



การเพิ่มประสิทธิภาพในระบบการคัดสรรบุคลากรอิเล็กทรอนิกส์

ด้วยเทคนิคการจับคู่เชิงความหมาย

The enhancement of efficiency in e-recruitment system using

semantic matching technique

ปองพล นิลพฤกษ์^{1*} และ รักษารัตน์ ขนานขาว¹

¹สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

อำเภอธัญบุรี จังหวัดปทุมธานี 12110

*E-mail: pongpon_n@rmUTT.ac.th

บทคัดย่อ

กว่าทศวรรษที่ผ่านมาอัตราการใช้งานอินเทอร์เน็ตมีแนวโน้มเพิ่มมากขึ้นอย่างรวดเร็ว ก่อให้เกิดผลกระทบโดยตรงต่อการเปลี่ยนแปลงอย่างมีนัยสำคัญต่อการคัดสรรบุคลากรอิเล็กทรอนิกส์ อย่างไรก็ตามการคัดสรรบุคลากรให้ตรงกับเป้าหมายที่เหมาะสมยังเป็นประเด็นที่ท้าทายอยู่เสมอ ในงานวิจัยนี้ได้ทำการศึกษการเพิ่มประสิทธิภาพในระบบการคัดสรรบุคลากรอิเล็กทรอนิกส์ด้วยเทคนิคการจับคู่เชิงความหมาย โดยนำเอาเทคโนโลยีเว็บเชิงความหมาย ฐานความรู้ออนโทโลยี และเทคนิคการจับคู่เชิงความหมายมาประยุกต์ใช้ โดยมีวัตถุประสงค์เพื่อออกแบบกรอบแนวคิดของระบบต้นแบบ ออกแบบต้นแบบฐานความรู้ออนโทโลยี และเพิ่มประสิทธิภาพในกระบวนการจับคู่เชิงความหมายด้วยเทคนิคกระบวนการหาความคล้ายคลึงกัน รวมถึงอัตราการเรียนรู้และรูปแบบผลรวมน้ำหนัก สุดท้ายผู้วิจัยได้นำกระบวนการจับคู่เชิงความหมายดังกล่าวมาทดสอบด้วยกลุ่มข้อมูลตัวอย่างผู้สมัครงานจำนวน 10 คน สรุปได้ว่าถ้าเปรียบเทียบกับวิธีการแบบจับคู่ด้วยพจน์และการหาความคล้ายคลึงกันแบบดั้งเดิมแล้ว กระบวนการจับคู่เชิงความหมายของงานวิจัยนี้สามารถแยกความแตกต่างได้มีประสิทธิภาพดีกว่า โดยวัดจากผลการคำนวณความเกี่ยวข้องของข้อมูลผู้สมัครกับความต้องการ ตลอดจนการนำเอาเทคโนโลยีเว็บเชิงความหมายมาประยุกต์ใช้ในระบบต้นแบบทำให้สารสนเทศสามารถถูกแบ่งปันในระบบอินเทอร์เน็ตได้ง่ายและเป็นมาตรฐานเดียวกัน อีกทั้งยังเป็นการลดความซ้ำซ้อนในการฝากประวัติส่วนตัว

Received: May 30, 2015

Revised: Jun 30, 2015

Accepted: Jun 30, 2015

ของผู้สมัครงานด้วยเช่นกัน แต่ถึงกระนั้นการออกแบบออนไลน์ที่มีประสิทธิภาพที่หลาย ประกอบกับการเพิ่มปัจจัยอื่นเสริมเข้าไปในกระบวนการจับคู่เชิงความหมาย สามารถเพิ่มประสิทธิภาพในการค้นหาบุคลากรได้เหมาะสมยิ่งขึ้น

คำสำคัญ: การคัดสรรบุคลากรอิเล็กทรอนิกส์, การจับคู่เชิงความหมาย, เทคโนโลยีเว็บเชิงความหมาย, ออนไลน์

Abstract

In the last ten years, global trend of the Internet access increasingly grows very fast. As a result, it significantly impacts on the changes of e-Recruitment process directly. However, finding the most suitable applicant is still very difficult and always challenges in a recruitment process. This research intends to study about the enhancement of efficiency in e-recruitment system using semantic matching technique. The Semantic Web technology, ontology knowledge base, and semantic matching technique are mainly used in this research in order to design the conceptual framework of the prototype system, design the prototype of ontology knowledge based, and improve the efficiency in matching process using the similarity and learning curve method. Moreover, Weight Sum Model (WSM) is also used to improve the efficiency in matching process. Finally, 10 sample data of applicants was tested on the proposed semantic matching, then the result expresses that the proposed semantic matching is more suitable than traditional terminological and similarity matching in case of distinguishing among applicants by calculated scores. The final scores from proposed semantic matching can nicely distinguish the differences among target applicants as well as the prototype system applying the Semantic Web technology can easily enable information sharing and reduce the duplicated information on the Internet because it is a standard. However, in future works, if we add more the good relationship of ontology or other factors, those can also enhance the efficiency in recruitment process to find more suitable person.

Keywords: e-Recruitment, Semantic Matching, Semantic Web Technology, Ontology

1. บทนำ

ในยุคดิจิทัล (digital age) อินเทอร์เน็ตมีผลกระทบโดยตรงต่อการเปลี่ยนแปลงอย่างมีนัยสำคัญในขั้นตอนการสมัครงานและการปรับปรุงช่องทางการสื่อสารระหว่างผู้สรรหาบุคลากร (recruiters) และผู้สมัครงาน (applicants) ในทุกวันนี้มีจำนวนผู้ใช้งานอินเทอร์เน็ตเพิ่มขึ้น

มากมายอย่างต่อเนื่อง [1] โดยเฉพาะอย่างยิ่งในกระบวนการสรรหาบุคลากรในหลายองค์กร ต่างนิยมนำมาใช้เป็นช่องทางในการรับสมัครงานเนื่องจากช่วยทำให้เกิดความสะดวกสบาย ลดต้นทุนและเวลาในการคัดเลือกผู้สมัคร ปัจจุบันทั้งผู้สรรหาบุคลากรและผู้สมัครต่างใช้อินเทอร์เน็ตเป็นเครื่องมือในกระบวนการนี้ [2] ดังนั้นจึงมีการ

นำข้อมูลเก็บไว้บนเว็บไซต์สมัครงานซึ่งมีอยู่หลายเว็บ ทั้งนี้เพื่อเป็นการเพิ่มโอกาสในการได้งานของผู้สมัคร ประกอบกับทำให้ผู้สรรหาบุคลากรมีตัวเลือกผู้สมัครมากขึ้น ทั้งคู่จึงฝากข้อมูลดังกล่าวไว้ในหลาย ๆ ที่ แต่อย่างไรก็ตาม ปัญหาที่ตามมาคือการปรับปรุงข้อมูลส่วนตัวหรือข้อมูลตำแหน่งงาน ที่จำเป็นต้องตามไปแก้ไข เนื่องจากข้อมูลได้กระจัดกระจายออกไป ส่งผลให้เสียเวลาในการปรับปรุงข้อมูล รวมทั้งรูปแบบโครงสร้างข้อมูลส่วนใหญ่ในแต่ละเว็บไซต์ไม่เหมือนกัน ทำให้มีปัญหาในเรื่องของการแลกเปลี่ยนหรือแบ่งปันข้อมูล ตลอดจนขาดความเชื่อมโยงของข้อมูล นอกจากนี้ระบบการค้นหาผู้สมัคร/ตำแหน่งงานส่วนมากยังใช้กลไกในการค้นหาข้อมูลด้วยคำสำคัญ (keyword) ซึ่งวิธีการค้นหาแบบนี้ เครื่องคอมพิวเตอร์ไม่สามารถทำความเข้าใจและประมวลความหมายหรือความสัมพันธ์ของข้อมูลได้อย่างตรงประเด็น ตัวอย่างเช่น ผู้สรรหาบุคลากรต้องการผู้สมัครที่มีความเชี่ยวชาญในด้านการเขียนโปรแกรมด้วยภาษา Java สมมุติว่ามีผู้สมัครอยู่ 2 คน คนหนึ่งมีประสบการณ์ในการเขียนโปรแกรมด้วยภาษา C# และ PHP ในส่วนผู้สมัครอีกคนหนึ่งซึ่งมีประสบการณ์ในการเขียนโปรแกรมด้วยภาษา C++ และ Ruby ในมุมมองของผู้สรรหาบุคลากร ผู้สมัครทั้ง 2 คน มีประสบการณ์และความเชี่ยวชาญใกล้เคียงกับความต้องการ ถ้าผู้สรรหาบุคลากรค้นหาข้อมูลผู้สมัครจากเว็บสมัครงานทั่วไป อาจจะไม่ได้พบข้อมูลของผู้สมัครทั้งคู่ เนื่องจากผลลัพธ์ที่ได้จากการค้นหาจากคำสำคัญนั้นไม่ตรงกับข้อมูลความเชี่ยวชาญของผู้สมัครทั้งสอง (Java ไม่เท่ากับ C# PHP C++ หรือ Ruby) ซึ่งกระบวนการค้นหาดังกล่าวถูกเรียกว่าการจับคู่ด้วยพจน์ (terminological matching) [3]สาเหตุเหล่านี้

