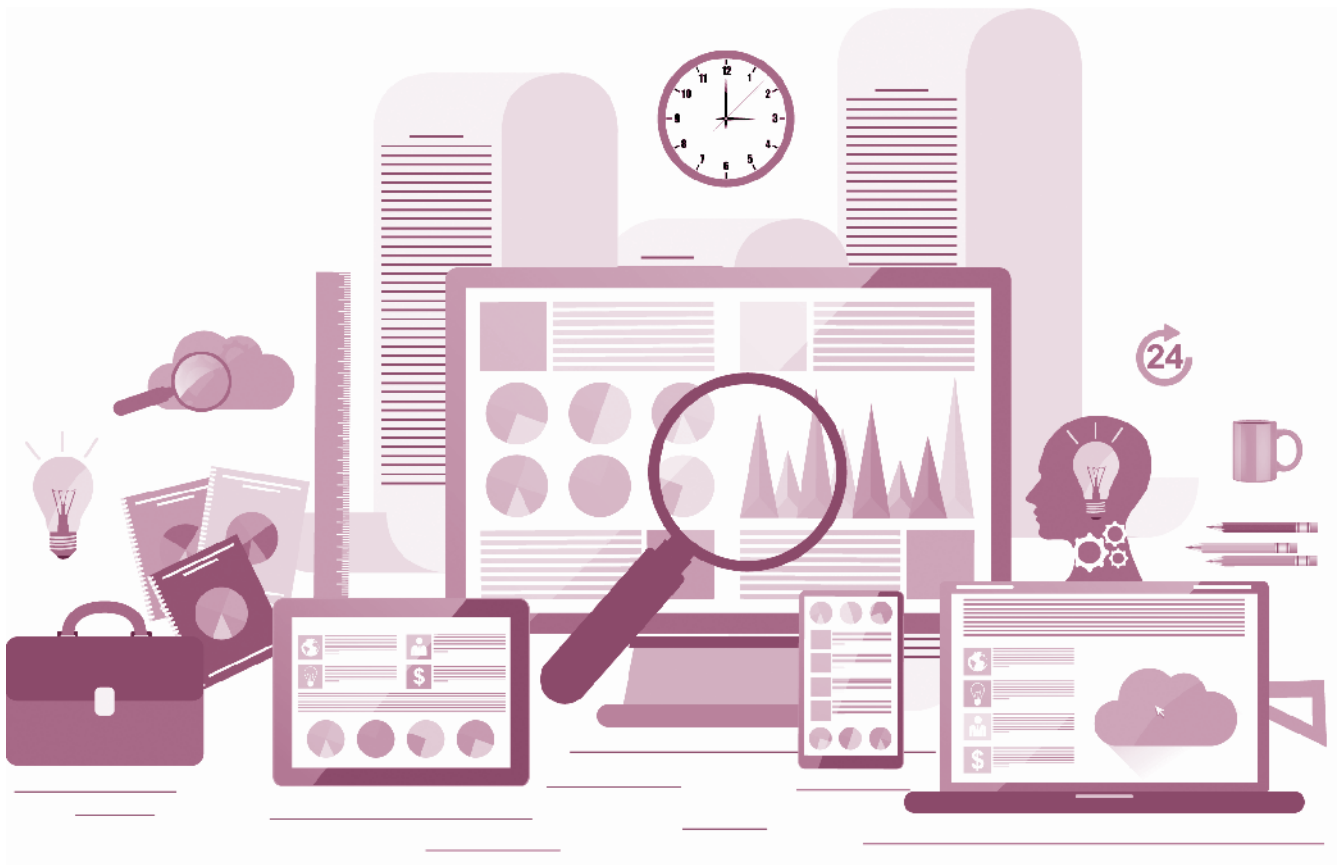




เตรียมรับมือ Big Data Crisis ด้วย Data Management

ต่อ จากฉบับที่แล้ว

คุณสมบัติของข้อมูลที่ดี

1. **ถูกต้องแม่นยำ (accuracy)** ข้อมูลที่ดีควรมีความถูกต้องแม่นยำสูง หรือถ้ามีความคลาดเคลื่อน (errors) อยู่บ้างก็ควรที่จะสามารถควบคุมให้มีความคลาดเคลื่อนน้อยที่สุดหรือในระดับที่ยอมรับได้

