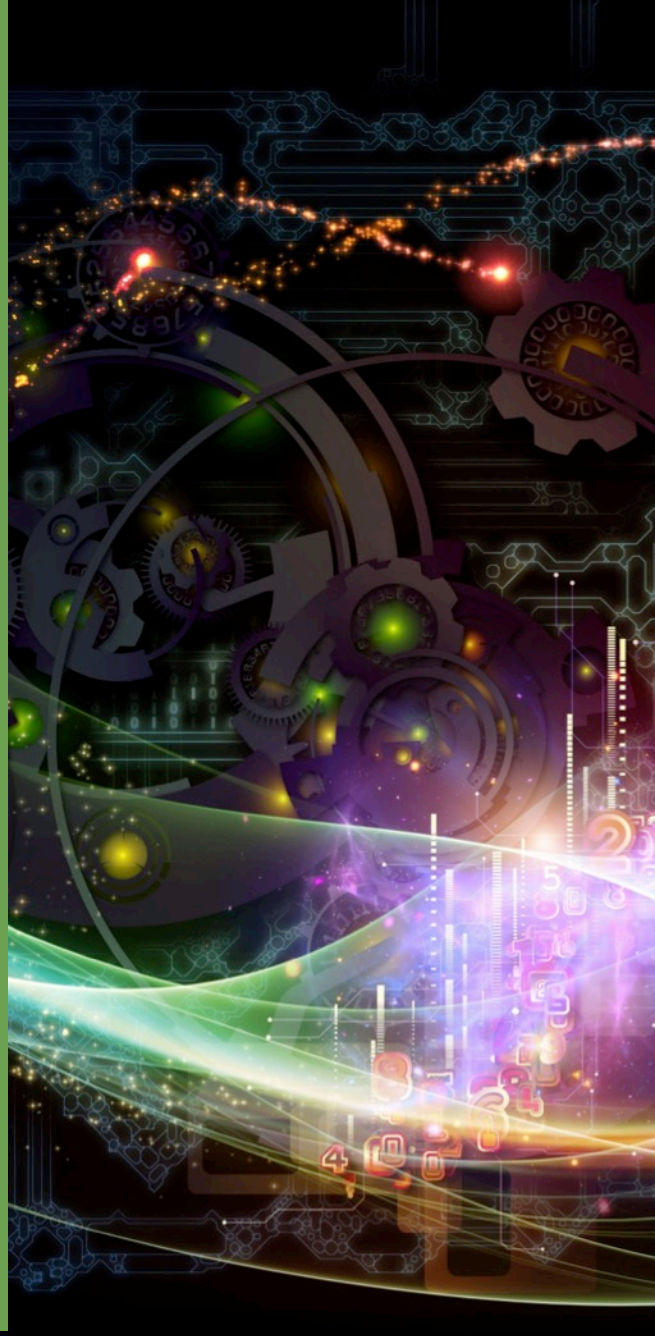
Software Solutions Symposium 2017

March 20–23, 2017

6 Things You Need to Know About Data Governance

John Klein

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213



This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM-0004340

1. **A data set produces benefits *only* when it is used to make decisions.**

2. value ≡ ∑benefits – ∑costs

3. Moving Parts

Producer

- Sensor
- Open source feed (e.g., Twitter)
- 3rd Party Source (e.g., geo map)
- External system

