

บทที่ 1

บทนำ

1. ความเป็นมาและความสำคัญของปัญหา

ในยุคที่อินเทอร์เน็ตมีบทบาทในชีวิตประจำวันทำให้เกิดเว็บแอปพลิเคชันมากมาย อีกทั้งการใช้งานที่หลากหลายของเว็บแอปพลิเคชัน ทำให้เว็บที่มีลักษณะข้อมูลที่คล้ายคลึงกัน มีการใช้งานที่แตกต่างกัน เมื่อรูปแบบการใช้งานของเว็บแอปพลิเคชันประจำแหล่งข้อมูลนั้นๆ มีลักษณะการใช้งานที่แตกต่างกัน จึงทำให้ผู้ใช้งานต้องเรียนรู้การใช้งานเว็บแอปพลิเคชันที่ไม่คุ้นเคยใหม่ รวมทั้งบางแหล่งข้อมูลยังมีความล่าช้าในการนำข้อมูลกลับมาประมวลผล ทำให้ผู้ใช้ไม่ได้รับความสะดวกในการเข้าถึงข้อมูล และใช้เวลามากในการรวมหรือเปรียบเทียบข้อมูลจากแหล่งต่าง ๆ ด้วยตัวเอง

ยกตัวอย่างกรณีศึกษาข้อมูลชีวสารสนเทศศาสตร์ ผู้ใช้ที่เป็นนักวิจัยต้องการข้อมูลจีโนมจากหลาย ๆ ฐานข้อมูล เพื่อนำมาประมวลผลและตัดสินใจ โดยที่ฐานข้อมูลจีโนมเหล่านั้นเป็นฐานข้อมูลที่มีการพัฒนาเครื่องมือที่จะนำเข้ามาใช้ในการวิเคราะห์ข้อมูลจีโนมที่เก็บอยู่ เพื่อผลิตข้อมูลที่มีความจำเพาะมากขึ้นและใช้ในการตอบคำถามพื้นฐานทางชีววิทยาต่อไป ถึงแม้ฐานข้อมูลเหล่านี้จะได้รับการสนับสนุนจากศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติ (National Center for Genetic Engineering and Biotechnology: BIOTEC) หรือไบโอเทค แต่ข้อมูลหลักยังอยู่ตามหน่วยงานต่าง ๆ ที่รับการสนับสนุนทุนวิจัย ซึ่งมีรูปแบบของการดำเนินงานที่แตกต่างกัน ดังนั้นการเข้าถึงฐานข้อมูลของจีโนมต่าง ๆ จะทำได้โดยเข้าไปที่เว็บไซต์ของโครงการนั้นๆ เพื่อเรียนรู้วิธีการดึงข้อมูลจำเพาะออกมา การเข้าถึงข้อมูลจีโนมที่แตกต่างกันจึงเป็นไปได้แต่ไม่สะดวกนัก และยังติดอยู่กับรูปแบบของการใช้งานที่ไม่เหมือนกัน รวมถึงข้อมูลที่ได้ออกมาอาจมีรูปแบบที่แตกต่างกัน ทำให้นักวิจัยเกิดความสับสนได้ ดังนั้นการสร้างระบบกลางที่มีมาตรฐานการเข้าถึงที่สะดวกให้กับนักวิจัย จะช่วยให้นักวิจัยสามารถใช้ข้อมูลจีโนมดังกล่าวในรูปแบบเดียวกัน ซึ่งสามารถเข้าใจได้ง่าย สามารถนำมาเปรียบเทียบกันได้อย่างสะดวก และเป็นประโยชน์ในงานวิจัยต่อไปในอนาคต

การบูรณาการข้อมูลที่ใช้ต้องการจากแหล่งข้อมูลต่าง ๆ ที่มีความเป็นอิสระต่อกันให้อยู่ในระบบศูนย์กลางที่มีโครงสร้างหรือส่วนติดต่อมาตรฐาน จะช่วยให้ผู้ใช้งานข้อมูลที่ต้องการไปใช้ได้ดียิ่งขึ้น เนื่องจากไม่ต้องทำการศึกษาจากเว็บแอปพลิเคชันอื่นที่ระบบศูนย์กลางมีข้อมูลอยู่แล้ว เช่น ระบบศูนย์กลางของแหล่งข้อมูลจีโนม สามารถช่วยให้ผู้ใช้ไม่ต้องเข้าใจถึงโครงสร้างของฐานข้อมูลหรือการเรียนรู้การใช้งานเว็บแอปพลิเคชันของแหล่งข้อมูลจีโนมของโครงการต่าง ๆ และนอกจากนี้ผู้ใช้ไม่จำเป็นต้องทราบว่าฐานข้อมูลของจีโนมนั้นอยู่ที่ใด ผู้ใช้สามารถใช้คำถามที่อยู่ในลักษณะที่สั้น กระชับ เพื่อถามถึงข้อมูลที่ใช้ต้องการโดยที่ผู้ใช้ไม่จำเป็นต้องถามเกี่ยวกับจีโนมของข้อมูลประเภทเดียวกันในแต่ละคำถาม นั่นคือ ผู้ใช้สามารถถามคำถามเกี่ยวกับจีโนมของข้อมูลที่แตกต่างกันในหนึ่งคำถาม โดยที่แหล่งข้อมูลเสมือนจะทำหน้าที่ไปหาคำตอบนั้นให้กับผู้ใช้เอง นอกจากนี้การใช้แหล่งข้อมูลเสมือนไม่จำเป็นต้องใช้งานผ่านทางเว็บเบราว์เซอร์เพียงอย่างเดียว ผู้ใช้สามารถเขียนโปรแกรมเพื่อที่จะร้องขอข้อมูลและนำข้อมูลไปประมวลผลตามที่ใช้ต้องการต่อไป

