



บันทึกข้อความ

กองนโยบายและแผนการใช้ที่ดิน
เลขที่รับ ๗๖๕
วันที่ ๑๔. กพ. ๖๕
เวลา ๑๑.๐๐ - ๑๑.๐๐ น.

ส่วนราชการ กลุ่มวางแผนบริหารจัดการพื้นที่ชุ่มน้ำ กองนโยบายและแผนการใช้ที่ดิน โทร. ๒๒๐๘

ที่ กษ ๐๘๓๗.๐๕/๖๗๑

วันที่ ๑๑ กุมภาพันธ์ ๒๕๖๘

เรื่อง ขอส่งรายงานสรุปการอบรมการพัฒนาความรู้ ครั้งที่ ๑/๒๕๖๘

เรียน ผู้อำนวยการกลุ่มวางแผนบริหารจัดการพื้นที่ชุ่มน้ำ

ตามที่ ข้าพเจ้านายดิเรก คงแพ ตำแหน่ง นักวิเคราะห์นโยบายและแผนชำนาญการพิเศษ ได้เข้าร่วมอบรมพัฒนาความรู้ และอบรมพัฒนาความรู้ผ่านระบบ e-Learning ซึ่งใช้เป็นตัวชี้วัดรายบุคคลตามแบบประเมินผลสัมฤทธิ์ รอบที่ ๑ ปีงบประมาณ ๒๕๖๘ จำนวน ๒ เรื่อง ได้แก่ (๑) ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ และ (๒) การสร้างความเข้าใจกรอบธรรมาภิบาลข้อมูล (Data Governance Framework Awareness) ของสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) ผ่านระบบการฝึกอบรมผ่านสื่ออิเล็กทรอนิกส์ (https://e-learning.dga.or.th/xlms_ega/userroom/elearning.jsp) จึงขอส่งรายงานสรุปการอบรมเรื่องดังกล่าว ตามเอกสารแนบ

จึงเรียนมาเพื่อโปรดพิจารณา

(นายดิเรก คงแพ)

นักวิเคราะห์นโยบายและแผนชำนาญการพิเศษ

เรียน ผอ.นผ.

เพื่อโปรดพิจารณา

(นางสาวพิมพ์พร พรพรหมินทร์)

นักวิเคราะห์นโยบายและแผนชำนาญการพิเศษ
ผู้อำนวยการกลุ่มวางแผนบริหารจัดการพื้นที่ชุ่มน้ำ

รายงานสรุปการอบรม/สัมมนา/พัฒนาความรู้/ประชุมเชิงปฏิบัติการ/และเป็นวิทยากร
กองนโยบายและแผนการใช้ที่ดิน กรมพัฒนาที่ดิน

ส่วนที่ 1 ข้อมูลทั่วไป

ชื่อ.....นายดิเรก.....นามสกุล.....คงแพ.....
ตำแหน่ง.....นักวิเคราะห์นโยบายและแผนชำนาญการพิเศษ.....กลุ่ม/ฝ่าย.....กลุ่มวางแผนบริหารจัดการพื้นที่ชุ่มน้ำ
หลักสูตร/หัวข้อเรื่องอบรม/สัมมนา/พัฒนาความรู้.....
.....หลักสูตร.....ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ
สถานที่อบรม/สัมมนา/พัฒนาความรู้.....
.....ระบบการฝึกอบรมผ่านสื่ออิเล็กทรอนิกส์.....ของ.....สถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล
(https://e-learning.dga.or.th/xlms_ega/userroom/elearning.jsp).....
หน่วยงานที่จัดฝึกอบรม/สัมมนา/พัฒนาความรู้.....
.....สถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล.....สำนักงานพัฒนารัฐบาลดิจิทัล(องค์การมหาชน) (สพร.).....
ตั้งแต่วันที่.....11.....เดือน.....กุมภาพันธ์.....พ.ศ.....2568.....ถึงวันที่.....11.....เดือน.....กุมภาพันธ์.....พ.ศ.....2568 (เช้า)

เพื่อ ☒ อบรม ☐ สัมมนา ☐ อื่นๆ ระบุ.....

ส่วนที่ 2 สิ่งที่ได้รับจากการอบรม/สัมมนา/พัฒนาความรู้

2.1 รายงานสรุปเนื้อหาสาระสำคัญในการอบรม/ สัมมนา/พัฒนาความรู้.....
.....การอบรมหลักสูตร.....ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ มีวัตถุประสงค์ (1) เพื่อให้มีความรู้เกี่ยวกับประเภทข้อมูลต่าง ๆ (2) เพื่อให้มีความรู้ความเข้าใจเกี่ยวกับการวิเคราะห์ข้อมูลและระดับการวิเคราะห์ข้อมูล และ (3) เพื่อให้เกิดความเข้าใจและสามารถปรับตัวในการใช้ข้อมูลสำหรับปฏิบัติงาน

.....อธิบายแยกตามหลักสูตร.....แบ่งได้เป็น.....

บทเรียนที่ 1 Big Data

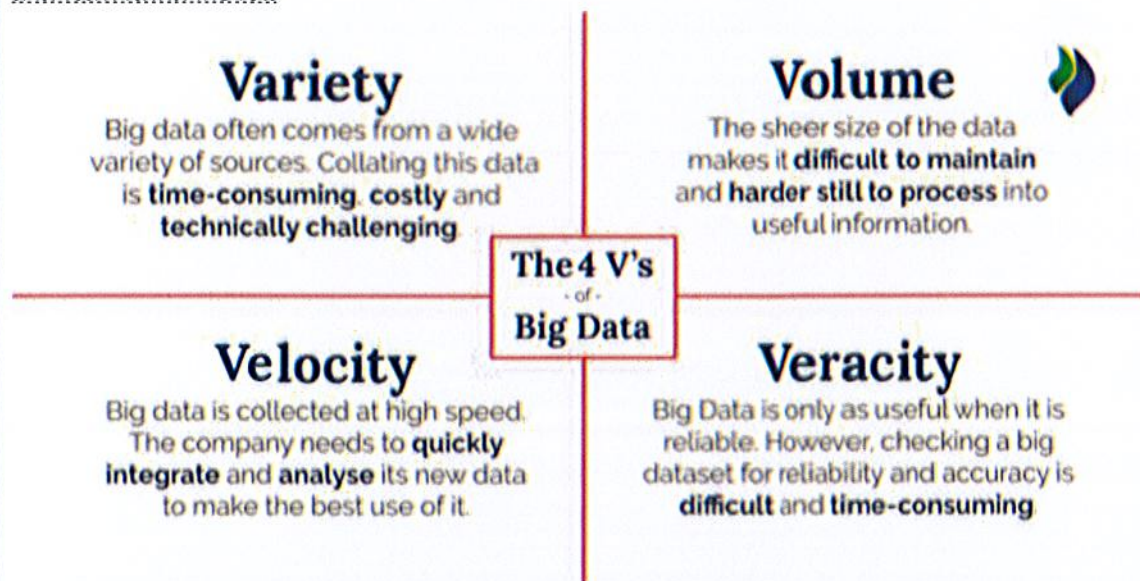
.....What is Big Data? Big Data คือข้อมูลที่มีขนาดใหญ่ มีทั้งแบบโครงสร้างปกติและโครงสร้างข้อมูลที่ไม่ใช่รูปแบบ ซึ่งทั้งหมดเป็นข้อมูลที่ใช้ในเชิงธุรกิจ มักจะถูกนำไปใช้กับงานที่ต้องวิเคราะห์ ข้อมูลที่มีความซับซ้อน และไม่สามารถประเมินขนาดข้อมูลได้



