

---

# Introducción a Tabula y OpenRefine

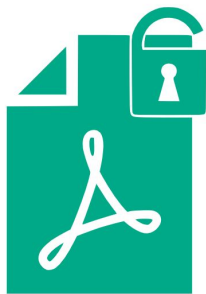
Extracción y limpieza de datos

---

# Tabula

**Extrae datos de tablas guardadas en PDF**

- Es fácil de usar
- Tiene limitaciones
- Es gratis



# Open Refine

**Simplifica tareas de limpieza**

- No es tan intuitivo
- Útil para datasets grandes
- Es gratis





## Claves de Tabula

- Funciona mejor cuanto más limpia sea la estructura de los datos.
- Si nos interesan páginas concretas, conviene hacer limpieza antes (Small PDF)

# Claves de OpenRefine

Ideal para pulir conjuntos de datos que se han creado a mano.

Más potente en limpieza que en análisis (Cluster)

Permite abrir archivos que superan la capacidad de Excel.

*Esta historia es solo para fines ilustrativos.*



