

การจัดทำข้อมูลเปิดของกรมพัฒนาธุรกิจการค้าที่ให้บริการตามมาตรฐานด้วย โปรแกรมโอเพนซอร์ส CKAN

OPEN DATA GENERATION OF THE BUSINESS DEVELOPMENT DEPARTMENT THAT PROVIDES STANDARD SERVICES WITH THE OPENSOURCE PROGRAM CKAN

ภูวิช ไม้แก้ว¹

ผศ.ดร. ชัยพร เขมะภาคะพันธ์²

บทคัดย่อ

ข้อมูลของกรมพัฒนาธุรกิจการค้าเป็นข้อมูลภาครัฐที่มีความสำคัญ ซึ่งข้อมูลเหล่านี้เกิดจากการให้บริการภายใต้ภารกิจของหน่วยงาน และได้มีการจัดทำชุดข้อมูลเปิดเผยเพื่อนำเสนอสำหรับ ภาครัฐ เอกชน และประชาชนผู้ที่สนใจ สามารถติดตามข้อมูลข่าวสารผ่านหน้าเว็บไซต์ เพื่ออ้างอิงเนื้อหาและแหล่งที่มาซึ่งใน กระบวนการวิเคราะห์ข้อมูลขนาดใหญ่ ผู้ใช้ข้อมูลจำเป็นต้องสำรวจหาแหล่งข้อมูล คัดเลือกข้อมูลที่ตรงตาม ความต้องการ ร้องขอข้อมูลไปยังหน่วยงาน และทำความเข้าใจกับข้อมูลนั้นเป็นเรื่องที่ซับซ้อนและยุ่งยาก จำเป็นต้องมีการบริหารจัดการข้อมูลภาครัฐอย่างเป็นระบบเพื่อให้ผู้ใช้งานสามารถเลือกใช้ข้อมูลได้อย่างสะดวก และสามารถเข้าถึงข้อมูลภาครัฐเพื่อวิเคราะห์และตัดสินใจได้ทันต่อสถานการณ์และความต้องการใช้งาน

งานวิจัยนี้มุ่งเน้นไปที่การศึกษาการจัดทำชุดข้อมูลที่สามารถเปิดเผยต่อสาธารณะโดยมีการจัดทำ คำอธิบายข้อมูลหรือเมทาดาทา (Metadata) สำหรับชุดข้อมูลเปิด ที่เป็นพื้นฐานสำคัญ ซึ่งการจัดทำดังกล่าว เป็นไปตาม มาตรฐานรัฐบาลดิจิทัลด้วยแนวทางการจัดทำบัญชีข้อมูลภาครัฐ (Government data catalog guideline) มาตรฐานรัฐบาลดิจิทัลด้วยแนวทางการเปิดเผยข้อมูลเปิดภาครัฐในรูปแบบดิจิทัลต่อสาธารณะ (Open Government Data Guideline) และธรรมาภิบาลข้อมูลภาครัฐ (Data governance for government) โดยใช้ เครื่องมือซอฟต์แวร์โอเพนซอร์ส Comprehensive Knowledge Archive Network (CKAN) และส่วนเสริม CKAN Open-D เพื่อสนับสนุนการจัดทำชุดข้อมูลเปิดของกรมพัฒนาธุรกิจการค้า ซึ่งทำให้ลดกระบวนการจัดทำข้อมูล เปิดที่หน่วยงานจัดทำ และเพิ่มความถูกต้องของชุดข้อมูลที่มีการกำหนดโครงสร้างและระยะเวลาปรับปรุงข้อมูล ทำให้เกิดการเผยแพร่ชุดข้อมูลเปิดตามคุณลักษณะของข้อมูลเปิดที่ดี โดยมีข้อมูลเปิดจากแต่ละส่วนงานทั้งหมด

¹ นักศึกษาหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต วิศวกรรมคอมพิวเตอร์

² ที่ปรึกษาสารนิพนธ์

11 ชุดข้อมูล ภายใต้ธรรมาภิบาลข้อมูลให้อยู่ในระบบ CKAN ซึ่งระบบสามารถจัดการผู้ที่เกี่ยวข้องกับการจัดทำชุดข้อมูลทั้งหมด แต่ตั้งเริ่มต้นจัดทำข้อมูลไปจนถึงเผยแพร่ข้อมูล พร้อมทั้งนำเสนอชุดข้อมูลได้ยืดหยุ่นต่อการใช้งานทั้งผ่านเว็บไซต์ เชื่อมต่อกับโปรแกรมภายนอก หรือเชื่อมโยงผ่าน API

คำสำคัญ : ข้อมูลเปิดเผย , มาตรฐานรัฐบาลดิจิทัล , Metadata , CKAN

ABSTRACT

Information from the Department of Business Development is business sector information and is important. This information is generated by providing services under the mission of the agency. And has prepared a set of disclosure information to present to the public sector, private sector, and the public who are interested. You can follow news updates through the website. To reference content and sources, which are used in the process of analyzing big data. Data users need to explore sources of information. Select the information that meets your needs. request information from the agency And understanding that data is complex and difficult. It is necessary to manage government data systematically so that users can choose to use data conveniently and can access government data for analysis and decision-making in time for situations and usage needs.

This research focuses on the study of the preparation of publicly available datasets by providing metadata for open datasets. That is fundamental, and such preparation is in accordance with Digital Government Standards on Government Data Catalogue Guidelines, Digital Government Standards on Open Government Data Guidelines, and Data Governance Government (Data governance for government) using open source software tools. Comprehensive Knowledge Archive Network (CKAN) and the CKAN Open- D extension to support the creation of the Department of Business Development's open datasets This reduces the amount of open data that the agency produces. and increases the accuracy of structured datasets and data update intervals. This results in the propagation of open data sets based on the characteristics of good open data. There are 11 open data sets from each division under good governance in the CKAN system. from the beginning of the preparation of the information until the dissemination of the information, as well as offering flexible data sets for use both through the website, connecting to external programs, or linked via API.

