A Reference Architecture for Probabilistic Ontology Development

Semantic Technologies for Intelligence, Security and Defense 2013

Richard J. Haberlin
Paulo C.G. da Costa
Kathryn B. Laskey

November 13-15, 2013

EMSolutions

Background

Ontology Defined

An ontology is an explicit, formal representation of knowledge about a domain of application. This includes

- Types of entities that exist in the domain;
- Properties of those entities;
- Relationships among entities;
- Processes and events that happen with those entities;

where the term entity refers to any concept (real or fictitious, concrete or abstract) that can be described and reasoned about within the domain of application [Costa, 2005].

"An ontology is an explicit specification of a conceptualization [Gruber, 95]*."*

Ontology

- Ontologies provide a hierarchical structure of entity classes and a formal way of expressing their relationships
  - First-order expressivity
  - Supports logical reasoning
- There is significant literature on engineering traditional ontologies
- Ontologies lack built-in, principled support to adequately account for uncertainty
  - Annotating ontologies with simple probability annotations fails to convey structure of probabilistic representation
  - Less expressive probability schemes do not convey ontology structure, and so are inadequate

Probabilistic Ontology Defined

Ontology

Uncertainty

A *probabilistic* ontology is an explicit, formal representation of knowledge about a domain of application. This includes

- Types of entities that exist in the domain;
- Properties of those entities;
- Relationships among entities;
- Processes and events that happen with those entities;
- **Statistical regularities that characterize the domain;**
- **Inconclusive, ambiguous, incomplete, unreliable, and dissonant knowledge related to entities of the domain;**
- **Uncertainty about all the above forms of knowledge;**

where the term entity refers to any concept (real or fictitious, concrete or abstract) that can be described and reasoned about within the domain of application [Costa, 2005].

A probabilistic ontology extends a traditional ontology to represent uncertainty.

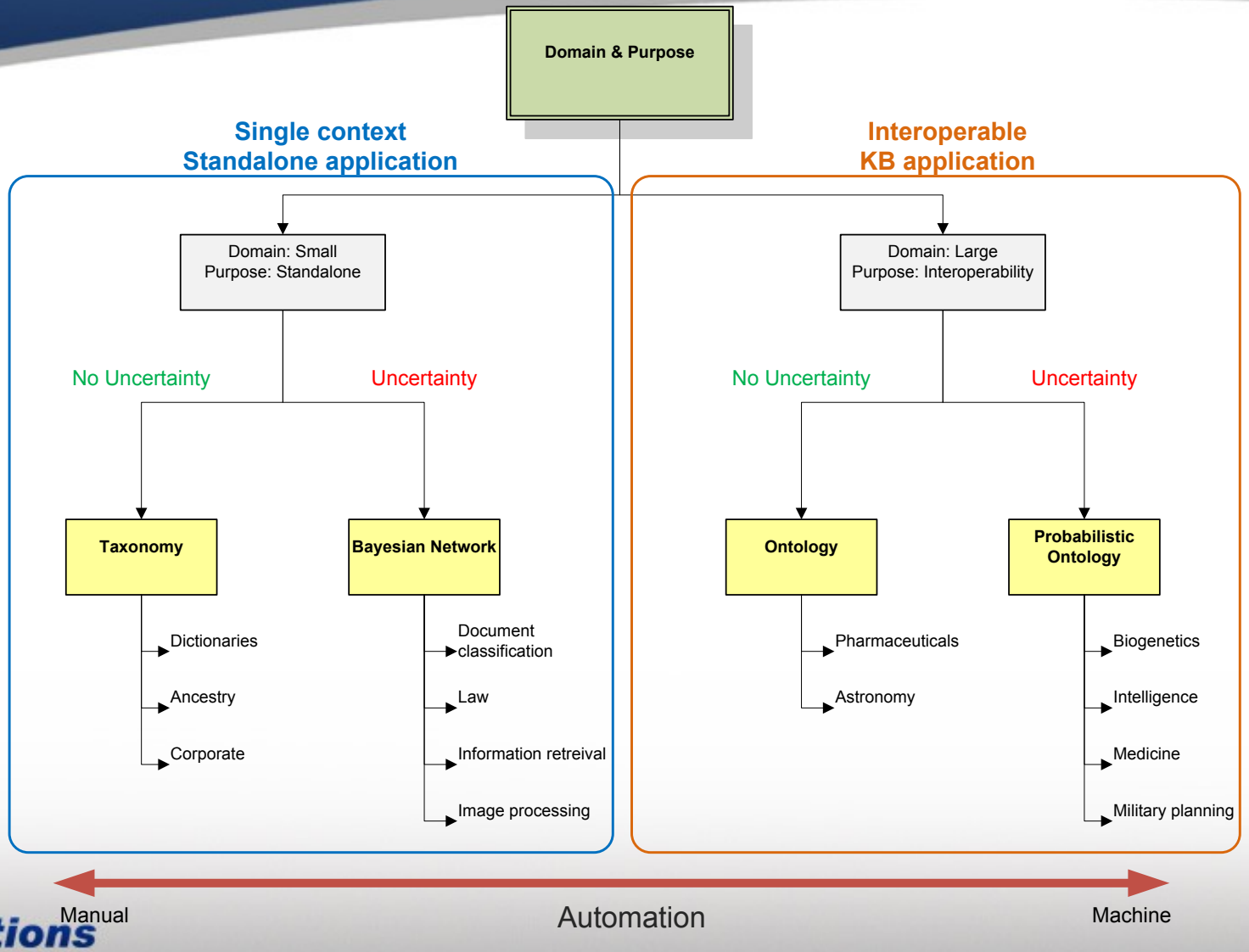
Probabilistic Ontology

- Integrates inferential reasoning power of probabilistic representations with first-order expressivity of ontologies
- Provides a means to represent and reason with uncertainty
- Limited literature on construction

Comprehensively describes knowledge about a domain and the uncertainty embedded in that knowledge in a principled, structured and sharable way [Brisset, 2003].

