

UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

Definición de Dominio y Aplicación para el Proyecto

INTEGRANTES:

- Katherine Ximena Morocho
- María Fernanda Ordóñez

1. DEFINICIÓN DEL DOMINIO DE TRABAJO

Nuestro proyecto va a ser realizado dentro del dominio de Educación, específicamente a los metadatos de OREs (Recursos Educativos Abiertos), ya que en la actualidad es un tema de suma importancia y con mucha fluidez de la información que podríamos utilizar para el aprendizaje inteligente.

2. SELECCIÓN DE AL MENOS DOS FUENTES DE DATOS DEL DOMINIO ELEGIDO

Después de una lectura y análisis de las fuentes de datos para nuestro dominio, hemos decidido usar la extracción de datos de perfiles de Google Scholar en la que utilizamos un algoritmo en el lenguaje R, con la que haremos la minería de datos. Para la extracción de los perfiles se utilizará el método Web Scraping online, ya que él se ha comprobado que es el método con el mejor tiempo de extracción de perfiles.

A continuación, citamos un ejemplo donde Web Scraping online selecciona los perfiles de un tema determinado; como universidades y el número de publicaciones que tenía cada universidad.

Universidad	País	#Perfiles	#Publicaciones
UFM	Guatemala	14	393
ESPOL	Ecuador	67	1061
FURB	Brasil	38	1360
UTP	Panamá	77	1434
UH	Cuba	79	2758

	TIEMPO POR MÉTODO WEB SCRAPING Perfiles / Publicaciones (minutos)			
Universidad	Copiar / Pegar	Local Browser	Local Software	Online
UFM	8 / 130	2 / 35	3 / 50	2 / 35
ESPOL	42 / 354	9 / 95	14 / 140	9 / 94
FURB	24 / 445	5 / 122	8 / 179	5 / 120

3. SELECCIÓN DE APLICACIÓN QUE SE CONSTRUIRÁ PARA EXPLOTAR EL GRAFO DE CONOCIMIENTO.

Hemos optado por realizar un buscador semántico y la aplicación que utilizaremos para la construcción del grafo de conocimiento trataremos de incorporar el Elasticsearch, ya que es un software libre en su versión de Apache Lucenen. En esta aplicación la búsqueda se realizará mediante un índice de documentos creado anteriormente en el que todo el contenido se almacena de forma preparada. Además, se puede construir con elasticsearch un servidor de búsqueda estable en poco tiempo.



Hemos analizado el tipo de buscador que queremos implementar concluyendo que nuestro dominio es muy extenso, y de mucha utilidad para muchas actividades dentro del campo de la educación, además que Elasticsearch es de fácil implementación, rápido y escalable.

Bibliografía

- <https://revistas.utp.ac.pa/index.php/id-tecnologico/article/view/1807/pdf>