

Exploring A Lexicon with Adapting NLP Technology for Understanding Human Language
by Professor Kyoko Kanzaki from Otemon Gakuin University, Japan



ANDAMAN INTELLIGENT TOURISM AND SERVICE INFORMATICS CENTER (AI-TASI)
COLLEGE OF COMPUTING, PRINCE OF SONGKLA UNIVERSITY, PHUKET CAMPUS
PRESENT

EXPLORING A LEXICON WITH ADAPTING NLP TECHNOLOGY FOR UNDERSTANDING HUMAN LANGUAGE



by

PROFESSOR KYOKO KANZAKI
Otemon Gakuin University, Japan

What you will learn:

- Introduction of NLP and the Lexicons
- Building a lexicon by using NLP technology
- What a lexicon needs for language understanding and generating
- Case studies

WEDNESDAY
19 JAN 2022

14:00-16:00

THAILAND TIME UTC+7

Registration: <https://forms.gle/ZQNdPg4UcjXtZtEH7>



Zoom link: <https://psu-th.zoom.us/j/97216974243?pwd=SlUzTEVBNltaGJyeDRuZVVFcE1Jdz09>



Meeting ID: 972 1697 4243
Passcode: 350654



Date: 19 January 2022

Time: 14:00-16:00

Location: Zoom

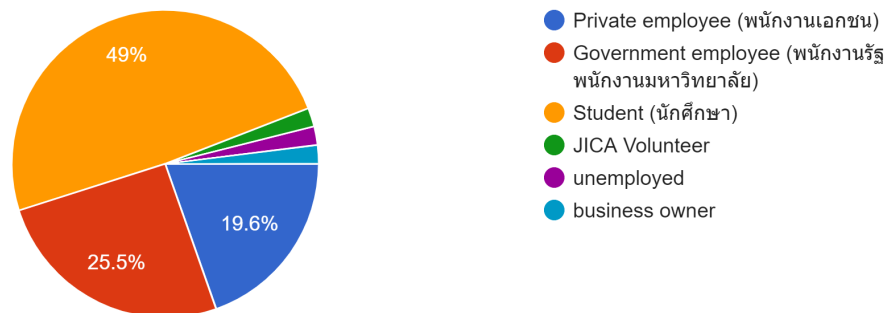
Host: Andaman Intelligent Tourism and Service Informatics Center (AI-TASi), College of Computing, PSU Phuket

Registration Link: <https://computing.psu.ac.th/th/13428/>

No.of Registrations: 51 participants

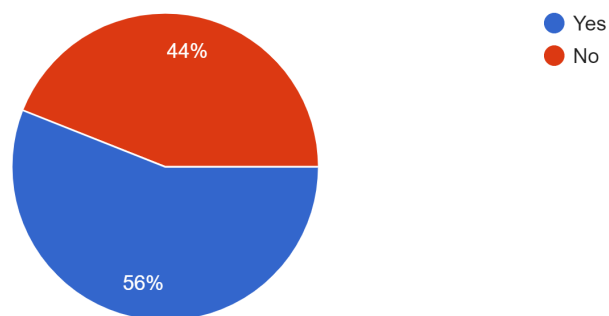
Occupation

51 responses



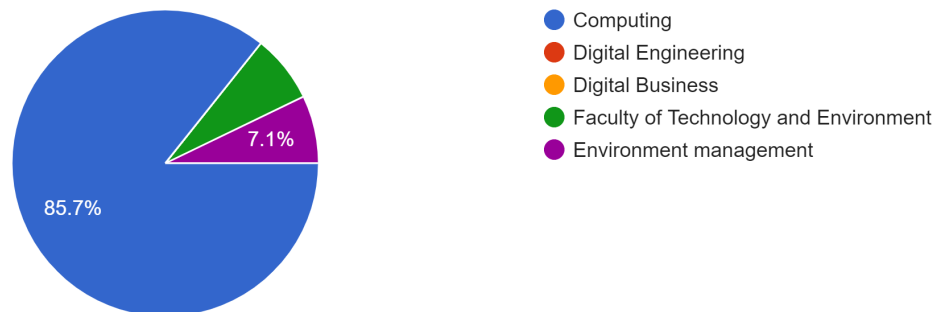
Are you PSU Student?

25 responses



Major

14 responses



What have you learned from this seminar?

