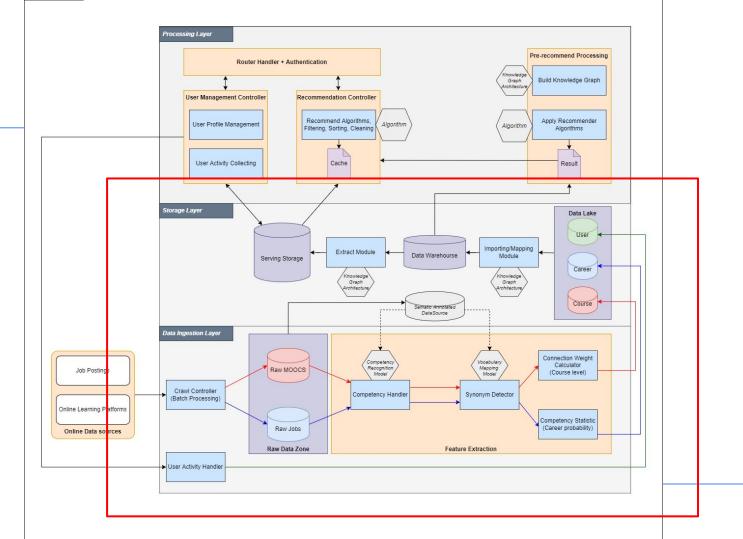
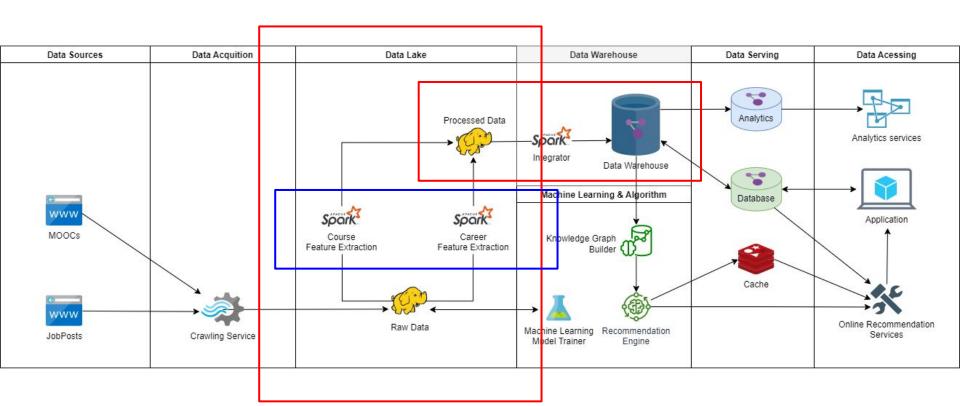
## **COURSE SEARCH SYSTEM**

