



Streams and graphs

Behind the scenes of building a doctor in your pocket

Domenico Corapi

Senior Software Engineer at Babylon Health – AI Cognitive



Hi Alex, how can I help?

I've got a really bad headache
and I don't know what to do...



No problem, let me ask
you a few questions



Chatting to your virtual doctor

- In this presentation 😊
 - Healthcare data modelling and **knowledge bases**
 - **Connecting data** about patients/conditions semantically using a *DSE Graph*
 - **Event sourcing** through *Apache Kafka*
- Not in this presentation
 - Chatbot architecture and natural language processing
 - Symptom triage

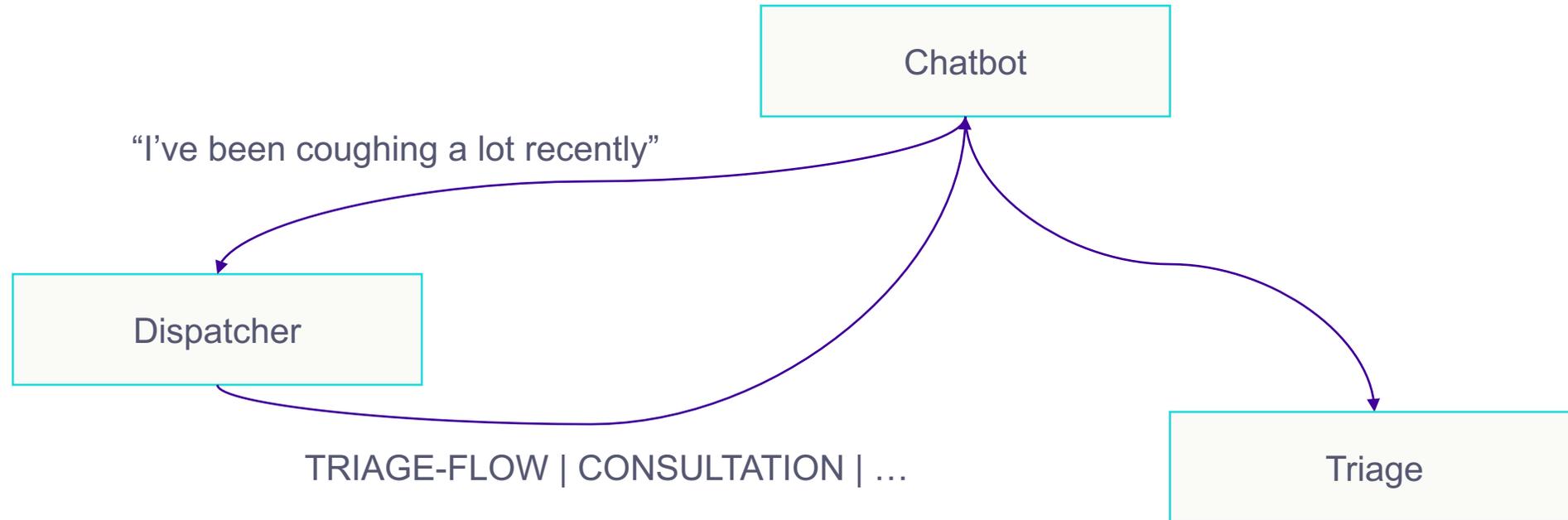
Babylon, I'm sick...

Hello, how can I help you?

I have been coughing a lot recently

Ask Babylon

From text to triage



Babylon, I'm sick...

Firstly, I'd like to check that I understand you correctly. Is your main symptom: **Cough?**

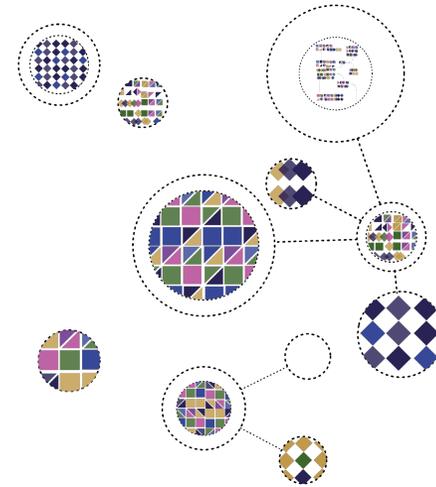
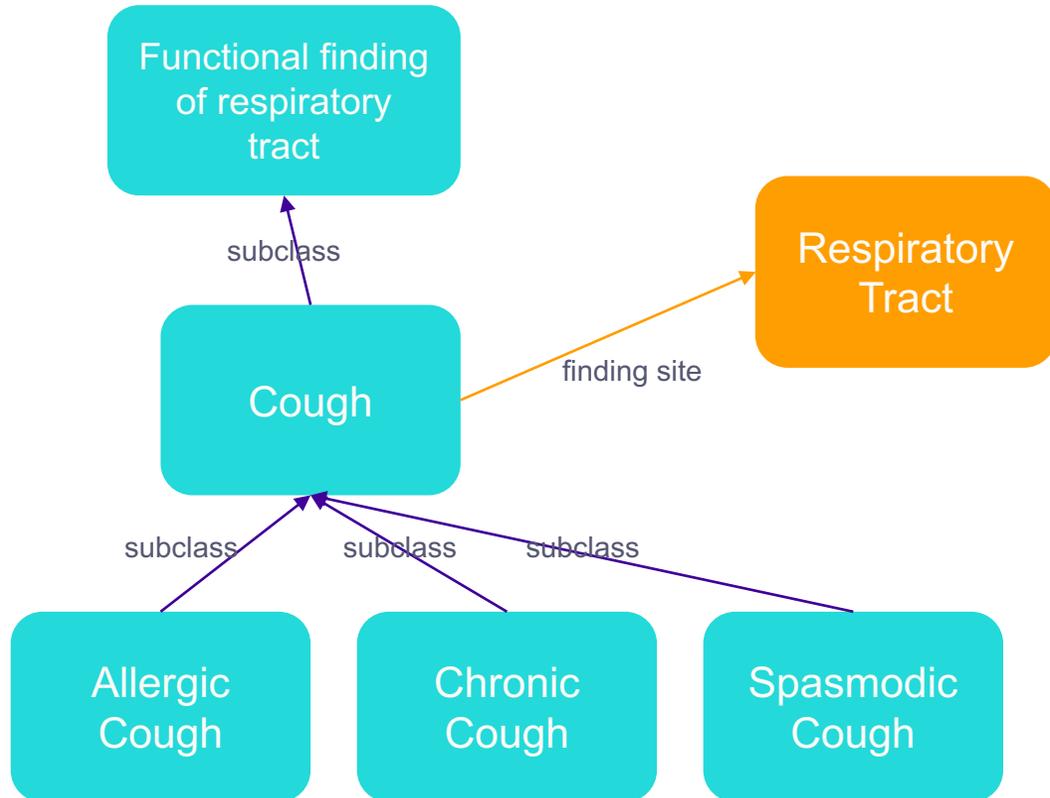
Yes



