



Big Data

มวลมหาข้อมูลที่ถูกซ่อนอยู่รอบกระแสนวัตกรรม

วิษณุฤทธิ์ เมาระพงษ์

ที่ปรึกษาโครงการสารสนเทศของหน่วยงานภาครัฐ
สังกัดสถาบันวิจัยเพื่อการพัฒนาประเทศไทย
แห่งมหาวิทยาลัยธรรมศาสตร์

ต่อ จากฉบับที่แล้ว

การพัฒนา Hadoop ในช่วงแรกๆ เกิดจาก Facebook และ Twitter เหตุเพราะว่าทั้ง 2 องค์กรขนาดใหญ่ด้าน Internet ของโลกมีผู้ใช้งานอยู่เป็นจำนวนมาก และเป็นแหล่งผลิตข้อมูลจำนวนมากในแต่ละวันข้อมูลเหล่านี้มีค่ามากเพราะมันสามารถบอกได้ว่าผู้ใช้งานชอบไม่ชอบอะไร บริโภคอะไร และสามารถนำมาใช้ประกอบการคัดเลือกเนื้อหา (target content) ให้เหมาะกับผู้ใช้แต่ละคนได้ ตัวบริการเองก็จะสร้างประโยชน์กับผู้ใช้งานได้มากขึ้น และการลงโฆษณาก็แม่นยำมากขึ้น ดังนั้น Hadoop ในยุคแรกๆ จึงออกแบบมาสำหรับการใช้งานลักษณะนี้เป็นหลัก ภายหลัง Hadoop ถูกพัฒนาขึ้นให้สามารถนำมาใช้วิเคราะห์ข้อมูลลักษณะอื่นๆ ได้ เช่น

- สามารถวิเคราะห์ข้อมูลขนาดใหญ่ๆ ได้ทั้งหมด ไม่ต้องแบ่งวิเคราะห์เฉพาะกลุ่มตัวอย่างขนาดจำกัด
- สามารถวิเคราะห์ข้อมูลจากหลายๆ แหล่ง (Multiple data sources) ได้เลย ไม่จำกัดเฉพาะฐานข้อมูลที่จัดเก็บข้อมูลอย่างเป็นระบบ (Structured records) แบบที่องค์กรเคยใช้ในอดีต
- สามารถวิเคราะห์ข้อมูลที่ไม่มี Schema ได้ เหมาะสำหรับงานที่เราไม่รู้ฟอร์แมตหรือแพทเทิร์นของข้อมูลจนกว่าจะเริ่มการวิเคราะห์

ด้วยการพัฒนาดังกล่าวประกอบการนำมาประยุกต์ใช้งานได้ อย่างเป็นรูปธรรม และประสบผลสำเร็จ ส่งผลให้ Hadoop กลายเป็นเครื่องมือที่ขาดไม่ได้สำหรับภาคธุรกิจที่ต้องตัดสินใจโดยอิงกับแนว

โน้มใหม่ๆ พฤติกรรมของผู้บริโภค ที่สกัดมาได้จากแหล่งข้อมูลที่หลากหลายในปริมาณที่มหาศาล

ในปัจจุบันหลายๆ องค์กรใช้ Hadoop Technology ในการพัฒนา Big Data ของตน อาทิ Facebook Yahoo และ Twitter นอกจากนี้ Technology Vendor ต่างๆ อาทิ Oracle IBM EMC² หรือแม้แต่ Microsoft ต่างก็นำ Hadoop มาประยุกต์ใช้ในเทคโนโลยีของตัวเองเพื่อพัฒนาผลิตภัณฑ์ทางด้าน Big Data ทั้งนี้ Hadoop ไม่ได้มาแทนที่ระบบฐานข้อมูลในรูปแบบเดิมแต่เป็นการใช้งานร่วมกัน ทั้งแบบเดิมที่เป็น Structure Data และการนำ Unstructure Data ขององค์กรที่อาจเก็บไว้ในระบบอย่าง Hadoop เข้ามาพิจารณาร่วมกับข้อมูลอื่นๆ ภายนอก เช่น Social Network App อย่าง Facebook Tweeter แล้วนำมาวิเคราะห์ข้อมูลโดยใช้เครื่องมืออย่าง Business Intelligence



