

Impacts of User Modeling on Personalization of Information Retrieval: *An Evaluation with Human Intelligence Analysts*

Eugene Santos Jr., Qunhua Zhao,
Hien Nguyen, and Hua Wang
University of Connecticut





Outline

- **Motivation & Objectives**
- **Our approach**
- **Experiment**
- **Results**
- **Conclusion**





Outline

- **Motivation & Objectives**
- Our approach
- Experiment
- Results
- Conclusion



Motivation

- The need of having a personalized information retrieval system for intelligence analysts
- The challenges caused by the availability and time constraints of intelligence analysts



Objectives

- Evaluate external effectiveness of our user model with target users and compare with keyword-based system
- Study the impacts of our model on augmenting personalization in an information retrieval process
- Collect users' feedback





Outline

- Motivation & Objectives
- **Our approach**
- Experiment
- Results
- Conclusion





An example

Query: Banking transaction

Retrieved document:

Report 1: date 1 April, 2003

Report 14: date 21 April, 2003

Report 16: date 27 April, 2003

Report 7: date 15 April, 2003

Report 8: date 19 April, 2003



An example

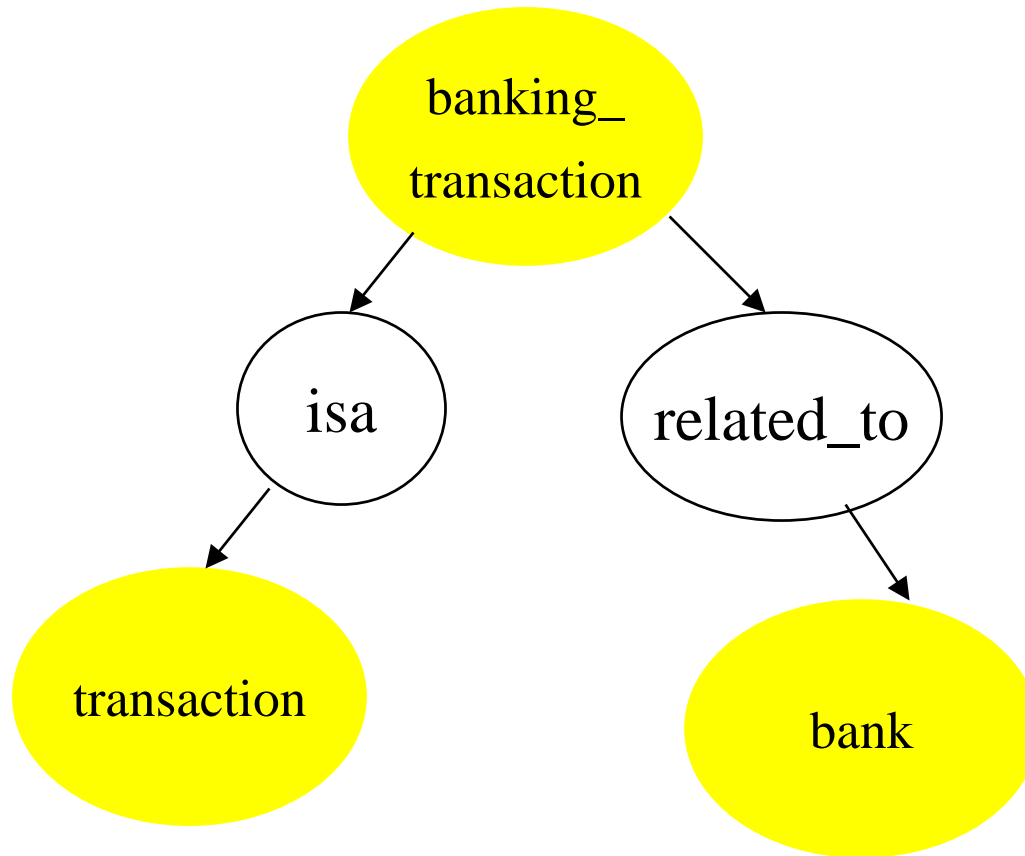
- ✓ Report 1: date 1 April, 2003
- ✓ Report 14: date 21 April, 2003
- ✓ Report 16: date 27 April, 2003
- Report 7: date 15 April, 2003
- Report 8: date 19 April, 2003

