

Estado actual y tendencias de las Tecnologías del lenguaje en Sanidad



Vicomtech

Vicomtech

- Centro de investigación aplicada en interacción e inteligencia artificial
- Cuenta con un equipo de 15 investigadores especializados en procesamiento del habla y lenguaje natural
- Lleva más de 15 años realizando I+D+i para el sector de la Salud, especialmente en análisis y visualización de imágenes médicas
- Con el auge de la demanda del Big Data en Salud, ha empezado a trabajar en tecnologías del lenguaje en Sanidad



¿Qué demanda el sector Sanitario?

- Herramientas de codificación automática
- Herramientas para estructurar texto libre en la Historia Clínica Electrónica (HCE)
- Herramientas que faciliten la anonimización de textos clínicos

Herramientas de codificación automática

- Motivadas por la necesidad de agilizar el proceso de codificación de la HCE
- Existen varias soluciones en el mercado, basadas principalmente en variantes enriquecidas de búsquedas léxicas partiendo del diagnóstico determinado por un médico
- No es viable por el momento dar el salto hacia paradigmas de clasificación basados en aprendizaje automático que exploten toda la información disponible en la HCE, dada la cantidad de datos etiquetados de entrenamiento necesarios para cada código

Herramientas para estructurar texto libre en la HCE

- Motivadas por el auge del Big Data en salud, tanto para investigación clínica como para soporte a la decisión
- MetaMap¹ es la herramienta de referencia en este ámbito para texto clínico en inglés. Es capaz de mapear texto biomédico al Metatesauro UMLS. Combina un enfoque intensivo en conocimiento con técnicas de procesamiento de lenguaje natural.
- Existen algunas herramientas de este tipo para español en el mercado.

¹MetaMap, a Tool for Recognizing UMLS Concepts in Text
<https://metamap.nlm.nih.gov>

Herramientas para estructurar texto libre en la HCE

UMLSmapper²

The screenshot displays the UMLSmapper interface, which is used for biomedical term normalization. It is divided into three main sections:

- Concepts by semantic group:** A vertical sidebar on the left with color-coded buttons for different semantic groups: Disorders (red), Procedures (orange), Anatomy (green), Living Beings (teal), Chemicals and Drugs (blue), and Objects (purple).
- Annotated text:** The central panel shows a sample medical text with terms highlighted in colored boxes corresponding to the semantic groups. The text includes: "Se trata de una paciente de 37 años con enfermedad de Raynaud de larga duración y vasculopatía necrotizante.", "Alergias: látex, penicilina y neomicina. Dermatitis alérgica de contacto por níquel, colofonia, formaldehído, etilendiamina, mercaptobenzotiazol, y budesonida.", "No tabaquismo. Hematuria recurrente desde febrero de 2006 con presentación de 6 - 8 episodios con estudio urológico negativo. En sedimentos de orina presencia habitual de microalbuminuria y microhematuria. Pericarditis en octubre de 2007. Miopericarditis en noviembre de 2007. Osteocondritis en 4ª costilla izquierda (gammagrafia sugestiva de patología benigna costochondral) que requirió tratamiento corticoideo.", "En abril de 2008 presenta bacteriemia por E. coli, precisando tratamiento con Linezolid y Ciprofloxacino con aparición de hepatitis aguda medicamentosa y pancitopenia secundarios, precisando cambio de antibiótico con Aztreonam con buena evolución.", "IQs: lesión meniscal, articulación temporomandibular en dos ocasiones precisando injerto de cresta iliaca."
- Concept details:** The right panel provides detailed information for the selected concept, "enfermedad de Raynaud". It includes:
 - Parent Concepts:** enfermedad de Raynaud
 - UMLS ID:** C0034734
 - Semantic group:** Disorders
 - Semantic type:** Disease or Syndrome
 - Table of Term, Source, and ID:**

Term	Source	ID
Raynaud	MDRSPA	
enfermedad de Raynaud	SCTSPA	195295006
 - Description:** A scrollable text area containing a definition: "blue. When the blood flow returns, the skin turns red and throbs or tingles. In severe cases, loss of blood flow can cause sores or tissue death. Primary Raynaud's happens on its own. The cause is not known. There is also secondary Raynaud's, which is caused by injuries, other diseases, or certain medicines. People in colder climates are more likely to develop Raynaud's. It is also more common in women, people with a family history, and those over age 30. Treatment for Raynaud's may include drugs to keep the blood vessels open. There are also simple things you can do yourself, such as..."
 - Subconcepts:**
 - fenómeno de Raynaud
 - síndrome CREST

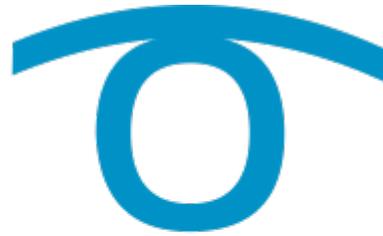
²UMLSmapper, Biomedical term normalization of EHRs with UMLS, LREC 2018
<https://snlt.vicomtech.org/u/mlsmapper>

Herramientas que faciliten la anonimización de textos clínicos

- Motivadas por la nueva ley de protección de datos de carácter personal europea, GDPR
- La mayoría de herramientas de anonimización disponibles en el mercado están orientadas a anonimizar datos estructurados.
- Se pueden combinar técnicas de procesamiento de lenguaje natural (p.ej. detección de entidades nombradas) y aprendizaje dinámico para facilitar la identificación y anonimización de datos de carácter personal en los textos clínicos

Otras tendencias tecnológicas

- Detección y desambiguación de abreviaturas clínicas
- Detección de negación y especulación en textos clínicos
- Enriquecimiento automático de terminologías y tesauros
- Identificación de relaciones entre conceptos clínicos



www.vicomtech.org
info@vicomtech.org