"It would be interesting to have a tool guiding the user on the steps necessary to create a probabilistic ontology and link this documentation to its implementation [Carvalho, 2011]."

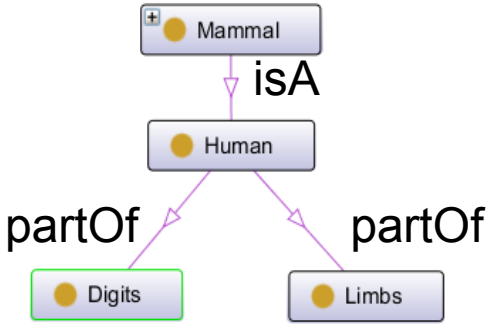
Selecting a Representation

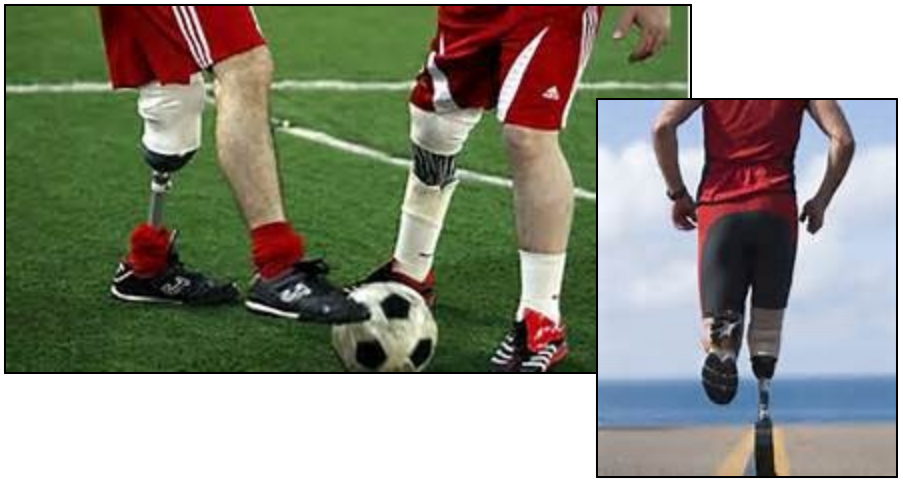


EMSolutions

Why Probabilistic Ontologies?

- Suppose an ontology of organisms contains the following classes and relationships:





- Humans *usually* have:
  - 2 arms & 2 legs
  - 10 fingers & 10 toes
- However, if a man loses a limb….
  - Is he no longer human?

Premise of an argument can be uncertain (e.g. Humans have 2 legs): (in)validity of the argument imposes no condition on the certainty of the conclusion (an amputee is Human).

EMSolutions

The Problem

- The Semantic Technologies (ST) community needs a comprehensive methodology for the development, implementation, and evaluation of probabilistic ontologies
  - Ontology use is on the rise
  - A means to incorporate uncertainty is a necessity
  - Limited literature on production of probabilistic ontologies
- Ontological engineering ensures ontologies developed for knowledge-sharing and reuse are explicit, logical and defensible
- Standard ontological engineering methods provide insufficient support for complexity of probabilistic ontology development