นำมาซึ่งปัญหาของผู้สมัครงานและผู้ที่ต้องการหางานในการสืบค้นและการนำข้อมูลมาใช้ในการตัดสินใจเลือกตำแหน่งงานที่เหมาะสมกับตัวผู้สมัคร ดังนั้นถ้าคอมพิวเตอร์สามารถเข้าใจโครงสร้างความหมายและความสัมพันธ์ของคำค้นระบบจะสามารถนำเสนอข้อมูลที่เกี่ยวข้องและสอดคล้องกับความสนใจของผู้ใช้ได้มีประสิทธิภาพมากขึ้น

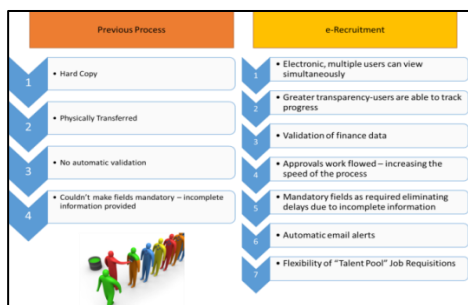
เทคโนโลยีเว็บเชิงความหมาย (Semantic Web Technology) ได้ถูกแนะนำอย่างกว้างขวางสำหรับนำมาประยุกต์ใช้งานในกระบวนการนิยามฐานความรู้ออนโทโลยี (ontology knowledge base) การค้นหาเชิงความหมาย (semantic searching) และมาตรฐานการแบ่งปันข้อมูลในเครือข่ายอินเทอร์เน็ต มีงานวิจัยหลายงานที่นำเทคโนโลยีดังกล่าวมาประยุกต์ใช้ในระบบสรรหาบุคลากรอิเล็กทรอนิกส์ (e-Recruitment) โดยเฉพาะอย่างยิ่งกระบวนการค้นหาผู้สมัครที่เหมาะสมที่สุด ยังเป็นประเด็นที่ทำให้อยู่เสมอ

ทั้งนี้จากปัญหาที่กล่าวมาข้างต้น ผู้วิจัยจึงได้นำแนวคิดของเทคโนโลยีเว็บเชิงความหมาย รวมถึงการสร้างฐานความรู้ออนโทโลยี เพื่อนำมาช่วยเพิ่มประสิทธิภาพในกระบวนการสรรหาและคัดเลือกบุคลากรแบบอิเล็กทรอนิกส์ โดยวัตถุประสงค์ประกอบด้วย การออกแบบกรอบแนวคิดของระบบต้นแบบบนพื้นฐานของเทคโนโลยีเว็บเชิงความหมาย การออกแบบต้นแบบฐานความรู้ออนโทโลยี การออกแบบกระบวนการจับคู่เชิงความหมาย (semantic matching) โดยใช้คุณสมบัติของฐานความรู้ออนโทโลยีที่ถูกสร้างขึ้น และสุดท้ายเป็นการนำกระบวนการจับคู่เชิงความหมายที่ถูกสร้างขึ้นมาทดสอบด้วยกลุ่มข้อมูลผู้สมัครตัวอย่าง

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบสรรหาบุคลากรอิเล็กทรอนิกส์

ระบบสรรหาบุคลากรอิเล็กทรอนิกส์ คือ กระบวนการสรรหาบุคคลซึ่งใช้ช่องทางเชิงอิเล็กทรอนิกส์ โดยทั่วไปแล้วกระบวนการนี้จะกระทำผ่านอินเทอร์เน็ต เพื่อให้เกิดประสิทธิภาพสูงที่สุดในแง่ของความเร็วในการตัดสินใจหาบุคคลที่เหมาะสมกับงาน จากรูปที่ 1 เป็นการเปรียบเทียบลักษณะการทำงานระหว่างกระบวนการแบบเก่ากับระบบสรรหาบุคลากรอิเล็กทรอนิกส์ [4] โดยกระบวนการแบบเก่าใช้วิธีการสื่อสารแบบใช้เอกสารจริง ส่งผ่านช่องทางการขนส่ง เช่น ทางไปรษณีย์ เป็นต้น รวมถึงยังไม่มีระบบการตรวจเช็คคำผิดของข้อมูล ตรงกันข้ามกับกระบวนการสรรหาบุคลากรเชิงอิเล็กทรอนิกส์ ที่ใช้ข้อมูลแบบออนไลน์ มีระบบตรวจเช็คคำผิด และผู้ใช้งานสามารถสื่อสารกันได้ตลอดเวลา เป็นต้น



รูปที่ 1 การเปรียบเทียบลักษณะการทำงานระหว่างกระบวนการแบบเก่ากับระบบสรรหาบุคลากรอิเล็กทรอนิกส์

จากข้อมูลงานวิจัยพบว่าแนวโน้มของอัตราการค้นหางานหรือการสรรหาบุคลากรทั้งภาคธุรกิจและภาครัฐ ได้ใช้สื่อกลางแบบออนไลน์เป็นหลัก โดยพบว่าช่องทางการหาบุคลากรออนไลน์ยอดนิยม ยังคงเป็นของเว็บไซต์หางาน

(job site) ตามมาด้วยการลงโฆษณาในเว็บไซต์ของบริษัท (company website) ส่วนสังคมออนไลน์อยู่เป็นอันดับ 3 [5]

2.2 เทคโนโลยีเว็บเชิงความหมาย

เทคโนโลยีเว็บเชิงความหมาย คือ ส่วนขยายของโครงข่ายบนเว็บ (World Wide Web) ในปัจจุบัน ที่ถูกพัฒนาขึ้นเพื่อให้คอมพิวเตอร์สามารถสืบค้น รวบรวม และประมวลผลสารสนเทศได้อย่างชาญฉลาด โดยใช้พื้นฐานรูปแบบของสารสนเทศเชิงความหมาย ซึ่งระบบคอมพิวเตอร์สามารถเข้าใจความหมายของนิยามต่าง ๆ ที่มนุษย์นิยามขึ้นได้ [6] คำว่าเทคโนโลยีเว็บเชิงความหมายถูกคิดค้นโดย Tim Berners-Lee โดยมาตรฐานที่เกี่ยวข้องเกิดจากร่วมมือขององค์กร W3C หรือ World Wide Web Consortium [7] นอกจากนี้การนิยามความหมายของสารสนเทศที่เหมาะสม (well-defined meaning) ยังเป็นส่วนสำคัญที่ทำให้ระบบคอมพิวเตอร์ประมวลได้อย่างถูกต้องอีกด้วย [8] การประยุกต์ใช้ประโยชน์จากเทคโนโลยีเว็บเชิงความหมาย มีส่วนสำคัญในการทำให้อินเทอร์เน็ตสามารถเชื่อมโยงความสัมพันธ์กันอย่างเป็นระบบ พร้อมทั้งสามารถอธิบายนิยามของสิ่งที่มนุษย์สนใจได้ดียิ่งขึ้น [9] ดังนั้นงานวิจัยนี้ จึงได้นำประโยชน์จากเทคโนโลยีดังกล่าวมาประยุกต์ใช้ในการออกแบบกรอบแนวคิดของระบบสรรหาบุคลากรอิเล็กทรอนิกส์ เพื่อเพิ่มประสิทธิภาพในการค้นหาคandidateให้ตรงกับความต้องการได้ดียิ่งขึ้น

2.3 ฐานความรู้ออนโทโลยี

ในศาสตร์ของวิทยาการคอมพิวเตอร์ ฐานความรู้ออนโทโลยี คือ การอธิบายรูปแบบโครงสร้างชนิด คุณสมบัติ และความสัมพันธ์ระหว่างข้อมูล ในขอบเขตที่ผู้ออกแบบสนใจ

(domain) โดยสามารถใช้โครงร่างพื้นฐานความสัมพันธ์ของเทอม (term) สำหรับเพื่อใช้เป็นฐานความรู้ได้ [10,11] ปกติแล้วฐานความรู้ออนโทโลยีสามารถถูกนิยามด้วยข้อความหรือภาษาใด ๆ ก็ได้ แต่อย่างไรก็ตามในมาตรฐานของเทคโนโลยีเว็บเชิงความหมาย (หัวข้อที่ 2.2) กรอบมาตรฐานของ Resource Description Framework (RDF) [12] และ Web Ontology Language (OWL) [13] ได้ถูกนำมาใช้เป็นมาตรฐานในการสร้างฐานความรู้ออนโทโลยี รวมถึงยังมีข้อดีในการแบ่งปันสารสนเทศได้ ดังนั้นการออกแบบออนโทโลยีในงานวิจัยนี้ จึงได้สร้างออนโทโลยีต้นแบบจาก RDF และ OWL เพื่อนิยามฐานความรู้และความสัมพันธ์ของข้อมูลที่เกี่ยวข้อง (กล่าวถึงในหัวข้อที่ 3.2)

