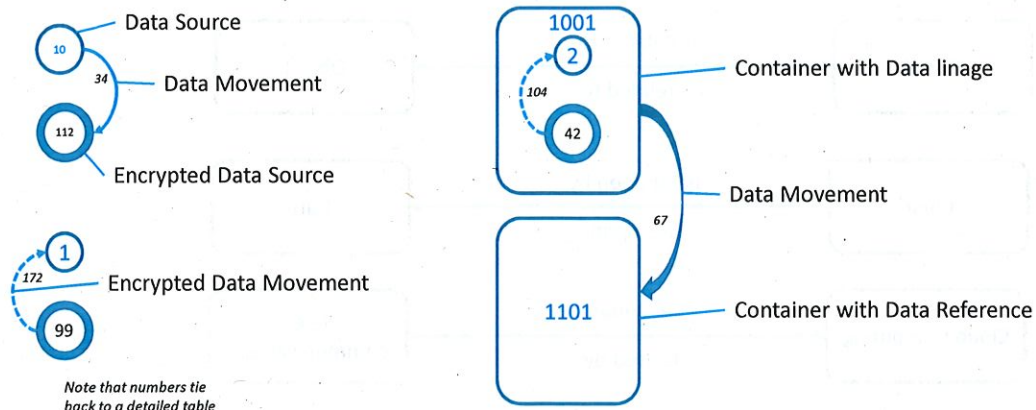# Creating Data Management Diagrams
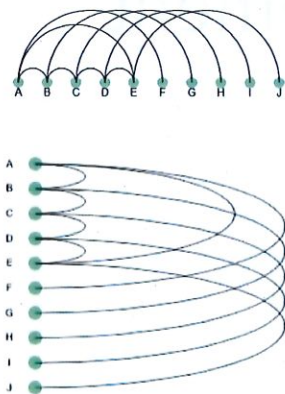
VCU Capstone Project Guidance

Create a visual grammar to convey your point

- Data Source
- Data Movement
- Encrypted Data Source
- Encrypted Data Movement

Note that numbers tie back to a detailed table

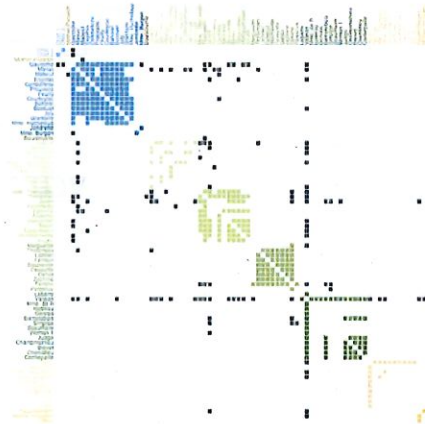- Container with Data linage
- Data Movement
- Container with Data Reference

Arc diagrams can be used to show data linage
Visualize the complexity

http://bl.ocks.org/mayblue9/dcc49ef6e3888f37f755177c4a248f2c

https://homes.cs.washington.edu/~jheer/files/zoo/
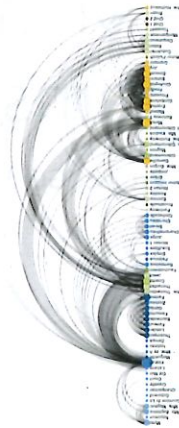
FORML: Formal Object Role Modeling Language
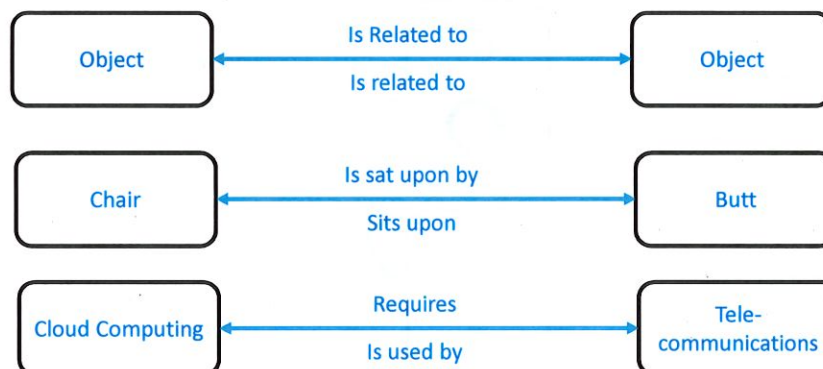
Dependency Diagram

What does the data table look like?