.....What is Big Data?
.....รูปแบบของข้อมูล Big Data สามารถเป็นได้หลากหลาย ได้แก่

.....1. Behavioral Data ได้แก่ ข้อมูลเชิงพฤติกรรมการใช้งานต่าง ๆ เช่น server log, พฤติกรรมการคลิกดูข้อมูล หรือข้อมูลการใช้งาน ATM	
.....2. Image & Sounds เช่น ภาพถ่าย วิดีโอ รูปจาก Google Street View ภาพถ่ายทางการแพทย์ ลายมือ ข้อมูลเสียงที่บันทึกไว้	
.....3. Languages เช่น Text Message ข้อความที่ถูก Tweet เนื้อหาต่าง ๆ ในเว็บไซต์	
.....4. Records เช่น ข้อมูลทางการแพทย์ ข้อมูลผลการสำรวจที่มีขนาดใหญ่ ข้อมูลภาษี	
.....5. Sensors เช่น ข้อมูลอุณหภูมิ ข้อมูลภูมิศาสตร์	
ลักษณะสำคัญของ Big Data	
Big Data จะต้องมียกย่องสำคัญ 4V ดังต่อไปนี้ จึงจะเรียกได้ว่าเป็น Big Data	
.....1. ปริมาณ (Volume)	
คือ ปริมาณข้อมูลที่สามารถผลิตและจัดเก็บไว้ได้จะต้องขนาดที่ใหญ่มากเพียงพอ ซึ่งปริมาณของข้อมูลจะเป็นข้อบ่งบอกได้ถึงคุณภาพและประสิทธิภาพของข้อมูลภายในนั้น และสามารถนำไปพิจารณาต่อได้ว่าข้อมูลที่มีอยู่เหล่านั้น ถือเป็น Big Data หรือไม่	
.....2. ความหลากหลาย (Variety)	
คือ ความหลากหลายของประเภทของข้อมูล โดยสามารถเพิ่มประสิทธิภาพในการนำข้อมูลไปวิเคราะห์ต่อยอดได้ ไม่ว่าจะเป็นข้อมูลที่เป็นตัวหนังสือ, รูปภาพ, ข้อมูลเสียงที่บันทึกไว้, วิดีโอหรือไฟล์ประเภทอื่นจากหลากหลายแหล่งที่มา ก็สามารถเป็นส่วนหนึ่งในการเพิ่มประสิทธิภาพการทำงานของ Big Data ได้	
ซึ่งจะเป็นข้อมูลจากทุกฝ่ายไม่ว่าจะเป็นการเงิน ฝ่ายขาย การตลาด ลูกค้าสัมพันธ์ บัญชี รวมถึงตัวหนังสือที่อาจเป็นบทสนทนาระหว่างแผนก หรือระหว่างบริษัท ซึ่งอาจเป็นข้อความจาก Social Media รวมถึง URLs ที่มีข้อมูลเข้ามาอยู่ในทุกวัน	
.....3. ความเร็ว (Velocity)	
คือ ความเร็วในการประมวลผลและผลิตข้อมูลขึ้นมาเพื่อให้ทันกับความต้องการของผู้ใช้งาน ซึ่ง Big Data คือข้อมูลที่เข้ามาแบบ Real-Time และประมวลผลอยู่ตลอดเวลา แตกต่างจาก Small Data ที่ไม่สามารถทำได้	
Big Data จะมีความถี่ในการประมวลผลที่มากกว่า การบันทึกข้อมูลที่มากกว่า และเผยแพร่ข้อมูลที่มากกว่า ข้อมูลของ Big Data ที่มีอัตราการเพิ่มขึ้นของข้อมูลที่มีการเพิ่มขึ้นได้อย่างรวดเร็ว โดยไม่ว่าจะเป็น	
.....-ข้อมูลตัวอักษรการสนทนา	
.....-ข้อมูลการบันทึกเสียง	
.....-ข้อมูลการถ่ายภาพวิดีโอ	
.....-ข้อมูลอัตราการสั่งซื้อสินค้า	
.....-ข้อมูลโปรโมชันต่างๆ	
ซึ่งหากมานั่งดูแล้วจะถือว่าเป็นข้อมูลที่มีอยู่มาก และมีการอัปเดตเคลื่อนไหวอยู่ตลอดเวลา	
.....4. คุณภาพของข้อมูล (Veracity)	
คือ คุณภาพของข้อมูลที่สามารถนำไปวิเคราะห์ข้อมูลต่อไปได้อีก เป็นข้อมูลที่ยังไม่ผ่านการประมวลผลอยู่ในรูปแบบของข้อมูลดิบซึ่งสามารถนำไปประมวลผลต่อไปได้ และเป็นข้อมูลที่มาจากหลากหลายแหล่งไม่ว่าจะเป็น Facebook, Youtube, Twitter ซึ่งข้อมูลจากแหล่งเหล่านี้ยากที่จะสามารถควบคุมคุณภาพรวมถึงการคัดกรองข้อมูล และความน่าเชื่อถือของข้อมูล จึงต้องนำข้อมูลเหล่านี้เข้าสู่กระบวนการทำ Data Cleansing	
นอกจาก 4V ที่เป็นลักษณะสำคัญของ Big Data นี้แล้วก็ยังมียกย่องข้ออื่นๆ ที่สามารถบ่งบอกได้ว่าเป็น Big Data เช่นเดียวกัน ได้แก่	
Scalability คือ ขนาดของข้อมูลทั้งหมดที่มีที่ต้องสามารถขยายขนาดได้อย่างรวดเร็ว	

Relational คือ ความเกี่ยวข้องกันของข้อมูล ข้อมูลที่มีความเกี่ยวข้องกันอยู่จะสามารถทำให้การประมวลผลสามารถทำได้มากยิ่งขึ้น



