

หัวข้อวิจัย ระบบต้นแบบการสร้างฐานข้อมูลสารสนเทศบนฐานกฎแบบอัตโนมัติสำหรับ  
เอกสารราชการไทย

ผู้วิจัย ผู้ช่วยศาสตราจารย์อุษานาฏ เอื้ออภิสิทธิ์วงศ์

หน่วยงาน มหาวิทยาลัยราชภัฏนครราชสีมา

ปีที่ทำการวิจัยเสร็จเรียบร้อย พ.ศ. 2558

### บทคัดย่อ

เอกสารทางราชการเป็นเอกสารที่ถูกใช้ในหน่วยงานราชการเป็นจำนวนมากและสามารถแบ่งลักษณะเอกสารที่แตกต่างกัน คือ เอกสารที่ใช้ภายในหน่วยงานหรือเอกสารที่ใช้ภายนอกหน่วยงาน โดยรูปแบบคุณลักษณะของเอกสารมีความแตกต่างทั้งชนิดของไฟล์ โครงสร้างภายในเอกสาร การจัดวางตำแหน่งหรือลักษณะของตัวอักษรทำให้ยากต่อการสกัดข้อมูลออกจากเอกสาร เนื่องจากเอกสารราชการเป็นเอกสารเชิงข้อความที่ไม่มีโครงสร้าง ดังนั้นในงานวิจัยชิ้นนี้นำเสนอระบบต้นแบบสำหรับการสกัดข้อมูลสารสนเทศบนฐานของกฎออกจากเอกสารราชการไทยและเก็บลงในฐานข้อมูลแบบอัตโนมัติ ข้อมูลที่นำมาทดลองเป็นข้อมูลเอกสารราชการไทยของคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมาจำนวน 150 เอกสาร โดยผลการประเมินความพึงพอใจจากผู้ใช้งานในระดับพอใช้ (ค่าเฉลี่ยของความพึงพอใจเท่ากับ 3.20) และมีค่าความถูกต้องในการสกัดข้อมูลเท่ากับร้อยละ 80

**Title** The Prototype System of Automatic Database Construction based on Rules for the Thai Official Documents.

**Researcher** Assist. Prof. Usanad Ua-apisitwong

**Institute** Nakhon Ratchasima Rajabhat University

**Year** 2015

### **Abstract**

There are a lot of usage for official documents in government agencies. They can be separated to two types: Internal agency documents and External agency documents. They have many differences of document's features such as file format, document structure, document alignment, or fonts. From these reason, it's so hard to extract data from a document because of the official document is an unstructured data. This paper aims to develop a prototype system for automatic extracting any information from Thai government documents based on generated rules. The data are 150 Thai official documents from the faculty of science and technology, Nakhon Ratchasima Rajabhat University. For the experimental results, the satisfaction level of the user is fair (3.20 on average) and the extracting accuracy is 80 percentage.