Emerging Semantic Web Commercialization Opportunities

Ken Baclawski
Northeastern University

The Semantic Web: Its not just for searching anymore!

- The web is a versatile infrastructure for basic data availability.
- The main emphasis was on human-mediated interactions via web browsers but new uses are rapidly increasing.
- These new uses can benefit from the ontology based techniques and tools of the Semantic Web.

Identifying an opportunity

- Domain knowledge
  - Technical background
  - Community organization
- Identify urgent needs
- Understand the trends
  - Short-term evolution
  - Possible paradigm shifts
- Recognize an opportunity

Outline

- Interoperability and integration of legacy systems
- Web services and composite applications
- Records management
- Uncertain, incomplete and conflicting information
- Decision and policy making
- Collaboration tools
- Recent developments

Interoperability of legacy systems

- Legacy systems and databases are characterized by:
  - A large variety of formats
  - High degree of complexity
  - Many technologies of various ages
- Need to interoperate and integrate
- Trend is toward encoding more semantics in the data representation itself
- Opportunity to develop products and services for interoperability and integration.

Web services and composite applications

- The web is being used not only for retrieval of data but also for using tools and services.
- The need is to find the required services, and to get them to communicate with each other.
- The trend is to use semantic annotation to describe/advertise services, to express requests, and to represent the responses, but very unevenly.
- The opportunity is to built agile workflow management tools that can deal with the differing levels of semantic annotation.

Simple Semantic Web Architecture and Protocol (SSWAP)

- SSWAP is a protocol for semantic web services. See http://sswap.info
- Unlike other protocols, SSWAP uses a single format and protocol for description, registration, discovery and invocation.
- SSWAP was developed using OWL as its basis, and OWL inference is fundamental to its operation.

Records management

- Solving the electronic health record problem will add little to the existing paper-based records if the systems are not interoperable.
- Simply automating paper-based processes has relatively little impact on productivity.
- Gains in efficiency and improved patient care require a change in the overall process of medical care delivery.

Records Opportunity

- Develop event ontologies that:
  - Support interoperability
  - Are independent of workflows and processes
  - Are compatible with existing processes
- Develop products that:
  - Assist organizations to evolve toward electronic data management
  - Serve the interests of many stakeholders

Reasoning with uncertainty

- The Semantic Web is an extension of the current web in which information is given well defined meaning… (Berners-Lee, Hendler & Lassila)
- The Semantic Web is based on formal logic for which one can only assert facts that are unambiguously certain.
- Unfortunately, there are many sources of uncertainty, such as measurements, unmodeled variables, and subjectivity.

The Bayesian Web

- The challenge is to develop a full-featured stochastic reasoning infrastructure, comparable to the logical reasoning infrastructure of the Semantic Web.
- The *Bayesian Web* is a proposal to add reasoning about uncertainty to the Semantic Web.

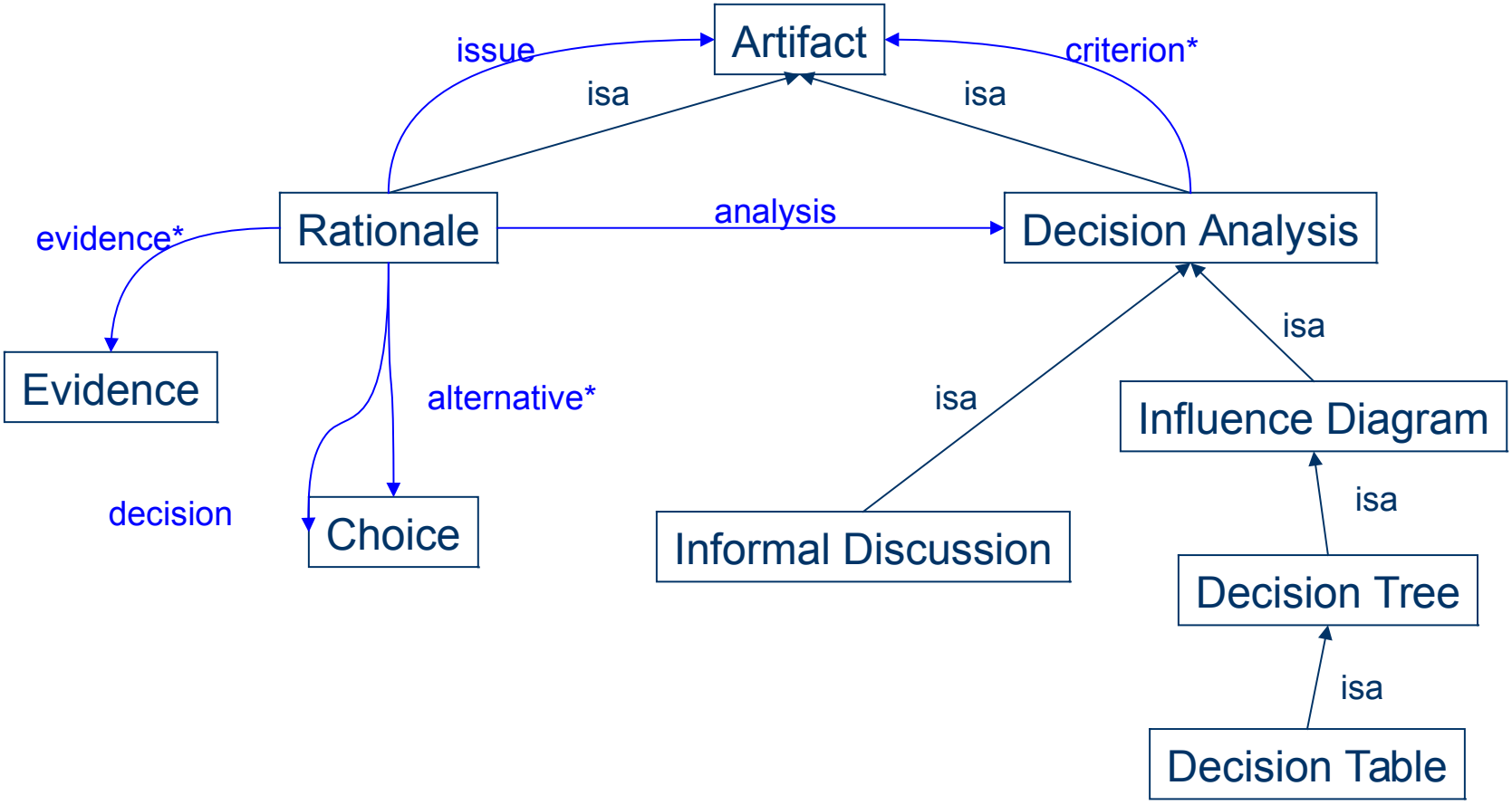
Bayesian Web facilities

- Common interchange format
- Ability to refer to common variables
- Context specification
- Authentication and trust
- Open hierarchy of probability distribution types
- Component based construction of BNs
- BN inference engines
- Meta-analysis services

