TAC 2012
Cold Start

Knowledge Base Population

Many slides borrowed
from Jim Mayfield and
Paul McNamee, JHU

When Lisa's mother Marge Simpson went to a weekend getaway at Rancho Relaxo, the movie The Happy Little Elves Meet Fuzzy Snuggleduck was one of the R-rated european adult movies available on their cable channels.

After two years in the academic quagmire of Springfield Elementary, Lisa finally has a teacher that she connects with. But she soon learns that the problem with being middle-class is that

When Lisa's mother Marge Simpson went to a weekend getaway at Rancho Relaxo, the movie The Happy Little Elves Meet Fuzzy Snuggleduck was one of the R-rated european adult movies available on their cable channels.

Springfield

After two years in the academic quagmire of Springfield Elementary, Lisa finally has a teacher that she connects with. But she soon learns that the problem with being middle-class is that

Springfield Elementary

Marge Simpson

Homer Simpson

Lisa Simpson

Bart Simpson

When Lisa's mother Marge Simpson went to a weekend getaway at Rancho Relaxo, the movie The Happy Little Elves Meet Fuzzy Snuggleduck was one of the R-rated european adult movies available on their cable channels.

After two years in the academic quagmire of Springfield Elementary, Lisa finally has a teacher that she connects with. But she soon learns that the problem with being middle-class is that

Bottomless Pete, Nature's Cruelest Mistake

Marge Simpson

per:cities_of_residence

per:children

per:children

per:spouse

Homer Simpson

per:alternate_names

Springfield Elementary

per:schools_attended

Lisa Simpson

Bart Simpson

Cold Start

Schema

Entity-Valued Relations

Relation	Inverse(s)
per:children	per:parents
per:other_family	per:other_family
per:parents	per:children
per:siblings	per:siblings
per:spouse	per:spouse
per:employee_of	{org,gpe}:employees*
per:member_of	org:membership*
per:schools_attended	org:students*
per:city_of_birth	gpe:births_in_city*
per:stateorprovince_of_birth	gpe:births_in_stateorprovince*
per:country_of_birth	gpe:births_in_country*
per:cities_of_residence	gpe:residents_of_city*
per:statesorprovinces_of_residence	gpe:residents_of_stateorprovince
per:countries_of_residence	gpe:residents_of_country*
per:city_of_death	gpe:deaths_in_city*
per:stateorprovince_of_death	gpe:deaths_in_stateorprovince*
per:country_of_death	gpe:deaths_in_country*
org:shareholders	{per,org,gpe}:holds_shares_in*
org:founded_by	{per,org,gpe}:organizations_founded*
org:top_members_employees	per:top_member_employee_of*
{org,gpe}:member_of	org:members
org:members	{org,gpe}:member_of
org:parents	{org,gpe}:subsidiaries
org:subsidiaries	org:parents
org:city_of_headquarters	gpe:headquarters_in_city*
org:stateorprovince_of_headquarters	gpe:headquarters_in_stateorprovince*
org:country_of_headquarters	gpe:headquarters_in_country*

String-Filled Relations

per:alternate_names

per:date_of_birth

per:age

per:origin

per:date_of_death

per:cause_of_death

per:title

per:religion

per:charges

org:alternate_names

org:political_religious_affiliation

org:number_of_employees_members

org:date_founded

org:date_dissolved

org:website

Cold Start

Schema

per:children
per:other_family
per:parents
per:siblings
per:spouse
per:employee_of
per:member_of
per:schools_attended
per:city_of_birth
per:stateorprovince_of_birth
per:country_of_birth
per:cities_of_residence
per:statesorprovinces_of_reside
nce
per:countries_of_residence
per:city_of_death
per:stateorprovince_of_death
per:country_of_death
org:shareholders
org:founded_by

The Task

You are given:

Schema

per:children
per:other_family
per:parents
per:siblings
per:spouse
per:employee_of
per:member_of
per:schools_attended
per:city_of_birth
per:stateorprovince_of_birth
per:country_of_birth
per:cities_of_residence
per:statesorprovinces_of_residence
per:countries_of_residence
per:city_of_death
per:stateorprovince_of_death
per:country_of_death

When Lisa's mother Marge Simpson went to a weekend getaway at Rancho Relaxo, the movie The Happy Little Elves Meet Fuzzy Snuggleduck was one of the R-rated european adult movies available on their cable channels.

