Shamsuddeen Hassan Muhammad, PhD

✉ s.muhammad@imperial.ac.uk   📞 +447 707-515-926   📱 +447 707-515-926

🐦 @shmuhammadd   in linkedin.com/in/shmuhammad   ○ github.com/shmuhammadd

🌐 shmuhammadd.github.io   🎓 Google Scholar   📒 Semantic Scholar

Academic and Research Experience

2025 – ⋯   **Visiting Lecturer**
African Institute for Mathematical Sciences (AIMS) Cameroon.

2023 – ⋯   **Advanced Research Fellow**
Imperial College London, UK

Google DeepMind Academic Fellow
Imperial College London, UK

Visiting Lecturer
York St John University,UK

2020 – 2023   **NLP Researcher, INESCTEC Porto, Portugal**
INESCTEC Porto, Portugal

2012 – ⋯   **Lecturer,**
Department of Computer Science, Bayero University, Kano–Nigeria.

2010 – 2012   **Lecturer,**
Department of Computer Science, Saadatu Rimi College of Education, Kano State-Nigeria

Education

2023   **Ph.D. in Computer Science, University of Porto, Portugal**.
Thesis Title: *Domain-Specific and Context-Aware Approaches to Sentiment Analysis*.
Supervisor: *Pavel Brazdil*

2013   **M.Sc. in Computer Science, University of Manchester, UK**.
Dissertation Title: *Error-Correcting Codes in Wireless Communication*.
Supervisor: *Barry Cheetham*

2009   **B.Sc. in Computer Science, Bayero University, Kano**.

Research Interest

Computational Social Science: Sentiment Analysis, Hate Speech Detection, Emotion Analysis

Low-Resource Languages: Dataset Creation, Data Annotation, LLM Evaluation, and Adaptation

Multilingual NLP: Enhancing Models for Underrepresented Languages and Cultures

Grants and Awards

Grants

2023   Principal-Investigator: Nigerian AI Research Grant for Automatic Offensive Language Detection for Nigerian Languages (5,000,000 Naira)

2022   Principal-Investigator: Artificial Intelligence for Education Innovation Research Network in Africa (40,000 USD)

Principal-Investigator: Lacuna Fund 2021 Language Datasets Award: Development of Hate Speech Dataset for African Languages (147,000 USD)

Co-Investigator: Lacuna Fund Grant for Developing Nigerian Speech Corpus (120,000 USD)

2021   Principal-Investigator: Lacuna Fund 2020 Language Datasets Award: Development of Sentiment Analysis Dataset for African Languages (60,000 USD)

Principal-Investigator: Lacuna Fund Grant for SemEval Shared Task (8,500 USD)

2023   Co-PI: Google Research exploreCSR award (18,000 USD)

Co-PI: Google Research TensorFlow award (10,000 USD)

2022   Principal Investigator: Oracle OCI Award (50,000 USD)

Co-PI: Google Research exploreCSR award (18,000 USD)

Co-PI: TensorFlow award (10,000 USD)

Awards

2024   Google DeepMind Academic Fellow at Imperial College London, UK

Honourable Mention: Best Paper Award for the SemEval Task Description Paper on Semantic Textual Relatedness.

2023   Area Chair Award at IJCNLP-AACL for the MasakhaNEWS Paper

Best Paper Award, AfricaNLP workshop 2023, Kigali Rwanda.

IRCAI, Global Top 100 list 2022, international list of AI solutions for reaching the 17 SGDs.

2022   Google Africa Travel Award. (3,000 USD)

RStudio Conference Travel Award. (1,500 USD)

EMNLP Conference Travel Award. (1,500 USD)

Faculty of Science, UPorto, PhD Scholarship Award. (52,000 EURO)

2021   Best Paper Award at the AfricaNLP Workshop @EACL for the MasakhaNER Paper

2022   Jiangsu University(JSU), China, PhD scholarship, Award (Declined).

2018   Nigerian TetFund Ph.D. Scholarship Award.

NITDA MSc Scholarship Award (25,000 USD).

Black in AI Travel Award to NeurIPS 2022, Value of Award (1,100 USD)

Kano State Govt-Nigeria , Postgraduate Scholarship to Study MSc in Computer Science.

Second Best Computer Science Graduating Student Award, Bayero University, Kano-Nigeria

Teaching Experience

- Artificial Intelligence Concept: Lecturer, York St John University, London Campus (2024)
- Machine Learning: Lecturer, York St John University, London Campus (2024)
- Machine Learning: Teaching Assistant, Imperial College London (2024)
- Natural Language Processing: Teaching Assistant, Imperial College London (2025)
- Deep Learning: Lead Teaching Assistant, NeuroMatch Academy (2023)
- Python Programming: Lecturer, Bayero University, Kano (2017)
- Data Structures and Algorithms: Lecturer, Bayero University, Kano (2016)
- Data Mining: Lecturer, Bayero University, Kano (2016)

Teaching Certificates

2019   **Teaching in Higher Education, University of Porto, Portugal**.
Certificate for teaching in higher education.

2020   **Certified Instructor, The Carpentries**.
Certified instructor in teaching foundational computational and data skills to researchers using R and Python programming languages.

Leadership and Community Involvement in Data Science

2022   Co-founder, Arewa Data Science: `https://arewadatascience.github.io`
A community for empowering marginalized students to reach their full potential in data science.

Co-founder, HausaNLP: `https://hausanlp.github.io`
HausaNLP is an open-source community dedicated to advancing natural language processing in the Hausa language, which is the second most widely spoken language in Africa.

Active Member of MasakhaneNLP and Black in AI Community.

