



# บันทึกข้อความ

ส่วนราชการ... สถานีพัฒนาที่ดินพังงา ตำบลนาเตย อำเภอท้ายเหมือง จังหวัดพังงา โทร. ๐-๗๖๔๖-๑๕๓๒

ที่ กษ ๐๘๑๘.๐๙/..... วันที่ ๓ สิงหาคม ๒๕๖๖

เรื่อง... สรุปบทเรียนที่ผ่านการพัฒนาทักษะด้านดิจิทัล

เรียน ผู้อำนวยการสถานีพัฒนาที่ดินพังงา

ตามที่ กรมฯ ได้กำหนดให้ข้าราชการทั่วไป ดำเนินการจัดทำตัวชี้วัดรายบุคคลด้านการพัฒนาบุคลากร “ระดับความสำเร็จของการพัฒนาบุคลากรในหน่วยงาน” รอบการประเมินที่ ๒ (๑ มีนาคม – ๓๐ กันยายน ๒๕๖๖) ของปีงบประมาณ พ.ศ.๒๕๖๖ โดยให้มีการพัฒนาทักษะด้านดิจิทัล ๑ เรื่อง ตามเงื่อนไขหลักสูตรและพัฒนาความรู้ ๑ เรื่อง รวมทั้งสรุปบทเรียน ๑ เรื่อง ส่งให้ผู้บังคับบัญชาทราบ นั้น

บัดนี้ ข้าพเจ้าได้เข้ารับการฝึกอบรมหลักสูตรด้านดิจิทัล ของสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล Thailand Digital Government Academy หรือ TDGA ซึ่งเป็นหลักสูตรพื้นฐานสำหรับบุคลากรที่ต้องการความรู้ด้านดิจิทัลในการปฏิบัติงาน ผ่านระบบออนไลน์ (TDGA E-Learning) และได้สรุปบทเรียนดังกล่าวเรียบร้อยแล้ว จำนวน ๑ หลักสูตร ดังนี้

## หลักสูตร ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ

BIG DATA คือ ข้อมูลขนาดใหญ่ มีทั้งแบบโครงสร้างปกติและโครงสร้างข้อมูลที่ไม่มีความสัมพันธ์ ซึ่งทั้งหมดเป็นข้อมูลที่ใช้ในเชิงธุรกิจ มักจะถูกใช้กับงานพวกที่ต้องวิเคราะห์ ข้อมูลที่มีความซับซ้อน และไม่สามารถประเมินขนาดข้อมูลได้

### ๑. รูปแบบของข้อมูลของ Big Data แบ่งเป็น ๕ รูปแบบ ได้แก่

- ๑.๑. Behavioral Data ได้แก่ ข้อมูลเชิงพฤติกรรมการใช้งานต่างๆ เช่น Server Log พฤติกรรมการคลิกดูข้อมูล หรือ ข้อมูลการใช้ ATM
- ๑.๒. Image & Sound ตัวอย่าง เช่น ภาพถ่าย วีดีโอ รูปจาก Google Street View การถ่ายภาพทางการแพทย์ ลายมือ ข้อมูลเสียที่ถูกบันทึกไว้
- ๑.๓. Languages ตัวอย่างเช่น Text Message ข้อความที่ถูก Tweet เนื้อหาต่างๆ ในเว็บไซต์
- ๑.๔. Records ตัวอย่างเช่น ข้อมูลทางการแพทย์ ข้อมูลผลสำรวจที่มีขนาดใหญ่ ข้อมูลทางภาษี
- ๑.๕. Sensors ตัวอย่างเช่น ข้อมูลอุณหภูมิ, Accelerometer, ข้อมูลทางภูมิศาสตร์

### ๒. Big Data ประกอบด้วยคุณลักษณะ ๔ ประการ คือ

- ๒.๑. Volume คือ ข้อมูลที่มีขนาดใหญ่ มีปริมาณข้อมูลมากซึ่งสามารถเป็นได้ทั้งข้อมูลแบบ offline หรือ Online
- ๒.๒. Variety คือ ข้อมูลมีความหลากหลาย สามารถเป็นได้ทั้งที่มีโครงสร้างและข้อมูลที่ไม่สามารถจับ Pattern ได้
- ๒.๓. Velocity คือ ข้อมูลมีการเปลี่ยนแปลงตลอดเวลา อย่างรวดเร็ว มีการส่งผ่านข้อมูลอย่างต่อเนื่อง ในลักษณะ Streaming ทำให้การวิเคราะห์ข้อมูลแบบ Manual มีข้อจำกัด
- ๒.๔. Veracity คือ ข้อมูลมีความไม่ชัดเจน (Untrusted, Uncleaned)

