

CURRICULUM VITAE

DO Thi-Ngoc-Diep

2020

PERSONAL INFORMATION

Full name **ĐỖ THỊ NGỌC DIỆP**
Sex Female
Date of birth 1982
Place of birth Hanoi - Vietnam
Nationality Vietnamese
Home address Le Thanh Nghi Street, Hanoi, Vietnam
Personal email diepdn@gmail.com
Phone number 0912606900



CAREER HISTORY

2012 - Present Lecturer at Speech Communication Department, International Research Institute of Multimedia Information, Communication and Applications (MICA), Hanoi University of Science and Technology (HUST) - carrying out teaching and research duties.

Office: Building B1- Room 1004, No 1, Dai Co Viet Street, Hanoi, Vietnam

Email: ngoc-diep.do@mica.edu.vn

diep.dothingoc@hust.edu.vn

Homepage: <http://www.mica.edu.vn/perso/Do-Thi-Ngoc-Diep>

2005-2011 Junior researcher at Speech Communication Department, International Research Institute MICA – HUST.

EDUCATION

2011 Grenoble University, Grenoble – France, PhD.
- Specialty: Computer Science
- Thesis: *Extraction of parallel corpus for machine translation from/to an under-resourced language*
- Advisor: Prof. Laurent BESACIER
- Thesis in French

2007 Hanoi University of Science and Technology, Hanoi – Vietnam, Master of Science
- Specialty: Information Processing and Communication, School of Information and Communication Technology
- Thesis: *Modeling and use of vocalic gestures in Automatic Speech*

Recognition

- Advisor: Prof. Eric CASTELLI
- Thesis in English

- 2005 Hanoi University of Science and Technology, Hanoi – Vietnam, Engineer
- Specialty: Software Engineering, School of Information and Communication Technology
 - Thesis: “*E-learning and an application of SCORM 2004 to the Learning Content Management System*”
 - Advisor: Prof. NGUYEN Ngoc Binh
 - Graduation grade: Excellent
- 2000 Baccalaureate from Viet Duc Senior High School, Hanoi – Vietnam
- Graduation grade: Excellent

CERTIFICATE AND AWARD

Mar 2015	Award of National Institute of Information and Communications Technology (NICT) Japan for a 2-month internship research fellowship at Multilingual Translation Laboratory - NICT, Kyoto, Japan. Research project on Vietnamese-Japanese machine translation, Supervisor Prof. Masao UTIYAMA (working in English)
May 2015	Certificate of Hanoi University of Science and Technology for completing the pedagogical training program
2007-2010	PhD scholarship from French Embassy in Vietnam for the academic years 2007-2008, 2008-2009, 2009-2010, studying in France
Jun 2006	Grant from EGIDE organization - France for a 3-month summer training internship at Laboratory CLIPS-LIG, Grenoble, France. Research project on automatic video content indexing, Supervisor Dr. Georges QUENOT (working in English)
2005-2007	Grant from Hanoi University of Science and Technology for Early admission master student in the academic years 2005-2006, 2006-2007
Aug 2005	Early admission certificate from Hanoi University of Science and Technology designed for very-good student to become master student
Apr 2005	Certificate of Merit from the Vietnam Post and Telecoms Ministry’s for Distinguished Information Technology Student of the academic year 2003-2004
Apr 2005	Certificate of Vietnamese Ministry of Science and Technology for achieving the Fundamental Information Technology Engineer Skill Standards (equivalent to the JITEC’s Fundamental Information Technology Engineer Skill Standards of Japan)
2000-2005	Grant from Hanoi University of Science and Technology for very-good student in 5 academic years from 2000 to 2005
Aug 2000	Early admission certificate from National Economics University reserved for excellent high school student

TEACHING EXPERIENCE

List of past and current courses	Level	Place	Language
Data structure and Algorithms	College	Industrial Vocational College Hanoi	Vietnamese
Data structure and Algorithms	Master	HUST	Vietnamese
Introduction to Database	Engineer	HUST	Vietnamese
Programing technique	Engineer	HUST	Vietnamese
Object-oriented system analysis and design	Engineer	HUST	Vietnamese
Pack Office	Master	French University Poles Hanoi	English
Machine translation	Engineer	Institute of Technology of Cambodia	English
Machine translation	Master	HUST	English

Other interested courses, but not limited: C/C++ programming, Operating systems, Database management system, Object-oriented programming technique, UNIX tools and scripting, Software engineering, Machine learning, Natural language processing, Speech processing, etc.

