

Seid Muhie Yimam (Ph.D.)

📍 Albert-Einstein-Ring 8-10 22761, Hamburg, Germany

✉ seid.muhie.yimam@uni-hamburg.de

☎ +49 40 42883 2383

🐦 @seyyaw

📘 @Seyaw

📄 [seid.muhie.yimam](https://www.linkedin.com/in/seyaw)

🌐 <https://www.linkedin.com/in/seyaw>

🌐 <https://www.hcds.uni-hamburg.de/en/hcds/head-hcds/yimam.html>





Employment History

- 04/2022 – current 📌 **Technical Lead**, Hub of Computing and Data Science (HCDS), Universität Hamburg, Germany.
Leading and consulting research on digital humanities, big data processing of textual content, teaching NLP and Data science courses, supervising students on interdisciplinary AI and data science research topics.
- 01/2020 – 03/2022 📌 **Research Associate**, LT Group, Universität Hamburg, Germany.
Leading research in language technology.
- 10/2016 – 12/2019 📌 **Postdoctoral researcher**, LT Group, Universität Hamburg, Germany.
Writing project proposals, student supervision, shared task and workshop management, low-resource research projects
- 10/2016 – 12/2019 📌 **Doctoral researcher**, Language Technology Group, Universität Hamburg, Germany. Research on semantic writing aid (Par4Sem/Sim, Par4Acad), dissertation, student supervision, Teaching (Natural Language Processing and the Web (NLP4Web) and Statistical Methods of Language Technology (SMoLT))
- 09/2012 – 09/2016 📌 **Scientific software engineer + doctoral researcher**, Language Technology Lab + UKP lab, Technische Universität Darmstadt, Germany.
Software development (WebAnno, GermaNER, Network of the Day, New/s/leak), teaching assistant for NLP4Web, Scientific activities (review, conference, workshop attendance)
- 09/2011 – 08/2012 📌 **Semantic web software developer**, OKKAM srl, Italy.
Work on the project of 'Entity Naming System (ENS)' for managing and interlinking entities in networks of data; Integration of semantic technologies at OKKAM which include RDF, SPARQL, and RDF stores; Develop reconciliation services into 'Open Refine RTL' tool for the OKKAM ENS; Develop the full stack of WebID using Java.
- 12/2009 – 09/2010 📌 **Lecturer**, Adama University, Adama Science and Technology University, Ethiopia.
Teach advanced programming (Java) and advanced networking courses, participate in the development and evaluation of faculty and department curriculum; participate in requirement analysis and evaluation of different campus-based software applications.
- 08/2004 – 01/2007 📌 **Lecturer**, Haramaya University, Ethiopia.
Provide basic and advanced course lectures for undergraduate students and coordinate the different activities in the Department of Information Technology.






Education

- 2014 – 2019 📌 **Department of Computer Science, Universität Hamburg, Germany.**
Dr. rer. nat. in Computer Science (Informatik)
Dissertation: Adaptive Approaches to Natural Language Processing in Annotation and Application
- 2010 – 2011 📌 **Department of Computer Science, University of Trento, Italy.**
Advanced master in Human language technologies and interfaces
Thesis with internship: Sentence based sentiment analysis from Telecom Italia texts








Education (continued)

- 2007 – 2009  **Department of Computer Science, Addis Ababa University, Ethiopia.**
M.Sc. in Computer Science
Thesis: TETEQ - Amharic Question Answering System for Factoid Questions
- 1999 – 2004  **Department of Computer Science, Addis Ababa University, Ethiopia.**
B.Sc. in Computer Science
Project: Allocations of students to higher education in Ethiopian universities, Ministry of Education, Ethiopia

Research Directions



- Social NLP  Hate speech, Sentiment analysis, fake news and misinformation
- Semantics  Text similarity, complex word identification, word embedding
- ML/AI  Machine learning for NLP applications, adaptive learning
- Annotation  Web-based annotation tool (WebAnno), Social network-based annotation (ASAB), crowdsourcing (MTurk + Toloka)
- Low-resource  Research for EthioNLP, collect datasets, build models