Volunteer Service

- Area Chair: ACL Rolling Review (ARR)
- Reviewer: ACL, EMNLP, NAACL, COLING, LREC, AfricaNLP
- Organizer: AfricaNLP, SemEval Shared Tasks

Academic Service

Service to the University

- Imperial College, Grant Review Member
- Senate Member, Bayero University, Kano
- Faculty Examination Lead, Bayero University, Kano
- BSc Computer Science with Economics course Leader

List of Publications

Shamsuddeen Hassan Muhammad

March 21, 2025

Publications

1. **Muhammad, S. H.**, Ousidhoum, N., Abdulmumin, I., *et al.* (2025). *BRIGHTER: Bridging the Gap in Human-Annotated Textual Emotion Recognition Datasets for 28 Languages.* arXiv preprint. Retrieved from `https://arxiv.org/abs/2502.11926`. [31]

2. **Muhammad, S. H.**, Ousidhoum, N., Abdulmumin, I., *et al.* (2025). *SemEval Task 11: Bridging the Gap in Text-Based Emotion Detection.* In Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval-2025). Vienna, Austria: Association for Computational Linguistics. [32]

3. **Muhammad, S. H.**, Abdulmumin, I., Ayele, A., Ousidhoum, N., Adelani, D., et al. (2023). *AfriSenti: A Twitter Sentiment Analysis Benchmark for African Languages.* In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (pp. 13968–13981). Singapore: Association for Computational Linguistics. [22]

4. **Muhammad, S. H.**, Abdulmumin, I., Ayele, A., Adelani, D. I., Ahmad, I. S., et al. (2025). *AfriHate: A Multilingual Collection of Hate Speech and Abusive Language Datasets for African Languages.* arXiv preprint, arXiv:2501.08284. Retrieved from `https://arxiv.org/abs/2501.08284`.

5. **Muhammad, S. H.**, Adelani, D. I., Ruder, S., Ahmad, I. S., Abdulmumin, I., et al. (2022). *NaijaSenti: A Nigerian Twitter Sentiment Corpus for Multilingual Sentiment Analysis.* In Proceedings of the Thirteenth Language Resources and Evaluation Conference (pp. 590–602). Marseille, France: European Language Resources Association. Retrieved from `https://aclanthology.org/2022.lrec-1.63/`. [26]

6. Vargas, F. A., Guimarães, S., **Muhammad, S. H.**, Alves, D., Ahmad, I. S., et al. (2024). *HausaHate: An Expert Annotated Corpus for Hausa Hate Speech Detection.* In Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH 2024). [41]

7. Parida, S., Bojar, O., Abdulmumin, I., **Muhammad, S. H.**, Ahmad, I. S., et al. (2024). *Findings of WMT2024 English-to-Low Resource Multimodal Translation Task.* In Conference on Machine Translation. [40]

8. Myung, J., Lee, N., Zhou, Y., Jin, J., Putri, R. A., et al. (2024). *BLEnD: A Benchmark for LLMs on Everyday Knowledge in Diverse Cultures and Languages.* [34]

9. Ousidhoum, N. D., **Muhammad, S. H.**, Abdalla, M., Abdulmumin, I., Ahmad, I. S., et al. (2024). *SemEval Task 1: Semantic Textual Relatedness for African and Asian Languages.* In International Workshop on Semantic Evaluation. [38]

10. Ousidhoum, N. D., **Muhammad, S. H.**, Abdalla, M., Abdulmumin, I., Ahmad, I. S., et al. (2024). *SemRel2024: A Collection of Semantic Textual Relatedness Datasets for 14 Languages.* In Annual Meeting of the Association for Computational Linguistics. [37]

11. Adelani, D. I., Ojo, J., Azime, I. A., Jian, Z. Y., Alabi, J. O., et al. (2024). *IrokoBench: A New Benchmark for African Languages in the Age of Large Language Models.* [9]

12. Wang, J., Adelani, D. I., Agrawal, S., Rei, R., Briakou, E., et al. (2023). *AfriMTE and AfriCOMET: Empowering COMET to Embrace Under-resourced African Languages.* [42]

13. Adewumi, T. P., Adeyemi, M., Anuoluwapo, A., Peters, B., Buzaaba, H., et al. (2023). *AfriWOZ: Corpus for Exploiting Cross-Lingual Transfer for Dialogue Generation in Low-Resource, African Languages.* [11]

14. **Muhammad, S. H.**, Brazdil, P., Jorge, A. M. (2023). *Combining Symbolic and Deep Learning Approaches for Sentiment Analysis.* In Compendium of Neurosymbolic Artificial Intelligence. [28]

15. Dione, C. M. B., Adelani, D. I., Nabende, P., Alabi, J. O., Sindane, T., et al. (2023). *MasakhaPOS: Part-of-Speech Tagging for Typologically Diverse African Languages.* In Annual Meeting of the Association for Computational Linguistics. [18]

16. Nekoto, W. O., Marivate, V., Matsila, T., Fasubaa, T. E., Kolawole, T., et al. (2020). *Participatory Research for Low-resourced Machine Translation: A Case Study in African Languages.* arXiv preprint, arXiv:2010.02353. [35]

17. **Muhammad, S. H.** (2019). *An Overview of Sentiment Analysis Approaches.* [23]

18. **Muhammad, S. H.**, Abdulmumin, I., Ahmad, I. S., Gwadabe, T. R., et al. (2023). *AfriSenti: A Twitter Sentiment Analysis Benchmark for African Languages.* arXiv preprint. [24]

19. Aliyu, S. M., Abdulmumin, I., **Muhammad, S. H.**, Ahmad, I. S., Salahudeen, S. A., et al. (2023). *HausaNLP at SemEval-2023 Task 10: Transfer Learning, Synthetic Data and Side-information for Multi-level Sexism Classification.* arXiv preprint. [13]

20. Ousidhoum, N. D., **Muhammad, S. H.**, Abdalla, M., Abdulmumin, I., Ahmad, I. S., et al. (2023). *AfriQA: Cross-lingual Open-Retrieval Question Answering for African Languages.* In Conference on Empirical Methods in Natural Language Processing. [36]

21. Emezue, C. C., Nigatu, H. H., Thinwa, C., Zhou, H., **Muhammad, S. H.**, et al. (2023). *The African Stopwords Project: Curating Stopwords for African Languages.* arXiv preprint. [19]

22. Abdulmumin, I., Beukman, M., Alabi, J. O., Emezue, C. C., **Muhammad, S. H.**, et al. (2022). *Separating Grains from the Chaff: Using Data Filtering to Improve Multilingual Translation for Low-Resourced African Languages.* arXiv preprint. [1]

23. Adelani, D. I., Neubig, G., Ruder, S., Rijhwani, S., Beukman, M., et al. (2022). *MasakhaNER 2.0: Africa-centric Transfer Learning for Named Entity Recognition.* In Conference on Empirical Methods in Natural Language Processing. [8]

24. Meyer, J., Adelani, D. I., Casanova, E., Oktem, A., **Muhammad, S. H.**, et al. (2022). *BibleTTS: A Large, High-Fidelity, Multilingual, and Uniquely African Speech Corpus.* arXiv preprint. [21]

