WP4
LLM for medical triage

Dr Ivan Lerner

Assistant Hospitalier Universitaire
Département d'informatique médicale, HEGP
Université Paris Cité

Team

- LORIA
    - Emmanuel Vincent, Senior Research Scientist
    - Gaël Guibon, Associate Professor
- AP-HP
    - Medical Informatics
        - Dr Ivan Lerner
    - Emergency Department
        - Dr Gustave Toury
    - SAMU 93
        - Pr Frédéric Lapostolle
- LINAGORA

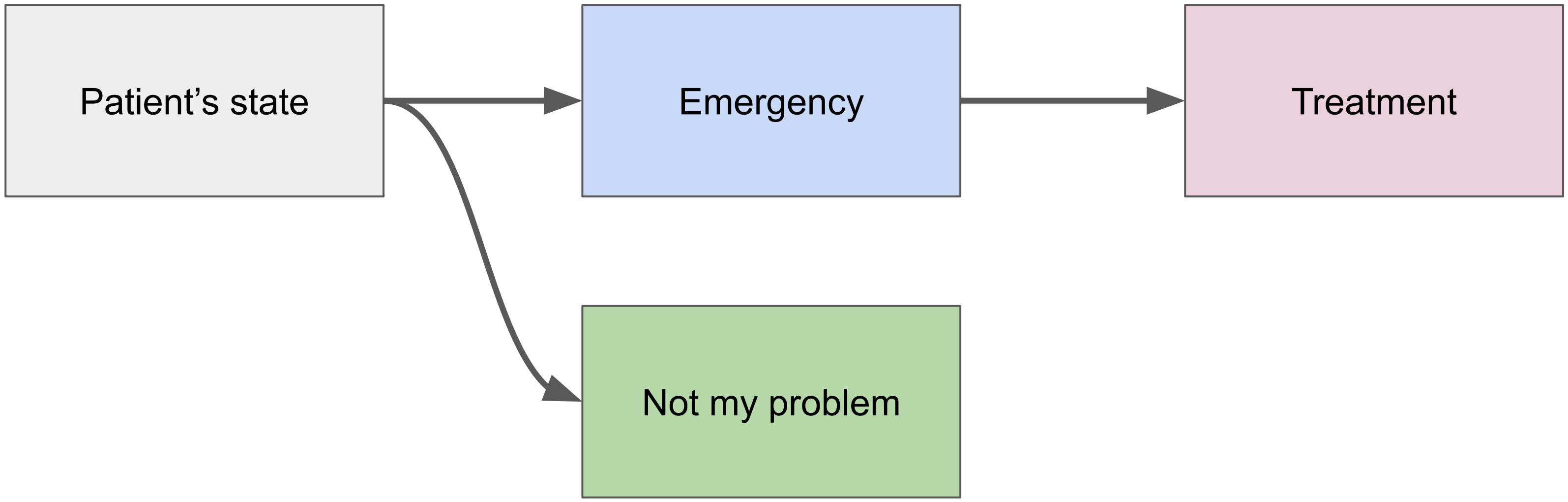
Plan

- About emergency medicine
- About call triage
- The long term objective
- The problem we can solve
- How we plan to do it