suspicious
banking
transactions
involving *Abdul
Ramazi.*



Example of Query Graph

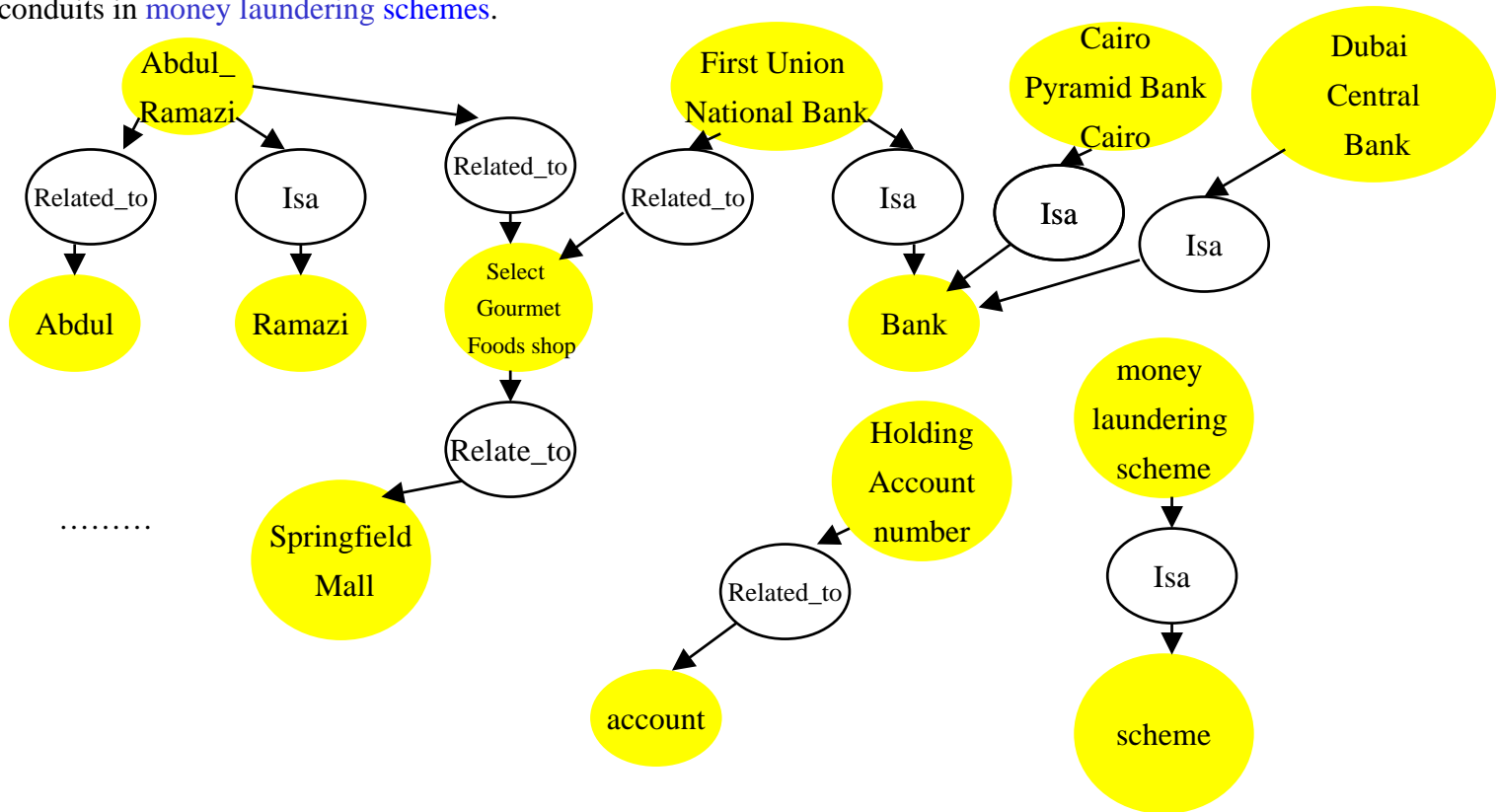
