Microsoft Research

Predictions, Decisions, and Intelligence in the Open World

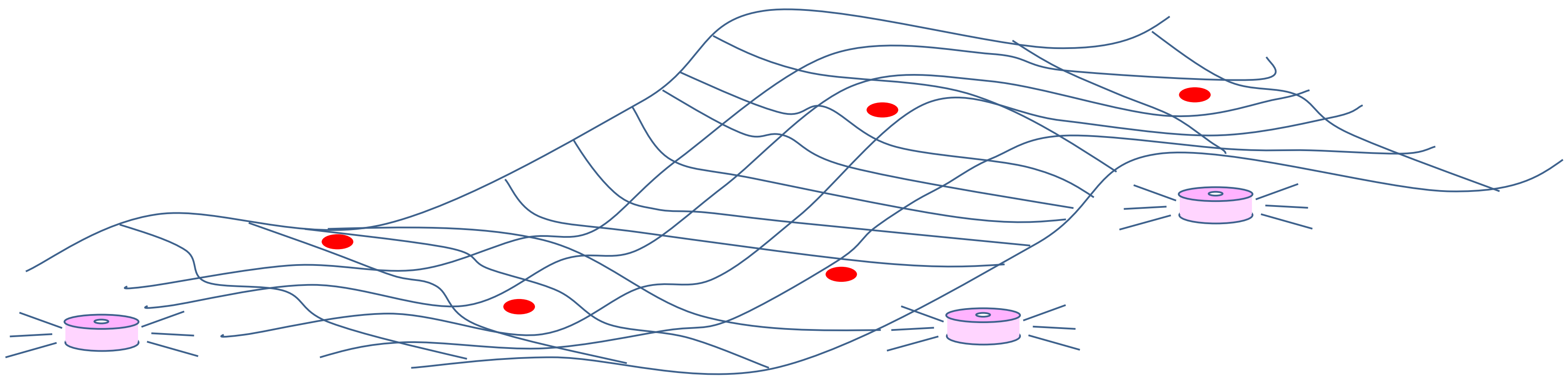
Eric Horvitz
Microsoft Research

Microsoft Research Asia
Faculty Summit 2012

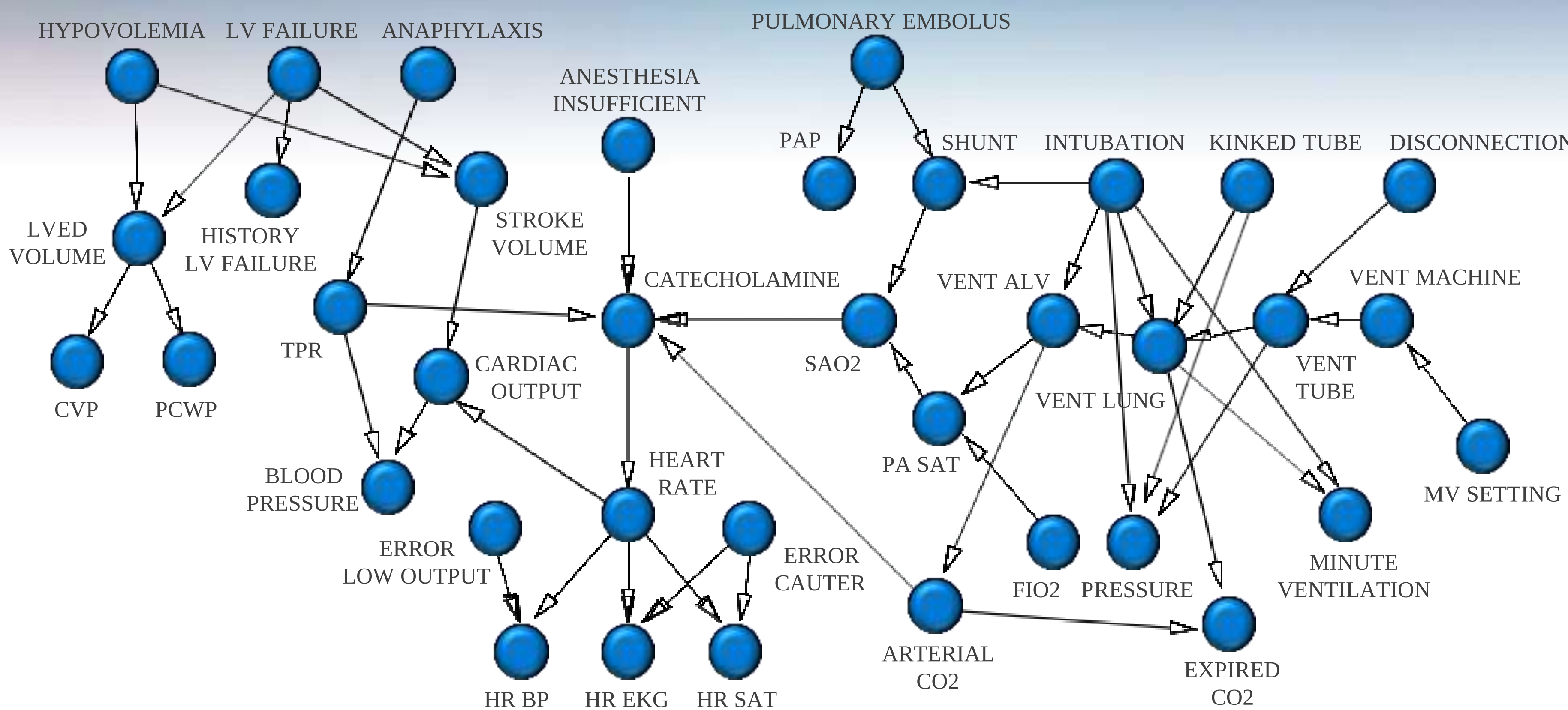
Exciting Times

- ↑ Computation & memory
- ↑ Data & connectivity
- ↑ Learning & reasoning prowess

