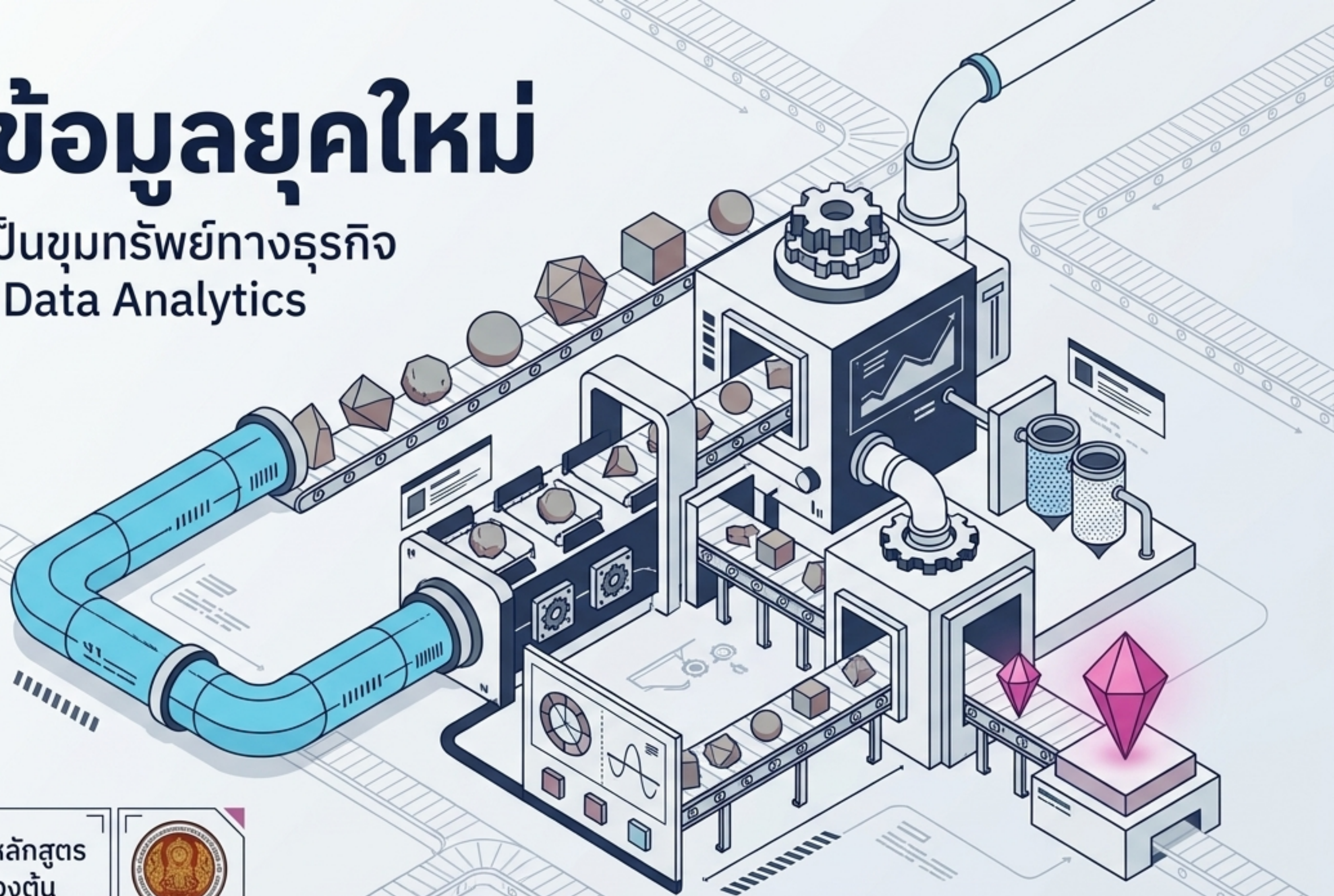


โรงงานข้อมูลยุคใหม่

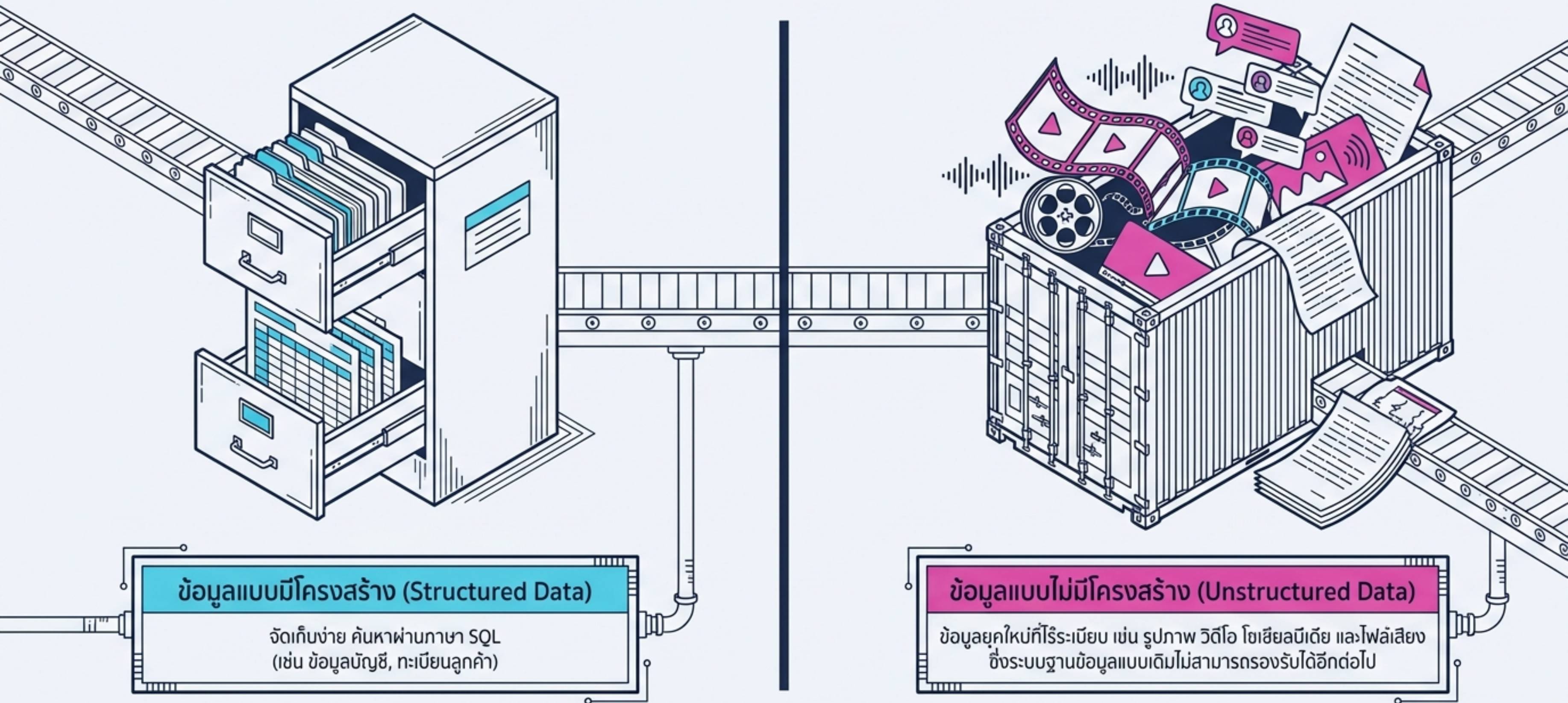
เปลี่ยนข้อมูลดิบให้เป็นยุทธศาสตร์ทางธุรกิจ
ด้วย Big Data และ Data Analytics