Query: banking transaction:



Example of Document Graph

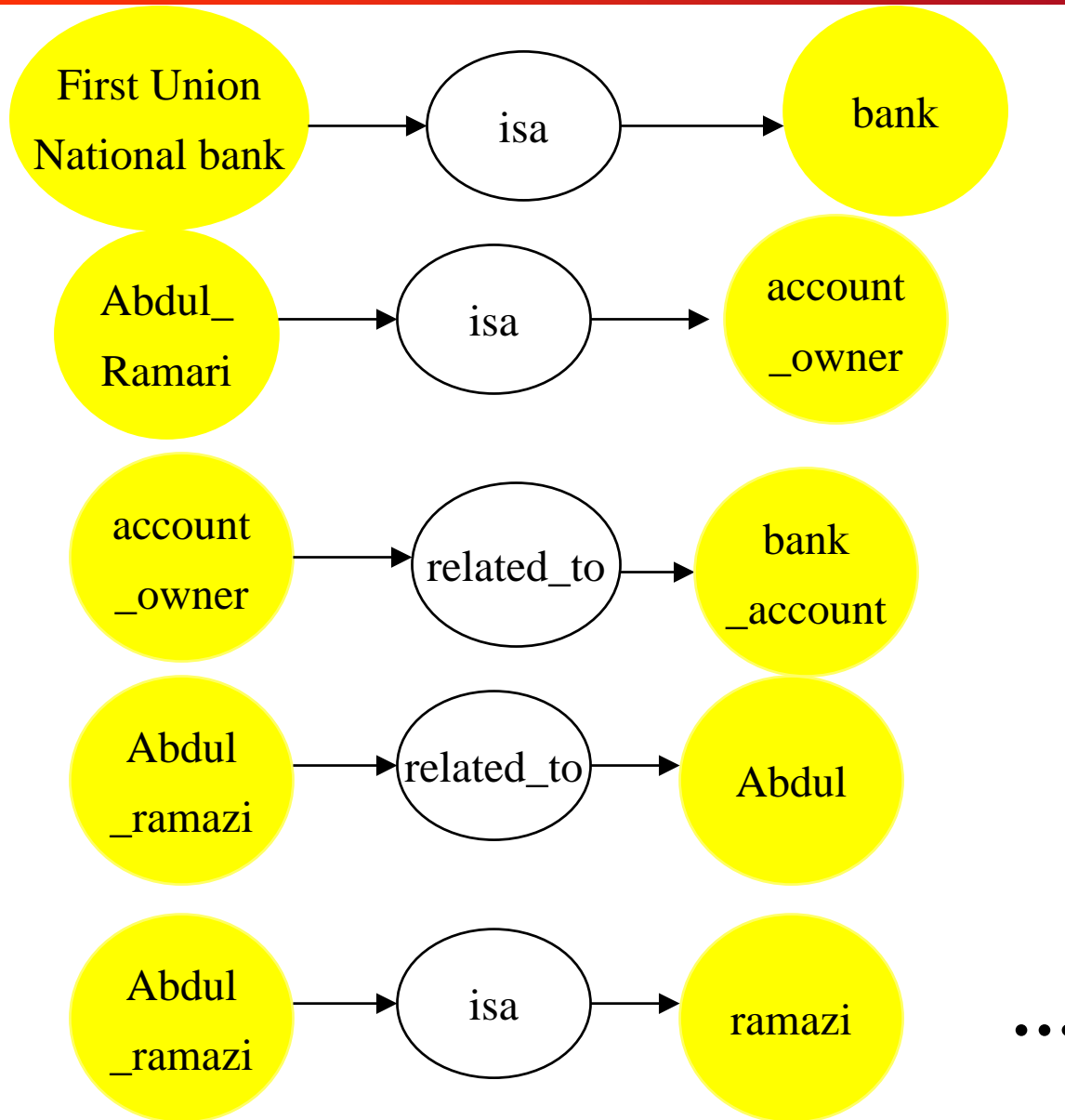
FBI 1) Report Date: 1 April, 2003.

