

Integración de IA en Sistemas de Gestión Documental de RTVE

Marcos Pérez

Tabla de contenidos

- ▶ **Introducción**
- ▶ **Integración**
- ▶ **Resultados**
- ▶ **Futuro**

INTRODUCCIÓN

Base documental - actual

- ▶ **Almacena metadata y objetos asociados**
- ▶ **Alta / Baja / Modificación**
- ▶ **Búsqueda**

Base documental - actual

The screenshot shows the ARCA search interface with a grid of search results. The results are categorized by date and include various cultural and historical events, such as 'OBSERVATORIO DEL ROCK DE LOS MUCHACHOS', 'CAMARÓN DE LA ISLA', and 'ANTONIO SAINT EXUPÉRY'. Each result includes a thumbnail, a title, a date, and a 'Ver video' button.

Bienvenida

The screenshot shows the ARCA document viewer interface. It features a video player displaying a concert by the group Celtas Cortos. The interface includes a detailed metadata section on the left, a video player in the center, and a storyboard on the right. The metadata includes information such as the document title, production details, and technical specifications.

Documento

Base documental - Desafíos

- ▶ **Demasiado expertos en la base para buscar**
- ▶ **Mucha labor manual**
- ▶ **Cada vez se genera más objetos que documentar**

Base documental - Objetivos

- ▶ **Reducir tiempos en tener algo**
- ▶ **Usabilidad en general, para acercar a otros usuarios**
- ▶ **Mayor información sobre lo que hay**

