

IMPROVING REQUIREMENTS TRACING VIA INFORMATION RETRIEVAL

Jane Huffman Hayes, Alex Dekhtyar, James Osborne



Thanks to Jesse Yannelli for artwork, Jane Huang for some slides



OUTLINE

- ❖ Summary/Contributions of RE 2003 paper
- ❖ Influence
 - Requirements tracing
 - Requirements engineering
 - Model-based software engineering
 - Miscellany
- ❖ Closing thoughts



The Year is 2003

FUN FACTS

Middle of post-Y2K technology slump

Is software industry in the US doomed because of
imminent outsourcing?

Google is 5 years old, already a verb

Four more years 'til iPhone



FUN FACTS

... and a NASA IV&V facility in Fairmont, WV
wants researcher help with Software Assurance...

(and gives \$\$\$!)



5

CAN SOFTWARE ENGINEERING RESEARCHERS HELP LAUNCH SPACESHIPS?

(i.e.: what is the most annoying thing about IV&V?)





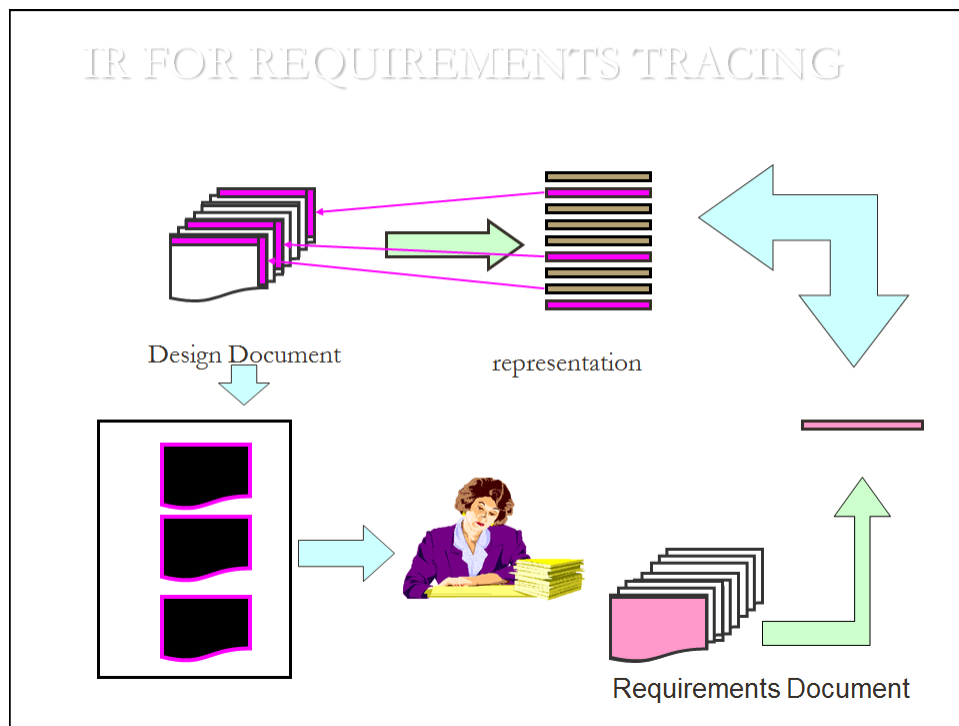
Why, REQUIREMENTS
TRACING, of course!

<insert voices of hundreds of IV&V analysts here>

SO, WE NOTICED TWO THINGS...

Best way to help IV&V crowd:
automate tracing!

Tracing requirements is just like Google



N - number of low-level requirements;
M - number of high-level requirements;

Hits - number of correct candidate links
Strikes - number of false positives
Misses - number of missed links

Precision = $\frac{\text{Hits}}{\text{Hits} + \text{Strikes}}$ **Recall** = $\frac{\text{Hits}}{\text{Hits} + \text{Misses}}$

MODIS Dataset

- 10 x 10
- 19 x 50

WHAT WE MEASURED IN 2003

OUR RESULTS SLIDE FROM 2003

• Human analyst vs. Thesaurus retrieval

	SuperTracePlus Tool	Analyst[1]	Retrieval with thesaurus algorithm
Correct links	41	41	41
Correct links found	26	18	35
Total number of answers	67	39	86
Missed requirements[2]	3	6	4
Average recall[3]	69.37%	53.30%	71.69%
Average precision[3]	56.48%	53.55%	32.76%
Overall recall[4]	63.41%	43.90%	85.36%
Overall precision[4]	38.80%	46.15%	40.69%
Performance (hours)	N/A – included in analyst performance	9	Seconds for algorithm, 0.33 for thesaurus building

[1] Analyst was working with the results of SuperTracePlus analysis

[2] "Missed requirements" refers to high-level requirements with existing links for which the method failed to return a correct match

[3] Average recall and precision were computed by first computing the recall and precision for each individual requirement and then taking the average over all requirements

[4] Overall recall and precision were computed by dividing the total number of retrieved links over the number of existing links (recall) and the total number of retrieved links over the total number of candidate links (precision)

WHAT WE LEARNED (2003)

- “vanilla” IR : keyword matches
- Languages of Requirements and Design documents are different
 - Leads to low recall
- Many keyword matches are circumstantial
 - Leads to low precision
- Analysts: high recall but low precision
(on small datasets)
- Thesaurus : recall improves significantly
precision improves some

WHAT WE LEARNED (2013)

