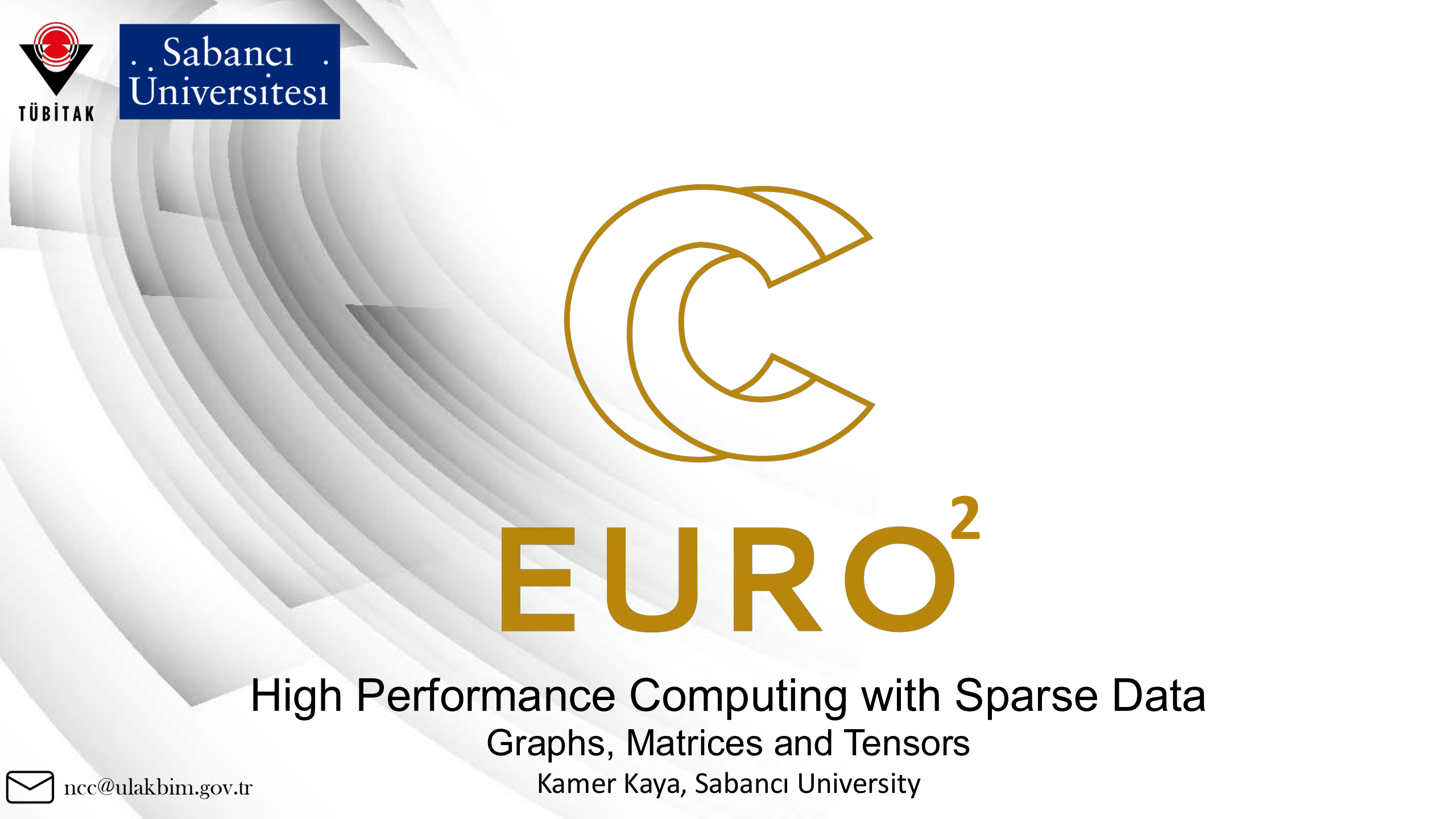


High Performance Computing with Sparse Data

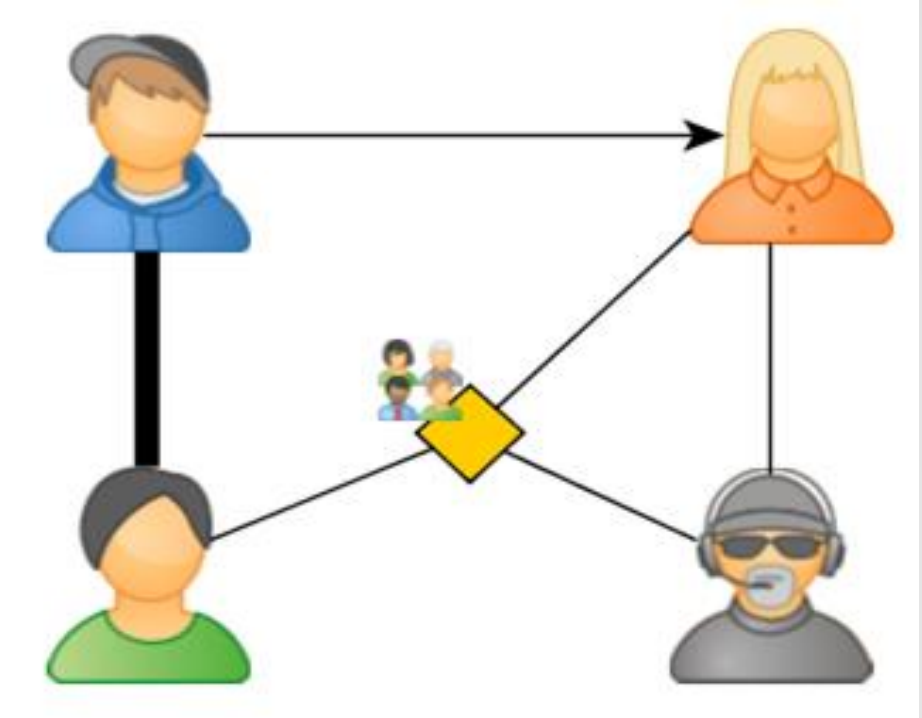
Graphs, Matrices and Tensors

ncc@ulakbim.gov.tr

Kamer Kaya, Sabancı University

Graphs

- Graphs are structures used to model the relationships between elements in data.
- Different types of vertices and edges can be used to model different types of relationships.



Graphs in Real Life

EURO²

Network	Vertices	Vertex Metadata	Edges	Edge Metadata
Airlines	Airports	Terminals, city, population, national/international	Uçaklar, rotalar	Flight frequency, number of passengers, distance, gasoline usage, capacity
Banking and Finance	Customers, accounts	Person, ID, products, account balance, loan amount, demographic data	Transactions	Type, quantity, safe/unsafe, place, time, vehicle
Social Networks	Users	Name, demographic data, likes, posts, memberships	Interactions	Time, duration, subject, frequency
Medical Networks	Doctors	Demographic data, specialty, workplace information, daily-weekly average number of patients	Patients	Demographics, diagnosis, treatment, frequency of visits, insurance
Supply Chain Networks	Storage locations	Location, size, capacity, download upload speed, automatic/manual maintenance	Trucks, routes	Load amount, return, distance, driver, maintenance costs