2.4 งานวิจัยที่เกี่ยวข้อง

ในปัจจุบันเว็บไซต์หางานชั้นนำหลายเว็บไซต์ ได้แก่ JobsDB, JobTopGun, Craigslist, CareerBuilder, Indeed, และ Monster ต่างก็มีเครื่องมือที่ใช้ในการค้นหางาน ซึ่งพบกับปัญหาต่าง ๆ เช่น การหาคนที่เหมาะสม การใช้ข้อมูลร่วมกัน และการขาดมาตรฐานสำหรับการเก็บรักษาข้อมูล มีงานวิจัยหลากหลายงานที่พยายามแก้ปัญหาดังกล่าวโดยนำเทคนิคหลายเทคนิคมาใช้ร่วมกัน งานวิจัย [14] ได้กล่าวถึงการใช้เทคโนโลยีเว็บเชิงความหมาย ในขั้นตอนการสรรหาออนไลน์ให้มีประสิทธิภาพ งานวิจัยนี้แนะนำงานออกเป็น 3 ส่วนหลัก ส่วนแรกเสนอแนวคิดเชิงภาพรวมในการนำเทคโนโลยีเชิงความหมายมาช่วยสนับสนุนกระบวนการสรรหาบุคลากรแบบออนไลน์ในแต่ละขั้นตอนที่แตกต่างกัน ส่วนที่สองนำเสนอวิธีการทางด้านเทคนิค ซึ่งแสดงให้เห็นถึงการนำออนโทโลยีมาใช้ ได้แก่ วิธีการออกแบบ

โครงสร้างพื้นฐาน วิธีการเชื่อมโยงและบูรณาการข้อมูล สถาปัตยกรรมระบบที่พัฒนาขึ้นใช้ข้อกำหนดความต้องการตามสถานการณ์จริงในตลาดแรงงาน ตามบทบาทของผู้ที่มีส่วนเกี่ยวข้อง จุดเด่นที่สำคัญของงานวิจัยนี้ คือ การนำออนโทโลยีและเทคโนโลยีเว็บเชิงความหมายมาช่วยในขั้นตอนการสรรหาบุคลากร โดยเฉพาะการสร้างและพัฒนาเป็นออนโทโลยีสำหรับทรัพยากรบุคคล (HR ontology) โดยอ้างอิงตามมาตรฐานสากลในการสร้างรูปแบบของข้อมูลความต้องการทางด้านตำแหน่งงาน และข้อมูลประวัติส่วนตัวของผู้สมัครงาน รวมทั้งนำเสนอเทคนิคการคำนวณความคล้ายคลึงกันระหว่างความต้องการและผู้สมัคร โดยพิจารณาระยะทางระหว่างโหนด อย่างไรก็ตามงานวิจัยนี้ยังไม่ได้พูดถึงวิธีการคำนวณความคล้ายคลึงกันเชิงความหมายในองค์ประกอบของออนโทโลยีของโหนดที่อยู่ภายใต้บรรพบุรุษเดียวกัน ซึ่งเมื่อคำนวณหาความคล้ายคลึงกันออกมาแล้ว ค่าที่ได้จะมีค่าเท่ากันเสมอโดยไม่สามารถแยกความแตกต่างได้ รวมถึงยังไม่ได้กล่าวถึงช่องทางในการแบ่งปันข้อมูลให้กับระบบอื่น

ในงานวิจัย [15] ได้นำเสนอการแก้ปัญหาในเรื่องการนิยามข้อมูลที่ซ้ำซ้อน และแบ่งปันข้อมูลข้ามระบบได้ โดยมุ่งเน้นให้ผู้ที่มีงานสามารถแบ่งปันข้อมูลประวัติส่วนตัวของตัวเองให้กับระบบอื่น ผ่านมาตรฐาน “SPARQL Protocol and RDF Query Language” (SPARQL) แต่ถึงกระนั้น งานวิจัยดังกล่าวยังไม่ได้พูดถึงประเด็นการนำออนโทโลยีมาประยุกต์ใช้ในกระบวนการจับคู่เชิงความหมาย

หลังจากการศึกษางานวิจัย [16] เพิ่มเติมพบว่างานวิจัยดังกล่าว แนะนำแพลตฟอร์มการสรรหาบุคลากรอัจฉริยะโดยใช้ออนโทโลยีเชิง

ทักษะ (Skill ontology) และขั้นตอนวิธีในการจับคู่ระหว่างความต้องการกับเป้าหมายที่เป็นไปได้ โดยรูปของแบบจำลองออนโทโลยีเชิงทักษะประกอบด้วยโหนดทักษะ ที่มีเส้นเชื่อมทิศทางแสดงความสัมพันธ์ระหว่างโหนด แสดงในรูปของกราฟถ่วงน้ำหนัก (weighted graph) ซึ่งแทนปริมาณการสนับสนุนไปยังโหนดที่เกี่ยวข้อง (amount of contribution to related node) เช่น ถ้ามีประสบการณ์ด้านการเขียนโปรแกรมด้วยภาษา C จะมีปริมาณการสนับสนุนไปสู่ภาษา C++ ในระดับ 0.7 แต่ในทางกลับกัน ถ้ามีประสบการณ์ด้านการเขียนโปรแกรมด้วยภาษา C++ จะมีปริมาณการสนับสนุนภาษา C ในระดับแค่ 0.2 เนื่องจากภาษา C เป็นรากฐานของภาษา C++ ปริมาณการสนับสนุนไปสู่ภาษา C จึงมีน้อยกว่า (รู้ C++ ควรจะมีความรู้ภาษา C) อย่างไรก็ตามงานวิจัยนี้ ไม่ได้ระบุวิธีการได้มาของน้ำหนัก ดังนั้นผู้วิจัยจึงตั้งสมมติฐานว่า ความรู้ที่นิยามนั้นมีความถูกต้อง เช่น ผู้สมัครที่มีทักษะความชำนาญด้านภาษา C++ ควรจะมีความรู้ภาษา C เป็นต้น

โดยสรุปแล้ว ผู้วิจัยจึงได้นำแนวคิดจากงานวิจัยที่ได้ศึกษาดังกล่าว เพื่อเพิ่มประสิทธิภาพในระบบการคัดสรรบุคลากรอิเล็กทรอนิกส์ ซึ่งต่อไปจะเป็นขั้นตอนการดำเนินงานวิจัย

3. ขั้นตอนการดำเนินการวิจัย

ขั้นตอนการดำเนินงานวิจัยถูกแบ่งออกเป็น 4 ขั้นตอนหลัก ประกอบด้วย การออกแบบกรอบแนวคิดของระบบต้นแบบ (conceptual framework of prototype system design) การออกแบบต้นแบบฐานความรู้ออนโทโลยี (prototype of ontology knowledge base design) การออกแบบกระบวนการจับคู่เชิงความหมาย (semantic matching process design)

และการทดสอบด้วยกลุ่มข้อมูลตัวอย่างของผู้สมัครงาน (test by sample data of applicants)

3.1 การออกแบบกรอบแนวคิดของระบบต้นแบบ

กรอบแนวคิดของระบบต้นแบบในงานวิจัยนี้ ผู้วิจัยได้นำแนวคิดของงานวิจัยที่เกี่ยวข้อง รวมถึงประโยชน์ของเทคโนโลยีเว็บเชิงความหมายที่มีผู้วิจัยได้นำมาประยุกต์ใช้ที่แสดงได้ดังหัวข้อที่ 2.4 ซึ่งได้นำมาใช้สำหรับการนำเสนอกรอบแนวคิดของระบบต้นแบบ โดยแสดงได้ดังรูปที่ 2 จากรูปดังกล่าวได้แบ่งกรอบแนวคิดออกเป็น 3 ส่วนหลัก ๆ ได้แก่ ผู้สมัครงาน กลไกการจับคู่ (matching engine) และผู้สรรหาบุคลากร ซึ่งมีรายละเอียดดังต่อไปนี้

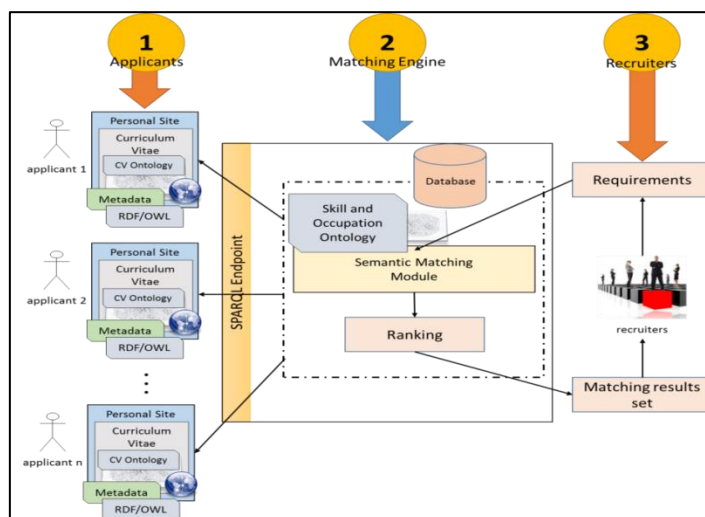
3.1.1 ผู้สมัครงาน

ส่วนของผู้สมัครงาน (ส่วนที่ 1) เป็นส่วนงานที่ผู้สมัครงานมีความเกี่ยวข้องเป็นหลัก ซึ่งในที่นี้ผู้สมัครงานจำเป็นต้องสร้างประวัติของตนเองในรูปแบบ RDF/OWL ตลอดจนต้องสอดคล้องกับมาตรฐานของ Europortfolio Europass [17], ISCED 2013 [18], ISCO 88 [19] และ NACE [20] ที่ถูกนำมาใช้เป็นกรอบมาตรฐานสำหรับประวัติส่วนตัวโดยสังเขป โดยอ้างอิงจากงานวิจัยที่เกี่ยวข้อง (หัวข้อที่ 2.4) การสร้างประวัติส่วนตัวเช่นนี้ ทำให้สารสนเทศของผู้สมัครงานถูกแบ่งปันได้ตามแนวคิดของ Linked Open Data (LOD) [21]

