

การบริการชุดข้อมูลงานวิจัยในสถาบันบริการสารสนเทศ

จัดทำโดย

นางสาวเบญจมาศ แสนหลวง รหัสนิสิต 65199160064

เสนออาจารย์ประจำรายวิชา

ผศ.ดร. ศศิพิมล ประพินพงศกร

บทความนี้เป็นส่วนหนึ่งของรายวิชา สศ631 สัมมนาวิชาชีพสารสนเทศ

หลักสูตรศศ.ม. สาขาวิชาสารสนเทศศึกษา

คณะมนุษยศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ปีการศึกษา 2565

# การบริการชุดข้อมูลงานวิจัยในสถาบันบริการสารสนเทศ

## Research dataset service in information institution

เบญจมาศ แสนหลวง<sup>1</sup>

### บทคัดย่อ

ชุดข้อมูลการวิจัยถือเป็นองค์ประกอบสำคัญของการวิจัย โดยขอบเขตของคำว่าชุดข้อมูลการวิจัยนั้นขึ้นอยู่กับวัตถุประสงค์ในการนำไปใช้ประโยชน์ขององค์กรหรือเจ้าของชุดข้อมูลซึ่งอาจอยู่ในรูปแบบของข้อมูลดิบหรือข้อมูลที่ถูกจัดจำแนกเฉพาะเรื่องก็ย่อมได้ (Renear, Sacchi, Wickett, 2010) ในปัจจุบันการแบ่งปันชุดข้อมูลงานวิจัยเป็นที่แพร่หลายมากขึ้นโดยเฉพาะในช่วงของการแพร่ระบาดของไวรัสโควิด-19 ที่ยิ่งสะท้อนให้เห็นถึงความสำคัญของการแบ่งปันชุดข้อมูลการวิจัยทางด้านการแพทย์ที่เกี่ยวข้องกับโรคและการรักษา รวมถึงชุดข้อมูลเกี่ยวกับวัคซีนเพื่อการพัฒนาาร่วมกัน และส่งผลให้ชุดข้อมูลการวิจัยในสาขาอื่นเกิดการพัฒนาและถูกเล็งเห็นถึงความสำคัญมากยิ่งขึ้น ชุดข้อมูลการวิจัยที่ให้บริการตามแหล่งสืบค้นหรือเว็บไซต์ของหน่วยงานมีหลายประเภทขึ้นอยู่กับวิธีการจัดจำแนกซึ่งสามารถแบ่งได้ตามลักษณะข้อมูล บริบทของข้อมูล โครงสร้างของข้อมูล และการนำไปใช้ประโยชน์ โดยวิธีการจัดจำแนกจะขึ้นอยู่กับลักษณะข้อมูลและความต้องการของผู้ใช้ บทความนี้แสดงตัวอย่างของการบริการชุดข้อมูลงานวิจัยในสถาบันบริการสารสนเทศที่สามารถพบได้ทั้งในหน่วยงานของไทยและต่างประเทศทั้งภาครัฐและเอกชน (Wipawin, 2019)

**คำสำคัญ:** ชุดข้อมูลการวิจัย, การบริการชุดข้อมูล, การจัดการชุดข้อมูลการวิจัย, สถาบันบริการสารสนเทศ

**Keyword:** Research dataset, Dataset service, Research data management, Information institute

### บทนำ

การให้บริการชุดข้อมูลงานวิจัยเป็นบริการที่เกิดขึ้นจากอิทธิพลของการวิจัยแบบเปิด (Open research) และข้อมูลแบบเปิด (Open data) ตลอดจนนโยบายการเปิดเผยข้อมูลภาครัฐของไทย (Open government data) ที่กล่าวถึงการเผยแพร่ข้อมูลการทำงานของหน่วยงานรัฐเพื่อให้ประชาชนสามารถเข้าถึงข้อมูลและตรวจสอบความโปร่งใสในการดำเนินงาน ดังนั้นถือได้ว่าการให้บริการชุดข้อมูล (Dataset service) เป็นกระแสการพัฒนางานวิจัยและกระบวนการวิจัยโดยใช้ชุดข้อมูลเป็นฐานซึ่งเป็นสิ่งที่กำลังได้รับความสนใจและถูกเล็งเห็นถึงความสำคัญมากยิ่งขึ้น ชุดข้อมูลงานวิจัยอาจเป็นข้อมูลดิบที่เกิดจากขั้นตอนหรือกระบวนการวิจัยหรือเป็นชุดข้อมูลที่ผ่านการวิเคราะห์และสังเคราะห์ขึ้นเพื่อใช้ประกอบการวิจัย ซึ่งชุดข้อมูล

---

<sup>1</sup> นิสิตระดับบัณฑิตศึกษา หลักสูตรศิลปศาสตรมหาบัณฑิต สาขาวิชาสารสนเทศศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ, บรรณารักษ์ ฝ่ายสารสนเทศ สำนักหอสมุด มหาวิทยาลัยเกษตรศาสตร์

การวิจัยถือเป็นข้อมูลที่สามารถนำไปใช้ประโยชน์ซ้ำเพื่อต่อยอดงานวิจัยใหม่ได้ อีกทั้งยังสามารถใช้เป็นเครื่องมือในการตรวจสอบย้อนกลับคุณภาพของงานวิจัย (Cox & Verbann, 2018)

