 ***points*** (0.8%): “now”, “today”, “morning”
- 622 ***ranges*** (26%): “recently”, “last week”, “lately”
- 716 ***recurring***: “throughout the day”, “regularly”, “before exercising”, “usually for about 30 minutes”, “every day”
- 1019 ***vague***: "in the meantime”, “ongoing”, “especially in the evenings”, “sometimes”, “frequent”

Events as series

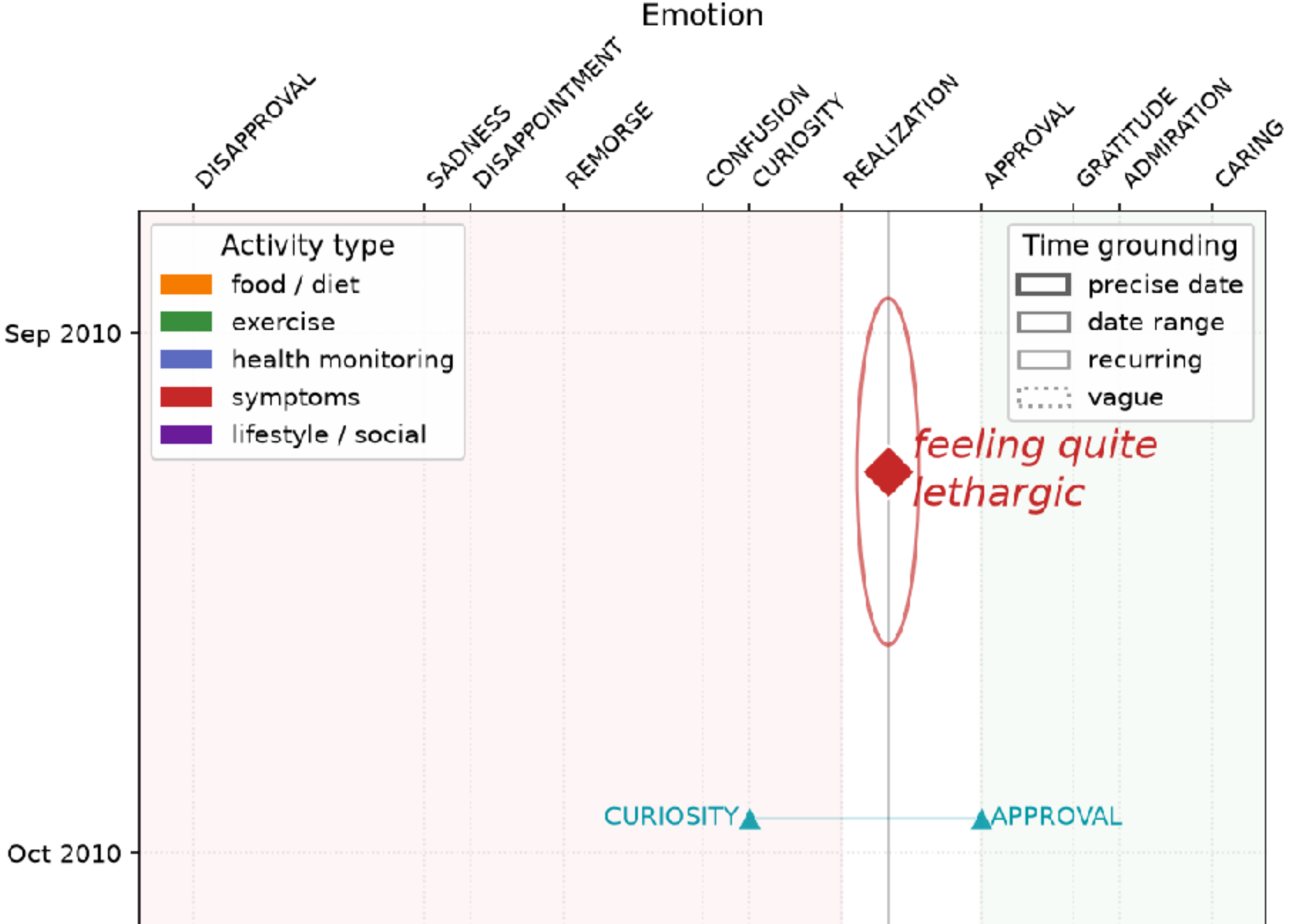
- https://schema.org/EventSeries
- Underspecification of events:
  - walking_series a n2mu:EventSeries, sem:Event,
    - dateTime, recurringTime, vagueTime
    - frequency: rare, regular, frequent
    - sem:hasSubevent