Teaching




- WS14-2022  Practical Classes : [Natural Language Processing and the Web](#)
- SS18-2021  Supervision: [Masterproject Web Interfaces for Language Processing Systems](#)
 Practice Classes: [Statistical Methods of Language Technology](#)
- WS2020  Lecture : [Natural Language Processing and the Web](#)
- WS2016/17  Practical Classes : [Übung Algorithmen und Datenstrukturen](#)
- SS 2014  Seminar coordinator : [Knowledge Engineering for Question Answering Systems, Seminars](#)
- 2004-2009  Lecture at the Department of Computer Science, **Haramaya University** and **Adama Science and Technology University**, Ethiopia. **Courses I taught:** Software Engineering, Object Oriented Programming, Advanced Programming (Java), Database Management System, Data Structure, Operating System.

Research Projects





Projects

- Google exploreCSR  I have secured the Google exploreCSR grant (2023) for a project entitled **Empowering Women in AI and NLP: Advancing Opportunities in Ethiopian Universities**. The project award amounts **\$18k**. Furthermore, I have received the TensorFlow Addons fund of amount **\$10k**.
- Google Inclusive AI  To further foster inclusive AI research, Prof. Chris Biemann, from the House of Computing and Data Science (HCDS), Universität Hamburg, Germany, in collaboration with Prof. Adem Chanie Ali from the Department of Journalism and Communications at Bahir Dar University, Ethiopia, has been awarded a **\$60k** grant from the Google Award for Inclusion Research Program for a project titled **AI-driven Monitoring of Attitude Polarization in Conflict-Affected Countries for Inclusive Peace Process and Women's Empowerment**. I operational manage this project.







Research Projects (continued)

- SAINT  Universität Hamburg provides extensive funding options for activities, mobility, and projects with international partners. In August 2021, the University of Leeds and Universität Hamburg established a Memorandum of Understanding to strengthen their cooperation in research and teaching. Among the first to receive the funding are Dr. Seid Muhie Yimam from UHH and Dr. Serge Sharoff from UL with the joint project entitled "SAINT - Sentiment Analysis INTerpretability". The project runs from January 1, 2023, till the end of May 2024, with a total amount of **10k€**.
- ARDIAS  AI-Enhanced Research Project Management, Discovery and Advisory System (ARDIAS) is a project funded by UHH for the **Ideas and Venture Fund**, which is part of the Excellence Strategy of the Federal and State Governments and facilitates short-term funding for especially promising research approaches. The project runs from September 1, 2022, till the end of September 2023, with a total amount of **50k€**.
- DAAD-FFM  **DAAD Fact-Finding Mission:** Fact-finding mission between Universität Hamburg and Ethiopian Universities: Capacity building and structural improvement for the establishment and improvement of AI courses in Computer Science and related disciplines, *interrupted* due to COVID-19 pandemic.

Projects I have Participated



- DIVID-DJ  Data Extraction and Interactive Visualization of Unexplored Textual Datasets for Investigative Data-Driven Journalism (Volkswagenstiftung)
- SEMSCH  Semantic Methods for Computer-supported Writing Aids (DFG)
- NoD  Network of the Day, Hochschulwettbewerb 2014
- CLARIN-D  Implementation of a web-based annotation platform for linguistic annotations (BMBF)

Skills





- Languages  Strong reading, writing and speaking competencies for **English** and **Amharic**. Intermediate level of speaking and reading for **German**
- Coding  Python | Java | Flutter | R | SQL | XML/XSL | JSON |
- ML & Tools  PyTorch, scikit-learn, spaCy, NLTK, WebAnno,
- Databases  MySQL | PostgreSQL | SPARQL | Elasticsearch
- Web Dev  Flask | FastAPI | HTML | CSS | JavaScript | Apache/Tomcat Web Server | WordPress
- Misc.  Academic research | teaching | training | consultation | Data science

Miscellaneous Experience




Awards and Scholarships

- 2022  **Oracle for Research Grant:** Oracle for Research Grant for the project entitled **AfriHate: A Multilingual Hate and Offensive Corpus for Conflict-Affected African Countries**
- 2021  **Merit Award:** GSCL Award for **best Doctoral Thesis** in the memory of Wolfgang Hoepfner, 2020.