2. **ทันเวลา (timeliness)** เป็นข้อมูลที่ทันสมัย (up to date) และทันต่อความต้องการของผู้ใช้ ถ้าเป็นข้อมูลที่มีความล่าช้า ก็อาจไม่มีคุณค่าถึงแม้จะเป็นข้อมูลที่ถูกต้องแม่นยำก็ตาม

3. **สมบูรณ์ครบถ้วน (completeness)** ข้อมูลที่เก็บรวบรวมมาต้องเป็นข้อมูลที่ให้ข้อเท็จจริง (facts) หรือข่าวสาร (information) ที่ครบถ้วนทุกด้านทุกประการ มิใช่ขาดส่วนหนึ่งส่วนใดไปทำให้นำไปใช้ประโยชน์ไม่ได้เนื่องจากอาจมีการตีความผิดพลาดหรือบิดเบือนจากข้อเท็จจริงบางประการ

4. **กะทัดรัด (conciseness)** ข้อมูลที่ได้รับส่วนใหญ่จะกระจัดกระจาย ควรจัดข้อมูลให้อยู่ในรูปแบบที่กระชับสะดวกต่อการใช้งาน และค้นหา ซึ่งผู้ใช้งานสามารถทำความเข้าใจได้ในทันที

5. **ตรงกับความต้องการของผู้ใช้ (relevance)** ควรเป็นข้อมูลที่ผู้ใช้ข้อมูลต้องการใช้งาน และจำเป็นต้องรู้ หรือรับทราบ หรือเป็นประโยชน์ต่อการจัดทำแผน การกำหนดนโยบายหรือตัดสินใจในเรื่องนั้นๆ มิใช่เป็นข้อมูลปริมาณมากแต่ไม่มีใครต้องการใช้หรือไม่ตรงกับความต้องการของผู้ใช้ข้อมูล



วิรัชฤทธิ์ เมาเรพงษ์

ที่ปรึกษาโครงการสารสนเทศบนหน่วยงานภาครัฐ
สังกัดสถาบันวิจัยและให้คำปรึกษา
แห่งมหาวิทยาลัยธรรมศาสตร์

นอกจากนี้ในการจัดเก็บข้อมูล ควรอย่างยิ่งที่จะต้องดำเนินการอย่างสม่ำเสมอ และต่อเนื่องในลักษณะของอนุกรมเวลา (time-series) เพื่อจะได้นำไปใช้ประโยชน์ในด้านการวิเคราะห์หรือหาแนวโน้มในอนาคตได้อย่างมีประสิทธิภาพ

แหล่งข้อมูล

แหล่งข้อมูล หมายถึง สถานที่หรือแหล่งกำเนิดข้อมูล ซึ่งจะแตกต่างกันไปตามประเภทของข้อมูล ข้อมูลบางอย่างเราอาจจะนำมาจากแหล่งข้อมูลหลายแหล่ง ซึ่งโดยปกติแล้ว ข้อมูลสำหรับการนำมาประมวลผลนั้นจะได้มาจาก 2 แหล่งที่มาด้วยกัน คือ

