

Data Engineer



Easily map data and processes' onto DVM

Data Scientist



Data Exploration and Query Formulation

Data Officer

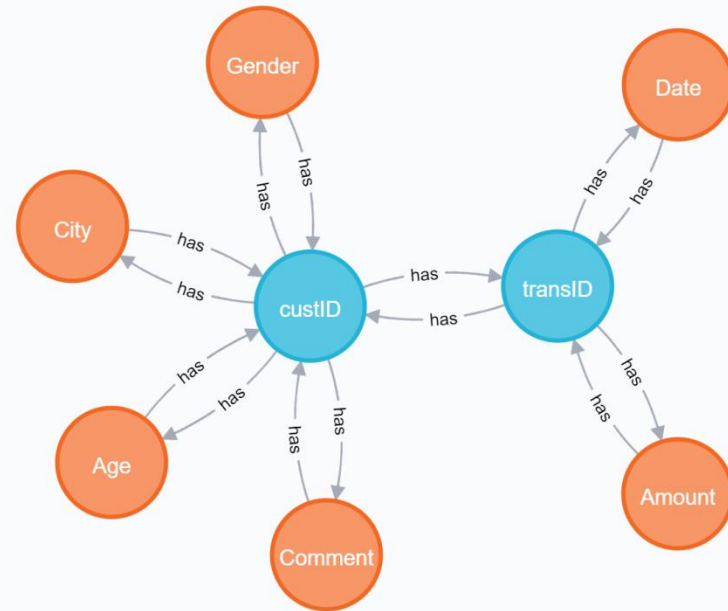


Data governance, Compliance

End Users



Personal Data Management Tools



Data Virtual Machines

A conceptual model in which all data stakeholders in Analytics environments collaboratively work

- *Data Engineers*
- *Data Scientists*
- *Data Officers*
- *End Users*

Data Virtual Machines

- Graph-based model, a modification of ER
- A computation (e.g. a query, a script) over existing data with output (u,v) , $u \in A_1$, $v \in A_2$ provides a mapping between attributes A_1 , A_2 and a mapping between A_2 , A_1
- Nodes in the graph represent attributes
- Edges represent computations (mappings) between nodes
- Nodes and edges are added agily, as needed
- Queries are subgraphs of DVM, algebraic framework for query evaluation, optimization