สรุปเนื้อหาและแนวปฏิบัติจากหลักสูตร
การจัดการข้อมูลขนาดใหญ่เบื้องต้น



สิ้นสุดยุคของตารางข้อมูล สู่ยุคไร้โครงสร้าง

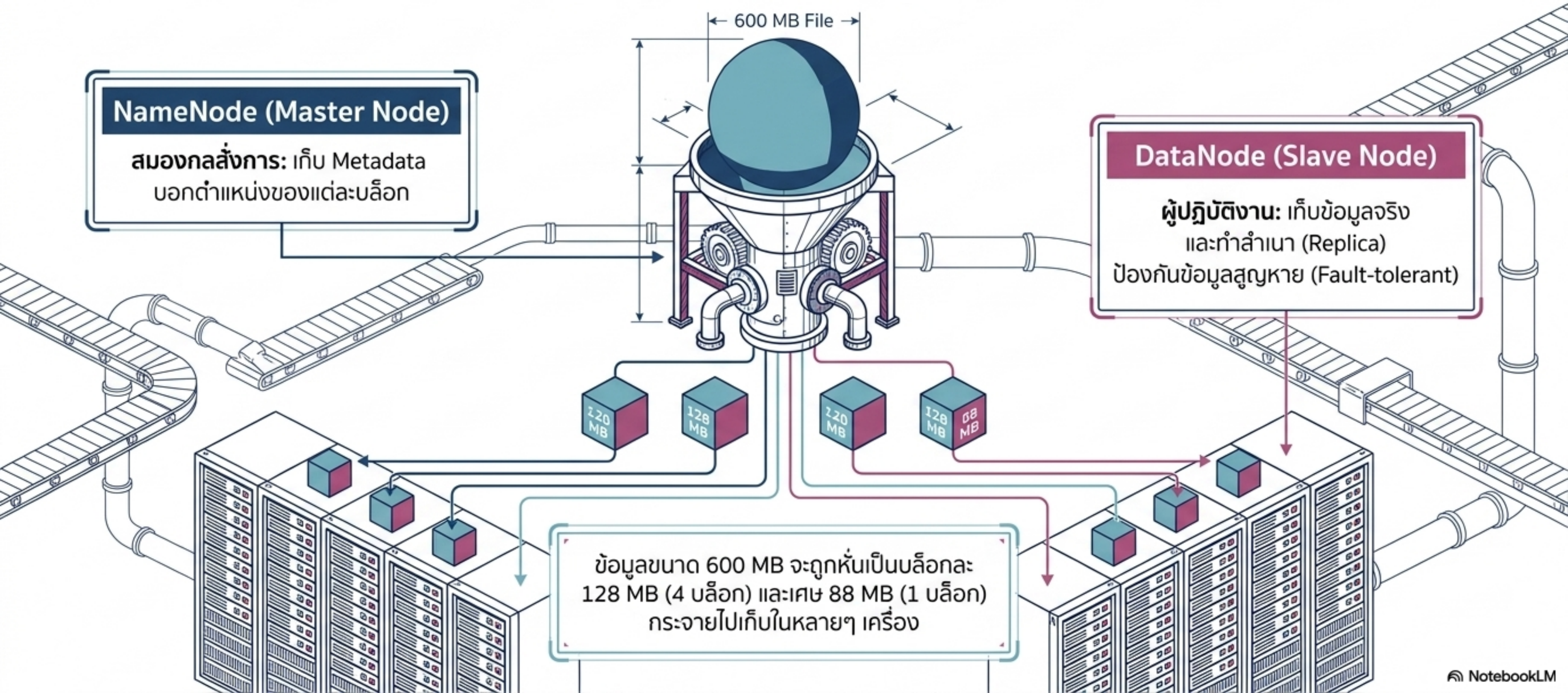


คุณลักษณะสำคัญ 4 ประการของ Big Data (The 4Vs)