Opportunities & directions



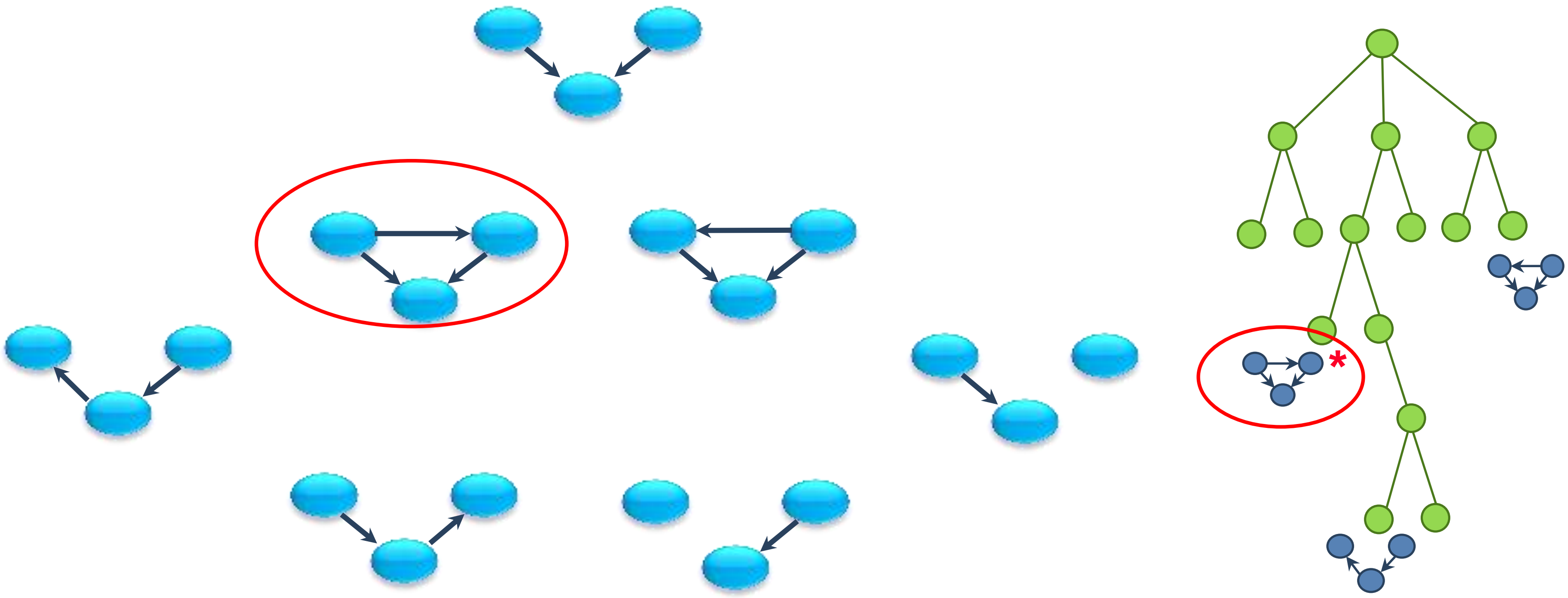
Advances in Representation & Reasoning



I. Beinlich, et al

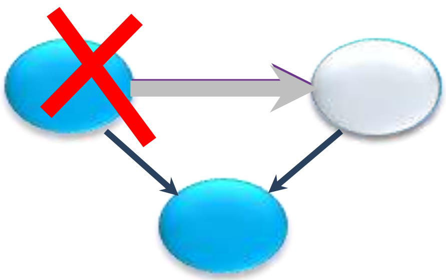
Learning Models from Data

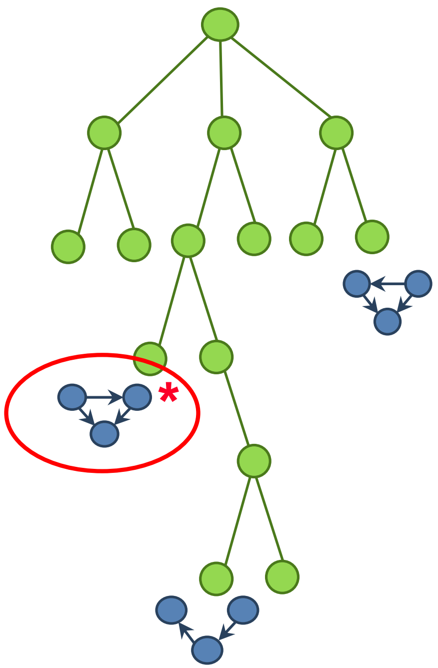
- New access to large amounts of data
- Procedures for learning predictive models



Causality & Hidden Variables

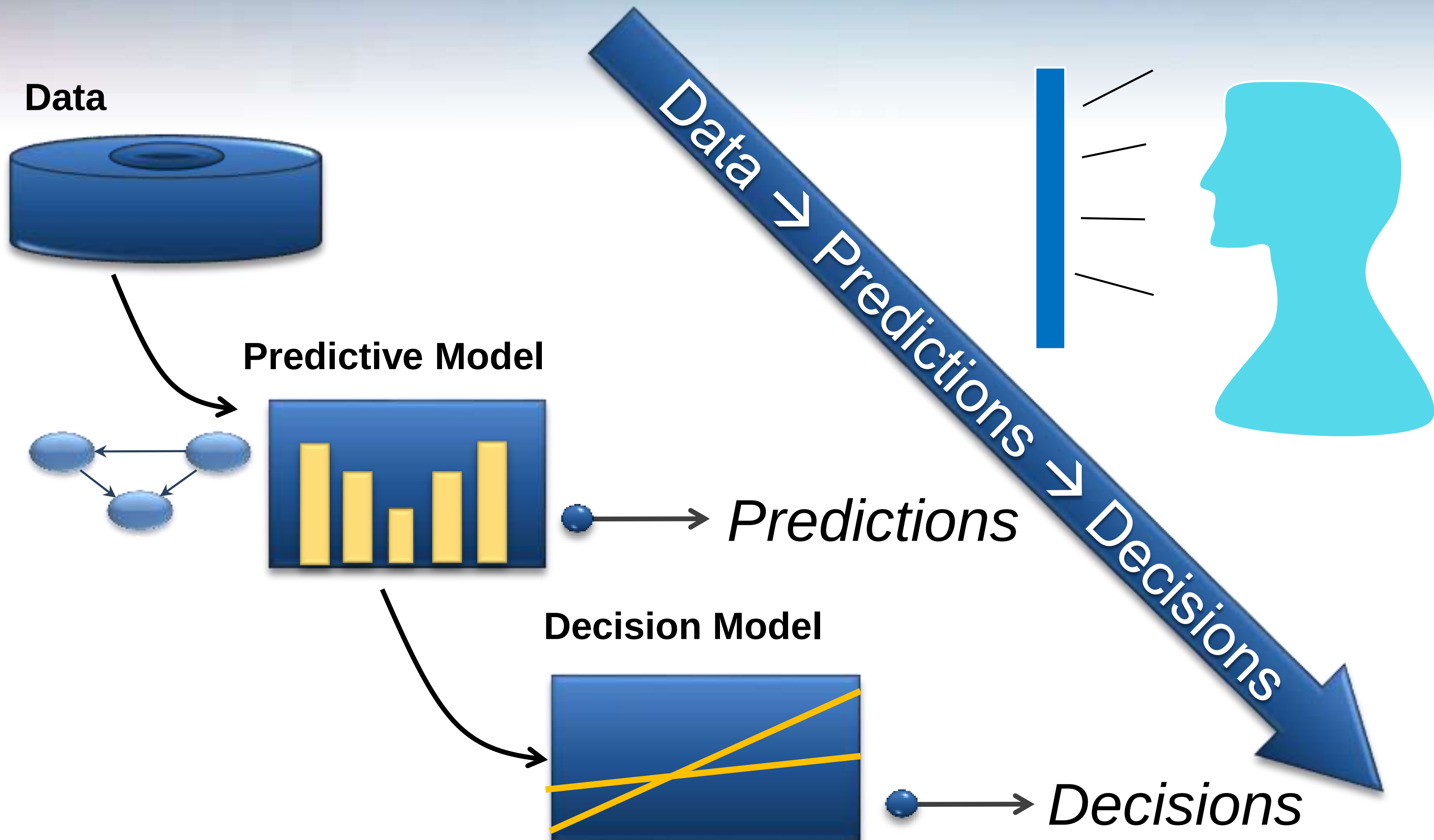
- Causality from observations (sometimes)