วิทยานิพนธ์นี้นำเสนอโครงสร้างพื้นฐานของระบบศูนย์กลางในลักษณะของแหล่งข้อมูลเสมือน (Virtual Resource) และใช้กรณีศึกษาของไบโอเทค โดยทำการรวบรวมข้อมูลจากแหล่งข้อมูลต่าง ๆ ที่มีลักษณะใกล้เคียงกันไว้ โดยอาศัยเทคโนโลยีเว็บเซอร์วิสช่วยในการบูรณาการข้อมูล (Integration Data) ผ่านภาษาเอกซ์เอ็มแอล ซึ่งจะช่วยให้นักวิจัยนั้นสามารถเข้าถึงข้อมูลและนำข้อมูลที่ต้องการได้จากแหล่งเดียว และด้วย

เทคโนโลยีเว็บเซิร์ฟเวอร์ นักวิจัยหรือนักพัฒนาสามารถพัฒนาโปรแกรมเพื่อเรียกใช้เว็บเซิร์ฟเวอร์เพื่อนำข้อมูลที่ได้ไปใช้งานตามความต้องการ โดยมีเทคโนโลยีเพียร์-ทู-เพียร์ช่วยจัดการและเพิ่มประสิทธิภาพในการเข้าถึงแหล่งข้อมูลให้รวดเร็ว และมีเสถียรภาพ

ประเด็นสำคัญของวิทยานิพนธ์นี้อยู่ที่ การศึกษา ออกแบบ และพัฒนาโครงสร้างพื้นฐานของระบบบูรณาการของแหล่งข้อมูลเสมือนโดยใช้เว็บเซิร์ฟเวอร์และเพียร์-ทู-เพียร์ การใช้เว็บเซิร์ฟเวอร์เป็นวิธีการเพิ่มความสามารถในการเข้าถึงแหล่งข้อมูลที่ใช้ต้องการผ่านแอปพลิเคชันที่ติดต่อกับเว็บเซิร์ฟเวอร์ศูนย์กลาง หรือเว็บเซิร์ฟเวอร์ใด ๆ เพื่อให้ได้ข้อมูลที่มีลักษณะโครงสร้างข้อมูลชุดเดียวกัน และเพียร์-ทู-เพียร์ช่วยเสริมประสิทธิภาพในการทำงานระหว่างเว็บเซิร์ฟเวอร์ในกรณีที่เว็บเซิร์ฟเวอร์ศูนย์กลางไม่สามารถให้บริการได้โดยผู้ใช้สามารถเข้าถึงข้อมูลได้ผ่านแหล่งข้อมูลใดข้อมูลหนึ่งเสมือนว่ากำลังใช้เว็บเซิร์ฟเวอร์ศูนย์กลางอยู่ ทั้งนี้ผลจากการใช้เว็บเซิร์ฟเวอร์และเพียร์-ทู-เพียร์ คาดว่าจะช่วยให้ผู้ใช้สามารถเข้าถึงข้อมูลได้รวดเร็ว มีเสถียรภาพ ข้อมูลที่ได้มีความถูกต้องตรงความต้องการ สามารถขยายระบบได้ง่าย และสามารถนำระบบนี้ไปใช้งานจริงต่อไป

2. วัตถุประสงค์ของวิทยานิพนธ์

- 2.1 เพื่อออกแบบและพัฒนาระบบโครงสร้างพื้นฐานในการเชื่อมต่อแหล่งข้อมูลของศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติให้การสนับสนุน โดยที่แหล่งข้อมูลต่าง ๆ มีความเป็นอิสระต่อกัน
- 2.2 เพื่อพัฒนาระบบกลางที่ให้บริการข้อมูลที่ได้รับจากการประมวลผลจากแหล่งข้อมูลย่อยต่าง ๆ
- 2.3 เพื่อพัฒนารูปแบบเว็บเซิร์ฟเวอร์และอินเทอร์เฟซของแหล่งข้อมูลให้ใช้งานได้สะดวก
- 2.4 เพื่อพัฒนาแหล่งข้อมูลเสมือนที่ผู้ใช้สามารถค้นหาข้อมูลจากแหล่งข้อมูลใด ๆ ก่อนเข้าถึงแหล่งข้อมูลจริง
- 2.5 เพื่อพัฒนาระบบกลางที่สามารถจัดการสิทธิ์การเข้าถึงแหล่งข้อมูลของผู้ใช้ภายในระบบได้

