

รายงานการศึกษาส่วนบุคคล
(Individual Study)

เรื่อง แนวทางการบริหารจัดการข้อมูลขนาดใหญ่
ของกรุงเทพมหานคร โดยใช้ ซอฟต์แวร์จัดการข้อมูล
ขนาดใหญ่ ประเภท Open Source
: Hadoop Technology

จัดทำโดย นายณที เทียนประทีป
ตำแหน่ง นักวิชาการคอมพิวเตอร์ชำนาญการ
สังกัด กองบริการระบบคอมพิวเตอร์ สำนักยุทธศาสตร์และประเมินผล

รายงานนี้เป็นส่วนหนึ่งของการฝึกอบรม
หลักสูตรนักบริหารมหานครระดับต้น รุ่นที่ ๒๐
สถาบันพัฒนาข้าราชการกรุงเทพมหานคร
ประจำปีงบประมาณ พ.ศ. ๒๕๕๗

บทสรุปผู้บริหาร

ปัจจุบันปริมาณข้อมูลที่เกิดจากการทำงานประจำวัน รวมทั้งข้อมูลที่ได้จากสื่อออนไลน์ ผ่านเครือข่ายสังคมที่มีปริมาณมากขึ้นอย่างต่อเนื่อง ส่งผลให้มีข้อมูลจำนวนมากที่มีขนาดใหญ่เกินกว่าที่อุปกรณ์จัดเก็บข้อมูลจะสามารถจัดการได้ หากไม่มีการวางแผนและการบริหารจัดการที่ดี ย่อมส่งผลกระทบต่อประสิทธิภาพการทำงานโดยรวมของกรุงเทพมหานคร เพื่อเตรียมความพร้อมสำหรับปัญหาดังกล่าว จึงได้จัดทำรายงานการศึกษาส่วนบุคคล เรื่อง แนวทางการบริหารจัดการข้อมูลขนาดใหญ่ของ กรุงเทพมหานคร โดยใช้ ซอฟต์แวร์จัดการข้อมูลขนาดใหญ่ ประเภท Open Source : Hadoop Technology โดยมีวัตถุประสงค์หลักเพื่อบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานคร ให้เป็นไปอย่างมีประสิทธิภาพ เพิ่มความสามารถในการคัดกรองความรู้เชิงลึกจากข้อมูลที่มีอยู่ ให้เกิดประโยชน์สูงสุดต่อกรุงเทพมหานคร โดยมีเป้าหมายเพื่อให้กรุงเทพมหานคร มีระบบการบริหารจัดการเก็บข้อมูลขนาดใหญ่ โดยใช้ ซอฟต์แวร์จัดการข้อมูลขนาดใหญ่ ประเภท Open Source : Hadoop Technology

โดยรายงานการศึกษานี้ได้อธิบายถึงความหมายและที่มาของปริมาณข้อมูลขนาดใหญ่ รวมถึงการจัดการข้อมูลขนาดใหญ่ด้วย Hadoop Technology ซึ่งสามารถแยกประเภทของข้อมูลได้ว่า ข้อมูลใดจัดเป็นข้อมูลขนาดใหญ่ ตามคุณสมบัติเบื้องต้น คือ ปริมาณ ความเร็ว ความหลากหลาย นอกจากนี้ยังอธิบายถึงกระบวนการ ในการจัดการข้อมูลขนาดใหญ่ เทคโนโลยีที่ใช้ในการจัดการข้อมูลขนาดใหญ่คือ Hadoop Technology ซึ่งเป็น Freeware ที่ไม่มีภาระค่าใช้จ่ายในการใช้งาน มีความยืดหยุ่นสูง และสามารถประมวลผลกับข้อมูลขนาดใหญ่ให้ได้ผลลัพธ์แบบทันที (Real-time) กรอบหรือระยะเวลาในการดำเนินการ ประมาณ ๑๘๐ วัน หรือ หกเดือน โดยมีผู้ที่เกี่ยวข้องทั้งในส่วนผู้บริหารด้านเทคโนโลยีสารสนเทศของกรุงเทพมหานคร (CIO) และเจ้าหน้าที่ด้านเทคโนโลยีสารสนเทศของกรุงเทพมหานคร มีการวัดความสำเร็จทั้งในระดับผลผลิต (Output) และในระดับผลลัพธ์ (Outcome) ด้วยการสำรวจความคิดเห็นของผู้ที่เกี่ยวข้อง ผ่านแบบสำรวจความพึงพอใจระดับไม่น้อยกว่าร้อยละ ๙๐

ประโยชน์ที่กรุงเทพมหานครจะได้จากการนำ Hadoop Technology มาใช้ในการบริหารจัดการข้อมูลขนาดใหญ่ คือ ช่วยลดการสูญเสียด้านงบประมาณเกี่ยวกับการลงทุนด้านซอฟต์แวร์ที่ใช้ในการจัดการฐานข้อมูลที่หลากหลายของกรุงเทพมหานคร เพิ่มประสิทธิภาพการบริหารจัดการสินทรัพย์ด้านซอฟต์แวร์ เป็นเครื่องมือที่ช่วยเพิ่มขีดความสามารถให้บริการด้านข้อมูล ช่วยเพิ่มความปลอดภัยให้กับระบบ สร้างความเชื่อมโยงระหว่างข้อมูลข่าวสารที่มีอยู่ในระบบ เป็นการเพิ่มมูลค่าให้กับข้อมูลที่มีอยู่เดิม เพิ่มขีดความสามารถในการแข่งขันให้กรุงเทพมหานคร ในระดับมหานครทั่วโลก ทำให้กรุงเทพมหานครมีระบบฐานข้อมูลที่มั่นคง มีประสิทธิภาพ เป็นพื้นฐานสำหรับการพัฒนาไอทีต่อไปในอนาคต

๑. ชื่อเรื่อง

แนวทางการบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานคร โดยใช้ ซอฟต์แวร์จัดการข้อมูลขนาดใหญ่ ประเภท Open Source : Hadoop Technology