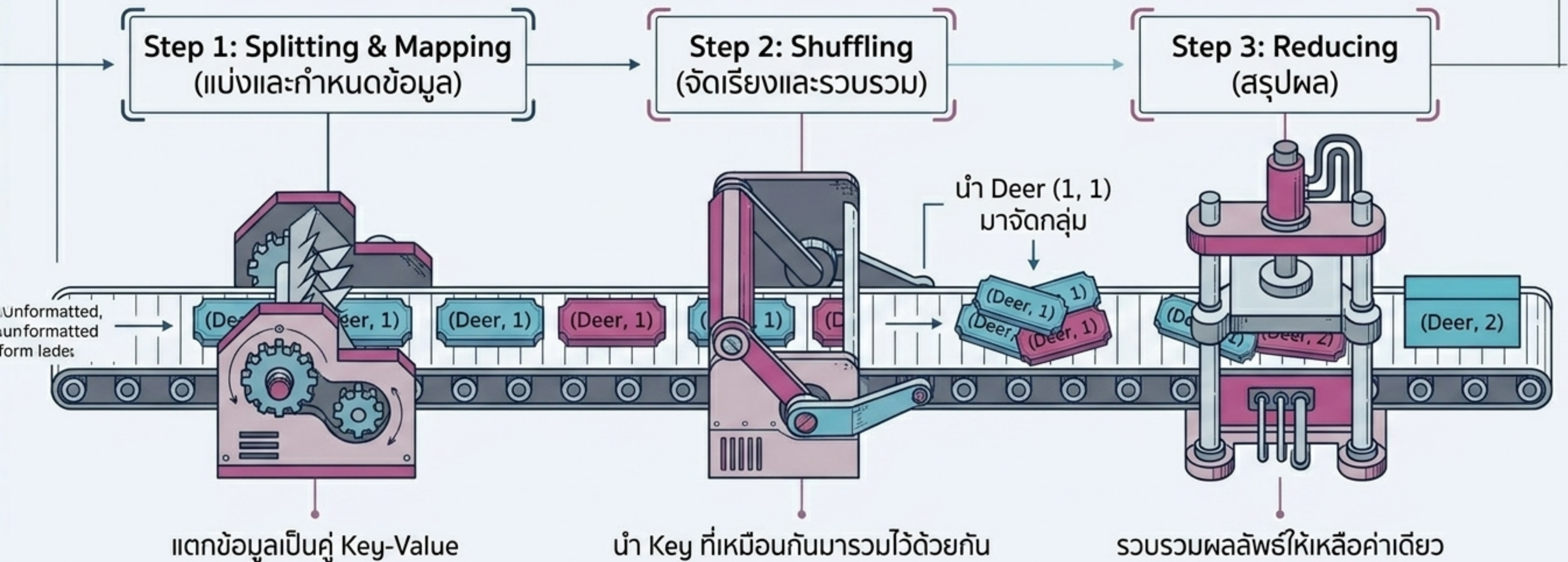


การจัดเก็บข้อมูลด้วย HDFS (Hadoop Distributed File System)

กระบวนการ: กระจายข้อมูลใหญ่ ให้กลายเป็นบล็อกย่อย (Block)



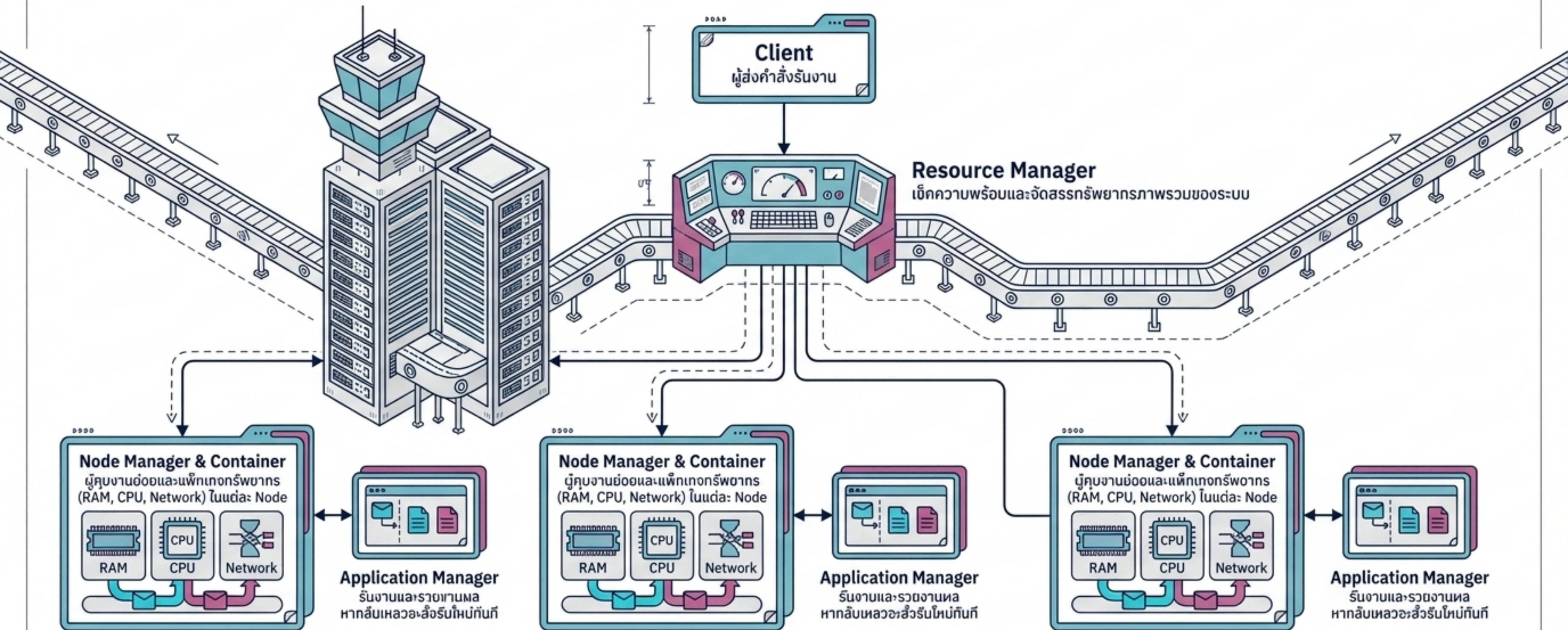
MapReduce: ประมวลผลข้อมูลแบบคู่ขนาน



ประมวลผลแบบขนาน (Parallel Processing) ทำให้วิเคราะห์ข้อมูลขนาดใหญ่ได้ไวขึ้นและทนทานต่อความเสียหาย