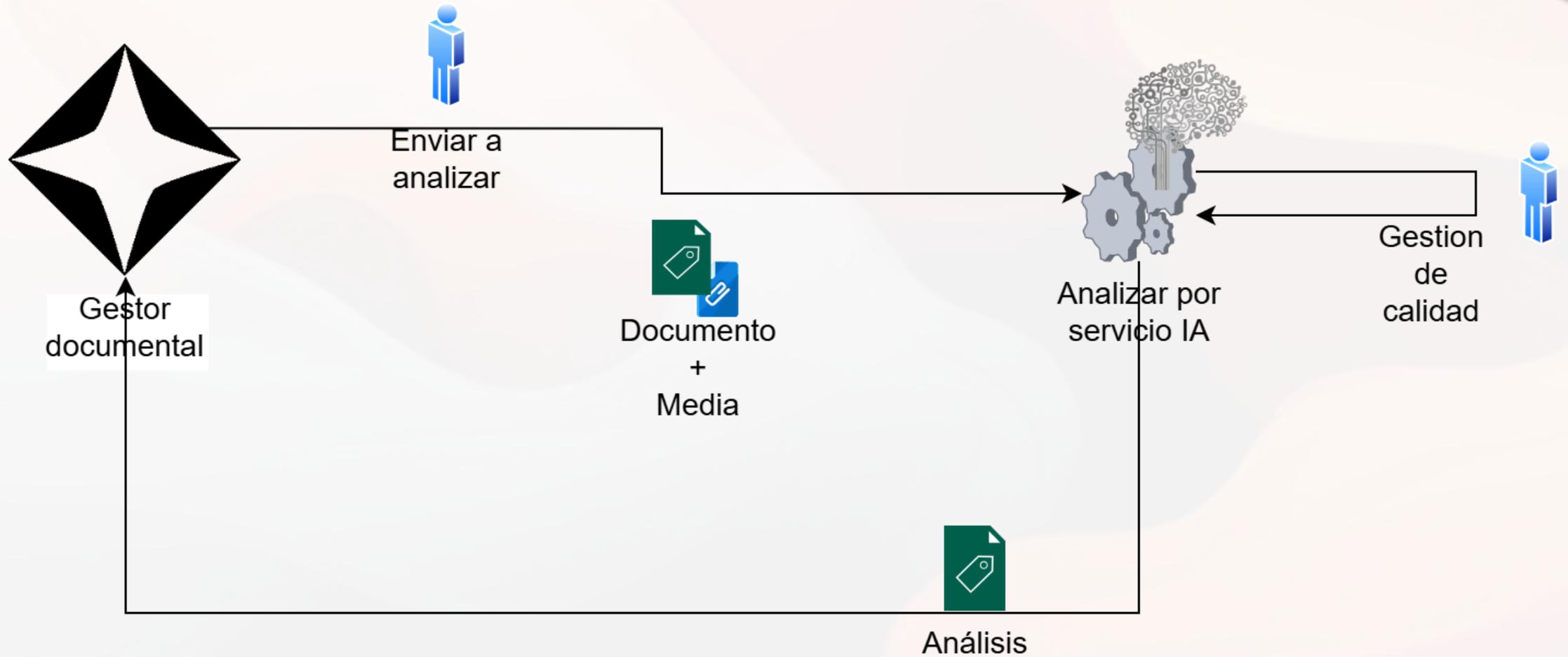
Inteligencia Artificial

- ▶ **Realizar tareas típicas de humano**
- ▶ **Aprendizaje y adaptación**
- ▶ **Mejor interacción hombre-máquina**