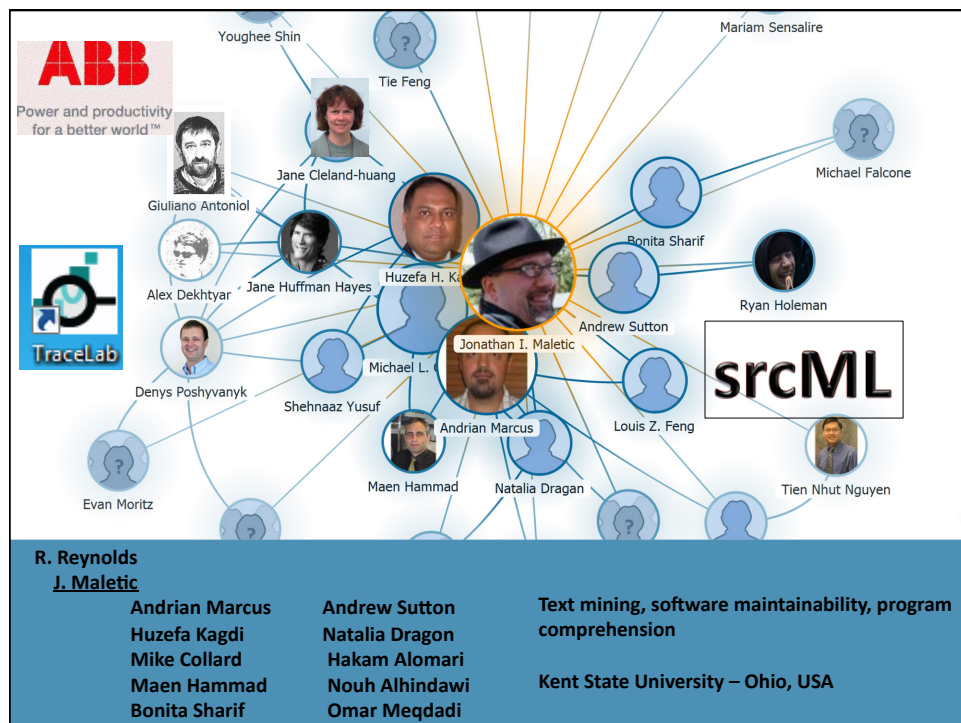
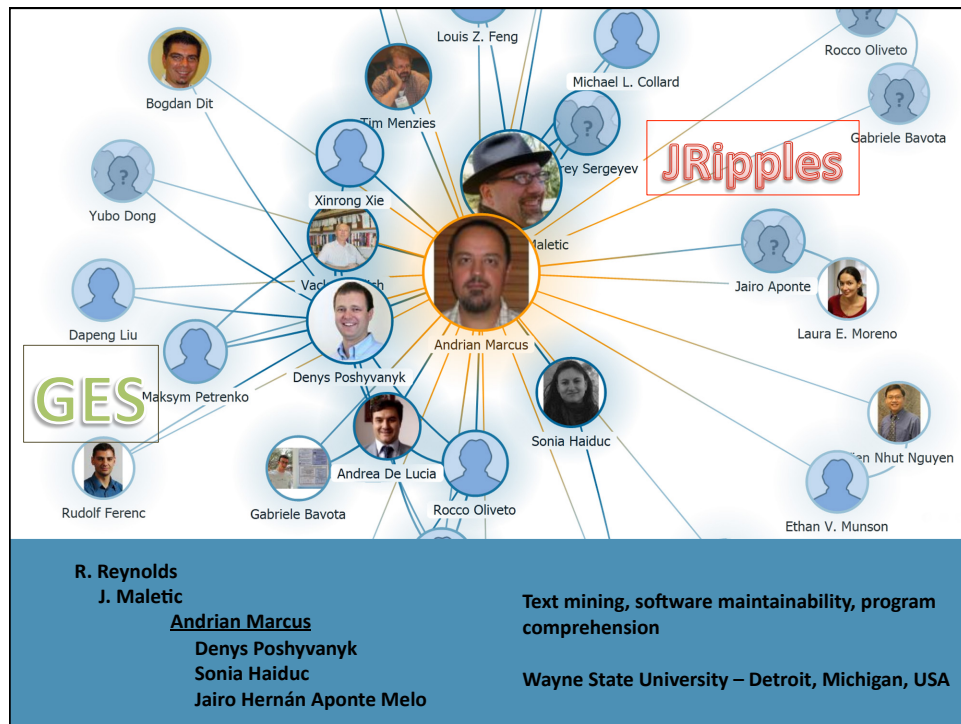
High Recall, Low Precision is a norm

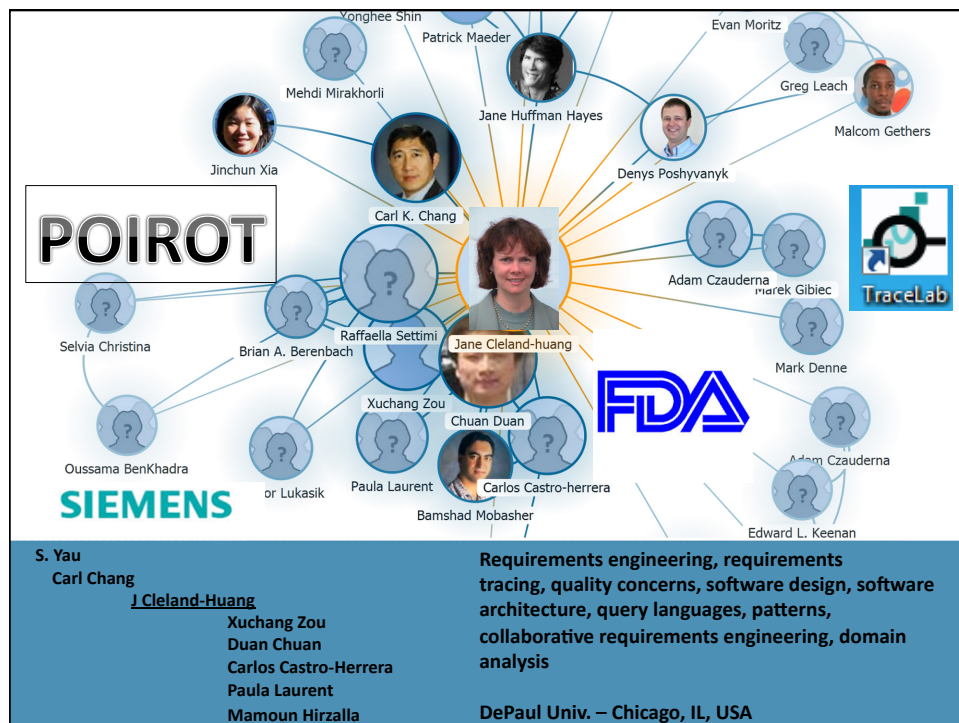
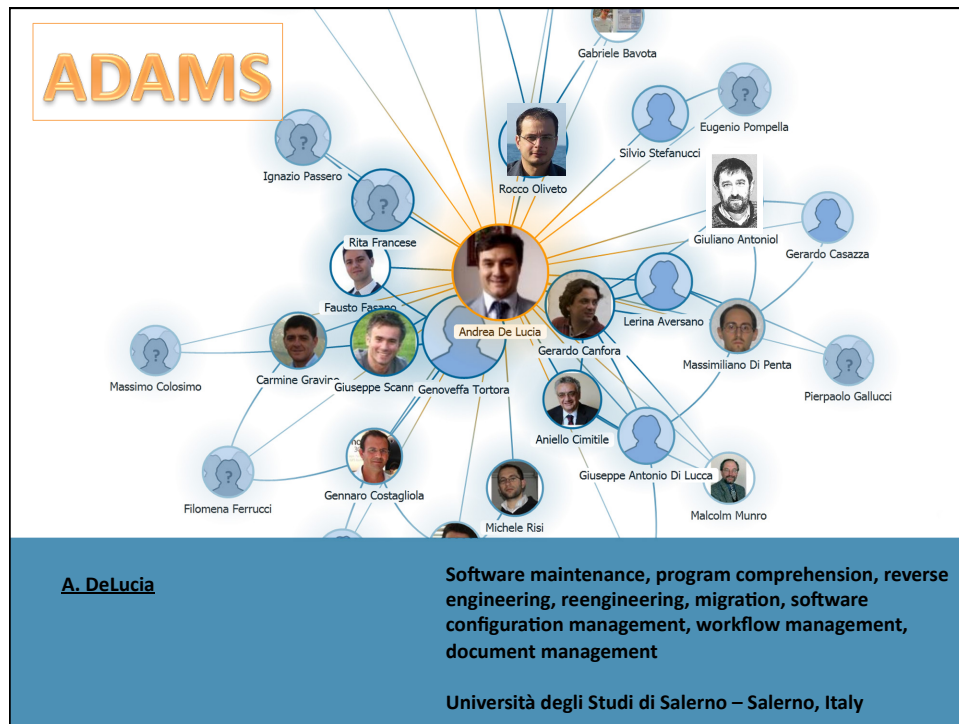
Move from MODIS to “real” datasets