FBI: [Abdul Ramazi](#) is the owner of the [Select Gourmet Foods shop](#) in Springfield Mall. [First Union National Bank](#) lists Select Gourmet Foods as holding [account number](#). Six checks totaling \$35.000 have been deposited in this account in the past four months and are recorded as having been drawn on accounts at the [Pyramid Bank of Cairo](#), Egypt and [the Central Bank of Dubai](#), United Arab Emirates. Both of these banks have just been listed as possible conduits in [money laundering schemes](#).





Intersection of retrieved relevant documents





Existing Interest Set

Interest concept	Interest level
money_launders	0.875
deposit	0.8125
withdraw	0.8
bank_account	0.75
.....	



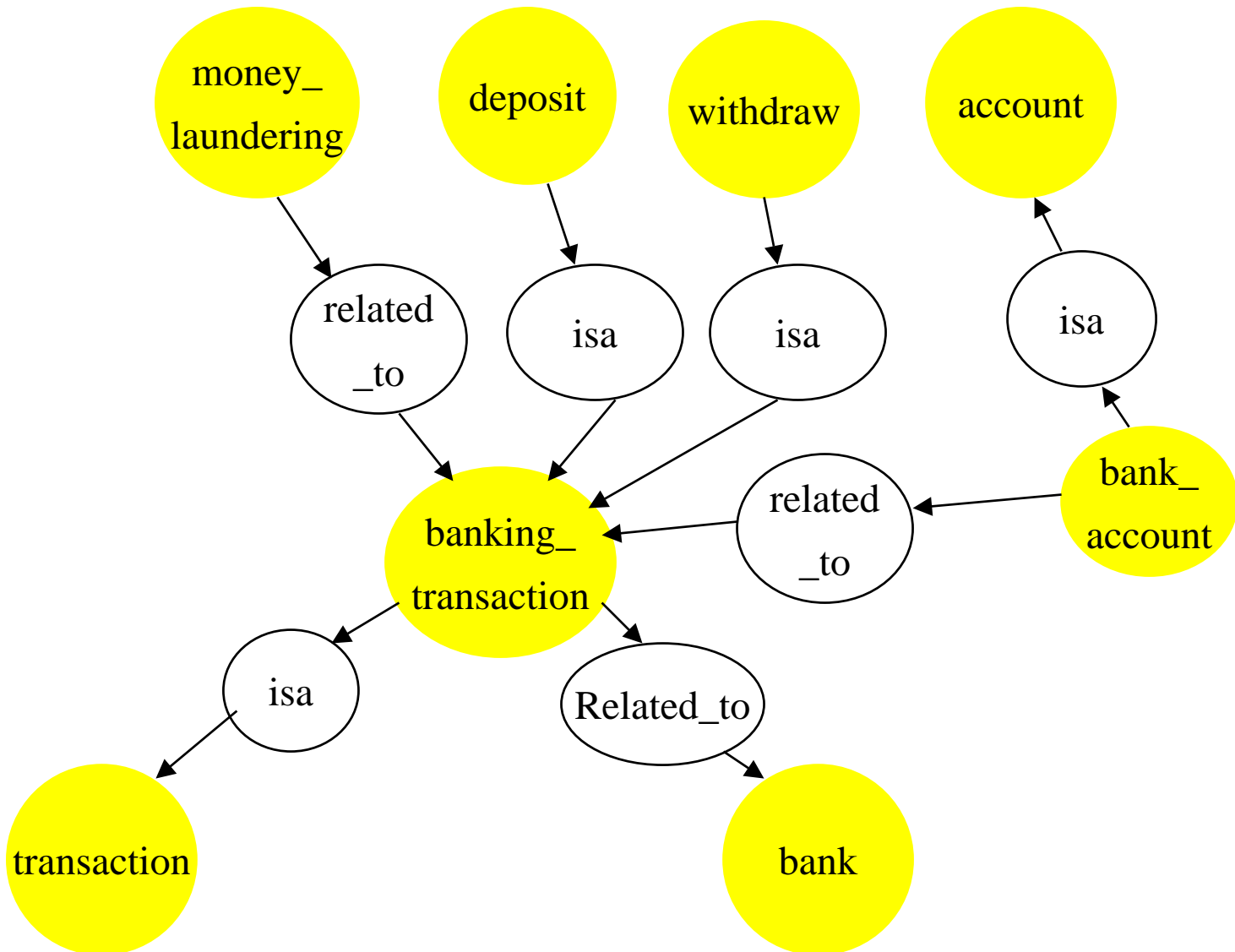


Updated Interest Set

Interest concept	Interest level
abdul_ramazi	1.0
chicago	1.0
bank_account	0.6125
first_union_national_bank	1.0
.....	

