

David Ifeoluwa ADELANI

PERSONAL DATA

PLACE OF BIRTH: Abeokuta, Nigeria
ADDRESS: Saarbrücken, Germany
EMAIL: d.adelani@ucl.ac.uk

WORK EXPERIENCE

SEP 2022 - PRESENT	Senior Research Fellow (DeepMind Academic Fellow) <i>Department of Computer Science, University College London, UK.</i> Supported by DeepMind
JAN 2019 - SEP 2022	PhD Researcher (Group: Spoken Language Systems Group) <i>Department of Language Science & Technology, Saarland University</i> Supported by the EU H2020 COMPRISE project
JAN 2019 - APR 2019	Research Intern (Group: Echizen's Content Security Lab) <i>National Institute of Informatics (NII)</i> Internship on Adversarial Text Generation with Prof. Isao Echizen
MAR 2017 - DEC 2018	Research Assistant (Group: Networked Systems Research Group) <i>Max Planck Institute for Software System (MPI-SWS)</i> Supported by MPI-SWS Pre-Doctoral Scholarship for 22 months for working on topic modeling and causal inference in social media with Dr. Przemyslaw Grabowicz
AUG 2016- MAR 2017	Research Assistant (Group: Intelligent User Interfaces) <i>German Research Centre for Artificial Intelligence (DFKI)</i> Deep learning and Case-Based Reasoning for Energy Power Forecast [10 hours/week]
JUL 2012 - APR 2015	Computer Scientist II (Lead of Database Group) <i>Federal Ministry of Communication Technology, Nigeria</i>
APR 2011 - JUL 2012	Professional Service Consultant (Junior) <i>NCR (NIGERIA) PLC.</i>

EDUCATION

JAN 2015 - SEP 2022	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 GPA: 1.6 (Best 1.0, Worst 5.0)
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

RESEARCH INTERESTS

Natural Language Processing (NLP)
NLP for low-resource languages
NLP for social good
Privacy-preserving machine learning
Computational social science

TEACHING EXPERIENCE

TEACHING ASSISTANT: Neural Networks: Implementation and Application - Winter Semester 17/18
Saarland University (with Prof. Dr. Dietrich Klakow)
TEACHING ASSISTANT: Statistical Natural Language Processing - Summer 2017
Saarland University (with Prof. Dr. Dietrich Klakow)

OTHER RESEARCH PROJECTS

OCT 2022 - SEP 2023 Conversational AI and Benchmark datasets for African languages
(A Masakhane Project supported by [Lacuna Fund 2021](#))
MAY 2021 - DEC 2022. Named Entity Recognition and Parts of Speech for African Languages
(A Masakhane Project supported by [Lacuna Fund 2020](#))
OCT 2016 - JAN 2017. Detecting Events in Social Media using Non-parametric Bayesian Methods
(with Dr. Przemyslaw Grabowicz @ MPI-SWS)
MAY 2016 - SEP 2016. Neural Networks for Semantic Role Prediction
(with Prof. Dr. Vera Demberg)

AWARDS, SCHOLARSHIPS AND CERTIFICATES

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
JUN. 2020 AI4D Fellowship to create English-Yorùbá Corpus for Machine Translation
OCT. 2019 Attended the PRAIRIE AI summer school (fully funded) in Paris (October 3 - 5)
MAY. 2019 Invited to Google's 3rd European NLP Summit in Zurich (June 24 - 26)
MAY. 2019 Best Poster Presentation Award at the Web Conference (formerly WWW) 2019
OCT. 2018 PyTorch Scholarship Challenge 2018/2019 from Facebook and Udacity
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)
DEC. 2012 Oracle Database 11g Certified Professional (OCP)
AUG. 2012 Oracle Database 11g Certified Associate (OCA)

COMPUTER SKILLS

Programming: Python, JAVA, R, HTML, JavaScript, and PHP
Frameworks: [Pandas](#), [Statsmodel](#), [Gensim](#), [TensorFlow](#), [Keras](#), [PyTorch](#), and [scikit-learn](#)
Databases: Oracle Database 11g, Microsoft SQL Server, MySQL, and Access

VOLUNTEER SERVICE

Reviewer for The Web Conference 2018 (sub-reviewer), AfricaNLP Workshop (2020, 2021)
EACL SRW 2021, WeaSUL @ICLR 2021, NAACL 2021, ARR 2021 - present,
COLING 2022, EMNLP 2022
Mentor at [Data Science Nigeria](#)
Active Member of [Masakhane](#) NLP for African languages leading the [Lacuna NER project](#)