๒. หลักการและเหตุผล

แนวโน้มทิศทางเทคโนโลยีสารสนเทศของโลกในยุคปัจจุบัน ต่างมุ่งไปสู่ทิศทางเดียวกัน คือเน้นการขับเคลื่อนกระบวนการผ่านเครือข่ายอินเทอร์เน็ต มีหลายปัจจัยที่เอื้อให้ปริมาณข้อมูลมีจำนวนเพิ่มมากขึ้น ทั้งข้อมูลที่เกิดจากการทำงานประจำวัน รวมทั้งข้อมูลที่ได้จากสื่อออนไลน์ผ่านเครือข่ายสังคม (Social Network) ที่มีปริมาณมากขึ้นอย่างต่อเนื่อง ส่งผลให้มีข้อมูลจำนวนมากที่มีขนาดใหญ่เกินกว่าที่อุปกรณ์จัดเก็บข้อมูลจะสามารถจัดการได้ หากไม่มีการวางแผนและการบริหารจัดการที่ดี ย่อมส่งผลกระทบต่อประสิทธิภาพการทำงานโดยรวมของกรุงเทพมหานคร การแก้ไขปัญหาทำได้โดย กำหนดแนวทางการบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานคร โดยใช้ ซอฟต์แวร์จัดการข้อมูลขนาดใหญ่ ประเภท Open Source : Hadoop Technology ซึ่ง Hadoop Technology เป็นซอฟต์แวร์ที่ช่วยบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานครที่มีจำนวนมาก ให้เป็นไปอย่างมีประสิทธิภาพ และช่วยเพิ่มขีดความสามารถในการคัดกรองความรู้เชิงลึกจากข้อมูลเดิมที่มีอยู่ในระบบ ให้เกิดประโยชน์สูงสุด

ปัญหาดังกล่าวส่งผลกระทบโดยตรงต่อการเรียกใช้ข้อมูล และเมื่อคนจำนวนมากต่างสร้างข้อมูลจำนวนมากมายมมหาศาล จึงทำให้เกิดการไหลเวียนของข้อมูลจำนวนมาก จนข้อมูลเหล่านั้นกลายเป็นปริมาณข้อมูลขนาดใหญ่หรือที่เรียกว่า Big Data การที่มีข้อมูลใหม่ๆ เกิดขึ้นและเปลี่ยนแปลงอย่างรวดเร็ว และความต้องการข้อมูลเป็นแบบทันที (Real-time) นอกจากนี้การพัฒนาและใช้งานระบบงานสารสนเทศทั้งของหน่วยงานภาครัฐและภาคธุรกิจ ล้วนทำให้เกิดข้อมูลต่างๆ ซึ่งถูกจัดเก็บสะสมอยู่ในรูปแบบของระบบฐานข้อมูลจากอดีตจนถึงปัจจุบัน ข้อมูลจึงมีปริมาณมากขึ้นเรื่อยๆ ทำให้ระบบงานต่างๆ เหล่านั้น เข้าถึงข้อมูลได้ช้าและประสิทธิภาพในการประมวลผลข้อมูลก็ประมวลผลช้าตามไปด้วย ถึงแม้จะมีการปรับแต่ง (Tuning) ฐานข้อมูลเพื่อเพิ่มการประมวลผลและการเข้าถึงข้อมูลแล้วก็ตาม แต่ก็สามารถเพิ่มประสิทธิภาพได้เพียงระดับหนึ่งและระยะหนึ่งเท่านั้น ในขณะที่ระบบงานยังคงดำเนินงานต่อไป ข้อมูลก็จะค่อยๆ เพิ่มขึ้นเรื่อยๆ จนทำให้มีปริมาณมากขึ้นกว่าเดิม ซึ่งปัญหาการเข้าถึงและการประมวลผลข้อมูลก็ยังคงเป็นปัญหาที่จะต้องได้รับการบริหารจัดการอย่างถูกต้องต่อไป

กรุงเทพมหานคร เป็นองค์กรที่มีหน่วยงานภายในจำนวนมาก ประกอบด้วย หน่วยงานระดับสำนักหรือเทียบเท่า จำนวน ๒๑ หน่วยงาน หน่วยงานระดับสำนักงานเขต จำนวน ๕๐ สำนักงานเขต หน่วยงานพาณิชย์กรุงเทพมหานคร จำนวน ๖ หน่วยงาน ฯลฯ ซึ่งต้องให้บริการประชากรในกรุงเทพมหานครประมาณ ๕ ล้านกว่าคน (สำนักงานสถิติแห่งชาติ) บวกประชากรแฝงอีกหนึ่งเท่าตัว หรือประมาณ ๕ ล้านกว่าคน ด้วยจำนวนหน่วยงานที่มากมายเหล่านี้ ล้วนเป็นเหตุที่ทำให้เกิดปริมาณข้อมูล (Volume) ที่มีจำนวนเพิ่มขึ้นอย่างมหาศาล ไม่ว่าจะเป็นข้อมูลที่เกิดขึ้นจากการติดต่อในแต่ละวัน (Transaction) ข้อมูลเก่าที่ถูกจัดเก็บไว้เป็นระยะเวลาหลายปี ข้อมูลในรูปข้อความ (Text) รวมถึงรูปภาพ (Image) ที่เกิดขึ้นอย่างต่อเนื่องจากสื่อสังคมออนไลน์ และการเพิ่มขึ้นของจำนวนข้อมูลที่เกิดจากเครื่องที่ใช้ตรวจจับสถานะต่าง (Sensor) ที่มีการจัดเก็บ รวบรวมไว้ จนปริมาณข้อมูลได้ระดับจาก กิกะไบต์ (Gigabyte) เทระไบต์ (Terabyte) เพตะไบต์ (Petabyte) เอกซะไบต์ (Exabyte) ไปจนถึง เซตตะไบต์