ทั้งนี้เพื่อเป็นการลดความซ้ำซ้อนในการฝากประวัติส่วนตัวของผู้สมัครงาน จากกรอบแนวคิดดังกล่าวผู้สมัครต้องฝากประวัติส่วนตัวไว้ที่ไซต์ส่วนตัว (personal site) หรือตำแหน่งที่สามารถระบุช่องทางในการเข้าถึงข้อมูลได้โดยทั่วไปตำแหน่งดังกล่าวถูกเรียกในเชิงเทคนิคว่า SPARQL Endpoint [22] เพื่อเปิดช่องทางให้กับ

ระบบอื่น ๆ มาเรียกดูข้อมูลได้ ดังนั้นการปรับปรุงข้อมูลของผู้สมัครงานสามารถกระทำได้ ณ ตำแหน่งนี้ที่เดียว ซึ่งเป็นประโยชน์อย่างมากต่อการลดความซ้ำซ้อนของข้อมูลผู้สมัครงาน เนื่องจากลักษณะการสมัครงานของระบบรับสมัครงานออนไลน์โดยทั่วไป ผู้สมัครงานจำเป็นต้องกรอกประวัติส่วนตัวทุกไว้ในระบบนั้น ซึ่งโดยส่วนใหญ่แล้วผู้สมัครจะฝากข้อมูลไว้ในระบบรับสมัครงานออนไลน์มากกว่า 1 ระบบขึ้นไป เพื่อเพิ่มโอกาสให้ตนเองได้รับความสนใจต่อผู้สรรหาบุคลากรมากยิ่งขึ้น ส่งผลให้ข้อมูลประวัติส่วนตัว

นั้นกระจัดกระจายไปยังระบบรับสมัครงานต่าง ๆ ถ้าหากผู้สมัครมีเหตุจำเป็นต้องปรับปรุงข้อมูลส่วนตัว ผู้สมัครจำเป็นต้องเข้าไปปรับปรุงในทุก ๆ ระบบที่ผู้สมัครได้ฝากข้อมูลไว้ ด้วยเหตุนี้กรอบแนวคิดในส่วนของผู้สมัครงานที่ผู้วิจัยได้นำเสนอ จึงสามารถแก้ปัญหาความซ้ำซ้อนของข้อมูลผู้สมัครงาน ตลอดจนสามารถแบ่งปันสารสนเทศให้กับระบบภายนอกให้เข้าถึงข้อมูลตามมาตรฐานที่ได้กล่าวไว้แล้ว



รูปที่ 2 กรอบแนวคิดของระบบต้นแบบ

3.1.2 กลไกการจับคู่

ในส่วนของกลไกการจับคู่ (ส่วนที่ 2) ผู้วิจัยได้ออกแบบกรอบแนวคิดโดยเน้นในส่วนของการกระบวนการจับคู่ผู้สมัครงานให้ตรงกับความต้องการของผู้สรรหาบุคลากรให้ได้มากที่สุด ในส่วนนี้ประกอบด้วย

- ออนโทโลยี สำหรับเก็บความสัมพันธ์ของข้อมูลที่เกี่ยวข้องกันซึ่งรายละเอียดได้ถูกอธิบายในหัวข้อที่ 3.2

- ฐานข้อมูล สำหรับจัดเก็บข้อมูลที่จะนำมาใช้ในการคำนวณหาความสัมพันธ์ระหว่างความต้องการของผู้สรรหาบุคลากรกับข้อมูลของผู้สมัครงาน เช่น ตำแหน่งที่ต้องการ ทักษะที่ต้องการ และข้อมูลประสบการณ์ เป็นต้น

- กระบวนการจับคู่เชิงความหมาย ในที่นี้ถูกอธิบายในหัวข้อที่ 3.3

- การจัดเรียงลำดับ เป็นขั้นตอนสำหรับการจัดเรียงรายการของผู้สมัครงานที่ตรงกับความต้องการของผู้สรรหา โดยการจัดเรียงจะ

เรียงจากลำดับเป้าหมายที่ใกล้กับความต้องการมากที่สุด ถูกอธิบายในหัวข้อที่ 3.3 เช่นเดียวกัน

3.1.3 ผู้สรรหาบุคลากร

ส่วนผู้สรรหาบุคลากร (ส่วนที่ 3) เป็นส่วนสำหรับป้อนความต้องการในการค้นหาผู้สมัครงานที่เป็นเป้าหมายเข้าไปในระบบ ถือว่าเป็นส่วนหลักสำหรับผู้สรรหาบุคลากร และความต้องการดังกล่าวจะถูกนำไปใช้ในกระบวนการจับคู่เชิงความหมายต่อไป

ตารางที่ 1 มาตรฐานที่ใช้สำหรับ CV Ontology

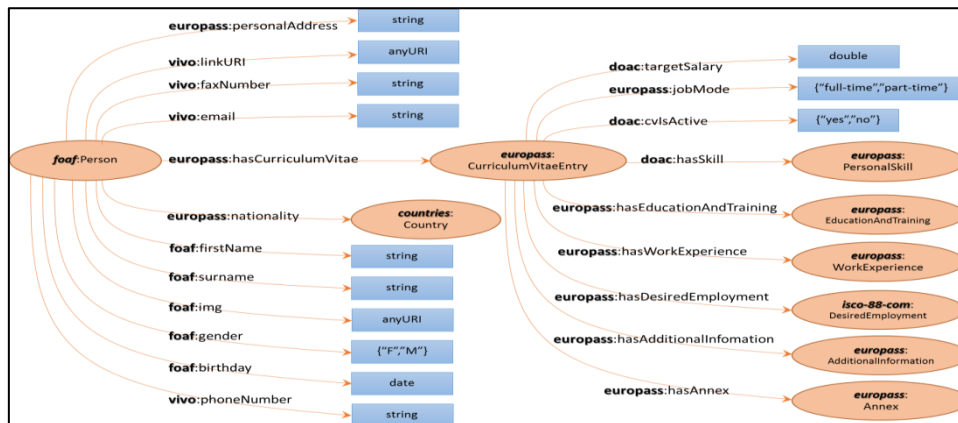
มาตรฐาน	รายละเอียด
Europortfolio Europass	มาตรฐานสำหรับการออกแบบประวัติส่วนตัวโดยสังเขปของผู้สมัครงาน หรือโดยทั่วไปถูกเรียกว่า “CV”
ISCED 2013	มาตรฐานสากลสำหรับการจัดประเภทของสาขาการศึกษา (International Standard Classification of Education) ถูกนำมาใช้เพื่อจัดประเภทของสาขาการศึกษาให้เป็นมาตรฐานเดียวกัน
ISCO 88	มาตรฐานสากลสำหรับการจัดประเภทของอาชีพ (International Standard Classification of Occupations) ถูกนำมาใช้เพื่อจัดประเภทอาชีพที่เป็นเป้าหมายของผู้สมัครงานและผู้สรรหาบุคลากร
NACE	มาตรฐานยุโรปสำหรับการจัดประเภทของอุตสาหกรรม (European Standard for Industry Classifications) ถูกนำมาใช้เพื่อจัดประเภทของอุตสาหกรรมต่าง ๆ ให้เป็นมาตรฐานเดียวกัน

จากมาตรฐานที่ได้กล่าวมา ผู้วิจัยได้ออกแบบต้นแบบออนโทโลยีสำหรับ CV Ontology ให้สอดคล้องกับมาตรฐานดังกล่าว โดยโดยภาพรวมของต้นแบบออนโทโลยีสำหรับ CV Ontology แสดงได้ดังรูปที่ 3 ส่วนที่ผู้วิจัยได้นำมาขกตัวอย่างในการออกแบบกระบวนการจับคู่เชิงความหมาย เป็นรายละเอียดของทักษะส่วนบุคคล (personal skill) และอาชีพเป้าหมาย (desired employment) ของผู้สมัครงาน ดังนั้นจึงได้นำฐานความรู้เกี่ยวกับความสัมพันธ์ของทักษะรวมถึงอาชีพ มาประยุกต์ใช้ในกระบวนการจับคู่เชิงความหมายเพื่อให้ได้ประสิทธิภาพสูงสุด ขั้นตอนต่อมา ผู้วิจัยได้ออกแบบต้นแบบออนโทโลยี

3.2 การออกแบบต้นแบบฐานความรู้ออนโทโลยี

ในส่วนของฐานความรู้ออนโทโลยี ผู้วิจัยได้แบ่งออกเป็น 2 ส่วนหลัก ได้แก่ต้นแบบออนโทโลยีสำหรับผู้สมัครงาน (หัวข้อที่ 3.1.1) และออนโทโลยีสำหรับเก็บความสัมพันธ์ของทักษะรวมทั้งประเภทอาชีพ ซึ่งต้นแบบออนโทโลยีสำหรับผู้สมัครงานในงานวิจัยนี้ถูกเรียกว่า “CV Ontology” ใช้มาตรฐานตามตารางที่ 1

