UML-Based Ontology Toolset (UBOT)

Lockheed Martin Integrated Systems & Solutions

PI: Paul Kogut

Key Researchers:

Jeff Heflin, Zhengxiang Pan – Lehigh University

Mitch Kokar, Ken Baclawski, Chris Matheus, Jerzy Letkowski - VIS/Northeastern Univ.

Richard Waldinger - Kestrel/SRI

Prasanta Bose - Lockheed Martin Advanced Technology Center

Yui Leung, Bill Holmes, Ted Mielczarek, Kathleen Ryan LM IS&S

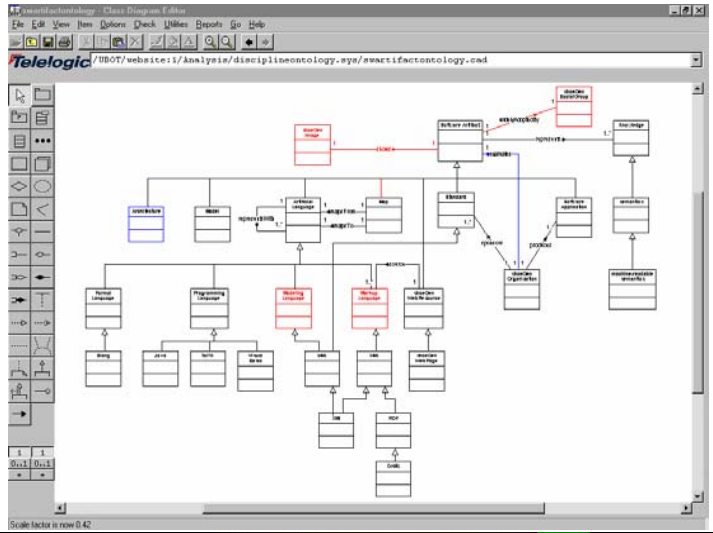
November 2004

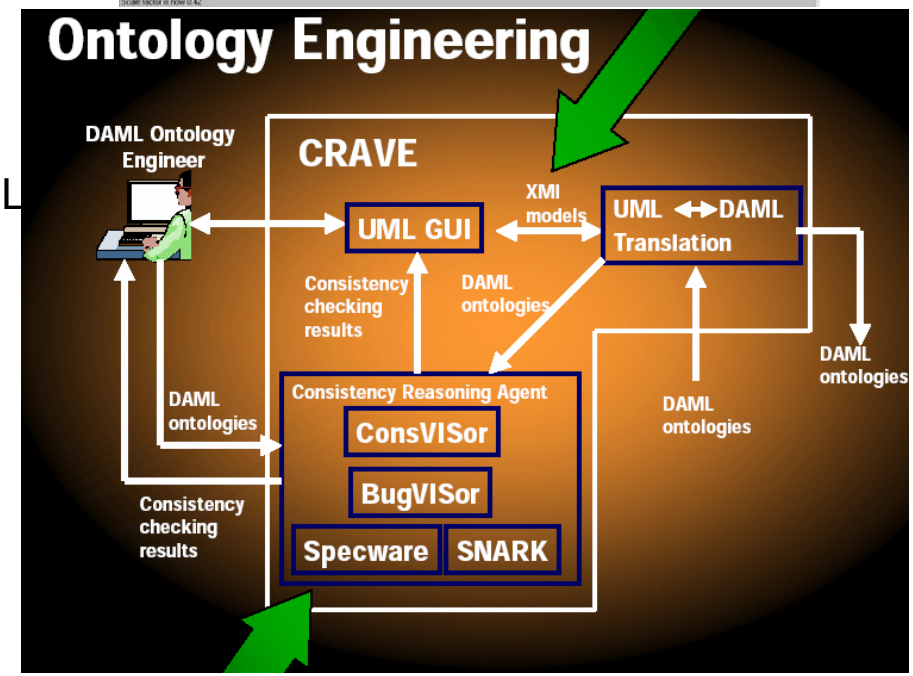
UBOT
Overall Program Summary

- What is the basic problem you are trying to solve?
  - Transition of Semantic Web technology to DoD practitioners is difficult
    - **Ontology engineering is hard**
    - **Markup creation is time consuming**
    - **Application architectures and engineering tradeoffs are not well understood**
- What was the technical solution strategy?
  - Apply software engineering principles:
    - **Automated tools**
    - **Formal methods**
    - **Software architecture**
- What were the basic elements of the research and program approach?
  - Ontology engineering - graphical tools and consistency reasoning
  - Automated markup generation – natural language processing
  - Architectures – quantitative reasoning benchmarks and design patterns

