From Business Requirements to Ontology Evaluation

How to get from extrinsically defined needs to meaningful and effective evaluation of ontology suitability

Amanda Vizedom

Ontology Summit 2013

Track B: Extrinsic Aspects of Ontology Evaluation

28 February, 2013

Focus & Agenda

- Focus:
  - Less on particular ontology characteristics and corresponding evaluation methods and metrics
  - More on figuring out what kinds of evaluation should be applied to ontologies for particular applications.

- Agenda
  - Outline of a process:
    - Starting with business*/operational* requirements
    - Resulting in ontology evaluation plan that guides development and/or selection toward most suitable, high quality ontologies
  - Example with process walk-through
  - Review

Outline of a Process

1. Identify Business Requirements (received)
2. Refine Business Requirements
3. Derive Technical Requirements (System)
4. Identify Ontology (and other component) Requirements
5. Identify Ontology Evaluation Elements
6. Develop Recommended Ontology Evaluation Details
7. Incorporate Ontology Evaluation Plan into general project plan (lifecyles, relationships to other project elements; resource needs & use; schedule …)

NB:

- This is a simplified outline, emphasizing distinct tasks & how they enable each other.
- In reality, may not single or linear; may have feedback & other cycles, iterations.
- Additional structure may come from specific Software Engineering, Systems Engineering, Requirements Engineering, QA practices

Expertise Required (who does what?)

1. Identify business requirements

2. Refine business requirements (elicit and develop detail to drive technical design)

3. Derive Technical Requirements (System)

4. Derive Ontology Requirements ⟺ identify other component & sub-system requirements

5. Identify Ontology Evaluation, QA Elements relevant to requirements — identify component, system, and sub-system QA elements

6. Develop Recommended Ontology Evaluation Metrics, Methods, when they should be applied in lifecycle*, document risks by omission

7. Incorporate into project plan, including lifecycles, resource needs and use, schedule, etc.)

Requires: - understanding of business / operational context and issues
- technical understanding of ontology engineering & architecture
- understanding of project management practices and issues

Too Common: Not for Ontology

1. Identify business requirements

2. Refine business requirements (elicit and develop detail to drive technical design)

3. Derive Technical Requirements (System)

4. Derive Ontology Requirements ⟺ identify other component & sub-system requirements

5. Identify Ontology Evaluation, QA Elements relevant to requirements — identify component, system, and sub-system QA elements

6. Develop Recommended Ontology Evaluation Metrics, Methods, when they should be applied in lifecycle*, & risks by omission

