CURRICULUM VITAE
MASHA MEDVEDEVA

PERSONAL INFORMATION

NAME:	Masha Medvedeva
PLACE AND DATE OF BIRTH:	Moscow, Russia | 21 December 1993
CURRENTLY LIVING IN :	The Hague, the Netherlands
PHONE (PRIVATE):	(+31) (0)6 31237186
EMAIL:	medvmr@gmail.com
WEBSITE:	masha-medvedeva.github.io

PROFILE

I am a legal tech engineer and researcher. My work lies within analysing and implementing solutions for automation in legal domain, as well as analysing products that are currently on the market. In the past much of my work lied within automatic detection of linguistic patterns within legal data, with particular focus on predicting court decisions. A big part of my experience is working with judgements of the European Court of Human Rights as well as Dutch case law. My professional interests include but are not limited to improving accessibility to justice, legal analytics, interpretability of NLP models, automation through AI, and related areas.

RELEVANT WORK EXPERIENCE

2022-PRESENT **Posdoctoral researcher** in Legal Technology
COHUBICOL, Department of Digital Security, Radboud University, Nijmegen

- Evaluating legal tech tools available on the market and researching their claimed and actual functionality
- Developing an educational tool to help legal experts understand technology behind legal tech services
- Publishing academic work in AI and Law
- Building Typology of Legal Technologies

2017-2022 **PhD candidate** in NLP and Law
Centre for Language and Cognition Groningen & Department of Legal Science, University of Groningen

- Researched a variety of topics in AI and Law
- Implemented machine learning systems for predicting and analysing court decisions in Dutch, French and English
- Developed systems for legal data analysis
- Consulted on developing legal tech tools
- Developed methodology and fundamental terminology in the field of predicting court decisions
- Successfully managed and budgeted 4,5-year project
- Supervised and mentored students, research assistants and interns
- Published 20+ articles
- Presented at 25+ international conferences
- Taught machine learning and natural language processing at bachelors and masters level

1

Media about my work

NRC	https://www.nrc.nl/nieuws/2020/12/30/robot-weet-welke-uitspraak-het-hof-zal-doen-a4025683 (Dutch)
Mr. Online	https://www.mr-online.nl/computer-voorspelt-ehrm-uitspraken/ (Dutch)
The Law of Tech Podcast	https://thelawoftech.com/podcast/using-machine-learning-to-predict-court-decisions-of-the-ecthr/

Education

2017-2022	**Doctor of Philosophy** in NLP & Law **University of Groningen**, the Netherlands
2015-2017	Joint Degree in Language and Communication Technologies, MA at **University of Groningen**, the Netherlands MSc at **Saarland University**, Germany
Fall 2013	Exchange Semester at **University of Bern**, Switzerland (Linguistics programme)
2011-2015	BA in Fundamental and Computational Linguistics *(cum laude)* **National Research University Higher School of Economics**, Moscow, Russia

Areas of Expertise

Python, Machine Learning, Natural Language Processing, Unix, Legal Technology, Legal Analytics, Judicial Diagnostics, Predictive (In)justice, Neural Networks, Information Retrieval, Information Extraction, Language Technology, Dutch Case Law, European Court of Human Rights
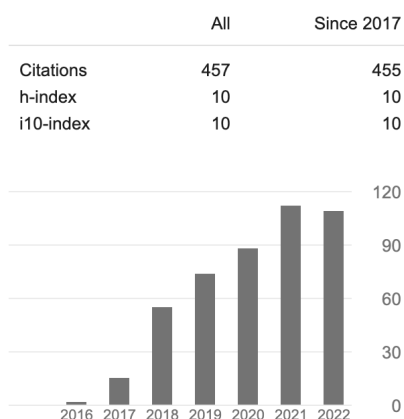
Languages

English (fluent), Russian (native), Dutch (limited working proficiency)

Publications

Diver, L., McBride, P., **Medvedeva, M**, Banerjee, A., D'hondt, E., Duarte Nicolau, T., Dushi, D., Gori, G., Van Den Hoven, E., Meessen, P., and Hildebrandt, M. (2022). Typology of legal technologies. In *Cross-disciplinary Research in Computational Law (CRCL): Computational 'Law' on Edge 2022*.

