



หน่วยที่ 1 ความรู้พื้นฐาน  
Big Data และประเภทข้อมูล  
รายวิชาการจัดการข้อมูลขนาดใหญ่ (Big Data)

กลุ่มผู้เรียน ปวส.1 สาขาวิชาเทคโนโลยีธุรกิจดิจิทัล  
ครูผู้สอน นางสุนันทา พุฒโต

# Big Data คืออะไร?

ข้อมูลที่มีปริมาณมหาศาล (Volume) มีความซับซ้อน และเพิ่มขึ้นอย่างรวดเร็วเกินกว่าที่ระบบฐานข้อมูลแบบเดิมจะจัดเก็บหรือประมวลผลได้

โลกเราสร้างข้อมูลใหม่สูงถึง 2.5 Quintillion byte ในแต่ละวัน!  
(เทียบเท่ากับข้อมูลตั้งแต่ปี ค.ศ. 1 ถึง 2000 รวมกัน)

หัวใจสำคัญไม่ได้อยู่ที่ปริมาณข้อมูล แต่อยู่ที่องค์กรจัดการและวิเคราะห์ (Analyze) ข้อมูลเหล่านั้นเพื่อสร้างความได้เปรียบทางธุรกิจอย่างไร

# วิวัฒนาการของ Big Data สู่ยุคดิจิทัล

## ยุค 1960s-70s (จุดเริ่มต้น)

- เริ่มมีการใช้คอมพิวเตอร์ประมวลผลข้อมูล
- พัฒนาศูนย์ข้อมูลแห่งแรกและฐานข้อมูลเชิงสัมพันธ์

## ปี 2005+ (ยุคข้อมูลระเบิด)

- การเติบโตของ Web 2.0 และ Social Media
- กำเนิด Hadoop เพื่อจัดการข้อมูลขนาดใหญ่

## ปัจจุบัน (ยุคแห่งการเชื่อมต่อ)

- การพัฒนาของ IoT และ Cloud Computing
- อุปกรณ์ทุกชนิดเชื่อมต่ออินเทอร์เน็ต ทำให้เกิดข้อมูลมหาศาลแบบ Real-time

# หัวใจสำคัญของ Big Data: ลักษณะ 6Vs

ข้อมูลที่จะถูกจัดว่าเป็น Big Data ที่มีคุณภาพสูง  
ควรมีลักษณะพื้นฐาน 6 ประการ ดังนี้:

**Volume**  
(ปริมาณ)



**Velocity**  
(ความเร็ว)



**Variability**  
(ความแปรผัน)



**Variety**  
(ความหลากหลาย)



**Value**  
(คุณค่า)



**Veracity**  
(ความถูกต้อง)



# เจาะลึก 3V แรก: โครงสร้างข้อมูลขนาดใหญ่

## Volume (ปริมาณ)

ข้อมูลมีจำนวนมากมหาศาลมากพอ (เช่น บันทึกรูทธรณมนับล้านรายการ) เพื่อให้การวิเคราะห์ได้ Insights ที่ตรงกับความเป็นจริง

## Velocity (ความเร็ว)

ข้อมูลถูกสร้างและรับส่งแบบ Real-time ทันต่อเหตุการณ์ (เช่น การรูดบัตรเครดิต, สัญญาณ GPS) ทำให้ตอบสนองได้ทันที

## Variety (ความหลากหลาย)

มีรูปแบบที่แตกต่างกัน ทั้งแบบมีโครงสร้าง (ตาราง) ทั้งโครงสร้าง และไม่มีโครงสร้าง (รูปภาพ, เสียง, ข้อความ)

# เจาะลึก 3V หลัง: คุณภาพและการนำไปใช้

## Veracity (ความถูกต้อง)



ความน่าเชื่อถือของแหล่งที่มา และคุณภาพของข้อมูล ต้องมีกระบวนการตรวจสอบยืนยันความถูกต้อง (Data Integrity)

## Value (คุณค่า)



ข้อมูลต้องนำไปเปลี่ยนเป็น 'Insights' ที่มีประโยชน์ สัมพันธ์กับธุรกิจ สร้างรายได้ หรือลดต้นทุนได้จริง

## Variability (ความแปรผันได้)



ข้อมูลสามารถเปลี่ยนแปลงรูปแบบหรือความหมายไปตามบริบทและการใช้งาน สามารถวิเคราะห์ได้จากหลายแง่มุม



# องค์ประกอบสำคัญในการทำ Big Data 5 ประการ



# วงจรชีวิตของข้อมูล (Data Lifecycle)



# ประเภทของข้อมูลขนาดใหญ่ (Types of Big Data)

**Structured Data**  
(ข้อมูลที่มีโครงสร้าง)  
ข้อมูลที่จัดเรียงเป็นระเบียบชัดเจน  
ค้นหาง่าย

**Semi-structured Data**  
(ข้อมูลกึ่งโครงสร้าง)  
ข้อมูลที่มีโครงสร้างบางส่วน  
มี Tag หรือ Hierarchy  
ช่วยจัดระเบียบ

**Unstructured Data**  
(ข้อมูลที่ไม่มีโครงสร้าง)  
ข้อมูลที่ไม่มีรูปแบบตายตัว  
ซับซ้อน และต้องใช้ AI  
ช่วยประมวลผล

# เจาะลึกประเภทข้อมูล (The Ultimate Comparison)

	ข้อมูลที่มีโครงสร้าง (Structured)	ข้อมูลกึ่งโครงสร้าง (Semi-structured)	ข้อมูลที่ไม่มีโครงสร้าง (Unstructured)
ลักษณะ (Definition)	จัดเก็บในตาราง (Table) มี Field อดจน (Relational DB) 	มี Tag/Hierarchy โครงสร้างยืดหยุ่น 	ไม่มีรูปแบบกำหนดล่วงหน้า ไม่สม่ำเสมอ 
การค้นหา (Query)	ค้นหา (Query) ได้ง่ายและรวดเร็ว 	ค้นหาได้ผ่าน Tag/Keyword 	ค้นหาตรงๆ ยาก ต้องผ่านการ Pre-process หรือ AI 
ตัวอย่าง (Examples)	MySQL, Excel, ไฟล์ CSV 	JSON, XML, Server Logs 	รูปภาพ, วิดีโอ, เสียง, ข้อความ (Text/Email) 