- อย่างแรกเลย เปิดมาก็ได้รู้เกี่ยว Lexicon โดยจากที่ฟังและจากที่อ่านเอา Lexicon คงจะเป็นคลังคำศัพท์ที่มีคำศัพท์ไว้มากมายเอาไว้ให้ใช้งาน และต่อมาก็คงจะเป็นโครงสร้างของภาษา ที่ต้องทราบว่าทำยังไงให้คอมพิวเตอร์เข้าใจภาษาของมนุษย์ได้ โดยจะนำประโยคป้อนเข้าไป และจะมี NP และ VP เมื่อได้แกรมมาที่ถูกต้องจากการเทรนนั้นแล้ว ก็จะทำสิ่งที่ได้ ไปทำ Deep Structure โดยจะเป็นการให้ความหมายของคำแต่ละคำ และนำไปเชื่อมเป็นประโยคที่สมบูรณ์ยิ่งขึ้น อย่าง Jones gives a present to the girl มันก็จะเป็นประโยคที่สมบูรณ์ เพราะคอมพิวเตอร์สามารถแยกคำแต่ละคำได้ชัดเจน ว่าควรจะอยู่ตรงตำแหน่งใด หรือถ้าให้พูดก็คือ ทำให้คอมพิวเตอร์เข้าใจเกี่ยวกับแกรมมาได้นั่นเอง
- Lexicon เป็นพจนานุกรมที่ใช้ประกอบการวิเคราะห์ความหมาย ซึ่งประกอบด้วยการสะกดคำที่ถูกต้อง ความหมาย และหน้าที่ของคำนั้น ๆ ในประโยค กรณีที่คำมีความหมายมากกว่าหนึ่งความหมาย จะแสดงความหมายทั้งหมดออกมาให้
- ได้เรียนรู้เกี่ยวกับ Lexicon ซึ่งเป็นคลังคำศัพท์ที่รวบรวมความรู้ของคน โดยคำศัพท์มีการเชื่อมโยงและถูกจัดเก็บโดยความสัมพันธ์ต่าง ๆ ทำให้เราสามารถเข้าใจความหมายของคำแต่ละคำ ตามความเข้าใจ คือการที่เรามีคลังคำศัพท์ เพื่อต้องการเข้าว่า หรือจับจุดให้ได้ว่า ข้อความนั้น ๆ ต้องการสื่อถึงอะไร เช่น ในประโยคมีการพูดถึง Cheese, pasta, หรือ rice ซึ่งคำเหล่านี้ก็สามารถจับประเด็นได้ว่าพูดถึงเรื่องของอาหารนั่นเอง หรือ มีข้อความว่า Red flower - red ก็คือ สี / flower ก็มีสี ดังนั้น ดอกไม้ที่มีสีแดง ก็มีความเชื่อมโยงความสัมพันธ์กัน
- วันนี้ได้เรียนรู้หลายเรื่อง เริ่มต้นก็จะเป็นเรื่องของ Lexicon ที่จะทำหน้าที่รวบรวมคำต่างๆ เหมือนเป็นพจนานุกรม แล้วก็ต่อด้วยเรื่องราวของคนในวงการ NLP คือคุณ Chomsky ที่มีโปรเจกต์ที่จะทำให้คอมพิวเตอร์นั้นสร้างประโยคต่างๆ ได้ด้วยตัวระบบเองในระดับโครงสร้างของภาษา คือ NP และ VP แล้วก็ใส่พวก Verb เข้าไประบบก็สามารถสร้างออกมาได้ แกรมมาถูกต้อง แต่ความหมายยังกำกวมเลยต้องมีการนำเทคโนโลยี Deep Structure เข้ามาช่วยทำให้ทำงานได้ดีขึ้นมาก เรื่องต่อมาก็คือ Frame เรื่องนี้ค่อนข้างที่จะ งง เหมือนว่าเป็น Room ที่ผมคิดไว้น่าจะเป็นแฟ้มๆ เก็บข้อมูลต่างๆ ทางด้าน สถิติ เหมือนแยกๆ ไว้ แล้วก็เทคโนโลยี Ontology ที่จะมาแยกอีกทีหนึ่ง
- NLP technology
- Realize on NLP implementation and case study from Japan