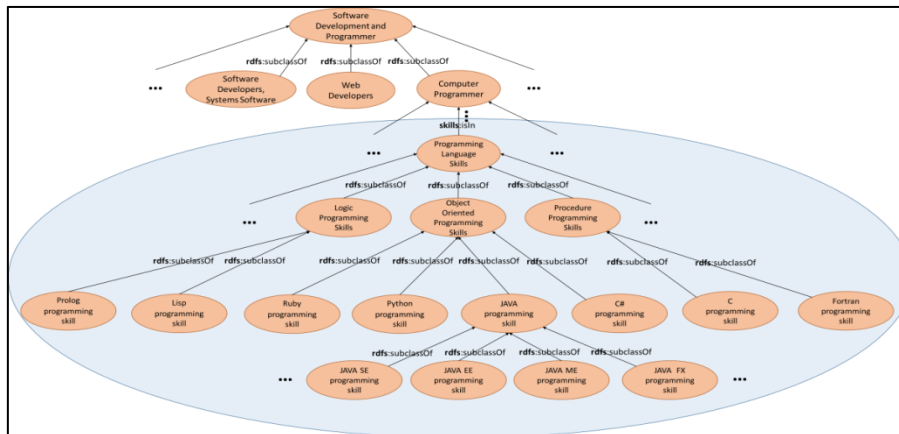
สำหรับเก็บความสัมพันธ์ของทักษะรวมทั้งอาชีพ ซึ่งถูกเรียกว่า “Skill and Occupation Ontology” เพื่อใช้เป็นฐานความรู้ในกระบวนการจับคู่เชิงความหมาย มาตรฐานที่ถูกนำมาใช้ในการออกแบบคือ ISCO 88 (ดูได้จากตารางที่ 1) และความสามารถที่เกี่ยวข้องกับอาชีพในด้านเทคโนโลยีสารสนเทศ (IT Occupational Competencies) [23] โดยจำกัดขอบเขตในศาสตร์ของสาขาวิชาวิทยาการคอมพิวเตอร์มาสร้างเป็นต้นแบบเพื่อทดสอบกระบวนการทำงานเท่านั้น หลังจากนั้นได้เพิ่มข้อมูลในส่วนของพื้นฐานการเขียนโปรแกรมในรูปแบบต่าง ๆ เพื่อใช้ในการทดสอบต่อไป



รูปที่ 3 ภาพรวมของต้นแบบออนโทโลยีสำหรับ CV Ontology

ตัวอย่างต้นแบบของ Skill and Occupation Ontology บางส่วนแสดงได้ดังรูปที่ 4 โดยแบ่งชนิดของกรอบความคิด (concept) ออกเป็น 2 ชนิด ประกอบไปด้วยอาชีพและทักษะที่มีความสัมพันธ์

กัน โดยตัวอย่างที่ผู้วิจัยได้นำไปใช้ในการอธิบายกระบวนการจับคู่เชิงความหมายอยู่ในส่วนออนโทโลยีย่อย (sub-ontology) ของ Programming Language Skills



รูปที่ 4 ต้นแบบของ Skill and Occupation Ontology บางส่วน

3.3 การออกแบบกระบวนการจับคู่เชิง

ความหมาย

ในกระบวนการจับคู่เชิงความหมาย ผู้วิจัยได้มาจากแนวคิดการจับคู่กราฟสำหรับการค้นหาเชิงความหมาย [24] มาประยุกต์ใช้ เริ่มต้นจากกระบวนการหาความคล้ายคลึงกัน (similarity) ระหว่างสองกรอบความคิด แสดงได้ดังสมการที่ (1):

$$\text{sim}_c(C_1, C_2) = 1 - \text{dis}_c(C_1, C_2);$$

where C = concept, sim = similarity, dis = distance,

$$\text{dis}_c(C_1, C_2) \leq 1 \quad \text{สมการที่ (1)}$$

ค่าของ C_1 และ C_2 หาได้จากความลึกของโหนด ในความสัมพันธ์ของทรีย่อย (sub-tree) แบบ superclass และ subclass โดยค่านี้ถูกเรียกว่า milestone แสดงได้ดังสมการที่ (2):

$$\text{milestone}(n) = \frac{1/2}{k^{l(n)}}$$

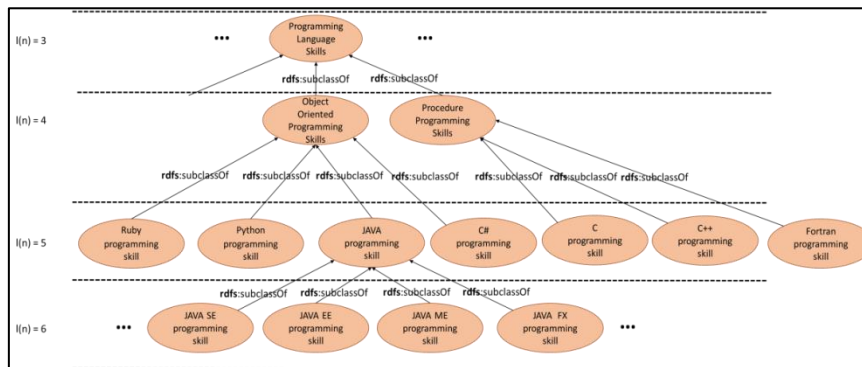
where k = decreased factor = 2, $l(n)$ = depth of

node

สมการที่ (2)

จากสมการที่ 2 ค่า k คือค่าที่บ่งบอกถึง factor ในการลดลงตามลำดับชั้นของทรี ซึ่งงานวิจัยนี้กำหนดเป็นค่าคงที่ไว้ที่ 2 และ $l(n)$ คือระดับความลึกของโหนด n ในลำดับชั้นของทรี ตัวอย่างเช่น ความลึกของ root ในชั้นทรีที่สนใจได้เป็น $l(\text{root}) = 0$ และระดับความลึกถัดไปจะเพิ่มขึ้น

ทีละ 1 สมมติกำหนดให้ C_R คือความต้องการของผู้สรรหาบุคลากร C_A คือข้อมูลที่ได้จากผู้สมัครงาน สมมุติ $C_R = \text{"C\# programming skill"}$ และ $C_A = \text{"C programming skill"}$ จากรูปที่ 5 สังเกตได้ว่า "C\#" และ "C programming skill" มีโหนดบรรพบุรุษร่วมกัน (co-parent node) คือ "Programming Language skill" ดังนั้นการคำนวณหาความคล้ายคลึงกันสามารถทำได้โดยหาความแตกต่างระหว่าง C_R และ C_A ก่อน โดยค่าความแตกต่างนี้จะมีค่าไม่เกิน 1 ซึ่งแสดงการคำนวณได้ดังต่อไปนี้:



รูปที่ 5 ส่วนของออนโทโลยีสำหรับทดสอบการหาความคล้ายคลึงกัน

$$\begin{aligned} \text{dis}_C(\text{C\# programming skill, C programming skill}) &= \text{dis}_C(\text{Programming Language skill, C\# programming skill}) \\ &\quad + \text{dis}_C(\text{Programming Language skill, C programming skill}) \\ &= (1/8 - 1/32) + (1/8 - 1/32) \\ &= 0.09375 + 0.09375 = 0.1875 \end{aligned}$$

จากนั้นนำค่าความแตกต่างดังกล่าวมาหาค่าความคล้ายคลึงกัน ได้ค่าเป็น:

$$\text{sim}_C(\text{C\# programming skill, C programming skill}) = 1 - 0.1875 = 0.8125$$

ถ้าเทียบความคล้ายคลึงกันระหว่างโหนดที่อยู่ระดับลึกซึ่งเปรียบเทียบกับกันแล้วจะมีความคล้ายคลึงมากกว่าโหนดในระดับบน ตัวอย่างเช่น $\text{sim}_C(\text{C programming skill, Fortran programming skill})$ ซึ่งได้ค่าเป็น 0.9375 จะมีค่าน้อยกว่า $\text{sim}_C(\text{JAVA SE programming skill, JAVA EE programming skill})$ ซึ่งได้ค่าเป็น 0.96875 เหตุผลที่

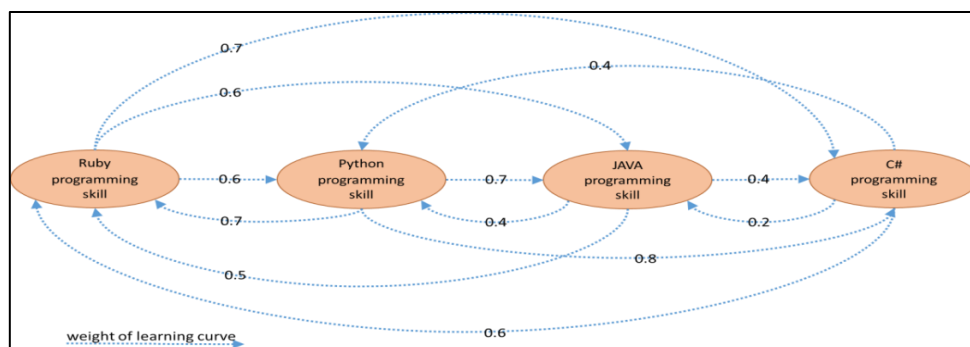
เป็นเช่นนี้เนื่องจากลักษณะทรีที่เป็น superclass และ subclass ยิ่งระดับโหนดลึกเท่าใดจะมีความเป็นลักษณะเฉพาะมากขึ้นเท่านั้น ด้วยเหตุนี้การเปรียบเทียบความคล้ายคลึงกันจะสมเหตุสมผลตามลักษณะโครงสร้างแบบลำดับชั้น