บทเรียนที่ 2 Data Lake

Data Lake เกิดขึ้นเนื่องจากการนำเอาข้อมูลจากแหล่งข้อมูลภายนอกองค์กร ข้อมูลจากเครือข่ายข้อมูลที่กระจายไปทั่วโลกมาใช้มากขึ้น ปริมาณข้อมูลจากแหล่งภายนอกเพิ่มอย่างต่อเนื่องและมีแนวโน้มที่จะเติบโตแบบก้าวกระโดดมากขึ้น Data Lake แก้ไขข้อจำกัดหลายอย่างของ Data Warehouse ที่ใช้กันมานาน

ข้อมูลที่จัดเก็บใน Data Lake คือ

- ข้อมูลดิบจำนวนมากและมีขนาดใหญ่
- ข้อมูลไม่มีรูปแบบแน่นอน
- การเข้าถึงข้อมูลไม่สามารถเข้าถึงได้ง่าย ต้องใช้ความสามารถของเจ้าหน้าที่วิเคราะห์ข้อมูล (Data Scientist)

ความแตกต่างระหว่าง Data Lake กับ Data Warehouse

Data Lake นั้นเก็บข้อมูลทั้งหมด สนับสนุนข้อมูลทุกชนิดไม่ใช่เพียงข้อมูลแบบ Structure ซึ่งผู้ใช้สามารถใช้ประมวลผลและวิเคราะห์ข้อมูลได้รวดเร็วกว่า

บทเรียนที่ 3 Big Data Analytic

Big Data Analytic คือการวิเคราะห์ข้อมูล Big Data ทำให้มีข้อมูลที่เป็นข้อเท็จจริงที่ผ่านการวิเคราะห์อย่างเป็นระบบเพื่อประกอบการตัดสินใจ โดยระดับการวิเคราะห์เป็นได้หลากหลายแล้วแต่รูปแบบการนำไปใช้งาน

ระดับการวิเคราะห์ข้อมูล

1. Descriptive Analytics เป็นการวิเคราะห์ในระดับที่นอกเหนือจากเกิดอะไรขึ้น เกิดเหตุการณ์สำคัญตอนไหน จำนวนเท่าไร ถัดไปไหน ตรงไหนบ้าง
2. Predictive Analytics เป็นการวิเคราะห์ในระดับที่ซับซ้อนยิ่งขึ้นหนึ่ง เป็นการประเมินว่าจะเกิดเหตุการณ์อะไรขึ้นต่อไป มีการให้ข้อมูลตัวชี้วัดของผลลัพธ์ที่อาจเกิดขึ้นถ้าแนวโน้มยังเป็นเช่นนี้ต่อไป
3. Prescriptive Analytics เป็นรูปแบบการวิเคราะห์ที่มีความซับซ้อนและยากที่สุด เพราะไม่เพียงพยากรณ์หรือทำนายว่าอะไรจะเกิดขึ้น แต่ยังให้คำแนะนำในทางเลือกต่าง ๆ และผลของทางเลือกต่าง ๆ ว่าจะมีผลดีและผลเสียอย่างไร โมเดลของ Prescriptive Analytics สามารถปรับเปลี่ยนได้ตามข้อมูลที่เพิ่มเติม

เข้ามามากขึ้นและ Prescriptive Analytics. นี่ยังเป็นการใช้ข้อมูลที่มากที่สุดและเกี่ยวกับเรื่อง Big Data อย่างมาก

รูปแบบการวิเคราะห์ แบ่งออกเป็น 3 ประเภท

..... 1. Data Mining คือการวิเคราะห์ข้อมูลจากข้อมูลจำนวนมาก (Big Data) เพื่อหาความสัมพันธ์ของข้อมูล โดยการจำแนกประเภท รูปแบบ เชื่อมโยงข้อมูลที่มีความสัมพันธ์กัน และความน่าจะเป็นที่เกิดขึ้น เพื่อให้ได้องค์ความรู้ใหม่ ที่สามารถนำไปใช้ประกอบการตัดสินใจในด้านต่าง ๆ เช่น ตลาดหลักทรัพย์ ธุรกิจ

..... 2. Text Mining เป็นเทคนิคเพื่อค้นหารูปแบบ (Pattern) จากข้อความมหาศาลโดยอัตโนมัติ โดยใช้ขั้นตอนวิธีจากวิชาสถิติ การเรียนรู้ของเครื่อง และการรู้จำแบบ หรือในอีกนิยามหนึ่ง การทำเหมืองข้อความ คือ กระบวนการที่กระทำกับข้อความ (โดยส่วนใหญ่จะมีจำนวนมาก) เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อความนั้น โดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง หลักคณิตศาสตร์ หลักการประมวลเอกสาร (Document Processing) หลักการประมวลผลข้อความ (Text Processing) และ หลักการประมวลผลภาษาธรรมชาติ (Natural Language Processing)

..... 3. Machine Learning เป็นศาสตร์ของโครงสร้างโมเดลคณิตศาสตร์ มุ่งเน้นที่จะสร้างองค์ความรู้ จากข้อมูลโดยเริ่มจากการสร้างโมเดลทางคณิตศาสตร์ที่มีความยืดหยุ่น และสามารถปรับตัวเองเข้ากับข้อมูล ที่ได้รับ เพื่อที่จะสามารถทำนายอนาคตได้

Descriptive Analytics เป็นการวิเคราะห์ในระดับที่บอกว่าเกิดอะไรขึ้น เพื่อบรรยายข้อมูลลักษณะข้อมูลนั้น ตัวอย่างเช่น

การเลือกตั้ง

- - จัดเก็บข้อมูลของผู้มีสิทธิเลือกตั้ง การเปลี่ยนแปลงที่อยู่ของผู้มีสิทธิเลือกตั้ง ผู้สมัครรับเลือกตั้ง
- - ข้อมูลเกี่ยวกับรายงานการทุจริตเลือกตั้งที่ส่งเข้ามา
- - ข้อมูลความคิดเห็นของผู้มีสิทธิเลือกตั้งเกี่ยวกับวิธีการจัดการเลือกตั้ง สถานที่และความสะดวก
- - ให้บริการตรวจสอบข้อมูลการเลือกตั้ง สถานที่จัดเลือกตั้ง ทางอินเทอร์เน็ต

e-commerce

..... - Cross-channel Analytics : ใช้ตรวจวัดความพึงพอใจของลูกค้าจากประสบการณ์ของลูกค้าที่มี ต่อผู้ให้บริการ

..... - Event Analytics : ช่วยในการตรวจวัดว่าผู้ใช้งานมีปฏิสัมพันธ์กันเนื้อหาในเว็บไซต์ของท่านอย่างไร เช่น ต้องการทราบจำนวนครั้งของปุ่มที่ถูกกดหรือมีขึ้นของข้อความหรือหัวข้อใดที่ถูกใช้งานมาก Predictive Analytics ช่วยให้สามารถพิสูจน์ทราบเหตุการณ์ก่อนที่จะเกิดขึ้นจริง เช่น การพิสูจน์ทราบ แบบแผนการขายจากกรอบของเวลาที่ผ่านมาเพื่อการพยากรณ์ปริมาณของสินค้าที่จะต้องจัดเก็บในสต็อกสำหรับขายรอบต่อไป