- ได้ความรู้เกี่ยวกับ Lexicon หรือจะเรียกอีกอย่างว่า MentalLexicon ได้รับพื้นฐานเกี่ยวกับLexicon ความรู้เกี่ยวกับ Chomsky's grammar โดยมีกฎการสร้างคือการสร้างไวยากรณ์เท่านั้น ได้ความรู้เกี่ยวกับศัพท์บัญญัติในภาษาธรรมชาติ
- Lexicon คือจำนวนความรู้ทั้งหมดที่ผู้พูดจำได้ในสมองของเขา ซึ่งเชื่อมโยงและจัดเก็บโดยความสัมพันธ์ต่าง ๆ บางครั้งหมายถึงฐานข้อมูลพจนานุกรม และเป็น keys สำคัญสำหรับการพัฒนา NLP ตัวอย่างคำที่สัมพันธ์กัน เช่น Attribute = สี value = ส้ม, แดง, ฟ้า เป็นต้น
- การใช้ NLP ในการทำ lexicon(พจนานุกรม) มีการบอกคุณสมบัติของคำในประโยคและบอกความหมายได้ชัดเจนเพื่อใช้ในการแปลประโยคได้อย่างถูกต้องสมบูรณ์
- รู้ว่าพจนานุกรมคือ จำนวนความรู้ทั้งหมดที่เกี่ยวกับคำศัพท์พื้นเมือง เป็นความรู้ที่เชื่อมโยงโดยความสัมพันธ์ต่างๆ พื้นฐานของ lexicon ไวยากรณ์Chomsky'sเป็นหนึ่งกฎในการพัฒนาภาษาศาสตร์และNLP เพื่ออธิบายความรู้ทางภาษาที่สร้างจากสมองมนุษย์ พยายามอธิบายกฎไวยากรณ์สากลเพื่อสร้างภาษาที่ไม่อิงจากภาษาต่างๆ โดยการใช้คอมพิวเตอร์ การใช้ไวยากรณ์เพื่ออธิบายการสร้างประโยค รู้เกี่ยวกับแกรมมาต่างๆที่ใช้ในการสร้างประโยค เช่น โครงสร้างแบบลึกกับพื้นผิว ความแตกต่างระหว่างสองแบบในประโยค เชื่อมกันโดยไวยากรณ์การเปลี่ยนแปลง ใน Lexicalism แนวคิดเป็นแกนนำของภาษา คำ เป็นส่วนหนึ่งของคำพูดและคำจำกัดความในภาษา สามารถใช้ในทางจิตวิทยาเชิงโน้มน้าวต่างๆได้
- รู้ว่าพจนานุกรมคือ จำนวนความรู้ทั้งหมดที่เกี่ยวกับคำศัพท์พื้นเมือง เป็นความรู้ที่เชื่อมโยงโดยความสัมพันธ์ต่างๆ พื้นฐานของ lexicon ไวยากรณ์Chomsky'sเป็นหนึ่งกฎในการพัฒนาภาษาศาสตร์ การใช้ไวยากรณ์เพื่ออธิบายการสร้างประโยค รู้เกี่ยวกับแกรมมาต่างๆที่ใช้ในการสร้างประโยค เช่น โครงสร้างแบบลึกกับพื้นผิว ความแตกต่างระหว่างสองแบบในประโยค
- NLP techniques
- ในกิจกรรมหัวข้อ Exploring A Lexicon with Adapting NLP Technology for Understanding Human Language ได้เรียนรู้ ความรู้ ข้อมูลของ ความหมายพื้นฐานของ Lexicon รวมไปถึงโปรแกรมของทางญี่ปุ่นที่เกี่ยวข้องกับ NLP เช่น Japanese WordNet หรือ Japanese Compound Verb Lexicon ที่มีกระบวนการขั้นตอนต่างๆ เช่น การเตรียมข้อมูล การแยกคำ เป็นต้น
- การปรับความเข้าใจกับภาษามนุษย์ NLP และ การเขียนยังงให้เข้าใจตรงกัน
- ได้เรียนรู้เกี่ยวกับNLP และ Lexicons ว่าสามารถนำไปใช้ได้หลากหลายไม่ว่าจะเอาไปทำการสร้างคำศัพท์โดยใช้เทคโนโลยี NLPและเอาไปใช้เป็นพจนานุกรมในการศึกษาได้
- เริ่มที่ Lexicon คืออะไร ตามสไลด์ของคนบรรยายเหมือนจะเขียนว่าเป็นคลังคำศัพท์ของบุคคล มั้งครับ ว่าด้วยเรื่องของพอลนัมที่ชื่อว่า Chomsky เหมือนว่านายคนนี้จะพยายามทำให้คอมพิวเตอร์สร้างประโยคได้ด้วยตัวเอง เลยทำความเข้าใจโครงสร้างภาษา ก็คือในประโยคจะมี NP กับ VP หลังจากที่โครงสร้างของประโยคถูกต้อง ก็ใส่คำศัพท์ตามโครงสร้าง NP ก็ใส่ Noun VP ก็ใส่ Verb ผลลัพธ์ที่ออกมาก็ถูกแกรมมาอยู่ แต่ความหมายมันมีน้อยอยู่ เลยต้องมี Deep Structure เข้ามาอีก เหมือนจะเกี่ยวกับการผสมคำโดยคำนึงถึงความหมายและการใช้งานของตัวเชื่อม ซึ่งก็คือ Verb นี่แหละ ถึงตรงนี้ประโยคเลยดูมีความหมายขึ้นมาหน่อย ที่สำคัญถัดมาจะเป็นเรื่องของ Frame เข้าใจว่า Frame เป็นเหมือนกรอบของคำ เช่น Frame คือ Room คำว่า Room ก็จะมีแบบ Bathroom, Kitchen หรือ Office อธิบายเป็นไทยไม่ถูก แต่ผมว่าจารย์น่าจะเข้าใจ แล้วก็มี Ontology อีกนิดนึง ที่เหมือนจะเป็นการแยกประเภทต่อไปจากหัวข้อใหญ่อีก สรุปเลยก็คือ Lexicon ต้องรู้คำศัพท์ โครงสร้างประโยค ความหมายของคำ Frame การเก็บสถิติ ประมาณนี้ ถัดมาจะเน้นไปที่เรื่องของ Adjective เป็นการศึกษาว่าคำขยายคำนี้ ควรจะเชื่อมกับคำถาม ตามตัวอย่างก็คือ Red จะไปเชื่อมกับ Size ก็แปลก ๆ มันต้องเชื่อมกับ Color เป็น Red color flower หลังจากตรงนี้มีแต่สูตรกับกราฟมัน ๆ ไม่เข้าใจเลยซักนิดเพราะภาษาญี่ปุ่นเริ่มจะเยอะ แต่เหมือนจะอธิบายเกี่ยวกับคำกับประโยคประเภทต่าง ๆ อีก พวกความรวมความซ้อน