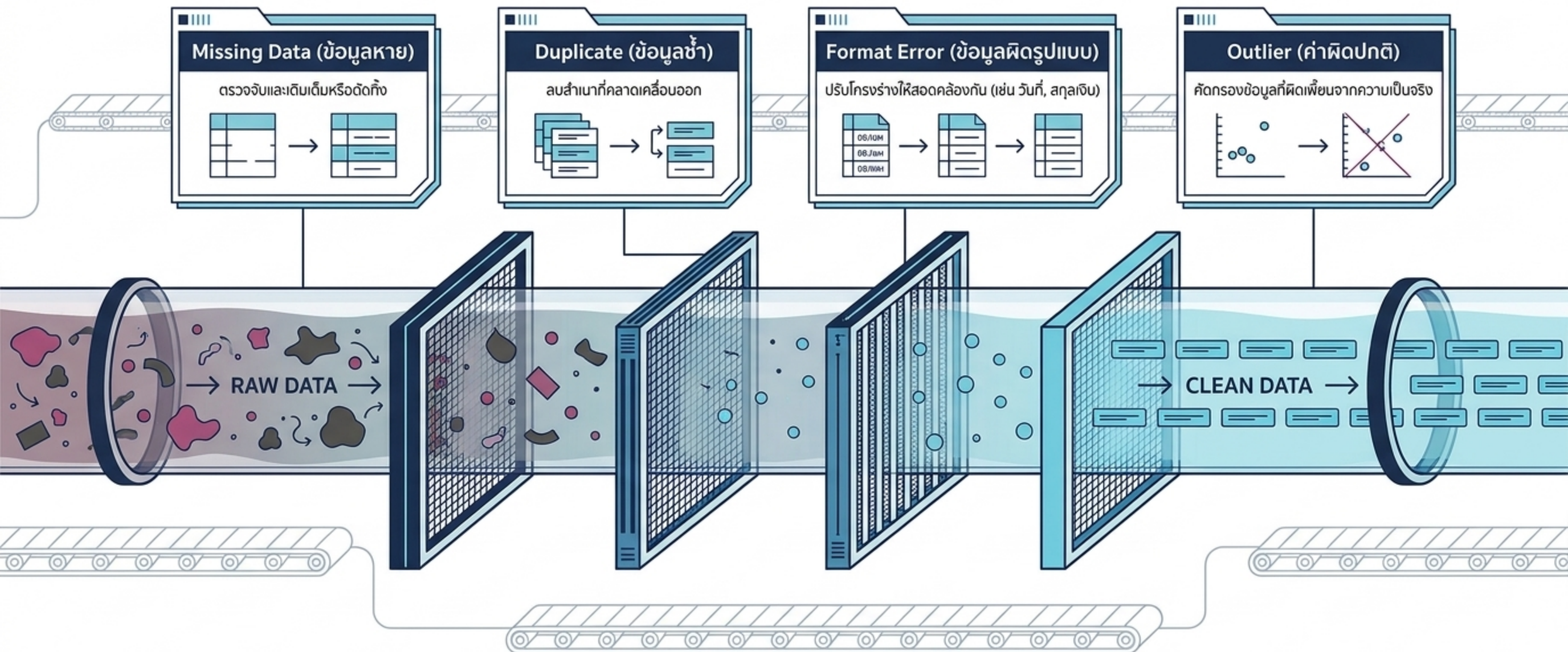
YARN: ผู้ควบคุมและจัดสรรทรัพยากร

Yet Another Resource Negotiator ทำหน้าที่จัดตารางงาน (Task) และบริหารทรัพยากร (Resource)



ระบบกรองคุณภาพข้อมูล (Data Quality Cleansing)

ข้อมูลที่ได้รับคุณภาพส่งผลกระทบต่อธุรกิจและการเงิน จึงต้องผ่านกระบวนการ Clean Data



เกราะป้องกันความเสี่ยง (Data Security & Risk Management)

ภัยคุกคาม (Threats)

