

การพัฒนา Extract Transform Loader สำหรับภาษีอิเล็กทรอนิกส์

Extract Transform Loader System Development for e-Tax Invoices

วาสนา แซ่โก้ย* สถาพร เทพสัมฤทธิ์พร ชนิกานต์ รอดมรรณ

ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

(*s6404053630211@email.kmutnb.ac.th)

บทคัดย่อ

งานวิจัยนี้มีมุ่งเน้นการพัฒนาและปรับปรุงกระบวนการ Extract-Transform-Load (ETL) สำหรับเอกสารภาษีอิเล็กทรอนิกส์ประเภท T02 (ใบแจ้งหนี้/ใบกำกับภาษี) โดยทำการออกแบบและพัฒนา Workflow รวมถึงพัฒนาระบบตรวจสอบความถูกต้องของข้อมูลก่อนการบันทึกลงฐานข้อมูล เครื่องมือที่ใช้ในการพัฒนาระบบประกอบด้วย 1) Pentaho Data Integration (PDI) ใช้สำหรับกระบวนการ ETL 2) PostgreSQL เป็นฐานข้อมูลหลัก 3) PgAdmin และ DataGrip ใช้ในการจัดการฐานข้อมูล 4) GitHub ใช้สำหรับบริหารจัดการซอร์สโค้ด และ 5) Visual Studio Code ใช้สำหรับการพัฒนาและปรับแต่งคำสั่ง ข้อมูลที่ใช้คือข้อมูลเอกสารภาษีอิเล็กทรอนิกส์ประเภท T02 ตั้งแต่ปี 2565 ถึง 2567 จำนวน 1,000 ฉบับ ทดสอบกับฐานข้อมูลที่ได้รับการออกแบบใหม่ ที่ได้รับรวบรวมมาจาก 12 แหล่ง ผลการพัฒนาคือ ระบบสามารถลดข้อผิดพลาดในการนำเข้าข้อมูลและการแปลงไฟล์ XML ได้ สามารถตรวจจับข้อผิดพลาดที่อาจเกิดขึ้นระหว่างการประมวลผลและแจ้งเตือนผู้ใช้ได้ทันที เช่น การตรวจจับรูปแบบเลขที่ใบแจ้งหนี้ที่ไม่ถูกต้อง หรือมีข้อมูลลูกค้าไม่ตรงกับฐานข้อมูลเดิม โดยผลการทดสอบแสดงให้เห็นว่าระบบมีความแม่นยำในการประมวลผลข้อมูลคิดเป็นร้อยละ 80 นอกจากนี้ การพัฒนาโครงสร้างฐานข้อมูลและการปรับปรุงกระบวนการทำงานดังกล่าว ทำให้ระบบสามารถรองรับการนำเข้าและจัดเก็บข้อมูลภาษีอิเล็กทรอนิกส์จำนวนมากได้ รวมไปถึงสามารถรองรับการขยายตัวของข้อมูลในอนาคต และรองรับการนำเข้าและส่งออกข้อมูลให้ถูกต้อง ครบถ้วนตามข้อกำหนดของกรมสรรพากร

คำสำคัญ: Extract-Transform-Load (ETL), Pentaho, e-Tax, ใบกำกับภาษีอิเล็กทรอนิกส์, T02 (ใบแจ้งหนี้/ใบกำกับภาษี)