A similar methodology is needed for development of probabilistic ontologies

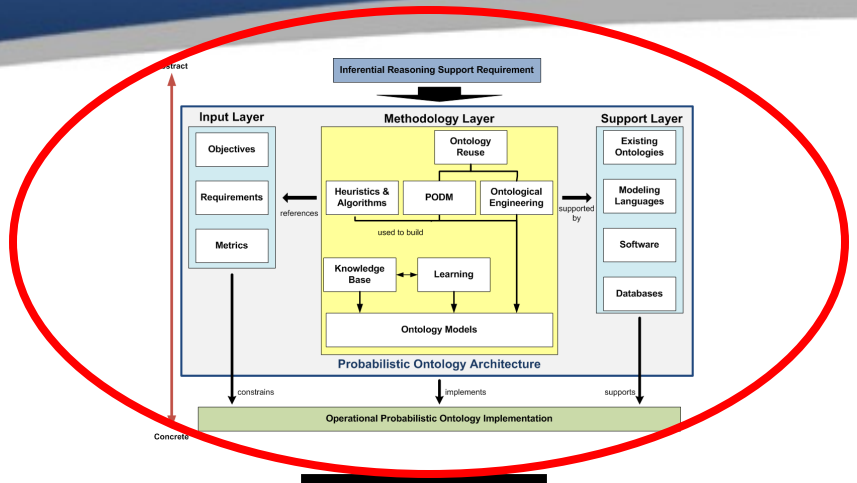
Our Solution

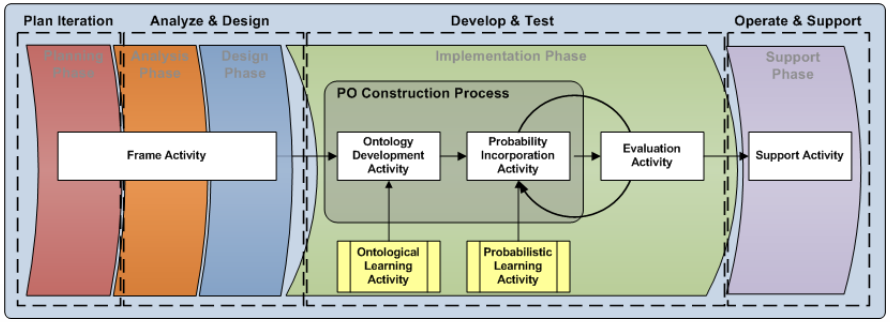
- Create a systematic approach to probabilistic ontology development
  - Facilitated through a reference architecture
    - Formalizes the application of the methodology
    - Extensible to various domains

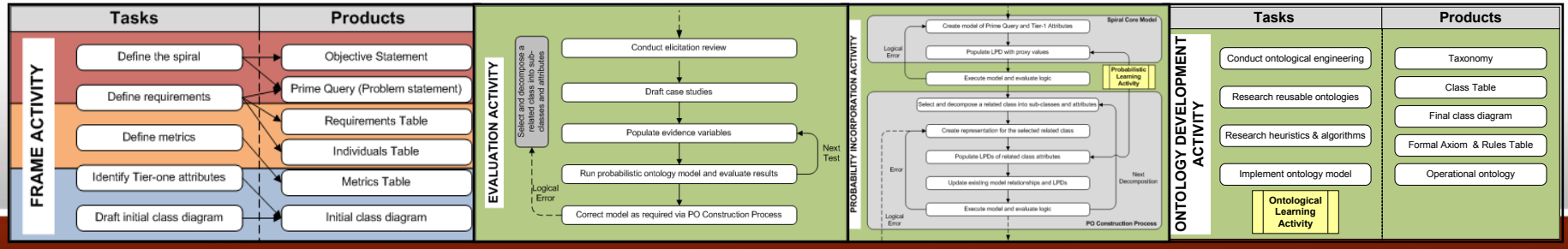
- Reference Architecture for Probabilistic Ontology Development
  - A generalized reference architecture designed to collect, catalogue, and define the components required for development of probabilistic ontologies and establish the criteria to be satisfied by any set of selected tools and methods

RAPOD provides a flexible solution

The RAPOD in Probabilistic Ontology Development







Reference Architecture for Probabilistic Ontology Development (RAPOD)

EMSolutions

Reference Architecture

- Provides a blueprint for architects to develop specific solution architectures within a defined domain.
  - Template for development
  - Defines integral components and their relationships
  - Reduces development time and project risk
- Standardizes language among participants
- Provides consistency of development within a domain
- Provides a reference for evaluation
- Establishes specifications and patterns

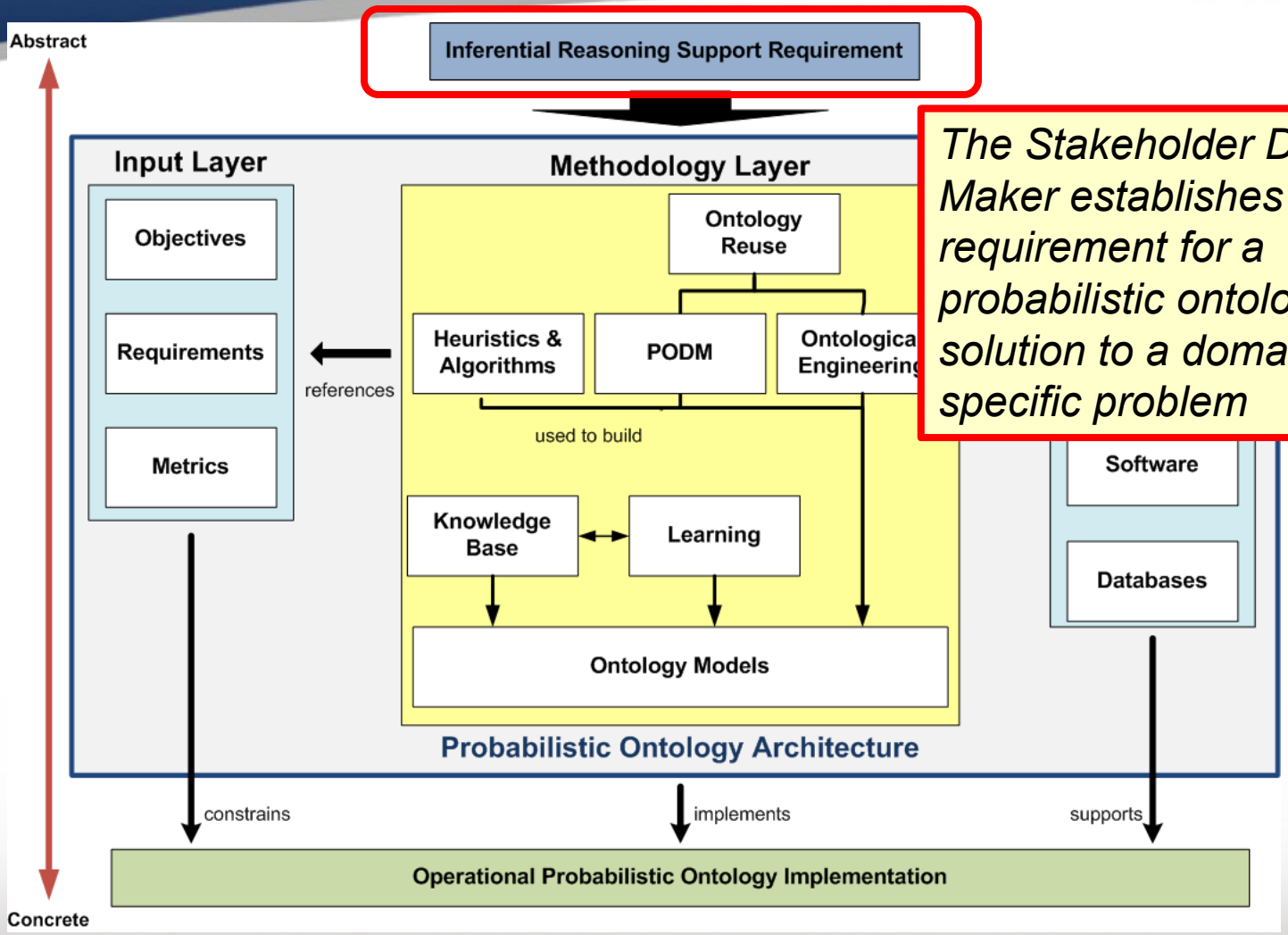
[A Reference Architecture is] "… an authoritative source of information about a specific subject area that guides and constrains the instantiations of multiple architectures and solutions [OASD/NII, 2010]*."*