But wait, there's more!!

The Task

You are given:

Schema

per:children
per:other_family
per:parents
per:siblings
per:spouse
per:employee_of
per:member_of
per:schools_attended
per:city_of_birth
per:stateorprovince_of_birth
per:country_of_birth
per:cities_of_residence
per:statesorprovinces_of_residence
per:countries_of_residence
per:city_of_death
per:stateorprovince_of_death
per:country_of_death

Lisa
Marge Simpson
Rancho Relaxo
Fuzzy Snuggleduck

When Lisa's mother Marge Simpson went to a weekend getaway at Rancho Relaxo, the movie The Happy Little Elves Meet Fuzzy Snuggleduck was one of the R-rated european adult movies available on their cable channels.

Schema

per:children
per:other_family
per:parents
per:siblings
per:spouse
per:employee_of
per:member_of
per:schools_attended
per:city_of_birth
per:stateorprovince_of_birth
per:country_of_birth
per:cities_of_residence
per:statesorprovinces_of_residence
per:countries_of_residence
per:city_of_death
per:stateorprovince_of_death
per:country_of_death

You Must Produce:

When Lisa's mother Marge Simpson went to a weekend getaway at Rancho Relaxo, the movie The Happy Little Elves Meet Fuzzy Snuggleduck was one of the R-rated european adult movies available on their cable channels.

Springfield

Bottomless Pete, Nature's Cruelest Mistake

per:cities_of_residence

per:alternate_names

Marge Simpson

per:spouse

Homer Simpson

per:children

per:children

Springfield Elementary

per:schools_attended

Lisa Simpson

Bart Simpson

How do you know that your KB is any good?

Marge Simpson

per:cities_of_residence

per:children

per:children

per:spouse

Homer Simpson

per:alternate_names

Bottomless Pete, Nature's Cruelest Mistake

Springfield Elementary

per:schools_attended

Lisa Simpson

Bart Simpson

How do you know that your KB is any good?

