บทที่ 2

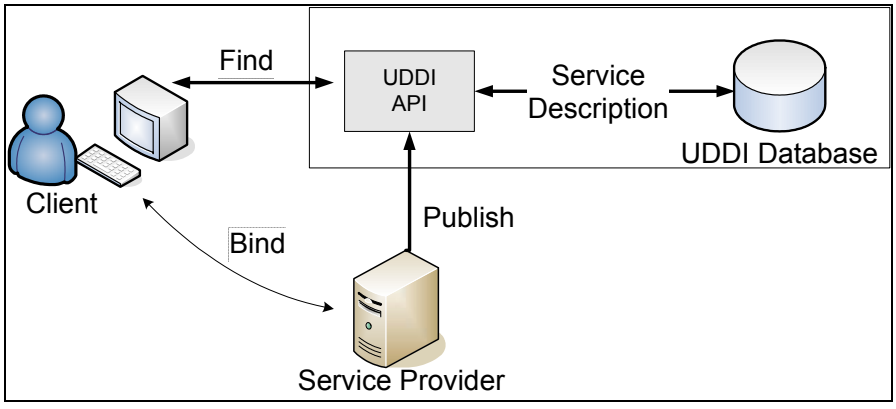
ทฤษฎีและหลักการ

บทนี้จะอธิบายหลักการค้นหาบริการแบบทั่วไป ซึ่งจะแสดงให้เห็นถึงปัญหาและวิธีการแก้ปัญหาของงานวิจัยที่ได้ศึกษามา นอกจากนี้ยังชี้ให้เห็นข้อจำกัดของวิธีการในงานวิจัยเหล่านั้น รวมไปถึงนำเสนอวิธีการแก้ปัญหาโดยวิธีของวิทยานิพนธ์นี้

2.1 การค้นพบบริการเว็บเซอร์วิส (Web Services Discovery)

เว็บเซอร์วิส คือ โปรแกรมที่ทำงานในระบบเครือข่าย และจะถูกเรียกใช้งานจากโปรแกรมอื่นในรูปแบบอาร์พีซี (Remote Procedure Call: RPC) แต่ละเว็บเซอร์วิสจะมีการอธิบายคุณลักษณะของบริการ (Service Description) เพื่อให้ผู้ใช้ทราบถึงการทำงานและวิธีการเรียกใช้เว็บเซอร์วิสนั้น ๆ คุณลักษณะของบริการจะถูกอธิบายโดยใช้ภาษาเอ็กซ์เอ็มแอล (XML: Extensible Markup Language) เช่นเดียวกับการส่งข้อมูลระหว่างเว็บเซอร์วิสที่อยู่ในรูปแบบเอ็กซ์เอ็มแอลบนโพรโทคอลเอชทีทีพี (Hyper Text Transport Protocol : HTTP) ทำให้เราสามารถเรียกใช้เว็บเซอร์วิสใด ๆ ได้โดยไม่ขึ้นอยู่กับฮาร์ดแวร์ ซอฟต์แวร์ หรือภาษาที่ใช้โปรแกรม [29]

ซึ่งในปัจจุบันส่วนที่ผู้ให้บริการทำการประกาศบริการและผู้ใช้ค้นหาบริการจะนิยมใช้กับระบบยูดีดีไอ (Universal Description, Discovery and Integration : UDDI) [5] ซึ่งมีรูปแบบการทำงานดังในรูปที่ 2.1



รูปที่ 2.1 กรอบระบบการทำงานของเว็บเซอร์วิส

รูปที่ 2.1แสดงกรอบระบบของเว็บเซอร์วิส การทำงานเริ่มต้นจากผู้ให้บริการส่งคำสั่งร้อง

ขอการประกาศบริการของตนตามโครงสร้างข้อมูลของยูดีดีไอผ่านส่วนติดต่อโปรแกรมประยุกต์ยูดีดีไอ (UDDI API) จากนั้นส่วนติดต่อโปรแกรมประยุกต์ยูดีดีไอจะจัดเก็บคำอธิบายบริการไว้ในฐานข้อมูล ตัวอย่างการใช้งานส่วนติดต่อโปรแกรมประยุกต์ในการประกาศบริการได้แก่ save_business, save_service, save_binding, save_tModel เมื่อมีผู้ใช้บริการส่งคำสั่งร้องขอการค้นหาบริการมายังยูดีดีไอ ส่วนติดต่อโปรแกรมประยุกต์ยูดีดีไอจะเป็นตัวรับคำสั่งและค้นหาบริการที่ตรงกับเงื่อนไขจากฐานข้อมูล จากนั้นก็จะส่งผลลัพธ์กลับไปยังผู้ใช้บริการเพื่อให้ผู้ใช้บริการใช้ข้อมูลที่ได้รับในการเรียกใช้บริการ ตัวอย่างการใช้งานส่วนติดต่อโปรแกรมประยุกต์ในการค้นหาบริการได้แก่ find_business, find_service, find_binding, find_tModel เป็นต้น

สรุปขั้นตอนการพัฒนาเว็บเซอร์วิสมีดังนี้

1) ทะเบียนการให้บริการ (Registry) เมื่อนักพัฒนาทำการสร้างเว็บเซอร์วิสแล้ว จะต้องเตรียมรายละเอียดของเซอร์วิส เช่น คำอธิบายของเซอร์วิส, ที่อยู่ที่ใช้ติดต่อ, วิธีการเข้าถึง เป็นต้น เพื่อไปทำการลงทะเบียนที่ศูนย์กลางการเก็บข้อมูลเซอร์วิส

2) การค้นพบหรือการค้นหาบริการ (Discovery/Find) ผู้พัฒนาหรือโปรแกรมที่มีการติดต่อกับเว็บเซอร์วิส จะทราบทราบที่อยู่ของบริการบนเว็บเซอร์วิส โดยการค้นหาข้อมูลจากศูนย์กลางการเก็บข้อมูลเซอร์วิส เมื่อพบเซอร์วิสที่ต้องการแล้ว ก็จะได้ทราบถึงรายละเอียดการทำงานของบริการและที่อยู่ที่จะเรียกใช้เว็บเซอร์วิส

3) การเรียกและประมวลผล (Invoke and Execute) หลังจากได้รายละเอียดของเซอร์วิสที่ต้องการ ผู้พัฒนาจะสร้างโปรแกรมตามที่เว็บเซอร์วิสกำหนดและเรียกใช้ฟังก์ชันบนเว็บเซอร์วิส

ส่วนที่สำคัญที่พูดถึงในการค้นหาบริการมีสองส่วนคือ ส่วนติดต่อโปรแกรมประยุกต์ยูดีดีไอ (UDDI API) และฐานข้อมูลยูดีดีไอ (UDDI Database) ในรูปที่ 2.1 ซึ่งมีหน้าที่ในการค้นหาและเก็บข้อมูลคำอธิบายของบริการตามลำดับ ซึ่งการค้นหาบริการที่ประกาศไว้ในยูดีดีไอโดยส่วนใหญ่เงื่อนไขที่ใช้ในการค้นหาได้แก่

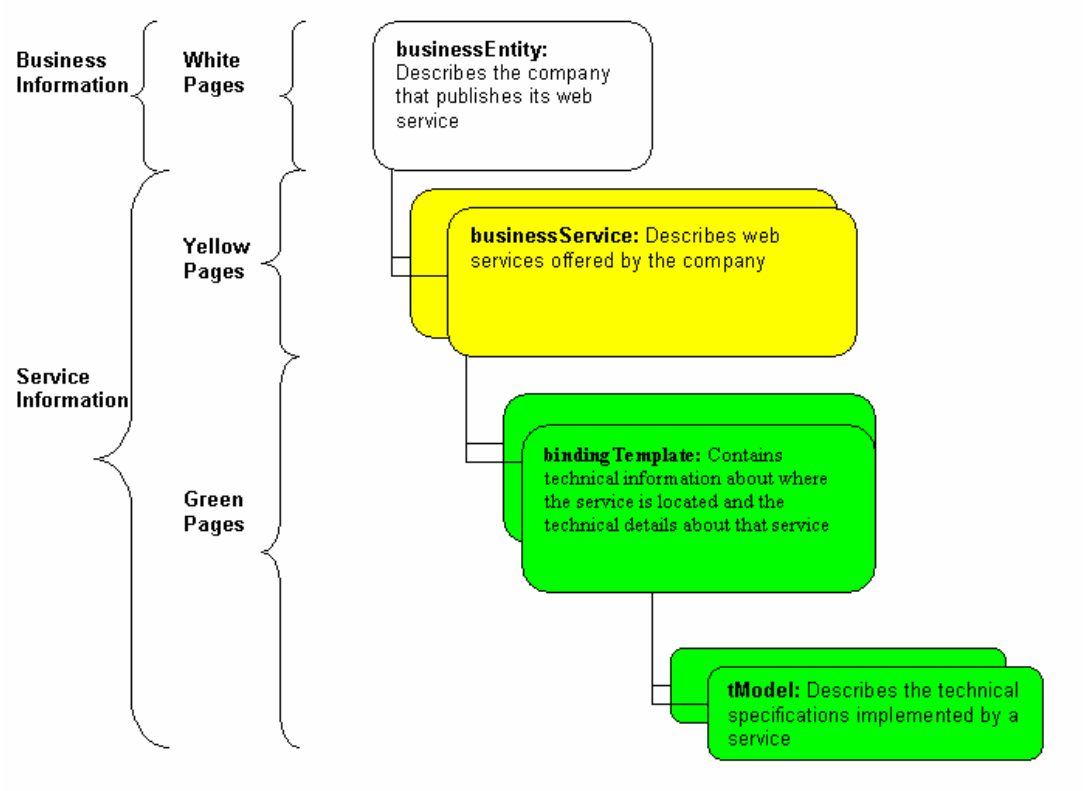
- ค้นหาโดยธุรกิจ เป็นการค้นหาสิ่งที่มีอยู่จริงของธุรกิจ (BusinesEntity) ที่มีลักษณะตามต้องการโดยมิวิธีการค้นหาได้ 3 แบบ คือ
  - การค้นหาโดยใช้ชื่อธุรกิจ (Business Name) ทำได้โดยการใช้ชื่อของธุรกิจในการค้นหา
  - การค้นหาโดยใช้ประเภทธุรกิจ (Business Category) ทำได้โดยการระบุรหัสประเภทธุรกิจตามมาตรฐาน ดังต่อไปนี้ คือ
    - Ntis-gov:naics:1997 (North American Industry Classification System)
    - Unspsc-org:unspsc (Universal Standard Products and Services Classification)
    - Ntis-gov:sic:1987 (Standard Industrial Classification)

- Uddi-org:iso-ch:3166-1999 (Codes for Geographic Location)

- การค้นหาโดยใช้ตัวระบุธุรกิจ (Business Identifier) โดยใช้รหัสดังต่อไปนี้ คือ
  - Thomasregister-com:supplierID
  - Dnb-com:D-U-N-S (Dun & Bradstreet Number Identifier System)

- ค้นหาโดยบริการ เป็นการค้นหาบิสิเนสเซอร์วิสโดยอาศัยชื่อของบริการ หรือกุญแจบริการ (Service Key) เช่น การค้นหาบริการจองตั๋ว เป็นต้น
- ค้นหาโดยทีโมเดล ทำได้โดยอาศัยชื่อของทีโมเดล หรือกุญแจทีโมเดล (tModel Key)

ส่วนข้อมูลคำอธิบายบริการในฐานข้อมูลยูดีดีไอแบ่งออกเป็น 2 ส่วนหลักคือ คำอธิบายธุรกิจที่ให้บริการ และคำอธิบายตัวบริการเอง โดยคำอธิบายธุรกิจที่ให้บริการได้แก่สิ่งที่มีอยู่จริงของธุรกิจ (BusinessEntity) หรือที่ถูกเรียกว่าเอกสารหน้าขาว (White Page) และคำอธิบายบริการ ซึ่งภายในคำอธิบายข้อมูลบริการเองก็ถูกแบ่งออกเป็น 2 ส่วน คือข้อมูลทั่วไปของตัวบริการ ได้แก่บิสิเนสเซอร์วิส (Business Service) หรือที่ถูกเรียกว่าเอกสารหน้าเหลือง (Yellow Page) และข้อมูลเกี่ยวกับการเรียกใช้บริการ ได้แก่ต้นแบบผูกมัด (BindingTemplate) และทีโมเดล (tModel) หรือที่ถูกเรียกรวมกันว่าเอกสารหน้าเขียว (Green Page) โดยโครงสร้างดังกล่าวของยูดีดีไอแสดงในรูปที่ 2.2



รูปที่ 2.2 โครงสร้างยูดีดีไอ

โดยสรุปเงื่อนไขการค้นหาแต่ละแบบมีทางเลือกในการค้นหาดังตารางที่ 1

ตารางที่ 1 แสดงรูปแบบการค้นหาของยูดีดีไอ

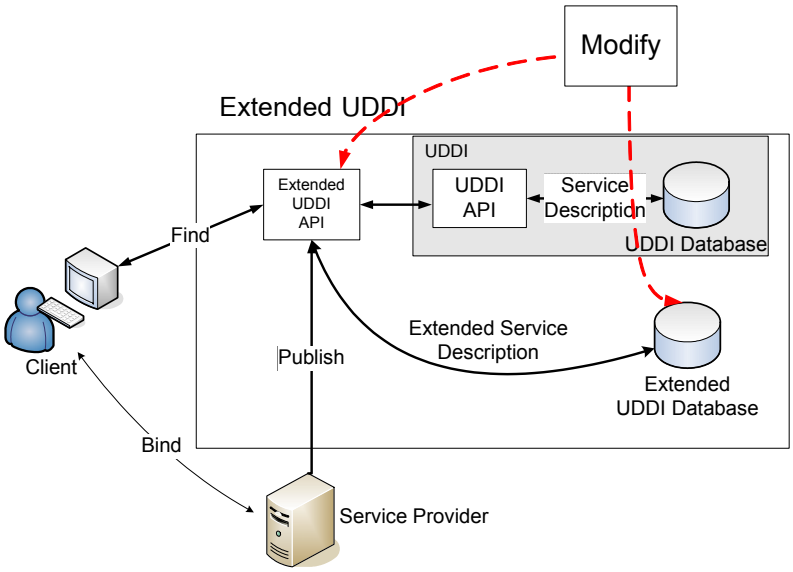
Service	Business	TModel
Service Name	Business Name	TModel Name
Category Bag	DiscoveryURL	Identifier Bag
TModel Bag	Identifier Bag	Category Bag
	TModel Bag	

จะเห็นได้ว่ายูดีดีไอนั้นมีข้อจำกัดในสองส่วนสำคัญที่ได้กล่าวในเบื้องต้นคือ ส่วนที่ทำการค้นหาจะระบุการสืบค้นได้จากเงื่อนไขที่จำกัด และส่วนเก็บข้อมูลคำอธิบายบริการสามารถอธิบายรายละเอียดของบริการได้เพียงในระดับหนึ่งซึ่งเป็นข้อมูลแบบกว้าง ๆในรูปแบบข้อความซึ่งยากแก่การวิเคราะห์ถึงความสัมพันธ์ของรายละเอียดของบริการในข้อมูลเหล่านั้น

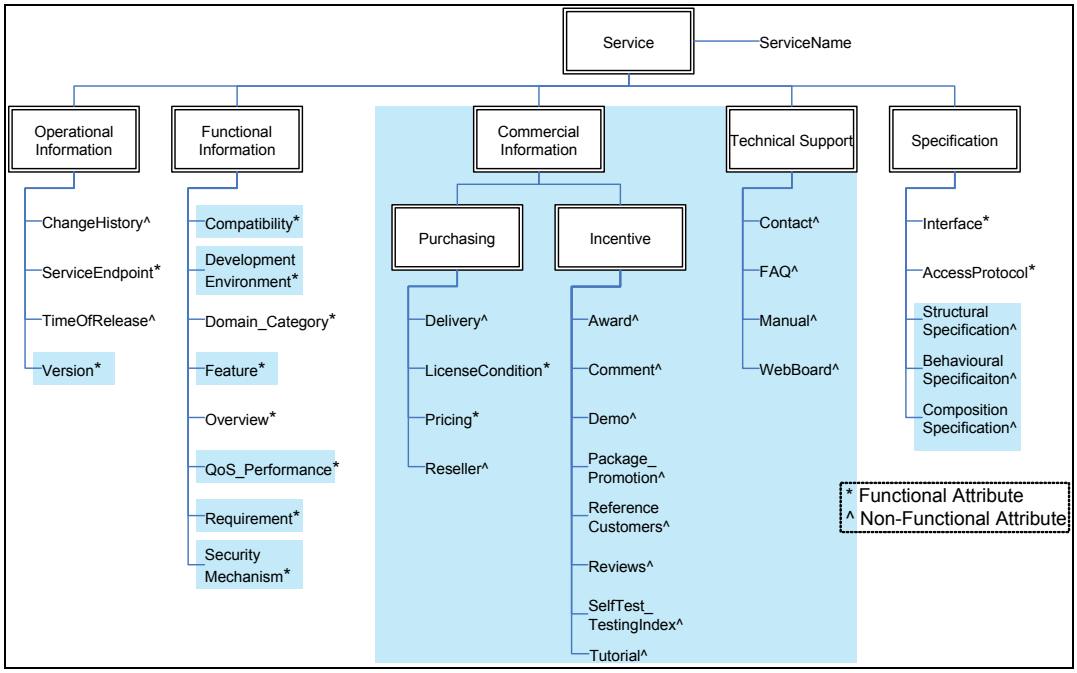
2.1.1 ตัวอย่างระบบการค้นหาบริการ

จากในบทความ [7] ได้นำเสนอต้นแบบของระบบจดทะเบียนยูดีดีไอ ซึ่งเป็นระบบจัดการเพื่อลงทะเบียนและค้นหารายละเอียดเกี่ยวกับเว็บเซอร์วิสต่าง ๆนั้น เป็นรูปแบบของยูดีดีไอมาตรฐาน

ข้อมูลคุณลักษณะของบริการที่มีการโฆษณาอยู่ในระบบยูดีดีไอทั่วไปนั้นมีน้อยเกินไป บทความ [8] จึงได้ปรับปรุงในส่วนของการเก็บข้อมูลคำอธิบายบริการให้มีคุณลักษณะสำหรับคำอธิบายบริการมากขึ้น และเพิ่มเงื่อนไขในการค้นหาจากส่วนที่ทำการค้นหาได้มากขึ้นดังในรูปที่ 2.3 โดยคุณลักษณะที่ได้เพิ่มเติมจากยูดีดีไอมาตรฐานเป็นไปดังในส่วนที่แรเงาในรูปที่ 2.4



รูปที่ 2.3 แสดงส่วนที่บทความ [8] เพิ่มเติมจากระบบยูดีดีไอปกติ



รูปที่ 2.4 การผสานคุณลักษณะเกี่ยวกับบริการ [8]

จะเห็นได้ว่าการเพิ่มคุณลักษณะเป็นเพียงการเพิ่มทางเลือกให้ผู้ใช้ในการค้นหา แต่ไม่สามารถค้นหาในเชิงความหมายได้ อีกทั้งยังไม่ช่วยเพิ่มคุณภาพของการค้นหาบริการได้ ซึ่งการวัดคุณภาพของการค้นหาบริการจะอธิบายถึงในหัวข้อถัดไป

2.1.2 การวัดคุณภาพในการค้นหาบริการ

ในการค้นหาบริการ สิ่งที่ผู้ใช้ต้องการคือบริการที่ตรงกับความต้องการซึ่งข้อมูลที่ได้รับกลับมาจากการค้นหาเป็นสิ่งที่ช่วยวัดประมาณคุณภาพของการค้นหาได้ จากงานวิจัย [10] ได้อธิบายถึงกุญแจที่ใช้ในการวัดคุณภาพการค้นคืนสารสนเทศไว้ 2 ค่า ซึ่งนำมาประยุกต์ใช้ในการวัดคุณภาพการค้นหาบริการได้แก่