นี้เป็นเพียงการพัฒนาที่เกิดขึ้นในปัจจุบัน แต่การพัฒนายังดำเนินไปอย่างต่อเนื่อง และรวดเร็ว ทำให้ Big Data กลายเป็นประเด็นที่ส่งผลกระทบต่อการทำงานขององค์กรต่างๆ รวมทั้งในแง่ของกฎหมายที่อาจยังไล่ตามเทคโนโลยีอยู่ห่างๆ ซึ่งบางครั้งในบางกรณีอาจสร้างช่องโหว่ที่อาจกลายเป็นอุปสรรคสำคัญต่อการพัฒนาการรักษาความเป็นส่วนตัว ข้อมูลส่วนบุคคล ทรัพย์สินทางปัญญา ความมั่นคงปลอดภัยด้านข้อมูล



ในอีกแง่มุมหนึ่ง คือ การพึ่งพาข้อมูลอย่างเข้มข้น จนอาจล้มค้ำจนถึงคุณภาพของข้อมูล แม้ผลการวิเคราะห์จะแม่นยำเพียงใด แต่ความน่าเชื่อถือของข้อมูลก็ยังคงเป็นดาบ 2 คมที่ต้องให้ความใส่ใจ อาจต้องพึ่งพาวิจารณ์ญาณ และทักษะการตัดสินใจของผู้บริหารองค์กรการเป็นสำคัญ

บทสรุป

ในทุกๆ วัน โลกเราสร้างข้อมูลมากถึง 2.5 quintillion byte (2.5×10^{18} byte) ข้อมูลเหล่านี้มาจากทุกที่ ไม่ว่าจะเป็นจากระบบ Sensor ต่างๆ ทั้งใน Mobile Device ใน Gadget ต่างๆ Sensor ที่ใช้ในการรวบรวมข้อมูลสภาพภูมิอากาศ การเคลื่อนตัวของเปลือกโลก Sensor จับการเคลื่อนไหว ฯลฯ การ Post ข้อความ รูปภาพ ที่ส่งไปยัง Social Media ต่างๆ Image Capture Video Streaming บันทึกการขาย การซื้อขาย (purchase transaction) ต่างๆ สัญญาณ GPS จากโทรศัพท์มือถือ รวมทั้งข้อมูลที่เกิดจากการใช้บริการระบบสารสนเทศขององค์กรต่างๆ เป็นต้น ข้อมูลที่ถูกนิยามว่าเป็น Big Data เหล่านี้วนเวียนและเกี่ยวข้องกับกระแสของธุรกิจในปัจจุบัน

ด้วยปัจจัยต่างๆ และการพัฒนาด้านเทคโนโลยีทำให้องค์กรต่างๆ เข้าถึง และสามารถนำเอา Big Data ที่ต้องการมาจัดเก็บ และวิเคราะห์ เพื่อใช้เป็นข้อมูลการคาดการณ์ต่างๆ ที่สำคัญต่อการก้าวเดินขององค์กร สิ่งที่ต้องคำนึงถึงก็คือ ศักยภาพขององค์กรว่าจะสามารถจัดการกับข้อมูลในระดับใดได้ด้วยตนเอง หรือในบางส่วนจำเป็นต้องลงทุนเพื่อให้ได้ผลลัพธ์จากการนำเอา Big Data ที่องค์กรให้ความสนใจมาใช้เสริมศักยภาพในการประเมินสถานการณ์การแข่งขันในตลาด เพราะประโยชน์ที่ได้จาก Big Data คือการผสมผสานกันระหว่างการใช้ข้อมูลวิเคราะห์ด้วยข้อมูลในกระบวนการทางธุรกิจหลักขององค์กรที่มีรูปแบบเฉพาะ ซึ่งองค์กรมีความชำนาญในการเข้าถึง และสกัดเพื่อนำมาใช้ประโยชน์ซึ่งถือเป็นประโยชน์จากการประเมินจากภายใน ส่วนผลการวิเคราะห์ในมิติของ Big Data ถือว่าเป็นการประเมินในแง่มุมที่องค์กรให้ความสนใจ เป็นข้อมูล และประเด็นสำคัญที่เกิดขึ้นในสถานการณ์ที่เปลี่ยนแปลง ณ ขณะนั้น ของสภาวะแวดล้อมภายนอก พฤติกรรมผู้บริโภค การปฏิสัมพันธ์