Align it to a ground truth KB

Springfield

Bottomless Pete, Nature's Cruelest Mistake

per:alternate_names

per:cities_of_residence

Marge Simpson

per:spouse

Homer Simpson

per:children

per:children

Springfield Elementary

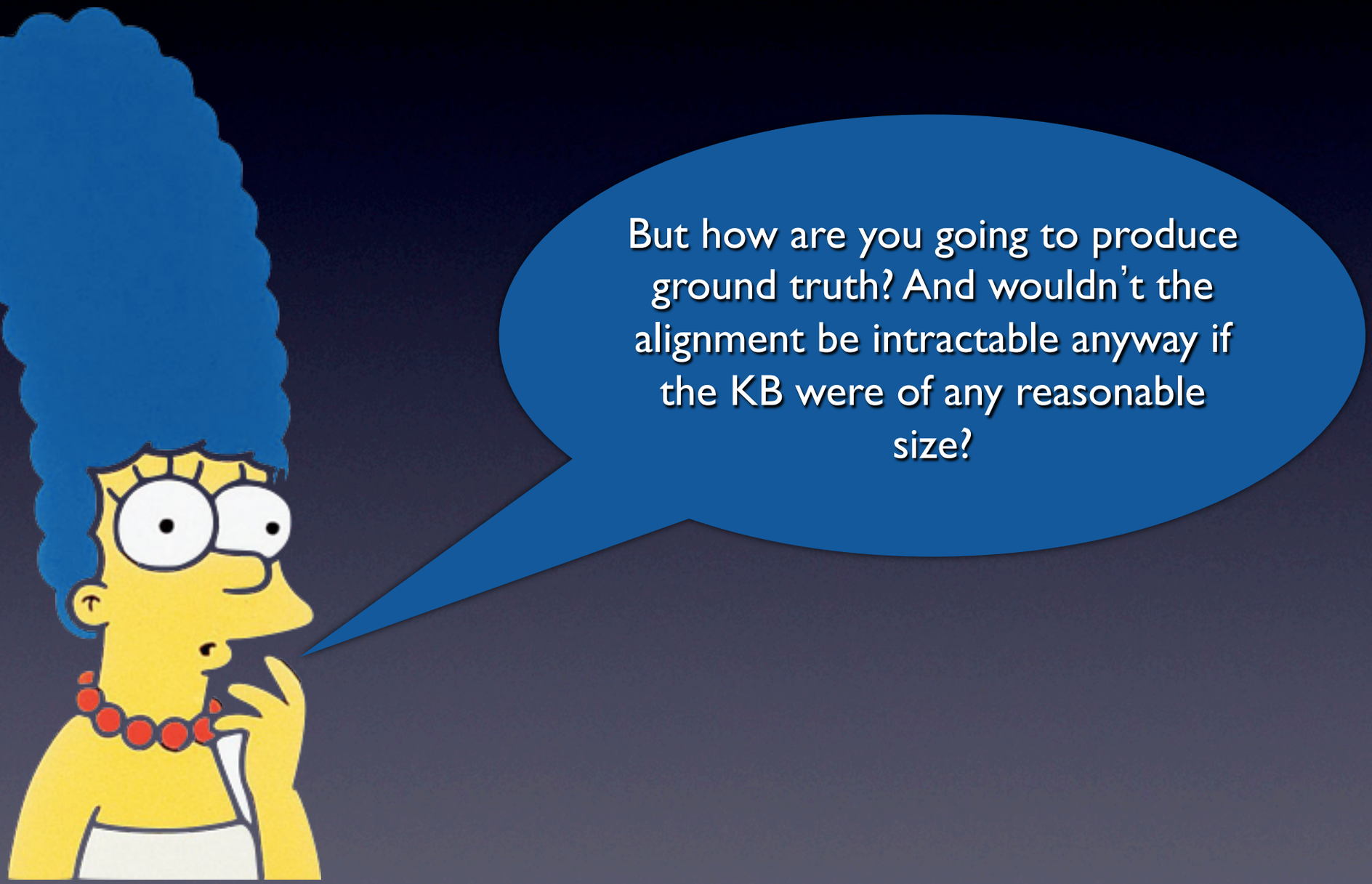
per:schools_attended

Lisa Simpson

Bart Simpson

How do you know that your KB is any good?

Align it to a ground truth KB



Where did the children of Marge Simpson go to school?

per:children
per:schools_attended

Bottomless Pete, Nature's Cruelest Mistake

per:alternate_names

per:cities_of_residence

Marge Simpson

per:spouse

Homer Simpson

per:children

per:children

Springfield Elementary

per:schools_attended

Lisa Simpson

Bart Simpson

When Lisa's mother Marge Simpson went to a weekend getaway at Rancho Relaxo, the movie The Happy Little Elves Meet Fuzzy Snuggleduck was one of the R-rated european adult movies available on their cable channels.

After two years in the academic quagmire of Springfield Elementary, Lisa finally has a teacher that she connects with. But she soon learns that the problem with being middle-class is that

