

Carlos Utrilla Guerrero

☎ +34 629 95 44 70 • ✉ c.utrilla.guerrero@gmail.com • 🌐 carlosug.github.io

Beside the projects listed in the CV, during my career, I have also participated in the following projects:

Main Research projects contribution

Distributed Knowledge Graphs for food (KGFOOD): The project started in *Aug 2022*, and aims at building a knowledge graphs on Food Composition Open Datasets and food price sources for identifying sustainable food choices using Knowledge Graphs Embeddings (KGE) as part of EU COST Action on Distributed Knowledge Graphs. **ROLE:** Daily coordinator of the project, leading data conversion to RDF, generation of standard APIs to query the converted datasets and usecase analysis (e.g. link prediction, completing KG with embedding techniques or network analysis) in '*Food and Health Domain*'. It is expected to submit a conference paper at SWAT4HCLS -Semantic Web and FAIR data in agriculture, food and nutrition domain. <https://github.com/MaastrichtU-IDS/kg-food>

- Semantic web: RDF, SPARQL, Ontologies, RDFLib, Protege
- Languages: Python and Bash
- Research line: Ontology engineering, Knowledge publication and Knowledge graph completion

FAIR Data for traditional games (PLAYFAIR): Since *July 2021*, I collaborated with the Digital Ludeme Project (DLP) team through a joint research initiative called the PLAYFAIR project. This project focused on implementing semantic web in digital humanities, had a *budget of 15,000 euros*, and ended in *March 2022*. **ROLE:** I took full initiative of the project, from its original formulation, development of the proposal and its execution. Presented as demo at the Conference Annual DARIAH 2022: Storytelling. <https://github.com/MaastrichtU-IDS/play-fair>

- Semantic web: RDF, SPARQL, Ontologies, RDFLib
- Languages: Python and Bash
- Tools: TriplyDB based on CLARIAH infrastructure
- Research line: Semantic interoperability, Knowledge publication and Network science

LD Wizard for Life and Social sciences (LDWizard): This small contribution built upon the *PLAYFAIR* project (above) to technically evaluate CLARIAH tools. As a result, we proposed an alternative and built an easy-to-use software to convert tabular data into Linked Data. **ROLE:** I co-developed the tool and inserted OWL ontologies that are relevant to digital humanities domain. <https://humanities.wizard.semanticscience.org/>

- Semantic web: RDF, SPARQL, Ontologies, RDFLib
- Languages: Python, TypeScript and Bash
- Web dev: HTML/CSS and React
- Research line: Knowledge capture, mappings and metadata editor tools

DSRI Research projects contribution

Medical Informatics:

- Distributed analytics and privacy preserving techniques in social-emotions and biomedical domain (ODISSEI): With colleagues from Inspectorate Directorate and Postdoc. Dr. Chang, I helped to design a library to generate synthetic tabular or RDF data using Conditional Generative Adversary Networks (GANs) (Python). <https://odissei-data.nl/en/privacy-preserving-techniques/>
- Visual Interface for Data Analytics on neonatal sepsis in preterm infants (VIDA) – I provided inputs on the semantic solution framework for a grant proposal E JP Rd-European Join Programme 2022 on Rare Diseases with Dr. Jose Luis Garcia Gimenez from CIBERer and Prof. Dr. Villamor Zambrano Eduardo from (MUMC+), which was focused on reconciling data provenance, predictive performance on the sepsis prediction and classification tasks.
- Network science for Mental Disorder (2022) – Estimating Time-Varying Models in High-Dimensional EMA Data: Co-supervision of bachelor student Julian Adam that focused on data mapping variables and data analysis using Generalized Additive Modeling (GAM) framework (Python).
- Netherlands Consortium of Dementia Cohorts (NCDC) 2021 – I worked on data harmonization WP Activities to implement FAIR data management for the mental health research. *Website NCDC cohort*

- BReIN WP5 - 2020: FAIR data infrastructure – Developed a working plan to create FAIR data and metadata, collect and storage patient data, classification of disease with ML, investigate molecular determinants of negative environment factors in AD Patients. I contributed in the system design and implementation of semantic metadata and data models, formalized using OWL and SHACL. www.maastrichtuniversity.nl/brein

○

Semantics:

- KG for Open Government Data (ESGREEN) – I helped to generate a RDF knowledge graph for green spaces infrastructure data from cities (Python, RML Mappings). <https://github.com/carlosug/esgreen-kg>. A master student was going to further explore the feasibility of using SIO Ontology towards interoperability in a research project started in Oct 2022.
- IDS Project Dashboard – I co-developed a web service to insert projects information from git repositories, alongside with generating JSON file using DOAP schema (Python, React/TypeScript). <https://github.com/MaastrichtU-IDS/projects>
- INDEX – I defined the user requirements for the system developed by Vincent Emonet with a web service to index standard knowledge representations, such as OWL ontologies and SHACL shapes generating metadata for active SPARQL endpoints (Python, React/TypeScript). <http://index.semanticscience.org/>
- A master student research project about Mercedes Benz Claims Automatic Recommendation system powered by semantic technologies (CAR) – (Python, React/TypeScript). <http://index.semanticscience.org/>

Open Science, FAIR and community standards:

- FAIR Enough – I helped define the metrics for FAIR evaluation tool of resources online (Python, React/TypeScript). <https://github.com/MaastrichtU-IDS/fair-enough>
- European Open Science Cloud-Life (EOSC-Life) - I reported on the user requirements and personas for the 'FAIRAssist' tool that was designed for EOSC-Life actors and profile educational resources (e.g. Generate an ontology for digital skills and competences for EOSC Personas). Check slides
- FAIRness evaluation reports for the Netherlands Consortium of Dementia Cohorts (NCDC) - I utilized FAIR Maturity Evaluation services for evaluating different online resources within NCDC and suggested future directions to ensure resources are compliant with FAIR principles.

Industry and user/department support:

- Data Sharing Coalition (DSC) – DSC is a large interdisciplinary research project involving several industry partners. The goal is to promote data sharing practices that are inputs for the development of generic data sharing standards. ROLE: I was particularly involved in building a knowledge base from texts about human trafficking data partnered with Sustainable Rescue NGO and Rosaka. Check an interview in DSC social media about expected contribution to DSC.
- Data Science Research Infrastructure (DSRI) – I provided technical support to researchers on the HPC cluster, and maintained project website and documentation. I started to learn Kubernetes for cluster management system and containerization. <https://dsri.maastrichtuniversity.nl/>
- Analysis Ready Data (KLEOS.space) – I resolved issues on data fusion and conducted statistical analysis of different countries with different geolocating datasets using geo standards and R for the European Agency Space
- Data governance and Financial Knowledge model (KNEIP Communications) – I helped resolve data collection and data integration issues in different projects related to EU regulations using ORACLE products. I played with FuzzyWuzzy python package (<https://pypi.org/project/fuzzywuzzy/>) to create a tool for financial term similarities and substitutions.