CSCI-548:
Information Integration on the Web
Spring 2015

Jose Luis Ambite

Craig Knoblock

University of Southern California

Course Overview

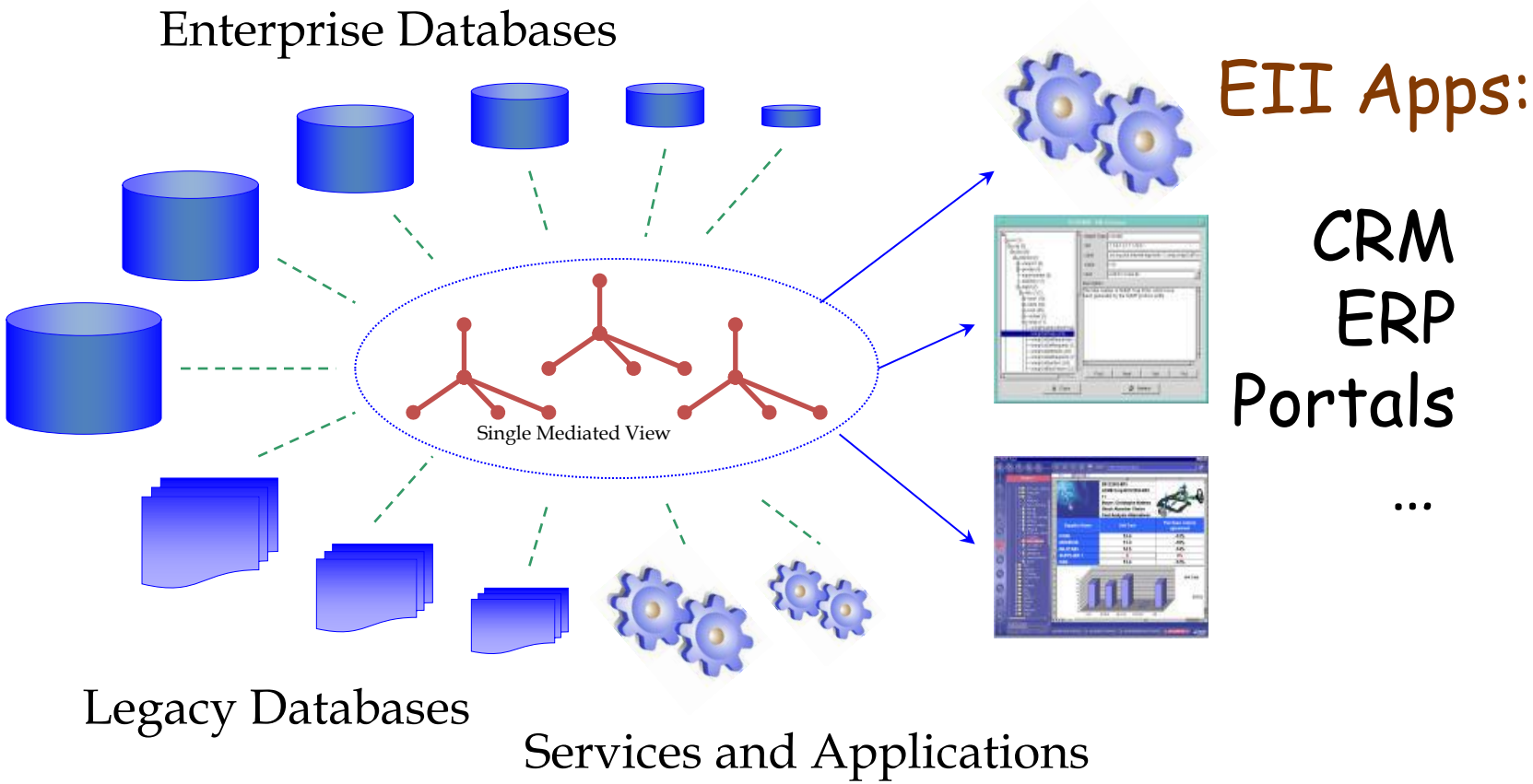
Information Integration Challenges

- Accessing the data
- Understanding the data
- Resolving differences
  - Schema-level
  - Data-level
- Efficiency
  - Query evaluation
  - Data transformation

…on the Web

- The Web is an incredible source of data
- New challenges arise:
  - Need to turn web pages into structured data
  - Don't have control over the data
  - Sources have input/output constraints
  - Distributed nature of the web can make integration slow
  - How to relate information across sources

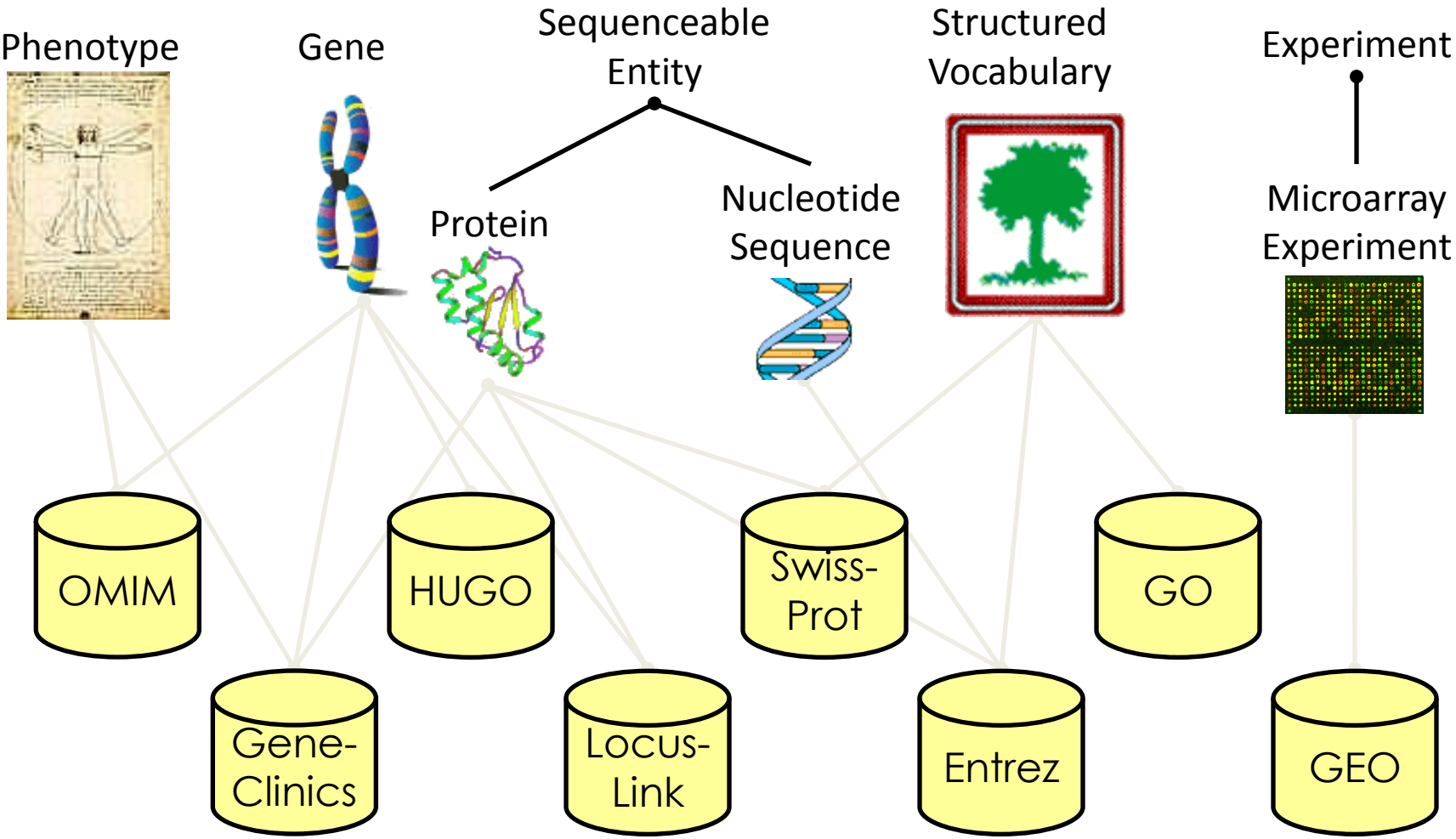
Application Area 1: Business



50% of all IT $$$ spent here!

[from Doan, Halevy, Ives, Principles of Data Integration, 2012]

Application Area 2: Science



Hundreds of biomedical data sources available; growing rapidly!

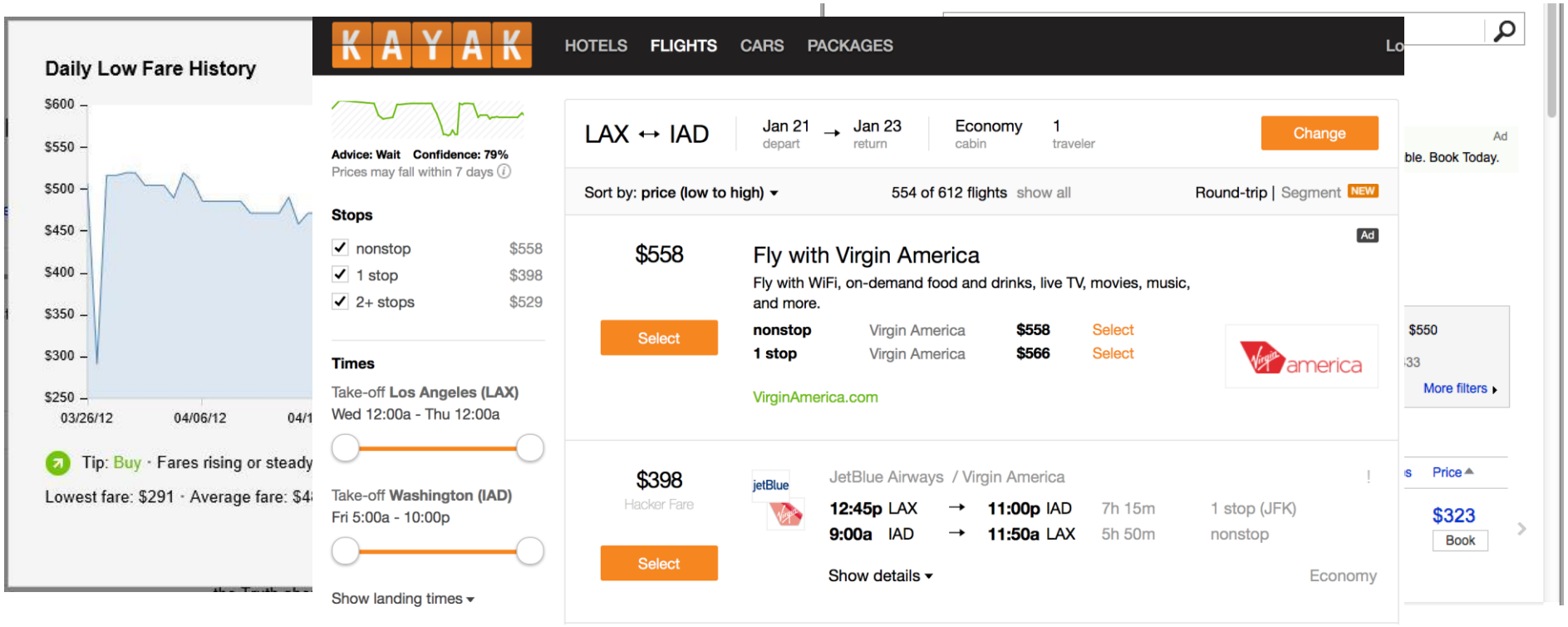
[from Doan, Halevy, Ives, Principles of Data Integration, 2012]