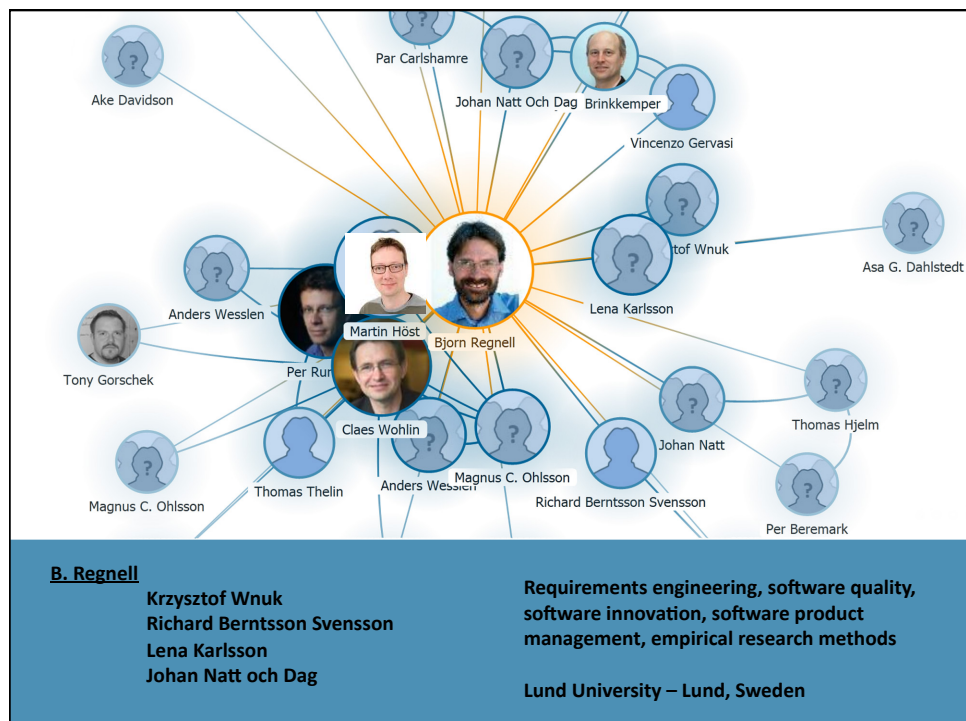
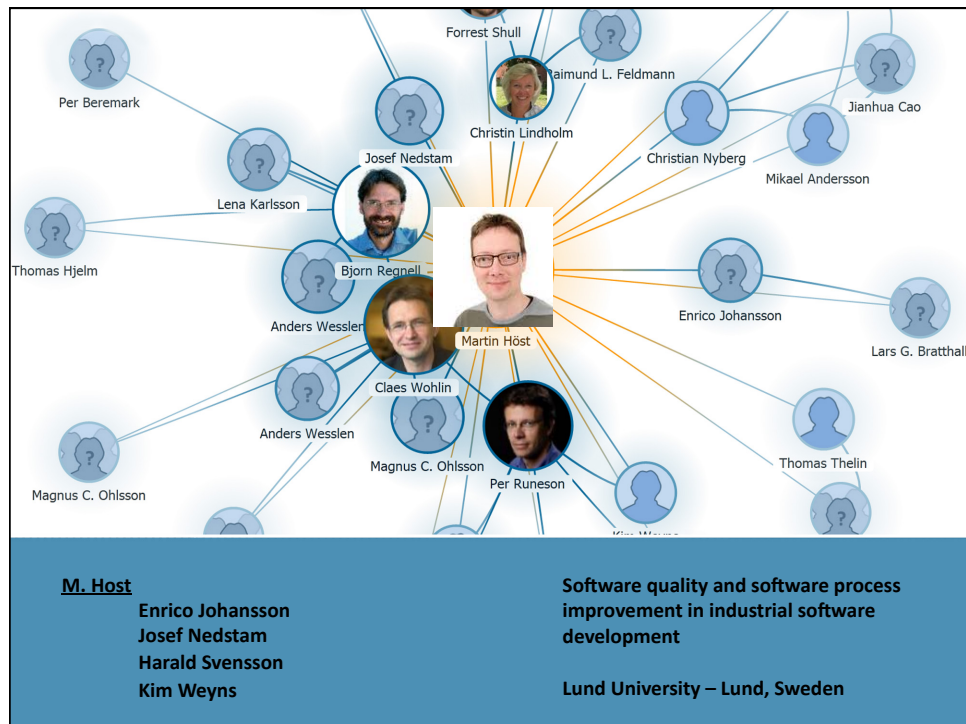
Human Analysts: can't live with'em
can't live without'em