(Zettabyte) ซึ่งเป็นระดับที่ยังไม่มีระบบการเก็บข้อมูลใดที่สามารถจัดการข้อมูลขนาดเซตตะไบต์ได้นอกจากปริมาณข้อมูลจำนวนมากแล้วข้อมูลเหล่านั้นยังมีความหลากหลาย (Variety) ที่แตกต่างกัน จากลักษณะของข้อมูลประกอบด้วยข้อมูลแบบมีโครงสร้างแน่นอน (Structured) แบบไม่มีโครงสร้าง (Unstructured) หรือสุดท้ายแบบกึ่งโครงสร้าง (Semi-structured) ซึ่งการจะวิเคราะห์ข้อมูลปริมาณมหาศาล ที่มีความหลากหลายข้างต้น ย่อมประมวผลได้ยาก ถึงแม้จะประมวผลได้ ก็ต้องใช้ขั้นตอนและใช้เวลาค่อนข้างมาก สิ่งเหล่านี้ย่อมส่งผลต่อความเร็ว (Velocity) ในการเรียกใช้ข้อมูลภายในเวลาที่ต้องการ เพื่อตอบสนองต่อความต้องการของประชาชนที่เปลี่ยนไป ตามการดำเนินชีวิต (Life Style) ไม่นับรวมความแปรผัน (Variability) อันเกิดจากปริมาณข้อมูลที่เพิ่มขึ้นในแต่ละวัน ที่ยากต่อการบริหารจัดการ ในขณะที่ข้อมูลรูปแบบเดิมมีการเปลี่ยนแปลงอย่างช้าๆ แต่ข้อมูลรูปแบบใหม่ๆ กลับมีการเปลี่ยนแปลงอย่างรวดเร็ว สิ่งต่างๆเหล่านี้ ถือเป็นปัญหาของกรุงเทพมหานครที่ต้องได้รับการแก้ไขอย่างเร่งด่วน

การนำ Hadoop Technology มาใช้ในการบริหารจัดการข้อมูลขนาดใหญ่ กรุงเทพมหานคร เป็นแนวทางในการแก้ไขปัญหาดังกล่าวที่เหมาะสม เนื่องจาก Hadoop Technology จะช่วยลดการสูญเสียด้านงบประมาณเกี่ยวกับการลงทุนด้านซอฟต์แวร์ที่ใช้ในการจัดการฐานข้อมูลที่หลากหลายของกรุงเทพมหานคร ช่วยเพิ่มประสิทธิภาพการบริหารจัดการสินทรัพย์ด้านซอฟต์แวร์ของกรุงเทพมหานคร เป็นเครื่องมือที่ช่วยเพิ่มขีดความสามารถให้บริการด้านข้อมูลให้มีประสิทธิภาพมากขึ้น ช่วยเพิ่มความปลอดภัยให้กับระบบ Hadoop ยังสร้างความเชื่อมโยงระหว่างข้อมูลข่าวสารที่มีอยู่ในระบบ ช่วยเพิ่มมูลค่าให้กับข้อมูลที่มีอยู่เดิมให้สามารถนำไปใช้เพื่อนำไปสู่การตัดสินใจที่ดีที่สุด ช่วยให้ผู้บริหารกรุงเทพมหานครสามารถกำหนดนโยบาย ที่สามารถตอบสนองต่อความต้องการที่แท้จริงของประชาชน สอดคล้องกับวิสัยทัศน์ของกรุงเทพมหานครที่ต้องการเป็นศูนย์กลางในภูมิภาคเอเชียตะวันออกเฉียงใต้ และมุ่งสู่มหานครน่าอยู่อย่างยั่งยืน ในปี พ.ศ. ๒๕๕๙ มีการกำหนดเป็น กลยุทธ์หลักตามแผนพัฒนากรุงเทพมหานคร ๑๒ ปี ระยะที่ ๒ (พ.ศ. ๒๕๕๖ – ๒๕๕๙) ประเด็นยุทธศาสตร์ที่ ๗ เพิ่มประสิทธิภาพการดำเนินงานด้านเทคโนโลยีสารสนเทศสำหรับการบริหารและพัฒนามหานคร เพิ่มขีดความสามารถในการแข่งขันให้กรุงเทพมหานคร ในระดับมหานครทั่วโลก ทำให้กรุงเทพมหานครมีระบบฐานข้อมูลที่มีมั่นคง มีประสิทธิภาพ เป็นพื้นฐานสำหรับการพัฒนาไอทีต่อไปในอนาคต

๓. วัตถุประสงค์

๑. เพื่อให้การบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานครเป็นไปอย่างมีประสิทธิภาพ เพิ่มการพัฒนาความสามารถในการคัดกรองความรู้และข้อมูลเชิงลึกจากข้อมูลดิจิทัล ที่มีจำนวนมากและอยู่ในรูปแบบที่ซับซ้อน ให้เกิดประโยชน์สูงสุดต่อกรุงเทพมหานคร
๒. เพื่อพัฒนาคุณภาพการบริการให้ตอบสนองความต้องการที่แท้จริง ในลักษณะทันที (Real-time) ทุกที่ ทุกเวลา ผ่านสังคมเครือข่าย (Social Network) ด้วยการบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานคร
๓. เพื่อพัฒนาการใช้เทคโนโลยีสารสนเทศที่เหมาะสม เป็นการใช้ทรัพยากรอย่างคุ้มค่า และเกิดประโยชน์สูงสุด ทั้งในแง่พัฒนาการบริการและการกำหนดกลยุทธ์ ของกรุงเทพมหานคร

๔. เป้าหมาย

กรุงเทพมหานคร มีระบบการบริหารจัดการข้อมูลขนาดใหญ่ ที่มีประสิทธิภาพ ด้วยการนำซอฟต์แวร์ จัดการข้อมูลขนาดใหญ่ ประเภท Open Source : Hadoop Technology มาประยุกต์ใช้ ทำให้การบริหารจัดการข้อมูล เกิดประโยชน์สูงสุดในทุกภาคส่วนที่เกี่ยวข้อง

๕. ปัจจัยความสำเร็จ

การบริหารจัดการข้อมูลขนาดใหญ่ ให้เกิดประโยชน์สูงสุด เพื่อให้เกิดความได้เปรียบทางการแข่งขัน มีปัจจัยความสำเร็จ ตามทฤษฎี Using information to achieve a competitive advantage หรือ การใช้ข้อมูลข่าวสารให้บรรลุและเกิดประโยชน์ทางการแข่งขัน ของ Dr. Steven Gordon ประกอบด้วย

๑. Differentiation หรือ “ความแตกต่าง” ความแตกต่างในที่นี้หมายถึง ความยืดหยุ่นของระบบ ที่สามารถทำงานได้กับข้อมูลทั้งแบบมีโครงสร้างแน่นอน แบบไม่มีโครงสร้าง หรือแบบกึ่งโครงสร้าง การมีบริการที่ดี ตอบสนองต่อความต้องการที่แท้จริง ของผู้ใช้บริการระบบ ด้วยความรวดเร็วในการใช้งานแบบทันทีทันใด (Real-time) ความถูกต้องของข้อมูล ความง่ายในการใช้งาน ความง่ายในการบำรุงรักษา รวมถึงความคงทน ความประหยัด และคุณลักษณะอื่นที่ดีกว่า ทำให้ผู้ใช้งานรู้สึกได้ถึงความแตกต่าง จากองค์กรอื่น

