

Group work on Medical IR

Georgios Peikos, Wojciech Kusa, Annisa Maulida Ningtyas

Agenda

- Introduction of Information Extraction
- Information Extraction Related Task
- Health Ontology
- MedCAT – Medical Concept Annotation Tool

From unstructured to structured form



- Allergies could makes fatigue
- Author: Erica Cirino
- Reviewer: Elaine K. Luo MD

Can Allergies Make You Tired?



Medically reviewed by Elaine K. Luo, M.D. — By Erica Cirino — July 23, 2018

Allergies happen when your immune system has a strong reaction that generally shouldn't cause a reaction. These substances are called

Most of the time, allergens simply cause mildly uncomfortable symptoms

- coughing
- itching
- sneezing
- skin irritation
- runny nose

Luckily most people with allergies have only mild discomfort. But not everyone complains of feeling tired. Can allergies make you sleepy?

How we reviewed this article:

[SOURCES](#) [HISTORY](#)

Our experts continually monitor the health and wellness space, and we update our articles when new information becomes available.

Current Version

Jul 23, 2018

Written By

Erica Cirino

Edited By

Frank Crooks

Medically Reviewed By

Elaine K. Luo, MD

What is Information Extraction?

Information Extraction – *identify and extract instances of fields of interest from unstructured or semi-structured text and convert to structured representation*

- Field of interest could be **named entities**, relations, or events
- Known by many others name - *named entity extraction, knowledge base population, attribute extraction, ..*

Information Extraction Related Tasks

Identify **medical concepts**: Named Entity Recognition (NER)

`chest pain`

`MRI`

Identify **medical concepts** link to **medical KB**: Medical Concept Normalization

`chest pain` -
`chest pain
(SCUI:
29857009)`

`MRI` -
`Magnetic
Resonance
Imaging (SCUI:
113091000)`

Identify **relationships** **between concepts**: Relationship Extraction

Treats(Prozac,
clinical
depression)

Causes(cataract,
decreased
vision)

Extract **assertions**, **attributes** values

Blood pressure:
high, low, normal

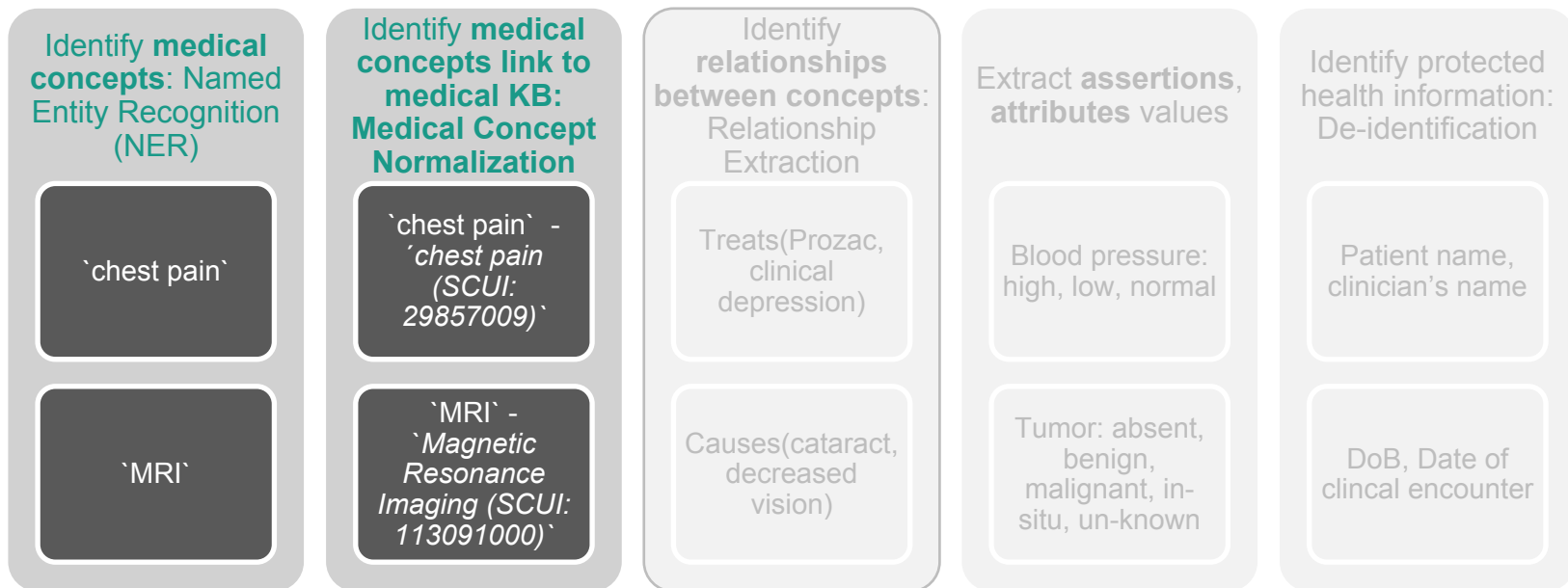
Tumor: absent,
benign,
malignant, in-
situ, un-known

