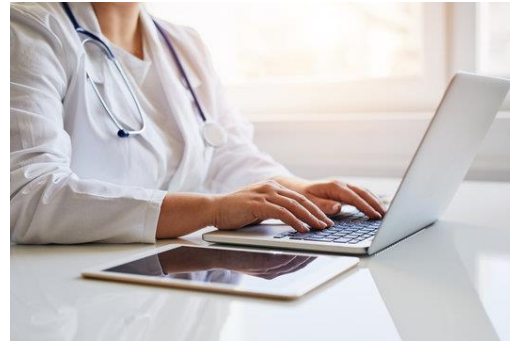


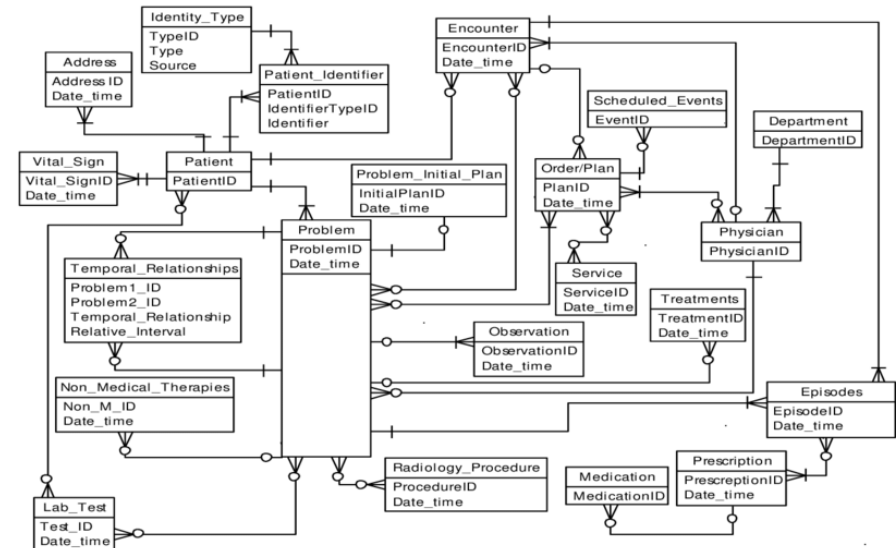
Automating structuring clinical texts

Stig Hellemans, MD
PhD-Student

Evolution healthcare information systems



61-years old female with Crohn's disease. Attended cycling event in Salo, flu prirorlry. Arfter cycling, experienced breathing difficulties and went to the emergency department and elevated herart enzymes and incompensation were found. Was admitted to the ICU for care of incompensation and pneumonia. In UKG 2.6. ef 30%. In coronary angiography, significant stenoses in RCA, LCX and mmarin. Trordray, elective quadruple bypass LITA-LAD, Ao-LOM-LPL and Ao-RBD, in which goord flow. Pre.op. the posterior wall of the left ventricle and the septum contract lamely, ef about 35%, mitral valve 1-2/4 leak. Aortic cross-clamp time 1 h 32 min. Post.op. ef over 40%. On basis of the UKG-finding pre.op. Simdax-infusion was initiated. On arrival to ICU, haemodynamics was stable, norepinephrine administered. Cardiac index 3.2. Warming-up and weaning in ventilator.



Digitalization

Structuring

➔ More administrative work & time lost!
➔ WHY DO WE DO THIS?

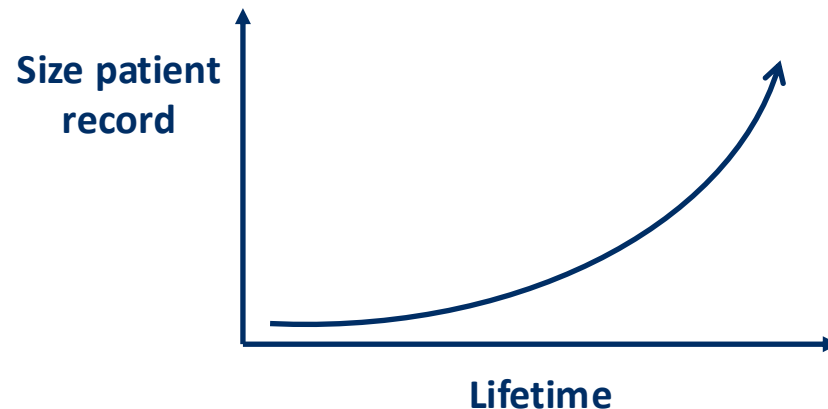
Why are we on this path?