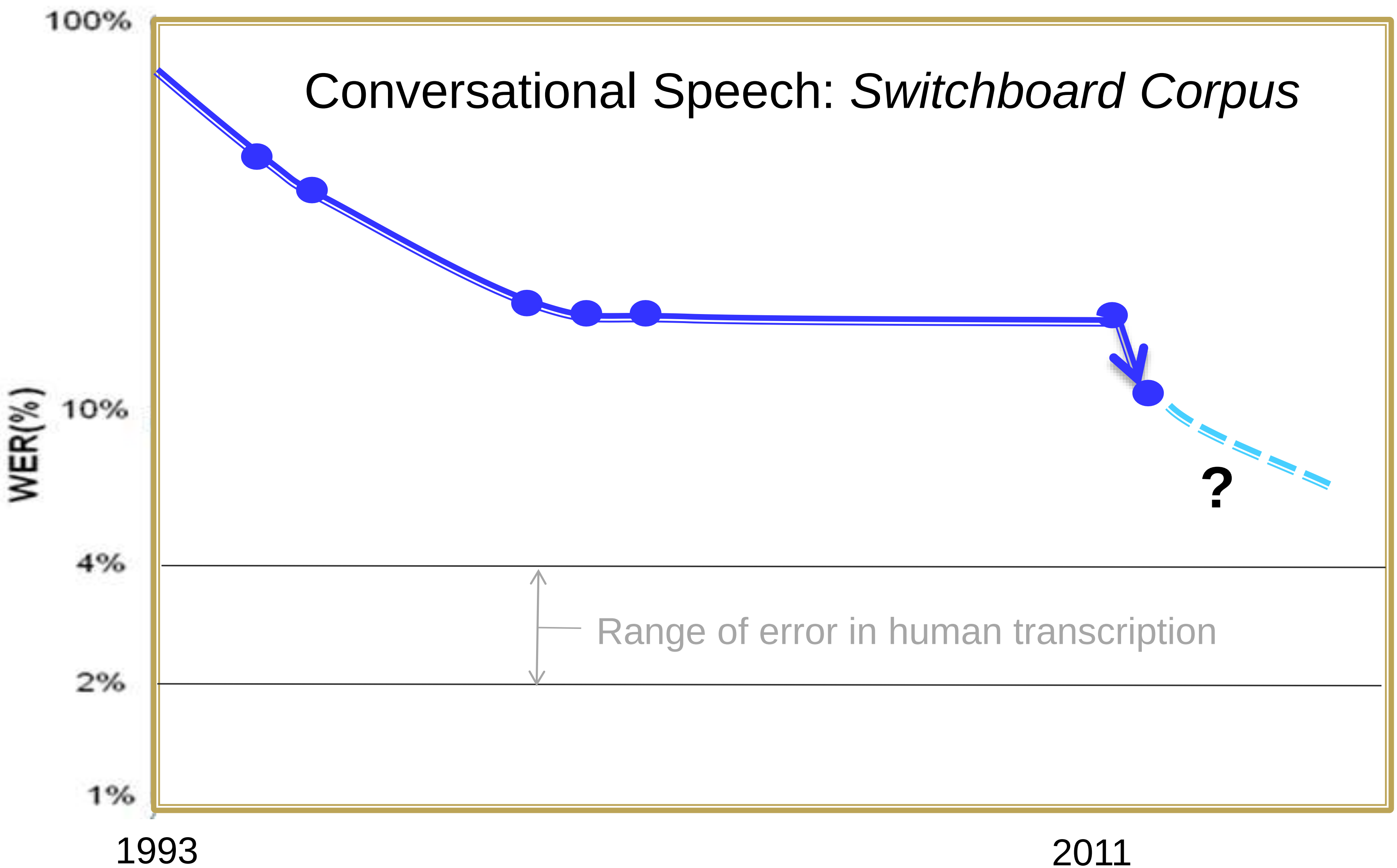


Data → Predictions → Decisions

- Best actions under uncertainty



Exciting Directions

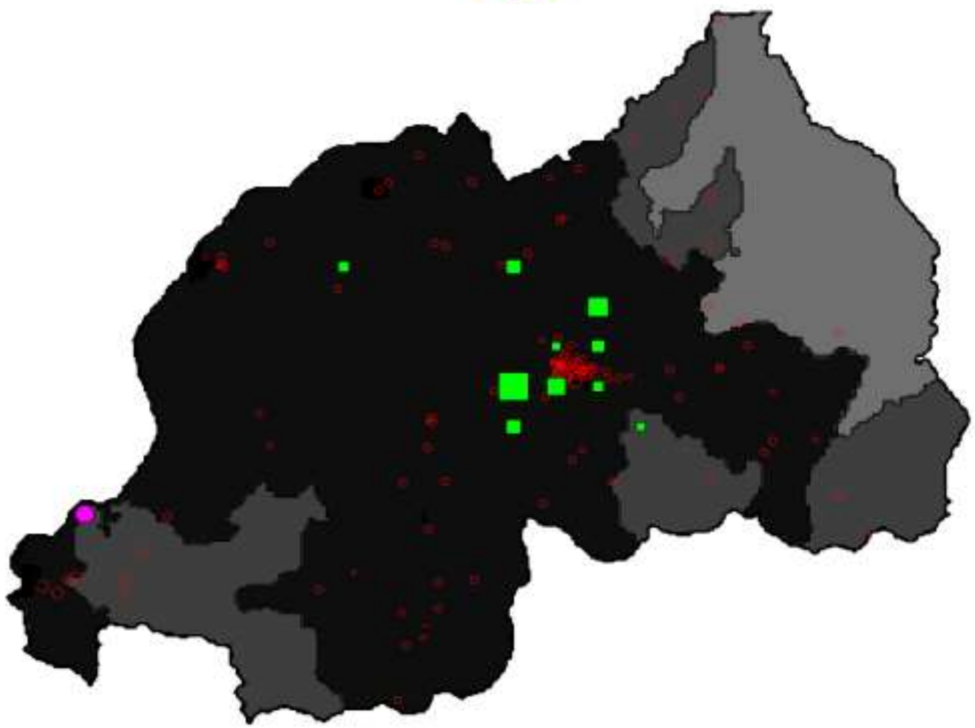


Exciting Directions

Ambient, “in-stream” data resources

Example: *Lac Kivu* earthquake, Congo

Rwandan call densities: 6 days, 140 towers, 10.5m calls



Behind the Scenes… in Daily Life

Learning & prediction in daily use

- Desktop …and on the go
- In car
- Living room

Time Period	TOP4	TOP3	TOP2	TOP1	Number of app launches
Weekdays (WD)	89.62%	85.06%	77.02%	60.51%	90,073,391
Weekends (WE)	92.25%	87.99%	79.99%	62.87%	31,447,609