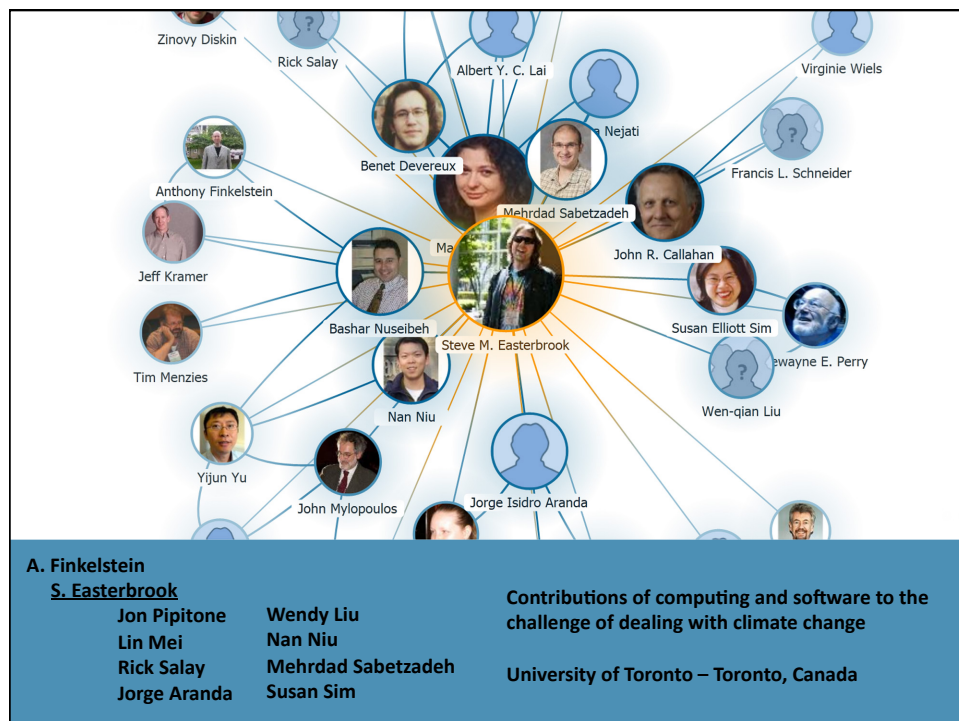
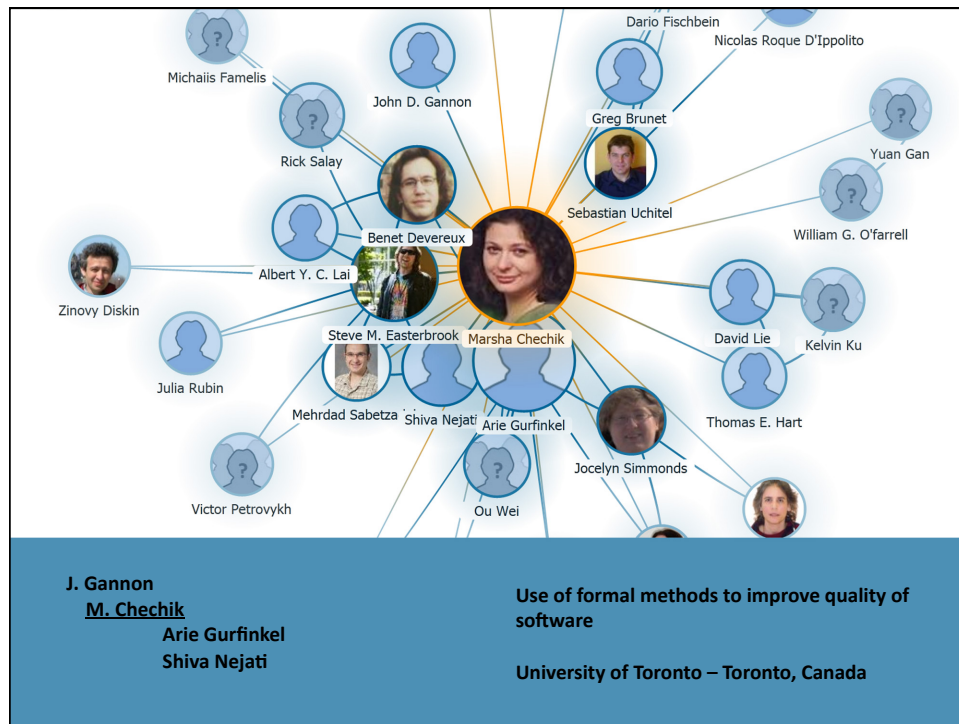
This award belongs to our community...