Making clinical data more FAIR (Findable, Accessible, Interoperable, Reusable)



Better & more efficient healthcare:

- Data is readily available without having to read the whole patient record
- Patient groups can be monitored & managed collectively
- Close collaboration across care institutions is greatly improved



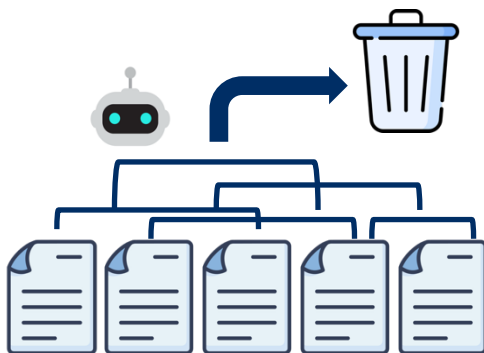
Solutions to make unstructured texts more FAIR

1.

Keep the source texts unstructured

→ Use search + AI agents to get the right data out each time

‘Structure on the fly’



2.

Build structure on top of the source texts

→ Needs validation + data standard

‘Structure once to reuse’



3.

Everything in forms
WITHOUT freeform texts

→ Problem would be solved
BUT data input problem

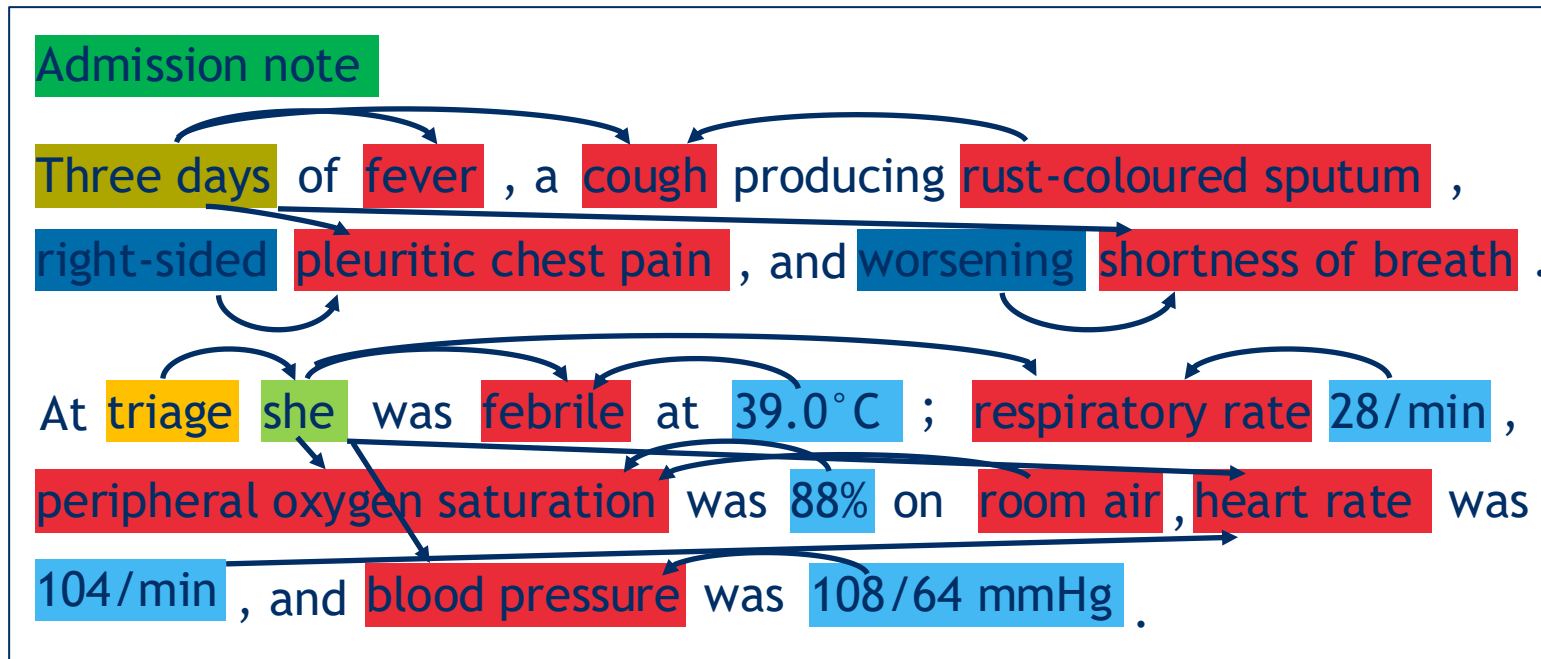
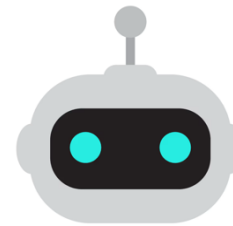
‘Structure is the goal’

