Email: edouard.fouche@kit.edu
Blog: edouardfouche.com

Find me on:
LinkedIn
Google Scholar
GitHub
Instagram

EXPERTISE

Data Science / Data Mining
Machine Learning
Algorithms
Visualization

PROGRAMMING

Python (~6 years)
Scala (~4 years)
R (~3 years)
Also: Java, SQL, Bash, C, Web…

SOFTWARE

Git, Linux, Database Systems, LateX, Adobe & MS Office Suite

SOFT SKILLS

Scientific Writing
Leadership
Agile Methodology
Cultural Awareness
Self-improvement

LANGUAGES

French (Native)
English (Fluent, C2)
German (Fluent, C2)
Spanish (Basic, A2)

PERSONAL INTERESTS

Photography (Analog/Digital), Hiking, Travelling, Cooking, Blogging, Roller Derby (Coach and Board at Roller Derby Karlsruhe)

Dr.-Ing. Edouard Fouché

Postdoctoral Researcher in Data Science

Experience

2020-now **Postdoc in Data Science, Karlsruhe Institute of Technology (KIT).**
- Conducting independent research, collaborations and student theses.
- Co-supervision of 4+ doctoral students, Data Science lecturer.
- Organizer for chair-wide meetings, coaching and mentoring.

2016-2020 **Researcher & Ph.D. Candidate in Data Science, KIT.**
- Published 8 peer-reviewed research papers and a doctoral thesis.
- Supervised 10+ student theses (Master/Bachelor level).
- Worked with 2 research labs (international) and 5 industry partners.

2014-2016 **Working student, IBM R&D GmbH, Böblingen.**
- Created ibmdbpy, an open-source interface between Python and DB2, which then became the topic of my Master's thesis at KIT.
- Co-supervision of two fellow interns. Python coaching.
- Talk at LWA 2015 conference and PyData Berlin 2016 meet-up.

Education

2020-2021 **MBA, Quantic School of Business and Technology.** (fka "Smartly")
Online, 7% admission rate. Topics: Accounting, Market & Economies, Data & Decisions, Leading Organizations, Marketing & Pricing, Finance, Supply Chain & Operations, Strategy, Entrepreneurship.

2016-2020 **Ph.D. (Dr.-Ing.) in Data Science, KIT.**
- Thesis: Estimating Dependency, Monitoring and Knowledge Discovery in High-Dimensional Data Streams. (with Prof. Klemens Böhm)
- Summa cum laude. Helmholtz Doctoral Prize. German GPA: 1.0

2015-2016 **M.Sc. Computer Science, KIT.**
- Thesis: Fast In-Database Feature Selection. (with Prof. Emmanuel Müller)
- German GPA: 1.4 (Very Good)

2010-2015 **Diplôme d'Ingénieur (Dipl.-Ing.), ESIEE Paris, France.**
- Major: Computer Science. Final internship at IBM R&D (6 months).
- "Félicitations du Jury" and "Parcours d'Excellence" award.

Awards

2021 **Helmholtz Doctoral Prize 2020.**
The prize recognizes outstanding achievements during the doctoral phase. The Helmholtz Association is the largest german scientific organisation.

2019 **SSDBM Best Paper Award.** Paper: Monte Carlo Dependency Estimation Scientific and Statistical Database Management (Core A conference).

2017 **Software Campus Member:** 100K€ research grant, training, mentoring. Co-founded by BMBF and TRUMPF GmbH + Co. KG.

2015 **Parcours d'Excellence ESIEE Paris.** Top 1% of the cohort.

2014 **Data Mining Cup, Prudsys AG, 2nd place:** €1000 cash price.

Research Interests

- Unsupervised Learning in High-Dimensional Data Streams
- Correlation Analysis, Dependency Estimation
- Bandit Algorithms, Sequential Decision Making
- Outlier Detection, Clustering

List of Publications (Peer-Reviewed)

10. Edouard Fouché, Florian Kalinke and Klemens Böhm. 'Efficient Subspace Search in Data Streams'. In: *Information Systems* 97 (2021), p. 101705. ISSN: 0306-4379. DOI: 10.1016/j.is.2020.101705
9. Edouard Fouché. 'Estimating Dependency, Monitoring and Knowledge Discovery in High-Dimensional Data Streams'. **Summa cum laude**. PhD thesis. Karlsruher Institut für Technologie (KIT), July 2020. DOI: 10.5445/IR/1000127232
8. Edouard Fouché, Yu Meng, Fang Guo, Honglei Zhuang, Klemens Böhm and Jiawei Han. 'Mining Text Outliers in Document Directories'. In: *ICDM*. **Acceptance Rate: 9.8%**. IEEE, 2020, pp. 152–161. DOI: 10.1109/ICDM50108.2020.00024
7. Edouard Fouché, Alan Mazankiewicz, Florian Kalinke and Klemens Böhm. 'A Framework for Dependency Estimation in Heterogeneous Data Streams'. In: *Distributed and Parallel Databases* (2020). DOI: 10.1007/s10619-020-07295-x
6. Avipsa Roy, Edouard Fouché, Rafael Rodriguez Morales and Gregor Möhler. 'In-Database Geospatial Analytics using Python'. In: *ARIC@SIGSPATIAL*. ACM, 2019, pp. 17–24. DOI: 10.1145/3356395.3365598
5. Hasan Ümitcan Yilmaz, Edouard Fouché, Thomas Dengiz, Lucas Krauß, Dogan Keles and Wolf Fichtner. 'Reducing energy time series for energy system models via self-organizing maps'. In: *Information technology* 61.2-3 (2019). 37.06.01; LK 01, pp. 125–133. ISSN: 2196-7032, 1611-2776. DOI: 10.1515/itit-2019-0025
4. Daniel Popovic, Edouard Fouché and Klemens Böhm. 'Unsupervised Artificial Neural Networks for Outlier Detection in High-Dimensional Data'. In: *ADBIS*. Lecture Notes in Computer Science. Springer, 2019, pp. 3–19. DOI: 10.1007/978-3-030-28730-6_1
3. Edouard Fouché, Junpei Komiyama and Klemens Böhm. 'Scaling Multi-Armed Bandit Algorithms'. In: *KDD*. **Acceptance rate: 9.0%**. ACM, 2019. DOI: 10.1145/3292500.3330862
2. Edouard Fouché and Klemens Böhm. 'Monte Carlo Dependency Estimation'. In: *SSDBM*. **Best Paper Award**. ACM, 2019. DOI: 10.1145/3335783.3335795
1. Edouard Fouché, Alexander Eckert and Klemens Böhm. 'In-Database Analytics with ibmdbpy'. In: *SSDBM*. ACM, 2018, 31:1–31:4. DOI: 10.1145/3221269.3223026

International Mobility

- Sept-Dec 2019. University of Illinois at Urbana-Champagin (Prof. Jiawei Han). IL, USA.
  - 3-month research collaboration in the group of Prof. Jiawei Han. This has lead to the publication of "Mining Text Outliers in Document Directories" at ICDM2020 (see above).
- June 2019. RIKEN AIP (Asst. Prof. Junpei Komiyama). Tokyo, Japan.
  - 1-month research collaboration at RIKEN AIP. Following up on the work with Junpei Komiyama ("Scaling Multi-Armed Bandit Algorithms", KDD19). Manuscript currently in preparation.
- Before 2014: Growing up and studying in Paris, France.

Teaching

Supervised Theses

12. Change Detection in High-Dimensional Data Streams. Mar21 – Sep21. Master's thesis. Tanja Fenn.
11. Analysis and Visualization of Semantics from Massive Document Directories. Mar21 – Jul21. Bachelor's Thesis. Klevia Ulqinaku.
10. Subspace Search in Data Streams. May19 – Oct19. Master's thesis. Florian Kalinke. GitHub. Publication.
9. Anytime Tradeoff Strategies for Multiple Targets. May19 – Nov19. Master's thesis. Marco Heyden. GitHub/Publication (in progress).
8. Interactive Visualization of Correlations in High-Dimensional Streams. Mar19 – Jun19. Bachelor's thesis. Yimin Zhang. GitHub.
7. Statistical Generation of High-Dimensional Data Streams with Complex Dependencies. Jun18 – Nov18. Bachelor's thesis. Alexander Poth. GitHub.
6. Adaptive Variational Autoencoders for Outlier Detection in Data Streams. Sep18 – Feb18. Master's thesis. Florian Pieper. GitHub.
5. On the Interpretability of Anomaly Detection via Neural Networks. Apr18 – Sep18. Master's thesis. Marco Sturm. With Daimler AG.
4. Energy Time Series Reduction. Jan18 – Apr18. Bachelor's thesis. Lucas Krauß. Publication.
3. Relevance-Driven Feature Engineering. Jul17 – Dec17. Master's thesis. Rosina Kazakova. With BMW AG.
2. High-Dimensional Neural-Based Outlier Detection. Apr17 – Sep17. Master's thesis. Daniel Popovic. GitHub. Publication.
1. Quality of High-Contrast Subspace Search in Data Streams with Missing Values. Feb17 – Apr17. Bachelor's thesis. Jonathan Bechtle. Co-advised with Georg Steinbuss.

