KAK SOKY

National IT/MIS Consultant

📞 +855-89-808-991   ✉ soky.kak@gmail.com

🌐 https://ksoky.github.io/   📍 Phnom Penh, Cambodia

EDUCATION

Doctor of Informatics
Kyoto University
Major: Speech Recognition
2019 - 2023

Master of IT Engineering
Royal University of Phnom Penh
Major: Software Engineering
2013 - 2016

Teacher of Higher Education
National Institute of Education
Major: Pedagogy in IT
2014 - 2015

Bachelor of Computer Science
Svay Rieng University
Major: Software Engineering
2006 - 2010

SKILLS

- System analysis and Design
- IT Training and Support
- IT Project Management
- Technology Solution Development
- Documentation Management
- AI Research and Development
- Data Science and Big data analysis

ABOUT ME

I graduated with a PhD in Informatics from Kyoto University, Japan. I am an adjunct lecturer at the Cambodia Academy of Digital Technology (CADT). I have been a software developer for more than 4 years and have more than 10 years of experience in academics including teaching, conducting research, coordinating research and development projects, master's programs in data science, and advising student final projects.

Adjunct Lecturer   2023 - Present
Cambodia Academy of Digital Technology

- Providing lectures on Python for Data Science, Object-Oriented Programming with Java, and Database Management Systems.
- Advisory to internship students expecting to graduate in computer science and data science specialists. The projects include a wide range of applications from Web Applications to Data Science Projects.

Digital Consultant   Mar 2023 - Dec 2023
Cambodia Academy of Digital Technology

Coordinated bachelor and master programs in Data Science by preparing the curriculum and arranging the internship program. Supervised students' internships in software engineering and data science specialist and advised the research projects related to automatic speech recognition (ASR) and natural language processing (NLP). Gave lectures on object-oriented programming (OOP) in Java and Database administration.

Research Coordinator   2017 - 2019
National Institute of Post, Telecoms and ICT

Collected text, speech, and image data for Khmer ASR and optical character recognition (OCR). Worked on post-processing of Khmer OCR, Improved Khmer ASR system and integrated into mobile application.

IT Officer   2015 - 2019
Ministry of Education, Youth and Sports

Joint the national examination committee. Trained the computer literacy and examination system to officer under ministry of education, youth and sports.

TECHNOLOGIES

Program

- Java
- Python
- C#
- C/C++
- SQL
- Shell Script

Database Management System

- Microsoft SQL Server
- MySQL
- Oracle
- PostgreSQL

Tool and Frameworks

- Laravel
- Pytorch
- ArcGIS Pro
- QGIS

LANGUAGES

- KHMER
- ENGLISH

PUBLICATIONS

Journal

K. Soky, S. Li, C. Chu, T. Kawahara, Finetuning Pretrained Model with Embedding of Domain and Language Information for ASR of Very Low-Resource Settings. International Journal of Asian Language Processing, Vol 33, no. 04, 2023.

K. Soky, M. Mimura, T. Kawahara, C. Chu, S. Li, C. Ding, S. Sam, TriECCC: Trilingual Corpus of the Extraordinary Chambers in the Courts of Cambodia for Speech Recognition and Translation Studies. International Journal of Asian Language Processing Vol. 31, no. 03&04 2250007, 2022.

Internship Researcher	2016 - 2017

National Institute of Information Communications Technology

Collected data and built Khmer G2P, Khmer ASR, and Khmer Text-to-Speech. Wrote a research paper and technical reports.

Adjunct Lecturer	2015 - 2016

Kirirom Institute of Technology

Lectured fundamental of computer and advised students in digital marketing project.

Adjunct Lecturer	2012 - 2016

Svay Rieng University

Lectured android programming, network security, and mobile communication courses. Supervised students to do a final thesis.

Java Software Developer	2013 - 2014

Moblie C&C Cambodia Co., LTD

Developed range of web applications using Java Spring Framework and developed mobile app using hybrid technology.

Software Developer	2011 - 2013

Biz Solution Co., LTD

Developed some desktop applications including a student management system, an Inventory management system using C# programming, and SQL Server as a database administration.

REFERENCES

H.E Kith Chankrisna

Under Secretary of State
Ministry of Land Management, Urban Planning and Construction
Email : kithchankrisna@gmail.com

Prof. Tastuya Kawahara

Professor at School of Informatics
Kyoto University
Email : kawahara@i.kyoto-u.ac.jp

Conference

K. Soky, S. Li, C. Chu, T. Kawahara, Domain and Language Adaptation Using Heterogeneous Datasets for Wav2vec2.0-Based Speech Recognition of Low-Resource Language. In proc. ICASSP 2023, Rhodes Island, Greece, 2023.

K. Soky, S. Li, C. Chu, T. Kawahara, Domain and Language Adaptation of Large-scale Pretrained Model for Speech Recognition of Low-resource Language. In proc. IEICE-SP, 2022

K. Soky, Z. Gong, S. Li, NICT-Tib1: A Public Speech Corpus of Lhasa Dialect for Benchmarking Tibetan Language Speech Recognition Systems. In proc. O-COCOSDA, 2022.

K. Soky, M. Mimura, S. Li, C. Chu, T. Kawahara, Leveraging Simultaneous Translation for Enhancing Transcription of Low-resource Language via Cross Attention Mechanism. In proc. Interspeech, 2022.

K. Soky, M. Mimura, T. Kawahara, S. Li, C. Ding, C. Chu, S. Sam, Khmer Speech Translation Corpus of the Extraordinary Chambers in the Courts of Cambodia (ECCC). In Proc. O-COCOSDA2021, 2021

K. Soky, S. Li, M. Mimura, C. Chu, T. Kawahara, On the Use of Speaker Information for Automatic Speech Recognition in Speaker-imbalanced Corpora. In Proc. APSIPA2021, 2021

K. Soky, S. Li, M. Mimura, C. Chu, T. Kawahara, Comparison of End-to-End Models for Joint Speaker and Speech Recognition. In Proc. IEICE-SP, 2021

K. Soky, S. Li, T. Kawahara, S. Seng, Multi-lingual transformer training for Khmer automatic speech recognition. In Proc. APSIPA ASC, 2019.

Saly, K., Soky, K., Jinfu, N., Yoshinori, S., Kato, H., Kawai, H. Hidden Markov Model based Speech Synthesis system for Khmer Language. In Proc. ONA, 2018.

K. Soky, C. Vichet, S. Sam, Khmer Automatic Speech Recognition based on DNN models. In Proc. ONA, 2017.

Hour, K., Soky, K., Vichet, C., Sethserey, S. OCR post-processing for Khmer language: Error detection using Conditional Random Field. In Proc. ONA, 2017.

Soky, K., Lu, X., Shen, P., Kato, H., Kawai, H., Vanna, C., & Chea, V. Building WFST based Grapheme to Phoneme Conversion for Khmer. 3rd Khmer Natural Language Processing, 2016.