แต่อย่างไรก็ตามวิธีการดังกล่าวยังไม่สามารถบอกความแตกต่างระหว่างข้อมูลที่อยู่ในระดับเดียวกันตามที่ได้กล่าวไว้ในงานวิจัยที่เกี่ยวข้อง ตัวอย่างเช่น ถ้าผู้สรรหางานมีความต้องการ "Ruby" โดยสมมุติผู้สมัครมีอยู่ 2 คน เมื่อคำนวณค่าความคล้ายคลึงกันระหว่างข้อมูลจะได้

ว่า $\text{sim}_C(\text{Ruby programming skill, JAVA programming skill})$ สำหรับผู้สมัครคนที่ 1 และ $\text{sim}_C(\text{Ruby programming skill, C\# programming skill})$ สำหรับผู้สมัครคนที่ 2 ซึ่งทั้งสองคนได้คะแนนค่าความคล้ายคลึงกันเท่ากันพอดี จึงไม่สามารถแยกความแตกต่างได้ในกรณีนี้

ดังนั้นผู้วิจัยจึงได้ประยุกต์ใช้แนวคิดโดยได้กำหนดค่าเพิ่มเติมคือ อัตราการเรียนรู้ (learning curve) จากแนวคิดของงานวิจัยที่เกี่ยวข้องในการกำหนดปริมาณการสนับสนุนไปยังโหนดที่เกี่ยวข้อง ซึ่งในแต่ละโหนดจะมีอัตราการเรียนรู้ไปสู่โหนดอื่น ๆ แตกต่างกันไป เช่น ถ้ามีประสบการณ์ด้านการเขียนโปรแกรมด้วยภาษา C จะมีอัตราการเรียนรู้ภาษา C++ เพิ่มขึ้นในระดับ 0.7 แต่ในทางกลับกัน ถ้ามีประสบการณ์ด้านการ

เขียนโปรแกรมด้วยภาษา C++ จะมีอัตราการเรียนรู้ภาษา C ในระดับ 0.2 ซึ่งใช้เวลาในการเรียนรู้เพิ่มเติมน้อยกว่า เนื่องจากภาษา C เป็นรากฐานของภาษา C++ [25] เพราะฉะนั้นการเรียนรู้จาก C++ ไป C จะใช้เวลาเรียนรู้มากกว่า ด้วยเหตุนี้ผู้วิจัยจึงได้กำหนดค่าอัตราการเรียนรู้ในแต่ละโหนดเพิ่มเติม ค่าที่ได้เกิดจากการรวบรวมข้อมูลจากโปรแกรมเมอร์จำนวน 20 คน ที่มีความเชี่ยวชาญแตกต่างกัน โดยใช้ลักษณะการสอบวัดผลเชิงทักษะเพื่อหาอัตราการเรียนรู้ แต่ละค่าจะมีค่าไม่เกิน 1 และถ้าเป็นโหนดเดียวกันจะมีค่าเป็น 0 ซึ่งในที่นี้ยกตัวอย่างความสัมพันธ์ในทรีออยของ “Object Oriented Programming Skills” แสดงได้ดังรูปที่ 6 เพื่อนำมาทดสอบกระบวนการในขั้นตอนถัดไป



รูปที่ 6 ความสัมพันธ์ของอัตราการเรียนรู้ในทรีออยของ “Object Oriented Programming Skills”

จากตัวอย่างก่อนหน้านี้ ถ้าเปรียบเทียบ $\text{sim}_C(\text{Ruby programming skill, JAVA programming skill})$ กับ $\text{sim}_C(\text{Ruby programming skill, C\# programming skill})$ จะได้ค่าเป็น 0.9375 ซึ่งไม่สามารถเลือกได้ว่าผู้สมัครคนใดเหมาะสมที่สุดจากที่ได้กล่าวไปแล้ว ผู้วิจัยจึงใช้รูปแบบผลรวมน้ำหนัก (WSM, Weight Sum Model) [26] เพื่อกำหนดค่าบรรทัดฐาน (criteria) เพิ่มขึ้น ในที่นี้ผู้วิจัยได้เพิ่มบรรทัดฐานของอัตราการเรียนรู้

เพื่อคำนวณหาผู้สมัครงานที่เหมาะสม แสดงได้ดังสมการที่ (3):

$$\text{suit}_C(C_R, C_A) = W_{\text{sim}} * \text{sim}_C(C_R, C_A) + W_{\text{learning_curve}} * (1 - \text{learning_curve_to}(C_A, C_R));$$

where $W_{\text{sim}} + W_{\text{learning_curve}} = 1.0$

สมการที่ (3)

สมมุติกำหนดให้ $W_{\text{sim}} = 0.5$ และ $W_{\text{learning_curve}} = 0.5$ ซึ่งค่าทั้งสองเป็นการกำหนดน้ำหนักความสำคัญให้กับความคล้ายคลึงและ

อัตราการเรียนรู้ตามลำดับ ค่าดังกล่าวไม่ได้ถูกกำหนดไว้ตายตัว แต่ขึ้นอยู่กับผู้ที่นำไปประยุกต์ใช้ต่อไป เช่น ค่าเหล่านี้อาจถูกป้อนโดยตรงจาก

ผู้ใช้งานระบบ เป็นต้น ดังนั้นเมื่อนำสมการที่ 3 มาคำนวณเพิ่มเติมจากขั้นตอนที่ผ่านมา ได้เป็น:

$$\begin{aligned} \text{suit}_c(\text{Ruby programming skill, JAVA programming skill}) &= 0.5 * 0.9375 + 0.5 \\ &\quad * [1 - \text{learning_curve_to}(\text{JAVA programming skill, Ruby programming skill})] \\ &= 0.5 * 0.9375 + 0.5 * 0.5 = 0.71875 \\ \text{suit}_c(\text{Ruby programming skill, C\# programming skill}) &= 0.5 * 0.9375 + 0.5 \\ &\quad * [1 - \text{learning_curve_to}(\text{C\# programming skill, Ruby programming skill})] \\ &= 0.5 * 0.9375 + 0.5 * 0.4 = 0.66875 \end{aligned}$$

จากตัวอย่างดังกล่าวใช้อัตราการเรียนรู้เพิ่มเติมเข้ามาเป็นบรรทัดฐานทำให้ผู้สมัครงานที่มีข้อมูล JAVA programming skill เหมาะสมมากกว่าจากคะแนนที่คำนวณได้ แต่ถึงกระนั้นความต้องการของผู้สรรหาและข้อมูลของผู้สมัครย่อมมีมากกว่าหนึ่ง ดังนั้นผู้วิจัยจึงกำหนดเป็นลักษณะของเซตข้อมูล แสดงได้ดังชุดต่อไปนี้:

$$\begin{aligned} R &= \{r_1, r_2, r_3, r_4, \dots, r_n\} \text{ where } r_i = \text{recruiter's requirement} \\ A &= \{a_1, a_2, a_3, a_4, \dots, a_n\} \text{ where } a_i = \text{applicant's profile} \end{aligned}$$

จากนั้นนำมาคำนวณเพิ่มเติม เพื่อจับคู่โดยหาค่าที่เหมาะสมที่สุดจากเซต R และ A โดยใช้สมการที่ (3) แสดงได้เป็นเมทริกซ์ต่อไปนี้:

$$\begin{aligned} \text{SUIT}_{\text{set}}(R, A) &= \\ \begin{bmatrix} \max[\text{suit}_c(r_1, a_1) & \text{suit}_c(r_1, a_2) & \dots & \text{suit}_c(r_1, a_n)] \\ \max[\text{suit}_c(r_2, a_1) & \text{suit}_c(r_2, a_2) & \dots & \text{suit}_c(r_2, a_n)] \\ \dots & \dots & \dots & \dots \\ \max[\text{suit}_c(r_m, a_1) & \text{suit}_c(r_m, a_2) & \dots & \text{suit}_c(r_m, a_n)] \end{bmatrix} &= \\ \begin{bmatrix} \text{suit}_{c_max}(r_1, a_{s1}) \\ \text{suit}_{c_max}(r_2, a_{s2}) \\ \dots \\ \text{suit}_{c_max}(r_m, a_{sn}) \end{bmatrix} \end{aligned}$$

เซต $\text{SUIT}_{\text{set}}(R, A) = \{\text{suit}_{c_max}(r_1, a_{s1}), \text{suit}_{c_max}(r_2, a_{s2}), \dots, \text{suit}_{c_max}(r_m, a_{sn})\}$ เป็นเซตคู่ของ r_i และ a_i ที่ถูกจับคู่แล้วได้ค่าสูงที่สุดในแต่ละแถวของเมทริกซ์ แต่อย่างไรก็ตาม ความสำคัญของเซตความต้องการของผู้สรรหาย่อมแตกต่างกัน ดังนั้นผู้วิจัยจึงเสนอการเพิ่มน้ำหนักในแต่ละความ

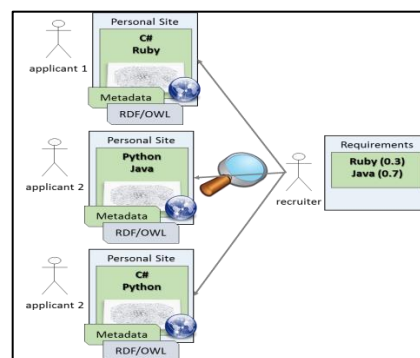
ต้องการของผู้สรรหา เพื่อให้ผู้สรรหาได้จัดลำดับความสำคัญของเซตความต้องการ ได้เป็นผลรวมสุดท้ายสำหรับคะแนนของผู้สมัครเป็นรายบุคคล แสดงได้ดังสมการที่ (4):

$$\begin{aligned} \text{SCORE}_{\text{applicant}}(\text{SUIT}_{\text{set}}(R, A)) &= \sum_{i=1}^n (W_i * \text{suit}_{c_max}(r_i, a_{s_i})); \\ \text{where } \sum_{i=1}^n (W_i) &= 1.0 \end{aligned}$$

