

Ngoc Tan Le

Curriculum vitae

*Travailler dur d'aujourd'hui pour un meilleur avenir !
Working hard today for a better tomorrow !*

Éducation

- 2014–
Maintenant **Doctorat**, *Université du Québec À Montréal*, Montréal, Canada.
3ème année du Doctorat en Informatique Cognitive
- 2009 **Master**, *Institut Scientifique et Technique d'Ingénieur de Lyon*, Lyon, France.
Maîtrise Informatique de gestion - option **MIAGE**
(Méthode Informatique Appliquée à la Gestion d'Entreprise)
- 2008 **Licence Informatique (L3)**, *Institut Scientifique et Technique d'Ingénieur de Lyon*, Lyon, France.
Licence Informatique (L3), option **MIAGE**
(Méthode Informatique Appliquée à la Gestion d'Entreprise)
- 2007 **DUT (Diplôme Université de Technologie)**, *IUT A de l'Université de Lyon 1*, Lyon, France.
Promotion **Supérieure Professionnalisant** – option **Génie Informatique**
- 2002 **Baccalauréat**, *Lycée de Luong Van Chanh*, Phu Yen, Vietnam.
Baccalauréat général

Expériences professionnelles

- 09/2016–
12/2016 **Responsable de cours**, *Université du Québec À Montréal*, Montréal, Canada.
Responsable de l'atelier du cours INF-3180, intitulé « Fichiers et Bases de données ».
- 12/2015–
05/2016 **Stagiaire scientifique**, *Programme Mitacs accéléré - Société de Metix Inc*, 1864 rue Suzor-Côté, Longueuil (Québec), J4N 1P6, Canada.
Superviseur : Professeure Fatiha Sadat
Extraction d'information dans les documents en finance.
Reconnaissance des entités nommées incluant les termes en finance.
- 06/2015–
08/2015 **Stagiaire scientifique**, *Laboratoire d'Informatique de Grenoble - LIG*, 41 rue des mathématiques, 38041, Grenoble cedex 9, France.
Superviseur : Professeur Laurent Besacier
Travail de recherche sur la traduction automatique statistique pour une paire de langues peu dotée, dont français-vietnamien.

160 rue des aubépines, App 202 – J7V 0Y3 Vaudreuil-Dorion – Canada

☎ +1 (514) 621 5932 • ✉ letan.dhcn@gmail.com

🌐 <https://lengoctan.wordpress.com/>

03/2011– **Enseignant universitaire**, *Université d'industrie de Ho Chi Minh ville*, Ho Chi Minh ville, Vietnam.
08/2014 Enseignement des cours en informatique au niveau du Baccalauréat.

08/2006– **Responsable informatique**, *Société ASTUS (28 salariés)*, 289 rue Garibaldi, 69007
09/2010 Lyon, France.
Maintenir le réseautage et les logiciels internes. Créer le site Web de la société. Concevoir un portail Web pour les clients.

04/2005– **Stagiaire**, *Institut d'Enseignement des Langues*, 20, boulevard Eugène Deruelle,
06/2005 69432 Lyon cedex 03, France.
Stagiaire en informatique: Création d'un site Web dynamique utilisant le langage de programmation de Web PHP-MySQL.

Collège– **Concours et Récompenses en Linguistique**, Vietnam.
Lycée

- Troisième prix national des meilleurs élèves de la langue française (Prix National - classe Terminale)
- Médaille bronze de la langue Française du concours d'Olympic 30-04 au lycée Le Hong Phong, Ho Chi Minh ville, Vietnam (classe Seconde)
- Médaille bronze de la langue Française du concours d'Olympic 30-04 au lycée Le Hong Phong, Ho Chi Minh ville, Vietnam (classe Première)
- Deuxième place du concours « Destination France » de la région du Centre du Vietnam (classe 3^{ème})
- Première place dans les concours des meilleurs élèves de la langue française de la ville Phu Yen, Vietnam (de la classe 4^{ème} à la classe Terminale)

Compétences linguistiques

vietnamien **Langue maternelle**

français **Niveau avancé**

anglais **Niveau intermédiaire**

Bilingue

Maîtrise anglais technique

Compétences informatiques

Programmation ADA, Visual Basic, C/C++/C#, JAVA et Python

Web HTML, JavaScript, XML et PHP/MySql

Base de données SERVER SQL et ORACLE (SQL*PLUS et PL/SQL)

