ถอดบทเรียนจากการอบรมผ่านระบบ TDGA E-learning
หลักสูตร "ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ"
ระหว่างวันที่ ๔ สิงหาคม ๒๕๖๖

.....

- **ผู้เข้าอบรม :** นายศักรินทร์ จันทรัตน์ นักวิชาการเกษตรปฏิบัติการ
- **หลักสูตร :** ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ
- **หน่วยงานจัดอบรม :** สถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล Thailand Digital Government Academy
- **ระยะเวลา :** ระหว่างวันที่ ๔ สิงหาคม ๒๕๖๖
- **สถานที่ :** ณ สถานีพัฒนาที่ดินมุกดาหาร
- **วัตถุประสงค์ของการฝึกอบรม**

๑. เพื่อให้ผู้เรียนมีความรู้พื้นฐานเกี่ยวกับข้อมูลขนาดใหญ่ (Big Data)
๒. เพื่อให้ผู้เรียนมีความรู้พื้นฐานเกี่ยวกับเครื่องมือวิเคราะห์ข้อมูล (Hadoop) เพื่อการทำงานเกี่ยวกับข้อมูลขนาดใหญ่
๓. เพื่อให้ผู้เรียนมีความเข้าใจพื้นฐานเกี่ยวกับการวิเคราะห์ข้อมูลขนาดใหญ่เพื่อการบริหารภาครัฐ

ถอดบทเรียน

ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ ปัจจุบันข้อมูลมีความหลากหลายและมีปริมาณมาก จนก่อให้เกิดปัญหาในการนำไปใช้งาน จึงจำเป็นต้องมีการจัดการจำแนก รวมไปถึงการวิเคราะห์ข้อมูล เพื่อให้องค์กรสามารถเลือกใช้ข้อมูลได้สะดวกขึ้น และสามารถนำเสนอข้อมูลให้เข้าใจได้ง่าย

๑. Big Data (ข้อมูลขนาดใหญ่)

Big Data คือ ข้อมูลทุกสิ่งทุกอย่างทั้งแบบโครงสร้างปกติและโครงสร้างข้อมูลที่ไม่มีรูปแบบ ซึ่งทั้งหมดเป็นข้อมูลที่ใช้ในเชิงธุรกิจ มักจะถูกใช้กับงานพวกที่ต้องวิเคราะห์ ข้อมูลที่มีความซับซ้อน และไม่สามารถประเมินขนาดข้อมูลได้ (> ๑ PB = ๑๐๐ ล้านล้านไบต์ =๑,๐๒๔ TB)

๑.๑ รูปแบบของข้อมูลของ Big Data แบ่งได้หลายชนิด ได้แก่

๑) Behavioral Data ได้แก่ ข้อมูลเชิงพฤติกรรมการใช้งานต่างๆ เช่น Server Log พฤติกรรมการคลิกดูข้อมูล หรือข้อมูลการใช้ ATM

๒) Image & Sounds ตัวอย่างเช่น ภาพถ่ายวีดีโอ รูปจาก Google Street View การถ่ายภาพทางการแพทย์ ลายมือข้อมูลเสียงที่ถูกบันทึก

๓) Languages ตัวอย่างเช่น Text Message ข้อความที่ถูก Tweet เนื้อหาต่างๆ ในเว็บไซต์

๔) Records ตัวอย่างเช่น ข้อมูลทางการแพทย์ ข้อมูลผลสำรวจที่มีขนาดใหญ่ ข้อมูลทางภาษี

๕) Sensors ตัวอย่างเช่น ข้อมูลอุณหภูมิ ข้อมูลทางภูมิศาสตร์

๑.๒ Big Data ประกอบด้วยคุณลักษณะ ๔ ประการ คือ

๑) Volume ข้อมูลมีขนาดใหญ่ มีปริมาณข้อมูลมากซึ่งสามารถเป็นได้ทั้งข้อมูลแบบ Offline หรือ Online

๒) Variety ข้อมูลมีความหลากหลาย สามารถเป็นได้ทั้งที่มีโครงสร้างและข้อมูลที่ไม่สามารถจับ Pattern ได้

