

**การค้นคืนสารสนเทศของเอกสารแบบข้อความ**  
(Information Retrieval on Textual Document)  
**สมเกียรติ เจริญตระกูลชัย , รัชต พิษวณิชย์**

**บทคัดย่อ**

รายงานฉบับนี้ เป็นการนำเสนอ ความรู้เกี่ยวกับระบบการค้นคืนสารสนเทศของเอกสารแบบข้อความ โดยมีการแสดงถึงความรู้ที่เกี่ยวข้องกับ การค้นคืนสารสนเทศของเอกสารแบบข้อความ เช่น การประมวลผลข้อความ การให้ค่าน้ำหนักของคำ เทคนิคหรือวิธีการในการพิจารณาคัดเลือกกลุ่มเอกสารที่เป็นคำตอบที่เกี่ยวข้องกับคำถามของผู้ใช้ วิธีการวัดประสิทธิภาพของระบบการค้นคืนสารสนเทศ และ การค้นคืนสารสนเทศของข้อมูลในระบบอินเทอร์เน็ต ในส่วนสุดท้ายของรายงานฉบับนี้ ได้มีการนำเสนองานวิจัยเรื่อง การจัดกลุ่มเอกสารข่าวภาษาไทย ด้วยเทคนิคการวิเคราะห์หาความหมายแอบแฝง โดยมีแนวคิดการทดลองการจัดกลุ่มเอกสารข่าวภาษาไทยด้วยเทคนิคการวิเคราะห์หาความหมายแอบแฝง รวมถึงมีการวัดผลความถูกต้องของการจัดกลุ่มเอกสารข่าว โดยการใช้เทคนิคการวิเคราะห์หาความหมายแอบแฝง เปรียบเทียบกับผลการจัดกลุ่มเอกสารข่าวภาษาไทยโดยมนุษย์

**Abstract**

This report presents a literature review on textual information retrieval. The contents of the report include topics on text processing, term weighting, matching techniques, system evaluation, and Internet-based information retrieval. Lastly, the report discusses a research proposal on Thai-language news text grouping using the Latent Semantic Analysis technique. The proposal seeks to compare the effectiveness of clustering by Latent Semantic Analysis against grouping assigned by humans.