

# IMAGE CAPTIONING

Senior Researcher,  
**Dr. Sanparith Marukatat**



## Chairperson

**Dr. Ye Kyaw Thu**

✉ [yekyaw.thu@nectec.or.th](mailto:yekyaw.thu@nectec.or.th)


## BIOGRAPHY

Sanparith Marukatat received a PhD degree in computer science from the Paris VI University (Université Pierre et Marie Curie), France, in 2004. His work concerned handwriting recognition and statistical models for sequence data. He is now with NECTEC, Thailand, where he works on machine learning methods for offline handwriting recognition and speech recognition.

## ABSTRACT

The resurgence of AI, powered by deep learning, has propelled the field of computer vision into a new era. The marriage of convolutional neural networks (CNNs) for image processing and recurrent neural networks (RNNs) and Transformers for sequence generation has yielded substantial progress in the realm of image understanding and interpretation. This has enabled the development of sophisticated models capable of not only recognizing intricate visual patterns but also generating linguistically informed captions that encapsulate the essence of an image. In this talk, we shall embark upon a comprehensive journey encompassing the theoretical underpinnings and practical implementations of deep learning models tailored to the task of image captioning. We will unravel the layers of innovation that have emerged, including attention mechanisms, Transformer architectures, and the fusion of multi-modalities.

 [meet.google.com/pmq-tukk-zzy](https://meet.google.com/pmq-tukk-zzy)

 9:00 – 12:00 (Thailand Time)

 September 2, 2023

