

Tabla de Contenido

Índice de Tablas	v
Índice de Ilustraciones	vi
Introducción	1
1. Estado del Arte	3
1.1. Wikidata	3
1.2. Sistemas de información y datos científicos	3
1.3. Herramientas de <i>entity linking</i>	5
1.4. Herramientas de Geolocalización de Texto	6
1.5. Herramientas de Visualización Geográfica	6
1.6. Bases de datos NoSQL	7
1.7. <i>Topic modelling</i>	8
2. Arquitectura de la solución	10
2.1. Datos, geolocalización y temas de investigación	10
2.2. Esquema de base de datos	12
2.3. <i>Frontend</i> y experiencia de usuario	12
2.4. Resumen de la Arquitectura	13
3. Datos, <i>entity linking</i> y geolocalización	14
3.1. Datos DBLP	14
3.2. Algoritmo inicial para asignar afiliaciones	18
3.3. Uso de herramienta <i>entity linking</i> para identificar entidades	18
3.4. Geolocalización usando SPARQL y Wikidata	20
4. Indexación y consulta de datos usando Elasticsearch	22
4.1. Elasticsearch	22
4.2. Configuración, <i>mapping</i> y campos de información	24
4.3. Indexación de datos	24
4.4. Consultas	26
5. <i>Frontend</i> y experiencia de usuario	29
5.1. Django y visión general	29
5.2. Barra de navegación	29

5.3. Página principal y búsqueda simple	30
5.4. Visualización en mapa de la investigación	30
5.5. Artículos asociados a una afiliación	31
5.6. Búsqueda avanzada	32
5.7. Toda la investigación en el mundo	33
6. Evaluación	35
6.1. Algoritmo para asignar afiliaciones	35
6.2. Herramientas de <i>entity linking</i>	36
6.3. Pruebas de rendimiento del sistema	36
6.3.1. Pruebas sobre Elasticsearch	37
6.3.2. Pruebas sobre el sistema en general	38
6.4. Percepción de usuarios	40
Conclusión	42
Bibliografía	44