

# David Ifeoluwa ADELANI

## PERSONAL DATA

---

PLACE OF BIRTH: Abeokuta, Nigeria  
ADDRESS: Montréal, Canada  
EMAIL: [david.adelani@mcgill.ca](mailto:david.adelani@mcgill.ca) or [david.adelani@mila.quebec](mailto:david.adelani@mila.quebec)  
WEBSITE: [dadelani.github.io](https://dadelani.github.io) or [On Google Scholar](#)

## EDUCATION

---

JAN 2015 - JUN 2023 Preparatory Phase (2015-2018) and Dissertation Phase of Doctoral Studies  
**Spoken Language Systems Group, Saarland Informatics Campus, Germany**  
Thesis: "Natural Language Processing for African Languages"  
ADVISOR: Prof. Dr. Dietrich Klakow  
GRADE: summa cum laude

JUL 2013 - DEC 2014 Master of Science in COMPUTER SCIENCE  
**African University of Science and Technology (AUST), Abuja, Nigeria**  
Thesis: "A DEVS-Based ANN Training and Prediction Platform" | Advisor: M.K. TRAORE  
GPA: 3.91/4

OCT 2006 - OCT 2010 Bachelor of Science in COMPUTER SCIENCE  
**Federal University of Agriculture, Abeokuta, Nigeria**  
Thesis: "Securing E-voting with EC-ELGAMAL" | Advisor: A.S. SODIYA  
GPA: 4.42/5

## WORK EXPERIENCE

---

AUG 2024 - Assistant Professor (Tenure-track)  
*School of Computer Science, Faculty of Science, McGill University*  
Core Academic Member at Mila & Canada CIFAR AI Chair

SEP 2022 - PRESENT Senior Research Fellow (DeepMind Academic Fellow)  
*Department of Computer Science, University College London, UK.*  
Supported by [DeepMind](#)

JAN 2019 - JUN 2023 PhD Researcher (Group: Spoken Language Systems Group)  
*Department of Language Science & Technology, Saarland University*  
Supported by the EU H2020 [COMPRISE](#) project

## RESEARCH INTERESTS

---

Multilingual Natural language processing (NLP)  
NLP for low-resource languages  
Large language model adaptation

## TEACHING EXPERIENCE

---

TEACHING Multilingual Representation Learning – Winter 25  
McGill University, Canada

CO-TEACHING Natural language Processing – Fall 24  
McGill University, Canada (with Prof. Jackie Cheung)

CO-TEACHING Natural language Processing (2 lectures) – Winter 24  
UCL, UK (with Dr. Pontus Stenetorp)

CO-TEACHING Natural language Processing (2 lectures) – Winter 23  
UCL, UK (with Dr. Tim Rocktäschel and Dr. Pontus Stenetorp)

## RESEARCH GRANTS

---

NOV 2025 - OCT 2028	Scaling Multilingual Multicultural Multimodal LLMs for LRLs US\$499,400 funded by schmidt sciences AI2050
JUNE 2025 - MAY 2027	IVADO Professor on NLP for low-resource languages CA \$56,000 funded by <a href="#">IVADO</a>
APR 2025 - MAR 2030	Multilingual, Multicultural and Multimodal Language Model Adaptation for Low-resource Languages CA \$192,500 (36K per year + 12.5K ECR supplement) funded by <a href="#">NSERC 2025 Discovery Grants</a>
APR 2025 - MAR 2027	LLM for machine translation of endangered and low-resource languages (Co-PI with Fatiha Sadat, Jackie Cheung, and James Crippen, CA\$60,000 funded by IVADO)
FEB 2025 - JAN 2026	End-to-end LLM-based speech translation for low-resource languages (Co-PI with Dr. Min Ma (Google), CA\$50,000 funded by Google via MILA)
SEP 2024 - AUG 2025	Adapting LLMs for MT evaluation in African languages (Co-PI with Dr. Colin Cherry (Google), CA\$50,000 funded by Google via MILA)
AUG 2024 - JUL 2029	Canada CIFAR AI Chair CA \$1,020,000 funded by <a href="#">Canadian Institute for Advanced Research (CIFAR)</a>
APR 2023 - SEP 2024	NiLE: Nigerian Languages for Education (Co-PI with Kola Tubosun, US\$45,000 funded by <a href="#">Africa AI4D-EduAI Hub</a> )
JUL 2023 - MAR 2024	Conversational AI and Benchmark datasets for African languages (PI, US \$207,814, A Masakhane project funded by <a href="#">Lacuna Fund 2021</a> )
MAY 2022 - APR 2023	YorùbáVoice: Developing Speech Corpus for Yoruba language (Co-PI with Kola Tubosun, US\$20,000 funded by <a href="#">Translated-Imminent Grant 2021</a> )
MAY 2021 - DEC 2022	Named Entity Recognition and Parts of Speech for African Languages (Co-PI with Peter Nabende, US\$156,500, A Masakhane Project funded by <a href="#">Lacuna Fund 2020</a> )