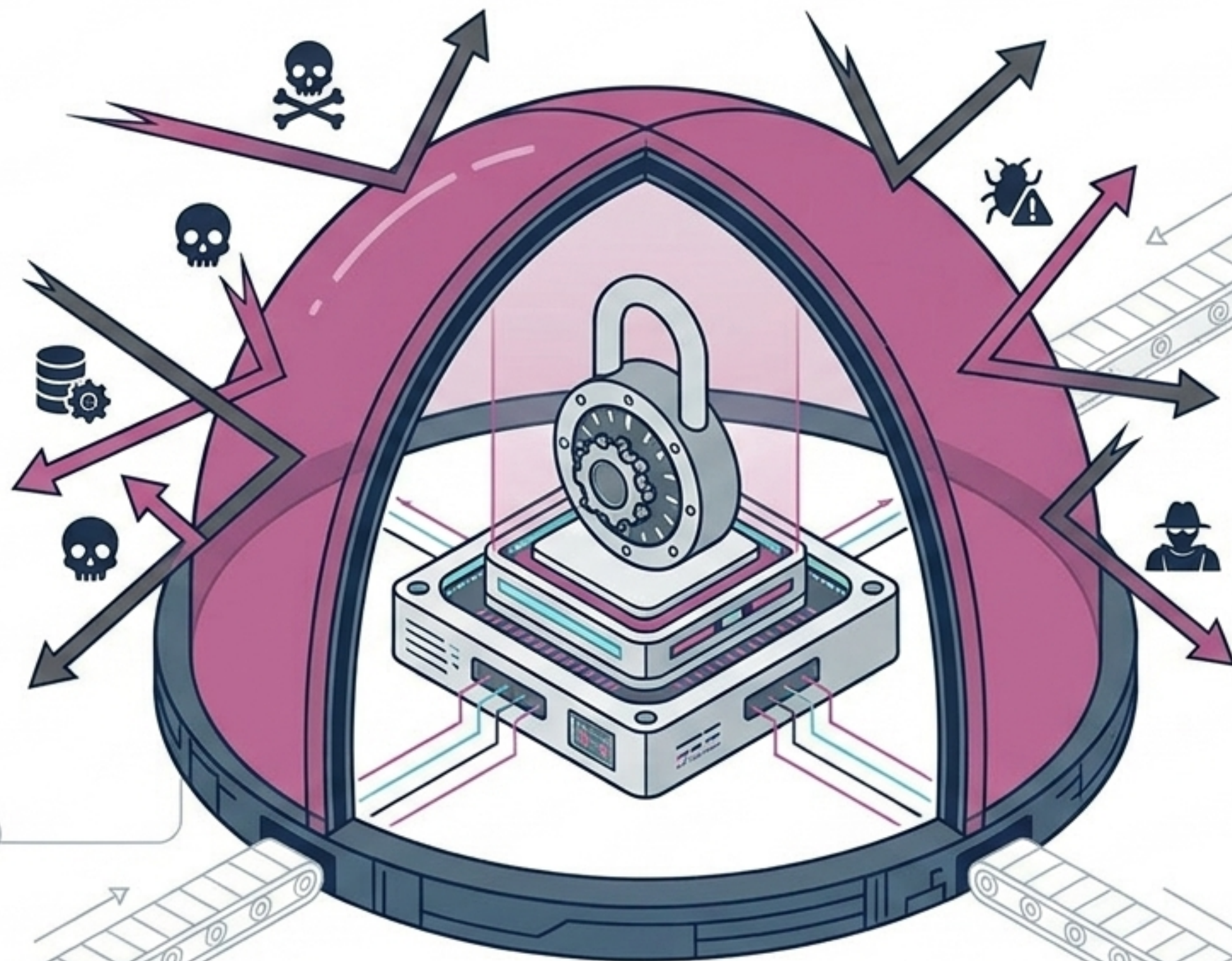
- แฮกเกอร์ (Hackers) และมัลแวร์



- การโจมตีผ่านช่องโหว่ เช่น SQL Injection



- ภัยจากบุคคลภายใน (Unauthorized Access)



หลักพื้นฐานการป้องกัน (CIA Triad)

- Confidentiality (ความลับ): ให้สิทธิ์เฉพาะผู้เกี่ยวข้อง



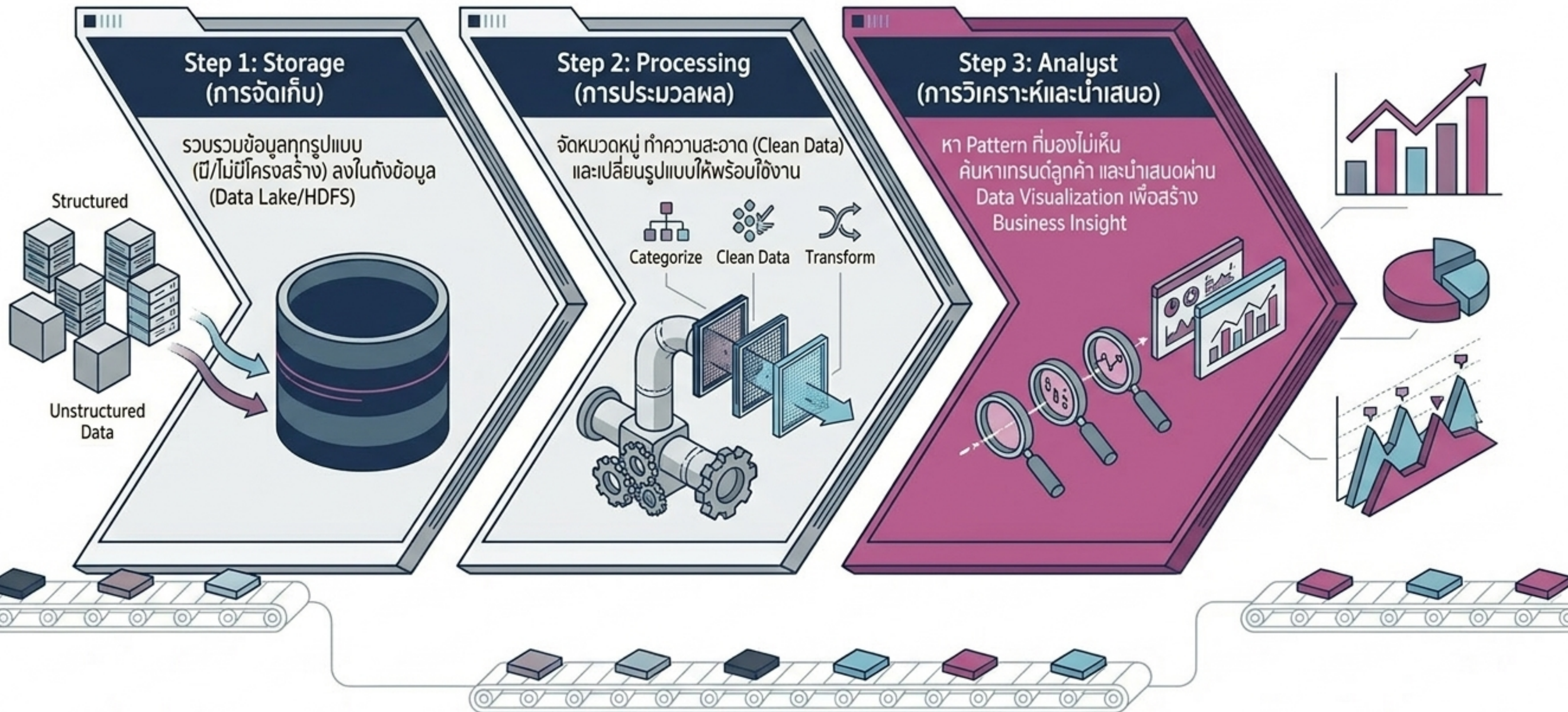
- Integrity (ความถูกต้อง): ข้อมูลต้องไม่ถูกเปลี่ยนแปลงโดยไม่ได้รับอนุญาต



