

สังเขปรายวิชา 344-563 การสืบค้นสารสนเทศ

Information Retrieval

ภาคการศึกษาที่ 1 ปีการศึกษา 2561

คำอธิบายรายวิชา (Course Description)

รูปแบบการสืบค้นสารสนเทศ การประเมินการสืบค้น ภาษาการสอบถามและปฏิบัติการ การดำเนินงานกับข้อมูล อักขระและข้อมูลสื่อผสม การสร้างดัชนีเพื่อการค้นหา การสืบค้นสารสนเทศแบบ ขนานและแบบกระจาย ระบบห้องสมุด และบรรณานุกรม ห้องสมุดดิจิทัล

Information retrieval models; retrieval evaluation; query language and operations; text and multimedia operations; indexed and searching; parallel and distributed information retrieval; libraries and bibliographical systems; digital libraries

ผลการเรียนที่คาดหวังของรายวิชา (Course Learning Outcomes : CLOs)

1. อธิบายการสกัดคุณลักษณะแทนข้อมูลที่ไม่เป็นโครงสร้าง เช่น ข้อความ
2. คำนวณหาค่าความคล้ายและความต่างของข้อมูล
3. แสดงการค้นคืนแบบ Boolean
4. แสดงการค้นคืนแบบ Vector Space Model
5. คำนวณหาค่าการวัดประเมินผลการค้นคืนข้อมูล
6. แสดงการจำแนกข้อความโดยใช้เทคนิคอย่างง่าย
7. อธิบายการสกัดคุณลักษณะพื้นฐานจากรูปภาพ
8. อธิบายแนวคิดการค้นคืนรูปภาพ
9. เขียนโปรแกรมดึงข้อความจาก Twitter
10. ทำการทดลอง เปรียบเทียบวิธีการค้นคืนข้อมูล
11. พัฒนาโปรแกรมคอมพิวเตอร์สร้างระบบการค้นคืน

ความสัมพันธ์ระหว่างผลการเรียนที่คาดหวังของรายวิชา (CLOs) กับ ELOs หลักสูตร วท.ม. วิทยาการคอมพิวเตอร์*

	ELOs หลักสูตร วท.ม. วิทยาการคอมพิวเตอร์*								
	1	2	3	4	5	6	7	8	9
1. อธิบายการสกัดคุณลักษณะแทนข้อมูลที่ไม่เป็นโครงสร้าง	✓		✓				✓		
2. คำนวณหาค่าความคล้ายและความต่างของข้อมูล	✓		✓				✓		
3. แสดงการค้นคืนแบบ Boolean	✓		✓				✓		
4. แสดงการค้นคืนแบบ Vector Space Model	✓		✓				✓		
5. คำนวณค่าการวัดประเมินผลการค้นคืนข้อมูล	✓		✓				✓		
6. แสดงการจำแนกข้อความโดยใช้เทคนิคอย่างง่าย	✓		✓				✓		
7. อธิบายการสกัดคุณลักษณะพื้นฐานจากรูปภาพ	✓		✓				✓		
8. อธิบายแนวคิดการค้นคืนรูปภาพ	✓		✓				✓		
9. เขียนโปรแกรมดึงข้อความจาก Twitter	✓	✓	✓		✓		✓		
10. ทำการทดลอง เปรียบเทียบวิธีการค้นคืนข้อมูล	✓	✓	✓		✓		✓	✓	✓
11. พัฒนาโปรแกรมคอมพิวเตอร์สร้างระบบการค้นคืน	✓	✓	✓		✓		✓	✓	✓

***ผลการเรียนรู้ที่คาดหวัง (Expected Learning Outcomes : ELOs) หลักสูตร วท.ม. วิทยาการคอมพิวเตอร์**

1. แก้ปัญหาด้วยกระบวนการทางวิทยาการคอมพิวเตอร์บนพื้นฐานความรู้ทางวิทยาศาสตร์และคณิตศาสตร์ โดยเลือกประยุกต์เทคโนโลยีคอมพิวเตอร์และบูรณาการกับศาสตร์อื่นได้อย่างเหมาะสม
2. พัฒนาซอฟต์แวร์ที่ถูกต้องตามหลักการวิศวกรรมซอฟต์แวร์
3. คิดอย่างมีวิจารณญาณและจัดการข้อมูลสารสนเทศอย่างเป็นระบบที่สนองต่อการใช้งานตาม วัตถุประสงค์อย่างเหมาะสมและสร้างสรรค์
4. ดูแลระบบคอมพิวเตอร์และเครือข่ายให้มั่นคงต่อการใช้งาน
5. แสดงพฤติกรรมที่สะท้อนความมีคุณธรรม จริยธรรม ความซื่อสัตย์ เคารพกฎระเบียบและกฎเกณฑ์ขององค์กร และสังคม มีความรับผิดชอบต่อตนเองและสังคมโดยยึดพระราชปณิธาน ของพระบรมราชชนก “ขอให้ถือประโยชน์ของเพื่อนมนุษย์เป็นกิจที่หนึ่ง”
6. ติดตามความก้าวหน้าทางวิชาการและผลกระทบของการเปลี่ยนแปลงทางเทคโนโลยีอย่างต่อเนื่อง และเสนอความสนใจที่ต้องการพัฒนาตนเองด้านวิทยาการคอมพิวเตอร์ผ่าน กระบวนการแก้ปัญหาและค้นคว้าทั้งในและนอกห้องเรียนด้วยตนเอง และภายใต้คำปรึกษาของอาจารย์ที่ปรึกษา สำหรับการพัฒนาตนเองและการวิจัยระดับปริญญาโทเพื่อวิทยานิพนธ์/ สารนิพนธ์ นำไปสู่การเรียนรู้ที่ยั่งยืนด้วยตนเอง
7. มีความรู้และความเข้าใจเกี่ยวกับหลักการและทฤษฎีที่สำคัญทางวิทยาการคอมพิวเตอร์ และ ถ่ายทอดความรู้ทางด้านวิทยาการคอมพิวเตอร์ไปยังผู้ที่เกี่ยวข้องได้อย่างถูกต้องและเหมาะสม
8. ทำงานเป็นทีม ทั้งในบทบาทผู้นำและผู้ร่วมทีม อย่างมีบุคลิกภาพที่เหมาะสมต่อจรรยาบรรณ ทางวิชาการและวิชาชีพ
9. สื่อสารทั้งภาษาไทย และภาษาอังกฤษ โดยใช้รูปแบบการสื่อสารได้อย่างถูกต้องและเหมาะสม

