

การจัดการข้อมูลด้วย ETL using Pentaho – PDI (OpenSource)

Connect to Generative AI (ChatGPT or Gemini)

DSA017

ฉบับแก้ไข วันที่ 12 กุมภาพันธ์ 2568

หลักการและเหตุผล:

การจัดการข้อมูลที่มีอยู่หลายประเภท และหลายระบบ ทำให้การรวบรวมข้อมูลต่าง ๆ ที่มีความหลากหลายต้องวิธีในการรวบรวม โดยส่วนใหญ่จะใช้วิธีการเขียนโปรแกรมเพื่อจัดการกับข้อมูล หรือใช้เครื่องมือภายในระบบ Database ช่วยในการจัดการซึ่งอาจจะมีข้อจำกัด โปรแกรมPDI เป็น ETL โปรแกรมหนึ่งที่สามารถเข้ามาช่วยในการทำงานและการจัดการกับข้อมูลเหล่านั้นได้อย่างมีประสิทธิภาพตัวหนึ่ง ในงานทางด้าน Data warehouse, Big Data

วัตถุประสงค์:

- เรียนรู้จัดการข้อมูลหลายประเภท
- เรียนรู้เทคนิคในการใช้เครื่องมือในการทำงานกับข้อมูล

หลักสูตรนี้เหมาะสำหรับ:

- Programmer
- นักวิเคราะห์ระบบ
- ผู้ที่สนใจทั่วไป

ความรู้พื้นฐาน:

- มีความรู้ความสนใจในระบบฐานข้อมูล
- มีความรู้ข้อมูลหลายประเภท
- มีการใช้งานข้อมูลหลากหลายลักษณะ

เนื้อหาหลักสูตร:

- | | |
|--|---|
| <ul style="list-style-type: none">• Introduction Pentaho Data Integration (PDI)• Pentaho Data Integration (PDI) Installation & Setup• Setting up environment and installing PDI• Opening Spoon - The Graphical UI• A Simple ETL Demonstration• The example problem statement• Demonstration of a PDI transformation and Job• Basic concepts - Theory for foundational understanding• What is ETL?• Data Warehouse, Ops Database and Data mart• ETL vs ELT• The ETL process: The practical part begins here• Data and the ETL process• DATA EXTRACTION: Extracting tabular data• Manually entering data into PDI• Inputting Data from a TXT (text) file• Input from multiple CSV files at the same time• Inputting Data from an Excel file• Extracting Data from Zipped files• Extracting from XML, Extracting from JSON• Extracting from an SQL table• Creating Sales table in SQL• Extracting from an SQL table• Storing and Retrieving Data from Cloud storage• Merging Data Streams | <ul style="list-style-type: none">• Replace in String Step• Time to check your understanding• Data Validation• Common Data Validation steps• Error Handling• Correcting the errors & merging with main stream• Time to check your understanding• Writing the errors to the log• Time to check your understanding• Writing the errors to a separate file• Time to check your understanding• Transformation and Analytics steps• Concatenating Address Fields• Data Aggregation using Group-by• Normalization and Denormalization• Number Range Step• PDI SQL Connection• Introduction to PDI - SQL connection• Reading and filtering data from DB into PDI• Updating and Inserting data into DB from PDI• Deleting data from SQL DB using PDI• Conceptual understanding for Loading Data• Facts and Dimensions tables• Time to check your understanding• Loading the data into a Data Mart• Running Java and Javascript• Scripting Steps |
|--|---|

- | | |
|---|--|
| <ul style="list-style-type: none"> • Time to check your understanding • Data Cleansing • Introduction to Data Cleansing • Value Mapper Step | <ul style="list-style-type: none"> • PDI Jobs vs Transformation • Scheduling a job for production environment • Metadata injection • Alternative to Pentah |
|---|--|

ค่าลงทะเบียนอบรม:	12,000 บาท (ไม่รวมภาษีมูลค่าเพิ่ม)
จำนวนชั่วโมงในการฝึกอบรม:	3 วัน (18 ชั่วโมง)
ช่วงเวลาฝึกอบรม:	9.00 - 16.00 น.
กำหนดการอบรม:	ตามตารางปฏิทินอบรมประจำปี https://www.career4future.com/trainingprogram

วิทยากร:



อาจารย์วิเชียร ยงรัตนกิจ

- วิทยากรรับเชิญ ประจำสถาบันพัฒนาบุคลากรแห่งอนาคต
- Microsoft Certified Professional (MCP)

หมายเหตุ

- สถาบันฯ มีการจัดเตรียมเครื่องคอมพิวเตอร์ (ในหลักสูตรที่ต้องใช้โปรแกรมคอมพิวเตอร์) เอกสารการอบรม อาหารว่าง และอาหารกลางวัน ให้กับผู้เข้าอบรม
- สถานที่อบรม ห้องอบรม ณ สถาบันพัฒนาบุคลากรแห่งอนาคต อาคาร สวทช. ชั้น 6 ถนนพระรามที่ 6 แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพฯ 10400
- เฉพาะหน่วยงานภาครัฐ และองค์กรของรัฐ ที่ไม่ใช่ธุรกิจและไม่แสวงหากำไร จะได้รับการยกเว้น ภาษีมูลค่าเพิ่ม
- สถาบันฯ เป็นหน่วยงานราชการ ได้รับการยกเว้นไม่ต้องหักภาษี ณ ที่จ่าย 3%
- ผู้เข้าร่วมอบรมจากหน่วยงานราชการสามารถเบิกค่าลงทะเบียนจากต้นสังกัดได้ ตามระเบียบกระทรวงการคลัง และไม่ถือเป็นวันลาเมื่อได้รับการอนุมัติจากผู้บังคับบัญชา
- ค่าใช้จ่ายในการส่งบุคลากรเข้าอบรมทางวิชาชีพของบริษัทหรือห้างหุ้นส่วนนิติบุคคล สามารถนำไปลดหย่อน ภาษีได้ 200%
- สถาบันฯ ขอสงวนสิทธิ์ในการเปลี่ยนแปลงเนื้อหาหลักสูตร วิทยากร รูปแบบการอบรม ตามความเหมาะสมและความจำเป็น เพื่อประโยชน์สูงสุดของผู้เข้ารับการอบรม
- สถาบันฯ ขอสงวนสิทธิ์ ไม่บันทึกภาพ วิดีโอ หรือบันทึกเสียง ตลอดระยะเวลาการอบรม เนื่องจากเป็นลิขสิทธิ์ร่วมระหว่างวิทยากรกับสถาบันฯ และเพื่อป้องกันการละเมิดข้อมูลส่วนบุคคล ตาม พ.ร.บ.คุ้มครองข้อมูลส่วนบุคคล
- ผู้เข้าอบรมต้องมีเวลาเรียนไม่ต่ำกว่า 80% และทำกิจกรรมทุกหัวข้อของหลักสูตร จึงจะได้รับวุฒิบัตรจาก สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.)

ติดต่อสอบถามรายละเอียด

สถาบันพัฒนาบุคลากรแห่งอนาคต (Career for the Future Academy)

73/1 อาคารสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) ชั้น 6

ถนนพระรามที่ 6 แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพฯ 10400

โทรศัพท์ 0 2644 8150 ต่อ 81886-7

โทรสาร 0 2644 8150

E-mail: training@nstda.or.th

www.career4future.com