The size and shape of the screen and printer
Paper should inform your design to be readable

Business Process

Criticality of constraint?

Types of constraints?

Business Process — IT Process — IT Process

Reference to table detail?

People — System — Facility — Third Party — Data Interface
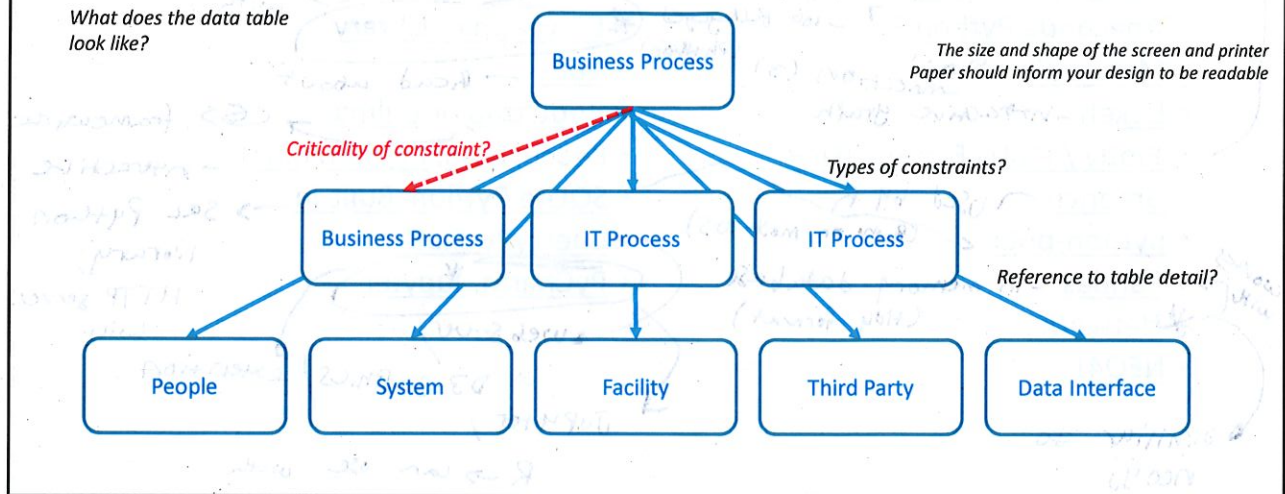
A Taxonomy is an effective way of organizing data to provide the appropriate viewpoint