No



Cough as a concept

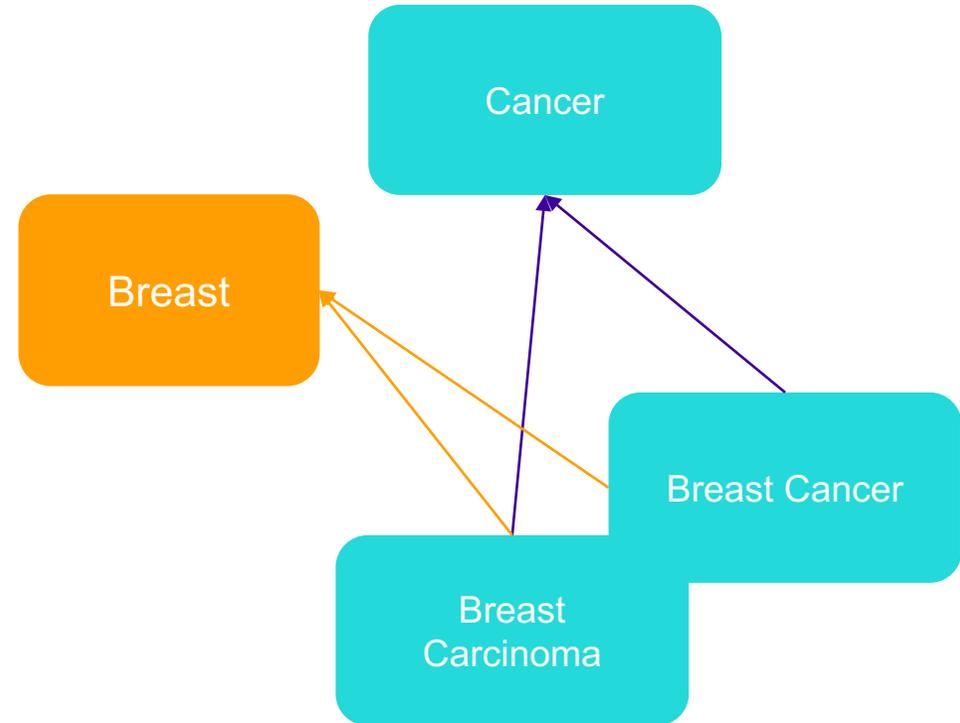


General information (collapse)