Recall : ได้จากจำนวนบริการที่เกี่ยวข้องที่ได้รับกลับมาหารด้วยจำนวนของบริการที่เกี่ยวข้องทั้งหมดที่มีอยู่ ซึ่งค่าสูงสุดในการวัด recall จะเกิดขึ้นเมื่อ บริการที่เกี่ยวข้องกับคำร้องขอทั้งหมดถูกส่งกลับมาดังสมการ (1)

$$\text{Recall} = \frac{\text{No. of relevant retrieval}}{\text{Total relevant service}} \quad (1)$$

Precision : ได้จากจำนวนของบริการที่เกี่ยวข้องที่ได้รับกลับมาหารด้วยจำนวนของบริการทั้งหมดที่ได้รับกลับมา ดังนั้นค่าสูงสุดของค่า precision จะเกิดขึ้นเมื่อ บริการที่ได้รับกลับมาเป็นบริการที่

เกี่ยวข้องกับคำร้องขอทั้งหมดดังสมการ (2)

$$\text{Precision} = \frac{\text{No. of relevant retrieval}}{\text{Total retrieval service}} \quad (2)$$

ซึ่งในการค้นหาบริการเว็บเซอร์วิสบนยูดีดีไอมาตรฐานรวมไปถึงการค้นหาในแบบของบทความ [8] จะให้ค่า *Recall* และ *Precision* ที่ค่อนข้างต่ำเนื่องจากการใช้พื้นฐานของคำสำคัญทำให้ระบบไม่สามารถสืบค้นในเชิงความหมายได้ข้อมูลที่ได้กลับมาจึงไม่ตรงกับความต้องการ ดังนั้นจึงมีการนำเทคโนโลยีเว็บเชิงความหมายเข้ามาประยุกต์ใช้กับระบบเว็บเซอร์วิสซึ่งรายละเอียดของเว็บเชิงความหมายและตัวอย่างงานวิจัยจะอธิบายในหัวข้อถัดไป

2.2 เว็บเชิงความหมาย (Semantic Web)

เว็บเชิงความหมาย คือเว็บที่มีความหมายในตัวเอง[6] วัตถุประสงค์หลักของเว็บเชิงความหมาย คือการเพิ่มส่วนที่เรียกว่า “เมต้าดาต้า” (Metadata) เข้าไปเป็นส่วนหนึ่งของข้อมูลออนไลน์เพื่อเครื่องคอมพิวเตอร์จะสามารถอ่านข้อมูลเหล่านั้นและเข้าใจได้ว่าข้อมูลนั้นคืออะไรหรือเป็นส่วนไหนของชุดข้อมูลและสามารถนำข้อมูลเหล่านั้นไปประมวลผลได้อย่างอัตโนมัติ ยกตัวอย่างเช่น การเชื่อมแคตาล็อกออนไลน์เข้ากับประวัติการสั่งซื้อสินค้าของลูกค้า เพื่อให้สามารถจัดส่งสินค้าได้อย่างเหมาะสมตามช่วงเวลาที่ลูกค้าพึงพอใจมากที่สุด

เว็บจะถูกออกแบบให้เป็นเหมือนกับแหล่งข้อมูลเอกสารต่าง ๆ ด้วยเป้าหมายซึ่งไม่ได้ถูกใช้ประโยชน์แค่เพียงสำหรับการติดต่อระหว่างมนุษย์ – มนุษย์เท่านั้น แต่เครื่อง (machine) ควรจะสามารถมีส่วนร่วมและช่วยเหลือด้วยเหมือนกัน หนึ่งในอุปสรรคหลัก ๆของเรื่องนี้ก็คือข้อมูลส่วนใหญ่ในเว็บถูกออกแบบมาให้สำหรับการบริโภคหรือใช้งานของมนุษย์เท่านั้น และถึงแม้ว่าข้อมูลจากฐานข้อมูล (Database) มีการออกแบบที่ดี (ชื่อ column บอกถึง ชนิด และความหมายของข้อมูล) แต่โครงสร้างของข้อมูลเหล่านั้นก็ยังไม่พอสำหรับการใช้เครื่องในการค้นหาเว็บ เพื่อหลีกเลี่ยงปัญหาของปัญญาประดิษฐ์ (Artificial Intelligent : AI) ในการฝึกฝนให้เครื่องเลียนแบบพฤติกรรมของมนุษย์ เว็บเชิงความหมายจึงเสนอการพัฒนาภาษาสำหรับข้อมูลที่มีความละเอียดพอสำหรับรูปแบบการประมวลผลของเครื่องโดยอาร์ดีเอฟ ช่วยให้สามารถจับคู่ข้อมูล (map data) ในรูปแบบใหม่ ในการใช้งานจริง ๆของเว็บเชิงความหมายนั้นอยู่ภายใต้การพัฒนาของภาษามาร์คอัพสำหรับปัญญาประดิษฐ์ แบบใหม่ ได้แก่ อาร์ดีเอฟ อาร์ดีเอฟเอส และโอดับเบิ้ลยูแอล (Web Ontology language : OWL) [16][17] ภาษาเหล่านี้กำหนดความหมายของคำไว้เป็นอย่างดี

2.2.1 อาร์ดีเอฟ (Resource Description Framework : RDF)

อาร์ดีเอฟ คือการกำหนดกรอบระบบเพื่ออธิบายทรัพยากร (resource) หรือข้อมูล เพื่อใช้แทนที่ทำให้ข้อมูลนั้นมีคำอธิบายความหมายที่สามารถเข้าใจได้โดยเครื่อง (Machine) ซึ่งประกอบด้วย 3 ส่วนคือ Resource , Property และ ค่าที่มาอธิบาย Property ข้อมูลที่นำมาแทนนั้นจะถูกเรียกว่า Metadata เช่น ชื่อเว็บเพจ วันที่ปรับปรุงเว็บเพจ ผู้เขียนเว็บเพจ เป็นต้น ข้อมูลเหล่านี้เรียกว่า “resource” โดย resource จะบรรจุการอธิบายความสัมพันธ์ของ resource เรียกว่า “property” และ ส่วนการอธิบายคุณลักษณะของ resource ด้วยชื่อของ property บวกกับค่าของ property ของ resource นั้นเรียกว่า “literal” ซึ่งค่าของ property จะเป็นข้อความตัวอักษรใด ๆ หรือเป็น resource อื่น ๆ ก็ได้ [13] พิจารณาตัวอย่าง

“Ryan is the customer of the resource http://www.amazon.com.”

สามารถแยกออกเป็นประเภทของส่วนประกอบได้ดังนี้

Subject (Resource) http://www.amazon.com
Predicate (Property) customer
Object (literal) ″Ryan″

จากโครงสร้างด้านบนสามารถเขียนออกเป็นอาร์ดีเอฟได้ดังนี้

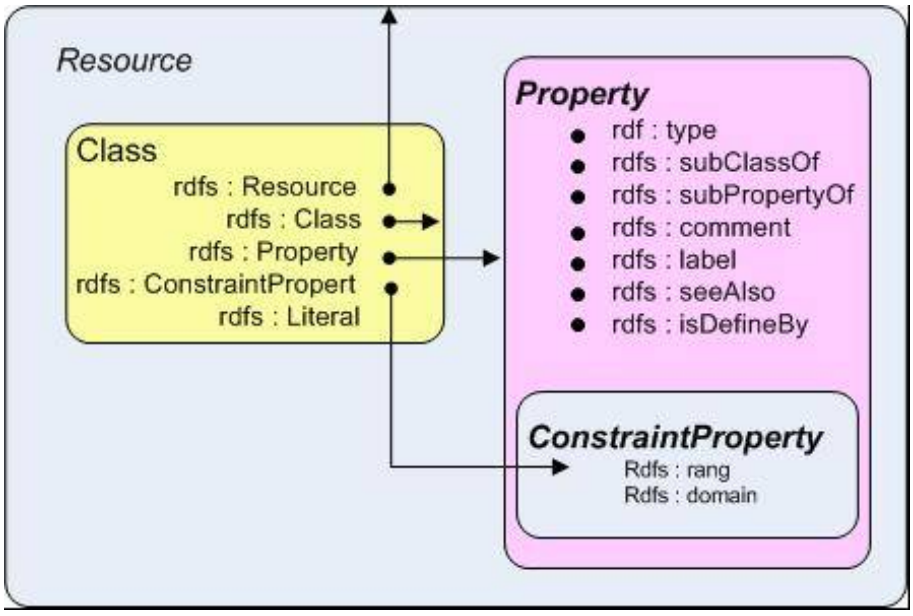
```
<?xml version="1.0"?>
<rdf:RDFxmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#
xmlns:s="http://description.org/schema/">
<rdf:Description about="http://www.amazon.com/">
<s:customer>Ryan</s: customer >
</rdf:Description>
</rdf:RDF>
```

รูปที่ 2.5 ตัวอย่างข้อมูลอาร์ดีเอฟ

จากตัวอย่างจะเห็นว่าอาร์ดีเอฟก็คือเอ็กซ์เอ็มแอลที่มีสกีมาเฉพาะเพื่อใช้อธิบาย resource หนึ่ง ๆ โดยเราสามารถประยุกต์ใช้อาร์ดีเอฟเพื่ออธิบาย resource ตามที่ต้องการได้ด้วยการปรับปรุงข้อมูลในสกีมา

2.2.2 อาร์ดีเอฟเอส (RDF Schema : RDF-S)

ในการบรรยายลักษณะเชิงความหมายตามมาตรฐานของอาร์ดีเอฟ นั้นยังไม่เพียงพอกับการกำหนดกลุ่มข้อมูลที่ต้องการความหมายและความสัมพันธ์ระหว่างกันเนื่องจากยังมีข้อจำกัด จึงต้องมีการสร้าง มาตรฐานที่สูงกว่าขึ้นมารองรับ นั้นก็คือ อาร์ดีเอฟเอส เพื่อใช้ในการบรรยายลักษณะของออนโทโลจี เพิ่มเติมจากอาร์ดีเอฟ โดยอาร์ดีเอฟเอสจัดเตรียมโครงสร้างสำหรับการบรรยายลักษณะของข้อมูลในรูปแบบคลาส (class) และ คุณสมบัติ (property) ซึ่งคลาสที่บรรยายนั้นจะเหมือนกับคลาส ในการเขียนโปรแกรมแบบการเขียนโปรแกรมเชิงวัตถุ (Object-Oriented Programming : OOP) จากคุณสมบัติดังกล่าวจะทำให้อาร์ดีเอฟเอสมีความสามารถในการบรรยาย resource เป็นคลาสได้ [14]



รูปที่ 2.6 ส่วนประกอบของอาร์ดีเอฟเอส [14]

รูปที่ 2.6 แสดงส่วนประกอบของอาร์ดีเอฟเอสนั้นจะประกอบด้วยส่วนที่ใช้ในการนิยามคลาส (Class), ส่วนที่ใช้ในการนิยามคุณสมบัติ (property) และส่วนที่เป็นค่า (value)

ตัวอย่าง การบรรยายลักษณะข้อมูลของมนุษย์ โดยใช้ไวยากรณ์ของอาร์ดีเอฟเอสจะแสดงดังใน รูปที่ 2.7

```
<?xml version="1.0"?>
    <rdf:RDF xmlns:rdf= "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"xml:base=  "http://www.human.fake/humans">
            <rdfs:Class rdf:ID="human" />
            <rdfs:Class rdf:ID="man">
            <rdfs:subClassOf rdf:resource="#human"/>
            </rdfs:Class>
            <rdfs:Class rdf:ID="woman">
            <rdfs:subClassOf rdf:resource="#human"/>
            </rdfs:Class>
    </rdf:RDF>
```

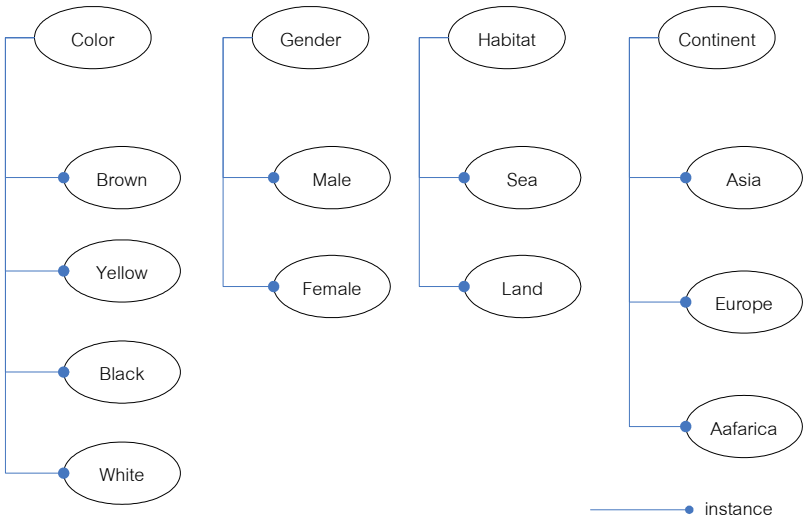
รูปที่ 2.7 ตัวอย่างข้อมูลอาร์ดีเอฟเอส [14]

2.2.3 โอดับเบิ้ลยูแอล (Web Ontology Language : OWL)

โอดับเบิ้ลยูแอลถูกพัฒนาขึ้นโดยดับเบิ้ลยูทรีซี (W3C) เพื่อให้เป็นภาษาที่สามารถอธิบายออนโทโลจีได้อย่างครอบคลุม เนื่องจากมีพื้นฐานมาจากอาร์ดีเอฟ โอดับเบิ้ลยูแอลจึงมีความสามารถของอาร์ดีเอฟและถูกเพิ่มเติมจำนวนคำศัพท์ที่แสดงถึงรูปแบบของความสัมพันธ์ระหว่างสิ่งที่ถูกอธิบาย ทำให้ภาษามีความกระชับและใช้ได้สะดวกกว่าอาร์ดีเอฟ ตัวอย่างของคำศัพท์ที่เพิ่มขึ้นจะอธิบายออนโทโลจีในลักษณะของคลาส (Class) การไม่มีส่วนร่วมระหว่างคลาส (Disjointness) การเท่ากันระหว่างคลาส (Equality) และการสมมาตรของคุณสมบัติ (Symmetry) เป็นต้น [16][17]

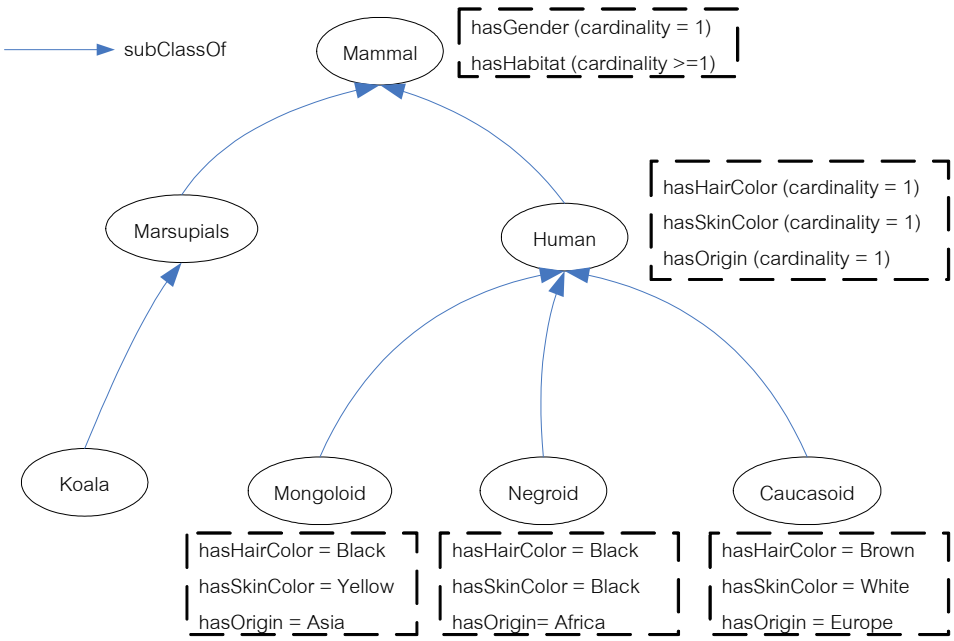
เนื่องจากโอดับเบิ้ลยูแอลมีพื้นฐานมาจากอาร์ดีเอฟ ดังนั้นรูปแบบการอธิบายทรัพยากรและลักษณะการเขียนจึงไม่แตกต่างกัน แต่ความสามารถหนึ่งที่โอดับเบิ้ลยูแอลมีเพิ่มขึ้นจากอาร์ดีเอฟคือ การที่โอดับเบิ้ลยูแอลสามารถอธิบายทรัพยากรในรูปแบบของคลาสที่มีคุณสมบัติต่าง ๆ การสืบทอดคุณสมบัติไปยังคลาสลูกและอินสแตนซ์ของคลาสเหล่านั้น ซึ่งในอาร์ดีเอฟจะอธิบายทรัพยากรได้ในรูปของอินสแตนซ์เท่านั้น ไม่สามารถอธิบายทรัพยากรในรูปแบบคลาสได้เอง

การอธิบายทรัพยากรโดยใช้โอดับเบิ้ลยูแอลจะเริ่มจากการสร้างคลาสของทรัพยากรที่ต้องการอธิบายขึ้นก่อน จึงกำหนดคุณสมบัติต่าง ๆ ของคลาสนั้น รวมถึงความสัมพันธ์ระหว่างคลาสนั้นกับคลาสอื่น ๆ ที่เกี่ยวข้องแล้วจึงสร้างอินสแตนซ์ของคลาสนั้น ตัวอย่างเช่น คลาสสี (Color) มีอินสแตนซ์ 4 ตัวคือ สีดำ (Black) สีขาว (White) สีเหลือง (Yellow) และสีน้ำตาล (Brown) คลาสเพศ (Gender) มีอินสแตนซ์ 2 ตัวคือ ชาย (Male) และหญิง (Female) คลาสที่อยู่ (Habitat) มีอินสแตนซ์ 2 ตัวคือ แผ่นดิน (Land) และทะเล (Sea) คลาสทวีป (Continent) มีอินสแตนซ์ 3 ตัวคือ เอเชีย (Asia) ยุโรป (Europe) และ แอฟริกา (Africa) ดังแสดงในรูปที่ 2.8



รูปที่ 2.8 คลาสและอินสแตนซ์ในโอดับเบิ้ลยูแอล

คลาสในรูปที่ 2.8 สามารถนำมาใช้เป็นพื้นฐานในการอธิบายคลาสที่มีความซับซ้อนยิ่งขึ้น เช่น คลาสสัตว์เลี้ยงลูกด้วยนม (Mammal) ซึ่งมีข้อกำหนดว่าสัตว์เลี้ยงลูกด้วยนมแต่ละตัวต้องมีเพศเดียว (hasGender (cardinality = 1)) และต้องมีที่อยู่อย่างน้อย 1 ค่า (hasHabitat (cardinality >=1)) โดยคลาสสัตว์เลี้ยงลูกด้วยนมมีซับคลาสเป็นมนุษย์ (Human) และสัตว์ที่มีกระเป๋าหน้าท้อง (Marsupials) ซึ่งมนุษย์มีคุณสมบัติที่สืบทอดมาจากคลาสสัตว์เลี้ยงลูกด้วยนม คือเพศและที่อยู่ นอกจากนี้ยังมีคุณสมบัติที่เพิ่มเติมขึ้นคือต้องมีสีผม (hasHairColor (cardinality = 1)) สีผิว (hasSkinColor (cardinality = 1)) และมีถิ่นกำเนิด (hasOrigin (cardinality = 1)) คลาสมองโกลอยด์ (Mongoloid) เป็นซับคลาสหนึ่งของคลาสมนุษย์ โดยมีผมสีดำ มีผิวเหลืองและมีถิ่นกำเนิดอยู่ในทวีปเอเชีย ดังแสดงในรูปที่ 2.9



