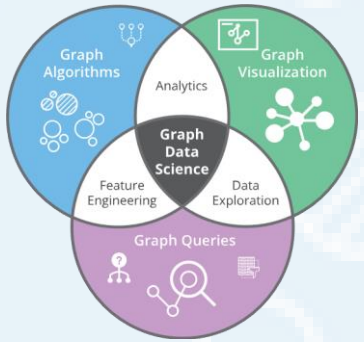


Habib University
shaping futures

CS343 Graph Data Science

Spring 2024

Graph Modelling

Chapter #2, Tomaz

Chapter #2, Ian Robinson

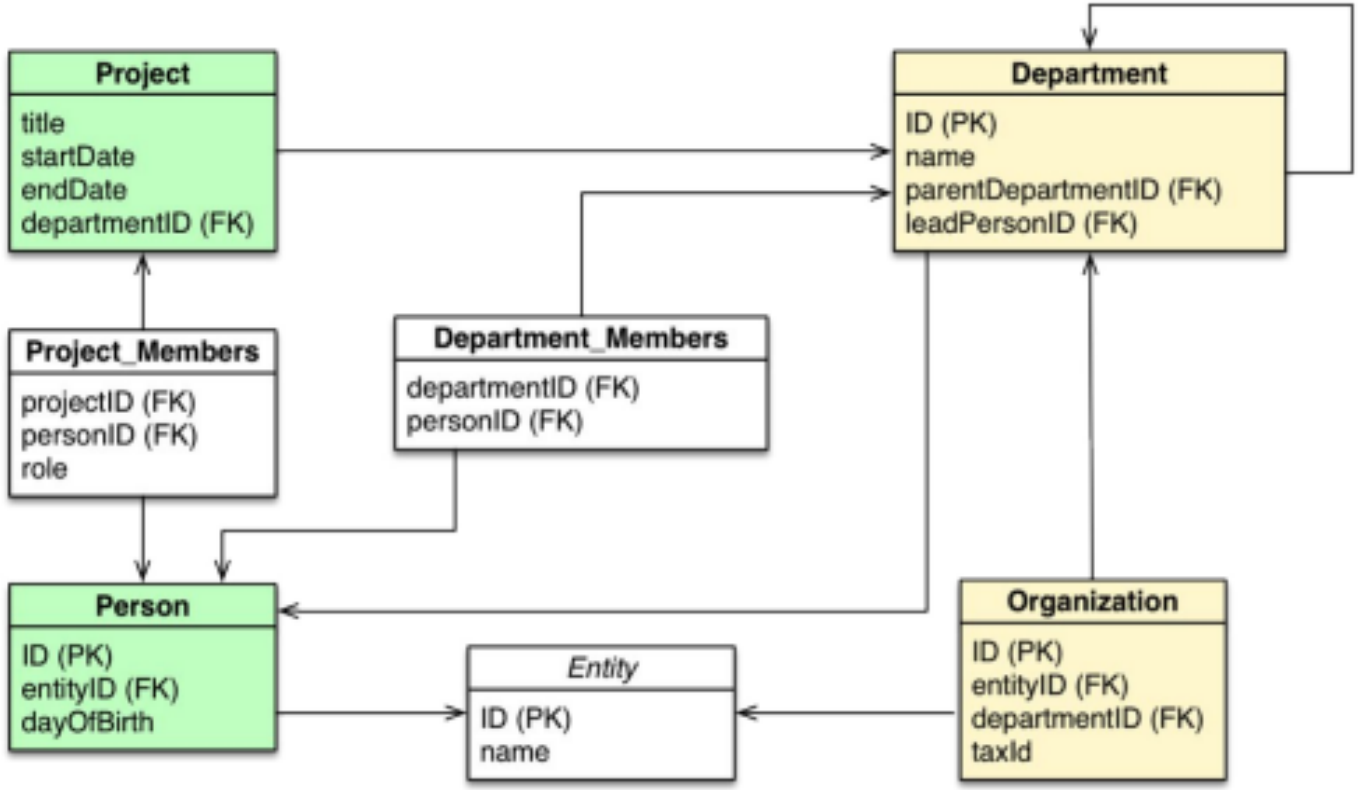
Muhammad Qasim Pasta

qasim.pasta@sse.habib.edu.pk

Slides are intended to be filled during the lectures. Certain details are intentionally omitted for in-class discussions. These slides are not meant to be used as reading material.