ในช่วงสถานการณ์การแพร่ระบาดของ COVID-19 การแบ่งปันชุดข้อมูลวิจัยเริ่มมีบทบาทและมีความสำคัญมากขึ้น น้าทิพย์ วิภาวิน (2563) กล่าวถึงชุดข้อมูลวิจัยแบบเปิดไวรัสโควิด-19 ที่เกิดจากการแบ่งปันชุดข้อมูลวิจัยทางการแพทย์ เพื่อทำความเข้าใจและศึกษาข้อมูลเกี่ยวกับไวรัส การพัฒนาวัคซีน อาการของโรค และวิธีการรักษาร่วมกัน เพื่อให้ได้มาซึ่งวิธีการรักษาโรคที่เหมาะสมและเป็นไปอย่างมีประสิทธิภาพ ซึ่งจากสถานการณ์ดังกล่าวสะท้อนให้เห็นว่าการแบ่งปันชุดข้อมูลมีส่วนช่วยสนับสนุนการแก้ปัญหาโรคระบาดที่เกิดขึ้น อีกทั้งนักวิจัยหรือสถาบันบริการสารสนเทศยังสามารถวิเคราะห์ข้อมูลจากชุดข้อมูลที่มีอยู่เพื่อพัฒนาวัคซีนและวิธีการรักษาที่เหมาะสมกับโรคที่เกิดจากไวรัสโควิด-19 (Tsung, et al., 2023)

บริบทของสถาบันบริการสารสนเทศในไทยให้บริการสารสนเทศงานวิจัย (Research Data Services) ในรูปแบบของบริการสนับสนุนการค้นคว้าและวิจัย ซึ่งเป็นบริการในลักษณะของการตอบคำถามและช่วยค้นคว้า หรือการฝึกอบรมใช้ฐานข้อมูลและเครื่องมือสนับสนุนการวิจัย (น้าทิพย์ วิภาวิน, 2559) จากการศึกษาของ จีรวรรณ และคณะ (2565) พบว่าสถาบันบริการสารสนเทศงานวิจัยของไทยยังไม่ได้มีการกำหนดนโยบายการให้บริการเฉพาะชุดข้อมูลวิจัยอย่างเป็นทางการเป็นรูปธรรม แต่มุ่งเน้นการบริการในลักษณะของการตอบคำถามและช่วยค้นคว้างานวิจัยเป็นหลัก หรือในบางหน่วยงานอาจมีการบริการชุดข้อมูลการวิจัยแต่ไม่ทราบว่าข้อมูลที่ให้บริการนั้นถือเป็นชุดข้อมูลงานวิจัยประเภทหนึ่ง ซึ่งถือเป็นข้อควรตระหนักในการสร้างความรู้ความเข้าใจเกี่ยวกับความสำคัญของชุดข้อมูลงานวิจัยให้แก่ผู้ให้บริการและผู้ใช้บริการทั้งในแง่ของการพัฒนา การนำไปใช้ประโยชน์ การเชื่อมโยงและการแบ่งปันชุดข้อมูลร่วมกัน

### ความหมายของชุดข้อมูล

Dataset หรือ Data set แปลว่า ชุดข้อมูล หมายถึง ข้อมูลที่ถูกรวบรวมหรือจัดกลุ่มตามความคล้ายคลึงกันหรือความเกี่ยวข้องกันและมีความเฉพาะเจาะจงในสาขาวิชาหรือเฉพาะเรื่อง (University of central florida, 2020) เช่น ชุดข้อมูลการแพร่ระบาดของไวรัสโควิด-19 ชุดข้อมูลสำมะโนประชากร ชุดข้อมูลการส่งออกสินค้า ชุดข้อมูลผลงานวิจัยด้านสิ่งแวดล้อม เป็นต้น โดยชุดข้อมูลจะประกอบด้วยข้อมูลหลายประเภท เช่น ข้อมูลตัวเลข ตัวอักษร รูปภาพ วิดีโอ และสื่อเสียง (Siman, 2023) ในขณะเดียวกันก็หมายรวมถึงคอลเล็กชันของข้อมูลดิบที่เกิดจากการศึกษาและวิจัย (National university library, 2023) สามารถจัดเก็บอยู่ในไฟล์รูปแบบต่างๆ เช่น CSV JSON และ SQL เป็นต้น ลักษณะทั่วไปของชุดข้อมูลมี 3 ประการ ได้แก่ มิติของชุดข้อมูล ระยะห่างหรือการกระจายตัวของชุดข้อมูล และความละเอียดของชุดข้อมูล หากเฉพาะเจาะจงเพียงชุดข้อมูลที่ปรากฏในงานวิจัย ชุดข้อมูลนั้นจึงถูกเรียกว่าชุดข้อมูลงานวิจัยหรือชุดข้อมูลวิจัย ซึ่งเป็นองค์ประกอบหนึ่งในข้อมูลการวิจัยและกระบวนการวิจัย

## ประเภทของชุดข้อมูลงานวิจัย

ประเภทของชุดข้อมูลงานวิจัยสามารถจัดแบ่งได้หลายแง่มุมขึ้นอยู่กับลักษณะข้อมูลและการนำไปใช้ประโยชน์ ยกตัวอย่างเช่น ในช่วงสภาวะการแพร่ระบาดของไวรัสโควิด-19 ชุดข้อมูลที่ถูกนำมาใช้ในการศึกษาวิจัยจะเป็นชุดข้อมูลที่มีความหลากหลาย เช่น ชุดข้อมูลตัวเลขที่แสดงจำนวนการติดเชื้อ การเสียชีวิต ชุดข้อมูลข้อความที่ให้รายละเอียดเกี่ยวกับอาการ และชุดข้อมูลเชิงพื้นที่ที่สะท้อนให้เห็นว่าปริมาณการติดเชื้อของแต่ละพื้นที่ เป็นต้น โดยชุดข้อมูลสามารถแบ่งได้ตามโครงสร้างข้อมูล ชนิดข้อมูล การใช้ประโยชน์ในเชิงสถิติ และชุดข้อมูลที่น่าไปใช้ประโยชน์กับ Machine Learning (Siman, 2023) รายละเอียดดังตารางที่ 1

