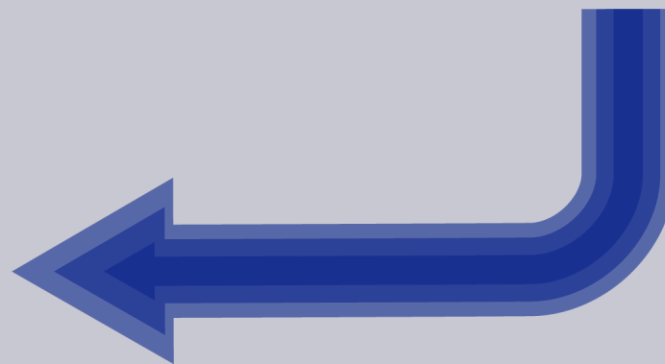


Intranet Mediator Inc.





Integrating the IIT Intranet Mediator with the United States Holocaust Memorial Museum

IPRO 311

Fall 2005



Team Members

Faculty Advisors

Ophir Frieder; CS

David Grossman; CS

Student Team Leader

Jordan Wilberding; CS

Coding Team

Tristan Sloughter; CS

John Quigley; CS

USHMM Team

Gideon Frieder

Data Integration Team

Kevin Harriss; CS

Jonathan Young; BME

Nicholas Lee; EE

Deliverables Team

Dale Player; EE

Evan Farrar; CS

Sponsor

Intranet Mediator
Inc.



Problem

- USHMM spent 10 years gathering data
- Data are in different languages
- Data are in different formats
- Need a way to efficiently search and distribute data



Solution

The Intranet Mediator

- The ultimate search engine



Goals for Fall '05

- Migrate to new Mediator
- Replace USHMM's search engine
 - Convert old rules to new Mediator format
 - Add new data sources
 - Implement new rules for additional data resources
 - Create intuitive user interface
- Improve the search quality of the Mediator
 - Stress test new Mediator
 - Spell correction



Key Accomplishments

- Interacted closely with the USHMM
- Expanded query log coverage to 75%
- Converted 250 rules to new Mediator format
- Added 100 queries to new Mediator format
- Successfully stress tested
- Created convenient website user interface
- Added spell corrector
- Extensive documentation
- Successfully migrated to new Mediator



The New Mediator

- What is the Mediator?

The IIT Intranet Mediator is a data mediator designed specifically for an intranet, capable of integrating a variety of data sources.

- Our New Design
 - Adheres to modern design standards
 - Components are distributable
 - RDF result return
 - Cached results



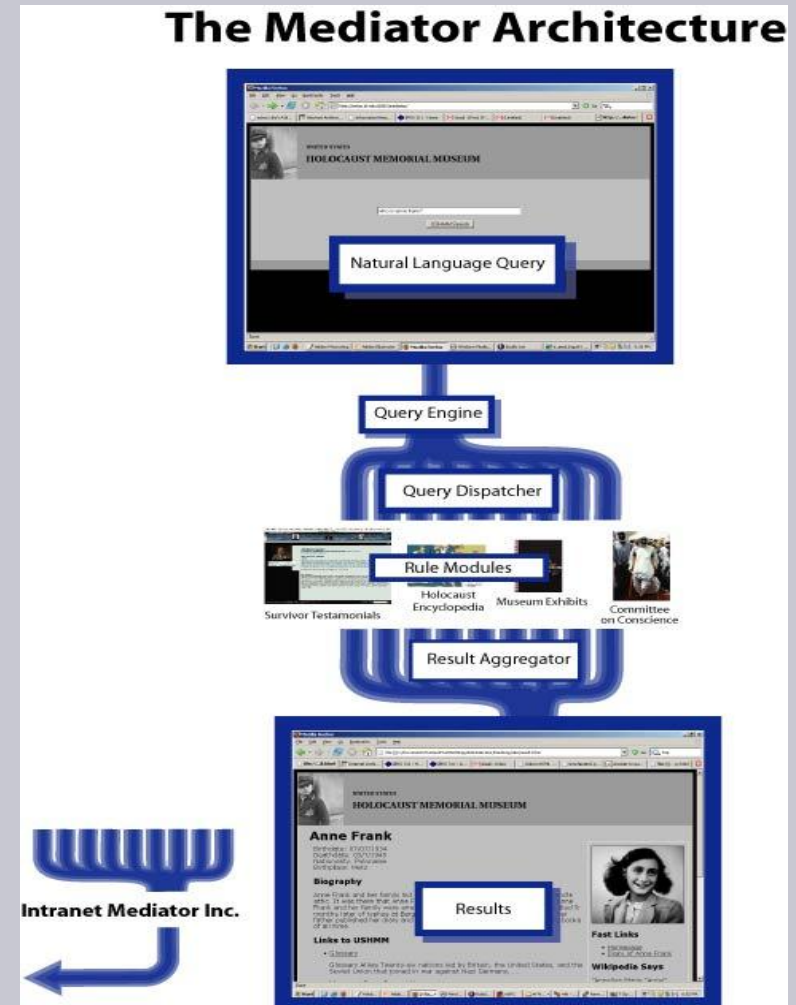
The Power of Mediator

- Simple, Google-esque user interface for query input
- Interprets natural language queries
- Speaks many data type 'languages'
- Tailors results to specifically answer questions
- Easy integration of new data
- Integrates static and dynamic data