รูปที่ 2.9 คลาสมนุษย์และซับคลาส

คลาสมองโกลอยด์สามารถเขียนในรูปโอดับเบิ้ลยูแอลได้ดังรูปที่ 2.10

```
<owl:Class rdf:ID=" Mongoloid ">
   <owl:Class>
    <owl:intersectionOf rdf:parseType="Collection">
     <owl:Restriction>
      <owl:onProperty>  <owl:ObjectProperty rdf:about="#hasHairColor"/>     </owl:onProperty>
      <owl:hasValue rdf:resource="#Black"/>
     </owl:Restriction>
     <owl:Restriction>
      <owl:hasValue>  <Continent rdf:ID="Asia"/>    </owl:hasValue>
      <owl:onProperty>  <owl:ObjectProperty rdf:about="#lhasOrigin"/>   </owl:onProperty>
     </owl:Restriction>
     <owl:Restriction>
      <owl:hasValue>   <Color rdf:ID="Yellow"/>    </owl:hasValue>
      <owl:onProperty>   <owl:ObjectProperty rdf:about="#hasSkinColor"/>    </owl:onProperty>
     </owl:Restriction>
    </owl:intersectionOf>
   </owl:Class>
  <rdfs:subClassOf rdf:resource="#Human"/>
 </owl:Class>
```

รูปที่ 2.10 คลาสมองโกลอยด์ในรูปแบบโอดับเบิ้ลยูแอล

โอดับเบิ้ลยูแอลจึงเป็นภาษาที่เหมาะสมที่ถูกเลือกมาใช้ในการนิยามออนโทโลจีใน

งานวิจัยนี้ซึ่งในหัวข้อต่อไปจะอธิบายถึงความหมายและและรายละเอียดของออนโทโลจีอย่างย่อแบบพอเข้าใจ

2.2.4 ออนโทโลจี (Ontology)

ออนโทโลจีเป็นแนวคิดที่ปรากฏอยู่ในหลายสาขาวิชา ซึ่งแต่ละสาขาวิชาก็มีความหมายของออนโทโลจีที่แตกต่างกันออกไป ในสาขาทางคอมพิวเตอร์เองก็มีผู้ให้คำจำกัดความของออนโทโลจีไว้หลากหลาย แต่คำจำกัดความหนึ่งที่ได้รับการยอมรับกันอย่างแพร่หลายคือ “ข้อกำหนดเกี่ยวกับแนวคิด” (“The specification of a conceptualization”) [20] โดยแนวความคิดของออนโทโลจีคือการพยายามที่จะบรรยายแนวคิดของโดเมนหรือขอบเขตความสนใจใด ๆ ในรูปของสิ่งต่าง ๆ ที่อยู่ภายในโดเมนและความสัมพันธ์ระหว่างสิ่งเหล่านั้น ซึ่งสามารถแสดงออกมาในรูปของระบบสัญลักษณ์ (Notation) ยกตัวอย่างเช่น คลาส (Class) อินสแตนซ์ (Instance) ความสัมพันธ์ (Relationship) คุณสมบัติ (Property) และ กฎ (Rule) โดยใช้ภาษาสำหรับแสดงความรู้ (Knowledge Representation Language) ซึ่งมีความชัดเจนและเที่ยงตรงมากกว่าการอธิบายโดยใช้ภาษาธรรมชาติ (Natural Language) ที่ใช้คำศัพท์มาเชื่อมต่อกันเป็นประโยคเพื่อบรรยายถึงสิ่งของในแง่มุมต่าง ๆ ทั้งนี้การใช้ระบบสัญลักษณ์จะช่วยสื่อความหมาย (Semantics) ให้ซอฟต์แวร์และเครื่องมือเข้าใจและสามารถนำไปใช้ประโยชน์ได้ โดยเฉพาะการค้นหาสิ่งที่ต้องการจากอินเทอร์เน็ตซึ่งเปรียบเสมือนคลังข้อมูลขนาดใหญ่ หากเราสามารถค้นหาสิ่งของได้จากความหมาย ไม่ใช่การค้นหาจากคำเหมือน (Name Matching) ก็จะช่วยให้เราได้ของที่ตรงตามความต้องการและสะดวกรวดเร็วยิ่งขึ้น ซึ่งตรงกับแนวความคิดของเว็บเชิงความหมาย (Semantic Web) [18]

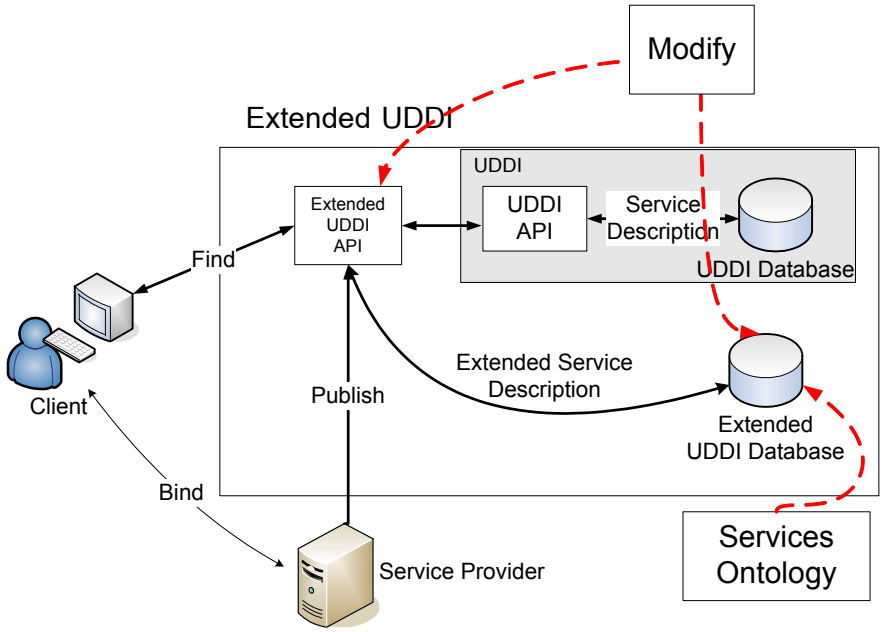
ซึ่งความหมายของออนโทโลจีในวิทยานิพนธ์นี้จะใช้แทนการนิยามข้อมูลที่มีความสัมพันธ์กันทั้งในรูปแบบของการสืบทอดลำดับชั้นและในรูปแบบของคุณสมบัติที่มีการเชื่อมต่อกันระหว่างข้อมูลในเชิงความหมาย เช่น การนิยามกลุ่มข้อมูลเกี่ยวกับการท่องเที่ยวมีการเชื่อมโยงความสัมพันธ์กันระหว่างสถานที่, ระยะทาง, หรือกิจกรรมต่าง ๆ กลุ่มข้อมูลส่วนนี้ก็จะถูกเรียกว่าออนโทโลจีการท่องเที่ยว เป็นต้น

เพื่อให้ออนโทโลจีสามารถอธิบายถึงสิ่งของต่าง ๆได้ จำเป็นจะต้องอาศัยภาษาที่มีความสามารถในการอธิบายถึงความหมายและความสัมพันธ์ระหว่างสิ่งของต่าง ๆ จึงมีการคิดค้นภาษาเพื่อรองรับความต้องการดังกล่าว โดยส่วนใหญ่เป็นภาษาที่มีพื้นฐานมาจากภาษาเอ็กซ์เอ็มแอล เช่น อาร์ดีเอฟ, โอดับเบิ้ลยูแอล, โอดับเบิ้ลยูแอล-เอส (Web Ontology Language for Services : OWL-S) เป็นต้น โดยอาร์ดีเอฟถือเป็นภาษาหนึ่งที่ได้รับความนิยมแพร่หลาย เพราะเป็นภาษาที่เข้าใจได้ง่ายและสามารถนำไปพัฒนาต่อให้มีความสามารถสูงขึ้นได้ วิทยานิพนธ์ฉบับนี้จะทำงานอยู่บนพื้นฐานของโอดับเบิ้ลยูแอลและอาร์ดีเอฟเป็นหลัก [18][19][20]

2.3 การใช้เทคโนโลยีเว็บเชิงความหมายเพื่อประยุกต์ใช้ในการค้นหาบริการเชิงความหมาย

การนำเทคโนโลยีเว็บเชิงความหมายเข้ามานิยามรายละเอียดของบริการในรูปแบบออนโทโลจีช่วยให้สามารถค้นหาบริการในเชิงความหมายได้ ซึ่งส่วนที่ทำการปรับปรุงเพื่อให้รองรับการค้นหาเชิงความหมายคือ ส่วนที่เก็บรายละเอียดของบริการและส่วนที่ทำหน้าที่ค้นหา ซึ่งจากในบทความ [9] ได้พัฒนาต่อมาจากในบทความ [8] โดยการเพิ่มการนิยามข้อมูลบริการเป็นออนโทโลจีด้วยภาษาโอดับเบิ้ลยูแอล เข้าไปในส่วนขยายฐานข้อมูลยูดีดีไอในรูปที่ 2.3 ซึ่งจะเป็นดังแสดงในรูปที่ 2.11

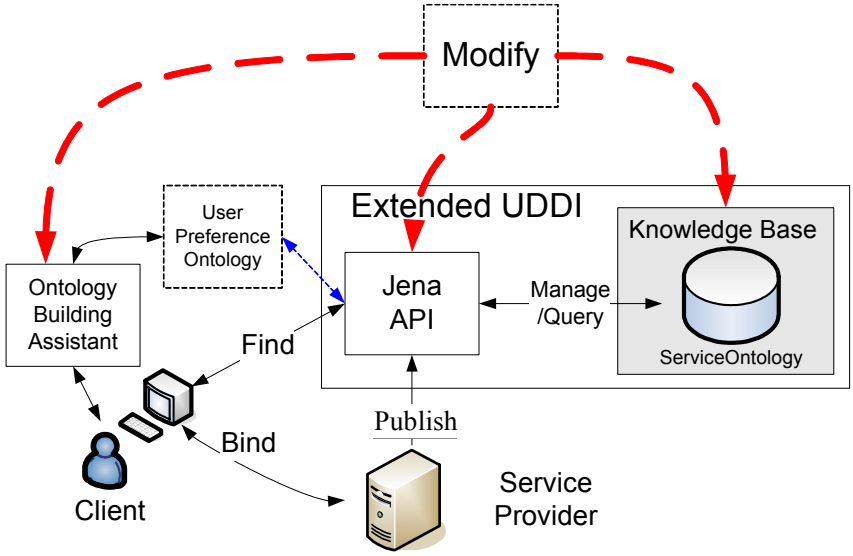
แต่ด้วยขั้นตอนการค้นหา จะทำการค้นหาจากการเปรียบเทียบคำสำคัญกับข้อมูลในยูดีดีไอมาตรฐานก่อนจากนั้นจึงนำบริการที่ได้มาเปรียบเทียบกับคุณลักษณะเชิงข้อกำหนดที่อยู่ในรูปแบบออนโทโลจี ซึ่งจากขั้นตอนของงานวิจัยนี้ทำให้เกิดจุดด้อยคือ ผลลัพธ์ที่ได้จากการค้นหาในขั้นตอนแรกอาจจะให้ผลลัพธ์ที่ผิดพลาดเนื่องจากเป็นค้นหาที่ไม่ใช่ในเชิงความหมาย ดังนั้นแม้ในขั้นตอนต่อไปจะทำการค้นหาด้วยออนโทโลจี แต่ก็เป็นการค้นจากบริการที่ได้จากผลลัพธ์ที่ผิดพลาดมาก่อนแล้ว ดังนั้นผลที่ได้จึงอาจยังมีข้อผิดพลาด



รูปที่ 2.11 การนิยามบริการเป็นออนโทโลจีในบทความ [9]

จากข้อด้อยดังกล่าว ในวิทยานิพนธ์นี้จึงได้นำเสนอการค้นหาบริการให้ได้คุณภาพการค้นหาที่ดีกว่าโดยการปรับปรุงเพิ่มเติมส่วนประกอบดังที่แสดงในรูปที่ 2.12 โดยการนิยามข้อมูลบริการแยกตามชนิดหรือโดเมนบริการในรูปแบบออนโทโลจี นอกจากนี้ยังนำเสนอการปรับปรุงส่วนที่ทำการค้นหาให้มีประสิทธิภาพ โดยอัลกอริทึมและขั้นตอนการค้นหาที่นำเสนอจะกล่าวถึงใน

ภายหลัง



รูปที่ 2.12 การปรับปรุงส่วนเก็บข้อมูลบริการและส่วนค้นหาบริการของวิทยานิพนธ์นี้

จากรูปที่ 2.12 มีรายละเอียดของแต่ละองค์ประกอบดังนี้

Client คือผู้ใช้หรือผู้ค้นหาบริการ

Service Provider คือผู้ให้บริการ หรือผู้พัฒนาบริการให้อยู่ในรูปของเว็บเซอร์วิส

Extended UDDI คือส่วนขยายยูดีดีไอซึ่งจำลองขึ้นโดยใช้เครื่องมือที่เรียกว่าเจยูดีดีไอ (jUDDI) [26] ประกอบไปด้วย

- ส่วนที่เก็บออนโทโลจีบริการ (Knowledge Base) และ
- ส่วนติดต่อโปรแกรมประยุกต์จีนา (Jena API) เป็นส่วนที่พัฒนาขึ้นจากจีนาใช้สำหรับการติดต่อระหว่างผู้ใช้และออนโทโลจี เพื่อจัดการค้นหาและเลือกบริการสำหรับผู้ใช้

User Preference Ontology คือออนโทโลจีความชอบของผู้ใช้ซึ่งถูกเก็บอยู่ในรูปแบบโอดับเบิ้ลยูแอล

Ontology Building Assistant คือเครื่องมือช่วยผู้ใช้ในการสร้าง *User Preference Ontology*

2.3.1 ผู้ใช้ (User/Client)

ผู้ใช้ในระบบของเว็บเซอร์วิสจะเป็นได้ทั้งผู้ให้บริการ (Service Provider) รายอื่น ผู้ใช้โดยตรง (End – Users) หรือ ผู้พัฒนาเว็บเซอร์วิส (Developer) ซึ่งความต้องการในการค้นหาบริการของแต่ละกลุ่มผู้ใช้จะแตกต่างกันได้แก่

- ผู้ให้บริการต้องหาค้นหาบริการเพื่อรวบรวมเป็นแหล่งให้บริการด้านต่าง ๆ อาจจะทำหน้าที่เหมือนตัวแทน (Agency) หรือเว็บท่า (Web Portal) ในการให้บริการที่หลากหลายหรือบริการในกลุ่มเดียวกัน สิ่งที่ผู้ใช้กลุ่มนี้ต้องป้อน

ให้แก่ระบบใช้ในการสืบค้นคือข้อมูลจำพวก ชนิดของบริการ เนื่องจากต้องการเพียงแค่รวบรวมบริการให้อยู่เป็นกลุ่มบริการ เพื่อให้ผู้ใช้อื่น ๆเข้ามาหาข้อมูลหรือใช้บริการ

- ผู้ใช้บริการต้องการค้นหาบริการเพื่อใช้งานบริการนั้นจริง เช่น ค้นหาบริการเพื่อจองตั๋วเครื่องบินหรือบริการจองที่พัก ผู้ใช้กลุ่มนี้จะต้องการรายละเอียดของบริการมากกว่ากลุ่มอื่น ๆ เนื่องจากต้องการค้นหาเพื่อใช้งานจริง ๆ นอกจากข้อมูลที่บอกว่าบริการการนั้นเป็นบริการชนิดไหนแล้ว ยังต้องการรู้ว่าบริการให้ผลลัพธ์เป็นอะไร เช่น ต้องการค้นหาบริการจองตั๋วเครื่องบินก็จะต้องการรู้ว่าบริการใดสามารถของตั๋วเที่ยวที่ออกจากสนามบินใดไปยังสนามบินใด เวลาที่เครื่องบินออกจากต้นทางถึงปลายทาง เป็นต้น ผู้ใช้กลุ่มนี้จะมีบทบาทต่อระบบบริการแบบธุรกิจกับลูกค้าซึ่งกำลังค่อย ๆมีบทบาทมากกว่าแบบุรกิจสู่ธุรกิจ ดังนั้นผู้ใช้กลุ่มนี้จึงจำเป็นต้องให้ข้อมูลรับเข้าที่มากขึ้นแก่ระบบ รวมถึงเงื่อนไขในการสืบค้นจะมีความซับซ้อนกว่ากลุ่มผู้ใช้กลุ่มอื่น
- ผู้พัฒนาเว็บเซอร์วิสจะต้องการค้นหาบริการเพื่อนำไปเป็นส่วนเพิ่มเติมหรือทำงานร่วมกับโปรแกรมหรือบริการของตนเอง ดังนั้นข้อมูลส่วนสำคัญที่ผู้ใช้กลุ่มนี้ต้องการคือ ข้อมูลวิธีการเข้าถึงและเรียกใช้งาน ข้อมูลรับเข้าและส่งออกของบริการ และที่อยู่ของบริการ

ในวิทยานิพนธ์นี้เน้นศึกษาสำหรับผู้ใช้โดยตรงและระบบบริการแบบธุรกิจกับลูกค้าเป็นหลัก ซึ่งจากระบบการค้นหาบริการแบบดั้งเดิมหรือระบบในงานวิจัยอื่น ๆที่ได้กล่าวไปแล้วใน บทที่ 1 มักจะมีปัญหาที่เกิดจากตัวผู้ใช้บริการเอง เช่น ใช้เงื่อนไขในการค้นหาที่กว้างเกินไป แต่ก็อาจเกิดจากการที่ผู้ใช้บริการไม่ทราบว่าผู้ให้บริการได้ระบุข้อมูลใดบ้างในโฆษณา ทำให้ต้องกำหนดเงื่อนไขกว้าง ๆ ไว้ก่อน ทำให้ได้บริการอื่นที่ไม่ต้องการมาด้วย ทำให้ผู้ใช้ต้องใช้เวลาในการคัดเลือกบริการมากตามไปด้วย หรือการระบุเงื่อนไขเจาะจงและละเอียดเกินไปทำให้ไม่ได้ผลลัพธ์กลับมาเลย เป็นต้น

2.3.2 ผู้ให้บริการ (Service Provider)

ผู้ให้บริการในระบบนี้มีหน้าที่นิยามและประกาศบริการของตนเองซึ่งมี 2 ส่วนที่ผู้ให้บริการต้องเป็นคนกำหนดได้แก่ นิยามบริการด้วยออนโทโลจีโดยอ้างอิงจากโครงสร้างของโอดับเบิ้ลยูแอลเอส และ กำหนดรายละเอียดของบริการตามโครงสร้างของโดเมนออนโทโลจีซึ่งในวิทยานิพนธ์นี้ได้ออกแบบเพิ่มเติมจากโครงสร้างออนโทโลจีการท่องเที่ยวของกลุ่มพันธมิตรการท่องเที่ยวแบบเปิด

2.3.3 ส่วนขยายยูดีดีไอ (Extended UDDI)

ส่วนขยายยูดีดีไอพัฒนาเพิ่มเติมจากเจยูดีดีไอซึ่งเครื่องมือนี้เป็นต้นแบบของระบบยูดีดี

ไอมาตรฐาน ซึ่งในวิทยานิพนธ์นี้จะมีการเปลี่ยนแปลงและเพิ่มเติมบางส่วนประกอบเข้าไปดังนี้

