

เรื่อง Google Scholar

จัดทำโดย

นาย ชนกฤษฏ์ ฤทธิ์เดช

เลขนักศึกษา 6209650404

เสนอ

อาจารย์ วสิศ ลิ้มประเสริฐ

รายงานนี้เป็นส่วนหนึ่งของวิชา DSI200

ภาคเรียนที่ 2 ปีการศึกษา 2563

มหาวิทยาลัยธรรมศาสตร์(ศูนย์รังสิต)

Project: Google Scholar

จากการศึกษาเกี่ยวกับการทำ web scraping เพื่อเก็บรวบรวมข้อมูล โดยใช้ HTTP request เก็บข้อมูลต่างๆ ชื่ออาจารย์ และบทวิจัยต่างๆ ของมหาวิทยาลัยธรรมศาสตร์ จากเว็บ Google Scholar ได้ข้อมูลมาดังนี้

1. ตารางเก็บข้อมูล User_ID ชื่อของผู้เขียน และสังกัด/คณะ

	user_id	author	affiliation
0	Re819VUAAAAJ	sandhya Babel	Professor, Sirindhorn International Institute ...
1	97Zz_TIAAAAJ	Sombat Muengtaweepongsa	Thammasat University
2	u-vl_alAAAAJ	Wutiphol Sintunavarat	Department of Mathematics and Statistics, Facu...
3	9bUei6wAAAAJ	Bunyarit Uyyanonvara	Associated Professor, SIIT, Thammasat University
4	UOeuXvQAAAAJ	Chanathip Namprempre	Assistant Professor, Thammasat University
...
293	NLshr9gAAAAJ	Ines Ayu Handayani	Master's Student, SIIT, Thammasat University
294	oALsv34AAAAJ	Kritsasith Warin	Lecturer, Faculty of Dentistry, Thammasat Univ...
295	kJaNuNMAAAAJ	Dulyaphab Chaturongkul	Lecturer in Political Science, Thammasat Unive...
296	7XmRxVUAAAAJ	Punnavich Khowrurk	Department of Computer Science, Thammasat Unive...
297	vvP6YwkAAAAJ	Adipon Euajarusphan	Lecturer , Thammasat University

298 rows x 3 columns

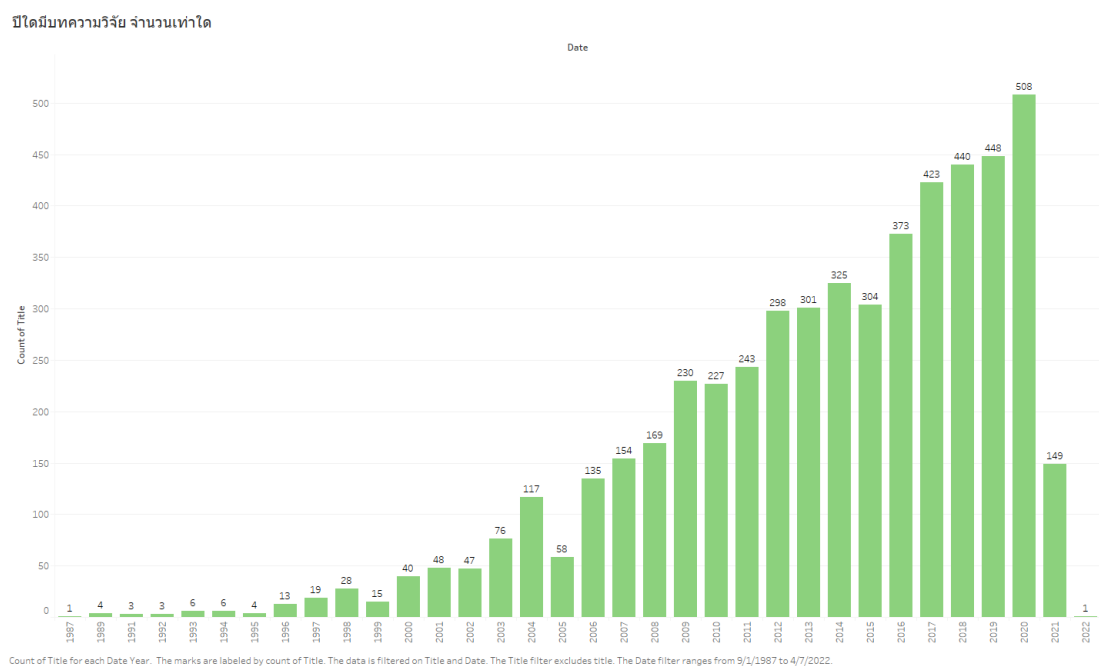
จากข้อมูลที่ได้มาทั้งหมดนั้น สามารถสรุปได้ว่า ในวันที่ทำการเก็บรวบรวมข้อมูลล่าสุดนั้น และข้อมูลสรุปรวมแถวที่ท้ายตาราง ทำให้ทราบว่าผู้เขียนบทวิจัย ที่อยู่ในมหาวิทยาลัยธรรมศาสตร์อยู่ที่ประมาณ 298 คน

2. ตารางเก็บข้อความ ผู้เขียน วันที่ และคำอธิบาย

	title	authors	publication_date	description	cite_by
0	Real-time people tracking in a camera network	Wasit Limprasert, Andrew Wallace, Greg Michaelson	4/24/2013	We present an approach to track several subjec...	19
1	Low-cost adsorbents for heavy metals uptake fr...	Sandhya Babel, Tonni Agustiono Kurniawan	2/28/2003	In this article, the technical feasibility of ...	4012
2	Physico-chemical treatment techniques for wast...	Tonni Agustiono Kurniawan, Gilbert YS Chan, Wa...	5/1/2006	This article reviews the technical applicabili...	1978
3	Cr (VI) removal from synthetic wastewater usin...	Sandhya Babel, Tonni Agustiono Kurniawan	2/1/2004	In this study, the technical feasibility of co...	949
4	Comparisons of low-cost adsorbents for treatin...	Tonni Agustiono Kurniawan, Gilbert YS Chan, Wa...	8/1/2006	In this article, the removal performance and c...	807
...
8224	พฤติกรรม การ ใช้ สื่อ ของ กลุ่ม เจิน เนอ เร ช...	เอ็ด จรัส พันธุ์, อติ พล	1/1/2018	This study aim to investigate how Generation X...	-
8225	กลยุทธ์ และ เครือ ข่าย การ สื่อสาร เพื่อ สร้าง...	อติ พล เอ็ด จรัส พันธุ์, ณัฐ นิ ชา กร กิ่ง มาลา	2018	การ วิจัย เรื่อง "กลยุทธ์ และ เครือ ข่าย การ ส...	1
8226	-	Adipon Euajarusphan	2017	-	-
8227	การ สื่อสาร เพื่อ การ พัฒนา เครือ ข่าย ตลาด สี...	เอ็ด จรัส พันธุ์, อติ พล	2015	บทความ ขึ้น นี้ จึง รัง นำ เสนอ ความ เข้าใจ ...	-
8228	การ สื่อสาร เพื่อ การ พัฒนา เครือ ข่าย ตลาด สี...	อติ พล เอ็ด จรัส พันธุ์	2015	บทความ ขึ้น นี้ จึง รัง นำ เสนอ ความ เข้าใจ ...	-
8229 rows x 5 columns					

จากข้อมูลที่ได้มาทั้งหมดนั้น สามารถสรุปได้ว่า ในวันที่ทำการเก็บรวบรวมข้อมูลล่าสุดนั้น และข้อมูลสรุปรวมแถวที่ท้ายตาราง ทำให้ทราบว่ามียุทธวิธีที่ถูกเขียนบน Google Scholar จากอาจารย์ในมหาวิทยาลัยธรรมศาสตร์เป็น เป็นจำนวน 8229 เรื่อง

และจากข้อมูลที่ได้มานั้น มีข้อมูลที่เป็นประเภทวันที่ จึงนำออกมาวิเคราะห์ และทำเป็นกราฟได้ จากการวิเคราะห์ทำให้เห็นว่า มียุทธวิธีกว่า 3000 เรื่อง ที่ไม่ถูกใส่วันที่เขียน หรือไม่สามารถดึงข้อมูลวันที่มาได้ และนำข้อมูลที่เหลือนั้น มาทำเป็นกราฟ เพื่อเปรียบเทียบว่า ในแต่ละปีนั้น มีผู้เขียนยุทธวิธีเป็นจำนวนเท่าใด ดังนี้



และจากข้อมูลที่เก็บมานั้น มีข้อมูลที่บอกถึงจำนวนผู้เข้าชมบทวิจัย (Cite by) จึงนำมาเปรียบเทียบ และเรียงลำดับ จากจำนวนผู้เข้าชม ตั้งแต่จำนวนผู้เข้าชมมากที่สุด จนไปถึงจำนวนผู้เข้าชมน้อยที่สุด และนำมาทำเป็นกราฟแท่งได้ ดังนี้