RAPOD Introduction

- Provides synergy of effort within the ST community
  - Identifies concepts, processes, languages, theories and tools
  - Synergizes effort of probabilists, logicians, decision analysts, computer scientists
- Spans knowledge, processes, models and tools necessary to engineer POs at a high level of abstraction
- Output defines a domain specific architecture that may be used to produce probabilistic ontologies in similar domain contexts

RAPOD output is an architecture

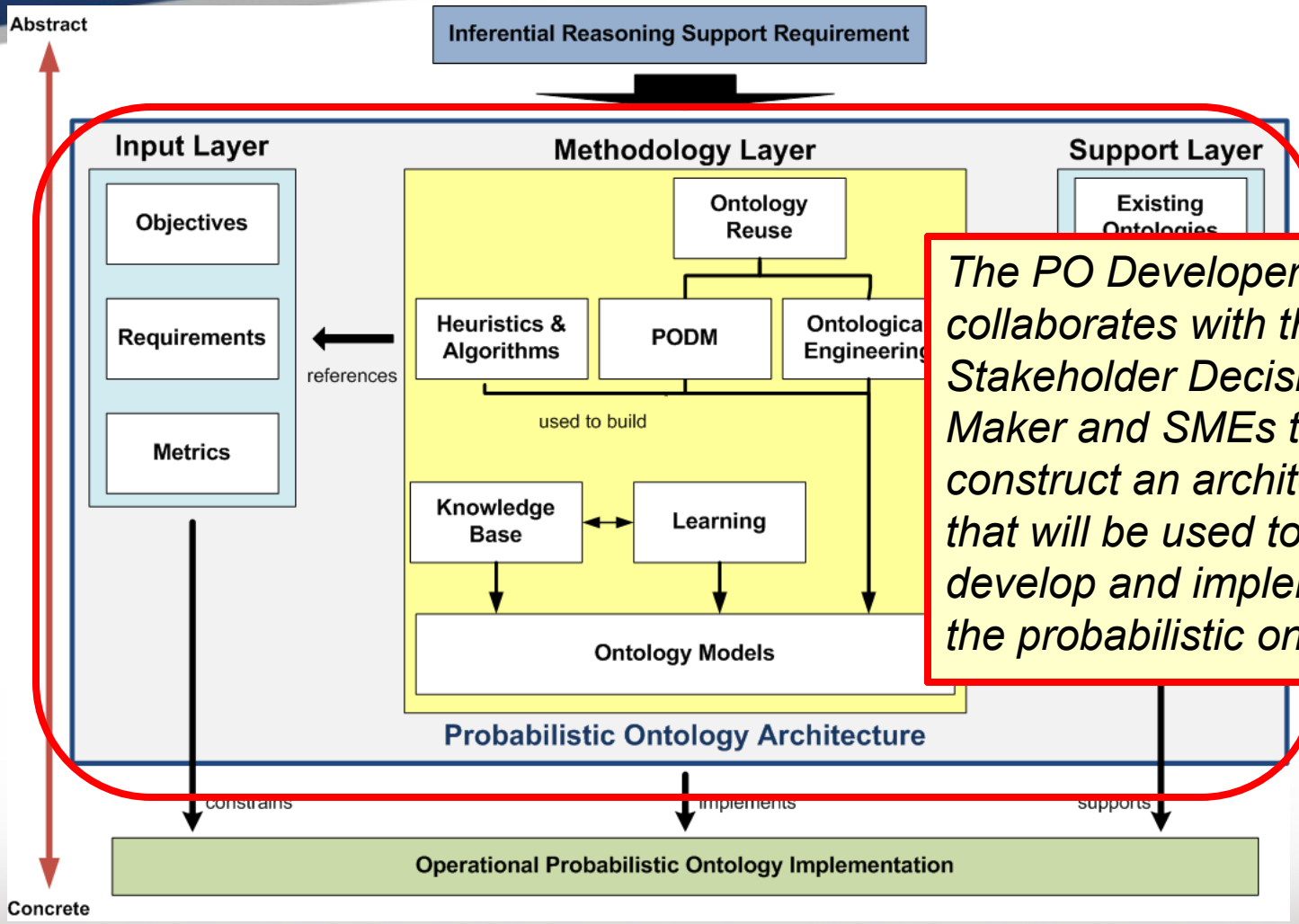
Reference Architecture for Probabilistic Ontology Development



