การพัฒนาฐานความรู้ออนโทโลยี สำหรับวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ

Developing an Ontology Knowledge Based for Automatic Online News Analysis

วิชุดา โชติรัตน์ (Wichuda Chotirat)* ผุสดี บุญรอด (Pudsadee Boonrawd)* และ
ศจีมาจ ณ วิเชียร (Sageemas Na Wichian)**

บทคัดย่อ

บทความวิจัยนี้มีวัตถุประสงค์เพื่อทำการศึกษา ออกแบบ พัฒนา และทดสอบฐานความรู้ออนโทโลยีสำหรับวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติที่เน้นการออกแบบตามหลักวงจรการออกแบบออนโทโลยี(Ontology Life Cycle) ซึ่งผู้วิจัยได้ทำการศึกษาปัจจัยที่เกี่ยวข้องกับการวิเคราะห์ข่าวออนไลน์ พิจารณาคำสำคัญซึ่งมีผลต่อการวิเคราะห์เนื้อหาข่าวออนไลน์ และทำการพัฒนาฐานความรู้ออนโทโลยีโดยใช้โปรแกรม Hozo-Ontology Editor ซึ่งเป็นโปรแกรมที่สามารถใช้งานได้โดยไม่เสียค่าใช้จ่ายสำหรับการออกแบบ พัฒนาออนโทโลยี จากนั้นทำการทดสอบฐานความรู้ออนโทโลยีที่พัฒนาขึ้น โดยผู้เชี่ยวชาญซึ่งมีผลการทดสอบอยู่ในระดับดี ($\bar{x}$ = 3.61 และ *S.D.*= 0.16) สามารถสรุปได้ว่าฐานความรู้ออนโทโลยีที่พัฒนาขึ้นมีการออกแบบที่ถูกต้อง และสอดคล้องสำหรับนำไปประยุกต์ใช้วิเคราะห์ข่าวออนไลน์โดยอัตโนมัติได้เหมาะสม

คำสำคัญ: ออนโทโลยี ข่าวออนไลน์ ระบบวิเคราะห์ข่าว การวิเคราะห์ข้อความ

Abstract

This paper presents a guideline for design development and ontology knowledge based testing for automatic analysis of news online. According to the Ontology Life Cycle, researchers have studied the factors related to the analysis of online news and considered keywords that affect analysis of content and developed an ontology knowledge based. This was developed using Hozo-Ontology Editor, this is a freeware program for infrastructure development and then evaluation was conducted. Results of the evaluation by experts show that the ontology knowledge based is a good level ($\bar{x}$ =3.61, *S.D.*= 0.16). Therefore this ontology knowledge base displays efficiency with this particular issue and could be applied for developing an automatic analysis of news online appropriately in further research.

Keyword: Ontology, News Online, News Analysis System, Text Analysis.

1. บทนำ

การเผยแพร่ข่าวสารในรูปแบบข่าวออนไลน์ (News Online) เป็นรูปแบบการนำเสนอเรื่องราวต่าง ๆ ที่เกิดขึ้นผ่านเครือข่ายอินเทอร์เน็ต และมีผลกระทบต่อคนจำนวนมากในสังคม ซึ่งหน่วยงานทั้งภาครัฐ และเอกชนตระหนักถึงความสำคัญของข้อมูลข่าวออนไลน์อย่างสูง และพบว่ามีการพัฒนาระบบจัดเก็บข้อมูลข่าวในรูปแบบฐานข้อมูลจำนวนมากเพื่อการค้นคืนข้อมูล และการนำข้อมูลมาใช้ประโยชน์ด้านการวิเคราะห์แนวโน้มสถานการณ์ที่เกิดขึ้น แต่เนื่องจากปริมาณข่าวที่มีจำนวนสูงจึงต้องอาศัยระยะเวลานานในการวิเคราะห์ข้อมูล