 **Too restrictive...**

Project 1: My moonshot goal

Structured representation is only useful if it is:

1. faithful enough to the text,
2. normalized enough to query
3. explicit enough to reason over



Chest radiography



Microbiology



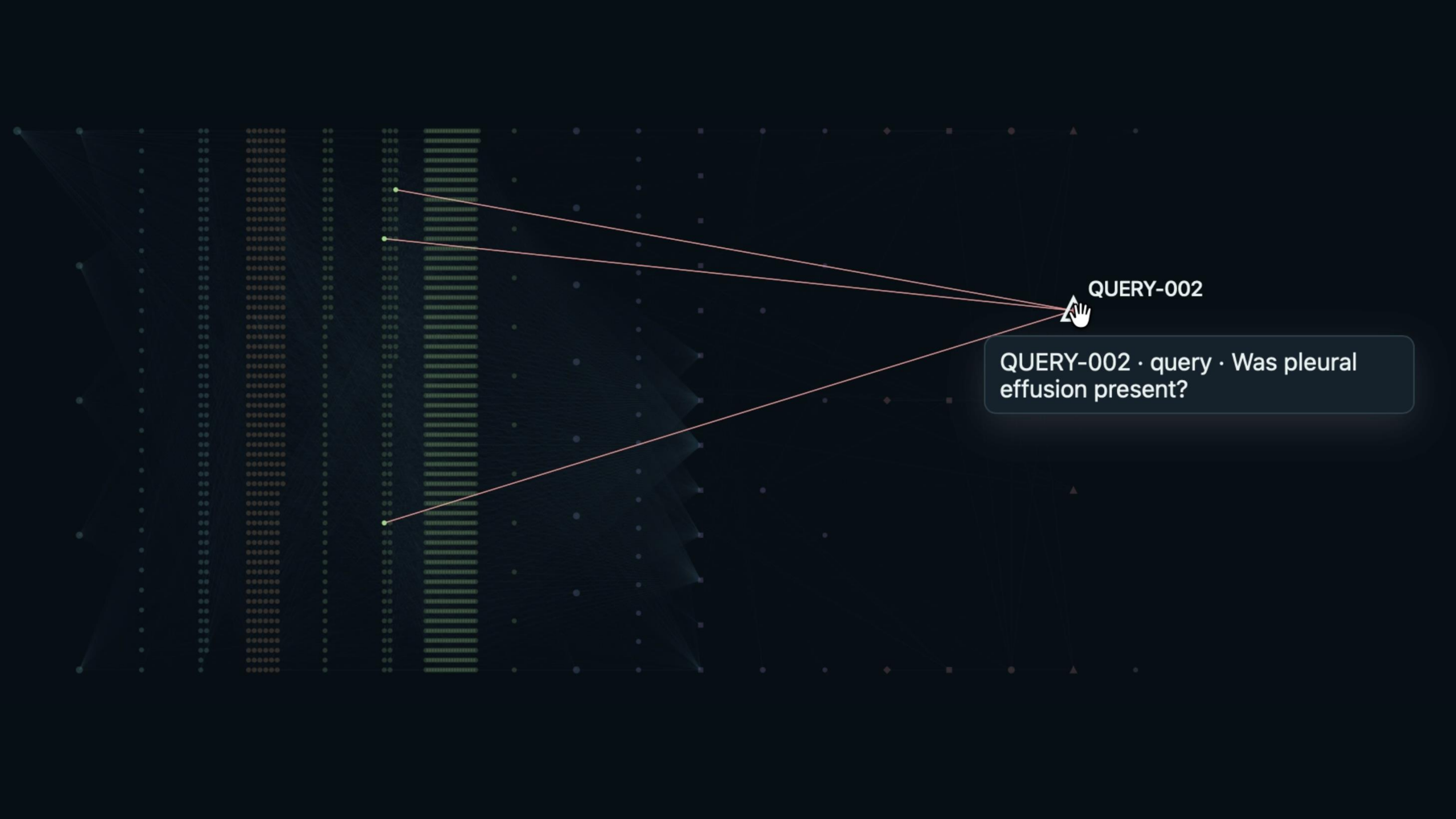
Ward progress



Discharge

SNOMEDCT

LOINC®



QUERY-002

QUERY-002 · query · Was pleural effusion present?