The Stakeholder Decision Maker establishes a requirement for a probabilistic ontology solution to a domain-specific problem

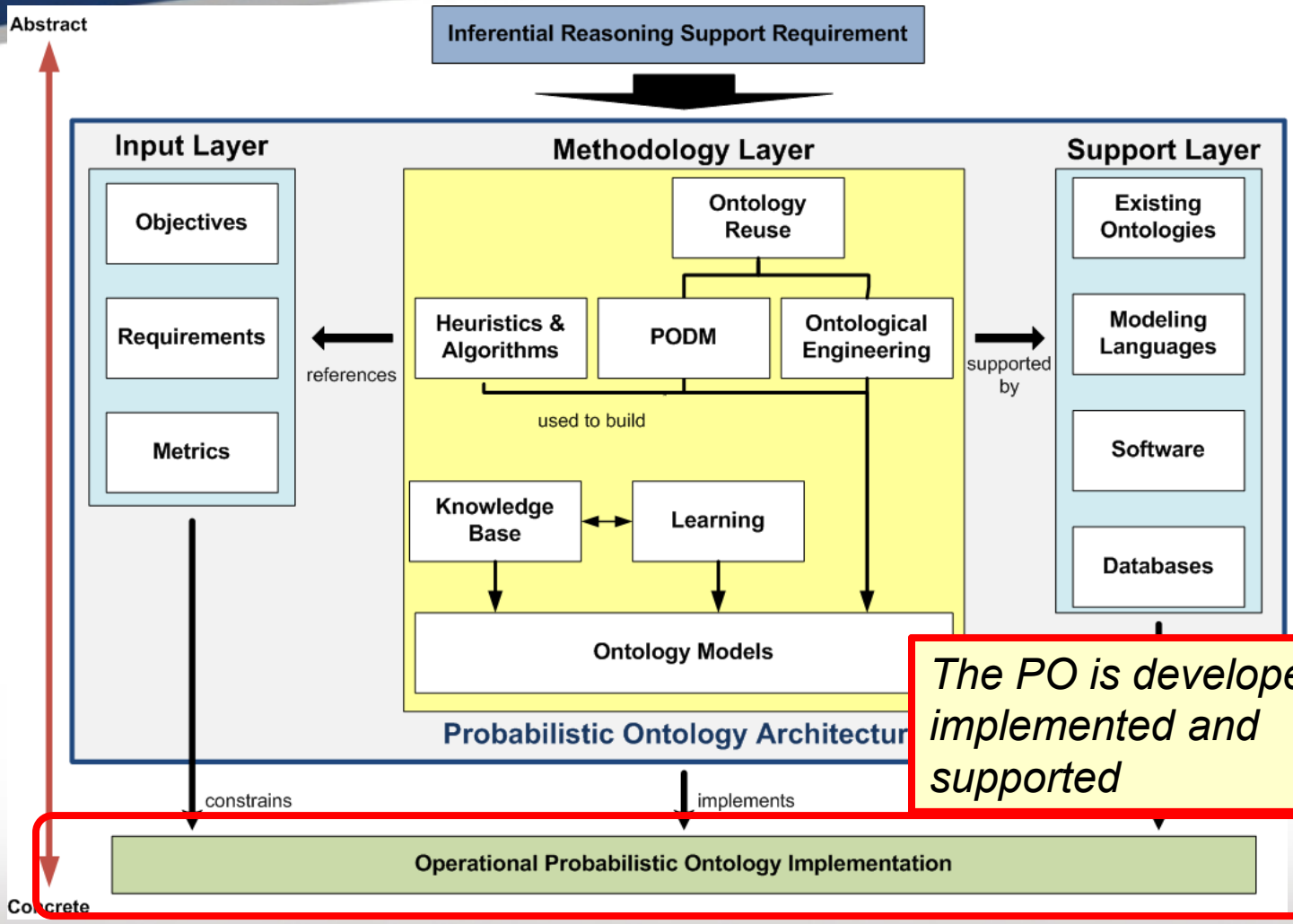
EMSolutions

Reference Architecture for Probabilistic Ontology Development



The PO Developer collaborates with the Stakeholder Decision Maker and SMEs to construct an architecture that will be used to develop and implement the probabilistic ontology

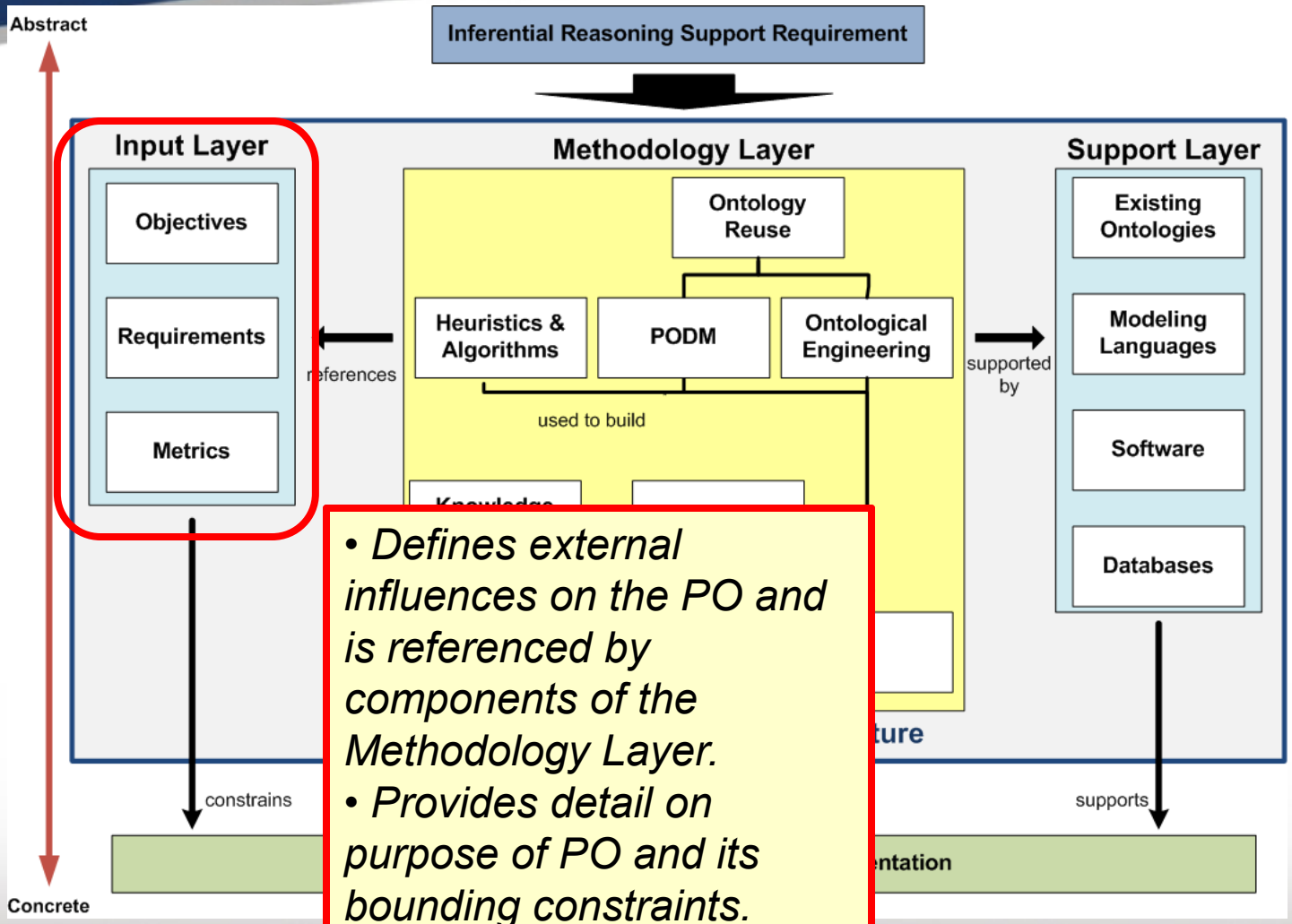
Reference Architecture for Probabilistic Ontology Development



The PO is developed, implemented and supported

EMSolutions

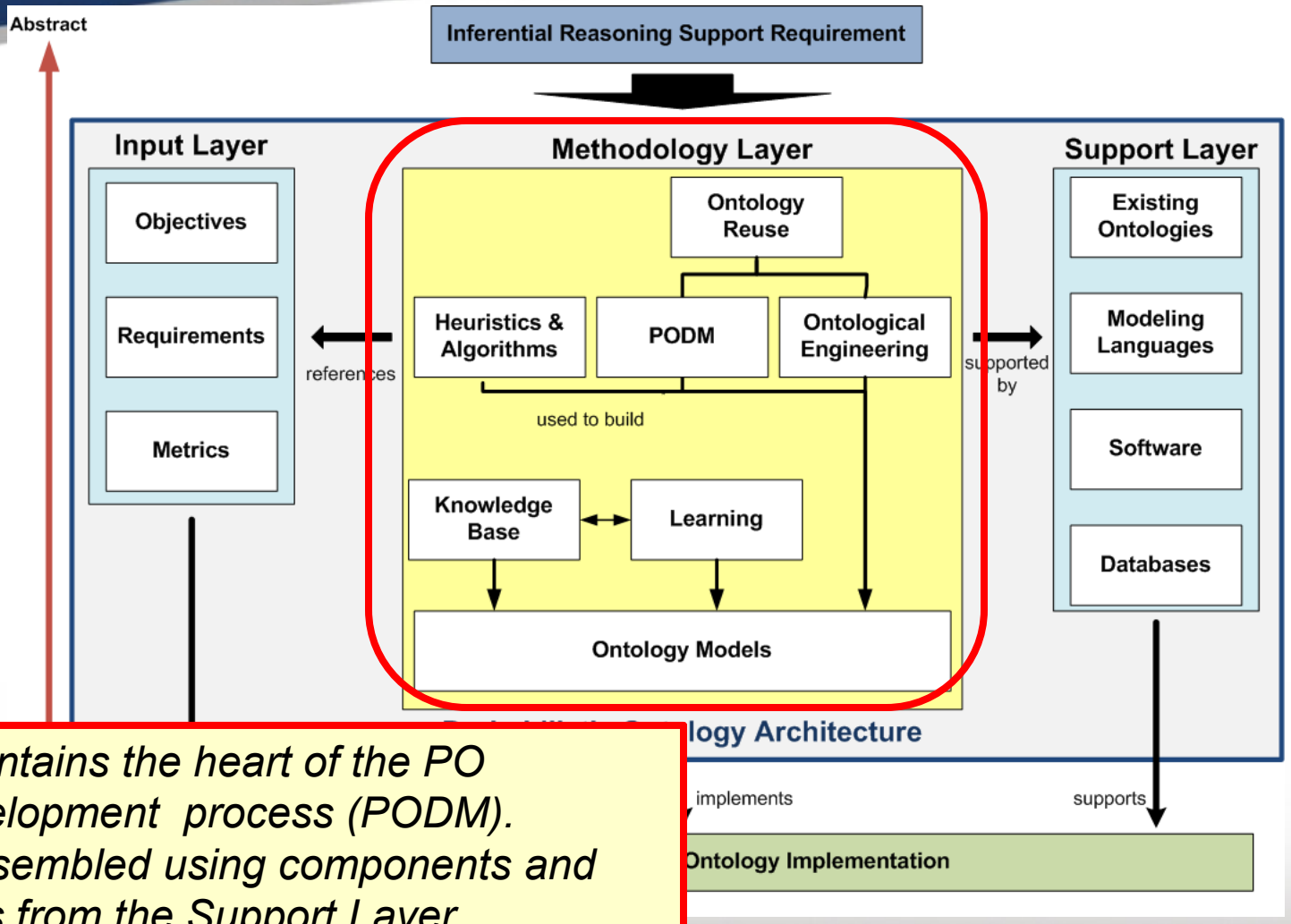
Reference Architecture for Probabilistic Ontology Development



- *Defines external influences on the PO and is referenced by components of the Methodology Layer.*
- *Provides detail on purpose of PO and its bounding constraints.*

EMSolutions

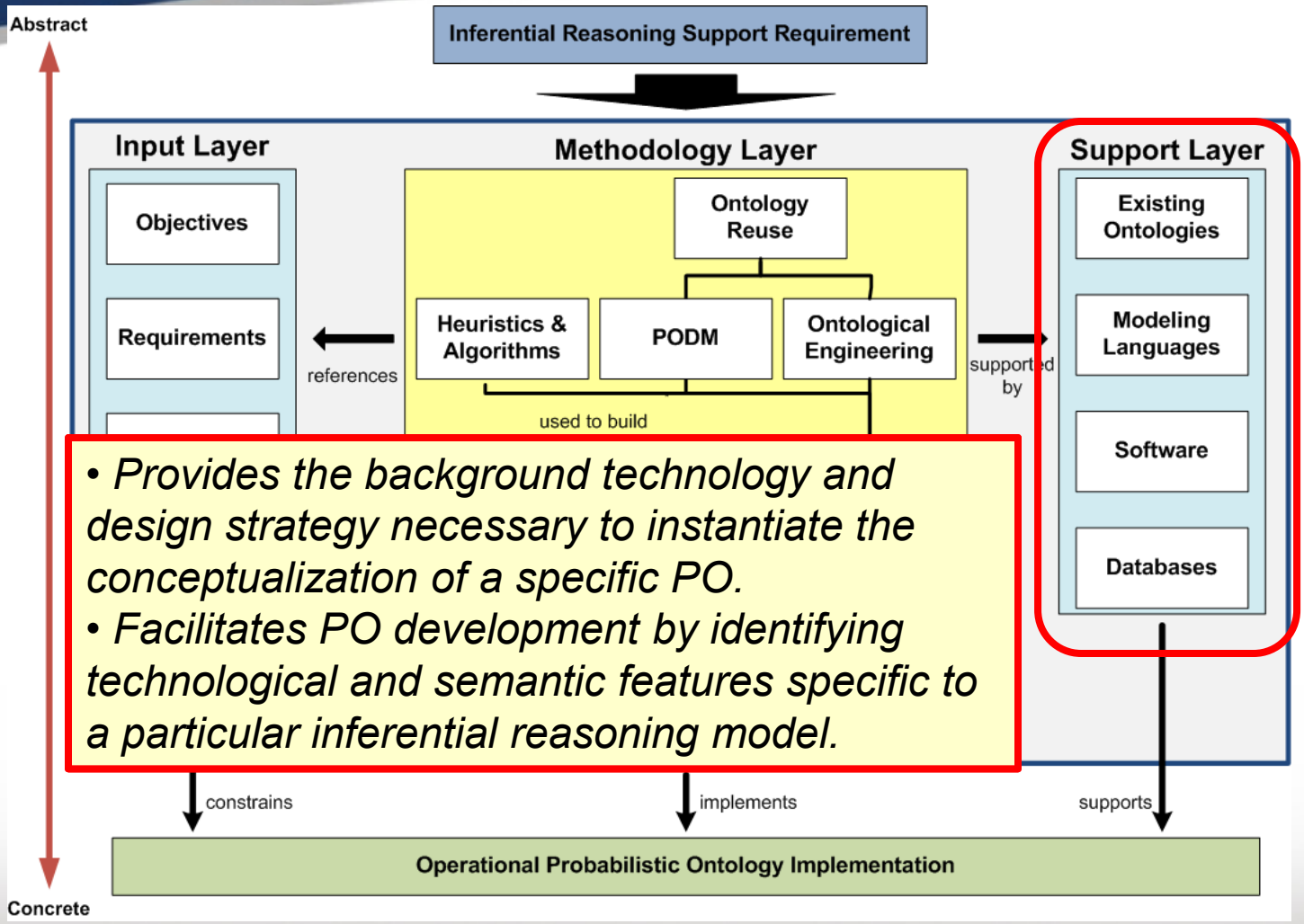
Reference Architecture for Probabilistic Ontology Development