Medvedeva, M (2022). *Identification, Categorisation and Forecasting of Court Decisions*. PhD thesis, University of Groningen.

Medvedeva, M., Wieling, M., and Vols, M. (2022). Rethinking the field of automatic prediction of court decisions. *Artificial Intelligence and Law*.

Medvedeva, M., Dam, T., Wieling, M., and Vols, M. (2021). Automatically identifying eviction cases and outcomes within case law of dutch courts of first instance. In *Legal Knowledge and Information Systems*, pages 13–22. IOS Press.

Medvedeva, M, Üstun, A., Xu, X., Vols, M., and Wieling, M. (2021). Automatic judgement forecasting for pending applications of the european court of human rights. In *Proceedings of the Workshops on Automated Semantic Analysis of Information in Legal Text (ASAIL 2021)*.

Medvedeva, M., Xu, X., Wieling, M., and Vols, M. (2020c). JURI SAYS: Prediction system for the European Court of Human Rights. In *Legal Knowledge and Information Systems: JURIX 2020: The Thirty-third Annual Conference, Brno, Czech Republic, December 9-11, 2020*, volume 334, page 277. IOS Press.

Wieling, M., Tiede, M., Rebernik, T., de Jong, L., Braggaar, A., Bartelds, M., **Medvedeva, M**, Heisterkamp, P., Offrede, T. F., Sekeres, H., et al. A novel paradigm to investigate phonetic convergence in interaction. In *Proceedings of 12th International Seminar on Speech Production*.

Medvedeva, M., Wieling, M., and Vols, M. (2020b). The danger of reverse-engineering of automated judicial decision-making systems. *arXiv preprint arXiv:2012.10301*.

Medvedeva, M., Vols, M., and Wieling, M. (2020a). Using machine learning to predict decisions of the European Court of Human Rights. *Artificial Intelligence and Law*, 28:237–266.

Haagsma, H., Kreutz, T., **Medvedeva, M.**, Walter, D., and Nissim, M. (2019). Overview of the Cross-Genre Gender Prediction Shared Task on Dutch at CLIN29. In *Proceedings of the Shared Task on Cross-Genre Gender Prediction in Dutch at CLIN29 (GxG-CLIN29)*.

Wieling, M., Blankevoort, C., Hukker, V., Jacobi, J., de Jong, L., Keulen, S., **Medvedeva, M.**, van der Ploeg, M., Pot, A., Rebernik, T., Veenstra, P., and Noiray, A. (2019). The Influence of Alcohol on L1 vs. L2 Pronunciation. In *Proceedings of ICPhS 2019*.

Medvedeva, M, Vols, M., and Wieling, M. (2018). Judicial Decisions of the European Court of Human Rights: Looking into the Crystall Ball. In *Proceedings of the Conference on Empirical Legal Studies in Europe 2018*.

Basile, A., Dwyer, G., **Medvedeva, M.**, Rawee, J., Haagsma, H., and Nissim, M. (2018). Simply the Best: Minimalist System Trumps Complex Models in Author Profiling. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 143–156. Springer.

Kroon, M., **Medvedeva, M.**, and Plank, B. (2018). When Simple N-gram Models Outperform Syntactic Approaches: Discriminating between Dutch and Flemish. In *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*, pages 244–253.

Basile, A., Dwyer, G., **Medvedeva, M.**, Rawee, J., Haagsma, H., and Nissim, M. (2017a). Is There Life beyond N-grams? A Simple SVM-based Author Profiling System. In *Working Notes of CLEF 2017-Conference and Labs of the Evaluation Forum.* **PAN Shared Task Winner**.

Medvedeva, M, Haagsma, H., and Nissim, M. (2017). An Analysis of Cross-Genre and In-Genre Performance for Author Profiling in Social Media. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 211–223. Springer. **Best Paper Award at CLEF 2017**.

Basile, A., Dwyer, G., **Medvedeva, M.**, Rawee, J., Haagsma, H., and Nissim, M. (2017b). N-gram: New Groningen Author-Profiling Model. *arXiv preprint arXiv:1707.03764.*

Medvedeva, M., Kroon, M., and Plank, B. (2017). When Sparse Traditional Models Outperform Dense Neural Networks: the Curious Case of Discriminating between Similar Languages. In *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects*, pages 156–163. Association for Computational Linguistics.