### ๓. Dato Lake

คือ ที่เก็บส่วนกลางซึ่งช่วยให้คุณจัดเก็บข้อมูลที่มีและไม่มีโครงสร้างในทุกขนาดได้ คุณสามารถจัดเก็บข้อมูลตามที่เป็นโดยไม่ต้องวางโครงสร้าง และยังสามารถใช้การวิเคราะห์ประเภทต่างๆ ได้ ตั้งแต่ แดชบอร์ดและการแสดงภาพไปจนถึงการประมวลผล Big Data การวิเคราะห์แบบเรียลไทม์ และ Machine Learning เพื่อสร้างแนวทางการตัดสินใจที่ดีขึ้น เป็นการนำเอาข้อมูลจากแหล่งข้อมูลภายนอกองค์กรข้อมูลจาก **เครือข่ายข้อมูล**จากเครือข่ายข้อมูลที่กระจายไปทั่วโลกใช้มากขึ้น เป็นข้อมูลจากแหล่งภายนอกเพิ่มอย่างต่อเนื่อง และมีแนวโน้มที่เติบโตแบบก้าวกระโดดมากขึ้น สามารถแก้ไขข้อจำกัดหลายอย่าง ของ Data Warehouse ที่ใช้กันมานาน

#### ๔. ข้อมูลที่จัดเก็บ ได้แก่

- ๔.๑ ข้อมูลดิบจำนวนมากและมีขนาดใหญ่
- ๔.๒ ข้อมูลไม่มีรูปแบบที่แน่นอน
- ๔.๓ การเข้าถึงข้อมูลไม่สามารถเข้าถึงได้ง่าย
- ๔.๔ ต้องใช้ความสามารถของเจ้าหน้าที่วิเคราะห์ข้อมูล (Data scientist)

### ๕. ความแตกต่างระหว่าง Data Lake เมื่อเทียบกับ Data Warehouse

- ๕.๑ เก็บข้อมูลทั้งหมด
- ๕.๒ สนับสนุนข้อมูลทุกชนิด ไม่ใช่เพียงข้อมูลแบบ Structure
- ๕.๓ ผู้ใช้ทุกประเภทสามารถใช้งานได้
- ๕.๔ ประมวลและวิเคราะห์ข้อมูลได้รวดเร็วกว่า

### ๖. Big Data Analytics

คือ การวิเคราะห์ข้อมูล Big Data ทำให้มีข้อมูลที่เป็นข้อเท็จจริงซึ่งผ่านการวิเคราะห์อย่างเป็นระบบเพื่อใช้ประกอบการตัดสินใจ โดยยกระดับของการวิเคราะห์ก็เป็นได้หลากหลายแล้วแต่รูปแบบการนำไปใช้งาน ระดับการวิเคราะห์ข้อมูลแบ่งเป็น ๓ ระดับได้แก่

๖.๑ Descriptive analytics คือการวิเคราะห์ข้อมูลแบบพื้นฐาน การวิเคราะห์เพื่อแสดงผลของรายการทางธุรกิจ เหตุการณ์ หรือกิจกรรมต่างๆที่เกิดขึ้น หรืออาจกำลังเกิดขึ้นในลักษณะที่ง่ายต่อการเข้าใจ หรือต่อการตัดสินใจ ตัวอย่างเช่น รายงานการขาย รายงานผลการดำเนินงาน

๖.๒ Data Analytics คือการนำข้อมูลที่มีอยู่มาเข้าสู่กระบวนการวิเคราะห์ เพื่อนำข้อมูลไปใช้ประโยชน์ในด้านต่าง ๆ ที่ครอบคลุมในทุกอุตสาหกรรม ทั้งภาครัฐและภาคเอกชน ซึ่งจะต้องอาศัยเทคโนโลยีหรือซอฟต์แวร์เฉพาะด้านที่เข้ามาช่วยให้การประมวลผลและวิเคราะห์ข้อมูลเป็นไปอย่างแม่นยำและนำไปใช้งานได้อย่างมีประสิทธิภาพ