25. **Muhammad, S. H.**, Brazdil, P., Jorge, A. M. (2022). *Symbolic Versus Deep Learning Techniques for Explainable Sentiment Analysis.* In Portuguese Conference on Artificial Intelligence. [29]

26. Adewumi, T. P., Adeyemi, M., Anuoluwapo, A., Peters, B., Buzaaba, H., et al. (2022). *AfriWOZ: Corpus for Exploiting Cross-Lingual Transferability for Natural Language Generation of Dialogues in Low-Resource, African Languages.* arXiv preprint. [10]

27. Abdulmumin, I., Dash, S. R., Dawud, M. A., Parida, S., **Muhammad, S. H.**, et al. (2022). *Hausa Visual Genome: A Dataset for Multi-Modal English to Hausa Machine Translation.* In International Conference on Language Resources and Evaluation. [2]

28. Aliyu, S. M., Wajiga, G. M., Murtala, M. T., **Muhammad, S. H.**, Abdulmumin, I., et al. (2022). *HERDPhobia: A Dataset for Hate Speech Against Fulani in Nigeria.* arXiv preprint. [14]

29. **Muhammad, S. H.**, Brazdil, P., Jorge, A. M. (2020). *Incremental Approach for Automatic Generation of Domain-Specific Sentiment Lexicon.* Advances in Information Retrieval, 12036, pp. 619-623. [27]

30. Ahmad, I. S., Abu Bakar, A., Yaakub, M. R., **Muhammad, S. H.** (2020). *A Survey on Machine Learning Techniques in Movie Revenue Prediction.* SN Computer Science, 1. [12]

31. Nekoto, W. O., Marivate, V., Matsila, T., Fasubaa, T. E., Kolawole, T., et al. (2020). *Participatory Research for Low-resourced Machine Translation: A Case Study in African Languages.* arXiv preprint, arXiv:2010.02353. [35]

32. Caswell, I., Kreutzer, J., Wang, L., Wahab, A., van Esch, D., et al. (2021). *Quality at a Glance: An Audit of Web-Crawled Multilingual Datasets.* Transactions of the Association for Computational Linguistics, 10, pp. 50-72. [17]

33. **Muhammad, S. H.**, Mustapha, A., Haruna, K. (2017). *Massive Open Online Courses: Awareness, Adoption, Benefits and Challenges in Sub-Saharan Africa.* In Proceedings of OcRI. [30]

34. Mustapha, A., **Muhammad, S. H.**, Salahudeen, S. A. (2016). *Massive Open Online Courses: A Success of Cloud Computing in Education.* In OcRI. [33]

35. Haruna, K., Baffa, M., **Muhammad, S. H.**, Abubakar, U. B. (2017). *A Framework for Implementation of E-Classroom System.* In OcRI. [20]

36. **Muhammad, S. H.**, Abdulmumin, I., Gwadabe, T. R., Alabi, J. O., et al. (2022). *NaijaSenti: A Nigerian Twitter Sentiment Corpus for Multilingual Sentiment Analysis.* In International Conference on Language Resources and Evaluation. [25]

37. Brazdil, P., **Muhammad, S. H.**, Oliveira, F. C. E., Cordeiro, J. P., Silva, F., et al. (2022). *Semi-Automatic Approaches for Exploiting Shifter Patterns in Domain-Specific Sentiment Analysis.* Mathematics. [16]

38. Adelani, D. I., Alabi, J. O., Fan, A., Kreutzer, J., Shen, X., et al. (2022). *A Few Thousand Translations Go a Long Way! Leveraging Pre-trained Models for African News Translation.* In North American Chapter of the Association for Computational Linguistics. [6]

39. **Muhammad, S. H.**, Brazdil, P., Jorge, A. M. (2023). *Symbolic Versus Deep Learning Techniques for Explainable Sentiment Analysis.* In Portuguese Conference on Artificial Intelligence. [29]

40. Abdulmumin, I., Khalid, A. A., **Muhammad, S. H.**, Ahmad, I. S., Aliyu, L. J., et al. (2023). *Leveraging Closed-Access Multilingual Embedding for Automatic Sentence Alignment in Low Resource Languages.* arXiv preprint. [3]

41. Adelani, D. I., Masiak, M., Azime, I. A., Alabi, J. O., Tonja, A. L., et al. (2023). *MasakhaNEWS: News Topic Classification for African Languages.* In International Joint Conference on Natural Language Processing. [7]

42. Wang, J., Adelani, D. I., Agrawal, S., Rei, R., Briakou, E., et al. (2023). *AfriMTE and AfriCOMET: Empowering COMET to Embrace Under-resourced African Languages.* arXiv preprint. [42]

43. Bayes, E., Azime, I. A., Alabi, J. O., Kgomo, J., Eloundou, T., et al. (2024). *Uhura: A Benchmark for Evaluating Scientific Question Answering and Truthfulness in Low-Resource African Languages.* arXiv preprint. [15]

44. Parida, S., Abdulmumin, I., **Muhammad, S. H.**, Ahmad, I. S., et al. (2023). *HaVQA: A Dataset for Visual Question Answering and Multimodal Research in Hausa Language.* In Annual Meeting of the Association for Computational Linguistics. [39]

45. Ousidhoum, N. D., **Muhammad, S. H.**, Abdalla, M., Abdulmumin, I., Ahmad, I. S., et al. (2024). *SemEval Task 1: Semantic Textual Relatedness for African and Asian Languages.* In International Workshop on Semantic Evaluation. [38]

46. Abdulmumin, I., Mkhwanazi, S., Mbooi, M. S., **Muhammad, S. H.**, Ahmad, I. S., et al. (2024). *Correcting FLORES Evaluation Dataset for Four African Languages.* In Conference on Machine Translation. [4]

47. Adewumi, T. P., Adeyemi, M., Anuoluwapo, A., Peters, B., Buzaaba, H., et al. (2022). *Itàkúròso: Exploiting Cross-Lingual Transferability for Natural Language Generation of Dialogues in Low-Resource, African Languages.* arXiv preprint. [10]

48. Adelani, D. I., Neubig, G., Ruder, S., Rijhwani, S., Beukman, M., et al. (2022). *MasakhaNER 2.0: Africa-centric Transfer Learning for Named Entity Recognition.* In Conference on Empirical Methods in Natural Language Processing. [8]