Lectures

6. Lecturer: Advanced Big Data Analytics. KIT, IPD Böhm. Summer 2021.
5. Teacher: Big Data Analytics Lab. KIT, IPD Böhm. Summer 2018.
4. Teacher: Seminar on Applied and Algorithmic Views on Machine Learning. KIT, IPD Böhm. Summer 2017.
3. Teacher: Big Data Analytics Lab. KIT, IPD Böhm. Summer 2017.
2. Tutor: Big Data Analytics Lab. KIT, IPD Böhm. Summer 2016.
1. Tutor: Customer Relation Management. KIT, IISM. Winter 2015.

Service

– Program Committee (PC) member / Reviewer:

- SIGKDD 2021 (ACM Special Interest Group on Knowledge Discovery and Data Mining)
- IJCAI 2021 (International Joint Conference on Artificial Intelligence)
- ICDM 2021 (IEEE International Conference on Data Mining)

– Steering Committee member: Karlsruhe House of Young Scientists (KHYS) at KIT (from 2021)

Communication

Talks

14. *Anomaly Detection for Accelerating Materials Development.* January 22, 2021. Workshop "Data Science for Materials Science". KIT Centers KCIST, MathSEE, Materials. Karlsruhe, Germany.
13. *Mining Text Outliers in Document Directories.* November 20, 2020. ICDM'20. Sorrento, Italy (virtual). Online: https://youtu.be/6dl3ZBxB3f0
12. *Estimating Dependency, Monitoring and Knowledge Discovery in High-Dimensional Data Streams: Doctoral Defence.* July 15, 2020. KIT. Karlsruhe, Germany.
11. *Scaling Multi-Armed Bandit Algorithms.* August 8, 2019. KDD'19. Anchorage, AK, USA.
10. *Monte Carlo Dependency Estimation.* July 23, 2019. SSDBM'19. Santa Cruz, CA, USA.
9. *Knowledge Discovery in High-Dimensional Streams.* July 8, 2019. BOSCH. Renningen, Germany.
8. *Knowledge Discovery in High-Dimensional Streams.* June 19, 2019. RIKEN AIP. Tokyo, Japan.
7. *Data Science for Large-Scale Pyrolysis.* May 09, 2019. KIT, Bioliq. Karlsruhe, Germany.
6. *Scaling Bandits for Energy Stream Monitoring.* October 10, 2018. KIT, Annual retreat of the DFG Research Training Group 2153: "Energy Status Data". Gleiszellen-Gleishorbach, Germany.
5. *Knowledge Discovery in High-Dimensional Streams .* April 11, 2018. TRUMPF GmbH + Co. KG, "Software Campus – Welcome Day". Ditzingen, Germany.
4. *A nice Bandit is like a nice Watch: It has some Complications.* August 17, 2018. KIT, Institute Retreat. Bühl, Germany.
3. *Knowledge Discovery in High-Dimensional Streams.* April 5, 2018. KIT, Outreach Workshop "Data Science for Engineers". Karlsruhe, Germany.
2. *Anomaly Detection in Streams of Energy Data.* October 10, 2017. KIT, Annual retreat of the DFG Research Training Group 2153: "Energy Status Data". Gleiszellen-Gleishorbach, Germany.
1. *Tutorial: Time Series Forecasting.* September 11, 2017. Uni Passau, Summer School on Future Energy Systems. Passau, Germany.

Posters

8. *Efficient Subspace Search in Data Streams.* January 4, 2021. Data Science Summer School 2020, Ecole Polytechnique, Paris, France (virtual).
7. *Knowledge Discovery in Energy Status Data.* December 4, 2019. Review of the DFG Research Training Group 2153: "Energy Status Data". Karlsruhe, Germany.
6. *Scaling Multi-Armed Bandit Algorithms.* August 8, 2019. KDD'19. Anchorage, AK, USA.
5. *Monte Carlo Dependency Estimation.* July 23, 2019. SSDBM'19. Santa Cruz, CA, USA.
4. *Unsupervised Outlier Detection in Streams of Energy Data.* September 11, 2017. Uni Passau, Summer School on Future Energy Systems. Passau, Germany.
3. *Time Series Reduction: Shortening time series without losing energy related key characteristics.* September 11, 2017. Uni Passau, Summer School on Future Energy Systems. Passau, Germany.
2. *Unsupervised Outlier Detection in High-Dimensional Data Streams.* August 28, 2017. Ecole Polytechnique, DS^3: Data Science Summer School. Paris, France.
1. *Development of Data Anaylsis Algorithms for the Pyrolysis Process.* April 21, 2017. KIT, 2nd Bioliq Symposium. Karlsruhe, Germany.