* *ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
** *วิทยาลัยเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อการออกแบบและพัฒนาระบบวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ โดยทำการศึกษาปัจจัยที่มีผลต่อการวิเคราะห์เนื้อหาข่าวออนไลน์ เพื่อการออกแบบ และพัฒนาฐานความรู้ออนโทโลยีสำหรับอธิบาย สิ่งต่างๆ ที่ถือเป็นองค์ประกอบสำคัญในเนื้อหาข่าว ส่งผลให้ระบบสามารถวิเคราะห์เนื้อหา และอนุมานข้อมูลต่างๆ ที่เกี่ยวข้องในเนื้อหาข่าวออนไลน์ได้อย่างมีประสิทธิภาพ ซึ่งงานวิจัยนี้ประกอบด้วยส่วนต่างๆ ที่สำคัญดังนี้ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง การออกแบบพัฒนาฐานความรู้ออนโทโลยี การทดลอง ผลการทดลอง สรุปผลและข้อเสนอแนะ เป็นต้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ข่าวออนไลน์ (News Online)

ข่าวออนไลน์จัดเป็นสารสนเทศที่มีอิทธิพลอย่างสูงต่อการรับรู้ข่าวสารในยุคปัจจุบันโดยนำเสนอเหตุการณ์ที่ได้รับความสนใจ และมีผลกระทบต่อคนส่วนใหญ่ผ่านช่องทางเครือข่ายอินเทอร์เน็ต ประกอบด้วยคำสำคัญ รูปภาพ และแหล่งที่มาของข้อมูลทำให้สามารถนำข้อมูลมาใช้ประโยชน์เพื่อการศึกษาข้อเท็จจริงสำหรับการตัดสินใจได้ โดยสามารถกำหนดกรอบที่ชัดเจนในการวิเคราะห์เนื้อหาได้ 4 ด้าน [1] คือ 1) เนื้อหาข่าวต้องนำเสนอแนวคิดที่ชัดเจน 2) ข้อมูลมาจากการรวบรวมและการสังเกตของผู้เขียน 3) เนื้อหาข่าวมีความเป็นไปได้ 4) ผู้รับข่าวสารต้องสามารถมองเห็นความเป็นจริงได้โดยไม่ได้เกิดจากการจินตนาการของผู้เขียนเท่านั้น

เนื้อหาข่าวออนไลน์ที่มีความสมบูรณ์ สามารถตอบคำถามพื้นฐาน (5W1H) ประกอบด้วย ใคร (Who) อะไร (What) ที่ไหน (Where) เมื่อไหร่ (When) ทำไม (Why) และอย่างไร (How) [2] ซึ่งหากสามารถทำการแยกคำสำคัญต่างๆ จากเนื้อหาข่าว จะช่วยให้สามารถวิเคราะห์เนื้อหาข่าวได้ ดังโครงการวิจัย Lydia [3] ซึ่งมีวัตถุประสงค์เพื่อทำการวิเคราะห์เนื้อหาข่าวออนไลน์โดยสร้างแบบจำลองเชิงสัมพันธ์ของสิ่งต่างๆ ได้แก่ บุคคล สถานที่ และอื่นๆ ที่เกี่ยวข้องด้วยเทคนิคการประมวลผลภาษาธรรมชาติ (NLP) ส่งผลให้ระบบสามารถทำการวิเคราะห์ข้อมูลทางสถิติเพื่อการทำนาย และอนุมานข้อมูลต่างๆ ที่เกี่ยวข้องกันได้อย่างอัตโนมัติ

2.2 ฐานความรู้ออนโทโลยี (Ontology Knowledge Based)

ฐานความรู้ออนโทโลยี คือ การอธิบายรูปแบบโครงสร้างความสัมพันธ์ระหว่างข้อมูลในขอบเขตที่สนใจ (Domain) เชิงมโนภาพ โดยสามารถใช้โครงร่างพื้นฐานความสัมพันธ์ของเทอม (Term) สำหรับเพื่อใช้เป็นฐานความรู้ได้ [4], [5] โดยฐานความรู้ออนโทโลยีมีองค์ประกอบดังนี้

1)แนวคิด (Concepts) คือ ขอบเขตของความรู้ หรือเรื่องใดเรื่องหนึ่ง และสามารถทำการอธิบายรายละเอียดได้

2) คุณสมบัติ (Properties) คือ คุณสมบัติต่างๆ ที่นำมาใช้อธิบายรายละเอียดแนวคิด

3) ความสัมพันธ์ (Relationships) คือ รูปแบบการแสดงความสัมพันธ์ระหว่างแนวคิด

4) ข้อกำหนดการสร้างความสัมพันธ์ (Axioms) คือ เงื่อนไขหรือตรรกะในการแปลงความสัมพันธ์ระหว่างแนวคิดกับแนวคิด หรือคุณสมบัติ เพื่อการแปลงความหมายที่ถูกต้อง