Keywords: Disclosure, Digital Government Standards, Metadata, CKAN

บทนำ

1. ที่มาและความสำคัญของปัญหา

กรมพัฒนาธุรกิจการค้าให้บริการข้อมูลเปิดเผยต่อสาธารณะโดยใช้หลักธรรมาภิบาลข้อมูลภาครัฐของหน่วยงานในการให้บริการข้อมูล กระบวนการในการจัดทำและนำเสนอมีขั้นตอนที่ซับซ้อนและเกี่ยวข้องกับคณะทำงานภายในหน่วยงานที่เกี่ยวข้องตามชุดข้อมูล ทำให้การประสานงานเพื่อเตรียมชุดข้อมูลเปิดในแต่ละส่วนงานอาจมีบางขั้นตอนที่แตกต่างกันไปตามการจัดเก็บข้อมูลการประมวลผลไปจนถึงการเผยแพร่ข้อมูล ส่งผลให้ข้อมูลที่ให้บริการนั้นไม่มีโครงสร้างที่ดี เช่นข้อมูลที่ให้บริการล้วนเป็นข้อมูลที่ผ่านการประมวลผลมาแล้วเป็นข้อมูลเชิงสรุป เนื้อหาและการปรับปรุงไม่มีประสิทธิภาพ เช่นกระบวนการปรับปรุงข้อมูลที่มีขั้นตอนจากผู้เป็นเจ้าของข้อมูลไปถึงผู้รวบรวมข้อมูลขออนุมัติผู้บังคับบัญชา และส่งให้กับผู้นำเสนอข้อมูล ซึ่งมีขั้นตอนที่ซับซ้อนและล่าช้าในการปรับปรุงและอาจเกิดความผิดพลาดได้ง่าย และการนำเสนอข้อมูลไม่มีรูปแบบที่สามารถนำไปประมวลผลหรือวิเคราะห์ข้อมูลต่อได้ เช่นข้อมูลจากแต่ละส่วนงานจะมีการกำหนดรูปแบบในการนำเสนอขึ้นมา ได้แก่ ไฟล์ .xls, .pdf หรือ .jpg จากสาเหตุดังกล่าวทำให้เกิดปัญหา 2 ด้าน ได้แก่

1) ผู้ให้บริการข้อมูลเปิดเผยมีการให้บริการข้อมูลตามหลักธรรมาภิบาลข้อมูลภาครัฐอยู่แล้ว แต่ในกระบวนการจัดการยังขาดระบบบริหารจัดการข้อมูลเปิดที่ทำให้การจัดทำข้อมูลที่มีจากหลายส่วนงานที่มีขั้นตอนที่ซับซ้อนซึ่งบางขั้นตอนไม่มีระบบช่วยจัดการทำให้ข้อมูลที่ต้องนำเสนอนั้นเป็นมาตรฐานที่ดีในการเปิดเผยข้อมูลเปิด

2) ผู้ให้บริการข้อมูลเปิดเผยอาจไม่สามารถหาข้อมูลได้จากแหล่งข้อมูลต้นทางที่น่าเชื่อถือ หรือหากต้องการค้นหาข้อมูลจากต้นทางก็อาจจะยากต่อการสืบค้นหาข้อมูลและใช้เวลาในการศึกษาชุดข้อมูลเนื่องจากไม่มีโครงสร้างชุดข้อมูลที่ชัดเจน หรือมีโครงสร้างแต่ก็อาจจะไม่สามารถนำข้อมูลนั้นมาประมวลผลวิเคราะห์ข้อมูลต่อได้ เพราะข้อมูลที่มีอยู่ส่วนใหญ่เป็นข้อมูลสรุป

งานวิจัยนี้จึงได้มุ่งเน้นที่การจัดทำข้อมูลเปิดของกรมพัฒนาธุรกิจการค้า โดยให้ระบบจัดการข้อมูลเปิดสามารถจัดการชุดข้อมูลจากหลายส่วนงานที่มีโครงสร้างที่ต่างกันมากำหนดชุดโครงสร้างข้อมูล และการดำเนินการจัดทำทรัพยากรข้อมูลในแต่ละส่วนงานให้มีกระบวนการจัดการด้วยระบบ

จัดการชุดข้อมูลเปิด โดยให้บริการตามมาตรฐานเป็นแนวทางที่กำหนดขึ้นเพื่อแก้ไขปัญหาเป็นระบบ ภายใต้กรอบแนวคิดการให้บริการข้อมูลภาครัฐ ที่มีวัตถุประสงค์ให้เป็นระบบสารสนเทศสำหรับการสืบค้นข้อมูลที่ช่วยให้ผู้ใช้สามารถสืบค้น ร้องขอ เข้าถึง และใช้ประโยชน์ข้อมูลอย่างเป็นรูปธรรม เพื่อให้หน่วยงานสร้างการขับเคลื่อนด้วยข้อมูล โดยมีการจัดทำคำอธิบายข้อมูลหรือเมตาดาตา (Metadata) สำหรับชุดข้อมูลภาครัฐ เพื่อลดขั้นตอนกระบวนการจัดเตรียมชุดข้อมูล การจัดทำชุดข้อมูล การตรวจสอบชุดข้อมูล ไปจนถึงการนำเสนอชุดข้อมูล ที่เป็นพื้นฐานสำคัญ ซึ่งการจัดทำดังกล่าวเป็นไปตาม มาตรฐานรัฐบาลดิจิทัลว่าด้วยแนวทางการจัดทำบัญชีข้อมูลภาครัฐ (Government data catalog guideline) มาตรฐานรัฐบาลดิจิทัลว่าด้วยแนวทางการเปิดเผยข้อมูลเปิดภาครัฐในรูปแบบดิจิทัลต่อสาธารณะ (Open Government Data Guideline) และธรรมาภิบาลข้อมูลภาครัฐ (Data governance for government) ด้วยซอฟต์แวร์โอเพนซอร์ส CKAN

วัตถุประสงค์ของการวิจัย

- (1) เพื่อศึกษาการจัดทำบัญชีข้อมูลเปิดและกำหนดชุดข้อมูลเปิด ของหน่วยงานกรมพัฒนาธุรกิจการค้า
- (2) เพื่อการพัฒนาชุดข้อมูลเปิดและทรัพยากรข้อมูลให้เป็นรูปแบบโครงสร้างและมาตรฐานที่ดีสำหรับการเปิดเผยข้อมูล โดยใช้ซอฟต์แวร์โอเพนซอร์ส CKAN

2. วิธีการดำเนินงานวิจัย

ปัจจุบันการจัดทำข้อมูลเปิดของกรมพัฒนาธุรกิจการค้ามีการให้บริการข้อมูลเปิดอยู่บนเว็บไซต์ของหน่วยงาน <http://www.dbd.go.th> ซึ่งข้อมูลที่สามารถให้บริการได้นั้นล้วนเป็นข้อมูลสาธารณะที่เปิดเผยได้โดยใช้หลักการธรรมาภิบาลข้อมูลภาครัฐ โดยหน่วยงานจะมีกระบวนการบันทึกข้อมูลตามภาระกิจหน้าที่ของแต่ละส่วนงานที่เป็นงานให้บริการ ด้าน จดทะเบียนนิติบุคคล ข้อมูลผู้ประกอบการที่ลงทุนภายในประเทศ การขอใบอนุญาตประกอบธุรกิจของคนต่างด้าว เงินทุนจดทะเบียนของนิติบุคคล เป็นต้น เมื่อข้อมูลถูกบันทึกลงสู่ระบบแล้ว จากนั้นผู้เป็นเจ้าของข้อมูลจะทำการประมวลผลข้อมูลตรวจสอบข้อมูลตามโครงสร้างรูปแบบข้อมูลที่ต้องการนำเสนอ และจัดทำข้อมูลนั้นให้กับผู้นำเสนอข้อมูล ก่อนจะขึ้นสู่กระบวนการนำเสนอข้อมูลก็จะต้องตรวจสอบข้อมูลโครงสร้างรูปแบบที่จะนำเสนออีกครั้งก่อนนำไปใช้นำเสนอในช่องทางต่างๆ ของหน่วยงาน ซึ่งแต่ละส่วนงานก็จะมีข้อมูล

เปิดที่ต้องทำการปรับปรุงและนำเสนอ แตกต่างกันไป เช่นบางข้อมูล นำเสนอจัดทำเป็นรายเดือน บางข้อมูลจัดทำนำเสนอเป็นรายปี เป็นต้น และยังมีเรื่องของประเภทไฟล์ข้อมูลที่นำเสนอออกไปก็มีหลายรูปแบบ เช่น Excel file , .pdf , .jpg ซึ่งก็เป็นไปตามแนวทางการความต้องการของการนำเสนอข้อมูล

และจากการศึกษางานวิจัยเพื่อแก้ไขปัญหาเกี่ยวกับการจัดการข้อมูลเปิดให้เป็นระบบพบว่า ปัจจุบันมีเครื่องมือเกี่ยวกับการจัดการข้อมูลเปิดชนิดโอเพ่นซอร์สอยู่หลายตัวซึ่งมีความแตกต่างกันไปตามลักษณะการจัดเก็บข้อมูล กระบวนการทำงาน การนำเสนอเนื้อหา และภาคธุรกิจที่นำไปใช้งาน เช่น โปรแกรม CKAN เป็นหนึ่งในซอฟต์แวร์โอเพ่นซอร์ส สำหรับจัดการข้อมูลเปิดที่เป็นที่นิยมและยังออกแบบให้รองรับการมีคุณสมบัติที่ดีของการเปิดเผยแพลตฟอร์มข้อมูลเปิด ตามมาตรฐานที่ได้มีการศึกษาอ้างอิงนำมาใช้ประกอบการเลือกเครื่องมือ ได้แก่มาตรฐานรัฐบาลดิจิทัลว่าด้วยแนวทางการจัดทำบัญชีข้อมูลภาครัฐ (Government data catalog guideline) มาตรฐานรัฐบาลดิจิทัลว่าด้วยแนวทางการเปิดเผยข้อมูลเปิดภาครัฐในรูปแบบดิจิทัลต่อสาธารณะ (Open Government Data Guideline) และธรรมาภิบาลข้อมูลภาครัฐ (Data governance for government) โดยตัวระบบ CKAN รองรับการจัดการข้อมูลนำเข้าและกำหนดรูปแบบได้เป็นไปตามธรรมาภิบาลข้อมูลของหน่วยงาน ช่วยในเรื่องของลำดับขั้นตอนและหน้าที่ของผู้ปฏิบัติงาน มีโครงสร้างข้อมูล และความถูกต้องของข้อมูล สามารถทำให้ข้อมูลที่อยู่ในโปรแกรมเป็นคุณลักษณะข้อมูลเปิดภาครัฐได้ครบ หน่วยงานยังต่อยอดระบบและข้อมูลเพื่อแชร์ข้อมูลระหว่างหน่วยงานและศูนย์กลางของชุดข้อมูลเปิดได้ สุดท้ายผู้ใช้บริการข้อมูลเปิดเผยได้รับประโยชน์จากชุดข้อมูลจากแหล่งข้อมูลต้นทางที่ได้มาตรฐาน และยืดหยุ่นกับการนำข้อมูลเปิดเผยไปใช้งานต่อได้

จากการค้นคว้างานวิจัยเครื่องมือจัดการข้อมูลเปิดโอเพ่นซอร์สที่ได้ทำการออกแบบระบบ ให้เหมาะสมและสามารถนำมาใช้งานกับหน่วยงานและสอดคล้องกับการพัฒนาระบบพบว่า มี 2 แพลตฟอร์มข้อมูลเปิดที่รองรับคุณสมบัติข้อมูลเปิดภาครัฐได้ดี คือ 1) CKAN และ 2) DKAN ทั้ง 2 แพลตฟอร์มข้อมูลเปิด มีคุณสมบัติที่ดีและคล้ายกันต่างกันที่หน่วยงานหรือองค์กรจะเลือกใช้และการพัฒนาระบบเชื่อมต่อกับระบบของหน่วยงาน หากเลือก CKAN โปรแกรมหลักจะถูกพัฒนาจากภาษา Python การนำไปใช้งานต่อหรือการพัฒนาหรือติดตั้งส่วนเสริมเพื่อรองรับการทำงานก็ต้องเป็นการเขียนโปรแกรมด้วยภาษา Python เป็นหลัก และ CKAN ยังได้เปรียบในแง่ของการใช้งานในหน่วยงานภาครัฐเนื่องจากมีส่วนเสริมที่สนับสนุนการพัฒนาระบบให้สอดคล้องและนำมาใช้งานได้ทันที ส่วนหากเลือก DKAN ก็จะแตกต่างที่ภาษาพื้นฐานในการพัฒนาระบบซึ่งใช้การเขียนโปรแกรมในภาษา PHP และมีการใช้งานร่วมกับโปรแกรม CMS ที่ชื่อ Drupal หากหน่วยงานมีการใช้งาน CMS ในส่วนนี้ DKAN ก็จะ