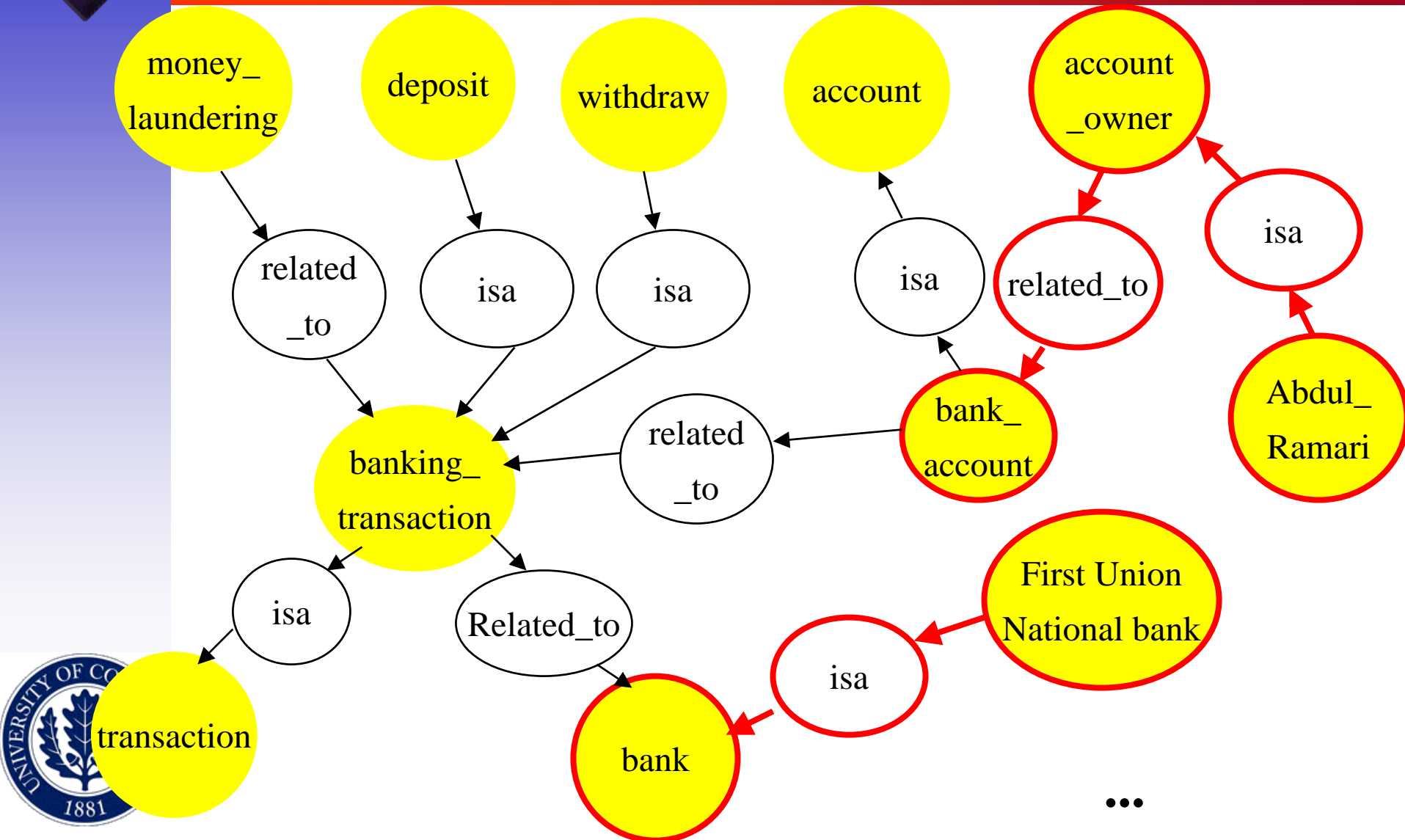


Existing Context Network

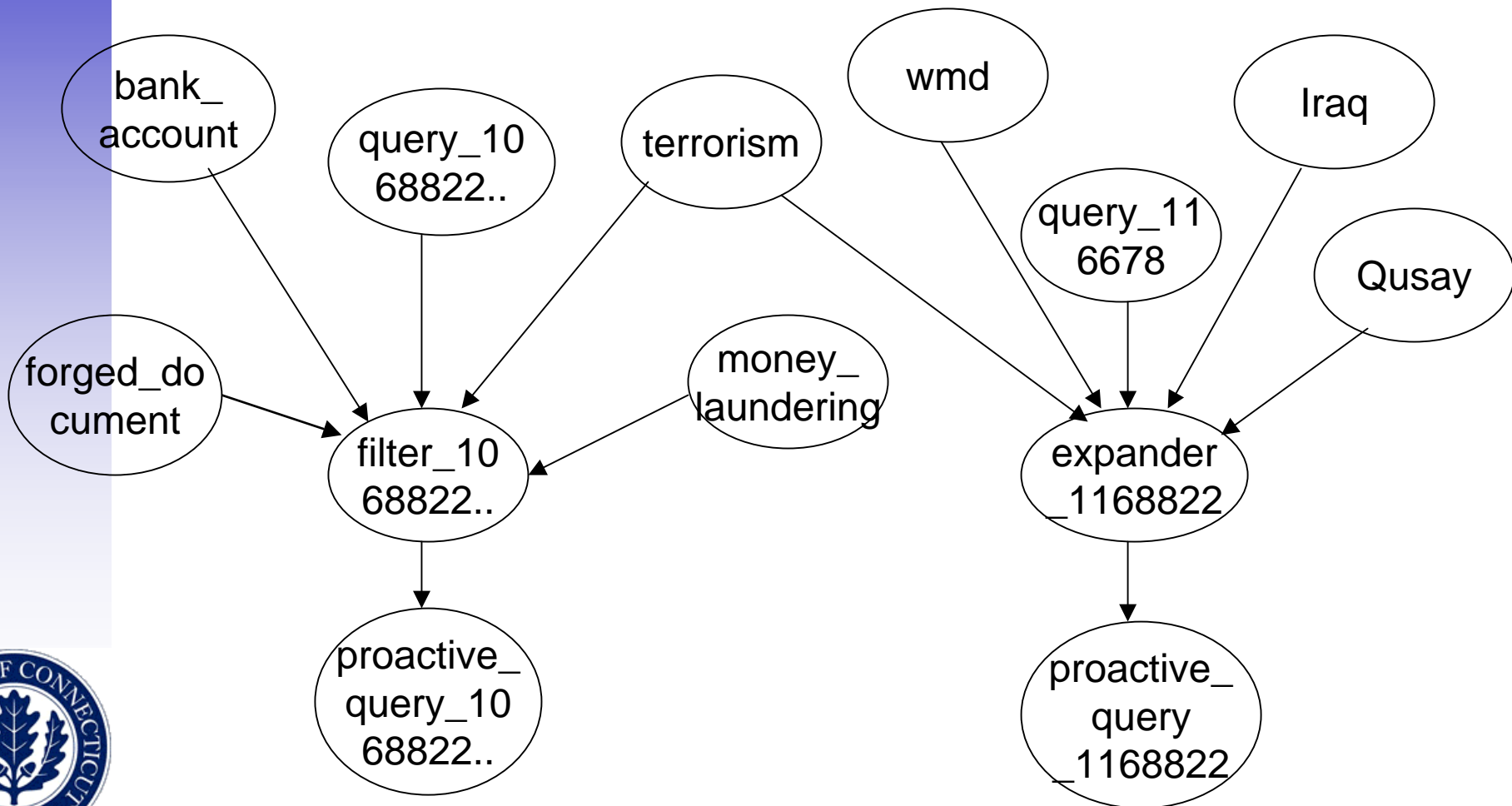




Updated Context Network

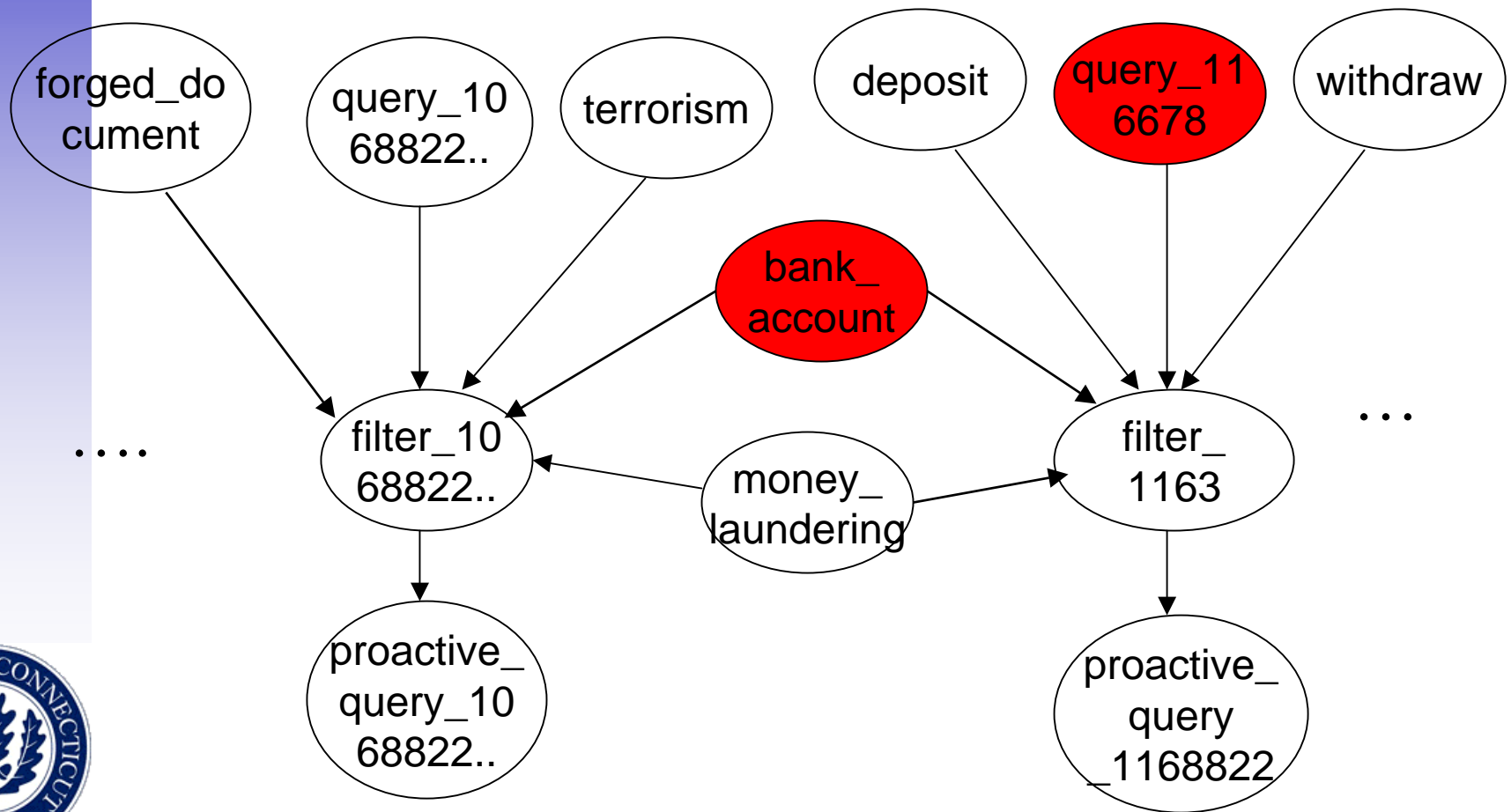


Existing Preference Network





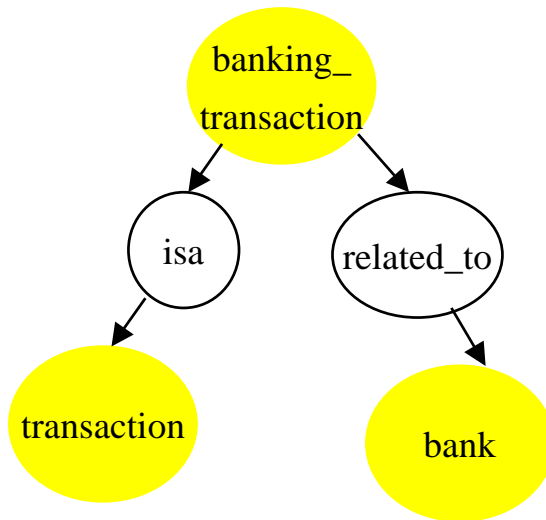
Updated Preference Network



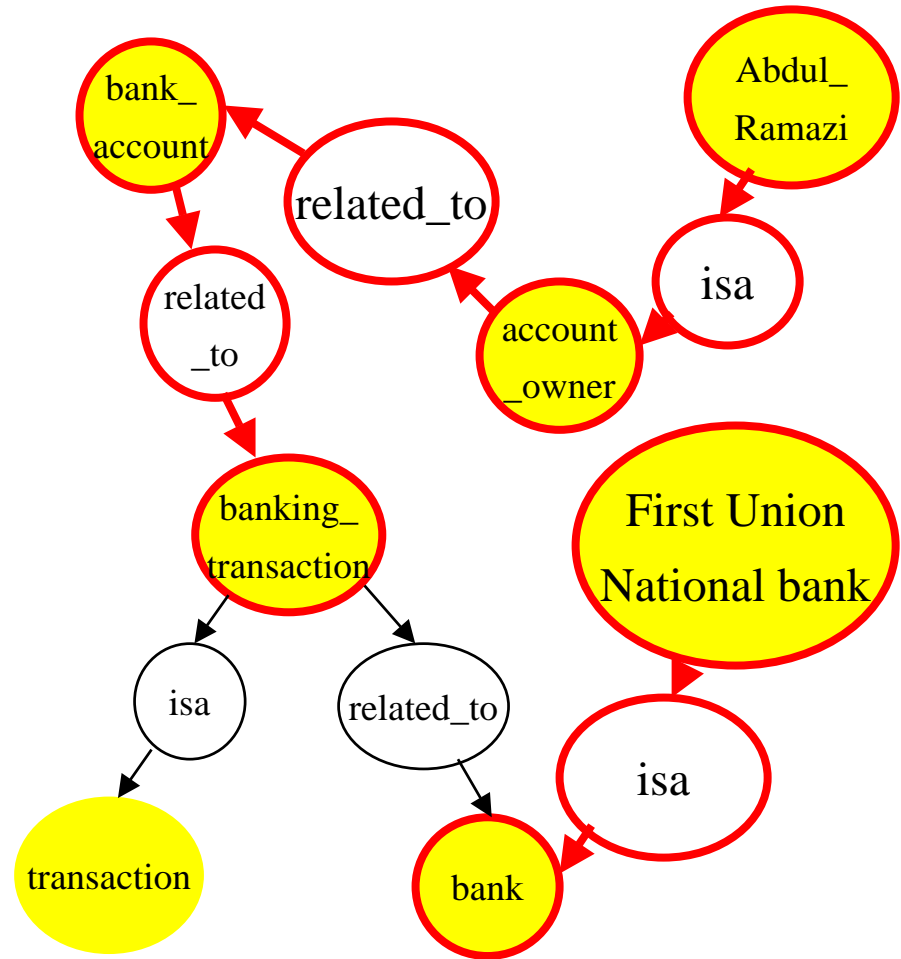


Modified Query Graph

Original query graph



Modified query graph





Outline

- Motivation & Objectives
- Our approach
- **Experiment**
- Results
- Conclusion



Experimental setup

- Evaluation took place in NIST laboratory
- Test bed: Center for Nonproliferation Studies (CNS) database (2003) on WMD, arms control and terrorism
- Three analysts took part in the experiment
- Our User Modeling (UM) approach vs. Verity Query Language



Procedure

- Each participant filled out an entry questionnaire when started and an exit questionnaire when finished
- Subjects use two different systems in parallel.
 - There are 10 scripted queries on “*Qumar R&D supporting BW*”.
 - Only first 10 documents have been reviewed for relevancy.
 - Marking relevant documents received after each query on intermediate questionnaire



Queries

1. Qumar research biological warfare
2. Qumar research institute, university biological warfare
3. Qumar biological research and biological warfare
4. Biological research facilities in Qumar
5. Intelligence assessment on Qumar biological research
6. Qumar foreign connections in biological weapons program
7. Bacu, Qumar, and Russia connections to WMD
8. Qumar's biologists visits Bacu
9. Russian biotechnology, missiles, aid to Qumar
10. China supply and Qumar biological weapons program



Users' profiles

	1	2	3
Highest degree	JD	MS	BA
Analysis experience	7 years	5 years	5 months
Computer expertise	novice	medium	medium
Use computer	no	yes	yes