- ส่วนที่เก็บข้อมูลของเว็บเซอร์วิส เมื่อพิจารณาจาก รูปที่ 2.1 ซึ่งเป็นยูดีดีไอมาตรฐาน กับรูปที่ 2.12 ซึ่งเป็นส่วนขยายยูดีดีไอ จะเห็นได้ว่ามีการเปลี่ยนจากการเก็บข้อมูลบริการในรูปแบบฐานข้อมูลทั่วไปเป็นฐานข้อมูลองค์ความรู้ในรูปแบบออนโทโลจีบริการ
- ส่วนติดต่อโปรแกรมประยุกต์จีนา เป็นส่วนที่พัฒนาขึ้นจากจีนา ใช้สำหรับเป็นส่วนติดต่อระหว่างผู้ใช้และออนโทโลจีในฐานข้อมูลองค์ความรู้ รวมถึงเพื่อจัดการค้นหาและเลือกบริการสำหรับผู้ใช้

2.3.4 ออนโทโลจีความชอบของผู้ใช้ (User Preference Ontology)

ในการสืบค้นบริการโดยทั่วไปจะได้ผลที่ออกมามากมายจนผู้ใช้ต้องเสียเวลาในการเลือก ในขณะเดียวกัน การระบุส่วนของคำสืบค้นมากเกินไปก็อาจจะทำให้ไม่ได้ผลลัพธ์กลับมาเลย ในการสืบค้นบริการโดยความชอบของผู้ใช้ เป็นการเพิ่มเติมส่วนพิเศษเข้าไปช่วยเพิ่มคุณภาพในการเลือกบริการที่ต้องการ ซึ่งจะถูกนำมาใช้เมื่อผู้ใช้เข้าสู่ระบบ ความชอบจะถูกนิยามเป็นเหมือนชุดของความต้องการของแต่ละบุคคลซึ่งสามารถรวบรวมได้หลายรูปแบบ เช่น ได้จากการเก็บข้อมูลแบบระยะยาว ได้รับโดยตรงซึ่งถูกระบุโดยผู้ใช้ในรูปแบบของเงื่อนไข และได้รับโดยตรงจากที่ผู้ใช้ให้ข้อมูลกับระบบเช่นการกรอกแบบสอบถามในขั้นตอนการลงทะเบียน เป็นต้น ในบทความ [27] เป็นตัวอย่างการนิยามความชอบของผู้ใช้ในรูปออนโทโลจีเพื่อช่วยจัดการเอกสารนำเสนอและข้อมูลการติดต่อให้โปรแกรมตัวแทน (Agent) สามารถนำไปพิจารณาถึงสถานะปัจจุบันของผู้ใช้ได้

วิทยานิพนธ์นี้ได้เสนอแนวทางการสอบถามผู้ใช้ถึงข้อมูลความชอบเกี่ยวกับบริการในการท่องเที่ยวเพื่อนำข้อมูลมาสร้างเป็นออนโทโลจีของความชอบผู้ใช้สำหรับช่วยระบุถึงสถานการณ์และเงื่อนไขเพื่อให้ระบบใช้เป็นข้อมูลเสริมในการตัดสินใจในการเลือกบริการ ซึ่งจะช่วยในการลดปัญหาการสืบค้นที่เกิดจากตัวผู้ใช้ที่ระบุรายละเอียดการสืบค้นได้ไม่ดีพอ ซึ่งข้อมูลความชอบของผู้ใช้นี้จะออกแบบโดยใช้ข้อมูลจากโครงสร้างความชอบลูกค้าที่กลุ่มพันธมิตรการท่องเที่ยวแบบเปิดได้นิยามไว้มาวิเคราะห์

2.3.5 Ontology Building Assistant (OBA)

ในวิทยานิพนธ์นี้ขอแทนด้วยคำย่อ “โอบีเอ : OBA” คือเครื่องมือช่วยผู้ใช้สร้าง User Preference Ontology เพื่อให้เกิดความสะดวกแก่ผู้ใช้ที่ไม่ได้มีความรู้ด้านออนโทโลจีและภาษาโอดับเบิ้ลยูแอลให้สามารถสร้างข้อมูลความชอบของตนเองขึ้นมาได้ ซึ่งรายละเอียดของส่วนนี้จะกล่าวถึงต่อไปในบทที่ 4

2.4 การค้นหาบริการโดยข้อมูลความชอบของผู้ใช้

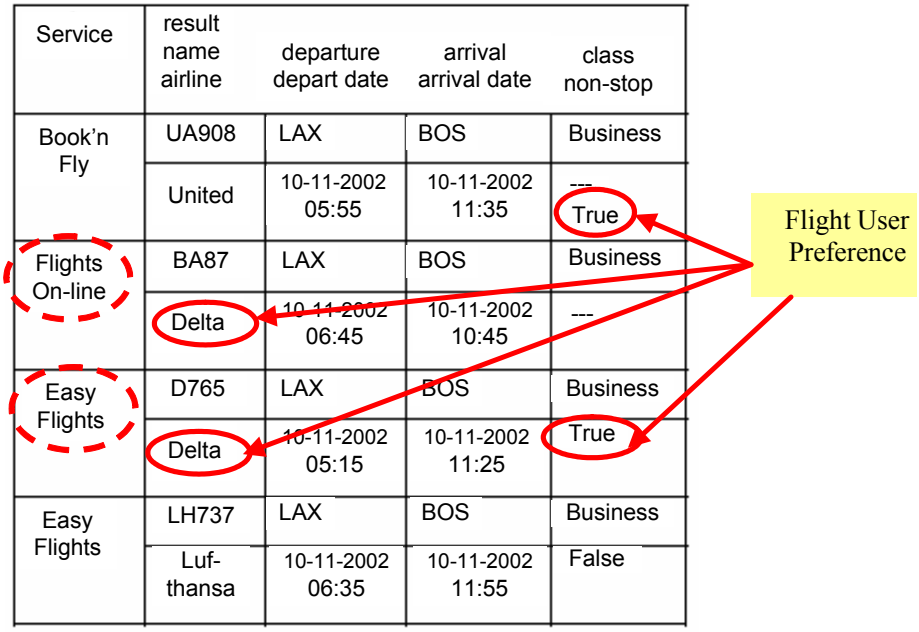
นอกจากข้อมูลรายละเอียดของแต่บริการและข้อมูลรับเข้าที่ถูกนำมาเปรียบเทียบเพื่อค้นหาบริการแล้ว ยังมีการนำข้อมูลอื่น ๆมาเสริมด้วยในกรณีที่ข้อมูลรับเข้าที่ผู้ใช้ป้อนมานั้นอาจไม่เพียงพอที่จะให้ระบบใช้เพื่อค้นหาและเลือกบริการ ในงานวิจัย [10] ได้เสนอการนิยามบริบทในรูปออนโทโลจีเพื่อกำหนดคุณลักษณะพิเศษของบริการสำหรับนำมาคัดเลือกบริการที่ผ่านขั้นตอนการสืบค้นจากข้อมูลรับเข้าและข้อมูลส่งออกของบริการที่ผู้ใช้ระบุแล้วให้ได้บริการที่ใกล้เคียงกับความต้องการของผู้ใช้มากยิ่งขึ้น

นอกจากข้อมูลแวดล้อมรอบตัวผู้ใช้และระบบแล้วบริบทยังสามารถใช้ข้อมูลส่วนอื่นรวมไปถึงข้อมูลของผู้ใช้บริการได้อีกด้วย ซึ่งงานวิจัย [11] ได้นำเสนออัลกอริทึมในการขยายคำร้องขอบริการที่ผู้ใช้ระบุและคาดหวังเพื่อสร้างทางเลือกแก่ผู้ใช้ เนื่องจากผลการสืบค้นอาจมีมากกว่า 1 บริการจึงมีการเพิ่มขั้นตอนการค้นหาขึ้น โดยใช้ข้อมูลความชอบของผู้ใช้ (User Preference) ซึ่งเป็นข้อมูลที่ผู้ใช้เคยระบุไว้กับระบบในรูปแบบข้อมูลผู้ใช้ระยะยาว (Long-Term Profile) เป็นข้อมูลรับเข้าเพิ่มเติมในคำสืบค้น หากบริการใดมีคุณสมบัติตรงกับข้อมูลรับเข้าเพิ่มเติมก็จะได้รับเลือกเนื่องจากใกล้เคียงสิ่งที่ผู้ใช้ต้องการมากกว่าบริการอื่น ดังตัวอย่างในรูปที่ 2.13 ซึ่งจากข้อมูลความชอบของผู้ใช้ระบุว่าชอบสายการบิน Delta และ class แบบ non – stop บริการ Easy Flights จึงถูกเลือก

Service	result name airline	departure depart date	arrival arrival date	class non-stop
Book'n Fly	UA908	LAX	BOS	Business
	United	10-11-2002 05:55	10-11-2002 11:35	---
Flights On-line	BA87	LAX	BOS	Business
	British Airways	10-11-2002 06:45	10-11-2002 10:45	---
Easy Flights	D765	LAX	BOS	Business
	Delta	10-11-2002 05:15	10-11-2002 11:25	True
Easy Flights	LH737	LAX	BOS	Business
	Luf-thansa	10-11-2002 06:35	10-11-2002 11:55	False

Flight User Preference

รูปที่ 2.13 การเลือกบริการที่ตรงกับความชอบของผู้ใช้ [11]

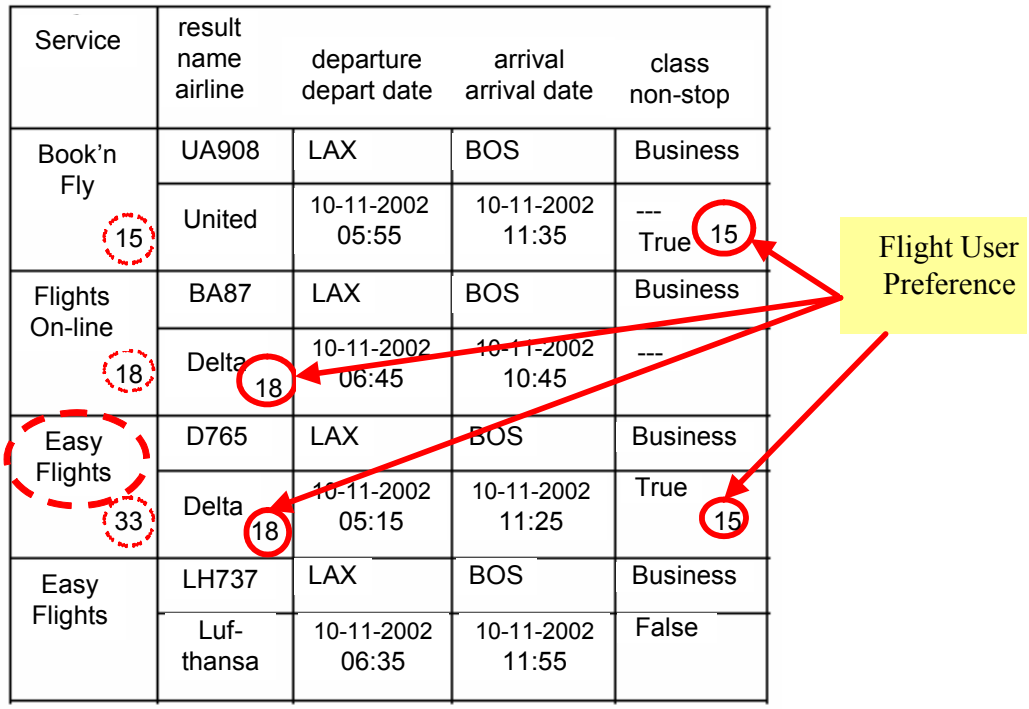


รูปที่ 2.14 กรณีที่มีหลายบริการที่มีความชอบตรงกับที่ผู้ใช้ระบุ [11]

แต่ในกรณีที่ข้อมูลความชอบของผู้ใช้มีหลายอย่างดังในรูปที่ 2.14 แต่ละบริการที่ผ่านการสืบค้นจากคำสืบค้นที่ผู้ใช้ระบุอาจมีคุณสมบัติตรงกับข้อมูลความชอบของผู้ใช้บางส่วนแตกต่างกัน ทำให้ระบบไม่อาจตัดสินใจว่าบริการที่มีข้อมูลความชอบใดใกล้เคียงความต้องการของผู้ใช้มากกว่ากัน ทุกบริการจะได้รับเลือกซึ่งอาจจะมีจำนวนมากและไม่สามารถจัดเรียงลำดับบริการที่ใกล้เคียงความต้องการของผู้ใช้มากที่สุด

ซึ่งในตัวอย่าง บริการอื่นก็มีสายการบิน Delta และ class แบบ non – stop เช่นกันซึ่งในบทความของงานวิจัย [11] นี้ไม่สามารถระบุได้ว่าควรเลือกบริการใด

วิธีในวิทยานิพนธ์นี้จึงกำหนดข้อมูลบางอย่างเพื่อให้ระบบได้ใช้เป็นข้อมูลเปรียบเทียบความสำคัญของแต่ละบริการในการเลือกบริการที่ใกล้เคียงความต้องการของผู้ใช้มากที่สุด ดังตัวอย่างในรูปที่ 2.15 ซึ่งรายละเอียดจะกลายถึงในบทที่ 4 ต่อไป

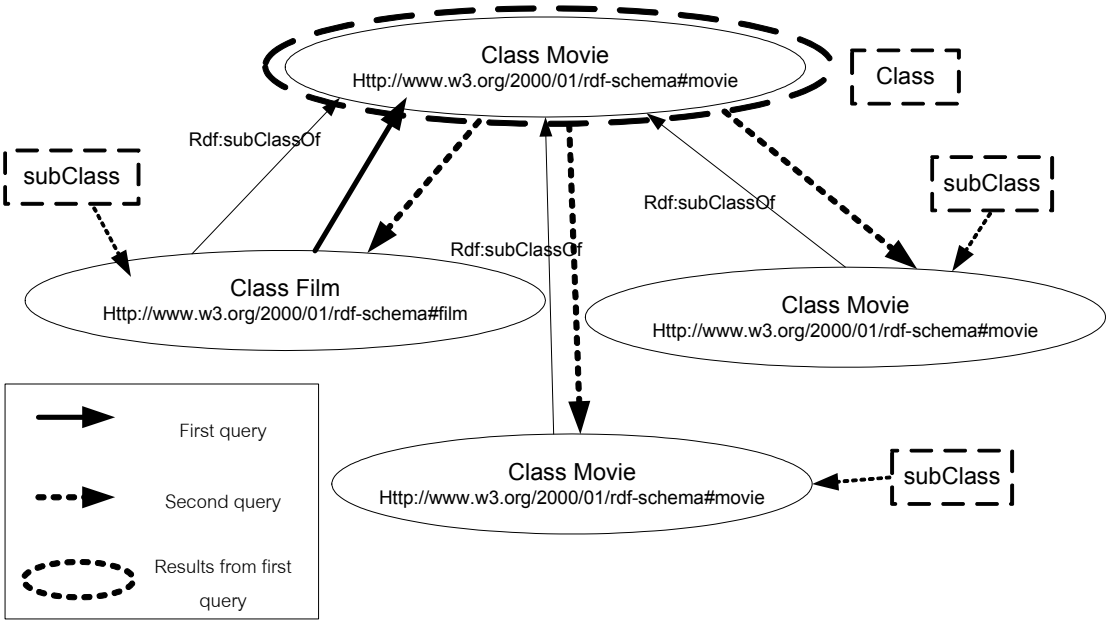


รูปที่ 2.15 การใช้คะแนนความชอบสำหรับเลือกบริการของวิทยานิพนธ์

2.5 การแก้ไขปัญหาการสืบค้นที่ไม่พบผลลัพธ์หรือได้ผลลัพธ์ที่ไม่ครอบคลุม

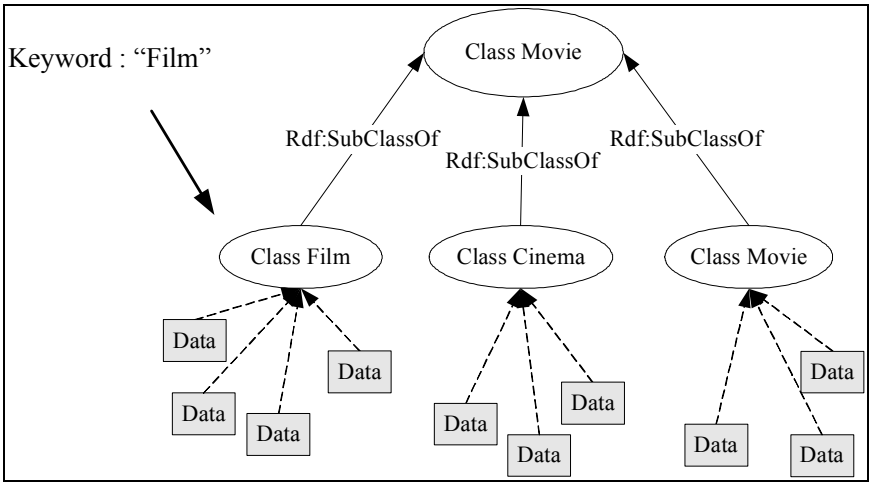
โดยปกติในการค้นหาข้อมูลแบบคำสำคัญจะพบกับปัญหาการสืบค้นที่ไม่พบผลลัพธ์หรือได้ผลลัพธ์ที่ไม่ครอบคลุมเนื่องจากข้อจำกัดของการจับคู่ระหว่างคำต่อคำ ซึ่งก็ได้มีงานวิจัยในบทความ [12] ได้ออกแบบข้อมูล เมต้าดาต้าเพื่ออธิบายข้อมูลให้มีความหมายโดยจัดเก็บข้อมูลและโครงสร้างข้อมูลด้วยอาร์ดีเอฟและอาร์ดีเอฟเพื่อให้การสืบค้นได้ผลลัพธ์ที่เป็นเอกสารที่มีความหมายของเนื้อหาตรงกับความหมายที่ผู้ใช้ต้องการ

หลักการคือสร้างกลุ่มคำที่มีความหมายเดียวกันขึ้นมาเพื่อแก้ไขปัญหาการสืบค้นที่ไม่พบผลลัพธ์หรือได้ผลลัพธ์ที่ไม่ครอบคลุมทั้งหมด เช่นตัวอย่างในรูปที่ 2.16 ได้นิยามคลาสชื่อ Movie เป็นคลาสหลักจากนั้นสร้างกลุ่มคำที่มีความหมายเหมือนกันคือคลาส Film, Cinema และ Movie เป็นคลาสย่อยของ Movie



รูปที่ 2.16 เส้นทางการสืบค้น [12]

จากรูปที่ 2.16เมื่อผู้ใช้ระบุคำสำคัญเข้ามาก็จะนำไปสืบค้นว่าตรงกับทรัพยากร (Resource) ใดบ้าง เช่น ใส่คำสืบค้นว่า “Film” ก็จะได้ตรงกับ Resource ของคลาสชื่อ Film จากนั้นทำการสืบค้นจากคลาส Film ขึ้นไปดูว่าคลาสหลักคืออะไรในที่นี้คือ Movie จากนั้นขั้นตอนต่อมาทำการสืบค้นเอาคลาสย่อยทั้งหมดของคลาสหลัก Movie เพื่อสืบค้นข้อมูลที่มาลงทะเบียนหรือมีความสัมพันธ์กับคลาสย่อยนั้น ๆ



