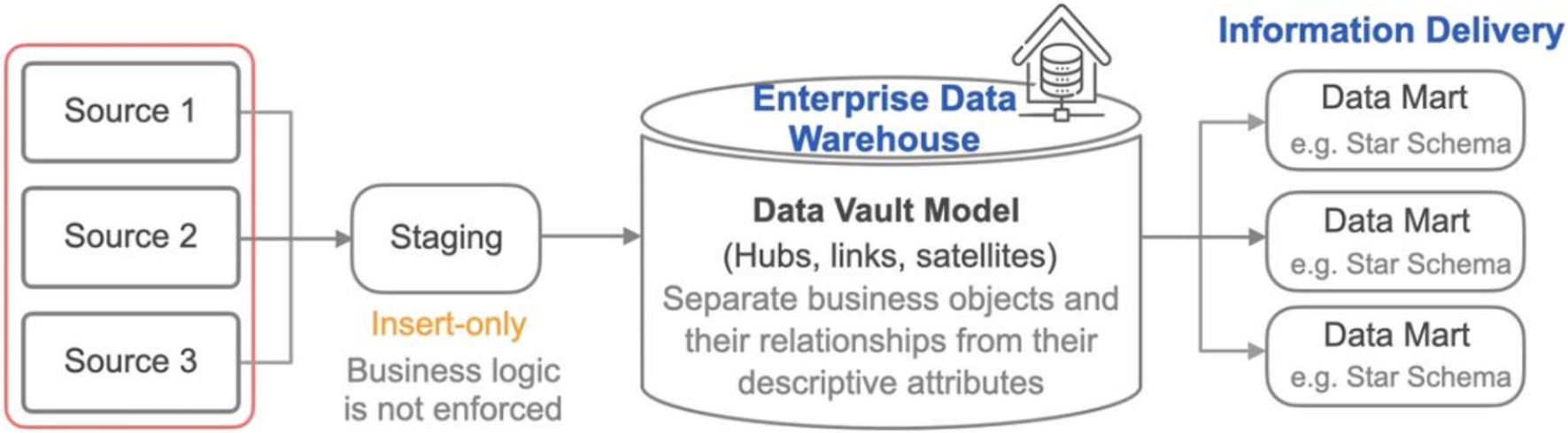
Data Vault

Dan Linstedt introduced Data Vault as a different approach to modeling the data in the data warehouse.



Data Vault Model

Three main types of tables:

Hub

Stores a unique list of business keys

Customers, products, employees, vendors

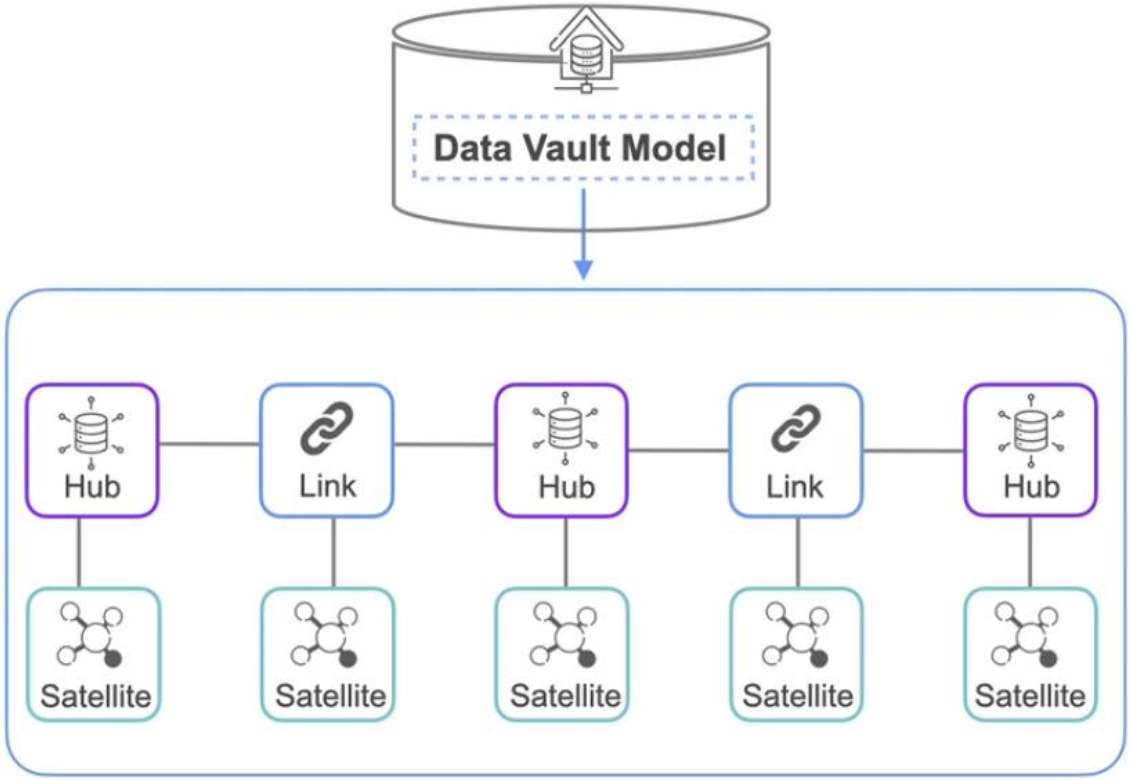
Link

Connects two or more hubs

Relationship, transaction, event

Satellite

Contains attributes that provide context for hubs and links



Data Vault - Step 1: Model the Hubs

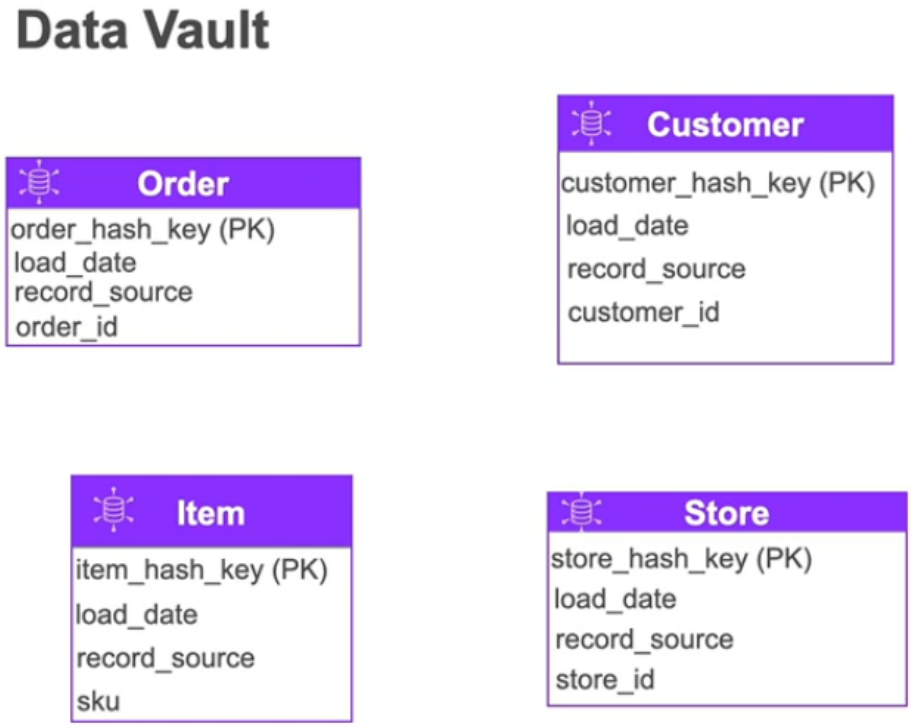
Data Vault

- What is the identifiable business element?
- How do users commonly look for data?
- A business key:
  - column(s) used by the business to identify and locate the data
  - not be a key generated in or tied to a particular source system

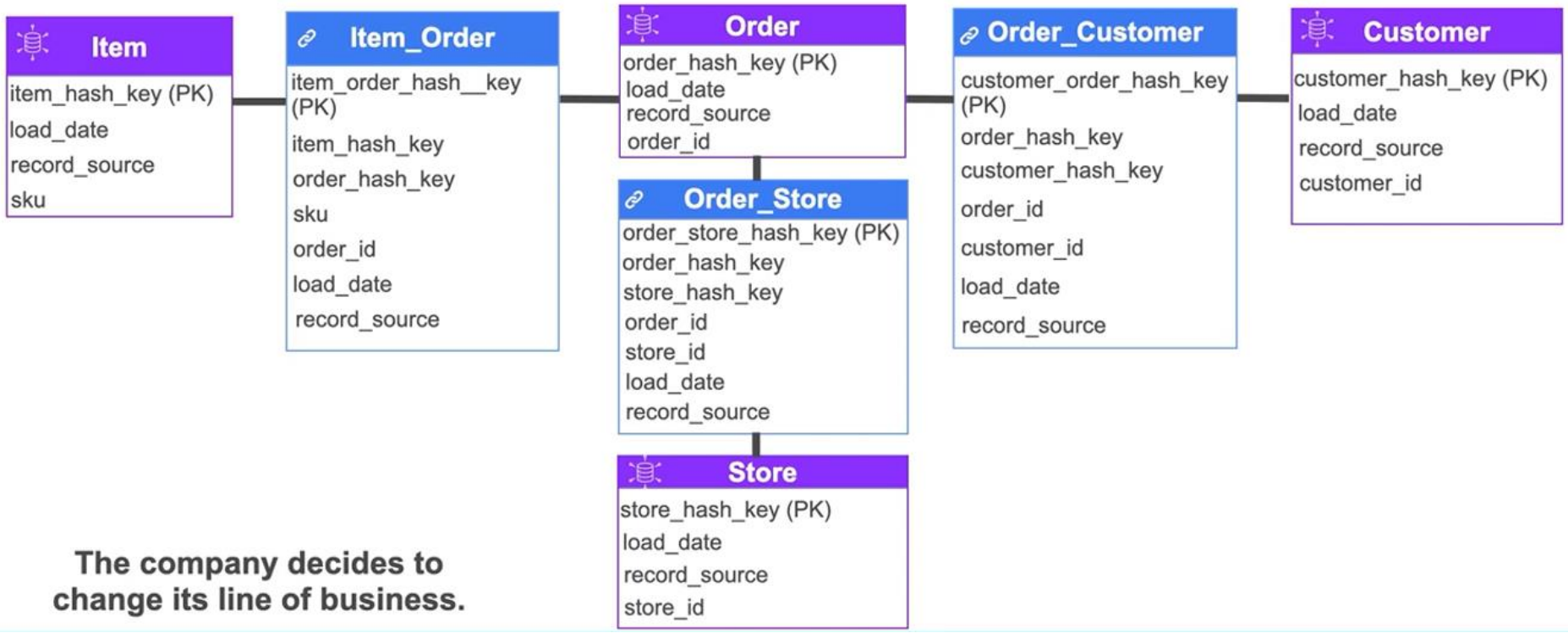
Data Vault - Step 1: Model the Hubs

A hub should contain:

- **The business key**
- **The hash key**:
  - Calculated as a hash of the business key
  - Used as the Hub primary key
- **The load date**: date on which the business key was first loaded
- **The record source**: the source of the business key



Data Vault - Step 2: Model the Links



The company decides to change its line of business.

Data Vault - Step 2: Model the Links

Data Vault
- Step 3: Satellites