Customers Consumers
Features
Products
Capabilities
Processes

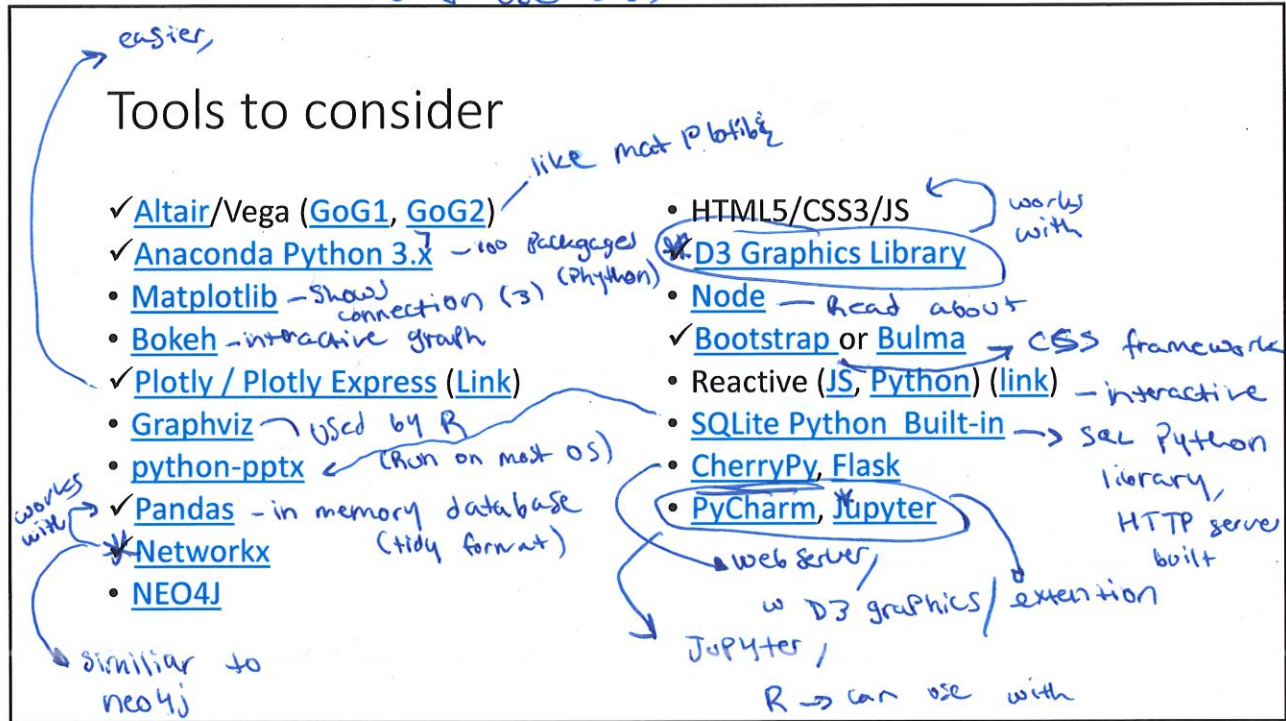| People	Processes	Technology
Abilities		Software
Experience		Hardware
Skills		Cloud

togoph → ideology for system thinking
ideology

*Pie charts   Stanford execution
framework — high-level
corporate level

3

Data lineage → changes being made
compliances
security
Exec
→ other use cases

harvey.summers@bbfa.com

---

easier

Tools to consider

like most plotlib

- ✓ Altair/Vega (GoG1, GoG2)
- ✓ Anaconda Python 3.x — ~100 packages (Python)
- • Matplotlib — show connection (3)
- • Bokeh — interactive graph
- ✓ Plotly / Plotly Express (Link)
- • Graphviz → used by R (run on most OS)
- • python-pptx
- ✓ Pandas — in memory database (tidy format)
- ✗ Networkx
- • NEO4J

works with → similiar to neo4j

- • HTML5/CSS3/JS — works with
- ✗ D3 Graphics Library
- • Node — read about
- ✓ Bootstrap or Bulma → CSS framework
- • Reactive (JS, Python) (link) — interactive
- • SQLite Python Built-in → SQL Python library, HTTP server built
- • CherryPy, Flask
- • PyCharm, Jupyter

web server w D3 graphics/ extention
Jupyter /
R → can use with

D3 - graph is integrated with Python and Front-end development.
HTML - JS

Tools: Timeline

Am a using data table
how is it useful

SEC Filing) (BAC)
(chord diagram)
influence (relationshi

Tips to consider

- ✓ Keep your data TIDY → every cell in table is filled out
  - • Format is tabular or json
  - • Denormalized
- • Tidy data can drive most graphics tools
- • Use Python 3.x with Multiprocessing for speed working with data (link)

  Python 3.7
  - • Async/await
  - • Yield not return
- (idiom changes with statement) • Dash can be used for very large files
- • Stackoverflow is your friend
- ✓ Useful references (link, link2, link3)

- • The presentation layer should be on the browser (JavaScript/HTML5/CSS3) or client (Graphviz)
- • The backend can be Python or there can be no backend (csv-parse)
- ✓ KEEP IT SIMPLE
  - • Fewest layer, tools, etc.
  - • Do one thing really well, then the next... don't do a dozen things badly
  - • Don't let purity prevent delivery
- • "import this" & "import antigravity"

Cofca → stream / real-time display

D3 - Network graph

Neo4j → too complicated, think about the problem work backwards, what the data looks like

D3 graphics library

Table/json

Python / networkX