Miscellaneous Experience (continued)

- 2020  **ENeTCollect Crowdfest 2020 travel grant**, Task 5, Preparing a shared task for aggregating crowdsourced boolean answers, University Of Coimbra, Coimbra, Portugal
- 2019  **LT4ALL Grant**, to the conference 'Language Technology for All (LT4All)': Enabling Linguistic Diversity and Multilingualism Worldwide, Paris, UNESCO Headquarters (4-6 December 2019)
- 2017  **Best Paper Award - Impact on Society**, [Paper - Yimam et al. 2017](#), by Fraunhofer Institute for Computer Graphics Research IGD
- 2010  **Full Scholarship**, International Master in Human Language Technology and Interfaces, Uni. Trento, Italy.




Conference and Workshop Related

- Scientific Committee  [ACL](#) | [EMNLP](#) | [NAACL](#) | [CoNLL](#) | [AAAI](#) | [ACL Rolling Review](#) | [EACL](#) | [LREC](#) | [RANLP](#) | [SemEval](#) | [BEA](#) | [NLDB](#) | [MWE](#) | [KONVENS](#) | [AfricaNLP](#) | [ICT4DA](#)
- Program Committee  The 3rd International Conference on ICT4DA
- Workshop Organizer  [SemEval 2024 Task 1: Semantic Textual Relatedness \(STR\)](#), [AfriSenti-SemEval Shared Task 12](#), [Multimodal and Multilingual Hate Speech Detection \(MMHS\) 2021](#) | [Complex Word Identification \(CWI\) Shared Task 2018](#)

Researcher Visit

- Sep 2019  IIT Kharagpur, India, Guest researcher
- Jan-Mar 2019  City University of Hong Kong, Hong Kong, Guest researcher
- Oct 2015  Center for Language Technology, University of Copenhagen, Research visit

Extracurricular Activities

- EUDiF  **Diaspora Professionals 4 Development**, Piloting digital learning on agri-food business and environmental informatics in Ethiopia, ongoing
- EthioNLP  Initiator of [EthioNLP](#), research on Ethiopian Languages NLP
- Masakhane  Member of [Masakhane](#), a grassroots NLP community for Africa, by Africans

Research Publications

Monographs

- **Yimam, S. M.** (2019): Adaptive Approaches to Natural Language Processing in Annotation and Application. Ph.D Thesis, University of Hamburg
- **Yimam, S. M.** (2009): TETEQ - AMHARIC QUESTION ANSWERING SYSTEM FOR FACTOID QUESTIONS. Master Thesis, Addis Ababa University

Book Chapters

- Biemann, C., Bontcheva, K., Eckart de Castilho, R., Gurevych, I., **Yimam, S. M.** (2017): Collaborative Web-based Tools for Multi-layer Text Annotation. In: N. Ide and J. Pustejovsky (Eds.): Handbook of Linguistic Annotation, Springer

Journal Publications

- Ayalew Kassahun and **Yimam, S. M.** Yonas Seifu Muanenda and Beshir Melkaw Ali and Seleshi Getahun Yalew(2024): Uncovering the priorities of scientific research on sustainable development goals: A case study in Ethiopia, Sustainable Development, published by ERP Environment and John Wiley & Sons Ltd. 2024;1–26 , DOI: 10.1002/sd.3020
- Jana, A., Venkatesh, G., **Yimam, S. M.**, and Biemann, C., Hypernymy Detection for Low-Resource Languages: A Study for Hindi, Bengali, and Amharic, ACM Transactions on Asian and Low-Resource Language Information Processing (2022).
- **Yimam, S. M.**, Ayele, A.A.; Venkatesh, G.; Gashaw I.; Biemann C. (2021): Introducing Various Semantic Models for Amharic: Experimentation and Evaluation with Multiple Tasks and Datasets. Future Internet 2021, 13, 275. <https://doi.org/10.3390/fi13110275>
- **Yimam, S. M.**, Biemann, C., Majnarić, L., Šabanović, Š., Holzinger, A. (2016): An adaptive annotation approach for biomedical entity and relation recognition. Brain Informatics, (online first), 10.1007/s40708-016-0036-4