 Pleural effusion

 Pleural effusion

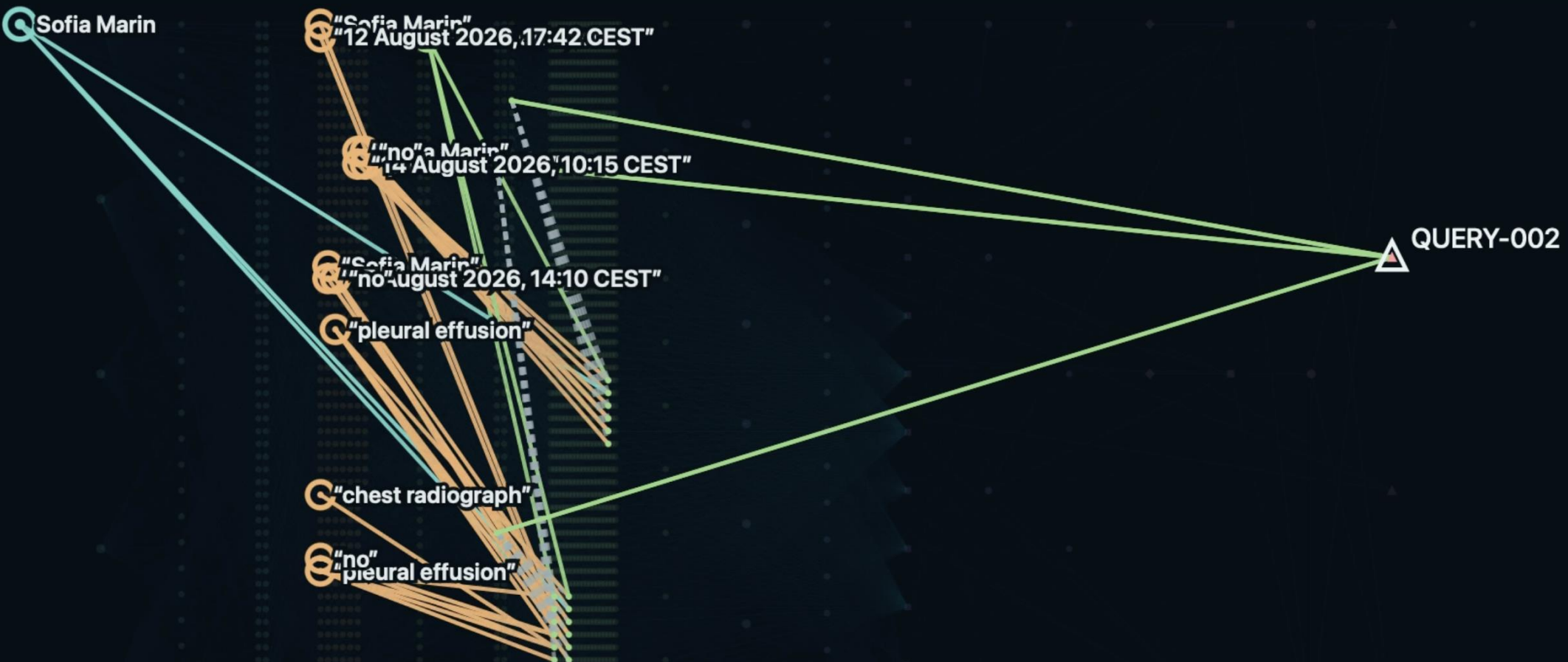
 Pleural effusion

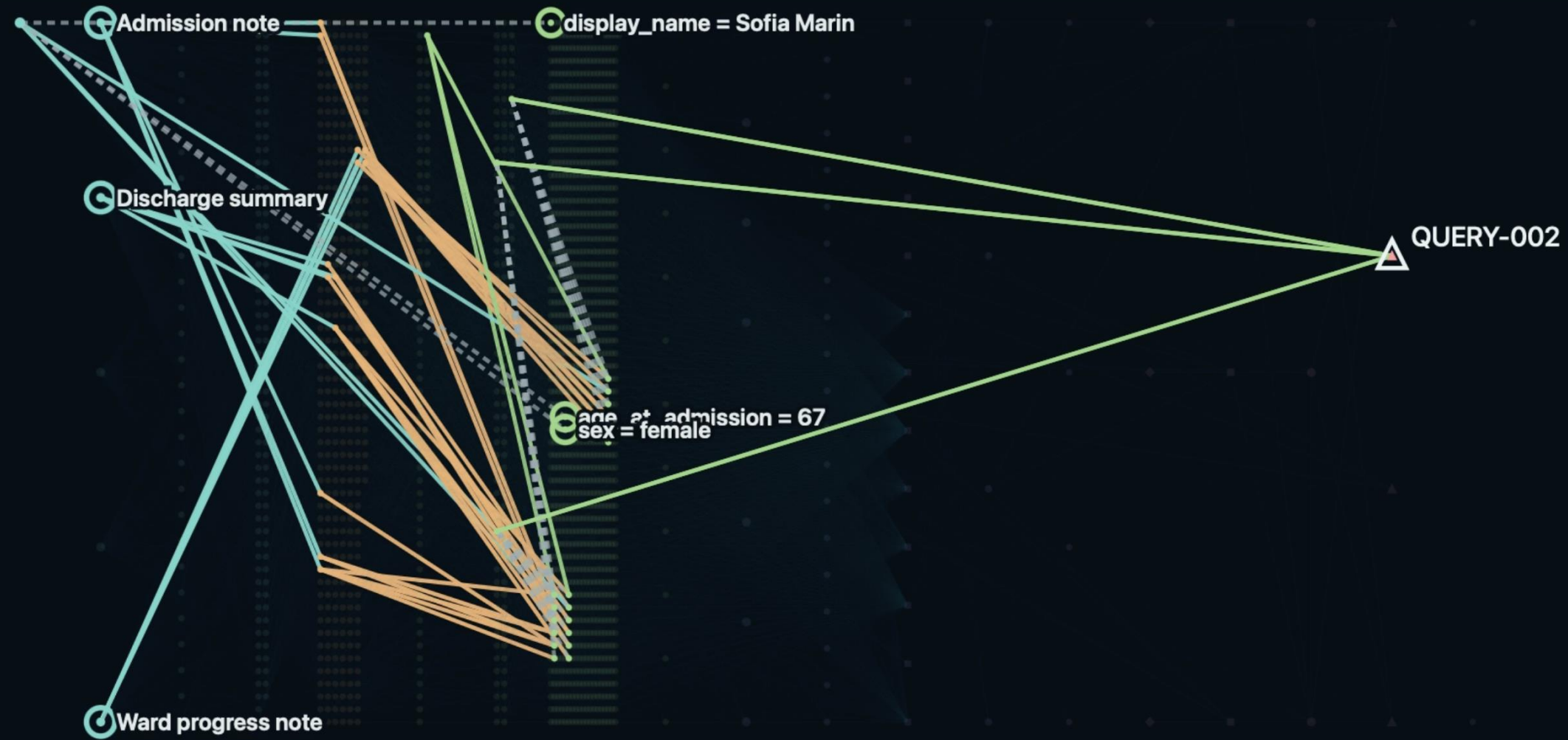
 QUERY-002

QUERY-002

normalized_focus = CON:PI:EURAL_EFFU...
experience = PAT:SYN:001
clinical_time_anchor = 2026-08-14T1...
polarity = negative
certainty = certain
clinical_status = current

normalized_focus = CON:PI:EURAL_EFFU...
experience = PAT:SYN:001
clinical_time_anchor = 2026-08-15T1...
polarity = negative
certainty = certain
clinical_status = current





Project 2: Normalizing posology texts (Dirk Broeckx)

Virtual Medicinal Product group
Paracetamol, Oral, 1g

Max 4x/day, 4-6h
in between

...



Maximum 4x per day

...



Semantic
normalization



```
[
  {
    "text": "Maximum 4x per day",
    "timing": {
      "repeat": {
        "frequency": 4,
        "period": 1,
        "periodUnit": "d"
      }
    },
    "doseAndRate": [
      {
        "doseQuantity": {
          "value": 1,
          "unit": "Tablet (unit of presentation)",
          "system": "http://snomed.info/sct",
          "code": "732936001"
        }
      }
    ],
    "additionalInstruction": [
      {
        "text": "Leave 4-6h in between"
      }
    ]
  }
]
```

Project 2: Pilot project



1 Pharmacy

40,412 unique texts



Per VMP group

33,667 unique texts included



{ 98.7% }
{ FHIR }

Automated structuring

10+ hours AI agents (~300€)

Take aways:

- 1) **Expert validation still needed**
- 2) **Long tail of rare patterns:**
 - only 9.5% texts share similar semantics with others
 - 80.1% of text usage is concentrated in 12.3% VMP groups (58% of patterns)
- 3) **Secondary focus on clarity of description for patients & caregivers**

Thank you for listening!

Please reach out:
stig.hellemans@uantwerpen.be