49. Adelani, D. I., Abbott, J. Z., Neubig, G., D'Souza, D., Kreutzer, J., et al. (2021). *MasakhaNER: Named Entity Recognition for African Languages.* Transactions of the Association for Computational Linguistics, 9, pp. 1116–1131. [5]

50. **Muhammad, S. H.**, Mustapha, A., Salahudeen, S. A. (2016). *Massive Open Online Courses: A Success of Cloud Computing in Education.* In OcRI. [33]

51. Haruna, K., Baffa, M., **Muhammad, S. H.**, Abubakar, U. B. (2017). *A Framework for Implementation of E-Classroom System.* In OcRI. [20]

References

[1] Idris Abdulmumin, Michael Beukman, Jesujoba Oluwadara Alabi, Chris C. Emezue, Everlyn Asiko, Tosin P. Adewumi, Shamsuddeen Hassan Muhammad, Mofetoluwa Adeyemi, Oreen Yousuf, Sahib Singh, and Tajuddeen Rabiu Gwadabe. Separating grains from the chaff: Using data filtering to improve multilingual translation for low-resourced african languages. *ArXiv*, abs/2210.10692, 2022.

[2] Idris Abdulmumin, Satya Ranjan Dash, Musa Abdullahi Dawud, Shantipriya Parida, Shamsuddeen Hassan Muhammad, Ibrahim Said Ahmad, Subhadarshi Panda, Ondrej Bojar, Bashir Shehu Galadanci, and Bello Shehu Bello. Hausa visual genome: A dataset for multi-modal english to hausa machine translation. In *International Conference on Language Resources and Evaluation*, 2022.

[3] Idris Abdulmumin, Auwal Abubakar Khalid, Shamsuddeen Hassan Muhammad, Ibrahim Said Ahmad, Lukman Jibril Aliyu, Babangida Sani, B.M. Abduljalil, and Sani Ahmad Hassan. Leveraging closed-access multilingual embedding for automatic sentence alignment in low resource languages. *ArXiv*, abs/2311.12179, 2023.

[4] Idris Abdulmumin, Sthembiso Mkhwanazi, Mahlatse S Mbooi, Shamsuddeen Hassan Muhammad, Ibrahim Said Ahmad, Neo Putini, Miehleketo Mathebula, Matimba Shingange, Tajuddeen Rabiu Gwadabe, and Vukosi Marivate. Correcting flores evaluation dataset for four african languages. In *Conference on Machine Translation*, 2024.

[5] David Ifeoluwa Adelani, Jade Z. 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 Hassan Muhammad, Chris C. Emezue, Joyce Nakatumba-Nabende, Perez Ogayo, Anuoluwapo Aremu, Catherine Gitau, Derguene Mbaye, Jesujoba Oluwadara Alabi, Seid Muhie Yimam, Tajuddeen Rabiu Gwadabe, Ignatius U. Ezeani, Andre Niyongabo Rubungo, Jonathan Mukiibi, Verrah Otiende, Iroro Orife, Davis

David, Samba Ngom, Tosin P. Adewumi, Paul Rayson, Mofetoluwa Adeyemi, Gerald Muriuki, Emmanuel Anebi, Chiamaka Ijeoma Chukwuneke, Nkiruka Bridget 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 N Faye, Blessing K. Sibanda, Orevaoghene Ahia, Bonaventure F. P. Dossou, Kelechi Ogueji, Thierno Ibrahima Diop, Abdoulaye Diallo, Adewale Akinfaderin, Tendai Munyaradzi Marengereke, and Salomey Osei. Masakhaner: Named entity recognition for african languages. *Transactions of the Association for Computational Linguistics*, 9:1116–1131, 2021.

[6] David Ifeoluwa Adelani, Jesujoba Oluwadara Alabi, Angela Fan, Julia Kreutzer, Xiaoyu Shen, Machel Reid, Dana Ruiter, Dietrich Klakow, Peter Nabende, Ernie Chang, Tajuddeen Rabiu Gwadabe, Freshia Sackey, Bonaventure F. P. Dossou, Chris C. 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, Tunde Oluwaseyi Ajayi, Yvonne Wambui Gitau, Jade Z. Abbott, Mohamed Ahmed, Millicent Ochieng, Anuoluwapo Aremu, Perez Ogayo, Jonathan Mukiibi, Fatoumata Ouoba Kabore, Godson Kalipe, Derguene Mbaye, Allahsera Auguste Tapo, Victoire Memdjokam Koagne, Edwin Munkoh-Buabeng, Valencia Wagner, Idris Abdulmumin, Ayodele Awokoya, Happy Buzaaba, Blessing K. Sibanda, Andiswa Bukula, and Sam Manthalu. A few thousand translations go a long way! leveraging pre-trained models for african news translation. In *North American Chapter of the Association for Computational Linguistics*, 2022.

[7] David Ifeoluwa Adelani, Marek Masiak, Israel Abebe Azime, Jesujoba Oluwadara Alabi, Atnafu Lambebo Tonja, Christine Mwase, Odunayo Ogundepo, Bonaventure F. P. Dossou, Akintunde Oladipo, Doreen Nixdorf, Chris C. Emezue, Sana Sabah Al-Azzawi, Blessing K. Sibanda, Davis David, Lolwethu Ndolela, Jonathan Mukiibi, Tunde Oluwaseyi Ajayi, Tatiana Moteu Ngoli, Brian Odhiambo, Abraham Toluwase Owodunni, Nnaemeka Obiefuna, Shamsuddeen Hassan Muhammad, Saheed Salahudeen Abdullahi, Mesay Gemeda Yigezu, Tajuddeen Rabiu Gwadabe, Idris Abdulmumin, Mahlet Taye Bame, Oluwabusayo Olufunke Awoyomi, Iyanuoluwa Shode, Tolulope Anu Adelani, Habiba Abdulganiy Kailani, Abdul-Hakeem Omotayo, Adetola Adeeko, Afolabi Abeeb, Anuoluwapo Aremu, Olanrewaju Samuel, Clemencia Siro, Wangari Kimotho, Onyekachi Raphael Ogbu, Chinedu E. Mbonu, Chiamaka Ijeoma Chukwuneke, Samuel Fanijo, Jessica Ojo, Oyinkansola F. Awosan, Tadesse Kebede Guge, Sakayo Toadoum Sari, Pamela Nyatsine, Freedmore Sidume, Oreen Yousuf, Mardiyyah Oduwole, Ussen Kimanuka, Kanda Patrick Tshinu, Thina Diko, Siyanda Nxakama, Abdulmejid Tuni Johar, Sinodos Gebre, Muhidin A. Mohamed, Shafie Abdi Mohamed, Fuad Mire Hassan, Moges Ahmed Mehamed, Evrard Ngabire, and Pontus Stenetorp. Masakhanews: News topic classification for african languages. In *International Joint Conference on Natural Language Processing*, 2023.