Bottomless Pete, Nature's Cruelest Mistake

Springfield

per:cities_of_residence

Marge Simpson

per:children

per:children

per:spouse

Homer Simpson

per:alternate_names

Springfield Elementary

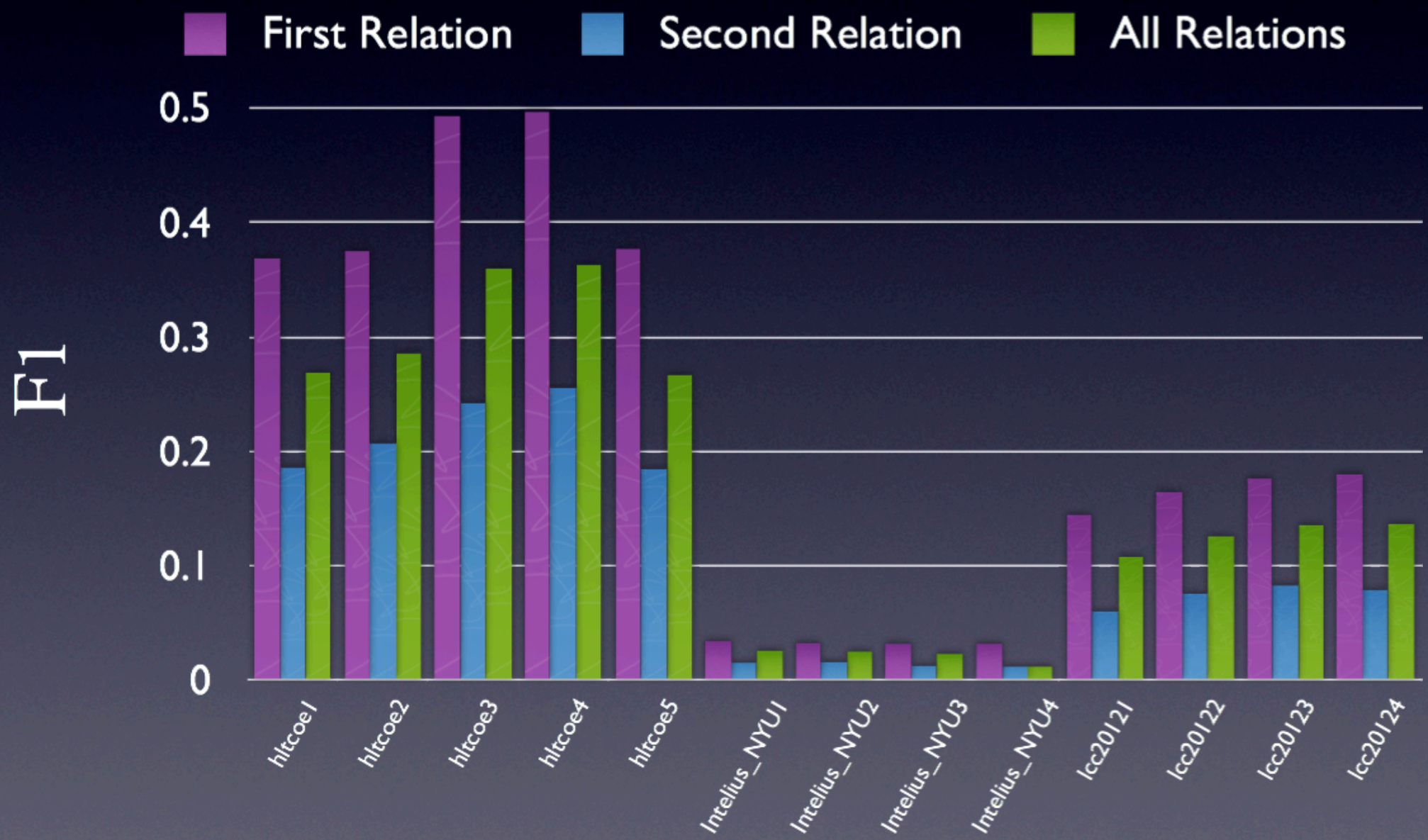
per:schools_attended

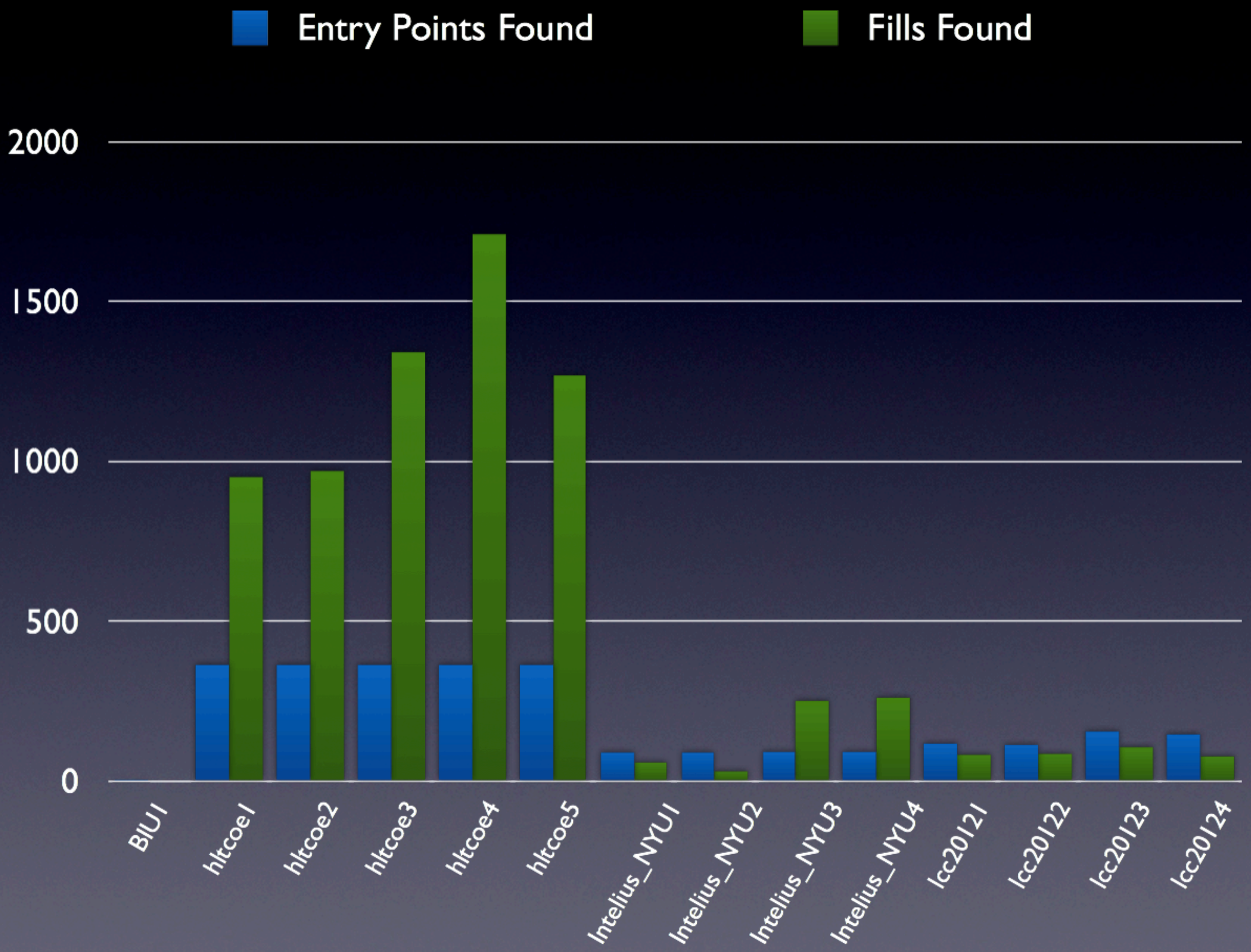
Lisa Simpson

Bart Simpson

Sample Evaluation Queries

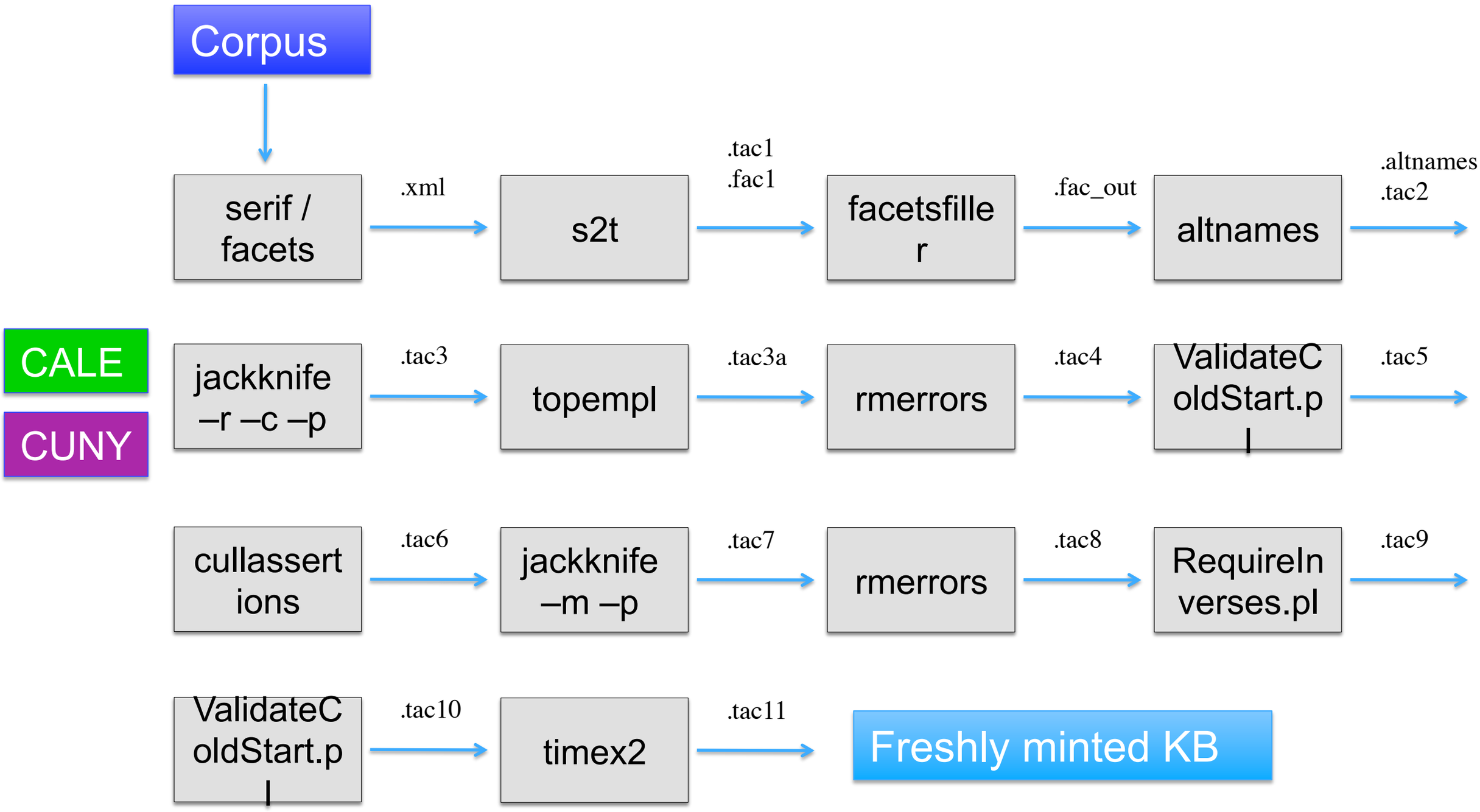
Query Entity	First Relation	Second Relation
Adriana Petryna	per:title	
Blackstone Group	org:founded_by	
William Shore	per:organizations_founded	org:date_founded
Wistar Institute	org:employees	per:title
Andrew W. Mellon	per:children	per:organizations_founded
Lycee Alliance Israelite Universelle	org:employees	per:schools_attended
Tsitsi Jaji	per:schools_attended	org:students


First Relation
Second Relation
All Relations
F1
0.5
0.4
0.3
0.2
0.1
0
hltcoe1
hltcoe2
hltcoe3
hltcoe4
hltcoe5
Intelius_NYU1
Intelius_NYU2
Intelius_NYU3
Intelius_NYU4
lcc20121
lcc20122
lcc20123
lcc20124


Entry Points Found
Fills Found
2000
1500
1000
500
0
BIUI
hltcoe1
hltcoe2
hltcoe3
hltcoe4
hltcoe5
Intelius_NYU1
Intelius_NYU2
Intelius_NYU3
Intelius_NYU4
lcc20121
lcc20122
lcc20123
lcc20124