๒. Cost Leadership หรือ การเป็นผู้นำด้านต้นทุน ด้วยการควบคุมค่าใช้จ่าย เพื่อสนองความต้องการลูกค้า ดูแลการให้บริการ การเลือกใช้เทคโนโลยี Hadoop ซึ่งเป็นซอฟต์แวร์ประเภท Open Source ทำให้ประหยัดเรื่องค่าใช้จ่ายในเรื่องลิขสิทธิ์การใช้งาน เป็นการลดต้นทุนในการบริหารจัดการ ให้มีความสำคัญกับการลดการสูญเสีย ควบคุมค่าใช้จ่ายอย่างระมัดระวัง ทำให้การปฏิบัติงานต่างๆเกิดประสิทธิภาพสูงสุด

๓. Focus เป็นการมุ่งเน้นการตอบสนองต่อความต้องการของลูกค้าหรือผู้ใช้บริการ โดยเน้นให้ความสำคัญต่อ “ความต้องการ” เป็นสำคัญ เพื่อตอบสนองต่อความต้องการที่แท้จริง โดยผู้ใช้บริการประกอบด้วย

- ประชาชน ทั่วไป ที่ต้องการข้อมูลข่าวสาร
- เจ้าหน้าที่ที่เกี่ยวข้องในการใช้ข้อมูลข่าวสาร

ซึ่งการมุ่งเน้นไปที่ความต้องการของลูกค้า เพื่อสร้างความพึงพอใจสูงสุดให้เกิดขึ้นเมื่อมาใช้บริการ เพื่อพัฒนาคุณภาพการบริการให้ตอบสนองความต้องการที่แท้จริง ในลักษณะทันที (Real-time) ทุกที่ ทุกเวลา ผ่านสังคมเครือข่าย (Social Network) ด้วยการบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานคร

๔. Linkage เป็นการสร้างความเชื่อมโยง เพื่อรับฟังส่วนสะท้อนกลับจากผู้ใช้บริการ นำกลับมาวิเคราะห์เพื่อปรับปรุงและพัฒนาคุณภาพให้ดียิ่งขึ้น เป็นการสร้างเครือข่ายความสัมพันธ์กับลูกค้าหรือผู้ใช้บริการ กับผู้ผลิตหรือผู้พัฒนา เพื่อให้การบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานครเป็นไปอย่างมีประสิทธิภาพ เพิ่มการพัฒนาความสามารถในการคัดกรองความรู้และ

ข้อมูลเชิงลึกจากข้อมูลดิจิทัล ที่มีจำนวนมากและอยู่ในรูปแบบที่ซับซ้อน ให้เกิดประโยชน์สูงสุดต่อกรุงเทพมหานคร

๕. Information Leadership เป็นผู้นำด้านข้อมูลข่าวสาร ด้วยข้อมูลข่าวสารถือเป็นองค์ประกอบที่สำคัญสำหรับองค์กรที่ประสบความสำเร็จ การประสบความสำเร็จในการบริหารระบบ Information Technology : IT ถือเป็นสิ่งสำคัญต่อการสร้างความได้เปรียบทางการแข่งขัน (Competitive Advantage) เป็นการสร้างการสนับสนุนการร่วมมือระหว่างภาคประชาชนกับภาครัฐ เพิ่มช่องทางการติดต่อสื่อสาร แพร่การใช้สารสนเทศร่วมกัน และการจัดการเชื่อมโยงข้อมูลให้เกิดประโยชน์สูงสุด

การวิเคราะห์สภาพแวดล้อมภายในโดยพิจารณาจากปัจจัยพื้นฐานหรือปัจจัยหลักในการบริหารจัดการ ๔ ด้าน ประกอบด้วย

๑. ทรัพยากรมนุษย์ (Man) มนุษย์เป็นเงื่อนไขที่สำคัญของการเกิดองค์การและการบริหาร กล่าวได้ว่าองค์การและการบริหารทำให้เกิดระบบการแบ่งหน้าที่กันทำงานโดยยึดหลักความสามารถและความถนัด ความรู้ในเรื่องการแบ่งหน้าที่กันทำงาน จัดเป็นเทคโนโลยีของสังคมอย่างหนึ่ง ส่วนความสามารถและความถนัดของมนุษย์เป็นสิ่งที่สามารถฝึกฝนและพัฒนาตามระดับสติปัญญา และระบบการศึกษาของสังคม ในที่นี้หมายถึง บุคลากรของกรุงเทพมหานคร ที่มีหน้าที่เกี่ยวข้องกับการดำเนินการบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานคร โดยเราแบ่งเป็น ๒ ระดับคือ

- ระดับบริหาร หมายถึงผู้บริหารของกรุงเทพมหานคร ทำหน้าที่กำหนดกลยุทธ์ นโยบาย แนวทางการดำเนินการ ไปสู่การปฏิบัติ ในการนำเทคโนโลยี Hadoop มาใช้ในการบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานคร
- ระดับปฏิบัติการ หมายถึง เจ้าหน้าที่ที่เกี่ยวข้องด้านเทคโนโลยีสารสนเทศในทุกละดับของกรุงเทพมหานคร เพื่อทำหน้าที่ศึกษาทำความเข้าใจกับนวัตกรรมด้านการบริหารจัดการข้อมูลขนาดใหญ่ ให้ความรู้ ทำความเข้าใจ ตลอดจนให้การสนับสนุนการปฏิบัติงาน เพื่อให้บรรลุวัตถุประสงค์