ตัวอย่างเช่น

..... **งานราชการ**

- • การตรวจสอบการทุจริต การตรวจสอบภัยคุกคาม
- • การรักษาความปลอดภัยในระบบไอที การปฏิบัติตามกฎระเบียบและการวิเคราะห์
- • การปฏิบัติตามกฎหมายภาษีอากร-การทุจริตและการละเมิด
- • การพยากรณ์แนวโน้มอาชญากรรมและการป้องกัน

..... **ด้านสุขภาพและสาธารณสุข**

..... • ข้อมูลประกอบการรักษา เช่น ข้อมูลประวัติผู้ป่วย ประวัติรักษาพยาบาล ประวัติการแพ้ยา หรือข้อมูลประวัติครอบครัวที่ใช้ประกอบการวิเคราะห์

..... • ผู้ผลิตยาและเวชภัณฑ์ใช้ข้อมูลมากและหลากหลาย เพื่อหาสาเหตุการเจ็บป่วยที่แท้จริง การวิเคราะห์เพื่อจำเพาะเจาะจงกลุ่มผู้ป่วยที่จะทดลองและติดตามผลการรักษา จากยาและเวชภัณฑ์ที่พัฒนาขึ้น และเพื่อการพัฒนากลยุทธ์ด้านการตลาดของยานั้น

- • การวิเคราะห์หลักและรูปแบบการแพร่เชื้อ เพื่อใช้ในงานวิจัยทางแพทย์
- • วิเคราะห์ปริมาณ Click Stream ขนาดใหญ่
- • การกำหนดเป้าหมายการโฆษณา การวิเคราะห์ พยากรณ์และแนวโน้มในเรื่องต่างๆ
- • การป้องกันการใช้งานผิดวิธี และการฉ้อโกงทางเว็บไซต์
- • การวิเคราะห์กราฟการใช้งานบนสังคมออนไลน์และการแบ่งส่วนแบ่งประเภทลูกค้า

.....การเงิน

- • การปฏิบัติตามระเบียบและการรายงาน
- • การวิเคราะห์และบริหารจัดการความเสี่ยง การวิเคราะห์การรักษาความปลอดภัย
- • การตรวจสอบทุจริตและบริหารจัดการความเสี่ยง การวิเคราะห์แบบแผนทางการค้าขายที่ผิดปกติ

.....บทเรียนที่ 4 Data Driven Business

..... Data Driven Business คือ Data = ข้อมูล Drive(n) = ขับเคลื่อน Business = ธุรกิจ เมื่อนำคำทั้ง 3 มารวมกัน Data Driven Business จึงหมายถึง การขับเคลื่อนธุรกิจด้วยข้อมูล การขับเคลื่อนธุรกิจด้วย Data คือ การตัดสินใจในกิจกรรมต่างๆ ด้วยการใช้ “ข้อมูล” ซึ่งต้องเป็นข้อมูลที่มี “ระบบ” ก่อให้เกิดการตัดสินใจได้อย่าง “ทันที” และ “อัตโนมัติ” กลยุทธ์ทางการตลาดที่ใช้ข้อมูลที่ได้จากลูกค้ามาจัดการ วิเคราะห์ เพื่อนำไปประกอบการตัดสินใจในการวางแผนกลยุทธ์การตลาดอื่นๆ ในอนาคต ซึ่งข้อมูลของลูกค้าสามารถศึกษาได้จากผลิตภัณฑ์ที่ลูกค้าเลือก พฤติกรรมการบริโภค การตอบสนองของลูกค้าต่อโฆษณาต่างๆ เมื่อได้ข้อมูลเหล่านี้มาแล้ว จะนำไปคาดการณ์ความต้องการของลูกค้า และใช้เป็นฐานข้อมูลในการวางแผนกลยุทธ์อื่นๆ เพื่อสร้าง Return on investment (ROI) ให้ได้มากที่สุดนั่นเอง

.....บทเรียนที่ 5 Big Data Analytics กับการบริหารภาครัฐ

..... Social Media Command Center หนึ่งในตัวอย่างเครื่องมือที่สร้างขึ้นมากเพื่อรองรับการนำ Big Data มาใช้ในเชิงธุรกิจ ภายใต้โซลูชันที่มีชื่อว่า TH3RE ซึ่งพัฒนาโดยบริษัท Computerlogy เครื่องมือดังกล่าวเปรียบเสมือนทีมงานที่ดูแลและมอนิเตอร์เรื่องราวต่างๆ บนโลกออนไลน์ให้แทนทีมงานของแบรนด์สินค้า ซึ่งมีคุณสมบัติหลักๆ คือ การดึงข้อมูลบนโลกออนไลน์ได้แบบเรียลไทม์, ข้อมูลที่คัดกรองมีคุณภาพ, แสดงผลข้อมูลที่เข้าใจง่ายเป็นภาพสวยงาม (Data Visualization) ซึ่งแบรนด์สามารถนำไปประยุกต์ใช้ได้หลากหลายรูปแบบ เช่น

- ● การดูภาพลักษณ์ของแบรนด์ และสินค้าต่างๆ บนออนไลน์ แยกแต่ละสื่อว่าเป็นอย่างไรบ้าง ในเชิงบวกหรือเชิงลบ
- ● ดูความเคลื่อนไหวแบบเรียลไทม์ ถ้าเกิดวิกฤตจะได้ป้องกันอย่างทันท่วงที
- ● วัดผลแคมเปญออนไลน์ หรืออีเวนท์ที่เปิดตัวไปว่าส่งผลกระทบมากน้อยเพียงใดในเชิงการตลาด
- ● ดู Insight ของลูกค้าในหลากหลายมิติมากขึ้น เช่น
 - ○ ดูได้ว่าบริเวณใด สถานที่ใด ช่วงเวลาไหนที่มีการพูดถึงแบรนด์หรือสินค้าบ่อยๆ
 - ○ เทรนด์คืออะไร คนพูดถึงอะไรกันบ่อยๆ และพูดถึงอะไรต่อ
- ● ข้อมูลสามารถนำไปต่อยอดวิจัยการตลาด (Market Research) ได้

.....บทเรียนที่ 6 Big Data Analytics กับการบริหารภาครัฐ

..... องค์การภาครัฐมีความจำเป็นที่จะต้องสร้างมูลค่าจากการวิเคราะห์ Big Data โดยมีแนวทางดังนี้