About emergency medicine

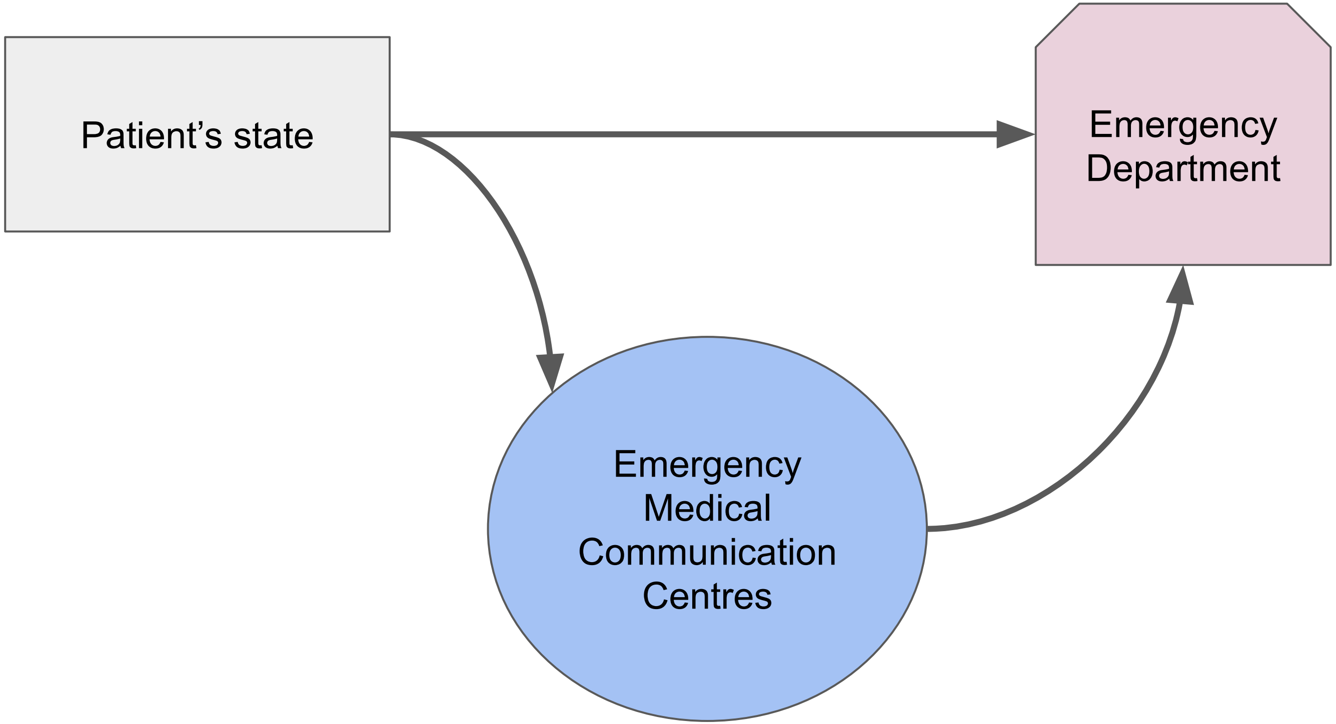
Classical medicine

Emergency medicine

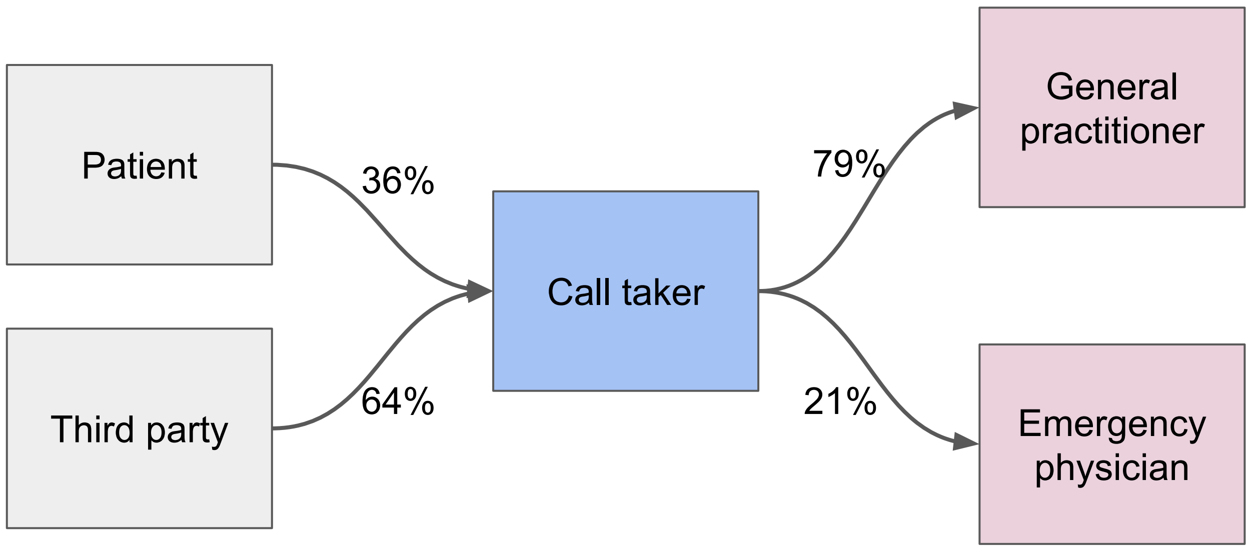


About call triage

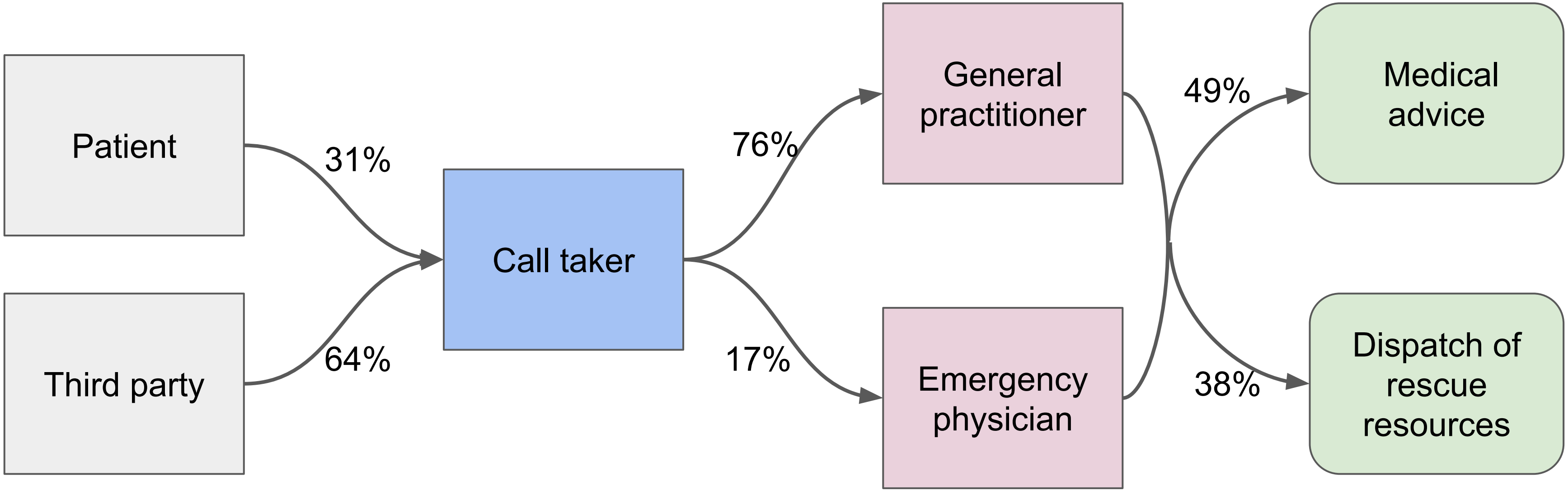
Patient pathways



Emergency Medical Communication Centres (France)



Emergency Medical Communication Centres (France)



Facteurs impactant la durée de communication dans un EMCC.M Bensoussan, F Lapostolle, Paul-Georges Reuter, under review

SAMU 92

1.6 millions inhabitants

9 months study

108 548 calls

58% women

aged 38 [18 – 66] years old

64% call made by a third party

84% call from home

Facteurs impactant la durée de communication dans un EMCC.M Bensoussan, F Lapostolle, Paul-Georges Reuter, under review

SAMU 92

Reason for calls	% Calls
Cardiology	37
Other	29
Medical other	19
Trauma	8
Respiratory	4
Neurology	2
Psychiatry	1
Intoxication	<1
Death	<1

Facteurs impactant la durée de communication dans un EMCC.M Bensoussan, F Lapostolle, Paul-Georges Reuter, under review