๒. ทรัพยากรในรูปเงินทุน (Money) เงินเป็นสื่อกลางในการแลกเปลี่ยนทรัพยากรอื่นๆ เงินเดือน ค่าจ้าง ค่าพัฒนาบุคลากร จัดเป็นค่าใช้จ่ายในการจัดหาทรัพยากรมนุษย์ ที่มีคุณภาพ เงินยังเป็นปัจจัยในการจัดหาเทคโนโลยีต่างๆรวมถึงเป็นค่าใช้จ่ายในการพัฒนาองค์ความรู้ในการบริหารวัตถุดิบ (Material) ในที่นี้หมายถึง งบประมาณที่ใช้ในการดำเนินการบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานคร ซึ่งในส่วนของต้นทุนด้านซอฟต์แวร์ที่ใช้ไม่มีเนื่องจากเป็น Open Source ซึ่งเป็น Freeware เปิดให้ใช้ฟรี คงมีก็แต่เพียง ค่าฝึกอบรมเจ้าหน้าที่ ค่าวัสดุ/ครุภัณฑ์ที่เกี่ยวข้อง ซึ่งถือได้ว่าช่วยประหยัดงบประมาณได้มากเมื่อเทียบกับการใช้เทคโนโลยีแบบอื่น

๓. ทรัพยากรในรูปวัสดุสิ่งของ (material) วัสดุสิ่งของจัดเป็นอุปกรณ์ที่อำนวยความสะดวกในการบริหาร อาจมองได้ว่าเป็นเทคโนโลยีทางการบริหาร พิจารณาได้ตั้งแต่อาคาร สถานที่ทำงาน อุปกรณ์เครื่องใช้ในสำนักงาน รถยนต์ เครื่องมือสื่อสาร ตลอดจนวัตถุดิบต่างๆ รวมไปถึงเครื่องจักรกลที่จำเป็น สิ่งเหล่านี้อาจพิจารณาได้ว่าเป็นวัฒนธรรมทางวัตถุที่มนุษย์มีการพัฒนาถ่ายทอด

สืบต่อกันมา จัดได้ว่าเป็นเทคโนโลยีในการบริหารงาน ในที่นี้ หมายถึง วัสดุ อุปกรณ์ ด้านเทคโนโลยีสารสนเทศ ที่เป็นส่วนหนึ่งของการบริหารจัดการข้อมูลขนาดใหญ่ เช่น เครื่องคอมพิวเตอร์ ที่ทำหน้าที่เครื่องแม่ข่ายส่วนกลางสำหรับผู้บริหาร เครื่องพิมพ์ และอื่นๆที่เกี่ยวข้อง แต่ส่วนใหญ่ใช้วัสดุ อุปกรณ์ด้านเทคโนโลยีสารสนเทศเดิมมีอยู่แล้ว

๔. ความรู้ในการจัดการ (management) ความรู้ในการจัดการเป็นความสามารถของผู้บริหารในอันที่จะดูแล และประสานกิจกรรมต่างๆ ให้ดำเนินลุล่วงไปตามวัตถุประสงค์หรือเป้าหมายที่กำหนดไว้ โดยเน้นหลักการเลือกใช้เทคโนโลยีที่เหมาะสม (Appropriate Technology) ที่มีอยู่ให้เกิดประสิทธิภาพ และประสิทธิผลสูงสุด สร้างความพึงพอใจสูงสุดให้กับผู้ใช้บริการ

เพื่อให้มีการกำหนดกลยุทธ์ที่เหมาะสม จึงนำกระบวนการวิเคราะห์ SWOT ซึ่งเป็นวิธีการและเทคนิคในการวิเคราะห์สภาวะขององค์กรและกระบวนการกำหนดกลยุทธ์ มาใช้ เพื่อการวิเคราะห์สภาพแวดล้อม (Environmental Analysis) ปัจจัยที่มีผลต่อการดำเนินการ ทั้งด้านดี และไม่ดี เพื่อหาศักยภาพ และความพร้อมของกรุงเทพมหานคร การทำ SWOT Analysis ภายในคือ จุดแข็ง (strength) จุดอ่อน (weakness) และภายนอกคือ โอกาส (opportunity) ภาวะคุกคาม (threat) ดังต่อไปนี้

จุดแข็ง/จุดเด่น (Strengths)	จุดอ่อน/จุดด้อย (Weakness)
- ซอฟต์แวร์ที่ใช้ในการบริหารจัดการข้อมูลเป็นแบบฟรี ไม่เสียค่าใช้จ่าย ประหยัดงบประมาณของกรุงเทพมหานคร - ความเร็วในการค้นหาข้อมูล เป็นแบบทันที (Real-time) - ฮาร์ดแวร์ ใช้ของที่มีในระบบ - บุคลากรของกรุงเทพมหานครด้านเทคโนโลยีสารสนเทศมีความพร้อม - ระบบโครงสร้างพื้นฐาน (Infrastructure) สามารถรองรับการทำงานได้	- กฎ ระเบียบบางอย่างของกรุงเทพมหานคร ไม่เอื้อต่อการดำเนินการ - เป็นเทคโนโลยีใหม่ ต้องใช้เวลาในการทำความเข้าใจ - บุคลากรของกรุงเทพมหานครบางส่วนยังขาดความรู้ความเข้าใจ
โอกาสในการดำเนินงาน (Opportunities)	อุปสรรคต่อการดำเนินงาน (Threats)
- เป็นทิศทางเทคโนโลยีสารสนเทศ (Trend Information Technology) ของโลก - สอดคล้องกับเทคโนโลยีเครือข่ายสังคมออนไลน์ ผ่านอุปกรณ์ Smart Medias ต่างๆ เช่น Smart Phone, Tablet ... - ความต้องการเข้าถึงข้อมูลข่าวสารภาคประชาชนสูงขึ้น	- ซอฟต์แวร์เป็นแบบ Open Source อาจขาดการสนับสนุนต่อเนื่อง - หน่วยงานภาครัฐยังไม่ให้ความสำคัญเท่าที่ควร

๖. การกิจที่ดำเนินการ

เนื่องจากขั้นตอนในการดำเนินการบริหารจัดการข้อมูลขนาดใหญ่ ของ กรุงเทพมหานคร โดยใช้ ซอฟต์แวร์ จัดการข้อมูลขนาดใหญ่ ประเภท Open Source : Hadoop Technology มีส่วนที่เกี่ยวข้อง ๒ ส่วน ดังนี้

ส่วนที่ ๑ ขั้นตอนการเตรียมการด้านเทคนิคสำหรับการบริหารจัดการข้อมูลขนาดใหญ่
ประกอบด้วยองค์ประกอบ ๓ องค์ประกอบดังต่อไปนี้