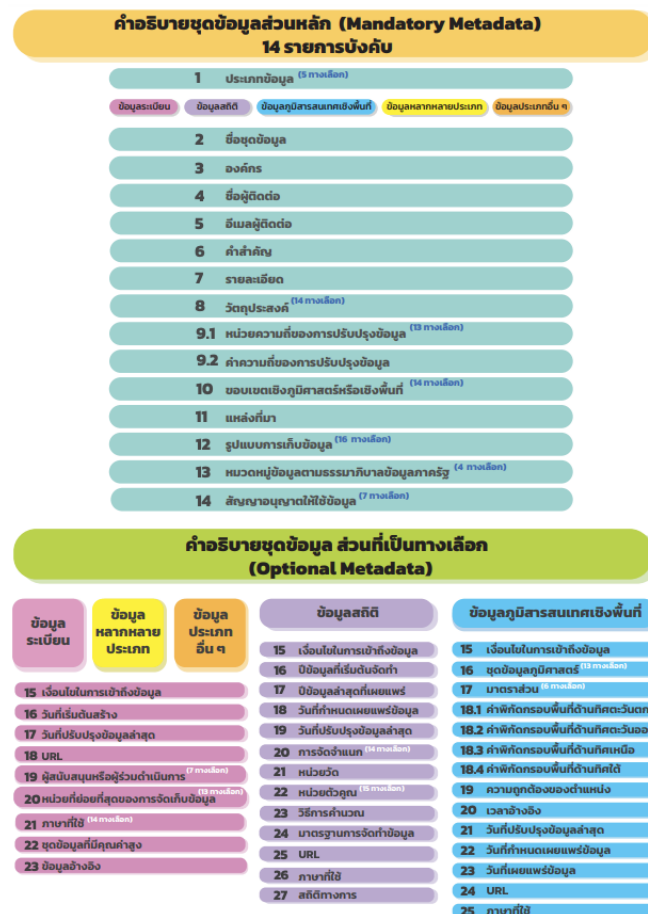
เหมาะกับการพัฒนาที่ใช้เวลาน้อยกว่าเนื่องจากตัวระบบสามารถเชื่อมต่อข้อมูลได้ทันที แต่ก็ยังต้อง
เสียเวลาปรับปรุงระบบให้รองรับการแชร์ข้อมูลข้ามแพลตฟอร์มกับหน่วยงานภาครัฐ ซึ่งรายละเอียด
ความสามารถรองรับคุณสมบัติที่ดีสำหรับข้อมูลเปิดภาครัฐสรุปได้ ดังนี้

ตารางที่ 1. ตารางคุณสมบัติของแพลตฟอร์มข้อมูลเปิด

คุณสมบัติ	คำอธิบาย	CKAN	DKAN	Github	Google Datasets	ODS	Socrata	Kaggle
เมทาดาทา (Metadata)	ข้อมูลโครงสร้าง ชุดข้อมูล คำอธิบาย อื่นๆ	+	+	+	+	+	+	+
เข้าถึงง่าย (Easy access)	มี UI ที่ใช้งานได้ ง่ายไม่ซับซ้อน	+	+		+	+	+	+
ความสามารถ ค้นหา (Searchability)	เข้าถึงได้จาก หลายวิธีการ ค้นหาชุดข้อมูล	+	+		+	+	+	+
อ่านด้วยเครื่อง (Machine- radable)	เป็นมาตรฐาน กลางรองรับชุด ข้อมูลสูงสุดระดับ 5 ดาว	+	+			+	+	
API	รองรับการ เชื่อมต่อชุดข้อมูล สู่ภายนอก	+	+		+	+	+	+
จำลองข้อมูลด้วย ภาพ (Data visualization)	วิเคราะห์ด้วย เครื่องมือภายใน ระบบได้	+	+			+	+	+
เชื่อมโยงระหว่าง แพลตฟอร์ม (Data linkage)	ทำงานร่วมกัน ระหว่าง แพลตฟอร์มได้	+	+			+	+	
ข้อมูลมีความ น่าเชื่อถือ (หน่วยงานกลาง)	ความน่าเชื่อถือ ของข้อมูลที่มี ให้บริการ (ภาครัฐ)	+	+			+	+	

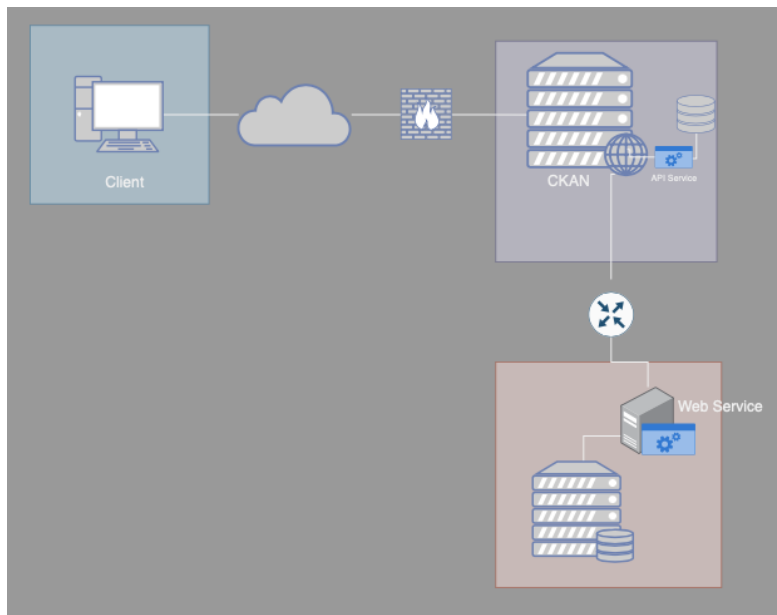
3. การออกแบบติดตั้งระบบ

จากการศึกษาวิจัยพบว่าจำเป็นต้องเตรียมองค์ประกอบ 3 ด้าน ได้แก่ 1) ข้อมูลเปิดเผยที่แปลงมาอยู่ในรูปแบบ raw data โดยกำหนดชุดข้อมูลตามโครงสร้าง Metadata และ Optional Metadata ซึ่งในหน่วยงานมีข้อมูลเปิดเผยทั้งหมด 11 ชุดข้อมูล และชุดทรัพยากรข้อมูล ตามธรรมชาติของกรมพัฒนาธุรกิจการค้า โดยการใช้ Template ในรูปแบบที่ CKAN สามารถนำชุดข้อมูลและรายละเอียดของชุดข้อมูลเข้าสู่ระบบ CKAN



ภาพที่ 1 คำอธิบายข้อมูลส่วนหลัก (Mandatory Metadata)

2) ทำการติดตั้งโปรแกรมสำหรับจัดการชุดข้อมูลเปิด ด้วยซอฟต์แวร์โอเพนซอร์ส CKAN และส่วนเสริม CKAN Open-D ผู้วิจัยได้กำหนดโครงสร้างและองค์ประกอบของระบบทั้งหมดเพื่อสนับสนุนการวิจัย



ภาพที่ 2 โครงสร้างและองค์ประกอบของระบบของการวิจัย

โดยระบบจะติดตั้งอยู่บน Virtual Machine (VM) บนระบบปฏิบัติการ Linux Ubuntu 20.04.5 LTS จากนั้นทำการติดตั้งโปรแกรมที่ทำงานในรูปแบบ Docker container และปรับแต่งระบบตัวระบบให้รองรับ การปรับปรุงข้อมูลด้วย API จากทรัพยากรข้อมูลที่มีการจัดเตรียมไว้ถูกส่งมายังเครื่องระบบ CKAN โดยระบบมีการติดตั้งพื้นที่สำหรับจัดเก็บข้อมูลสำหรับนำทรัพยากรข้อมูลที่ได้รับไปนำเข้าสู่ระบบ CKAN ดังนี้

2.1) ทำการจัดเตรียม ที่จัดเก็บข้อมูล API CKAN Script python และ Crontab เพื่อทำการปรับปรุงข้อมูลให้เป็นไปตามโครงสร้าง ชุดข้อมูล ทรัพยากรข้อมูล โดยมีการกำหนดดังนี้

2.1.1) กำหนด Path ที่จัดเก็บข้อมูล อยู่ใน /dataset และกำหนด Path ที่จัดเก็บ Script python สำหรับ เชื่อมต่อกับ API CKAN และไฟล์ข้อมูลที่จัดเก็บตามข้อ 1) โดยไฟล์จะอยู่ใน /home/bdexadmin/ckan-docker/script_crontab

2.1.2) API CKAN จะเป็นการสร้างด้วย Script Python เพื่อทำการ Post request ทำการสร้างชุดข้อมูลและทรัพยากรข้อมูลที่จัดเก็บตามข้อ 1)

2.1.3) การทำงานปรับปรุงชุดข้อมูลอาศัยการทำงานของ Crontab ซึ่งเป็นคำสั่งใน Linux Server สำหรับทำงานตามที่กำหนดซ้ำๆ โดยการทำงานจะเป็นลักษณะ Job งานตามที่กำหนดใน script ตัวอย่างเช่น ทรัพยากรข้อมูลของชุดข้อมูล “นิติบุคคลจดทะเบียนตั้งใหม่” ซึ่งมีการกำหนดข้อมูลเมทาดาตา (Metadata) ไว้ว่า ปรับปรุงข้อมูล ทุกๆ เดือน โดยจะต้องกำหนดการตั้งการทำงาน ดังนี้

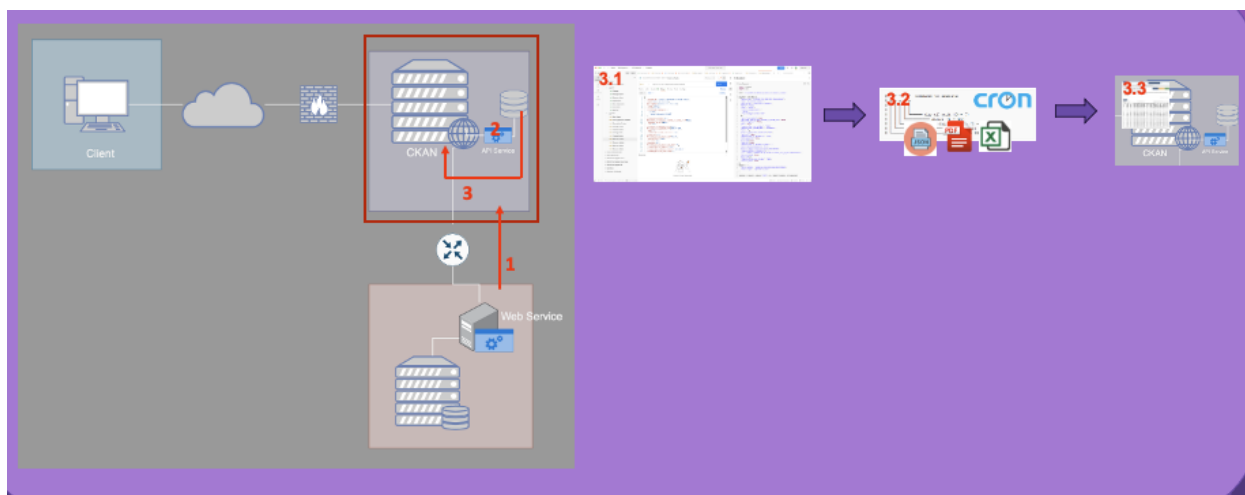
2.1.3.1) สร้าง Script โดยใช้คำสั่ง `crontab -e`

2.1.3.2) ทำการ เพิ่มคำสั่งนี้ลงไปในไฟล์ โดยกำหนดให้ ชุดข้อมูลนี้ที่ถูกกำหนดไว้

ใน `dataset_11_01.py` มีการเรียก Post request ด้วย Script python ของ API CKAN ไปทำการสร้างทรัพยากรข้อมูลที่ Path ที่กำหนดไว้ในข้อ 1) และจากนั้น ให้ทำการบันทึกการทำงานลงใน log file

```
* * * 1-12 * /usr/bin/python3 /home/bdexadmin/ckan-docker/script_crontab/dataset_11_01.py >>  
/home/bdexadmin/ckan-docker/script_crontab/log.txt >/dev/null 2>&1
```

2.1.3.3) ทรัพยากรข้อมูลจะถูกเพิ่มเข้าสู่ระบบ



ภาพที่ 3 แสดงขั้นตอนการทำงานของ API CKAN และ Script ในการสร้างทรัพยากรข้อมูล

เมื่อทำการติดตั้งเสร็จแล้วพบว่า ข้อมูลจากผู้เป็นเจ้าของข้อมูลทั้ง 11 ชุดข้อมูลสามารถทำการสร้างชุดข้อมูลที่เป็นมาตรฐานได้ กับอีกส่วนหนึ่งคือทรัพยากรข้อมูลก็จะถูกปรับแต่งให้ได้ข้อมูลตามโครงสร้าง Metadata และเพิ่มการปรับปรุงข้อมูลให้เป็นปัจจุบัน ทำให้ข้อมูลที่ต้องเปิดเผยสามารถเปิดเผยได้อย่างทันทีตามโครงสร้างที่มีการกำหนดร่วมกัน

4. ผลการวิจัย

ผลของการจัดทำชุดข้อมูลจากเดิมที่มีบางขั้นตอนไม่ได้มีระบบมาจัดการ ทำให้ข้อมูลเปิดของหน่วยงานอาจมีการพัฒนาเพื่อนำไปใช้งานจริงได้ไม่มีประสิทธิภาพ ได้แก่ การเปิดเผยข้อมูลตามธรรมาภิบาล หน่วยงานไม่มีรูปแบบข้อมูลที่เป็นมาตรฐาน ทำให้การจัดการนั้นมีความยากสำหรับผู้รวบรวมจัดทำและนำเสนอ และผู้ใช้งานก็ยังต้องทำความเข้าใจกับข้อมูลที่ได้รับ ข้อมูลเหล่านั้นก็อาจจะจะเป็นข้อมูลสรุปที่ไม่สามารถนำไปประมวลผลต่อได้รวมไปถึง ไม่ยืดหยุ่นกับการนำไปใช้งาน