สมการที่ (4)

ขั้นตอนต่อไปเป็นการทดสอบกระบวนการที่ผ่านมาด้วยกลุ่มข้อมูลตัวอย่างเพื่อทดสอบกระบวนการจับคู่เชิงความหมายที่ได้ออกแบบไว้

3.4 การทดสอบด้วยกลุ่มข้อมูลตัวอย่างของผู้สมัครงาน



รูปที่ 7 แผนภาพแสดงกลุ่มข้อมูลตัวอย่างสำหรับทดสอบกระบวนการจับคู่เชิงความหมาย

จากรูปที่ 7 สมมติมีผู้สมัครงานอยู่ทั้งหมด 3 คน มีข้อมูลแบ่งเป็นเซต $A_1 = \{\text{"C\#"}, \text{"Ruby"}\}$ $A_2 = \{\text{"Python"}, \text{"Java"}\}$ และ $A_3 =$

{“C#”, “Python”} ส่วนความต้องการของผู้สรรหาบุคลากรได้เป็นเซต $R_1 = \{\text{“Ruby”, “Java”}\}$ และนำมาเข้าสมการเพื่อหา $SUIT_{set}(R_i, A_i)$ ในแต่ละคู่ แสดงได้ดังตารางที่ 2

ต้องการแตกต่างกัน น้ำหนักของ “Ruby” และ “Java” คือ 0.3 และ 0.7 ตามลำดับ ซึ่งสรุปคะแนนและเรียงลำดับจากมากไปน้อยได้ดังตารางที่ 3

สุดท้ายเป็นกระบวนการหาคะแนนของผู้สมัครเป็นรายบุคคลโดยใช้สมการที่ (4) โดยในที่นี้ผู้สรรหาบุคลากรมีส่วนน้ำหนักของความ

ตารางที่ 2 ข้อมูลสรุปในการหาค่า $SUIT_{set}(R_i, A_i)$ ของกลุ่มตัวอย่างข้อมูลสำหรับทดสอบกระบวนการจับคู่เชิงความหมาย

R_i vs A_i	$sim_c(C_{R_i}, C_{A_i})$	$suit_c(R_i, A_i)$	$SUIT_{set}(R_i, A_i)$
R_1 vs A_1	$sim_c(\text{“Ruby”, “C#”}) = 0.9375$ $sim_c(\text{“Ruby”, “Ruby”}) = 1.0$ $sim_c(\text{“Java”, “C#”}) = 0.9375$ $sim_c(\text{“Java”, “Ruby”}) = 0.9375$	$suit_c(\text{“Ruby”, “C#”}) = 0.66875$ $suit_c(\text{“Ruby”, “Ruby”}) = 1.0$ $suit_c(\text{“Java”, “C#”}) = 0.86875$ $suit_c(\text{“Java”, “Ruby”}) = 0.66875$	$suit_{c_max}(\text{“Ruby”, “Ruby”}) = 1.0$ $suit_{c_max}(\text{“Java”, “C#”}) = 0.86875$
R_1 vs A_2	$sim_c(\text{“Ruby”, “Python”}) = 0.9375$ $sim_c(\text{“Ruby”, “Java”}) = 0.9375$ $sim_c(\text{“Java”, “Python”}) = 0.9375$ $sim_c(\text{“Java”, “Java”}) = 1.0$	$suit_c(\text{“Ruby”, “Python”}) = 0.61875$ $suit_c(\text{“Ruby”, “Java”}) = 0.71875$ $suit_c(\text{“Java”, “Python”}) = 0.61875$ $suit_c(\text{“Java”, “Java”}) = 1.0$	$suit_{c_max}(\text{“Ruby”, “Java”}) = 0.71875$ $suit_{c_max}(\text{“Java”, “Java”}) = 1.0$
R_1 vs A_3	$sim_c(\text{“Ruby”, “C#”}) = 0.9375$ $sim_c(\text{“Ruby”, “Python”}) = 0.9375$ $sim_c(\text{“Java”, “C#”}) = 0.9375$ $sim_c(\text{“Java”, “Python”}) = 0.9375$	$suit_c(\text{“Ruby”, “C#”}) = 0.66875$ $suit_c(\text{“Ruby”, “Python”}) = 0.61875$ $suit_c(\text{“Java”, “C#”}) = 0.86875$ $suit_c(\text{“Java”, “Python”}) = 0.61875$	$suit_{c_max}(\text{“Ruby”, “C#”}) = 0.66875$ $suit_{c_max}(\text{“Java”, “C#”}) = 0.86875$

ตารางที่ 3 ค่าคะแนนของผู้สมัครเป็นรายบุคคลเรียงจากมากไปน้อย

R_i vs A_i	$SCORE_{applicant}(SUIT_{set}(R_i, A_i))$
R_1 vs A_2	$(0.3)(0.71875) + (0.7)(1.0) = 0.915625$
R_1 vs A_1	$(0.3)(1.0) + (0.7)(0.86875) = 0.908125$
R_1 vs A_3	$(0.3)(0.66875) + (0.7)(0.86875) = 0.80875$

เมื่อทำการทดสอบด้วยข้อมูลตัวอย่างผู้สมัครงานจำนวน 10 คน (ใช้ข้อมูลและน้ำหนักความต้องการจากเซต R_1 ในตัวอย่างการทดสอบที่ผ่านมา) นำมาเปรียบเทียบกับวิธีการแบบจับคู่ด้วยพจน์ การหาความคล้ายคลึงกันแบบดั้งเดิม และค้นแบบกระบวนการจับคู่เชิงความหมายของงานวิจัยนี้ โดยเน้นเฉพาะทักษะด้านการเขียน

โปรแกรมคอมพิวเตอร์ แสดงได้ดังตารางที่ 4 สังกัดได้ว่าเทคนิคการจับคู่ด้วยพจน์ ถ้าพจน์หรือคำค้นไม่ตรงกับข้อมูล ระบบจะไม่สามารถตรวจพบข้อมูลได้เลย ในทางตรงกันข้าม เทคนิคการหาความคล้ายคลึงกันสามารถค้นคืนข้อมูลในระบบได้ทั้งหมดและคืนค่าเป็นลักษณะค่าความคล้ายคลึงกันในแต่ละพจน์

ตารางที่ 4 การเปรียบเทียบผลลัพธ์สำหรับเทคนิคต่าง ๆ ทดสอบด้วยข้อมูลตัวอย่างผู้สมัครงานในส่วนทักษะการเขียนโปรแกรมคอมพิวเตอร์

ลำดับ	ทักษะที่ปรากฏในประวัติของผู้สมัครงาน	ความต้องการของผู้สรรหา	เทคนิคที่ใช้			
			1) การจับคู่ด้วยพจน์	2) การหาความคล้ายคลึงกันแบบดั้งเดิม	3) กระบวนการจับคู่เชิงความหมายของผู้วิจัย	
1	C#	Ruby น้ำหนัก 0.3 และ Java น้ำหนัก 0.7	ไม่พบข้อมูล	$\text{sim}_c(\text{"Java"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"C\#"})=0.9375$	ความคล้ายคลึงเฉลี่ย 0.9375	คะแนน=0.80875
2	C#, Java		พบข้อมูล Java	$\text{sim}_c(\text{"Java"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Java"}, \text{"Java"})=1$ $\text{sim}_c(\text{"Ruby"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"Java"})=0.9375$	ความคล้ายคลึงเฉลี่ย 0.953125	คะแนน=0.915625
3	Ruby		พบข้อมูล Ruby	$\text{sim}_c(\text{"Java"}, \text{"Ruby"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"Ruby"})=1$	ความคล้ายคลึงเฉลี่ย 0.96875	คะแนน=0.768125
4	C#, Python		ไม่พบข้อมูล	$\text{sim}_c(\text{"Java"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Java"}, \text{"Python"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"Python"})=0.9375$	ความคล้ายคลึงเฉลี่ย 0.9375	คะแนน=0.80875
5	Python		ไม่พบข้อมูล	$\text{sim}_c(\text{"Java"}, \text{"Python"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"Python"})=0.9375$	ความคล้ายคลึงเฉลี่ย 0.9375	คะแนน=0.61875
6	C#		ไม่พบข้อมูล	$\text{sim}_c(\text{"Java"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"C\#"})=0.9375$	ความคล้ายคลึงเฉลี่ย 0.9375	คะแนน=0.80875
7	Ruby, Java		พบข้อมูล Ruby และ Java	$\text{sim}_c(\text{"Java"}, \text{"Ruby"})=0.9375$ $\text{sim}_c(\text{"Java"}, \text{"Java"})=1$ $\text{sim}_c(\text{"Ruby"}, \text{"Ruby"})=1$ $\text{sim}_c(\text{"Ruby"}, \text{"Java"})=0.9375$	ความคล้ายคลึงเฉลี่ย 0.96875	คะแนน=1.0
8	C#, Ruby		พบข้อมูล Ruby	$\text{sim}_c(\text{"Java"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Java"}, \text{"Ruby"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"Ruby"})=1$	ความคล้ายคลึงเฉลี่ย 0.953125	คะแนน=0.908125
9	Python		ไม่พบข้อมูล	$\text{sim}_c(\text{"Java"}, \text{"Python"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"Python"})=0.9375$	ความคล้ายคลึงเฉลี่ย 0.9375	คะแนน=0.61875
10	C#		ไม่พบข้อมูล	$\text{sim}_c(\text{"Java"}, \text{"C\#"})=0.9375$ $\text{sim}_c(\text{"Ruby"}, \text{"C\#"})=0.9375$	ความคล้ายคลึงเฉลี่ย 0.9375	คะแนน=0.80875