WD 6pm-12am	92.91%	89.00%	81.10%	63.33%	28,823,389
WD 12pm-6pm	92.76%	88.78%	80.84%	62.99%	24,631,593
WD 6am-12pm	93.17%	89.39%	81.61%	63.47%	14,697,524
WE 12pm-6pm	94.93%	91.35%	83.61%	65.55%	7,575,714
WE 6pm-12am	94.96%	91.34%	83.52%	65.28%	7,338,829
WD 12am-6am	94.65%	91.26%	84.00%	66.18%	5,537,338
WE 6am-12pm	95.53%	92.30%	85.02%	67.05%	3,437,670
WE 12am-6am	96.01%	92.87%	85.84%	68.64%	1,663,091
					121,521,000 Total launches

Bing Traffic-Sensitive Routing

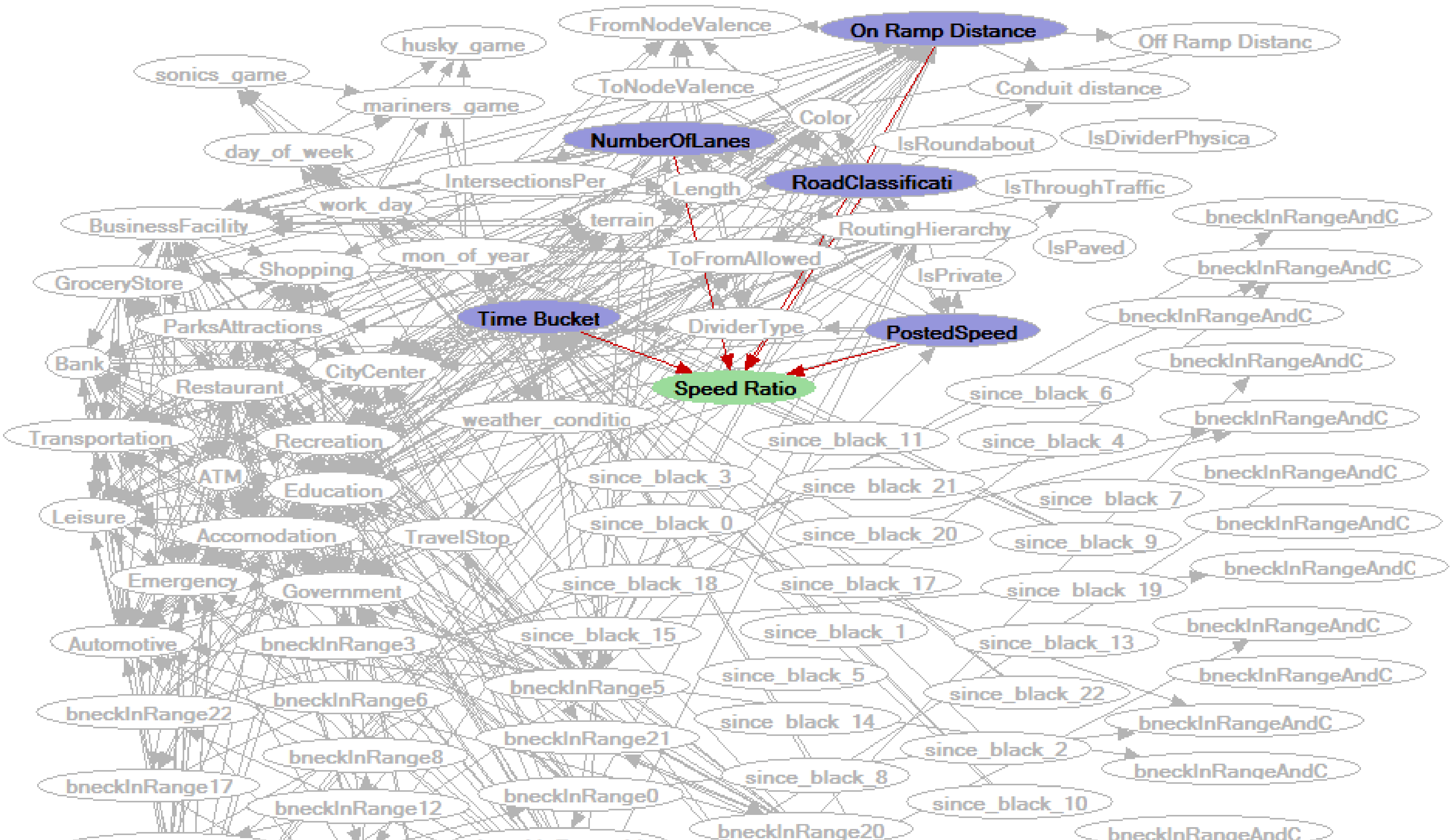
maps.bing.com * m.bing.com

- 72 cities across North America
- Flows assigned to ~60 million streets *every few minutes*