Conference Proceedings

- T. D. Belay, I. A. Azime, A. A. Ayele, G. Sidorov, D. Klakow, P. Slusallek, O. Kolesnikova **Yimam, S. M.** (2025): Evaluating the Capabilities of Large Language Models for Multi-label Emotion Understanding, in Proceedings of the 31st International Conference on Computational Linguistics: 3523-3540, 18 S., Abu Dhabi, UAE: Association for Computational Linguistics
- D. Dementieva, N. Babakov, A. Ronen, A. A. Ayele, N. Rizwan, F. Schneider, X. Wang, **Yimam, S. M.**, D. A. Moskovskiy, E. Stakovskii, E. Kaufman, A. Elnagar, A. Mukherjee, A. Panchenko (2025): Multilingual and Explainable Text Detoxification with Parallel Corpora, in: Proceedings of the 31st International Conference on Computational Linguistics: 7998-8025, 28 S., Abu Dhabi, UAE: Association for Computational Linguistics
- Myung, J., Lee, N., Zhou, Y., Jin, J., Putri, R. A., Antypas, D., Borkakoty, H., Kim, E., Perez-Almendros, C., Ayele, A. A., Gutierrez Basulto, V., Ibanez-Garcia, Y., Lee, H., Muhammad, S. H., Park, K., Rzaev, A. S., White, N., and **Yimam, S. M.**, Pilehvar, M. T., Ousidhoum, N., et al. (2024): BLEN: A Benchmark for LLMs on Everyday Knowledge in Diverse Cultures and Languages. NeurIPS 2024 Track Datasets and Benchmarks Poster. (Published: 26 Sept 2024, Last Modified: 16 Jan 2025, 2 additional authors not shown).
- Sewunetie, W., Beza, A., Abebe, H., Abuhay, T. M., Admass, W., Hassen, H., Haile, T., Hailemariam, H., Debebe, L., Moges, N., Bekele, N., Tilahun, S. L., Berta, M., Mammo, M., **Yimam, S. M.**, and Laszlo, K. (2024): Large Language Models for Sexual, Reproductive, and Maternal Health Rights. 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI).

Research Publications (continued)