7. Incorporate into project plan, including lifecycles, resource needs and use, schedule, etc.)

Better: SWE/SE practices encompass ontology requirements, evaluation, QA

1. Identify business requirements

2. Refine business requirements (elicit and develop detail to drive technical design)

3. Derive Technical Requirements (System)

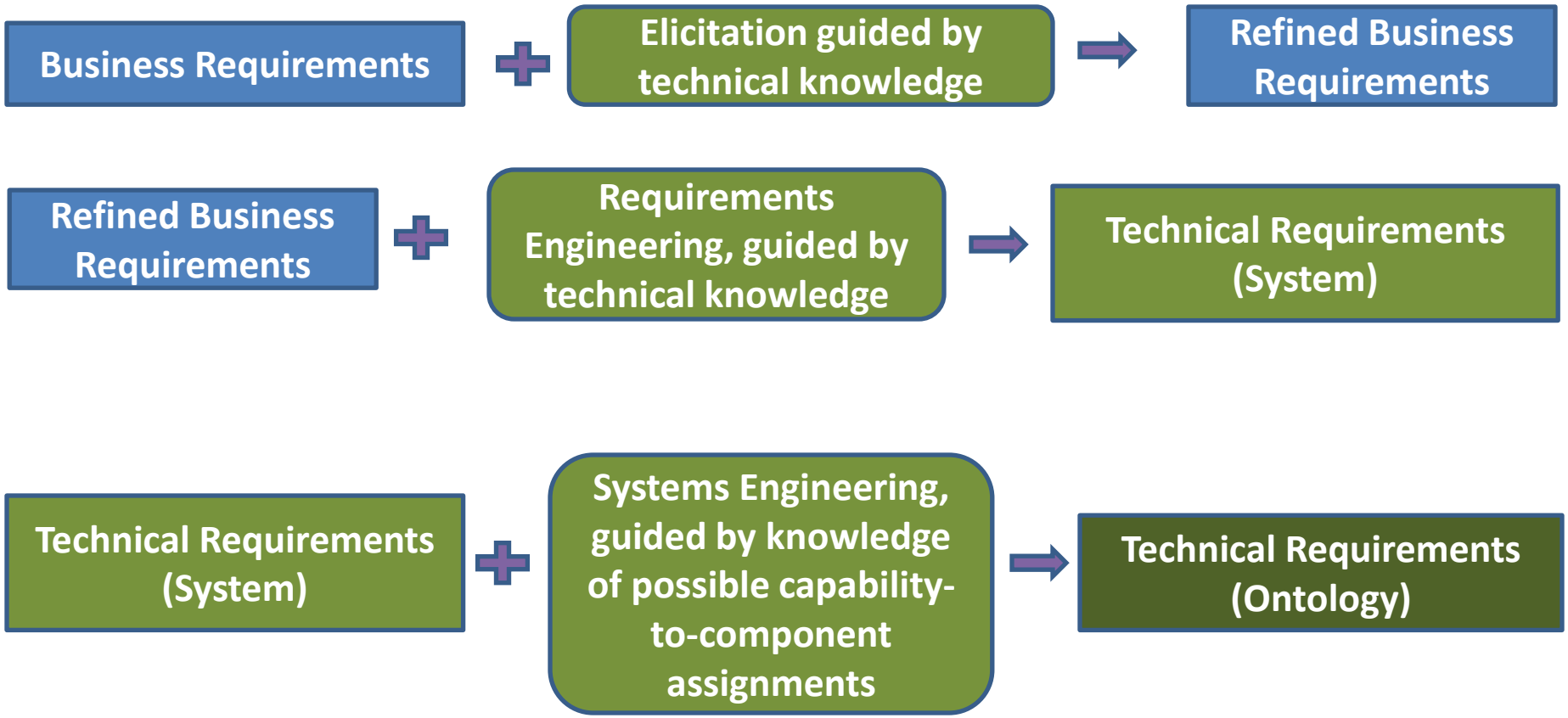
4. Derive Ontology Requirements ⟺ identify other component & sub-system requirements

5. Identify Ontology Evaluation, QA Elements relevant to requirements — identify component, system, and sub-system QA elements

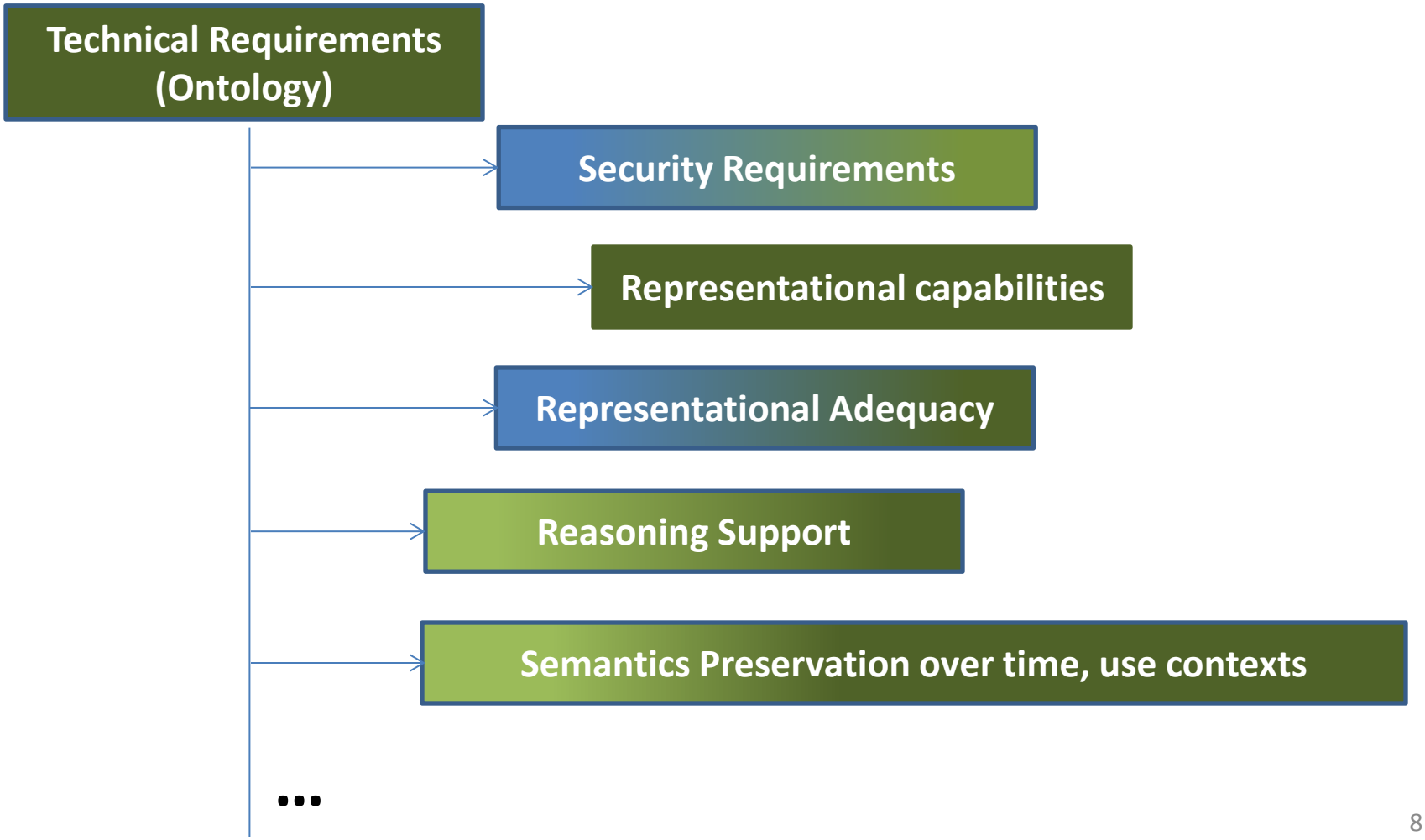
6. Develop Recommended Ontology Evaluation Metrics, Methods, when they should be applied in lifecycle*, & risks by omission