ตารางที่ 1 ประเภทชุดข้อมูลงานวิจัย

วิธีการแบ่งประเภทชุดข้อมูล	ประเภทข้อมูล	คำอธิบาย
แบ่งตามลักษณะข้อมูล	ชุดข้อมูลตัวเลข	ตัวเลขที่ถูกใช้สำหรับการวิเคราะห์เชิงปริมาณ
	ชุดข้อมูลข้อความ	ตัวอักษร ข้อความและเอกสาร
	ชุดข้อมูลมัลติมีเดีย	รูปภาพ วิดีโอ และไฟล์เสียง
	ชุดข้อมูลอนุกรมเวลา	ข้อมูลที่รวบรวมในช่วงเวลาหนึ่งเพื่อวิเคราะห์แนวโน้มและระยะเวลาของแต่ละเหตุการณ์
	ชุดข้อมูลเชิงพื้นที่	ข้อมูลอ้างอิงทางภูมิศาสตร์ เช่น ข้อมูล GPS
แบ่งตามบริบทข้อมูล (ข้อมูลเชิงสถิติ)	ชุดข้อมูลเชิงตัวเลข	ข้อมูลตัวเลขเท่านั้น
	ชุดข้อมูลสองตัวแปร	ชุดข้อมูลที่ประกอบด้วยข้อมูลสองชุดที่เชื่อมโยงกัน
	ชุดข้อมูลหลายตัวแปร	ชุดข้อมูลที่ประกอบด้วยข้อมูลมากกว่าสองชุดที่เชื่อมโยงกัน
	ชุดข้อมูลเชิงกลุ่ม	ชุดข้อมูลที่มีข้อความกำกับเพื่อแยกประเภทข้อมูล
	ชุดข้อมูลเชิงสัมพันธ์	ชุดข้อมูลที่ประกอบด้วยตัวแปรที่มีความสัมพันธ์กัน

แบ่งตามโครงสร้างข้อมูล	ชุดข้อมูลที่มีโครงสร้าง	ชุดข้อมูลที่ผ่านการจัดระเบียบข้อมูลอย่างมีโครงสร้างเพื่อให้ง่ายต่อการค้นหา การนำไปใช้ประโยชน์ต่อและการวิเคราะห์ข้อมูล
	ชุดข้อมูลที่ไม่มีโครงสร้าง	ชุดข้อมูลที่ไม่ได้จัดแบ่งอย่างชัดเจนซึ่งไม่สามารถนำไปใช้ได้ทันที เช่น ข้อมูลการสนทนา
	ชุดข้อมูลแบบผสมผสาน	ชุดข้อมูลที่มีทั้งข้อมูลที่มีโครงสร้างและไม่มีโครงสร้าง
แบ่งตามการใช้ประโยชน์ เช่น การฝึก Machine learning (ML)	ชุดข้อมูลสำหรับการฝึก	ชุดข้อมูลที่ใช้เพื่อฝึกโมเดลให้เรียนรู้จากชุดข้อมูล
	ชุดข้อมูลสำหรับการตรวจสอบ	ชุดข้อมูลที่ใช้เพื่อลดข้อผิดพลาดที่เกิดจากการเรียนรู้และทำให้โมเดลมีความแม่นยำมากขึ้น
	ชุดข้อมูลสำหรับการทดสอบ	ชุดข้อมูลที่ใช้สำหรับทดสอบผลลัพธ์สุดท้ายของโมเดลเพื่อยืนยันความถูกต้อง

### ความสำคัญของการบริการชุดข้อมูลงานวิจัย

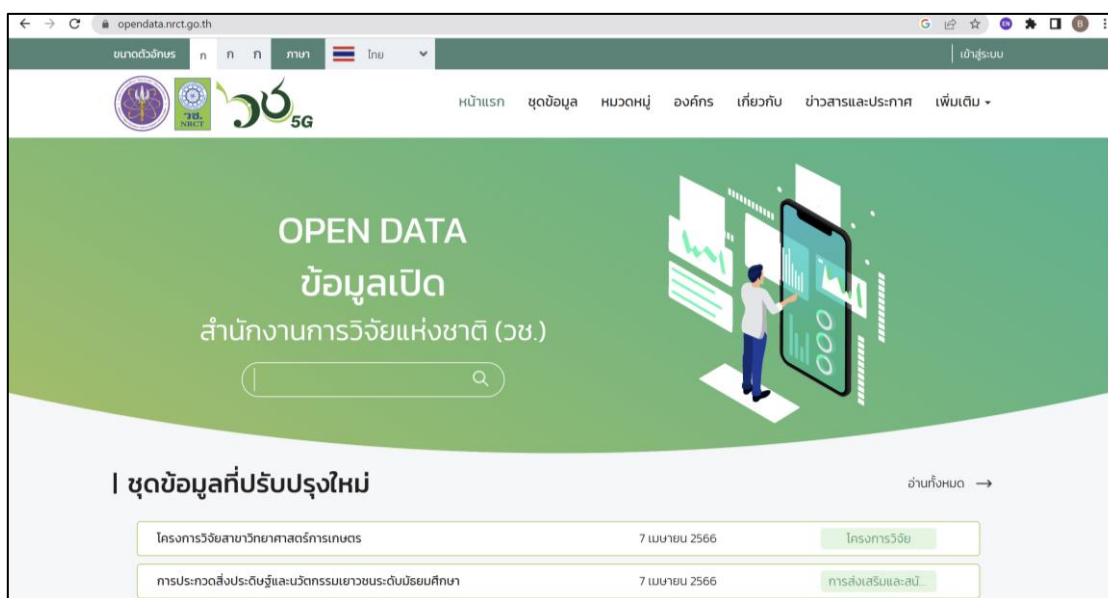
ในหน่วยงานภาครัฐการบริการชุดข้อมูลถือเป็นข้อกำหนดตามนโยบายของสำนักงานพัฒนารัฐบาลดิจิทัล (2564) กำหนดให้หน่วยงานภาครัฐเปิดเผยชุดข้อมูลเพื่อความโปร่งใสและเพื่ออำนวยความสะดวกและลดช่องว่างของการดำเนินงานร่วมกันทั้งระหว่างหน่วยงานและภายในหน่วยงาน ซึ่งสถาบันบริการสารสนเทศ ในฐานะของหน่วยงานสนับสนุนการศึกษาและวิจัยเห็นว่าการบริการชุดข้อมูลงานวิจัยเป็นการขยายผลการใช้ประโยชน์และต่อยอดชุดข้อมูลที่ถูกรวบรวมและผ่านการวิเคราะห์และสังเคราะห์เพื่อให้ข้อมูลนั้นถูกนำไปใช้ซ้ำในอนาคตซึ่งถือเป็นข้อมูลทฤษฎี (Gregory, K, 2020) การให้บริการชุดข้อมูลงานวิจัยนั้นเป็นการอำนวยความสะดวกให้แก่กวิจัยในการเก็บรวบรวมข้อมูล เนื่องจากชุดข้อมูลบางเรื่องสามารถเข้าร่วมกันได้ เช่น ชุดข้อมูลชื่อจังหวัด ชุดข้อมูลชื่อมหาวิทยาลัยของรัฐ ชุดข้อมูลเกี่ยวกับสถิติการใช้และความพึงพอใจ เป็นต้น นอกจากนี้การบริการชุดข้อมูลการวิจัยยังก่อให้เกิดกระบวนการตรวจสอบซ้ำคุณภาพหรือผลลัพธ์ของงานวิจัย ซึ่งถือเป็นการยกระดับความน่าเชื่อถือของงานวิจัยที่ถูกสร้างสรรค์ขึ้น