5 Best finished student projects

Building and deploying a Web page translation system for Vietnamese-English and Vietnamese-French.

Final year student project, 5-month internship at MICA, academic year 2011-2012

Student name: Antoine DAVID

Institute: Grenoble INP Esisar, France

Teaching Language: French, English.

Improving language model for Khmer automatic speech recognition.

Final year student project, 3-month internship at MICA, academic year 2012-2013

Student name: KAING Hour

Institute: Institute of Technology of Cambodia

Main tutor: Dr. SAM Sethserey - ITC

Teaching Language: English.

Research and integrating the speech synthesis module in a speech translation system on mobile phone, application for French-Vietnamese speech translation.

Fourth year student project, 3-month internship at MICA, academic year 2013-2014

Student name: Rodolphe FREBY

Institute: Polytech Grenoble, France

Teaching Language: French, English.

Building an offline speech recognition system on mobile phone.

Fourth year student project, academic year 2015-2016

Student name: NGUYỄN Văn Thịnh

Institute: Hanoi University of Science and Technology.

Research and building a speech translation system on mobile phone.

Final year student project, academic year 2016-2017

Student name: NGUYỄN Văn Thịnh

Institute: Hanoi University of Science and Technology.

Teaching assistant experiences

- Experience in developing, editing and implementing the International Master Program in Ambient Computing, Multimedia and Interaction (teaching in English), Hanoi University of Science and Technology as a committee/secretary member, in charge of designing and organizing courses, editing program and collecting relevant materials.
- Experience in research and designing of new courses and relevant materials.
- Experience in organizing projects for student/ group of students in training course to obtain knowledge from the real applications.

RESEARCH EXPERIENCE

Research Fields: Natural Language Processing

Interested Topics:

- Machine translation for under-resourced languages
- Machine translation for unwritten languages
- Collecting, mining and extracting multilingual data for machine translation
- Multilingual speech processing

Complete and current research projects	Year	Sponsor	Role
PI-languages: Languages and Speech processing technology for under-resourced languages	2009- 2012	ANR - France (LIG France, LIA France, MICA)	Researcher
Research and develop a parallel text corpus for Vietnamese and other under-resourced languages	2012	HUST	Promoter, researcher
Vietnamese Voice Assistant for Vision-impaired User – VIVAVU 1 st phase	2013	Ministry of Information and Communication	Promoter, researcher
Vietnamese Voice Assistant for Vision-impaired User – VIVAVU 2 nd phase	2014	Ministry of Information and Communication	Researcher
Advances in rapid prototyping for computerization of languages	2013-2014	PEPS HuMaIn CNRS – France (Lacito France, MICA)	Promoter in Vietnam, researcher
Intelligent human-machine interaction system using Vietnamese voice controlled in the noisy environment – National project	2012-2015	Ministry of Science and Technology	Researcher
Visually Impaired People Assistance using multimodal technologies	2014 -2016	VLIR Belgium (Ghent Belgium, INPG France, MICA)	Researcher
Research and develop a speech translation system on smartphone, application for French-Vietnamese language pair	2015	HUST	Promoter, researcher
Shoot My Mind	2015-2017	ASIAN-ITC (JFLI Japan,	Researcher

		NII Japan, LIG France, ITC Cambodia, MICA Vietnam)	
Research and develop a parallel fragment-based extraction method from comparable corpora	2016	HUST	Promoter, researcher
Research and development of an automatic translation system from Vietnamese text to Muong speech, applied to other unwritten minority languages in Vietnam – National project	2017-2020	Ministry of Science and Technology	Researcher

Reviewer/Scientific committee for several national/international journals, conferences and workshops (in English): IALP 2013, 2014, ICCE 2014, KhmerLP 2015, TALLIP 2015, ICTA 2016, JST 2016, REV-ECIT 2018, MAPR 2019

Publications

Publication links:

<https://scholar.google.com.vn/citations?user=VokAh3oAAAAJ&hl=vi>

<https://www.scopus.com/authid/detail.uri?authorId=54905450900>

<https://orcid.org/0000-0001-9197-7850>

1. Pham T.T.T., Nguyen DD., Ta B.H.P., Nguyen TB., **Do TND.**, Le TL. (2020) Person Search by Queried Description in Vietnamese Natural Language. In: Sitek P., Pietranik M., Krótkiewicz M., Srinilta C. (eds) Intelligent Information and Database Systems. ACIIDS 2020. Communications in Computer and Information Science, vol 1178. Springer, Singapore
2. Trần Thị Thu Thúy, **Đỗ Thị Ngọc Diệp**, Mạc Đăng Khoa, Phạm Văn Đồng, Cross-lingual phoneme recognition for familiar languages: applying to Vietnamese and Muong languages, in *Proceedings of 11th National Science Conference on Fundamental and Applied IT Research*, Hanoi – Vietnam, 2018 (paper in English).
3. Van-Thinh NGUYEN, **Thi-Ngoc-Diep DO**, Dang-Khoa MAC, Eric CASTELLI, Optimizing data transmission on mobile platform for speech translation system, *First Regional Conference on Optical character recognition and Natural language processing technologies for ASEAN languages (ONA 2017)*, Phnom Penh - Cambodia, 2017 (paper in English).
4. Viet-Son Nguyen, Trung-Kien Dao, Viet-Tung Nguyen, Do-Dat Tran, Dang-Khoa Mac, **DO Thi Ngoc Diep**, NGUYEN Tuan Ninh, A Domestic Solution with Vietnamese Speech Interaction, *Proceedings of the 9th regional conference on electrical and electronics engineering, pages 258-262*, 2016 (paper in English).
5. DAO Trung Kien, TRAN Thi Thanh Hai, LE Thi Lan, VU Hai, NGUYEN Viet Tung, MAC Dang Khoa, **DO Thi Ngoc Diep**, PHAM Thanh Thuy, Indoor navigation assistance system for visually impaired people using multimodal technologies, *The 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, IEEE issn: 978-1-5090-4757- 4, November 2016 (paper in English).

6. **Thi-Ngoc-Diep DO**, Masao UTIYAMA, Eiichiro SUMITA, Machine Translation from Japanese and French to Vietnamese, the Difference among Language Families, *In Proceedings of 2015 International Conference on Asian Language Processing*, pages 17-20, Suzhou – China, 2015 (paper in English).
7. **Do Thi-Ngoc-Diep**, Michaud Alexis, Castelli Eric. Towards the automatic processing of Yongning Na (Sino-Tibetan): developing a ‘light’ acoustic model of the target language and testing ‘heavyweight’ models from five national languages, *In proceedings of the 4th International Workshop on Spoken Language Technologies for Under-Resourced Languages (SLTU'14)*, ISBN 978-5-8088-0908-6, Publ: GUAP, pages 153-160, 2014 (paper in English).
8. **Thi Ngoc Diep DO**, Duy Binh NGUYEN, Dang Khoa MAC, Do Dat TRAN. Machine Translation Approach for Vietnamese Diacritic Restoration. *2013 International Conference on Asian Language Processing (IALP 2013)*, ISBN: 9781479910960, Publ: Institute of Electrical and Electronics Engineers (IEEE), pages 103-106, 2013 (paper in English).
9. Emmanuel FERREIRA, Pascal NOCERA, Maria GOUDI and **Thi-Ngoc-Diep DO**. YAST: A scalable ASR toolkit especially designed for under-resourced languages, *2012 International Conference on Asian Language Processing (IALP 2012)*, ISBN: 9781467361132, Publ: Institute of Electrical and Electronics Engineers (IEEE), pages 141-144, 2012 (paper in English).
10. **Thi-Ngoc-Diep Do**, Eric Castelli, Laurent Besacier, Mining Parallel Data from Comparable Corpora via Triangulation, *2011 International Conference on Asian Language Processing (IALP 2011)*, ISBN:9781457717338, Publ: Institute of Electrical and Electronics Engineers (IEEE), pages 185-188, 2011 (paper in English).
11. Laurent Besacier, Haithem Afli, **Do Thi Ngoc Diep**, Hervé Blanchon, Marion Potet, LIG Statistical Machine Translation Systems for IWSLT 2010, *In Proceedings of the International Workshop on Spoken Language Translation*, pp.99-104, Paris – France, 2010 (paper in English).
12. **Thi-Ngoc-Diep Do**, Laurent Besacier, Eric Castelli, Improved Vietnamese-French Parallel Corpus Mining Using English Language, *In Proceedings of the International Workshop on Spoken Language Translation*, pp.235-242, Paris – France, 2010 (paper in English).
13. **Thi-Ngoc-Diep Do**, Laurent Besacier, Eric Castelli, Apprentissage non supervisé pour la traduction automatique : application à un couple de langues peu doté, *Conférence sur le Traitement Automatique des Langues Naturelles*, Montréal – Canada, 2010 (paper in French).
14. **Thi-Ngoc-Diep Do**, Laurent Besacier, Eric Castelli, A Fully Unsupervised Approach for Mining Parallel Data from Comparable Corpora, *In Proceedings of the 14th Annual Conference of the European Association for Machine Translation*, Saint-Raphaël – France, 2010 (paper in English).
15. **Thi-Ngoc-Diep Do**, Laurent Besacier, Eric Castelli, Unsupervised SMT for a low-resourced language pair, *In Proceedings of the 2nd International Workshop on Spoken Languages Technologies for Under-resourced languages*, pages 130-135, Penang – Malaysia, 2010 (paper in English).
16. **Thi-Ngoc-Diep Do**, Viet-Bac Le, Brigitte Bigi, Laurent Besacier, Eric Castelli. Exploitation d’un corpus bilingue comparable pour la création d’un système de traduction probabiliste Vietnamien – Français, *Conférence sur le Traitement Automatique des Langues Naturelles*, Senlis – France, 2009 (paper in French).