เนื้อหาวิชา ประกอบด้วย

- บทที่ 1 Boolean Retrieval
- บทที่ 2 Term Vocabulary
- บทที่ 3 Vector Space Model and Ranked retrieval
- บทที่ 4 Web Scraping
- บทที่ 5 Evaluation in Information Retrieval
- บทที่ 6 Vector Space Classification
- บทที่ 7 Text Classification and Sentiment Analysis
- บทที่ 8 Image and Video Retrieval Concepts

วิธีการเรียนการสอน

เริ่มสอนเมื่อ 20 สิงหาคม 2560	จำนวนหน่วยกิต	3 (3-0-6) หน่วยกิต
สอนโดยการบรรยาย และ Active Learning สัปดาห์ละ 3 ชั่วโมง	รวม	45 ชั่วโมง
ศึกษาด้วยตนเองและทำงานที่ได้รับมอบหมาย สัปดาห์ละ 6 ชั่วโมง	รวม	90 ชั่วโมง

การวัดผล/ประเมินผล อิงเกณฑ์และอิงกลุ่ม จากสัดส่วนคะแนนของการวัดผลดังนี้

กลางภาค	35 %
ปลายภาค	35 %
แบบฝึกหัด การบ้าน ผลงานการพัฒนาโปรแกรม และการนำเสนอผลงาน	30 %

เกณฑ์การให้ระดับชั้นคะแนนเบื้องต้น

A	80 – 100	C	50 – 59
B+	75 – 79	D+	45 – 49
B	70 – 74	D	40 – 44
C+	60 – 69	E	0 - 39

ทั้งนี้ อาจมีการเปลี่ยนแปลงตามอิงเกณฑ์และอิงกลุ่ม

อาจารย์ผู้สอน อ.นิวรรณ วัฒนกิจรุ่งโรจน์ ห้องทำงาน CS313 E-mail: niwan.w@psu.ac.th

เอกสารประกอบการเรียนการสอน

- นิวรรณ วัฒนกิจรุ่งโรจน์, "เอกสารประกอบการสอนวิชาการสืบค้นสารสนเทศ" 2561
- Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, Introduction to Information Retrieval, Cambridge University Press. 2008.

<http://nlp.stanford.edu/IR-book/information-retrieval-book.html>

สังเขปแผนการสอน

รายวิชา 344-563 การสืบค้นสารสนเทศ

Information Retrieval

สัปดาห์ที่	รายละเอียดเนื้อหาวิชา
1	<p>ข้อตกลงเบื้องต้นเกี่ยวกับการเรียนการสอน</p> <ul style="list-style-type: none"> • วัตถุประสงค์ • ผลการเรียนรู้ที่คาดหวังของรายวิชา (CLOs) และความสอดคล้องกับผลการเรียนรู้ที่คาดหวังของหลักสูตรที่เรียน (ELOs) • รายละเอียดเนื้อหาวิชาที่จะสอน • การวัดผลและการประเมินผล • ข้อตกลงอื่น <p>0. แนะนำการใช้โปรแกรม R</p>
2-3	<p>1. Boolean Retrieval</p> <ul style="list-style-type: none"> • Unstructured vs. structured data • Term-document incidence matrix • Incidence vectors • Inverted Index
4	<p>2. Term Vocabulary</p> <ul style="list-style-type: none"> • Word segmentation, Tokenization • Stop words • Stemming

สัปดาห์ที่	รายละเอียดเนื้อหาวิชา
5-6	3. Vector Space Model and Ranked retrieval <ul style="list-style-type: none">• Ranked retrieval• Jaccard coefficient• Binary incidence matrix• Count matrix• Bag of words model• Term frequency• Document frequency• Tf-idf ranking• Documents as vectors• Distance vs. similarity
7-8	4. Web Scraping <ul style="list-style-type: none">• Scraping Twitter data With R
9	สอบกลางภาค

สัปดาห์ที่	รายละเอียดเนื้อหาวิชา
10-11	5. Evaluation in Information Retrieval <ul style="list-style-type: none"> • Unranked evaluation • Ranked evaluation
12-13	6. Vector Space Classification <ul style="list-style-type: none"> • kNN classification • Linear classifiers
14-15	7. Text Classification and Sentiment Analysis <ul style="list-style-type: none"> • Text Classification • Sentiment Analysis • Evaluating Classification
16	8. Image and Video Retrieval Concepts <ul style="list-style-type: none"> • Color features • Texture features • Shape features • Image retrieval • Extending image retrieval to video retrieval
17	สอบปลายภาค