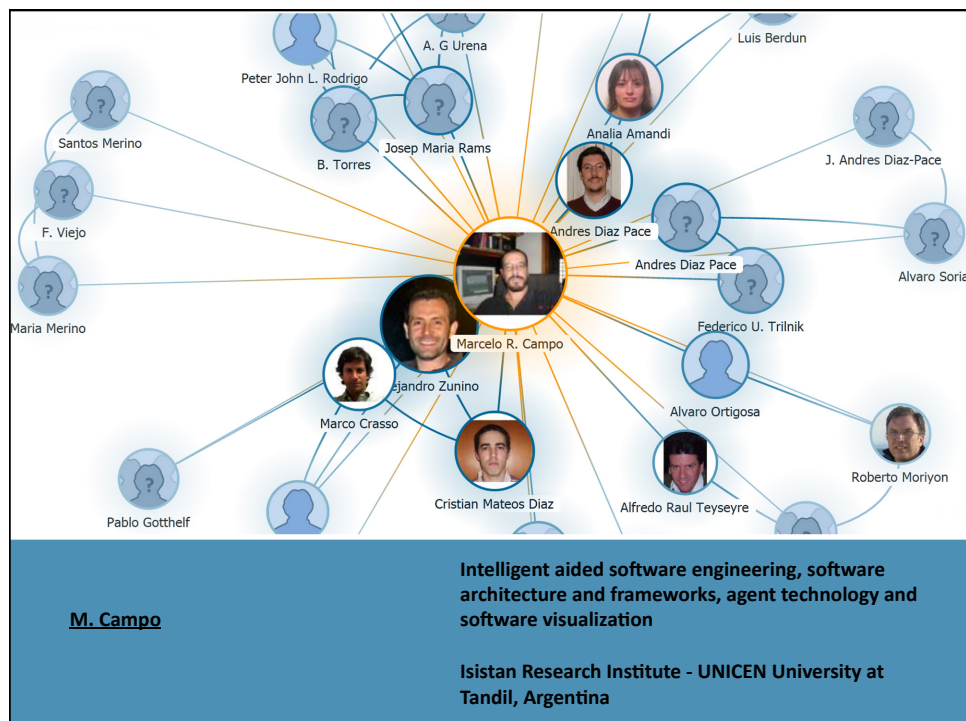
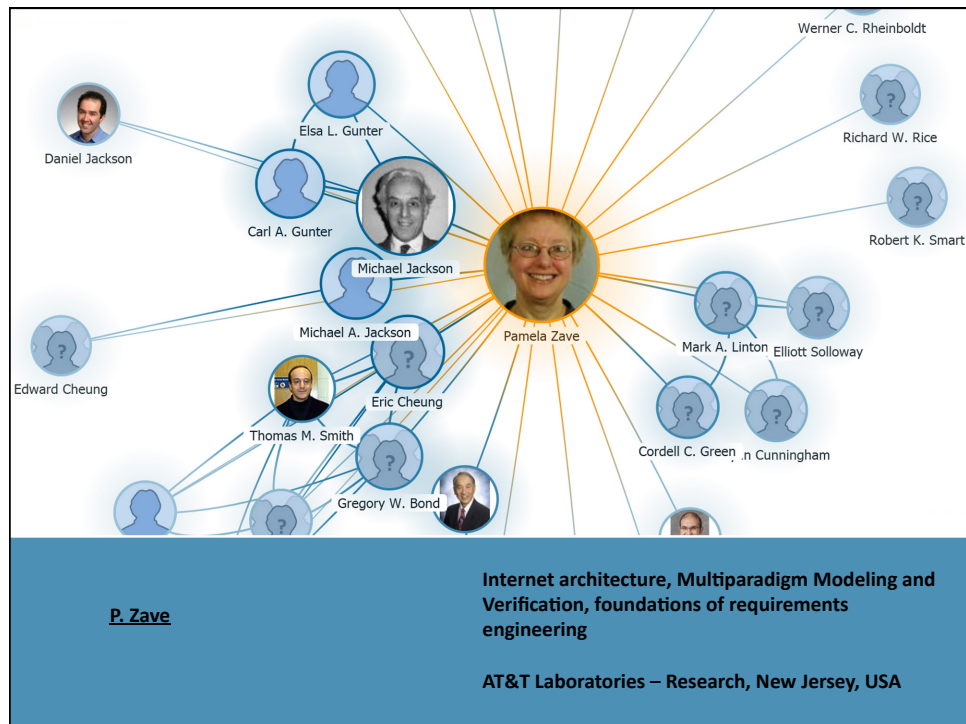
INFLUENCE

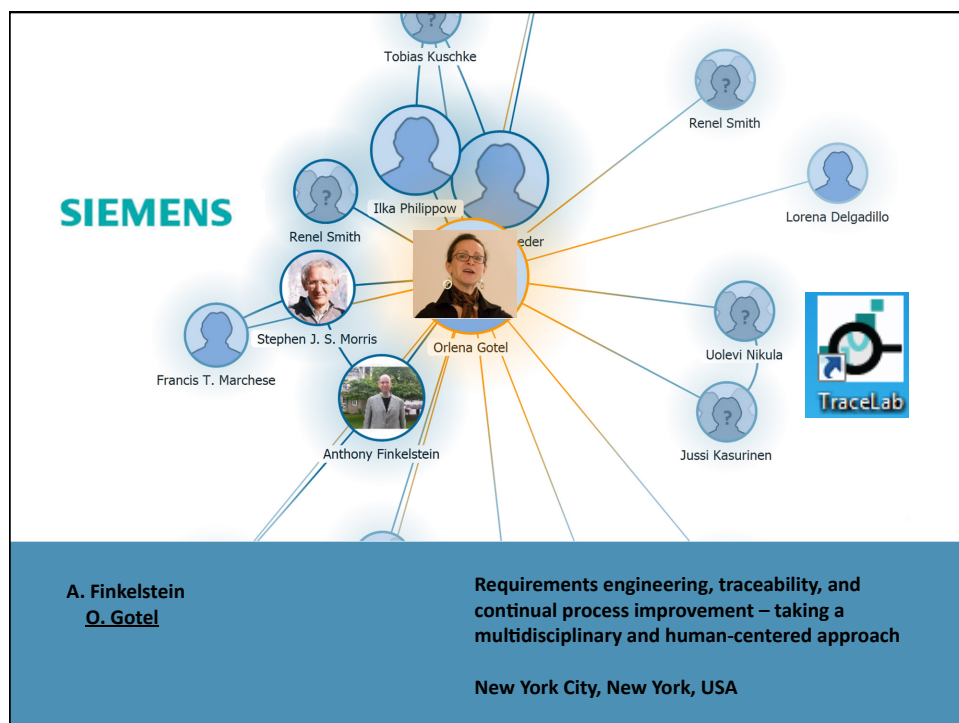
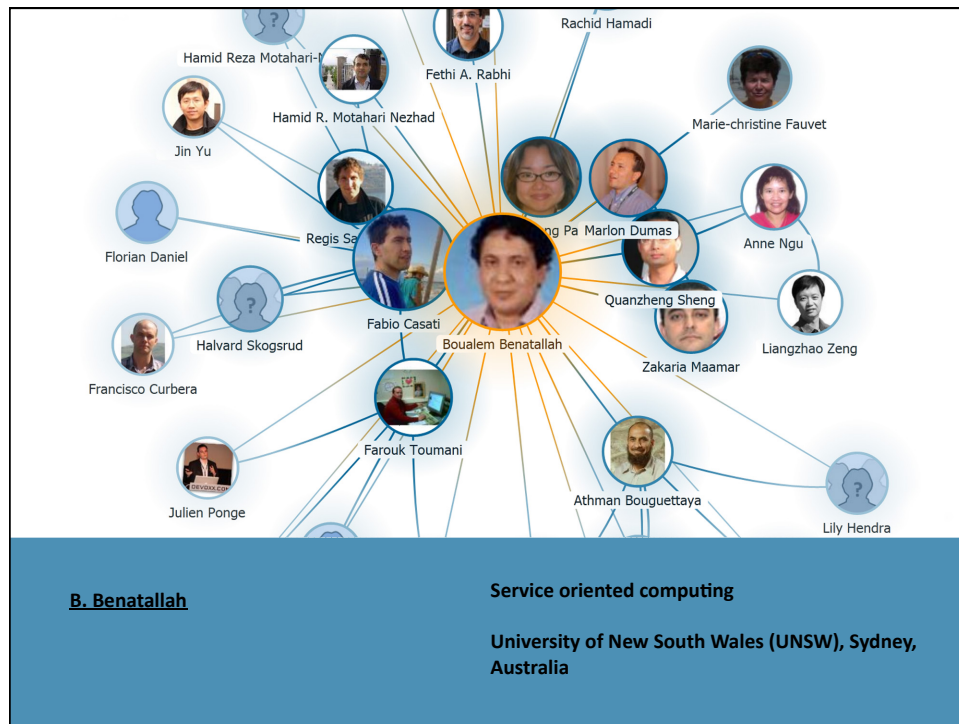