17. **Thi-Ngoc-Diep Do**, Viet-Bac Le, Brigitte Bigi, Laurent Besacier, Eric Castelli, Mining a comparable text corpus for a Vietnamese - French statistical machine translation system, *in Proceedings of the 4th EACL Workshop on Statistical Machine Translation, pages 165–172*, Athens - Greece, 2009 (paper in English).
18. Viet-Bac Le, Laurent Besacier, Sopheap Seng, Brigitte Bigi, **Thi-Ngoc-Diep Do**, Recent Advances in Automatic Speech Recognition for Vietnamese, *In Proceedings of the 1er International Workshop on Spoken Languages Technologies for Under-resourced languages, p47-52*, Ha Noi - Vietnam, 2008 (paper in English).
19. **Đỗ Thị Ngọc Diệp**, Mạc Đăng Khoa, Nguyễn Ngọc Bình, Application of SCORM 2004 in building an e-learning system, *in Proceedings of 2nd National Science Conference on Fundamental and Applied IT Research*, page 408-420, Ho Chi Minh city - Vietnam, 2006 (paper in Vietnamese).

SERVICE EXPERIENCE

2018	Organizer and manager of the website of the International Workshop on Spoken Language Technologies for Under-resourced Languages (SLTU) 2018 Organizer and manager of the website of the ACMI Master program
2012-2016	Organizer of regular meeting for Speech Communication Department, MICA Institute.
2012	Organizing committee of the International Conference on Asian Language Processing (IALP) 2012, Hanoi, Vietnam.

OTHER INFORMATION

Languages:

- Vietnamese: mother tongue
- Old certifications:
 - English: IELTS 5.5 (since 2000)
 - French: DELF A1 (since 2007), B1- en cours d'acquisition (since 2008)
- Currently:
 - English: proficiency
 - French: very good command

Personal strong points

- Strong knowledge and professional expertise
- Experience in living and working abroad (France, Japan)
- Working in international environment
- Teaching and researching in English
- Advisor of foreign internship students
- Participator of international research projects
- Participation in lot of international conferences, seminars, meetings abroad
- Strong relation with domestic/abroad lecturers/academic groups/researchers on speech and language processing
- Be a kindly and honest person; smart, agile and careful in work

Hobbies:

- Planting tree at home
- Making hand craft toys for child
- Entertainment: watching reality TV shows (The Amazing Race, Masterchef, Pékin Express, ...)