

Introducing I³CON

*The Information Interpretation and
Integration Conference*



Todd Hughes, Ph.D.
Senior Member, Engineering Staff
Advanced Technology Laboratories

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE AUG 2004		2. REPORT TYPE		3. DATES COVERED 00-00-2004 to 00-00-2004	
4. TITLE AND SUBTITLE Introducing I3CON. The Information Interpretation and Integration Conference				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Lockheed Martin Advanced Technology Laboratories,3 Executive Campus,Cherry Hill,NJ,08002				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES Proceedings of the 2004 Performance Metrics for Intelligent Systems Workshop (PerMIS '04), Gaithersburg, MD on August 24-26 2004					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 12	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

I³CON: Motivation



- **Semantic integration will be one of the first major accomplishments for ontology-based applications**
 - Heterogeneous information system and resource Interoperability is a major concern for military, government, industry
 - Many view this as the a fundamental technical challenge of the Semantic Web
- **To answer this challenge, there have been new developments in automated ontology and schema:**
 - Markup
 - Alignment
 - Merging
 - Translation
 - Learning
- **Much of this research has been funded by DARPA programs, but today the largest sponsors are EU programs**

I³CON: Observation



- **Semantic integration research community resembles the text retrieval community of 15 years ago**
 - Critical mass of globally distributed research programs
 - Large variety of technical approaches
 - Generally, but not universally, accepted metrics
 - No meaningful basis of evaluating one technical approach over another
- **The success of text retrieval technology was due in large measure to the Text Retrieval Conference (TREC)**
 - Promoted well-defined concepts for measuring success
 - Clarified metrics
 - Established realistic benchmarks
 - Created canonical challenge problems

The NIST TREC model has a proven record of success!

NIST TREC Model



- 1. Define the metrics**
- 2. Develop experiment format for easy participation by researchers**
- 3. Create development data sets and test data sets; publish the former**
- 4. Distribute test data sets to experiment participants**
- 5. Collect automatically generated results data**
- 6. Collate and compare results data**
- 7. Hold assessment workshop and end of cycle**

I³CON: Timeline



- **March 2004:** Met with NIST, “pilot” conference as PerMIS special session proposed
- **March-June 2004:**
 - Formed Organizational Committee
 - Recruited participants
 - Created ontology alignment format
 - Developed test ontology pairs
 - **May 25:** Gave presentation at DAML PI Meeting
- **June 15 2004:** Released test ontology pairs
- **July 16, 2004:** Collected alignment results data
- **July 16-August 20, 2004:** Compiled and analyzed results data
- **August 25, 2004:** I³CON special session at PerMIS

<http://www.atl.imco.com/projects/ontology/i3con.html>

I³CON and the TREC Model

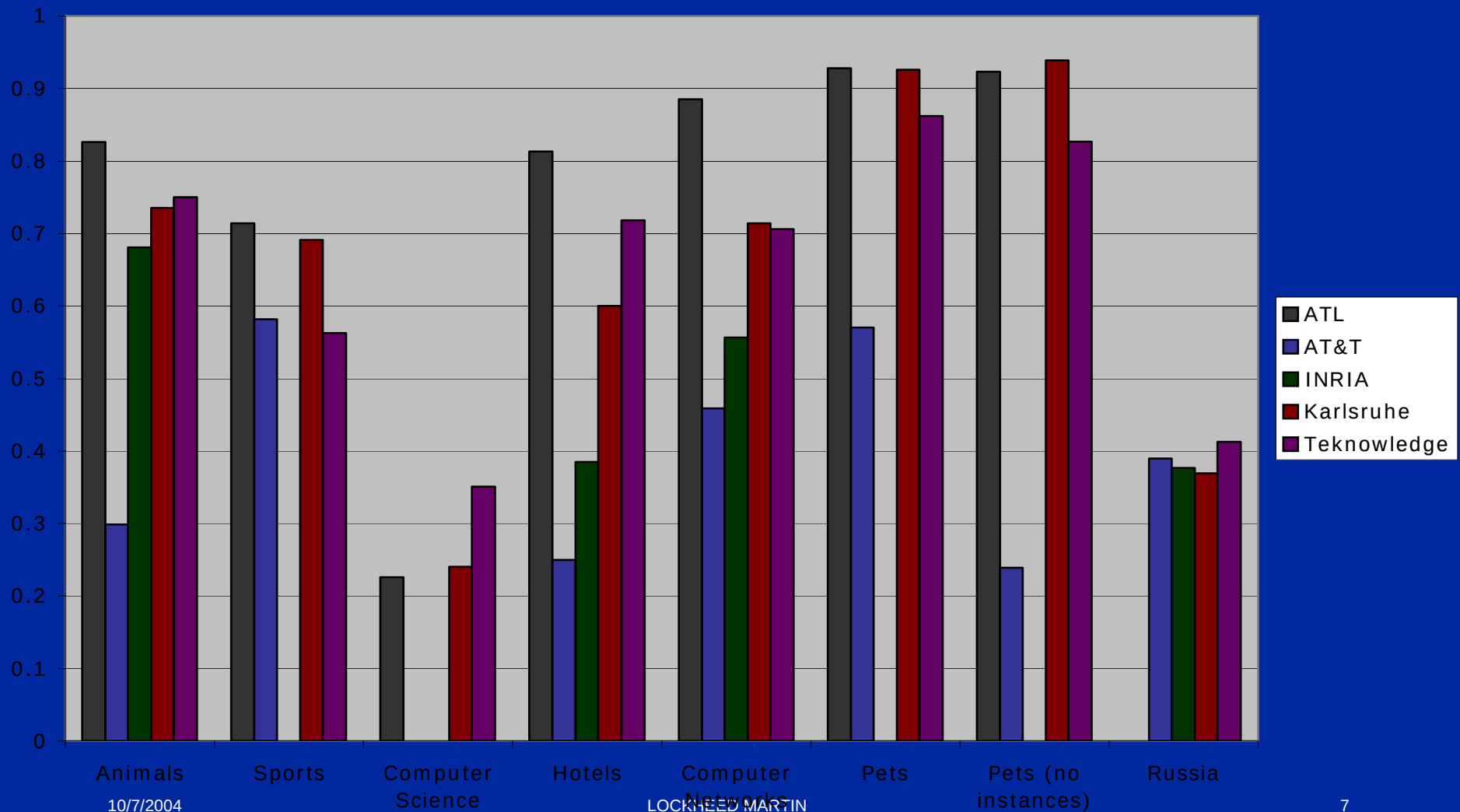


1. **Define the metrics** → Precision, Recall, fMeasure
2. **Develop experiment format for easy participation by researchers** → Ontology Alignment
Ontology; Experiment Set
Platform
3. **Create development data sets and test data sets; publish the former** → 2 development ontology
pairs; 8 test ontology pairs
4. **Distribute test data sets to experiment participants** → 5 participants
5. **Collect automatically generated results data** → Most participants submitted
alignment data for all
ontology pairs
6. **Collate and compare results data**
7. **Hold assessment workshop and end of cycle** → Where we are today

I³CON: Experiment Results Overview



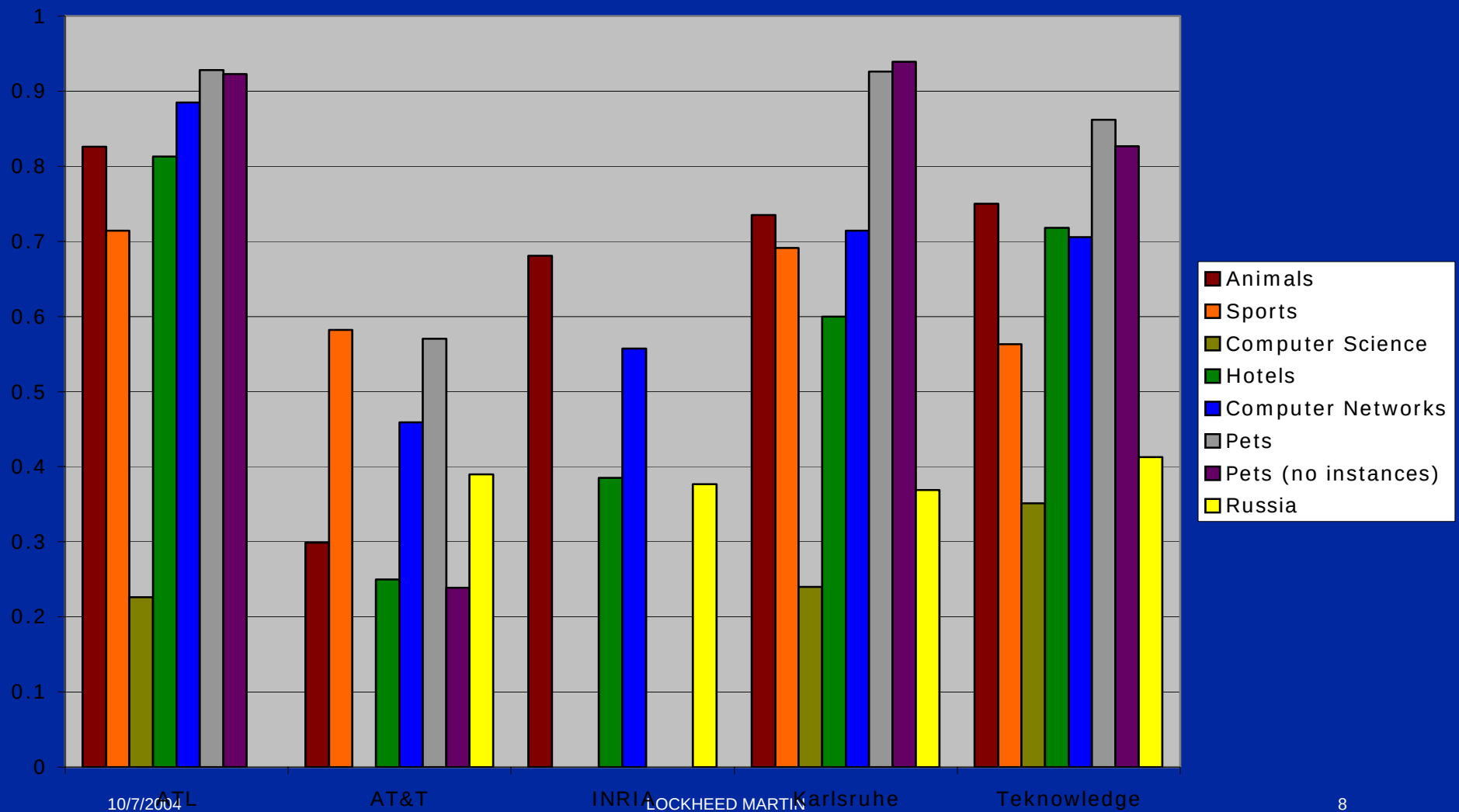
Ontology Pair vs. fMeasure



I³CON: Experiment Results Overview



Organization vs. fMeasure



I³CON Experiment: Lessons Learned



- No single technical approach performed best on *all* test ontology pairs
- No single ontology pair was best for *all* technical approaches
- All approaches performed >0.5 fMeasure on at least one ontology pair
- All approaches performed <0.5 fMeasure on at least one ontology pair

There is much more to be learned from the I3CON experiment data.

I³CON: Special Thanks



- **Organizational Support**

- Larry Reeker (NIST)
- Elena Messina (NIST)

- **Technology and Data**

- Ben Ashpole (ATL)
- Liz Palmer (ATL)
- Emil Macarie (ATL)
- Yun Peng (UMBC)
- Rong Pan (UMBC)

- **Experiment Participants**

- Jerome Pierson (INRIA)
- John Li (Teknowledge)
- Lewis Hart (AT&T)
- Marc Ehrig (University of Karlsruhe)

- **Guest Speakers**

- o Bill Andersen (Ontology Works)
- o Mike Pool (Information Extraction and Transport)
- o Yun Peng (University of Maryland Baltimore County)
- o Mike Gruningner (University of Maryland)

EON 2004



- **Evaluation of Ontology-based Tools 3rd International Workshop**
 - <http://km.aifb.uni-karlsruhe.de/ws/eon2004/>
- **Located at the 3rd International Semantic Web Conference (ISWC 2004)**
 - November 8, 2004
 - Hiroshima Prince Hotel, Hiroshima, Japan
- **EON Ontology Alignment Experiment**
 - Provides participants with a complete test base of ontology pairs
 - Test is based on one particular ontology dedicated to a very narrow domain and a number of alternative ontologies of the same domain