Busger op Vollenbroek, M., Carlotto, T., Kreutz, T., **Medvedeva, M.**, Pool, C., Bjerva, J., Haagsma, H., and Nissim, M. (2016). GronUP: Groningen User Profiling: Notebook for PAN at CLEF 2016. In Balog, K., Cappellato, L., Ferro, N., and Macdonald, C., editors, *Working Notes of CLEF 2016*, CEUR, pages 846–857. Workshop Proceedings. **PAN Shared Task Winner**.

Arkhangelshiy, T. and **Medvedeva, M.** (2016). Developing Morphologically Annotated Corpora for Minority Languages of Russia. In *Proceedings of Corpus Linguistics Fest 2016*, CEUR, pages 1–6.

	All	Since 2017
Citations	457	455
h-index	10	10
i10-index	10	10

https://scholar.google.com/citations?user=Z5ktfYMAAAAJ&hl=en

4

TALKS (SELECTED)

Nov 2022	CLCR22 conference (Brussels, Belgium) **Talk (with M.Hildebrandt, L. Diver, P.McBride): "Presentation of the COHUBICOL Typology of Legal Technologies"**
Sep 2022	iHub, Radboud University (Nijmegen, the Netherlands) **Talk (with M.Hildebrandt, L. Diver, P.McBride): "Typology of Legal Technologies"**
Mar 2022	Institute of Computing and Information Science, Radboud University (Nijmegen, the Netherlands) **Talk (invited): "Rethinking the field of automatic prediction of court decision"**
Mar 2022	EVICT Talks, University of Groningen (the Netherlands) **Talk (invited): "Automatically Identifying Eviction Cases and Outcomes Within Case Law of Dutch Courts of First Instance"**
Mar 2022	Computational Legal Studies Conference 2022 (Singapore, remote) **Talk: "Rethinking the field of automatic prediction of court decision"**
Feb 2022	Ministry of Justice and Security (the Hague, the Netherlands, remote) **Webinar (invited): "Myths and realities of automatic prediction of court decisions"**
Dec 2021	JURIX 2021 (Vilnius, Lithuania) **Talk: "Automatically Identifying Eviction Cases and Outcomes Within Case Law of Dutch Courts of First Instance"**
Jun 2021	4th International IT Law Summer School (Turkey, remote) **Talk: "Predicting Court Decisions: Myths and Realities"**
Jun 2021	SICSS-Law Panel (Maastricht, remote) **Talk (invited): "Panel discussion on the (many) paths to computational social science research in law."**
Jun 2021	ASAIL 2021 (Brazil, remote) **Talk: "Automatic judgement forecasting for pending applications of the European Court of Human Rights"**
Jan 2021	AI4Law Workshop series (Oxford, remote) **Talk (invited): "To predict or not to predict? Rethinking the field of automatic court decision classification"**
Dec 2020	JURIX 2020 (Czechia, remote) **Talk: "JURI SAYS: Prediction system for the European Court of Human Rights"**
May 2020	Stanford CodeX (USA, remote) **Talk (invited): "Using Machine Learning to Predict Decisions of the European Court of Human Rights"**