Outline

- Motivation & Objectives
- Our approach
- Experiment
- **Results**
- Conclusion





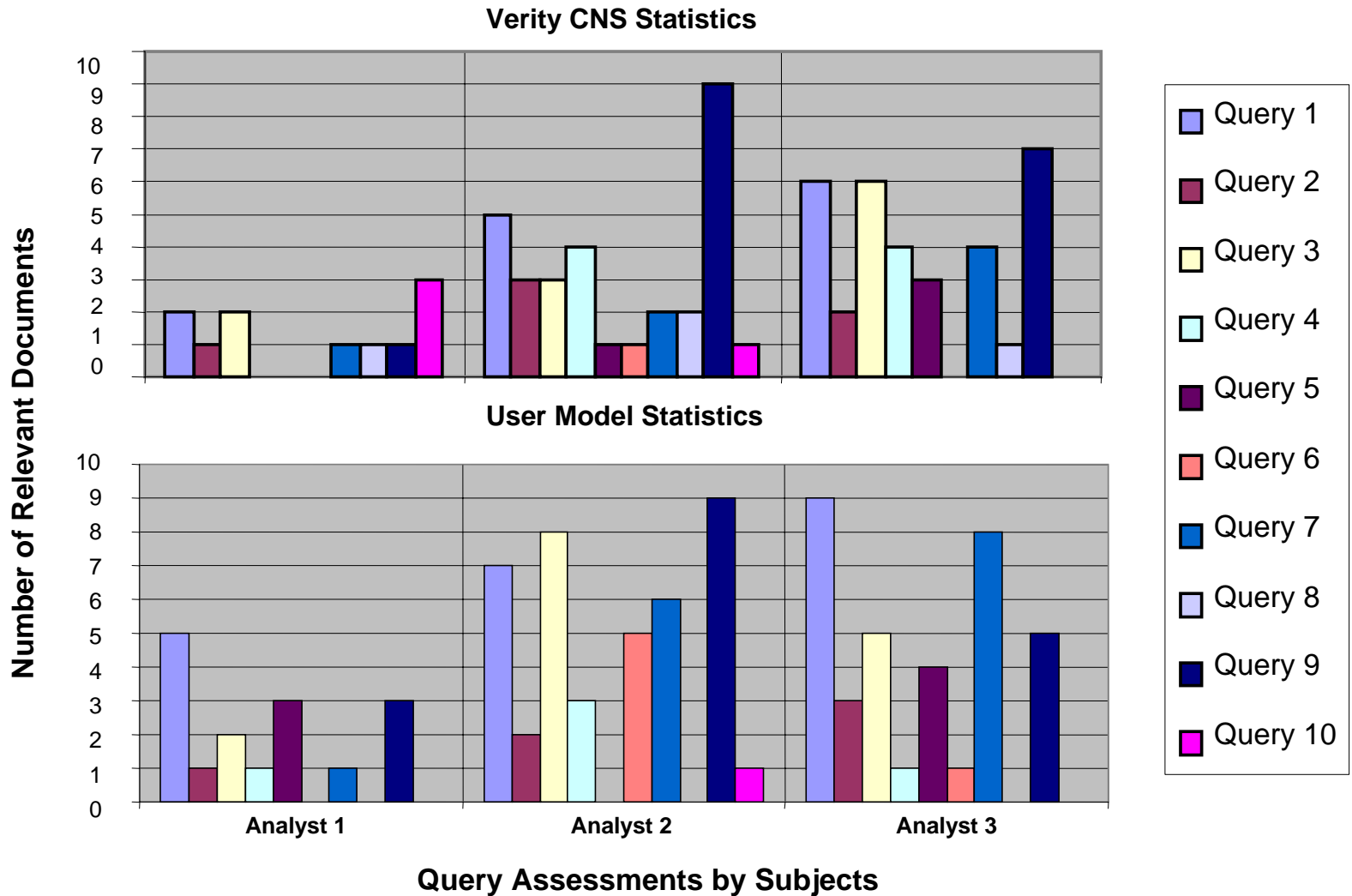
Entry Questionnaire

Question	1	2	3
What is your overall experience with systems using ranked outputs and full-text databases, such as Google? 1-7, 1 is very experienced, 7 is no experience	3	1	1
Have you ever used a system that asked you to indicate whether a document or other system response was relevant? Yes, No	yes	yes	no
When faced with a search problem do you tend to: Look at big picture first Look for details first Both	both	big picture	details
What is your knowledge of Terrorism 1-7, 1 very experienced, 7 no experience	2	3	2
What is your knowledge of WMD? 1-7, 1 very experienced, 7 no experience	3	2	2





Query Assessments





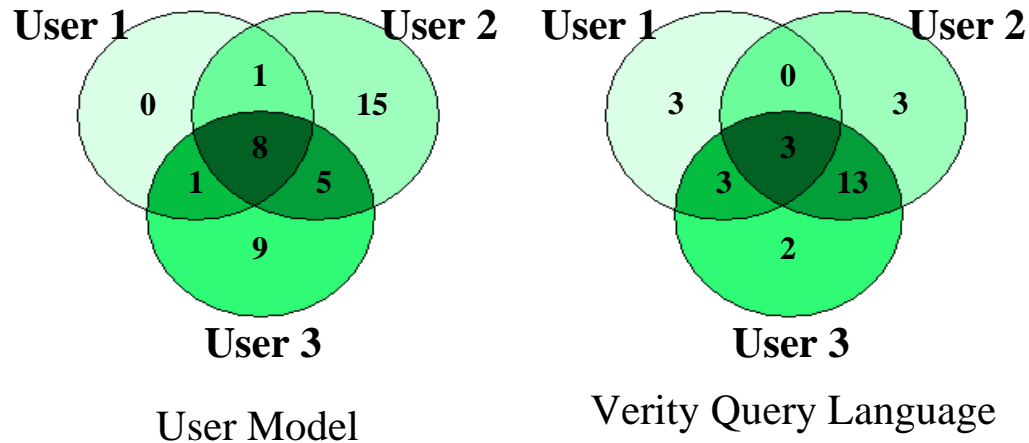
Results

	VQL system			UM system		
	1	2	3	1	2	3
Analyst						
Documents presented	100	100	90	100	100	90
Relevant documents	11	31	33	16	41	36
Unique document presented	49	49	45	67	72	54
Unique relevant documents	9	19	21	10	29	23
Precision	0.25			0.31		





Results (continue)



	User Model	Verity
Total unique relevant documents	39	27
Documents marked as relevant by all 3 analysts	8	3
Documents marked as relevant by more than 2 analysts	15	19
Documents marked as relevant by only 1 analyst	24	8



Summary

- Retrieves more relevant documents compared to Verity Query Language.
- Tracks individual differences

Question	Score
How satisfied are you with the overall results for this task using system with user model	4.3
How confidence are you with the results that they cover all possible aspects of the task	4.7
The regarding this task, do you think that user modeling approach helped you to retrieve critical document earlier in the process than the VQL	3.7
Ranking of mental demand	5.3
Ranking of physical demand	2.0
Ranking of temporal demand	5.0
Ranking of performance demand	4.7
Ranking of frustration	5.3
Ranking of effort	6.0



Conclusion

- UM assisted IR has retrieved more relevant documents than the Verity Query Language
- UM helps system track the interests and focus of individual analyst by retrieving more unique documents that relevant to each analyst' interests
- Overall, encouraging and positive feedback from intelligence analysts





Acknowledgement

This research has been funded by Advanced Research and Development Activity (ARDA) – US Government.

Special thanks to Dr. Jean Scholtz and Emile Morse who helped organize the evaluation, collect data and provide the preliminary data analyses.