๓) Velocity ข้อมูลมีการเปลี่ยนแปลงตลอดเวลาอย่างรวดเร็ว มีการส่งผ่านข้อมูลอย่างต่อเนื่องในลักษณะ Streaming ทำให้การวิเคราะห์ข้อมูลแบบ Manual มีข้อจำกัด

๔) Veracity ข้อมูลมีความไม่ชัดเจน (Untrusted, Uncleaned)

๒. Data lake

Data lake คือ การเก็บข้อมูลขนาดใหญ่ทุกรูปแบบจากหลายแหล่งโดยไม่ต้องมีการแปลงข้อมูล ข้อดีของ Data lake คือ มีการนำเอาข้อมูลจากแหล่งข้อมูลภายนอกองค์กรข้อมูลจากเครือข่ายข้อมูลที่กระจายไปทั่วโลกมาใช้มากขึ้น ปริมาณข้อมูลจากแหล่งภายนอกเพิ่มอย่างต่อเนื่องและมีแนวโน้มที่จะเติบโตแบบก้าวกระโดดมากขึ้น และแก้ไขข้อจำกัดหลายอย่างของ Data Warehouse ที่ใช้กันมานาน

ข้อมูลที่สามารถจัดเก็บได้ คือ

๑) ข้อมูลดิบจำนวนมากและมีขนาดใหญ่

๒) ข้อมูลไม่มีรูปแบบที่แน่นอน

๓) การเข้าถึงข้อมูลไม่สามารถเข้าถึงได้ง่ายต้องใช้ความสามารถของเจ้าหน้าที่วิเคราะห์ข้อมูล (Data Scientist)

ความแตกต่างระหว่าง Data Lake เมื่อเทียบกับ Data Warehouse

๑) เก็บข้อมูลทั้งหมด

๒) สนับสนุนข้อมูลทุกชนิด ไม่ใช่เพียงข้อมูลแบบ Structure

๓) ผู้ใช้ทุกประเภทสามารถใช้งานได้

๔) ประมวลและวิเคราะห์ข้อมูลได้รวดเร็วกว่า

๓. Big Data Analytics

การวิเคราะห์ข้อมูล Big Data ทำให้มีข้อมูลที่เป็นข้อเท็จจริง ซึ่งผ่านการวิเคราะห์อย่างเป็นระบบเพื่อใช้ประกอบการตัดสินใจ โดยระดับของการวิเคราะห์ก็เป็นได้หลากหลายแล้วแต่รูปแบบการนำไปใช้งาน ระดับการวิเคราะห์ข้อมูลแบ่งได้ ดังนี้

๓.๑ Descriptive Analytics เป็นการวิเคราะห์ในระดับที่บอกว่าเกิดอะไรขึ้น จำนวนเท่าไร แค่ไหน เกิดเหตุการณ์สำคัญตอนไหน ตรงไหนบ้าง

๓.๒ Predictive Analytics เป็นการวิเคราะห์ในลักษณะที่ซับซ้อนขึ้นไปอีกขั้นหนึ่ง คือ เป็นการประเมินว่าจะเกิดอะไรขึ้นต่อไป มีการให้ข้อมูลตัวชี้วัดของผลลัพธ์ที่อาจจะเกิดขึ้นถ้าแนวโน้มยังเป็นอย่างนี้ต่อไป

๓.๓ Prescriptive Analytics (วิเคราะห์เชิงทำนาย) เป็นรูปแบบการวิเคราะห์ข้อมูลที่มีความซับซ้อนและยากที่สุด เพราะไม่เพียงพยากรณ์หรือทำนายว่าอะไรจะเกิดขึ้น แต่ยังให้คำแนะนำในทางเลือกต่างๆ

และผลของทางเลือกต่างๆ ว่าจะมีผลดีและผลเสียอย่างไร โมเดลของ Prescriptive Analytics นั้นจะสามารถปรับเปลี่ยนได้ตามข้อมูลที่เพิ่มเติมเข้ามามากขึ้น และ Prescriptive Analytics นี้ยังเป็นการใช้ข้อมูลที่มากที่สุดและเกี่ยวพันกับเรื่อง Big Data เป็นอย่างมาก