[8] David Ifeoluwa Adelani, Graham Neubig, Sebastian Ruder, Shruti Rijhwani, Michael Beukman, Chester Palen-Michel, Constantine Lignos, Jesujoba Oluwadara Alabi, Shamsuddeen Hassan Muhammad, Peter Nabende, Cheikh M. Bamba Dione, An-

diswa Bukula, Rooweither Mabuya, Bonaventure F. P. Dossou, Blessing K. Sibanda, Happy Buzaaba, Jonathan Mukiibi, Godson Kalipe, Derguene Mbaye, Amelia Taylor, Fatoumata Ouoba Kabore, Chris C. Emezue, Anuoluwapo Aremu, Perez Ogayo, Catherine W. Gitau, Edwin Munkoh-Buabeng, Victoire Memdjokam Koagne, Allahsera Auguste Tapo, Tebogo Macucwa, Vukosi Marivate, Elvis Mboning, Tajuddeen Rabiu Gwadabe, Tosin P. Adewumi, Orevaoghene Ahia, Joyce Nakatumba-Nabende, Neo L. Mokono, Ignatius M Ezeani, Chiamaka Ijeoma Chukwuneke, Mofetoluwa Adeyemi, Gilles Hacheme, Idris Abdulmumin, Odunayo Ogundepo, Oreen Yousuf, Tatiana Moteu Ngoli, and Dietrich Klakow. Masakhaner 2.0: Africa-centric transfer learning for named entity recognition. In *Conference on Empirical Methods in Natural Language Processing*, 2022.

[9] David Ifeoluwa Adelani, Jessica Ojo, Israel Abebe Azime, Zhuang Yun Jian, Jesujoba Oluwadara Alabi, Xuanli He, Millicent Ochieng, Sara Hooker, Andiswa Bukula, En-Shiun Annie Lee, Chiamaka Chukwuneke, Happy Buzaaba, Blessing K. Sibanda, Godson Kalipe, Jonathan Mukiibi, Salomon Kabongo KABENAMUALU, Foutse Yuehgoh, Mmasibidi Setaka, Lolwethu Ndolela, Nkiruka Bridget Odu, Rooweither Mabuya, Shamsuddeen Hassan Muhammad, Salomey Osei, Sokhar Samb, Tadesse Kebede Guge, and Pontus Stenetorp. Irokobench: A new benchmark for african languages in the age of large language models. *ArXiv*, abs/2406.03368, 2024.

[10] Tosin P. Adewumi, Mofetoluwa Adeyemi, Aremu Anuoluwapo, Bukola Peters, Happy Buzaaba, Oyerinde Samuel, Amina Mardiyyah Rufai, Benjamin Ayoade Ajibade, Tajudeen Gwadabe, Mory Moussou Koulibaly Traore, Tunde Oluwaseyi Ajayi, Shamsuddeen Hassan Muhammad, Ahmed Baruwa, Paul Owoicho, Tolúlopé Ògúnrèmí, Phylis Ngigi, Orevaoghene Ahia, Ruqayya Nasir, Foteini Simistira Liwicki, and Marcus Liwicki. Ìtàkúròso: Exploiting cross-lingual transferability for natural language generation of dialogues in low-resource, african languages. *ArXiv*, abs/2204.08083, 2022.

[11] Tosin P. Adewumi, Mofetoluwa Adeyemi, Aremu Anuoluwapo, Bukola Peters, Happy Buzaaba, Oyerinde Samuel, Amina Mardiyyah Rufai, Benjamin Ayoade Ajibade, Tajudeen Gwadabe, Mory Moussou Koulibaly Traore, Tunde Oluwaseyi Ajayi, Shamsuddeen Hassan Muhammad, Ahmed Baruwa, Paul Owoicho, Tolúlopé Ògúnrèmí, Phylis Ngigi, Orevaoghene Ahia, Ruqayya Nasir, Foteini Simistira Liwicki, and Marcus Liwicki. Afriwoz: Corpus for exploiting cross-lingual transfer for dialogue generation in low-resource, african languages. *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2023.

[12] Ibrahim Said Ahmad, Adeela Abu Bakar, Mohd Ridzwan Yaakub, and Shamsuddeen Hassan Muhammad. A survey on machine learning techniques in movie revenue prediction. *SN Computer Science*, 1, 2020.

[13] Saminu Mohammad Aliyu, Idris Abdulmumin, Shamsuddeen Hassan Muhammad, Ibrahim Said Ahmad, Saheed Abdullahi Salahudeen, Aliyu Yusuf, and Falalu Ibrahim Lawan. Hausanlp at semeval-2023 task 10: Transfer learning, synthetic data and side-information for multi-level sexism classification. *ArXiv*, abs/2305.00076, 2023.

[14] Saminu Mohammad Aliyu, Gregory M. Wajiga, Muhammad Tukur Murtala, Shamsuddeen Hassan Muhammad, Idris Abdulmumin, and Ibrahim Said Ahmad. Herdphobia: A dataset for hate speech against fulani in nigeria. *ArXiv*, abs/2211.15262, 2022.

[15] Edward Bayes, Israel Abebe Azime, Jesujoba Oluwadara Alabi, Jonas Kgomo, Tyna Eloundou, Elizabeth Proehl, Kai Chen, Imaan Khadir, Naome A. Etori, Shamsuddeen Hassan Muhammad, Choice Mpanza, Igneciah Pocia Thete, Dietrich Klakow, and David Ifeoluwa Adelani. Uhura: A benchmark for evaluating scientific question answering and truthfulness in low-resource african languages. *ArXiv*, abs/2412.00948, 2024.

[16] Pavel Brazdil, Shamsuddeen Hassan Muhammad, Fátima De Cassia E. Oliveira, João Paulo Cordeiro, Fátima Silva, Purificação Silvano, and Antonio Leal. Semi-automatic approaches for exploiting shifter patterns in domain-specific sentiment analysis. *Mathematics*, 2022.