- 1. รับฟังความเห็น รวบรวมข้อมูล และปรึกษากับผู้มีส่วนได้ส่วนเสีย ทุกภาคส่วนเกี่ยวกับศักยภาพ การเพิ่มประสิทธิภาพและการกำหนดนิยามทางเทคนิคที่เกี่ยวข้องรวมถึงเป้าหมายที่จะทำให้เกิดประโยชน์ในการนำเอาข้อมูลมาร่วมใช้อย่างมีประสิทธิภาพในรูปแบบ Open Source ให้มากขึ้น
- 2. วางแผนการลงทุนในการจัดโครงสร้าง รวบรวมข้อมูลและการบริหารจัดการข้อมูลโดยอยู่บนพื้นฐานของความเชื่อมั่นต่อผู้มีส่วนได้ส่วนเสีย ซึ่งการจัดการข้อมูลต่างๆ จะต้องเป็นไปตามกฎหมาย และหลัก

จริยธรรมที่เกี่ยวข้อง รวมทั้งต้องสื่อสารให้ทุกภาคส่วนรวมทั้งประชาชนให้มีความเข้าใจ และมั่นใจในเรื่อง ความเป็นส่วนตัวของข้อมูล (Privacy)

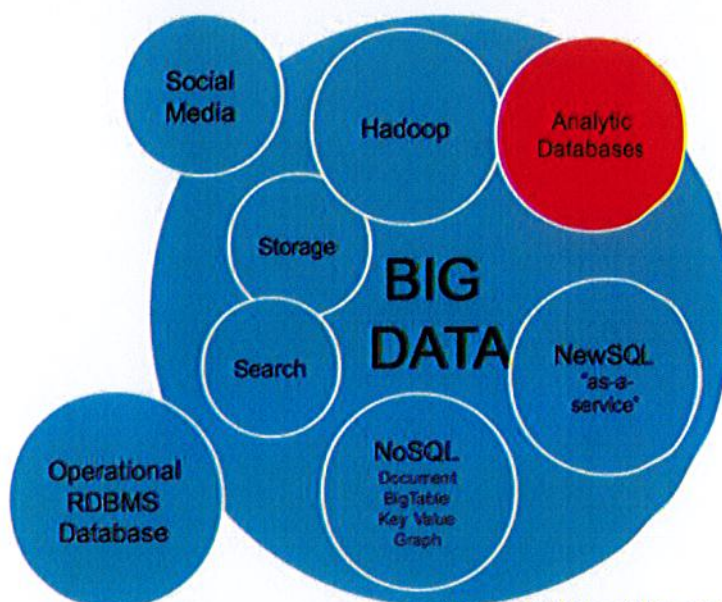
.....3. มีความเข้าใจและมีทักษะทางธุรกิจและทักษะทางเทคนิค ความท้าทายที่เกิดขึ้นจากพลังของผลการ วิเคราะห์ Big Data ในระดับชาตินั้น ผู้บริหารและเจ้าหน้าที่ในหน่วยงานภาครัฐจะต้องมีความเข้าใจ และมี ทักษะทางธุรกิจและทักษะทางเทคนิคอย่างมืออาชีพโดยจะต้องทำให้ Big Data สามารถเชื่อมโยง ให้เกิด ประโยชน์กับผู้ที่มีส่วนได้ส่วนเสียทุกกลุ่มโดยเฉพาะอย่างยิ่งประชาชน

.....4. เตรียมความพร้อมภายใต้ความเปลี่ยนแปลงของเทคโนโลยี เพื่อตอบสนองความต้องการของผู้มีส่วน ได้ส่วนเสียที่มีกฎระเบียบที่หลากหลาย จึงทำให้การออกแบบระบบ Big Data Analytics ต้องใช้ สถาปัตยกรรมด้านไอทีแบบผสมผสานที่สามารถปรับปรุ้งขนาดได้ (Scalability) และมีความยืดหยุ่น (Flexibility) โดยองค์กรภาครัฐที่ทำหน้าที่ดูแลระบบจะต้องเตรียมความพร้อมภายใต้การเปลี่ยนแปลงของ เทคโนโลยีอย่างรวดเร็ว

.....5. เจ้าหน้าที่ภาครัฐจะต้องปรับ Mindset ในการเข้าร่วมกับทุกภาคส่วน ต้องปรับตัวให้องค์กรภาครัฐเป็น ส่วนหนึ่งของระบบนิเวศอย่างกลมกลืน และจะต้องไม่เป็นผู้สร้างอุปสรรคให้เกิดขึ้นในการ สร้างระบบ Big Data Analytics เจ้าหน้าที่ภาครัฐจะต้องปรับ Mindset ในการเข้าร่วมกับทุกภาคส่วนในลักษณะพันธมิตร (Partnership) ให้ได้

.....6. ปรับปรุงวิธีคิดและกระบวนการเพื่อทำให้เกิดการแลกเปลี่ยนข้อมูลและการใช้ข้อมูลร่วมกันระหว่าง หน่วยงานภาครัฐ ภาครัฐจะต้องปรับปรุงวิธีคิดและกระบวนการเพื่อทำให้เกิดการแลกเปลี่ยนข้อมูลและการ ใช้ข้อมูลร่วมกันอย่างมีประสิทธิภาพ โดยไม่หวงข้อมูลว่าเป็นของหน่วยงานตัวเอง ซึ่งการแลกเปลี่ยนข้อมูล และใช้ข้อมูลร่วมกันจะเป็นกุญแจสำคัญที่นำไปสู่การมองเห็นความรู้ใหม่ที่เกิดจากข้อมูลในมิติใหม่และมีผล ต่อการสร้างแหล่งข้อมูล Big Data ได้อย่างรวดเร็ว โดยรัฐบาลจะต้องสร้างสภาพแวดล้อมด้านเทคโนโลยีที่ ปลอดภัยและสร้างกรอบการกำกับดูแลที่เหมาะสมในด้านการแลกเปลี่ยนข้อมูลและใช้ข้อมูลร่วมกัน ซึ่งอาจ วางนโยบาย Open Government เพื่อเปิดเผยข้อมูลภาครัฐให้มากขึ้นเพื่อนำไปสู่การทำ Big Data ให้ ประสบความสำเร็จอย่างแท้จริง

.....7. การกำหนดแนวทางและการบริการให้คำปรึกษาในด้าน Big Data Analytics ให้แก่ทุกภาคส่วน เป็น อีกหนึ่งปัจจัยแห่งความสำเร็จ ซึ่งองค์กรภาครัฐจะต้องมีส่วนร่วมให้มากขึ้นกว่าที่เป็นมาในอดีต ในรูปแบบการ ให้บริการที่ปรึกษา โดยจะต้องสร้างแนวทางและมาตรการเพื่อให้เกิดความร่วมมือ ในการให้ข้อมูลเชิงลึกก็ยิ่ง จะทำให้เกิดประโยชน์จาก Big Data Analytics มีความทรงพลังมากขึ้น