## AWARDS AND SCHOLARSHIPS

---

MAY. 2025	Best Theme Paper Award at the NAACL 2025, USA.
MAY. 2025	Outstanding Paper Award at the NAACL 2025, USA.
OCT. 2024	<a href="#">Dr.-Eduard-Martin-Preis</a> (Dissertation Prize of Fakultät für Mathematik und Informatik) of the Saarland University Society, Universität des Saarlandes
MAY. 2024	Best Paper Award at the AfricaNLP Workshop @ICLR 2024 for the AngOFA paper
NOV. 2023	Area Chair Award at the IJCNLP-AAACL 2023 for the MasakhaNEWS paper
JUN. 2023	summa cum laude. Ph.D. Graduation Honors, Saarland University, Germany.
MAY. 2023	Best Paper Award at the AfricaNLP Workshop @ICLR 2023 for the MasakhaNEWS paper
OCT. 2022	Best Paper Award (Grand Challenge KISTI category) at the COLING 2022, South Korea
APR. 2021	Best Paper Award at the AfricaNLP Workshop @EACL 2021 for the MasakhaNER paper
MAY. 2019	Best Poster Presentation Award at the Web Conference (formerly WWW) 2019
JUN. 2018	NII International Internship Program, Japan (February - April 2019)
MAR. 2017	MPI-SWS Pre-Doctoral Studies Scholarship (ended December 2018)
APR. 2015	Saarbrücken Graduate School of Computer Science Full Scholarship (ended Mar 2017)
JUL. 2013	Nelson Mandela Institute (NMI) Full Scholarship for MSc in Computer Science (\$22,500)

## STUDENT SUPERVISION

---

WINTER 2026	<b>MSc students:</b> Henok B Ademtew (M)
FALL 2025	<b>PhD students:</b> Marie Maltais (F), Damaris S. Ndjebayi (F), David Guzmán (M) <b>MSc students:</b> Victor Akinode (M)
FALL 2024	<b>PhD students:</b> Senyu Li (M) <b>MSc students:</b> Jessica Ojo (F), Hao Yu (M)

## INVITED TALKS

---

NOV 20, 2024	<a href="#">Invited talk at Network of AI Safety Institutes</a> at San Francisco How good are Large Language Models on African Languages?
APR 26, 2023	<a href="#">Invited talk @DeepMind</a> : NLP for under-resourced African languages
MAR 24, 2023	<a href="#">Invited talk @McGill</a> : NLP for under-resourced African languages
MAR 23, 2023	<a href="#">Invited talk @co:here</a> : NLP for under-resourced African languages
FEB 16, 2023	<a href="#">Invited talk @CMU</a> : NLP for under-resourced African languages
DEC 3, 2022	<a href="#">Transfer learning4NLP Workshop @NeurIPS</a> : Cross-lingual Transfer for NER
NOV 28, 2022	<a href="#">Black in AI Workshop @NeurIPS</a> : Cross-lingual Transfer for NER
JUL 14, 2022	<a href="#">Deeplo Workshop @NAACL</a> : Multilingual Language Model Adaptive Fine-Tuning: A Study on African Languages
DEC 6, 2021	<a href="#">Tutorial @NeurIPS</a> : A Journey Through the Opportunity of Low Resourced Natural Language Processing – An African Lens
MAY 27, 2021	<a href="#">UCREL CRS</a> : MasakhaNER: Named entity recognition for African Languages
APR 19, 2021	<a href="#">AfricaNLP Workshop @EACL</a> : MasakhaNER: NER for African Languages
MAR 26, 2021	<a href="#">CMU-LTI Colloquium</a> : MasakhaNER: NER for African Languages
SEP 23, 2020	<a href="#">NLP with Friends</a> : Development of NLP datasets and models for African Languages
JUL 7, 2020	<a href="#">Data Skeptic Podcast</a> : Sentiment Preserving Fake Reviews
JUL 3, 2020	<a href="#">AI4D Webinar</a> : Ensuring good text quality in African Language Datasets

## VOLUNTEER SERVICE

---

Director at	<a href="#">Masakhane Research Foundation</a> – grassroots NLP community for Africa
Organizer for	<a href="#">AfricaNLP 2022-2025</a> , <a href="#">Cross-Cultural Considerations in NLP Workshop 2023</a> , <a href="#">Multilingual Representation Learning Workshop 2023-2025</a>
SAC for	ACL 2025, EMNLP 2025
AC for	NAACL 2025, NeurIPS 2024, EMNLP 2024, ACL 2024, EACL 2024 & NAACL 2024 ACL 2023 & EMNLP 2023 (Multilingualism and/or Cross-Lingual NLP track)
Reviewer for	<a href="#">AfricaNLP Workshop</a> (2020, 2021), <a href="#">WeaSUL @ICLR</a> 2021, <a href="#">NAACL</a> 2021, <a href="#">ARR</a> 2021 - present, <a href="#">LREC</a> 2022, <a href="#">COLING</a> 2022, <a href="#">EMNLP</a> 2022, <a href="#">EACL</a> 2023

## PUBLICATIONS

---

(\*) Equal contributions

1. Kosei Uemura, Miaoran Zhang, and David Ifeoluwa Adelani. AfriMTEB and AfriE5: Benchmarking and Adapting Text Embedding Models for African Languages. In *EACL*, 2026
2. Kosei Uemura, David Guzmán, Quang Phuoc Nguyen, Jesujoba Oluwadara Alabi, Enshiun Annie Lee, and David Ifeoluwa Adelani. MERLIN: Multi-Stage Curriculum Alignment for Multilingual Encoder-LLM Integration in Cross-Lingual Reasoning. In *EACL*, 2026
3. Senyu Li, Jiayi Wang, Felermino D. M. A. Ali, Colin Cherry, Daniel Deutsch, Eleftheria Briakou, Rui Sousa-Silva, Henrique Lopes Cardoso, Pontus Stenetorp, and David Ifeoluwa Adelani. SSA-COMET: Do LLMs outperform learned metrics in evaluating MT for under-resourced African languages? In *EMNLP*, 2025
4. Jesujoba Oluwadara Alabi, Israel Abebe Azime, Miaoran Zhang, Cristina España-Bonet, Rachel Bawden, Dawei Zhu, David Ifeoluwa Adelani, and et al. AFRIDOC-MT: Document-level MT corpus for African languages. In *EMNLP*, 2025
5. Jesujoba Oluwadara Alabi, Michael A. Hedderich, David Ifeoluwa Adelani, and Dietrich Klakow. Charting the landscape of African NLP: Mapping progress and shaping the road ahead. In *EMNLP*, 2025
6. Jiayi Wang, Yao Lu, Maurice Weber, Max Ryabinin, David Ifeoluwa Adelani, Yihong Chen, Raphael Tang, and Pontus Stenetorp. Multilingual language model pretraining using machine-translated data. In *EMNLP*, 2025

7. Tadesse Destaw Belay, Israel Abebe Azime, Ibrahim Said Ahmad, David Ifeoluwa Adelani, Idris Abdulmumin, Abinew Ali Ayele, Shamsuddeen Hassan Muhammad, and Seid Muhie Yimam. AfroXLMR-social: Adapting pre-trained language models for African languages social media text. In *Findings of ACL: EMNLP*, 2025
8. Oluwadara Kalejaiye, Luel H. Beyene, David I. Adelani, Mmekut-mfon Gabriel Edet, Aniefon Daniel Akpan, Eno-Abasi Urua, and Anietie Andy. Ibom NLP: A step toward inclusive natural language processing for Nigeria’s minority languages. In *IJCNLP-AACL*, 2025
9. Jonah Dauvet, Min Ma, Jessica Ojo, and David Ifeoluwa Adelani. Reassessing speech translation for low-resource languages: Do LLMs redefine the state-of-the-art against cascaded models? In *Proceedings of the 5th Workshop on Multilingual Representation Learning (MRL 2025 at EMNLP)*, pages 149–160
10. Hao Yu and David Ifeoluwa Adelani. Training of LLM-based list-wise multilingual reranker. In *Proceedings of the 5th Workshop on Multilingual Representation Learning (MRL 2025) at EMNLP*, pages 652–663
11. Senyu Li, Felermينو Dario Mario Ali, Jiayi Wang, Rui Sousa-Silva, Henrique Lopes Cardoso, Pontus Stenetorp, Colin Cherry, and David Ifeoluwa Adelani. Evaluating WMT 2025 metrics shared task submissions on the SSA-MTE African challenge set. In *Proceedings of the Tenth Conference on Machine Translation*, pages 913–919, November 2025
12. Alon Lavie, Greg Hanneman, Sweta Agrawal, Diptesh Kanojia, Chi-Kiu Lo, Vilém Zouhar, ..., Jiayi Wang, David Ifeoluwa Adelani, and et al. Findings of the WMT25 shared task on automated translation evaluation systems: Linguistic diversity is challenging and references still help. In *WMT*, 2025
13. Luel Hagos Beyene, Vivek Verma, Min Ma, Jesujoba O Alabi, Fabian David Schmidt, Joyce Nakatumba-Nabende, and David Ifeoluwa Adelani. msteb: Massively multilingual evaluation of llms on speech and text tasks. In *ASRU*, 2025
14. Fabian David Schmidt, Ivan Vuli, Goran Glava, and David Ifeoluwa Adelani. Fleurs-SLU: A massively multilingual benchmark for spoken language understanding. In *Second Conference on Language Modeling (COLM)*, 2025
15. Chris Emezue, The NaijaVoices Community, Busayo Awobade, Abraham T. Owodunni, GMT Emezue ..., Sewade Ogun, Bunmi Akinremi, David Ifeoluwa Adelani, and Chris Pal. The NaijaVoices Dataset: Cultivating Large-Scale, High-Quality, Culturally-Rich Speech Data for African Languages. In *Interspeech 2025*, pages 1338–1342, 2025
16. Shamsuddeen Hassan Muhammad, Nedjma Ousidhoum, Idris Abdulmumin, Seid Muhie Yimam, Jan Philip Wahle, Christine De Kock, ..., David Ifeoluwa Adelani, and et al. SemEval-2025 task 11: Bridging the gap in text-based emotion detection. In *SemEval-2025*, pages 2558–2569, Vienna, Austria, July 2025. Association for Computational Linguistics
17. Senyu Li, Zipeng Sun, Jiayi Wang, Xue Liu, Pontus Stenetorp, Siva Reddy, and **David Ifeoluwa Adelani**. Warmup Generations: A Task-Agnostic Approach for Guiding Sequence-to-Sequence Learning with Unsupervised Initial State Generation. In *ACL*, 2025
18. Hao Yu, Jesujoba O Alabi, Andiswa Bukula, Jian Yun Zhuang, En-Shiun Annie Lee, ..., and **David Ifeoluwa Adelani**. INJONGO: A Multicultural Intent Detection and Slot-filling Dataset for 16 African Languages. In *ACL*, 2025
19. Jessica Ojo, Odunayo Ogundepo, Akintunde Oladipo, Kelechi Ogueji, Jimmy Lin, Pontus Stenetorp, and **David Ifeoluwa Adelani**. AfroBench: How Good are Large Language Models on African Languages? In *Findings of ACL*, 2025
20. Shivalika Singh, Angelika Romanou, Clémentine Fourrier, **David Ifeoluwa Adelani**, Jian Gang ... Ngui, and Sara Hooker. Global MMLU: Understanding and Addressing Cultural and Linguistic Biases in Multilingual Evaluation. In *ACL*, 2025
21. Shamsuddeen Hassan Muhammad, Nedjma Ousidhoum, Idris Abdulmumin, Jan Philip Wahle, ..., **David Ifeoluwa Adelani**, and et al. BRIGHTER: BRIdging the Gap in Human-Annotated Textual Emotion Recognition Datasets for 28 Languages. In *ACL*, 2025

22. **David Ifeoluwa Adelani**, Jessica Ojo, Israel Abebe Azime, Jian Yun Zhuang, Xuanli He Jesujoba O. Alabi, and Millicent Ochiengand Sara Hooker et al. IrokoBench: A New Benchmark for African Languages in the Age of Large Language Models. In *NAACL*, 2025
23. **David Ifeoluwa Adelani**, A. Seza Dogruöz, Iyanuoluwa Shode, and Anuoluwapo Aremu. Does Generative AI speak Nigerian-Pidgin?: Issues about Representativeness and Bias for Multilingualism in LLMs. In *Findings of NAACL*, 2025
24. Shamsuddeen Hassan Muhammad, Idris Abdulmumin, Abinew Ali Ayele, **David Ifeoluwa Adelani**, Ibrahim Said Ahmad, and et al. AfriHate: A Multilingual Collection of Hate Speech and Abusive Language Datasets for African Languages. In *NAACL*, 2025
25. Genta Indra Winata, Frederikus Hudi, Patrick Amadeus Irawan, David Anugraha, Rifki Afina Putri, ..., **David Ifeoluwa Adelani**, and et al. WorldCuisines: A Massive-Scale Benchmark for Multilingual and Multicultural Visual Question Answering on Global Cuisines . In *NAACL*, 2025
26. Shayne Longpre, Stella Biderman, Alon Albalak, Hailey Schoelkopf, Daniel McDuff, ..., Bertie Vidgen, Laura Weidinger, Arvind Narayanan, Victor Sanh, **David Adelani**, and et al. The Responsible Foundation Model Development Cheatsheet: A Review of Tools & Resources . In *Transactions on Machine Learning Research*, 2024
27. David Romero, Chenyang Lyu, Haryo Akbarianto Wibowo, Teresa Lynn, Injy Hamed, Aditya Nanda Kishore, Aishik Mandal, Alina Dragonetti, Artem Abzaliev, Atnafu L. Tonja, ..., **David Ifeoluwa Adelani**, David Le Meur, Emilio Villa-Cueva, and et al. CVQA: Culturally-diverse Multilingual Visual Question Answering Benchmark. In *NeurIPS D & B*, 2024
28. Orevaoghene Ahia, Anuoluwapo Aremu, Diana Abagyan, Hila Gonen, **David Ifeoluwa Adelani**, Daud Olamide Abolade, Noah A. Smith, and Yulia Tsvetkov. Voices unheard: Nlp resources and models for yorùbá regional dialects. In *EMNLP*, 2024
29. Kenza Benkirane, Laura Gongas, Shahar Pelles, Naomi Fuchs, Joshua Darmon, Pontus Stenetorp, **David Ifeoluwa Adelani**, and Eduardo Sánchez. Machine translation hallucination detection for low and high resource languages using large language models. In *EMNLP Findings*, 2024
30. Genta Indra Winata, Ruochen Zhang, and **David Ifeoluwa Adelani**. Miners: Multilingual language models as semantic retrievers. In *EMNLP Findings*, 2024
31. Senyu Li, Hao Yu, Jessica Ojo, and David Ifeoluwa Adelani. McGill NLP group submission to the MRL 2024 shared task: Ensembling enhances effectiveness of multilingual small LMs. In Jonne Sälevä and Abraham Owodunni, editors, *Proceedings of the Fourth Workshop on Multilingual Representation Learning (MRL 2024) @EMNLP*, 2024
32. Jiayi Wang, **David Ifeoluwa Adelani**, and Pontus Stenetorp. Evaluating WMT 2024 metrics shared task submissions on AfriMTE (the African challenge set). In *Proceedings of the Ninth Conference on Machine Translation*, 2024
33. Markus Freitag, Nitika Mathur, Daniel Deutsch, Chi-Kiu Lo, Eleftherios Avramidis, Ricardo Rei, Brian Thompson, Frederic Blain, Tom Kocmi, Jiayi Wang, **David Ifeoluwa Adelani**, and et al. Are LLMs breaking MT metrics? results of the WMT24 metrics shared task. In *Proceedings of the Ninth Conference on Machine Translation*, 2024
34. Akindele Michael Olawole, Jesujoba Oluwadara Alabi, Aderonke Busayo Sakpere, and **David Ifeoluwa Adelani**. YAD: Leveraging t5 for improved automatic diacritization of yorùbá text. In *5th Workshop on AfricaNLP @ICLR*, 2024
35. Osvaldo Luamba Quinjica and **David Ifeoluwa Adelani**. AngOFA: Leveraging OFA embedding initialization and synthetic data for angolan language model. In *5th Workshop on AfricaNLP @ICLR*, 2024
36. Anuoluwapo Aremu, Jesujoba Oluwadara Alabi, Daud Abolade, Nkechinyere Faith Aguobi, Shamsuddeen Hassan Muhammad, and **David Ifeoluwa Adelani**. NaijaRC: A multi-choice reading comprehension dataset for nigerian languages. In *5th Workshop on AfricaNLP @ICLR*, 2024

