





Irakaslegoak aurkeztutako GRADU AMAIERAKO LANAREN proposamena	 Universidad del País Vasco Euskal Herriko Unibertsitatea
Propuesta de TRABAJO FIN DE GRADO por parte del profesorado	 BILBOKO INGENIARITZA ESKOLA ESCUELA DE INGENIERÍA DE BILBAO

GEHIENEZKO LUZERA: ORRI BAT (2 ALDEAK) / LONGITUD MÁXIMA 1 HOJA (2 CARAS)

DATU OROKORRAK / DATOS GENERALES	
Lanaren izenburua / Título del Trabajo: Grafo de conocimiento (KG) para relaciones transparentes en Euskadi	
Tutorea (gehenez 2) / Tutor-a (máx. 2): Mikel Egaña	
e-mail: mikel.egana@ehu.eus	
Hizkuntza / Idioma:	EUSKERA o CASTELLANO
Izena emateko aurrebaldintzak / Requisitos para apuntarse:	
Sistemas Web	
Programación orientada a objetos	
Bases de Datos	

DESKRIBAPEN ZEHASTUTA / DESCRIPCION DETALLADA	
<p>El Gráfo de Relaciones Transparentes de las Instituciones del País Vasco (BCITR) es un Grafo de Conocimiento (Knowledge Graph) para integrar información sobre entidades y personas que puedan tener un conflicto de interés en las instituciones vascas, con el fin de documentar y analizar dichas relaciones. BCITR se basa en la Ontología de Relaciones Transparentes (TRO) que ofrece un vocabulario para publicar información sobre relaciones institucionales que deberían ser más transparentes y explícitas.</p> <p>El objetivo general de este proyecto es crear una infraestructura basada en Python para facilitar a los programadores y expertos en datos el procesamiento de nuevos conjuntos de datos, añadiéndolos al grafo siguiendo unas buenas prácticas ya establecidas.</p> <p>(TRO)https://github.com/mikel-egana-aranguren/Transparent-Relations-Ontology (BCITR)https://github.com/mikel-egana-aranguren/BasqueCountryInstitutionsTransparentRelationsGraph</p>	
GrAL-AREN HELBURUAK / OBJETIVOS DEL TFG	
<p>Obligatorios:</p> <ul style="list-style-type: none"> •Recopilar datasets de interés y documentar sus metadatos en RDF (DCAT, VOID,...). Implementar pipelines de transformación de datos a RDF desde diversas fuentes, mediante CWL y Morph-kgc. •Almacenar los datos en GraphDB. •Desarrollar una visualización web para los datos. •Crear e implementar una orquestación Docker con GraphDB, la visualización web, un servidor de Linked Data y un endpoint SPARQL. •Documentar buenas prácticas para metadatos, patrones URI, etc. •Crear módulos Python para tareas comunes: normalización de nombres, generación de URIs, reconciliación de entidades, descubrimiento de enlaces, validación de datos. <p>Opcionales:</p> <ul style="list-style-type: none"> •Realizar los cambios necesarios en TRO para acomodar los nuevos datos, incluída la validación a través de pruebas y documentación de ROBOT. 	

<p>Irakaslegoak aurkeztutako GRADU AMAIERAKO LANAREN proposamena</p>	 
<p>Propuesta de TRABAJO FIN DE GRADO por parte del profesorado</p>	

- Colaborar con otros desarrolladores.

GrAL-AREN EKARPENAK / CONTRIBUCIONES DEL TFG

- Descubrimiento de nuevos datasets que podrían apuntar a conflictos de interés.
- Estandarización de un vocabulario para representar dichos conflictos de interés.
- Creación de una infraestructura en Python/CWL sostenible para procesamiento de datos sobre conflictos de interés.

ERABILIKO DIREN TRESNAK / HERRAMIENTAS A USAR

- Lenguajes para datos: RDF, OWL, SHACL, SPARQL.
- Ontologías: TRO, DCAT, VoID, PROV-O, Schema, etc.
- Programación en Python:
 - RDFLib (<https://rdflib.dev/>).
 - Morph-KGC (<https://github.com/morph-kgc/morph-kgc>).
- Common Workflow Language (<https://www.commonwl.org/>).
- GraphDB (<https://www.ontotext.com/products/graphdb/>).
- Trifid (<https://github.com/zazuko/trifid>).
- YASGUI (<https://github.com/TriplyDB/Yasgui>).
- D3.js (<https://d3js.org/>).
- D3-SPARQL (<https://github.com/zazuko/d3-sparql>).
- Docker (<https://www.docker.com/>).
- Git (GitHub, GitHub actions).