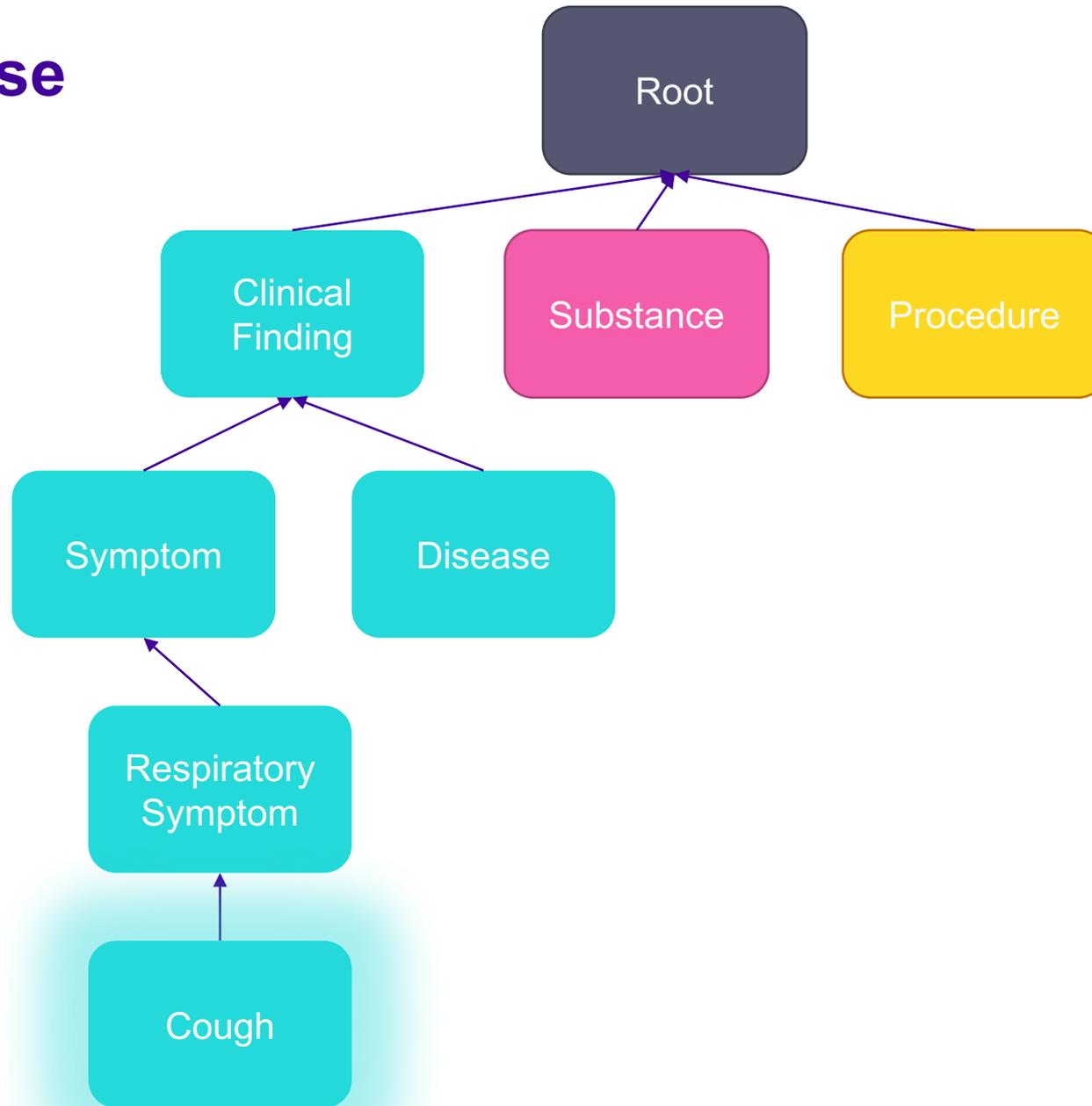
Concept IRI	http://kb.babylonhealth.com/snomedct-20170731T150000Z/49727002
Readable label	Cough
Semantic Types	Clinical finding
skos:prefLabel	Cough hoesten tos hosta toux
skos:altLabel	Observation of cough Cough (finding) tos (hallazgo) toux (constatation)

Maintaining a medical knowledge base

- **Knowledge engineering:** How do we improve the KB, machine reading and manual data entry
- **Stability/Versioning issues:** changes have costly implications
- **Redundancy and completeness:** multiple Identifiers for the same concept / absence of links or entities



Knowledge base Hierarchy



More questions...

Do you smoke cigarettes?

Yes, I currently smoke



I am an ex-smoker (previously smoked regularly but have now stopped)



