

Vincent Emonet

Knowledge graph engineer

Experiences

- Since July 2018 **Data science developer,**
Institute of Data Science, Maastricht University, Netherlands.
- Developed an automated standard knowledge indexing system (Python, SPARQL, React/TypeScript, GitHub Actions CI/CD): github.com/vemonet/shapes-of-you
 - Index standard knowledge representations, such as OWL ontologies and SHACL shapes, from git repositories hosted on GitHub, GitLab and Gitee.
 - Index active SPARQL endpoints, and generate qualitative and quantitative metadata about their content (classes, relations).
 - Provide user-friendly and well documented access points for the indexed knowledge and data: web UI, SPARQL, OpenAPI
 - Worked on the development of the Data Science Research Infrastructure, an OpenShift cluster for academic researchers to deploy Data Science services and jobs with 15 CPU nodes and 1 GPU node. Managed the cluster, provided support for researchers, led the effort to deploy applications for data science, such as JupyterHub, RStudio, Tensorflow/PyTorch on GPU via VSCode server or JupyterLab (Kubernetes, Docker)
 - Worked on the development of the OpenPredict API, a service to annotate and manage drug-disease predictions models using RDF metadata, and expose the predictions as part of the NIH NCATS Biomedical Data Translator network (Python)
 - Helped on the development of the Nanobench, a web UI to quickly write and publish structured claims in RDF to the Nanopublications network (Java)
 - Worked on the development of the Knowledge Collaboratory API to expose claims from the Nanopublication network to the NIH NCATS Biomedical Data Translator standards (Python).
 - Developed and published multiple GitHub Actions for the RDF ecosystem and data science: generate RDF data using RML mappings, execute SPARQL queries and upload RDF files using RDF4J, setup a local Apache Spark cluster in GitHub Actions runners
 - Developed a small system to create, index, and expose standard DOAP project descriptions for projects developed by the Institute of Data Science (Python, React/TypeScript, GitHub Actions CI/CD): github.com/MaastrichtU-IDS/projects
 - Developed Data2Services, a command line interface and documentation to help mapping structured data to RDF knowledge graphs using RML (Python, Docker). Published *Data2Services: enabling automated conversion of data to services* at SWAT4LS
 - Deployed collaborative, user-friendly and cost efficient documentation websites related to data science based on markdown and git (React, GitHub Pages), such as maastrichtu-ids.github.io/best-practices
 - Teaching assistant for the Bachelor degree course *Relational database* at the Department of Knowledge and Engineering (SQL)
 - Teaching assistant for the Master degree course *Building and mining knowledge graph* (RDF, SPARQL, ontologies, shapes, mapping languages)
 - Participated to multiple research conferences and events, such as the 2019 International Semantic Web Research School focusing on knowledge graph evolution and preservation.
 - Reviewed multiple publications for research journals, mostly about knowledge management.
- October 2017 to June 2018 **Conversational design engineer,**
Match Group, France.
- Involved in the development of a Natural Language Understanding (NLU) tool to build chatbots
 - Designed conversational patterns using API.ai framework
 - Analyzed conversations flow to improve user engagement using tools like Google Analytics
 - Managed the multilingual aspect and translations in 6 languages of the NLU tool developed
 - Our chatbot won the French 2018 Digital Innovation Award
- January 2015 to July 2017 **Computer science design engineer - Ontologies & Semantic Web,**
CNRS - Laboratory of Informatics, Robotics and Microelectronics of Montpellier, France.
- Administration and new features development on bioportal.lirmm.fr and agroportal.lirmm.fr platforms, used for ontology storage, exploration and text annotation (Ruby on Rails & Java): github.com/siffrproject
 - Kept compatibility with main version of the NCBO BioPortal (tests)
 - Packaged, integrated and documented Natural Language Processing programs, available at advanse.lirmm.fr (Java, AngularJS, Linux administration): gite.lirmm.fr/advanse
 - Enhanced and packaged the YAM++ ontologies matching tool (Java): gite.lirmm.fr/opendata
 - Developed a web interface for ontologies mappings validation (JEE) : yamplusplus.lirmm.fr

- March 2014 to September 2014 & June 2013 to September 2013
- Bioinformatics design engineer - Biomedical data processing,**
Université de Laval, Québec, Canada.
- Development of web services to access biomedical data in RDF format for the *bio2rdf.org* project: github.com/bio2rdf/bio2rdf-rest
 - Developed components for the Talend ETL for the Semantic Web (Java): github.com/vemonet/talend4sw
 - Converting biomedical databases to RDF format (PubMed, neXtProt) using Talend ETL and Java
 - Set up a Hortonworks Hadoop platform for Big Data
- June 2012 to August 2012
- Intern in bioinformatics in an ecology research laboratory,**
University of California, Berkeley, USA.
- Developed DNA sequence analysis programs using Python 2.7
 - Set up an Amazon Web Service machine image based on Linux

Education

- 2012 – 2014 **Master degree, Physiology, biotechnology and informatics,** *University of Poitiers, France*, Biotechnology.
- 2011 – 2012 **Bachelor's degree, Bioinformatics,** *University of Poitiers, France*, Biotechnology specialization.
- 2009 – 2011 **Technical College Diploma, Bioengineering,** *Toulon, France*, Biology and biochemistry analysis.

Skills

- Programming Python, Java, JavaScript, Bash (and ZSH)
- Knowledge standards SPARQL, RDF, ontologies (OWL, SKOS, OBO), Open API (Swagger), SQL
- Current web development React and TypeScript, using expo to create Progressive Web Apps with Material UI, basic experience using Google Analytics and ELK stack for monitoring
- Previous web development TypeScript Angular, AngularJS, Ruby on Rail, Java Tomcat, Bootstrap HTML/CSS
- Sysadmin OpenShift/Kubernetes cluster management, creation of Docker images for multiple open source projects, administration of Linux servers in production (Ubuntu, CentOS), basic knowledge in DNS and reverse proxy routing.

Languages

- English TOEIC Score 2014: 935. 2 years in the Netherlands, 6 months in the USA. Active participation in discussion, reading and writing documents
- Spanish Basic notions for communication

Interests

- Open Source In my work and free time, I enjoy discovering new projects, and innovative approaches to solve problems. When relevant, I also engage in discussions, improve the documentation, or the implementation. Visit my GitHub profile for more details about my contributions: github.com/vemonet
- Outdoor Bike touring, mountain biking, hiking, Ultimate frisbee

Publications

- Google Scholar: <https://scholar.google.fr/citations?user=G59f3woAAAAJ>
- Platforms
- HAL: https://hal.archives-ouvertes.fr/search/index/q/*/authFullName_s/Vincent+Emonet
 - dblp: <http://dblp.uni-trier.de/pers/hd/e/Emonet:Vincent>

References

Maastricht University
Michel Dumontier,
Professor of Data Science
(michel.dumontier@maastrichtuniversity.nl)

LIRMM lab
Computer science researchers
◦ Clément Jonquet (jonquet@lirmm.fr)
◦ Zohra Bellahsene (bella@lirmm.fr)
◦ Jérôme Azé (jerome.aze@lirmm.fr)

Bio2RDF project
François Belleau,
bio2rdf.org project co-founder
(francois.belleau@gmail.com)