อย่างไรก็ตามเทคนิคการหาความคล้ายคลึงกันเพียงอย่างเดียว ยังคงไม่สามารถบอกความแตกต่างข้อมูลของโหนดระดับเดียวกันได้ (สังเกตผู้สมัครลำดับที่ 1 และ 5 จากตารางที่ 4) ดังนั้นการนำเทคนิคดังกล่าวมาพัฒนาต่อเป็นกระบวนการจับคู่เชิงความหมายของผู้วิจัยที่ได้กล่าวไว้แล้ว สามารถระบุตัวเลขความแตกต่างได้

ชัดเจนกว่าเทคนิคการหาความคล้ายคลึงกันแบบดั้งเดิม (สังเกตค่าคะแนนที่ได้จากตารางที่ 4)

4. สรุปผลการวิจัยและข้อเสนอแนะ

งานวิจัยนี้ได้นำเสนอกระบวนการเพิ่มประสิทธิภาพในระบบการคัดสรรบุคลากรอิเล็กทรอนิกส์ด้วยเทคนิคการจับคู่เชิงความหมายโดยผู้วิจัยได้นำเสนอกระบวนการออกแบบกรอบ

แนวคิดของระบบต้นแบบ การออกแบบต้นแบบ ฐานความรู้ออนโทโลยี โดยนำมาประยุกต์ใช้กับ กระบวนการจับคู่เชิงความหมายที่ผู้วิจัยได้ออกแบบไว้ และนำมาทดสอบด้วยกลุ่มข้อมูลตัวอย่างของผู้สมัครงาน โดยสรุปได้ว่าถ้าเปรียบเทียบกับกระบวนการหาความคล้ายคลึงกันเพียงอย่างเดียวจากการนำเสนอในงานวิจัยที่เกี่ยวข้องแล้ว ผลคะแนนสุดท้ายจากกระบวนการจับคู่เชิงความหมายของงานวิจัยนี้สามารถแยกความแตกต่างได้มีประสิทธิภาพดีกว่าซึ่งเกิดจากการนำคำอธิบายการเรียนรู้มาประยุกต์ใช้เพิ่มเติมตลอดจนการนำเอาเทคโนโลยีเว็บเชิงความหมายมาประยุกต์ใช้ในระบบต้นแบบทำให้สารสนเทศสามารถถูกแบ่งปันในระบบอินเทอร์เน็ตได้ง่ายและเป็นมาตรฐานเดียวกัน รวมถึงเพื่อเป็นการลดความซ้ำซ้อนในการฝากประวัติส่วนตัวของผู้สมัครงานด้วยเช่นกัน

ในส่วนข้อเสนอแนะสำหรับงานวิจัยนี้ ประกอบไปด้วยประเด็นการกำหนดน้ำหนักของ อัตรการเรียนรู้ในการสร้างออนโทโลยีซึ่งได้มาจากผลสำรวจจากผู้เชี่ยวชาญในการรวบรวมองค์ความรู้ แต่อย่างไรก็ตามเนื่องจากในออนโทโลยีมีความสัมพันธ์ที่หลากหลาย จึงจำเป็นต้องอ้างอิงตามมาตรฐานและมีผู้เชี่ยวชาญในศาสตร์อื่น ๆ เพิ่มเติมจากแหล่งที่สามารถเชื่อถือได้ นอกจากนี้ควรเพิ่มปัจจัยอื่นเสริมเข้าไปในกระบวนการจับคู่เชิงความหมาย เช่น ประสบการณ์ ที่อยู่ การศึกษา และปัจจัยในด้านอื่นของผู้สมัครงาน เพื่อเพิ่มประสิทธิภาพให้กับระบบในการค้นหาข้อมูลที่เหมาะสมและน่าเชื่อถือได้ดียิ่งขึ้น สุดท้ายแล้วกระบวนการคัดสรรดังกล่าวยังสามารถทำได้ในมุมมองจากผู้สมัครงานไปค้นหาผู้คัดสรรหรือผู้

ว่าจ้างได้ ในลักษณะเดียวกันกับกระบวนการที่ผู้วิจัยได้นำเสนอไป

5. เอกสารอ้างอิง

- [1] Internet World Stats, "Internet Usage Statistics," [Online]. Available: <http://www.internetworldstats.com/stats.htm>. [Accessed 20 January 2015].
- [2] M. Sills, "E-recruitment: Comparison with traditional recruitment and the influences of social media," 2014.
- [3] S. Hussain, A. Mushtaque, H. Dhiman, Nida and S. N. Shivhare, "Approaches used for Ontology Matching in eRecruitment," *International Journal For Research in Applied Science and Engineering Technology (IJRASET)*, vol. 51, pp. 39-45, 2014.
- [4] The University of Sheffield, "Human Resources," [Online]. Available: <http://www.shef.ac.uk/hr/recruitment/erecruitment/atr>. [Accessed 20 December 2014].
- [5] M. Awang, N. R. Khan and M. A. Ghouri, "Impact of E-Recruitment and Job-Seekers Perception on Intention to Pursue the Jobs," *Management & Marketing*, vol. 11, no. 1, pp. 47-57, 2013.
- [6] P. Hitzler, M. Krotzsch and S. Rudolph, *Foundations of Semantic Web Technologies*, CRC Press, 2009.
- [7] T. Berners-Lee, J. Hendler and O. Lassila, "The Semantic Web," *Scientific American Magazine*, 2001.

- [8] W3Schools, "Semantic Web Tutorial," 2010. [Online]. Available: <http://w3schools.com/semweb/default.asp>. [Accessed 9 May 2014].
- [9] A. Medic and A. Golubovic, "Making secure Semantic Web," *Universal Journal of Computer Science and Engineering Technology*, pp. 99-104, 2010.
- [10] W3C Member, "Resource Description Framework," 15 March 2014. [Online]. Available: <http://www.w3.org/RDF/>. [Accessed 23 August 2014].
- [11] W3C Member, "Web Ontology Language," 11 December 2012. [Online]. Available: <http://www.w3.org/2001/sw/wiki/OWL>. [Accessed 23 August 2014].
- [12] C. Bizer, R. Heese, M. Mochol, R. Oldakowski, R. Tolksdorf and R. Eckstein, "The Impact of Semantic Web Technologies on Job Recruitment Processes," in *Proceedings of the 7th International Conference Wirtschaftsinformatik*, Berlin, 2005.
- [13] P. Nilaphruek, "A Job Recruitment System Using Semantic Web Technology," in *The 15th International Conference of International Academy of Physical Sciences Dec 9 - 13*, Pathumthani, Thailand, 2012.
- [14] H. Lv and B. Zhu, "Skill ontology-based semantic model and its matching algorithm," in *7th International Conference on Computer-Aided Industrial Design and Conceptual Design*, 2006, Hangzhou, 2006.
- [15] M. V. Coillie, "Europortfolio Europass CV HR-XML Application Profile," 2010.
- [16] UNESCO, "International Standard Classification of Education," UNESCO-UIS, 2006.
- [17] International Labour Organization, "International Standard Classification of Occupations," 18 September 2004. [Online]. Available: <http://www.ilo.org/public/english/bureau/stat/isco/isco88/index.htm>. [Accessed 12 November 2014].
- [18] Eurostat, "Statistical classification of economic activities in the European Community," Office for Official Publications of the European Communities, Luxembourg, 2008.
- [19] T. B. Lee, "HM Government," Opening up Government, 2009. [Online]. Available: <http://data.gov.uk/linked-data>. [Accessed 14 December 2014].
- [20] H. Paulheim and S. Hertling, "Discoverability of SPARQL Endpoints in Linked Open Data," in *The 12th International Semantic Web Conference and the 1st Australasian Semantic Web Conference*, Sydney, Australia, 2013.
- [21] United States Agency for International Development, Overview of Job Classifications and Certifications in the IT Sector, USAID, 2006.

-
- [22] J. Zhong, H. Zhu, J. Li and Y. Yu, "Conceptual Graph Matching for Semantic Search," in *10th International Conference on Conceptual Structures, ICCS 2002, Proceedings*, Borovets, 2002.
- [23] "C Programming and C++ Programming," CProgramming, 2011. [Online]. Available: <http://www.cprogramming.com/>. [Accessed 3 January 2015].
- [24] E. Triantaphyllou, "Multi-Criteria Decision Making: A Comparative Study," *The Netherlands: Kluwer Academic Publishers (now Springer)*, p. 320, 2000.
- [25] T. R. Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition* 5, pp. 199-220, 1993.
- [26] F. Arvidsson and A. Flycht-Eriksson, "Ontologies I," 2008.