Data

Publisher

- Acquires
- Stores
- Makes available

Data

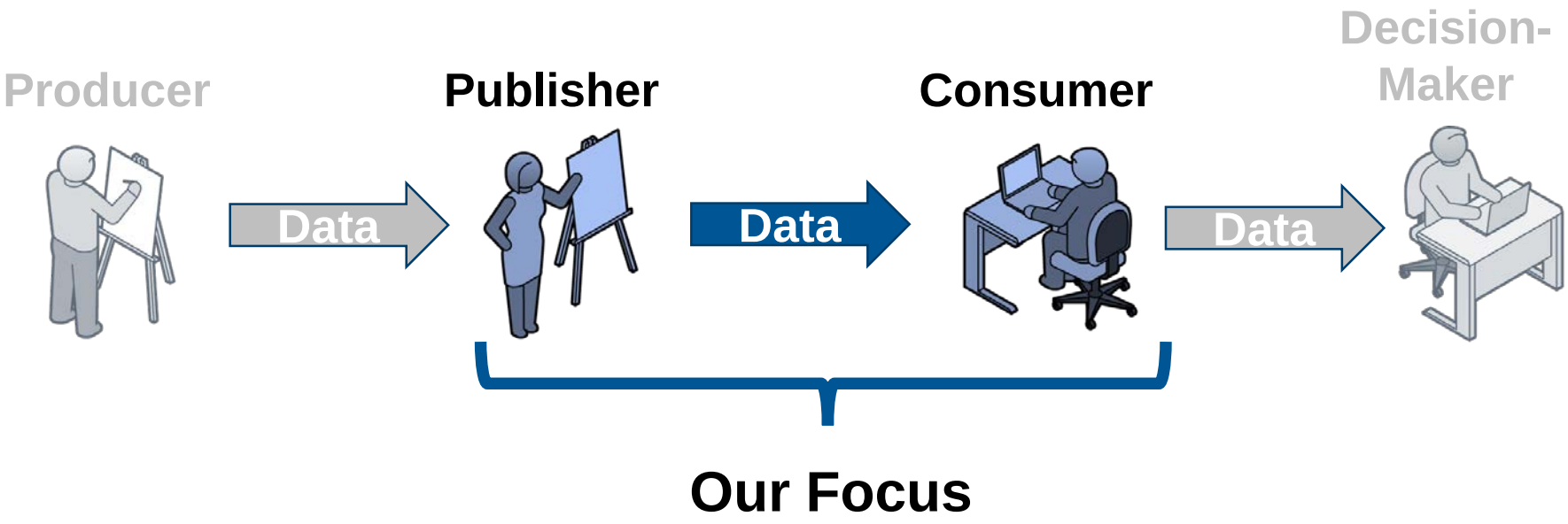
Consumer

- Develops decision support application

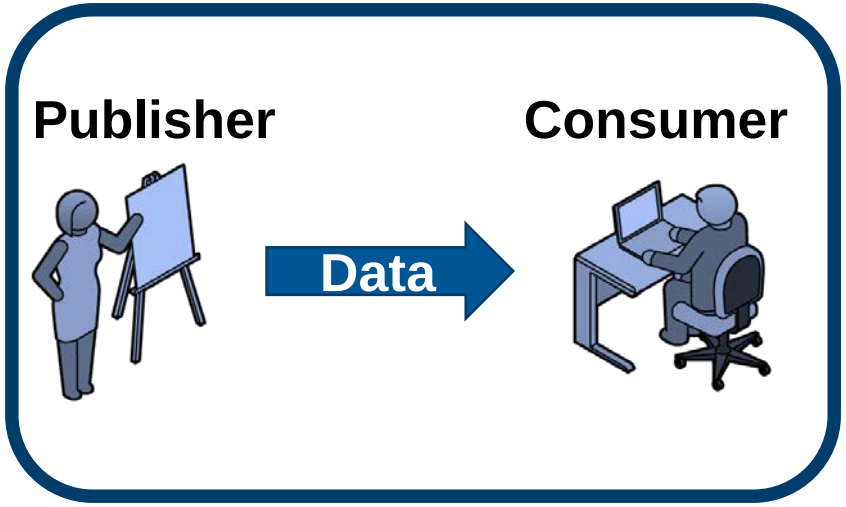
Data

Decision-Maker

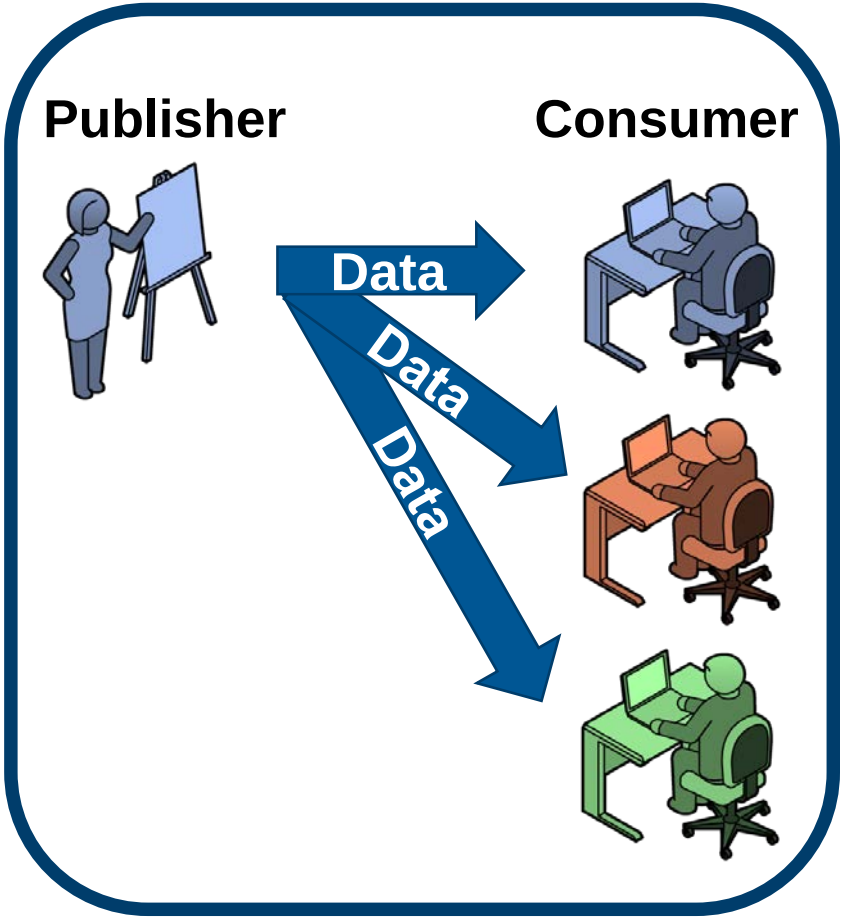
- Uses application to make decisions



Easy!



Challenging!



4. *Governance* constrains the data publisher, in order to help the data consumers

Constrains the publisher…

May *reduce* the costs to the publisher

- Reduce types and numbers of interfaces
- Restrict backward compatibility on interfaces
- Restrict technology options

Usually *increases* the costs to the publisher

- Different schema
- Data quality
- Quality of Service (e.g., availability)