1. แหล่งข้อมูลภายใน เป็นแหล่งข้อมูลที่อยู่ภายในองค์การ ข้อมูลที่ได้นั้นอาจได้มาจากการปฏิบัติงานของบุคลากรหรือมีอยู่แล้วในองค์การ เช่น ยอดขายประจำปี ข้อมูลผู้ถือหุ้น รายงานกำไรขาดทุน ข้อมูลเหล่านี้จะเกี่ยวข้องกับข้อเท็จจริงต่างๆ อาจเป็นข้อมูลที่สามารถเปิดเผยให้กับบุคคลภายนอกหรือรับทราบหรือไม่ก็ได้ หากข้อมูลนั้นเป็นข้อมูลที่เกี่ยวข้องกับการดำเนินงานหลักขององค์การ และมีความสำคัญมาก องค์การอาจมีกลไกในการบริหารจัดการเพื่อรักษาความลับ และป้องกันการรั่วไหลของข้อมูล

2. แหล่งข้อมูลภายนอก เป็นแหล่งข้อมูลที่อยู่ภายนอกองค์การ โดยทั่วไปแล้วสามารถนำข้อมูลต่างๆ เหล่านั้นมาใช้ประโยชน์ในองค์การในบริบทต่างๆ ได้ เช่น ข้อมูลลูกค้า เจ้าหนี้ อัตราดอกเบี้ยสถาบันการเงิน กฎหมาย และอัตราภาษีของรัฐบาล ข้อมูลขององค์การหรือหน่วยงานที่เกี่ยวข้อง ข้อมูลจากสื่อสังคมออนไลน์ เป็นต้น

การรวบรวม และตรวจสอบข้อมูล

การรวบรวม และตรวจสอบข้อมูล เป็นขั้นตอนที่มีความสำคัญประกอบด้วย

1. การเก็บรวบรวมข้อมูล เป็นเรื่องของการเก็บรวบรวมข้อมูลซึ่งมีจำนวนมาก และต้องเก็บให้ได้อย่างทันเวลา ปัจจุบันมีการใช้เทคโนโลยีที่ช่วยให้การจัดเก็บข้อมูลทำได้อย่างรวดเร็ว อาทิ การป้อนข้อมูลเข้าเครื่องคอมพิวเตอร์ การอ่านข้อมูลจาก Barcode, QR Code, RFID, Image Processing, Biometric ฯลฯ

2. การตรวจสอบข้อมูล เมื่อมีการเก็บรวบรวมข้อมูลแล้ว ก็จำเป็นต้องมีการตรวจสอบความถูกต้องของข้อมูลที่จัดเก็บเข้าสู่ระบบคอมพิวเตอร์ ข้อมูลจะต้องเชื่อถือได้ หากพบที่ผิดพลาดต้องได้รับการแก้ไข การตรวจสอบข้อมูลมีหลายวิธี อาทิ การเปรียบเทียบโครงสร้าง หรือรูปแบบเฉพาะของที่ต้องการจัดเก็บ เป็นต้น

การบริหารจัดการข้อมูลด้วยแฟ้มข้อมูล

การบริหารจัดการข้อมูลด้วยแฟ้มข้อมูลเป็นรูปแบบของการบริหารจัดการข้อมูลที่ได้รับคามนิยมจากองค์การต่างๆ ตั้งแต่อดีตก่อนที่จะมีการพัฒนาระบบฐานข้อมูล จนกระทั่งปัจจุบันการบริหารจัดการข้อมูลในรูปแบบนี้ยังถูกนำมาใช้งานอย่างต่อเนื่องในการปฏิบัติ และประสานการทำงานภายในองค์การระหว่างบุคคล และระหว่างแผนก โดยมีการเปลี่ยนรูปแบบจากการใช้งานเพื่อสร้างหรือนำเข้าข้อมูลเป็นไฟล์อิเล็กทรอนิกส์ด้วยโปรแกรมจัดการเอกสาร และสื่อสำเร็จรูป อาทิ Microsoft Office Open office เป็นต้น ซึ่งมีการบันทึกไฟล์ข้อมูลด้วยนามสกุลที่ได้รับการยอมรับเป็นมาตรฐานสากล