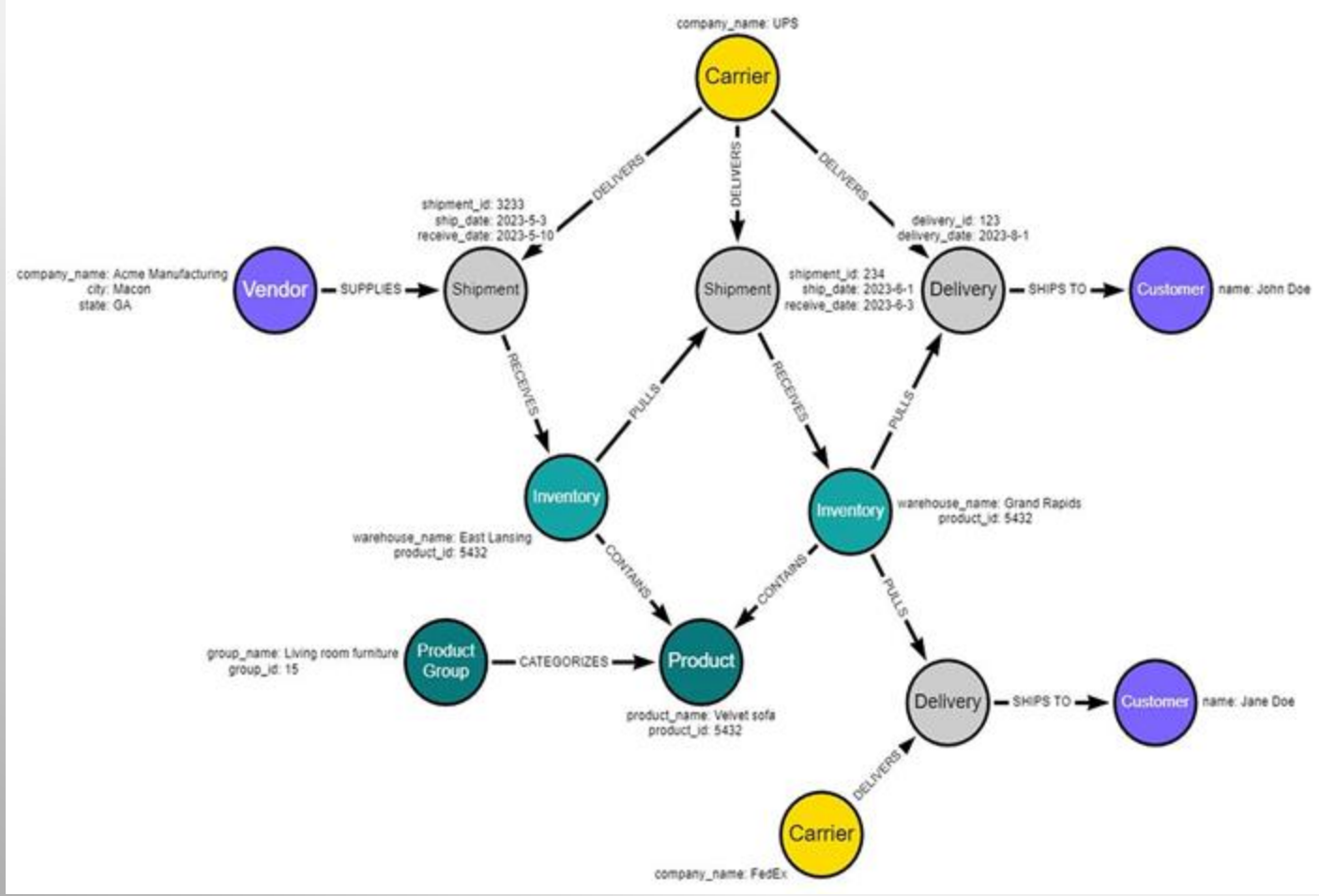
Graphs in Real Life

EURO²

Marketing Analytics – Graphs can be used to find the most influential people on a social network. By flowing their messages through the most influential people on a Social Network, Advertisers and Marketers can anticipate the biggest explosion of marketing revenue.

Banking Transactions – Graphs can be used to find unusual patterns, which helps reduce fraudulent transactions. There have been instances where illegal activity has been detected by analyzing the flow of money through interconnected banking networks.

