

# Proceso de Anonimización de texto libre

Proceso semiautomático cuya finalidad es eliminar cualquier texto que pueda comprometer la identidad de un paciente.

## Anonimización de texto libre en BIFAP

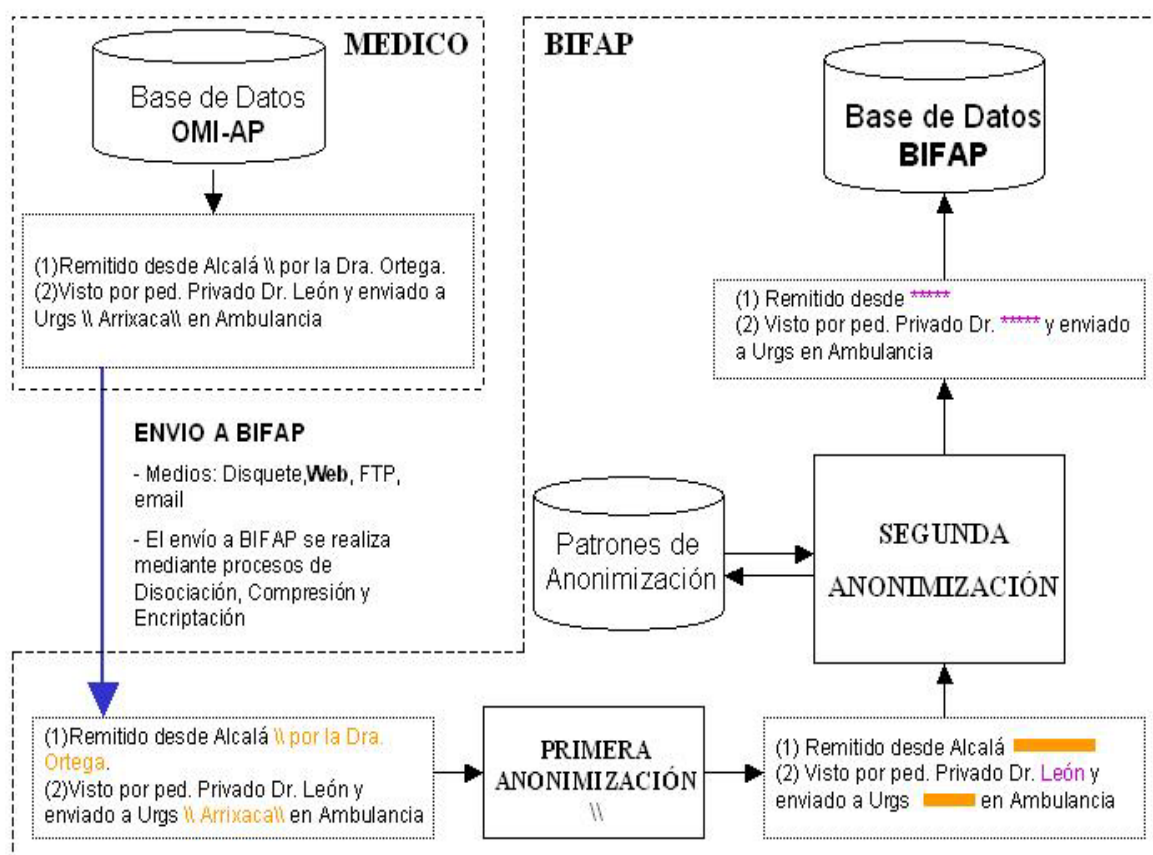


Figura 1. Proceso esquematizado de anonimización

### PRIMERA ANONIMIZACIÓN (con la colaboración del médico)

- Se eliminan datos que el médico estima son confidenciales y ha introducido entre dobles barras “\\ dato \\\” o después de una doble barra “\\ dato’.

### SEGUNDA ANONIMIZACIÓN

- Se basa en una base de datos de términos en los que se incluyen nombres, apellidos, hospitales, pueblos, países ...
- Actualmente existen cerca de 11.000 términos convenientemente clasificados para que la efectividad del algoritmo sea mayor.
- El proceso ‘aprende’. El proceso analiza las entradas y las salidas y reentrenándolo toma decisiones como añadir términos a la base de datos de términos, modificar la clasificación del término...

## **RESULTADOS**

Efectividad: deja de eliminar +/- unos 2 términos de baja sensibilidad por cada 1000 comentarios. Esta efectividad puede mejorarse ajustando los parámetros del algoritmo, con nuevos procesos de reaprendizaje...