37. Jiayi Wang, **David Ifeoluwa Adelani**, Sweta Agrawal, Ricardo Rei, Eleftheria Briakou, Marine Carpuat, Marek Masiak, Xuanli He, Sofia Bourhim, Andiswa Bukula, Muhidin A. Mohamed, ..., and Pontus Stenetorp. AfriMTE and AfriCOMET: Empowering COMET to Embrace Under-resourced African Languages. In *NAACL*, 2024
38. Comfort Ilevbare, Jesujoba Alabi, **Adelani, David Ifeoluwa**, Firdous Bakare, Oluwatoyin Abiola, and Oluwaseyi Adeyemo. EkoHate: Abusive language and hate speech detection for code-switched political discussions on Nigerian Twitter. In *Workshop on Online Abuse and Harms (WOAH) @NAACL*, 2024
39. **Adelani, David Ifeoluwa**, A. Seza Dogruoz, André Coneglian, and Atul Kr. Ojha. Comparing LLM prompting with cross-lingual transfer performance on indigenous and low-resource Brazilian languages. In *AmericasNLP Workshop @NAACL*, 2024
40. Tolulope Ogunremi, Kola Tubosun, Anuoluwapo Aremu, Iroro Orife, and **Adelani, David Ifeoluwa**. IròyinSpeech: A multi-purpose Yorùbá speech corpus. In *LREC-COLING*, 2024
41. Garry Kuwanto, Eno-Abasi E. Urua, Priscilla Amondi Amuok, Shamsuddeen Hassan Muhammad, Anuoluwapo ... Aremu, **Adelani, David Ifeoluwa**, Derry Tanti Wijaya, and Anietie Andy. Mitigating translationese in low-resource languages: The storyboard approach. In *LREC-COLING 2024*, 2024
42. **David Ifeoluwa Adelani**, Hannah Liu, Xiaoyu Shen, Nikita Vassilyev, Jesujoba Oluwadara Alabi, Yanke Mao, Haonan Gao, and Annie En-Shiun Lee. Sib-200: A simple, inclusive, and big evaluation dataset for topic classification in 200+ languages and dialects. In *EACL*, 2024
43. Yihong Chen, Kelly Marchisio, Roberta Raileanu, **David Ifeoluwa Adelani**, Pontus Stenetor, Sebastian Riedel, and Mikel Artetx. Improving Language Plasticity via Pretraining with Active Forgetting. In *NeurIPS*, 2023
44. Akintunde Oladipo, Mofetoluwa Adeyemi, Orevaoghene Ahia, Abraham Owodunni, Odunayo Ogundepo, **David Adelani**, and Jimmy Lin. Better Quality Pre-training Data and T5 Models for African Languages. In *EMNLP*, 2023
45. Sebastian Ruder, J. Clark, Alexander Gutkin, Mihir Kale, Min Ma, Massimo Nicosia, Shruti Rijhwani, ..., Brian Roark, Bidisha Samanta, Connie Tao, **David Ifeoluwa Adelani**, Vera Axelrod, and et al. XTREME-UP: A User-Centric Scarce-Data Benchmark for Under-Represented Languages. In *EMNLP*, 2023
46. Odunayo Ogundepo, Tajuddeen Rabi'u Gwadabe, Clara Rivera, J. Clark, Sebastian Ruder, **David Ifeoluwa Adelani**, Bonaventure F. P. Dossou, Abdoulahat Diop, Claytone Sikasote, Gilles Hacheme, Happy Buzaaba, and et al. AfriQA: Cross-lingual Open-Retrieval Question Answering for African Languages. In *EMNLP*, 2023
47. Shamsuddeen H. Muhammad, Idris Abdulmumin, Abinew A. Ayele, Nedjma Djouhra Ousidhoum, **David Ifeoluwa Adelani**, Seid M. Yimam, Ibrahim Said Ahmad, Meriem Beloucif, Saif M. Mohammad, Sebastian Ruder, and et al. AfriSenti: A Twitter Sentiment Analysis Benchmark for African Languages. In *EMNLP*, 2023
48. Francesco Tinner, **Adelani, David Ifeoluwa**, Chris Emezue, Mammad Hajili, Omer Goldman, ..., Rahul Aralikkatte, Najoung Kim, and Duygu Ataman. Findings of the 1st Shared Task on Multi-lingual Multi-task Information Retrieval at MRL 2023. In Duygu Ataman, editor, *MRL Workshop @EMNLP*, 2023
49. Kathleen Siminyu, Jade Abbott, Kola Tubosun, Aremu Anuoluwapo, Blessing K. Sibanda, Kofi Yeboah, **David Adelani**, and et al. Consultative engagement of stakeholders toward a roadmap for African language technologies. In *Patterns*, 2023
50. **David Ifeoluwa Adelani**, Marek Masiak, Israel Abebe Azime, Jesujoba Oluwadara Alabi, Atnafu Lambebo Tonja, Christine Mwase, Odunayo Ogundepo, Bonaventure F. P. Dossou, Akintunde Oladipo, ..., and Pontus Stenetorp. MasakhaNEWS: News Topic Classification for African languages. In *IJCNLP-AACL*, 2023
51. Iyanuoluwa Shode, **David Ifeoluwa Adelani**, Jing Peng, and Anna Feldman. NollySenti: Leveraging Transfer Learning and Machine Translation for Nigerian Movie Sentiment Classification. In *ACL*, 2023