Application Area 3: Web & Deep Web

[from Doan, Halevy, Ives, Principles of Data Integration, 2012]

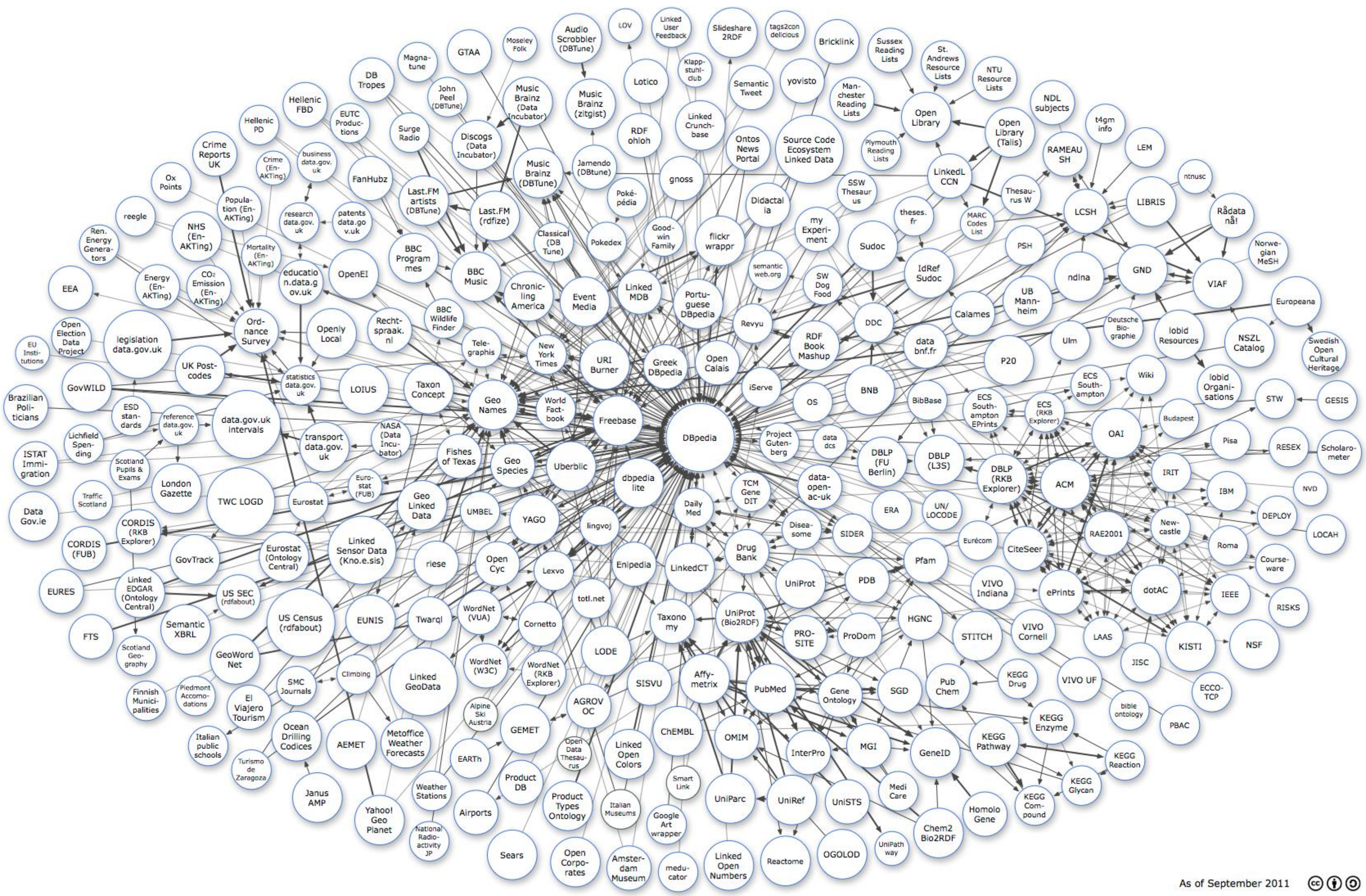
Web Example: Airfare Prediction

- Monitor, integrate flight prices across airlines
- Predict cheapest time to buy flight



- [Etzioni, Knoblock, et al. KDD 2003], lead to startup: Farecast, bought by Microsoft, incorported in travel.bing.com, now in Kayak

Application Area 4: Semantic Web & Linked Data



As of September 2011

Application Area 5: Big Data

- Dimension of Big Data:
  - Volume: size of data in bytes, files, or data objects.
  - Velocity: rate or frequency at which new data can be acquired or processed.
  - Variety: diversity in data sources, data models, data types: geospatial, genetic, imaging, relational, XML, web services, RDF, …
  - Variability: diversity in data interpretation or meaning.
- This course address Variety and Variability
- Without semantics, big data = big mess

Overview of Lecture Topics

Semantic Web

- A machine-readable semantic layer on the web
- Research Topics
  - Organizing knowledge
  - Formal languages: XML, RDF, OWL, FOL
  - Reasoning and querying over semantic web data
  - Provenance, Trust
  - Linked Data

Can You Google the Answer to These Questions?

- Which US cities have more than 500,000 people?
- Which congress members voted “No” on pro environmental legislation in the past four years, with high-pollution industry in their congressional districts?
- How many Twitter followers does Obama have in each state?
- Find all stores that sell a Canon 7D for less than $1,700


PALOS VERDES PENINSULA NEWS

GOOGLE TO HARVEST RDF-A

SATURDAY, MARCH 15, 2004

Search giant goes semantic

Norris Center for the Performing Arts, 27570 Crossfield Drive in RHE, on Saturday, March 20 at 2 and 4 p.m. Tickets are $38 for adults and $19 for students for the evening performance and $30 for adults and $15 for students at the matinee performance. For reservations, call 544-0403.

• UPCOMING — The Palos Verdes Peninsula Unified School District and Friends of School Music host the 15th Palos Verdes Elementary Choral Festival on March 23, 24 and 25 at the Norris Center for the Performing Arts, 27570 Crossfield Drive in RHE. All shows begin at 7:30 p.m. For tickets, call the Norris box office at 544-0403.

• ONGOING — The Distinctive Edge, 29050 S. Western Ave., Suite 113 in RPV, continues "Third Time's a Charm," an exhibit of 3-D collages by artist Steve Jacobsen, through March 30. For gallery hours, call 833-3613.

• ONGOING — "Natural Treasures" exhibition contin-

Two years ago, U.S. Navy personnel and their families assigned to the Atsugi Navy base, home of the U.S.S. Kitty Hawk, were treated to a rare experience when Terry Fleming and his local Irish/American band, Innisfree, traveled to the base to entertain them on St. Patrick's Day. Fleming and the other five members of Innisfree were delighted and honored to be able to go to Japan and lift the spirits, if only for a few hours, of the Navy personnel and their families.

For the third year in a row, Fleming — a local insurance broker in Rolling Hills by day and an entertainer by night — and the band traveled to entertain the Navy men, women and families at various bases throughout Japan.

Fleming, the leader of the band on accordion and harmonica, actually is the only member of the band from Ireland. Other members include lead singer Julie Delaney, a civil engineer in Newport Beach; Terry Doyle, guitar, a news director with CBS news; Denis Doyle, Celtic harpist, a professor at Glendale College; Kevin Wood, keyboards and bagpipes, music teacher and assistant director of the Orange County Symphony; and Mike Tiffany, bass, a computer engineer. The band has been playing the length and breath of California for the past 25 years. They have played at galas, wakes, weddings, birthdays and on occasions where there was little excuse for throwing a party.

Fleming says it was by coincidence the band got the opportunity to travel to Japan. Another band was unable to travel at the last minute and so he and his band were offered the opportunity to go in their place.

With some trepidation they made their first trip and with the overwhelming response they received at Atsugi, any fears they had were quickly allayed. On a damp St. Patrick's Day, hundreds of families, clad in many shades of green, whooped it up, sang their hearts out and danced up a storm. As the evening wore on, many in the audience were emboldened to try their hand or foot at the Irish gig, with much encouragement from the band.

Even though far from home, the Atsugi base — situated a few hours south of Tokyo — felt like home away from home, with its lush green rolling landscape and its multitude of cherry blossom trees. "Yet," Fleming says, "one was struck by the commitment and dedication of our men and women in uniform as they played their part in protecting and serving in an ever challenging and hostile world."

A couple of the families even took time out from their busy schedule to host Terry and the band members. They treated them to a guided tour of the base and accompanied them on a few exciting trips off the base, visiting beautiful ancient temples, monuments and revered giant Buddas.

On a visit to downtown Tokyo, the band came across what they assumed was a very rare sight, a place called "Scruffy Murphy's," an Irish pub located in the heart of a bristling downtown. Upon checking the establishment out, they discovered a real authentic Irish Pub with excellent Guinness and good pub grub. It also happened to be open mic night, so the band members took over the stage and entertained the locals for a few fun-filled hours. It turned out that it was just one of many establishments in the city.

A special bond developed between the band members and these families and already exchange visits have occurred when the same families were on leave in the United States.

For more information about the band, log on to www.InnisfreeMusic.com.

thebridge

http://googlewebmastercentral.blogspot.com/2009/05/introducing-rich-snippets.html

slide from Frank van Harmelen

slide from Frank van Harmelen

http://www.readwriteweb.com/archives/calais_4_linked_data.php

New York Times Index


The New York Times.

TITANIC SINKS FOUR HOURS AFTER HITTING ICEBERG; 866 RESCUED BY CARPATHIA, PROBABLY 1250 PERISH; ISMAY SAFE, MRS. ASTOR MAYBE, NOTED NAMES MISSING

The Lost Titanic Sailing Toward Out of Belfast Harbor.

PARTIAL LIST OF THE SAVED.