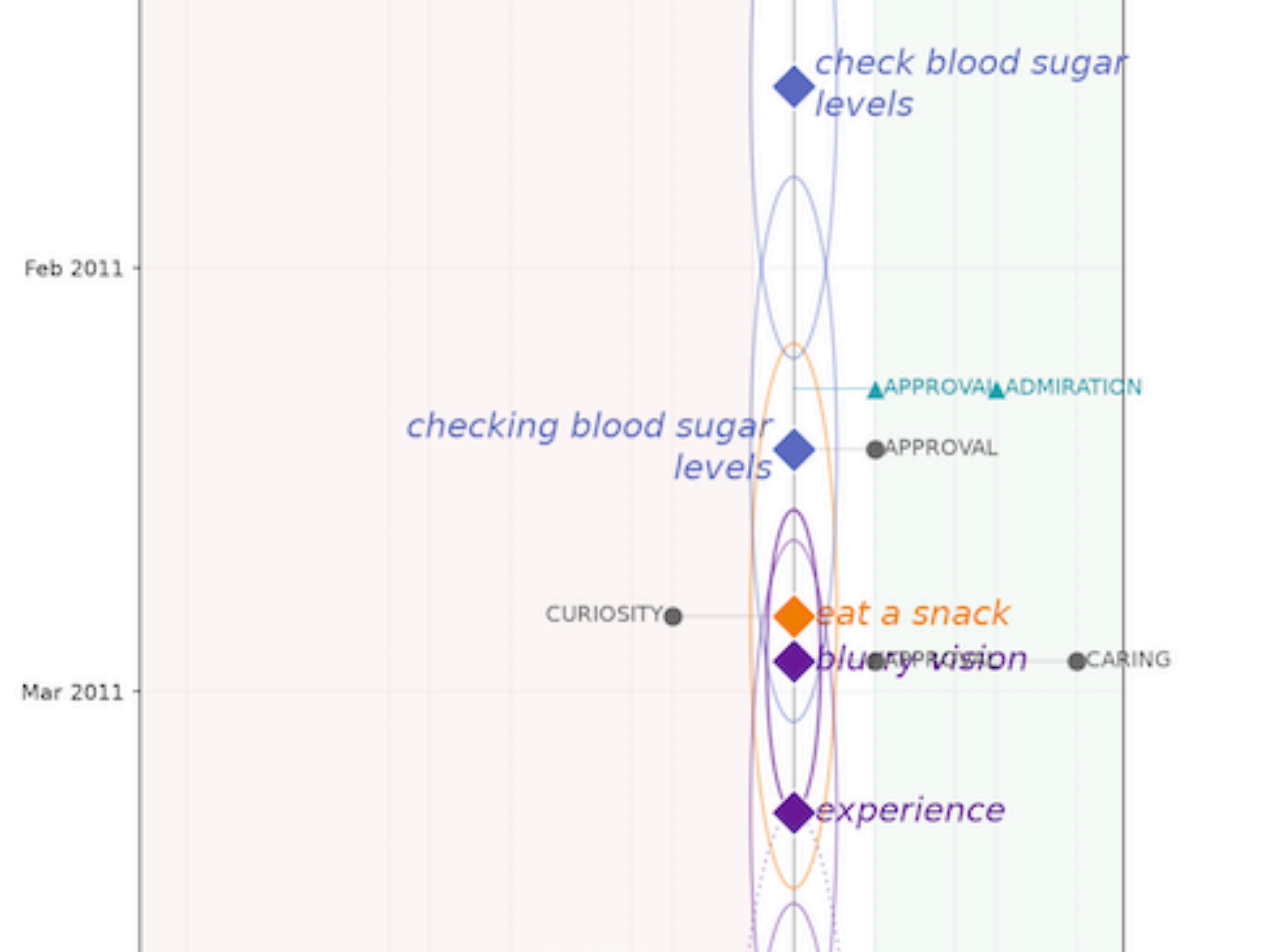
Grounding series in time

2023-4-22

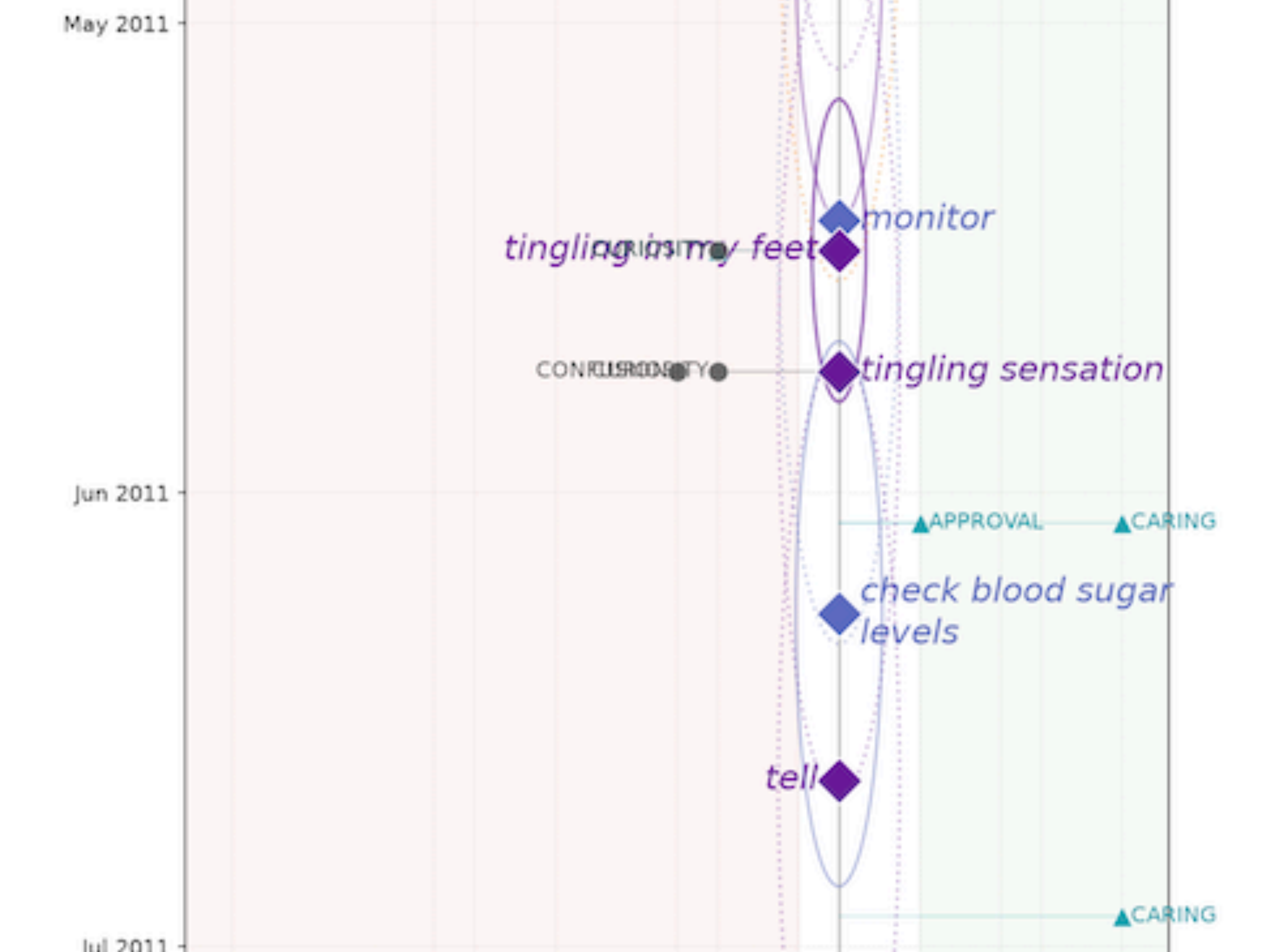


Activity Timeline — Emotion Perspectives by Speaker
ctivity (coloured by type) ○ = Rudolf △ = agent oval = grounded time (size encodes





Feb 2011
Mar 2011
check blood sugar levels
APPROVAL
ADMIRATION
checking blood sugar levels
APPROVAL
CURIOSITY
eat a snack
blurry vision
APPRAISAL
CARING
experience


May 2011
monitor
tingling in my feet
tingling sensation
Jun 2011
APPROVAL
CARING
check blood sugar levels
tell
CARING
Jul 2011

Open issues

- event instances remain an unsolved problem
- quantification:
  - how much of what type of things consumed
  - how long and how intensely exercised
  - degree and frequency of symptoms
- More closure is needed

Conclusions

- Semantic modeling of diabetes patient’s lifestyle
- Pro-active agent that responds given the explicit reasoning over the life style
- Natural conversations instead of synthetic
- Closure over semantics: what, how much, how frequent, constraints on event types and slot fillers
- Prompt LLMs with semantic model as closure
- Evaluation: semantic roles, instances of events, reasoning over timelines
- Responses by the agent and not by the LLM