52. Zheng X. Yong, Hailey Schoelkopf, Niklas Muennighoff, Alham F. Aji, **David Ifeoluwa Adelani**, Khalid Almubarak, M Saiful Bari, Lintang Sutawika, Jungo Kasai, Ahmed Baruwa, Genta I. Winata, Stella Rose Biderman, Dragomir R. Radev, and Vassilina Nikoulina. BLOOM+1: Adding Language Support to BLOOM for Zero-Shot Prompting. In *ACL*, 2023
53. Cheikh M. Bamba Dione\*, **David Ifeoluwa Adelani**\*, Peter Nabende, Jesujoba Alabi, Thapelo Sindane, Happy Buzaaba, Shamsuddeen Hassan Muhammad, Chris Chinenye Emezue, Perez Ogayo, and Dietrich Klakow. MasakhaPOS: Part-of-Speech Tagging for Typologically Diverse African languages . In *ACL*, 2023
54. Shamsuddeen H. Muhammad, Idris Abdulmumin, Seid M. Yimam, **David Ifeoluwa Adelani**, Ibrahim Said Ahmad, Nedjma Djouhra Ousidhoum, Abinew Ali Ayele, Saif M. Mohammad, Meriem Beloucif, and Sebastian Ruder. SemEval-2023 Task 12: Sentiment Analysis for African Languages (AfriSenti-SemEval). In *SEM Eval Workshop @ACL*, 2023
55. Idris Akinade, Jesujoba Oluwadara Alabi, **David Ifeoluwa Adelani**, Clement Odoje, and Dietrich Klakow.  $\epsilon$  KÚ: Integrating Yorùbá cultural greetings into machine translation. In *Cross-Cultural Considerations in NLP Workshop @EACL*, 2023
56. Rendani Mbuva, **David Ifeoluwa Adelani**, Tendani Mutavhatsindi, Tshimangadzo Rakhuu, Aluwani Mauda, Tshifhiwa Joshua Maumela, Andisani Masindi, Seani Rananga, Vukosi Marivate, and T. Marwala. MphayaNER: Named Entity Recognition for Tshivenda. In *4th Workshop on AfricaNLP @EACL*, 2023
57. **David Ifeoluwa Adelani**, Md Mahfuz Ibn Alam, Antonios Anastasopoulos, Akshita Bhagia, Marta R. Costa-jussà, and et al. Findings of the WMT’22 Shared Task on Large-Scale Machine Translation Evaluation for African Languages . In *Proceedings of the Seventh Conference on Machine Translation at EMNLP*, 2022
58. **David Ifeoluwa Adelani**, Graham Neubig, Sebastian Ruder, Shruti Rijhwani, Michael Beukman, Chester Palen-Michel, Constantine Lignos, Jesujoba O. Alabi, Shamsuddeen H. Muhammad, Peter Nabende, Cheikh M. Bamba Dione, ..., and Dietrich Klakow. MasakhaNER 2.0: Africa-centric Transfer Learning for Named Entity Recognition . In *EMNLP*, 2022
59. Jesujoba Oluwadara Alabi\*, **David Ifeoluwa Adelani**\*, Marius Mosbach, and Dietrich Klakow. Adapting Pre-trained Language Models to African Languages via Multilingual Adaptive Fine-Tuning. In *COLING*, 2022
60. Ernie Chang, Jesujoba Alabi, **David Ifeoluwa Adelani**, and Vera Demberg. Few-Shot Pidgin Text Adaptation via Contrastive Fine-Tuning. In *COLING*, 2022
61. Teven Le Scao, Angela Fan, Christopher Akiki, Ellie Pavlick, Suzana Ilić, Daniel Hesslow, Roman Castagn’e, Alexandra Sasha Luccioni, Francois Yvon, Matthias Gallé, Jonathan Tow, Alexander M. Rush, ..., **David Ifeoluwa Adelani**, and Thomas Wolf. BLOOM: A 176B-Parameter Open-Access Multilingual Language Model. *ArXiv*, abs/2211.05100, 2022
62. Hugo Laurençon, Lucile Saulnier, Thomas Wang, Christopher Akiki, Albert Villanova del Moral, Teven Le Scao, ..., **David Ifeoluwa Adelani**, ..., Margaret Mitchell, Sasha Luccioni, and Yacine Jernite. The BigScience ROOTS Corpus: A 1.6TB Composite Multilingual Dataset . In *NeurIPS (Datasets and Benchmarks Track)*, 2022
63. Josh Meyer, **David Ifeoluwa Adelani**, Edresson Casanova, Alp Öktem, Daniel Whitenack Julian Weber, Salomon Kabongo, Elizabeth Salesky, Irero Orife, Colin Leong, Perez Ogayo, Chris Emezue, Jonathan Mukiibi, Salomey Osei, and et al. BibleTTS: a large, high-fidelity, multilingual, and uniquely African speech corpus. In *Interspeech*, 2022
64. **David Ifeoluwa Adelani**, Jesujoba Oluwadara Alabi, Angela Fan, Julia Kreutzer, Xiaoyu Shen, Machel Reid, Dana Ruiter, Dietrich Klakow, Peter Nabende, Ernie Chang, Tajudeen Gwadabe, Freshia Sackey, , and et al. A Few Thousand Translations Go a Long Way! Leveraging Pre-trained Models for African News Translation. In *NAACL-HLT*, 2022
65. Miaoran Zhang, Marius Mosbach, **David Ifeoluwa Adelani**, Michael A. Hedderich, and Dietrich Klakow. MCSE: Multimodal Contrastive Learning of Sentence Embeddings. In *NAACL-HLT*, 2022
66. Shamsuddeen H. Muhammad, **David Ifeoluwa Adelani**, Sebastian Ruder, Ibrahim S. Ahmad, Idris Abdulmumin, Bello Shehu Bello, Monojit Choudhury, Chris C. Emezue, Sa-