Identify protected health information: De-identification

Patient name,
clinician's name

DoB, Date of
clinical encounter

Information Extraction Related Tasks



An Example: Find all medical concepts

The patient is a 63-year-old female with a three-year history of bilateral hand numbness and occasional weakness.

Within the past year, these symptoms have progressively gotten worse, to encompass also her feet.

She had a workup by her neurologist and an MRI revealed a C5-6 disc herniation with cord compression and a T2 signal change at the level

An Example: Find all medical concepts

The patient is a 63-year-old female with a three-year history of **bilateral hand numbness** and **occasional weakness**.

Within the past year, these symptoms have progressively gotten worse, to encompass also her feet.

She had a workup by her **neurologist** and an **MRI** revealed a **C5-6 disc herniation** with cord compression and a **T2 signal change** at the level

Relevant entities in medical record

Disease: symptoms, problems, signs, conditions

Treatment: medication, surgery, procedure, follow-up

Test: blood work, lab work

Anatomical sites: body parts, regions,

Chemical agent: drugs, substance, ..

.....

Health Ontology – SNOMED CT

- SNOMED CT is a clinical terminology
 - Systematized Nomenclature of Medicine (Clinical Terms)
 - Collection of names and words
 - Relevant to the field of medicine
 - Organised in a structured and logically consistent manner
- Standardized, multilingual vocabulary of clinical terminology used for electronic exchange of clinical health information

More ‘like’ a thesaurus than just a coding scheme

SNOMED CT Content

Clinical Findings

- Disease and deformity – scar
- Symptoms – difficulty breathing
- Social – finding of walking aid use
- Examination findings – tachycardia



Causes of Disease

- Forces – pressure change
- Events – road traffic accident
- Organisms – herpes simplex virus



Procedures

- Laboratory
- Therapy
- Clinical Investigation
- Surgical procedure



SNOMED CT Content

Anatomy

- Normal – knee joint
- Abnormal – ganglion cyst
- Lesions – bony callus



Observations

- Vital signs – blood pressure
- Body product observable – colour of urine
- Values - present