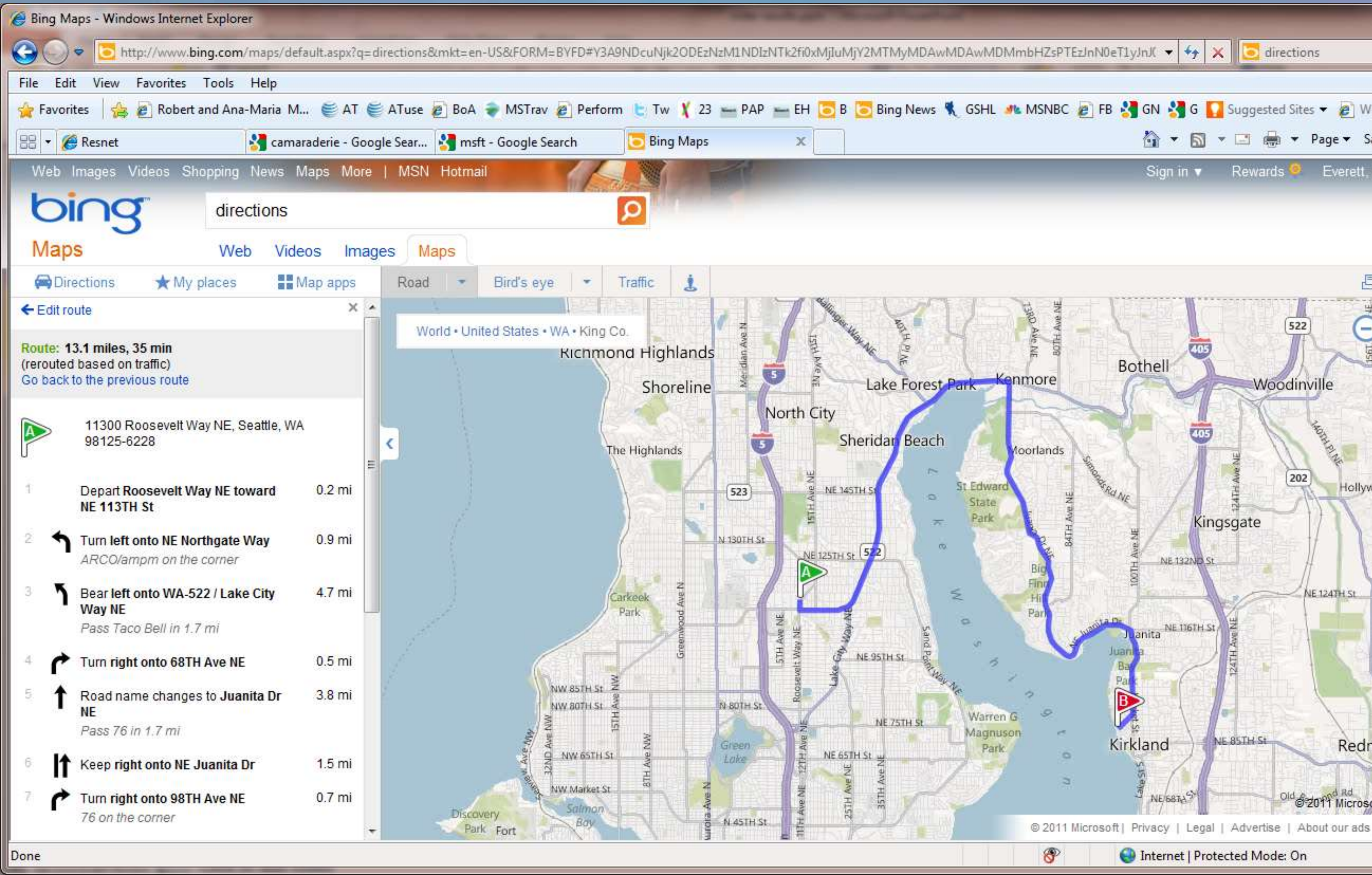
Case library

~1,000,000 km

~100,000 trips



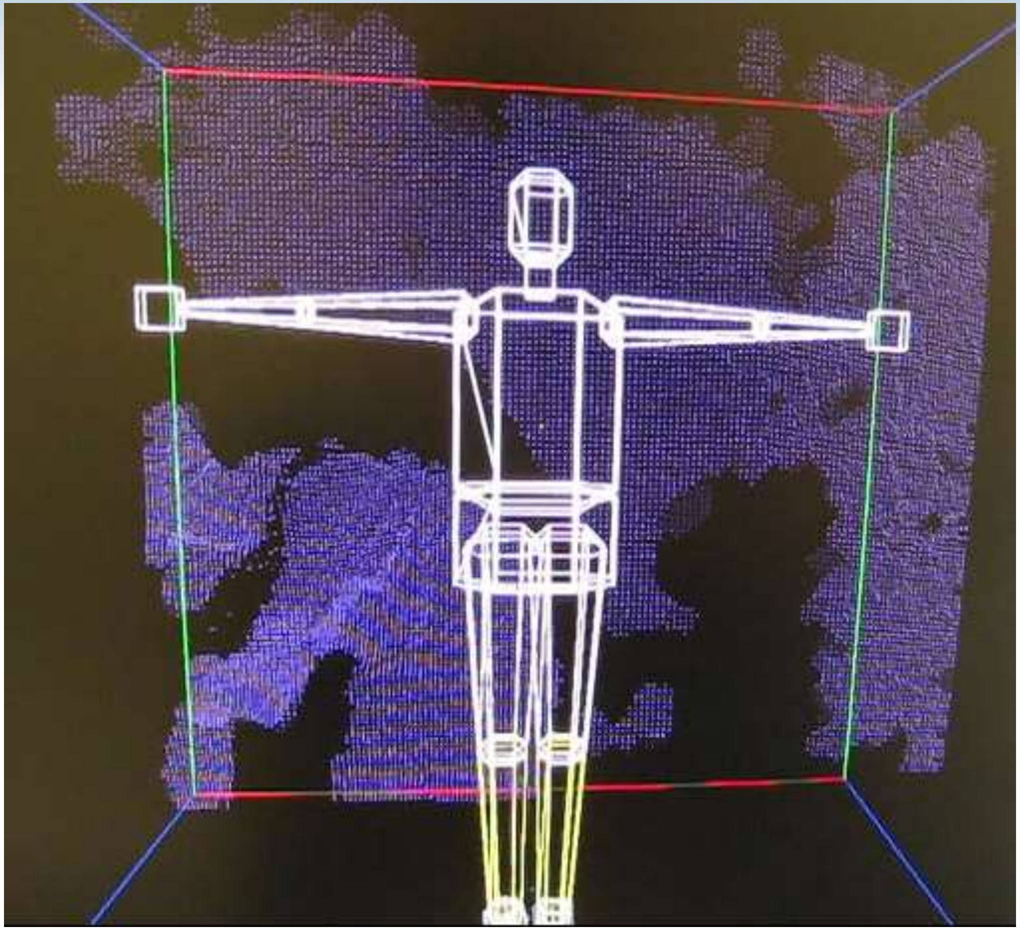
Bing Traffic-Sensitive Routing



And into the Living Room…

KINECT
for XBOX 360.