Source: Big Data and BI Best Practices: YellowFin

บทเรียนที่ 7 การจัดการ Big Data ด้วย Hadoop

วิธีที่จะช่วยให้องค์กรเห็นมุมมองแบบ end-to-end เกี่ยวกับ Data Pipeline ของพวกเรา Big Data มีศักยภาพในการสร้างมูลค่าทางธุรกิจเพื่อเพิ่มผลกำไรและสามารถใช้ในการแก้ปัญหาสำคัญๆ ของธุรกิจได้ ในขณะที่เครื่องมือทั้งหมดในระบบนิเวศ (Ecosystem) ได้ผุดขึ้นมาไล่เลี่ยกัน Hadoop เพื่อการวิเคราะห์และจัดการกับข้อมูล แต่เครื่องมือเหล่านั้นกลับมีความเชี่ยวชาญที่น้อยกว่าและช่วยได้เพียงส่วนหนึ่งเท่านั้น หากเป็นกระบวนการขนาดใหญ่ เมื่อองค์กรสามารถใช้ประโยชน์จาก Hadoop ได้อย่างมีประสิทธิภาพจะทำให้เกิดศักยภาพทางธุรกิจและผลกำไรทางด้านไอทีมากเป็นพิเศษเช่นเดียวกับเทคโนโลยีใดๆก็ตามที่มีการพัฒนาอย่างเต็มที่ก็จะทำให้เกิดอุปสรรคต่อการเข้ามาเป็นคู่แข่งรายใหม่ในตลาด โดยเฉพาะในเรื่องของการนำ Hadoop ไปใช้เป็นเครื่องมือในการวิเคราะห์ผลการดำเนินงานตามแนวคิดมูลค่าเพิ่ม (Value-Added Analytics) เพื่อให้ได้ประโยชน์สูงสุดจาก Hadoop ดังนั้นองค์กรต่างๆ จึงจำเป็นต้องทบทวนและทำการตรวจวิเคราะห์เกี่ยวกับ Data Pipeline ของพวกเขาโดยใช้มุมมองแบบ end-to-end ตามวิธีดังต่อไปนี้

..... 1: รับรองการส่งผ่านข้อมูลที่ยืดหยุ่นและปรับขนาดได้

ขั้นตอนแรกของ Data Pipeline ในองค์กรนั้นเกี่ยวข้องกับระบบต้นทางการสื่อสารข้อมูล (Source Systems) และข้อมูลดิบ (Raw Data) โดยจะส่งข้อมูลจากต้นทาง(Ingest)ผสมผสาน (Blended)และวิเคราะห์ข้อมูล (Analyze) การผสมผสานของข้อมูลที่หลากหลายจากโซโล่ทั่วทั้งองค์กรนั้นมักนำไปสู่ข้อมูลเชิงลึกที่สำคัญที่สุด ด้วยเหตุนี้ ความสามารถในการใช้ประโยชน์จากชนิดข้อมูลที่หลากหลายรูปแบบ และแหล่งที่มา จึงเป็นสิ่งจำเป็นที่สำคัญมากในโครงการข้อมูลและการวิเคราะห์ของ Hadoop ซึ่งในปัจจุบันองค์กรไม่เพียงแต่จะต้องเตรียมพร้อมข้อมูลที่พวกเขาวางแผนจะรวมเข้ากับ Hadoop แต่พวกเขายังจำเป็นต้องเตรียมข้อมูลสำหรับการใช้งานอื่นๆ ที่อาจเป็นไปได้ในอนาคตรวมถึงการวางแผนเพื่อลดภาระงานที่ต้องทำเองด้วยตนเอง วางแผนรูปแบบการนำเข้าข้อมูลที่สามารถใช้ซ้ำได้และการออกแบบเวิร์กโฟลว์การนำเข้าข้อมูลแบบไดนามิกที่เป็นส่วนสำคัญของสิ่งนี้

..... 2: ขับเคลื่อนการประมวลผลและการผสมผสานข้อมูลขนาดใหญ่

เมื่อองค์กรประสบความสำเร็จในการดึงข้อมูลอันหลากหลายเข้าสู่ Hadoop ในรูปแบบที่ยืดหยุ่นและปรับขนาดได้แล้วนั้น ขั้นตอนต่อไป ก็จะเป็นไปสู่ขั้นตอนของการประมวลผล (Processing) การแปลงสภาพข้อมูล (Transforming) และการผสมผสานข้อมูล (Data Blending) กับข้อมูลที่มีขนาดใหญ่บนคลัสเตอร์ของ Hadoop การวิเคราะห์ข้อมูลจาก Big data ยังต้องอาศัยหลักการพื้นฐานและเทคนิคบางอย่าง จึงมีระดับของความเป็นนามธรรมที่ต่างจาก Hadoop หรือเครื่องมืออื่นๆ ดังนั้นการบำรุงรักษาและพัฒนาแอปพลิเคชันที่ใช้ข้อมูลจำนวนมากจึงกลายเป็นสิ่งที่ทุกคนสามารถเข้าถึงได้ง่ายกว่า ไม่ได้จำกัดเพียงกลุ่มเล็กๆ ของผู้เชี่ยวชาญด้านการเขียนโค้ดเท่านั้น ในโลกที่ Big Data มีการพัฒนาไปอย่างรวดเร็ว แผนกไอทีก็ยังคงรักษาและออกแบบการแปลงสภาพของข้อมูล (Data Transformation) โดยไม่ต้องกังวลกับการเปลี่ยนแปลงโครงสร้างพื้นฐาน องค์กรควรพยายามหาแนวทางที่เป็นการผสมผสานระหว่างการควบคุม ความสามารถในการมองเห็นข้อมูลในเชิงลึกและการใช้งานที่ง่ายขึ้น แทนที่การใช้กล่องดำ (Black Box) ในการแปลงข้อมูลบน Hadoop

..... 3: ส่งมอบข้อมูลเชิงลึกเกี่ยวกับการวิเคราะห์ Big Data ที่สมบูรณ์

สิ่งสำคัญที่จะช่วยปลดล็อกการวิเคราะห์เพื่อให้ได้มาซึ่งข้อมูลที่เป็นประโยชน์สูงสุดจาก Hadoop นั่นก็คือการพิจารณาอย่างรอบคอบเกี่ยวกับกระบวนการทางธุรกิจแอปพลิเคชันและผู้ใช้ปลายทาง โดยข้อมูลผู้บริโภคที่แตกต่างกันนั้นก็อาจจำเป็นที่จะต้องใช้เครื่องมือและวิธีการที่แตกต่างกัน ทั้งนี้ก็ขึ้นอยู่กับว่าพวกเขานั้นต้องการข้อมูลอะไร แผนต่างๆที่พวกเขาวางไว้สำหรับข้อมูลเหล่านั้น และมีระดับความซับซ้อนแค่ไหน ในขณะที่นักวิทยาศาสตร์ด้านข้อมูลและนักวิเคราะห์ที่มีความเชี่ยวชาญขั้นสูง พวกเขาเริ่มทำการสืบค้นและสำรวจชุดข้อมูลใน Hadoop โดยพวกเขามักจะใช้คลังข้อมูล (Data Warehouse) และเทคโนโลยีประมวลผลเลเยอร์คล้าย SQL (SQL-like) เช่น Hive (ที่ทำหน้าที่ในการแปล SQL like ให้มาเป็น MapReduce) และ