๖.๓ Prescriptive Analytics คือการวิเคราะห์สิ่งที่จะเกิดขึ้นจากการใช้ทางเลือกต่าง ๆ ในเชิงแนะนำ

## ๗. Data Mining

คือ การวิเคราะห์ข้อมูลจากข้อมูลจำนวนมาก (Big Data) เพื่อหาความสัมพันธ์ของข้อมูลที่ซ่อนอยู่ โดยทำการจำแนกประเภท รูปแบบ เชื่อมโยงข้อมูลที่มีความสัมพันธ์กัน และหาความน่าจะเป็นที่จะเกิดขึ้น เพื่อให้ได้องค์ความรู้ใหม่ที่สามารถนำไปใช้ประกอบการตัดสินใจในด้านต่างๆ เช่น ตลาดหลักทรัพย์ ทางธุรกิจ ทางด้านการแพทย์ ยุทธศาสตร์ทหาร เป็นต้น

## ๘. Data Driven Business Data

คือ การขับเคลื่อนธุรกิจด้วยข้อมูล การขับเคลื่อนธุรกิจด้วย Data คือ การตัดสินใจในกิจกรรมต่างๆ ด้วยการใช้ “ข้อมูล” ซึ่งต้องเป็นข้อมูลที่มี “ระบบ” ก่อให้เกิดการตัดสินใจได้อย่าง “ทันที” และ “อัตโนมัติ” มีอยู่ ๔ หัวข้อ ดังนี้

๑. สินทรัพย์ในทางธุรกิจ
๒. เข้าถึงลูกค้าได้ดีขึ้น
๓. พัฒนาประสิทธิภาพของการทำงาน
๔. พัฒนาความพึงพอใจของลูกค้าและผลิตภัณฑ์

## ๙. Social Media Command Center

การบริโภคสื่อออนไลน์จะทำให้ธุรกิจเห็นพฤติกรรมของผู้บริโภคที่เป็นกลุ่มเป้าหมาย รูปแบบที่ผู้บริโภคเข้าไปมีปฏิสัมพันธ์กับธุรกิจในโลกออนไลน์ คือ Data ที่มีค่าของธุรกิจ ทั้งการคลิก การลดแชร์ การใช้เวลากับหน้าเว็บไซต์แต่ละแห่ง

## ๑๐. ธุรกิจสามารถใช้ประโยชน์จาก Social Media Command Center ได้อย่างไร

๑. Trend Watching ตรวจสอบกระแสในขณะนั้น เพื่อดูว่าคนกล่าวถึงเรื่องอะไรกันบ้าง
๒. Brand Health ตรวจสอบสุขภาพของธุรกิจ มีการกล่าวถึงธุรกิจในแง่ใดบ้างใน Social Network สามารถลงรายละเอียดได้ว่าเป็นการกล่าวถึงในแง่บวกหรือลบมากกว่ากัน
๓. Online Campaign Measurement วัดกระแสของแคมเปญออนไลน์ หรือ Viral ของธุรกิจได้
๔. Influencer Analysis ตรวจสอบการพูดถึงธุรกิจ โดย Influencer แต่ละคนหรือจะเรียกว่า ตรวจสอบ Performance ของ Influencer ก็ได้
๕. Event wish Live Steam ค้นหาจากแฮชแท็กของ Event แบบเรียลไทม์ ข้อมูลที่ได้จะมีลักษณะเป็นไทม์ไลน์ไหลลงมาเรื่อย ๆ ไม่ใช่ Report ตอนจบงาน
๖. Market Research เพื่อให้ธุรกิจสามารถวิเคราะห์คู่แข่งและผู้บริโภคได้อย่างมีประสิทธิภาพ
๗. Remapping ทำให้แบรนด์ตรวจสอบ Feedback จากผู้บริโภคที่อยู่ต่างสถานที่กันได้ เช่น ธุรกิจที่มีสาขาจำนวนมาก หรือ Counter Service
๘. Keyword Selection เป็นสิ่งจำเป็นเบื้องต้น ของการเริ่มต้นในการเขียนบทความ เช่นเดียวกับการที่เราจะลงทำสิ่งต่าง ๆ นั้น ต้องรู้เป้าหมาย