UBOT Ontology Engineering Technical Problem/Approach/Progress

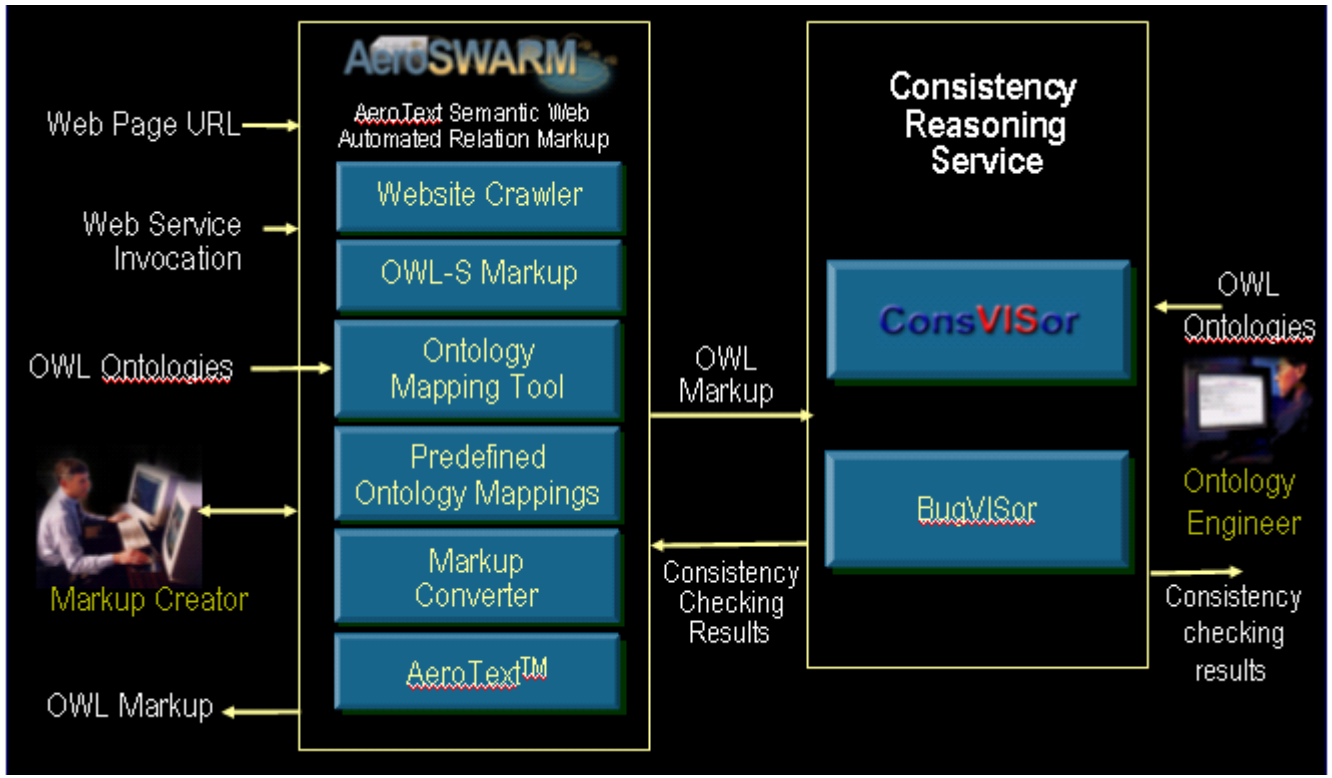
- How do we help mainstream software engineers move from OO development to Semantic Web?
- Build ontology engineering tools with team support based on COTS UML environments (Telelogic Tau)
- Build easy to use consistency reasoning tools (ConsVISor, BugVISor)
- Apply formal methods (Specware and SNARK) to verify DAML and OWL axiomatic semantics
- Influence OMG
  - Standard mapping between UML and DAML
- Metrics/success:
  - Partnership agreement with Telelogic
  - Initiated OMG Ontology WG
  - ConsVISor passed OWL semantics conformance tests and gives explanations of symptoms
  - Identified DAML semantics problems





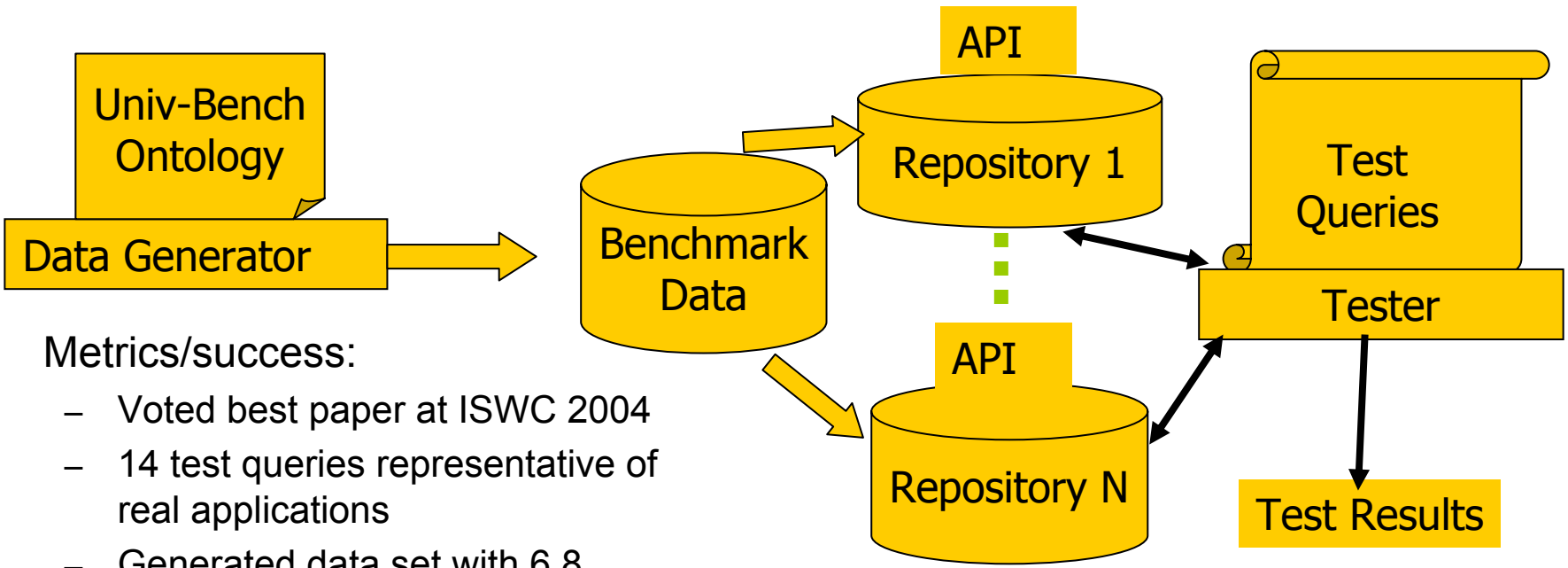
UBOT Markup Generation
Technical Problem/Approach/Progress

- How do we reduce the cost/effort required to mark up webpages and text?
- Develop open service for OWL markup generation: AeroSWARM
- Develop/implement approach to consistency checking of markup
- Apply AeroSWARM to real world knowledge management problems:
  - **GOWLgle - Semantic filtering of Google results**
  - **Cross-document co-reference**
- Metrics/success:
  - Generate markup for 44 common properties with around 80% precision
  - Helped many organizations implement this approach: UMBC, Horus, AT&T, Teknowledge, CIA, NGA, Tucana



UBOT Reasoning Architectures and Benchmarks
Technical Problem/Approach/Progress

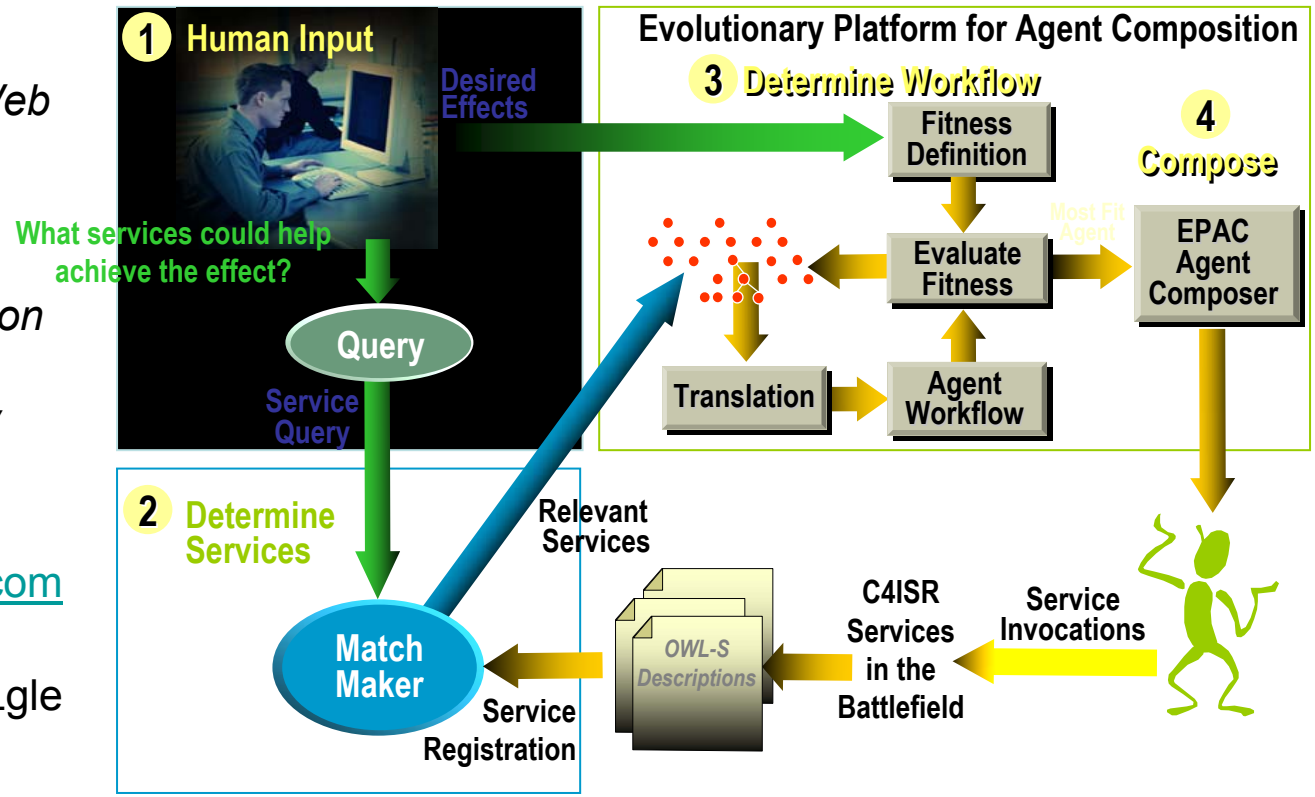
- How do we choose appropriate reasoning infrastructures for large OWL applications?
  - **Develop repeatable benchmark process and tools**
  - **Evaluate existing reasoning infrastructures**
- How can we develop scalable reasoning infrastructures?
  - **Develop approach for pre-computing all inferences and storing the results in a database for fast query response**



- Metrics/success:
  - Voted best paper at ISWC 2004
  - 14 test queries representative of real applications
  - Generated data set with 6.8 million triples

UBOT Architectures and Design Patterns Technical Problem/Approach/Progress

- How do we apply the Semantic Web to solve real problems for DoD, Intel community…?
  - **Develop prototype applications and disseminate design patterns:**
    - **Knowledge management/semantic filtering/assistant agents**
    - **Semantic web services/net-centric operations/sensors/fusion**
- Metrics/success:
  - IEEE conference paper: *Semantic Web technologies for Aerospace*
  - CACM paper: *Proactive Information Gathering for Homeland Security Teams*
  - NASA Sensorweb http://cosec/lmsal.com
  - Demos: DAML experiment, GOWLgle
  - Tech transition support: CIA, NGA