GVHD: TS.Nguyễn Trần Minh Thư







## Thiết kế Data Pipeline từ Raw data -> đổ vào thành Graph

## **Yêu cầu:** Từ dữ liệu raw là output của quá trình crawl dữ liệu -> dữ liệu có ý nghĩa lưu trong graph database

Name	Link	Rating	Enroll	Instructor	Time	Level	skillrequirement	SkillWillLeam	Description	SkillGain	Subject	Organization	Fee	Program
[Intermediate] Spatia	https://www.udemy.c	4.6	5.092	Minerva Singh	5 hours on-demand	video	Interest in spatial dat	Carry out the most co	PRACTICAL TRAIN	NG WITH REAL SPAT	Math		13.99	
1 day MVP 2.0   Go	https://www.udemy.c	4.7	42600	Evan Kimbrell	13.5 hours						Entrepreneurship		13.99	
3D Data Visualizatio	https://www.coursera	.org/leam/data-visual	3.01	Kalina Borkiewicz,A.	32 hours	Beginner		This course is an int	roduction to 3D scien	tific data visualization	Data Science	University of Illinois a	Enroll for Free	
70-462: Microsoft SC	https://www.udemy.c	4.5	17.549	Phillip Burton			You don't need any	Create Backups and	Please note:Microso	ft has withdrawn the-4	Other IT & Software			
A Beginners Guide to	https://www.udemy.c	3.8	2.124	Kshitij (DB Alchemis	Academy)		Oracle DBA's	Evolution and Journ	I will be looking to d	elve deeperinto Exad	Other IT & Software			
A Crash Course in Da	https://www.coursera	4.5	179.037	Jeff Leek, PhD,Brian	7 hours	Beginner		Describe Data Scien	ce's role in various co	Data Science	Data Science	Johns Hopkins Unive	Enroll for Free	Executive Data Sc
Accounting Data An	https://www.coursera	4.1	8.065	Ronald Guymon,Line	32 hours	Intermediate		Know how to operate	e software that will he	Python Programming	Data Science	University of Illinois a	Enroll for Free	Accounting Data
Administrador Cloud	https://www.udemy.c	4	575	Ricardo Párraga Zalo	5.5 hours on-deman	d video	Conocimientos de A	Conocer las ventajas	Ingresa en el mundo	de Cloud Computing	Network & Security		13.99	
Adobe InDesign CC -	https://www.udemy.c	4.7	49908	Daniel Walter Scott	7.5 hours						Graphic Design & III	ustration	13.99	
Advanced Business A	https://www.coursera	4.3	6.978	Manuel Laguna,Dan	19 hours			The analytics proces	ss is a collection of in	terrelated activities th	Data Science	University of Colorad	Enroll for Free	Advanced Busines
Advanced Clinical D	https://www.coursera	.org/leam/advanced-	clinical-data-science	Michael G. Kahn, M	4 hours	Intermediate		This course prepares	you to deal with adv	anced clinical data sc	Data Science	University of Colorad	Enroll for Free	Clinical Data Scie
Advanced Data Scie	https://www.coursera	4.6	13.148	Romeo Kienzler	9 hours	Advanced		This project comple	ter has proven a deep	understanding on ma	Data Science	IBM	Enroll for Free	Advanced Data So
Advanced Data Struc	https://www.coursera	4.8	73.266	Leo Porter, Mia Minn	29 hours	Intermediate		How does Google M	aps plan the best rout	Graphs Search	Computer Science	University of Californ	Enroll for Free	Object Oriented Ja
Advanced Data Visus	https://www.coursera	4.8	2.07	Collin Paschall	11 hours			Data visualization is	a critical skill for any	one that routinely usin	Data Science	Johns Hopkins Unive	Enroll for Free	Data Visualization
Advanced Database	https://www.edx.org/e	course/advanced-data	abase-queries	Aspen Olmsted	40 hours	Intermediate					Computer Science	NYUx	Free	Introduction to Da
Advanced Ethical Ha	https://www.udemy.c	4.4	140	JRDcademy Institution	9 hours						Network & Security		84.99	
Advanced Machine	https://www.udemy.c	4.1	24.36	Mammoth Interactive	20.5 hours on-dema	nd video	PyCham	Code in 3 programm	"Excellent! Thank yo	ou for all your hard wo	Business Analytics &	Intelligence	9.99	
Advanced SQL Serv	https://www.udemy.c	4.3	2.885	Mike West			You will need to hav	By the end of this co	COURSEREVIEWS:	"It's not what I though	SQL Server			
Advanced Styling wi	https://www.coursera	4.7	91555	Colleen van Lent, Ph	13 hours						Computer Science	University of Michiga	Free	Web Design for Ev
Advanced Topics an	https://www.edx.org/e	course/advanced_top	53661	Amando Fox, David	35 hours	Intermediate					Computer Science	Berkeley, Universty o	Free	
Agile Analytics	https://www.coursera	4.8	1758	Alex Cowan	15 hours	Beginner					Computer Science	University of Virginia	Free	
Agile Analytics	https://www.coursera	4.8	17906	Alex Cowan	15 hours	Beginner					Computer Science	University of Virginia	Free	Agile Developmen
Agile Leader Trainin	https://www.coursera	4.7	3423	Oxana Trotsenko	15 hours	Beginner					Business	University of Colorad	Free	Agile Leadership S
				100 0							127 17 12 32	Later to the state of	100	La de La Caración de Caración

- Xử lý trùng link?
- O Xử lý timestamp?
- Gọi NER model trong bước xử lý như thế nào?
- Hậu xử lý các từ chưa chuẩn hóa ví dụ: data warehouse/data warehouses,... (Xem qua Stemming vs Lemmatization)
- Phương pháp tính toán và thống kê tần suất xuất hiện của các năng lực nghề nghiệp để đánh giá mức độ phổ biến của nó trên một tiêu đề nghề nghiệp.
- Cách mapping dữ liệu vào Graph Database