Oct 2019	Maastricht Law & Tech Lab Launching Event (Maastricht, the Netherlands) **Talk (invited): "Potential and Implications of Using Machine Learning for Predicting and Making Court Decisions"**
Sep 2019	Law, Technology & Trust (Copenhagen) **Talk (invited): "Predictive Analytics v. Integrity & Trust in Legal Decision-Making"**
Jan 2019	CLIN29 (Groningen, the Netherlands) **Talk: "Tomorrow's Decisions Today: Using Press-Releases for the Automatic Prediction of Decisions of the European Court of Human Rights"**
Dec 2018	JURIX 2018 (Groningen, the Netherlands) **Presentation at Doctoral Consortium: "Automatic detection of linguistic patterns in legal big data"**
Oct 2018	Network Analysis and Machine Learning in Law - 2nd Annual Workshop on Law and Social Science Methods in Oslo (Oslo, Norway) **Talk (invited): "Looking into the Crystal Ball: Using Machine Learning to Predict Decisions of the European Court of Human Rights"**
Jun 2018	Conference for Empirical Legal Studies in Europe (Leuven, Belgium): **Talk: "Judicial Decisions of the European Court of Human Rights: Looking into the Crystal ball"**
May 2018	Conference Legal Aspects of Data Science (Tilburg, the Netherlands) **Talk (invited): "Using Machine Learning to Predict (Future) Judicial Decisions of the European Court of Human Rights"**
Jan 2018	NoLesLaw First Annual Conference 2018 (Copenhagen, Denmark) **Talk (invited): "Judicial Decisions of the European Court of Human Rights: Looking into the Crystal Ball" (with M. Vols)**
Dec 2017	Third annual conference Public Trust & Public Law (Groningen, the Netherlands) **Poster presentation: "Judicial Decisions of the European Court of Human Rights: Looking into the Crystal Ball"**
Oct 2014	4th Mikola Conference on Lexicology and Lexicography of the Uralic and Siberian languages (Szeged, Hungary) **Talk: "An Online Annotated Corpus of Udmurt Language" (with T. Arkhangelskiy)**
Sep 2014	Student Conference at School of Linguistcs at Higher School of Economics (Moscow, Russia) **Talk: "An Online Annotated Corpus of Udmurt Language"**
Mar 2013	Student conference on Linguistics at Faculty of Philology, Higher School of Economics (Moscow, Russia) **Talk: "Synonymy of verbs with and without 'sja' in Modern Russian language"**

STUDENT SUPERVISION

2022	Co-supervising (with M. Wieling, M.Vols & M. Bruijn) PhD student in NLP & Law Iris Schepers
2021	Co-supervising (with M. Wieling and M.Vols) master student in Information Science Iris Schepers Thesis: "Predicting Citations in Dutch Case Law with Natural Language Processing"
2021	Co-supervising (with M. Wieling and M.Vols) master student in Artificial Intelligence Chi Sam Mac Thesis: "Predicting Judicial Decisions for the Dutch Supreme Court Using Machine Learning"
2020	Co-supervising (with M. Wieling and G.Rigau) master student Xiao Xu Thesis: "Multilingual Classification for Legal Prediction"
2020	Co-supervising (with M. Wieling and M.Vols) bachelor student Thijmen Dam Thesis: "Predicting Judicial Decisions of the Court of First Instance and the Court of Appeal in Dutch Eviction Cases Using Machine Learning and Pre-Trial Information"
2020	Co-supervising (with M. Wieling and M.Vols) bachelor student Benjamin Kleppe Thesis: "Automatically predicting the decisions of Dutch eviction court cases"
2019	Co-supervising (with M. Wieling and M.Vols) bachelor student Willem Datema Thesis: "Predicting decisions of Dutch Rent Tribunal cases"
2019	Supervising research assistant Xiao Xu Project: Development of the online platform JURI SAYS (jurisays.com)
2018	Intern supervision of Nina Hosseini Research: "An analysis of dissenting opinions in the legal text of European Court of Human Rights"

Teaching Experience

Fall 2020	"Learning from data" - online masters course with Malvina Nissim *University of Groningen, the Netherlands*
Fall 2019	"Learning from data" - masters course with Malvina Nissim *University of Groningen, the Netherlands*
Fall 2019	Invited lecture for Data-wise minor: Natural Language Processing for Legal Data: Using Machine Learning for Predicting Court Decisions *University of Groningen, the Netherlands*
Fall 2015	Teaching Assistant (General and Russian Semantics) *Higher School of Economics, School of Linguistics, Moscow, Russia*
Spring 2014	Teaching Assistant (Academic Writing in Russian) *Higher School of Economics, Faculty of Philology, Moscow, Russia*
2012/2013	Teaching Assistant (German language for Beginner levels) *Higher School of Economics, Faculty of Philology, Moscow, Russia*

