Thomas Pellissier Tanon

Software Engineer

15B chemin de Charavel
38200 Vienne, France
✉ *thomas@pellissier-tanon.fr*
🌐 *thomas.pellissier-tanon.fr*
in *tpt93*
⌗ *Tpt*

Main Experiences

2016 – now **Vice President, Research and Innovation**, *Lexistems*
I am leading a small team of three people. We have built three products from scratch: SensibleTV, a voice TV assistant. SensibleSpeech, a multilingual speech-to-text system. SensibleSummary, an automated generative summarization system. The three systems are written in Python and Rust and make intensive use of modern machine learning (transformer neural networks using PyTorch, etc.).

2015 **Engineering internship**, *Google*
Creation of an automated importation tool from Freebase to Wikidata, two leading knowledge bases.

Education

2017 – 2020 **Ph.D. in Computer Science**, *Télécom Paris*
Ph.D. on knowledge graphs and data management. Lead to multiple publications in international research conferences (ISWC 2018, The Web Conference 2019, ESWC 2020 and 2021...). These publications got more than 400 citations according to Google Scholar.

2013 – 2016 **Computer science *licence* and *master***, *École normale supérieure de Lyon*
Normalien élève. I did my third year at Télécom ParisTech to focus on database and data management (*Master Data & Knowledge*). I got ranked 1st during this year. Internships at INRIA, Google, ENS Cachan, and Max-Planck-Institut für Informatik.

Activities

Wikimedia Volunteer
I maintain ProofreadPage, a MediaWiki extension used by Wikisource, a Wikipedia sister project. ProofreadPage is a plugin of MediaWiki written in PHP and JS to provide a book proofreading workflow. Tasks included code review of changes submitted by other volunteers and Wikimedia employees, refactoring, integration of modern best practices like unit tests, and deployment on the infrastructure of a top 10 website.

Oxigraph Maintainer
I created Oxigraph, a graph database written in Rust and based on the RocksDB key-value store. It focuses on the single-node use case and fully implements the SPARQL query language. It is still a work in progress but now provides competitive performances on simple OLTP workloads.

Languages

French Mother tongue
English Fluent

Skills

Programming Strong knowledge of Rust and Python. Good in PHP, JavaScript, Java, C, C++ and Lua.

Software MediaWiki, Git, Gerrit, GitLab, SQL, HTML, RDF and SPARQL.