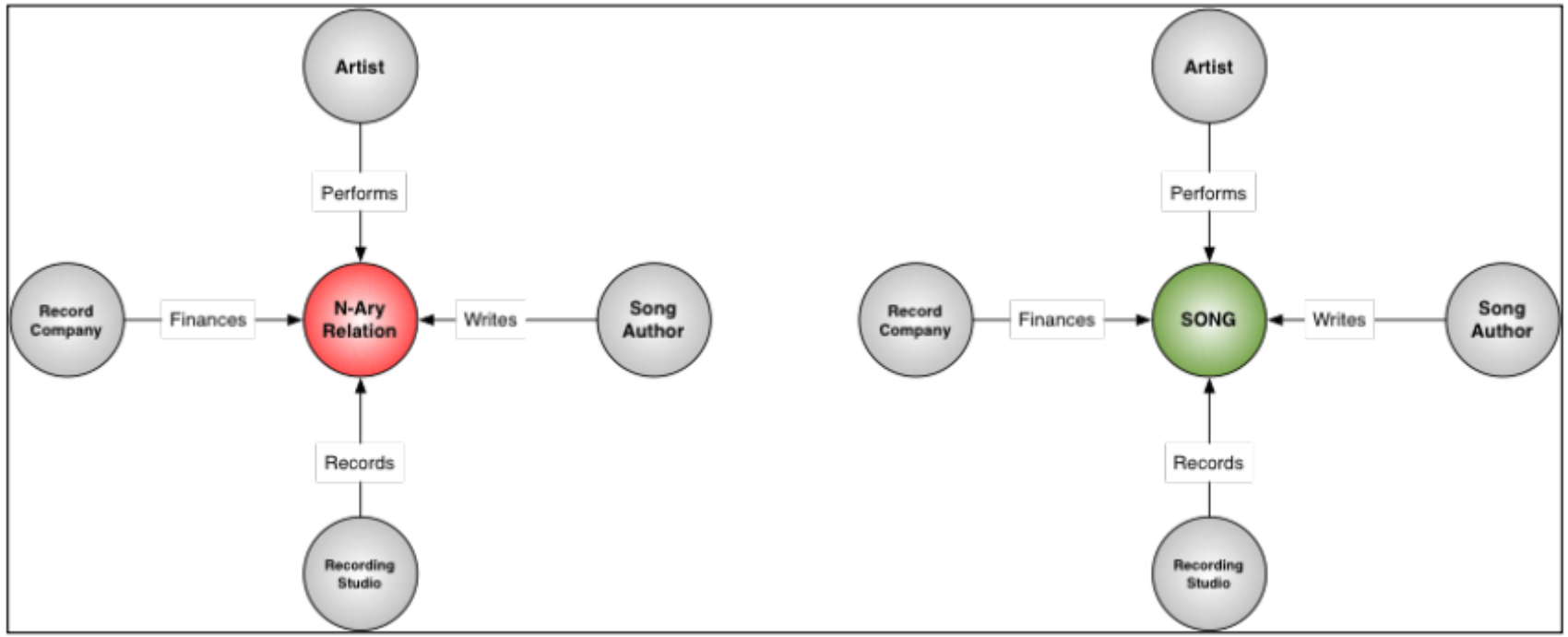
Labelled Property Graph Model

- Create a LPG model to store the data equivalent to that represented by the following ERD.