5) ตัวอย่างข้อมูล (Instance) คือ คำศัพท์ที่มีการกำหนดความหมายไว้ในออนโทโลยีเรื่องนั้นๆ

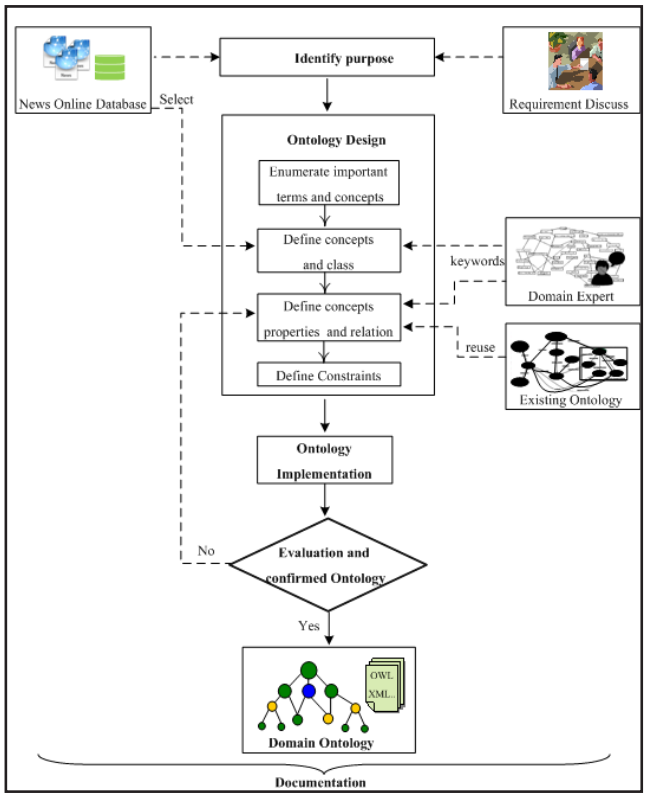
ออนโทโลยีจะถูกอธิบายคลาส (Class) และคุณสมบัติหรือคุณลักษณะด้วยสล็อต (Slot) ซึ่งอาจมีซับคลาส (Subclass) เพื่ออธิบายรายละเอียดของคลาสนั้นๆ โดยรูปแบบของการบรรยายจะขึ้นอยู่กับภาษาที่ใช้ ได้แก่ ภาษา RDFS และOWL เป็นต้น

2.3 โปรแกรมพัฒนาออนโทโลยี (Hozo-Ontology Editor)

การออกแบบและพัฒนาฐานความรู้ออนโทโลยีในงานวิจัยนี้ผู้วิจัยเลือกใช้โปรแกรม Hozo-Ontology Editor ที่พัฒนาขึ้นโดยมหาวิทยาลัยโอซากา (Osaka University) ประเทศญี่ปุ่น ซึ่งสามารถทำการจัดเก็บองค์ความรู้ในรูปแบบของออนโทโลยีได้อย่างสะดวกและไม่เสียค่าใช้จ่ายในการนำมาประยุกต์ใช้ [6]

3. กระบวนการพัฒนาฐานความรู้ออนโทโลยี

ในงานวิจัยนี้นำเสนอกระบวนการออกแบบ และพัฒนาฐานความรู้ออนโทโลยี โดเมนข่าวออนไลน์ ตามวงจรการออกแบบออนโทโลยี (Ontology Life Cycle) [7] และวงจรการพัฒนาซอฟต์แวร์ (Software Life Cycle Processes) ตามหลัก IEEE Standard 1074 -2006 [8], [9] ดังภาพที่ 1



ภาพที่ 1 *แผนภาพกระบวนการพัฒนาฐานความรู้ออนโทโลยี*

จากแผนภาพกระบวนการพัฒนาฐานความรู้ออนโทโลยี สามารถอธิบายรายละเอียดการดำเนินงานได้ดังนี้

3.1 การระบุวัตถุประสงค์งานวิจัย (Identify Purpose)

ผู้วิจัยมีวัตถุประสงค์เพื่อออกแบบพัฒนาระบบวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ อาศัยเทคนิคฐานความรู้ออนโทโลยีเพื่อการอธิบายองค์ประกอบของสิ่งต่างๆ ในเนื้อหาข่าวทำให้สามารถวิเคราะห์ และอนุมานข้อมูลที่เกี่ยวข้องได้

3.2 การออกแบบฐานความรู้อ้อนโทโลยี (Ontology Knowledge Based Design)

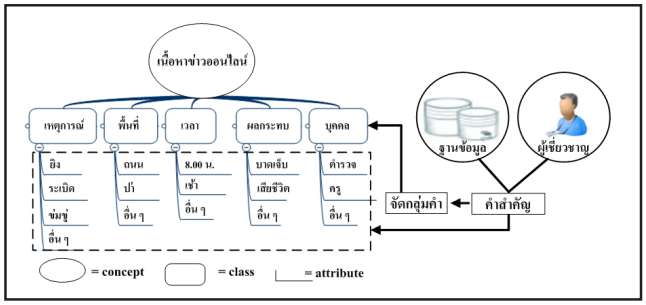
ผู้วิจัยได้ดำเนินการออกแบบออนโทโลยีตามหลักวงจรการออกแบบออนโทโลยี (Ontology Life Cycle) ดังนี้ [7], [9]

3.2.1) Determine Scope งานวิจัยนี้เน้นการใช้ประโยชน์จากข้อมูลข่าวออนไลน์เพื่อการวิเคราะห์เนื้อหาและอนุมานความสัมพันธ์ของสิ่งต่างๆ จากเนื้อหาข่าวออนไลน์

3.2.2) Consider Reuse การพิจารณาเลือกใช้ตัวแบบออนโทโลยีที่มีอยู่แล้วนำมาปรับใช้ซ้ำเพื่อเป็นแนวทางสำหรับการพัฒนาฐานความรู้ออนโทโลยีได้อย่างรวดเร็ว อาทิ ออนโทโลยีเหตุการณ์การก่อการร้าย [10] และออนโทโลยีสำหรับข่าวออนไลน์ [11] เป็นต้น

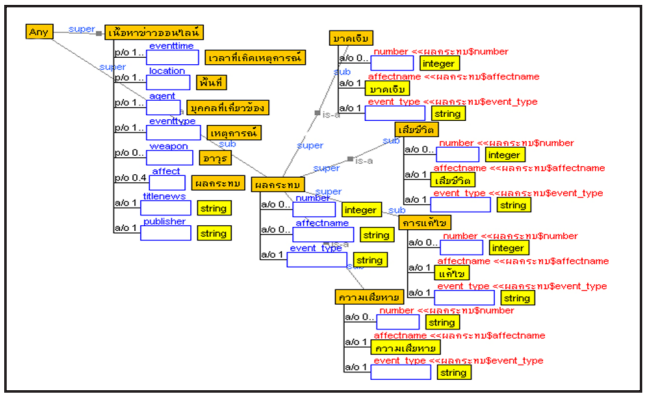
3.2.3) Enumerate Term ทำการระบุขอบเขตการพัฒนาฐานความรู้ออนโทโลยี โดเมนข่าวออนไลน์ที่เกี่ยวข้องกับพื้นที่จังหวัดชายแดนภาคใต้เน้นการวิเคราะห์เนื้อหาข่าวจากคำสำคัญที่ปรากฏในเนื้อหาข่าว

3.2.4) Define Class และ Concepts ทำการระบุคลาสและคอนเซพท์ที่เกี่ยวข้อง ซึ่งผู้วิจัยมีแนวทางการกำหนดคลาสแบบผสมผสาน (Combination development)โดยรวบรวมคำสำคัญที่ปรากฏในเนื้อหาข่าว จัดกลุ่มคำสำคัญ และทำการระบุคำสำคัญอื่นๆ ที่เกี่ยวข้องเพิ่มเติม ดังแสดงในภาพที่ 2



ภาพที่ 2 *การพิจารณาคำสำคัญเพื่อระบุคลาสออนโทโลยี*

3.2.5) Define Properties and Relations ระบุคุณลักษณะ หรือความสัมพันธ์ระหว่างคอนเซพท์ภายในขอบเขตที่สนใจ ดังแสดงในภาพที่ 3



ภาพที่ 3 *ภาพแสดงความสัมพันธ์ของคอนเซพท์*

ก) ความสัมพันธ์ระหว่างมโนทัศน์ “is-a” (จัดเป็น) ตัวอย่างเช่น ตาย is-a ผลกระทบ

ข) ความสัมพันธ์ชนิดคุณสมบัติแบบ “part-of” (การเป็นส่วนหนึ่งของ) ตัวอย่างเช่น ผลกระทบ part-of(p/o) เนื้อหาข่าวออนไลน์

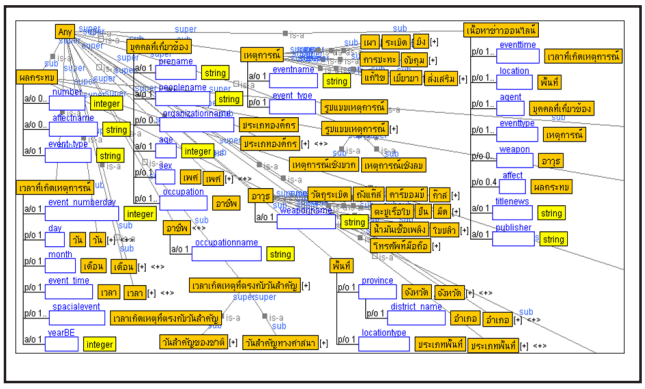
ค) ความสัมพันธ์ชนิดคุณสมบัติแบบ “attribute-of” (การเป็นคุณลักษณะ) ตัวอย่างเช่น Affect_number attribute-of (a/o) ผลกระทบ

3.2.6) Define Constraints ระบุข้อกำหนดต่าง ๆ ของข้อมูล เช่น คลาสของเพศมีการระบุข้อกำหนดของข้อมูลเพียง 2 ค่า คือ เพศชาย และ เพศหญิง

3.2.7) Create Instances กำหนดค่าตัวแทนข้อมูลของคอนเซพท์ต่าง ๆ ตัวอย่างเช่น ปืน เป็นอินสแตนซ์ของคลาสเหตุการณ์อาวุธ

3.3 การดำเนินการออนโทโลยี (Ontology Implement)

จากการศึกษาสภาพปัญหา และรวบรวมองค์ความรู้ที่เกี่ยวข้องเพื่อการออกแบบฐานความรู้ออนโทโลยีสำหรับวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ ผู้วิจัยจึงทำการสร้างฐานความรู้ออนโทโลยีด้วยโปรแกรม Hozo-Ontology Editor โดยมีผลลัพธ์โครงสร้างฐานความรู้ออนโทโลยีโดเมนข่าวออนไลน์ ภายใต้ขอบเขตของข่าวที่เกี่ยวข้องกับพื้นที่จังหวัดชายแดนภาคใต้ ดังแสดงในภาพที่ 4 ประกอบด้วยคลาสต่าง ๆ ที่สำคัญ ได้แก่ เวลาที่เกิดเหตุการณ์ พื้นที่ บุคคลที่เกี่ยวข้อง รูปแบบเหตุการณ์ อาวุธ และผลกระทบ เป็นต้น



ภาพที่ 4 *ภาพฐานความรู้ออนโทโลยีโดเมนข่าวออนไลน์*

จากนั้นทำการส่งออก (Export) เพื่อแปลงข้อมูลให้อยู่ในรูปแบบภาษา OWL (Web Ontology Language) ซึ่งเป็นภาษามาตรฐานสำหรับการพัฒนาโครงสร้างออนโทโลยีตามแนวทางเว็บเชิงความหมาย (Semantic Web) ที่มีแบบแผนอธิบายลำดับชั้น และความสัมพันธ์ระหว่าง Resource บนพื้นฐานภาษา RDF และ RDFS รวมทั้งเซทของกฎที่ได้จากเครื่องสามารถสร้างข้อสรุปเชิงตรรกะ และอนุญาตให้มีซับคลาสที่สืบทอดมาจากคุณสมบัติเดียวกันได้ ทำให้ผู้วิจัยสามารถนำมาปรับใช้เพื่อการเขียนโปรแกรมสำหรับอธิบายความหมาย และความสัมพันธ์ต่าง ๆ ของคำตามตรรกะได้ถูกต้อง ดังภาพที่ 5

```
<owl:Class rdf:ID = "เนื้อหาข่าวออนไลน์">
  <rdfs:label>DeepSouthNews</rdfs:label>
  <rdfs:subClassOf rdf:resource="#Any" />
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:minCardinality rdf:datatype= http://www.w3.org/2001
      XMLSchema#nonNegativeInteger">1</owl:minCardinality>
      <owl:onProperty rdf:resource="#has_eventtime" />
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
    <owl:onProperty rdf:resource="#has_eventtime" />
    <owl:allValuesFrom rdf:resource="#เวลาที่เกิดเหตุการณ์" />
    </owl:Restriction>
  </rdfs:subClassOf>
  …
</owl:Class>
```

ภาพที่ 5 *ฐานความรู้ออนโทโลยีในรูปแบบภาษา OWL*

4. การทดลองและผลการทดลอง

จากการศึกษาวัตถุประสงค์การพัฒนาระบบวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ ผู้วิจัยทำการสุ่มเลือกข่าวออนไลน์ภาษาไทยที่เกี่ยวข้องกับพื้นที่จังหวัดชายแดนใต้ จำนวน 200 ข่าว เพื่อใช้เป็นข้อมูลการทดลองโดยมีผลการทดลองดังนี้

4.1 การพิจารณาคำสำคัญจากเนื้อหาข่าวออนไลน์

ทำการพิจารณาความถี่การปรากฏคำต่าง ๆ ในเนื้อหาข่าวออนไลน์จากฐานข้อมูลข่าวภาษาไทย และจัดกลุ่มคำสำคัญที่มีความสอดคล้องกัน ดังตารางที่ 1

ตารางที่ 1 *ตารางการกำหนดคลาสในออนโทโลยี*

คลาส	คำอธิบาย และตัวอย่างข้อมูล
บุคคลที่เกี่ยวข้อง	บุคคลที่เกี่ยวข้อง เช่น ชาวบ้าน ตำรวจ
สถานที่เกิดเหตุ	สถานที่เกิดเหตุการณ์ เช่น ถนน จังหวัดยะลา
อาวุธที่ใช้	อาวุธที่ใช้ เช่น ปืน วัตถุระเบิด
ผลกระทบ	ผลกระทบที่เกิด เช่น เสียชีวิต
รูปแบบเหตุการณ์	เหตุการณ์ที่เกิด เช่น ยิง ระเบิด
เวลาที่เกิดเหตุ	วันและเวลาที่เกิดเหตุ เช่น 14 ก.พ. 54 08.30น.

4.2 ประเมินคำสำคัญที่มีผลต่อการวิเคราะห์ข่าวออนไลน์

การพิจารณาคำสำคัญ เพื่อกำหนดคลาสสำหรับฐานความรู้ออนโทโลยีโดเมนข่าวออนไลน์ที่เหมาะสม และสอดคล้องสำหรับการพัฒนาระบบวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ มีกระบวนการพิจารณาความสำคัญของคำในคลาสต่างๆ ดังนี้

1) การหาค่าความถี่ของคำที่ปรากฏในเนื้อหาข่าวออนไลน์ ดังสมการที่ (1), (2)

ก) การความถี่การปรากฏของคำสำคัญ (TF-IDF)

$$tf\text{-}idf_{t,d} = TF_{t,d} \times idf_t \quad (1)$$

ข) ค่าผกผันความถี่การปรากฏของคำสำคัญ (IDF)

$$idf_t = \log \frac{N}{df} \quad (2)$$

2) การหาค่าน้ำหนักของคำสำคัญในเนื้อหาข่าว W(t,d) ดังสมการที่ (3) [12]

ค) ค่าน้ำหนักของคำสำคัญในเนื้อหาข่าว

$$W_{(t,d)} = \frac{tf(t,d) \times \log_2(\frac{N}{N_t} + 0.01)}{\sqrt{\sum_{t \in d}[\, tf(t,d) \times \log_2(\frac{N}{N_t} + 0.01)\,]}} \quad (3)$$

โดย $W_{(t,d)}$ คือ ค่าน้ำหนักความถี่ของคำ (t) ที่ปรากฏ

tf คือ จำนวนการปรากฏคำสำคัญ (t)

N คือ จำนวนเอกสารข่าวออนไลน์ (200 ข่าว)

Nt คือ จำนวนเอกสารข่าวที่ปรากฏคำที่สนใจ

การพิจารณาค่า Term Frequency-Inverse Document Frequency (TF-IDF) เป็นการวัดค่าน้ำหนักเพื่อการประเมินค่าความสำคัญของคำนั้นๆ ในเนื้อหาข่าวออนไลน์ ร่วมกับการพิจารณาค่าน้ำหนักของคำ $W_{(t,d)}$ ซึ่งหากมีค่าน้ำหนักของคำสูงแสดงว่าคำสำคัญในคลาสดังกล่าวมีความถี่ของการปรากฏในเนื้อหาข่าวออนไลน์สูงเหมาะสำหรับกำหนดในคลาสดังกล่าว