Knowledge Graphs and Graph DBs



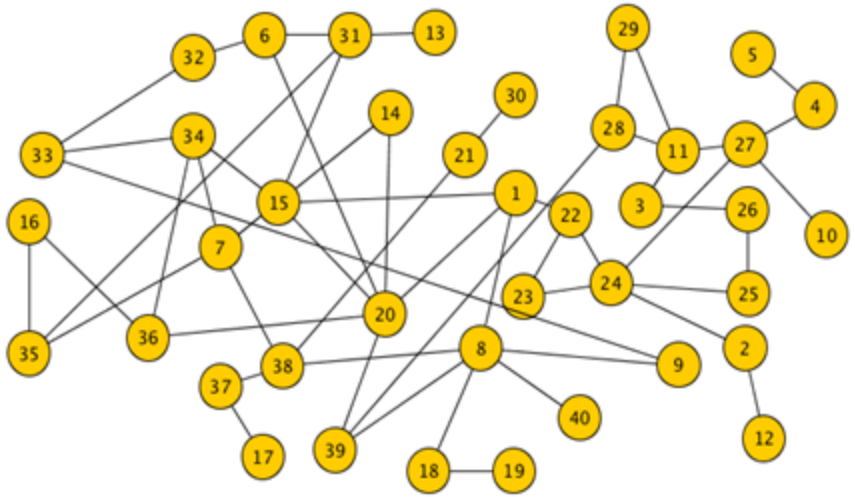
https://towardsdatascience.com/why-you-need-a-knowledgegraph-and-how-to-build-it-ac4f35cb75b7

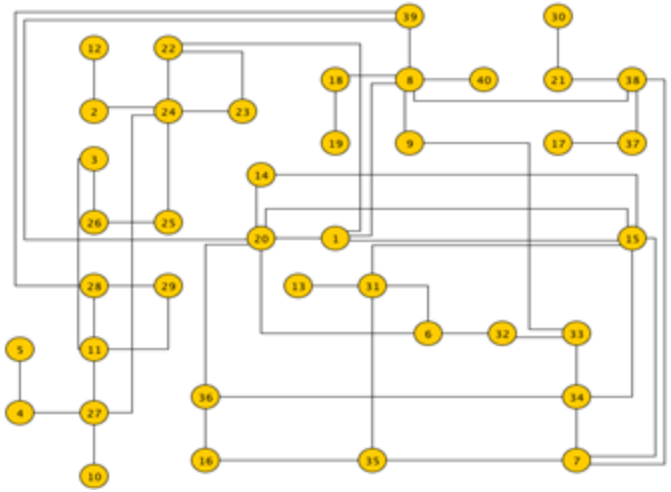
In real life, most data is organized as information/knowledge graphs.

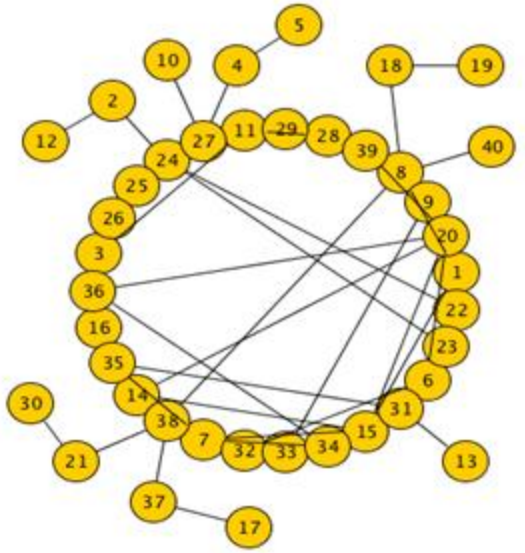
These graphs and the graph DBs used on the backend are different from classical DBs.

Organizations in banking, automobile industry, oil, healthcare, retail, publishing, media etc. use information graphs and graph databases to add value to their data.