SAMU 92

79% calls oriented toward general practitioner

49% calls led to medical advice

38% calls led to dispatch of rescue resources

Doctor decision made in 3 [2-4] minutes

Facteurs impactant la durée de communication dans un EMCC.M Bensoussan, F Lapostolle, Paul-Georges Reuter, under review

The accuracy of medical dispatch

Study	Country	Decision system	Outcome	Paramedics reference	Recall (%)	Precision (%)
Ball, 2016	Australie	MDPS	Life-threatening emergency	Triage decision	93.3	5.85
Dami, 2015	Suisse	CBD	Life-threatening emergency	NACA > 3	86	21.7
Moser, 2017	Suisse	CBD	Life-threatening emergency	NACA > 3	86.8	29.2
Leopardi, 2013	Italie	Autre	Vital emergency requiring a doctor's intervention	Local score	78	36.6
Ek, 2013	Suède	CBD	Urgence vitale et non vitale	METTS-A red/orange/yellow	94.5	88.5

Medical Priority Dispatch System (MDPS)
Criteria Based Dispatch System (CBDS)

Bohm K, Kurland L. The accuracy of medical dispatch - a systematic review. Scand J Trauma Resusc Emerg Med. 9 nov 2018;26(1):94.

Dialogue example

M : bonjour madame, je suis le docteur DAMANI le médecin régulateur heu votre fille nous a appelé parce que vous avez une douleur au niveau de la poitrine

P : c'est ça

M : qui est apparue depuis il y a quinze vingts minutes, heu est ce que vous avez des problèmes de santé?

P : heu vous savez un peu de cholestérol, un peu d'hypertension voilà, fin des trucs de base quoi hein

M : est ce que vous avez déjà fait un infarctus?

P : je pense pas. infarctus, c'est quoi ça?

M : les artères bouchées au niveau du cœur

P : heu non non non moi non. mais mon fils heu oui oui

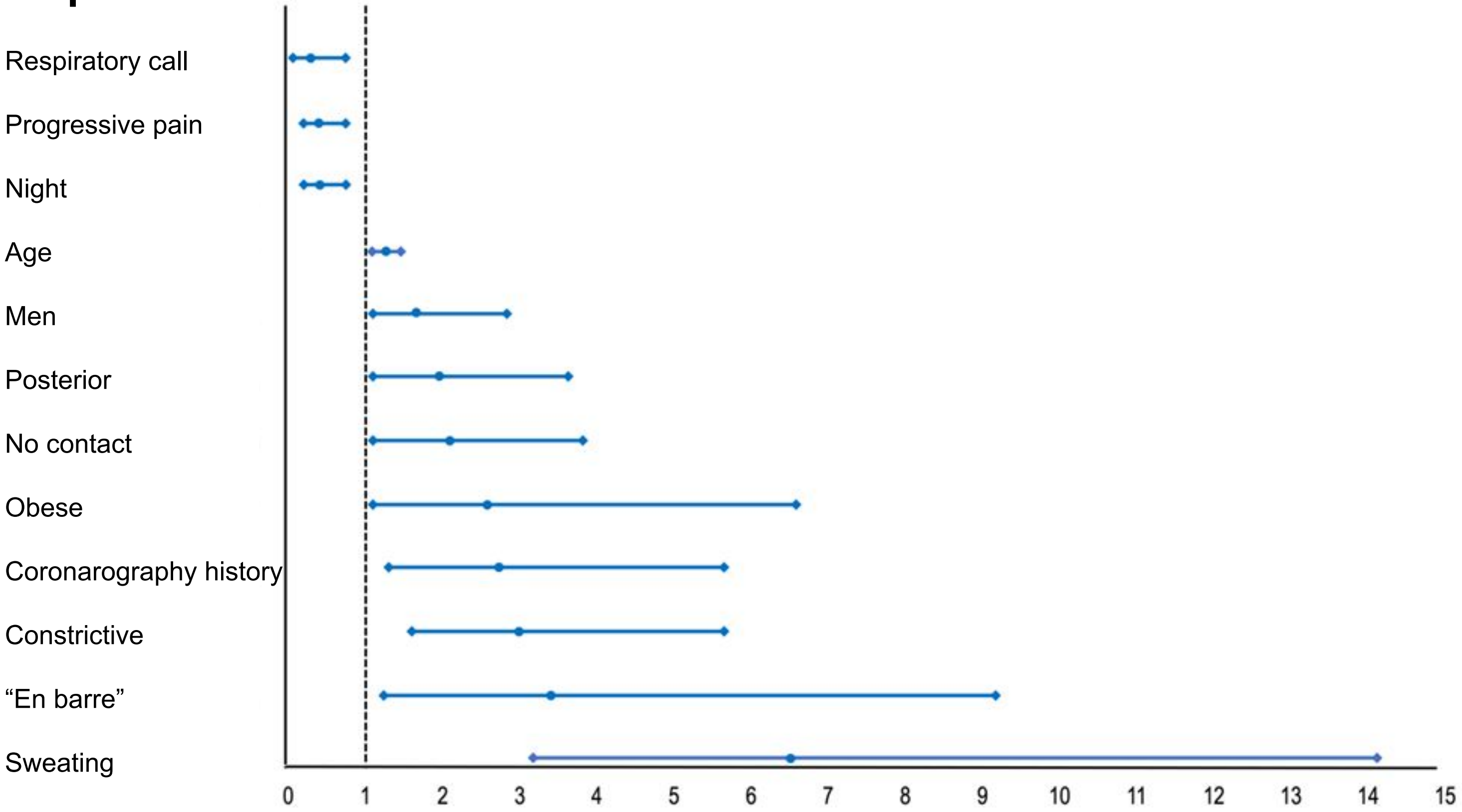
M : votr- heu votre fils. mais vous non?

P : heu et donc là vous avez mal. donc ça fait quinze minutes c'est toujours pas parti la douleur n'est n'est pas passée

M : heu est ce que ça serre au niveau de la poitrine?

