

*Making new things possible with search engines,
building bridges across technologies and between people*

Personal Data

Gender Male

Nationality Dutch

Education

2007–2011 **Doctorate (PhD)**, *University of Amsterdam*, Amsterdam, the Netherlands.

Dissertation: System Evaluation of Archival Description and Access.

Worked on the evaluation of information retrieval for digital archives of the *National Archives of the Netherlands*. Designed and implemented retrieval algorithms and multiple front-ends of a search engine. Processed and analyzed log files to understand user search behavior and to evaluate the search engine.

2006–2007 **Master of Science**, *University of Amsterdam*, Amsterdam, the Netherlands.

Human Centered Multimedia (Information Studies).

Major: multimedia systems, including video search. Master's degree thesis based on joined work with researchers of the *National Research Institute for Mathematics and Computer Science* on the credibility of information sources in search engines; work published and presented at major international venues.

2002–2006 **Bachelor, Master of Arts**, *University of Groningen*, Groningen, the Netherlands.

Information Science, note: Master's degree with honors, cum laude.

Major: language technology. Worked for Master thesis with the University Library to design and implement a Semantic Web system for their bibliographic catalogs; published and presented at international conferences.

Work Experience

May 2019 – **Sr. Search Engineer, Sr. Software Engineer**, *Elsevier*, Amsterdam, the Netherlands.

now Collaboration with infrastructure, back-end, front-end and QA engineers to port and improve search solutions in the Life Sciences domain. Design and implementation of the migration from MarkLogic database to Elasticsearch for EMBASE, which is a biomedical and pharmacological bibliographic search engine.

Technical lead on search solutions. Revised and extended indexing schemas to support search functionalities. Re-writing XQuery queries to equivalent but more efficient Elasticsearch DSL queries, which includes exact and phrase search, fulltext search, efficient query expansion with synonyms, proximity queries with search fields and wildcards. Implemented advanced (visual and nested) faceting and highlighting.

○ *Technologies used*: Java, Python, Bash scripting, Elasticsearch, Solr, Docker, Kubernetes, Kafka, Jenkins, Git, JIRA.

April 2017 – **Sr. Search Engineer, Sr. Software Engineer**, *Swisscom Directories AG*, Zürich, Switzerland.

April 2019 End-to-end responsibility for designing, planning, implementation and testing search algorithms on the local.ch online directory platform (phone book, Yellow Pages, finding restaurants and other SMEs). Collaboration in creating search functionalities to improve the SMEs' performance and relevance to users with a team of back-end engineers, product managers, user experience experts, designers and web and mobile developers.

Proposed, designed and implemented a pipeline from the Big Data platform to the main search engine by calculating near real-time trending, popular and top-rated businesses for delivering recommendations. Designed and implemented a framework to evaluate search effectiveness and investigate queries. Contributed to operational excellence by improving and optimizing the intricate matching and ranking of a highly customized search engine.

○ *Technologies used*: Java, Python, Bash scripting, Javascript, Solr, Elasticsearch, Docker, Kubernetes, MongoDB, Google BigQuery, Jupyter Notebook, Kafka, Bamboo, Git, JIRA.

Nov 2014 – **Sr. Search Engineer, Sr. Software Engineer**, *LexisNexis*, Leiden, the Netherlands.

Mar 2017 Technical lead on the design, implementation and evaluation of search engines (e.g. ingestion of terabytes of data, query parsing, retrieval algorithms) for bibliographic patent records and in cross-language full-text. Coordinating leading role in the delivery and technical evaluation of implementations by external consultants and contractors.

Designed infrastructure topology for superior search efficiency, designed and implemented expert patent (legal) search functionality such as stemmed vs. non-stemmed search, proximity search, case-sensitive search, cross-language retrieval (from English to CJK), document filtering based on document IDs, fields and user-generated content, advanced classification search, and patent family de-duplication algorithms.

○ *Technologies used*: Java, Node.js, AngularJS, Python, Bash scripting, XSLT, Elasticsearch, OpenStack, MongoDB, RabbitMQ, Jenkins/Bamboo, Git, JIRA.