การจัดทำข้อมูลเปิดของกรมพัฒนาธุรกิจการค้าที่ให้บริการตามมาตรฐานด้วยโปรแกรมโอเพนซอร์ส CKAN ประกอบไปด้วยชุดข้อมูลทั้ง 11 ชุด ดังนี้

- 1) นิติบุคคลจดทะเบียนตั้งใหม่
- 2) นิติบุคคลจดทะเบียนเลิกกิจการ
- 3) ข้อมูลทะเบียนนิติบุคคล
- 4) จำนวนการนิติบุคคลจดทะเบียนตั้งใหม่
- 5) ทุนจดทะเบียนนิติบุคคลตั้งใหม่
- 6) จำนวนการจดทะเบียนเลิกกิจการ
- 7) ทุนจดทะเบียนเลิกกิจการ
- 8) จำนวนนิติบุคคล
- 9) การลงทุนจากต่างชาติในนิติบุคคลไทย
- 10) จำนวนการอนุญาตให้คนต่างด้าวประกอบธุรกิจ (สะสม)
- 11) จำนวนการออกหนังสือรับรองการประกอบธุรกิจของคนต่างด้าว (สะสม)

เมื่อได้ชุดข้อมูลที่สามารถเปิดเผยได้จากหน่วยงานตามธรรมชาติของข้อมูลของหน่วยงาน กรมพัฒนาธุรกิจการค้า ซึ่งมีการแต่งตั้งคณะทำงานตามลำดับจากบนลงล่าง เพื่อกำหนด สิทธิ์หน้าที่ความรับผิดชอบ กระบวนการ มาตรการควบคุม การวัดผลการบริหารจัดการข้อมูล การจำแนกหมวดหมู่ข้อมูล และการกำหนด คำอธิบายชุดข้อมูล ซึ่งการนำระบบ CKAN เข้ามาใช้งานสำหรับการจัดการชุดข้อมูลเปิดเผยนั้น เกี่ยวข้องกับการ บริหารจัดการข้อมูล ซึ่งการบริหารจัดการข้อมูลนั้นครอบคลุม ขั้นตอนการสร้างข้อมูล การรวบรวมข้อมูล การจัดเก็บ การจัดเก็บถาวร การทำลายข้อมูล การประมวลผลข้อมูล การใช้ข้อมูล การแลกเปลี่ยน การ เชื่อมโยงข้อมูล และการเปิดเผยข้อมูลต่อสาธารณะ โดยตัวอย่างนี้จะเป็นการสรุปผลและเปรียบเทียบก่อนและ หลังใช้ระบบ CKAN ในการบริหารจัดการชุดข้อมูลเปิด ของชุดข้อมูลนิติบุคคลจดทะเบียนตั้งใหม่ ซึ่งเป็น 1 ใน 11 ชุดข้อมูล



ภาพที่ 4 สรุปผลเปรียบเทียบระหว่าง ระบบการจัดทำข้อมูลเปิด รูปแบบเดิม กับรูปแบบใหม่ โดยใช้ระบบ CKAN บริหารจัดการชุดข้อมูลเปิด ของชุดข้อมูลนิติบุคคลจดทะเบียนตั้งใหม่

จากภาพที่ 4 สามารถสรุปผลได้ ดังนี้

1) ขยายการกระจายของผู้จัดทำข้อมูลเปิดและเผยแพร่ข้อมูลเปิด ที่ประกอบไปด้วย กองทะเบียนธุรกิจ กองข้อมูลธุรกิจ และศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร มาจัดการผ่านระบบ CKAN โดยให้ระบบทำหน้าที่กำหนดสิทธิการเข้าถึงและใช้งานระบบที่เกี่ยวข้องในการจัดการชุดข้อมูลเปิด

2) นำธรรมชาติของข้อมูลของกรมพัฒนาธุรกิจการค้า ที่เดิมมีการนำมาใช้ปฏิบัติอยู่เดิมมาเพิ่มเติมมาตรฐานแนวทางการเปิดเผยข้อมูลเปิดภาครัฐในรูปแบบดิจิทัลต่อสาธารณะ มรด.-12001:2563 และมาตรฐานแนวทางการจัดทำบัญชีข้อมูลภาครัฐ มสพร. 1-2564 ซึ่งในประเทศไทย ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NETCTEC) โดยร่วมมือกับ สำนักงานสถิติแห่งชาติ และสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) (DGA) ปรับปรุงเพื่อให้ซอฟต์แวร์มีความสอดคล้องกับบริบทของประเทศ สำหรับใช้งานในประเทศ นำมาคิดปรับปรุงระบบ CKAN และส่วนเสริม CKAN Open-D รองรับการจัดทำชุดข้อมูลได้อย่างเป็นระบบ และสามารถนำเสนอและเชื่อมโยงข้อมูลระหว่างหน่วยงานภาครัฐได้ทันที

3) ลดภาระเจ้าของข้อมูลจะต้องเป็นผู้จัดทำ เนื่องจากชุดข้อมูลแต่ละชุดล้วนมีแหล่งที่มาจากหลายแหล่ง หลายโปรแกรม จัดทำ จัดเก็บ ตรวจสอบ ปรับปรุง รวมถึงการสำรองข้อมูล และทำลาย

ข้อมูล ในแต่ละส่วนงานของผู้เป็นเจ้าของข้อมูลทำให้หน่วยงานเดียวกันอาจมีหลายรูปแบบในขั้นตอนนี้ เมื่อมีระบบ CKAN สามารถรวมวิธีการปฏิบัติงานเป็นขั้นตอนที่เหมือนกันได้ ง่ายในการบริหารจัดการ

4) การรวบรวมแหล่งที่มาของข้อมูลพร้อมกับ รูปแบบชุดข้อมูลที่หน่วยงานจัดทำ ทำให้เกิดการประสานระหว่าง ผู้เป็นเจ้าของข้อมูลจัดทำ ผู้นำข้อมูลตรวจสอบนำเสนอ และเจ้าหน้าที่ด้านเทคนิคสามารถใช้ชุดข้อมูลร่วมกันเพื่อนำชุดข้อมูลของแต่ละหน่วยงานมาจัดเตรียมทรัพยากรข้อมูลและปรับปรุงข้อมูลได้ถูกต้องครบถ้วน