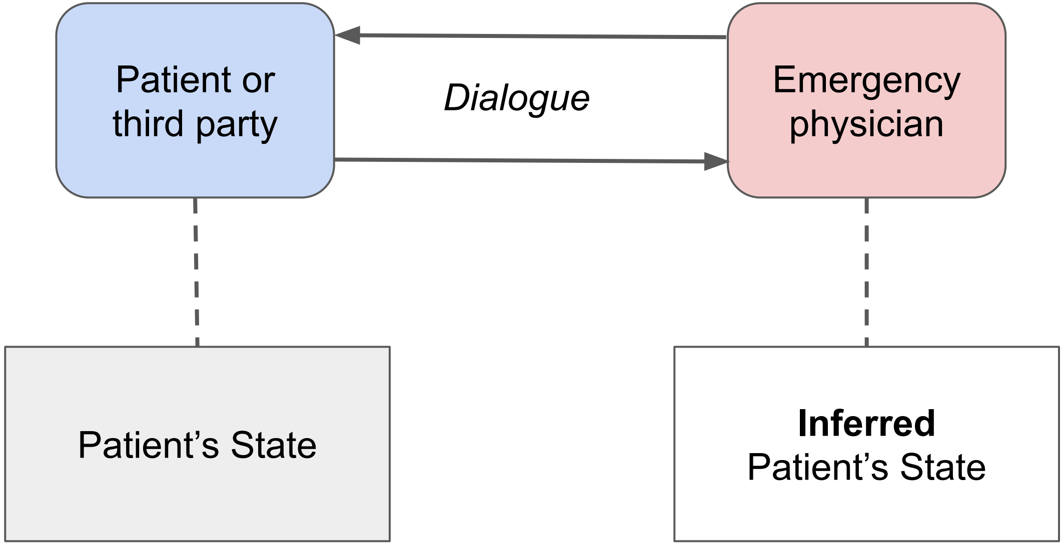
P : heu ça serre pas non ça serre pas

Thoracic pain characteristics and triage decision

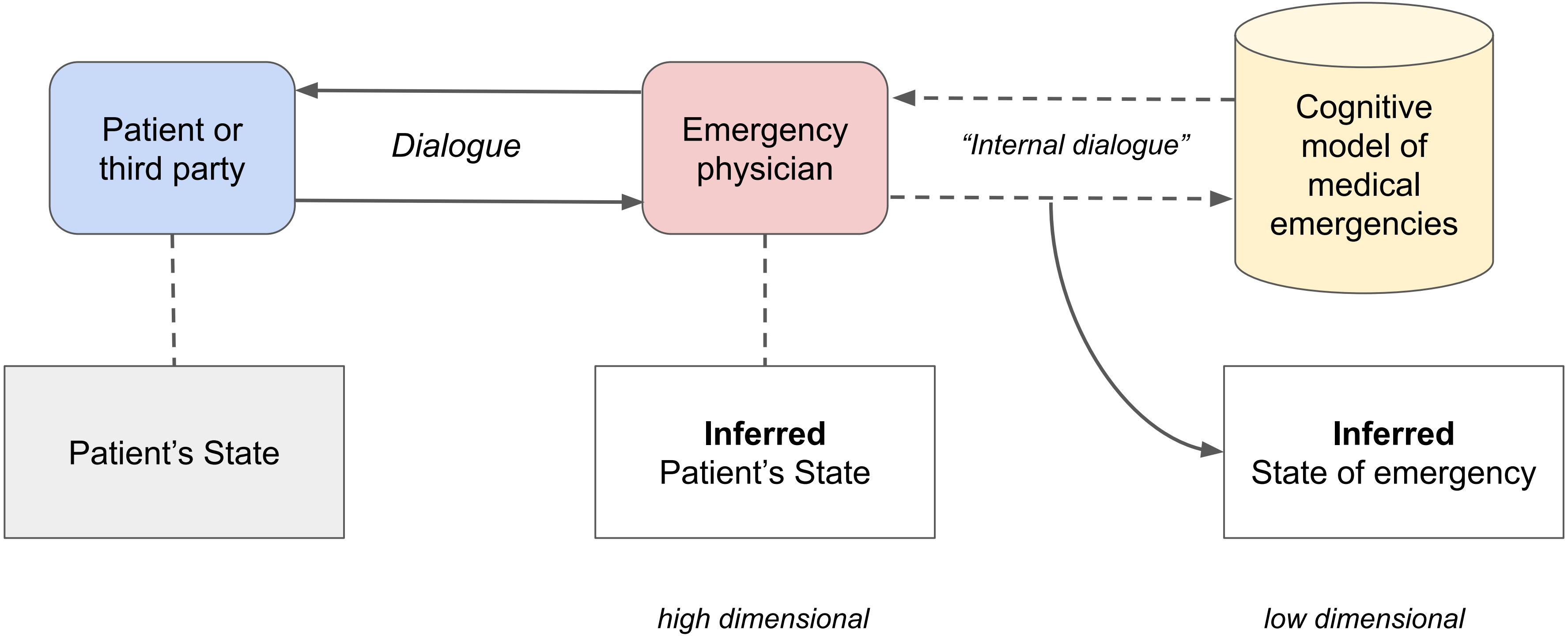


Decision making in EMS Emergency physician facing calls for chest pain. The REDOUT Study. Reuter Paul-Georges, Lapostolle Frédéric