จัดเก็บในรูปแบบของแฟ้มเอกสารด้วยระบบสารสนเทศเฉพาะด้าน หรือบนระบบจัดการไฟล์เอกสารของระบบปฏิบัติการ (Windows, Mac OS, Linux, Android, iOS, etc.) หรือแชร์ใช้งานร่วมกันบน Cloud Services อาทิ Google Drive, DropBox, SkyDrive, iCloud เป็นต้น โดยการบริหารจัดการข้อมูลด้วยแฟ้มข้อมูล มีองค์ประกอบพื้นฐานทางเทคนิค ดังนี้

แฟ้มข้อมูล (file) หมายถึง ข้อมูลที่จัดเก็บไว้ในสื่ออิเล็กทรอนิกส์ โดยข้อมูลจะถูกจัดเก็บในแฟ้มอิเล็กทรอนิกส์ที่มีการกำหนดชื่อแฟ้มจำแนกตามหมวด หรือเรื่อง อย่างเป็นระบบไม่มีความซ้ำซ้อนกัน

โครงสร้างแฟ้มข้อมูล (data structure) หมายถึงรูปแบบของการจัดระเบียบข้อมูล ประกอบด้วยโครงสร้างพื้นฐานที่เรียงลำดับจากหน่วยของข้อมูลที่เล็กที่สุดไปยังหน่วยที่ใหญ่ขึ้นตามลำดับดังต่อไปนี้

■ **บิต (Bit : Binary Digit)** คือหน่วยของข้อมูลที่เล็กที่สุดที่เก็บอยู่ในหน่วยความจำภายในคอมพิวเตอร์ ซึ่ง Bit จะแทนด้วยตัวเลขหนึ่งตัว คือ 0 หรือ 1 อย่างใดอย่างหนึ่ง เรียกตัวเลข 0 หรือ 1 ว่าเป็น บิต 1 บิต

■ **ไบต์ (Byte)** คือหน่วยของข้อมูลที่น่าบิตหลายๆ บิตมารวมกัน แทนตัวอักษรแต่ละตัว เช่น A, B, ..., Z, 0, 1, 2, ... ,9 และสัญลักษณ์พิเศษอื่นๆ เช่น \$, &, +, -, *, / ฯลฯ โดยตัวอักษร 1 ตัวจะแทนด้วยบิต 7 บิต หรือ 8 บิต ซึ่งตัวอักษรแต่ละตัวจะเรียกว่า ไบท์ เช่น ตัวอักษร A เมื่อเก็บอยู่ในคอมพิวเตอร์จะเก็บเป็น 1000001 ส่วนตัว B อักษรจะ จะเก็บเป็น 1000010 เป็นต้น

■ **เขตข้อมูล (Field)** คือ หน่วยของข้อมูลที่เกิดจากการนำตัวอักษรหลายๆ ตัวมารวมกันเป็นคำที่มีความหมาย เช่น รหัสลูกค้า ชื่อลูกค้า นามสกุล ที่อยู่ เบอร์ติดต่อ เป็นต้น

■ **ระเบียบ (Record)** คือ หน่วยของข้อมูลที่มีการนำเขตข้อมูลหลายๆ เขตข้อมูลที่มีความสัมพันธ์กันมารวมกัน หรือค่าของข้อมูลในแต่ละเขตข้อมูล เช่น ระเบียบของลูกค้า ประกอบด้วยเขตข้อมูลรหัสลูกค้า ชื่อ นามสกุล ที่อยู่ เบอร์ติดต่อ เป็นต้น

■ **แฟ้มข้อมูล (File)** คือ หน่วยของข้อมูลที่มีการนำระเบียบหลายๆ ระเบียบที่มีความสัมพันธ์กันมารวมกัน เช่น แฟ้มข้อมูลลูกค้า เป็นต้น

■ **ฐานข้อมูล (Database)** คือ หน่วยของข้อมูลที่มีการนำแฟ้มข้อมูลหลายๆ แฟ้มข้อมูลที่มีความสัมพันธ์กันมารวมกัน เช่น ฐานข้อมูลในระบบทะเบียนลูกค้าจะประกอบด้วยแฟ้มข้อมูลลูกค้า ผลิตภัณฑ์ การสั่งซื้อ เป็นต้น