7. Incorporate into project plan, including lifecycles, resource needs and use, schedule, etc.)

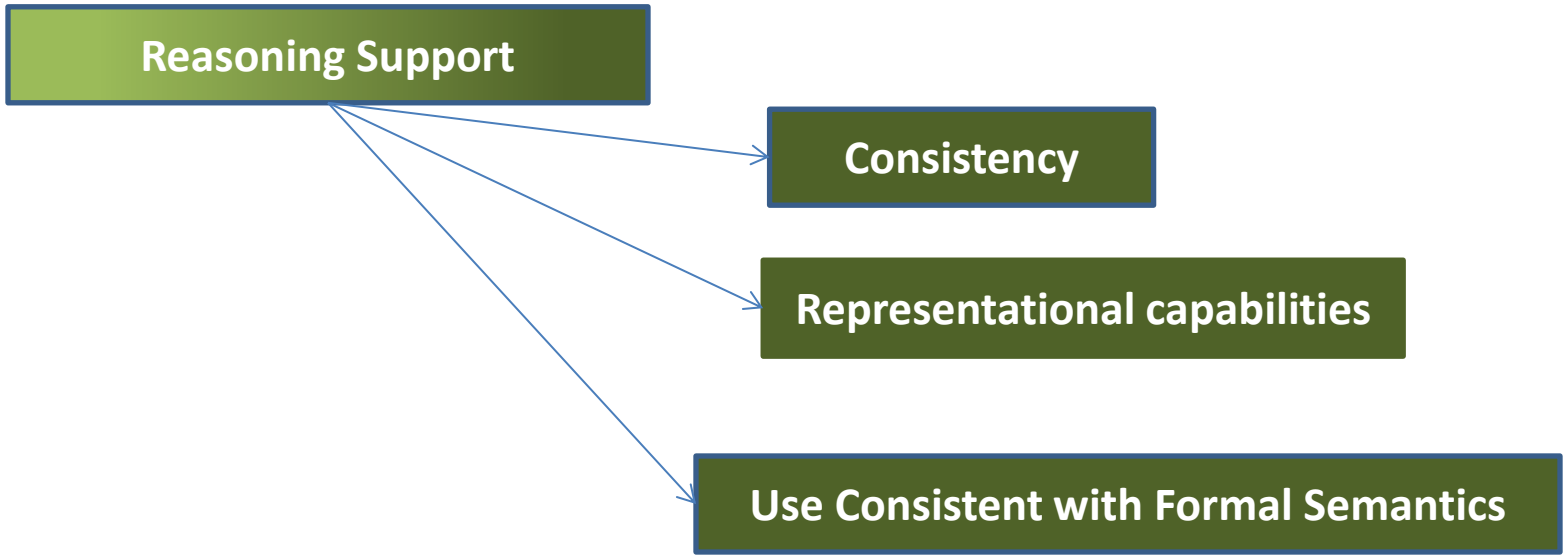
"Extrinsic" Considerations (from ontology POV) lead to Technical Requirements