The Mediator Architecture

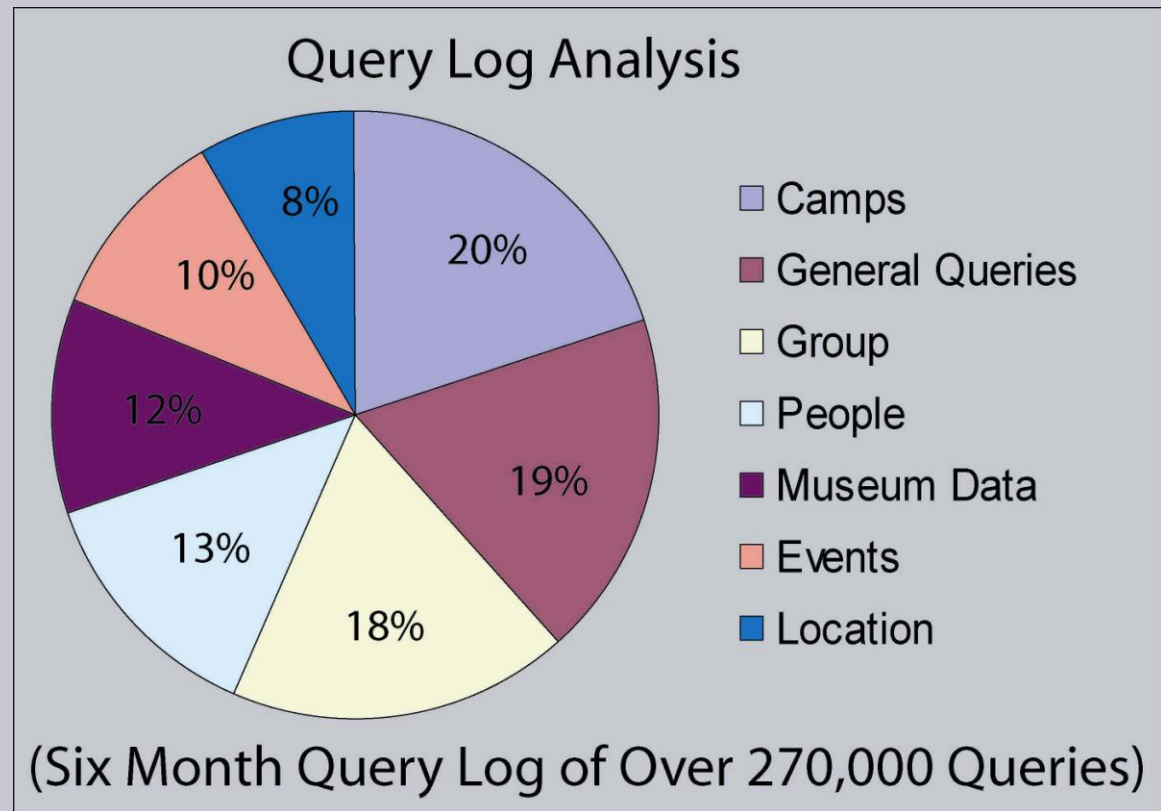
- Capabilities
 - Natural Language Queries
 - Multi-Structured Data Management
 - Exact Answers to a Query
- Essential Mediator Elements
 - User Interface
 - Query Engine
 - Rule Modules
 - Result Aggregator





Query Analysis

- Top Searches
 - Holocaust (6,409)
 - Auschwitz (5,190)
 - Anne Frank (4,799)
 - Concentration Camps (3,573)
 - Hitler (2,934)
 - Kristallnacht (2,369)
 - Children (2,183)
 - Pictures (1,991)
 - Photographs (1,915)
 - Dachau (1,534)
 - Adolf Hitler (1,463)





Data Sources

Structured

- Name Search
 - Holocaust Survivor Testimonials
- Learning Center
 - USHMM data set of maps, testimonials, and articles
- Fact Database
 - Common Query Information
- Wikipedia
 - Web-based encyclopedia
- Census Data
 - Slovakian Census
- Card Catalog
 - Information on Holocaust literature

Semi-Structured

- News Feeds
 - Articles from USHMM for Anti-Semitic and genocide awareness

Unstructured

- Web Crawl
 - Data contained on the USHMM website



Rule Development

- Rules from the past Mediator implementation were converted to the new format
- New rules were created based on additional questions created this semester



Rule Format

- Rules can consist of five parts:
 - Questions
 - Format of questions this rule matches
 - Required Fields
 - Fields required for this rule to be matched
 - Provided Fields
 - Fields this rule provides results for
 - Logic
 - How to retrieve the results
 - Answer
 - Format to return the results in



Rule Testing

- The new Mediator needed to be stress tested
- Rules needed to be tested for accuracy



Future Plans

- Extend Mediator Core Engine
 - Grammatical Correction
 - Cross-Lingual Queries
- Extend Data Collection
- Migrate to USHMM's Official Site
- Add Features Requested by USHMM



Acknowledgements

- Jeff Heard – Intranet Mediator Inc., CTO
- Larry Kane – Intranet Mediator Inc., CEO
- Jay Fisher – Director of Ed Kaplan
Entrepreneurial Studies