

Carlos Utrilla Guerrero

☎ +34 629954470 • 🌐 <https://orcid.org/0000-0002-9994-1462>
<https://carlosug.github.io/>

Education

University of Manchester, UK <i>M.Sc. in Social Research Methods and Statistics (CMIST), Dissertation: LINK – Assignments: LINK</i>	2014
University of Zaragoza, ES <i>BA in Business and Administration, Erasmus exchange program in Odense, DK (6 Months, 2012)</i>	2011

Research Interests

Knowledge graphs, Network science, Semantic web, Sustainable Research Software and Ontology Engineering

Career Profile

Research software engineer, Data scientists, Researcher

Computer skills

Languages: Python, R, TypeScript (React, ExpressJS, NextJS), LaTeX, SQL

Web: SPARQL, RDF, ontologies (OWL, SKOS, SIO), OpenAPI, JSON, XML, HTML&CSS, Material UI, D3.js

Tools: Bash, Git, Jupyter, Rstudio, Linux, Neo4J, Protégé, Docker

Languages

Spanish (Mother tongue), English (Fluent), French (Intermediate), Dutch (Basic).

Work Experience

Research Data & Software Services, TU Delft University, Delft, NL <i>Teacher Open Science</i> Participate in TU Delft DCC, Library and 4TU educational & research activities	<i>Since Jan 2023</i>
Institute of Data Science, Maastricht University, Maastricht, NL <i>Researcher data scientist</i> Provided support in teaching and R&D obligations	2020 – 2023
Kleos Space S.A, Luxembourg, LU <i>Data scientist</i> Provided support in data management and statistical analysis for the ESA projects	2018 – 2019
Kneip Communications, Luxembourg, LU <i>Data manager</i> Participated in data governance for the Financial knowledge base	2015 – 2018
European Investment Bank (EIB), Luxembourg, LU <i>Graduate research assistant (6 months)</i> International scholarship in the IT department focused in data modeling, definition and implementation	2014
Imagniarium Group S.A, Istanbul, TR <i>Trainee data analyst (8 months)</i> International scholarship in a retail company to develop and implement the Forecasting Model	2013
Association of Young Entrepreneurs of Zaragoza, Zaragoza, ES <i>Bachelor trainee marketing (8 months)</i> Scholarship in a Spanish public administration to monitor the operational cash-flow and write the annual budgeting	2010
Ibercaja saving Bank, Zaragoza, ES <i>Bachelor trainee finance (6 months)</i> Scholarships in a Spanish bank to study the probability of default using logistic regression	2010

Volunteer Experience

Turkiye I.S.I. Tme Engelliler Spor Federasyonu, Ankara, TR

2010

European Volunteer Service (7 months)

Collaborated in logistic and administrative activities

Relevant research projects contribution

Scientific software Knowledge graph (SOFTKG): A project initiated recently in TU Delft Library to explore how semantic web technologies enhance the interoperability of 4TUResearchRepository and wide range of applications and user experience. GOAL: Understand how to describe software dependencies and its relationship using SW (e.g. *LinkML schemas – using author schemas in YAML that describe the structure of your data and validating data in a variety of formats -JSON, RDF, TSV-*) to computationally reproduce software environment of objects deposited in 4TUResearchData repository

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 Humanities 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
- Research line: Cross-domain interoperability, NLP and data sharing infrastructure

Community Data Driven Insights (CDDI): This project consisted of a diverse research project ("Showcases") from Maastricht University, exploring the technical, ethical and social complexity of implementing FAIR. ROLE: I developed ETL processes and workflows for data analysis, and data visualisation, as well as generate conceptual data models and standards for some showcases 'FAIRifing showcases', which focuses on the interoperability by annotating the data with ontologies and standard vocabularies. This work has been accepted for the Fair Digital Object (FDO) International Conference in Oct 2022. Conf. article: <https://doi.org/10.3897/rio.8.e96082>