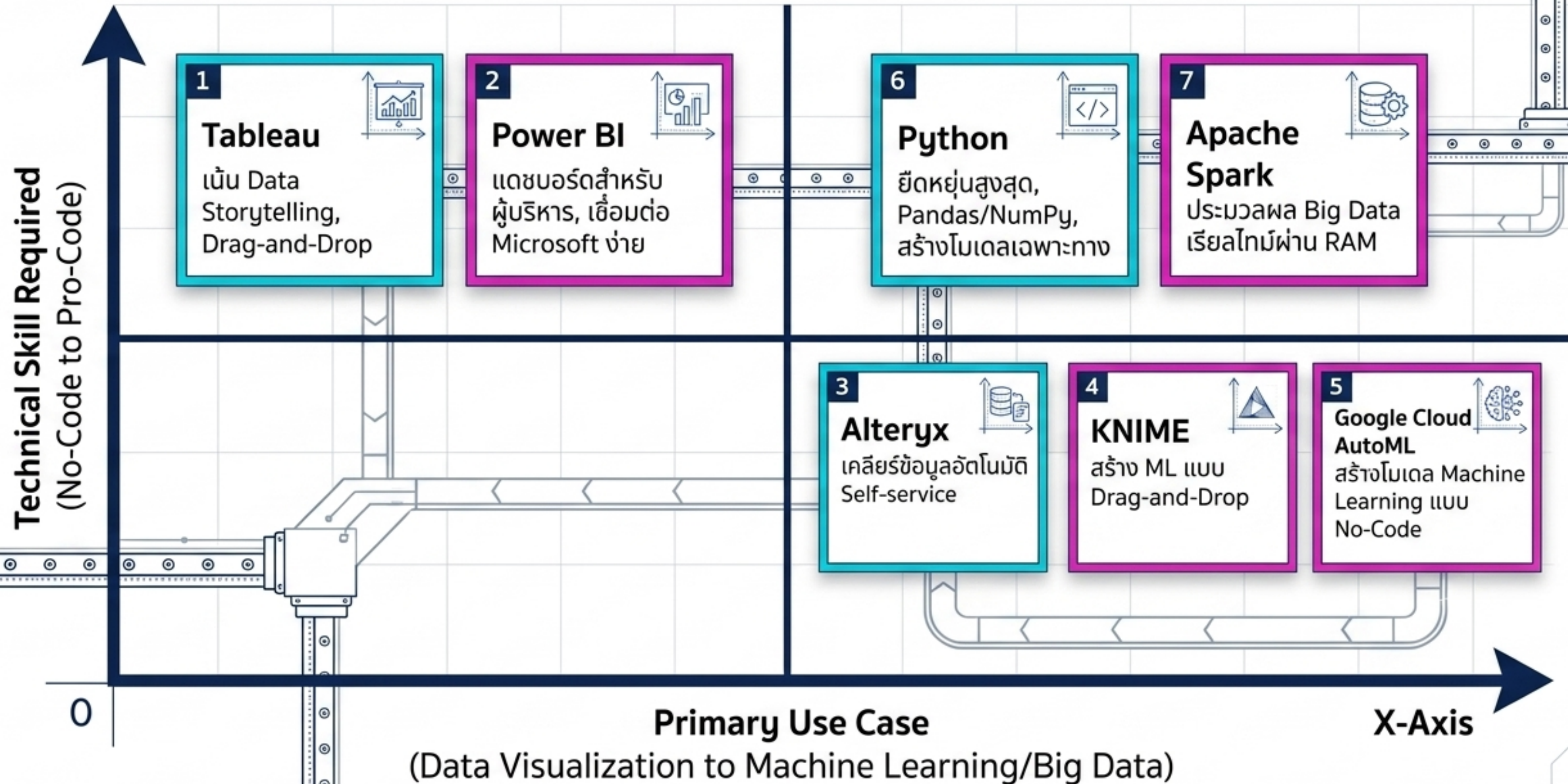
- Availability (ความพร้อมใช้งาน): เข้าถึงได้เสมอเมื่อต้องการ มีระบบสำรองข้อมูล



จากการจัดการ สู่การวิเคราะห์ (The Analytics Pipeline)



เมทริกซ์เครื่องมือวิเคราะห์ข้อมูล (The 7 Analytics Tools)



5 เช็กลิสต์ก่อนเลือกใช้เครื่องมือ (Tool Selection Funnel)

Level 1: Business Needs (ความต้องการธุรกิจ)

ต้องการดูแบบเรียลไทม์? พยากรณ์อนาคต? หรือแค่แสดงผล?

Level 2: Budget (งบประมาณ)

ซอฟต์แวร์ลิขสิทธิ์, ระบบ Cloud หรือ Open Source ฟรี?

Level 3: Scalability (การขยายตัว)

