

CURRICULUM VITÆ

SOPHIA KATRENKO

WWW: <http://sophia.katrenko.com/>

E-mail: sophia@katrenko.com

Other: DBLP , Google Scholar

EXPERIENCE

Head of Data Science at DEUS, 10/2020-
Senior Manager Projects at Cognizant, 02/2019-
2020

Head of Smart Content Capabilities at Elsevier, 10/2016-
12/2018

Amsterdam, the Netherlands

project lead for NLP/ML projects for Research Products: gathering requirements, planning, execution, stakeholder management, driving from PoC to productionalisation and going live • projects with internal stakeholders to improve content quality (Data Science squad lead; ETL and analysis on Databricks; Scrum), engaging with the 3rd party providers, reporting on KPIs • team management (NLD and IND)

Team Lead Data Science at Elsevier, 04/2015-
09/2016

Amsterdam, the Netherlands

tech lead (hands-on) within Research Products for commercial bibliometrics projects (30+ individual contribution, 100+ as a team) a year for EMEA, APAC, and AMER, meeting financial targets (e.g., interdisciplinary research for HEFCE, UK: methodology and execution; International comparative performance of the UK research base 2016 for the Department for Business, Energy and Industrial Strategy, UK) • World university ranking • data for national assessments • team management (NLD and JPN)

Data Scientist at Elsevier, 08/2014-
03/2015

Amsterdam, the Netherlands

data extraction and analysis of bibliometric data (Scopus and external sources) • serving governmental and funding bodies

Quantitative Analyst at Capital Tool Company, 04/2013-
06/2014

G-20 SME Finance Challenge winner

Amsterdam, the Netherlands

default prediction for SME and programming for the TREFI platform (Java, Weka, MySQL, shell scripting) • monthly product updates • time series analysis (R) • reporting to the World Bank

Postdoctoral Researcher at Utrecht University, 2010 - 2013
Utrecht, the Netherlands

“Between logic and common sense” NWO project • reasoning in language (Java) • publishing • teaching • student supervision

Visiting Researcher at University of Oslo, Oslo, Norway <i>building a system for detecting inferences in text (Java)</i>	08 - 09/2011
Researcher at University of Amsterdam, Amsterdam, the Netherlands <i>research on information extraction (Java, Perl) • publishing • teaching</i>	2008 - 2010
PhD Researcher at University of Amsterdam, Amsterdam, the Netherlands <i>Adaptive Information Disclosure project (part of Virtual Lab e-science) • programming (Perl, Java) • publishing • teaching</i>	2004 - 2008
Researcher at Lviv Polytechnic University, Lviv, Ukraine <i>research on resolving ambiguities in Ukrainian (Perl, XSLT)</i>	2002 - 2004
Visiting Researcher at University of Tübingen, Tübingen, Germany <i>Virtual Patient project (Perl) • morpho - syntactic disambiguation in German</i>	11/2002 - 11/2003
Visitor at Bergakademie Freiberg, Freiberg, Germany <i>literature review on vagueness and ambiguities in language</i>	01/2002
Freelance journalist at Cnews <i>writing articles (IT, analytics)</i>	2001 - 2002
Analyst at Golden Rain Advertising, Lviv, Ukraine <i>IS design and implementation, predictive models in advertising</i>	1999 - 2001

EDUCATION

<i>PhD in Computer Science,</i> University of Amsterdam, Amsterdam, the Netherlands	10.09.2009
<i>Title: "A Closer Look at Learning Relations from Text"</i> <i>Supervisor:</i> Prof. Dr. P. W. Adriaans <i>Committee:</i> Prof. Dr. Colin de la Higuera, Prof. Dr. Guus Schreiber, Prof. Dr. Peter Sloot, Prof. Dr. Maarten de Rijke, Prof. Dr. Bob Wielinga	
<i>MSc in Computer Science (with distinction),</i> Department of Information Systems and Networks, National University Lviv Polytechnics, Lviv, Ukraine	2000-2001
<i>BSc in Computer Science (with distinction),</i> National University Lviv Polytechnics, Lviv, Ukraine	1996-2000
School 8 with advanced course in German, Lviv, Ukraine	1986-1996

PARTICIPATION IN CHALLENGES

Genic Interaction Extraction Challenge (2005) • 2nd Recognising Textual Entailment Challenge (2006) • BioCreative II (2006) • TREC Genomics (2007) • SemEval (2007) • Concrete Noun Categorization (2008, top performing system on 6-way classification)

TECHNICAL SKILLS

platforms: *nix systems, Windows
languages: shell scripting, Python, Java, Perl, C++, Prolog
cluster-computing frameworks: Apache Spark
IDE: PyCharm, Eclipse, NetBeans
database management: MySQL, Amazon DynamoDB, MS Access
database design tools: DFD and ERD in ERWin
data mining/ML tools and libraries: libsvm, keras, scikit-learn, R, Weka (API)
cloud computing: AWS
version control: svn, Git
other: Scrum (Certified Scrum Master 2015), Databricks, HPC XML, XSLT, HTML, \LaTeX , Microsoft Office