- ข้อมูล (Data)
- กระบวนการ (Process)
- เครื่องมือหรือเทคโนโลยี (Tools/Technologies)

ข้อมูล (Data)

ต้องสามารถบอกได้ว่าข้อมูลประเภทใดเป็นข้อมูลขนาดใหญ่ (Big Data) คุณลักษณะของข้อมูลขนาดใหญ่ (Big Data) ประกอบด้วย ๓ ลักษณะ

๑. ปริมาณ (Volume) หมายถึงปริมาณข้อมูลจำนวนมาก ซึ่งโครงสร้างฐานข้อมูลเดิมไม่สามารถจัดเก็บได้
๒. ความเร็ว (Velocity) หมายถึงเวลาในการประมวลผล หรือการเข้าถึงข้อมูลช้าจนไม่สามารถรองรับได้เนื่องจากอัตราการเพิ่มขึ้นของข้อมูลเข้าสู่ฐานข้อมูลเป็นไปอย่างรวดเร็ว เพิ่มปริมาณข้อมูลมากขึ้น
๓. ความหลากหลาย (Variety) หมายถึงรูปแบบของข้อมูลที่มีความหลากหลาย ทั้งที่มีโครงสร้าง ไม่มีโครงสร้าง และแบบกึ่งโครงสร้าง

ดังนั้นเมื่อแหล่งข้อมูลต่างๆขยายตัวและเพิ่มขึ้นอย่างต่อเนื่อง ผ่านสังคมเครือข่าย เช่น Facebook, Line, Twitter ล้วนเป็นสาเหตุให้เกิดข้อมูลที่มีขนาดและความซับซ้อนมาก อยู่ในรูปแบบที่หลากหลายและแตกต่างกัน การนำข้อมูลเหล่านั้นมาใช้ให้เกิดประโยชน์และสามารถตอบคำถามได้หลากหลายมุมมอง เราเรียกว่า ระบบการบริหารจัดการข้อมูลขนาดใหญ่ หรือ Big Data

กระบวนการ (Process)

เป็นการบริหารจัดการข้อมูลขนาดใหญ่ หรือ Big Data ซึ่งเป็นการทำงานในลักษณะการกระจายข้อมูล เนื่องจากข้อมูลมีขนาดใหญ่ ไม่สามารถเคลื่อนย้ายได้ จึงใช้วิธีการประมวลผลที่แหล่งข้อมูลและส่งผลลัพธ์กลับมาตามที่ต้องการ

โดยการจัดการข้อมูลมีองค์ประกอบดังต่อไปนี้

๑. การจัดเก็บข้อมูล (Storage)
๒. การประมวลผล (Processing)
๓. การวิเคราะห์ (Analysis)
๔. การรายงานสรุป (Visualization)

สรุปกระบวนการบริหารจัดการข้อมูล Big Data ในลักษณะการกระจายข้อมูล หมายถึง การเคลื่อนย้ายกลุ่มของข้อมูลเข้าสู่ระบบฐานข้อมูลเพื่อการประมวลผลและนำไปเก็บในส่วนจัดเก็บข้อมูล ซึ่งจะเก็บข้อมูลปริมาณมหาศาลที่ไม่สามารถเคลื่อนย้ายได้ โดยการประมวลผลจะเป็นการส่งคำร้องขอ (Request) เพื่อเข้ามาประมวลผลข้อมูลในจุดที่จัดเก็บข้อมูล และส่งผลลัพธ์ (Result) ให้อยู่ในมุมมองต่างๆตามที่ใช้ต้องการ ในปัจจุบันนี้ มีการพัฒนาซอฟต์แวร์ Hadoop เข้ามาช่วยในการบริหารจัดการข้อมูล Big Data โดยที่ซอฟต์แวร์ Hadoop เป็นซอฟต์แวร์แบบ Open Source ที่พัฒนาขึ้นมาเพื่อเป็นระบบวิเคราะห์และจัดการกับข้อมูลประเภท Big Data

ตัวอย่างการบริหารจัดการข้อมูลบนสังคมเครือข่าย ประเภท Facebook ของ ๕๐ สำนักงานเขต ซึ่งจัดเป็นข้อมูลประเภท Big Data เนื่องจากมีคุณลักษณะครบทั้ง ๓ อย่างคือมีการเพิ่มขึ้นของปริมาณอย่างรวดเร็ว (Volume) ข้อมูลเหล่านั้นมีความหลากหลายทางโครงสร้างของข้อมูล (Variety) ส่งผลกระทบต่อความเร็วในการเข้าถึงข้อมูล (Velocity) เมื่อผู้บริหารอยากทราบว่า ขณะนี้มีพื้นที่ใดบ้างที่อยู่ในเขตกรุงเทพมหานคร ที่ประชาชนประสบปัญหาไม่ได้รับความสะดวกบนทางเท้า ผู้บริหารหรือส่วนกลางจะส่งคำร้องขอ (Request) ในเรื่องปัญหานั้นผ่านทางระบบบริหารจัดการข้อมูล Big Data ระบบจะส่งคำร้องขอไปยังสำนักงานเขตทุกสำนักงานเขต เพื่อให้ประมวลผลจากฐานข้อมูลของแต่ละสำนักงานเขต โดยที่ฐานข้อมูลเหล่านั้นดึงข้อมูลจาก Facebook มาเก็บไว้ที่ฐานข้อมูลแบบทันทีทันใด (Real-time) ตามปกติ เมื่อประมวลผลเสร็จ แต่ละสำนักงานเขตจะส่งผลลัพธ์ (Result) ไปให้ส่วนกลางตามที่ร้องขอ ส่วนกลางจะรวบรวมผลลัพธ์ที่ได้ อาจจะเป็น ข้อความคำร้องเรียนเกี่ยวกับความไม่สะดวกบนทางเท้าของแต่ละเขต รูปภาพที่เกี่ยวข้อง วิดีโอ หรือคลิปเสียง มาเพื่อทำการวิเคราะห์ข้อมูล พร้อมรายงานสรุปให้กับผู้บริหารทั้งในรูปแบบทางสถิติ ตาราง หรือกราฟประกอบข้อมูล ต่างๆ

เครื่องมือหรือเทคโนโลยี (Tools/Technologies)

