

# Thibault Sellam

Tolstraat 50-3  
Amsterdam 1073SC, the Netherlands  
+31 6 52 73 43 05  
[thibault.sellam@gmail.com](mailto:thibault.sellam@gmail.com)  
<http://homepages.cwi.nl/~sellam>

## RESEARCH INTERESTS

My mission is to design statistical models and software to help users explore and understand their data, regardless of their technical skills. Relevant fields include ***data exploration***, ***query recommendation***, ***unsupervised learning***, ***subspace search*** and ***visualization***.

## PROFESSIONAL EXPERIENCE

- Oct. 2011 - **PhD Candidate, Dutch National Research Center for Mathematics and Computer Science (CWI);** Amsterdam, the Netherlands  
Now Database Architectures group - advisor: Prof. Martin Kersten
- *Query Advisors For Data Exploration*: I designed four algorithms to help users write queries with machine learning. One is currently being developed by a commercial spin-off, MonetDB Solutions. I co-lead the engineering team.
  - *Traffic Data Analysis*: I co-developed systems to query, model and simulate GPS data, in collaboration with TomTom.
  - I am involved in the development of MonetDB, an Open Source DBMS.
  - I write and review articles for data management venues (e.g., SIGMOD, VLDB, CIKM, TKDE)
- May 2014 - **Research Intern, Microsoft Research;** Mountain View, California  
Aug 2014 Bing Social group - advisor: Dr. Omar Alonso.
- I developed a library to cluster graphs in a large scale, distributed, environment. I used this library to enhance expert search on social media.
  - I co-developed a pipeline to extract quantitative information from Twitter. The internship led to one patent and several publications.
- Sept. 2010 - **Software Engineer, JP Morgan;** London, UK  
Aug. 2011 I worked for the Credit Hybrids Trading division (Front-Office).
- I developed and supported the risk, pricing and analytics pipelines.
  - Ad-hoc projects: migrations, bulk trading operations, reporting.
- June 2008 - **Intern Software Engineer, JP Morgan;** London, UK  
Sept. 2008 I interned for the Structured Funds department. I devised a market data retrieval application.

## EDUCATION

- Oct. 2016 **PhD in Computer Science, University of Amsterdam**  
(expected) Amsterdam, the Netherlands
- Sept. 2010 **Master's - Artificial Intelligence, Université Paris Dauphine;** Paris, France  
Master's thesis at the Uppsala Database Laboratory (Uppsala, Sweden).
- Aug. 2008 **Bachelor's - Computer Science, Université Paris Dauphine;** Paris, France
- Aug. 2007 **Bachelor's - Economics, Université Paris Dauphine;** Paris, France

## PUBLICATIONS

1. Thibault Sellam, Emmanuel Müller, and Martin Kersten. Semi-Automated Exploration of Data Warehouses. In ACM Conference on Information and Knowledge Management (CIKM), Knowledge Management Track, 2015.
2. Thibault Sellam, Omar Alonso. Raimond: Quantitative Data Extraction from Twitter to Describe Events. International Conference on Web Engineering (ICWE), 2015.
3. Thibault Sellam. Fast Cartography for Data Explorers, PhD workshop of Very Large Data Bases (VLDB) 2013.
4. Thibault Sellam, Martin Kersten. Meet Charles, Big Data Query Advisor. The biennial Conference on Innovative Data Systems Research (CIDR), 2013.
5. Holger Pirk, Thibault Sellam, Stefan Manegold, Martin L. Kersten. X-device query processing by bitwise distribution, International Workshop on Data Management on New Hardware (DaMoN), SIGMOD, 2012.
6. Konrad Bösch, Thibault Sellam, Holger Pirk, René Beier, Peter Mieth, Stefan Manegold. Scalable Generation of Synthetic GPS Traces with Real-Life Data Characteristics. TPC Technology Conference, (TPCTC), VLDB, 2012.

## PUBLICATIONS UNDER REVIEW

1. Thibault Sellam and Martin Kersten. Cluster-Driven Navigation of the Query Space, submitted to IEEE TKDE in Feb 2015.
2. Thibault Sellam, Martin Hentschel, Vasilis Kandyas, Omar Alonso. e#: Sharper Expertise Detection from Microblogs, submitted to EDBT 2016.
3. One patent pending submission since January 2015, with Microsoft.

## SKILLS

|           |   |
|-----------|---|
| Languages | <i>French</i> : native; <i>English</i> : fluent written and spoken (TOEFL: 106/120, IELTS: 8.0/9); <i>Dutch</i> : intermediate (European level B2); <i>German</i> : basic                           |
| IT        | OS : Linux, MacOS<br>Analytics: R (ggplot/dplyr), Pig, D3, Tableau<br>Parallelism: MapReduce, OpenCL<br>Development : Java, C, Python, JavaScript<br>Databases : MonetDB, PostgreSQL, MySQL, Sybase |

## PERSONAL

As a hobby, I practice several instruments and I produce music at a semi-professional level. During the period Jan.-Sept. 2015, I toured full-time with the French music band FAUVE. We sold more than 150,000 albums (three gold records), achieved top position in several national charts and received wide media coverage.

## REFERENCES

Prof. Martin Kersten (University of Amsterdam, CWI) - [martin.kersten@cwi.nl](mailto:martin.kersten@cwi.nl)  
Prof. Stefan Manegold (Leiden University, CWI) - [stefan.manegold@cwi.nl](mailto:stefan.manegold@cwi.nl)  
Dr. Omar Alonso (Microsoft) - [omalonso@microsoft.com](mailto:omalonso@microsoft.com)