- heed S. Abdullahi, Anuoluwapo Aremu, Alipio George, and Pavel Brazdil. NaijaSenti: A Nigerian Twitter Sentiment Corpus for Multilingual Sentiment Analysis. In *LREC*, 2022
67. Ali Davody, **David Ifeoluwa Adelani**, Thomas Kleinbauer, and Dietrich Klakow. TOKEN is a MASK: Few-shot Named Entity Recognition with Pre-trained Language Models. In *In Proceedings of the 25th International Conference on Text, Speech and Dialogue (TSD)*, 2022
  68. En-Shiun Annie Lee, Sarubi Thillainathan, Shravan Nayak, Surangika Ranathunga, **David Ifeoluwa Adelani**, and Ruisi Su and Arya D. McCarthy. Pre-Trained Multilingual Sequence-to-Sequence Models: A Hope for Low-Resource Language Translation? In *Findings of ACL*, 2022
  69. Dawei Zhu, Michael A. Hedderich, Fangzhou Zhai, **David Ifeoluwa Adelani**, and Dietrich Klakow. Task-Adaptive Pre-Training for Boosting Learning With Noisy Labels: A Study on Text Classification for African Languages. In *Workshop on Insights from Negative Results in NLP 2022 at ACL*, 2022
  70. Iyanuoluwa Shode, **David Ifeoluwa Adelani**, and Anna Feldman. YOSM: A new Yorùbá Sentiment Corpus for Movie reviews. In *3rd Workshop on AfricaNLP @ICLR*, 2022
  71. **David Ifeoluwa Adelani**, Miaoran Zhang, Xiaoyu Shen, Ali Davody, Thomas Kleinbauer, and Dietrich Klakow. Preventing author profiling through zero-shot multilingual back-translation. In *EMNLP*, 2021
  72. **David Ifeoluwa Adelani** and Dana Ruiter and Jesujoba O. Alabi and Damilola Adebajo and Adesina Ayeni and Mofe Adeyemi and Ayodele Awokoya and Cristina España-Bonet . The Effect of Domain and Diacritics in Yorùbá-English Neural Machine Translation. In *MT Summit*, 2021
  73. **David Ifeoluwa Adelani**, Jade Abbott, Graham Neubig, Daniel D'souza, Julia Kreutzer, Constantine Lignos, Chester Palen-Michel, Happy Buzaaba, Shruti Rijhwani, Sebastian Ruder, Stephen Mayhew, and et al. MasakhaNER: Named Entity Recognition for African Languages . In *TACL*, 2021
  74. Kathleen Siminyu, Godson Kalipe, Davor Orlic, Jade Abbott, Vukosi Marivate, Sackey Freshia, Prateek Sibal, Bhanu Neupane, **David I. Adelani**, Amelia Taylor, Jamiil Toure ALI, Kevin Degila, Momboladji Balogoun, Thierno Ibrahima DIOP, and et al. AI4D – African Language Program. In *2nd Workshop on AfricaNLP @EACL*, 2021
  75. Michael A. Hedderich, **David Ifeoluwa Adelani**, Dawei Zhu, Jesujoba O. Alabi, Udia Markus, and Dietrich Klakow. Transfer Learning and Distant Supervision for Multilingual Transformer Models: A Study on African Languages. In *EMNLP*, 2020
  76. **David Ifeoluwa Adelani**, Ryota Kobayashi, Ingmar Weber, and Przemyslaw A. Grabowicz. Estimating community feedback effect on topic choice in social media with predictive modeling. In *EPJ Data Science*, 2020 (Also accepted at IC2S2 2021 as Extended Abstract)
  77. **David Ifeoluwa Adelani**, Ali Davody, Thomas Kleinbauer, and Dietrich Klakow. Privacy guarantees for de-identifying text transformations. In *Interspeech*, 2020
  78. Aleena Thomas, **David Ifeoluwa Adelani**, Ali Davody, Aditya Mogadala, and Dietrich Klakow. Investigating the impact of pre-trained word embeddings on memorization in neural networks. In *In Proceedings of the 23rd International Conference on Text, Speech and Dialogue (TSD)*, 2020
  79. **David Ifeoluwa Adelani\***, Michael A. Hedderich\*, Dawei Zhu\*, Esther van den Berg, and Dietrich Klakow. Distant supervision and noisy label learning for low resource named entity recognition: A study on Hausa and Yorùbá, Practical ML for Developing Countries Workshop at ICLR 2020
  80. Iroro Orife, **David I. Adelani**, Timi Fasubaa, Victor Williamson, Wuraola Fisayo Oyewusi, Olamilekan Wahab, and Kola Tubosun. Improving Yorùbá diacritic restoration, AfircaNLP Workshop at ICLR 2020
  81. Ernie Chang, **David Ifeoluwa Adelani**, Xiaoyu Shen, and Vera Demberg. Unsupervised Pidgin Text Generation By Pivoting English Data and Self-Training, AfircaNLP Workshop at ICLR 2020