RESEARCH INTERESTS

applied machine learning • explainable AI • sustainable ML

AWARDS AND FELLOWSHIPS

Yggdrasil grant (29% selection rate; 26,000 NOK)	2011
EACL grant for participation in ESSLLI 2004	2004
Volkswagen Foundation stipend	2002/2003
Leonhard Euler fellowship	2001/2002
Lviv Municipality Stipend	1999

SERVICE

Mentor at "Hack the crisis NL"	2020
Mentor at Smart Tribe	2020
Startup mentoring	2019
Secretary of the Standing Committee of European Summer School in Logic, Language and Information, held in a different European city yearly • around 500 participants	2007 - 2018
Reviewer for NLDB (2014, 2012), LREC (2020, 2018, 2014, 2012), ACL 2013, *SEM 2013, IEEE Transactions on Neural Networks and Learning Systems (2013), EACL Student Session 2012, PC CBMS 2012, PC Workshop on Relational Models of Semantics (RELMS) at ACL 2011, PC CogSci Conference 2010, IEEE CBMS (2010, 2009, 2008), HLIE workshop at ECML PKDD 2008, ESSLLI Student Session 2007, MICAI 2006	

ORGANIZATION

- | | |
|---|------|
| Belgian-Dutch Conference in Machine Learning,
together with Maarten van Someren and Pieter Adriaans
Amsterdam, the Netherlands | 2007 |
| Organizer of ESSLLI Student Session, together with Janneke Huitink
Malaga, Spain | 2006 |
| Co-organizer of the “Language and Computation” track
at ESSLLI Student Session 2005, together with Tamás Biró
Edinburgh, United Kingdom | 2005 |

TEACHING

- | | |
|---|---------------------|
| Learning in Computational Linguistics (lecturer, UU)
student evaluation 2012: lecturer (overall)=4, lecturer (expertise)=4.5,
course content=4.3, increased interest in subject=4.5 (max=5) | 2011, 2012 |
| Taal- en spraaktechnologie (co-lecturer, UU) | 2011, 2012 |
| Artificial Intelligence, Master in AI project (UvA) | 2010 |
| Capita in Machine Learning, Language and Speech (UvA) | 2009, 2008,
2007 |
| Language Learning (TA, UvA) | 2007, 2006 |

PUBLICATIONS (SELECTION OUT OF 30+ PUBLICATIONS)

IN JOURNALS AND REFEREED PROCEEDINGS

Sophia Katrenko. *Could you make me a favour and do coffee, please?: Implications for Automatic Error Correction in English and Dutch*. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics (*SEM)*, 2012.

Sophia Katrenko, Pieter Adriaans, and Maarten van Someren. Using Local Alignments for Relation Extraction. *International Journal of Artificial Intelligence Research (JAIR)*, impact factor 1.981 (ISI Web), 38:1–48, 2010.

M. Roos, M. S. Marshall, A. P. Gibson, M. Schuemie, E. Meij, S. Katrenko, W.R. van Hage, K. Krommydas, and P.W. Adriaans. Structuring and Extracting Knowledge for the Support of Hypothesis Generation in Molecular Biology. *BMC Bioinformatics*, 10(Supplement 10), 2009.

Sophia Katrenko and Pieter Adriaans. A Local Alignment Kernel in the Context of NLP. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling)*, Manchester, UK, 2008. [25% acceptance rate].

Sophia Katrenko and Pieter Adriaans. Semantic Types of Some Generic Relation Arguments: Detection and Evaluation. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL/HLT)*, Columbus, USA, 2008. [26,8% acceptance rate].

THESES

Sophia Katrenko. *A Closer Look at Learning Relations from Text*. PhD thesis, University of Amsterdam, the Netherlands, 2009.

Sophia Katrenko. *An Information System in the Advertising Branch*. Master's thesis, National University Lviv Polytechnics, Ukraine, 2002.

EDITED VOLUMES

Janneke Huitink and Sophia Katrenko, editors. *Proceedings of the ESLLI 2006 Student Session*. Malaga, Spain, 2006.

Pieter Adriaans, Sophia Katrenko, and Maarten van Someren, editors. *Proceedings of the Dutch-Belgian Conference on Machine Learning*. Amsterdam, the Netherlands, 2007.

NATURAL LANGUAGES

Ukrainian (*native*) • English (*fluent*) • Dutch (*very good*, passed Korte Vrijstellingstoets in 2010) • German (*fluent*, 10 years at school - from primary to high school, subjects: Deutsche Sprache, Deutsche Literatur, Landeskunde, Marketing, and Technische Übersetzung) • Polish (*very good*)

November 3, 2020