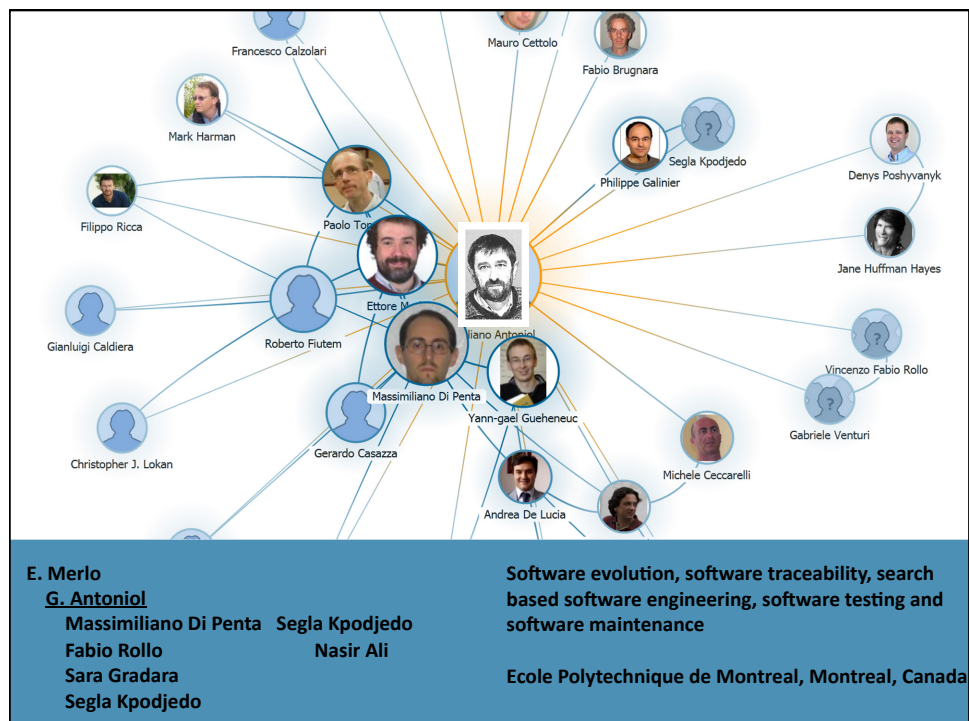
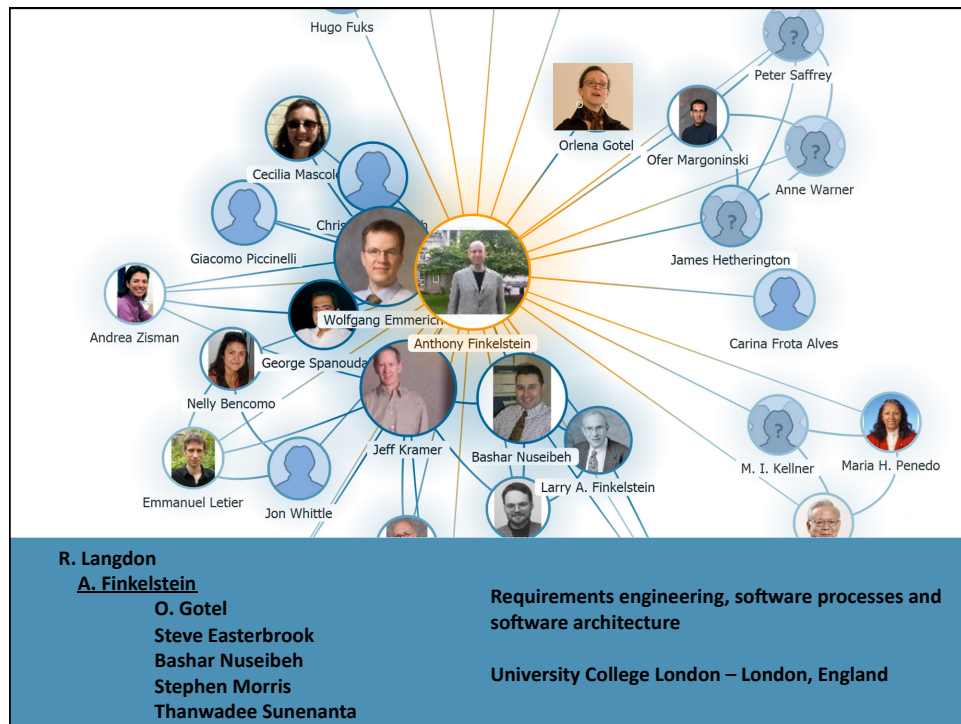