- *Contains the heart of the PO development process (PODM).*
- *Assembled using components and tools from the Support Layer.*

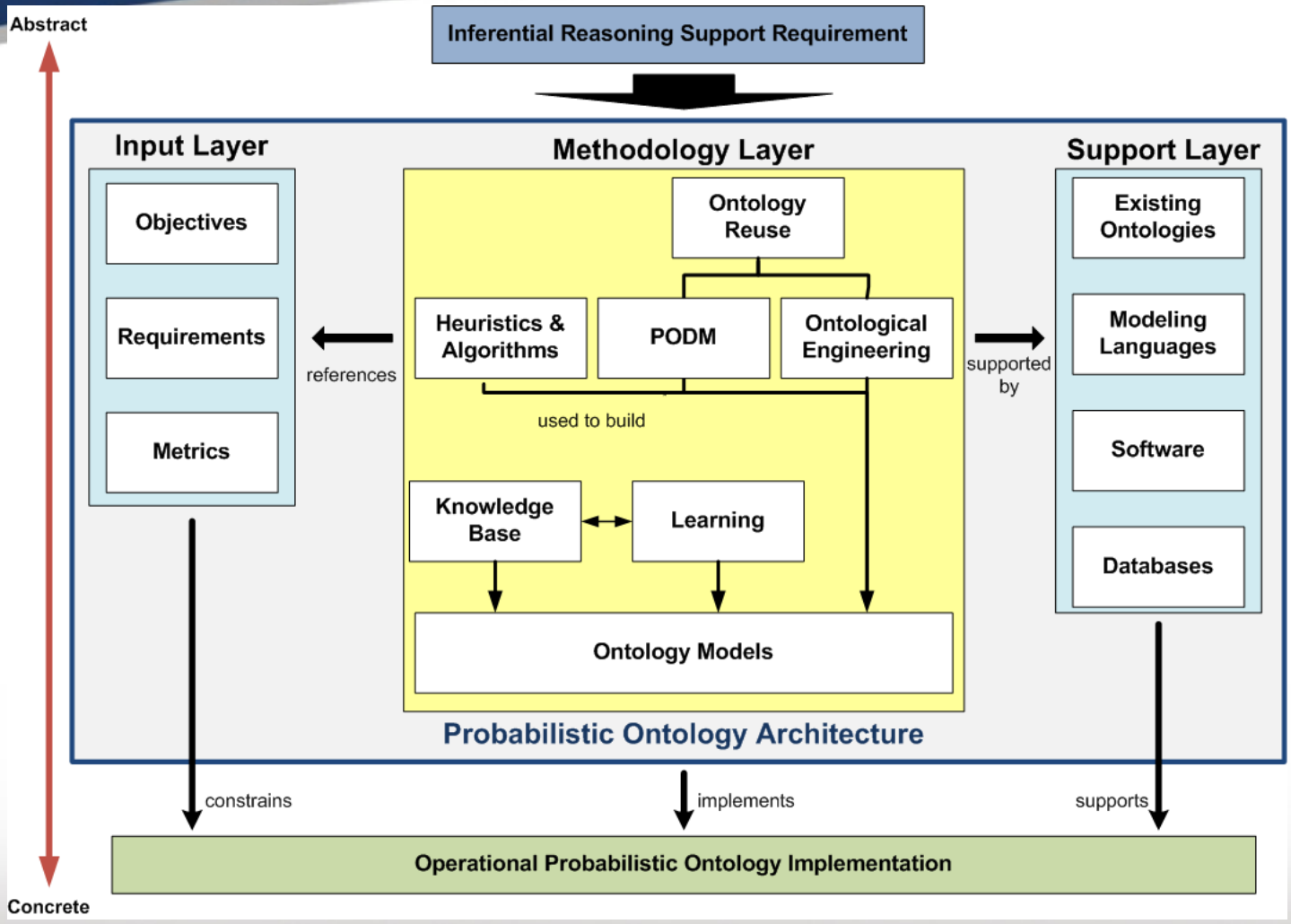
EMSolutions

Reference Architecture for Probabilistic Ontology Development



EMSolutions

Reference Architecture for Probabilistic Ontology Development



EMSolutions

A Reference Architecture for Probabilistic Ontology Development

Richard Haberlin
Paulo Costa
Kathy Laskey

EMSolutions