Impala (ซึ่งเป็นเครื่องมือที่คล้ายกับ Hive) นับว่าโชคดีที่เครื่องมือเหล่านี้ใช้เวลาในการเรียนรู้ไม่นาน เนื่องจากมันเป็นภาษาที่ใช้สำหรับค้นหาข้อมูลจากฐานข้อมูลหรือ Query Language (QL) ที่คุ้นเคย ฐานข้อมูล NoSQL ที่มีประสิทธิภาพสูงและปรับขนาดได้นั้นมีการใช้งานควบคู่กับ Hadoop มากขึ้น การรวบรวม Big Data จากเว็บมือถือและปริมาณงานของ IoT นั้นได้ถูกสร้างในระบบ NoSQL ก่อนที่จะถูกส่งเข้าสู่ซอฟต์แวร์ Hadoop ในทางกลับกัน ปริมาณงานแบบดัดและสตรีมมิ่งการวิเคราะห์ที่ประมวลผลโดย Hadoop ก็สามารถแชร์กับระบบ NoSQL ได้ การเพิ่มขึ้นของระบบ NoSQL ที่เข้าคู่กับมูลค่าของ Big Data ทำให้องค์กรต่างๆ เริ่มแสวงหาผู้เชี่ยวชาญด้านไอทีที่มีทักษะเกี่ยวกับ NoSQL และ Hadoop เพื่อให้ได้ประโยชน์สูงสุดจาก Big Data ของพวกเขา ดังนั้นการพิจารณาให้ Hadoop เป็นส่วนในการขยาย Analytic Pipeline จึงเป็นสิ่งสำคัญมาก ธุรกิจจำนวนมากคงคุ้นเคยกับฐานข้อมูลประสิทธิภาพสูงที่ปรับให้เหมาะสำหรับการวิเคราะห์ที่ใช้ปลายทางหรือที่รู้จักกันทั่วไปว่า 'ฐานข้อมูลการวิเคราะห์' (Analytic Databases) องค์กรเหล่านี้ได้พบว่าการส่งมอบชุดข้อมูลที่ได้รับการแก้ไขจาก Hadoop ไปยังฐานข้อมูลที่กำลังพัฒนานั้น เป็นวิธีที่มีประสิทธิภาพสูงในการปลดปล่อยพลังการประมวลผลของ Hadoop

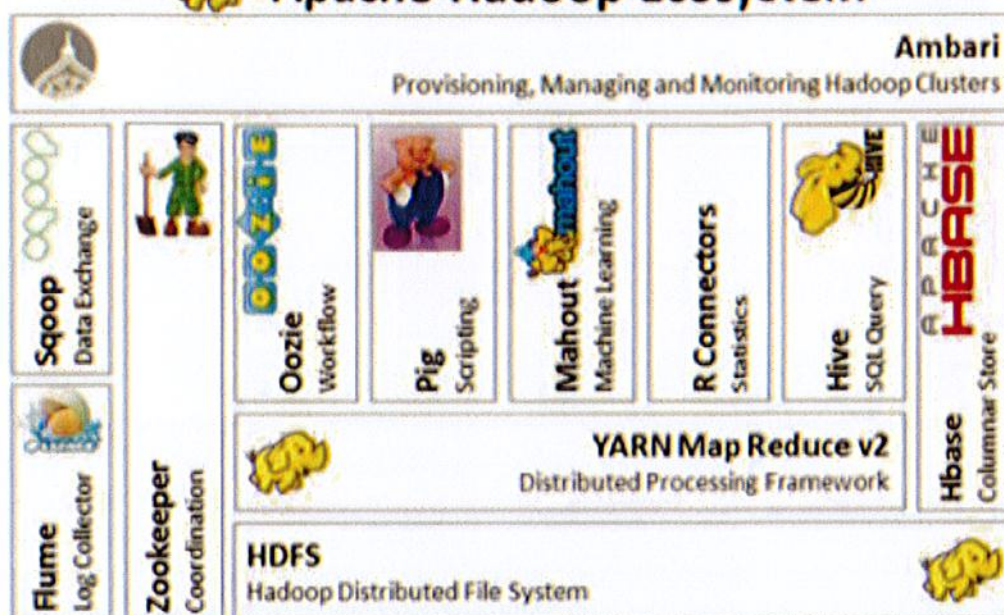
..... 4. ใช้แนวทางเชิงโซลูชัน (Solution-Oriented)

..... ในช่วงไม่กี่ปีที่ผ่านมา ความก้าวหน้าหลายๆ อย่างได้ถูกสร้างขึ้นในระบบนิเวศ (Ecosystem) Hadoop และยังคงเติบโตในฐานะของแพลตฟอร์มเพื่อการปรับใช้กับองค์กรการผลิต (Production Enterprise) ในส่วนของข้อกำหนดสำหรับความคิดริเริ่มด้านเทคโนโลยีนั้น มีแนวโน้มที่จะได้รับการพัฒนาและอยู่ในช่วงของการ "กำลังดำเนินการ" ซึ่ง Hadoop ก็ได้แสดงให้เห็นถึงองค์ประกอบใหม่ใน Data Pipeline ที่กว้างขึ้น ผลที่ตามมาคือความคิดริเริ่มที่เกี่ยวข้องใหม่นั้นมักจะมียุทธวิธีแบบค่อยเป็นค่อยไป (Phased Approach) เมื่อคำนึงถึงเรื่องนี้แล้ว ผู้ประเมินซอฟต์แวร์ต่างก็รู้ว่าพวกเขาไม่สามารถพบกับอุปกรณ์ที่มีทุกอย่างครบครัน (Off-the-shelf) เพื่อตอบสนองความต้องการด้านข้อมูลและการวิเคราะห์ของ Hadoop ได้ ทั้งในปัจจุบันและในอนาคตข้างหน้า ซึ่งในปัจจุบันนี้คำว่า "รองรับอนาคต" (Future-proof) ได้ถูกนำมาใช้มากขึ้นในเรื่องที่เกี่ยวกับ Big Data แต่ยังมีสิ่งที่จะต้องคำนึงถึงนั่นก็คือ ความสามารถในการขยายระบบและความยืดหยุ่นเพื่อรองรับการทำงานที่สูงขึ้น ซึ่งทั้งสองอย่างนี้ควรจะเป็นส่วนสำคัญของแบบสำรวจรายการ (Checklist) ของโครงการทั้งหมด การแปลงพอร์ตเพื่อให้ทำงานได้อย่างราบรื่นผ่านการกระจาย Hadoop ที่แตกต่างกันเป็นจุดเริ่มต้นของศักยภาพที่ดี แต่ความยั่งยืนที่แท้จริงนั้น องค์กรจำเป็นต้องมีแพลตฟอร์มเพื่อความยืดหยุ่นอันสอดคล้องกับนวัตกรรมแบบเปิดที่ขับเคลื่อนระบบนิเวศ (Ecosystem) Hadoop