รูปที่ 2.17 ข้อมูลที่มีความสัมพันธ์กับแต่ละคลาสย่อย

การนิยามกลุ่มคำที่เหมือนกันแล้วสืบค้นจากคลาสย่อยแล้วย้อนขึ้นไปคลาสหลัก จากนั้น

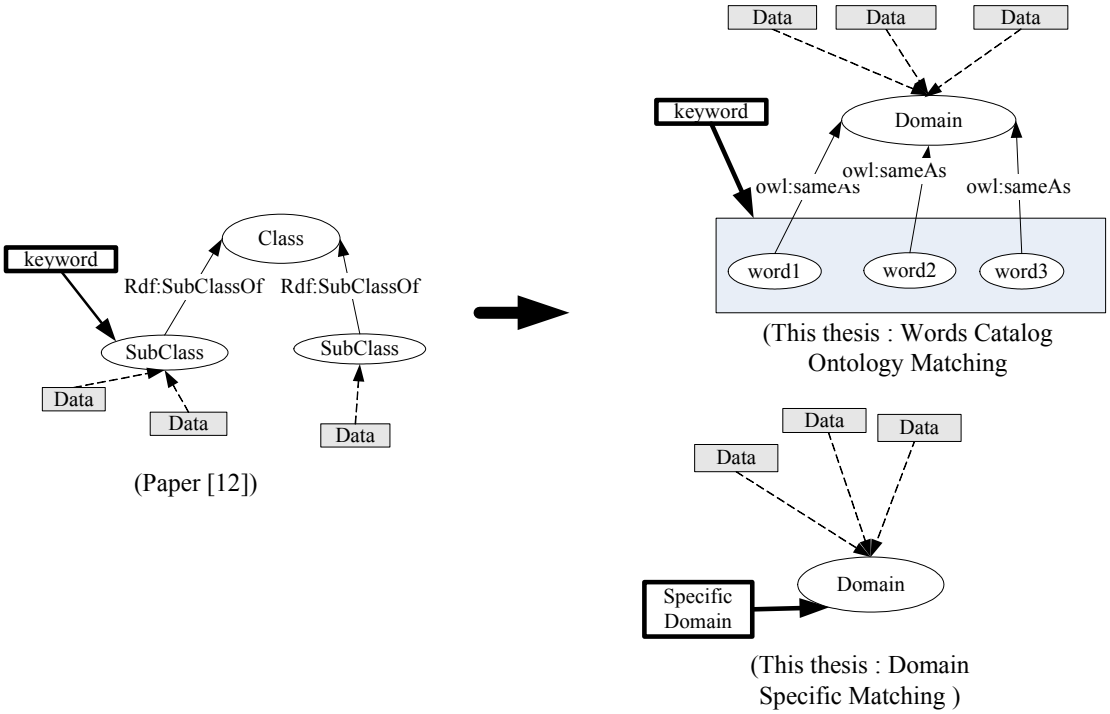
กระจายไปยังคลาสย่อยทั้งหมดดังรูปที่ 2.17 มีข้อด้อยดังนี้

- การสืบค้นมีมากขั้นตอนเนื่องจากนำคำสำคัญไปสืบค้นจากคลาสย่อย แล้วหาคลาสหลักแล้วสืบค้นกลับไปแต่ละคลาสย่อยอีกที
- ในการใช้งานจริงผู้ใช้อาจเกิดความสับสนเนื่องจากไม่รู้จะกำหนดความสัมพันธ์ของข้อมูลให้กับคลาสไหนดีเพราะทุกคลาสมีความหมายเหมือนกัน
- ข้อจำกัดการนิยามข้อมูลแบบอาร์ดีเอฟ อนุญาตให้นิยามข้อมูลที่ชื่อซ้ำกันได้ เช่น Movie ผู้ใช้หรือระบบอาจสับสนว่าอันไหนคือคลาสหลักหรือคลาสย่อย
- หากแต่ละคำมีคำที่มีความหมายเหมือนกันอยู่มากจะต้องนิยามคลาสย่อยเพิ่มขึ้นมากมาย ทำให้ข้อมูลที่มากำหนดความสัมพันธ์ไม่เป็นระเบียบและเกิดความสับสน

ซึ่งจากข้อจำกัดดังกล่าวนั้นเมื่อเปรียบเทียบกับวิธีที่เสนอในวิทยานิพนธ์นี้จะเห็นความแตกต่างดังในรูปที่ 2.18 ดังนี้

- วิทยานิพนธ์นี้ใช้การนิยามข้อมูลด้วยโอดับเบิ้ลยูแอลแทนอาร์ดีเอฟเนื่องจากคุณสมบัติที่ไม่อนุญาตให้นิยามข้อมูลที่ชื่อซ้ำกัน และคุณสมบัติ Same as ของโอดับเบิ้ลยูแอลช่วยในการนิยามข้อมูลที่เป็นเรื่องเดียวกันหรือมีความหมายเหมือนกันได้ยืดหยุ่นกว่าการใช้คุณสมบัติ Sub class ของอาร์ดีเอฟ
- ใช้การลงทะเบียนข้อมูลกับโดเมนหลักของข้อมูลนั้น ๆแทนการลงทะเบียนข้อมูลกับข้อมูลซึ่งเป็นคลาสย่อยของโดเมนนั้น

นอกจากนี้เมื่อนำวิธีทั้งสองในวิทยานิพนธ์นี้ไปประยุกต์ใช้กับการค้นหาบริการยังช่วยเพิ่มคุณภาพการค้นหาซึ่งรายละเอียดจะอธิบายในบทที่ 3 ต่อไป



รูปที่ 2.18 เปรียบเทียบการแก้ไขปัญหาการสืบค้นที่ไม่พบผลลัพธ์หรือได้ผลลัพธ์ที่ไม่ครอบคลุมจากวิธีของบทความ [12] กับวิทยานิพนธ์นี้

นอกจากปัญหาการสืบค้นไม่พบผลลัพธ์หรือไม่ครอบคลุมซึ่งเป็นที่มาของคุณภาพการค้นหาที่ต่ำแล้วการหาข้อมูลอื่นมาทดแทนในกรณีที่ไม่พบผลลัพธ์นั้นได้จึงเป็นสิ่งที่จำเป็นอีกด้วย ซึ่งจะกล่าวถึงในหัวข้อถัดไป

2.6 การหาข้อมูลทดแทน

ในการค้นหาข้อมูลหรือค้นหาบริการมักจะพบปัญหาไม่พบข้อมูลใดตรงกับที่ระบุ วิทยานิพนธ์นี้จึงได้นำเสนอวิธีการหาข้อมูลที่มีความใกล้เคียงโดยเปรียบเทียบจากคุณสมบัติในออนโทโลจีของข้อมูลเหล่านั้นมาเสนอเพื่อทดแทนข้อมูลที่ไม่สามารถหาได้ ซึ่งรายละเอียดจะกล่าวถึงในบทที่ 5

2.7 ภาพรวมของระบบของวิทยานิพนธ์

ในหัวข้อนี้จะอธิบายถึงภาพรวมของวิธีที่ใช้ในวิทยานิพนธ์นี้ในการค้นพบบริการเว็บเซอร์วิสเชิงความหมายและเหมาะสมสำหรับความต้องของการผู้ใช้ โดยเริ่มจากการแสดงถึงสถานะของงานวิจัยนี้เมื่อเปรียบเทียบกับวิธีของงานวิจัยอื่น ๆ จากนั้นเป็นการอธิบายถึงวิธีและกระบวนการพื้นฐานที่ใช้ในการจับคู่ระหว่างข้อมูลรับเข้ากับคำอธิบายของบริการ

2.7.1 การแก้ปัญหาการค้นหาบริการด้วยการใช้ข้อมูลความชอบของผู้ใช้และออนโทโลจีในการค้นหาบริการเชิงความหมาย

จากเนื้อหาในบทที่ 1 ได้กล่าวถึงประเด็นของข้อจำกัดเกี่ยวกับคุณภาพของการค้นหาบริการเว็บเซอร์วิสในหลาย ๆวิธี ซึ่งวิทยานิพนธ์นี้พยายามที่จะปรับปรุงคุณภาพผลลัพธ์ในการค้นหาบริการจากปัญหาในประเด็นต่าง ๆ ซึ่งสรุปได้ ดังนี้

เมื่อผู้ใช้ต้องการร้องขอบริการระบบการค้นหาบริการเว็บเซอร์วิสแบบดั้งเดิมจึงเกิดขึ้นเพื่อช่วยผู้ใช้ในการค้นหาได้แก่ระบบยูดีดีไอดังที่ในบทความ [7] ได้กล่าวไว้ ซึ่งโดยพื้นฐานนั้นยูดีดีไอสามารถตอบสนองผู้ใช้ได้จำกัด จึงมีการเพิ่มเติมรายละเอียดของผู้ให้บริการและบริการเพื่อให้ผู้ใช้มีทางเลือกในการค้นหามากขึ้นดังในบทความ [8] แต่เนื่องจากยังเป็นพื้นฐานของการใช้คำสำคัญและตารางจึงยังพบกับปัญหาการสืบค้นที่ได้ผลลัพธ์ที่มีคุณภาพต่ำ ซึ่งหมายถึงไม่เป็นการค้นหาในเชิงความหมาย ระบบไม่เข้าใจคำสืบค้นที่ผู้ใช้ให้มาเนื่องจากอาจจะให้มามากหรือน้อยจนเกินไป และไม่พบผลลัพธ์ที่ต้องการหรือไม่คลอบคลุมในสิ่งที่ผู้ใช้คาดหวัง การนิยามบริการด้วยออนโทโลจีดังในบทความ [9] และการใช้บริบท ดังในบทความ [10] มาช่วยเสริมคำร้องขอหรือข้อมูลรับเข้าจะช่วยแก้ปัญหาการค้นหาในเชิงความหมายและคำสืบค้นที่ไม่ชัดเจนของผู้ใช้ได้ แต่ก็ยังประสบปัญหาการได้ผลลัพธ์ไม่ครอบคลุมหรือไม่พบตามต้องการ ซึ่งในบทความ [12] ได้พยายามเสนอวิธีการนิยามกลุ่มข้อมูลด้วยอาร์ดีเอฟเพื่อแก้ปัญหานี้แต่วิธีการและข้อด้อยของอาร์ดีเอฟยังทำให้แก้ปัญหาได้ไม่ดีพอดังที่กล่าวไปแล้ว ส่วนในบทความ [11] ได้นำข้อมูลส่วนบุคคลของผู้ใช้ซึ่งเป็นส่วนย่อยของบริบทมาช่วยในการกรองบริการที่ผู้ใช้ต้องการ แต่วิธีการดังกล่าวยังแก้ปัญหาได้แค่ระดับหนึ่งและยังมีข้อจำกัดดังที่อธิบายไว้ในหัวข้อ 2.4 ซึ่งวิธีของวิทยานิพนธ์นี้ ได้แก้ปัญหาเหล่านี้โดยการ

- นิยามบริการด้วยออนโทโลจี ซึ่งจะช่วยแก้ปัญหาการสืบค้นในเชิงความหมาย
- ใช้โดเมนและการนิยามกลุ่มคำในการกรองชนิดของบริการเพื่อเพิ่มคุณภาพของการค้นหาให้ได้ผลลัพธ์ตรงและครอบคลุมกับที่คำสืบค้นระบุด้วย โอดับเบิ้ลยูแอลซึ่งจะช่วยลดข้อด้อยของ อาร์ดีเอฟได้อีกด้วย
- ใช้การผ่อนคลายในการเลือกข้อมูลทดแทนจะช่วยแก้ปัญหาการไม่ได้ผลลัพธ์ดังที่ระบุไว้

ในคำสืบค้น โดยเสนอข้อมูลที่มีความใกล้เคียงกับข้อมูลที่ต้องการมาเสนอให้เลือกแทน

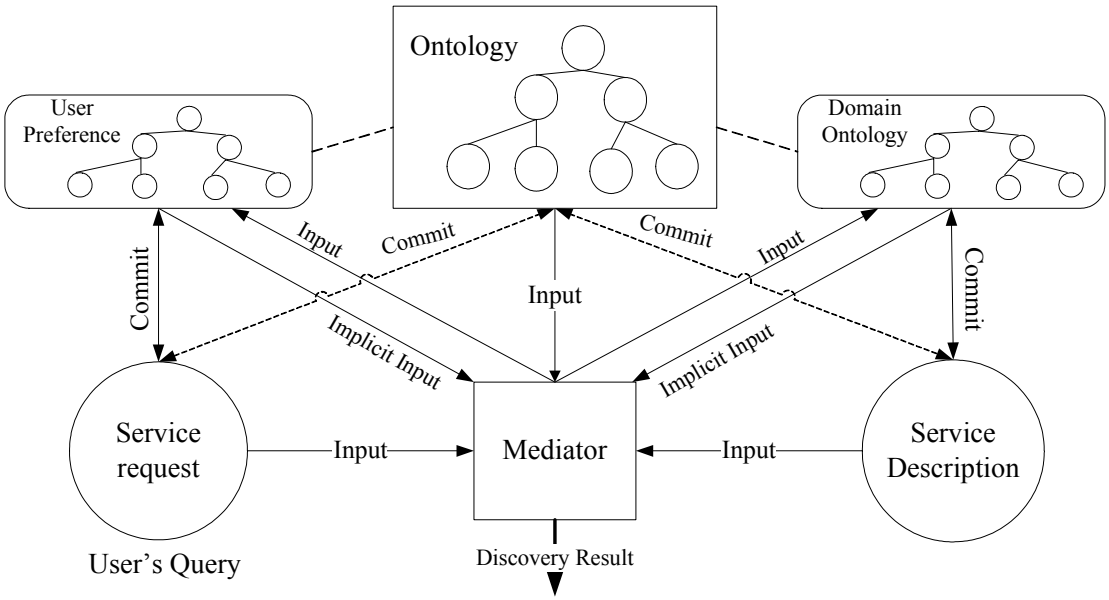
- ใช้ข้อมูลความชอบของผู้ใช้ช่วยเสริมคำสืบค้นเพื่อลดปัญหาจากการที่สิ่งที่ผู้ใช้ระบุมาไม่ชัดเจนและไม่เพียงพอ นอกจากนี้ยังช่วยเลือกบริการที่มีคุณสมบัติใกล้เคียงกับสิ่งที่ผู้ใช้ชอบหรือต้องการมากขึ้นในระดับหนึ่ง

2.7.2 รูปแบบโดยย่อของระบบ

จาก รูปที่ 2.19 เป็นรายละเอียดของการทำงานในระบบโดยย่อ โดยมีเดียเตอร์ (Mediator) เป็นเหมือนตัวกลางที่คอยจัดการเรื่องการค้นหาบริการ ในการทำงานจะรับข้อมูลรับเข้า (input) ซึ่งถูกส่งเข้ามาจาก 3 ส่วนได้แก่

1. จากคำร้องขอบริการ (Service request) จากผู้ใช้ซึ่งอยู่ในรูปของคำสืบค้นที่ผู้ใช้ให้มา
2. จากคำอธิบายหรือคำโฆษณาของบริการ (Service) ซึ่งผู้ให้บริการได้ประกาศไว้
3. จากออนโทโลจีซึ่งเป็นส่วนที่เป็นเหมือนความเข้าใจร่วมกัน (Common Understanding) ขององค์ความรู้ที่มีเดียเตอร์จะนำไปใช้เปรียบเทียบระหว่างคำร้องขอบริการกับคำอธิบายบริการ

นอกจากนี้ยังมีส่วนที่เป็นเสมือนข้อมูลรับเข้าโดยนัย (Implicit Input) ได้แก่ออนโทโลจีของความชอบของผู้ใช้ (User preference) และออนโทโลจีโดเมนบริการ (Domain Ontology) ซึ่งจะถูกใช้เป็นข้อมูลเสริมให้กับคำสืบค้นในการค้นหาบริการเว็บเซอร์วิส



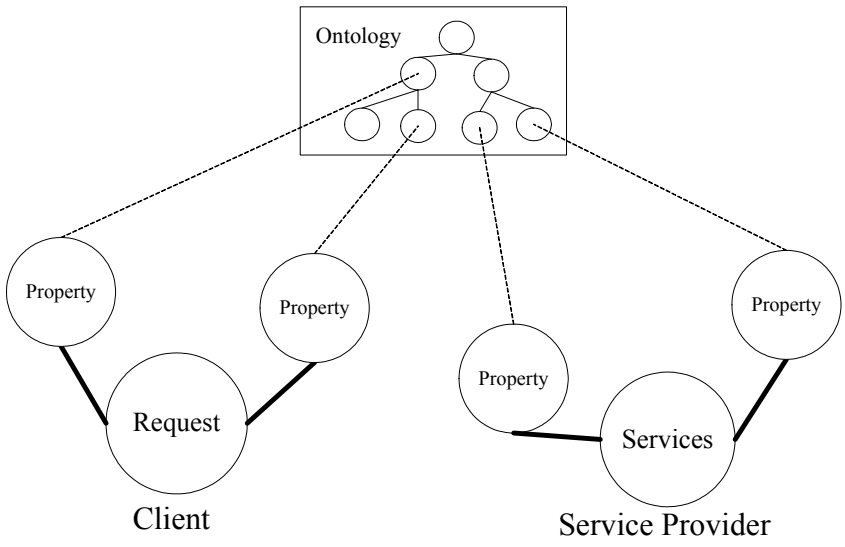
รูปที่ 2.19 ภาพรวมโดยย่อของระบบที่ออกแบบ

ข้อมูลความชอบของผู้ใช้ที่ถูกเก็บไว้ในรูปแบบออนโทโลจีจะรับข้อมูลมาจากมีเดียเตอร์ซึ่ง

ทำการแปลงข้อมูลส่วนบุคคลและความชอบที่ผู้ใช้กรอกให้แก่ระบบให้อยู่ในรูปออนโทโลจี

ส่วนข้อมูลโดเมนบริการนั้นจะมาจากที่ผู้ให้บริการกำหนดโดเมนให้กับบริการตามโครงสร้างออนโทโลจีที่ระบบให้กำหนดไว้

ในการจับคู่ (Matching) ระหว่างข้อมูลสืบค้นจากผู้ใช้กับคำอธิบายบริการนั้นจะเรียกว่าการจับคู่เชิงความหมาย (Semantic Matching) โดยการใช้ออนโทโลจีเป็นข้อมูลส่วนกลางในการเปรียบเทียบระหว่างสิ่งที่ผู้ใช้ระบุกับสิ่งที่ผู้ให้บริการนิยามไว้ ตัวอย่างเช่น บริการที่นิยามไว้ว่าเป็น 'Flight Service' ซึ่งจะหมายความรวมไปถึง 'Air Ticketing', 'Flight', 'Seat' หรือ 'Airport' ด้วย ซึ่งหากผู้ใช้ระบุว่าต้องการบริการที่จองตั๋วเครื่องบิน (Air Ticketing) หากใช้วิธีแบบคำสำคัญ ระบบจะไม่สามารถเข้าใจได้ว่า 'Flight Service' กับ 'Air Ticketing' หมายถึงเรื่องเดียวกันการจับคู่ก็จะผิดพลาด ซึ่งหากเรามีการกำหนดในออนโทโลจีถึงความสัมพันธ์ระหว่าง 'Flight Service' กับ 'Air Ticketing' ว่าหมายถึงเรื่องเดียวกันก็จะเป็นการจับคู่เชิงความหมาย ด้วยเหตุนี้การระบุคำร้องขอบริการหรือคำสืบค้นและคำอธิบายบริการจึงควรถูกกำหนดความสัมพันธ์ไปยังออนโทโลจีดัง รูปที่ 2.20



รูปที่ 2.20 การแสดงแบบของการร้องขอและบริการ

โดยพารามิเตอร์คุณสมบัติ (Property) ของบริการที่ได้จากคำสืบค้นของผู้ใช้และคำอธิบายบริการของแต่ละบริการนั้นได้ถูกแบ่งออกเป็น 4 ชนิดได้แก่