Graphs – Structure is important

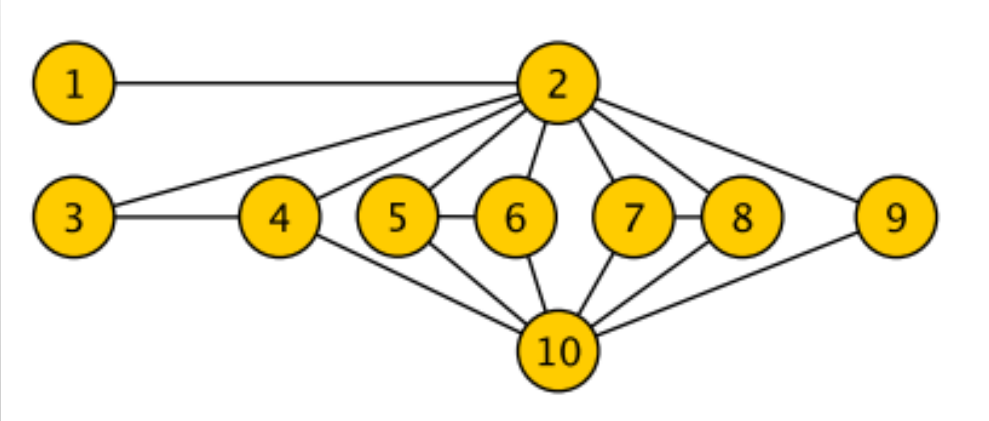






Data Structures to Store/Process Graphs

EURO²

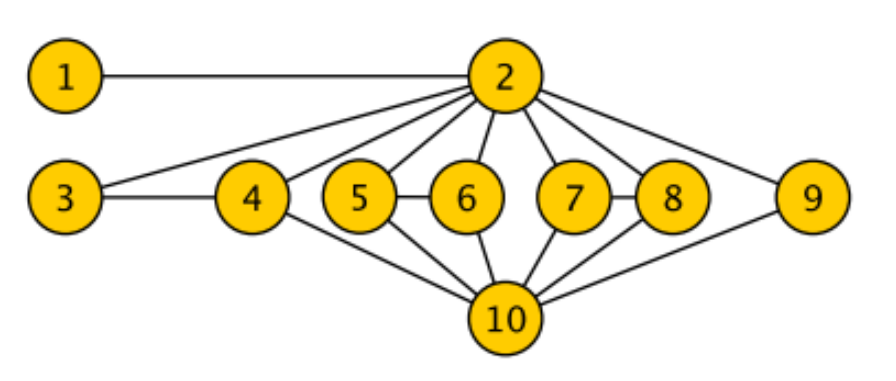


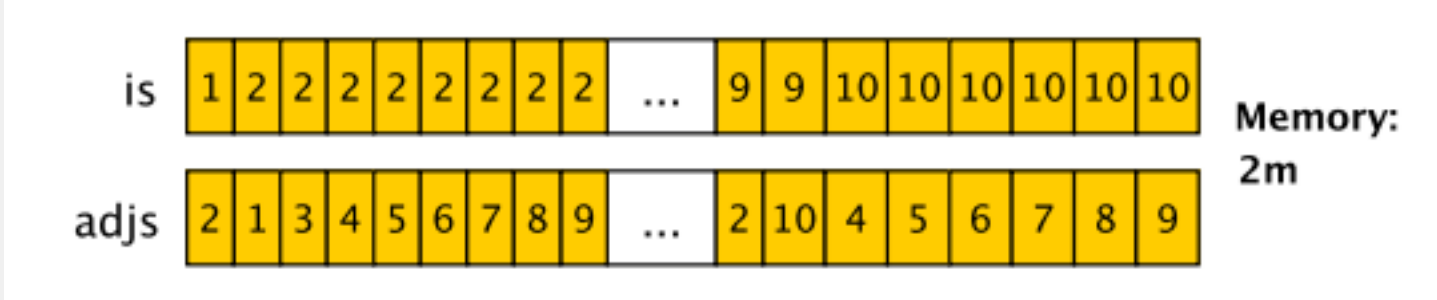
	1	2	3	4	5	6	7	8	9	10
1		■								
2	■		■	■	■	■	■	■	■	
3		■	■							
4		■	■							
5		■				■				■
6		■			■					■
7		■						■		■
8		■					■			■
9		■								■
10				■	■	■	■	■	■	

Matrix-like storage

- 1M vertex 10^{12} values.
- 1TB values – 4TB memory

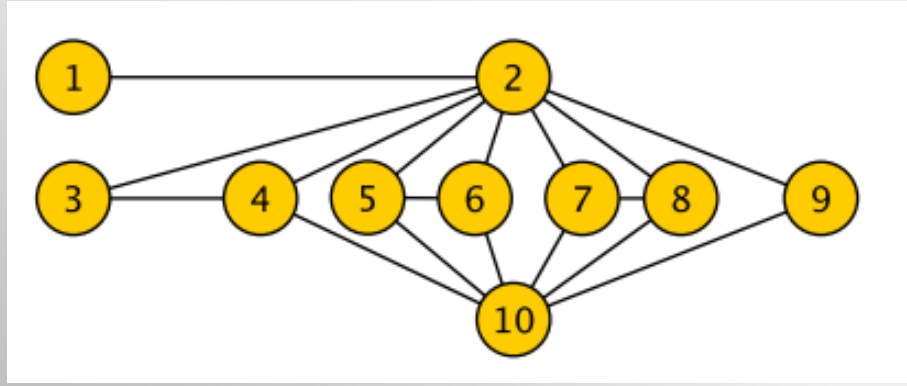
Sparse Data Structures (COO)