[17] Isaac Caswell, Julia Kreutzer, Lisa Wang, Ahsan Wahab, Daan van Esch, Nasanbayar Ulzii-Orshikh, Allahsera Auguste Tapo, Nishant Subramani, Artem Sokolov, Claytone Sikasote, Monang Setyawan, Supheakmungkol Sarin, Sokhar Samb, Benoît Sagot, Clara Rivera, Annette Rios Gonzales, Isabel Papadimitriou, Salomey Osei, Pedro Ortiz Suarez, Iroro Orife, Kelechi Ogueji, Andre Niyongabo Rubungo, Toan Q. Nguyen, Mathias Muller, Andre Matthias Muller, Shamsuddeen Hassan Muhammad, Nanda Firdausi Muhammad, Ayanda Mnyakeni, Jamshidbek Mirzakhalov, Tapiwanashe Matangira, Colin Leong, Nze Lawson, Sneha Kudugunta, Yacine Jernite, M. Jenny, Orhan Firat, Bonaventure F. P. Dossou, Sakhile Dlamini, Nisansa de Silva, Sakine cCabuk Balli, Stella Biderman, Alessia Battisti, Ahmed Baruwa, Ankur Bapna, Pallavi N. Baljekar, Israel Abebe Azime, Ayodele Awokoya, Duygu Ataman, Orevaoghene Ahia, Oghenefego Ahia, Sweta Agrawal, and Mofetoluwa Adeyemi. Quality at a glance: An audit of web-crawled multilingual datasets. *Transactions of the Association for Computational Linguistics*, 10:50–72, 2021.

[18] Cheikh M. Bamba Dione, David Ifeoluwa Adelani, Peter Nabende, Jesujoba Oluwadara Alabi, Thapelo Sindane, Happy Buzaaba, Shamsuddeen Hassan Muhammad, Chris C. Emezue, Perez Ogayo, Anuoluwapo Aremu, Catherine Gitau, Derguene Mbaye, Jonathan Mukiibi, Blessing K. Sibanda, Bonaventure F. P. Dossou, Andiswa Bukula, Rooweither Mabuya, Allahsera Auguste Tapo, Edwin Munkoh-Buabeng, Victoire Memdjokam Koagne, Fatoumata Ouoba Kabore, Amelia Taylor, Godson Kalipe, Tebogo Macucwa, Vukosi Marivate, Tajuddeen Rabiu Gwadabe, Mboning Tchiaze Elvis, Ikechukwu E. Onyenwe, Gratien Gualbert Atindogbé, Tolulope Anu Adelani, Idris Akinade, Olanrewaju Samuel, Marie-Rosette Nahimana, Th'eogene Musabeyezu, Emile Niyomutabazi, Ester Chimhenga, Kudzai Gotosa, Patrick Mizha, Apelete Agbolo, Seydou T. Traoré, Chinedu Uchechukwu, Aliyu Yusuf, Muhammad Abubakar Abdullahi, and Dietrich Klakow. Masakhapos: Part-of-speech tagging for typologically diverse african languages. In *Annual Meeting of the Association for Computational Linguistics*, 2023.

[19] Chris C. Emezue, Hellina Hailu Nigatu, Cynthia Thinwa, He Zhou, Shamsuddeen Hassan Muhammad, Lerato Louis, Idris Abdulmumin, Samuel Oyerinde, Benjamin Ayoade Ajibade, Olanrewaju Samuel, Oviawe Joshua, Emeka Onwuegbuzia,

Handel Emezue, Ifeoluwatayo A. Ige, Atnafu Lambebo Tonja, Chiamaka Ijeoma Chukwuneke, Bonaventure F. P. Dossou, Naome A. Etori, Mbonu Chinedu Emmanuel, Oreen Yousuf, Kaosarat Aina, and Davis David. The african stopwords project: curating stopwords for african languages. *ArXiv*, abs/2304.12155, 2023.

[20] Khalid Haruna, Baffa, Shamsuddeen Hassan Muhammad, and U. B. Abubakar. A framework for implementation of e-classroom system. 2017.

[21] Josh Meyer, David Ifeoluwa Adelani, Edresson Casanova, Alp Oktem, Daniel Whitenack Julian Weber, Salomon Kabongo KABENAMUALU, Elizabeth Salesky, Iroro Orife, Colin Leong, Perez Ogayo, Chris C. Emezue, Jonathan Mukiibi, Salomey Osei, Apelete Agbolo, Victor Akinode, Bernard Opoku, Samuel Olayemi Olanrewaju, Jesujoba Oluwadara Alabi, and Shamsuddeen Hassan Muhammad. Bibletts: a large, high-fidelity, multilingual, and uniquely african speech corpus. *ArXiv*, abs/2207.03546, 2022.

[22] Shamsuddeen Muhammad, Idris Abdulmumin, Abinew Ayele, Nedjma Ousidhoum, David Adelani, Seid Yimam, Ibrahim Ahmad, Meriem Beloucif, Saif Mohammad, Sebastian Ruder, Oumaima Hourrane, Alipio Jorge, Pavel Brazdil, Felermino Ali, Davis David, Salomey Osei, Bello Shehu-Bello, Falalu Lawan, Tajuddeen Gwadabe, Samuel Rutunda, Tadesse Destaw Belay, Wendimu Messelle, Hailu Balcha, Sisay Chala, Hagos Gebremichael, Bernard Opoku, and Stephen Arthur. AfriSenti: A Twitter sentiment analysis benchmark for African languages. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 13968–13981, Singapore, December 2023. Association for Computational Linguistics.

[23] Shamsuddeen Hassan Muhammad. An overview of sentiment analysis approaches. 2019.

[24] Shamsuddeen Hassan Muhammad, Idris Abdulmumin, Abinew Ali Ayele, Nedjma Djouhra Ousidhoum, David Ifeoluwa Adelani, Seid Muhie Yimam, Ibrahim Said Ahmad, Meriem Beloucif, Saif M. Mohammad, Sebastian Ruder, Oumaima Hourrane, Pavel Brazdil, Felermino D'ario M'ario Ant'onio Ali, Davis C. Davis, Salomey Osei, Bello Shehu Bello, Falalu Ibrahim, Tajuddeen Rabiu Gwadabe, Samuel Rutunda, Tadesse Destaw Belay, Wendimu Baye Messelle, Hailu Beshada Balcha, Sisay Adugna Chala, Hagos Tesfahun Gebremichael, Bernard Opoku, and Steven Arthur. Afrisenti: A twitter sentiment analysis benchmark for african languages. *ArXiv*, abs/2302.08956, 2023.

