IOMED

Bringing Intelligence to Clinical Research

www.iomed.es

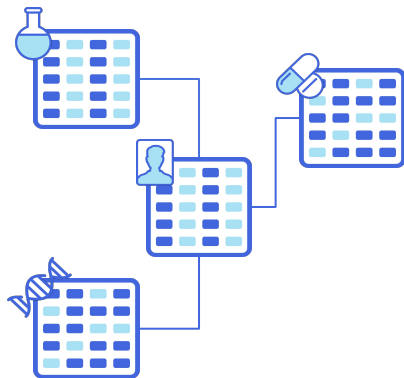# Our goal at IOMED

Accelerating clinical

research…..

by automatizing search

in hospital data

¿What kind of data can we find in the hospitals?

Structured data

Diagnostics, procedures, medication, laboratories, demographic data, etc.

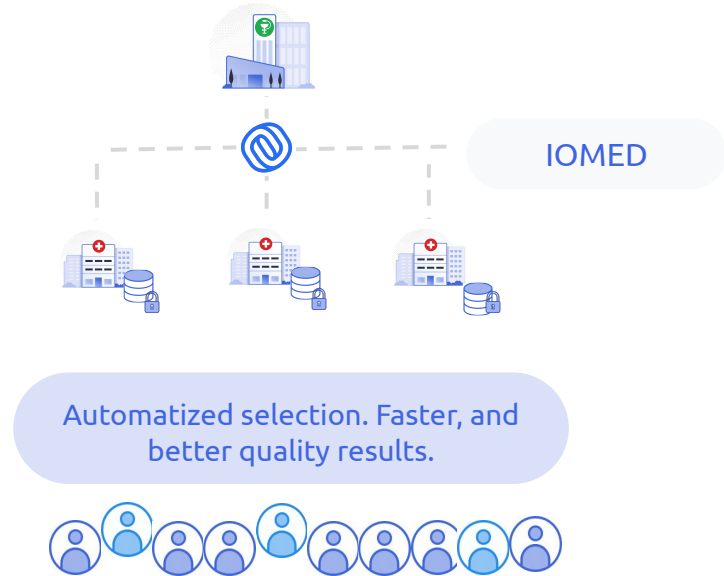## Unstructured data

Text
eg clinical records

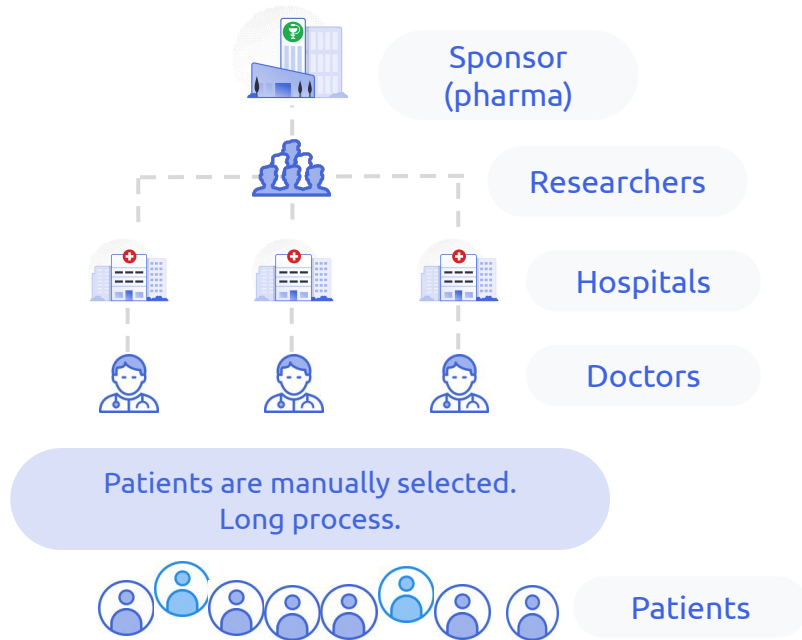**Images**
RMN, TAC, ecos, etc.

80-85% of all data!

Clinical studies

Sponsor (pharma)

Researchers

Hospitals

Doctors

IOMED

Patients are manually selected. Long process.

Automatized selection. Faster, and better quality results.

Patients

Our goal at IOMED

① Clinical Health Records   ② Structured database   ③ Information used for clinical research