- Jan 2014 – **Software Engineer, Information Analyst**, *Vancis Advanced ICT Services B.V.*, Amsterdam, the Netherlands.
 Nov 2014
 Technical lead on software solutions for Cloud services. Business development for these services and introduction of new Cloud services, in particular Platform-as-a-Service (PaaS). Set up MoSCoW analyses as part of RFIs and RFPs. Identified business processes and KPIs by interviewing colleagues and extracting information from systems, storing and integrating different information streams for metering and billing of cloud services. Design, recommendation and implementation of changes with automation.
 ◦ *Technologies used*: Hadoop, OpenStack, OpenShift, Bash scripting, Perl, PHP, Python, jQuery, (My)SQL, SOAP, CLI-based APIs, UML, SVN, Git.
- Nov 2012 – **Information Specialist**, *NIOD Institute for War, Holocaust and Genocide Studies*, Royal Netherlands Academy of Arts and Sciences, Amsterdam, the Netherlands, 0.2 FTE.
 Dec 2013
 Worked on the European Holocaust Research Infrastructure (EHRI) project, working with archivists and (international) IT specialists on a digital research environment for historians.
 ◦ *Technologies used*: Perl, JSON, XML, XSLT, Apache Solr, Git.
- Oct 2011 – Dec 2013 **Software Engineer**, *Meertens Institute*, Royal Netherlands Academy of Arts and Sciences, Amsterdam, the Netherlands.
 Design and development of advanced search engines and user interfaces for a plethora of complex and large amounts of structured data (i.e. Dublin Core, TEI, WSDL, DIDL, custom markup) for scientists. I delivered software to provide digital access to historical editorial cartoons in newspapers.
 ◦ *Technologies used*: Apache Solr, PHP, Perl, jQuery, XML, XSLT, Bash scripting, SQL, R, SVN.
- Aug. 2011 – **Consultant**, *VLC Projects B.V.*, Amsterdam, the Netherlands.
 Sept. 2011
 Developed ideas on enterprise search and worked on technical solutions. Worked on content management in Alfresco and Drupal. Gave trainings and did presentations.
- Oct. 2007 – **Lecturer**, *University of Amsterdam*, Amsterdam, the Netherlands, 0.25 FTE.
 June 2011
 Independently teaching the course “Representation and Retrieval” – mainly on using XML for describing, presenting, and gaining access to personal document collections, including preparing slides, giving lectures, supervising and grading course work.
- July 2007 – **Junior Researcher**, *University of Amsterdam*, Amsterdam, the Netherlands, 0.75 FTE.
 June 2011
 Member of the Retrieving Encoded Archival Descriptions More Effectively (README) project funded by the Netherlands Organisation for Scientific Research (NWO), which resulted in numerous publications and presentations at large top-tier scientific conferences in Europe, the United States and Asia.
 ◦ *Technologies used*: XML databases, Lucene, Indri, Perl, SPSS, Bash scripting, L^AT_EX en BibT_EX.
- Feb. 2006 – **Intern**, *University Library Groningen*, Groningen, the Netherlands.
 July 2006
 Worked on the development of a Semantic Web system based on bibliographic information stored in MARC21 records at the Digital Library department.
 ◦ *Technologies used*: XML, XSLT, Perl, Protégé, Sesame, SeRQL.

Professional and Technical Skills

Research	Experienced with quantitative and qualitative research following the scientific method, including experimental user studies, reporting on and presenting the results.	Application Stack	Experienced with working in Linux; comfortable with creating and working with SOAP, RESTful services, web and CLI-based APIs and libraries.
Programming	Java, Python, Perl, PHP, Bash scripting, R, Javascript, XSL, XQuery, SQL, L ^A T _E X.	Other	European driving license B.

Natural Languages

Dutch	Fluent, both oral and in writing.	German	Limited working proficiency.
English	Fluent, both oral and in writing.	Chinese	Limited working proficiency.

Grants and Awards

- *Association for Computers in the Humanities* (ACH) Young Scholar bursary in 2008; \$1000.
- *European Conference on Digital Libraries* (ECDL) Student Scholarship in 2009; €360.
- *e-Humanities* grant in 2012 by the Netherlands Academy of Arts and Sciences to work on access to historical editorial cartoons in newspapers; €30K
- *Reed Tech Innovation Award 2016* for innovative contributions and subsequent commercial release of the patent research tool TotalPatent One; handed out one time per year to an employee.

Miscellaneous

- Additional information with descriptions of my completed projects, selected publications, and recommendations can be found on my LinkedIn profile: www.linkedin.com/in/juntezhang.