

# Simon Lia-Jonassen

• (+47) 906 77 403 • [simon.jonassen@gmail.com](mailto:simon.jonassen@gmail.com) • <skype:simon.jonassen> • <http://s-j.github.io>

---

**Summary:** I am an engineer with experience in development of large-scale analytics and data management systems (Cxense), search engine research (NTNU, Yahoo!), and building innovative Web-applications (Sun, UNINETT). I am passionate about making things possible and solving problems that require just the right combination of analytical skills, technical knowledge, creativity and determination. Most of all I enjoy helping others. I am a team player who takes the initiative and likes to explore new areas. I strive for clear and ef-

ficient solutions, seeing the whole picture and the details, delivering my best and always moving forward.

**Specialties:** research and development; system design and implementation; performance evaluation and optimization; technical documentation; customer support and troubleshooting; giving talks and lectures.

**Interests:** information retrieval, data mining and machine learning; algorithms; distributed systems; architecture, efficiency and scalability; visualization and graphics.

---

## Experience

Cxense

OSLO, NORWAY

### Technical Lead

Oct '16 – present

Working with data analysis and machine learning.

### Senior Software Engineer

Oct '12 – Sep '16

Focused on online analytical query processing and data mining, having most impact on the API capable of importing, exporting and crunching millions of breadcrumbs users leave on the Internet in matter of milliseconds. This includes the underlying design, implementation, orders-of-magnitude performance optimizations, technical documentation and support, helping colleagues to solve their puzzles and close new deals, and presenting the technology and the challenges we are working with.

Norwegian University of Science and Technology (NTNU)

TRONDHEIM, NORWAY

### PhD Fellow/Researcher

Aug '08 – Aug '12

Did research on query processing efficiency in distributed search engines contributing to 10+ scientific articles published in good conference proceedings and journals. These publications tackle different problems within partitioned query processing, pruning and caching, propose novel techniques and demonstrate substantial improvements over state-of-the-art methods. The experimental methodology in use ranges from actual and highly optimized implementation to simulation and analytical modeling.

### Teaching Assistant/Lead TA

Jan '05 – Jun '12

Assisted [Introduction to databases](#), [Management of very large data volumes](#), [Introduction to algorithms](#), [Advanced algorithms](#) and [Linear algebra and differential equations](#). Tasks varied from helping students and grading exercises/exams to making exercises, giving lectures, leading TA's, system and course administration. Have also worked as a recruiter for NTNU.

Yahoo! Research

BARCELONA, SPAIN

### Visiting Researcher/Intern

Sep '11 – Feb '12

Worked closely with [Ricardo Baeza-Yates](#), [B. Barla Cambazoglu](#) and [Fabrizio Silvestri](#) focusing on caching and query processing optimizations in distributed Web-search engines. This collaboration resulted in four publications, including a journal article (TWEB) and a full SIGIR paper.

Sun Microsystems (Oracle)

TRONDHEIM, NORWAY

### Summer Intern

Jun '07 – Aug '07

Wrote a tool for aggregation and visualization of performance logs from overnight benchmarks.

UNINETT

TRONDHEIM, NORWAY

### Summer Intern

Jun '05 – Aug '05, Jun '06 – Aug '06

Maintained [interactive network maps](#) and wrote a prototype for real-time aggregation of network topology and statistics from heterogeneous sources and visualization with Google Maps.

## Education

Norwegian University of Science and Technology

TRONDHEIM, NORWAY

### Doctor of Philosophy, Computer Science

2008 – 2013

Thesis: [Efficient Query Processing in Distributed Search Engines](#); GPA: 5.75/6.

### Master of Technology, Computer Engineering

2003 – 2008

Thesis: [Distributed Inverted Indexes](#); GPA: 5.47/6.