- Azime, I. A., Tonja, A. L., Belay, T. D., Fuge, M. Y., Wassie, A. K., Jada, E. S., Chanie, Y., Sewunetie, W. T., and **Yimam, S. M.** (2024): Walia-LLM: Enhancing Amharic-LLaMA by Integrating Task-Specific and Generative Datasets. EMNLP 2024, Miami, Florida, USA.
- Abinew Ali Ayele, Nikolay Babakov, Janek Bevendorff, Xavier Bonet Casals, Berta Chulvi, Daryna Dementieva, Ashaf Elnagar, Dayne Freitag, Maik Fröbe, Damir Korenčić, Maximilian Mayerl, Daniil Moskovskiy, Animesh Mukherjee, Alexander Panchenko, Martin Potthast, Francisco Rangel, Naqee Rizwan, Paolo Rosso, Florian Schneider, Alisa Smirnova, Efstathios Stamatatos, Elisei Stakovskii, Benno Stein, Mariona Taulé, Dmitry Ustalov, Xintong Wang, Matti Wiegmann, **Seid Muhie Yimam**, Eva Zangerle. (authors are listed in alphabetical order) (2024): Overview of PAN 2024: Multi-Author Writing Style Analysis, Multilingual Text Detoxification, Oppositional Thinking Analysis, and Generative AI Authorship Verification, 15th International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF 2024), Grenoble, France.
- Nedjma Ousidhoum, Shamsuddeen Hassan Muhammad, Mohamed Abdalla, Idris Abdulmumin, Ibrahim Said Ahmad, Sanchit Ahuja, Alham Fikri Aji, Vladimir Araujo, Abinew Ali Ayele, Pavan Baswani, Meriem Beloucif, Chris Biemann, Sofia Bourhim, Christine de Kock, Genet Shanko Dekebo, Oumaima Hourrane, Gopichand Kanumolu, Lokesh Madasu, Samuel Rutunda, Manish Shrivastava, Tamar Solorio, Nirmal Surange, Hailegnaw Getaneh Tilaye, Krishnapriya Vishnubhotla, Genta Indra Winata, **Yimam, S. M.**, Saif M. Mohammad (2024): SemRel2024: A Collection of Semantic Textual Relatedness Datasets for 13 Languages. Findings of the Association for Computational Linguistics (ACL 2024), Bangkok Thailand.
- Atnafu Lambebo Tonja, Israel Abebe Azime, Tadesse Destaw Belay, Mesay Gemedo Yigezu, Moges Ahmed Mehamed, Abinew Ali Ayele, Ebrahim Chekol Jibril, Michael Melese Woldeyohannis, Olga Kolesnikova, Philipp Slusallek, Dietrich Klakow and **Yimam, S. M.**: EthioLLM: Multilingual Large Language Models for Ethiopian Languages with Task Evaluation
- Ayele A.A., **Yimam, S. M.**, Belay T.D., Asfaw T. and Biemann C. (2023): Exploring Amharic Hate Speech Data Collection and Classification Approaches, in the 14th Conference RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING, Varna, Bulgaria
- Ayele A.A., Dinter S., **Yimam, S. M.** and Biemann C. (2023): Multilingual Racial Hate Speech Detection Using Transfer Learning, in the 14th Conference RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING, Varna, Bulgaria
- Schneider, F., **Yimam, S. M.**, Petersen-Frey, F., Biemann, C., von Nordheim, G., Kleinen-von Königslöw, K., (2023): CodeAnno: Extending WebAnno with Hierarchical Document Level Annotation and Automation. The 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023), System Demonstrations Track, Dubrovnik, Croatia
- Belay T. D., Tonja A. L., Kolesnikova O., **Yimam, S. M.**, Ayele A. A., Haile S.B., Sidorov G., Gelbukh A. (2022): The Effect of Normalization for Bi-directional Amharic-English Neural Machine Translation, International Conference on Information and Communication Technology for Development for Africa (ICT4DA 2022), Bahir Dar, Ethiopia
- Ayele A. A., Belay T. D., Asfaw, T. T., Dinter S., **Yimam, S. M.** (2022): The 5Js in Ethiopia: Amharic Hate Speech Data Annotation Using Toloka Crowdsourcing Platform, International Conference on Information and Communication Technology for Development for Africa (ICT4DA 2022), Bahir Dar.
- Beloucif M., **Yimam, S. M.**, Stahlhacke S. and Biemann C. (2022): Elvis vs. M. Jackson: Who has More Albums? Classification and Identification of Elements in Comparative Question. In the 2022 International Conference on Language Resources and Evaluation (LREC 2022), Marseille, France.
- Belay, T. D., Ayele, A.A., Gelaye, G., **Yimam, S. M.**, Biemann, C. (2021): Impacts of Homophone Normalization on Semantic Models for Amharic. Proceedings of the Third International Conference on ICT for Development for Africa (ICT4DA 2021), Bahir Dar, Ethiopia

Research Publications (continued)