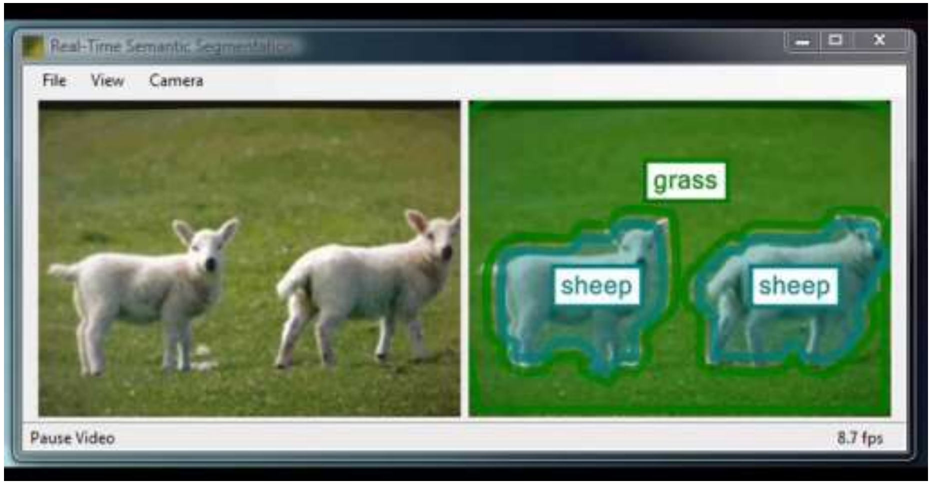
Pursuing Consumer-Centric Robustness



Shotton, J., Fitzgibbon, A. ; Cook, M. ; Sharp, T. ; Finocchio, M. ; Moore, R. ; Kipman, A. ; Blake, A.

Pursuing Consumer-Centric Robustness

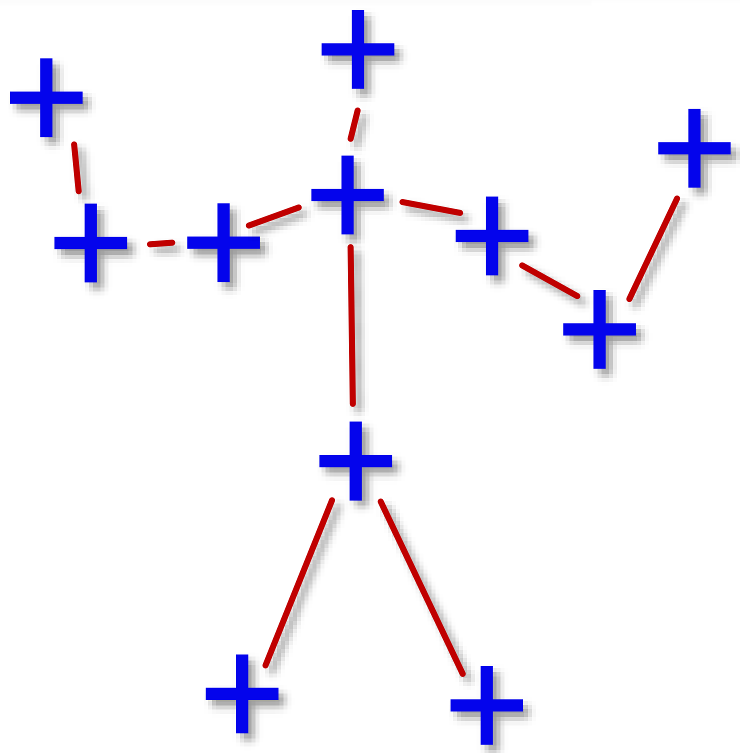
Prior work on segmentation & object recognition