## Qualifications

- **Development:** Adept programmer and problem solver, extensive user of Java and GNU/Linux. Fluent with continuous development process. Experienced with customer support and system maintenance.
- **Management:** Strong communication, organization, writing and presentation skills. Experienced with managing small teams and getting things done.
- **Research:** Familiar with experimental research and publication process. Experienced with leading independent research, working within a group and with geographically remote co-authors.
- **Natural languages:** Norwegian, Russian – *bilingual proficiency*, English – *full professional proficiency*.

## Selected Publications

- **Improving dynamic index pruning via linear programming** – *Proc. of the 12th International Workshop on Large-Scale and Distributed Systems for Information Retrieval (LSDS-IR), ACM 2015.*
- **Improving the performance of pipelined query processing with skipping – and its comparison to document-wise partitioning** – with Svein Erik Bratsberg. *World Wide Web Journal*, 17(5), pp 949-967, Springer 2014.
- **A term-based inverted index partitioning model for efficient distributed query processing** – with B. Barla Cambazoglu, Enver Kayaaslan and Cevdet Aykanat. *Transactions on the Web*, 7(3)-15, 23p, ACM 2013.
- **Prefetching query results and its impact on search engines** – with B. Barla Cambazoglu and Fabrizio Silvestri. *Proc. of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp 631-640, ACM 2012.
- **Modeling static caching in Web search engines** – with Ricardo Baeza-Yates. *Proc. of the 34th European Conference on Information Retrieval (ECIR)*, pp 436-446, Springer 2012.
- **Intra-query concurrent pipelined processing for distributed full-text retrieval** – with Svein Erik Bratsberg. *Proc. of the 34th European Conference on Information Retrieval (ECIR)*, pp 413-425, Springer 2012.
- **Efficient processing of top-k spatial keyword queries** – with João B. Rocha-Junior, Orestis Gkorgkas and Kjetil Nørvåg. *Proc. of the 12th International Symposium on Spatial and Temporal Databases (SSTD)*, pp 205-222, Springer 2011.
- **Efficient compressed inverted index skipping for disjunctive text-queries** – with Svein Erik Bratsberg. *Proc. of the 33rd European Conference on Information Retrieval (ECIR)*, pp 530-542, Springer 2011.
- **A combined semi-pipelined query processing architecture for distributed full-text retrieval** – with Svein Erik Bratsberg. *Proc. of the 11th International Conference on Web Information Systems Engineering (WISE)*, pp 587-601, Springer 2010. *Received the Best Paper Award.*

*Please refer to my Google Scholar profile for the complete list of publications.*

## Selected Talks

- **Large-Scale Real-Time Data Management for Engagement and Monetization** – Invited Talk at LSDS-IR, 2015.
- **Efficient Query Processing in Web Search Engines** – Guest Lecture at UiO, 2015.
- **A brief introduction to Apache Spark** – Brown Bag Session at Cxense, 2014.
- **Hardware Developments and Algorithm Design: "What should be done to IR algorithms to meet current, and possible future, hardware trends?"** – Trial Lecture at NTNU, 2013.
- **Efficient Query Processing in Distributed Search Engines** – PhD Defense at NTNU, 2013.

## Honors, Awards and Volunteering

- Guest lecturer: UiO (Apr '16, Apr '15), NTNU (Oct '15, Apr '15) and Westerdals ACT (Oct '15, Nov '14).
- Reviewer: Information Processing & Management (2016), LSDS-IR (2015, 2014), ECIR (2012).
- External examination and thesis reviewer: UiO (Aug '16, Jun '16, Aug '15), NTNU (Jun '14).
- ACM collegiate programming contests: 14<sup>th</sup> out of 51 teams at the North-Western European Contest 2007 (2<sup>nd</sup> out of 6 Norwegian teams) and 10<sup>th</sup> out of 86 at the Nordic Contest 2005 (4/36 Norway, 2/18 NTNU).
- Won a Photoshop competition at a LAN party in 2002, two prizes for solving Calculus 3 puzzles in 2004, 1337-of-the-class award at the graduation dinner in 2008, and a Cxense OWL Award ('Wizard') in Q3 2016.

*Diplomas, transcript records and references are available on request.*