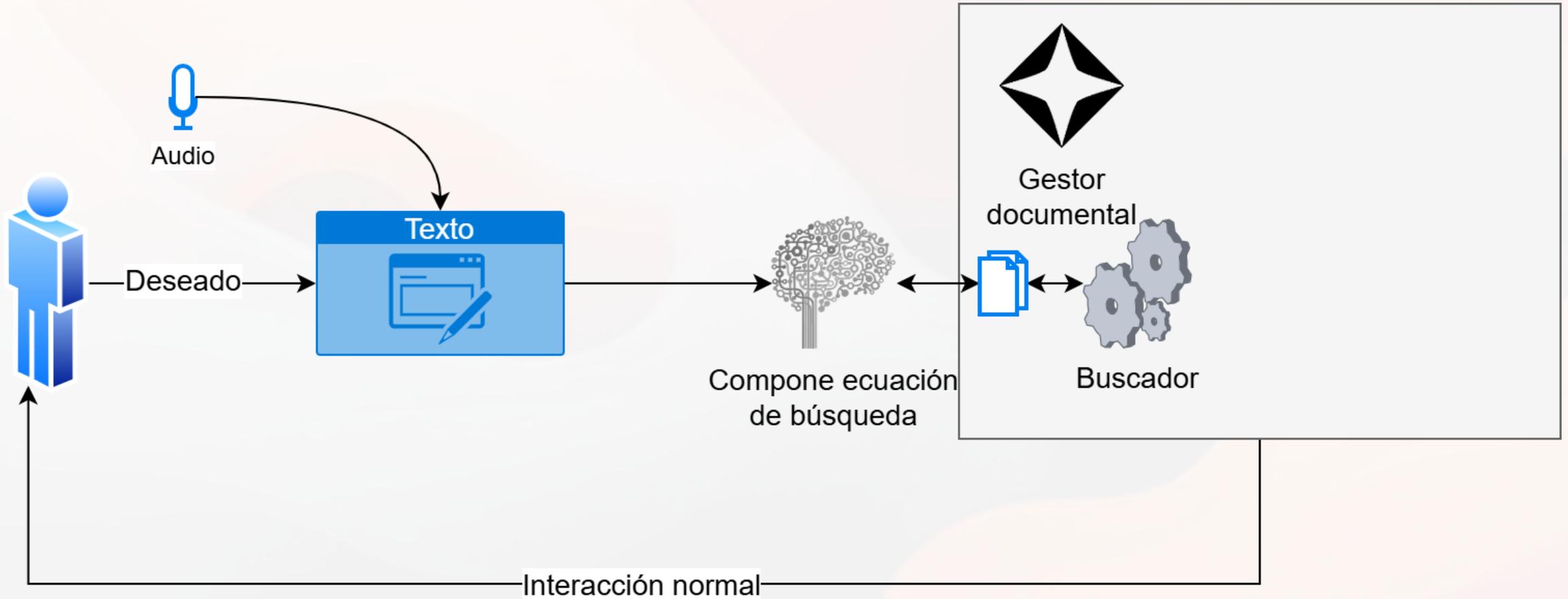


INTEGRACIÓN

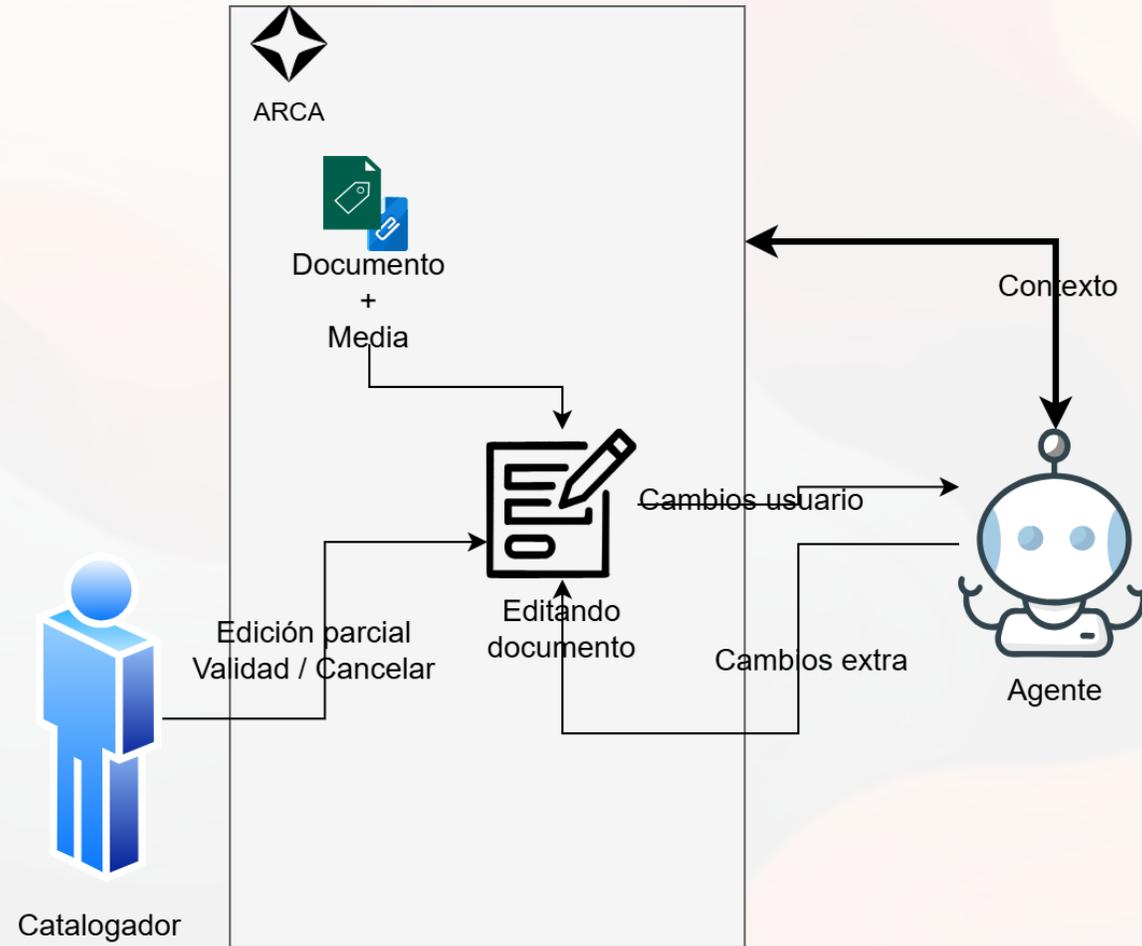
Uso - Análisis automático de objetos