จากตารางที่ 2 พบว่าคำสำคัญที่กล่าวถึงบุคคล หรือองค์กรที่เกี่ยวข้อง รูปแบบของเหตุการณ์ สถานที่เกิดเหตุการณ์ ผลกระทบจากเหตุการณ์ ช่วงเวลาของเหตุการณ์ และอาวุธที่ใช้ในการเกิดเหตุ มีค่าความถี่ TF-IDF และค่าน้ำหนักของการปรากฏคำสำคัญดังกล่าวในเนื้อหาข่าวออนไลน์ $W_{(t,d)}$ ตามลำดับ

ตารางที่ 2 *ตารางการพิจารณาค่าความสำคัญของคลาส*

คลาสที่สนใจ	TF-IDF	$W_{(t,d)}$
1. บุคคล หรือองค์กรที่เกี่ยวข้อง	1.74	12.08
2. รูปแบบของเหตุการณ์	6.10	12.08
3. สถานที่เกิดเหตุการณ์	6.69	12.33
4. ผลกระทบจากเหตุการณ์	23.11	11.63
5. ช่วงเวลาของเหตุการณ์	35.23	11.20
6. อาวุธที่ใช้ในที่เกิดเหตุ	39.04	9.89

สามารถนำคลาสดังกล่าวมากำหนดเป็นคลาสสำคัญสำหรับระบบวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติได้อย่างสอดคล้อง และเหมาะสม

4.3 การทดสอบฐานความรู้ออนโทโลยี

ทดสอบฐานความรู้ออนโทโลยีโดยพิจารณาจากขอบเขตวัตถุประสงค์การวิจัย ผู้วิจัยได้แบ่งประเด็นการทดสอบออกเป็น 4 ด้าน คือ 1) การสื่อความหมาย (Interpretability: I) 2) ความถูกต้อง (Accuracy: A) 3) ความชัดเจน (Clarity: C) 4) ความครอบคลุม (Comprehensiveness: O) ซึ่งค่าตัวแปรทั้ง 4 ด้านผ่านการพิจารณาประเมินผลโดยผู้เชี่ยวชาญ ซึ่งมีระดับคะแนนตั้งแต่ 0 คะแนน (เห็นด้วยน้อยที่สุด) จนถึงระดับ 5 คะแนน (เห็นด้วยมากที่สุด) ทำการหาผลการทดสอบโดยรวมจากการคำนวณค่าตัวแปรทั้ง 4 ดังสมการที่ (4)

ก) $\sum Score = I \times w_i + A \times w_a + C \times w_c + O \times w_o \quad (4)$

กำหนดค่าถ่วงน้ำหนัก Weight (w) = 0.25 โดย $\sum w = 1$

ตารางที่ 3 *ผลการทดสอบฐานความรู้ออนโทโลยี*

ผลการทดสอบฐานความรู้ออนโทโลยีโดเมนข่าวออนไลน์			
ประเด็นการทดสอบ	$\bar{x}$	*Score*	*S.D.*
การสื่อความหมาย (I)	3.50	$w_i = 0.88$	0.09
ความถูกต้อง (A)	3.76	$w_a = 0.94$	0.30
ความชัดเจน (C)	3.50	$w_c = 0.88$	0.00
ความครอบคลุม (O)	3.61	$w_o = 0.92$	0.25
Total	3.61	$\sum Score = 3.62$	0.16

จากตารางการประเมินฐานความรู้ออนโทโลยีโดเมนข่าวออนไลน์ที่พัฒนาขึ้นโดยผู้เชี่ยวชาญ พบว่าอยู่ในระดับดีทั้ง 4 ด้าน คือ 1) การสื่อความหมาย 2) ความถูกต้อง 3) ความชัดเจน 4) ความครอบคลุม กล่าวคือ $\bar{x}$ = 3.61 และ *S.D.*= 0.16

