Semantics and Services Enabled Problem Solving Environment for Trypanosoma cruzi

Amit P. Sheth  
*Wright State University - Main Campus*, amit.sheth@wright.edu

Rick L. Tarleton

Mark Musen

Satya S. Sahoo  
*Wright State University - Main Campus*

Prashant Doshi

See next page for additional authors

Follow this and additional works at: https://corescholar.libraries.wright.edu/knoesis

Part of the Bioinformatics Commons, Communication Technology and New Media Commons, Databases and Information Systems Commons, OS and Networks Commons, and the Science and Technology Studies Commons

Repository Citation
https://corescholar.libraries.wright.edu/knoesis/987

This Presentation is brought to you for free and open access by the The Ohio Center of Excellence in Knowledge-Enabled Computing (Kno.e.sis) at CORE Scholar. It has been accepted for inclusion in Kno.e.sis Publications by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu, library-corescholar@wright.edu.
Authors
Amit P. Sheth, Rick L. Tarleton, Mark Musen, Satya S. Sahoo, Prashant Doshi, and Natasha Noy
Semantics and Services enabled Problem Solving Environment for Trypanosoma cruzi

PI: Amit Sheth, LexisNexis Ohio Eminent Scholar, Wright State U. with Rick Tarleton, Distinguished Professor, Univ. of Georgia, and Mark Musen, Professor (Biomedical Informatics); Division Head (BMIR)

Semantic & Services enabled PSE (ssPSE)

Funding expected 1Q2008:
Collaboration with National Centers for Biomedical Computing (R01)
National Heart, Lung, and Blood Institute; National Institutes of Health
Partners & Partner Institutions

• **Kno.e.sis Center** (Dr. Amit Sheth, key contributor: Satya Sahoo)
• **Tarleton Lab**, Cellular Biology, University of Georgia (Dr. Rick Tarleton)
• Computer Science, University of Georgia (Dr. Prashant Doshi)
• **NCBO**/Stanford Medical Informatics Center (Dr. Mark Musen, Dr. Natasha Noy)
Project Overview

• Facilitate *T. cruzi* research through an ontology driven Problem Solving Environment

• Biological Objectives:
  o Identification of vaccine, diagnostic, and therapeutic targets for *T. cruzi* caused Chagas disease
  o Chagas disease affects 18 million people in Latin America

• Computer Science Objectives
  o Dynamic integration of local and public multimodal data to answer biological questions
  o Suite of intuitive and comprehensive interfaces for usability and easy adoption by biologists
Proposed *T. cruzi* PSE System Architecture

**OPEN BIOMEDICAL ONTOLOGIES (OBO)**
- **OWL**
  - GO
  - ProPreO
  - *T. cruzi* Lifecycle
  - Metabolic Pathways

**OPEN BIOMEDICAL DATABASE (OBD)**
- **OWL**
  - Multiple Ontology Alignment mappings
  - *T. cruzi* Semantic Data annotations

**SEMANTIC WEB SERVICES**
- SAWSDL
- OWL-S

**T. cruzi** Problem Solving Environment (PSE)
- **UDDI**
- **ONTOGRAPHY**
- **RESOURCE**
  - Data & Service Annotation
  - *T. cruzi* Query Interface
  - Web Process Composition

**BioPORTAL**
• Phased development of a suite of ontologies for parasitic organisms such as *T. cruzi*
  o Extension of BioPAX ontology for T.cruzi human specific life cycle pathways
  o Extension of ProPreO ontology to incorporate RT-PCR procedure and incorporate microarray data analysis
• Investigate novel methods for ontology alignment and merging
  o Extension PROMPT tool with our work on semantic mappings
Computer Science Research (contd.)

- Semantic annotation of experimental data with ontological concepts
  - Extension of Integrated Semantic Information and Knowledge System (ISiS) to *T.cruzi* specific experimental datasets
  - Extension of NCBO Open Biological Data initiative (OBD) and integration with ISiS
  - Semantic Web services (SWS) based scientific workflows
    - SWS discovery, composition
- Ontology driven entity identification and relationship extraction from biomedical text
- Complex query processing and visualization
  - SPARQ2L for partially defined patterns in the search query
  - Intuitive Web based query interface
**T. Cruzi PSE Query Interface – Semantic Annotation of Experimental Data**

*Build your query to the database*

By navigating through the ontology schema (i.e., the definition of the knowledge base), the system will guide you throughout the process of, e.g., "Gene -> codes for -> Protein -> expressed in -> Epimastigote (Life Cycle)". The interface includes search capabilities with options for Amino Acid Sequence, Protein Expression, and life cycle stage.
Preliminary Work to build upon

• Ontologies
  o ProPreO ontology
  o EnzyO Ontology
• Standards
  o SAWSDL
• Tools/Systems
  o SAWSDL4J & Woden APIs
  o SemBOWSER
• Past projects:
  o The Integrated Technology Resource for Biomedical Glycomics