### ประโยชน์ของการบริการชุดข้อมูลงานวิจัย

1. การให้บริการชุดข้อมูลงานวิจัยถือเป็นการเปิดเผยข้อมูลอันเป็นวัตถุดิบสำคัญในการศึกษาและวิจัย ดังนั้นการเปิดเผยชุดข้อมูลจะช่วยเป็นเครื่องกระตุ้นถึงความเชื่อถือได้ของงานวิจัย รวมถึงเป็นการวางกรอบของการดำเนินการวิจัยและการเก็บข้อมูลให้มีคุณภาพมากขึ้น

2. ถือเป็นการใช้ประโยชน์ซ้ำจากชุดข้อมูลงานวิจัยเพื่อการต่อยอดงานวิจัยซึ่งหนึ่งชุดข้อมูลการวิจัยอาจถูกนำไปต่อยอดด้วยกระบวนการใหม่ที่ก่อให้เกิดนวัตกรรมและผลลัพธ์ที่แตกต่าง
3. เกิดสังคมที่ใช้สารสนเทศเป็นฐานซึ่งเกิดจากจากการแบ่งปันชุดข้อมูลร่วมกันเพื่อใช้ประกอบการตัดสินใจ เช่น การวางแผนธุรกิจ การหาแนวโน้มในการพัฒนาสินค้าและบริการ การวางแผนทางการเงิน (Gebru, et al, 2018) การศึกษาค้นคว้าเกี่ยวกับการรักษาโรคที่เกิดจากไวรัสโควิด-19
4. ช่วยให้การทำงานและการแลกเปลี่ยนข้อมูลระหว่างหน่วยงานหรือภายในหน่วยงานเป็นไปอย่างมีระบบและสะดวกมากยิ่งขึ้น
5. อำนวยความสะดวกให้นักวิจัยในการเก็บรวบรวมข้อมูลที่มีความซ้ำซ้อนกับข้อมูลเดิมที่มีอยู่และเป็นข้อมูลที่มีความถี่ในการเปลี่ยนแปลงน้อยมาก เช่น ข้อมูลรายชื่อจังหวัด ข้อมูลรายชื่อประเทศ เป็นต้น

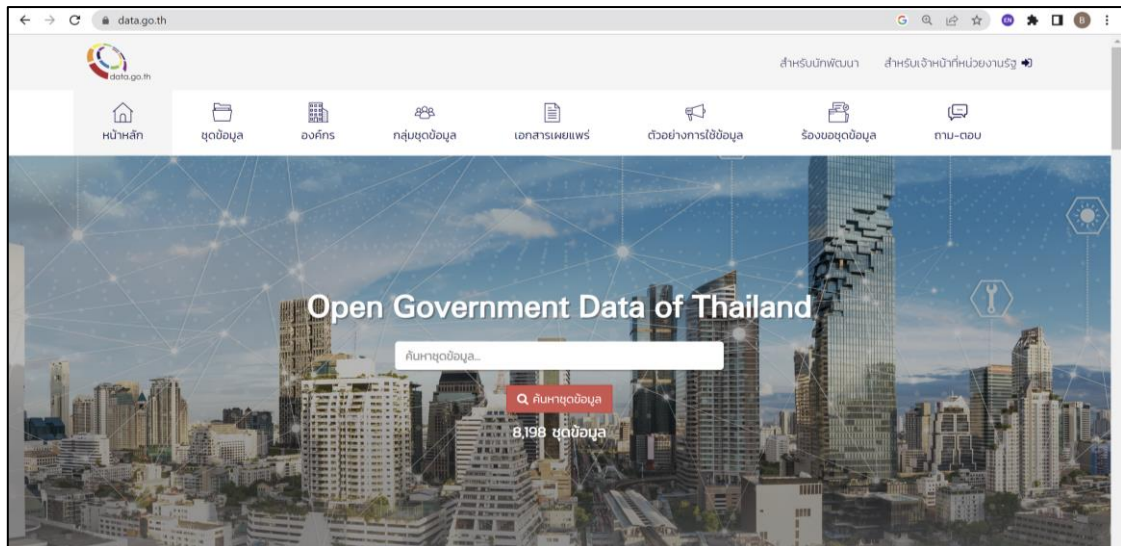
### การบริการชุดข้อมูลงานวิจัยในสถาบันบริการสารสนเทศ/หน่วยงาน/องค์กร



ภาพที่ 1 เว็บไซต์ข้อมูลเปิดของสำนักงานการวิจัยแห่งชาติ (วช.)

ที่มา: <https://opendata.nrct.go.th/>

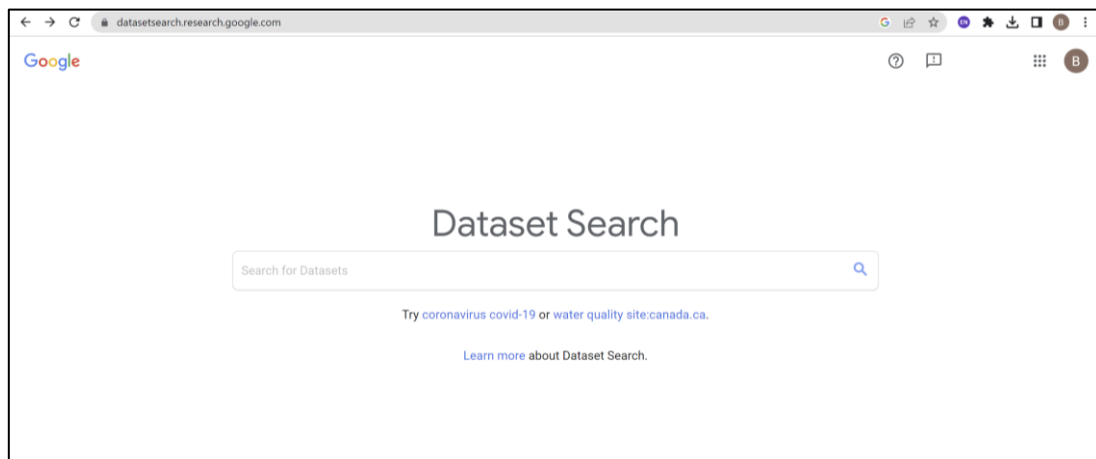
เว็บไซต์ข้อมูลเปิดของสำนักงานการวิจัยแห่งชาติรวบรวม เผยแพร่และให้บริการชุดข้อมูลงานวิจัยอันเป็นข้อมูลที่เกิดขึ้นตามอำนาจหน้าที่ของสำนักงานการวิจัยแห่งชาติ (วช.) ที่ว่าด้วยการให้ทุน การวิจัย การจัดทำฐานข้อมูลและดัชนีด้านวิทยาศาสตร์วิจัยและนวัตกรรม การสนับสนุนการดำเนินงานของโครงการวิจัยและนวัตกรรมของประเทศ ฯลฯ ซึ่งข้อมูลข้างต้นถือเป็นสินทรัพย์สำคัญที่สามารถนำไปใช้สนับสนุนการวิจัยและการคิดค้นนวัตกรรมในอนาคตได้ โดยชุดข้อมูลถูกจัดหมวดหมู่ตามลักษณะของข้อมูลที่สำนักงานการวิจัยแห่งชาติจัดเก็บ เช่น โครงการวิจัย กรอบการวิจัย มาตรฐานการวิจัย และผลงานวิจัย เป็นต้น ในปัจจุบันมีชุดข้อมูลการวิจัยให้บริการอยู่ประมาณ 172 ชุด (ข้อมูล ณ วันที่ 11 เม.ย. 2566)



ภาพที่ 2 ชุดข้อมูลเปิดของภาครัฐของไทย

ที่มา: <https://data.go.th/>

Open Government Data of Thailand อยู่ในความรับผิดชอบของสำนักงานพัฒนารัฐบาลดิจิทัล (สปร.) มีหน้าที่ในการบริหารจัดการเว็บไซต์เพื่อให้บริการแก่ประชาชน ชุดข้อมูลที่จัดเก็บมาจากหน่วยงานและองค์กรภาครัฐซึ่งพบว่ามีทั้งชุดข้อมูลที่เกิดจากการดำเนินงานและชุดข้อมูลที่เกิดจากการวิจัยและถูกจัดแบ่งเป็นหมวดหมู่ ตามเนื้อหาของชุดข้อมูล เช่น ศาสนา ศิลปะและวัฒนธรรม การศึกษา สาธารณสุข เกษตรกรรม การเมืองและการปกครอง ทรัพยากรธรรมชาติและสิ่งแวดล้อม เป็นต้น ปัจจุบันมีชุดข้อมูลให้บริการจำนวน 8,200 ชุด (ข้อมูล ณ วันที่ 11 เม.ย. 2566)