- von Boguszewski, N., Moin, S., Bhowmick, A., **Yimam, S. M.**, Biemann, C. (2021): How Hateful are Movies? A Study and Prediction on Movie Subtitles. Proceedings of KONVENS, Düsseldorf, Germany
- Wiechmann, M., **Yimam, S. M.**, Biemann, C. (2021): ActiveAnno: General-Purpose Document-Level Annotation Tool with Active Learning Integration. The 2021 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies - System Demonstrations, Mexico City, Mexico (online)
- Gooding, S., Kochmar, E., **Yimam, S. M.**, Biemann, C. (2021): Word Complexity is in the Eye of the Beholder. The 2021 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT), Mexico City, Mexico.
- Mathew, B., Saha, P., **Yimam, S. M.**, Biemann, C., Goyal, P., Mukherjee, A. (2021): HateXplain: A Benchmark Dataset for Explainable Hate Speech Detection. Proceedings of AAAI-21, Virtual Conference.
- Haase, C., Anwar, S., **Yimam, S. M.**, Friedrich, A., Biemann, C. (2021): SCoT: Sense Clustering over Time: a tool for the analysis of lexical change. The 2021 Conference of the European Chapter of the Association for Computational Linguistics - System Demonstrations. Kyiv, Ukraine (Online)
- **Yimam, S. M.**, Alemayehu H. M., Ayele A. A. and Biemann C. (2020): Exploring Amharic Sentiment Analysis from Social Media Texts: Building Annotation Tools and Classification Models. The 28th International Conference on Computational Linguistics (COLING 2020), Barcelona, Spain (poster)
- **Yimam, S. M.**, Venkatesh, G., Lee, J. Biemann, C. (2020): Automatic Compilation of Resources for Academic Writing and Evaluating with Informal Word Identification and Paraphrasing System, The International Conference on Language Resources and Evaluation (LREC 2020), Marseille, France.
- Wiedemann G., **Yimam, S. M.**, and Biemann C. (2018) : A Multilingual Information Extraction Pipeline for Investigative Journalism. In Proceedings of 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018). Brussels, Belgium
- **Yimam, S. M.**, Biemann C. (2018): Demonstrating Par4Sem - A Semantic Writing Aid with Adaptive Paraphrasing. In Proceedings of 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018). Brussels, Belgium .
- Wiedemann G., **Yimam, S. M.**, and Biemann C. (2018) : New/s/leak 2.0 – Multilingual Information Extraction and Visualization for Investigative Journalism. In: Proceedings of the 10th International Conference on Social Informatics (SocInfo 2018). St.Petersburg, Russia
- **Yimam, S. M.**, Biemann C. (2018): Par4Sim – Adaptive Paraphrasing for Text Simplification. In Proceedings of The 27th International Conference on Computational Linguistics (COLING 2018). Santa Fe, New-Mexico, USA .
- **Yimam, S. M.**, Štajner S., Riedl M., Biemann C. (2017): CWIG3G2 - Complex Word Identification Task across Three Text Genres and Two User Groups. In Proceedings of The 8th International Joint Conference on Natural Language Processing (IJCNLP 2017). Taipei, Taiwan
- **Yimam, S. M.**, Štajner S., Riedl Martin, Biemann C. (2017): Multilingual and Cross-Lingual Complex Word Identification. In Proceedings of The 2017 International Conference on Recent Advances in Natural Language Processing (RANLP). Varna, Bulgaria
- **Yimam, S. M.**, Ulrich, H., von Landesberger, T., Rosenbach, M., Regneri, M., Panchenko, A., Lehmann, F., Fahrer, U., Biemann, C. and Ballweg, K. (2016): new/s/leak – Information Extraction and Visualization for Investigative Data Journalists. ACL 2016 Demo Session, Berlin, Germany
- **Yimam, S. M.**, Biemann, C., Majnaric, L., Šabanović, Š., Holzinger, A. (2015): Interactive and Iterative Annotation for Biomedical Entity Recognition, International Conference on Brain Informatics and Health (BIH'15), London, UK
- Benikova, D., **Yimam, S. M.**, Biemann C. (2015). GermaNER: Free Open German Named Entity Recognition Tool. In: Proceedings of the GSCL 2015. Essen, Germany

Research Publications (continued)

- **Yimam, S. M.**, Eckart de Castilho, R., Gurevych, I., Biemann C. (2014): Automatic Annotation Suggestions and Custom Annotation Layers in WebAnno. Proceedings of ACL 2014 System Demonstrations, Baltimore, MD, USA
- **Yimam, S. M.**, Gurevych, I., Eckart de Castilho, R., and Biemann C. (2013): WebAnno: A Flexible, Web-based and Visually Supported System for Distributed Annotations. Proceedings of ACL-2013, demo session, Sofia, Bulgaria