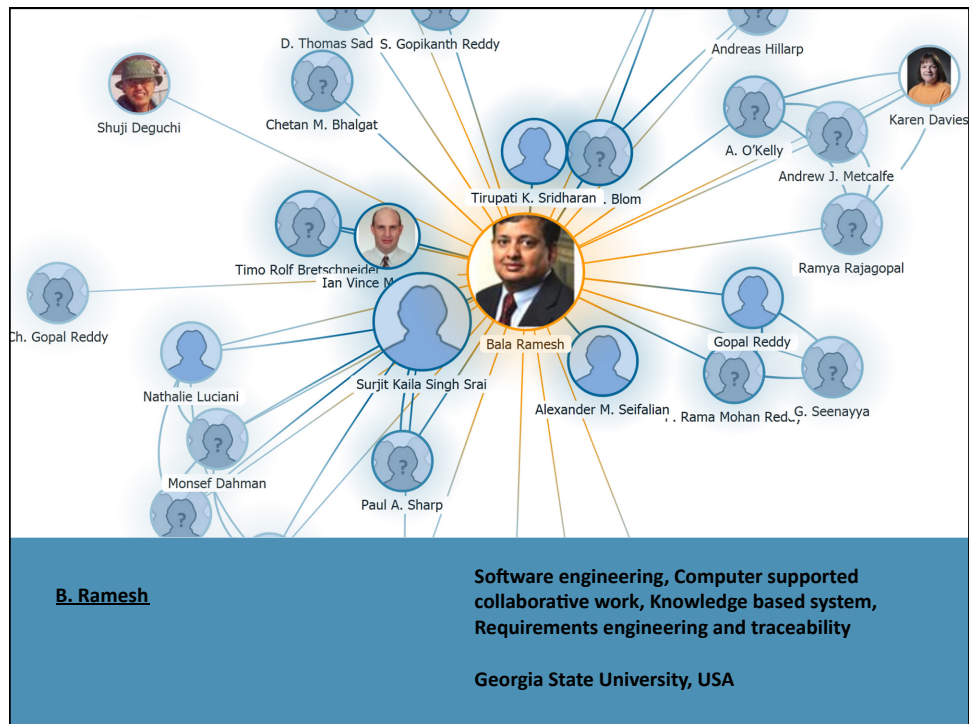


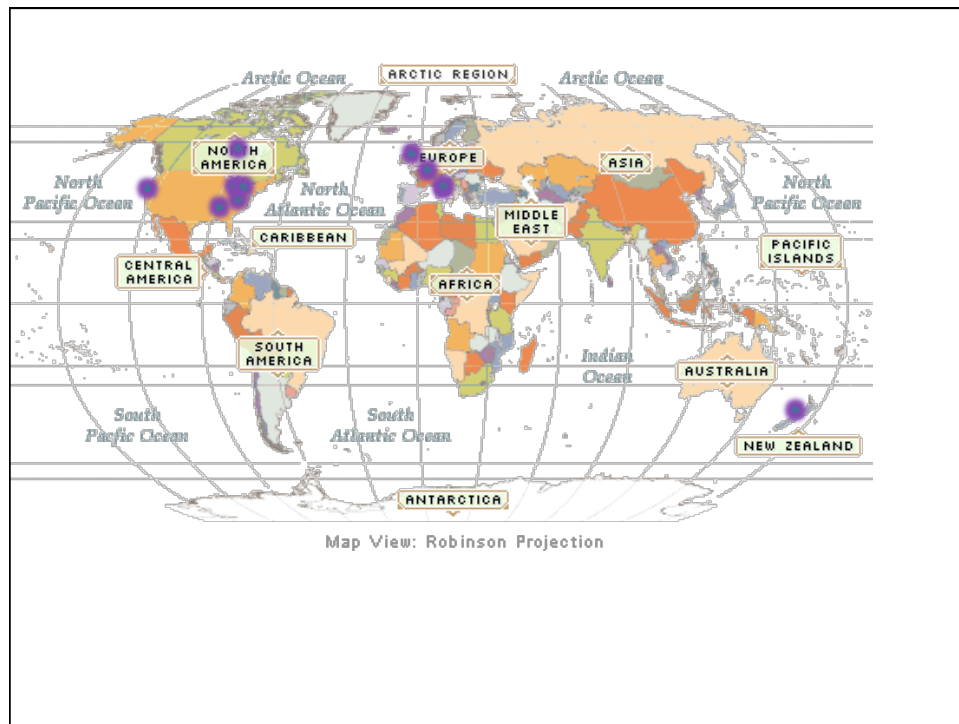


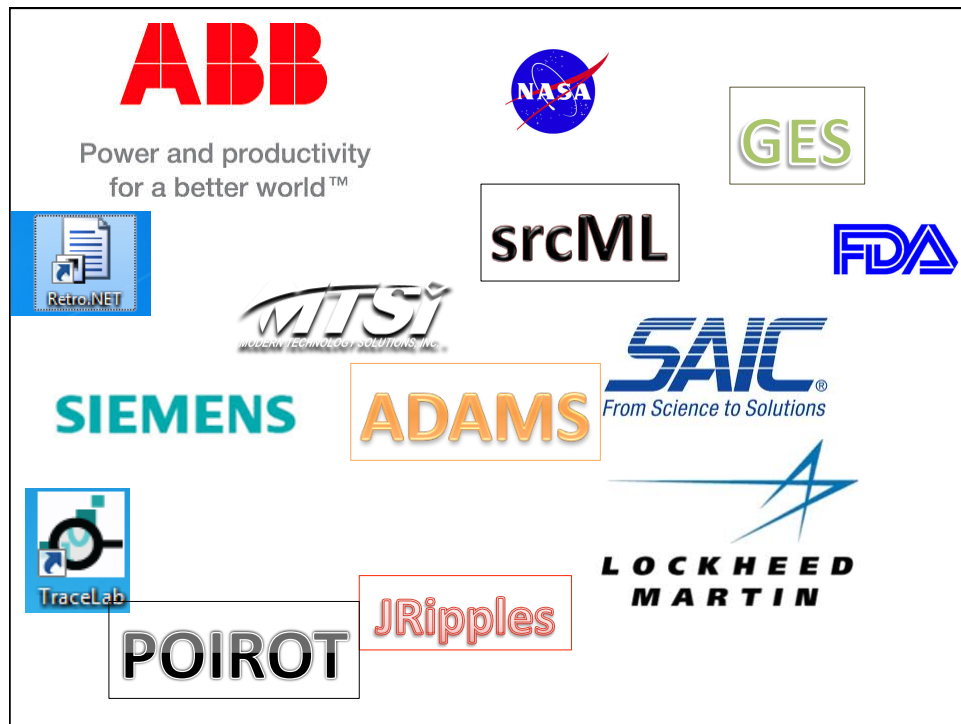






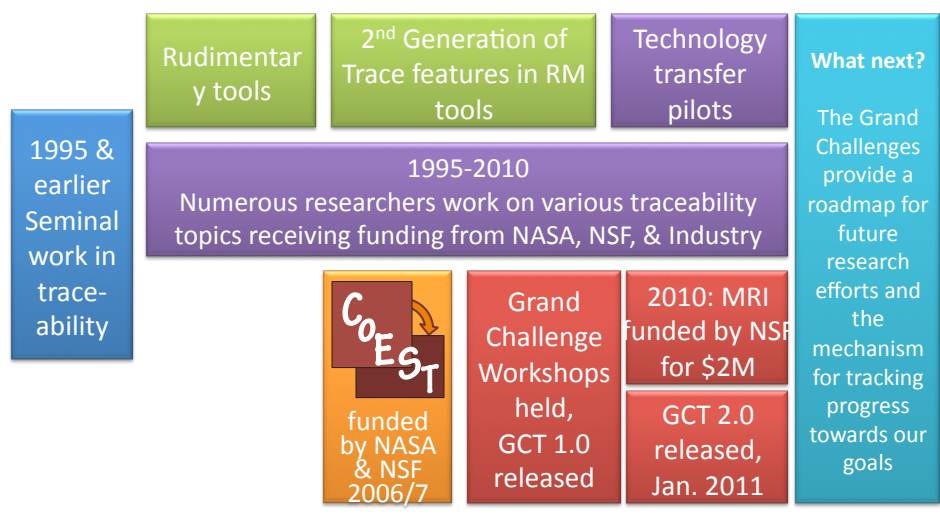


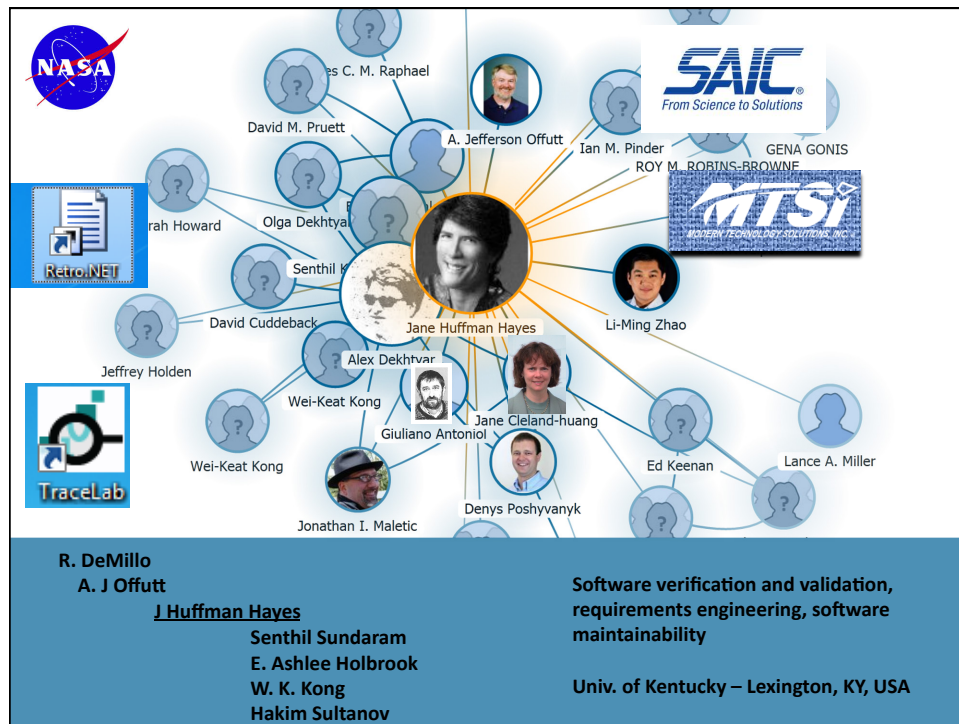




QUESTIONS?

PAST, PRESENT, FUTURE

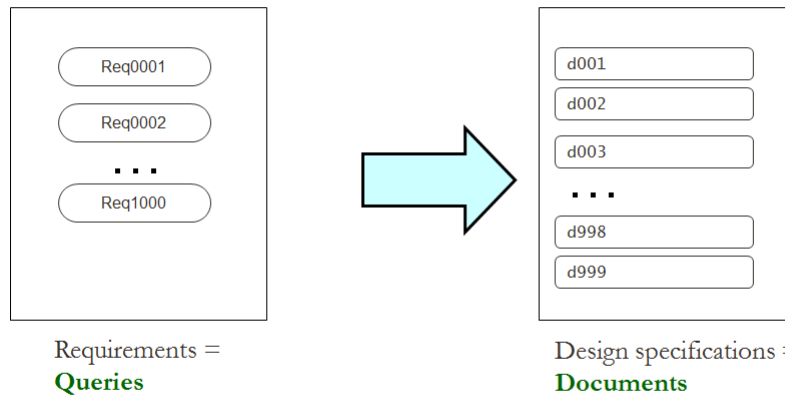




FUTURE DIRECTIONS

- ❖ Ubiquity
- ❖ Assisted methods
- ❖ Application-specific

TRACING AS INFO RETRIEVAL PROBLEM



WHAT WE LEARNED (2003)

Monterey, CA is AWESOME!



WHAT WE TRIED IN 2003: VECTOR (TF-IDF) MODEL

- Vocabulary of keywords determined $V = \{k_1, \dots, k_n\}$
- Each document (*and query*) is represented as a vector

$$\mathbf{d} = (w_1, w_2, \dots, w_n)$$

of keyword weights

- Weights:

$$w_i = \text{freq}(k_i, d) \cdot \log\left(\frac{1}{\text{FREQ}(k_i)}\right)$$

+ Simple Thesaurus

Frequency of keyword in document

Inverse frequency of keyword in collection