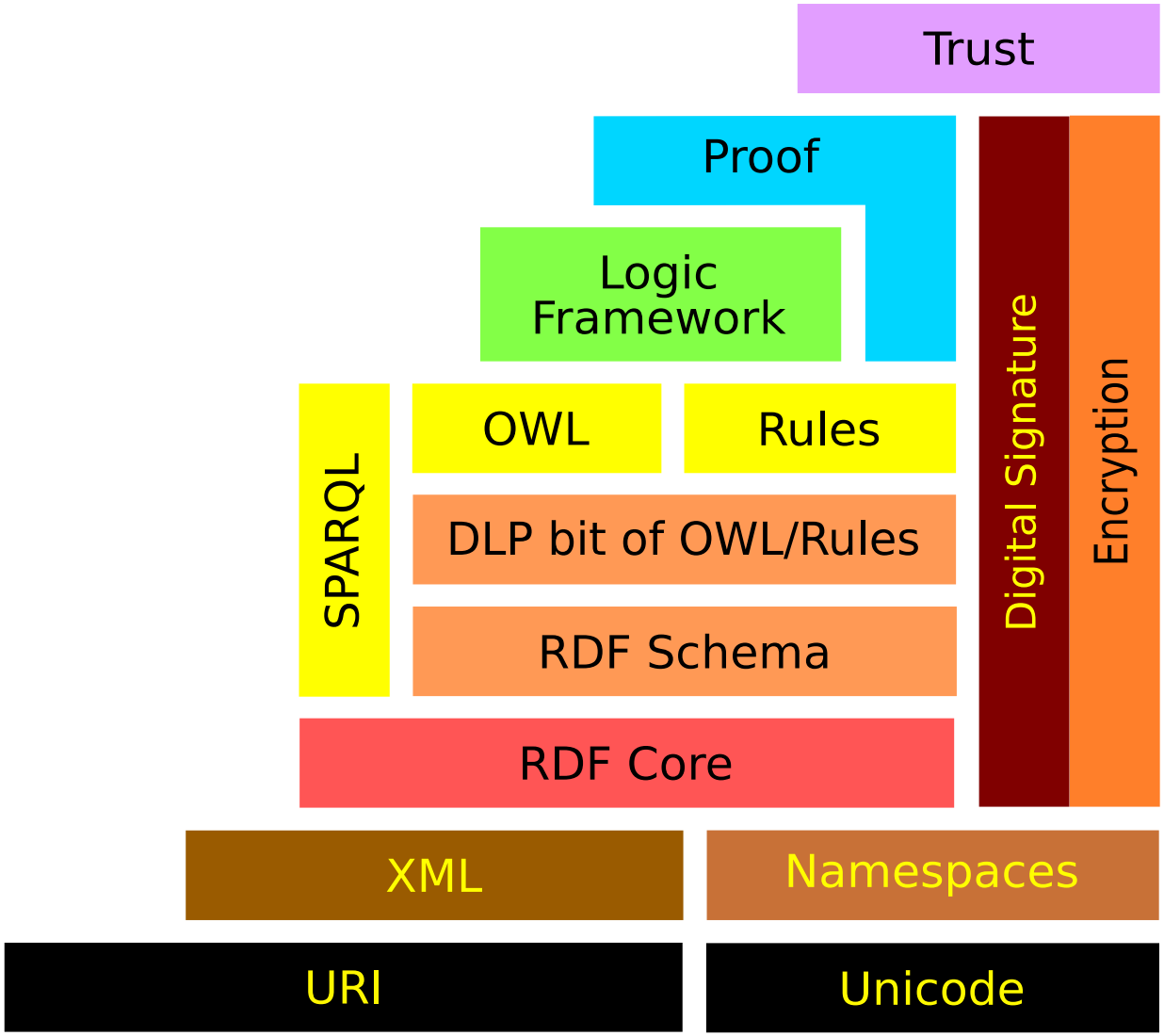
- Started in 1851,
  cross ref'd every NYT article,
  published since 1913
- 10k tags on all kinds of topics
- 3500 in very frequent use
- Open License, usable as
  reference ontology for anybody
- BTW: strong internal use of semantic
  technologies for … targeting advertising

slide from Frank van Harmelen

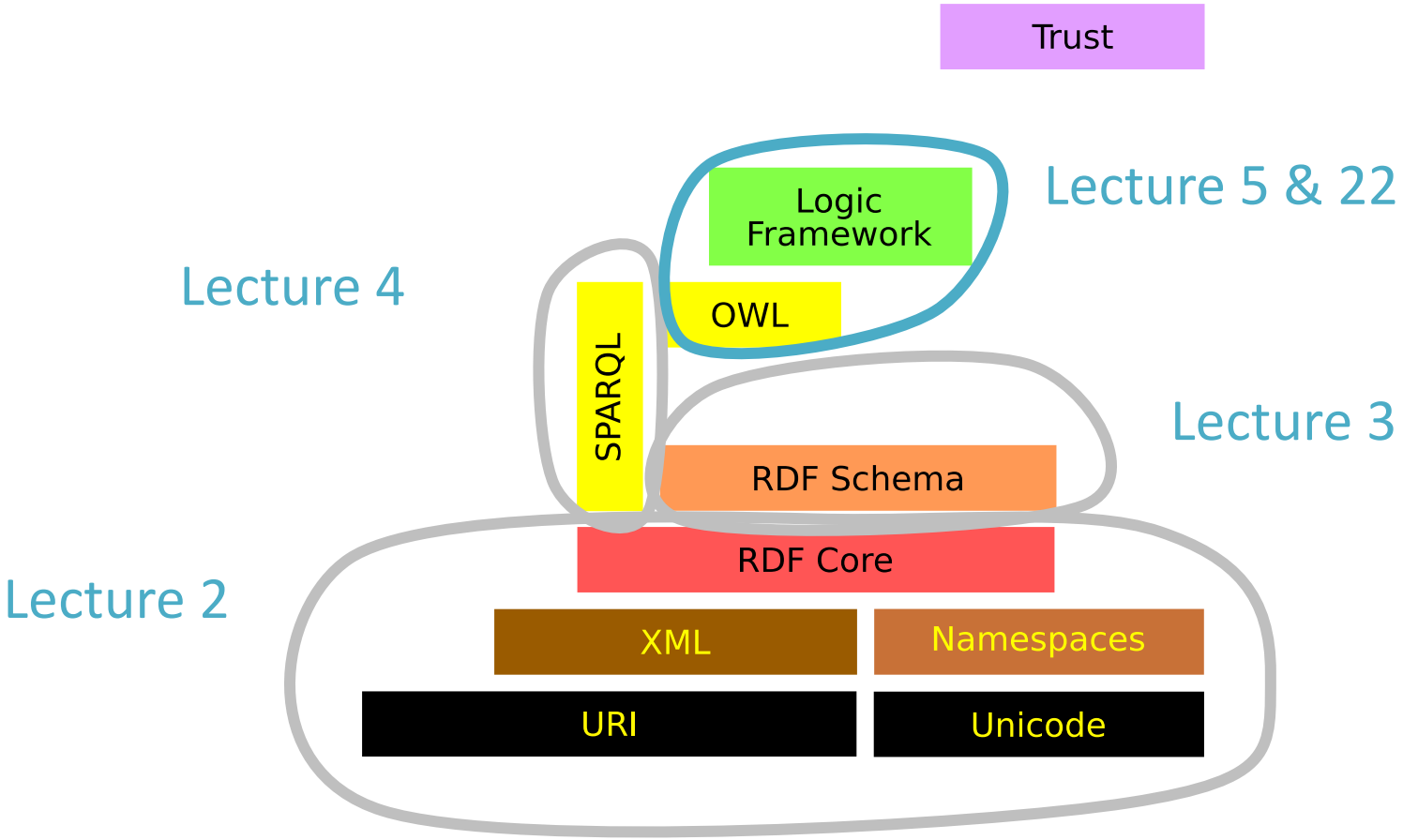
BBC World Cup Web site

slide from Frank van Harmelen

Semantic Web Layer Cake



Semantic Web Layer Cake



Semantic Web: Description Logics, Ontology Web Language (OWL2)

formal underpinning of the semantic web

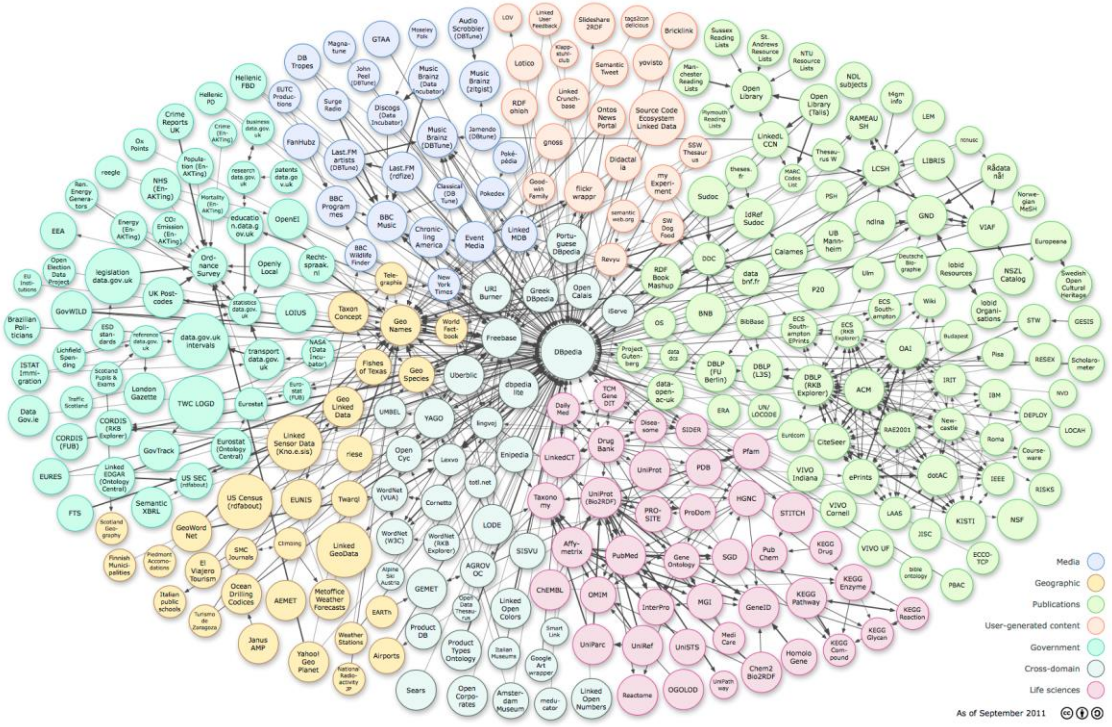
Constructor	DL Syntax	Example	FOL Syntax
intersectionOf	$C_1 \sqcap \ldots \sqcap C_n$	Human $\sqcap$ Male	$C_1(x) \wedge \ldots \wedge C_n(x)$
unionOf	$C_1 \sqcup \ldots \sqcup C_n$	Doctor $\sqcup$ Lawyer	$C_1(x) \vee \ldots \vee C_n(x)$
complementOf	$\neg C$	$\neg$Male	$\neg C(x)$
oneOf	$\{x_1\} \sqcup \ldots \sqcup \{x_n\}$	{john} $\sqcup$ {mary}	$x = x_1 \vee \ldots \vee x = x_n$
allValuesFrom	$\forall P.C$	$\forall$hasChild.Doctor	$\forall y.P(x,y) \rightarrow C(y)$
someValuesFrom	$\exists P.C$	$\exists$hasChild.Lawyer	$\exists y.P(x,y) \wedge C(y)$
maxCardinality	$\leqslant nP$	$\leqslant$1hasChild	$\exists^{\leqslant n} y.P(x,y)$
minCardinality	$\geqslant nP$	$\geqslant$2hasChild	$\exists^{\geqslant n} y.P(x,y)$

HappyParent = Person $\wedge$ $\forall$hasChild.(Doctor$\vee$Rich))

Linked Data

a giant Web-scale graph of data about everything

Wikipedia: "… recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF."



