



## บันทึกข้อความ

กองนโยบายและแผนการใช้ที่ดิน  
เลขที่รับ..... C. ๗๒๓  
วันที่..... ๙ ก.พ. ๖๖  
เวลา..... ๑๕.๑๓ น.

ส่วนราชการ กลุ่มนโยบายและวางแผนการใช้ที่ดิน กองนโยบายและแผนการใช้ที่ดิน โทร ๒๑๗๗  
ที่ กษ ๐๘๓๗.๐๔/๑๐๖ วันที่ ๙ กุมภาพันธ์ ๒๕๖๖  
เรื่อง สรุปเนื้อหาจากการฝึกอบรม  
เรียน ผอ.กลุ่มนโยบายและวางแผนการใช้ที่ดิน

ตามที่กลุ่มนโยบายและวางแผนการใช้ที่ดิน กองนโยบายและแผนการใช้ที่ดิน ได้กำหนดให้มีการพัฒนาทักษะด้านดิจิทัล ๑ เรื่อง และการพัฒนาความรู้ ๑ เรื่อง และทำสรุปบทเรียน ซึ่งเป็นตัวชี้วัดตามแบบประเมินผลสัมฤทธิ์ ปี ๒๕๖๖ ครั้งที่ ๑ นั้น

กระผมได้เข้าอบรมด้วยระบบการเรียนออนไลน์ ( e-learning) หลักสูตร ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ ของสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) และเข้าการฝึกอบรมหลักสูตร การใช้งานระบบให้บริการคลังข้อมูลดินและที่ดิน ของกรมพัฒนาที่ดินเสร็จสิ้นแล้ว ซึ่งสามารถสรุปเนื้อหา หลักสูตรความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ ตามเอกสารที่แนบมาพร้อมนี้ จำนวน ๑ ฉบับ

จึงเรียนมาเพื่อโปรดพิจารณา

เรียน ผอ.กนผ.

เพื่อโปรดพิจารณา

(นางสาวกรรณิศา สฤกษ์ศิริ)

นักวิเคราะห์นโยบายและแผนชำนาญการพิเศษ  
ผู้อำนวยการกลุ่มนโยบายและวางแผนการใช้ที่ดิน

๙ ก.พ. ๖๖

อดิศร ใจชื่น

(นายอดิศร ใจชื่น)

นักสำรวจชำนาญการพิเศษ

ลงนามแล้ว

(นางสาวพิมพ์ฉัตร นวลตะออง)

นักวิเคราะห์นโยบายและแผนชำนาญการพิเศษ  
ผู้อำนวยการกลุ่มวางแผนการจัดการที่ดินในพื้นที่เสี่ยงภัยทางทะเล  
รักษาการแทน ผู้อำนวยการกองนโยบายและแผนการใช้ที่ดิน

- ๙ ก.พ. ๒๕๖๖

รายงานสรุปการอบรม/สัมมนา/พัฒนาความรู้/ประชุมเชิงปฏิบัติการ/และเป็นวิทยากร  
กองนโยบายและแผนการใช้ที่ดิน กรมพัฒนาที่ดิน

\*\*\*\*\*

ส่วนที่ ๑ ข้อมูลทั่วไป

ชื่อ นายอดิสร นามสกุล ใจชื่น

ตำแหน่ง นักสำรวจดินชำนาญการพิเศษ กลุ่ม นโยบายและวางแผนการใช้ที่ดิน

หลักสูตร/หัวข้อเรื่องอบรม/สัมมนา

การอบรมด้วยระบบการเรียนรู้ออนไลน์ (e-learning) หลักสูตร ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับ  
ข้าราชการและบุคลากรภาครัฐทุกระดับ

สถานที่อบรม/สัมมนา/พัฒนาความรู้

กรมพัฒนาที่ดิน

หน่วยงานที่จัดฝึกอบรม/สัมมนา/พัฒนาความรู้

สถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

ตั้งแต่วันที่ ๑ เดือน กุมภาพันธ์ พ.ศ. ๒๕๖๖ ถึงวันที่ ๑ เดือน กุมภาพันธ์ พ.ศ. ๒๕๖๖

เพื่อ  อบรม  สัมมนา  อื่นๆ ระบุ.....