I have never smoked



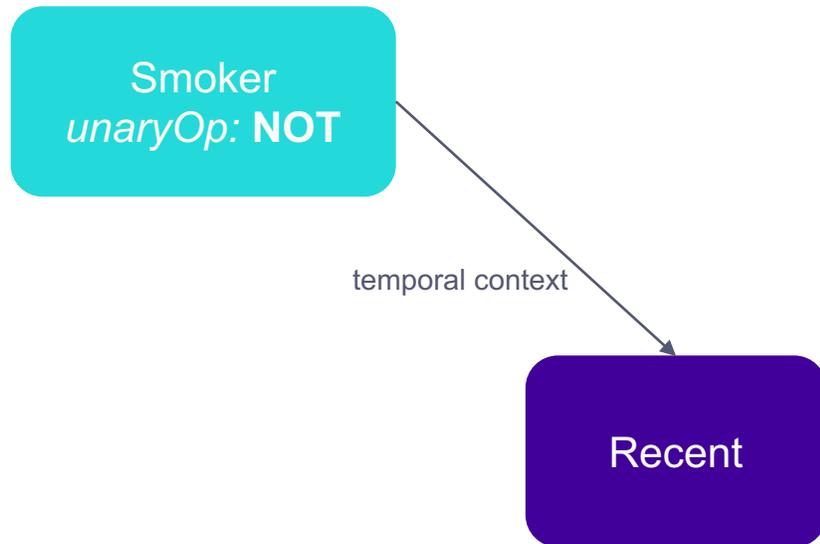
I don't know



Complex concepts

Ex smoker

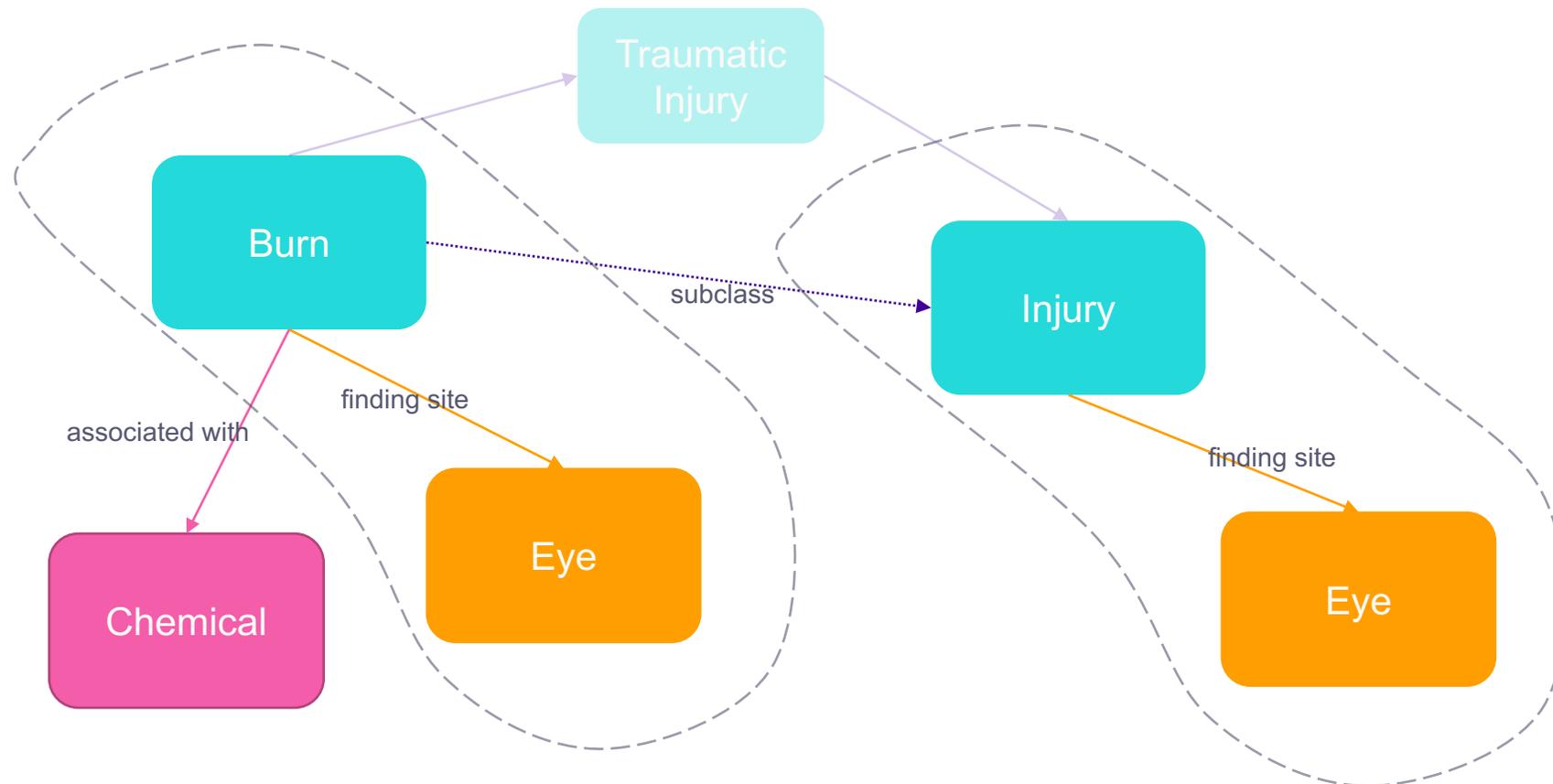
NOT (base_concept(smoker) AND temporal_context(recent))



```
{  
  "unaryOp": "NOT",  
  "baseConcept": {  
    "label": "Smoker",  
    "iri": "http://kb.../umls2017aa_snomedct-2017_03_01/22253000"  
  },  
  "modifiers": [  
    {  
      "type": {  
        "label": "Temporal context", ...  
      },  
      "value": {  
        "baseConcept": { "label": "Recent", ... }  
      }  
    }  
  ]  
}
```

Knowledge algorithms: reasoning about complex concepts

Is “Chemical Burn of Eye” a “Eye Injury?”



...eventually



Talk to a GP preferably within 1 week

If symptoms persist, worsen, or you are concerned, seek further advice.

[Book an appointment](#)

Why do I suggest this?

There is a chance of: Flu, Acoustic neuroma. Remember it's only a possibility. A medical professional can help you decide if it should be taken seriously.

Possible causes

Allergic rhinitis

Likely ●●●○

Inflammation of the inside of the nose due to an allergy, for example hay fever caused by pollen. >