FAVOURITE QUOTE

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

PUBLICATIONS

1. 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 (Also accepted at AfricaNLP Workshop at ICLR 2022)*, 2022
2. Ernie Chang, Jesujoba Alabi, David Adelani, and Vera Demberg. Few-Shot Pidgin Text Adaptation via Contrastive Fine-Tuning. In *COLING (Also accepted at AfricaNLP Workshop at ICLR 2022)*, 2022
3. 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, Apelete Agbolo, Victor Akinode, Bernard Opoku, Samuel Olanrewaju, Jesujoba Alabi, and Shamsuddeen Muhammad. BibleTTS: a large, high-fidelity, multilingual, and uniquely African speech corpus. In *Interspeech*, 2022
4. 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
5. David Ifeoluwa Adelani, Jesujoba Oluwadara Alabi, Angela Fan, Julia Kreutzer, Xiaoyu Shen, Machel Reid, Dana Ruiter, Dietrich Klakow, Peter Nabende, Ernie Chang, Tajuddeen Gwadabe, Freshia Sackey, Bonaventure F. P. Dossou, Chris Chinenye Emezue, Colin Leong, Michael Beukman, Shamsuddeen Hassan Muhammad, Guyo Dub Jarso, Oreen Yousuf, Andre Niyongabo Rubungo, Gilles HACHEME, Eric Peter Wairagala, Muhammad Umair Nasir, Benjamin Ayoade Ajibade, Oluwaseyi Ajayi Ajayi, Yvonne Wambui Gitau, Jade Abbott, Mohamed Ahmed, Millicent Ochieng, Anuoluwapo Aremu, Perez Ogayo, Jonathan Mukiibi, Fatoumata Ouoba Kabore, Godson Koffi KALIPE, Derguene Mbaye, Allahsera Auguste Tapo, Victoire Memdjokam Koagne, Edwin Munkoh-Buabeng, Valencia Wagner, Idris Abdulmumin, and Ayodele Awokoya. A Few Thousand Translations Go a Long Way! Leveraging Pre-trained Models for African News Translation. In *NAACL-HLT*, 2022
6. Miaoran Zhang, Marius Mosbach, David Ifeoluwa Adelani, Michael A. Hedderich, and Dietrich Klakow. MCSE: Multimodal Contrastive Learning of Sentence Embeddings. In *NAACL-HLT*, 2022
7. Shamsuddeen Hassan Muhammad, David Ifeoluwa Adelani, Sebastian Ruder, Ibrahim Said Ahmad, Idris Abdulmumin, Bello Shehu Bello, Monojit Choudhury, Chris Chinenye Emezue, Saheed Salahudeen Abdullahi, Anuoluwapo Aremu, Alipio George, and Pavel Brazdil. NaijaSenti: A Nigerian Twitter Sentiment Corpus for Multilingual Sentiment Analysis. In *LREC*, 2022
8. 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
9. 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 (Also accepted at AfricaNLP Workshop at ICLR 2022)*, 2022

10. Iyanuoluwa Shode, David Ifeoluwa Adelani, and Anna Feldman. YOSM: A new Yorùbá Sentiment Corpus for Movie reviews. In *AfricaNLP 2022 Workshop at ICLR*, 2022
11. 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
12. David I. 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
13. 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, Israel Abebe Azime, Shamsuddeen Muhammad, Chris Chinenye Emezue, Joyce Nakatumba-Nabende, Perez Ogayo, Anuoluwapo Aremu, Catherine Gitau, Derguene Mbaye, Jesujoba Alabi, Seid Muhie Yimam, Tajuddeen Gwadabe, Ignatius Ezeani, Rubungo Andre Niyongabo, Jonathan Mukiibi, Verrah Otiende, Iroro Orife, Davis David, Samba Ngom, Tosin Adewumi, Paul Rayson, Mofetoluwa Adeyemi, Gerald Muriuki, Emmanuel Anebi, Chiamaka Chukwuneke, Nkiruka Odu, Eric Peter Wairagala, Samuel Oyerinde, Clemencia Siro, Tobius Saul Bateesa, Temilola Oloyede, Yvonne Wambui, Victor Akinode, Deborah Nabagereka, Maurice Katusiime, Ayodele Awokoya, Mouhamadane MBOUP, Dibora Gebreyohannes, Henok Tilaye, Kelechi Nwaike, Degaga Wolde, Abdoulaye Faye, Blessing Sibanda, Orevaoghene Ahia, Bonaventure F. P. Dossou, Kelechi Ogueji, Thierno Ibrahima DIOP, Abdoulaye Diallo, Adewale Akinfaderin, Tendai Marengereke, and Salomey Osei. MasakhaNER: Named Entity Recognition for African Languages . In *TACL*, 2021 (Also accepted at AfricaNLP Workshop @ EACL 2021)
14. David I. Adelani, Dana Ruiter, Jesujoba O. Alabi, Damilola Adebajo, Adesina Ayeni, Mofe Adeyemi, Ayodele Awokoya, and Cristina España-Bonet. MENYO-20k: A Multi-domain English-Yorùbá Corpus for Machine Translation and Domain Adaptation. In *AfricaNLP Workshop @EACL*, 2021
15. 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, Davis David, Chayma Fourati, Hatem Haddad, and Malek Naski. AI4D – African Language Program. In *AfricaNLP Workshop at EACL*, 2021
16. 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
17. 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)
18. David Ifeoluwa Adelani, Ali Davody, Thomas Kleinbauer, and Dietrich Klakow. Privacy guarantees for de-identifying text transformations. In *Interspeech*, 2020
19. 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
20. 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

21. 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
22. 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
23. 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
24. 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
25. 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 (Also accepted at IC2S2 2019 as Extended Abstract)
26. 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
27. 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
28. Adesina Simon Sodiya, SA Onashoga, and DI Adelani. A secure e-voting architecture. In *Information Technology: New Generations (ITNG), 2011 Eighth International Conference on*, pages 342–347. IEEE, 2011