



ทำความเข้าใจกับ Data Analytics โลกของการแข่งขันด้วยข้อมูล

ต่อ จากฉบับที่แล้ว

วิษณุคุณท์ เมาเรพพ์

ที่ปรึกษาโครงการสารสนเทศคอมพิวเตอร์หน่วยงานภาครัฐ

สำนักตลาดทุนวิจัยและให้คำปรึกษา

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

ปัจจุบันเทคโนโลยีด้านข้อมูลทั้งในแง่ของการจัดเก็บ และการนำมาใช้ประโยชน์นั้นได้รับการพัฒนาไปอย่างรวดเร็วแบบก้าวกระโดด ส่งผลต่อการปรับเปลี่ยนพฤติกรรมการใช้งานข้อมูลเพื่อวัตถุประสงค์ต่างๆ ทั้งที่เป็นข้อมูลส่วนบุคคลที่เกิดขึ้นในชีวิตประจำวัน อาทิ ประวัติการใช้บริการ ซื้อ-ขายแลกเปลี่ยน การทำธุรกรรมต่างๆ ข้อมูลการเข้าถึงสื่อความบันเทิง เครือข่ายสังคม ฯลฯ ข้อมูลจาก Sensor ของอุปกรณ์ Internet of thing รวมถึงข้อมูลส่วนตัวอย่างเช่น ข้อมูลสุขภาพ เป็นต้น ทำให้เกิดเทคโนโลยีในช่วงเวลานี้คงจะคุ้นหูคุ้นตาหลายๆ ท่าน และได้ลองสัมผัสมาบ้างแล้วอย่าง “Cloud Services” ซึ่งเป็นบริการหลากหลายรูปแบบที่เกิดขึ้นจากเทคโนโลยี Cloud Computing ให้เราได้ใช้งานกันโดยบริการต่างๆ ก็ผ่นกลไกกลืนไปกับวิถีชีวิตในปัจจุบันจนแทบจะไม่ได้รู้สึกว่ามันมีการเปลี่ยนแปลงไปมากนักน้อยเพียงใด ยกเว้นก็แต่ความรู้สึกว่าสะดวก และรวดเร็วที่เพิ่มมากขึ้น

อีกเทคโนโลยีก็คือ “Big Data” ซึ่งถ้าจะให้นิยามกันให้เข้าใจง่ายๆ ก็คือข้อมูลปริมาณมหาศาลที่อยู่ในลักษณะ และรูปแบบต่างๆ ทั้งตัวเลข ข้อความ เสียง รูปภาพ วิดีโอ ฯลฯ ซึ่งส่วนใหญ่เกิดจากการให้บริการต่างๆ ที่มีการเชื่อมต่อกันของผู้ให้บริการบนเครือข่าย Internet ทั้งที่เป็น Private และ Public ทั้งที่เป็นส่วนตัว และขององค์กร ถูกจัดเก็บไว้บนทรัพยากรทางสารสนเทศที่เป็นส่วนตัวใน

อุปกรณ์สื่อสารพกพา อุปกรณ์พวก Wearable ต่างๆ ในคอมพิวเตอร์ และที่เป็นขององค์กร อยู่ในเครื่องแม่ข่ายที่ให้บริการ เครื่องแม่ข่ายที่ใช้จัดเก็บข้อมูลประวัติการทำธุรกรรมต่างๆ รวมไปถึงที่อยู่ในรูปแบบของ Cloud Storage ซึ่งมีการเชื่อมต่อกับเครือข่าย Internet ข้อมูลมหาศาลเหล่านี้มีคุณลักษณะพิเศษ คือ มันมีปริมาณที่มากมายมหาศาล และมีรูปแบบที่หลากหลายหรือแม้กระทั่งไม่มีรูปแบบที่ตายตัว จนการจัดเก็บ และบริหารจัดการข้อมูลในรูปแบบ



