

# H PPY BUZ B

Postdoctoral Researcher Princeton University

Email:hbuaaba@gmail.com

## RESEARCH INTEREST

My current research focuses on developing technologies for African languages. In the past, I spent time investigating the use of natural-gradient Bayesian methods to improve uncertainty estimation in pre-trained language models and how to make them robust and adaptable with a keen interest in low resource African languages.

## EMPLOYMENT

### Postdoctoral Research Associate

Princeton University

From September 2023

Princeton NJ, USA

### Postdoctoral Researcher, Approximate Bayesian Inference Team

RIKEN Center for Advanced Intelligence Project

From June 2022 - August 2023

Tokyo, Japan

### Ph.D Student, Systems and Information Engineering

University of Tsukuba

April 2018 - March 2022

Tsukuba, Japan

## EDUCATION

### Ph.D Computer Science, University of Tsukuba

Supervisor: Prof. Toshiyuki Amagasa.

April 2018 - March 2022

### Master of Regional Development, Toyo University

Supervisor: Prof. Nakabasami Chieko.

Sept 2015 – Sept 2017

### Bachelor of Applied Physics, UR College of Science and Technology (Former KIST)

Thesis Supervisor: Prof. Pogrebnoi Alexander.

Jan 2008 - Mar 2012

## LIST OF PUBLICATIONS

1. D. Adelani, J. Ojo, I. Azime, J. Zhuang, J. Alabi, X. He, M. Ochieng, S. Hooker, A. Bukula, E. Lee, C. Chukwuneke, **H. Buzaaba**, B. Sibanda, G. Kalipe, J. Mukiibi, S. Kabongo, F. Yuehgoh, M. Setaka, L. Ndolela, N. Odu, R. Mabuya, S. Muhammad, S. Osei, S. Samb, T. Guge, T. Sherman, P. Stenetorp. IrokoBench: A New Benchmark for African Languages in the Age of Large Language Models *In the international Conference of the Nations of Americas Chapter of the Association for Computational Linguistics NAACL), 2025*. To appear
2. N. Ermolaev, U. Inyanga, T. Olatoye, **H. Buzaaba**., Computational Approaches to Nigerian Literature: Analyzing Texts in Yoruba and E k. *Alliance of Digital Humanities Organizations international professional/disciplinary conference for 2024 ADHO2024*).
3. O. Ogundepo, T. Gwadabe, C. Rivera, J. Clark, S. Ruder, D. Adelani, B. Dossou, A. DIOP, C. Sikasote, G. Hacheme, **H. Buzaaba**, I. Ezeani, R. Mabuya, S. Osei, C. Emezue, A. Kahira, S. Muhammad, A. Oladipo, A. Owodunni, A. Tonja, I. Shode, A. Asai, T. Ajayi, C. Siro, S. Arthur, M. Adeyemi, O. Ahia, A. Anuoluwapo, O. Awosan, C. Chukwuneke, B. Opoku, A. Ayodele, V. Otiende, C. Mwase, B. Sinkala, A. Rubungo, D. Ajisafe, E. Onwuegbuzia, H. Mbow, E. Niyomutabazi, E. Mukonde, F. Lawan, I. Ahmad, J. Alabi, M. Namukombo, M. Chinedu, M. Phiri, N. Putini, N. Mngoma, P. Amuok, R. Iro, S. Adhiambo. AfriQA: Cross-lingual Open-Retrieval Question Answering for African Languages. *In the International Conference on Empirical Methods in Natural Language Processing EMNLP2023*).
4. C. Bamba, D. Adelani, P. Nabende, J. Alabi, T. Sindane, **H. Buzaaba**, S. Muhammad, C. Emezue, P. Ogayo, A. Aremu, C. Gitau, D. Mbaye, J. Mukiibi, B. Sibanda, B. F. P. Dossou, A. Bukula, R. Mabuya, A. Tapo, E. Munkoh-Buabeng, V. Koagne, F. Kabore, A. Taylor, G. Kalipe, T. Macucwa, V. Marivate, T. Gwadabe, E. Mboning, I. Onyenwe, G. Atindogbe, T. Adelani, I. Akinade, O. Samuel, M. Nahimana, T. Musabeyezu, E. Niyomutabazi, E. Chimhenga, K. Gotosa, P. Mizha, A. Agbolo, S. Traore, C. Uchechukwu, A. M. Abdullahi, D. Klakow. MasakhaPOS: Part-of-Speech Tagging for Typologically Diverse African languages. *In the 61st Annual Meeting of the Association for Computational Linguistics ACL2023*).