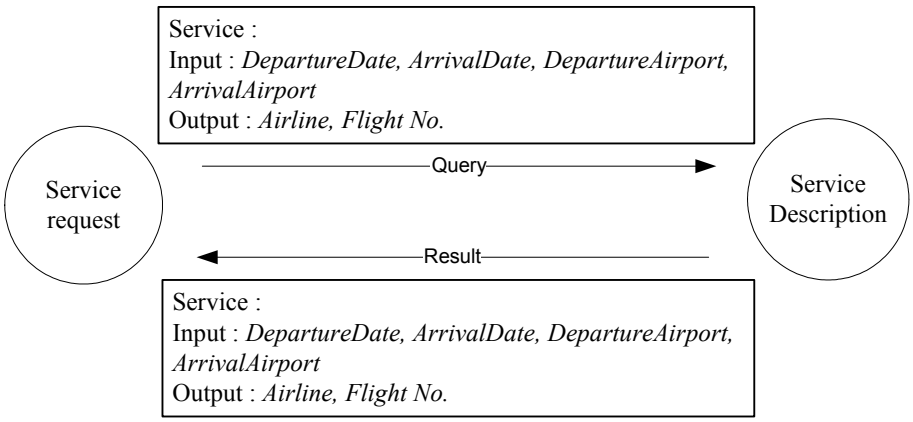
- โดเมนบริการ (Service Domain): จะบอกถึงชนิดหรือขอบเขตของบริการว่าเกี่ยวข้องกับเรื่องใด ใช้เป็นกุญแจในการกรองเอาบริการที่เกี่ยวข้องกับที่ผู้ใช้ระบุเพื่อลดจำนวนบริการที่ไม่เกี่ยวข้องออกไป
- พารามิเตอร์ในเชิงหน้าที่ของบริการ (Functional Parameter): จะเป็นรายละเอียด

เกี่ยวกับหน้าที่การทำงานของบริการว่าทำอะไรได้บ้างซึ่งแบ่งออกเป็น 2 ส่วนหลัก ๆ คือข้อมูลส่งออก (Output) และข้อมูลรับเข้า (Input)

- *ข้อมูลส่งออก (Output)* เป็นข้อมูลผลลัพธ์ที่ผู้ใช้ร้องขอและเป็นสิ่งที่บริการสามารถตอบสนองได้ เช่น ข้อมูลส่งออกของบริการหนึ่งเป็นตั๋วเครื่องบิน (Ticket) เป็นต้น
- *ข้อมูลรับเข้า (Input)* เป็นข้อมูลที่ผู้ใช้สามารถให้กับระบบและบริการสามารถรับเข้าไปเพื่อประมวลผลเป็นข้อมูลส่งออกได้ เช่น วันเวลาการเดินทาง (DepartureDate, ArrivalDate) ต้นทาง-ปลายทาง (DepartureAirport, ArrivalAirport) ในการจองตั๋วเครื่องบิน เป็นต้น

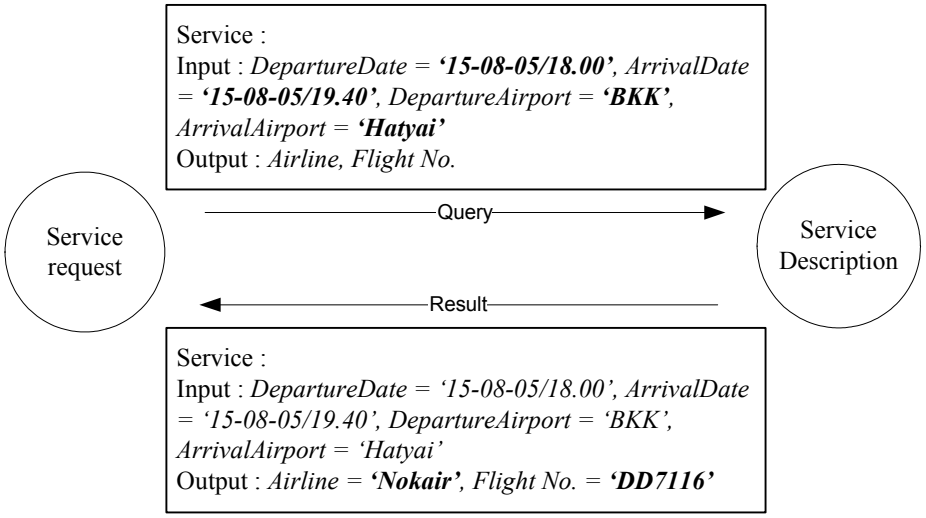
○ พารามิเตอร์ที่ไม่ใช่หน้าที่ของบริการ (Non - Functional Parameter): ได้แก่ข้อมูลซึ่งไม่ได้บอกถึงหน้าที่การทำงานของบริการโดยตรงแต่เป็นข้อมูลเสริมที่ให้ผู้ใช้ใช้ในการตัดสินจากข้อมูลอื่น ๆ เช่น ชื่อหรือรายละเอียดของผู้ให้บริการ บริบท (Context) หรือ คุณภาพของการให้บริการ (QoS of Service) เป็นต้น

○ คุณสมบัติในเชิงความหมายของบริการ (Semantic Service Property): โดยปกติบริการเว็บเซอร์วิสจะระบุคุณสมบัติเพียงว่าบริการเหล่านั้นรับและส่งข้อมูลออกมาเป็นอะไรบ้างแต่ไม่ได้บอกถึงคุณสมบัติในเชิงความหมาย เช่น บริการ 'Flight Service' มีหน้าที่ในการจองตั๋วเครื่องบินซึ่ง ข้อมูลรับเข้าที่สามารถรับได้ คือ วันเวลาการเดินทาง (DepartureDate,ArrivalDate), ต้นทาง-ปลายทาง (DepartureAirport, ArrivalAirport) ซึ่งจะให้ข้อมูลส่งออกเป็น สายการบิน (Airline) และ เที่ยวบิน (Flight No.) เป็นต้น ซึ่งการระบุค่าให้แก่พารามิเตอร์เหล่านี้จะทำให้การค้นหาเชิงรูปแบบ (Syntactic) กลายเป็นเชิงความหมายได้ดังความแตกต่างที่เห็นจาก รูปที่ 2.21 และ รูปที่ 2.22



รูปที่ 2.21 ตัวอย่างการสืบค้นบริการแบบทั่วไป

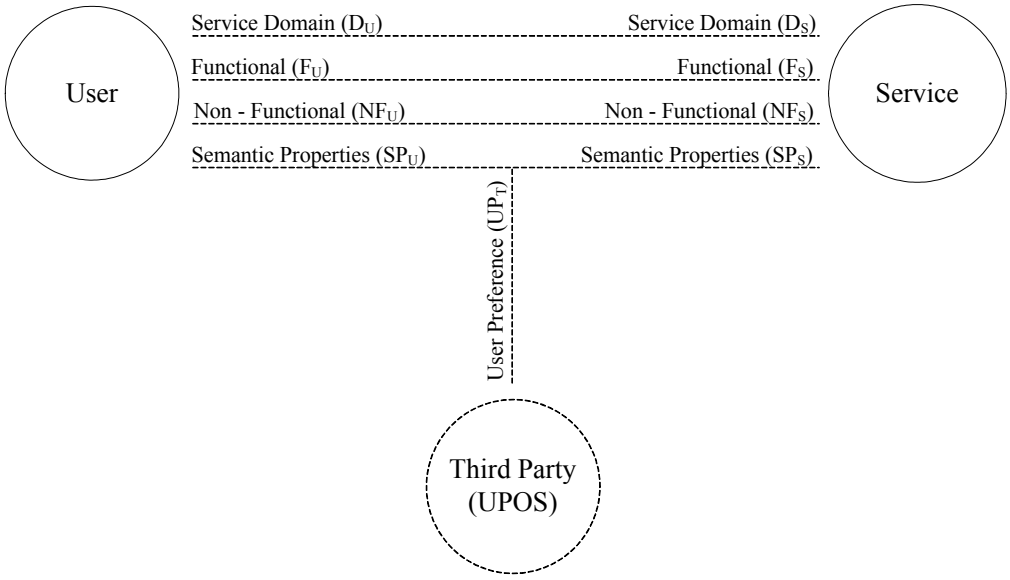
จากรูปที่ 2.21 เป็นตัวอย่างของการสืบค้นบริการเว็บเซอร์วิสแบบทั่วไปที่ได้รับการพัฒนาเพิ่มจากยูดีดีไอมาตรฐาน (สืบค้นโดยการระบุชื่อหรือชนิดของบริการเพียงอย่างเดียว ไม่สามารถระบุข้อมูลรับเข้าหรือข้อมูลส่งออกของบริการได้) ซึ่งสามารถระบุถึงคุณสมบัติว่าต้องการสืบค้นเพื่อหาบริการที่มีพารามิเตอร์ข้อมูลรับเข้าและข้อมูลส่งออกดังที่ผู้ใช้ต้องการเพียงแต่ไม่สามารถกำหนดค่าให้กับพารามิเตอร์เหล่านั้นได้เนื่องจากคำอธิบายของบริการที่เป็นดับเบิ้ลยูเอสดีแอลนั้นไม่ได้ถูกออกแบบมาให้สนับสนุนเพื่อจัดเก็บและนำเสนอข้อมูลเหล่านี้โปรแกรมหรือระบบที่ทำหน้าที่ค้นหาจึงไม่สามารถเข้าใจหรือวิเคราะห์ได้ว่าบริการเหล่านั้นมีความสามารถในการตอบสนองตามค่าที่ระบุมาในพารามิเตอร์หรือไม่ แต่หากมีการนิยามคำอธิบายคุณลักษณะของบริการในรูปแบบออนโทโลจีส่วนที่ทำหน้าที่ค้นหาจะสามารถเข้าใจและนำค่าพารามิเตอร์ที่ผู้ใช้ระบุไปเปรียบเทียบกับค่าพารามิเตอร์ที่แต่ละบริการระบุไว้ว่าสามารถตอบสนองได้ดังรูปที่ 2.22 ซึ่งจะเป็นหนทางเบื้องต้นที่จะนำไปสู่การสืบค้นในเชิงความหมาย ในการสืบค้นแบบนี้เหมาะสำหรับสนับสนุนระบบบริการแบบลูกค้าสู่ลูกค้า



รูปที่ 2.22 ตัวอย่างการสืบค้นเชิงความหมาย

นอกจากคุณสมบัติของบริการทั้ง 4 แบบจากที่ได้กล่าวเบื้องต้นซึ่งผู้ใช้เป็นผู้ระบุเพื่อให้ระบุเปรียบเทียบกับคำอธิบายบริการแล้ว ในวิทยานิพนธ์นี้ยังได้นำข้อมูลความชอบของผู้ใช้ (User Preference) ซึ่งผู้ใช้ไม่ได้ระบุมาในคำร้องขอบริการหรือคำสืบค้น แต่เคยให้ข้อมูลเหล่านี้ไว้กับระบบ เมื่อผู้ใช้ทำการค้นหาบริการ ระบบจะดึงเอาข้อมูลส่วนนี้มาเป็นข้อมูลช่วยเสริมกับคุณสมบัติบริการที่ผู้ใช้ระบุในการค้นหา ดังในรูปที่ 2.23

การจับคู่ระหว่างคุณสมบัติของบริการซึ่งถูกร้องขอโดยคำสืบค้นของผู้ใช้กับคำอธิบายบริการของผู้ให้บริการจะถูกแสดงในรูปที่ 2.23 โดยเริ่มจากชนิดของบริการที่ผู้ใช้ระบุ (D_U) จะถูกนำไปเปรียบเทียบกับชนิดของบริการที่ผู้ให้บริการระบุไว้ในโดเมนออนโทโลจี (D_S) ต่อมาเป็นส่วนของการจับคู่ระหว่างพารามิเตอร์ในเชิงหน้าที่ของบริการโดย (F_U) เป็นส่วนที่ผู้ใช้ต้องการในขณะที่ (F_S) เป็นส่วนแต่ละบริการสามารถทำได้ ซึ่งพารามิเตอร์นี้ยังถูกแบ่งออกเป็นส่วนของข้อมูลรับเข้าและข้อมูลส่งออกอีกด้วยซึ่งจะกล่าวถึงในบทถัดไป พารามิเตอร์ตัวต่อมาคือ(NF_U) และ (NF_S) ซึ่งเป็นการจับคู่กันระหว่างพารามิเตอร์ที่ไม่ใช่หน้าที่ของบริการที่ผู้ใช้ต้องการและที่แต่ละบริการระบุตามลำดับ และสุดท้ายเป็นค่าหรือข้อมูลที่ผู้ใช้ระบุว่าต้องการให้กับพารามิเตอร์แต่ละตัว (SP_U) ซึ่งจะถูกนำไปเปรียบเทียบกับความสามารถในเชิงความหมายของแต่ละบริการที่สามารถทำได้ (SP_S) ซึ่งหากข้อมูลที่ระบุจากผู้ใช้มีไม่เพียงพอหรือบริการที่ได้มาจากการสืบค้นมีมากเกินไป ข้อมูลความชอบของผู้ใช้ (UP_T) จะถูกนำเป็นเป็นข้อมูลเสริมในการเปรียบเทียบด้วย

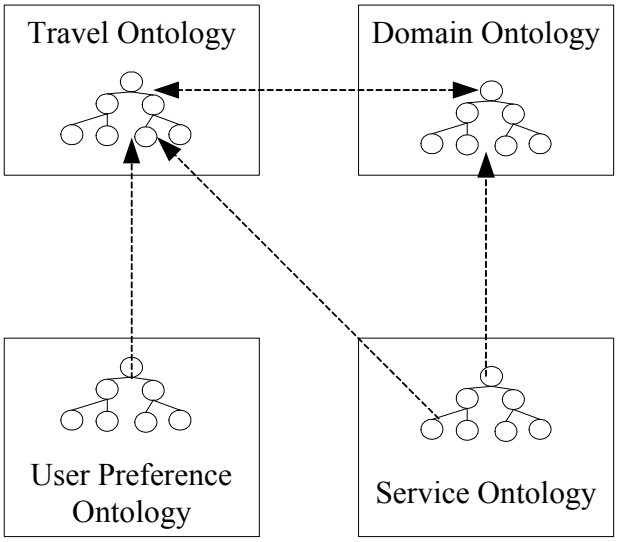


รูปที่ 2.23 การจับคู่ระหว่างคุณสมบัติของบริการ

โดยที่กำหนดให้ตัวอักษร $_U$ แทนการแสดงถึงพารามิเตอร์ที่ได้รับจากฝั่งผู้ใช้ และตัวอักษร $_S$ แทนการแสดงถึงพารามิเตอร์ที่ถูกระบุไว้ในแต่ละบริการ

2.7.3 ออนโทโลจีที่ใช้ในวิทยานิพนธ์

ออนโทโลจีเป็นหนึ่งในส่วนประกอบหลักของระบบที่ใช้ในวิทยานิพนธ์นี้ เพื่อการสร้างความเข้าใจร่วมกันและการสร้างเหตุผลในเชิงความหมายของบริการ ซึ่งในหัวข้อนี้จะอธิบายถึงรายละเอียดของการออกแบบแต่ละออนโทโลจีที่ใช้ ซึ่งเพื่อให้ง่ายแก่การเข้าใจและทดสอบจึงได้ทำการจำกัดขอบเขตของบริการไว้สำหรับเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว ซึ่งประกอบไปด้วยออนโทโลจีหลัก ๆได้แก่ ออนโทโลจีการท่องเที่ยว (Travel Ontology), ออนโทโลจีโดเมนบริการ (Domain Ontology), ออนโทโลจีบริการ (Service Ontology) และออนโทโลจีข้อมูลความชอบของผู้ใช้ (User Preference Ontology) ดังที่ปรากฏในรูปที่ 2.24



รูปที่ 2.24 ออนโทโลจีที่ใช้ในวิทยานิพนธ์

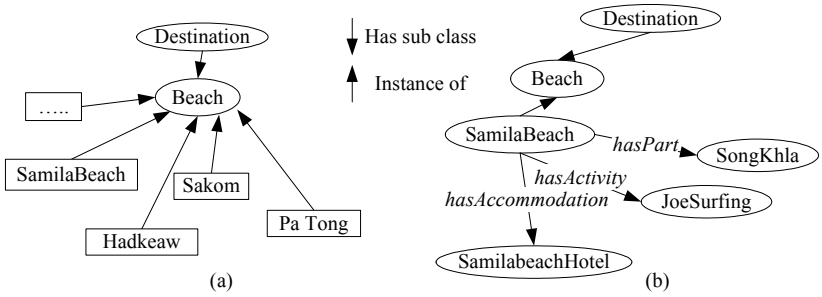
จากรูปที่ 2.24 แสดงถึงความสัมพันธ์ระหว่างออนโทโลจีที่ใช้ในงานวิจัยนี้ ซึ่งมีการเชื่อมโยงอ้างถึงกันสำหรับนำข้อมูลที่ใช้ในการค้นหามาเปรียบเทียบ

ออนโทโลจีการท่องเที่ยว (Travel Ontology)

เป็นออนโทโลจีที่นิยามข้อมูลที่เกี่ยวกับการท่องเที่ยวไว้ใช้สำหรับเป็นองค์ความรู้ส่วนกลางที่จะนำข้อมูลมาเปรียบเทียบ ซึ่งข้อมูลในออนโทโลจีนี้ได้แก่ข้อมูลจำพวก องค์กร บริษัทที่ให้บริการเกี่ยวกับการท่องเที่ยว สายการบิน โรงแรม สถานที่ รวมไปถึงข้อมูลอื่น ๆ ที่เกี่ยวข้องการบริการการท่องเที่ยว ในงานวิจัยนี้ได้อ้างอิงจากออนโทโลจีของ Holger Knublauch [25] ซึ่งเป็นตัวอย่างออนโทโลจีการท่องเที่ยวที่นิยมใช้ในการศึกษากันอย่างแพร่หลายในด้านเว็บเชิงความหมาย มีรายละเอียดดังใน รูปที่ 2.25 ซึ่งเป็นตัวอย่างที่ถูกพัฒนาด้วยโปรแกรมโพรทีเจ

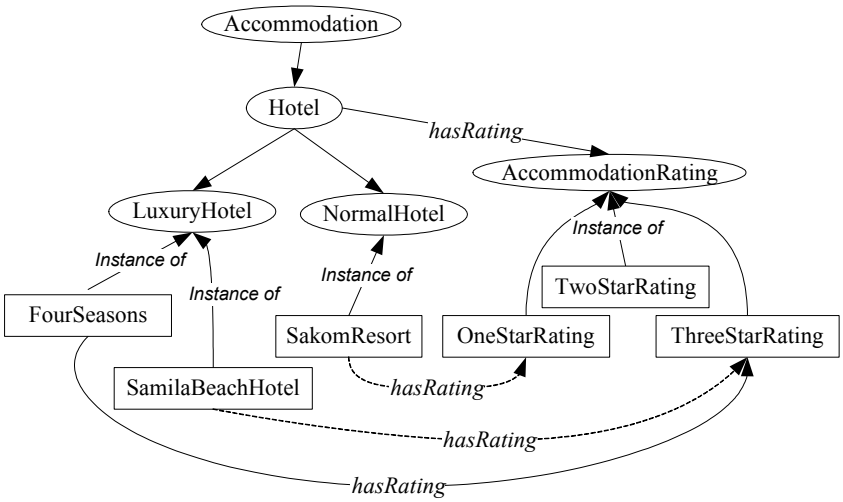
รูปที่ 2.25 ออนโทโลจีการท่องเที่ยว

ซึ่งตัวอย่างรายละเอียดความสัมพันธ์ระหว่างข้อมูลในออนโทโลจีนี้เป็นดังใน รูปที่ 2.26 และ รูปที่ 2.27



รูปที่ 2.26 ออนโทโลจีของ Beach

จาก รูปที่ 2.26 (a) เป็นตัวอย่างของออนโทโลจีชายหาด เช่น Sakom, Hadkeaw, Pa Tong เป็นต้น และใน รูปที่ 2.26 (b) แสดงถึงคุณสมบัติที่คลาสของชายหาดมีได้แก่ *hasPart*, *hasActivity* และ *hasAccommodation* เป็นต้น

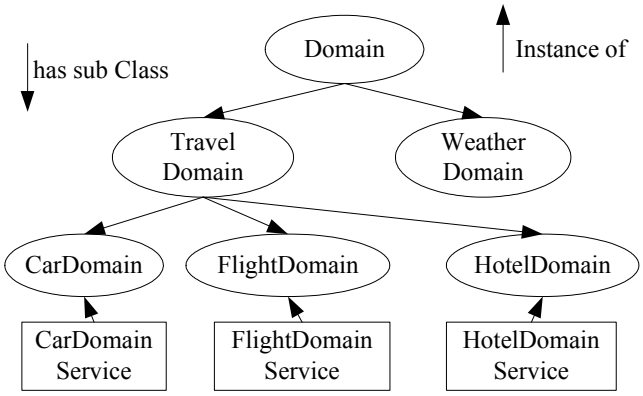


รูปที่ 2.27 ออนโทโลจีของ Accommodation

ใน รูปที่ 2.27 เป็นตัวอย่างคลาสในออนโทโลจีที่พัก Accommodation

ออนโทโลจีโดเมนบริการ (Domain Ontology)

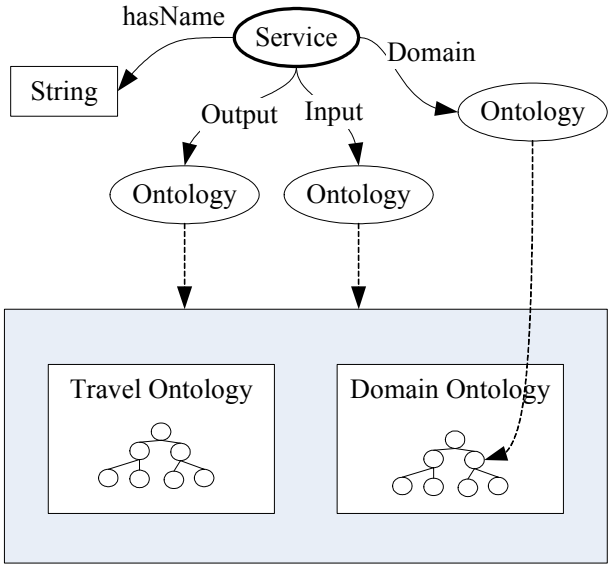
เป็นออนโทโลจีซึ่งรวบรวมโดเมนบริการในขอบเขตต่าง ๆ ซึ่งจะถูกใช้ในการระบุชนิดของบริการซึ่งรายละเอียดของออนโทโลจีนี้ใช้ออนโทโลจีที่ออกแบบมาจากกลุ่มพันธมิตรการท่องเที่ยวแบบเปิดซึ่งมีตัวอย่างออนโทโลจีของโดเมนบริการดังใน รูปที่ 2.28



รูปที่ 2.28 ตัวอย่างออนโทโลจีโดเมนบริการ

ออนโทโลจีบริการ (Service Ontology)

ออนโทโลจีบริการเป็นออนโทโลจีที่งานวิจัยนี้ได้ออกแบบขึ้นเพื่อเป็นโครงสร้างออนโทโลจีส่วนบน (Upper Ontology) ของแต่ละบริการ



รูปที่ 2.29 ออนโทโลจีบริการ

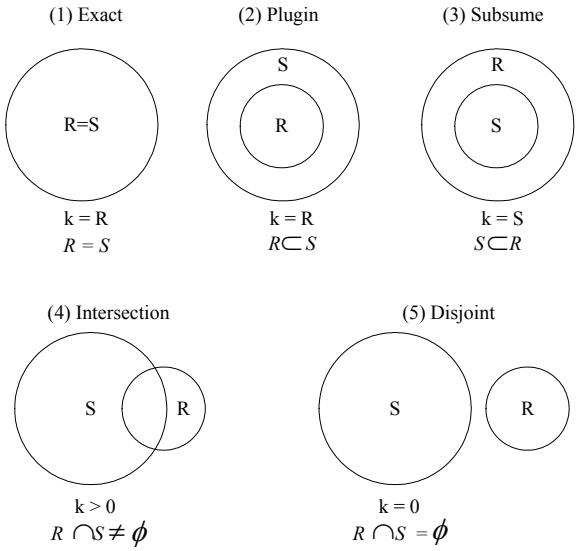
จาก รูปที่ 2.29 เป็นโครงสร้างออนโทโลจีบริการซึ่งประกอบไปด้วยชื่อของบริการซึ่งเป็นสายอักขระ (String) ส่วนต่อมาจะบอกถึงชนิดหรือโดเมนบริการซึ่งจะชี้ไปยังออนโทโลจีส่วนกลางคือออนโทโลจีโดเมนบริการ และสุดท้ายจะบอกถึงความสามารถในเชิงหน้าที่ของบริการว่าต้องการรับข้อมูลรับเข้าและข้อมูลส่งออกอะไรบ้างส่วนนี้จะชี้ไปยังข้อมูลส่วนกลางทั้งออนโทโลจีการท่องเที่ยวและออนโทโลจีโดเมนบริการ

ออนโทโลจีข้อมูลความชอบของผู้ใช้ (User Preference Ontology)

เป็นข้อมูลผู้ใช้และความชอบในการใช้บริการต่าง ๆ ซึ่งระบุถึงสถานการณ์และเงื่อนไข เพื่อให้ระบบใช้เป็นข้อมูลเสริมในการตัดสินใจ ซึ่งงานวิจัยนี้ได้ออกแบบขึ้นโดยใช้ข้อมูลจากโครงสร้างบริการการท่องเที่ยวของกลุ่มพันธมิตรการท่องเที่ยวแบบเปิดซึ่งจะกล่างถึงรายละเอียดต่อไปในบทที่ 4

2.7.4 ระดับความเหมือนของการจับคู่ (Matching Degree)

ในการสืบค้นข้อมูลโดยทั่วไปต้องการให้ได้ผลลัพธ์ที่เกี่ยวข้องกับที่ต้องการมากที่สุดนั่นคือมีค่า recall และ precision สูงสุด ในการวัดคุณภาพการจับคู่ระหว่างคุณสมบัติของคำร้องขอกับคำอธิบายบริการในงานวิจัยนี้จะคิดจากจำนวนคุณสมบัติที่เหมือนกันของคุณสมบัติที่ผู้ใช้ต้องการและที่แต่ละบริการระบุ เช่น ผู้ใช้ต้องการค้นหาบริการและได้ระบุข้อมูลรับเข้า ข้อมูลส่งออก รวมไปถึงคุณสมบัติและพารามิเตอร์อื่น ๆ รวมแล้ว 7 อย่าง ในขณะที่แต่ละข้อมูลที่จะถูกนำมาจับคู่นั้น มีคุณสมบัติที่สามารถตอบสนองตามที่ผู้ใช้ต้องการได้ไม่เท่ากันและแตกต่างกันไป ดังเช่น service1 มีคุณสมบัติที่ตรงกับที่ผู้ใช้ระบุไว้ 6 อย่างแต่ไม่ตรง 1อย่าง ในขณะที่ service2 มีคุณสมบัติที่ตรงกับที่ผู้ใช้ระบุไว้ 4 อย่างแต่ไม่ตรง 3 อย่าง และservice3 มีคุณสมบัติที่ตรงกับที่ผู้ใช้ระบุไว้ 3 อย่างแต่ไม่ตรง 4 อย่าง นั่นแสดงว่า service1 มีคุณสมบัติที่ตรงกับผู้ใช้ต้องการมากกว่า service2และ3 ในขณะที่ service2 มีคุณสมบัติที่ตรงกับผู้ใช้ต้องการมากกว่า service3 เป็นต้น ซึ่งชนิดของความเหมือนของการจับคู่จะแบ่งดังในรูปที่ 2.30



รูปที่ 2.30 ชนิดของการจับคู่

จากรูปที่ 2.30 เป็นการแบ่งชนิดของการวัดค่าความเหมือนระหว่างข้อมูล โดยที่ R เป็นจำนวนคุณสมบัติที่ฝั่งร้องขอระบุ โดยจะใช้ค่านี้เป็นฐานในการเปรียบเทียบ กับค่า S ที่เป็นจำนวน

คุณสมบัติของข้อมูลที่จะนำมาเปรียบเทียบ (เช่น จำนวนคุณสมบัติของแต่ละบริการที่จะนำมาเปรียบเทียบกับจำนวนคุณสมบัติที่ผู้ใช้ระบุ) กับค่า R ส่วนค่า k เป็นจำนวนของคุณสมบัติที่ S มีเหมือนกับ R เช่นผู้ใช้ร้องขอบริการได้ระบุคุณสมบัติของบริการไว้ในคำสืบค้น request1 ว่าต้องการบริการที่มีคุณสมบัติ 5 อย่าง ($R = 5$) และบริการที่ชื่อ service1 ได้ประกาศรายละเอียดไว้ซึ่งประกอบด้วยคุณสมบัติ 6 อย่าง ($S = 6$) แต่คุณสมบัติที่ service1 มีเหมือนกับ request1 มีอยู่ 4 อย่าง ดังนั้นค่า k = 4 เป็นต้น ซึ่งจำแนกชนิดของการจับคู่อธิบายได้ดังนี้

การจับคู่ที่แม่นยำอย่างแท้จริง (Exact) : หมายความว่าบริการและคำสืบค้นที่ผู้ใช้ร้องขอมีคุณสมบัติตรงกันทุกประการ จึงทำให้ค่า $R = S$ นั่นคือค่า k = R เป็นการจับคู่ที่ดีที่สุด

การจับคู่แบบเสริม (Plugin): เป็นการจับคู่ที่ดีที่สุดอันดับสองรองจาก Exact นั่นคือคุณสมบัติที่บริการมีอยู่มีมากกว่าที่คำสืบค้นระบุและสามารถตอบสนองกับที่ผู้ใช้ร้องขอได้ทั้งหมด ทำให้ค่า k= R เนื่องจาก R เป็นกลุ่มย่อย (subset) ของ S

การจับคู่แบบจัดเป็นหมวดหมู่ (Subsume): คือคุณสมบัติที่บริการมีทั้งหมดเป็นส่วนหนึ่งหรือกลุ่มย่อยของที่ผู้ใช้ร้องขอทำให้ค่า k = S เนื่องจาก S เป็นกลุ่มย่อย (subset) ของ R

การจับคู่แบบตัดผ่านกัน (Intersection): จะมีความคล้ายคลึงกับแบบ Subsume คือเป็นการจับคู่ที่ไม่ดีมากนักเนื่องจากบางส่วนของคุณสมบัติของบริการตรงกับบางส่วนของคุณสมบัติที่ผู้ใช้ร้องขอเท่านั้น ค่า k จึงมีค่ามากกว่าศูนย์แต่น้อยกว่า S หรือ R

การจับคู่แบบไม่มีส่วนต่อ (Disjoint): นั่นคือบริการและคำร้องขอไม่มีส่วนใดที่เหมือนกันเลย ทำให้ค่า k = 0 บริการที่เป็นชนิดนี้จะไม่ได้ถูกนำไปพิจารณาในขั้นตอนต่อไป

ดังนั้นบริการที่เป็นแบบ Exact และ Plugin จะถูกจัดอยู่ในกลุ่มที่มีความถูกต้องสูง ในขณะที่บริการที่เป็นแบบ Subsume และ Intersection จะถูกจัดให้อยู่ในกลุ่มที่มีความใกล้เคียงกับที่ต้องการเนื่องจากยังมีบางส่วนที่ตรงกัน แต่บริการที่เป็นแบบ Disjoint จะถูกจัดให้เป็นกลุ่มที่เป็นการจับคู่ที่ไม่เหมาะสมจึงถูกตัดออกไป

ถึงแม้ว่าบริการในกลุ่มที่มีความใกล้เคียง (Subsume และ Intersection) จะถูกนำไปพิจารณาต่อ แต่เพื่อลดจำนวนข้อมูลบริการที่นำมาเปรียบเทียบมีมากเกินไปจึงตัดบริการที่มีคุณสมบัติตรงกับที่ผู้ใช้ระบุน้อยกว่าครึ่งออก (k < 0.5R) และนำข้อมูลบริการส่วนที่มีคุณสมบัติตรงกับที่ผู้ใช้ระบุมากกว่าครึ่งไปทำการสืบค้นในขั้นตอนต่อไป

2.8 ขั้นตอนการค้นหาบริการที่นำเสนอในวิทยานิพนธ์

ในหัวข้อนี้จะอธิบายถึงอัลกอริทึมที่ใช้ในการค้นหาบริการ ซึ่งเป็นขั้นตอนวิธีที่จะแยกบริการออกจากส่วนที่ไม่เกี่ยวข้องทีละส่วน จนเหลือเพียงบริการที่ใกล้เคียงกับที่ผู้ใช้ต้องการมากที่สุด ซึ่งอัลกอริทึมนี้จะเน้นการกรองบริการทีละขั้นโดยยึดจากพารามิเตอร์ที่มีการจัดเรียง

ความสำคัญเป็นเงื่อนไขในการสืบค้น ซึ่งลำดับความสำคัญของพารามิเตอร์ที่จะนำมาสืบค้นนั้นจะช่วยให้ได้คุณภาพการค้นหาสูงยิ่งขึ้น โดยขั้นตอนการค้นหาจะเรียงตามลำดับความสำคัญของการค้นหาบริการ คือ

1. โดเมนบริการ (Service Domain) ที่ผู้ใช้ระบุ (D_U)
2. พารามิเตอร์ในเชิงหน้าที่ของบริการ (Functional Parameter) ซึ่งแบ่งออกเป็น 2 ส่วนหลัก ๆคือข้อมูลส่งออก (Output) และข้อมูลรับเข้า (Input)
    a. ข้อมูลส่งออก (Output) ของบริการที่ผู้ใช้ระบุ (Fo_U)
    b. ข้อมูลรับเข้า (Input) ของบริการที่บริการระบุ (Fi_S)
3. พารามิเตอร์ที่ไม่ใช่หน้าที่ของบริการ (Non – Functional Parameter) ที่ผู้ใช้ระบุ (NF_U)
4. คุณสมบัติในเชิงความหมายของบริการ (Semantic Service Property) ที่ผู้ใช้ระบุ (SP_U)
5. ข้อมูลทดแทน
6. ความชอบของผู้ใช้ (User Preference) ที่ไม่ได้ระบุในคำสืบค้น (UP_T)

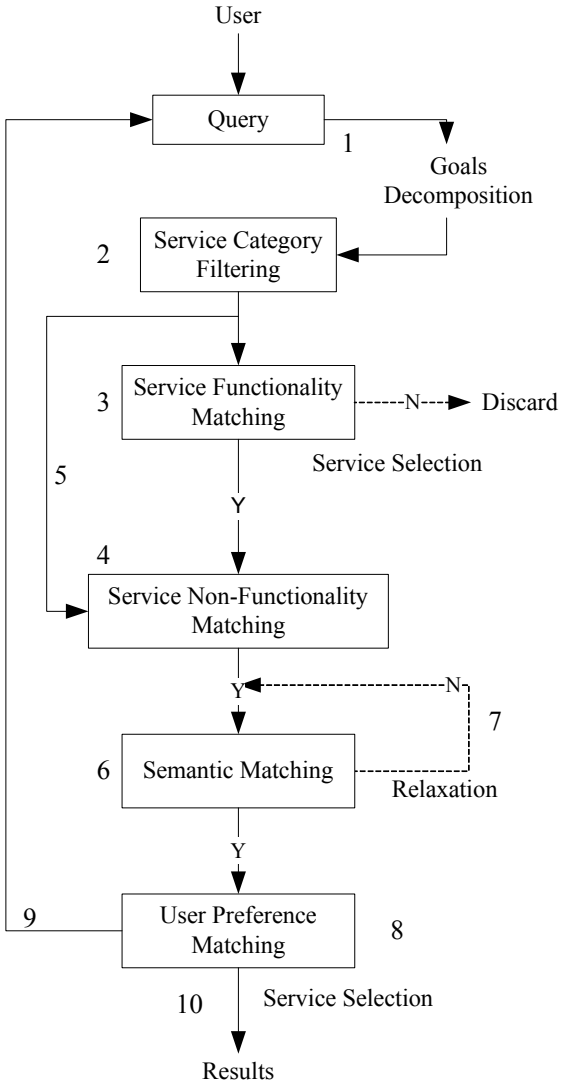
เนื่องจากปัญหาของการค้นหาแบบคำสำคัญคือระบบไม่สามารถเข้าใจความหมายของคำอธิบายบริการที่อยู่ในรูปข้อความได้จึงไม่สามารถจำแนกได้ว่าบริการใดเป็นบริการชนิดใด ดังนั้นการจับคู่ระหว่างคำสำคัญกับข้อความอาจทำให้ได้ข้อมูลที่ผิดพลาดเนื่องจากคำที่เขียนเหมือนกันแต่ความหมายต่างกัน ทำให้คุณภาพของการค้นหาต่ำเพราะได้บริการที่ไม่เกี่ยวข้องกับที่ผู้ใช้ต้องการกลับมามากมาย เมื่อนำไปค้นหาในขั้นตอนต่อไปจะทำให้ได้ข้อมูลที่ผิดพลาด ดังนั้นการใช้โดเมนหรือชนิดของบริการสำหรับกรองบริการที่เกี่ยวข้องกับคำร้องขอจะช่วยให้ได้บริการที่ไม่ผิดพลาดอีกทั้งช่วยลดจำนวนบริการในการนำไปค้นหาในขั้นตอนต่อไปอีกด้วย

จากนั้นเมื่อมั่นใจได้ว่าบริการที่นำมาคัดเลือกนั้นไม่ผิดไปจากชนิดหรือโดเมนที่ต้องการแล้ว ส่วนต่อมาคือจะคำนึงถึงหน้าที่หรือความสามารถของบริการว่าสามารถตอบสนองคำร้องขอได้เพียงใดพารามิเตอร์ที่บอกถึงความสามารถของบริการว่าสามารถให้ผลลัพธ์หรือข้อมูลส่งออกเป็นอะไร หรือต้องการรับข้อมูลรับเข้าอะไรบ้างจึงมีความสำคัญเป็นลำดับต่อมา ซึ่งโดยปกติข้อมูลส่งออกมักจะเป็นตัวบอกถึงลักษณะการทำงานของบริการนั้น ๆ ดังนั้นในการพิจารณาพารามิเตอร์ส่วนนี้จึงให้ความสำคัญกับข้อมูลส่งออกที่ผู้ใช้ระบุ แล้วจึงพิจารณาถึงข้อมูลรับเข้าที่บริการระบุ

เมื่อได้บริการที่มีคุณสมบัติการทำงานตรงกับที่ต้องการแล้วจึงมาพิจารณาต่อถึงคุณสมบัติเสริมที่ไม่ได้เกี่ยวข้องกับการทำงานของบริการ

หลักจากได้บริการที่มีพารามิเตอร์ตรงกับที่ผู้ใช้ระบุแล้วจึงนำมาพิจารณาค่าของพารามิเตอร์แต่ละตัวในเชิงความหมาย ซึ่งหากพบว่าไม่มีบริการใดมีคุณสมบัติดังที่ได้ระบุไว้ข้อมูลทดแทนจะถูกนำมาเป็นพารามิเตอร์ในการสืบค้นแทน