..... 5: เลือกผู้ขายที่เหมาะสม

..... กระแสของ Big Data ส่งผลให้ผู้ให้บริการโซลูชันหลังไหลเข้าพื้นที่ตลาดกันอย่างท่วมท้น แพคเกจที่พวกเขาเสนอนั้นอาจแตกต่างกันอย่างหลากหลาย ซึ่งมีตั้งแต่เครื่องมือทางสถิติอย่างง่ายไปจนถึงแอปพลิเคชันการเรียนรู้เครื่องจักรขั้นสูง องค์กรจึงควรระบุประเภทข้อมูลที่พวกเขาจะประมวลผลเพื่อเลือกเทคโนโลยีที่รองรับพวกเขา แพลตฟอร์มที่น่าพึงพอใจนั้นจำเป็นต้องเข้ากันได้ดีกับเครื่องมือวิเคราะห์ที่มีอยู่ โดยให้การเข้าถึงกับพนักงานที่พวกเขาต้องการและมีความยุ่งยากน้อยที่สุดต่อเวิร์กโฟลว์ ผู้ให้บริการ NoSQL และ Hadoop บางรายกำลังร่วมมือกันจัดทำข้อเสนอที่ครอบคลุม รวบรวมระบบของพวกเขาเพื่อเพิ่มความคล่องตัวในการเคลื่อนไหวระหว่างระบบโครงสร้างและซอฟต์แวร์ สิ่งนี้ยังช่วยลดความยุ่งยากให้กับลูกค้าเนื่องจากพวกเขาสามารถจัดการกับทุกอย่างได้ในจุดเดียว

Apache Hadoop Ecosystem



The Apache Hadoop project

2.2 ประสพการณ์/ประโยชน์ที่ได้รับ/การประยุกต์ใช้กับหน่วยงาน

☒ ต่อตนเอง

ทำให้สามารถเรียนรู้และเข้าใจระบบนิเวศ (Ecosystem) Hadoop นำไปทำ Data Analytics ใน Artificial Intelligence การประยุกต์ใช้อย่างกว้างขวาง เช่น การมองเห็นของคอมพิวเตอร์ (Computer Vision) การประมวลผลภาษาธรรมชาติ (Natural Language Processing) การประมวลผลภาษาธรรมชาติ (Natural Language Processing) เป็นต้น.....

☒ ต่อหน่วยงาน / การนำมาประยุกต์ใช้กับหน่วยงาน

Hadoop รองรับ Big Data การวิเคราะห์การเรียนรู้เสริมกำลังเชิงลึก (DRL: Deep Reinforcement Learning) เป็นสาขาย่อยขั้นสูงของปัญญาประดิษฐ์ (AI) และ Machine Learning (ML) ที่ผสมผสานเทคนิค Deep Learning หรือ Recurrent Neural Network (RNN) สามารถแก้โจทย์ที่ไม่ใช่คนแยกแยะสิ่งมีชีวิต.....

2.3 ปัญหาและอุปสรรคในการอบรม/สัมมนา/พัฒนาความรู้

- การใช้เครื่องมือเป็นการเรียนรู้โปรแกรมใหม่ ๆ จึงต้องมีการศึกษาเพิ่มเติมในโปรแกรมที่ใช้.....
 - การวิเคราะห์ การประยุกต์ใช้ Hadoop ในการทำงานแบบ AI deep learning ยังต้องใช้เวลาในการศึกษา อีกมาก.....

2.4 ข้อคิดเห็นและข้อเสนอแนะ

- ควรพัฒนาฐานข้อมูล Hadoop และเครื่องมือ - AI deep learning ให้มากขึ้นในการอำนวยความสะดวกให้แก่ผู้ปฏิบัติงาน รวมทั้งตอบโจทย์เชิงนโยบายได้ อย่างรวดเร็ว....
 - ควรมีเวิร์กช็อปกลางในการทำงาน GPU มากพอทำระบบแบบจำลองได้ เตรียมความพร้อมในงานวิจัย รองรับงานวางแผนการใช้ที่ดิน.....
 - ควรสร้างอัตลักษณ์ งานด้าน Big Data กับ AI ของกลุ่ม กอง กรม ให้เป็นที่รู้จักในระบบออนไลน์ ผ่านการนำเสนอผลงาน และข้อมูลมีมาตรฐาน

ลงชื่อ.....

(...นายดิเรก...คงแพ...)

ตำแหน่ง...นักวิเคราะห์นโยบายและแผนชำนาญการพิเศษ.....

ผู้รายงาน

วันที่...๑๑...เดือน...ก.พ...พ.ศ. ...๖๘

ส่วนที่ 3 ความเห็นของผู้บังคับบัญชา

(✓) ทราบ

ลงชื่อ.....

(..นายปราบพล โตหิวัน..)

ตำแหน่ง...ผู้อำนวยการกองนโยบายและแผนการใช้ที่ดิน.....

วันที่...๒๕...เดือน...กุมภาพันธ์...พ.ศ. ...๒๕๖๓

ประกาศนียบัตร

ให้ไว้เพื่อแสดงว่า

ดิเรก คงแพ

ได้ผ่านการอบรมด้วยระบบการเรียนออนไลน์ในบทเรียน
ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการ
และบุคลากรภาครัฐทุกระดับ

รวมระยะเวลาทั้งสิ้น 1:0 ชั่วโมง

โดยสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล
ภายใต้การดำเนินงานของสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
ให้ไว้ ณ วันที่ 11 ก.พ. 2568

A.L.

(นางไอรดา เหลืองวิไล)

รองผู้อำนวยการสำนักงานพัฒนารัฐบาลดิจิทัล

สำเนาถูกต้อง

[Signature]

รักษาการแทนผู้อำนวยการสถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล



afbd03dc

Signed by สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) (สพร.)

Date: 2025-02-11T20:36:04.548+07:00

Reason: Confirm Certificate