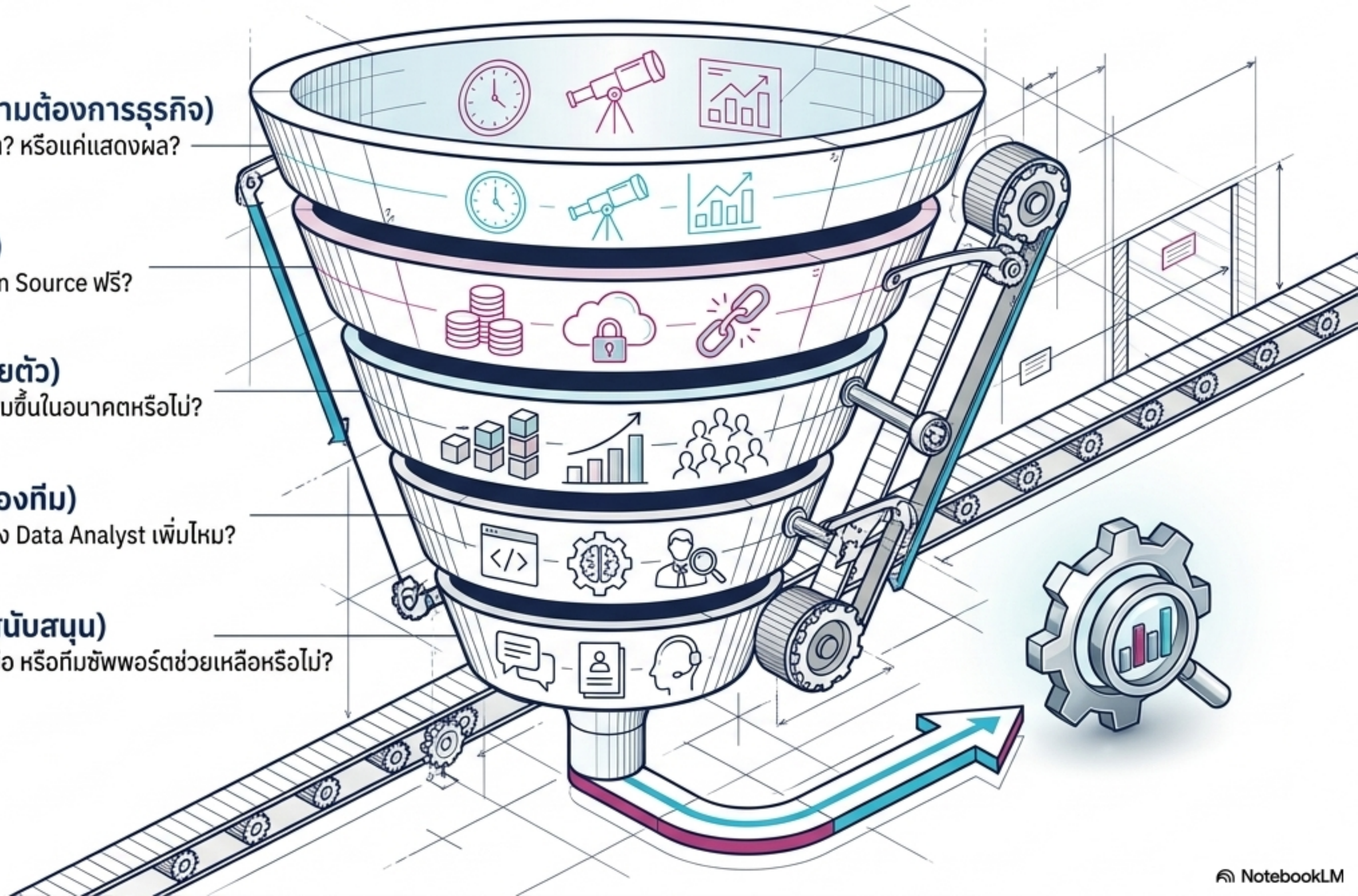
รองรับปริมาณข้อมูลและจำนวนผู้ใช้ที่จะเพิ่มขึ้นในอนาคตหรือไม่?

Level 4: Team Skills (ทักษะของทีม)

ทีมมีความรู้เขียนโปรแกรมหรือไม่? ต้องจ้าง Data Analyst เพิ่มไหม?

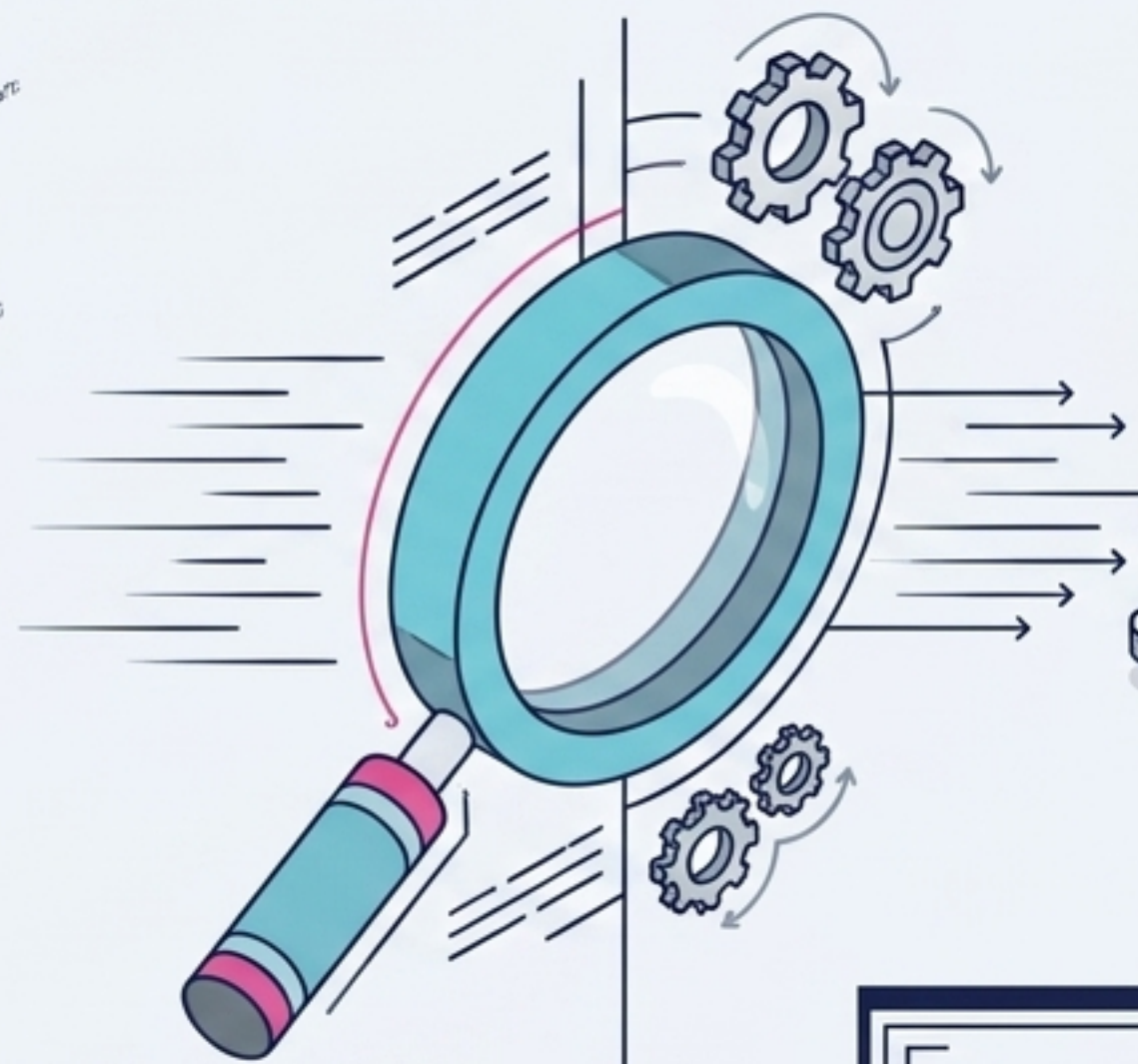
Level 5: Tech Support (การสนับสนุน)