# รูปแบบของข้อมูลที่พบบ่อย (Common Big Data Formats)

**Behavioral Data**  
(ข้อมูลเชิงพฤติกรรม)

Server logs,  
พฤติกรรมการใช้เว็บไซต์,  
การใช้งานตู้ ATM

**Images & Sounds**  
(ข้อมูลภาพและเสียง)

ภาพถ่ายทางอากาศ, 5ดีโอ,  
ภาพ Google Street View, ไฟล์เสียง

**Languages**  
(ข้อมูลภาษาและข้อความ)

Text message,  
โพสต์บน Social Media,  
บทความในเว็บไซต์

**Records**  
(ข้อมูลบันทึกธุรกรรม)

ข้อมูลประวัติการรักษาพยาบาล,  
ข้อมูลภาษี, ผลสำรวจขนาดใหญ่

**Sensors**  
(ข้อมูลเซนเซอร์)

สัญญาณ GPS, ข้อมูลอุณหภูมิ,  
ตัวจับความเร็วรถยนต์



# บุคลากรเบื้องหลังงาน Big Data (Big Data Personnel)

**System Administrator**  
(ผู้ดูแลระบบ)

รับผิดชอบโครงสร้างพื้นฐาน  
การรวบรวมข้อมูล และวางแผน  
นำเข้าข้อมูลจากภายนอกสู่ระบบ

**Data Engineer**  
(วิศวกรข้อมูล)

สร้างระบบ จัดเตรียม และ  
ทำความสะอาดข้อมูล (Data Pipeline)  
ให้อยู่ในสภาพพร้อมใช้งาน

**Data Analyst / Scientist**  
(นักวิเคราะห์ / นักวิทยาศาสตร์ข้อมูล)

นำข้อมูลไปวิเคราะห์ด้วยสถิติหรือสร้าง  
Machine Learning Model เพื่อหา  
insights และทำรายงานเพื่อธุรกิจ

# ประโยชน์ของ Big Data ต่อธุรกิจ (Business Benefits)

## Personalization (การนำเสนอแบบรู้ใจ)

เข้าใจความต้องการของลูกค้าทำรายบุคคล  
เสนอสินค้าและโปรโมชั่นได้ตรงใจที่สุด  
(เช่น การแนะนำหนังของ Netflix)

## Risk Management (การจัดการความเสี่ยง)

คาดการณ์ความเสี่ยงล่วงหน้า  
เช่น การตรวจจับการทุจริต  
(Fraud Detection) ในระบบธนาคาร

## Operational Efficiency (เพิ่มประสิทธิภาพการทำงาน)

ลดความสูญเสียในกระบวนการผลิต  
วางแผน Supply Chain  
และใช้ทรัพยากรอย่างคุ้มค่า

## New Product Innovation (สร้างนวัตกรรมใหม่)

ค้นพบโอกาสทางการตลาดใหม่ๆ  
จากการวิเคราะห์เทรนด์และ  
พฤติกรรมผู้บริโภค

# Case Study: การใช้ Big Data ขับเคลื่อนการท่องเที่ยวไทย

## Context:

การร่วมมือระหว่าง  
Dtac, จุฬาฯ, และภาครัฐ

## Action:

นำ **'Mobility Data'**  
(ข้อมูลการเคลื่อนที่ของประชากร  
จากสัญญาณเครือข่ายมือถือ)  
นับหมื่นล้านรายการ  
มาจัดเรียงและประมวลผล  
โดยไม่ระบุตัวตน



## Result:

- แปลงผลเป็น **Insights**  
พฤติกรรมนักท่องเที่ยวเชิงลึก
- ช่วยให้ภาครัฐวางแผนฟื้นฟู  
และสร้างนโยบายส่งเสริมการ  
ท่องเที่ยวได้อย่างแม่นยำ
- เอกชนจัดแคมเปญสอดคล้อง  
กับพฤติกรรมจริงของนักท่องเที่ยว

# บทสรุปหน่วยที่ 1: พื้นฐาน Big Data



## Key Takeaways:

- ✓ Big Data คือชุดข้อมูลมหาศาลที่ระบบแบบเดิมจัดการไม่ได้
- ✓ ประเมินคุณภาพข้อมูลด้วย 6Vs: Volume, Velocity, Variety, Veracity, Value, Variability
- ✓ ข้อมูลแบ่งเป็น 3 ประเภทหลัก: มีโครงสร้าง, กึ่งโครงสร้าง, และไม่มีโครงสร้าง
- ✓ เป้าหมายสูงสุดคือการแปลง 'ข้อมูลดิบ' เป็น 'Insights' เพื่อขับเคลื่อนและสร้างความได้เปรียบทางธุรกิจ

'Data is the new oil, แต่มูลค่าที่แท้จริงเกิดจากการนำไปกลั่นใช้งาน!'  
— ครุสุมินทา พุฒโต (แผนกเทคโนโลยีธุรกิจดิจิทัล)