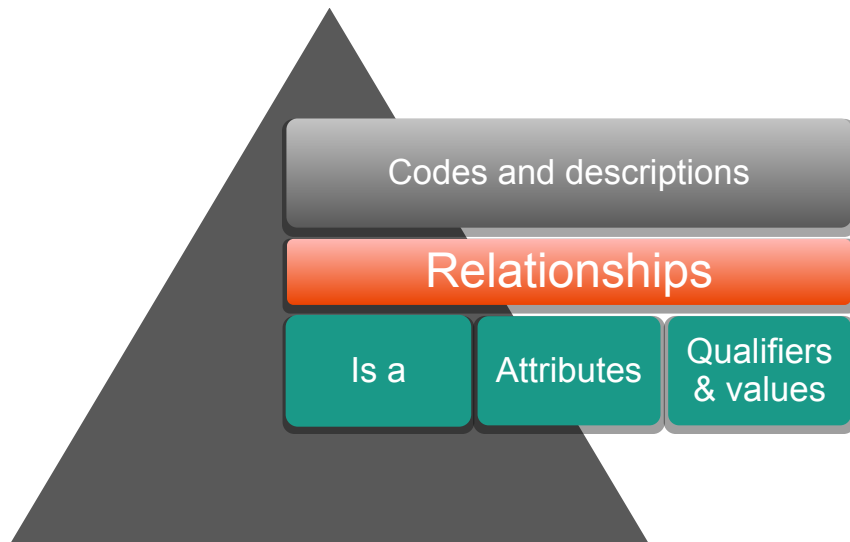


Products

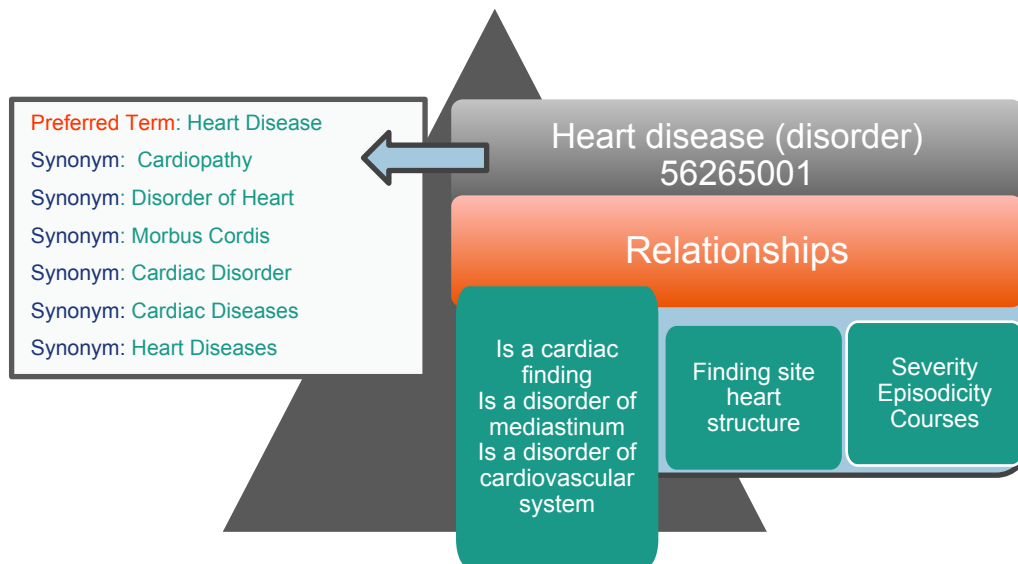
- Drugs – paracetamol
- Substances – latex
- Devices – bedpan



A Concept in SNOMED CT



A Concept in SNOMED CT

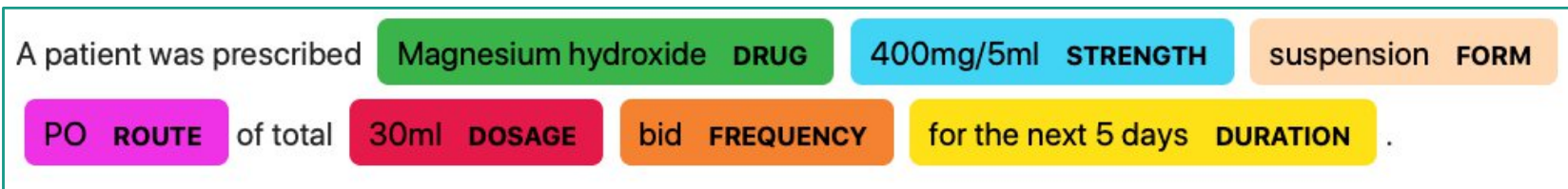


SNOMED-CT Browser tool

<https://browser.ihtsdotools.org/?perspective=full&conceptId1=404684003&edition=MAIN/2021-07-31&release=&languages=en>

IE Techniques - NER

A patient was prescribed Magnesium hydroxide 400mg/5ml suspension PO of total 30ml bid for the next 5 days



Medical NER Tools

- Spacy¹
- Scispacy²
- METAMAP³
- **MedCAT⁴**
- ...