การสื่อสารต่างๆ ที่เกิดขึ้น ทั้งนี้เพื่อมาเติมเต็มมุมมอง ในการคาดการณ์ พยากรณ์ก้าวที่จะเดินต่อไปขององค์กรด้วยความมั่นคง และเชื่อถือได้ ซึ่งส่วนใหญ่จะเป็นการลงทุนในเชิง Outsource ทั้งในแง่ของทรัพยากร ผู้เชี่ยวชาญ ที่ปรึกษา รวมถึงตัว Big Data เองด้วย

นอกจากนี้การพัฒนา Big Data ที่สำคัญประการหนึ่ง นอกเหนือจากการปรับปรุงโครงสร้างระบบเทคโนโลยีสารสนเทศขององค์กรด้านข้อมูล (Information Infrastructure) แล้ว ยังต้องมีการพัฒนาบุคลากรให้เข้าใจถึงเทคโนโลยีด้าน Big Data และต้องมีการวางแผนในการนำข้อมูลทั้ง Structure และ Unstructure จากภายใน

และภายนอกองค์กรมาใช้งาน รวมทั้งการมองหาผู้เชี่ยวชาญทางด้านข้อมูลที่เป็น Data Scientist มาร่วมทำงาน

การที่องค์กรก้าวเข้าไปสู่จุดที่นำเอา Big Data มาใช้งานจึงไม่ใช่การลงทุนใน Scale ทั่วไปแต่เป็นการลงทุนที่ใช้งบประมาณมากพอสมควร เพื่อหวังผลของความได้เปรียบในเชิงธุรกิจอย่างแท้จริง และไม่ใช่ในมิติด้าน IT เพียงด้านเดียว แต่คือการเปลี่ยนแปลงรูปแบบการทำธุรกิจ หมายความว่าองค์กรที่จะใช้งาน Big Data อย่างจริงจังต้องมีองค์ประกอบ 3 ประการ ได้แก่

- **Data** ต้องมีการจัดเก็บข้อมูลไว้เพื่อการวิเคราะห์ ซึ่งในอนาคตจะต้องอยู่ในรูปแบบของ digital form
- **IT infrastructure** ต้องจัดเตรียมโครงสร้างพื้นฐานทางด้าน IT เพื่อรองรับการจัดเก็บ และวิเคราะห์ข้อมูลขนาดใหญ่ (หมายรวมถึงไปถึงการ Outsourcing)
- **Analytical Skills** การวิเคราะห์ข้อมูล ซึ่งจำเป็นต้องใช้เครื่องมือ และผู้เชี่ยวชาญที่เหมาะสม

ทั้งนี้ เมื่อก้าวโดยสรุปแล้ว Big Data สำหรับองค์กรก็คือการวางยุทธศาสตร์ในการจัดเก็บข้อมูล การกำหนดรูปแบบวิธีการวิเคราะห์ข้อมูล และการนำข้อมูลมาใช้กับกระบวนการทำธุรกิจ ซึ่งเดิมองค์กรเองอาจไม่ได้ให้ความสำคัญมากเท่าไรนัก ไม่ได้มีการวางแผนไว้อย่างเป็นทางการทั้งกระบวนการ แต่ในยุคปัจจุบัน “ข้อมูล” นั้นถือเป็นสิ่งสำคัญที่สุดในการชิงความได้เปรียบในการทำธุรกิจไปแล้ว

แหล่งข้อมูลอ้างอิง

- eCommerce ฉบับที่ 182, Feb 2014 in How to by ดร.เสกฐิติวิทย์ เกิดผล
- Graham M. 2012. Big data and the end of theory?. The Guardian
- Shah, Home and Capella. 2012. Good Data Won't Guarantee Good Decisions. Harvard Business Review
- “What is big data?”, IBM
- Social business, analytics, mobility and cloud computing
- “Data, data everywhere,” The Economist Feb 25th, 2010
- How ‘big data’ is transforming business, McKinsey
- Big data: The next frontier for innovation, competition, and productivity, McKinsey
- Big Data Now: Current Perspectives from O'Reilly Radar
- BIG DATA และ ANALYTICS เอาชนะตลาดด้วยฐานข้อมูล Cloud Computing, ดร.อธิป อัครวานันท์ 