และสุดท้ายเมื่อพารามิเตอร์ที่จะนำมาเป็นคำสืบค้นเพื่อเลือกบริการคือข้อมูลความชอบของผู้ใช้ซึ่งผู้ใช้เคยระบุไว้ในระบบ ข้อมูลนี้จะถูกดึงขึ้นมาเพื่อเปรียบเทียบกับคุณสมบัติที่บริการเหล่านั้นมีเพื่อให้ได้บริการตรงกับที่ผู้ใช้ต้องการมากที่สุด โดยขั้นตอนในการค้นหาบริการทั้งหมดจะถูกแสดงดังในรูปที่ 2.31



รูปที่ 2.31 แผนภาพแสดงขั้นตอนการค้นหาของอัลกอริทึม

จากขั้นตอนการค้นหาบริการดังในรูปที่ 2.31มีรายละเอียดดังนี้

1. Goals Decomposition จากต้นแบบจะทำการแยกเป้าหมายของผู้ใช้ว่าระบุรายละเอียดในค้นหาในขอบเขตใด เช่น ต้องการรายละเอียดแค่หน้าที่การทำงานของบริการโดยระบุแค่ ชนิดของบริการ (Domain), ข้อมูลรับเข้า (input) และข้อมูลส่งออก (output) หรือผู้ใช้อาจจะระบุถ้า

รายละเอียดเชิงความหมายของบริการ (ตัวอย่างเช่น ต้องการบริการที่สามารถจองโรงแรมที่อยู่ติดชายหาด)

2. Service Category Filtering นำข้อมูลคำร้องขอมาทำการกรองประเภทโดเมนบริการและตัดทุกบริการที่ไม่ตรงกับข้อมูลที่ผู้ใช้ระบุในคำสืบค้น โดยโดเมนที่ผู้ใช้ระบุจะถูกเปรียบเทียบกับ โดเมนย่อยที่อยู่กลุ่มเดียวกันเช่น hotel accommodation และ room ถูกจัดอยู่ใน HotelDomain

ขั้นตอนนี้เป็นการแก้ปัญหาการได้บริการที่ไม่เกี่ยวข้องหรือไม่ครอบคลุม (Low Recall and Low Precision Solution) ซึ่งเป็นผลให้ค่า Recall และ Precision มีค่าต่ำ การกรองบริการที่เกี่ยวข้องเฉพาะในโดเมนที่ต้องการ แบ่งได้เป็น 2 วิธีคือ

Domain Specific Matching ให้ผู้ใช้ระบุถึงโดเมนที่ชัดเจนเช่น เลือกโดเมนจาก combo box ที่เป็น Flight ระบบก็จะเลือกเอาทุกบริการที่เป็น Flight domain ออกมา วิธีนี้จะง่ายและถูกต้องแน่นอน ระบบไม่จำเป็นต้องวิเคราะห์ถึงคำสืบค้นเนื่องจาก มีตัวเลือกที่คงตัวให้ผู้ใช้เลือก จากวิธีนี้ค่า recall และ precision ที่วัดได้จะเป็น 1 หรือถูกต้อง 100 %

Words Catalog Ontology Matching ผู้ใช้ระบุโดเมนในคำสืบค้น วิธีนี้จำเป็นต้องมีการสร้างชุดคำศัพท์ของคำที่มีความหมายเหมือนกันไว้สำหรับเปรียบเทียบกับคำสืบค้น เช่นกลุ่มคำที่มีความหมายเดียวกับ Flight domain คือ Flight, airline, air ticketing, air service และอื่นๆ เนื่องจากผู้ใช้แต่ละคนอาจจะระบุคำสืบค้นที่แตกต่างกันไป ความถูกต้องและครอบคลุมของวิธีนี้จึงขึ้นอยู่กับความละเอียดและครอบคลุมเพียงพอของกลุ่มคำที่มีความหมายเหมือนกัน

3. Service Functionality Matching หากผู้ใช้ระบุขอบเขตของบริการแค่หน้าที่การทำงานของบริการ เช่น ข้อมูลรับเข้า (input) และ ข้อมูลส่งออก (output) ให้ทำการนำบริการที่ได้จากการกรองมาเปรียบเทียบส่วนของหน้าที่การทำงานของบริการโดยการจับคู่ระหว่างรายละเอียดในคำสืบค้นกับออนโทโลจีบริการของแต่ละบริการที่ระบุไว้ ทำการตัดทิ้งบริการที่ไม่ถูกต้อง และเก็บผลลัพธ์เอาไว้

4. Service Non-Functionality Matching หากผู้ใช้ระบุรายละเอียดที่ไม่ใช่หน้าที่การทำงานของบริการ ใช้สิ่งที่ผู้ใช้ระบุมาสืบค้นกับออนโทโลจีเชิงความหมายของบริการ เก็บผลลัพธ์ที่ได้ไว้สืบค้นในขั้นตอนต่อไป

5. หากผู้ใช้ไม่ได้ระบุดังข้อที่ 3 ให้ข้ามมาข้อที่ 4 ต่อไป และหากไม่ได้ระบุดังข้อที่ 4 ให้ทำข้อที่ 3 และข้ามไปข้อที่ 6 ต่อไป

6. Semantic Matching ทำการสืบค้นเชิงความหมายโดยการนำค่าของพารามิเตอร์ที่ผู้ใช้ระบุมาเปรียบเทียบกับที่บริการนิยามไว้

7. Relaxation ไม่มีบริการที่ผู้ใช้คาดหวังและผู้ใช้ยืนยันจะใช้การสืบค้นด้วยคำสืบค้นเดิมกลับไปทำในขั้นตอนที่ 6 โดยการใช้การผ่อนคลายออนโทโลจีบริการ หากผู้ใช้ไม่ยืนยันในการใช้

คำสืบค้นเดิมให้ผู้ใช้ทำการระบุรายละเอียดการสืบค้นใหม่

8. User Preference Matching หากในการสืบค้นในขั้นตอนต้น ๆ ได้บริการที่ผ่านการสืบค้นอีกหลายบริการ และผู้ใช้มีข้อมูลส่วนตัวที่นิยามขึ้นในรูปแบบออนโทโลจีซึ่งรายละเอียดระบุถึงข้อมูลความชอบจากออนโทโลจี ให้นำบริการเหล่านั้นเปรียบเทียบกับข้อมูลความชอบที่ผู้ใช้ระบุไว้ในเงื่อนไขข้อมูลส่วนตัว เพื่อนำไปจัดลำดับบริการที่เหมาะสมและเสนอกับผู้ใช้ ซึ่งหลักและวิธีการนำข้อมูลความชอบของผู้ใช้ไปคิดคำนวณค่าคะแนนความชอบและคะแนนรวมของบริการจะกล่าวถึงในบทที่ 4

2.9 เครื่องมือที่ใช้ในวิทยานิพนธ์

2.9.1 กลุ่มพันธมิตรการท่องเที่ยวแบบเปิด (Open Travel Alliance : OTA)

กลุ่มพันธมิตรการท่องเที่ยวแบบเปิด [21] เป็นการร่วมมือกันระหว่างองค์กรที่ทำงานเกี่ยวกับระบบการท่องเที่ยวในการออกแบบข้อกำหนดรายละเอียดกระบวนการการให้บริการในเชิงท่องเที่ยวซึ่งแบ่งโดเมนหลัก ๆได้แก่ บริการโรงแรม (Hotal Service) บริการเที่ยวบิน (Flight Service) และบริการรถ (Car Service) เป็นต้น ซึ่งรายละเอียดจะแสดงไว้ในภาคผนวก ข

2.9.2 จีนา (Jena)

จีนาเป็นกรอบการทำงานสำหรับสร้างโปรแกรมประยุกต์ทางด้านเว็บเชิงความหมาย (Semantic Web) [23] โดยสนับสนุนสภาพแวดล้อมในการเขียนโปรแกรมสำหรับ อาร์ดีเอฟ โอดับเบิ้ลยูแอล รวมถึง การอนุมานโดยใช้กฎ (Rule-Based Inference Engine) โดยจีนามีส่วนประกอบดังนี้

- ส่วนติดต่อโปรแกรมประยุกต์อาร์ดีเอฟ (RDF API)
- การนำเข้าและส่งออกอาร์ดีเอฟโมเดลในรูปแบบ อาร์ดีเอฟเอ็กซ์เอ็มแอล โนเทชันทรี และ เอ็นทริปเปิล
- ส่วนติดต่อโปรแกรมประยุกต์โอดับเบิ้ลยูแอล (OWL API)
- การสนับสนุนการจัดการอาร์ดีเอฟโมเดลทั้งในหน่วยความจำ (Memory) และจากฐานข้อมูล (Database)

ภาษาอาร์ดีคิวแอล (RDQL – Query Language) เป็นภาษาสำหรับการค้นหาคลาสหรืออินสแตนซ์ที่อธิบายโดยใช้อาร์ดีเอฟ โดยใช้ความสัมพันธ์ระหว่างซับเจกต์ เพรดิเคทและออบเจกต์เป็นเงื่อนไขในการค้นหา ตัวอย่างการค้นหาซับเจกต์ดังใน รูปที่ 2.32 คือการสืบค้นข้อมูลบริการซึ่งมีคุณสมบัติคือมีจุดหมายปลายทาง (toDestination) ไปยังหาดสมิหลา (SamilaBeach) ซึ่งข้อมูลเหล่านี้ถูกนิยามขึ้นด้วยความสัมพันธ์ในเชิงออนโทโลจีและเก็บไว้ใน

http://localhost:8080/testtravel.owl

```
SELECT ?Service
WHERE  (?Service, <p:toDestination>, p:SamilaBeach)
USING  p FOR <http://localhost:8080/testtravel.owl#>
```

รูปที่ 2.32 ตัวอย่างการใช้ภาษาอาร์ดีคิวแอลในการค้นหาข้อมูลจากออนโทโลจี

2.9.3 โพรทีเจ (Protege)

โพรทีเจเป็นเครื่องมือสำหรับการพัฒนาออนโทโลจีและฐานความรู้ (Knowledge-based) โดยสนับสนุนการสร้างออนโทโลจีด้วยภาษา อาร์ดีเอฟ และโอดับเบิ้ลยูแอล เนื่องจากโพรทีเจได้นำไลบรารีของจีนามาใช้ ทำให้โพรทีเจรองรับการอนุมานโดยใช้กฎและการค้นหาโดยใช้ภาษาอาร์ดีคิวแอลได้ นอกจากนี้ยังใช้งานได้ง่ายอีกด้วย [24]

ซึ่งในงานวิจัยนี้ได้นำโพรทีเจมาใช้สำหรับออกแบบและสร้างข้อมูลให้เป็นออนโทโลจี

2.10 กรณีศึกษาการท่องเที่ยว (Travel Case)

ในหัวข้อนี้จะอธิบายถึงกรณีศึกษาซึ่งจะนำเสนอในวิทยานิพนธ์นี้ ซึ่งจะยกตัวอย่างการค้นหาของบริการในโดเมนการท่องเที่ยวสำหรับทดสอบอัลกอริทึมและกระบวนวิธีการของวิทยานิพนธ์ เพื่อให้ง่ายแก่การเข้าใจจึงได้ย่อขอบเขตการทดสอบลงมาเป็นการค้นหาบริการเกี่ยวกับการจองตั๋วเครื่องบินดังตัวอย่างต่อไปนี้

พิจารณาตัวอย่าง ผู้ใช้ Warakorn ต้องการค้นหาบริการที่สามารถจองตั๋วเครื่องบินจาก กทม.ไปยังหาดใหญ่ ซึ่งต้องการถึงสนามบินหาดใหญ่ก่อนเวลา 21.00น. ในวันที่ 15 สิงหาคม 2005 เมื่อผู้ใช้สอบถามมายังระบบหรือโปรแกรมประยุกต์ที่ช่วยค้นหาบริการ พารามิเตอร์ที่เกี่ยวข้องกับคำสืบค้นได้แก่ Domain, DepartureAirport, ArrivalAirport, ArrivalDate ซึ่งในกรณีนี้จะสรุปพารามิเตอร์ต่าง ๆได้ดังนี้

- โดเมนบริการที่ผู้ใช้ระบุ (D_U) - Domain
- พารามิเตอร์ในเชิงหน้าที่ของบริการ
    - ข้อมูลส่งออกของบริการที่ผู้ใช้ระบุ (Fo_U) - DepartureAirport, ArrivalAirport, ArrivalDate, Service
    - ข้อมูลรับเข้าของบริการที่ผู้ใช้ระบุ (Fi_U) - DepartureAirport, ArrivalAirport, ArrivalDate
- พารามิเตอร์ที่ไม่ใช่หน้าที่ของบริการที่ผู้ใช้ระบุ (NF_U) - None
- คุณสมบัติในเชิงความหมายของบริการที่ผู้ใช้ระบุ (SP_U) - Domain = “Flight Service”,

DepartureAirport = “BKK”, ArrivalAirport = “Hatyai”, ArrivalDate “< 15-08-05 21.00” เป็นต้น

เริ่มจากสืบค้นบริการจากโดเมนบริการซึ่งจะเลือกเอาทุกบริการที่มีความหมายของโดเมนบริการเหมือนกับคำว่า “FlightService” ตัวอย่างการสืบค้นด้วยภาษาอาร์ดีคิวแอลจะเป็นดังในรูปที่ 2.33

```
SELECT     ?srv
WHERE    (?services, <owl:sameAs>, <p:FlightService>)
                (?srv, <rdf:type>, ?services)
USING     p FOR <http://localhost:8080/ServiceDomainOntology.owl#>
```

รูปที่ 2.33 ตัวอย่างการสืบค้นด้วยโดเมนบริการ

จากนั้นจะสืบค้นบริการจากพารามิเตอร์หน้าที่ของบริการดังตัวอย่างในรูปที่ 2.34

```
SELECT     ?srv
WHERE        (?service, <p:hasOutput>, ?out)
                (?out, <owl:sameAs>, <p:DepartureAirport>)
                        .
                        .
                (?out, <owl:sameAs>, <p:ArrivalAirport>)
                        .
                        .
                (?out,<owl:sameAs>, <p:ArrivalDate>)
                 (?srv, <rdf:type>, ?services)
USING     p FOR <http://localhost:8080/travelontology.owl#>
```

รูปที่ 2.34 ตัวอย่างการสืบค้นในเชิงหน้าที่ของบริการ

ส่วนต่อมาเป็นตัวอย่างการสืบค้นเชิงความหมายให้แก่บริการที่ผ่านการสืบค้นขั้นต้นมาแล้วตัวอย่างการสืบค้นด้วยภาษาอาร์ดีคิวแอลจะเป็นดังในรูปที่ 2.35

```
SELECT ?service, ?date
WHERE       (?service, <p:DepartureAirport>, p:BKK)
            (?service, <p:ArrivalAirport>, p:Hatyai)
            (?service, <p:ArrivalDate>, ?date)
              AND    ?date <  '15-08-05 19:40'
USING    p FOR <http://localhost:8080/travelontology.owl#>
```

รูปที่ 2.35 ตัวอย่างการสืบค้นที่ระบุความหมายของพารามิเตอร์

หากผู้ใช้มีการระบุความชอบไว้ในออนโทโลจีความชอบ เช่น สมมุติว่าในตัวอย่างนี้ผู้ใช้เคยระบุไว้ว่าชอบที่นั่งเครื่องบินแบบริมหน้าต่าง ระบบจะนำบริการทั้งหมดที่ผ่านขั้นตอนที่แล้วมาสืบค้นจากความชอบของผู้ใช้ดังตัวอย่างในรูปที่ 2.36

```
SELECT ?service
WHERE       (?service, <p:AirSeat>, p:Window)
USING    p FOR <http://localhost:8080/travelontology.owl#>
```

รูปที่ 2.36 ตัวอย่างการสืบค้นจากความชอบของผู้ใช้

จากตัวอย่างที่ได้กล่าวไปแล้วนี้เป็นตัวอย่างกรณีทดสอบแบบหนึ่งซึ่งกรณีทดสอบแบบอื่น ๆ จะถูกกล่าวถึงในบทต่อไป

ตัวอย่างบริการที่ใช้ทดสอบ (Services Example)

ตัวอย่างบริการที่จะใช้ในวิทยานิพนธ์นี้จะถูกนิยามขึ้นซึ่งจะระบุถึงโดเมนและพารามิเตอร์ต่าง ๆที่แต่ละบริการสามารถทำได้

ตารางที่ 2 ตัวอย่างบริการและพารามิเตอร์ที่ใช้ในการทดสอบ

Service Name	Domain	Parameter
AirlineService	Flight	DepartureAirport, ArrivalAirport,DepartureDate, ArrivalDate, AirSrvcClass, Airline
FlightBooking	Flight	DepartureAirport, ArrivalAirport, ArrivalDate, FlightType, DepartureDate, AirSrvcClass, Price
Air Travel	Flight	DepartureAirport, ArrivalAirport,DepartureDate, ArrivalDate, AirSeat, Airline
FlightService	Flight	DepartureAirport, ArrivalAirport, ArrivalDate DepartureDate, AirSrvcClass,FligtType

ตารางที่ 3 ตัวอย่างบริการและคุณสมบัติที่บริการสามารถให้ได้

Service Name	Airline	DptAirpt DptDate	ArvAirpt ArvDate	ClassType	Seat
AirlineService	OneTwo-GO	Bkk	Hatyai	First	Aisle
		15-08-05 18.00	15-08-05 19.40	-	
FlightBooking	Nok air	Bkk	Hatyai	Economy	Window
		15-08-05 19.00	15-08-05 20.35	Direct	
AirTravel	Air Asia	Bkk	Hatyai	Business	Non Smoking
		15-08-05 17.00	15-08-05 18.45	Nonstop	
FlightService	ThaiAirline	Bkk	Hatyai	First	Window
		15-08-05 13.00	15-08-05 14.50	Nonstop	

จาก ตารางที่ 2และ ตารางที่ 3 เป็นส่วนหนึ่งของบริการที่นำมาทดสอบในวิทยานิพนธ์นี้ ซึ่งรายละเอียดเพิ่มเติมของบริการที่ใช้จะกล่าวถึงในส่วนต่อไป

2.11 สรุป

ในบทนี้ได้กล่าวถึงหลักการพื้นฐานของเว็บเซอร์วิสและเว็บเชิงความหมายซึ่งในการแก้ปัญหาที่เกิดจากการค้นหาบริการในยูดีดีไอจำเป็นต้องใช้เทคโนโลยีเว็บเชิงความหมายเข้ามาช่วย ทำให้เกิดระบบที่เรียกว่าเว็บเซอร์วิสเชิงความหมาย (Semantic Web Service) คือ เว็บเซอร์วิสที่เน้นการอธิบายความหมายรายละเอียดของบริการที่เครื่องสามารถเข้าใจได้โดยสิ่งที่นอกเหนือจากเมต้าดาต้าของมันเอง โดยใช้การสร้างบริการเป็นออนโทโลจีทำให้สามารถค้นหาบริการได้อย่างถูกต้องตามความต้องการ

นอกจากการใช้เทคโนโลยีเว็บเชิงความหมายในการนิยามบริการให้มีรายละเอียดเชิงความหมายมากขึ้นแล้วซึ่งเป็นการแก้ปัญหาของฝั่งผู้ให้บริการแล้ว ยังจำเป็นต้องมีกระบวนการในการแก้ปัญหาการค้นหาบริการที่เกิดจากฝั่งผู้ใช้อีกด้วย ซึ่งเนื้อหาเกี่ยวกับวิธีการนิยามบริการด้วยออนโทโลจีและวิธีที่ช่วยเพิ่มประสิทธิภาพการค้นหาบริการให้สูงขึ้นจะถูกกล่าวอยู่ในบทที่ 3