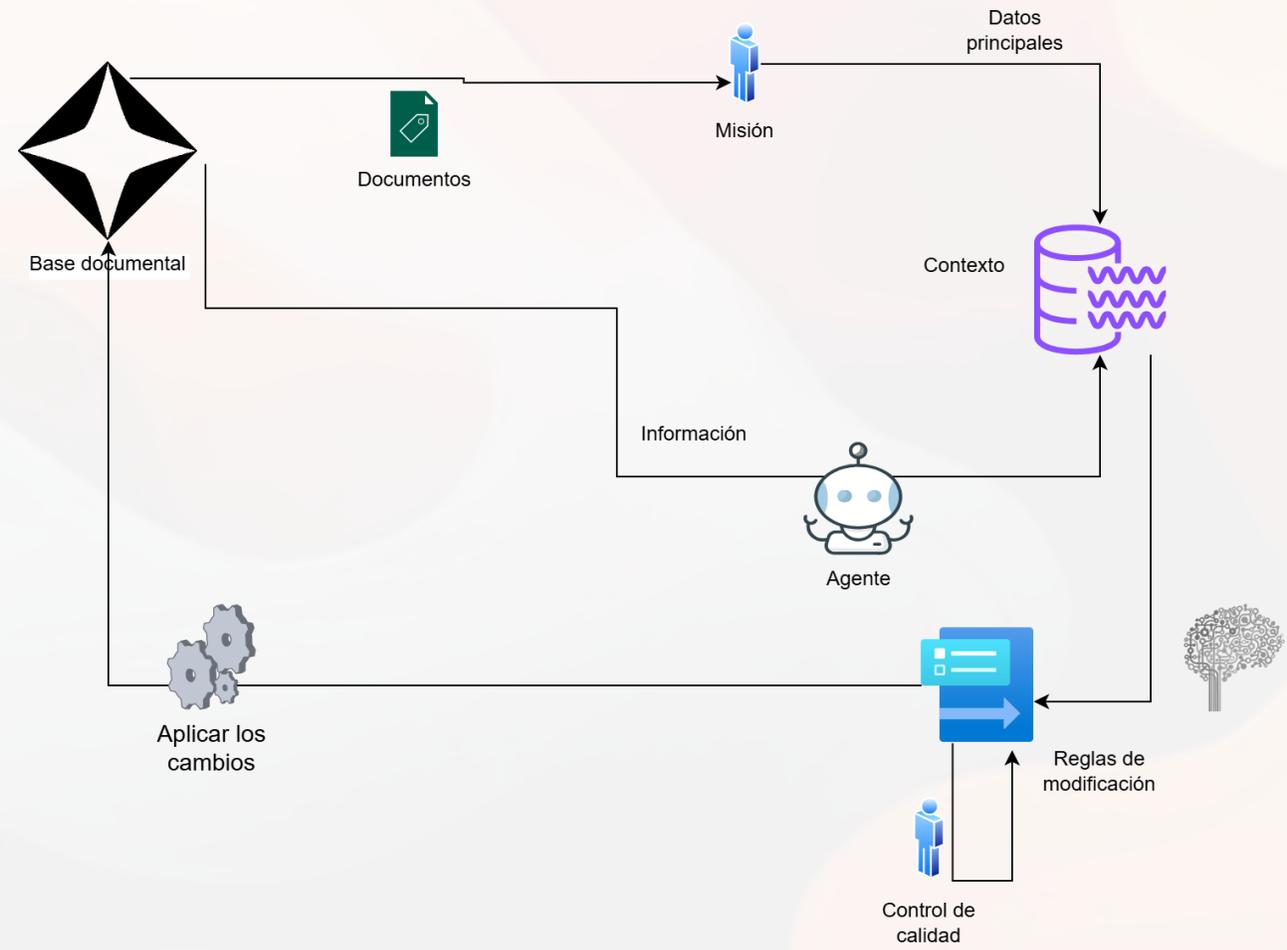
Uso - Ayuda a la búsqueda



Uso - Ayuda al análisis



Uso - Modificación masiva de documentos

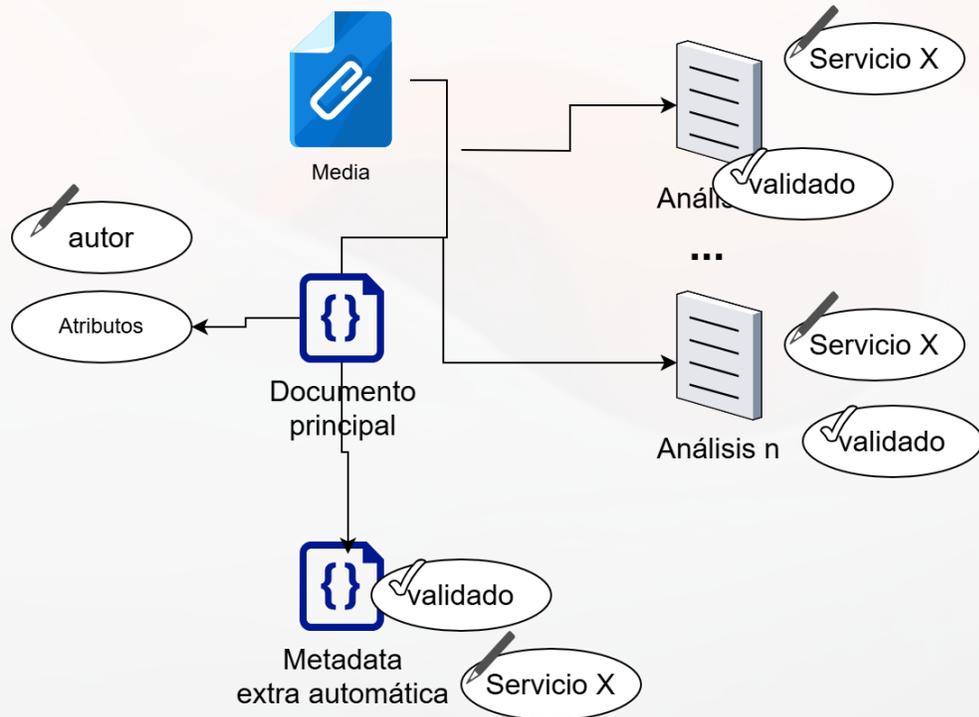