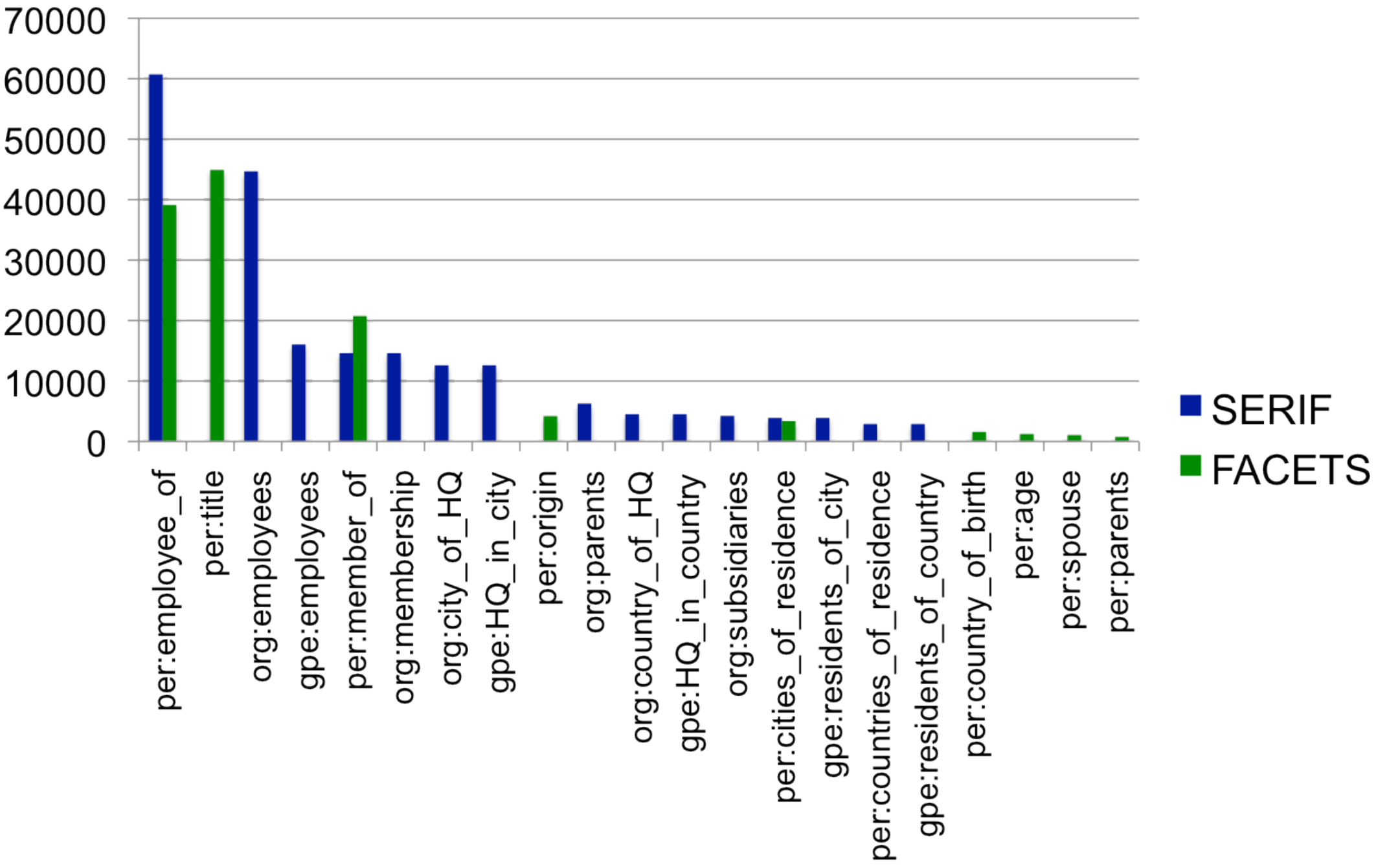
human language technology
center of excellence

Pipeline



human language technology
center of excellence

Practice: Washington Post articles



26k 2010 Washington Post articles (194k assertions)

human language technology
center of excellence

Demo: Query Engine

- **Since no training data was available, we assessed our practice KBs qualitatively**

26k 2010 Washington Post articles (194k assertions)

human language technology
center of excellence

Good KELVIN

- Jared Fogle is an employee of Subway
- Freeman Hrabowski works for UMBC, founded the Meyerhoff Scholors Program, and attended both Hampton University and the University of Illinois
- Elena Kagan attended Oxford, Harvard, and Princeton
- The Applied Physics Laboratory is a subsidiary of Johns Hopkins University
- Southwest Airlines is headquartered in Texas

26k 2010 Washington Post articles (194k assertions)

human language technology
center of excellence

Bad KELVIN

- **Harry Reid is an employee of the Republican Party. (KELVIN also learns he is employee of the Democratic Party)**
  - **Tend to see more trouble with popular entities**
- **Big Foot is an employee of Starbucks**
- **Steven Spielberg lives in Iran**
- **Jill Biden is married to Jill Biden**

26k 2010 Washington Post articles (194k assertions)

human language technology
center of excellence

1 year of WP = 1 computer scientist

- **KELVIN learns:**
  - **("ian soboroff") per:employee_of ("national institute of standards and technology")**
  - **("ian soboroff") per:title "computer scientist"**