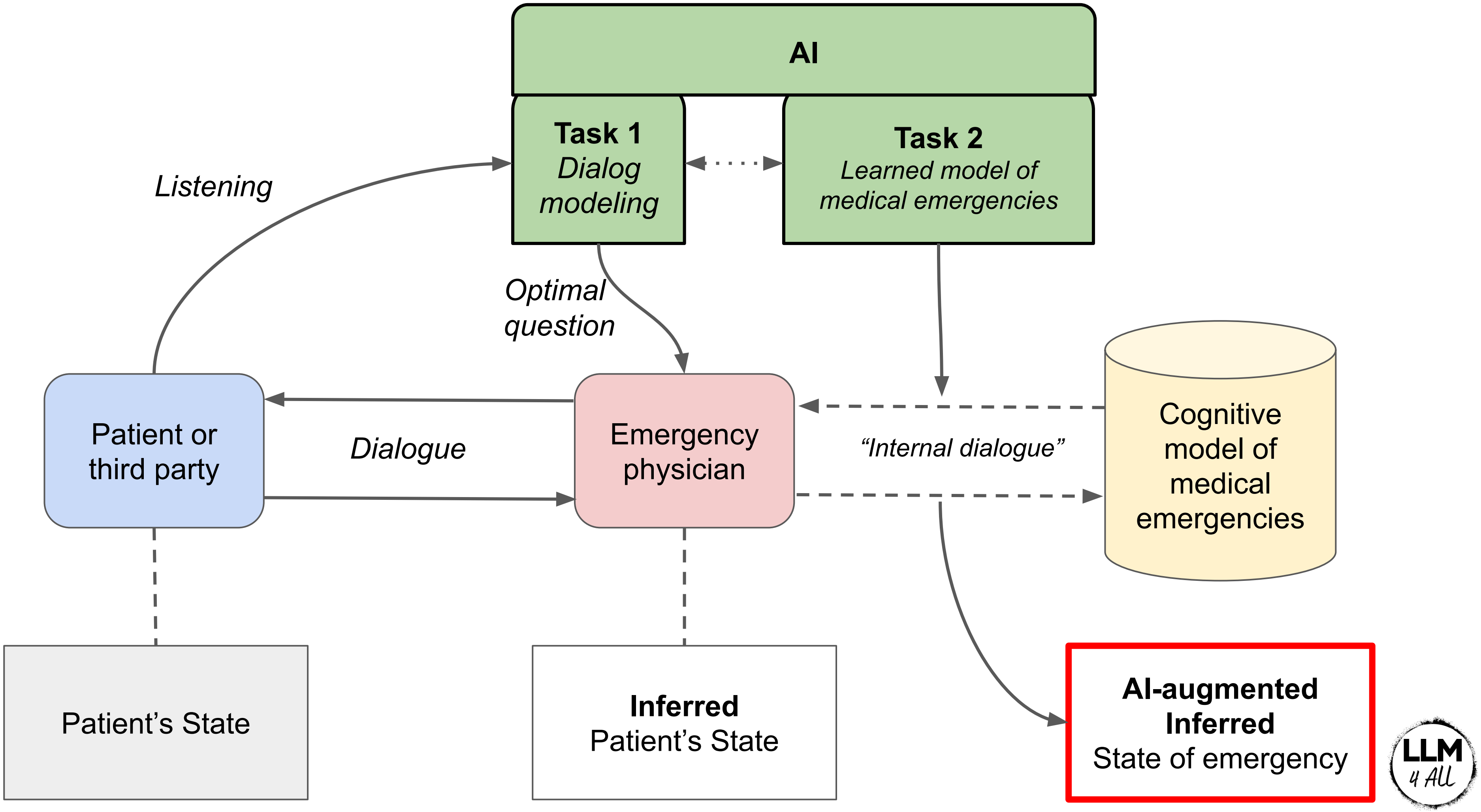
A high-level model of medical triage



A high-level model of medical triage



The long term objective


AI
Task 1
Dialog modeling
Task 2
Learned model of medical emergencies
Listening
Optimal question
Patient or third party
Dialogue
Emergency physician
"Internal dialogue"
Cognitive model of medical emergencies
Patient's State
Inferred
Patient's State
AI-augmented
Inferred
State of emergency
LLM 4 ALL

AI tasks

Task 1

- *What is the set of optimal questions, given the context (previous utterances), that maximises the accuracy of the inferred state of emergency with minimum cardinality*

Task 2

- *What is the level of emergency from observations of a given patient's state*

Ideal dataset and best dataset

Ideal : infinitely large dataset of triage calls

- each dataset item is: a dialogue, a ground truth "level of emergency"
- learn a function **f**, **f**(dialogue) = ground truth "level of emergency"
- multi-class supervised problem

Best : large multicentric dataset of triage calls, linked with SNDS

- proxy for a ground truth "level of emergency"
  - medical triage decision → *Average physician cognitive model*
  - call back
  - death, hospitalisation within 2 weeks (probabilistic link with SNDS) → *Beyond physician cognitive model*
- each dataset item is: a dialogue, a proxy "level of emergency"