เดิมๆ ใช้ไม่ได้ผล ต้องมีการพัฒนารูปแบบ และวิธีการจัดเก็บรวมถึงการนำมาใช้ใหม่ เป็นข้อมูลที่มีการเปลี่ยนแปลงอยู่ตลอดเวลาถึงขั้นในระดับเสียวินาทีทำให้การจัดเก็บต้องมีความยืดหยุ่นสูงเพื่อรองรับการเพิ่มขึ้นหรือเปลี่ยนแปลงของข้อมูลได้ และอุปสรรคสำคัญคือการนำมาใช้ประโยชน์เพราะข้อมูลมหาศาลที่จัดเก็บมาในทุกวันนี้ จะต้องผ่านกระบวนการวิเคราะห์ที่เข้มข้นก่อนนำมาใช้ประโยชน์ได้อย่างน่าเชื่อถือ (มีข้อมูลอื่นๆ ที่ไม่จำเป็นปะปนมาด้วย) ซึ่งหากองค์กรต้องการจะใช้ประโยชน์จากข้อมูลเหล่านี้อาจจำเป็นต้องวางแผนกันตั้งแต่การจัดเก็บกันเลยทีเดียว แต่ก็ไม่ได้หมายความว่าต้องจัดการกับข้อมูลทั้งหมดที่จะนำมาใช้ เพียงแต่ต้องเข้าใจว่าจะนำข้อมูลส่วนไหนมาใช้ประโยชน์อะไร และพิจารณาว่าข้อมูลที่นำมาใช้นั้นเป็นข้อมูลที่ต้องจัดการจัดเก็บเอง อาทิ ข้อมูลการใช้บริการของลูกค้า คู่ค้า พันธมิตรหรือหน่วยงานภายในองค์กรหรือเป็นข้อมูลที่มาจากการซื้อขายสังคัมต่างๆ ข้อมูลการใช้บริการจากผู้ให้บริการอื่นๆ ที่มีความเกี่ยวเนื่องซึ่งเป็นข้อมูลที่ Public หรือสามารถร้องขอเพื่อนำมาใช้ประโยชน์ได้ตามข้อตกลงต่างๆ เหล่านี้คือสิ่งที่ต้องทบทวน และดำเนินการก่อนที่จะเริ่มนำเอา Big Data มาใช้ประโยชน์ สรุปง่ายๆ คือ ควรทำความเข้าใจก่อนว่าข้อมูลอะไรบ้างที่จะต้องใช้ จะต้องจัดเก็บไว้ในรูปแบบใด เก็บไว้ที่ไหน สามารถเก็บไว้ได้นานแค่ไหนต้องปรับปรุงทุกๆ ช่วงเวลาใด รองรับปริมาณได้มากน้อยเพียงใดหรือจะนำข้อมูลที่จำเป็นต้องใช้จากแหล่งจัดเก็บข้อมูลจากที่อื่นๆ ออกมาใช้ได้อย่างไรบ้าง

เมื่อเรามี Big Data แล้วขั้นตอนต่อมาก็คือการนำมาวิเคราะห์และใช้ประโยชน์ ซึ่งมันก็ไม่ได้ง่ายอย่างที่คิด เนื่องจากเทคโนโลยีและรูปแบบที่ใช้ในการจัดเก็บ รวมถึงแหล่งที่มาของข้อมูลมี

หลากหลายมาก หากจะนำข้อมูลออกมาใช้งานก็ต้องดำเนินการให้สอดคล้องกับเทคโนโลยีการจัดเก็บนั้นๆ ต้องอาศัยเทคนิควิธีการเฉพาะ ซึ่งขั้นตอนนี้ก็จำเป็นต้องอาศัยเจ้าหน้าที่สารสนเทศที่มีความรู้เฉพาะทางเข้ามาดำเนินการ และเมื่อสามารถเตรียมข้อมูลพร้อมสำหรับการวิเคราะห์ได้แล้ว จึงจะเข้าสู่ขั้นตอนของการทำ “Data Analytics”

ก่อนที่เราจะมีการทำ Data Analytics นั้น ในยุคที่ยังมีการจัดเก็บข้อมูลการบริการ ข้อมูลลูกค้า ข้อมูลธุรกรรม ข้อมูลทางการตลาด การสั่งซื้อวัตถุดิบ การขนส่ง ข้อมูลคู่ค้า ฯลฯ ไว้ที่ระบบภายในขององค์กรหรืออาศัยข้อมูลจากแหล่งภายนอกซึ่งมีปริมาณไม่มากนัก และมีรูปแบบการจัดเก็บที่ชัดเจน มีความน่าเชื่อถือเพราะสามารถควบคุมการเกิดขึ้น และรูปแบบในการจัดเก็บข้อมูลได้ เราจะนำข้อมูลดังกล่าวมาวิเคราะห์ด้วย Software ที่เรียกกันว่า Business Intelligence หรือ BI เพื่อให้ได้ผลลัพธ์เป็นแนวโน้ม ทิศทางต่างๆ หรือข้อมูลเชิงลึกเพื่อประกอบการตัดสินใจ ซึ่งอาจจะมีการผนวก BI เชื่อมโยงเข้ากับระบบ MIS (Management Information System) หรือ DSS (Decision Support System) โดยข้อมูลจะถูกวิเคราะห์ และจัดทำเป็นลักษณะของ Cube เพื่อประโยชน์ในการทำ Report ที่แสดงมุมมองของข้อมูลในมิติต่างๆ ที่สามารถบิดมุมมองได้ตามที่ต้องการ และมีความซับซ้อนพอสมควร แต่เมื่อจะต้องวิเคราะห์ Big Data วิธีการที่กล่าวมาไม่สามารถนำมาใช้งานได้โดยง่าย ต้องมีการพัฒนา และปรับปรุงเทคโนโลยีในการทำ Analytics ให้รองรับรูปแบบของ Big Data และต้องมีการลงทุนเพื่อจัดหาทรัพยากรที่ใช้ในการประมวลผล ซึ่งปัจจุบันเรามีเทคโนโลยี Virtual machine หรือเรียกสั้นๆ ว่า VM มี Cloud Computing ก็ช่วยลด

ค่าใช้จ่ายในส่วนดังกล่าวได้มากพอสมควร มีบริการให้เช่าทรัพยากรเพื่อการประมวลผลข้อมูลโดยมีข้อตกลงในการรักษาความลับ (Confidential disclosure agreement) และ SLA (Services Level Agreement) ที่เหมาะสม

อ่านต่อฉบับหน้า