Uso - Datos Meta información

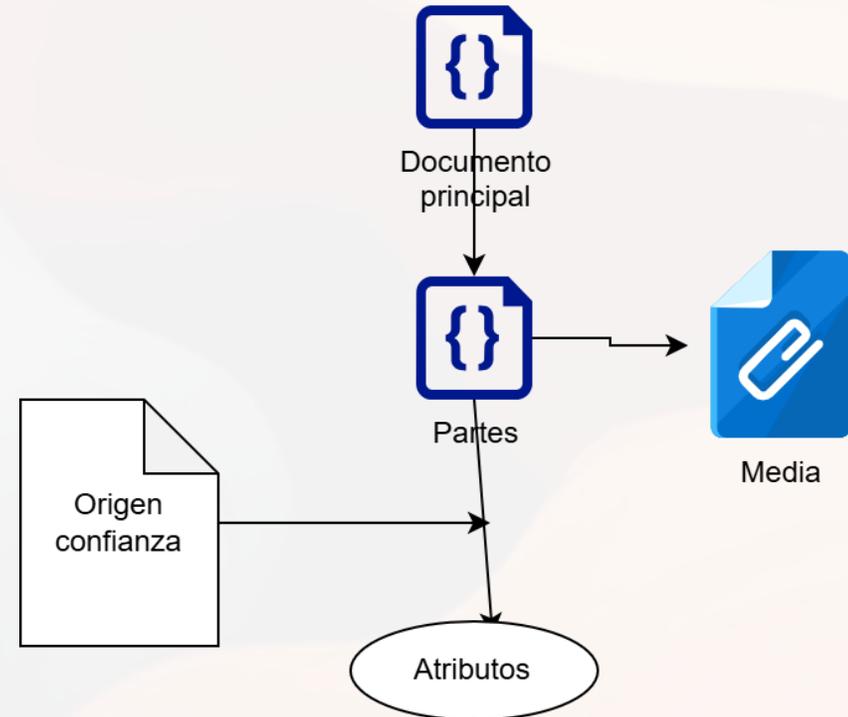
- ▶ **Confianza**
- ▶ **Autor (versión, configuración)**
- ▶ **Validado**