Workshop Proceedings

- Sewunetie, W., Tonja, A., Belay, T., Nigatu, H. H., Gebremeskel, G., Mossie, Z., Seid, H., and **Yimam, S.** (2024): Gender Bias Evaluation in Machine Translation for Amharic, Tigrigna, and Afaan Oromoo. Proceedings of the 2nd International Workshop on Gender-Inclusive Translation Technologies, Sheffield, United Kingdom. European Association for Machine Translation (EAMT). *[Online]. Available:*
- Daryna Dementieva, Daniil Moskovskiy, Nikolay Babakov, Abinew Ali Ayele, Naqee Rizwan, Florian Schneider, Xintong Wang, **Seid Muhie Yimam**, Dmitry Ustalov, Elisei Stakovskii, Alisa Smirnova, A Elnagar, Animesh Mukherjee, Alexander Panchenko. (2024): Overview of the Multilingual Text Detoxification Task at PAN 2024, Working Notes of CLEF 2024, Grenoble, France *[Online]. Available:*
- Melese Ayichlie Jigar, Abinew Ali Ayele, **Yimam, S. M.** and Chris Biemann (2024): Detecting Hate Speech in Amharic Using Multimodal Analysis of Social Media Memes. Proceedings of The Fourth Workshop on Threat, Aggression & Cyberbullying, Torino, Italy
- Abinew Ali Ayele, Esubalew Alemneh Jalew, Adem Chanie Ali, **Yimam, S. M.**, Chris Biemann (2024): Exploring Boundaries and Intensities in Offensive and Hate Speech: Unveiling the Complex Spectrum of Social Media Discourse. Proceedings of The Fourth Workshop on Threat, Aggression & Cyberbullying, Torino, Italy
- Shamsuddeen Hassan Muhammad, Idris Abdulmumin, **Yimam, S. M.**, David Ifeoluwa Adelani, Ibrahim Sa'id Ahmad, Nedjma Ousidhoum, Abinew Ayele, Saif M Mohammad, Meriem BELOUCIF, Sebastian Ruder. SemEval-2023 Task 12: Sentiment Analysis for African Languages (AfriSenti-SemEval): arXiv preprint arXiv:2304.06845. 2023.
- Shamsuddeen Hassan Muhammad, Idris Abdulmumin, Abinew Ali Ayele, Nedjma Ousidhoum, David Ifeoluwa Adelani, **Yimam, S. M.**, Ibrahim Sa'id Ahmad, Meriem BELOUCIF, Saif Mohammad, Sebastian Ruder, Oumaima Hourrane, Pavel Brazdil, Felermine Dário Mário António Ali, Davis Davis, Salomey Osei, Bello Shehu Bello, Falalu Ibrahim, Tajuddeen Gwadabe, Samuel Rutunda, Tadesse Belay, Wendimu Baye Messelle, Hailu Beshada Balcha, Sisay Adugna Chala, Hagos Tesfahun Gebremichael, Bernard Opoku, Steven Arthur. Afrisenti: A Twitter sentiment analysis benchmark for African languages: arXiv preprint arXiv:2302.08956. 2023.
- Tonja A. L., Belay T. D., Azime I. A., Ayele A. A., Mehamed M. A., Kolesnikova O., **Yimam, S. M.** (2023): Natural Language Processing in Ethiopian Languages: Current State, Challenges, and Opportunities, In the fourth workshop on Resources for African Indigenous Languages (RAIL) at EAACL2023, Dubrovnik, Croatia
- Banerjee D., **Yimam, S. M.**, Awale S. and Biemann C (2023), ARDIAS: AI-Enhanced Research Management, Discovery, and Advisory System, The AAAI-23 Workshop on Scientific Document Understanding at the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23), Washington, DC, USA.
- Ayele A. A., Belay T. D., **Yimam, S. M.**, Dinter S., Asfaw, T. T., Biemann C. (2022): The 5Js in Ethiopia: Amharic Hate Speech Data Annotation Using Toloka Crowdsourcing Platform, The Sixth Widening Natural Language Processing Workshop (WiNLP 2022) in conjunction with EMNLP 2022, Abu Dhabi, UAE

Research Publications (continued)