How will you apply the knowledge to your task?

- จากเท่าที่ดูมาแล้วก็เข้าใจคร่าวๆว่า NLP มันช่วยให้คอมพิวเตอร์เข้าใจภาษามนุษย์ของเรา โดยที่เราต้องสอนคอมพิวเตอร์ให้มันเข้าใจถึงคำศัพท์ ความหมาย และโครงสร้างของประโยค นอกตามตรง ผมยังไม่เข้าใจคำถามของคำว่า ช่วยงาน แต่ถ้าเอาตามความคิดของผมคือ ผมอาจจะนำ NLP ไปประยุกต์ใช้กับเทคโนโลยีหรือไม่ก็พวกซอฟต์แวร์ต่างๆได้ ถ้าหากเรียน NLP เยอะพอสมควร
- นำเทคนิค วิธีการคิด ไปใช้วิเคราะห์คำซึ่งมีรูปหรือเสียงเหมือนกันจะพิจารณาว่าเป็นคนละคำ , คำมีความหมายใกล้เคียงกัน, ขอบเขตคำกำหนดอย่างไร
- อาจจะนำไปใช้งาน Machine Translation เพื่อให้สามารถเข้าใจและแปลได้อย่างถูกต้องมากยิ่งขึ้น
- นำไปใช้งานได้ยาก แค่นี้เป็นตัวอย่างได้ในอนาคต เพราะตอนนี้ยังเรียนแค่ NLP พื้นฐาน เข้าใจหลักการทำงาน แต่ยังไม่ได้ลงมือทำที่
- develop our models
- May apply for the research
- นำภาษาธรรมชาติไปใช้ประโยชน์ในการทำโครงการในอนาคตได้อย่างเข้าใจมากยิ่งขึ้น
- ทำให้ระบบ computer เข้าใจภาษามนุษย์มากขึ้น
- การเรียนในวิชาสาขาที่กำลังศึกษา เป็นความรู้เพิ่มเติมในการจัดทำโปรเจกต์จบการศึกษา
- นักจิตวิทยา ,นักวิเคราะห์
- สามารถไปพัฒนางานที่เราจะทำในอนาคตได้หลายๆงานและสามารถไปแนะนำคนอื่นได้อีกด้วยจะทำให้คนมีความรู้ด้วย
- SOM with semantics
- นำความรู้ที่ได้จากกิจกรรมนี้ไปปรับใช้ในการเรียน รวมไปถึงการประยุกต์ใช้ในอนาคดด้วย
- จะะลึกในการเข้าใจในภาษามนุษย์มากขึ้น ทำให้เราเข้าใจตรงกันเวลาจะพูดในสิ่งที่เราคิดจะทำมันด้วยกัน
- นำไปต่อยอดจากสิ่งที่ได้รับความรู้มา
- ตอบไม่ได้ครับ ไม่ได้เข้าใจมากขึ้นขนาดที่จะเอามาใช้ประโยชน์ได้ สิ่งที่ได้จากการดูคลิปก็คือเข้าใจการทำงานของ NLP ขึ้นมานิดหน่อยเท่านั้นเอง

List of participants:

https://docs.google.com/spreadsheets/d/1wvefmbnPRc_ou1PriZNxyj-D0-f3UmwQ8qq4tA4hltQ/edit#gid=1797608951

YouTube: <https://www.youtube.com/watch?v=m4kQ-5txhss>

Facebook Live: <https://fb.watch/aDzZGPVFat/>

Evaluation by PSU: <https://forms.gle/ab6cLzS6nmqFNayT9>

Evaluation Result by PSU:

https://docs.google.com/forms/d/1h5md9jHmo4piCKcX8Ps9HSIYXQcCm2R_Y2dhkAUQ2A0/edit?usp=sharing