5) การนำเสนอข้อมูล สามารถบริการได้ผ่านระบบ CKAN ซึ่งผู้นำเสนอข้อมูลสามารถตรวจสอบ ตรวจสอบ รวมถึงปรับแต่งการนำเสนอได้ผ่านระบบ และผู้ใช้ข้อมูลก็สามารถนำข้อมูลไปใช้งานได้ อย่าง ถูกต้องตามโครงสร้างและคำอธิบายข้อมูลอีกทั้งระบบ CKAN ยังยืดหยุ่นในการนำไปใช้งาน โดยสามารถเชื่อมต่อกับ โปรแกรมภายนอก หรือเชื่อมโยงข้อมูลผ่าน API ได้

5. สรุปผลการวิจัย

การปรับปรุงการเปิดเผยข้อมูลของกรมพัฒนาธุรกิจการค้า โดยใช้มาตรฐานต่างๆที่เกี่ยวข้อง และการเลือกใช้ระบบ CKAN ที่ติดตั้งและมีการปรับปรุงพัฒนาขึ้นนี้มีศักยภาพในการจัดการข้อมูล ตั้งแต่ต้นทางที่ต้องกำหนดโครงสร้างในระดับหน่วยงานทั้งหมดเข้าสู่ระบบ และการปรับปรุงข้อมูลทรัพยากรข้อมูลให้สอดคล้องกับโครงสร้างโดยได้มีการจัดทำชุดคำสั่ง API ให้ระบบ CKAN รับข้อมูลเข้าสู่ระบบเพื่อนำไปเปิดเผยและยังสามารถให้บริการข้อมูลเปิดเผยได้ผ่านตัวระบบ CKAN โดยตรง หรือ Download ข้อมูล หรือ เชื่อมต่อเข้ากับ API ของระบบ ทำให้การเปิดเผยข้อมูลมีความยืดหยุ่นและสะดวกต่อการนำไปใช้งาน ซึ่งผู้ใช้งานก็สามารถศึกษาข้อมูลทั้งหมดได้จากโครงสร้างของชุดข้อมูลที่ต้องการนำไปใช้งาน ทั้งหมดนี้สามารถสรุปผลการวิจัยได้ดังนี้

ตัวชี้วัด	ก่อนงานวิจัย	ผลลัพธ์	หมายเหตุ
1. การจัดการข้อมูลเปิดตามรูปแบบมาตรฐาน - (มรค.-12001:2563) - (มสพร. 1-2564) - ธรรมชาติของข้อมูลภาครัฐของหน่วยงาน	- ชุดข้อมูลทั้ง 11 ชุดข้อมูล อยู่บนเว็บไซต์ กรมพัฒนาธุรกิจการค้า ยังไม่มีการจัดการข้อมูลให้อยู่ในรูปแบบข้อมูลเปิดที่เป็นมาตรฐาน	- มีการกำหนดโครงสร้างชุดข้อมูลเปิด ตามคุณลักษณะของข้อมูลเปิด ได้แก่ 1) ข้อมูลมีความพร้อมใช้ซึ่งเป็นชุดข้อมูลที่มาจากระบบการรวมมาภิบาลข้อมูลของหน่วยงาน 2) เป็นทรัพยากรข้อมูล ที่ได้จากแหล่งข้อมูลโดยตรง จากผู้ดูแลข้อมูล	การจัดทำชุดข้อมูลเป็นไปตามธรรมชาติของมาภิบาลข้อมูลภาครัฐของหน่วยงาน

		3) มีการกำหนดความถี่ในการปรับปรุงข้อมูลตามโครงสร้างเมทาดาดา โดยการใช้เครื่องมือ CKAN และ CKAN API สำหรับการจัดการ 4) ระดับการเปิดเผยข้อมูลที่สามารถอ่านได้ด้วยเครื่อง อยู่ในระดับ 3 ดาวเป็นอย่างน้อย 5) มีการจัดการแปลงข้อมูลให้มีความถูกต้อง เช่น ข้อมูลไม่ซ้ำกัน ไม่มีค่าว่าง 6) ไม่มีข้อมูลส่วนบุคคล 7) นำชุดข้อมูลของหน่วยงาน เข้าสู่ Template ตามมาตรฐานเมทาดาดา	
	1. นิติบุคคลจดทะเบียนตั้งใหม่ และนิติบุคคลจดทะเบียนเลิกกิจการ ชุดข้อมูลเดิมเป็น Excel File มีการระบุ เมทาดาดา แต่ยังไม่มีการปรับปรุงให้ข้อมูลเป็นปัจจุบัน แหล่งข้อมูลเดิม https://dbd.go.th/news_view.php?nid=469423520	ทำการแยกข้อมูลเป็น 2 ชุดข้อมูล 1) นิติบุคคลจดทะเบียนตั้งใหม่ 2) นิติบุคคลจดทะเบียนเลิกกิจการ	
	2. ข้อมูลทะเบียนนิติบุคคล แหล่งข้อมูลเดิม ชุดข้อมูลเดิมไม่มีให้บริการ	เรียกใช้ชุดข้อมูลที่หน่วยงานจัดทำขึ้นใหม่ในรูปแบบ API Data Method GET https://openapi.dbd.go.th/api/v1/juristic_person/{เลขทะเบียนนิติบุคคล} 3) ข้อมูลทะเบียนนิติบุคคล	
ตัวชี้วัด	ก่อนการวิจัย	ผลลัพธ์	หมายเหตุ
	3. จำนวนการจดทะเบียนนิติบุคคลตั้งใหม่ ทุนจดทะเบียนนิติบุคคลตั้งใหม่ และ จำนวนการจดทะเบียนเลิกกิจการ ทุนจดทะเบียนเลิกกิจการ ชุดข้อมูลนิติบุคคลรายจังหวัด เป็น Excel File ในรูปแบบข้อมูลสรุป ไม่สามารถนำมาประมวลผลใช้งานต่อได้ แหล่งข้อมูลเดิม https://dbd.go.th/more_news.php?cid=1459	ทำการแยกชุดข้อมูลออกเป็น 3 ชุดข้อมูล 4) จำนวนการจดทะเบียนนิติบุคคลตั้งใหม่ 5) ทุนจดทะเบียนนิติบุคคลตั้งใหม่ 6) จำนวนการจดทะเบียนเลิกกิจการ 7) ทุนจดทะเบียนเลิกกิจการ	
	4. จำนวนนิติบุคคล แหล่งข้อมูลเดิม ชุดข้อมูลเดิมไม่มีให้บริการ	เรียกใช้ชุดข้อมูลที่หน่วยงานจัดทำขึ้นใหม่ในรูปแบบ API Data Method GET https://openapi.dbd.go.th/api/v1/	