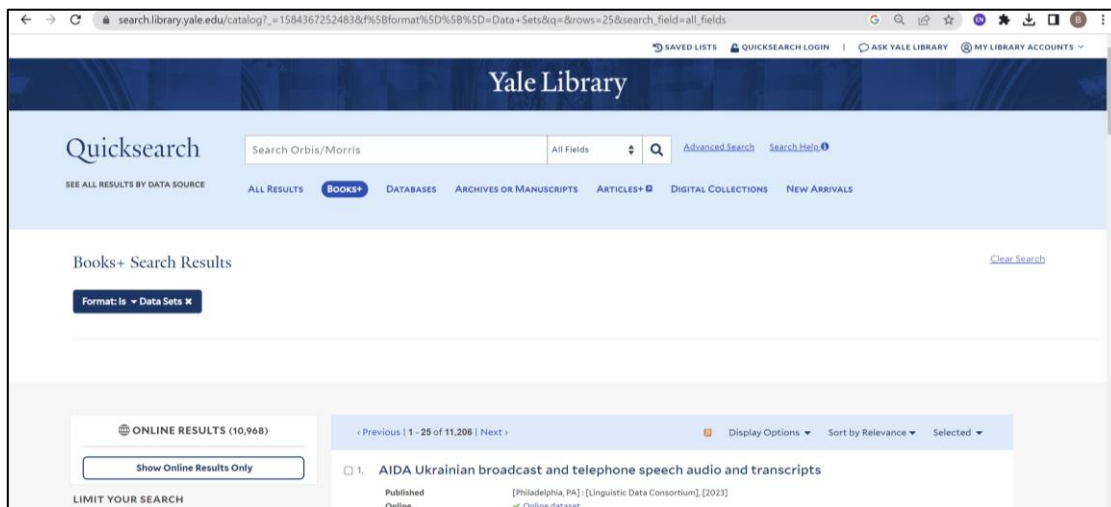


ภาพที่ 3 เครื่องสืบค้น Dataset search ของ Google

ที่มา: <https://datasetsearch.research.google.com/>

Dataset search เป็นเครื่องมือช่วยสืบค้นชุดข้อมูลที่พัฒนาขึ้นโดยบริษัท Google วิธีการใช้งานจะใช้หลักการเดียวกันกับการสืบค้นด้วย Search engines ทว่าผลลัพธ์ที่ได้นั้นจะให้เฉพาะข้อมูลที่เป็นชุดข้อมูลเท่านั้น ซึ่งชุดข้อมูลที่ให้บริการจะเป็นข้อมูลสหสาขาวิชาที่นำเข้าโดยเจ้าของชุดข้อมูลให้บริการทั้งข้อมูลในเชิง

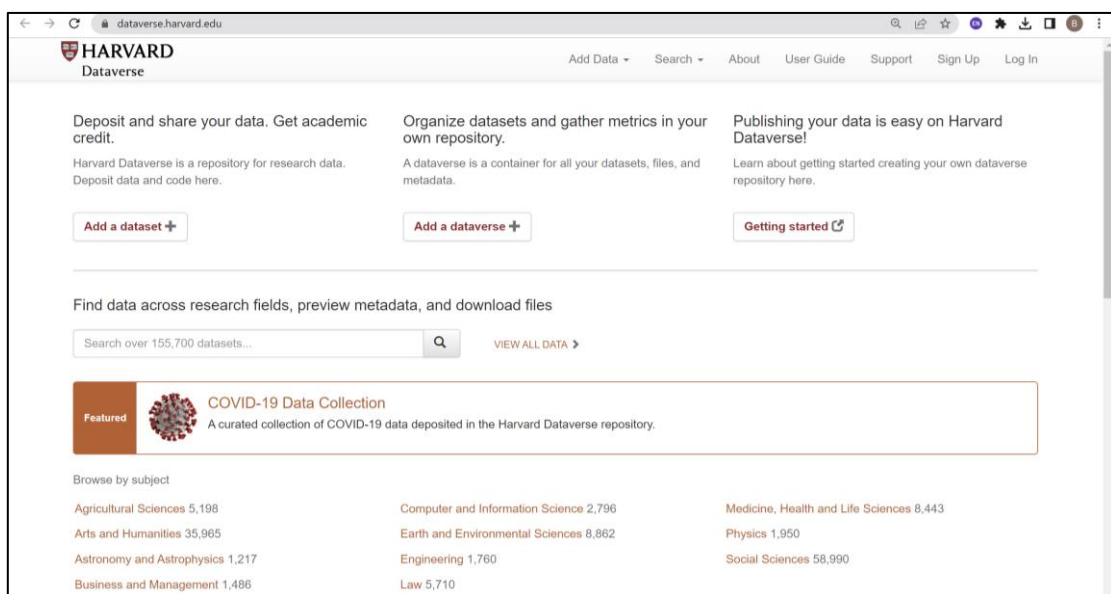
วิชาการ ข้อมูลทางวิทยาศาสตร์ ข้อมูลของรัฐบาล และข้อมูลจากองค์กรเอกชน (Brickley, Burgess, and Noy, 2019) เช่น ชุดข้อมูลไวรัสโควิด-19 ชุดข้อมูลเรื่องคุณภาพน้ำ ชุดข้อมูลเรื่อง SDG เป็นต้น



ภาพที่ 4 Dataset service of Yale library

ที่มา: <https://search.library.yale.edu/catalog>

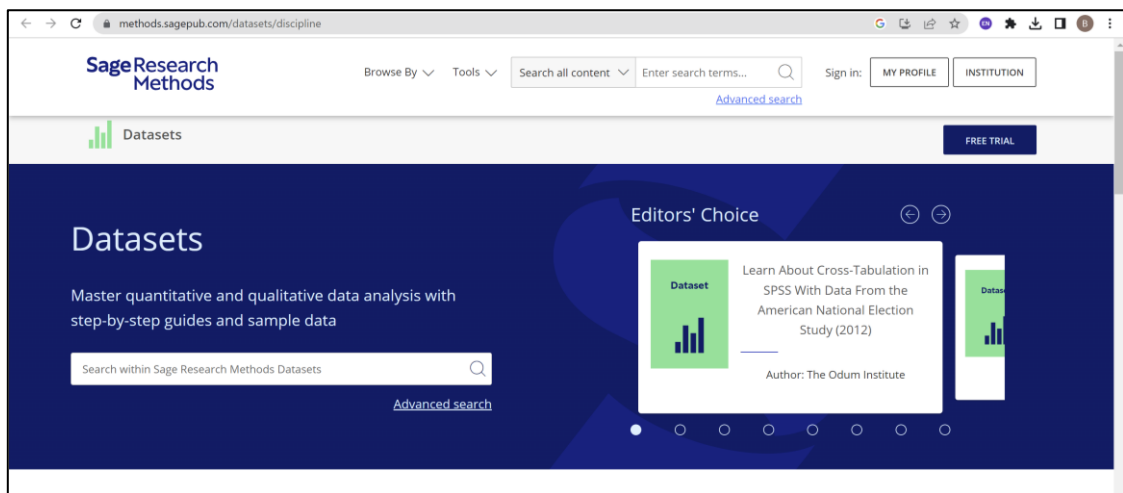
ห้องสมุดของมหาวิทยาลัยเยล (Yale university library) ให้บริการชุดข้อมูลในลักษณะของการเป็นคอลเล็กชันย่อยของห้องสมุดจัดอยู่ในคอลเล็กชัน Books+ ซึ่งสามารถเข้าถึงได้ในรูปแบบชุดข้อมูลออนไลน์และรูปแบบเอกสาร โดยรูปแบบชุดข้อมูลออนไลน์นั้นจะสามารถใช้งานได้เฉพาะนักศึกษาและบุคลากรของมหาวิทยาลัยที่มี Yale NetID เท่านั้น ในขณะที่รูปแบบเอกสารจะให้ชุดข้อมูลในรูปแบบไฟล์ข้อความ (.txt) ด้วยบริบทของห้องสมุดทำให้วิธีการสืบค้นที่แตกต่างจากแหล่งอื่นโดยสามารถสืบค้นได้จากหมวดหมู่ LC ได้