Viral upper respiratory tract infection

Moderately likely ●●○○

Viral infection of the throat, nose and sinuses. >

Flu

Moderately likely ●●○○

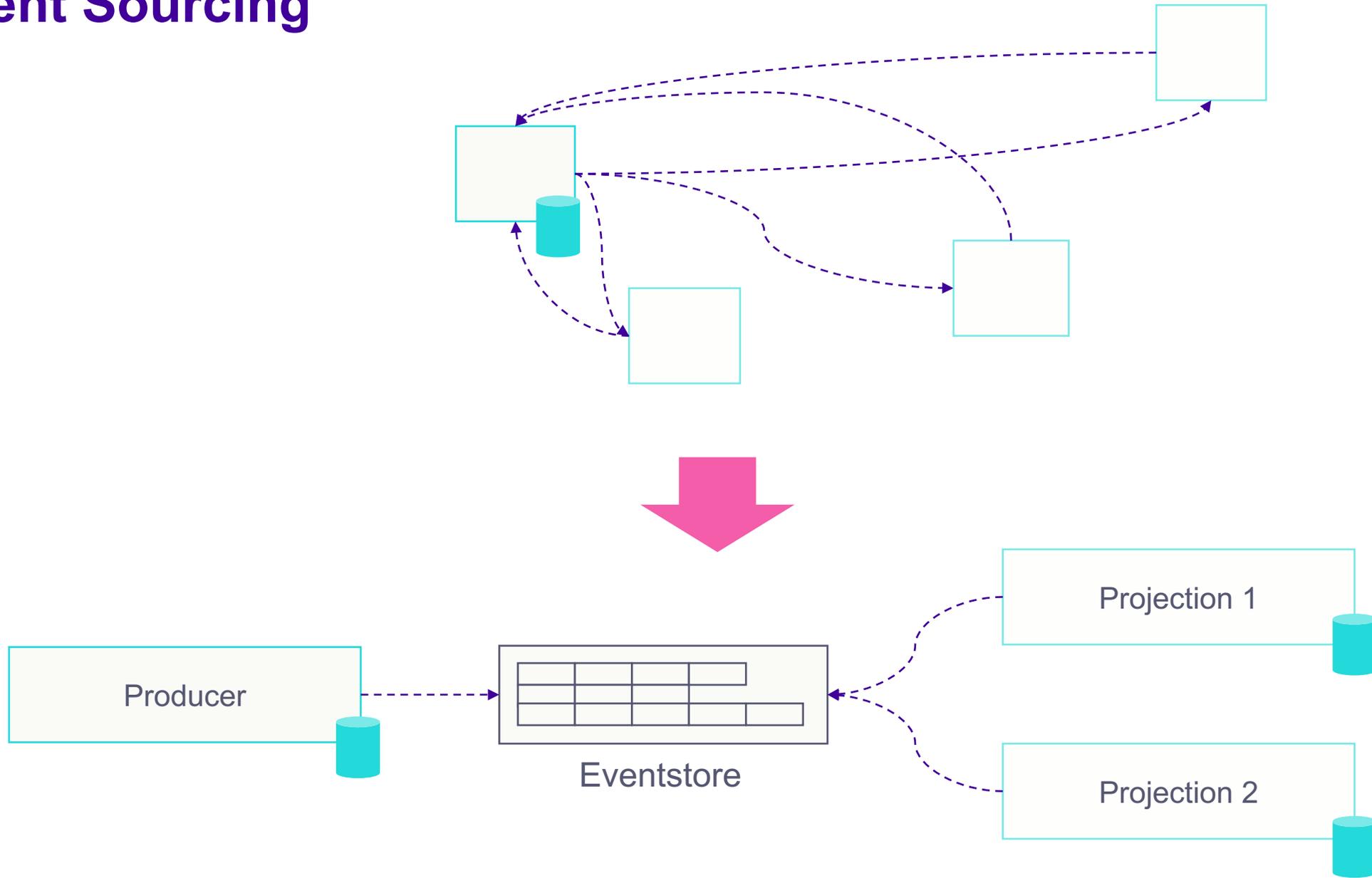
Infection with Influenza virus. >

Acoustic neuroma

Less likely ●○○○

Non-cancerous growth of the nerve that connects the inner ear to the brain. >

Event Sourcing

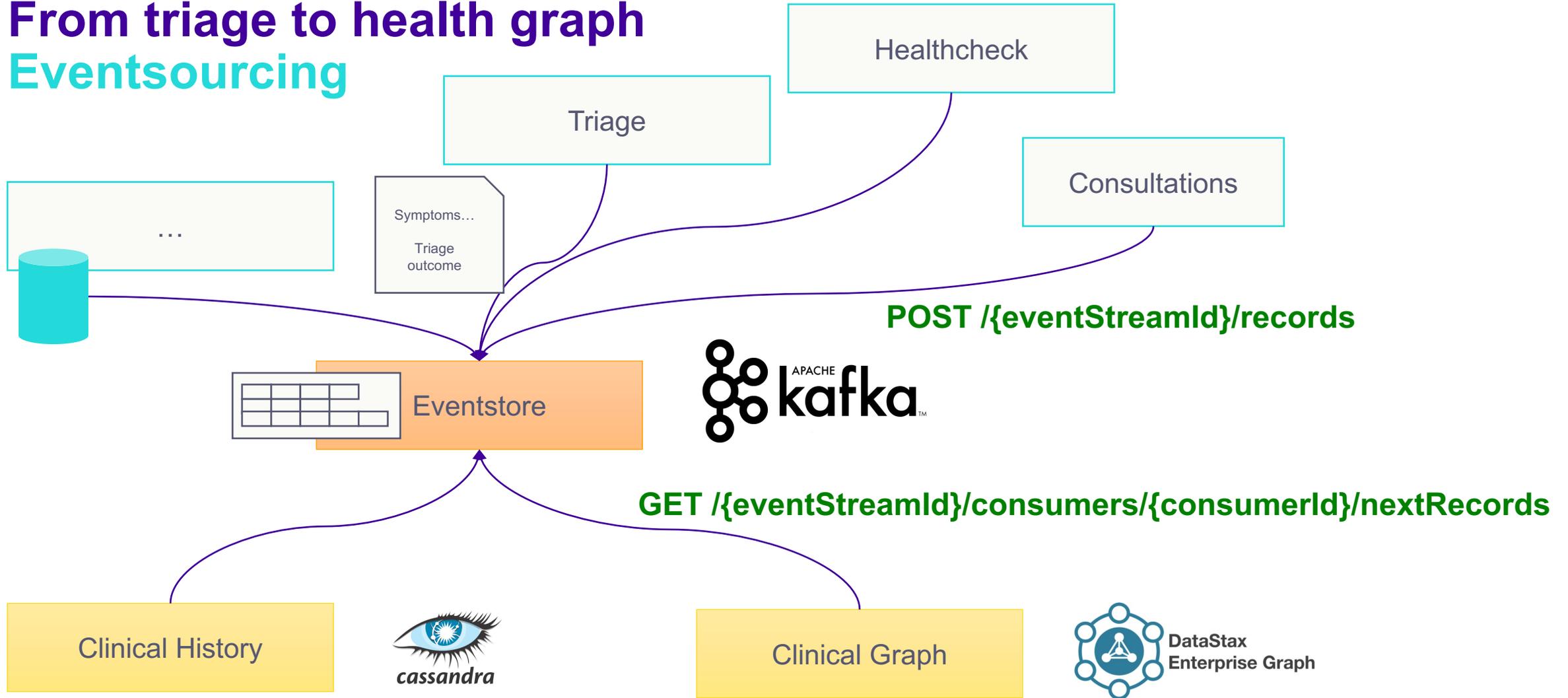


Concepts inside events

```
{
  context: {...},
  payload: {
    {
      "baseConcept":{
        "label":"Pain",
        "iri":"http://kb.../umls2017aa_snomedct-2017_03_01/22253000"
      },
      "modifiers":[
        {
          "type":{
            "label":"Finding site",...
          },
          "value":{
            "baseConcept":{ "label":"Lower leg structure", ... },
            "modifiers":[
              {
                "type":{ "label":"Laterality", ... },
                "value":{
                  "baseConcept":{ "label":"Left", ... }
                }
              }
            ]
          }
        }
      ]
    }
  ]
}
```