## ๑๑. Hadoop

Hadoop คือ Java programming framework ที่รองรับการทำงานที่ต้องประมวลผลและเก็บข้อมูลขนาดใหญ่ เป็นส่วนหนึ่งของ Apache project โดย Hadoop ถูกออกแบบมาให้เป็น application ที่สามารถทำงานได้บนระบบแบบ node หรือมี hardware จำนวนหลายๆเครื่องพร้อมกัน เพื่อรองรับข้อมูลขนาดใหญ่ ซึ่งใช้การแตก file system ออกมากระจายตาม node ให้สามารถทำงานได้รวดเร็วพร้อมทั้งส่งข้อมูลหากันระหว่าง node ทั้งหมด รวมถึงมีความสามารถในการจัดการ node มีเสียหายได้โดยไม่ทำให้เกิดข้อมูลสูญหาย ปัจจุบัน Hadoop ถูกนำมาใช้ในงาน big data จำพวก การคำนวณข้อมูลทางวิทยาศาสตร์เฉพาะทาง, เชิงธุรกิจ รวมถึงวางแผนการขาย และประมวลผลข้อมูล sensor จำนวนมาก หรือ internet of things (IOT)

HDFS นั้นยังประกอบไปด้วย ๒ ส่วนย่อย

๑. Name Node (Master Node) คือ เป็นส่วนเก็บ Metadata หรือข้อมูลที่อธิบายรายละเอียดเกี่ยวกับข้อมูล อย่างเช่น จำนวนของบล็อก เครื่องที่จัดเก็บ ข้อมูลอันไหนถูกเก็บไว้ที่ไหน เป็นต้น
๒. data Node (Slave Node) คือ เป็นส่วนที่เก็บข้อมูลจริงๆของ HDFS ทำหน้าที่อ่านและเขียนตามคำสั่ง นอกจากนี้ยังทำหน้าที่ทำสำเนาข้อมูล (Replica) ของแต่ละบล็อกเก็บไว้ในอีก Data Node เพื่อป้องกันการสูญหายของข้อมูลอีกด้วย ทำให้เมื่อเกิดเหตุฉุกเฉินขึ้นมา ข้อมูลก็ยังอยู่กับเราครบ เรียกสถานการณ์แบบนี้ว่า มี fault-tolerant

### ส่วนประกอบของ Hadoop

อย่างที่บอกว่าเป็น Software framework ทำให้ Hadoop เองประกอบไปด้วย module จำนวนมาก ซึ่งขั้นต่ำที่จะต้องมียกคือ

- Hadoop Common: เป็น library ที่จำเป็นในการใช้งานเปรียบได้กับ kernel
- Hadoop Distributed File System (HDFS): สำหรับเก็บข้อมูลภายใต้ server node จำนวนมาก ซึ่งต้องอาศัย bandwidth ที่สูงมากในการทำงาน
- Hadoop Yet Another Resource Negotiator (YARN): ทำหน้าที่จัดการทรัพยากรที่มีอยู่ รวมถึงรอบการทำงานของ application
- Hadoop MapReduce: เป็นรูปแบบ program ที่จัดการกับ ข้อมูลขนาดใหญ่เพื่อสร้างความสัมพันธ์ระหว่างกัน

### ประโยชน์ของ Hadoop MapReduce

๑. ความเร็ว (Speed): สามารถประมวลผลข้อมูลทั้งแบบมีโครงสร้างและไม่มีโครงสร้างได้ภายในเวลาอันสั้น
๒. ทนทานต่อความเสียหาย (Fault-tolerance): สามารถทำงานต่อได้แม้จะเกิดความขัดข้องขึ้น
๓. ข้อมูลพร้อมใช้งานตลอดเวลา (Data Availability): มีข้อมูลที่ทำสำเนาและกระจายไปยัง Node อื่นๆ ทำให้มั่นใจได้ว่าข้อมูลจะไม่หายไปไหน
๔. ประมวลผลแบบขนาน (Parallel Processing): ประมวลผลข้อมูลแบบขนานกับไปกับงานอื่นๆได้ในชุดเดียวกัน ทำให้ทำงานได้ไวขึ้น