มีชุมชนผู้ใช้งาน (Community) เอกสารคู่มือ หรือทีมซัพพอร์ตช่วยเหลือหรือไม่?



การนำเสนอภาพข้อมูล (Data Visualization)

การทำให้ข้อมูลเป็น 'นามธรรม' (Abstraction) สู่ 'แบบแผน' (Pattern)



Business Impact:

- เปลี่ยนสถิติซับซ้อนให้เข้าใจง่ายผ่านสถิติอย่างง่าย
- หาคความสัมพันธ์และเทรนด์การตลาด
- ช่วยผู้นำตัดสินใจเชิงกลยุทธ์ได้อย่างแม่นยำและรวดเร็ว

The Big Data Ecosystem: กลไกขับเคลื่อนธุรกิจยุคใหม่

Phase 1: Ingestion (แหล่งที่มา)

โซเชียลมีเดีย, เซ็นเซอร์ IoT (ข้อมูลมหาศาล ไร้โครงสร้าง)

Phase 2: Storage & Engine (จัดเก็บและประมวลผล)

ระบบ HDFS และ YARN
กระจายทรัพยากรทำงานแบบคู่ขนาน (MapReduce)

Phase 3: Governance (ควบคุมคุณภาพ)

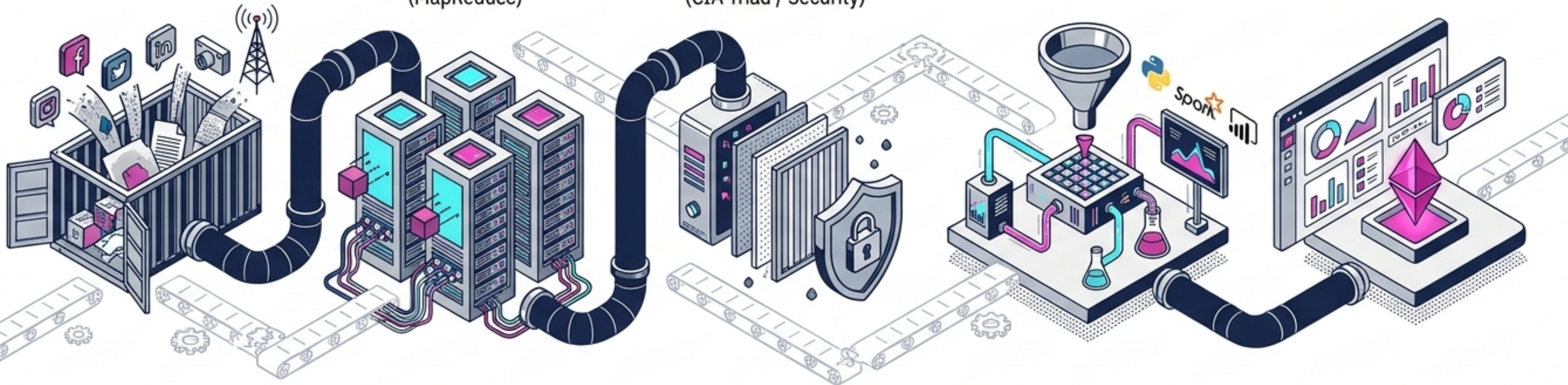
คัดกรองข้อมูลเสีย (Clean Data) และทางโล่ป้องกันภัย (CIA Triad / Security)

Phase 4: Analytics Lab (วิเคราะห์)

สกัด Insight ผ่านเครื่องมือระดับโลก (Python, Spark, Power BI)

Phase 5: Actionable Insight (ผลลัพธ์)

กราฟสรุปผลเพื่อการตัดสินใจทางธุรกิจที่เหนือกว่าคู่แข่ง



Phase 1: Ingestion (แหล่งที่มา)

โซเชียลมีเดีย, เซ็นเซอร์ IoT (ข้อมูลมหาศาล ไร้โครงสร้าง)

Phase 2: Storage & Engine (จัดเก็บและประมวลผล)

ระบบ HDFS และ YARN
กระจายทรัพยากรทำงานแบบคู่ขนาน (MapReduce)

Phase 3: Governance (ควบคุมคุณภาพ)

คัดกรองข้อมูลเสีย (Clean Data) และทางโล่ป้องกันภัย (CIA Triad / Security)

Phase 4: Analytics Lab (วิเคราะห์)

สกัด Insight ผ่านเครื่องมือระดับโลก (Python, Spark, Power BI)

Phase 5: Actionable Insight (ผลลัพธ์)

กราฟสรุปผลเพื่อการตัดสินใจทางธุรกิจที่เหนือกว่าคู่แข่ง

ข้อมูลไม่ใช่แค่ 'ทรัพยากร' แต่คือ "พลังขับเคลื่อนธุรกิจ"



องค์กรที่นำ Data Analytics Tools มาใช้อย่างถูกต้อง สามารถก้าวนำหน้าคู่แข่ง
เปลี่ยนความเสี่ยงเป็นโอกาส และตัดสินใจได้อย่างแม่นยำ

เริ่มต้นสร้างสถาปัตยกรรมข้อมูลของคุณตั้งแต่วันนี้ เพื่องานวิเคราะห์แห่งอนาคต