ประเภทของแฟ้มข้อมูล (File)

แฟ้มข้อมูลจะถูกแบ่งแยกประเภทตามการใช้งาน ซึ่งแบ่งออกเป็น 5 ประเภทต่างๆ ดังนี้

■ **แฟ้มข้อมูลรายการหลัก (Master File)** ทำหน้าที่จัดเก็บข้อมูลที่ไม่มีการเปลี่ยนแปลงหรือมีสภาพค่อนข้างคงที่ เช่นแฟ้ม



ข้อมูลลูกค้า ซึ่งการปรับปรุงแก้ไขข้อมูลใน Master File ให้ทันสมัยสามารถทำได้ 3 รูปแบบ คือการเพิ่ม (add) การลบออก (delete) และการแก้ไข (modify) เช่น การเพิ่มระเบียบของลูกค้าในกรณีที่เป็นลูกค้าใหม่ การแก้ไขระเบียบของลูกค้าในกรณีที่ลูกค้ามีการเปลี่ยนแปลงที่อยู่ เป็นต้น

■ **แฟ้มข้อมูลรายการเปลี่ยนแปลง (Transaction File)** ทำหน้าที่จัดเก็บข้อมูลที่มีการเคลื่อนไหวหรือมีการเปลี่ยนแปลงอยู่เสมอ เช่น แฟ้มข้อมูลการสั่งซื้อสินค้าที่มีการเพิ่มเติมเปลี่ยนแปลงอยู่อย่างสม่ำเสมอ เป็นต้น ซึ่งมีความเชื่อมโยงกับ แฟ้มข้อมูลลูกค้า

■ **แฟ้มข้อมูลรายงาน (Report File)** ทำหน้าที่เก็บรายงานที่ได้จากการประมวลผลผ่านระบบคอมพิวเตอร์ ซึ่งจัดเก็บได้สะดวกและดีกว่าการเก็บเป็นเอกสารกระดาษอีก รวมทั้งสามารถสั่งพิมพ์เมื่อใดก็ได้ตามที่ต้องการ

■ **แฟ้มข้อมูลเก็บผลลัพธ์ (Output File)** โปรแกรมคอมพิวเตอร์จะมีการรับข้อมูลเข้ามาประมวลผล และได้ผลลัพธ์เป็นข้อมูลใหม่ออกมา ข้อมูลใหม่อาจแสดงออกทางหน่วยแสดงผลหรือจัดเก็บไว้ในแฟ้มข้อมูลก็ได้ ซึ่งจะเรียกแฟ้มข้อมูลที่เกิดข้อมูลใหม่นี้ว่าแฟ้มข้อมูลเก็บผลลัพธ์ สามารถนำแฟ้มข้อมูลนี้ไปเป็นข้อมูลนำเข้าให้กับโปรแกรมอื่นๆ ได้ตามลำดับขั้นตอนของการประมวลผลที่ต้องการต้องการใช้งาน

■ **แฟ้มข้อมูลสำรอง (Backup File)** ใช้จัดเก็บสำรองข้อมูลในแฟ้มข้อมูลที่มีความสำคัญ การสำรองข้อมูลเป็นสิ่งสำคัญมากในการใช้งานระบบคอมพิวเตอร์ เนื่องจากสื่อที่ใช้จัดเก็บข้อมูลต่างๆ อาจเกิดปัญหาได้โดยที่ผู้ใช้คาดไม่ถึง ดังนั้นควรจัดเก็บข้อมูลลงบนสื่อบันทึกข้อมูลอื่นๆ เพื่อเป็นการสำรองข้อมูลในกรณีที่ปัญหาซึ่งจะช่วยสามารถนำข้อมูลที่สูญหายกลับมาใช้ใหม่ได้