

## HBGDki Ontology Development & Knowledge Integration Architecture

**John S. Erickson** ([erickj4@rpi.edu](mailto:erickj4@rpi.edu)), **Deborah L. McGuinness** ([dlm@cs.rpi.edu](mailto:dlm@cs.rpi.edu)), **John Sheehan** ([sheehj4@rpi.edu](mailto:sheehj4@rpi.edu))

Rensselaer Polytechnic Institute / 110 8<sup>th</sup> Street / Troy, NY, 12180 USA

### Introduction

<http://tw.rpi.edu>

Ontologies will enable the HBGDki infrastructure to “understand” what terms mean, how they interconnect, and where information came from. will provide a family of program relevant ontologies that are also used to power a set of semantic services and that will be used to communicate unambiguously. Success is realized by allowing program members to and integrate relevant data to support a wide range of analyses, thus enabling more efficient, higher quality, and better supported results. Our work is focused on semantic support for Agile HBGDki exploration and visualization and will support reporting, application interfaces, smart searches, context-aware browsing, filtering, reasoning, connections to content management systems. We will design an automated semantic conversion process to port HBGDki data, in particular data from a visual interactive knowledge graph representation (especially DebateGraph) and program database dump(s) into a "curated" knowledge graph.

### HBGDki Ontology Goals

Ensure that...

- All data elements are traceable
- The process of making decisions is documented
- Definitions clearly specified; provenance accessible
- The process of selecting and changing definitions over time is recorded
- All assumptions are specified and clarified w/ provenance
- The process is transparent and reproducible

### Concept: Semantic Markup of DebateGraph

### Glossary:

**RPI** – Rensselaer Polytechnic Institute  
**TWC** – Tetherless World Constellation at Rensselaer Polytechnic Institute  
**HBGDki** – Healthy Birth Growth and Development Knowledge Integration

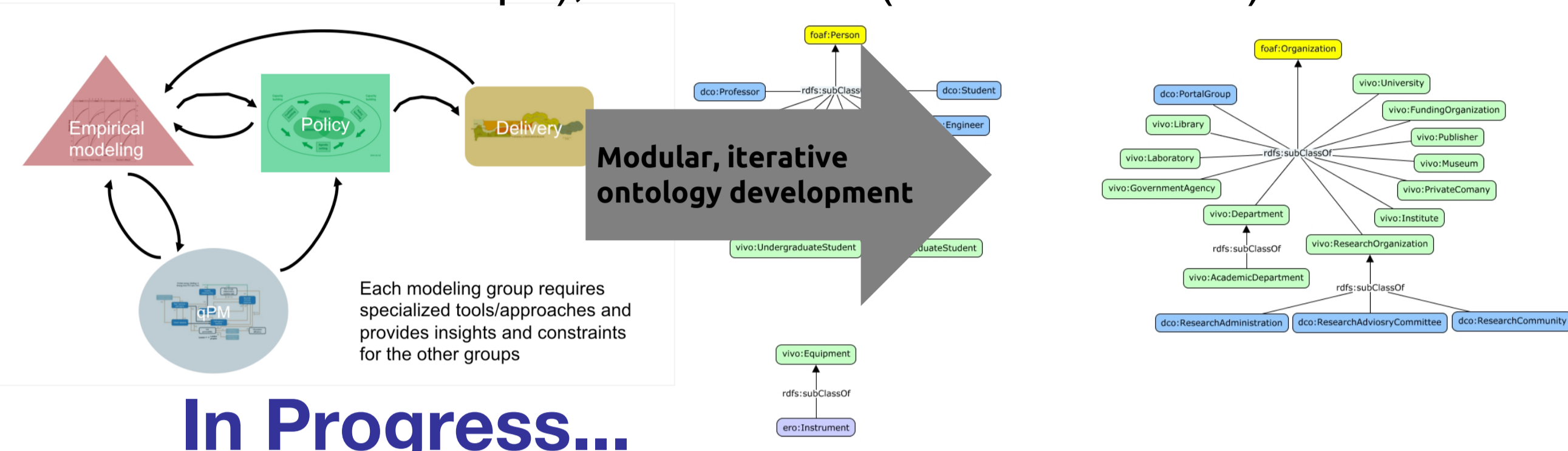
### Acknowledgments:

David Price - DebateGraph



### Strategy: Models for...

- HBGDki entities at all levels of hierarchy
- Data assets & provenance
- Program Questions, methods & tools, conclusions
- Surge team activities: "inputs," "outputs"
- Organizational structures, definitions
- Communications pathways
- Approach is: **Top-down** (from literature and relevant ontologies/vocabularies); **Bottom-up** (from data and DebateGraph); **Middle-out** (from use cases)



### In Progress...

#### Bottom-up: Inventory of HBGDki concepts

- Where do they fit in relation to each other
- Especially: How are they relevant to HBGDki work?
- What HBGDki questions, methods, answers are they associated with?
- What assets do they relate to?
- What communications pathways do they characterize?

#### Middle-out: Use case driven ontology discovery and grounding

#### Top-down: Ontology/Vocabulary collection

### Sources (Input requested!)

Grant Folder	Analysis Folder	Study Name	Data Contributor	DB
OPP1101329	Ameba201501	AMC (Gambia)	AMC (Gambia)	Yes
CPP	20141221	Collaborative Perinatal Project	Public (NIH)	Yes
CMC/Malere	PHC3894229	CMC Malere Birth Cohort Study (M202) with Anthropometry and Diarrheal Episodes	CMC Vellore	Yes
OPPI112895	PHM20668253	Bangladesh Rice Study	UC Davis, Aresnault	Yes
OPPI112895	PHM20713702	Bangladesh Diarrhea Study	UC Davis, Aresnault	Yes
OPPI112895	PHM20646919	Peru HoSAR	UC Davis, Aresnault	Yes

**Longitudinal clinical trials**  
**Interim analysis ongoing studies (blinded)**  
**Clinical trial simulation**  
**Omics**  
**Data Science**  
**Cross-sectional survey data**  
**Meta-analyses**

**Policy**  
**Media analyses**

**Gestational age**  
 Gestation is the period of time between conception and birth.  
 Gestational age is the common term used during pregnancy to describe the woman's last menstrual cycle to the current date.  
 Infants born before 37 weeks are considered premature.

### Medical Subject Headings

**Preferred Name**: Gestational Age  
**Synonyms**: Age, Fetal; Ages, Gestational; Chronologic Fetal Maturity; Age, Gestational; Maturity, Chronologic Fetal; Fetal Maturity, Chronologic; Fetal Age; Fetal Age, Gestational; Gestational Ages  
**Definitions**: The age of the conceptus, beginning from the time of FERTILIZATION; in clinical contexts, the gestational age is often estimated as the time from the last day of the last MENSTRUATION which is about 2 weeks before OVIULATION and fertilization.  
<http://nltx.bioontology.org/ontology/MESH/D006965>

### RPI HBGDki Virtual Observatory

- Leverage semantics to support ontology-aware discovery
- Support agile generation/curation of HBGDki Knowledge Graph: Semantic Mediation layer
- Provide lightweight, low-impact infrastructure for ontology development and instance data management
- Demonstrate applications (visualizations) based on ontology-driven markup of DebateGraph-hosted content
- Leverage web standards for embedded markup, linked data and ontologies to embed rich metadata essential to HBGDki program in DebateGraph map
- Harvest embedded metadata using standard libraries, load into cloud triplestore, build query-based apps
- Complement with related supporting linked data

### RPI HBGDki Studies Browser Prototype

Searchable facets based on vocabulary terms characterizing studies

### Supporting HBGDki activities with Ontology