Uso - Modelo de datos Nuestro caso

Deuda técnica



Modelo heredado



Modelo generalizado

Proceso de integración

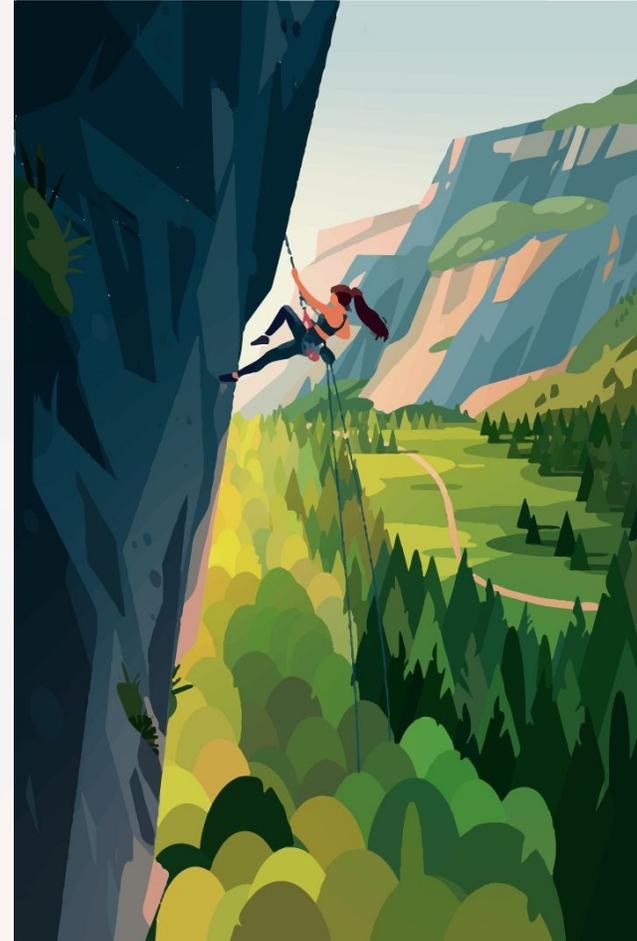
- ▶ **Gradual**
- ▶ **Interno VS Externo**
- ▶ **Expertos**
- ▶ **¿Commodity?: S2T, OCR**
- ▶ **Uso de tecnologías conocidas en la casa**

IAs utilizadas

- ▶ **Productos instalables (Opentext Media server, easyocr)**
- ▶ **Pesos abiertos (S2T Whisper)**
- ▶ **Servicios externos (Mediante integradores)**

Desafíos

- ▶ **Cantidad de datos
manejados**
- ▶ **Control tareas**
- ▶ **Control de calidad**
- ▶ **Departamentos
involucrados**



Privacidad y seguridad

- ▶ **Confidencialidad**
- ▶ **Datos dentro / fuera U.E.**
- ▶ **Cuidado con Caras y voces (reconocer famosos)**