- Belay T. D., Tonja A. L., Kolesnikova O., **Yimam, S. M.**, Ayele A. A., Haile S.B., Sidorov G., Gelbukh A. (2022): The Effect of Normalization for Bi-directional Amharic-English Neural Machine Translation, The Sixth Widening Natural Language Processing Workshop (WiNLP 2022) in conjunction with EMNLP 2022, Abu Dhabi, UAE
- Belay, T. D., **Yimam, S. M.**, Ayele, A. A., and Biemann, C. (2022): Question Answering Classification for Amharic Social Media Community Based Questions, The 1st Annual Meeting of the ELRA/ISCA Special Interest Group on Under-Resourced Languages (SIGUL 2022), Marseille, France.
- Belay, T. D., Ayele A.A. and **Yimam, S. M.** (2021): The Development of Pre-processing Tools and Pre-trained Embedding Models for Amharic. Proceedings of The fifth WiNLP (“Widening NLP”) Workshop held in conjunction with EMNLP 2021, Punta Cana, Dominican Republic. .
- V (2021): MasakhaNER: Named Entity Recognition for African Languages. Transactions of the Association for Computational Linguistics, 2021, 9 1116–1131.
- Wiedemann G, **Yimam, S. M.**, Biemann C. (2020): UHH-LT at SemEval-2020 Task 12: Fine-Tuning of Pre-Trained Transformer Networks for Offensive Language Detection. Proceedings of The 14th International Workshop on Semantic Evaluation (SemEval), Barcelona, Spain. (poster).
- **Yimam, S. M.**, Ayele, A. A., Biemann C. (2019): Analysis of the Ethiopic Twitter Dataset for Abusive Speech in Amharic. In Proceedings of International Conference On Language Technologies For All: Enabling Linguistic Diversity And Multilingualism Worldwide (LT4ALL 2019). Paris, France p. 210-214 .
- **Yimam, S. M.**, Biemann, C., Malmasi, S., Paetzold, G.H., Speica, L., Štajner, S., Tack, A., Zampieri, M., (2018): A Report on the Complex Word Identification Shared Task 2018. Proceedings of the 13th Workshop on Innovative Use of NLP for Building Educational Applications, New Orleans, LA, USA
- **Yimam, S. M.**, Remus S., Panchenko A., Holzinger A., Biemann C. (2017): Entity-Centric Information Access with the Human-in-the-Loop for the Biomedical Domains. Biomedical NLP Workshop associated with RANLP 2017. Varna, Bulgaria
- Müller, M., Ballweg, K. von Landesberger, T., **Yimam, S. M.**, Fahrer, U., Biemann, C., Rosenbach, M., Regneri, M., Ulrich, H. (2017). Guidance for Multi-Type Entity Graphs from Text Collections. EuroVis Workshop on Visual Analytics 2017, Barcelona, Spain
- Nandi, T., Biemann, C., **Yimam, S. M.**, Gupta, Deepak., Kohail, S., Ekbal, A., Bhattacharyya, Pushpak. (2017): IT-UHH at SemEval-2017 Task 3: Exploring Multiple Features for Community Question Answering and Implicit Dialogue Identification, In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval 2017), Vancouver, Canada.
- Eckart de Castilho, R. Mújdricza-Maydt, E., **Yimam, S. M.**, Hartmann, S., Gurevych, I., Frank, A. and Biemann, C. (2016): A Web-based Tool for the Integrated Annotation of Semantic and Syntactic Structures. Proceedings of the COLING workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH), Osaka, Japan
- Ballweg K., Zouhar F., Wilhelmi-Dworski P., von Landesberger T., Fahrer U., Panchenko A., **Yimam, S. M.** Biemann C., Regneri M., Ulrich H. (2016) new/s/leak – A Tool for Visual Exploration of Large Text Document Collections in the Journalistic Domain, Baltimore, MD, USA,
- **Yimam, S. M.**, Martínez Alonso, H., Riedl M. and Biemann, C. (2016): Learning Paraphrasing for Multiword Expressions. The 12th Workshop on Multiword Expressions (MWE 2016), co-located with ACL 2016, Berlin, Germany
- **Yimam, S. M.** (2015): Narrowing the Loop: Integration of Resources and Linguistic Dataset Development with Interactive Machine Learning. NAACL 2015 Student Research Workshop, p. 88–95, Denver, Colorado
- Eckart de Castilho, R., Biemann, C., Gurevych, I., **Yimam, S. M.** (2014): WebAnno: a flexible, web-based annotation tool for CLARIN. CLARIN Annual Conference 2014, Soesterberg, The Netherlands

Research Publications (continued)

- Benikova, D., Fahrer, U., Gabriel, A., Kaufmann, M., **Yimam, S. M.**, von Landesberger, T., Biemann, C. (2014): Network of the Day: Aggregating and Visualizing Entity Networks from Online Sources. KONVENS 2014 Workshop proceedings: NLP4CMC, pp. 48-52, Hildesheim, Germany
- **Yimam, S. M.**, Libse, M. (2009): TETEYEQ: Amharic Question Answering For Factoid Questions, SE-PLNo9. SALTMIL workshop - Information Retrieval and Information Extraction for Less Resourced Languages (IE-IR-LRL), p. 17-25

References

Available on Request