- **paulmac@hpcc1: more WPB_ENG_20100506.0012.sgm**
- **One scholar Anderson has been talking with is Dan Cohen, the director of the Center for History and New Media at George Mason University; another is Ian Soboroff, a computer scientist with the National Institute of Standards and Technology, who studies how well technology works with human language. Soboroff hopes that in the future, Twitter can be used to help researchers understand how information was dispersed and what human networks in the early 21st century looked like.**

human language technology
center of excellence

Experimental Results

			0-hop slots			1-hop slots		
	CUNY	CALE	P	R	F1	P	R	F1
hltcoe1			0.5256	0.2843	0.3690	0.1620	0.2182	0.1859
hltcoe2	YES		0.4929	0.3031	0.3753	0.1818	0.2406	0.2071
hltcoe3		YES	0.4865	0.5000	0.4932	0.1849	0.3510	0.2423
hltcoe4	YES	YES	0.4799	0.5155	0.4971	0.1842	0.4168	0.2555
hltcoe5	YES		0.4937	0.3053	0.3773	0.1453	0.2531	0.1846

CALE (better entity linking) materially improves recall

Precision (unsurprisingly) is lower for 0-hop vs. 1-hop

Jena RDF database

Text search interface

Fuseki RDF triple store query service

Pubby RDF linked data

SPARQLer Update Validator

eb4.cs.umbc.edu:3031/update-v...

SPARQLer Update Validator

```
PREFIX xsd:   <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf:   <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:  <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl:   <ht
PREFIX fn:    <htt
PREFIX apf:   <ht
```

Input syntax:

Line numbers:

Validate SPARQL Upda

Fuseki Query

Dataset: /ds

SPARQL Query

Output: JSON

XSLT style sheet (

Force the accep

Get Results

SPARQL Update

Perform update

File upload

File: Choose File No

aph: default

entity...

KELVINpedia

kelvin-pedia.appspot.com

KELVINpedia

...the TAC 2012 ColdStart System

A B C D E F G

Hello there! w

White House | Tackbp

eb4.cs.umbc.edu:8899/page/res...

kbp:org-alternate_names	"House" "White House"
kbp:org-employees	ent:Alan_Simpson ent:Clinton ent:George_W_Bush
kbp:org-members	
kbp:org-membership	
kbp:org-subsidiaries	
kbp:org-top_members_empl	
rdf:type	

ws information

RDF/XML | Brov

White House | Tackbp

eb4.cs.umbc.edu:8899/page/res...

White House at Tackbp

http://eb4.cs.umbc.edu:8899/resource/White_House

Property	Value
kbp:canonical_mentionString	"White House"
kbp:entity_id	:e_WPB_ENG_20101202_0081_20
rdfs:label	White House
kbp:mentionString	"House" "White House" "White house"
kbp:mentioned_in	document-offset:WPB_ENG_20101201.0018 document-offset:WPB_ENG_20101201.0029 document-offset:WPB_ENG_20101201.0034 document-offset:WPB_ENG_20101201.0042 document-offset:WPB_ENG_20101201.0043 document-offset:WPB_ENG_20101201.0051 document-offset:WPB_ENG_20101201.0052 document-offset:WPB_ENG_20101201.0069 document-offset:WPB_ENG_20101201.0086 document-offset:WPB_ENG_20101201.0089 document-offset:WPB_ENG_20101201.0090 document-offset:WPB_ENG_20101202.0006 document-offset:WPB_ENG_20101202.0009 document-offset:WPB_ENG_20101202.0010 document-offset:WPB_ENG_20101202.0015 document-offset:WPB_ENG_20101202.0025 document-offset:WPB_ENG_20101202.0033 document-offset:WPB_ENG_20101202.0042 document-offset:WPB_ENG_20101202.0054 document-offset:WPB_ENG_20101202.0075 document-offset:WPB_ENG_20101202.0077 document-offset:WPB_ENG_20101202.0078

human language technology
center of excellence

Some reasoning usecases

- **Deriving new relations and properties**

  husband_of(X,Y) => male(X)

- **Default reasoning**

  religion(X,R) ^ member(X,C) ^ church(C) ^
  unknown(affiliation(C,?)) => affiliation(C,R)

- **Detecting inconsistencies**

  birthDate(X,D1) ^ deathDate(X,D2) ^ before(D2,D1)

- **Detecting unusual or anomalous data**

  A person with more than 10 children

- **Providing evidence for/against entity merging**