## ๑๒. Big data มีกระบวนการทำงาน มีอยู่ ๓ ขั้นตอนหลักๆ ดังนี้

### ๑. จัดเก็บข้อมูล (Storage)

เป็นขั้นตอนการจัดเก็บรวบรวมข้อมูลทั้งหมดจากแหล่งต่างๆ ไม่ว่าจะเป็นข้อมูลที่มีคุณภาพ รวมถึงข้อมูลที่คาดว่าอาจจะเป็นประโยชน์ ไม่ว่าจะเป็นข้อมูลที่เป็นตัวอักษร ไฟล์เอกสาร ไฟล์รูปภาพ ไฟล์วีดีโอ ไฟล์เสียงที่ถูกบันทึก จะถูกเก็บรวบรวมไว้ที่นี้

### ๒. การประมวลผลข้อมูล (Processing)

การประมวลผลข้อมูล หลังจากที่นำข้อมูลมารวบรวมไว้ได้ในที่เดียวแล้ว ข้อมูลต่างๆ จะถูกนำไปจัดหมวดหมู่ให้อยู่ในกลุ่มที่มีความเกี่ยวข้องกันหรือความสัมพันธ์ใกล้เคียงกัน ให้ผลลัพธ์คล้ายคลึงกันมากที่สุด แล้วจึงนำมาเปลี่ยนเป็นรูปแบบข้อมูลเพื่อนำเอาข้อมูลที่มีอยู่เหล่านี้เข้าระบบข้อมูลผ่านการประมวลผลแล้ว

### ๓. การวิเคราะห์ข้อมูล (Analyst)

การวิเคราะห์ข้อมูลและนำเสนอข้อมูล หลังจากที่ข้อมูลทั้งหมดได้ถูกจัดกลุ่มและแยกประเภทเรียบร้อยแล้วนั้น ต่อจากนั้นจะนำมาวิเคราะห์หา Pattern ความเกี่ยวข้องกันทั้งหมด ที่อาจมองไม่เห็นได้เลยด้วยตา ไม่ว่าจะเป็นการทำ แนวโน้มของการตลาด ความต้องการของลูกค้า กระแสที่อาจเกิดขึ้นได้ในอนาคต และข้อมูลด้านอื่นที่เป็นประโยชน์ และจัดมานำเสนอในรูปแบบต่างๆ ไม่ว่าจะเป็นรูปภาพหรือกราฟ

จึงเรียนมาเพื่อโปรดทราบ



(นางเกียร คำคง อากาศสุวรรณ)

นักวิชาการเกษตรชำนาญการ

# ประกาศนียบัตร

ให้ไว้เพื่อแสดงว่า

เกยร อากาศสุวรรณ

ได้ผ่านการอบรมด้วยระบบการเรียนออนไลน์ในบทเรียน  
ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการ  
และบุคลากรภาครัฐทุกระดับ

รวมระยะเวลาทั้งสิ้น 1 : 0 ชั่วโมง

โดยสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล  
ภายใต้การดำเนินงานของสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)  
ให้ไว้ ณ วันที่ 30 ก.ค. 2566

*A. H.*

( นางไอรดา เหลืองวิไล )

รองผู้อำนวยการสำนักงานพัฒนารัฐบาลดิจิทัล

รักษาการแทนผู้อำนวยการสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล

Signed by สำนักงานพัฒนารัฐบาลดิจิทัล(องค์การมหาชน) (สพร.)

Digital Government Development Agency (Public

Organization) (DGA)

Date: 2023-07-30T18:06:01.166+07:00



0718490d

# ประกาศนียบัตร

ให้ไว้เพื่อแสดงว่า

เกยूर อากาศสุวรรณ

ได้ผ่านการอบรมด้วยระบบการเรียนออนไลน์ในบทเรียน  
มาตรฐานกรอบธรรมาภิบาลข้อมูลภาครัฐ  
(Data Governance Framework Standard)

รวมระยะเวลาทั้งสิ้น ๐ : 30 ชั่วโมง

โดยสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล  
ภายใต้การดำเนินงานของสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)  
ให้ไว้ ณ วันที่ 30 ก.ค. 2566

*A. H.*

( นางไอรดา เหลืองวิไล )

รองผู้อำนวยการสำนักงานพัฒนารัฐบาลดิจิทัล

รักษาการแทนผู้อำนวยการสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล

Signed by สำนักงานพัฒนารัฐบาลดิจิทัล(องค์การมหาชน) (สพร.)

Digital Government Development Agency (Public

Organization) (DGA)

Date: 2023-07-30T18:06:01.538+07:00



a3981f99