Graph Modelling

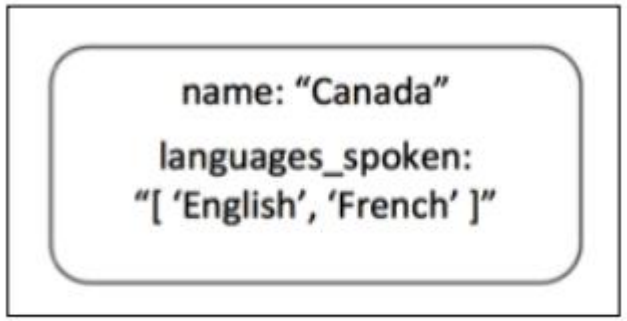
- Look for n-array relationships



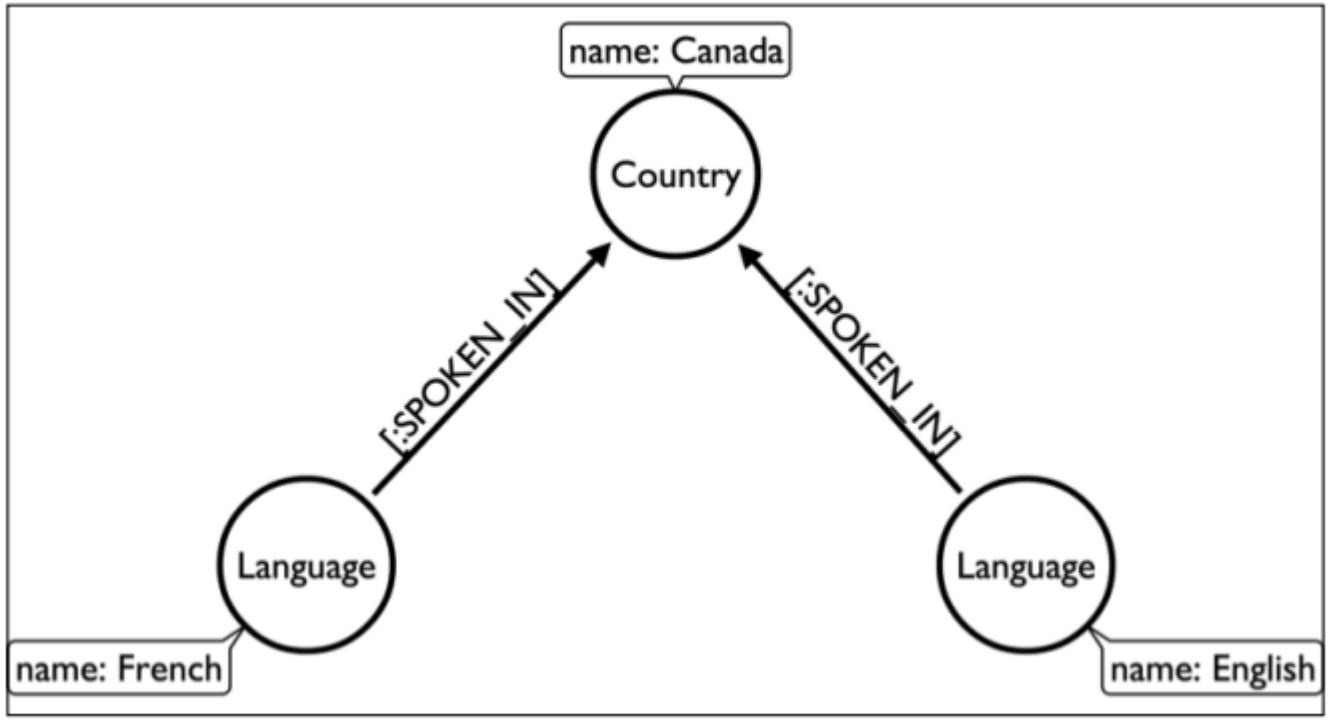
Transforming n-ary relationships into nodes