๔. Data Driven Business

การขับเคลื่อนธุรกิจด้วยข้อมูล คือ การตัดสินใจในกิจกรรมต่างๆ ด้วยการใช้ “ข้อมูล” ซึ่งต้องเป็นข้อมูลที่มี “ระบบ” ก่อให้เกิดการตัดสินใจได้อย่าง “ทันที” และ “อัตโนมัติ”

จุดเด่นของการขับเคลื่อนธุรกิจด้วยข้อมูล

๑) สินทรัพย์ในทางธุรกิจ ข้อมูลถือเป็นทรัพย์สินที่มีมูลค่า ข้อมูลจะแสดงให้เห็นทั้งจุดอ่อนและจุดแข็งจากการดำเนินงานที่ผ่านมา และช่วยให้เกิดการพัฒนาธุรกิจ

๒) เข้าถึงลูกค้าได้ดีขึ้น การเก็บข้อมูลของลูกค้าเพื่อให้ได้มาซึ่งบริการและสินค้าที่ตรงต่อความต้องการของลูกค้า ได้แก่ ต้องการช่องทางซื้อขาย และช่องทางการชำระเงินที่สะดวก นอกจากนี้ยังทำให้บริษัทเกิดความเปลี่ยนแปลงในด้านของการจัดการ โดยเฉพาะข้อมูลที่ต้องมีการพัฒนาระบบจัดเก็บและระบบป้องกันรักษาข้อมูล นอกจากนี้ทางบริษัทยังต้องตื่นตัวในการสร้างและปรับปรุงนโยบายที่เกี่ยวข้องกับการใช้ระบบข้อมูลให้มีความรัดกุมและทันสมัยอยู่เสมอ

๓) พัฒนาประสิทธิภาพและการทำงาน Big Data มีความสามารถที่จะช่วยด้านการพัฒนาประสิทธิภาพและการทำงานภายในของธุรกิจเกือบทุกประเภท เช่น ติดตามการทำงานของพนักงาน การติดตามด้านสุขภาพและความเครียดที่อาจเกิดขึ้นในระหว่างการทำงาน สามารถช่วยพัฒนาทรัพยากรบุคคลและการจ้างงานได้ด้วย นอกจากนี้ทางบริษัทยังต้องตื่นตัวในการสร้างและปรับปรุงนโยบายที่เกี่ยวกับการใช้ระบบข้อมูลให้มีความรัดกุมและทันสมัยอยู่เสมอ

๔) พัฒนาความพึงพอใจของลูกค้าและผลิตภัณฑ์ บริษัทจะสามารถใช้ข้อมูลที่เก็บมาได้จากลูกค้าในการพัฒนาผลิตภัณฑ์และประสบการณ์การใช้สินค้าได้ เพื่อคาดการณ์ถึงปัญหาที่อาจเกิดขึ้นกับในอนาคต

ตัวอย่างการใช้ Big Data

๑) การเลือกตั้ง

๒) e-Commerce

๓) งานราชการ

๔) สุขภาพและสาธารณสุข

๕) Digital Media

๖) การเงิน

๗) การค้า

๕. Social Media Command Center

การบริโภคสื่อออนไลน์จะทำให้ธุรกิจเห็นพฤติกรรมของผู้บริโภคที่เป็นกลุ่มเป้าหมายรูปแบบที่ผู้บริโภคเข้าไปมีปฏิสัมพันธ์กับธุรกิจในโลกออนไลน์ คือ Data ที่มีค่าของธุรกิจทั้งการคลิก การกดแชร์ การใช้เวลากับหน้าเว็บไซต์แต่ละแห่ง

ข้อมูลที่รวบรวมมาจากออนไลน์

- ข้อมูลด้าน Demographic หรืออายุ เพศ การศึกษา หรืออาชีพ

- ข้อมูลด้านไลฟ์สไตล์ และความสนใจ ซึ่งธุรกิจจะนำไปใช้ในการเลือกกลุ่มเป้าหมายในการโฆษณาได้แม่นยำมาก

- Social Media Command Center กำลังเป็นเครื่องมือสำคัญและกำลังได้รับความนิยมเป็นอย่างสูงที่จะคอยดูแลความเคลื่อนไหวธุรกิจองค์กรที่เกิดขึ้นบนโลกออนไลน์