Other work experience

Jan 2019	Co-organizer of Shared Task on Cross-Genre Gender Prediction in Dutch (GxG-CLIN29) co-located with the 29th Conference on Computational Linguistics in the Netherlands *Groningen, the Netherlands*
Jan 2019	Member of the organizing commitee for Computational Linguistics in the Netherlands conference 2019 (CLIN29) *Groningen, the Netherlands*
Jun 2018	Member of the organizing commitee for the 39th International Linguistics Conference TABU Dag 2018 *Groningen, the Netherlands*
Fall 2014	Internship in Information Retrieval *Yandex Search Engine, Moscow, Russia*
Aug 2014	Internship in Teaching Russian as a Foreign Language *Russian Language Centre, Higher School of Economics, Moscow*
2013-2015	Russian-English, English-Russian translator, organizer of charity events *Association of Social Marketing, Moscow, Russia*
Jul 2013	Internship in Teaching Russian as a Foreign Language *Russian Language Centre, Higher School of Economics, Moscow*
Mar 2012	Assistant to prof. Johanna Nichols *Higher School of Economics, Faculty of Philology, Moscow, Russia*

Projects (selected)

2022	Contributing to building the Typology of Legal Technologies: `https://publications.cohubicol.com/typology/`
2020	Building JURI SAYS: `jurisays.com`, an online platform for forecasting future decisions of the European Court of Human Rights
2018	Participation in VarDial Evaluating Campaign 2018: Discriminating between Similar Languages Shared Task co-located with COLING 2018 *(with Martin Kroon and Barbara Plank)* **(4th place)**
2017	Participation in PAN Author Profiling shared task at CLEF 2017 *(with Angelo Basile, Gareth Dwyer, Josine Rawee, Hessel Haagsma and Malvina Nissim)* **(1st place)**
2017	Participation in VarDial Evaluating Campaign 2017: Discriminating between Similar Languages Shared Task co-located with EACL 2017 *(with Martin Kroon and Barbara Plank)* **(2nd place)**
2016	Udmurt→Russian Machine Translation with Apertium at Summer School in Rule-Based Machine Translation at Universitat d'Alacant, Alacant, Spain
2016	Participation in PAN Author Profiling shared task at CLEF 2016 *(with Mart Busger op Vollenbroek, Talvany Carlotto, Tim Kreutz, Chris Pool, Johannes Bjerva, Hessel Haagsma and Malvina Nissim)* **(1st place)**
2014-2015	Corpora of Texts Generated by Multilingual Speakers: Development of Principles and Creation of Prototype
Aug 2014	Development of an open-source speech recognition system for Icelandic at "Coding for Language Communities" Summer School at CIDLeS, Minde, Portugal
2014-2015	Development of a Speech Recognition System for Russian Speakers with Distinctive Speech Impairments
2013-2014	Development of an Online Annotated Corpus of Udmurt Language `http://web-corpora.net/UdmurtCorpus/search/` **(currently used for Yandex.translate system)**
2012-2013	Automatic Extraction of Grammar-Related Data from Texts for the Corpus of Modern Greek
2011-2012	Developing a Prototype of a Chat-Bot Emulating a Representative of Geek Subculture (LeoNerd)

Scholarships and Awards (selected)

Sep '19	Participated in successful application for ERC Starter Grant for EVICT project (PI: prof. Michel Vols)
Oct '17-Mar '22	**4,5-year scholarship** towards pursuing a doctorate degree at University of Groningen, the Netherlands
Sep '17	**Best Paper Award** at CLEF 2017 (with Hessel Haagsma & Malvina Nissim)
Apr '17	**Winner of CLEF 17 PAN Author Profiling Shared Task** (with Angelo Basile, Gareth Dwyer, Josine Rawee, Hessel Haagsma, and Malvina Nissim)
Jul '16	**Full Scholarship** for working on the Udmurt→Russian Machine Translation system using Apertium at Summer School in Rule-Based Machine Translation at Universitat d'Alacant, Alacant, Spain
Apr '16	**Winner of CLEF 16 PAN Author Profiling Shared Task** (with Mart Busger op Vollenbroek, Talvany Carlotto, Tim Kreutz, Chris Pool, Johannes Bjerva, Hessel Haagsma and Malvina Nissim)
Sep '15-Aug '17	**Erasmus+ Category A Scholarship** for Language and Communication Technologies Joint Degree
Sep '13-Jan '14	**Full financial support from National Research University Higher School of Economics** for an exchange semester at University of Bern, Switzerland
Sep '12-Jun '13	**Increased scholarship of the Government of the Russian Federation** awarded for public and creative cultural activities