5. T. Adewumi, M. Adeyemi, A. Anuoluwapo, B. Peters, **H. Buzaaba**, O. Samuel, A. M. Rufai, B. Ajibade, T. Gwadabe, M. K. Traore, T. Ajayi, S. Muhammad, A. Baruwa, P. Owoicho, T. Ogunremi, P. Ngigi, O. Ahia, R. Nasir, F. Liwicki, M. Liwicki. AfriWOZ: Corpus for Exploiting Cross-Lingual Transferability for Generation of Dialogues in Low-Resource, African Languages. *In the International Joint Conference on Neural Networks IJCNN 2023*).
6. D. Adelani, G. Neubig, S. Ruder, S. Rijhwani, M. Beukman, C. Palen-Michel, C. Lignos, J. Alabi, S. Muhammad, P. Nabende, C. M. B. Dione, A. Bukula, R. Mabuya, B. F. P. Dossou, B. Sibanda, **H. Buzaaba**, J. Mukiibi, G. Kalipe, D. Mbaye, A. Taylor, F. Kabore, C. C. Emezue, A. Aremu, P. Ogayo, C. Gitau, E. Munkoh-Buabeng, V. M. Koagne, A. A. Tapo, T. Macucwa, V. Marivate, M. T. Elvis, T. Gwadabe, T. Adewumi, O. Ahia, J. Nakatumba-Nabende, N. L. Mokono, I. Ezeani, C. Chukwuneke, M. O. Adeyemi, G. Q. Hacheme, I. Abdulmumin, O. Ogundepo, O. Yousuf, T. Moteu, D. Klakow. MasakhaNER 2.0: Africa-centric Transfer Learning for Named Entity Recognition. *In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP*), 2022.
7. D. Adelani, J. Alabi, A. Fan, J. Kreutzer, X. Shen, M. Reid, D. Ruiters, D. Klakow, P. Nabende, E. Chang, T. Gwadabe, F. Sackey, B. F. P. Dossou, C. Emezue, C. Leong, M. Beukman, S. Muhammad, G. Jarso, O. Yousuf, A. N. Rubungo, G. Hacheme, E. P. Wairagala, M. U. Nasir, B. Ajibade, T. Ajayi, Y. Gitau, J. Abbott, M. Ahmed, M. Ochieng, A. Aremu, P. Ogayo, J. Mukiibi, F. O. Kabore, G. Kalipe, D. Mbaye, A. A. Tapo, V. M. Koagne, E. Munkoh-Buabeng, V. Wagner, I. Abdulmumin, A. Awokoya, **H. Buzaaba**, B. Sibanda, A. Bukula, S. Manthalu. A Few Thousand Translations Go a Long Way! Leveraging Pre-trained Models for African News Translation. *In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics NAACL*), 2022.
8. **H. Buzaaba**, T. Amagasa. A scheme for efficient question answering with low dimension reconstructed embeddings. *In The 23rd international conference on information integration and web intelligence iiWAS*) 2021.
9. D. I. Adelani, J. Abbott, G. Neubig, D. D'souza, J. Kreutzer, C. Lignos, C. Palen-Michel, **H. Buzaaba**, S. Rijhwani, S. Ruder, S. Mayhew, I. A. Azime, S. H. Muhammad, C. C. Emezue, J. Nakatumba-Nabende, P. Ogayo, A. Anuoluwapo, C. Gitau, D. Mbaye, J. Alabi, S. M. Yimam, T. R. Gwadabe, I. Ezeani, R. A. Niyongabo, J. Mukiibi, V. Otiende, I. Orife, D. David, S. Ngom, T. Adewumi, P. Rayson, M. Adeyemi, G. Muriuki, E. Anebi, C. Chukwuneke, N. Odu, E. P. Wairagala, S. Oyerinde, C. Siro, T. S. Bateesa, T. Oloyede, Y. Wambui, V. Akinode, D. Nabagereka, M. Katusiime, A. Awokoya, M. MBOUP, D. Gebreyohannes, H. Tilaye, K. Nwaike, D. Wolde, A. Faye, B. Sibanda, O. Ahia, B. F. P. Dossou, K. Ogueji, T. I. DIOP, A. Diallo, A. Akinfaderin, T. Marengereke, S. Osei. MasakhaNER: Named Entity Recognition for African Languages. *Transactions of the Association for Computational Linguistics* 2021) 9: 1116–1131.
10. **H. Buzaaba**, T. Amagasa. Question Answering Over Knowledge Base: A Scheme for Integrating Subject and the Identified Relation to Answer Simple Questions. *SN COMPUT. SCI.* 2, 25 2021).
11. **H. Buzaaba**, T. Amagasa. A modular approach for efficient simple question answering over knowledge base. *In the 30th international conference on Database and expert systems applications DEXA*), 2019.

---

## TEACHING

**FRS159 Teaching Computers to Understand African Languages**

Course Instructor Princeton University

Fall2024

Princeton, NJ

**Data Engineering 1 for master's students**

Teaching Assistant University of Tsukuba

Oct – Dec 2019

Tsukuba, Japan

**Database systems for undergraduate students**

Teaching Assistant University of Tsukuba

Oct – Dec 2018

Tsukuba, Japan

---

## REVIEWING

<b>2025</b>	COLING
<b>2024</b>	LREC-COLING
<b>2023</b>	NeurIPS
<b>2022</b>	NeurIPS, ACML
<b>2021</b>	ICLR

## OTHER ACTIVITIES

---

<b>2025</b>	Co-organizing the ACL2025 workshop on AfricaNLP.
<b>2024</b>	Co-organizing the ICLR2024 workshop on AfricaNLP.
<b>2023</b>	Co-organizing the ICLR2023 workshop on AfricaNLP.

## PRIZES      AWARDS

---

1. 2019: **Merit Award**, Students award for best presenter at the 11th Forum on Data Engineering and Information Management DEIM 2019.
2. 2017: **Mastuo Award**, an award for best master thesis of Regional Development Studies, by Toyo University.
3. 2015: **Abe-Initiative Scholarship**, a fully funded masters scholarship awarded to young Africans by the Japanese government.