RESULTADOS DE INTEGRACIÓN

Grandes cantidades de información

- ▶ **“Media” que sale**
- ▶ **Información que vuelve**

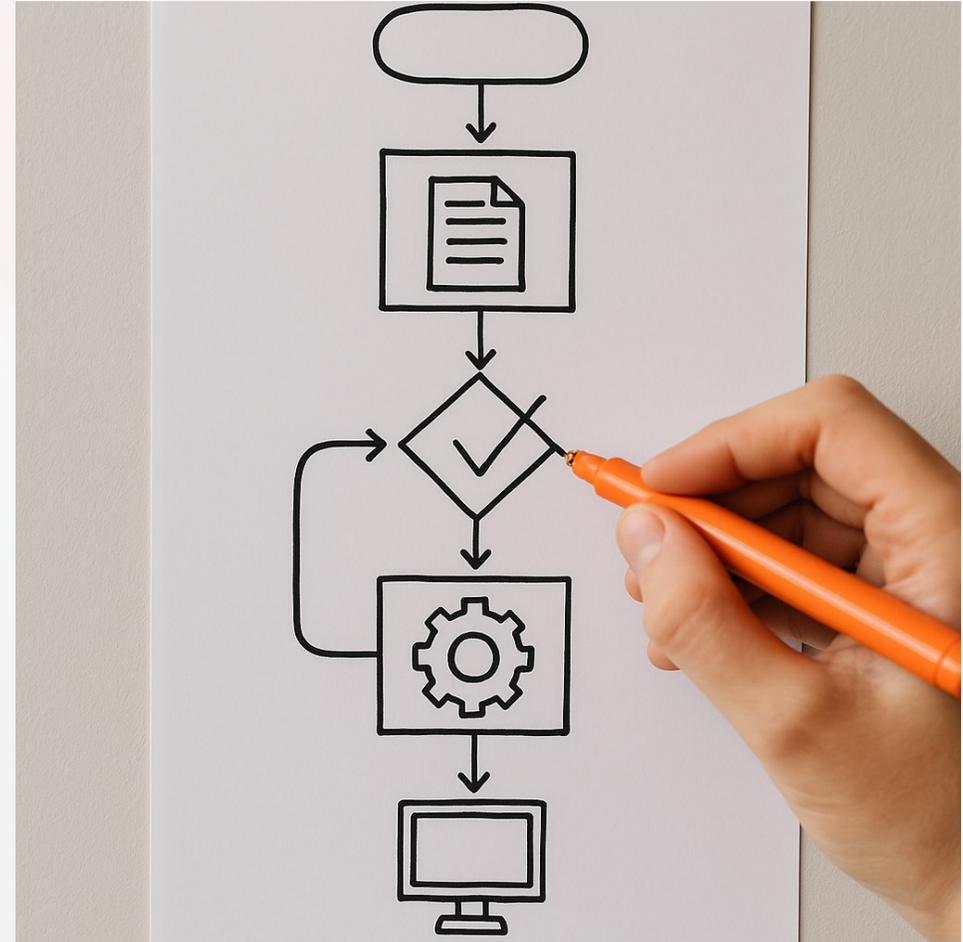


Volumetría RTVE

- ▶ **Horas de audio / video gestionadas 3M**
- ▶ **Número de diferentes análisis definidos (12)**
- ▶ **Número de diferentes análisis almacenados 100K**

Cambios en los flujos de trabajo

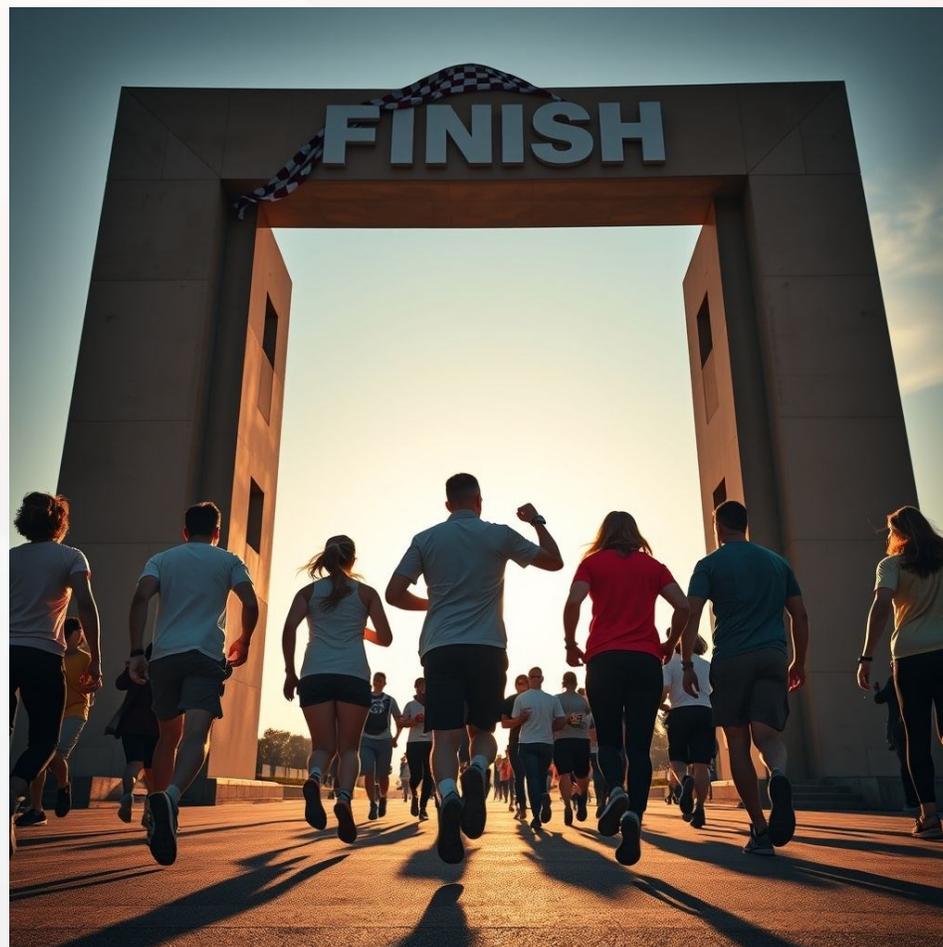
- ▶ **Automatizar llamadas a analizar**
- ▶ **Documentalistas pueden partir de algo**
- ▶ **Expertos revisando tareas realizadas**