82. Jesujoba O. Alabi, Kwabena Amponsah-Kaakyire, **David I. Adelani**, and Cristina España-Bonet. Massive vs. curated word embeddings for low-resourced languages. the case of Yorùbá and Twi. In *12th International Conference on Language Resources and Evaluation (LREC)*, 2020
83. **David Ifeoluwa Adelani**, Haotian Mai, Fuming Fang, Huy H. Nguyen, Junichi Yamagishi, and Isao Echizen. Generating sentiment-preserving fake online reviews using neural language models and their human- and machine-based detection. In *34th International Conference on Advanced Information Networking and Applications (AINA)*, 2020
84. Zijian Wang, Scott Hale, **David Ifeoluwa Adelani**, Przemyslaw Grabowicz, Timo Hartmann, Fabian Flöck, and David Jurgens. Demographic inference and representative population estimates from multilingual social media data. In *The Web conference*. ACM, 2019
85. **David Ifeoluwa Adelani**, Fabian Flöck, Przemyslaw Grabowicz, Scott Hale, Timo Hartmann, and David Jurgens. Agendas on social media: Inferring policy attention from non-representative data. *International Conference on Computational Social Science (IC2S2)* 2018. Extended Abstract
86. **David Ifeoluwa Adelani** and Mamadou Kaba Traore. Enhancing the reusability and interoperability of artificial neural networks with DEVS modeling and simulation. *International Journal of Modeling, Simulation, and Scientific Computing*, 7(3), 2015
87. A.S. Sodiya, S.A. Onashoga, and **D.I. Adelani**. A secure e-voting architecture. In *Information Technology: New Generations (ITNG), 2011 Eighth International Conference on*, pages 342–347. IEEE, 2011

## FAVOURITE QUOTE

---

Anything you find your hands doing, do it with all your might