



Lydia Nishimwe


DOCTORANTE EN TRADUCTION AUTOMATIQUE

Ma thèse porte sur la **robustesse** de la **traduction automatique neuronale** face aux **textes non standards** issus des **contenus générés par les utilisateurs** sur les réseaux sociaux.

Domaines de recherche

- Traduction automatique
- Normalisation lexicale
- Modèles de langue
- Apprentissage profond

 lydia.nishimwe@inria.fr

 lydianishimwe.info

 [linkedin.com/in/lydia-nishimwe](https://www.linkedin.com/in/lydia-nishimwe)

 twitter.com/LydiaNishimwe

 github.com/lydianish

Formation académique

■ Doctorat en Informatique

Centre Inria de Paris / Sorbonne Université

Octobre 2021 – Présent

Paris, France

- Sujet de thèse : **Traduction automatique neuronale robuste**
- Directeurs de thèse : **Benoît Sagot** et **Rachel Bawden**

■ Diplôme d'ingénieure

École Centrale de Nantes

Septembre 2017 – Mai 2021

Nantes, France

Options disciplinaires :

- 3ème année : **Mathématiques et Applications**
 - Filière : **Statistique et Science des données**
 - Option professionnelle : **R&D**
- 2ème année : **Informatique**
 - Semestre d'échange à l'Université Polytechnique de Madrid : **Organisation Industrielle**

■ Licence Mathématiques et Informatique

Université Grenoble Alpes

Septembre 2014 – Juillet 2017

Grenoble, France

Expériences professionnelles

■ Stage – Recherche et Développement

Orange Labs

Juin 2020 – Décembre 2020

Lannion, France

- Intitulé : **Inférence de séquences masquées**
- Environnement : LaTeX, Python, TensorFlow, Keras
- Mots-clés : **modèles seq2seq (non) autorégressifs et (non) monotones, démasquage de logs**

■ Stage – Développement Logiciel

Mean-In-Full

Mai 2017 – Juillet 2017

Meylan, France

- Intitulé : **Intégration de la plateforme Opencast avec le système RoCamRoll**
- Environnement : Linux, LaTeX, Erlang, HTTP

■ Stage – Programmation Assembleur

Laboratoire TIMA

Grenoble, France

Mai 2016 – Juin 2016

- Intitulé : **Vérification fonctionnelle d'un microprocesseur ARM7**
- Environnement : Linux, LaTeX, VHDL, C, ARM

■ Stage – Programmation Fonctionnelle

Laboratoire VERIMAG

Mai 2015

Grenoble, France

- Intitulé : **Simulation de traces GPS**
- Environnement : Linux, LaTeX, Lutin, Lustre

Langues

Anglais

Bilingue

Français

Bilingue

Espagnol

Avancé

Allemand

Intermédiaire

Swahili

Intermédiaire

Kinyarwanda

Élémentaire

Langages de programmation

Python

(PyTorch, TensorFlow, Scikit-Learn)

Matlab

R

C

Java

JavaScript

SQL

Compétences transversales

Leadership

Esprit d'équipe

Autonomie

Résilience

Rigueur

Rédaction

Art oratoire

Centres d'intérêt

Chorale

Guitare

K-Dramas

Danses folkloriques

Associations

■ Groupe Sainte Marie

Interprète anglais-français

Mars 2016 - Mars 2019

Paris, France

■ Chrétiens en Grandes Écoles (CGE)

Responsable du réseau CGE pour la ville de Nantes

Avril 2018 - Janvier 2019

Nantes, France

■ Entraide Scolaire Amicale (ESA)

Soutien scolaire pour des lycéens en milieu défavorisés

Février 2018 - Janvier 2019

Nantes, France

■ Chorale Jeunes de Tous Pays (JTP)

Présidente d'une chorale d'étudiants internationaux

Septembre 2016 - Juillet 2017

Grenoble, France

Références

Benoît Sagot

Responsable de l'équipe ALMAnaCH, Inria Paris

benoit.sagot@inria.fr

Rachel Bawden

Chargée de recherche, équipe ALMAnaCH, Inria Paris

rachel.bawden@inria.fr

Publications

Lydia Nishimwe, Benoît Sagot, Rachel Bawden (2024). *Making Sentence Embeddings Robust to User-Generated Content*. Acceptée à LREC-COLING. [↗](#)

Lydia Nishimwe (2023). *Normalisation lexicale de contenus générés par les utilisateurs sur les réseaux sociaux*. TALN-RÉCITAL. [↗](#) 🏆 Best Paper Award 🏆

Omer Goldman, Francesco Tinner, Hila Gonen, Benjamin Muller, Victoria Basmov, Shadrack Kirimi, Lydia Nishimwe, Benoît Sagot, Djamel Seddah, Reut Tsarfaty, Duygu Ataman (2022). *The MRL 2022 Shared Task on Multilingual Clause-level Morphology*. MRL. [↗](#)

Jesujoba O Alabi, Lydia Nishimwe, Benjamin Muller, Camille Rey, Benoît Sagot, Rachel Bawden (2022). *Inria-ALMAnaCH at the WMT 2022 shared task: Does Transcription Help Cross-Script Machine Translation?* WMT. [↗](#)