- Data Visualization การแสดงผลข้อมูลในรูปแบบที่เข้าใจได้ง่าย

- Real Time Monitoring ไม่ใช่การรวบรวมเป็นรายงานสรุป แต่แสดงผลออกมาแบบเรียลไทม์

- Quality Data นอกจากจะบอกปริมาณแล้ว ยังบอกทิศทางและรายละเอียดว่าคนกล่าวถึงธุรกิจในแง่มุมบวกหรือลบ

๖. Big Data Analytics กับการบริหารภาครัฐ

Big Data Analytics เป็นเครื่องมือที่มีความจำเป็นของภาครัฐในทุกประเทศในการขับเคลื่อนการบริหารราชการแผ่นดินอย่างมีประสิทธิภาพ และขับเคลื่อนเศรษฐกิจให้ทันต่อการเปลี่ยนแปลงทางเทคโนโลยี อย่างก้าวกระโดดในศตวรรษที่ ๒๑

๑) รับฟังความเห็น รวบรวมข้อมูล และปรึกษากับผู้มีส่วนได้ส่วนเสีย

๒) วางแผนการลงทุนในการจัดโครงสร้าง

๓) มีความเข้าใจและมีทักษะทางธุรกิจและทักษะทางเทคนิค

๔) เตรียมพร้อมภายใต้การเปลี่ยนแปลงของเทคโนโลยี

๕) เจ้าหน้าที่ภาครัฐจะต้องปรับ Mindset ในการเข้าร่วมกับทุกภาคส่วน

๖) ปรับปรุงวิธีคิดและกระบวนการเพื่อทำให้เกิดการแลกเปลี่ยนข้อมูล และการใช้ข้อมูลร่วมกันระหว่างหน่วยงานภาครัฐ

๗) กำหนดแนวทางและการบริการให้คำปรึกษาในด้าน Big Data Analytics ให้แก่ทุกภาคส่วน

ประสบการณ์/ประโยชน์ที่ได้รับ/การประยุกต์ใช้กับหน่วยงาน

ต่อตนเอง

ทำให้ได้เรียนรู้เกี่ยวกับความรู้พื้นฐานเกี่ยวกับข้อมูลขนาดใหญ่ (Big Data) เครื่องมือวิเคราะห์ข้อมูล (Hadoop) เพื่อการทำงานเกี่ยวกับข้อมูลขนาดใหญ่ ตลอดจนการวิเคราะห์ข้อมูล เพื่อสามารถนำไปใช้เป็นแนวทางในการจัดการข้อมูลที่มีอยู่ได้อย่างเป็นระบบ

ต่อหน่วยงาน/การนำมาประยุกต์ใช้กับหน่วยงาน

สามารถนำความรู้เรื่อง Big Data และการนำไปใช้ให้เกิดประโยชน์ กระบวนการในการจัดเก็บ Big Data รวมถึงเรียนรู้เกี่ยวกับเครื่องมือที่ใช้ในการวิเคราะห์ Big Data และแนวทางในการวิเคราะห์ Big Data เพื่อการทำงานภาครัฐ ตลอดจนสามารถนำไปใช้เป็นแนวทางในการจัดการข้อมูลที่มีอยู่ของหน่วยงานให้สามารถเลือกใช้ข้อมูลได้สะดวกขึ้น และสามารถนำเสนอข้อมูลให้เข้าใจได้ง่าย

ประกาศนียบัตร

ให้ไว้เพื่อแสดงว่า

ศักรินทร์ จันทรัตน์

ได้ผ่านการอบรมด้วยระบบการเรียนออนไลน์ในบทเรียน
ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการ
และบุคลากรภาครัฐทุกระดับ

รวมระยะเวลาทั้งสิ้น 1 : 0 ชั่วโมง

โดยสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล
ภายใต้การดำเนินงานของสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
ให้ไว้ ณ วันที่ 4 ส.ค. 2566

(นางไอรดา เหลืองวิโล)
รองผู้อำนวยการสำนักงานพัฒนารัฐบาลดิจิทัล
รักษาการแทนผู้อำนวยการสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล