



Unstructured Information Management Architecture

Ernesto Alonso Cortés Prada

Unstructured Information



De **mayor tamaño**



Más común o la **más presente actualmente.**



De **mayor crecimiento**



Texto



Voz o audio



Video o imagen

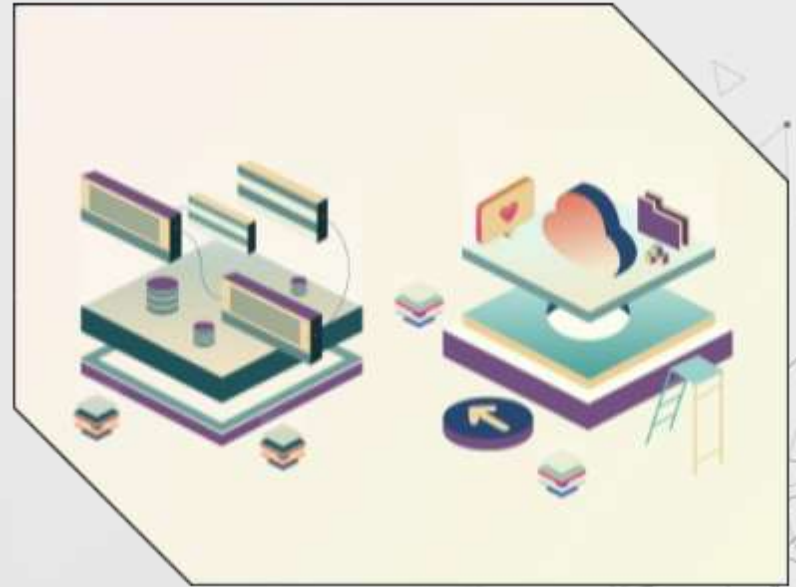


El contenido de alto valor está enterrado entre mucho **ruido, datos atípicos** o de **poca utilidad.**

Unstructured Information Management (UIM)

Aplicaciones o software que analizan de grandes volúmenes de información no estructurada para:

- 01** Descubrir.
- 02** Organizar.
- 03** Entregar conocimiento relevante.





**PROCESAMIENTO DE
LENGAJE NATURAL (NLP)**

01

RECUPERACIÓN DE
INFORMACIÓN (IR)

02

**APRENDIZAJE
AUTOMÁTICO**

03

TECNOLOGÍAS USADAS POR UIM

04

ONTOLOGÍAS

05

**RAZONAMIENTO
AUTOMATIZADO**

06

FUENTES DE
CONOCIMIENTO

Apache UIMA

Apache UIMA es una implementación de código abierto del estándar UIMA, desarrollado por originalmente por IBM y escrito en Java y C++. Este es un marco de arquitectura y software que permite la comunicación entre la información estructurada y no estructurada.



**Unstructured
Information
Management
Architecture**

An Apache Project

UIMA

Analysis engines

Analiza e infiere y registra atributos descriptivos sobre un documento (completo o partes).

Annotators

Componente básico que se incluye en los analysis engines

Analysis results

Resultado del análisis hecho en los motores de análisis.

01 The topic of document D102 is "CEOs and Golf"

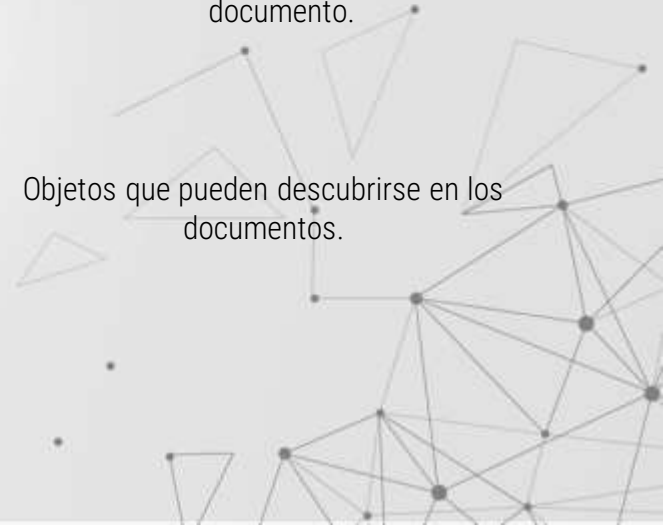
02 The span from position 101 to 112 in document D102 denotes a person

Common analysis structure

Contiene en análisis de la estructura de un documento.

Type systems

Objetos que pueden descubrirse en los documentos.



Bibliografia

1. *Apache. (s. f.). Apache UIMA. Apache UIMA. Recuperado 17 de mayo de 2022, de <https://uima.apache.org>*
2. *IBM. (s. f.). UIMA concepts. Recuperado 17 de mayo de 2022, de <https://www.ibm.com/docs/en/db2/9.7?topic=examples-uima-concepts>*