UBOT Milestones, Accomplishments, Transition/Handoff

2000

- homework 1 software engineering ontology
- homework 2 NLP and machine learning for DAML queries
- present to OMG Agent SIG *UML models for DAML ontologies*
- homework 3 NLP for text markup

2001

- demos: UML ontology engineering, ontology consistency checking, AeroDAML
- discover errors in DAML axiomatic semantics via theorem proving
- AeroDAML sent to UMBC for IT Talks
- sign partnership agreement with Telelogic UML vendor
- submit 3 HotDAML proposals: web accessible AeroDAML, ConsVISor, DAML VisuaLinks
- Lockheed Martin Joint Symposium *XML+Semantics = DAML*
- demo 3 HotDAML proposals
- organized OMG Software Services Grid Workshop - W3C OMG summit
- AeroDAML sent to Horus
- AeroDAML wins HotDAML award
- UML 2001 conf. *Extending UML to support Ontology Engineering for the Semantic Web*

UBOT Milestones, Accomplishments, Transition/Handoff

2001

- presented *AeroDAML: Applying Info Extraction to Generate DAML Annotation* to:
  - **Int. Conf. Knowledge Capture workshop**
  - **DARPA EELD kickoff**
  - **Semantic Web for Military Users**

2002

- Lehigh University joins team
- Demo: AeroDAML front end for AT&T ontology driven knowledge dissemination
- NASA Sensorweb prototype - DAML-S applied to solar science
- Demo: AeroDAML integrated with Teknowledge MS Word markup tool
- OMG Web Services workshop – *Applying UML to model Semantic Web services*
- 1st ISWC – *Consistency Checking of Semantic Web Ontologies*
- Demo: DAML-S for ISR integrated with SONAT
- Demo: DLDB precompute inferences and store in database
- Worked with NG on Horus automated markup tool

UBOT Milestones, Accomplishments, Transition/Handoff

2003

- IEEE Aerospace conf. *Semantic Web Technologies for Aerospace*
- Demo: AeroSWARM
- Demo: Lehigh University benchmark
- 2nd ISWC *Benchmarking DAML+OIL repositories*
  - *DLDB: Extending Relational Databases to support Semantic Web Queries*
- CIA Developers Forum *Semantic Web Technologies and Applications*

2004

- Demo: GOWLgle
- Demo: cross document co-reference
- Formally verified semantics for OWL full
- 3rd ISWC *An Evaluation of Knolwedge Base Systems for Large OWL datasets*
  - *Towards a Symptom Ontology for Semantic Web Applications*

UBOT
Remaining Issues

Issue	Remediation
Semantic filtering of text documents	Continue GOWLgle research
Cross-document co-reference	Utilize more clues and reasoning
Need more reasoner benchmarks and data	Apply Lehigh benchmark to COTS reasoners, benchmark SWRL and matchmakers
How do we do matchmaking in heterogeneous dynamic hostile environments like net-centric operations?	Investigate fault tolerant peer to peer approaches
How does service composition and discovery interact?	More large scale experiments
Consistency checking of SWRL	Extend ConsVISor
Ontology engineering is still hard – average practitioner not equivalent to DAML researcher	Machine learning, cognitive science, advanced NLP, intelligent assistants, reasoners with explanation

UBOT Summary

- We did our part to change the world by contributing a software engineering perspective:
  - **Ontology engineering is hard**
    - **We developed practical semantic consistency checking tools**
    - **We helped bridge OMG and W3C**
  - **Markup creation is time consuming**
    - **We pioneered application of NLP for markup generation**
  - **Application architectures and engineering tradeoffs are not well understood**
    - **We developed and disseminated design patterns**
    - **We established a framework for choosing appropriate Semantic Web reasoning infrastructures**
  - **AeroSWARM, ConsVISor and Lehigh Univ. Benchmark are available on semwebcentral.org**

(Program Name)
Backup