ภาพที่ 5 Harvard Dataverse

ที่มา: <https://dataverse.harvard.edu/>

Harvard Dataverse เป็นคลังข้อมูลการวิจัยที่พัฒนาโดยมหาวิทยาลัยฮาร์วาร์ดร่วมกับนักวิจัยทั่วโลก โดยเป็นคลังข้อมูลเปิดที่อนุญาตให้นักวิจัยจากทุกสาขาวิชาทั้งภายในและภายนอกของ Harvard สามารถนำฝากข้อมูลเพื่อการจัดเก็บถาวร เพื่อเพิ่มการเข้าถึงและการอ้างอิงทางวิชาการและการแบ่งปันข้อมูลร่วมกัน ซึ่งผู้นำฝากข้อมูลสามารถเป็นได้ทั้งตัวบุคคล องค์กรรวมถึงสำนักพิมพ์ เนื่องด้วยสถานการณ์แพร่ระบาดของไวรัสโคโรนา-19 ที่ผ่านมาทาง Harvard Dataverse ได้รวบรวมข้อมูลและพัฒนาคอลเล็กชันพิเศษที่เรียกว่า Covid-19 Data Collection ขึ้นซึ่งเป็นหัวข้อเดียวที่ Harvard Dataverse รวบรวมเป็นพัฒนาคอลเล็ก เพื่อให้ง่ายต่อการเข้าถึงข้อมูลของนักวิจัย และปัจจุบันมีชุดข้อมูลให้บริการจำนวน 155,700 ชุดข้อมูล



ภาพที่ 6 Sage research method

ที่มา: <https://methods.sagepub.com/>

สำนักพิมพ์ Sage ได้พัฒนาเครื่องมือช่วยสืบค้นชุดข้อมูลที่เรียกว่า Sage research method ซึ่งจัดให้ชุดข้อมูลเป็นส่วนหนึ่งของคอลเล็กชันที่สามารถสืบค้นได้ โดยวิธีการสืบค้นจะใช้หลักการเดียวกันกับการสืบค้นของฐานข้อมูล Sage และสามารถกรองชุดข้อมูลได้จาก สาขาวิชา วิธีการวิจัย ข้อมูลเชิงปริมาณ ข้อมูลเชิงคุณภาพและคอลเล็กชัน

### บทสรุป

การบริการชุดข้อมูลการวิจัยในสถาบันบริการสารสนเทศมีให้บริการทั้งในรูปแบบของคอลเล็กชันหนึ่งในห้องสมุดและการแยกออกมาเป็นเว็บไซต์เฉพาะที่ให้บริการชุดข้อมูลเพียงอย่างเดียว ชุดข้อมูลที่ให้บริการอยู่ในปัจจุบันมีทั้งชุดข้อมูลที่เกิดจากการวิจัย เช่น ชุดข้อมูลทางการแพทย์เรื่อง การแพร่ระบาดของไวรัสโคโรนา-19 ในรูปแบบของภูมิสารสนเทศ ชุดข้อมูลการนำเข้า-ส่งออกสินค้า ชุดข้อมูลความพึงพอใจ เป็นต้น และยังมีชุดข้อมูลในลักษณะของการจัดกลุ่มแบ่งหมวดหมู่ของเนื้อหา เพื่อใช้ในการศึกษาวิจัย ประกอบการตัดสินใจ และการสำรวจแนวโน้ม ซึ่งประเภทของชุดข้อมูลที่พบเห็นได้บ่อยมีทั้งประเภทตัวเลขข้อความ สื่อมัลติมีเดีย และข้อมูลทางภูมิศาสตร์ ซึ่งการเลือกใช้ชุดข้อมูลแต่ละประเภทจะขึ้นอยู่กับวัตถุประสงค์ของการนำไปใช้ประโยชน์

ในปัจจุบันชุดข้อมูลได้รับความสนใจและถูกเล็งเห็นถึงความสำคัญมากยิ่งขึ้น ส่งผลให้ Google และสำนักพิมพ์ที่ให้บริการฐานข้อมูลออนไลน์ได้พัฒนาเครื่องมือช่วยสืบค้นชุดข้อมูล เพื่อช่วยสนับสนุนให้นักวิจัยหรือผู้ใช้บริการของตนสามารถเข้าถึงชุดข้อมูลที่ต้องการได้ง่ายและสะดวก สำหรับประเทศไทยการบริการชุดข้อมูลปรากฏให้เห็นในหน่วยงานภาครัฐ เช่น กระทรวง กรม และหน่วยงานวิจัยของรัฐเป็นไปตามระเบียบเรื่องการเปิดเผยข้อมูลรัฐตามที่กฎหมายกำหนด ซึ่งหน่วยงานรัฐของต่างประเทศเองก็ต้องเปิดเผยข้อมูลการดำเนินงานตามกฎหมายเช่นเดียวกัน และในบริบทของหน่วยงานที่เป็นสถาบันบริการสารสนเทศและหน่วยงานบริการวิชาการของต่างประเทศ เช่น ห้องสมุดและมหาวิทยาลัยให้ความสำคัญกับการบริการชุดข้อมูลเพื่อการสนับสนุนการศึกษาและวิจัยของนักศึกษาประชาชนและนักวิจัย เพื่อการเรียนรู้ ต่อยอดและสร้างสรรค์นวัตกรรม (Wittmann, et al., 2019) เช่น ห้องสมุดมหาวิทยาลัยเยล (Yale university library) ให้บริการชุดข้อมูลในลักษณะของการเป็นคอลเล็กชันย่อยของห้องสมุดใหญ่บริการทั้งแบบชุดข้อมูลออนไลน์และไฟล์ข้อความ มหาวิทยาลัยฮาร์วาร์ดพัฒนาคอลเล็กชันชุดข้อมูลการวิจัยเพื่อให้นักวิจัย องค์กรและสำนักพิมพ์สามารถนำเข้าข้อมูล จัดเก็บ เผยแพร่และแบ่งปันชุดข้อมูล ทว่าในบริบทของห้องสมุดในไทยพบว่ามีการช่วยสืบค้นตอบคำถาม บริการสนับสนุนการวิจัยและให้บริการเครื่องมือช่วยสนับสนุนการวิจัย เช่น โปรแกรมตรวจการคัดลอก โปรแกรมการจัดการรายการบรรณานุกรม แต่ยังไม่พบว่ามีบริการชุดข้อมูลการวิจัยรวมด้วย

## รายการอ้างอิง

- Brickley.D, Burgess.M, and Noy.N. (2019). Google Dataset Search: Building a search engine for datasets in an open Web ecosystem. In *The World Wide Web Conference (WWW '19)*. Association for Computing Machinery, New York, NY, USA, 1365–1375. <https://doi.org/10.1145/3308558.3313685>.
- Cox, A., & Verbann, E. (2018). *Exploring research data management*. Facet.
- Geburu, T., Morgenstern, J.H., Vecchione, B., Vaughan, J.W., Wallach, H.M., Daumé, H., & Crawford, K. (2018). Datasheets for datasets. *Communications of the ACM*, 64, 86 - 92.
- Google technology company. (2018). *Google dataset search*. Retrieved from <https://datasetsearch.research.google.com/>.
- Gregory, K. (2020). A dataset describing data discovery and reuse practices in research. *Sci Data*. <https://doi.org/10.1038/s41597-020-0569-5>.
- Harvard College. (2023). *Harvard Dataverse*. Retrieved from <https://dataverse.harvard.edu/>.
- National university library. (2023). *Research process: dataset*. Retrieved from <https://resources.nu.edu/researchprocess/datasets>.
- Renear, A.H., Sacchi, S. and Wickett, K.M. (2010). Definitions of dataset in the scientific and technical literature. *Proc. Am. Soc. Info. Sci. Tech.*, 47, 1-4. <https://doi.org/10.1002/meet.14504701240>.
- Sage Publications. (2023). *Sage research method*. Retrieved from <https://methods.sagepub.com/>.
- Siman, E. (2023). *What Is a Dataset? Definitive Guide*. Retrieved from <https://brightdata.com/blog/web-data/what-is-a-dataset>.
- Tsueng, G., Mullen, J.L., Alkuzweny, M. et al. (2023). Outbreak.info Research Library: a standardized, searchable platform to discover and explore COVID-19 resources. *Nature Methods*. <https://doi.org/10.1038/s41592-023-01770-w>.
- University of central florida. (2020). *Research data and dataset*. Retrieved from <https://guides.ucf.edu/metadata/dataset#s-lg-box-1549549>.

Wipawin N. (2019). Dataset search in research data services. *J Chulabhorn Royal Acad* 1(2), 1-15. Retrieved from <https://he02.tci-thaijo.org/index.php/jcra/article/view/233180>.

Wittmann, R., Neatrou, A., Cummings, R., & Myntti, J. (2019). From Digital Library to Open Datasets. *Information Technology and Libraries*, 38(4), 49-61. <https://doi.org/10.6017/ital.v38i4.11101>.

Yale university library. (2022). *Dataset service of Yale library*. Retrieved from <https://search.library.yale.edu/catalog>.

จิรวรรณ ศรีวงษ์, น้าทิพย์ วิภาวิน, อรอนงค์ นัยวิกุล และ อารีย์ วัฒนกิจจานุกิจ. (2565). การบริการข้อมูลวิจัยด้านข่าวในไทย. *วารสารศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์*, 22(2), 179–194. <https://doi.org/10.14456/lartstu.2022.19>.

น้าทิพย์ วิภาวิน. (2559). บริการสารสนเทศงานวิจัย. *วารสารห้องสมุด*, 60(2), 55-65.

น้าทิพย์ วิภาวิน. (2563). ชุดข้อมูลวิจัยแบบเปิดไวรัสโควิด-19: ข้อมูลเปิดสำหรับชุมชนวิจัยทั่วโลก. *วารสารห้องสมุด*, 64(1), 36–46.

สำนักงานการวิจัยแห่งชาติ. (2564). *เว็บไซต์ข้อมูลเปิดของสำนักงานการวิจัยแห่งชาติ (วช.)*. เข้าถึงได้จาก <https://opendata.nrct.go.th/>.

สำนักงานพัฒนารัฐบาลดิจิทัล. (2564). *มาตรฐานและหลักเกณฑ์การเปิดเผยข้อมูลเปิดภาครัฐ*. เข้าถึงได้จาก <https://www.dga.or.th/policy-standard/standard/dga-005/54063-2/>.

\_\_\_\_\_. (2558). *เว็บไซต์ชุดข้อมูลเปิดของภาครัฐของไทย*. เข้าถึงได้จาก <https://data.go.th/>.