The problem we can solve

The datasets we have now

Simulated triage calls

- 3 hours 24 minutes
- 61 calls
- 3077 utterances

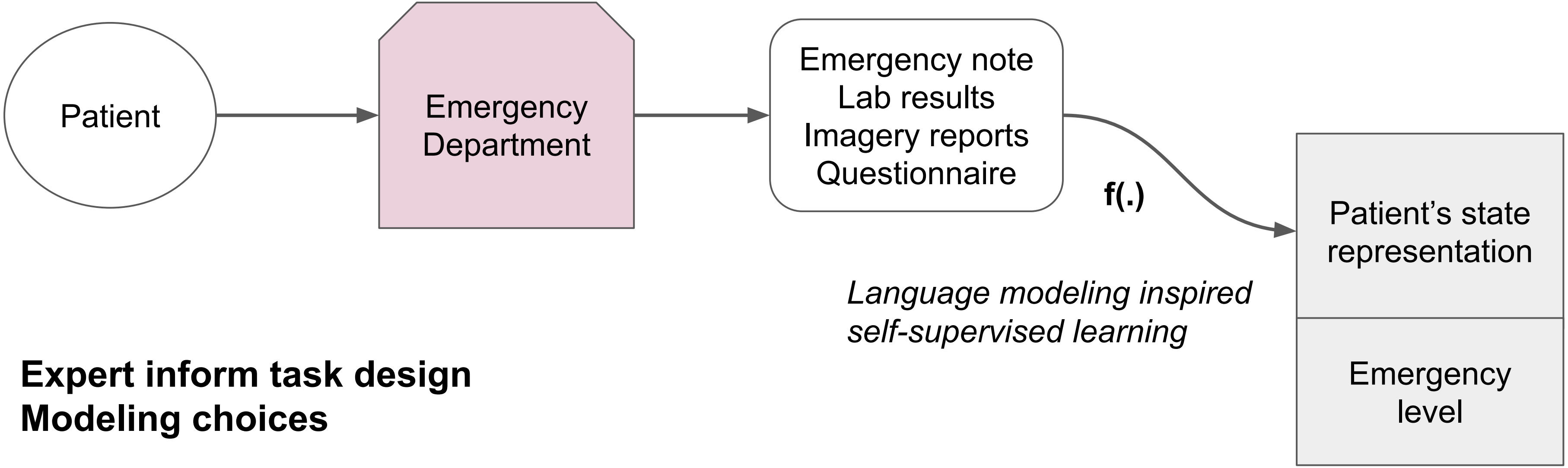
Relevant for NLP

HEGP Emergency Visits

- > 100k visits
- patient’s state
  - emergency note
  - lab results
  - imagery reports
  - questionnaire
- level of emergency
  - physician decision from encounter
  - hospitalisation, death

Relevant for medical questions

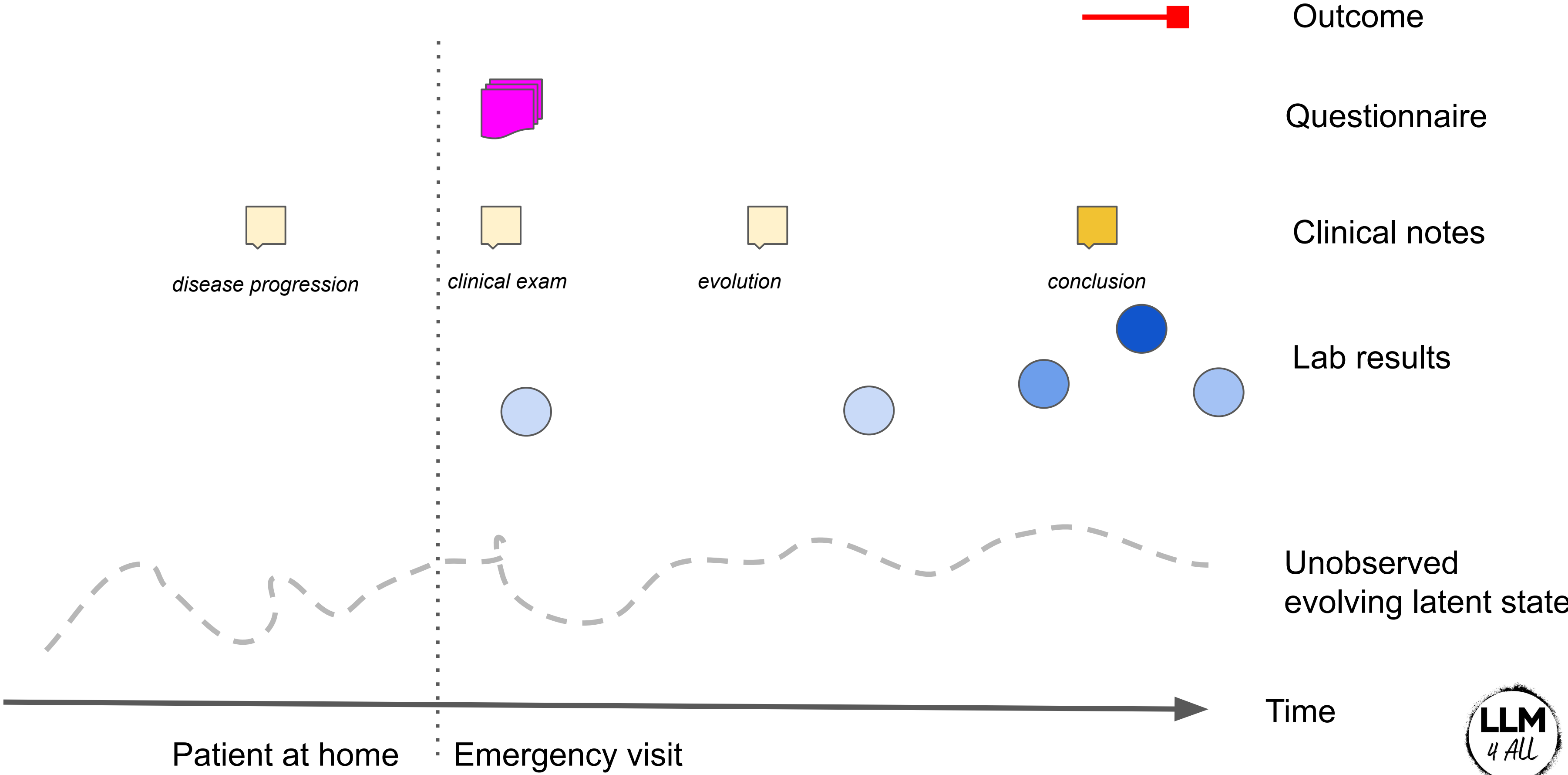
Problem 1: Learning a patient’s state representation



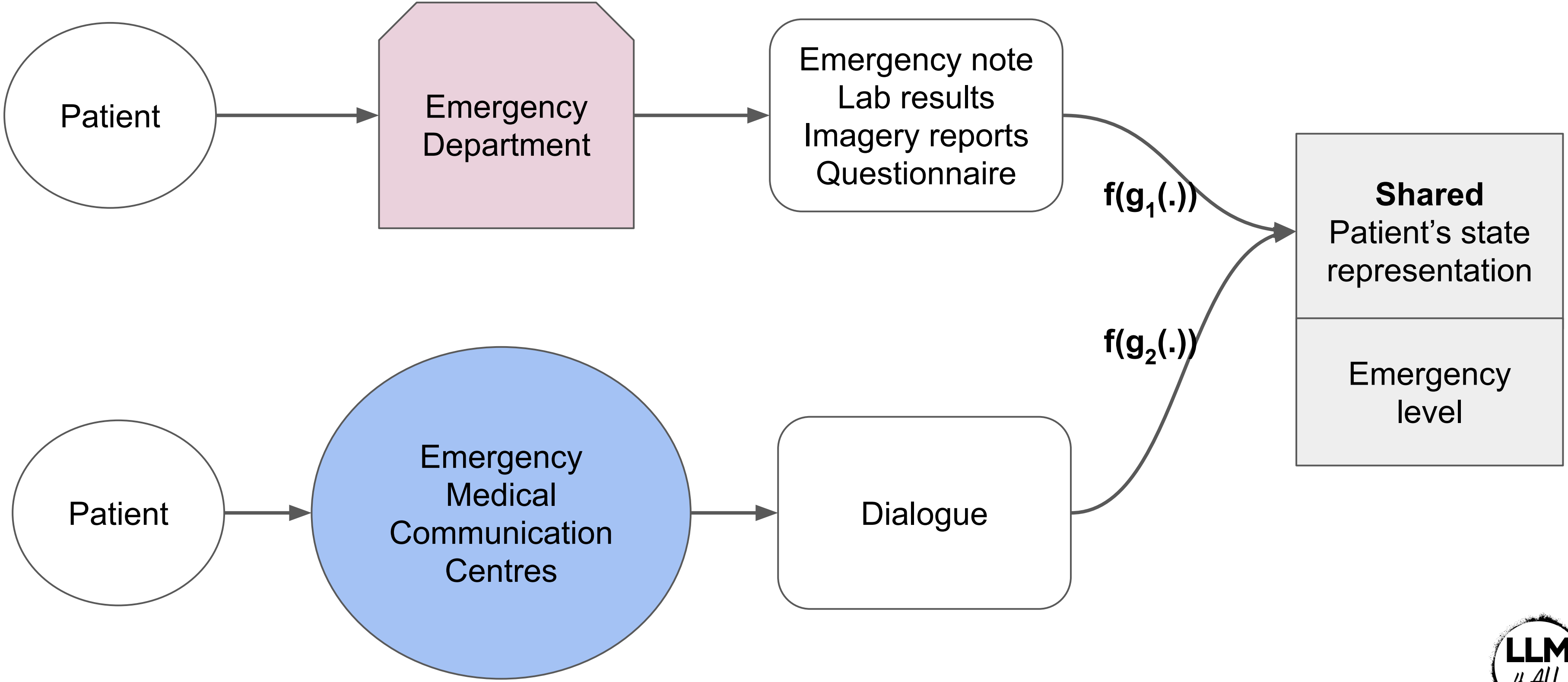
Expert inform task design
Modeling choices

Miotto, Riccardo, et al. "Deep patient: an unsupervised representation to predict the future of patients from the electronic health records." Scientific reports 6.1 (2016): 1-10.

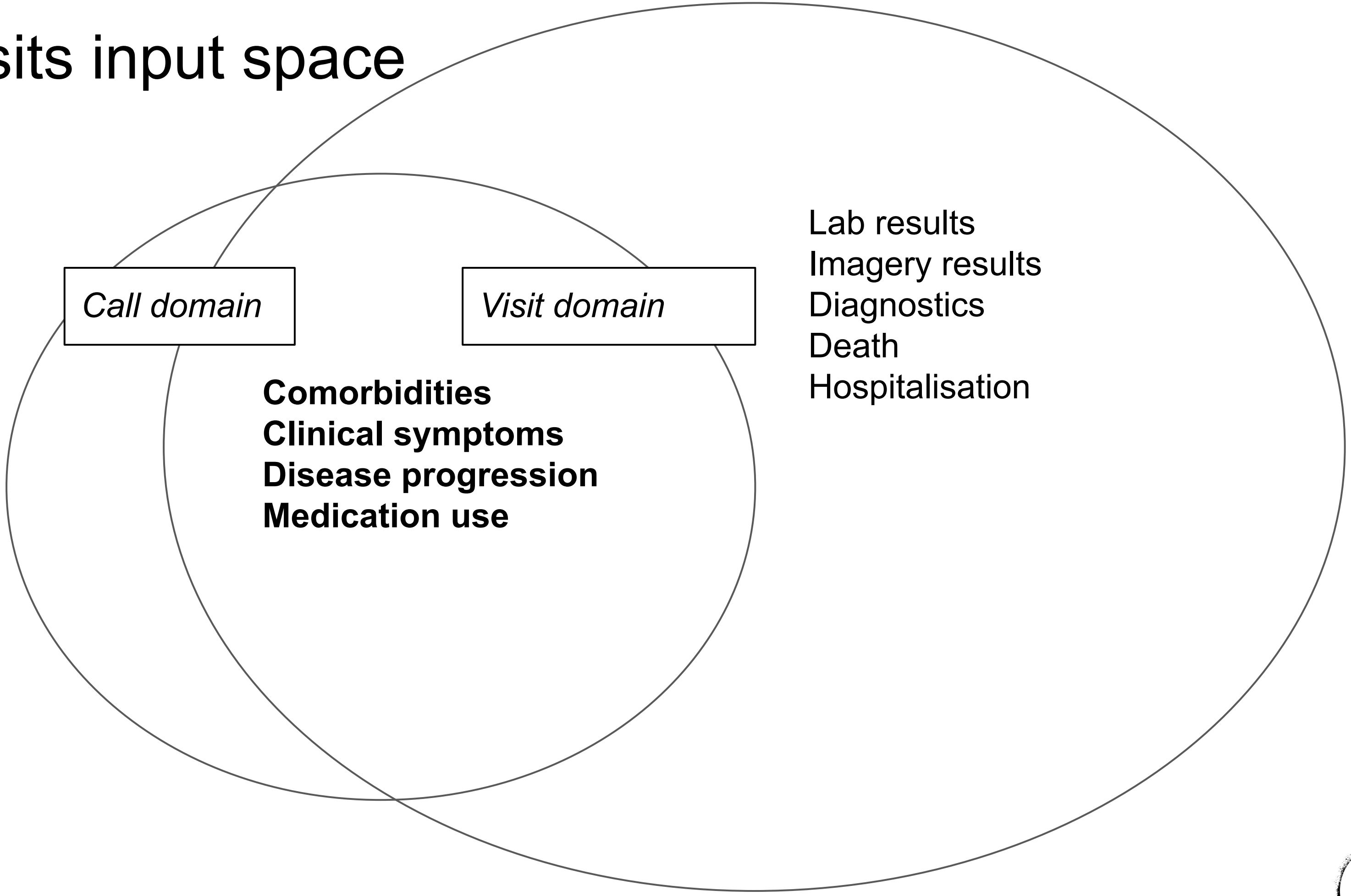
Next-token or masked token prediction



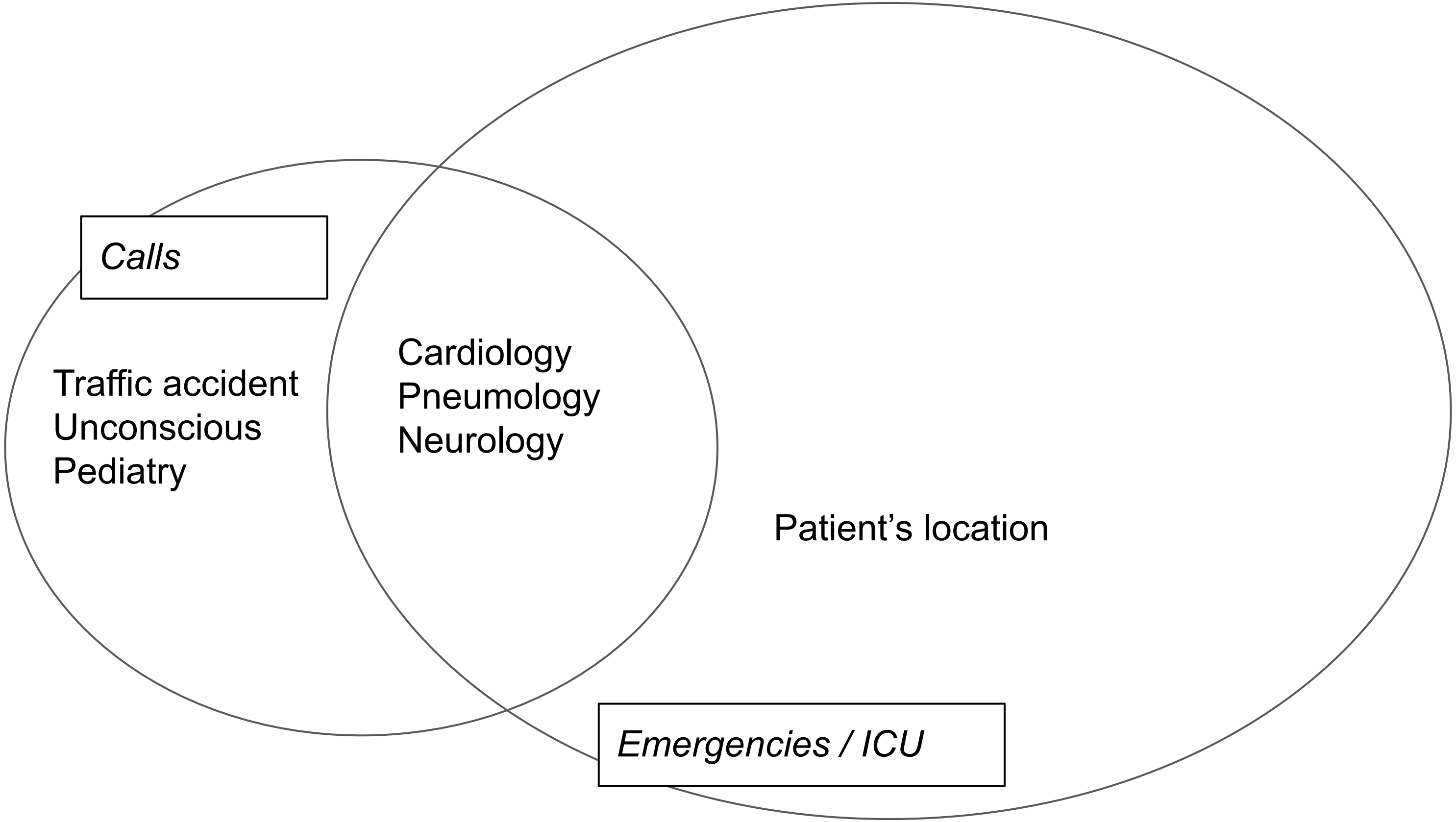
Problem 1bis: Aligning calls / visits representations



Calls / Visits input space



Calls / Visits patient’s state distribution overlap



Problem 2: augmented call triage dialogue modeling

Learning a dialogue model

- learn **m(**previous utterances**) =** next utterance
- in the call triage domain

Augment the dialogue model with the patient's state model

- *inferred patient's state* = $\mathbf{f(g_2(}$previous utterances$\mathbf{))}$
- learn **m(**previous utterances, *inferred patient's state***) =** next utterance

How we plan to do it

Some ideas, open discussion

Everything is language modeling: one model, one self-supervised task

Transform all data into tokens

- discretize continuous variables…
- engineer sentences for UMLS graph relations

Add temporal embeddings, and variable type embeddings

Jointly fine tune an LLM on:

- pseudo medical tokens datasets
- medical call triage
- open source medical ressources datasets (e.g. from Dr Bert paper)

2 models:
1 dialog model, 1 patient's state model

Learn a patient's state model, jointly on the call / visit domain

Infer patient's state on the call dataset

Finetune a dialog model with extended input inferred patient's state

Engineer everything

Define an expert-based set of variables to extract, structure, normalize

Information extraction and entity linking:

- visit domain (notes)
- call domain

Feature engineering, feature selection and statistical modeling

Thank you for your attention !