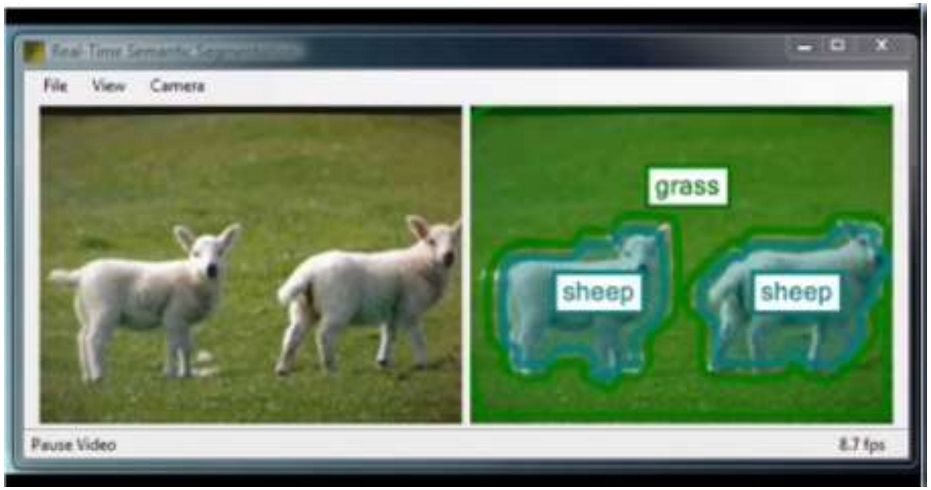


J. Shotton, J. Winn, C. Rother, A. Criminisi

Pursuing Consumer-Centric Robustness

Prior work on segmentation & object recognition





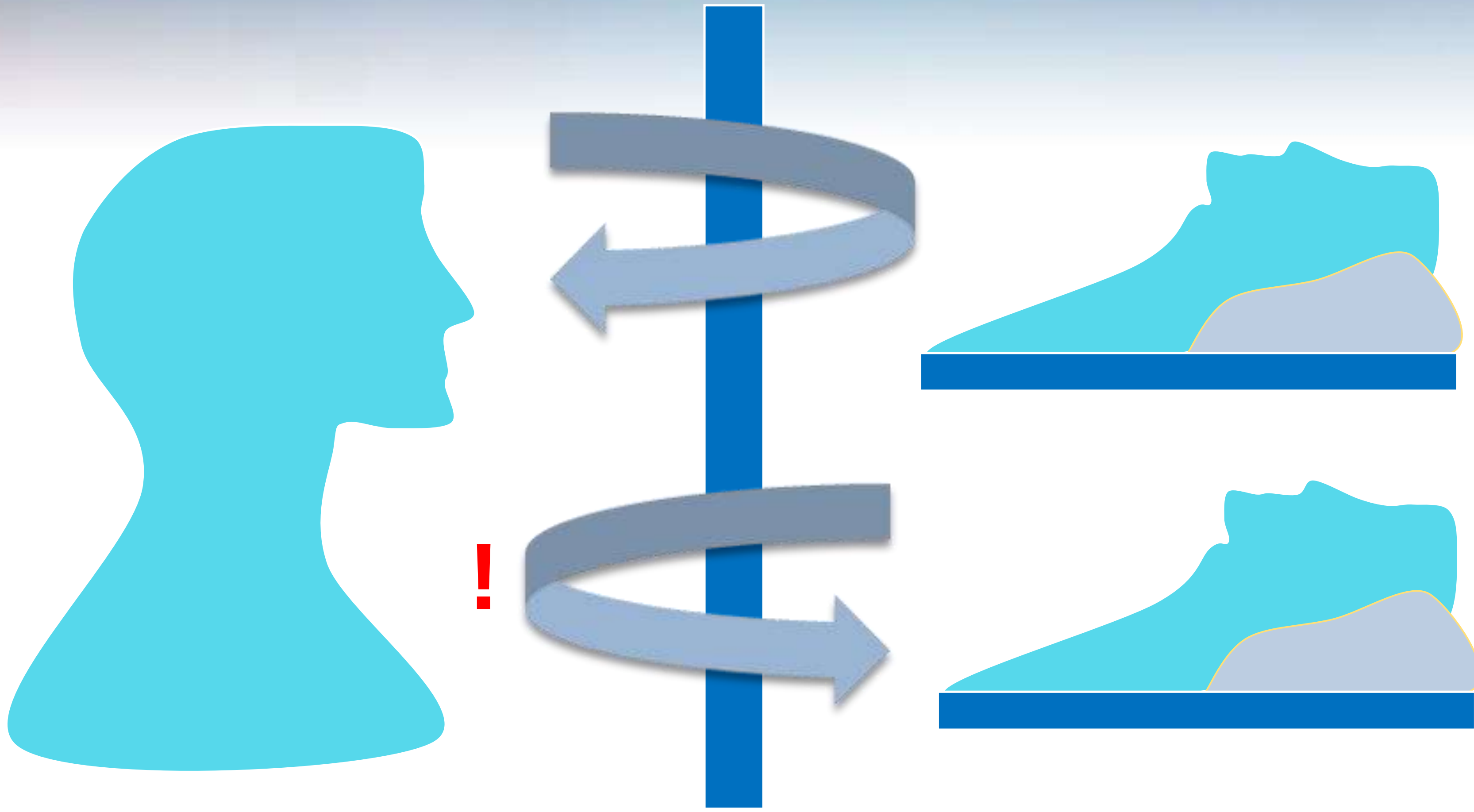
J. Shotton, J. Winn, C. Rother, A. Criminisi

Several Directions

- Healthcare
- Complementary computing
- Integrative intelligence

Focus: Healthcare

Predicting Readmission



With M. Bayati, M. Braverman, P. Koch, K. Mack, M. Gillam, M. Smith, R. Cazangi, J. Gatewood, PD S

Readmissions Challenge

The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL ARTICLE

◄ Previous Volume 360:1418-1428 April 2, 2009 Number 14 Next ►

Rehospitalizations among Patients in the Medicare Fee-for-Service Program

Stephen F. Jencks, M.D., M.P.H., Mark V. Williams, M.D., and Eric A. Coleman, M.D., M.P.H.

ABSTRACT

Background Reducing rates of rehospitalization has attracted attention from policymakers as a way to improve quality of care and reduce costs. However, we have limited information on the frequency and patterns of rehospitalization in the United States to aid in planning the necessary changes.

Methods We analyzed Medicare claims data from 2003–2004 to describe the patterns of

- **~20% within 30 days**
- **~35% in 90 days**

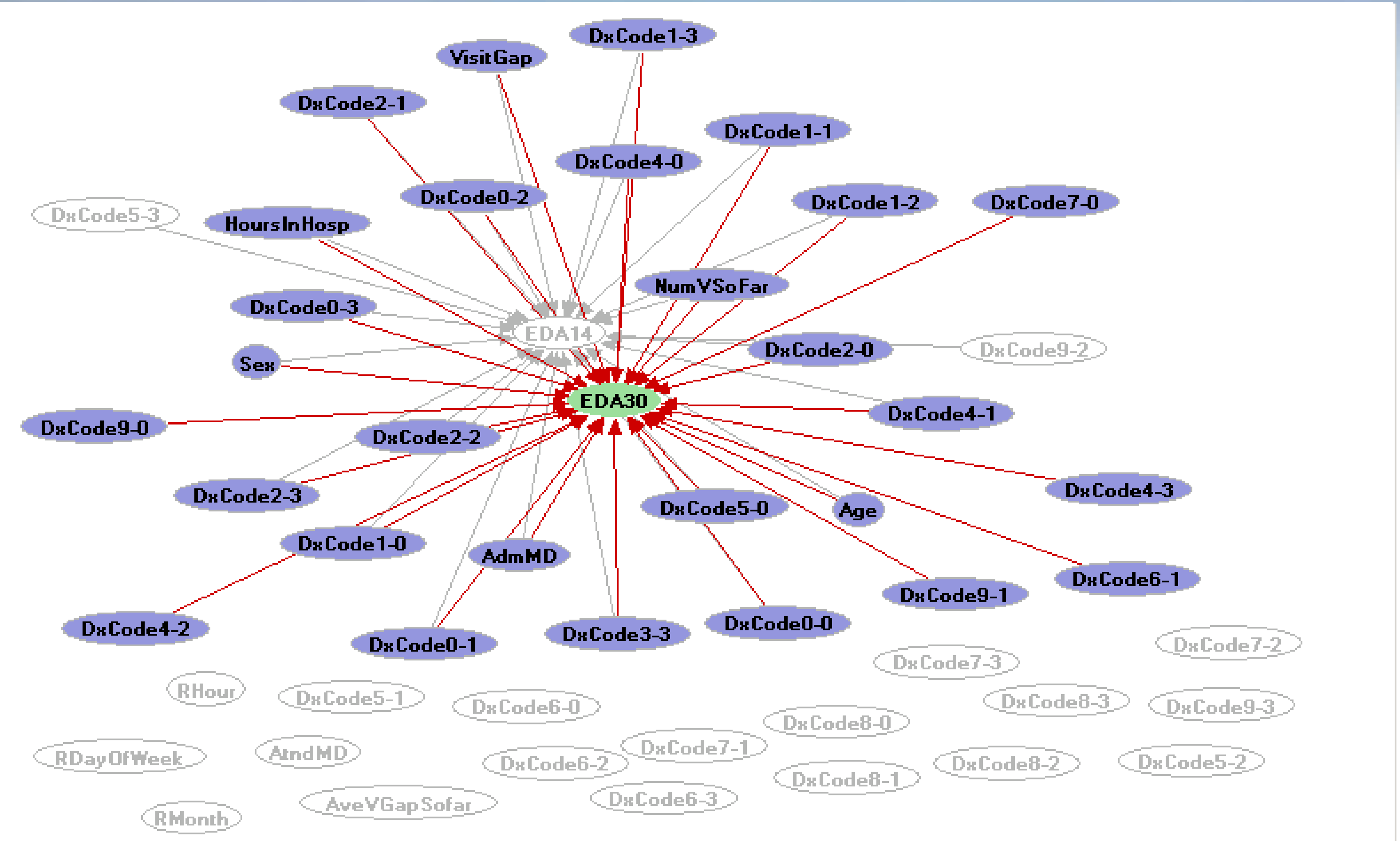
➢ ***Estimated cost to Medicare in 2004: $17.4 billion***