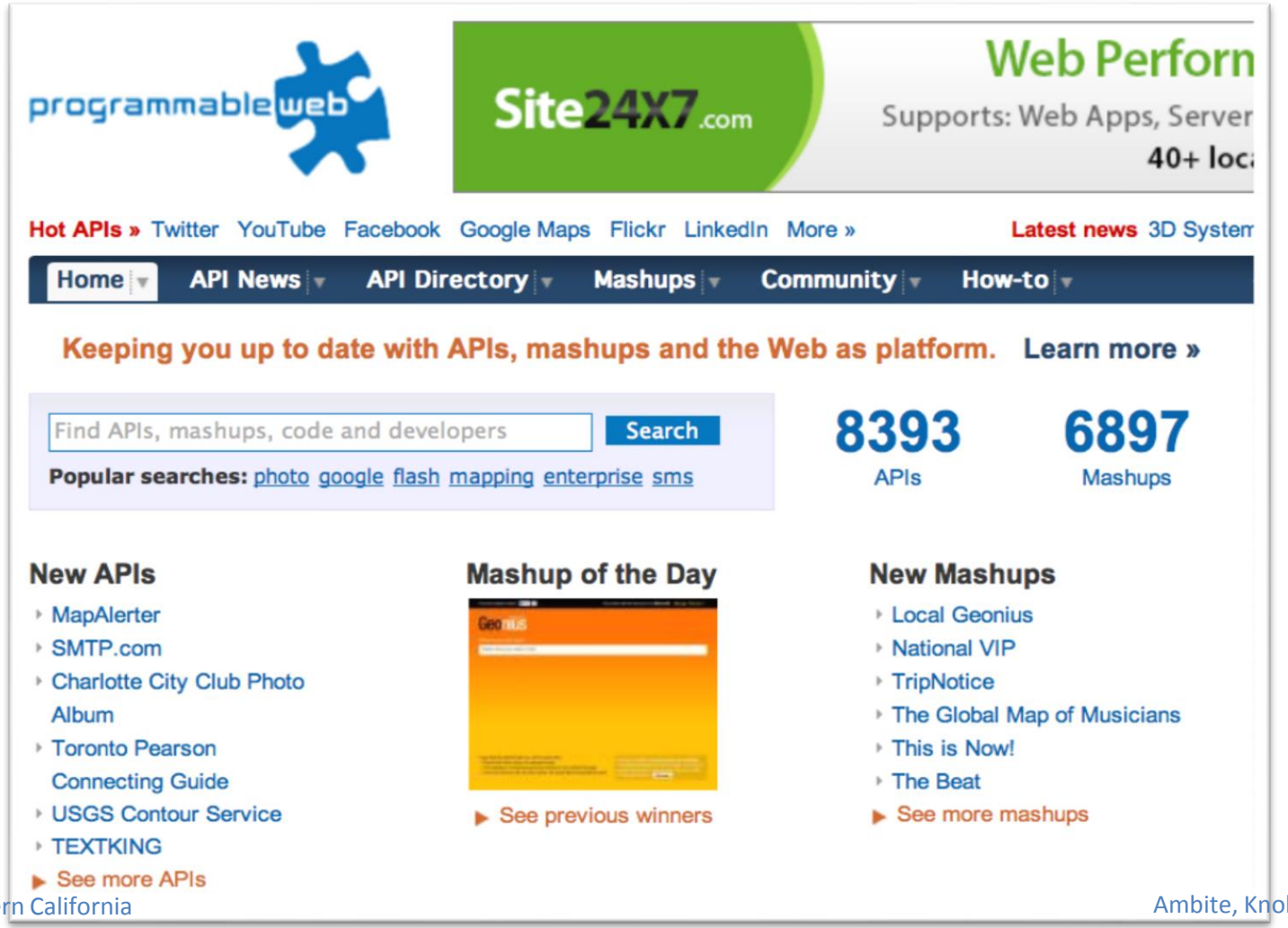
31+ billion assertions

Linked Data

- Why Linked Data
- The Linked Data principles
- Web of data
- How to Publish Linked Data on the Web
- Tools
- Applications

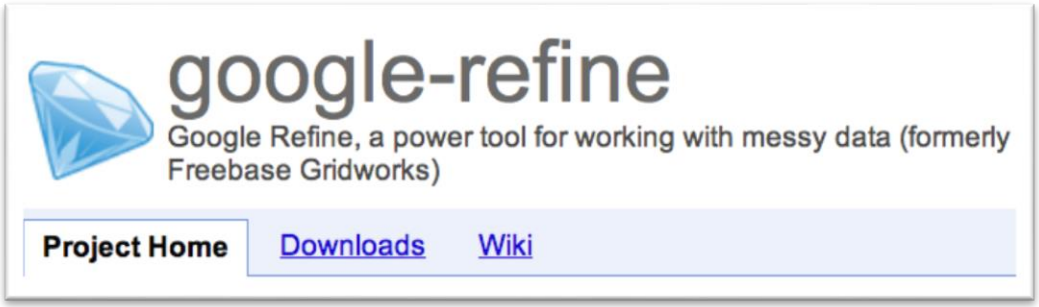
Services Over Linked Data

http://www.programmableweb.com

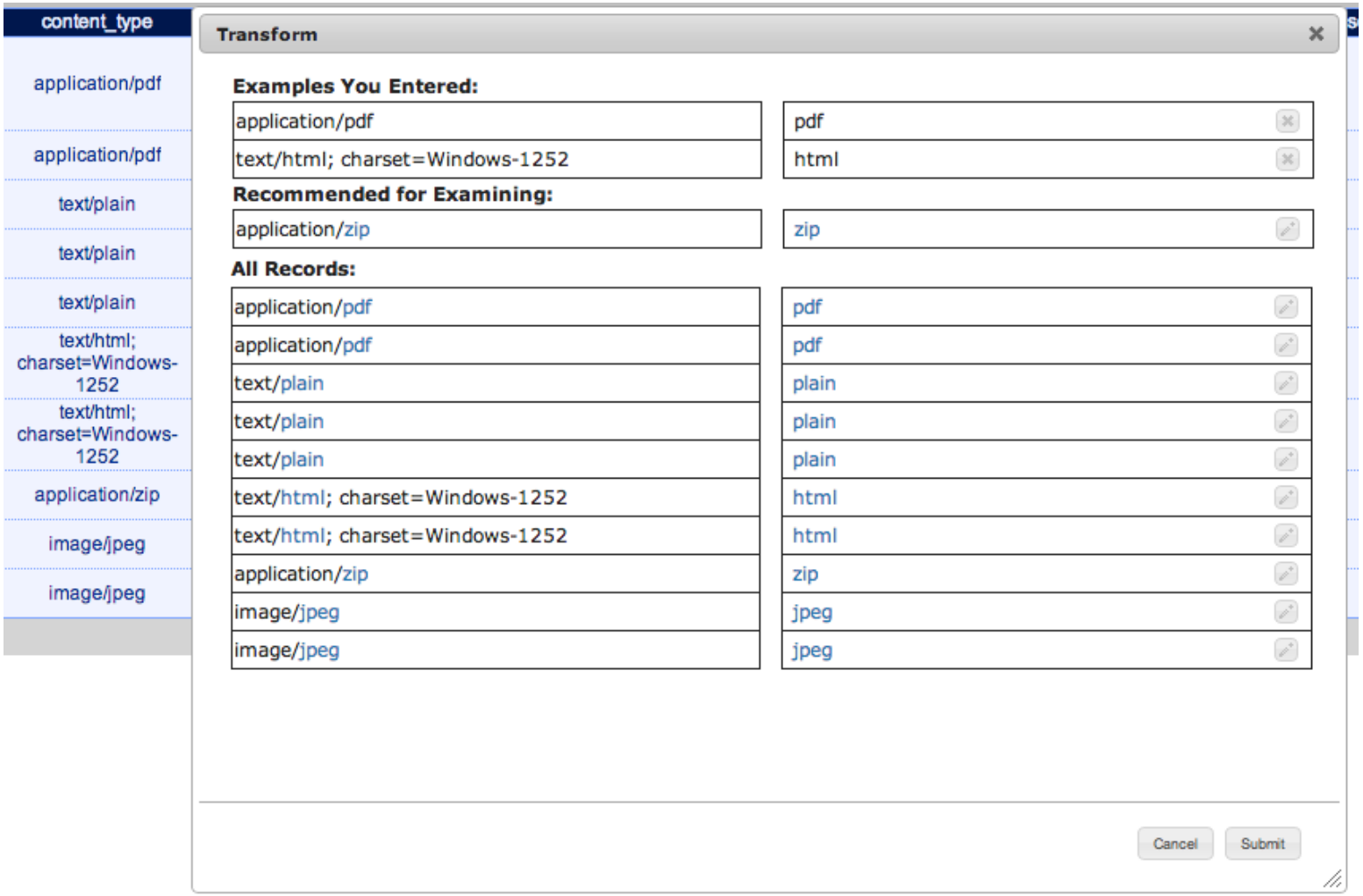


Data Cleaning

- Types of dirty data
- Challenges
  - Detection
  - Fixing
  - Large data sets
- Techniques
- Research problems



Data Transformation By Example



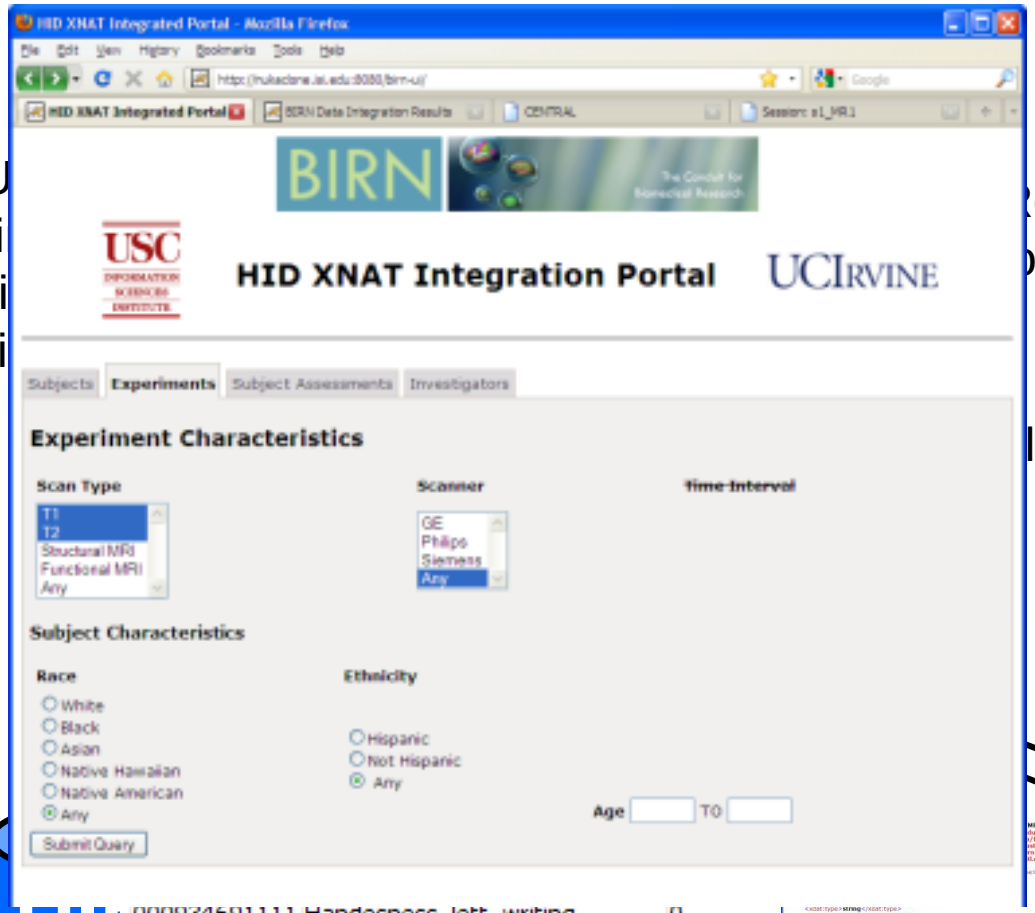
Database theory basics: queries, query containment, Datalog

- Conjunctive queries
- Query containment
  - q1(x) :- r(x,x), r(x,y1), r(y1, y2), r(y2, y3), r(y3,x)
  - q2(x) :- r(x,x) [select * from R(a,b) where R.a=R.b]
  - q1 = q2 ???
- Recursive logic programs: Datalog
  - path(X, Y) :- arc(X, Y)
  - path(X, Y) :- path(X, Z), path(Z, Y).