From triage to health graph

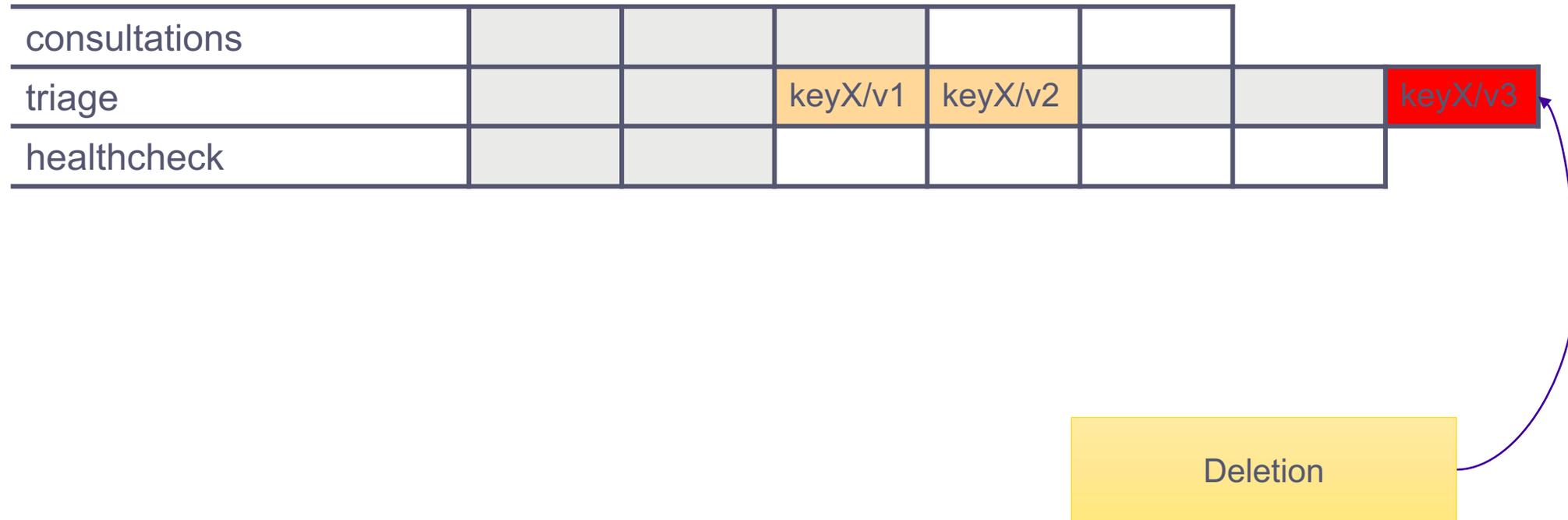
Eventsourcing



Projections



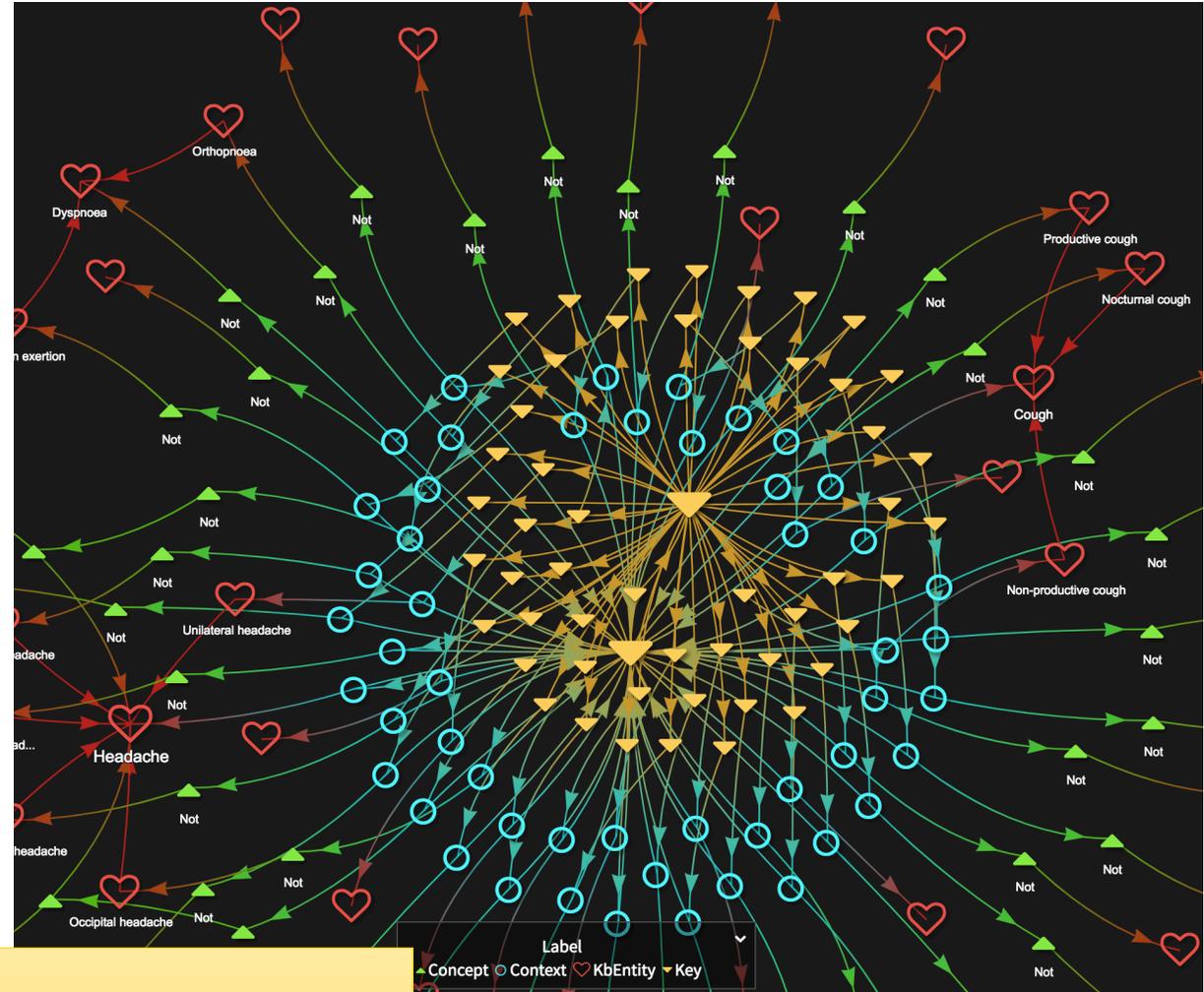
Projections: deletion through compaction