5. สรุปผลการวิจัย และข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อการออกแบบและพัฒนาฐานความรู้ออนโทโลยีสำหรับวิเคราะห์ข่าวออนไลน์ภาษาไทยโดยอัตโนมัติ ซึ่งผู้วิจัยได้ดำเนินการจัดเตรียมข้อมูลข่าวออนไลน์ พิจารณาคำสำคัญที่มีผลต่อการวิเคราะห์เนื้อหาข่าว โดยหาค่าความถี่ของคำที่ปรากฏในเนื้อหาข่าว (TF-IDF) ส่งผลให้สามารถรวบรวมคำสำคัญที่เหมาะสมสำหรับการพัฒนาฐานความรู้ออนโทโลยี โดเมนข่าวออนไลน์ที่เกี่ยวข้องกับพื้นที่ 5 จังหวัดชายแดนภาคใต้ จากนั้นทำการทดสอบฐานความรู้ออนโทโลยีที่พัฒนาขึ้นโดยผู้เชี่ยวชาญ ซึ่งมีผลการทดสอบโดยรวมอยู่ในระดับดี($\bar{x}$ = 3.61, *S.D.* = 0.16)

จากผลการวิจัยสามารถนำฐานความรู้ออนโทโลยีที่พัฒนาขึ้นไปประยุกต์ใช้ร่วมกับการวิเคราะห์ข้อมูลด้วยเทคนิคการทำเหมืองข้อมูล หรือเทคนิคอื่นๆ เพื่อประโยชน์การวิเคราะห์ข้อมูลที่รวดเร็ว และการศึกษาความสัมพันธ์ของสิ่งต่างๆ ที่สนใจจากเนื้อหาข่าวออนไลน์ภาษาไทยได้ในอนาคต

6. เอกสารอ้างอิง

[1] H. Clase de Vreese, "News framing : Theory and typology," *Information Design Journal + Document Design,* vol.13, no. 1, pp. 51-62, 2005.

[2] วีรพงษ์ พลนิกรกิจ, *การเขียนข่าววิทยุ กระจายเสียงและวิทยุโทรทัศน์เบื้องต้น*. กรุงเทพ : เอดิสันเพรส, 2545.

[3] L. Lloyd, D. Kechagias and S. Skiena, "Large Scale Online Text Analysis Using Lydia," APSA 2010 Annual Meeting Paper. [Online]. Available from: http://paper.ssrn.com/sol13/papers.cfm? [10 January, 2011].

[4] T. Gruber, *Ontology*, 2007 [Online]. Available from: http://tomgruber.org/writing/ontology-definition- 2007.htm [06 August, 2010].

[5] Seartout and et al., 1996 อ้างถึงใน "ระบบสนับสนุนการบำรุงรักษาออนโทโลยี," (ออนไลน์). สืบค้นวันที่ 20 ธันวาคม 2553. จาก http://naist.cpe.ku.ac.th/iknow/report/report2007/six.pdf

[6] K. Kozaki, E. Sunagawa and et al, "Hozo: an Ontology Development Environment-Treatment of Role Concept and Dependency Management," *Knowledg Engineering and Knowledge Management : Ontologies and the Semantic Web Lecture Notes in Computer Science*, pp. 155-163, 2002.

[7] N.F. Noy and D.L.Mc.Guinness, *Ontology Development 101 : A Guide to Creating Your First Ontology* [Online]. Available from:http://ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html [10 January, 2011].

[8] IEEE Standards Board. 1997. I*EEE Standard for Developing Software Life Cycle Processes* [Online]. Available from: http://www.softwareresearch.net/fileadmin/src/docs/teaching/SS05/PM/1074_DevLife CycleProcesses.PDF [10 January, 2011].

[9] G. Beydoun, "Reflecting on Ontologies: Towards Ontology-based Agent-oriented Software Engineering," *Proc. of the 5th Australasian Ontologies Workshop*, pp. 7-16, 2009.

[10] U. Inyaem, "Ontology-Based Terrorism Event Extract," *the1stInternational Conference on Information Science and Engineering (ICISE* 2009*)*, pp. 912-915, 2009.

[11] P. Jukub, "Hristo Tanev and Pinar Oezden Wennerberg. Extracting Violent Event From On-Line News for Ontology Population," *LNCS* 4439, pp. 287-300, 2007.

[12] Y. Li, J. Yuan, and X. Ye, "Method for Feature Word Weight Calculating," *Intelligent Computing and Intelligent Systems IEEE International Conference on ICIS* 2009, pp. 309-312, 2009.