



# Aman Berhe

**Date of birth:** 08/03/1989 | **Phone number:** (+33) 667836342 (Mobile) |

**Email address:** [bamanzaid@gmail.com](mailto:bamanzaid@gmail.com) |

**Address:** 8 Rue Vernier, 7501, Paris, France (Work)

## WORK EXPERIENCE

01/03/2022 – CURRENT Paris, France

**POSTDOC** IRD/UMMISCO AND QUINTEN, SORBONNE UNIVERSITY

- NLP for health. Development of French language models for biomedicine ([AliBERT](#))
- Medical documents structurization leveraging LLMs
- Automatic Systematic Literature Review (ASLR) for bacteria signatures in gut microbiome
- Medical code (ICD code) builder
- Self-Aligned Pretraining for French Biomedical Pretrained Language Models B-PLMs
- Medical Entity Mapping (MEL) to UMLS codes in French

01/09/2020 – 31/07/2021 Orsay, France

**ATER** UNIVERISTY PARIS SACLAY UNIVERSITY PARIS SACLAY IN IUT D'ORSAY

Teaching computer science: medium of interaction in French and English

01/09/2018 – 31/07/2020 Orsay, France

**PHD-TEACHER** UNIVERSITY PARIS SACLAY IN IUT D'ORSAY

Teaching computer science: medium of interaction in French and English

01/02/2017 – 31/07/2017 Nancy, France

**INTERNSHIP** INRIA/LORIA

Black-box optimization of DNNs for acoustic modeling. AM for ASR based on DNNs, including LSTM, highway networks and wide residual networks, hyper-parameter optimization, speech processing

01/04/2016 – 31/07/2016 Caen, France

**INTERNSHIP** GREYC-CNRS UMR 6072

Classifying sentences with symptoms of diseases which includes tree tagging, scrapping, text processing, sequence mining

01/09/2013 – 31/08/2015 Mekelle, Ethiopia

**ASSISTANT LECTURER** MEKELLE UNIVERSITY

Teaching computer science

## EDUCATION AND TRAINING

01/01/2018 – 28/10/2021 Orsay, France

**PHD. COMPUTER SCIENCE** LIMSI/CNRS Université Paris-Saclay

Machine Learning and Natural Language Processing (NLP)

**Address** Campus Universitaire bâtiment 507, Rue du Belvédère, 91400, Orsay, France |

**Field of study** Information and Communication Technologies |

**Thesis** Extraction and Description of the Narrative Structure from TV Series: supervision by Claude Barras and Camille Guinaudeau

18/02/2019 – 27/02/2019 Orsay, France

**SLURM SEQUENCER TO ACCESS COMPUTING CLUSTERS** Sciences et Technologies de l'Information et de la Communication

---

**Address** 91400, Orsay, France

01/09/2015 – 31/07/2017 Saint-Étienne, France

**M.SC. COMPUTER SCIENCE** Université Jean Monnet

---

**Address** 10 Rue Tréfilerie, 42100, Saint-Étienne, France

01/09/2007 – 31/07/2012 Mekelle, Ethiopia

**B.SC. COMPUTER SCIENCE AND ENGINEERING** Mekelle Institute of Technology

---

01/09/2005 – 30/06/2007 Mekelle, Ethiopia

**ETHIOPIAN HIGHER EDUCATION ENTRANCE EXAMINATION (EHEEE)** Aste-Yohannes Preparatory School

---

## ● LANGUAGE SKILLS

---

Mother tongue(s): **TIGRIGNA** | **AMHARIC**

Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken production	Spoken interaction	
<b>ENGLISH</b>	C2	C2	C1	C1	C1
<b>FRENCH</b>	C1	C1	B2	B2	B2

*Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user*

## ● DIGITAL SKILLS

---

Deep Learning frameworks: PyTorch, Tensorflow, Keras | Programming: Python, Java, C++, HTML, CSS, Angular JS, ..

## ● PUBLICATIONS

---

2023

**[AlIBERT: A Pre-trained Language Model for French Biomedical Text](#)**

---

Write here the description...

man Berhe, Guillaume Draznieks, Vincent Martenot, Jean-Daniel Zucker et al.; In BioNLP, ACL.

**Link** <https://aclanthology.org/2023.bionlp-1.19/#:~:text=AliBERT has achieved state of,and with much smaller corpora.>

2019

**Video Scene Segmentation of TV Series Using Multimodal Neural Features**

---

Aman Berhe, Camille G., Claude B., Series-International Journal of TV Serial Narratives

**Link** [https://scholar.google.com/citations?](https://scholar.google.com/citations?view_op=view_citation&hl=en&user=Aw6sC1AAAAAJ&citation_for_view=Aw6sC1AAAAAJ:u5HHmVD_uO8C)

[view\\_op=view\\_citation&hl=en&user=Aw6sC1AAAAAJ&citation\\_for\\_view=Aw6sC1AAAAAJ:u5HHmVD\\_uO8C](https://scholar.google.com/citations?view_op=view_citation&hl=en&user=Aw6sC1AAAAAJ&citation_for_view=Aw6sC1AAAAAJ:u5HHmVD_uO8C)

2021

**Scene linking annotation and automatic scene characterization in tv series**

---

Aman Berhe, Camille G., Claude B., CEUR Workshop Proceedings

**Link** [https://scholar.google.com/citations?](https://scholar.google.com/citations?view_op=view_citation&hl=en&user=Aw6sC1AAAAAJ&citation_for_view=Aw6sC1AAAAAJ:d1gkVwhDpl0C)

[view\\_op=view\\_citation&hl=en&user=Aw6sC1AAAAAJ&citation\\_for\\_view=Aw6sC1AAAAAJ:d1gkVwhDpl0C](https://scholar.google.com/citations?view_op=view_citation&hl=en&user=Aw6sC1AAAAAJ&citation_for_view=Aw6sC1AAAAAJ:d1gkVwhDpl0C)

2024

**Onco-AliBERT: Extracting Oncological Concepts From French Clinical Notes (SUBMITTED)**

---

Aman Berhe, Pauline Guilin, Vincent Martinet, et.al

2024

**DeepASLR: Automatic Systematic Literature Review Based on LLMs for Microbiome Signature Evidence (UNDER REVIEW)**

---

Performing automatic systematic literature review for bacteria signature and building knowledge graph between different bacteria species, diseases and chemicals describing their relationship.

Aman Berhe, Jean-Daniel, Eugeni Belds, et al.

2021

**Détection de scènes remarquables dans un contexte des séries TV**

---

Aman Berhe, Camille G., Claude B., CORIA