		juristic_person_operating/{ปี พ.ศ. เช่น 2566} 8) จำนวนนิติบุคคล	
	5. การลงทุนจากต่างชาติในนิติบุคคลไทย แหล่งข้อมูลเดิม ชุดข้อมูลเดิมไม่มีให้บริการ	เรียกใช้ชุดข้อมูลที่หน่วยงานจัดทำขึ้นใหม่ ในรูปแบบ API Data Method GET https://openapi.dbd.go.th/api/v1/invest_val_by_nat/{ปี พ.ศ. เช่น 2566} 9) การลงทุนจากต่างชาติในนิติบุคคลไทย	
	6. จำนวนการอนุญาตให้คนต่างด้าวประกอบธุรกิจ (สะสม) และจำนวนการออกหนังสือรับรองการประกอบธุรกิจของคนต่างด้าว (สะสม) ชุดข้อมูลเดิมเป็น .pdf File จัดทำเป็นข้อมูลสรุปรายปี แหล่งข้อมูลเดิม https://dbd.go.th/more_news.php?cid=721	ทำการแยกชุดข้อมูลออกเป็น 2 ชุดข้อมูล 10) จำนวนการอนุญาตให้คนต่างด้าวประกอบธุรกิจ (สะสม) 11) จำนวนการออกหนังสือรับรองการประกอบธุรกิจของคนต่างด้าว (สะสม)	
2. การเลือกใช้เครื่องมือจัดการชุดข้อมูลเปิด	- มีระบบแยกกันแต่ละส่วนงานสำหรับจัดการ ใช้วิธีการจัดเก็บข้อมูลตามธรรมชาติของข้อมูลของหน่วยงานในแต่ละส่วนงาน และจัดทำข้อมูลเพื่อให้บริการบนเว็บไซต์ของกรมพัฒนาธุรกิจการค้า ผู้ใช้บริการต้องค้นหา และทำความเข้าใจข้อมูลตามรูปแบบที่จัดทำ	- มีการเลือกใช้เครื่องมือจัดการข้อมูลเปิดโอเพนซอร์ส CKAN และส่วนเสริม CKAN Open-D ที่มี การปรับปรุงระบบให้สอดคล้องกับบริบท เป็นไปตามมาตรฐาน รองรับภาษาไทย เพิ่มประสิทธิภาพ เช่นการจัดการข้อมูล เครื่องมือสนับสนุนการ กำหนดสิทธิ การจัดการข้อมูลและการนำเสนอ ข้อมูล และมีกระบวนการปรับปรุงข้อมูลให้เป็นไป ตามความถี่ตามโครงสร้างเมทาดาทาข้อมูล	
ตัวชี้วัด	ก่อนการวิจัย	ผลลัพธ์	หมายเหตุ
3. การเรียกใช้ชุดข้อมูล	- ไม่สะดวกต่อการค้นหาข้อมูล และการนำข้อมูลไปใช้คือ บางชุดข้อมูลเป็นข้อมูลรายละเอียด บางชุดข้อมูลเป็นข้อมูลสรุป และบางชุดข้อมูลอยู่ในรูปแบบ .pdf File เนื่องจากข้อมูลที่ใช้บริการมีการจัดทำที่ไม่ได้อยู่ในรูปแบบมาตรฐานข้อมูลเปิดที่ง่ายต่อการเผยแพร่ และสามารถแบ่งปันให้บริการชุดข้อมูล และทรัพยากรข้อมูล หรือไม่สามารถอ่านได้ด้วยเครื่อง เพื่อนำไปประมวลผลต่อได้	1) สามารถนำเสนอชุดข้อมูล โครงสร้างข้อมูลเมทาดาทา พจนานุกรมข้อมูล ให้ผู้ใช้งานสามารถศึกษา และทำความเข้าใจก่อนนำข้อมูลไปใช้งานต่อ 2) สามารถนำเสนอชุดข้อมูลทรัพยากรข้อมูลในระดับข้อมูลเปิด 3 ดาว ที่สามารถนำข้อมูลไปประมวลผลต่อได้ 3) มีช่องทางให้บริการข้อมูลได้แก่ - ผ่านหน้าระบบ CKAN - ผ่านเรียกใช้ Data API สามารถเชื่อมต่อทดสอบผ่าน Postman เพื่อทดสอบผลการเรียกดู	

		ทรัพยากรข้อมูล หรือ โปรแกรม Power BI สำหรับ การนำเสนอข้อมูลแบบ Data Visualization - Data Visualization รูปแบบ Data Playground บนระบบ CKAN	
--	--	--	--

บรรณานุกรม

ภาษาไทย

ธรรมาภิบาลข้อมูลภาครัฐ (Data Governance for Government)

<https://standard.dga.or.th/government-website-standard/1925/>

ธรรมาภิบาลข้อมูลภาครัฐ ของกรมพัฒนาธุรกิจการค้า. Retrieved from

https://www.dbd.go.th/more_news.php?cid=1889

มาตรฐานรัฐบาลดิจิทัล ว่าด้วยแนวทางการเปิดเผยข้อมูลเปิดภาครัฐในรูปแบบดิจิทัลต่อสาธารณะ (Open Government Data Guideline)(มรด-12001:2563) <https://standard.dga.or.th/dg-std/2028/>

มาตรฐานของสำนักงานพัฒนารัฐบาลดิจิทัล ว่าด้วยแนวทางการจัดทำบัญชีข้อมูลภาครัฐ. (2564)

Retrieved from https://www.dga.or.th/wp-content/uploads/2021/11/1.-Final_GD-Catalog-Guideline-v.1.0_16032564.pdf

ระบบบัญชีข้อมูลภาครัฐ (Government data catalog). Retrieved from <https://gdhelppage.nso.go.th/index.html>

ภาษาต่างประเทศ

A comprehensive review of open data platforms, prevalent technologies, and functionalities Retrieved from

https://www.researchgate.net/publication/365355169_A_comprehensive_review_of_open_data_platforms_prevalent_technologies_and_functionalities
<https://ckan.org/government>

Joshua Tauberer. (2019). The 8 Principles of Open Government Data. Retrieved from

<https://opengovdata.org>

Technology Options. Commonly Used Open Data Platforms

Retrieved from <https://opendatatoolkit.worldbank.org/en/technology.html>

World Wide Web Foundation. (2015) . ODB Global Report. 3rd ed.

Retrieved from <http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-GlobalReport.pdf>