Same input, different output Projections

Date	Concept
2018-06-07T11:46:11.898411Z	NOT Unintentional weight gain
2018-06-07T11:46:11.898411Z	Cough
2018-06-07T11:46:11.898411Z	NOT Productive cough
2018-06-07T11:46:11.898411Z	NOT Nocturnal cough
2018-06-07T11:46:11.898411Z	NOT Dyspnea
2018-06-07T11:46:11.898411Z	NOT Dyspnea on exertion
2018-06-07T11:46:11.898411Z	NOT Orthopnea
2018-06-07T11:46:11.898411Z	NOT Paroxysmal nocturnal dyspnea
2018-06-07T11:46:11.898411Z	NOT Abnormal breath sounds
2018-06-07T11:46:11.898411Z	Ex-smoker
2018-06-07T11:46:11.898411Z	NOT Smoker
2018-06-07T11:46:11.898411Z	Family history of asthma
2018-06-07T11:46:11.898411Z	NOT Family history of malignant neoplasm of lung
2018-06-07T11:46:11.898411Z	NOT H/O: asbestos exposure
2018-06-07T11:46:11.898411Z	Sore throat symptom
	...

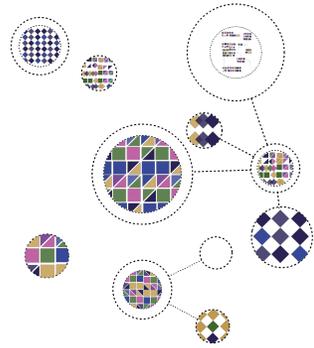
Clinical History



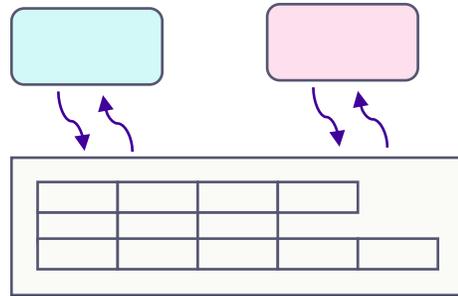
Clinical Graph



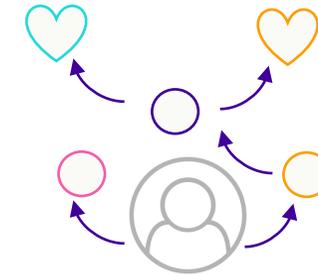
Connecting data with...



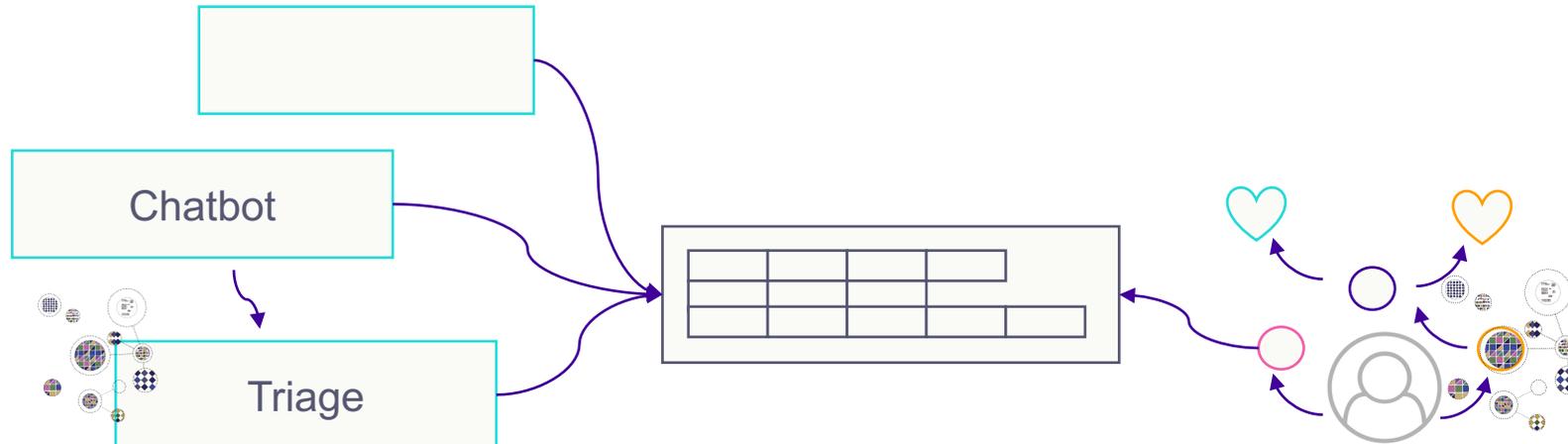
Babylon Knowledge Base
Medical source of truth and shared language