NLP
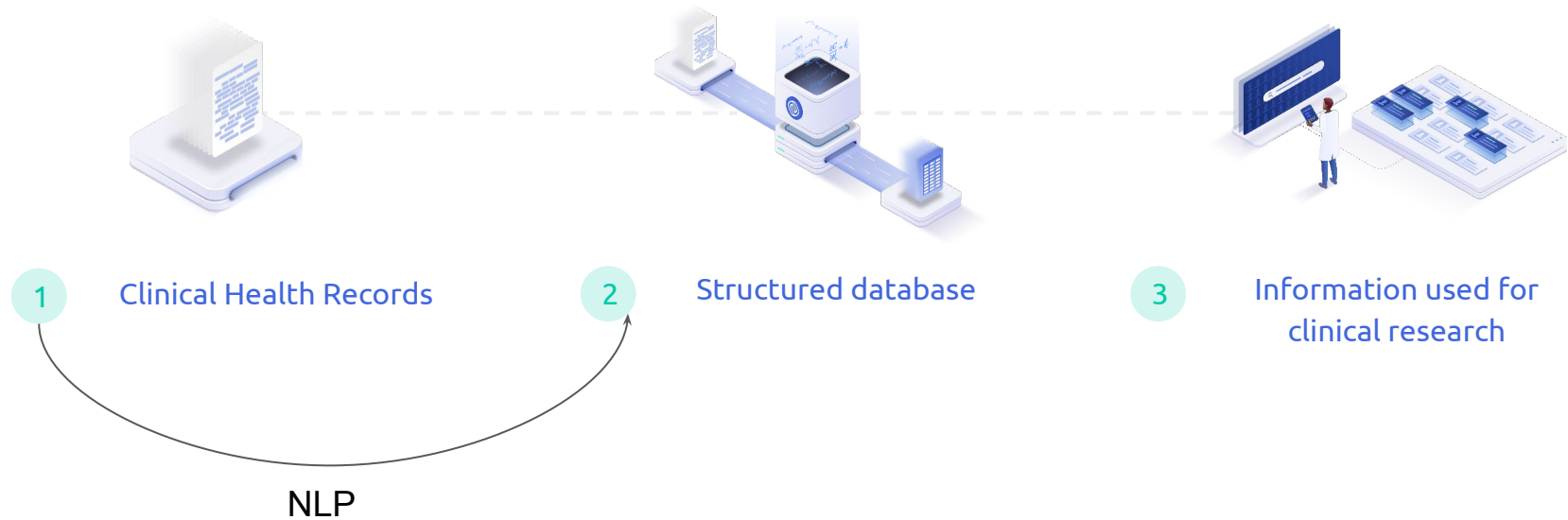
Consulta pro **dolr brazo izquierdo** desde **ayer.**
Est a**mañana persiste dolor.**
Tuvo dolor similar cando tuvo **IAM**.
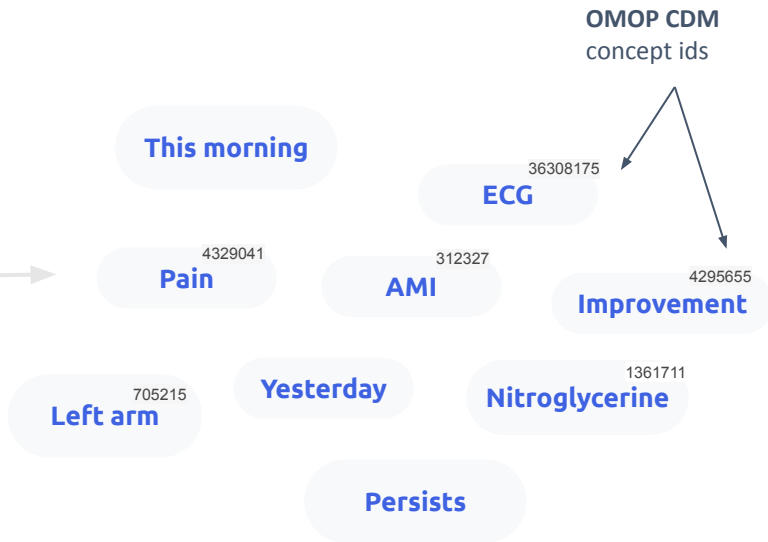Ha tomado **cafinitrina** sin **mejora**.
Solicito **ECG**

	A	B	C	D	E	ECG
1	NIH	IAM	Dolor	Tto	Antecedente	
2	N1087491	03/05/2012	TRUE	1	TRUE	04/
3	N1492453	04/05/2012	TRUE	2	TRUE	NO
4	N3108749	05/05/2012	FALSE	1	TRUE	08/
5	N4941087	06/05/2012	TRUE	1	TRUE	
6	N0874951	07/05/2012	TRUE	2	FALSE	
7	N1087496	08/05/2012	FALSE	4	TRUE	

Named Entity Recognition

OMOP CDM
concept ids

Visits for **lft arm pain** since **yestrday**.
This morning the **pain persists**.
She had similar **pain** when she had **AMI**.
She has taken **nitroglycerine** without
improvement.
I request **ECG**

NER

This morning

ECG
36308175

Pain
4329041

AMI
312327

Improvement
4295655

Left arm
705215

Yesterday

Nitroglycerine
1361711

Persists

Context Detection

Visits for **lft arm pain** since **yestrday**.
This morning the **pain persists**.
She had similar **pain** when she had **AMI**.
She has taken **nitroglycerine** without **improvement**.
I request **ECG**

Pain [present]   **ECG** [future]   **AMI** [past]

Nitroglycerine [past]

improvement [negated]

Paper available at SEPLN: ContextMEL. http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6275

Relation Extraction

Visits for **lft arm pain** since **yestrday**.
This morning the **pain persists**.
She had similar **pain** when she had **AMI**.
She has taken **nitroglycerine** without **improvement**.
I request **ECG**