Logical Data Integration

- Warehouses: Extract-transform-load (ETL)
- Virtual Integration: mediators
  - Automatically select and compose information across sources
  - Research Topics
    - Modeling language: Relational, XML, Description Logics (OWL 2 & Profiles)
    - Schema mappings: Logical formulas that relate source and global schemas
      - Global-as-View (GAV), Local-as-View (LAV) integration, more general formulas (GLAV)
    - Scalability: efficient query rewriting, query optimization
- Sample Tools
  - Research: BIRN mediator
  - Commercial: IBM InfoSphere Federation Server, Oracle Data Integrator, Talend, Informatica, Expressor, Infobright, …

FBIRN Data Integration



Results

Source	Subject ID	Age	Gender	Scan Type
XNAT	OAS1_0266_MR1	51	M	t1
XNAT	OAS1_0034_MR1	51	M	t1
XNAT	OAS1_0207_MR1	51	M	t1
HID	000998262706	51	M	t1
HID	000998262706	52	M	t1
HID	000998041611	53	M	t1
HID	000913186207	53	M	t1
HID	000960528669	53	M	t1
HID	000913186207	54	M	t1
HID	000947193547	55	M	t1
XNAT	OAS1_0389_MR1	55	M	t1

result
(XML

```
<cell>76</cell>
<cell>active</cell>
</row>
- <row>
<cell>OAS1_0016_MR1</cell>
<cell><CENTRAL_OASIS_CS></cell>
<cell>82</cell>
<cell>active</cell>
</row>
```

000934691111	Handesness_left_writing	0
000934691111	Handesness_right_writing	2
000934691111	Mother's Education	(null)
000934691111	Number of Children	(null)

XNAT

. . .

Human **I**maging **D**atabase(s)
Oracle DB

[Front. NeuroScience 2010]

E**X**tensible **N**euroimaging **A**rchive **T**oolkit
Web service API

Schema Mappings

BIRN mediator uses logical schema mappings

- Source-to-target tuple-generating dependencies (st-tgds), aka global/local as view (GLAV) mappings

$$\forall \vec{X}, \vec{Y}, \Phi_S(\vec{X}, \vec{Y}) \rightarrow \exists \vec{Z}, \Psi_G(\vec{X}, \vec{Z})$$

- Intuitively: query over source predicates (antecedent Φ_S), whose answers populate the domain/target predicates (consequent Ψ_G)

S1_EXPSEGMENT(eID, projectID, subjectID, prtclID, date …) ^
S2_PROTOCOL(prtclID, protocolDescription, …) →

Scan(eID, subjectID, prtclID, date …) ^ hasDisease(subjectID, 'Bipolar')

Subject(subjectID, age, 'Male', …) ^ T1(prtclID) ^ age >= 40

- Mappings used to answer queries or to generate ETL plans

Query Rewriting

- Given Query Q for T1 scans of Male subjects over 50 (over domain schema)

Q(subjectID, age, diseaseID, eID) ←
Subject(subjectID, age, 'Male', …) ^ hasDisease(subjectID, diseaseID) ^
Scan(eID, subjectID, prtclID, date …) ^ T1(prtclID) ^ age > 50

- and Schema Mappings m1 and m2:

m1: S1_EXPSEGMENT(eID, projectID, subjectID, prtclID, date …) ^
S2_PROTOCOL(prtclID, protocolDescription, …) →
Scan(eID, subjectID, prtclID, date …) ^ hasDisease(subjectID, 'Bipolar')
Subject(subjectID, age, 'Male', …) ^ T1(prtclID) ^ age >= 40

m2: S3_SUBJECT(subjectID, age, gender, …) → Subject(subjectID, age, gender, …)

- System rewrites query Q to an executable query Q' (over source schemas)

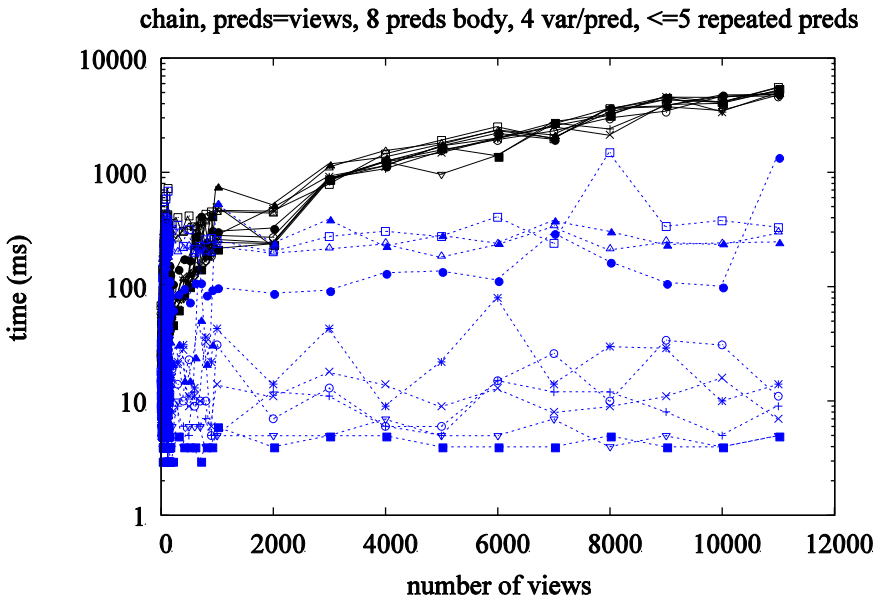
Q'(subjectID, age, 'Bipolar', eID) ←
S1_EXPSEGMENT(eID, projectID, subjectID, prtclID, date …) ^
S2_PROTOCOL(prtclID, protocolDescription, …) ^
S3_SUBJECT(subjectID, age, gender, …) ^ age > 50

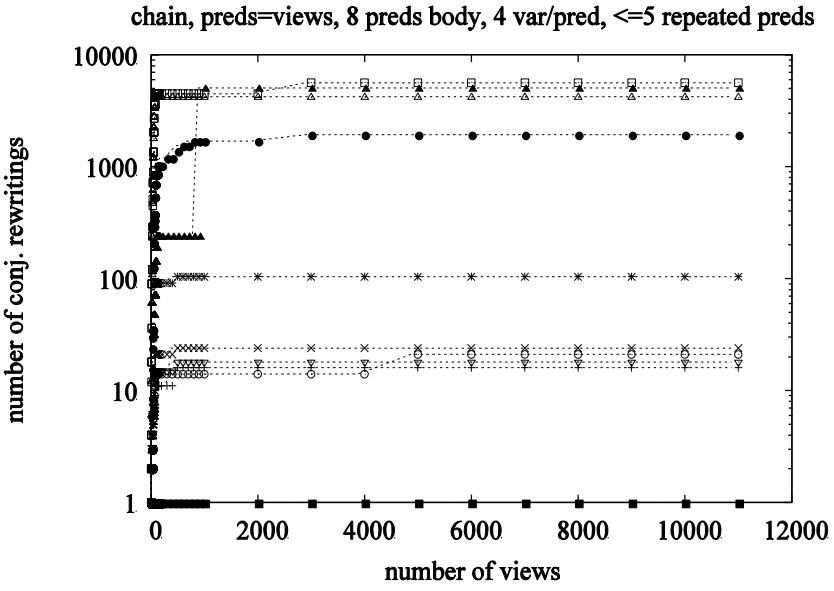
- Q' is translated to evaluation plan or ETL workflow (next slide)

Scalable Data Integration

Query rewriting is NP-hard, but

GQR [Konstantinidis & Ambite, SIGMOD 2011] can rewrite user query over 11,000 source mappings (i.e., 100 tables in 110 DBs) in under 1 second



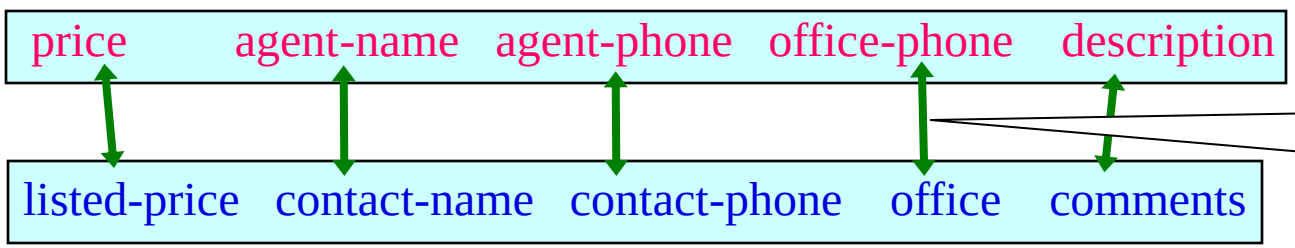


Schema Mapping

- Given two different sources with different schemas, how do we automatically align the information
- Research Topics
  - Automatic schema alignment based on structure and naming
  - Automatic alignment based on the source contents

Schema Mapping

Mediated schema



If "office" occurs in name => office-phone

Schema of realestate.com

realestate.com

listed-price	contact-name	contact-phone	office	comments
$250K	James Smith	(305) 729 0831	(305) 616 1822	Fantastic house
$320K	Mike Doan	(617) 253 1429	(617) 112 2315	Great location
.....	.....	.....	.....	.....

If "fantastic" & "great" occur frequently in data instances => description

homes.com

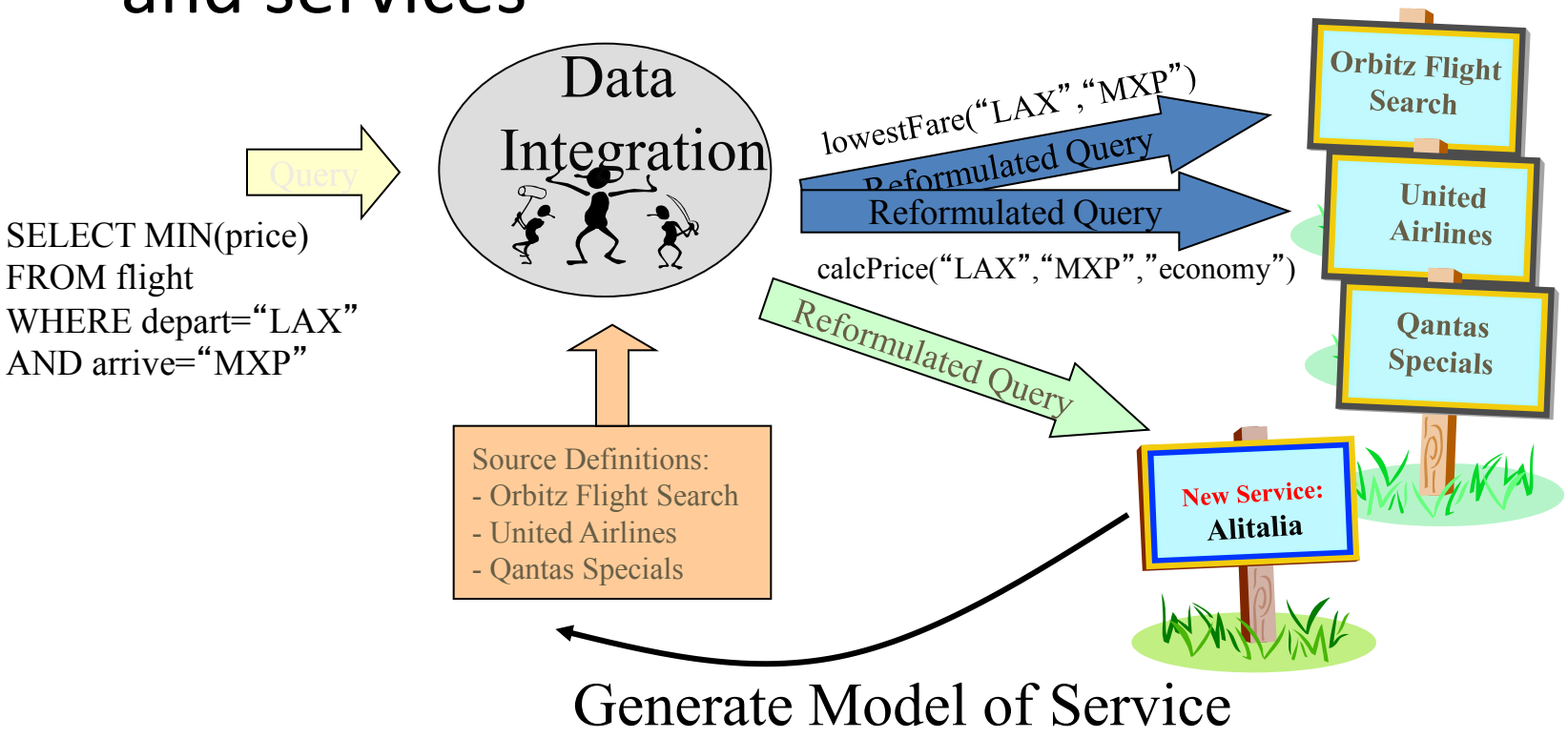
sold-at	contact-agent	extra-info
$350K	(206) 634 9435	Beautiful yard
$230K	(617) 335 4243	Close to Seattle