…to help the consumers

Reduce local and global costs of consuming the data

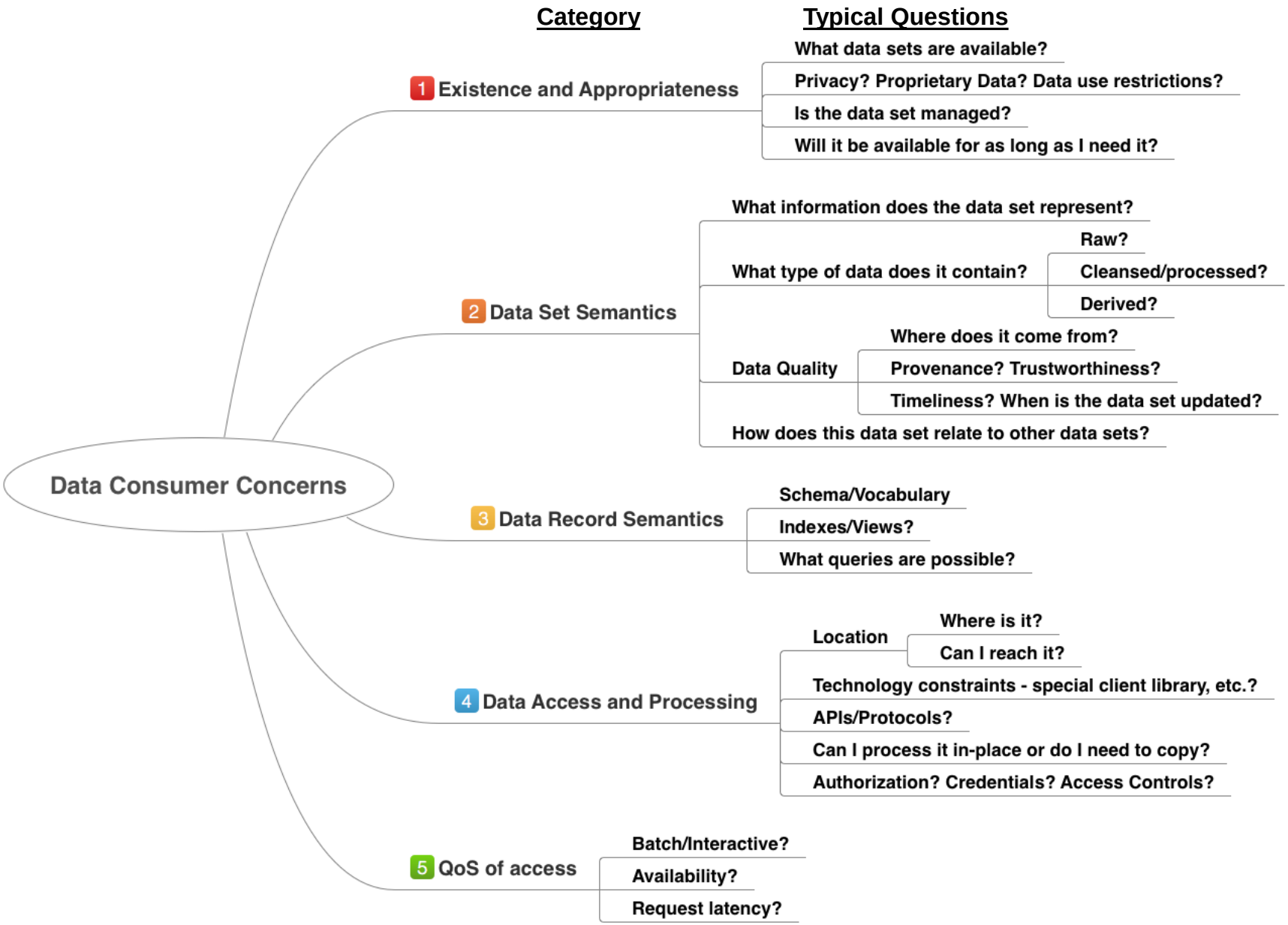
- Standard interfaces
- High quality data
- Service level agreements

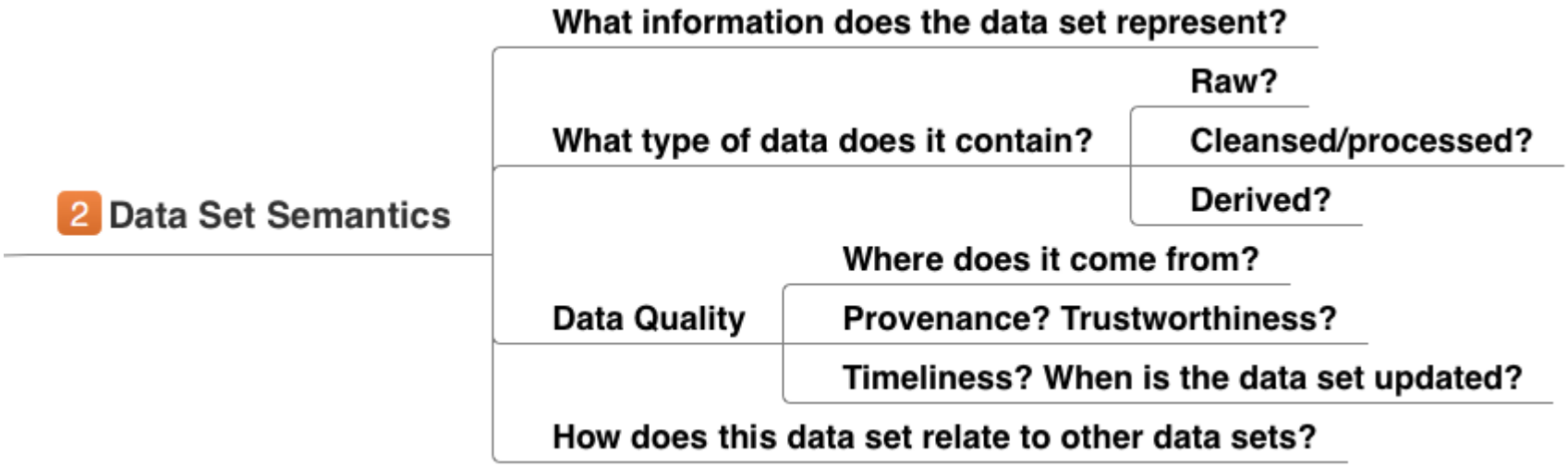
Enable consumers to *deliver more benefits*

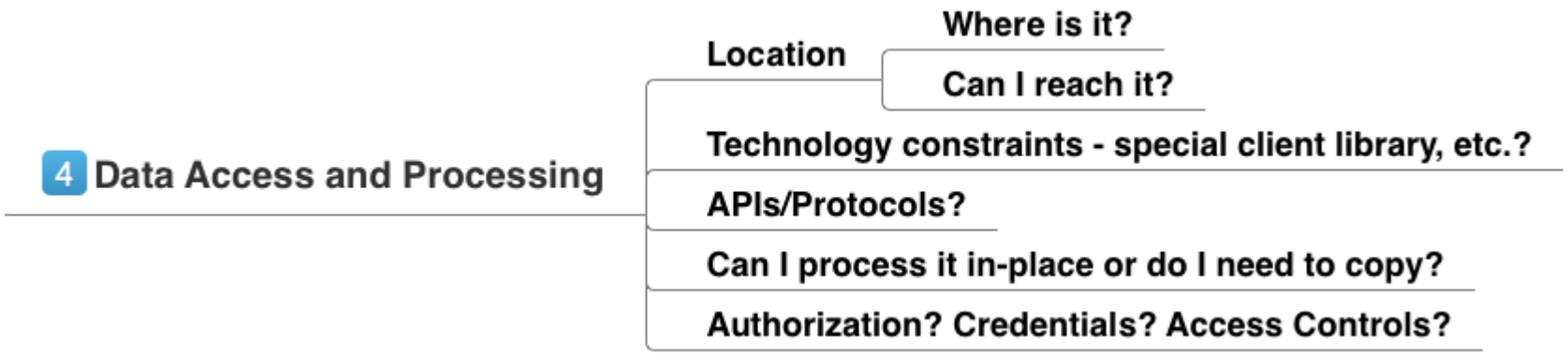
- New decision support applications
- Higher quality decisions

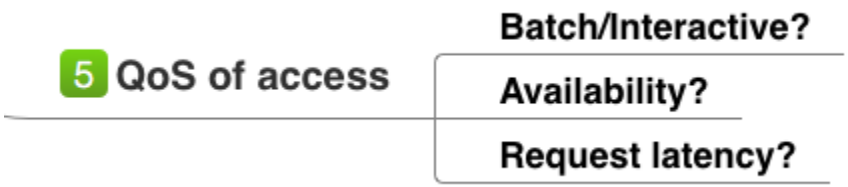
5. Apply governance only when it increases value (benefits > costs)

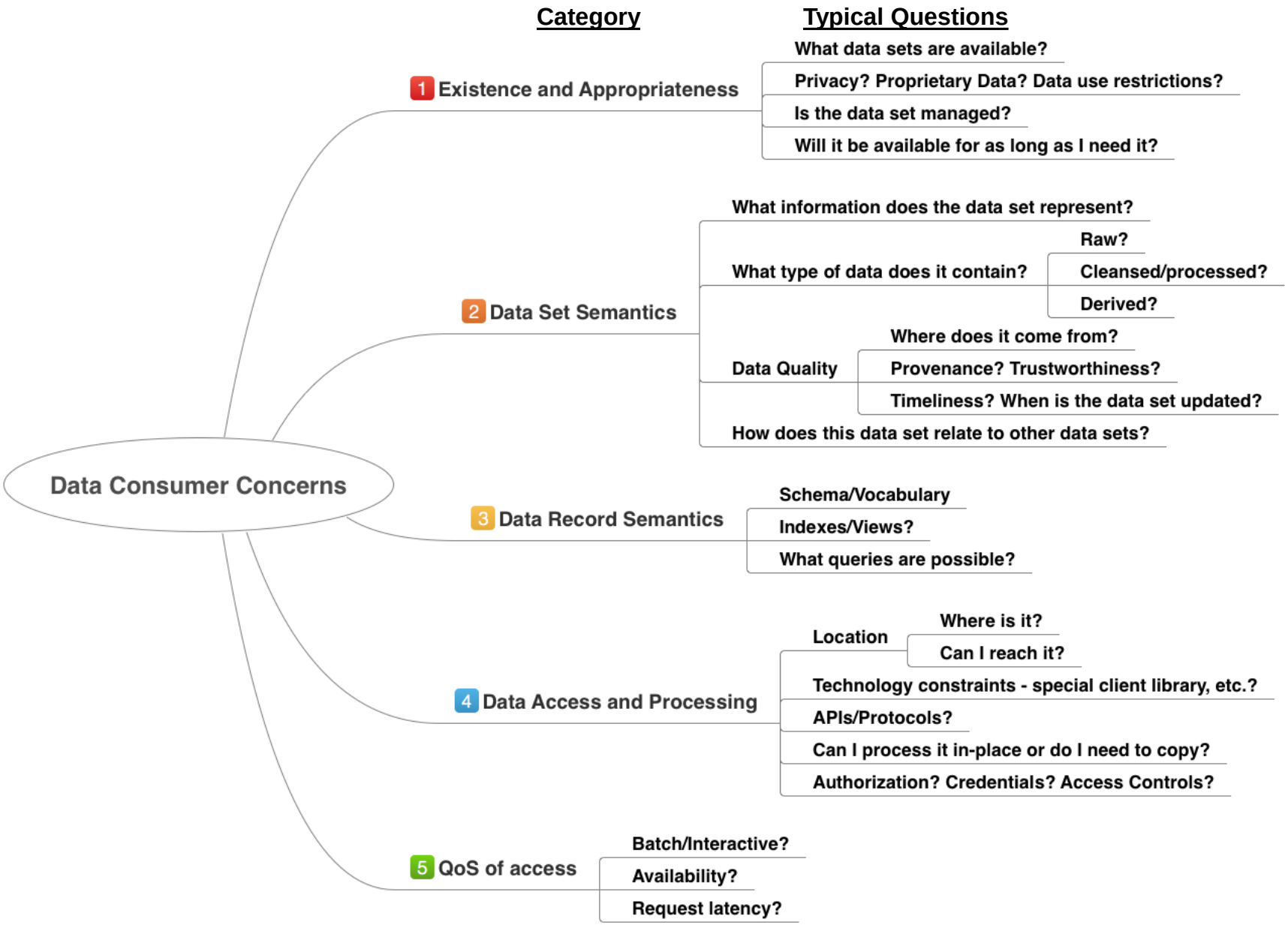
6. Focus your governance on the things that data consumers want











Playbook for Data Governance

Identify your high benefit decisions

- Infrequent but high impact, high frequency but low impact, …

Identify the data sets that support your highest benefit decisions

What is the producer/consumer relationship for each high benefit data set?

- 1-1 → Probably no governance needed
- 1-Many, Many-Many → Consider applying governance

What to govern = what constraints to impose on the producer?

- Use Data Consumer Concerns checklist – where are your gaps?
- Balance *cost* and *benefit* to keep *value* positive

Work down the list of high benefit decisions

Review periodically – have your high benefit decisions changed?

Contact Information

John Klein

Senior Member of the Technical Staff
Architecture Practices Initiative
+1 412-268-7378
jklein@sei.cmu.edu

U.S. mail:

Software Engineering Institute
Customer Relations
4500 Fifth Avenue
Pittsburgh, PA 15213-2612
USA

World Wide Web:

www.sei.cmu.edu
www.sei.cmu.edu/contact.html

Customer Relations

customer-relations@sei.cmu.edu
+1 412-268-5800