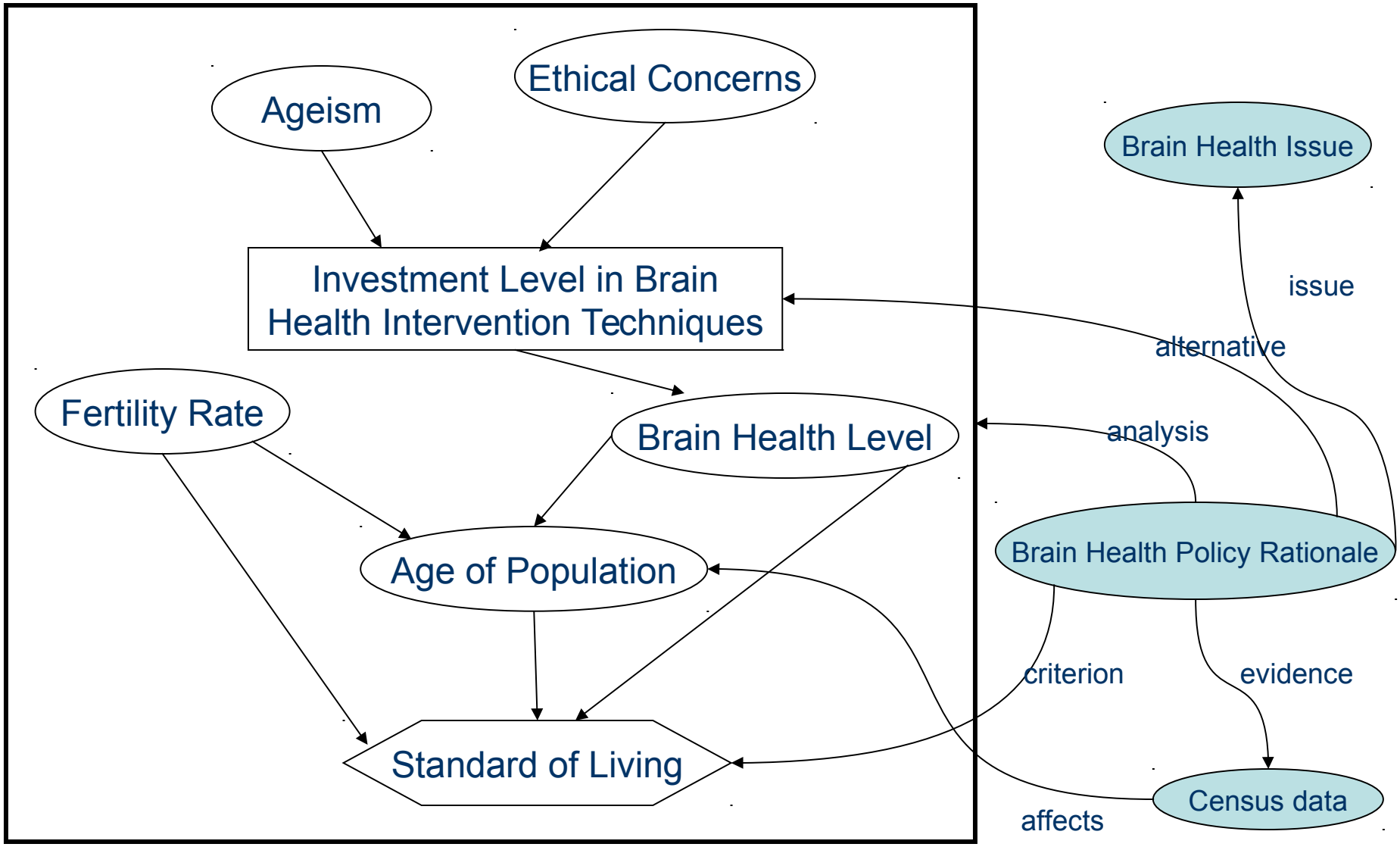
Decision Analysis

- Important part of policy and development processes.
- Formal annotation of decisions and their analyses can have many benefits.
  - Integration with the process
  - Recognition of need to reconsider when circumstances evolve
  - Decisions can be delayed
  - Decisions can be reused for other situations
- An annotated decision is called a **rationale**

Rationale Ontology



Policy Decision Example



Collaboration tools

- People need to collaborate to solve problems.
- The need is to support rapid team formation and problem solving even when the people are geographically dispersed.
- The trend is to use wikis and blogs rather than face-to-face meetings.
- The opportunity is to develop tools that facilitate collaboration over the web without losing the advantages of face-to-face meetings that make them desirable.

Wikis

- Wikis are a popular tool for collaboration.
- They have been used for rapid team formation and collaboration.
- They have a number of disadvantages:
  - Mix of natural language and untyped links.
  - Focus is on simplicity and presentation, not structure and semantics.

Semantic Wikis

- A wiki with an underlying knowledge model (ontology) is a *semantic wiki.*
- Data in the wiki is annotated with meta-data in RDF or OWL.
- Links are typed and annotated, also in RDF or OWL.
- Machines can infer new facts from the explicitly asserted facts.
- Search and retrieval are facilitated by the semantics.
- Interoperability is greatly improved.

Recent Developments

- RDF storage provided by database vendors
  - Oracle has both a product and an active Database Semantic Technologies Group
  - Many RDF stores are layered on a general purpose RDBMS: Jena, Sesame, RDQL, …
- Non-relational RDF storage products
  - Siderean, Tucana, OWLIM, Allegro Graph, …

Open Ontology Repository (OOR)

- Recent initiative of the Ontolog Forum
- The purpose of the initiative is to promote the global use and sharing of ontologies by:
  - 1. establishing a hosted registry-repository;
  - 2. enabling and facilitating open, federated, collaborative ontology repositories;
  - 3. establishing best practices for expressing interoperable ontology and taxonomy work in registry-repositories.