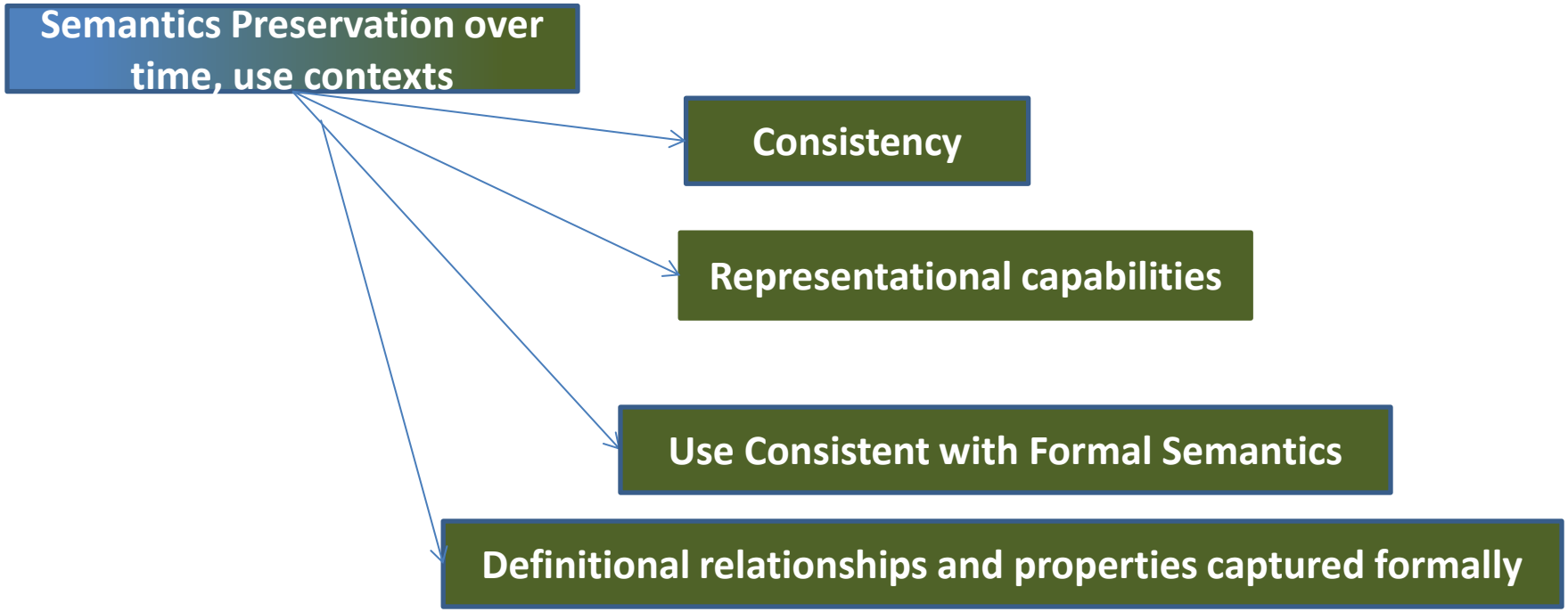
Technical Requirements are mixed "extrinsic" "combined-relational" and "intrinsic" (from ontology POV)



“Intrinsic” Ontology Aspects Matter from “Extrinsically”-Grounded perspective



“Intrinsic” Ontology Aspects Matter from “Extrinsically”-Grounded perspective



Process Walk through Example

- FredCorp: An Enterprise Search Story

FredCorp: The Story

- FredCorp has lots of internal documents, email, intranet portal posts, records kept by various internal departments ...
- FredCorp has clients & offices in the US, Brazil, France, & Japan.
- FredCorp staff often can't find the documents they are looking for. They tried intranet environments such as Sharepoint. They tried keyword-based search applications. They want a better search solution for internal enterprise search and retrieval.
- Some key players at FredCorp saw a semantic search solution being demonstrated at a conference.
- They followed up with the vendor.
- They decided that they like some things about it, but they don't think it exactly fits their needs, e.g.,
  - Only English & French editions; search & docs must be in same language;
  - annual license fees are very steep;
  - maintenance and configuration changes require ongoing vendor contracts
- They have a lot of top-notch technical folks, and they think they can probably build something better suited for less money.

FredCorp: Business Needs (received)

1. Identify business requirements

- Bottom line: FredCorp wants a system that support semantic search and retrieval of documents on FredCorp Intranet.
- It should do better than keyword search or high-structure portals they have tried.
  - Users should find the right documents more of the time
  - Users shouldn't have to spend so much time wading through the wrong documents
- Any FredCorp user at any office should be able to use the system to find any document (except for certain kinds of documents that are protected by law or policy)
- It should be maintainable and configurable by FredCorp staff; potentially to include a few extra dedicated specialists, but not too many.

FredCorp: Refined Business Needs

2. Refine business requirements (elicit and develop detail to drive technical design)

Added Requirements and details

- Support current office language localization & future non-disruptive expansion
- Require and utilize intranet sign-on and profiles
- Support restricted access to current, specified document types & future non-disruptive revision
- Use existing data (user accounts, profiles), from existing data sources, to inform access behavior.
- [Desired, not required] Support restricted access based on document content as well as categories
- Index, handle queries, & retrieve documents using concepts, terminology, & relevance factors from:
  - FredCorp's primary business domain, including any industry standard concepts and vocabulary
  - FredCorp's business activities and processes, as may be referenced in internal documents.
  - FredCorp's internal, supporting units and their activities (e.g., IT, security, HR, AP, AR, facilities, sales, marketing, customer support, engineering, finance, …) including any activity-standard concepts and vocabulary
- Handle all current document formats & support future non-disruptive expansion
- Handle all current document stores & support future non-disruptive expansion
- Be effectively useable by FredCorp employees with little to no special training.
- Cost no more than $X to build, $Y/year to maintain (including any special personnel & equipment)
- Use cases showing expected user/system/document interactions
- [Desired, not required] Noting coming challenges for company in area of regulatory compliance self-monitoring, support system, sub-system, or component reuse and/or extension in support of compliance monitoring solution down the road.
- Allow users to set up monitors that automatically alert them to new documents matching specified queries
- Be stable and reliable, sufficient for mission-critical functions.

…

FredCorp: System Technical Requirements

3. Derive Technical Requirements (System)

- Perform cross-lingual indexing and query-based search & retrieval of FredCorp unstructured, semi-structured, and structured documents
  - Access documents anywhere on intranet
  - Analyze contents (structure analysis, text analysis, disambiguation, matching of document elements to language-independent concepts)
  - Create index of document occurrences of concepts
  - Analyze queries to match to concepts in index
  - Retrieve documents based on relationship between query concepts and indexed concepts
  - Refine retrieval based on likely relevance factors
  - Be usable and intuitive by FredCorp employees at all levels and in all departments
  - Be feasible given cost, schedule, and resource constraints
  - Exceed current base P/R of search using prior or current systems.
  - Return results within n seconds
  - [lower priority] support future reuse or extension for compliance monitoring
  - …

FredCorp: Ontology Requirements

4. Derive Ontology Requirements

- Representation Expressiveness sufficient to capture concepts referenced in FredCorp documents & relationships between those concepts that affect relevance.
- Covers above concepts (e.g., FredCorp primary business domain, internal unit support activity concepts, customer concepts)
- Includes multi-lingual concept lexicalization, supports use by NLP components
- Includes relationships between concepts as needed for relevance reasoning
- Architecture supports low-cost concept & lexical coverage extension
- …

FredCorp: Ontology Evaluation Elements

5. Identify Ontology Evaluation, QA Elements relevant to requirements

- Coverage sufficiency, lexical sufficiency
- Representational accuracy (concepts are modeled in keeping with operator/SME usage and domain object relationships
- Logical consistency
- …
- Relevance-affecting relationships specified

FredCorp: Ontology Evaluation Recommendations

6. Develop Recommended Ontology Evaluation Metrics, Methods, when they should be applied in lifecycle*, & risks by omission

- Measure coverage against FredCorp corpus using independent NLP tools and/or system indexing performance (may vary across stages)
- Measure representational accuracy
  - Directly through validation by operators/SMEs, if sound interface for such validation available
  - Indirectly through performance of integrated system, using operator/SME validation of retrieval relevance
  - Indirectly through performance of integrated system, using click-through data gathered automatically
- Check logical consistency continuously
- Check for concepts without lexicalization, key relationships

FredCorp: Incorporation of Ontology Evaluation Recommendations into project plan

7. Incorporate into project plan, including lifecycles, resource needs & use, schedule, …)

- Handled by managers, but should include input & validation by technical leads, including ontology
- Systems and ontology lifecycles should be in synch.
- Routine testing (nightly, pre-build, regression, pre-deployment) should include ontology testing
- Ontologists should be expected to support testing and be accountable for results
- Plans, including resources (staff & tools), informed by needs.

Review & Research Needs

1. Identify business requirements

2. Refine business requirements (elicit and develop detail to drive technical design)

3. Derive Technical Requirements (System)

4. Derive Ontology Requirements

identify other component & sub-system requirements

5. Identify Ontology Evaluation, QA Elements relevant to requirements

identify component, system, and sub-system QA elements

6. Develop Recommended Ontology Evaluation Metrics, Methods, when they should be applied in lifecycle*, document risks by omission

7. Incorporate into project plan, including lifecycles, resource needs and use, schedule, etc.)