- Semantic web: RDF and Ontologies
- Languages: Python and R
- Database: SPARQL, MongoDB, SQL
- Research line: Cross-domain interoperability, knowledge validation and Open Science

Other projects contribution

Semantics:

- Visual Interface for Data Analytics on neonatal sepsis in preterm infants (VIDA) – I provided inputs on the semantic technology framework for a grant proposal EJP RD - European Joint Program on Rare Diseases call, focused on reconciling data provenance, data sharing and predictive performance on the sepsis prediction and classification tasks

- 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/>

Open Science, FAIR and community standards:

- 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.
- 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

Quantitative methods and statistics:

- Developing distributed analytics and privacy preserving techniques in social-emotions domain (ODISSEI) – With colleagues from Inspectorate Directorate, 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/>
- 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)

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

Teaching Experience

2023: Graduate school: Research data and software management: <https://doi.org/10.5281/zenodo.6325919>

2020; 2021; 2022: Graduate course: Applied Research II: <https://maastrichtu-ids.github.io/AppliedRR/>

2020; 2021, 2022: Bachelor course: Quantitative Research Methods. <https://maastrichtu-ids.github.io/global-studies/>

Supervision

2022: Co-supervision of one master student in estimating Time-Varying Models in High-Dimensional EMA Data

List of Abstract Publications

•M. Paneque, C. Utrilla Guerrero, J. Garcia Nieto, Maria del Mar Roldán-García, Remzi Celebi (2023) "Towards Nutrition Information and Food Prices linked in a Knowledge Graph"

•C. Utrilla Guerrero, Maria Vivas Romero, Mark Dolman and Michel Dumontier (2022). "FAIR begins at home: Implementing FAIR via the Community Data Driven Insights". First International Conference on FAIR Digital Objects 2022. Leiden, NL. <https://arxiv.org/abs/2303.07429>

•**C. Utrilla Guerrero** and **V. Emonet** (2022). “PLAYFAIR Project: Challenges and solutions towards building and visualising FAIR data for traditional games”. *Storytelling - DARIAH Annual Event. Athens, GR*. <https://doi.org/10.5281/zenodo.6588450>

Selected Awards and Scholarships

2021: Research Fellowship grant in the national infrastructure CLARIAH Research Fellows 2021. *Budget 15K euros*

2013: Leonardo Da Vinci Global Scholarship from Spain's Ministry of Education

Events (co)Organized

2020 – present: Co-founder of FAIR Coffee Lectures with Open Science Community Maastricht

2022: In-person Data Science Research Infrastructure (DSRI) Community Event

Workshops, Presentations and Conferences

Workshops:

- o 2021: Workshop on FAIR Essentials. <https://github.com/MaastrichtU-IDS/fair-workshop>
- o 2021: Workshop on semantic web - DOAP <https://github.com/MaastrichtU-IDS/doap-workshop>
- o 2021: Coding for researchers in Python <https://github.com/MaastrichtU-IDS/coding4all>
- o 2020: R Carpentry Workshop <https://esciencecenter-digital-skills.github.io/2020-10-19-Data-Carpentry-with-R/>

Invited talks and conferences:

- o 2022: Storytelling - DARIAH Annual Event 2022, Athens, (GR) – <https://github.com/MaastrichtU-IDS/coding4all>
- o 2022: Metadata 4 DKE 2022, Lyon, (FR) – <https://cost-dkg.github.io/metadata4dkg2022/>
- o 2022: FAIR Coffee Lectures, Maastricht, (NL) – <https://www.openscience-maastricht.nl/events/fair-coffee-lecture-25-may-2022/>
- o 2020: Online M-BIC Open Science Symposium, Maastricht, (NL) – <https://www.maastrichtuniversity.nl/events/online-m-bic-open-science-symposium>

Interests

Sports: Tennis, Padel, Swimming

Others: Open source enthusiast, Traveling, History, Neonatology and Cooking