Eventstore
Event sourcing implementation and stream of events



Clinical Graph
Graph representation of clinical facts



Semantics use case. Talking to a patient. Show me all recent concepts reported

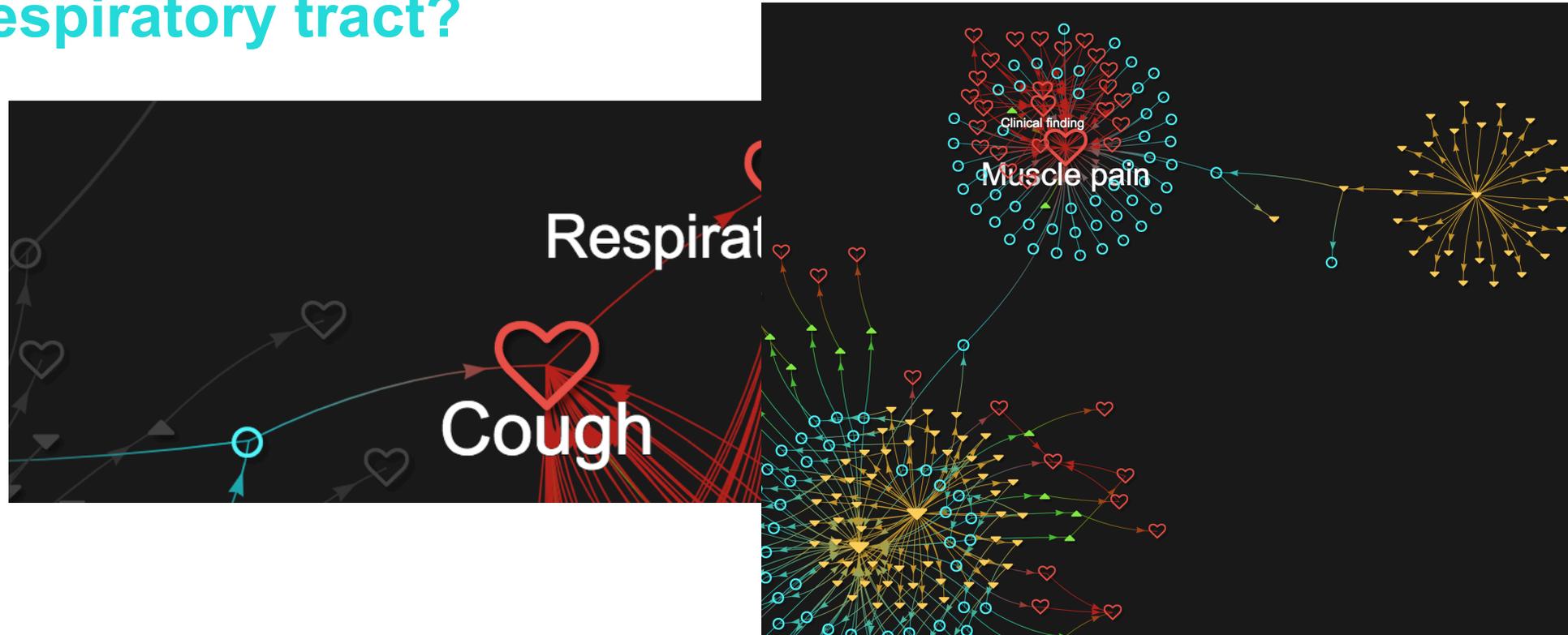
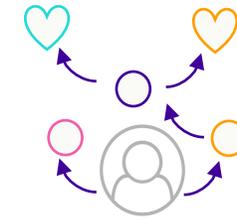
Date	Concept
2018-06-07T11:46:11.898411Z	NOT Unintentional weight gain
2018-06-07T11:46:11.898411Z	Cough
2018-06-07T11:46:11.898411Z	NOT Productive cough
2018-06-07T11:46:11.898411Z	NOT Nocturnal cough
2018-06-07T11:46:11.898411Z	NOT Dyspnea
2018-06-07T11:46:11.898411Z	NOT Dyspnea on exertion

```
SELECT * FROM clinicalhistory WHERE  
Date > ... AND Date < ...
```

Filter by “respiratory tract”?

Knowledge Graph as glue

Has the user recently experienced symptoms in the respiratory tract?



Respiratory tract ← [finding site] Cough ← Context {timestamp} ← Patient

Challenges and future directions

- **Knowledge Representation and Reasoning:**
 - Maintaining and improving a medical Knowledge Base
 - Rule based inference, integrity constraints, uncertain information
- **Scaling:** Scaling up, relying on DSE graph for high volumes of data
- **Integration:** Making it easier for clients to create events (Change Data Capture)
- **Machine Learning**
 - Graph embeddings for concept and users
 - Risk prediction



Thank you
For listening

domenico.corapi@babylonhealth.com