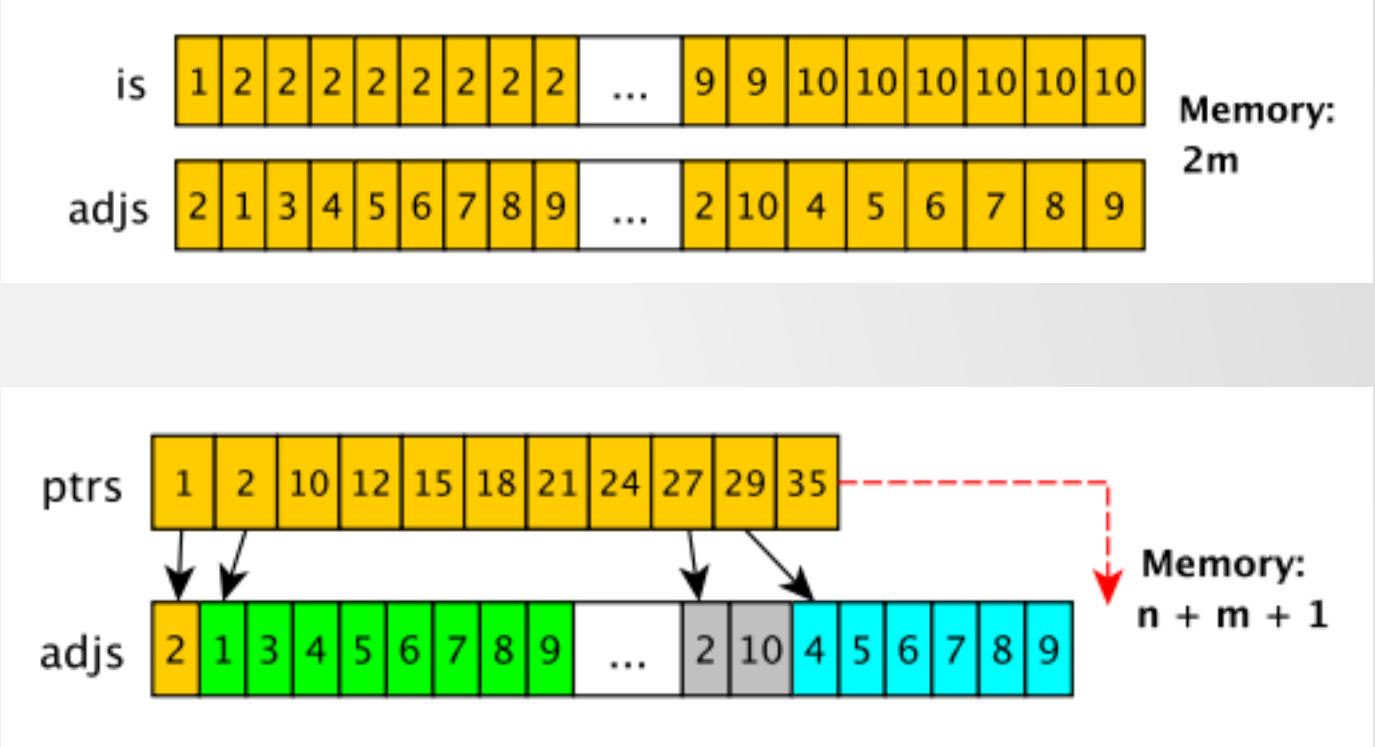




Coordinates

Sparse Data Structures (CRS)





Compressed graph data structure

Thanks

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101101903. The JU receives support from the Digital Europe Programme and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia