Natural Ontologies:
Understanding and Addressing Alignment Issues

Kurt Conrad

The 9th International Protégé Conference
July 23rd 2006, Stanford Campus, Palo Alto, California

Introduction

- What is the Ontology Management Team?
- Issue: Semantics are Naturally Dynamic.
- Approach: Analyze Transformational Origins of Semantic Changes
- Strategy: Formalize MetaKnowledge
- Approach: Analyze Impact of Human Values on Semantic Association
- Method: Semantic Consensus Acceleration

The Ontology Management Team

- OMT is a spin-off of the Ontolog Forum
- Particpants

Kurt Conrad	Conrad@SagebrushGroup.com (408) 247-0454
Bob Smith	Bob@1TallTrees.com (714) 536-1084
Bo Newman	Bo.Newman@KM-Forum.org (540) 459-9592
Joe Beck	Joe.Beck@EKU.edu (859) 622-6359

OMT Mission and Focus

- Clearly differentiate Ontology Management and Engineering activities.
  - Understand factors driving effective management of ontology engineering projects.
- Bridge theoretical, problem, and engineering domains
  - Produce improved methodologies for ontology development.
  - Develop reliable methods for driving (natural) ontological alignment within working groups.
  - Identify and develop methods to ensure quality and alignment of needed conceptual models.
- Leverage new insights and methodologies to deal with more general issues of policy development within large enterprises.

OMT Policy Vector #1

- Issue: Implicit policymaking by technologists.
  - Management's abdication of policymaking responsibilities.
  - Occurs when the policy implications of design decisions are poorly understood.
  - Risk increases when dealing with emerging technologies.

- Strategy: Make policy decisions transparent.
  - Identify critical decisions that have significant downstream policy impacts.
  - Develop analytical and governmental processes to enable understanding and resolution of associated policy issues by the appropriate stakeholders.

OMT Policy Vector #2

- Issue: Disruptive impact of incomplete consensus and Dynamic Semantics on ontology engineering efforts.

- Strategy: Expand ontology development methodologies to deal effectively with Dynamic Semantics.
  - Understand the role of natural ontologies in sensemaking.
  - Develop methods to assess semantic distance, volatility, and drift.
  - Develop procedures for negotiating and “narrowing-the-gap” when semantic conflicts are encountered.
  - Develop a “language” for defining semantic policy that is usable by both policy makers and ontological engineers.

OMT Policy Vector #3

- Issue: Other areas of organizational behavior are increasingly experiencing the alignment issues typically associated with ontological engineering initiatives.

- Strategy: Generalize ontology management practices to:
  - Deal with broader issues of organizational meaning.
  - Resolve semantic issues in other policymaking domains.
  - Balance and prioritize semantic alignment efforts across initiatives.
  - Establish benchmarks for semantic accountability.

Issue: Semantics are Naturally Dynamic

Semantics are Naturally Dynamic

- Initial uncertainties and inevitable semantic changes undermine the alignment of formalization efforts.
- Dynamic Semantics (DS) results from the interplay of three agent types and their associated ontologies
  - Individual Agents
  - Social Agents
  - Automated Agents

Formalization Doesn't Stabilize Semantics

- Meaning isn't an inherent property
  - Ultimately the product of human imagination and creativity
- Complex set of mechanisms both drive and limit changes to perceived meaning
- DS ultimately impact automated systems
  - From individuals (changes to conceptualization)
  - From groups (dynamic / evolving consensus)
- Making semantics explicit doesn't necessarily limit or slow upstream change

Approach: Analyze Transformational Origins of Semantic Changes

Analyzing Dynamic Semantics

- While analysis of agents and agent types can be used to identify DS, it doesn't provide enough detail to drive specific responses
- Need more sophisticated models that let us look beyond agent types to specific Semantic Classes and properties
  - Knowledge Vector Model
  - MetaKnowledge Continuum Model

Knowledge Vector Model

$$D \; f \; I \; f \; K \Rightarrow KU$$

- Describes a continuum of Knowledge Artifacts
  - Data (D) is transformed (*f*) into Information
  - Information (I) is transformed into Knowledge
  - Knowledge (K) represents the point of actionable synthesis of all event-specific (K_E) and prior (K_P) knowledge
  - K enables a Knowledge Utilization Event (KU)
  - KU is either an action or a decision
- Boundaries are not necessarily discrete

MetaKnowledge Continuum Model

$$D \; f \; I \; f \; K \Rightarrow KU$$

mD mI mK

- Also describes a continuum without discrete boundaries
- MetaKnowledge (MK) comprises a range of KAs
  - Like the generic term MetaData
  - The term MetaKnowledge is used deliberately
- mD, mI, and mK (specific to this model) document
  - Critical semantic properties of each class of KA
  - Knowledge about the transformations that produced them

Interpretive Semantics

- Deal with the interpretation and meaning of symbols
- Knowledge about how symbols map to concepts
  - Potentially ambiguous (multiple meanings, double entendres)
  - Result of observational and symbol selection behaviors
- Answers the question: What is it?
- Maps to mD in the MK Continuum Model

Contextual Semantics

- Deal with context and pattern recognition
- Knowledge about how KAs (or parts) relate to
  - Transformation and representation behaviors
  - Other KAs
  - The real world
- Answers the questions
  - What kind? What is it about?
  - Who, where, when?
  - How, especially "How does this fit?"
- Maps to mI in the MK Continuum Model

Contextual Semantics Principles

- Appears to be the “first place” that values play a significant role in sense-making
- Meaning often expressed in historical terms that imply future semantics
  - “You always..”
  - “They always…”
  - “It always…”
- Contextual Semantics often implied, incomplete, or missing
  - Supplied by the interpreting agent

Aspirational Semantics

- Deals with underlying motivations, drivers, rationality
- Knowledge about how KAs are synthesized and optimized to enable behavior
- Answers the question “Why?”
  - Infrequently documented
  - Often tacit and implicit
- Individuals provide, when missing
  - Routine source of semantic breakdowns
- Maps to mK in the MK Continuum Model

Behavioral Semantics

- Meaning, as described in behavioral terms
- Often involves complex semantic chains (scenarios) which comprise
  - Events
  - Conditions
  - Other behaviors
- Representations range from tacit to explicit
  - Culture
  - Law

Strategy: Formalize MetaKnowledge

Addressing Dynamic Semantics

- Ignore semantic volatility
  - Leave critical semantic properties largely implicit and tacit
  - Appropriate for non-critical issues
- Stomp
  - Make semantic properties more explicit
  - Formalize without addressing sources of volatility
  - Good when you can get away with it
- Embrace
  - Understand change vectors
  - Balance explicit, implicit, and tacit representations
  - Implement mechanisms to identify and/or leverage “natural” misalignments as they emerge

Semantic Nirvana / Artificial Utopia

- Formalization is optimized for computerized inference
  - First Order Logic typically considered the most expressive representation
- Two-step semantic resolution model
  - Symbol interpretation — Maps symbol to concept
  - Axiomatic component — Used to document, communicate, and potentially infer behavioral implications
- Limitations
  - Limited Contextual and Aspirational Semantics
  - Practical issues limit “expressiveness”
    - » Availability of subject matter experts
    - » Truth is unbounded but resources are limited
    - » Can you read KIF?

Dynamic Semantic Findings

- Formalizing meaning in the face of uncertainty is
  - Difficult, at best
  - Potentially chaotic
- Therefore, an improved ability to identify and understand Dynamic Semantics enables
  - More resiliency to be built in, in the first place
  - Less remediation. Fewer false starts
  - Changes to be more easily anticipated and reacted to
  - Semantic change to be used as a resource for enhancing delivered value

MetaKnowledge Formalization Strategy

- Identify areas of potential semantic conflict
  - Resolve conflict, as appropriate
  - Integrate semantic models, where each has value
- Avoid sub-optimization around machine-processable semantics
  - Evaluate each Semantic Class for potential value
  - Consciously balance or optimize the explicit representations
  - Document implicit and tacit MK triggers that are only usable by individuals and groups
- Where practical, expand the range of targeted KUs and associated behavioral semantics

MetaKnowledge Formalization Benefits

- Doesn't have to be difficult or expensive
- Supports and enables rapid prototyping
  - Scalable from small projects to large knowledge architectures
  - Reporting model starts "right" and improves through time
- Leverages implicit and tacit knowledge within the organization
  - Enables re-contextualization and Knowledge Perpetuation
  - Less brittle than other approaches
- Doesn't preclude use of logic-based formalizations
  - Can speed and document emergent consensus
  - Helps ensure alignment of human behavior with axioms
- Expected to be the foundation for many emerging Ontology Management practices

Approach: Analyze Impact of Human Values on Semantic Association

Decisionmaking is Expensive

- Relies on a series of complex and sophisticated activities:
  - Data collection and interpretation.
  - Organization and pattern recognition.
  - Identification of motivations, causalities, and implications.
  - Synthesize actionable knowledge.

- Evolutionary advantages to be gained by reducing cost (time and effort) of decisionmaking.
  - e.g., OODA Loops.

Values Cut Decisionmaking Costs…

- Usually abstracted (decontextualized) in order to be applied across a variety of behavioral contexts.
- Often comprise significant implicit and tacit knowledge components.
  - Including "truly" tacit knowledge that defies articulation.
- Represents a form of bounded rationality.
  - Doesn't just impact the amount of knowledge used in decisionmaking (sufficiency).
  - Impacts pattern recognition and other transformative behaviors.
- Enables much of the decisionmaking process to remain implicit and tacit, operating at a subconscious level.
- Values act as a "technology".

…But at the Risk of Poor Decisions

- Values can drive sub-optimal decisions.
  - Limit perception of contrary and mitigating evidence.
  - Reinforce understood and accepted interpretations.
  - Constrain associated semantics to those consistent with established behavioral patterns.
- Risk increases with changes across or within specific behavioral contexts.
- Risk increases when the impact of values on decisionmaking goes unrecognized.

Potential Organizational Issues

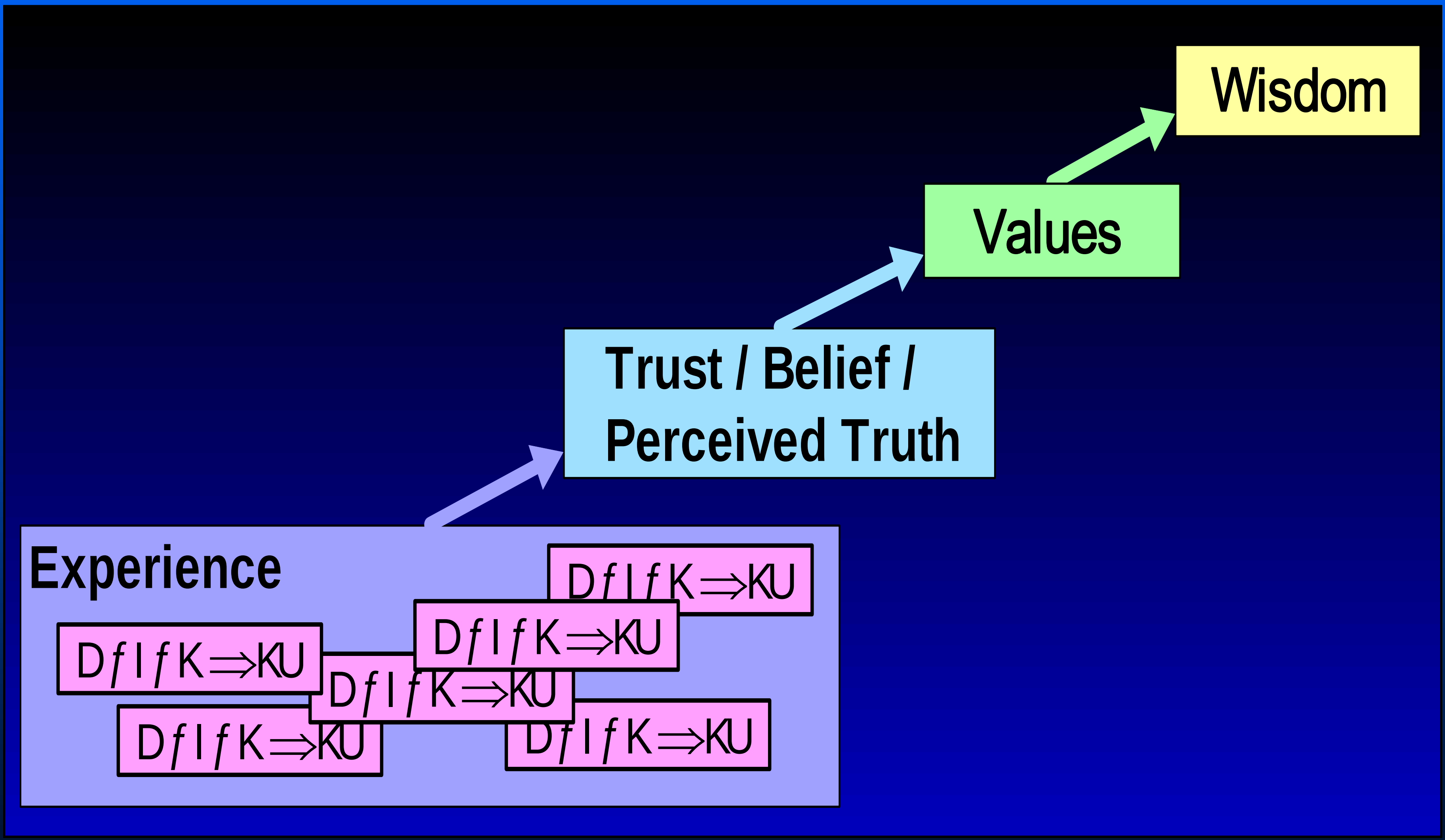
- When unaccounted for, implicit and tacit values in systems can drive:
  - Semantic breakdowns.
  - Polarization and conflict.
  - Group think.
    - Perpetuation of hidden biases.
    - Missed opportunities.
    - Inability to perceive risks.

Knowledge Vector Model

$$D \; f \; I \; f \; K \Rightarrow KU$$

- Provides basis for understanding the origin of values and the impact of values on semantic association.
- Agents use knowledge to execute behaviors.
  - Automated agents require all elements of the Knowledge Vector Model (KA, *f*, & KU) to be explicit.
  - Individuals and organizations can leverage implicit and tacit knowledge elements to "skip steps".
- People have a strong incentive to skip steps to make things easier.
- The scientific method does not come naturally.

Impact of Knowledge on Values



Values

- Values (including Principles, Interests, and Expectations) are KAs that represent networks of prior knowledge.
- Most importantly, values are routinely disassociated with major portions of the originating knowledge network.
- Disassociation is a form of knowledge compression.
  - Reduces processing time, communications time, recall time.
  - Allows values to be used as abstract KAs that can be applied in a variety of behavioral contexts.
- Wisdom relates to judgment and the ability to work intelligently with multiple sets of values.
  - Conflict avoidance, conflict resolution.
  - Balance short-term and long-term perspectives.
  - Avoid unary values-based, sub-optimal decisionmaking.

Semantic Segmentation Model

$$D \ f \ I \ f \ K \Rightarrow KU$$

Interpretive — Contextual — Behavioral — Aspirational

- Sensemaking is a multi-step process.
- Knowledge Vector elements are associated with a distinct classes of semantic properties.
  - Interpretive semantics.
  - Contextual semantics.
  - Aspirational semantics.
  - Behavioral semantics.
- Values directly impact the association of increasingly sophisticated meanings throughout the Knowledge Vector.

Values Impact Semantic Associations

- Perception.
  - What KA are considered irrelevant, noisy?
  - What KA are considered important, valuable, potentially useful?
- Interpretation.
  - What concepts are associated with inputs / symbols?
  - What referents are associated with the concepts?
  - How is ambiguity to be resolved?

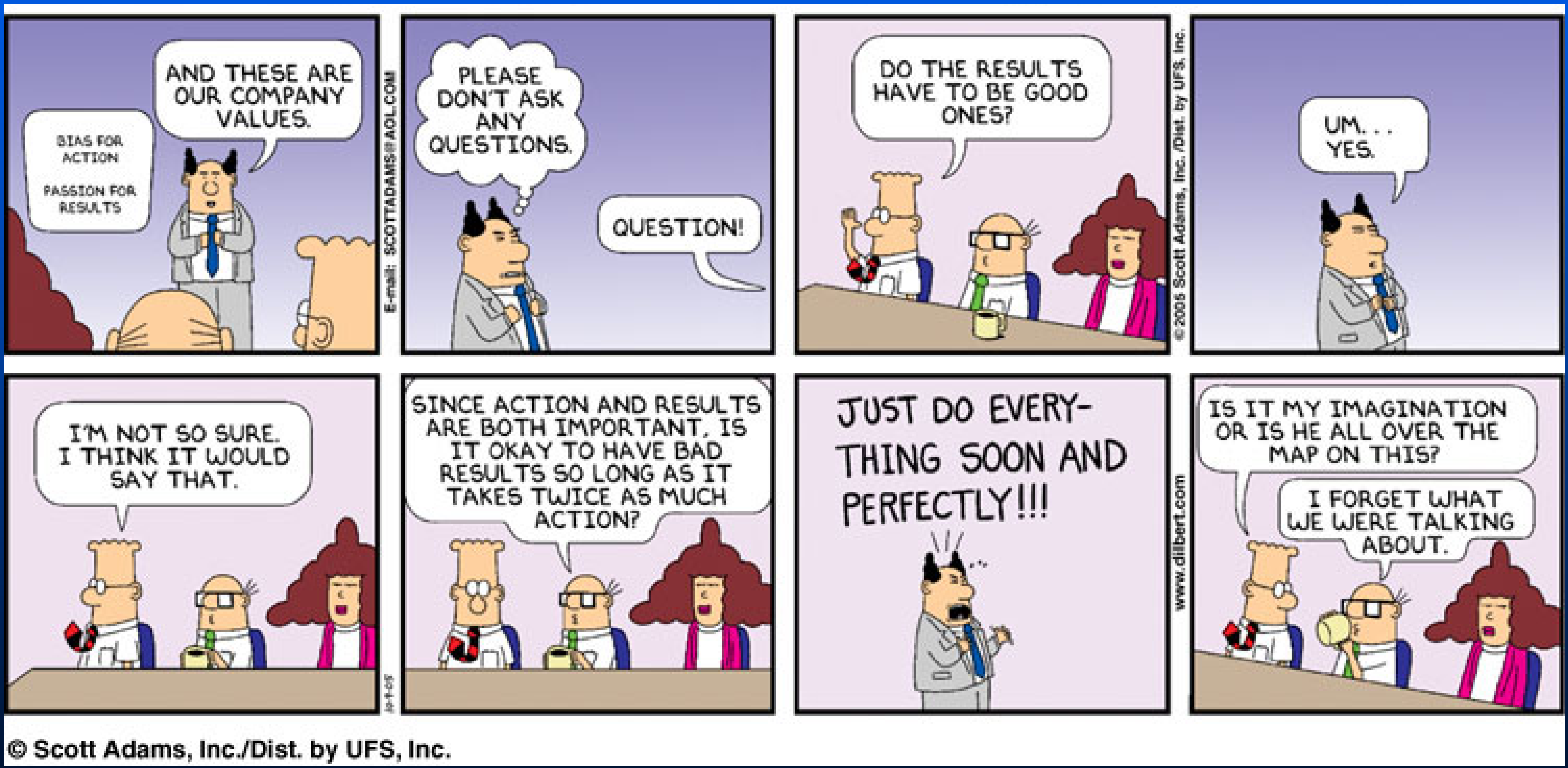
Values Impact Semantic Associations

- Contextualization.
  - Which context / sensemaking structure is most appropriate?
  - How does this KA fit? How should it be positioned?
  - What rules should be used to organize and relate KAs?
  - What patterns emerge? Are they useful or distractions?
  - What can be inferred? What implicit K is relevant / important?
  - Are the sources credible? Has this K proved its value in the past?
  - Is the K applicable at this time, in this situation?
  - Is this consistent with history and trends?

Values Impact Semantic Associations

- Aspirations.
  - How do the KAs relate to goals and objectives?
  - What are the motivations, intentions, and strategies of the relevant agents?
  - Do the agents support my interests or are they in conflict?
  - How does potential conflict influence K acquired from those sources?
  - Is the K actionable, or do critical K gaps exist?
- Behaviors.
  - What casual models apply?
  - What behaviors / results can be expected?
  - What are the risks and probabilities?
  - What alternatives and contingencies exist?

Ambiguous Semantics of Values



Conclusions

- Values, not facts, drive decisions.
- Values may represent the ultimate semantic technology.
- Values-based analysis methods:
  - Facilitate better ontological alignment across individuals and organizations.
  - Improve the quality of policymaking.
  - Stabilize organizational context, requirements, and specifications for engineering efforts.
  - Improve the quality of semantic formalization and articulation.

On the Horizon

- Values-based ontology development process.
- Introduction of values-based logic to complement existing axiomatic models of semantic formalization.
- Extend values-based conceptual alignment methods to address a broad range of policy development and governance issues.

Method: Semantic Consensus Acceleration (SCA)

SCA Origins

- Methodology developed to address a range of issues and initiatives.
  - Metadata framework to improve semantic interoperability.
  - Web search, retrieval, and navigation.

- Cross-cultural teams
  - Comprised marketing, engineering, technical documentation personnel.
  - Leveraged multiple value systems to illuminate semantic conflicts.
  - Given a Semantic Workpackage, comprising a set of related terms.
  - Identified and differentiated individual concepts and key relationships (including organizational scope).
  - Documented the level of consensus and any outstanding knowledge gaps.
  - Delivered conceptual “requirements” and “specifications” to engineering team.

SCA Process

- Step 1: Research Terms
  - List any additional terms and sources
- Step 2: Inventory Semantics
  - List concepts and draft / identify working definitions
  - Identify “core concept” and document supporting rationale
  - Map related concepts & relationships to core, document behavioral relevance
- Step 3: Normalize Semantics
  - Refine & validate working definitions to produce candidate definitions
  - Document supporting rationales
- Step 4: Normalize Terminology
  - Select normative term for each concept
  - Document term’s origins and history
  - Document any alternative semantics
  - List proper and improper aliases