ส่วนที่ ๒ สิ่งที่ได้รับจากการอบรม/สัมมนา/พัฒนาความรู้

๒.๑ รายงานสรุปเนื้อหาสาระสำคัญในการอบรม/สัมมนา/พัฒนาความรู้

ข้อมูลขนาดใหญ่ (Big Data)

Big Data คือ ปริมาณข้อมูลจำนวนมากมหาศาลที่มีอยู่ในองค์กรทุกรูปแบบ ไม่ว่าจะแหล่งที่มาจะมาจากภายใน  
หรือภายนอกองค์กรก็ตาม ทั้งนี้แบ่งออกเป็นข้อมูลที่มีโครงสร้างชัดเจน (Structured Data) และข้อมูลที่มี  
โครงสร้างไม่ชัดเจน (Unstructured Data) จากข้อมูลที่มีจำนวนมากการนำข้อมูลมาดังกล่าวมาใช้ประโยชน์  
ต้องอาศัยเทคโนโลยีในการบริหารจัดการจัดเก็บและการประมวลผลข้อมูล เพื่อให้ได้ผลลัพธ์ถูกต้อง รวดเร็ว น่าเชื่อถือ  
เพื่อใช้ในการวิเคราะห์ข้อมูลใหญ่ในเชิงลึก เพื่อทำนายหรือคาดการณ์อนาคตได้

รูปแบบของข้อมูล Big Data ได้แก่

๑. ข้อมูลเชิงพฤติกรรมการใช้งานต่างๆ (Behavioral Data) เช่น พฤติกรรมการคลิกดูข้อมูล ข้อมูลการใช้  
ATM เป็นต้น

๒. ข้อมูลภาพและเสียง (Image & sounds Data) ได้แก่ ภาพถ่าย วิดีโอ ลายมือ ข้อมูลเสียงที่ถูกบันทึกไว้  
 เป็นต้น

๓. ข้อมูลทางภาษา (Languages Data) เช่น ข้อความต่างๆ เนื้อหาต่างๆในเว็บไซต์ เป็นต้น

๔. ข้อมูลจากการบันทึก (Records Data) เช่น ข้อมูลทางการแพทย์ ข้อมูลผลสำรวจที่มีขนาดใหญ่ ข้อมูล  
ทางภาษี เป็นต้น

๕. ข้อมูลจากอุปกรณ์ตรวจจับสัญญาณต่างๆ (Sensors Data) เช่น ข้อมูลอุณหภูมิ ข้อมูลจากการเคลื่อนที่  
ของวัตถุ (accelerometer) เป็นต้น

## คุณลักษณะของข้อมูล Big Data

ประกอบด้วยคุณลักษณะ ๔ ประการ ได้แก่

๑. ปริมาณ (Volume) หมายถึง เป็นข้อมูลที่มีขนาดใหญ่หรือเป็นข้อมูลที่มีปริมาณมากมายมหาศาล ไม่ว่าจะอยู่ในรูปแบบออนไลน์หรือออฟไลน์ โดยข้อมูลเหล่านี้จะมีปริมาณมากกว่าหน่วย TB (Terabyte) ขึ้นไป

๒. ความหลากหลาย (Variety) หมายถึง ข้อมูลที่มีรูปแบบหลากหลาย ได้แก่ รูปภาพ วิดีโอ เพลง ข้อมูลระบบฐานข้อมูล หรือข้อมูลจากสื่อสังคมออนไลน์

๓. ความเร็ว (Velocity) หมายถึง ข้อมูลที่เข้าสู่ระบบจะมีอัตราการเพิ่มขึ้นอย่างรวดเร็วและต่อเนื่อง เช่น ข้อมูลที่เกิดจากโทรศัพท์เคลื่อนที่ที่ผู้ใช้เก็บเป็นข้อมูลประเภทภาพถ่าย ข้อมูลการพิมพ์การสนทนา

๔. ความจริง (Veracity) หมายถึง ข้อมูลที่มีเป็นความจริง แม่นยำ ถึงข้อมูลจะมาจากแหล่งข้อมูลที่ต่างกัน หรือคนละประเภท ต้องมีการจัดระเบียบและวิเคราะห์ข้อมูลใดมีความถูกต้องแม่นยำมากที่สุด

## การจัดเก็บข้อมูลแบบ Data Lake

หมายถึง พื้นที่เก็บข้อมูลส่วนกลางจากทุกแหล่งข้อมูล ทุกรูปแบบ ข้อมูลสามารถอยู่ทั้งในรูปแบบของที่มีโครงสร้างชัดเจน (Structured Data) เป็นการจัดเก็บในระบบฐานข้อมูล และโครงสร้างไม่ชัดเจน (Unstructured Data) เช่น รูปภาพ วิดีโอ เป็นต้น ข้อมูลที่จะใช้หรือไม่คิดว่าจะได้ใช้ก็เก็บได้หมด ไม่ต้องมีการแปลงข้อมูลใดๆ ก่อนเก็บ สามารถใช้เครื่องมือหลายๆ อย่าง เพื่อให้รวดเร็วในการดึงข้อมูลมาใช้ตัดสินใจในองค์กร

## การวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data Analytics)

การวิเคราะห์ข้อมูล Big data ต้องอาศัยหลักการพื้นฐานบางอย่างเพื่อพัฒนาเป็นเทคนิคในการดึงข้อมูลสำคัญออกจากชุดฐานข้อมูลขนาดใหญ่ที่เต็มไปด้วยข้อมูลหลากหลายรูปแบบ เพื่อนำมาหารูปแบบของข้อมูลที่มองไม่เห็นด้วยตาเปล่า หารูปแบบความสัมพันธ์ที่ซ่อนอยู่ ผลจากการวิเคราะห์ข้อมูล Big data ทำให้มีข้อมูลที่ เป็นข้อเท็จจริง ซึ่งผ่านการวิเคราะห์อย่างเป็นระบบเพื่อใช้ประกอบการตัดสินใจ โดยระดับของการวิเคราะห์ ก็เป็นได้หลากหลาย แล้วแต่รูปแบบการนำไปใช้งาน แบ่งออกเป็น ๓ ระดับ ได้แก่

๑. การวิเคราะห์ข้อมูลแบบพื้นฐาน (Descriptive analytics) เป็นการวิเคราะห์ผลลัพธ์จากการกระทำต่างๆ ที่บอกว่าเกิดอะไรขึ้น จำนวนเท่าไร เกิดเหตุการณ์สำคัญตอนไหน เช่น ปัจจัยที่มีผลต่อพฤติกรรมการตัดสินใจของลูกค้า การรายงานการขาย หรือรายงานผลการดำเนินการ เป็นต้น

๒. การวิเคราะห์แบบพยากรณ์ (Predictive analytics) เป็นการวิเคราะห์โดยนำข้อมูลที่เคยเกิดขึ้นแล้วในอดีตมาพยากรณ์ ทำนายในสิ่งที่กำลังจะเกิดขึ้นในอนาคตด้วยแบบจำลองทางสถิติ หรือการใช้ AI เทคโนโลยี ปัญญาประดิษฐ์มาช่วยในการวิเคราะห์ข้อมูลใน Big data ที่มีจำนวนมหาศาล

๓. การวิเคราะห์แบบให้คำแนะนำ (Prescriptive analytics) การวิเคราะห์ในลักษณะนี้ค่อนข้างมีความซับซ้อนมากที่สุด เพราะเป็นการพยากรณ์สิ่งที่กำลังจะเกิดขึ้นเพื่อหาสาเหตุ ข้อดี ข้อเสีย และมีเรื่องของระยะเวลาเข้ามาเกี่ยวข้อง เพื่อให้คำแนะนำหรือเลือกผลลัพธ์ที่ดีที่สุด อีกทั้งยังวิเคราะห์ได้ว่าการเลือกแต่ละแนวทาง จะให้ผลลัพธ์เป็นไปในทิศทางใดบ้าง

## รูปแบบของการวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data Method) แบ่งเป็น ๓ รูปแบบ ได้แก่

๑. Data Mining เป็นการวิเคราะห์ข้อมูลจากข้อมูลจำนวนมาก (Big Data) เพื่อหาความสัมพันธ์ของข้อมูลที่ซ่อนอยู่ โดยทำการจำแนกประเภท รูปแบบ เชื่อมโยงข้อมูลที่มีความสัมพันธ์กัน และหาความน่าจะเป็นที่จะเกิดขึ้น เพื่อให้ได้องค์ความรู้ใหม่ที่สามารถนำไปใช้ประกอบการตัดสินใจในด้านต่างๆ เช่น ตลาดหลักทรัพย์ ทางธุรกิจ ทางด้านการแพทย์ ยุทธศาสตร์ทหาร เป็นต้น

๒. Text Mining เป็นเทคนิคเพื่อค้นหารูปแบบจากข้อความจำนวนมากโดยอัตโนมัติ ซึ่งเป็นกระบวนการที่กระทำกับข้อความ เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อความนั้น โดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง หลักคณิตศาสตร์ หลักการประมวลเอกสาร (Document

Processing) ข้อความ (Text Processing) และภาษาธรรมชาติ (Natural Languages Processing)

๓. Machine Learning เป็นศาสตร์ของการสร้างโมเดลคณิตศาสตร์ มุ่งเน้นที่จะสร้างองค์ความรู้จากข้อมูล โดยเริ่มจากการสร้างโมเดลคณิตศาสตร์ที่มีความยืดหยุ่นและสามารถปรับตัวเองเข้ากับข้อมูลที่ได้รับ เพื่อที่จะสามารถทำนายอนาคตได้ เช่น เทคโนโลยีที่อยู่เบื้องหลังหุ่นยนต์อัจฉริยะต่างๆ หรือเทคโนโลยีการบินแบบไร้คนขับหรือระบบวิเคราะห์ภาพ เสียง ภาพเคลื่อนไหวและภาษามนุษย์

### การวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data Analytics) กับการบริหารภาครัฐ

องค์กรภาครัฐในยุคดิจิทัล มีความจำเป็นที่จะต้องสร้างมูลค่าจาก Big Data ทั้งข้อมูลภายในและภายนอกหน่วยงาน เพื่อให้สามารถใช้และวิเคราะห์ข้อมูลได้รวดเร็ว ทันต่อสถานการณ์ และสร้างนวัตกรรมใหม่ๆ ขององค์กร โดยมีแนวทางในการดำเนินงานดังนี้

๑. รับฟังความคิดเห็น รวบรวมข้อมูล และปรึกษากับผู้มีส่วนได้ส่วนเสียเกี่ยวกับศักยภาพ การเพิ่มประสิทธิภาพและการกำหนดนิยามทางเทคนิคที่เกี่ยวข้อง รวมไปถึงเป้าหมายที่จะทำให้เกิดประโยชน์ในการนำเอาข้อมูลมาร่วมใช้อย่างมีประสิทธิภาพในรูปแบบ Open Source ให้มากขึ้น

๒. วางแผนการลงทุนในการจัดโครงสร้าง รวบรวมข้อมูลและการบริหารจัดการข้อมูลโดยอยู่บนพื้นฐานของความเชื่อมั่นต่อผู้มีส่วนได้ส่วนเสีย ซึ่งการจัดการข้อมูลต่างๆ จะต้องเป็นไปตามกฎหมายและหลักจริยธรรมที่เกี่ยวข้อง รวมทั้งต้องสื่อสารให้ทุกภาคส่วนรวมทั้งประชาชนให้มีความเข้าใจและมั่นใจในเรื่องความเป็นส่วนตัวของข้อมูล (Privacy)

๓. มีความเข้าใจ มีทักษะทางธุรกิจและทางเทคนิค โดยจะต้องทำให้ Big Data สามารถเชื่อมโยงประโยชน์ให้เกิดขึ้นกับผู้มีส่วนได้ส่วนเสียทุกกลุ่มโดยเฉพาะอย่างยิ่งประชาชน ด้วยการสร้างพันธมิตรในรูปแบบ win-win เพื่อที่จะเกิดความสำเร็จได้อย่างรวดเร็ว

๔. เตรียมพร้อมภายใต้การเปลี่ยนแปลงของเทคโนโลยี เพื่อตอบสนองความต้องการของผู้มีส่วนได้ส่วนเสียที่มีกฎระเบียบที่หลากหลาย จึงทำให้การออกแบบระบบ Big Data Analytics จะต้องใช้สถาปัตยกรรมด้านไอทีแบบผสมผสานที่สามารถปรับขนาดได้ (Scalability) และมีความยืดหยุ่น (Flexibility)

๕. เจ้าหน้าที่ของรัฐจะต้องปรับ Mindset ในการเข้าร่วมกับทุกภาคส่วนในลักษณะพันธมิตร (Partnership) ให้ได้ จะต้องปรับตัวให้องค์กรภาครัฐเป็นส่วนหนึ่งของระบบนิเวศอย่างกลมกลืน และจะต้องไม่เป็นผู้สร้างอุปสรรคให้เกิดขึ้นในการสร้างระบบ Big Data Analytics

๖. ภาครัฐจะต้องปรับปรุงวิธีคิดและกระบวนการ เพื่อทำให้เกิดการแลกเปลี่ยนข้อมูลและการใช้ข้อมูลร่วมกันระหว่างหน่วยงานภาครัฐกันเองอย่างมีประสิทธิภาพ

๗. กำหนดแนวทางและการบริการให้คำปรึกษาในด้าน Big Data Analytics ให้แก่ทุกภาคส่วนให้มากขึ้นกว่าที่เป็นมาในอดีต

### ตัวอย่างโปรแกรมสำหรับการวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data)

อาปาเชฮาดูป (Apache Hadoop) เป็นซอฟต์แวร์ตัวหนึ่งที่ยิมนำมาใช้ในการทำ Big Data เป็นซอฟต์แวร์เฟรมเวิร์ค (Framework) ที่ถูกออกแบบมาเพื่อรองรับการทำงานบนระบบคอมพิวเตอร์แบบกระจาย (Distributed Computing) และสนับสนุนการประมวลผลแบบขนาน (Parallel Processing) ที่มีความเสถียรสูง และสามารถเพิ่มขยายจำนวนเครื่องในระบบได้อย่างมหาศาล จึงเหมาะกับการประมวลผลข้อมูลใหญ่ (Big Data) โดยฮาดูปเป็นซอฟต์แวร์แบบ Open Source ของมูลนิธิอาปาเชซอฟต์แวร์ (Apache Software Foundation) ที่เปิดโอกาสให้บุคคลอื่นนำเอาซอฟต์แวร์นั้นไปพัฒนาต่อได้ ซึ่งฮาดูปประกอบด้วยกลุ่มของชุดคำสั่งต่างๆ (Libraries) เพื่อช่วยอำนวยความสะดวกแก่นักพัฒนาแอปพลิเคชัน ที่จะสร้างระบบหรือวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data Analytics) ได้อย่างมีประสิทธิภาพ

## หลักการทํางานของฮาดูป

หลักการทํางานจะประกอบด้วย ๔ ส่วนหลัก ได้แก่

๑. ฮาดูปทั่วไป (Hadoop Common) เป็นกลุ่มข้อมูลของคลาส (Class) หรือไลบรารี (Libraries) จำนวนมาก เพื่อรองรับการทํางานของฮาดูป เช่น การกำหนดค่าหรือการปรับเปลี่ยนค่าข้อมูล

๒. ระบบแฟ้มข้อมูลแบบกระจายฮาดูป (Hadoop Distributed File System: HDFS) เป็นการนำข้อมูลเข้า (Input Data) ที่มีขนาดใหญ่จำนวนมาก มาทำการแบ่งข้อมูลขนาดใหญ่ให้มีขนาดเล็กลง (Data Splitting) เพื่อกระจายข้อมูลขนาดเล็กๆ เหล่านี้ ไปประมวลผลในระบบเครื่องคอมพิวเตอร์คลัสเตอร์ โดยกระบวนการทํางานของระบบแฟ้มข้อมูลแบบกระจายฮาดูป จะประกอบด้วย ๒ ส่วนหลัก ได้แก่ Master Node ซึ่งจะมีเพียงโหนดเดียว และ Slave Nodes ซึ่งจะมีได้หลายโหนด

๓. ฮาดูปยาน (Hadoop YARN) โดยยาน (YARN) มาจากคำว่า "Yet Another Resource Negotiator" ที่ทำหน้าที่เป็นผู้บริหารทรัพยากร (Resource Management) ของแมปรีดิวซ์ กำหนดและควบคุมการประมวลผล สนับสนุนการทํางานกับระบบที่เกี่ยวข้องกับฮาดูป และรองรับการประมวลผลข้อมูลแบบทันทีทันใดหรือเวลาจริง (Realtime) อีกด้วย

๔. แมปรีดิวซ์ (Map Reduce) เป็นโปรแกรมการทํางานที่ทํางานอยู่บนฮาดูปที่นำข้อมูลที่ได้แบ่งให้เป็นข้อมูลเล็ก (Data Splitting) เข้าสู่ขั้นตอนการทํางานของแมปรีดิวซ์ โดยแมปรีดิวซ์จะประกอบด้วย ๒ ส่วน ได้แก่ ขั้นตอนแมปเป็นการกำหนดคู่คีย์ (Key Pair) ของข้อมูล เพื่อทำการกระจายข้อมูลไปประมวลผลยังโหนดต่างๆ ในระบบเครื่องคอมพิวเตอร์คลัสเตอร์ตามคำสั่งการทํางาน และขั้นตอนรีดิวซ์เป็นการนำผลที่ได้จากการประมวลผลแบบกระจายของขั้นตอนแมปในแต่ละโหนด กลับมาทำการจัดเรียงลำดับและสรุปผลข้อมูลเพื่อแสดงผลลัพธ์ การทํางานที่รวดเร็วในการประมวลผลข้อมูลขนาดใหญ่

### ๒.๒ ประสพการณ์/ประโยชน์ที่ได้รับ /การประยุกต์ใช้กับหน่วยงาน

ต่อตนเอง

ทำให้มีความรู้ความเข้าใจเกี่ยวกับข้อมูลขนาดใหญ่ (Big Data) เพิ่มขึ้น

ต่อหน่วยงาน / การนำมาประยุกต์ใช้กับหน่วยงาน

ศึกษาของข้อมูลเพื่อการจำแนกประเภทข้อมูล หารูปแบบที่เชื่อมโยงกันของข้อมูลที่มีความสัมพันธ์กัน เพื่อใช้ในการงานด้านการวางแผนการใช้ที่ดิน

### ๒.๓ ปัญหาและอุปสรรคในการอบรม/สัมมนา/พัฒนาความรู้

.....

### ๒.๔ ข้อคิดเห็นและข้อเสนอแนะ

.....

ลงชื่อ



( นายอดิศร ใจชื่น )

ตำแหน่ง นักสำรวจดินชำนาญการพิเศษ

ผู้รายงาน

วันที่.....๙.....เดือน...กุมภาพันธ์.....พ.ศ....๒๕๖๖....

ส่วนที่ ๓ ความเห็นของผู้บังคับบัญชา

(✓) ทราบ

.....  
.....  
.....

ลงชื่อ.....



(นางสาวพิมพ์ลีย์ นวลระยอง)

นักวิเคราะห์นโยบายและแผนชำนาญการพิเศษ

ผู้อำนวยการกลุ่มวางแผนการจัดการที่ดินในพื้นที่เสี่ยงภัยทางการเกษตร

ตำแหน่ง.....

วันที่.....เดือน.....พ.ศ.....

รักษาราชการแทน ผู้อำนวยการกองนโยบายและแผนการใช้ที่ดิน

๕ ก.พ. ๖๖