

# Desarrollo de componentes de tecnologías del lenguaje clínico y biomédico para datos en español y lenguas co-oficiales

Maite Oronoz  
(en nombre del grupo  
de investigación)



Universidad  
del País Vasco  
Euskal Herriko  
Unibertsitatea

**HiTZ**  
Hizkuntza Teknologiko Zentroa  
Basque Center for Language Technology



# Grupo de investigación IXA: Personal



Grupo de investigación

- multidisciplinar
- de la UPV/EHU
- formado por **54 miembros**:

- 27 mujeres (15 Ingenieras, 10 lingüistas, y 2 técnicas)
- 27 hombres (21 ingenieros, 5 lingüistas y 1 técnico)
- 8 miembros trabajando en el dominio médico

Considerado como **grupo consolidado** por la UPV/EHU

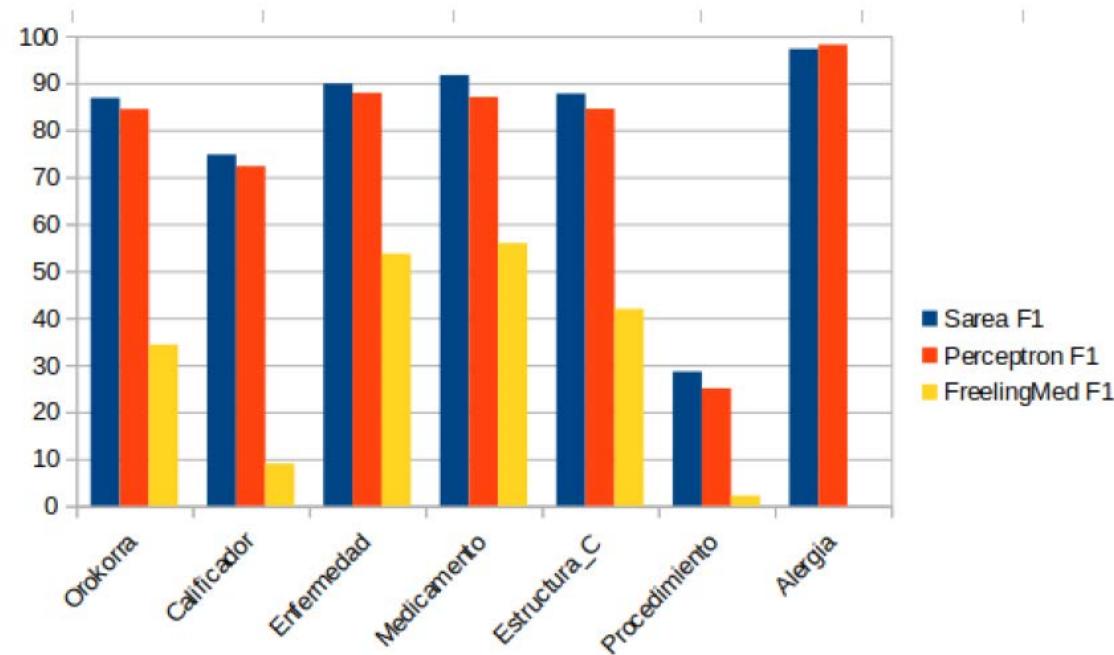




## Subgrupo medicina y análisis textos médicos

### Trabajo en español

- FreeLingMed: Analizador Freeling adaptado (inclusión de SNOMED CT, CIE-10, abreviaturas médicas etc.)
- Etiquetador de conceptos médicos (3 aproximaciones):
  - FreeLingMed (es, en)
  - Perceptron
  - Redes Neuronales

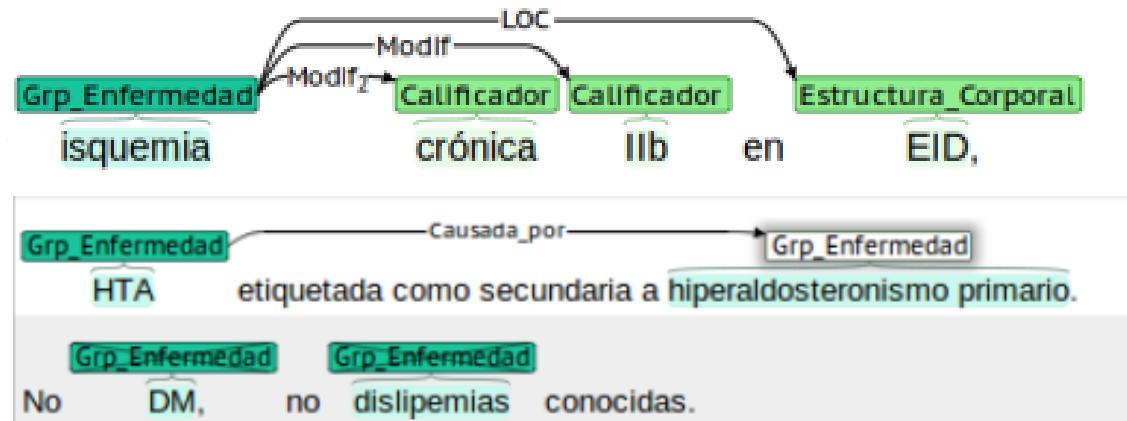




## Subgrupo medicina y análisis textos médicos

### Trabajo en español

Entidades y relaciones médicas (incluidas ADRs)



DTs expressed by doctors in MRs	ICD
Adenocarcinoma de prostata	185
Ca. prostata	185
CÁNCER DE PRÓSTATA	185
adenocarcinoma de pulmon estadio IV	162.9
CA pulmón estadio 4	162.9
ADENOCARCINOMA PÁNCREAS	157.9

Clasificación CIE-10

Standard DTs from ICD-reference	ICD
NEOPLASIA MALIGNA PROSTATA	185
N. MALIGNA DE BRONQUIO Y PULMON SIN ESPECIFICAR	162.9
N. MALIGNA DE PANCREAS PARTE SIN ESPECIFICAR	157.9





## Subgrupo medicina y análisis textos médicos

### Trabajo en euskera

- Traducción automática de terminología médica:
  - en-eu → SNOMED (sin validar)
  - en-eu → CIE-10-MC (parte validado en Osakidetza)
- Traducción automática de informes médicos reales
  - Contrato con Osakidetza
  - eu → es
  - Traductor automático neuronal
    - Recoger corpus paralelo de informes médicos en centros de salud
    - Generar traductor
    - Post-edición