Validación resultados

- ▶ **Gestión de calidad, con gente dedicada a evaluar y proponer mejoras**
- ▶ **Cada vez se analiza más**
- ▶ **Cada vez se quiere tener analizado antes**
- ▶ **Se busca por los resultados**

¿Final?



FUTURO

Orientado a la aplicación

- ▶ **Modelo de datos más moderno y generalizado**
- ▶ **Facilidad de añadir nuevos análisis**
- ▶ **Gestión ciclo petición análisis y regreso resultados**

Orientado a la búsqueda

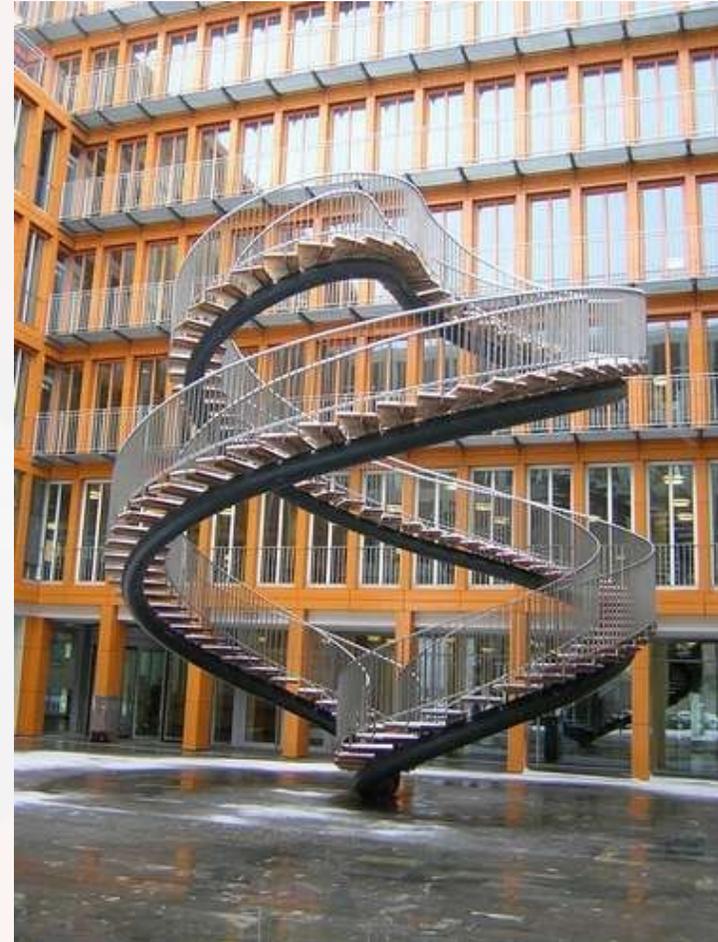
- ▶ **Búsqueda vectorial**
- ▶ **Asistente a la búsqueda**
- ▶ **Refinar búsqueda**
- ▶ **Lenguaje natural**

Orientado a la edición

- ▶ **Asistencia más personalizada a la edición**
- ▶ **Afianzar relaciones entre documentos**

Orientado a la mejora continua

- ▶ **Modelos adaptándose a novedades**
- ▶ **Incorporar recalcuro de análisis**
- ▶ **Escalado, requiere cambios profundos en arquitectura**
- ▶ **Cambios en modelo de datos**



FINAL

Final

Gracias

●▶ **Marcos Pérez** (marcos.perez@rtve.es)

– <https://www.linkedin.com/in/marcos-p%C3%A9rez-294626b>