Pain

ECG

AMI

Located in
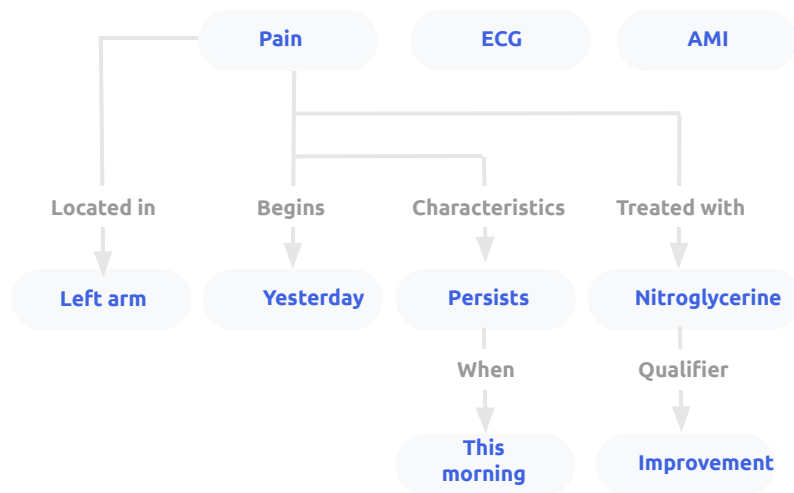
Begins

Characteristics

Treated with

Left arm

Yesterday

Persists

Nitroglycerine

When
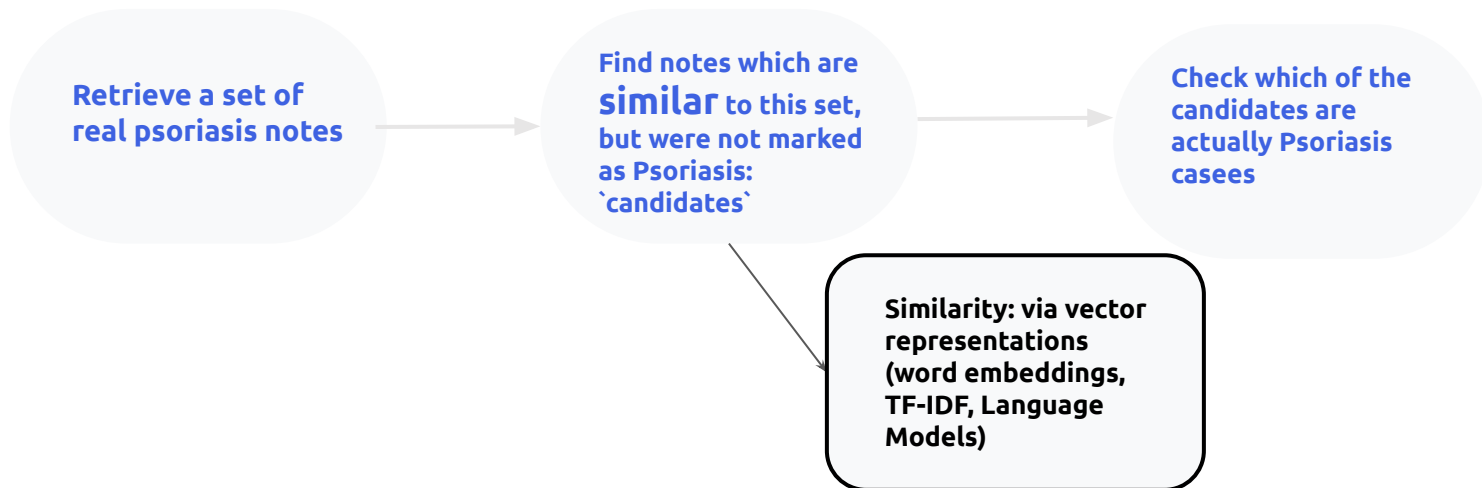
Qualifier

This morning

Improvement

Anonimization

Patient **Hernán Díaz** Visits for lft arm pain since yestrday.
This morning the pain persists.
She had similar pain when she had AMI.
She has taken nitroglycerine without improvement. I request ECG

Patient **Juan González** Visits for lft arm pain since yestrday.
This morning the pain persists.
She had similar pain when she had AMI.
She has taken nitroglycerine without improvement. I request ECG

Semantic Search

- Imagine we are trying to identify notes with Psoriasis, so we run our system on a dataset
- How can we find the notes our system misses?
- Idea: use semantic similarity

Retrieve a set of real psoriasis notes → **Find notes which are similar to this set, but were not marked as Psoriasis: `candidates`** → **Check which of the candidates are actually Psoriasis casees**

Similarity: via vector representations (word embeddings, TF-IDF, Language Models)

Some of the tools we use

- Deep learning models for general problems (NER, contexts, etc)
- Rule-based methods and hybrids for more specific solutions
- Stack: python, Spacy, PyTorch, SQL

Some of the challenges we face

- Processing complete hospitals in a decent time
- Processing remotely, in the hospitals
- Reducing annotation times: active learning
- Transferring our models to other languages

Interested? Any questions?

paula.chocron@iomed.es