3. ขอบเขตและข้อจำกัดของวิทยานิพนธ์

- 3.1 พัฒนาเว็บเซิร์ฟเวอร์เพื่อเป็นส่วนเชื่อมต่อกับแหล่งข้อมูลจีโนมที่ศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติให้การสนับสนุน
- 3.2 พัฒนาเว็บเซิร์ฟเวอร์กลางที่เป็นทะเบียนของเว็บเซิร์ฟเวอร์ที่เชื่อมต่อกับแหล่งข้อมูลจีโนมที่ศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติให้การสนับสนุน
- 3.3 ใช้เพียร์-ทู-เพียร์ในการขยายระบบและแบ่งปันข้อมูลเพื่อเพิ่มประสิทธิภาพในการเข้าถึงแหล่งข้อมูล
- 3.4 ใช้ศูนย์กลางของเพียร์-ทู-เพียร์ในการจัดการและควบคุมสิทธิ์การเข้าถึงแหล่งข้อมูลของผู้ใช้

4. สถานที่ทำงานวิจัย

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น และศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติ อุทยานวิทยาศาสตร์ ประเทศไทย

5. ขั้นตอนและวิธีการดำเนินงาน

5.1 ศึกษาค้นคว้าเรื่องเทคโนโลยีเว็บเซอวิส และเพียร์-ทู-เพียร์ ในการนำไปใช้งานบูรณาการข้อมูลจากวรรณกรรม และเอกสารต่าง ๆ ที่เกี่ยวข้อง

5.2 ออกแบบและพัฒนาเว็บเซอวิสเพื่อเป็นส่วนติดต่อในการเข้าถึงข้อมูลในแหล่งข้อมูล เรียกว่า ศูนย์กลางแหล่งบริการข้อมูลเว็บเซอวิส

5.3 ออกแบบและพัฒนาเว็บเซอวิสกลางหรือศูนย์กลางการค้นหาและจัดการแหล่งข้อมูล เพื่อใช้เป็นทะเบียนจัดเก็บแหล่งบริการข้อมูลเว็บเซอวิส และควบคุมสิทธิการใช้งานของผู้ใช้

5.4 พัฒนาศูนย์กลางการค้นหาและจัดการแหล่งข้อมูล และศูนย์กลางแหล่งบริการข้อมูลเว็บเซอวิสให้ทำงานร่วมกันโดยอัตโนมัติด้วยหลักการเพียร์-ทู-เพียร์

5.5 พัฒนาโปรแกรมทดสอบ สำหรับประเมินประสิทธิภาพของระบบในกรณีต่าง ๆ

5.6 ออกแบบและพัฒนาส่วนติดต่อผู้ใช้ เพื่อเป็นต้นแบบในการใช้งานจริง

5.7 สรุป และวิเคราะห์ผลการทดสอบ

5.8 จัดทำ และนำเสนอวิทยานิพนธ์

6. ประโยชน์ที่คาดว่าจะได้รับ

6.1 มีแหล่งข้อมูลจีโนมกลางที่รวมข้อมูลจากแหล่งข้อมูลที่ศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติให้การสนับสนุน โดยที่การเข้าถึงข้อมูลมีความปลอดภัยตามสิทธิในการเข้าถึงข้อมูลของผู้ใช้

6.2 มีการนำเสนอเทคโนโลยี และแนวทางที่สามารถนำไปประยุกต์ใช้ในการบูรณาการข้อมูลเสมือนจากแหล่งข้อมูลต่าง ๆ และแนวทางในการค้นหาข้อมูลที่ต้องการจากหลาย ๆ แหล่งข้อมูลได้ที่แหล่งเดียว

6.3 มีการนำเสนอแนวทางในการเข้าถึงข้อมูลต่าง ๆ ได้รวดเร็ว มีเสถียรภาพ ข้อมูลที่ได้มีความถูกต้อง และสามารถขยายระบบได้ง่าย

6.4 มีการนำเสนอโครงสร้างพื้นฐานสำหรับงานวิจัยทางชีวสารสนเทศศาสตร์ในอนาคต

6.5 มีการนำเสนอโครงสร้างพื้นฐานที่ใช้ร่วมกับโปรแกรมสร้างกระบวนการทำงาน (Work Flow Application) เพื่อเชื่อมต่อกับเว็บเซอวิสอื่นๆ ได้