[25] Shamsuddeen Hassan Muhammad, David Ifeoluwa Adelani, Ibrahim Said Ahmad, Idris Abdulmumin, Bello Shehu Bello, Monojit Choudhury, Chris C. Emezue, Anuoluwapo Aremu, Saheed Abdul, and Pavel Brazdil. Naijasenti: A nigerian twitter sentiment corpus for multilingual sentiment analysis. In *International Conference on Language Resources and Evaluation*, 2022.

[26] Shamsuddeen Hassan Muhammad, David Ifeoluwa Adelani, Sebastian Ruder, Ibrahim Sa'id Ahmad, Idris Abdulmumin, Bello Shehu Bello, Monojit Choudhury, Chris Chinenye Emezue, Saheed Salahudeen Abdullahi, Anuoluwapo Aremu, Alípio Jorge, and Pavel Brazdil. NaijaSenti: A Nigerian Twitter sentiment corpus for multilingual sentiment analysis. In Nicoletta Calzolari, Frédéric Béchet, Philippe Blache,

Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 590–602, Marseille, France, June 2022. European Language Resources Association.

[27] Shamsuddeen Hassan Muhammad, Pavel Brazdil, and Alípio Mário Jorge. Incremental approach for automatic generation of domain-specific sentiment lexicon. *Advances in Information Retrieval*, 12036:619 – 623, 2020.

[28] Shamsuddeen Hassan Muhammad, Pavel Brazdil, and Alípio Mário Jorge. Combining symbolic and deep learning approaches for sentiment analysis. In *Compendium of Neurosymbolic Artificial Intelligence*, 2023.

[29] Shamsuddeen Hassan Muhammad, Pavel Brazdil, and Alípio Mário Jorge. Symbolic versus deep learning techniques for explainable sentiment analysis. In *Portuguese Conference on Artificial Intelligence*, 2023.

[30] Shamsuddeen Hassan Muhammad, Abdulrasheed Mustapha, and Khalid Haruna. Massive open online courses: Awareness, adoption, benefits and challenges in sub-saharan africa. 2017.

[31] Shamsuddeen Hassan Muhammad, Nedjma Ousidhoum, Idris Abdulmumin, Jan Philip Wahle, Terry Ruas, Meriem Beloucif, Christine de Kock, Nirmal Surange, Daniela Teodorescu, Ibrahim Said Ahmad, David Ifeoluwa Adelani, Alham Fikri Aji, Felermino D. M. A. Ali, Ilseyar Alimova, Vladimir Araujo, Nikolay Babakov, Naomi Baes, Ana-Maria Bucur, Andiswa Bukula, Guanqun Cao, Rodrigo Tufino Cardenas, Rendi Chevi, Chiamaka Ijeoma Chukwuneke, Alexandra Ciobotaru, Daryna Dementieva, Murja Sani Gadanya, Robert Geislinger, Bela Gipp, Oumaima Hourrane, Oana Ignat, Falalu Ibrahim Lawan, Rooweither Mabuya, Rahmad Mahendra, Vukosi Marivate, Andrew Piper, Alexander Panchenko, Charles Henrique Porto Ferreira, Vitaly Protasov, Samuel Rutunda, Manish Shrivastava, Aura Cristina Udrea, Lilian Diana Awuor Wanzare, Sophie Wu, Florian Valentin Wunderlich, Hanif Muhammad Zhafran, Tianhui Zhang, Yi Zhou, and Saif M. Mohammad. Brighter: Bridging the gap in human-annotated textual emotion recognition datasets for 28 languages, 2025.

[32] Shamsuddeen Hassan Muhammad, Nedjma Ousidhoum, Idris Abdulmumin, Seid Muhie Yimam, Jan Philip Wahle, Terry Ruas, Meriem Beloucif, Christine De Kock, Tadesse Destaw Belay, Ibrahim Said Ahmad, Nirmal Surange, Daniela Teodorescu, David Ifeoluwa Adelani, Alham Fikri Aji, Felermino Ali, Vladimir Araujo, Abinew Ali Ayele, Oana Ignat, Alexander Panchenko, Yi Zhou, and Saif M. Mohammad. SemEval task 11: Bridging the gap in text-based emotion detection. In *Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval-2025)*, Vienna, Austria, 2025. Association for Computational Linguistics.

[33] Abdulrasheed Mustapha, Shamsuddeen Hassan Muhammad, and Saheed Abdullahi Salahudeen. Massive open online courses: A success of cloud computing in education. In *OcRI*, 2016.

[34] Junho Myung, Nayeon Lee, Yi Zhou, Jiho Jin, Rifki Afina Putri, Dimosthenis Antypas, Hsuvas Borkakoty, Eunsu Kim, Carla Pérez-Almendros, Abinew Ali Ayele, V'ictor Guti'errez-Basulto, Yazm'in Ib'anez-Garc'ia, Hwaran Lee, Shamsuddeen Hassan Muhammad, Kiwoong Park, Anar Rzayev, Nina White, Seid Muhie Yimam, Mohammad Taher Pilehvar, Nedjma Djouhra Ousidhoum, José Camacho-Collados, and Alice Oh. Blend: A benchmark for llms on everyday knowledge in diverse cultures and languages. *ArXiv*, abs/2406.09948, 2024.

[35] Wilhelmina Onyothi Nekoto, Vukosi Marivate, Tshinondiwa Matsila, Timi E. Fasubaa, T Kolawole, Taiwo Helen Fagbohungbe, Solomon Oluwole Akinola, Shamsuddeen Hassan Muhammad, Salomon Kabongo KABENAMUALU, Salomey Osei, Sackey Freshia, Andre Niyongabo Rubungo, Ricky Macharm, Perez Ogayo, Orevaoghene Ahia, Musie Meressa, Mofetoluwa Adeyemi, Masabata Mokgesi-Selinga, Lawrence Okegbemi, Laura Martinus, Kolawole Tajudeen, Kevin Degila, Kelechi Ogueji, Kathleen Siminyu, Julia Kreutzer, Jason Webster, Jamiil Toure Ali, Jade Z. Abbott, Iroro Orife, Ignatius U. Ezeani, Idris Abdulkabir Dangana, Herman Kamper, Hady ElSahar, Goodness Duru, Ghollah Kioko, Espoir Murhabazi, Elan Van Biljon, Daniel Whitenack, Christopher Onyefuluchi, Chris C. Emezue, Bonaventure F. P. Dossou, Blessing K. Sibanda, Blessing Itoro Bassey, Ayodele Olabiyi, Arshath Ramkilowan, Alp Oktem, Adewale Akinfaderin, and Abdallah Bashir. Participatory research for low-resourced machine translation: A case study in african languages. *ArXiv*, abs/2010.02353, 2020.