Source Modeling

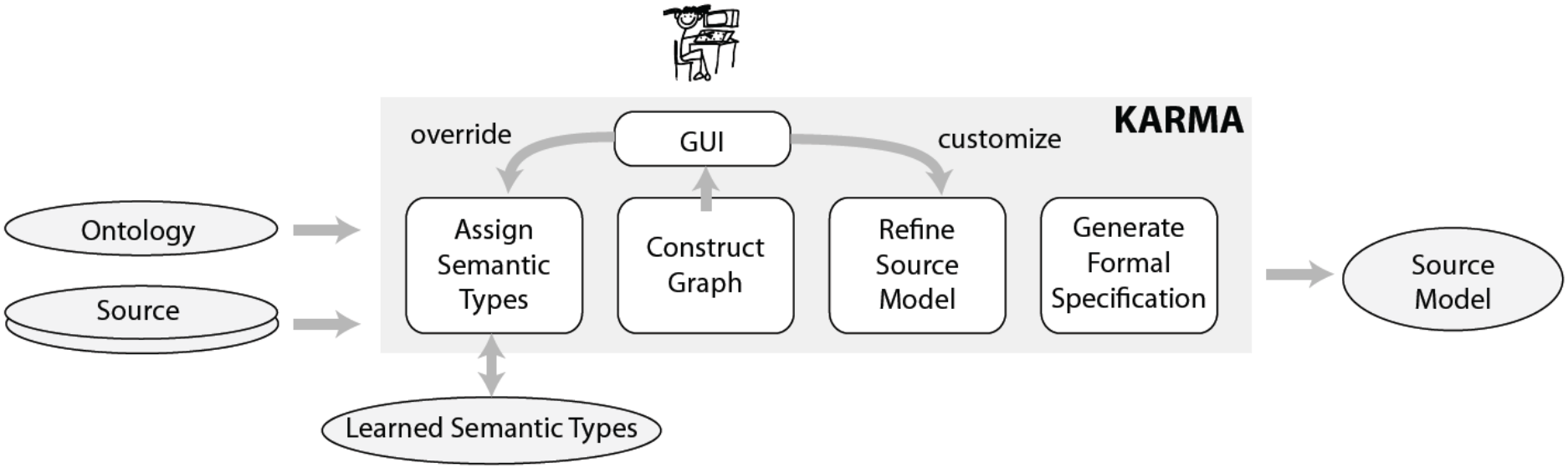
- Semantic typing
- Source discovery
- Automatic source modeling
- Interactive source modeling

Automatic Source Modeling

- How to learn semantic descriptions of sources and services



Semi-Automatic Source Modeling



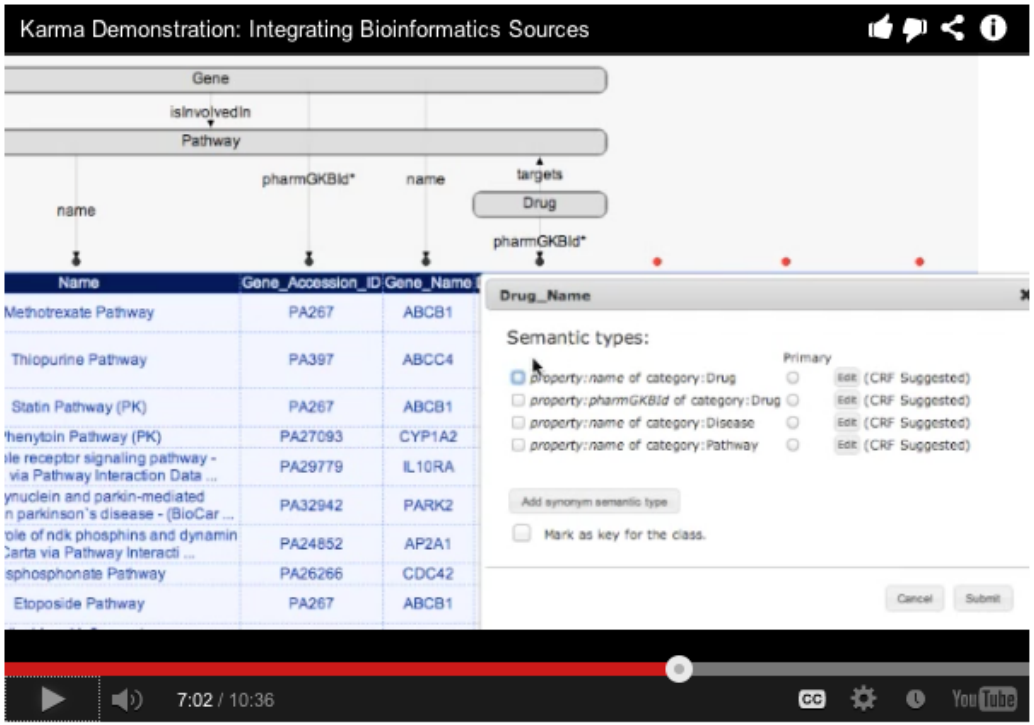
Relational database to RDF mapping

Anything to RDF mapping

Karma

A Data Integration Tool

Karma is available as open source (Apache 2 License) Download »



Principal Investigators

Craig Knoblock

Pedro Szekely

http://www.isi.edu/integration/karma/

String Similarity: Why Strings Don't Match Perfectly?

typos "Joh" vs "John"

OCR errors "J0hn" vs "John"

formatting conventions "03/17" vs "March 17"

abbreviations "J. S. Sargent" vs "John Singer Sargent"

nick names "John" vs "Jock"

word order "Sargent, John S." vs "John S. Sargent"

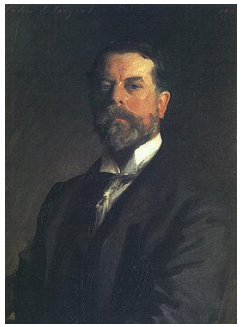
String Similarity Problem Definition

Given X and Y sets of strings

Find pairs (x, y)
such that both x and y
refer to the same real world entity

"John S. Sargent"

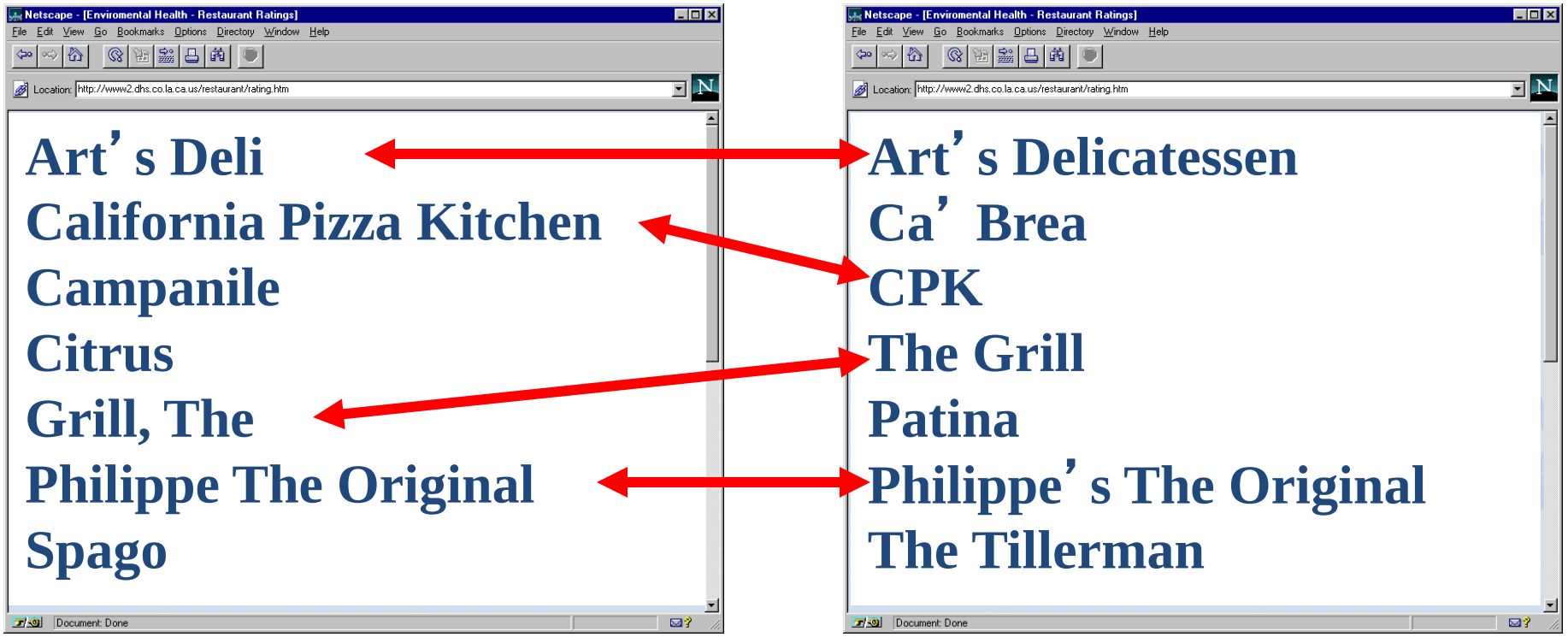
"John Singer Sargent"