Graph Modelling

- Avoid “rich” properties
- Property vs Node



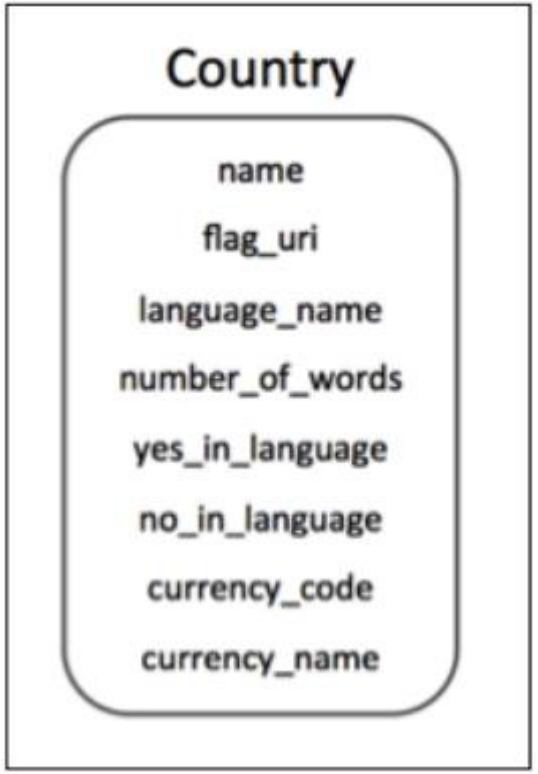
Using "rich" properties

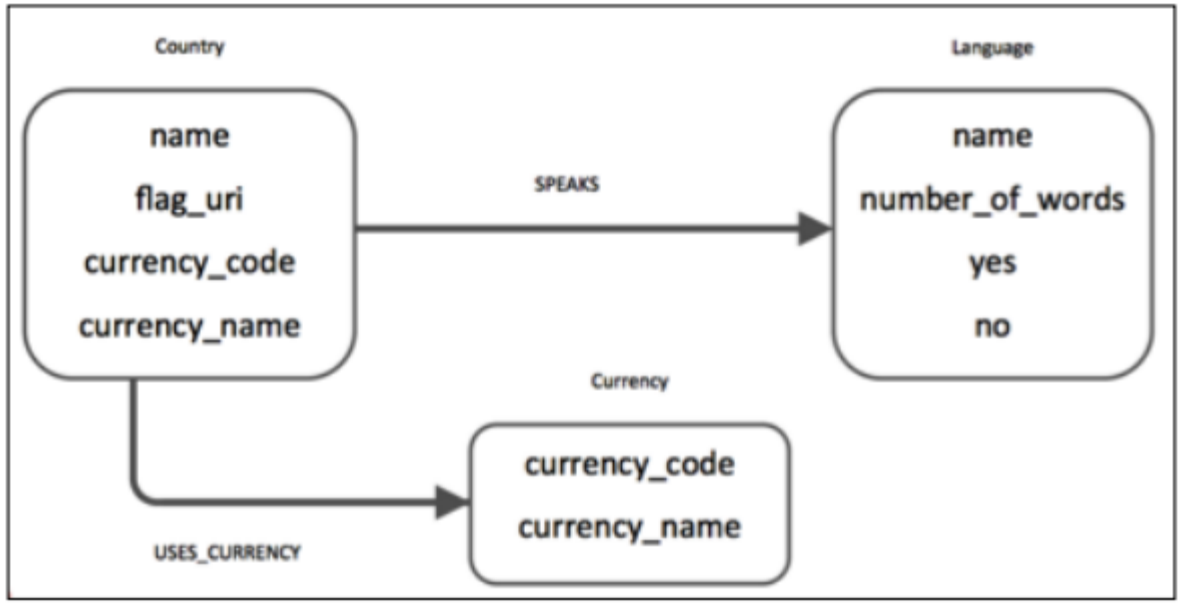


Granulating "rich" properties

Graph Modelling

- Avoid nodes representing multiple concepts





Graph Modelling

- Super node or “dense” node problem
  - Super node: node that happens to have a lot of relationships.

- Performance issues: travers all path to find a possible answer
  - If the super node is at the end of a path, and you are not traversing through it, it will not have much impact.
  - The super node can be harmless if it isn’t part of a high-transaction rate query, or multi-user update.
- Super node refactoring
  - last resort

Read: https://medium.com/neo4j/graph-modeling-all-about-super-nodes-d6ad7e11015b

Case Study: RDBMS to Graph

- Your are provided a case study to create Labelled Property Graph Model

CourseNumber	CourseName	CreditHours	Type	DepartmentID	PreRequisiteNumber	FacultyLead	MinorID	SemesterNumber
CS101	Algorithmic Problem Solving	3	Foundation	1	1	Dr. Waqar Saleem	1	1
CS101L	Algorithmic Problem Solving Lab	1	Foundation	1	1	Dr. Waqar Saleem	1	1
CS102, CE171	Data Structures and Algorithms	3	Foundation	1	1	Dr. Qasim Pasta	1	2
CS102L, CE171L	Data Structures and Algorithms Lab	1	Foundation	1	1	Dr. Qasim Pasta	1	2
CS113, MATH113	Discrete Mathematics	3	Foundation	1, 3	1	Dr. Waqar Saleem	1, 2	2
CS355, CE373	Database Systems	3	Kernel	1, 2	2	Dr. Qasim Pasta	1	3
CS355L, CE373L	Database Systems Lab	1	Kernel	1, 2	3	Dr. Qasim Pasta	1	3
CS224, CE272	Object Oriented Programming	3	Kernel	1, 2	2	Nadia Nasir	1	2
CS224L, CE272L	Object Oriented Programming Lab	1	Kernel	1, 2	3	Nadia Nasir	1	2
CS201	Data Structures II	3	Kernel	1	2, 3, 4	Dr. Faisal Alvi	1	4
CS232, CE324	Operating Systems	3	Kernel	1, 2	2, 3, 5, 6	Dr. Muhammad Mobeen Movania	1	5
CS232L, CE324L	Operating Systems Lab	1	Kernel	1, 2	2, 3, 5, 6	Dr. Muhammad Mobeen Movania	1	5
CS212	Nature of Computation	3	Kernel	1	4	Dr. Waqar Saleem	1	5
CS412	Algorithms: Design and Analysis	3	Kernel	1	7, 8	Dr. Shah Jamal Alam	1	6