เทคโนโลยีที่ใช้ในการบริหารจัดการข้อมูลขนาดใหญ่ คือ Apache Hadoop เป็นซอฟต์แวร์เฟรมเวิร์ค (Framework) ถูกออกแบบมาเพื่อทำงานบนระบบคอมพิวเตอร์แบบกระจาย (Distributed Computing i.e. Cloud Computing) และสนับสนุนการทำงานแบบขนาน (Parallel) โดยมีชุดคำสั่ง (API) เพื่อช่วยอำนวยความสะดวกแก่นักพัฒนาแอปพลิเคชันที่จะสร้างระบบค้นหาหรือวิเคราะห์ข้อมูลขนาดใหญ่ (Large Dataset, Internet Scale Dataset)

Hadoop สามารถแบ่งออกเป็น ๒ ส่วนใหญ่คือ MapReduce Framework ซึ่งเป็นส่วนของการปฏิบัติงานในการวิเคราะห์ข้อมูล และ Hadoop File System (HDFS) ซึ่งมีความสามารถในการจัดเก็บข้อมูลขนาดใหญ่ ใน ปัจจุบัน Hadoop ได้รับการสนับสนุนจากองค์กรต่างๆ มากมายไม่ว่าจะเป็น Adobe, Alibaba, Amazon Web Services, Baidu, Microsoft Bing, China Mobile, eBay, Facebook, Hulu, IBM, Twitter, Rackspace และ Yahoo!

ข้อดีของ Hadoop

- ต้องการ
๑. มีความยืดหยุ่นสูงทำให้สามารถเพิ่มหรือลดจำนวนคอมพิวเตอร์เมื่อใดก็ได้ตามที่ต้องการ
 ๒. สามารถติดตั้งบนคอมพิวเตอร์ชนิดใดก็ได้ รวมทั้งสามารถทำการติดตั้งระบบได้อย่างรวดเร็ว
 ๓. สามารถเพิ่มหรือลดประสิทธิภาพได้ตามความต้องการของแต่ละงาน
 ๔. Reliable เชื่อถือได้และสามารถพยากรณ์ล่วงหน้าได้
 ๕. เฟรมเวิร์ค Hadoop เป็นระบบการจัดการข้อมูลขนาดใหญ่แบบกระจายระบบหนึ่งที่ Facebook เลือกใช้ ซึ่งมีระบบนิเวศต่าง ๆ ให้พร้อมสรรพ เช่น ระบบโครงสร้างไฟล์ HDFS, ฐานข้อมูล Hbase, ระบบวิเคราะห์และประมวลผลข้อมูล Hive, และโครงสร้างภาษาโปรแกรม Hadoop MapReduce เป็นต้น

ส่วนที่สอง ขั้นตอนในการดำเนินการบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานคร โดยใช้ซอฟต์แวร์ จัดการข้อมูลขนาดใหญ่ ประเภท Open Source : Hadoop Technology ประกอบด้วยขั้นตอนการดำเนินการดังนี้

- ศึกษารายละเอียด การใช้งานของ เทคโนโลยี Hadoop
- การจัดเตรียม ความพร้อม Hardware Software ระบบเครือข่ายที่ต้องใช้
- จัดการติดตั้งเทคโนโลยี Hadoop ทั้งส่วนกลาง และหน่วยงานที่เกี่ยวข้อง
- ทดสอบการทำงานของระบบ
- ฝึกอบรมสร้างทักษะการใช้งานให้กับบุคลากรของกรุงเทพมหานคร
- รายงานสรุปและประเมินผล ส่งให้ผู้บริหาร

๗. ผู้มีส่วนเกี่ยวข้อง

๑. ผู้บริหารกรุงเทพมหานคร กำหนดกลยุทธ์ นโยบาย แนวทางการดำเนินการ ไปสู่แนวทางการบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานคร ให้ไปในทิศทางเดียวกัน
๒. บุคลากรด้านเทคโนโลยีสารสนเทศของกรุงเทพมหานคร ทำหน้าที่ ขับเคลื่อนแนวทางให้เกิดผลในทางปฏิบัติ ทั้งในด้านการให้ความรู้ การสนับสนุน การช่วยเหลือ การอำนวยความสะดวกให้แนวทางการบริหารจัดการข้อมูลขนาดใหญ่ ของกรุงเทพมหานคร เป็นไปอย่างมีประสิทธิภาพและประสิทธิผล
๓. บุคลากรด้านอื่นของกรุงเทพมหานคร ที่ทำหน้าที่สนับสนุน ข้อมูลข่าวสารเข้าสู่ระบบ ในฐานะเจ้าของข้อมูล ต้องมีความรู้ความเข้าใจ เกี่ยวกับคุณลักษณะของข้อมูลขนาดใหญ่ (Big Data) กระบวนการบริหารจัดการ ตลอดจนเครื่องมือหรือเทคโนโลยีที่เกี่ยวข้อง เพื่อให้การทำงานเป็นไปอย่างมีประสิทธิภาพ

๘. ระยะเวลาการดำเนินการ

ระยะเวลาการดำเนินการแนวทางการบริหารจัดการข้อมูลขนาดใหญ่ ของ กรุงเทพมหานคร โดยใช้ ซอฟต์แวร์ จัดการข้อมูลขนาดใหญ่ ประเภท Open Source : Hadoop Technology จำนวน ๑๘๐ วัน

ขั้นตอนการดำเนินการ	ระยะเวลาดำเนินการ					
	เดือน ๑	เดือน ๒	เดือน ๓	เดือน ๔	เดือน ๕	เดือน ๖
๑. ศึกษารายละเอียด การใช้งาน ของ เทคโนโลยี Hadoop	↔					
๒. การ จัดเตรียม ความพร้อม Hardware Software ระบบ เครือข่ายที่ต้องใช้		↔				
๓. จัดการ ติดตั้งเทคโนโลยี Hadoop ทั้ง ส่วนกลาง และหน่วยงานที่เกี่ยวข้อง			↔			
๔. ทดสอบ การทำงานของระบบ				↔		
๕. ฝึกอบรม สร้างทักษะการใช้งานให้กับ บุคลากรของกรุงเทพมหานคร					↔	
๖. รายงาน สรุปและประเมินผล ส่งให้ ผู้บริหาร						↔

๙. แนวทางการประเมินผล

๙.๑ ตัวชี้วัดความสำเร็จแยกเป็น

- ระดับผลผลิต (Output)