Record Linkage

Zagats Restaurant Guide Source

Department of Health Restaurant Source

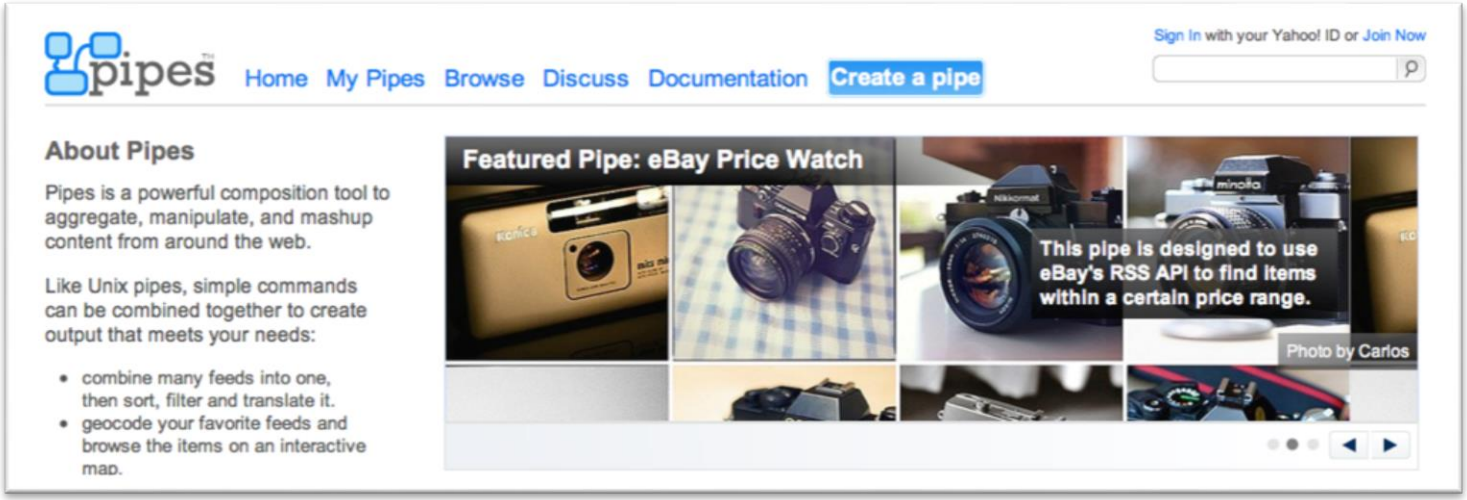


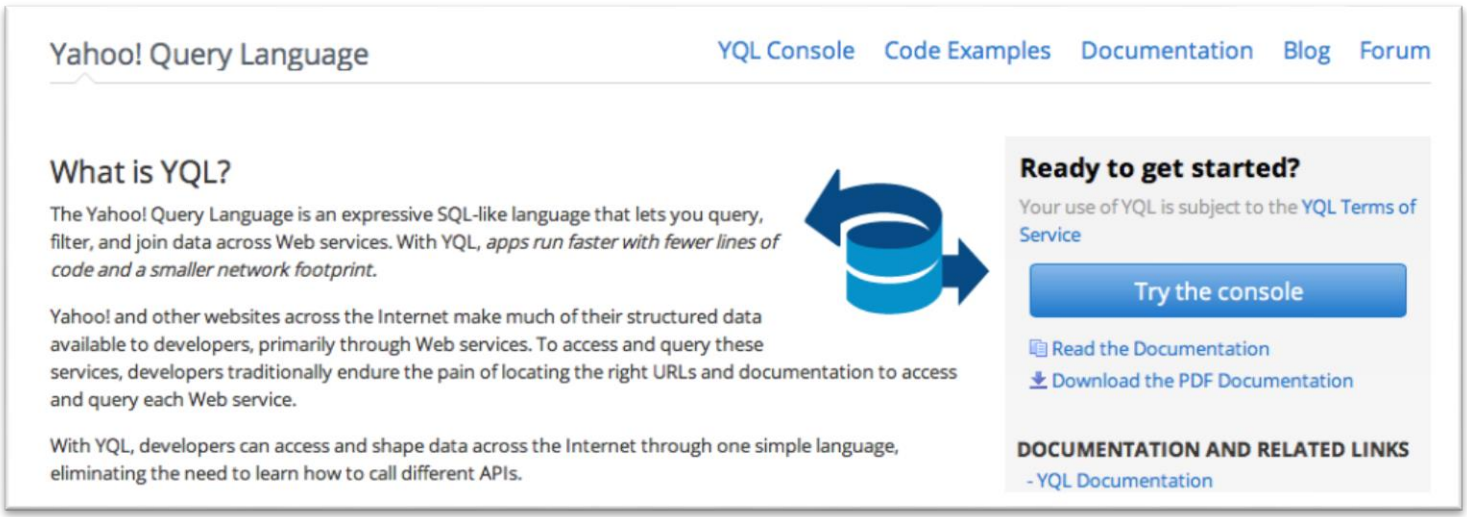
How can the same objects be identified when they are stored in inconsistent text formats?

Record Linkage

- Align entities across sources
- Research Topics:
  - Blocking
  - Matching individual attributes
  - Matching records
  - Matching entities

Mashup Construction





Information Extraction (IE)

"1988 Honda Accrd for sale! Only 80k miles, Runs Like New, V6, 2WD... $2,500 obo. SUPER DEAL."

Information Extraction

- Finding structure in unstructured text
- Research Topics
  - Extraction using NLP techniques
  - Extraction with Conditional Random Fields
  - Exploiting reference sets for extraction

Ontology-based data access and integration

- Use ontology language as domain model
  - OWL2 profiles
- Answering queries under description logic constraints
  - Unions of conjunctive queries
  - Datalog

Ontology-based data access and integration: Example Rewriting

Ontology:

$$\text{Professor} \sqsubseteq \exists \text{teaches},$$
$$\exists \text{teaches}^{-} \sqsubseteq \text{Student}.$$

Clauses:

$$\text{teaches}(x, f(x)) \leftarrow \text{Professor}(x),$$
$$\text{Student}(x) \leftarrow \text{teaches}(y, x).$$

Query: $Q(x) \leftarrow \text{teaches}(x, y) \wedge \text{Student}(y).$

Rewriting:

$$Q(x) \leftarrow \text{teaches}(x, y) \wedge \text{Student}(y).$$
$$Q(x) \leftarrow \text{teaches}(x, y) \wedge \text{teaches}(z, y),$$
$$Q(x) \leftarrow \text{Professor}(x).$$

Main idea: compile the ontology inferences into the query

Wrapper Generation and Learning

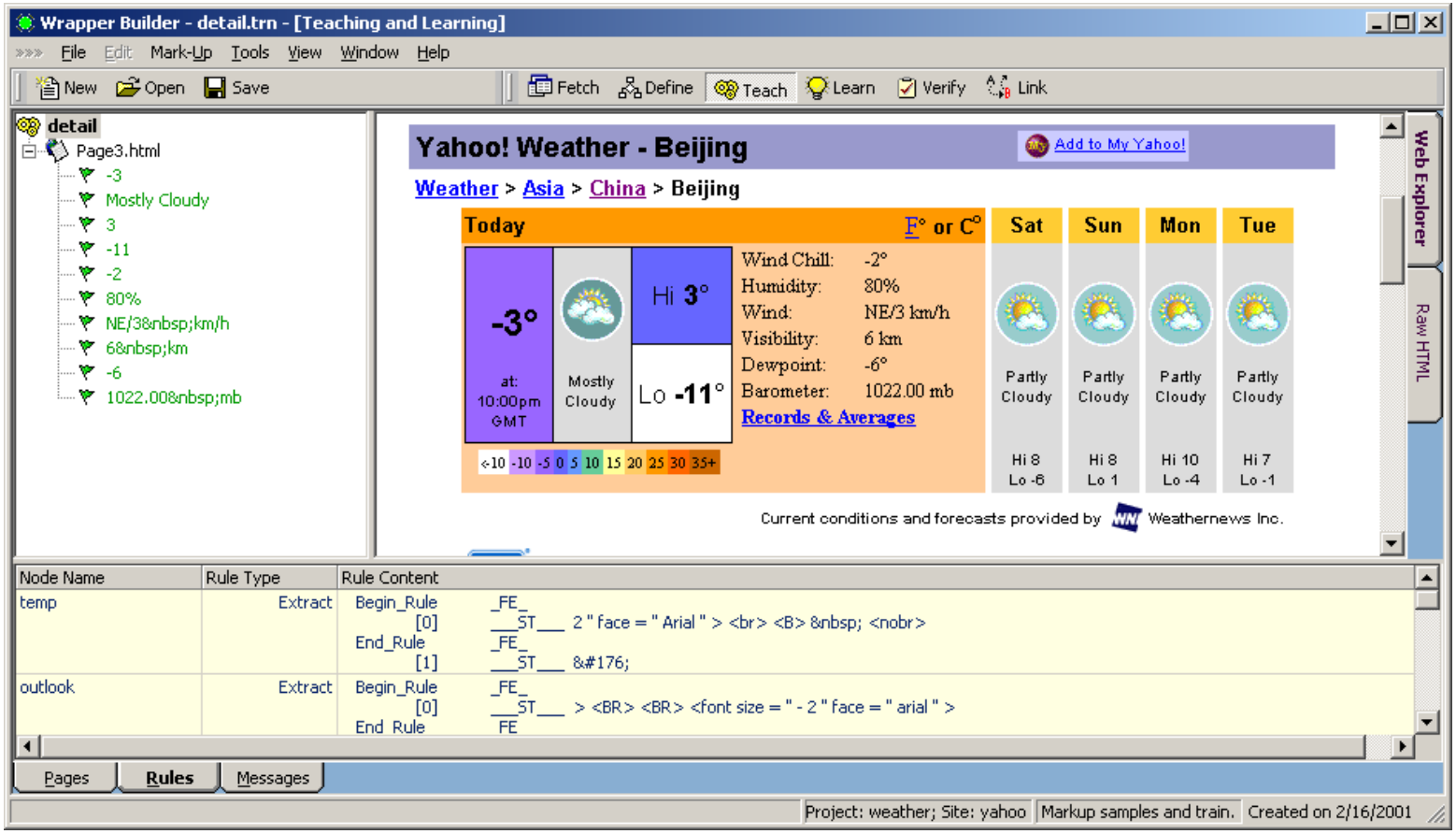
- Turning online sources into structured information
- Research Topics
  - Wrapper Learning
  - Automatic Wrapper Generation
  - Wrapper Maintenance

Wrapper Generation and Learning

NAME	Casablanca Restaurant
STREET	220 Lincoln Boulevard
CITY	Venice
PHONE	(310) 392-5751

Wrapper generation and learning

- Wrapper automatically learned based on a few examples

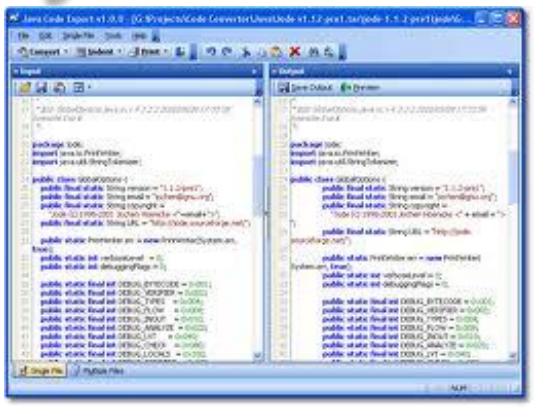


→ **Startup: Fetch Technologies (www.fetch.com)**

[Muslea, Minton & Knoblock, 1999] [Knoblock, Lerman, Minton & Muslea, 2000]

Intellectual Property

Patents



Copyrights

®

Trademarks

Trade Secrets

Intellectual Property

- What types of things can you patent
- When do you copyright something
- When would you file a trademark
- What is a trade secret
- Are we allowed to extract data from online sources
- When you do, who owns the data
- ..and so on…

Course Mechanics

Where to find Professor Ambite…

Research Assistant Professor, Computer Science

Project Leader, Information Sciences Institute

- Office hours:
  - Main campus: Outside ZHS 163 or THH 114 (Immediately after class)
  - Information Sciences Institute, Marina del Rey 310-448-8472 (by appointment)
- Email: ambite@isi.edu

Where to find Professor Knoblock…

- Research Professor
  Computer Science and Spatial Sciences
- Office location:
  - Spatial Sciences AHF B55a
- Office hours:
  - Wed 10-11am in AHF B55a (Spatial Sciences)
  - By appointment at ISI: 310-448-8786
- Email: knoblock@isi.edu

Teaching Assistant

- TA: Bo Wu
  - Office Hours: Tuesday, 10am-12pm
- Location: plaza between RTH café and EEB
- Email: bowu@isi.edu

Course Materials and Web Pages

- Dropbox folder:
  - https://www.dropbox.com/sh/5si94aql8inubb1/AABXH91b2iQHaPhidUo1Z7nXa?dl=0
  - All readings, slides, homeworks, etc will be posted on the shared Dropbox folder
- Google discussion board
  - All questions should be posted
    - Expect response from TA in 24 hours, otherwise let Profs know
    - If you know the answer to a posted question, please try to provide helpful suggestions
    - But please don't post answers to homeworks!
- Blackboard – blackboard.usc.edu
  - Submit homeworks on Blackboard

Prerequisites & Recommended

- Prerequisites
  - CS561 -- Introduction to AI
- Recommended
  - CS585 – Database Systems
  - Some programming experience

Grading (CSCI 548)

- Homework: 50%
  - Must be turned in on time, but
  - You can turn in a homework assignment up to 1 week late with a 20% grade penalty
- Quizzes: 25%
  - One per week, at the beginning of the class
    - Questions based on the lectures, readings, and homeworks from previous week
  - No make ups
  - We will drop the lowest grade
- Final Exam: 25%
  - 3:30pm class: Friday, May 8, 2-4 p.m.
  - You must take the final for your class – no exceptions!