¹<https://github.com/explosion/spaCy>

²<https://allenai.github.io/scispacy/>

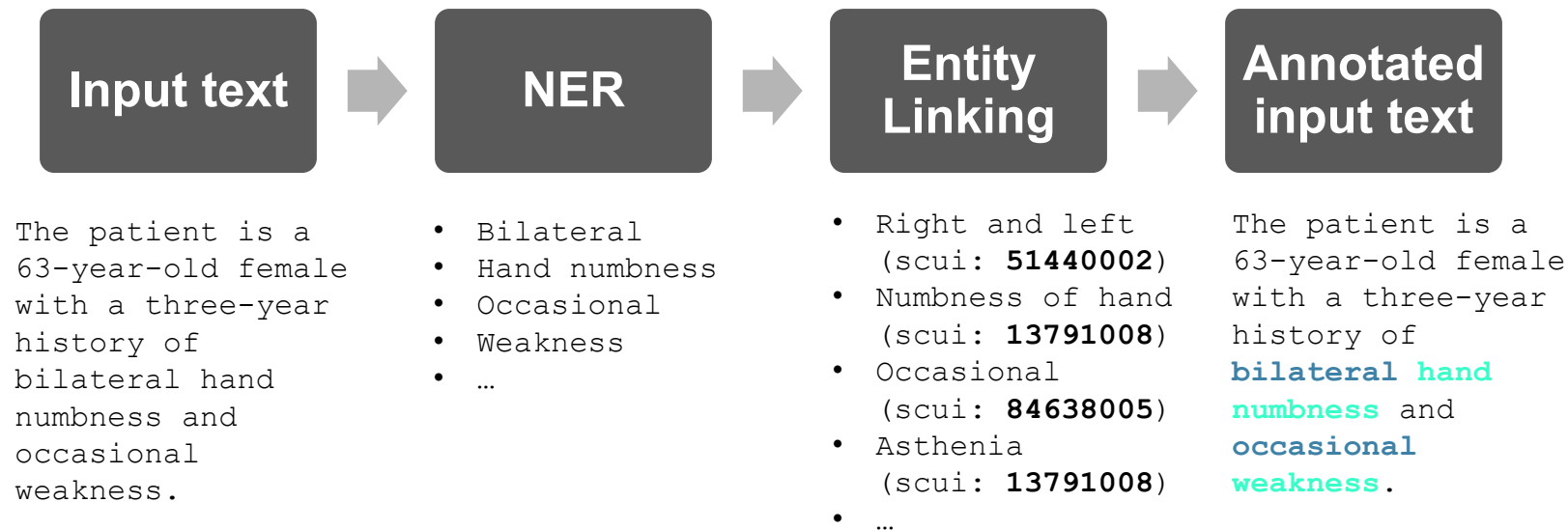
³<https://lhncbc.nlm.nih.gov/ii/tools/MetaMap.html>

⁴<https://github.com/CogStack/MedCAT>

MedCAT - Medical Concept Annotation Tool

- An open-source NER and Linking (NER+L) and an annotation tool and online learning training interface.
- Use to extract information from Electronic Health Records (EHRs) and link it to biomedical ontologies like SNOMED-CT and UMLS.

MedCAT Workflow



MedCAT Demo

<https://medcat.rosalind.kcl.ac.uk/>

Reference

- The content of this slides are taking from coursera course: '**Information Extraction from Free Text Data in Health**'