

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 Publicaciones científicas, específicamente a los metadatos de las **tecnologías emergentes(nuevas tecnologías)**, 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 Semantic Scholar(otra fuente), google scholar) librería scholarly utilizando específicamente el API RESTful(revisar gran cantidad de datos) ya que es una programación basada en Python, que permite extraer información de registros individuales según lo necesario.

Para la extracción de las publicaciones científicas se utilizará los atributos del API para sacar información como: **Id, abstract, titulo, autor, año de publicación.**

Publicaciones de los cinco años 2015

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.

MARCO TEORICO

TECNOLOGÍAS EMERGENTES

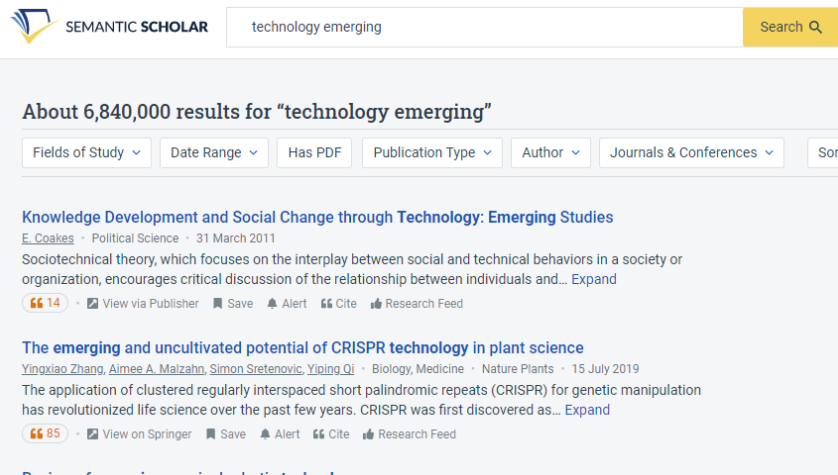
Hablamos de tecnología emergentes o también llamadas convergentes si se caracterizan por la totalidad de posibilidades y el impacto que provocan a corto, mediano y largo plazo concerniente a proceso, ya que en lo que se refiere a personas e instituciones aun no son bien comprendidas.

Estas tecnologías cumplen con el objetivo de traer mejoras infalibles, y provocar un desarrollo empresarial muy grande.

Podemos hablar de ejemplos concretos dentro de esta tecnología como sería el caso de las criptomonedas, con la herramienta blockchain que promete en el futuro un control mayoritario de las finanzas en el mundo.

SEMANTIC SCHOLAR

Podemos decir que es un buscador académico con Inteligencia Artificial que fue creado para llevar a cabo ingeniería e investigación que permitiera tener un alto impacto en el campo de la inteligencia artificial.



The screenshot shows the Semantic Scholar search interface. At the top, the Semantic Scholar logo is on the left, and a search bar contains the text 'technology emerging' with a search button. Below the search bar, a banner indicates 'About 6,840,000 results for "technology emerging"'. A row of filters is visible: 'Fields of Study', 'Date Range', 'Has PDF', 'Publication Type', 'Author', 'Journals & Conferences', and 'Sort'. Two search results are displayed:

- Knowledge Development and Social Change through Technology: Emerging Studies** by E. Coakes, Political Science, 31 March 2011. The abstract mentions sociotechnical theory. It has 14 citations and includes options to view via publisher, save, alert, cite, and follow the research feed.
- The emerging and uncultivated potential of CRISPR technology in plant science** by Yingxiao Zhang, Aimee A. Malzahn, Simon Sretenovic, Yiping Qi, Biology, Medicine, Nature Plants, 15 July 2019. The abstract discusses the application of CRISPR for genetic manipulation. It has 85 citations and includes similar options to the first result.

El API RESTful nos proporciona atributos que permiten la extracción de datos importantes que vamos a necesitar para nuestro buscador semántico como:

- **id** string: ID del paper de investigación
- **title** string: Título del paper.
- **paperAbstract** string: Se extrae el abstract que tenga el paper.
- **authors** list: Se genera los autores de los paper validando Id y título.
- **year** int: El año de publicación del paper.

Example Paper Record

```
{
  "id": "4cd223df721b722b1c40689caa52932a41fcc223",
  "title": "Knowledge-rich, computer-assisted composition of Chinese couplets",
  "paperAbstract": "Recent research effort in poem composition has focused on the use of automatic language generation...",
  "entities": [
  ],
  "fieldsOfStudy": [
    "Computer Science"
  ],
  "s2Url": "https://semanticscholar.org/paper/4cd223df721b722b1c40689caa52932a41fcc223",
  "pdfUrls": [
    "https://doi.org/10.1093/11c/fqu052"
  ],
  "s2PdfUrl": "",
  "authors": [
    {
      "name": "John Lee",
      "ids": [
        "3362353"
      ]
    },
    "...",
  ],
  "inCitations": [
    "c789e333fddb963883a0b5c96c648bf36b8cd242"
  ],
  "outCitations": [
    "abe213ed63c426a089bdf4329597137751dbb3a0",
    "...",
  ],
  "year": 2016,
  "venue": "DSH",
  "journalName": "DSH",
  "journalVolume": "31",
  "journalPages": "152-163",
  "sources": [
    "DBLP"
  ]
}
```

Bibliografía

- <https://revistas.utp.ac.pa/index.php/id-tecnologico/article/view/1807/pdf>
- [Semantic Scholar Open Research Corpus \(s2-public-api-prod.us-west-2.elasticbeanstalk.com\)](#)
-