Analyse et Conception MERISE et UML

Bureautique Microsoft ACCESS, WORD, EXCEL et POWERPOINT

160 rue des aubépines, App 202 – J7V 0Y3 Vaudreuil-Dorion – Canada

☎ +1 (514) 621 5932 • ✉ letan.dhcn@gmail.com

🌐 <https://lengoctan.wordpress.com/>

Domaines de recherche

- Traduction automatique statistique (**TAS**)
- Reconnaissance des entités nommées bilingues
- Translittération des entités nommées bilingues
- Relations entre les langues **français, anglais** et **vietnamien**

Projet de Thèse

Traduction automatique pour une paire de langues peu dotée

Résumé Parmi les langues peu dotées, le vietnamien, quoique pratiqué dans le monde par environ 80 millions de locuteurs, cette langue fait partie des langues encore peu représentées au sein de la communauté scientifique internationale, tant du point de vue des sciences humaines qu'informatique, ou le domaine du **TALN** est situé. Au Vietnam, les travaux en **TALN** sont encore rares, et sont souvent entrepris sans la participation des linguistes, le terrain de la linguistique informatique est peu étudié (Do, 2011). Les récents travaux portant sur la gestion de ressources multilingues et l'alignement automatique de textes parallèles sont une opportunité pour la valorisation de la langue vietnamienne dans cette société où l'information joue un rôle essentiel et a un impact considérable sur les plans économique, social et développement humain. C'est dans ce contexte multilingue que nous avons élaboré notre projet de thèse, qui consiste à construire des outils et des ressources linguistiques indispensables pour la traduction automatique depuis une langue bien dotée de ressources linguistiques, le français vers une langue peu dotée, le vietnamien.

Publications

- 2017** * Fatma Mallek, **Ngoc Tan Le**, Fatiha Sadat. **Automatic Machine Translation for Arabic Tweets**. *Intelligent Natural Language Processing: Trends and Applications. Scientific Research Group in Egypt (SRGE) editor. Springer, 2017, book chapter.*
- * **Ngoc Tan Le**, Fatiha Sadat. **Information Extraction in Real-World Business Documents**. *CICLing : International Conference on Computational Linguistics and Intelligent Text Processing, April 17-23, 2017 Budapest, Hungary.*
- 2016** * **Ngoc Tan Le**, Fatiha Sadat. **UQAM-NL: Named entity recognition in Twitter messages**. *COLING 2016, the 26th International Conference on Computational Linguistics, Association for Natural Language Processing (ANLP), the 2nd Workshop on Noisy User-generated Text Osaka, Japan, 11-16 December 2016, p.197-202.*
- * **Ngoc Tan Le**, Fatiha Sadat. **Construction de ressources linguistiques pour un corpus annoté bilingue franco-vietnamien et traduction automatique des entités nommées pour Français-Vietnamien**. *84ème Congrès de l'ACFAS – Association francophone pour le savoir, Université du Québec à Montréal, Québec, Canada, 9-13 Mai 2016.*

- 2015 * **Ngoc Tan Le**, Fatiha Sadat. **Building a Bilingual Vietnamese-French Named Entity Annotated Corpus through Cross-Linguistic Projection**. *22ème Traitement Automatique des Langues Naturelles, Caen, France, 2015*.
- * **Ngoc Tan Le**, Fatiha Sadat. **Error Analysis of Named Entity Translation output for Poor-Resourced Bilingual Vietnamese-French Pair**. *7th Language Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, November 27-29, 2015, Poznań, Poland, p.304-308*.
- 2013 * **Ngoc Tan Le**, Ngoc Tien Le, Dien Dinh. **An Approach of Chunk Alignment for French-Vietnamese Bilingual Corpora**. *IJCSI International Journal of Computer Science Issues, ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784, Vol. 10, Issue 2, No 3, March 2013, p.111-117*.
- * **Ngoc Tan Le**, Ngoc Tien Le, Dien Dinh. **French-Vietnamese statistical machine translation combining with chunk alignments**. *Journal of Computer Science and Cybernetics ISSN: 1813-9663, Vol 29, No 4, 2013, p.358-368*.

Intérêts

- Sports Badminton, Tennis, Soccer
Lecture Scientifique, Documentaire, Policier
Cinéma Comédie, Action, Arts martiaux