More on Grading

- This is a hard class, but you will learn a lot!
  - Principles and theory
    - Technical readings and lectures (quizzes, final exam)
  - Putting principles into practice
    - Homeworks
  - We do give B's, C's, and even D's and F's
- Grade distribution
  - Roughly half A's and B's (consider a C a failing grade)

94 - 100 = A	74 - 76 = C
90 – 93 = A-	70 - 73 = C-
87 – 89 = B+	67 - 69 = D+
84 – 86 = B	64 - 66 = D
83 – 83 = B-	60 - 63 = D-
77 – 79 =C+	Below 60 is an F

Readings

- Posted on Dropbox each week
  - You can read it online or print them
- Read all required readings before class
  - You will get more out of the lectures
- Optional readings go in more depth
  - Read them to learn more

Slides

- Available online before the lecture
- Not intended as a replacement for the lecture
- You can bring them to class
- Final version of the slides available after class

Working Together

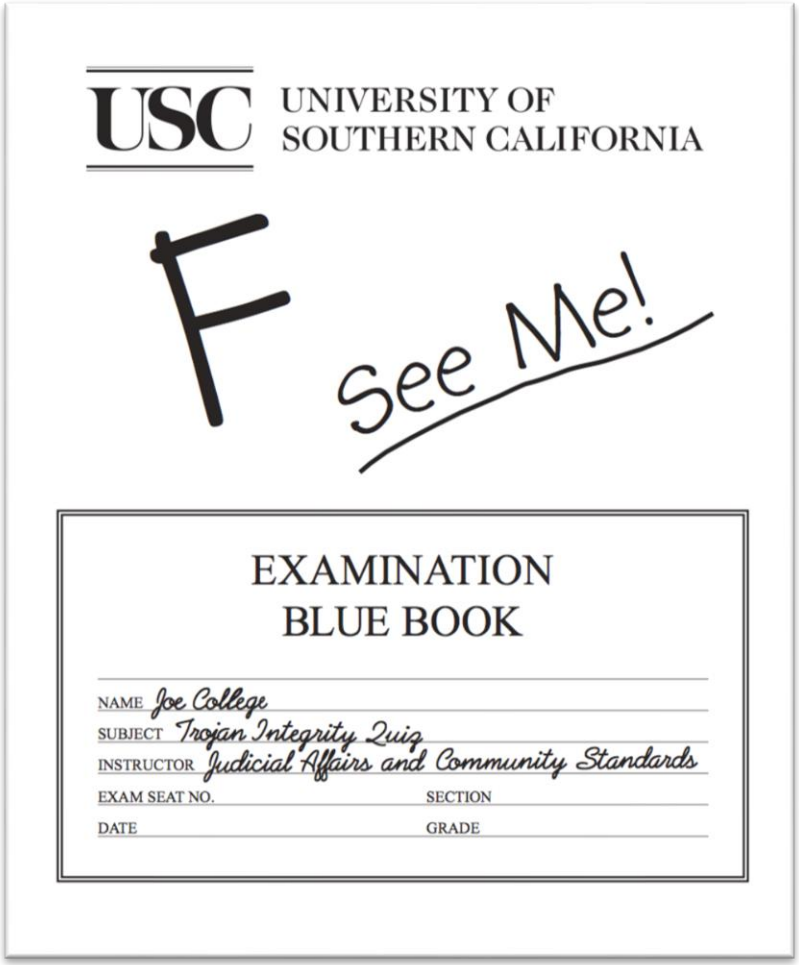
- Each person must do their own homework
  - We will check for overlap in homeworks
  - If we find any plagiarism, all parties lose credit so
    - Don’t share your answers
    - Don’t leave printouts in the trash with your answers
    - Don’t give out your password
    - Don’t copy others (they may have the wrong answer anyway!)
- You can ask the professors or TAs for help

Cheating

- Not tolerated!
- All infractions will be reported

- Examples:
  - Turning in someone else's homework
  - Doing the homework in collaboration with someone else and then turning in your own copy
  - Copying from someone else during a quiz or exam

We Follow USC Policies

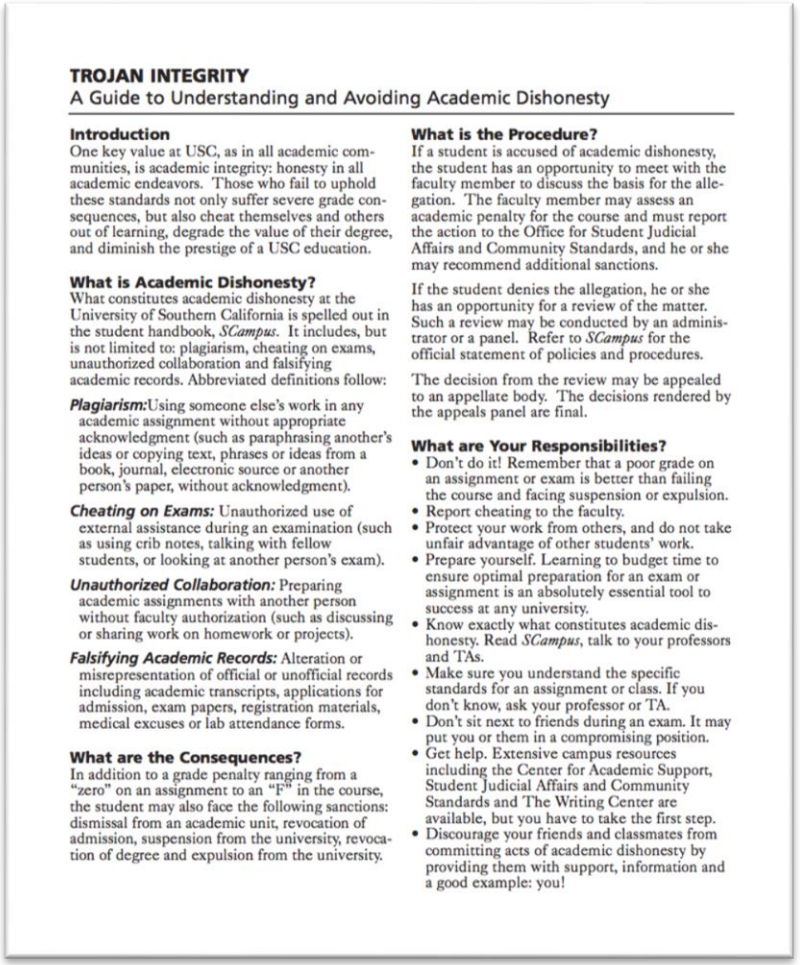


USC UNIVERSITY OF SOUTHERN CALIFORNIA

F See Me!

EXAMINATION
BLUE BOOK

NAME Joe College
SUBJECT Trojan Integrity Quiz
INSTRUCTOR Judicial Affairs and Community Standards
EXAM SEAT NO. SECTION
DATE GRADE



TROJAN INTEGRITY
A Guide to Understanding and Avoiding Academic Dishonesty

Introduction
One key value at USC, as in all academic communities, is academic integrity: honesty in all academic endeavors. Those who fail to uphold these standards not only suffer severe grade consequences, but also cheat themselves and others out of learning, degrade the value of their degree, and diminish the prestige of a USC education.

What is Academic Dishonesty?
What constitutes academic dishonesty at the University of Southern California is spelled out in the student handbook, *SCampus*. It includes, but is not limited to: plagiarism, cheating on exams, unauthorized collaboration and falsifying academic records. Abbreviated definitions follow:

Plagiarism: Using someone else's work in any academic assignment without appropriate acknowledgment (such as paraphrasing another's ideas or copying text, phrases or ideas from a book, journal, electronic source or another person's paper, without acknowledgment).

Cheating on Exams: Unauthorized use of external assistance during an examination (such as using crib notes, talking with fellow students, or looking at another person's exam).

Unauthorized Collaboration: Preparing academic assignments with another person without faculty authorization (such as discussing or sharing work on homework or projects).

Falsifying Academic Records: Alteration or misrepresentation of official or unofficial records including academic transcripts, applications for admission, exam papers, registration materials, medical excuses or lab attendance forms.

What are the Consequences?
In addition to a grade penalty ranging from a "zero" on an assignment to an "F" in the course, the student may also face the following sanctions: dismissal from an academic unit, revocation of admission, suspension from the university, revocation of degree and expulsion from the university.

What is the Procedure?
If a student is accused of academic dishonesty, the student has an opportunity to meet with the faculty member to discuss the basis for the allegation. The faculty member may assess an academic penalty for the course and must report the action to the Office for Student Judicial Affairs and Community Standards, and he or she may recommend additional sanctions.

If the student denies the allegation, he or she has an opportunity for a review of the matter. Such a review may be conducted by an administrator or a panel. Refer to *SCampus* for the official statement of policies and procedures.

The decision from the review may be appealed to an appellate body. The decisions rendered by the appeals panel are final.

What are Your Responsibilities?
- Don't do it! Remember that a poor grade on an assignment or exam is better than failing the course and facing suspension or expulsion.
- Report cheating to the faculty.
- Protect your work from others, and do not take unfair advantage of other students' work.
- Prepare yourself. Learning to budget time to ensure optimal preparation for an exam or assignment is an absolutely essential tool to success at any university.
- Know exactly what constitutes academic dishonesty. Read *SCampus*, talk to your professors and TAs.
- Make sure you understand the specific standards for an assignment or class. If you don't know, ask your professor or TA.
- Don't sit next to friends during an exam. It may put you or them in a compromising position.
- Get help. Extensive campus resources including the Center for Academic Support, Student Judicial Affairs and Community Standards and The Writing Center are available, but you have to take the first step.
- Discourage your friends and classmates from committing acts of academic dishonesty by providing them with support, information and a good example: you!

Print it (from Blackboard), read it, sign it and return it to us

Waiting List

- Even if you are in the waiting list, come to the classes and do the quizzes/homeworks
- Likely there will be additional openings in the next couple of weeks

When the Course is Over

- Directed research (MS or Phd Students)
- M.S. Thesis
- Summer interns (MS or Phd)
- Research Assistantships (1-2 Phd Students)
  - We can also recommend you for positions in other groups
- Graders/Course Producers (MS)
- Recommendation letters (anyone that gets at least an A-)
- Positions at related companies
  - Companies are often looking for recommendations of students

Questions?