[36] Odunayo Ogundepo, Tajuddeen Rabiu Gwadabe, Clara Rivera, J. Clark, Sebastian Ruder, David Ifeoluwa Adelani, Bonaventure F. P. Dossou, Abdoulahat Diop, Claytone Sikasote, Gilles Hacheme, Happy Buzaaba, Ignatius M Ezeani, Rooweither Mabuya, Salomey Osei, Chris C. Emezue, Albert Njoroge Kahira, Shamsuddeen Hassan Muhammad, Akintunde Oladipo, Abraham Toluwase Owodunni, Atnafu Lambebo Tonja, Iyanuoluwa Shode, Akari Asai, Tunde Oluwaseyi Ajayi, Clemencia Siro, Steven Arthur, Mofetoluwa Adeyemi, Orevaoghene Ahia, Aremu Anuoluwapo, Oyinkansola Awosan, Chiamaka Ijeoma Chukwuneke, Bernard Opoku, Awo Maria Ayodele, Verrah Otiende, Christine Mwase, Boyd Sinkala, Andre Niyongabo Rubungo, Daniel Ajisafe, Emeka Onwuegbuzia, Habib Mbow, Emile Niyomutabazi, Eunice Mukonde, Falalu Ibrahim Lawan, Ibrahim Said Ahmad, Jesujoba Oluwadara Alabi, Martin Namukombo, Mbonu Chinedu, Mofya Phiri, Neo Putini, Ndumiso Mngoma, Priscilla Amuok, Ruqayya Nasir Iro, and Sonia Adhiambo34. Afriqa: Cross-lingual open-retrieval question answering for african languages. In *Conference on Empirical Methods in Natural Language Processing*, 2023.

[37] Nedjma Djouhra 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, Christian Biemann, Sofia Bourhim, Christine de Kock, Genet Shanko Dekebo, Oumaima Hourrane, Gopichand Kanumolu, Lokesh Madasu, Samuel Rutunda, Manish Shrivastava, Thamar Solorio, Nirmal Surange, Hailegnaw Getaneh Tilaye, Krishnapriya Vishnubhotla, Genta Indra Winata, Seid Muhie Yimam, and Saif Mohammad. Semrel2024: A collection of semantic textual relatedness datasets for 14 languages. In *Annual Meeting of the Association for Computational Linguistics*, 2024.

[38] Nedjma Djouhra Ousidhoum, Shamsuddeen Hassan Muhammad, Mohamed Abdalla, Idris Abdulmumin, Ibrahim Said Ahmad, Sanchit Ahuja, Alham Fikri Aji, Vladimir Araujo, Meriem Beloucif, Christine de Kock, Oumaima Hourrane, Manish Shrivastava, Thamar Solorio, Nirmal Surange, Krishnapriya Vishnubhotla, Seid Muhie Yimam, and Saif Mohammad. Semeval task 1: Semantic textual relatedness for african and asian languages. In *International Workshop on Semantic Evaluation*, 2024.

[39] Shantipriya Parida, Idris Abdulmumin, Shamsuddeen Hassan Muhammad, Aneesh Bose, Guneet Singh Kohli, Ibrahim Said Ahmad, Ketan Kotwal, Sayan Deb Sarkar, Ondrej Bojar, and Habeebah Adamu Kakudi. Havqa: A dataset for visual question answering and multimodal research in hausa language. In *Annual Meeting of the Association for Computational Linguistics*, 2023.

[40] Shantipriya Parida, Ondrej Bojar, Idris Abdulmumin, Shamsuddeen Hassan Muhammad, and Ibrahim Said Ahmad. Findings of wmt2024 english-to-low resource multimodal translation task. In *Conference on Machine Translation*, 2024.

[41] Francielle Alves Vargas, Samuel Guimarães, **Shamsuddeen Hassan Muhammad**, Diego Alves, Ibrahim Said Ahmad, Idris Abdulmumin, Diallo Mohamed, Thiago A. S. Pardo, and Fabrício Benevenuto. Hausahate: An expert annotated corpus for hausa hate speech detection. *Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH 2024)*, 2024.

[42] Jiayi Wang, David Ifeoluwa Adelani, Sweta Agrawal, Ricardo Rei, Eleftheria Briakou, Marine Carpuat, Marek Masiak, Xuanli He, Sofia Bourhim, Andiswa Bukula, Muhidin A. Mohamed, Temitayo Olatoye, Hamam Mokayede, Christine Mwase, Wangui Kimotho, Foutse Yuehgoh, Anuoluwapo Aremu, Jessica Ojo, Shamsuddeen Hassan Muhammad, Salomey Osei, Abdul-Hakeem Omotayo, Chiamaka Chukwuneke, Perez Ogayo, Oumaima Hourrane, Salma El Anigri, Lolwethu Ndolela, Thabiso Mangwana, Shafie Abdi Mohamed, Ayinde Hassan, Oluwabusayo Olufunke Awoyomi, Lama Alkhaled, Sana Sabah Al-Azzawi, Naome A. Etori, Millicent Ochieng, Clemencia Siro, Samuel Njoroge, Eric Muchiri, Wangari Kimotho, Lyse Naomi Wamba Momo, Daud Olamide Abolade, Simbiat Ajao, Tosin P. Adewumi, Iyanuoluwa Shode, Ricky Macharm, Ruqayya Nasir Iro, Saheed Salahudeen Abdullahi, Stephen E. Moore, Bernard Opoku, Zainab Akinjobi, Abeeb Afolabi, Nnaemeka Obiefuna, Onyekachi Raphael Ogbu, Sam Brian, Verrah Otiende, Chinedu E. Mbonu, Sakayo Toadoum Sari, and Pontus Stenetorp. Afrimte and africomet: Empowering comet to embrace under-resourced african languages. *ArXiv*, abs/2311.09828, 2023.