มีการนำเทคโนโลยี Hadoop มาประยุกต์ใช้เพื่อการบริหารจัดการข้อมูลขนาดใหญ่ (Big Data) ของกรุงเทพมหานคร ภายในระยะเวลาที่กำหนด

- ระดับผลลัพธ์ (Outcome)

สร้างความพึงพอใจให้เกิดขึ้นกับผู้ให้บริการ ทั้งในระดับผู้บริหารของ กรุงเทพมหานคร บุคลากรของกรุงเทพมหานคร ตลอดจนประชาชน โดยสามารถค้นหาข้อมูลได้อย่าง

ทันทีทันใด (Real-time) จากแหล่งข้อมูลโดยตรงที่ครอบคลุมทุกส่วนที่เกี่ยวข้อง ผลลัพธ์ที่ได้มีหลากหลายรูปแบบ สามารถกำหนดแนวทางในการแก้ไขปัญหาได้อย่างครอบคลุม ครบถ้วน รวมถึงติดตามประเมินผล จากความพึงพอใจของประชาชนผ่านระบบเครือข่าย ประเภทต่างๆ

๙.๒ วิธีการในการประเมินผล

วิธีการในการประเมินผลการนำเทคโนโลยี Hadoop มาประยุกต์ใช้เพื่อให้การบริหารจัดการข้อมูลขนาดใหญ่ (Big Data) ของกรุงเทพมหานคร โดยการสำรวจความคิดเห็นของผู้ใช้บริการ ซึ่งแบ่งระดับผู้ให้บริการเป็น ๒ ส่วน

ส่วนที่หนึ่ง คือ ผู้บริหารของกรุงเทพมหานครที่ใช้ระบบการบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานคร

ส่วนที่สอง คือ บุคลากรด้านเทคโนโลยีสารสนเทศของกรุงเทพมหานครที่ใช้ระบบการบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานคร

๙.๓ เครื่องมือที่ใช้ในการประเมินผล

เครื่องมือที่ใช้ในการประเมินผล ที่สอดคล้องกับวิธีการประเมินผลคือ แบบสำรวจความพึงพอใจต่อผู้ให้บริการระบบการบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานคร แบบสำรวจความพึงพอใจ ๒ ส่วนคือ

ส่วนที่หนึ่ง คือ แบบสำรวจความพึงพอใจไม่น้อยกว่าร้อยละ ๙๐ สำหรับผู้บริหารของกรุงเทพมหานครที่ใช้ระบบการบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานคร

ส่วนที่สอง คือ แบบสำรวจความพึงพอใจไม่น้อยกว่าร้อยละ ๙๐ สำหรับบุคลากรด้านเทคโนโลยีสารสนเทศของกรุงเทพมหานครที่ใช้ระบบการบริหารจัดการข้อมูลขนาดใหญ่ของกรุงเทพมหานคร

๑๐. ข้อเสนอแนะ

สำหรับกรุงเทพมหานครควรเตรียมความพร้อม สำหรับความท้าทายในการนำเทคโนโลยีที่ใช้จัดการข้อมูลขนาดใหญ่ (Big Data) หรือเครื่องมือวิเคราะห์ต่างๆมาใช้ในการบริหารจัดการข้อมูลขนาดใหญ่ (Big Data) ไม่ว่าจะเป็นการแปลงข้อมูลปริมาณมากที่ไม่มีโครงสร้างไปสู่ข้อมูลที่มีโครงสร้างด้วยการทำงานของเทคโนโลยี Hadoop จากนั้นจึงนำข้อมูลไปวิเคราะห์ด้วยเครื่องมือหรือโปรแกรมวิเคราะห์ทางสถิติเพื่อให้ได้สารสนเทศที่ช่วยสนับสนุนการตัดสินใจสำหรับผู้บริหารในการวางแผนการทำงาน การทำนายพฤติกรรมประชาชนผู้ให้บริการ รวมทั้งการจัดทำแผนยุทธศาสตร์ที่สอดคล้องกับความต้องการที่แท้จริงของประชาชน ที่ได้จากงานประจำวันหรือจากพฤติกรรมการบริโภคของประชาชนที่ใช้บริการผ่านสื่อสังคมออนไลน์ต่างๆ

เพื่อสร้างมูลค่าและนำไปสู่โอกาสในการแข่งขันของกรุงเทพมหานคร จากการนำผลการวิเคราะห์ข้อมูลมหาศาลที่เกิดขึ้นมาใช้อย่างคุ้มค่าร่วมกับเทคโนโลยีอื่น ด้วยปริมาณข้อมูลที่เกิดขึ้นอย่างมหาศาลทั้งจากข้อมูลงานประจำวันและข้อมูลจากเครือข่ายสังคมผ่านอุปกรณ์ต่างๆทำให้หลาย

หน่วยงานกำลังเผชิญกับความท้าทายในการจัดการข้อมูลขนาดใหญ่ จำนวนมากมายมหาศาลที่มีแนวโน้มว่าจะมีมากขึ้นอย่างต่อเนื่อง และเพื่อให้เกิดประโยชน์ต่อการนำข้อมูลขนาดใหญ่มาใช้วิเคราะห์การดำเนินงานของกรุงเทพมหานคร จึงเป็นแรงผลักดันที่ทำให้กรุงเทพมหานครต้องหาแนวทางในการพัฒนาการบริการโดยการนำเทคโนโลยีและนวัตกรรมใหม่ๆมาใช้วิเคราะห์ข้อมูลขนาดใหญ่เพื่อหาผลจากการวิเคราะห์ที่ได้มาปรับปรุงและวางแผนการทำงานของกรุงเทพมหานคร ตลอดจนเพื่อลดความเสี่ยงในการทำงานและลดค่าใช้จ่ายที่สูงมากในการจัดซื้อและการปรับปรุงฮาร์ดแวร์ที่จำเป็นต่อการรองรับข้อมูลขนาดใหญ่จำนวนมาก การจัดการและการวิเคราะห์ข้อมูลขนาดใหญ่หรือ Big Data จะทำให้เกิดโอกาสในการดำเนินงานในภาครัฐที่สอดคล้องกับความเปลี่ยนแปลงของข้อมูลที่มีอย่างทันทีและตลอดเวลาได้อย่างมีประสิทธิภาพ