Learning from a Case Library

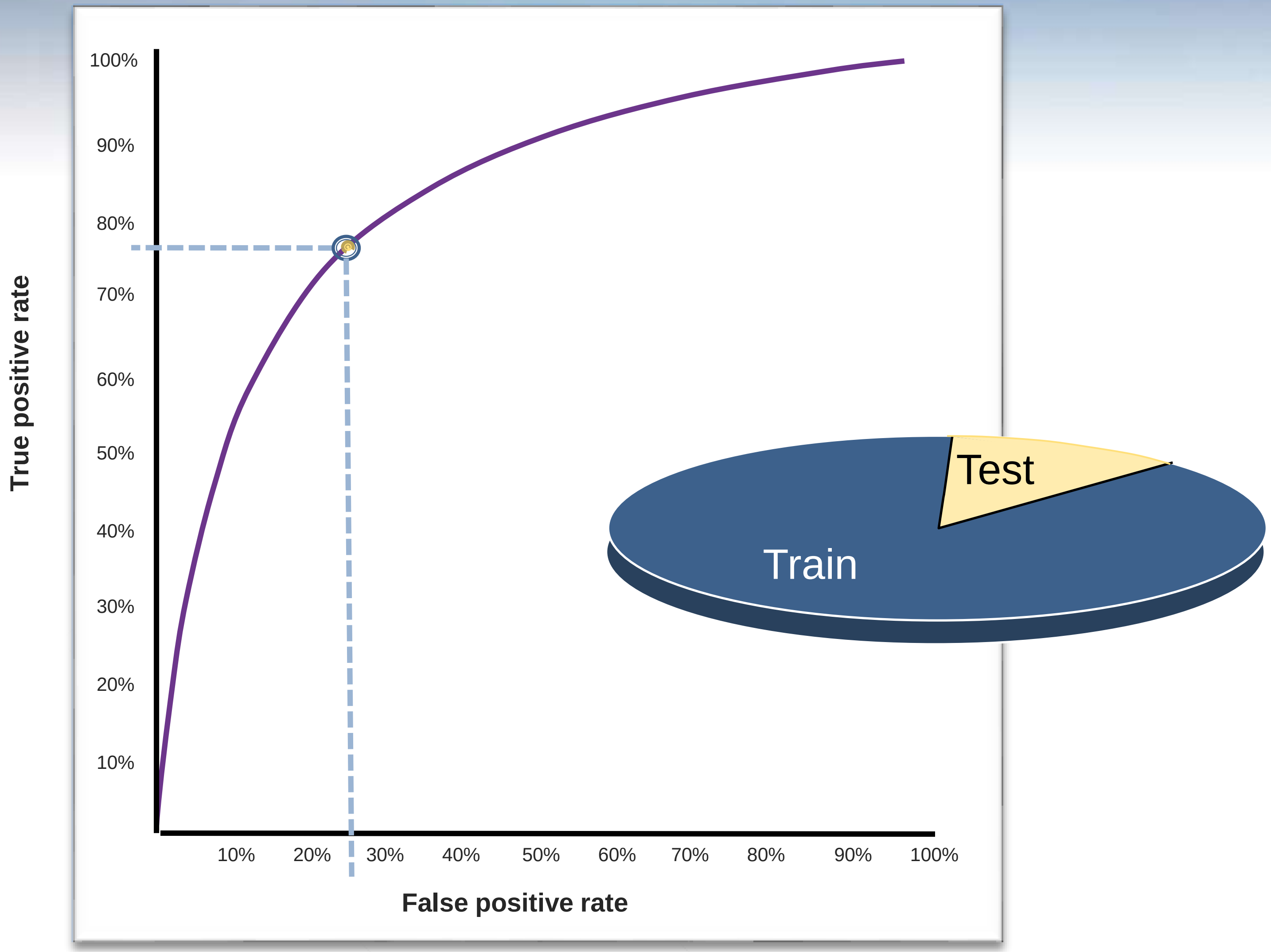
- Washington Hospital Center (Wash DC)
- All visits during the years 2001 to 2009
  (e.g., ~300,000 ED visits)
  - Admissions, discharge, transfer (ADT)
  - Chief complaint in free text
  - Age, gender, demographics
  - Diagnosis codes (ICD-9)
  - Lab results and studies
  - Medications
  - Vital signs
  - Procedures
  - Locations in hospital
  - Admitting and attending MD codes
  - Fees and billing

~25,000 variables considered in dataset

Building a Predictive Model for Readmission



Performance of Classifier for Readmission



Identifying Evidential Relevance

Weight	Feature description	Frequency
0.68398	Dx0->2 = Excessive vomiting in pregnancy	0.31%
0.61306	Dx3->2 = Personal history of malignant neoplasm	0.28%
0.58281	Dx0->2 = Heart failure	0.30%
0.56708	Dx0->1 = Nephritis, nephrotic syndrome, and nephrosis	0.09%
0.56649	Dx3->2 = Heart failure	0.28%
0.54663	Complaint sentence contains ''suicidal''	0.17%
0.48415	Dx1->2 = Disorders of function of stomach	0.07%
0.47257	Dx5->0 = Diseases Of The Genitourinary System	0.15%
0.46136	Dx0->2 = Chronic airway obstruction, not elsewhere classified	0.10%
0.44555	Dx4->2 = Depressive disorder, not elsewhere classified	0.10%
0.44257	Stayed 14 hours in the ER	0.10%
0.43890	Dx0->1 = Other psychoses	0.32%
0.43513	Dx0->0 = Diseases Of The Blood And Blood-Forming Organs	0.46%
0.42582	Complaint sentence contains ''dialysis''	0.19%
0.41888	Dx0->2 = Depressive disorder, not elsewhere classified	0.27%
0.41302	Dx1->1 = Nephritis, nephrotic syndrome, and nephrosis	0.29%
0.38506	Complaint sentence contains ''fluid''	0.10%
0.37474	69 < Age	9.22%

Translation: Research → Open World

Readmissions Manager for Microsoft Amalga

Reducing Hospital Readmissions is an Impending Priority

Overview

One in five Medicare inpatients is readmitted within 30 days. The Centers for Medicare and Medicaid Services (CMS) considers 40%-75% of these readmissions to be preventable.

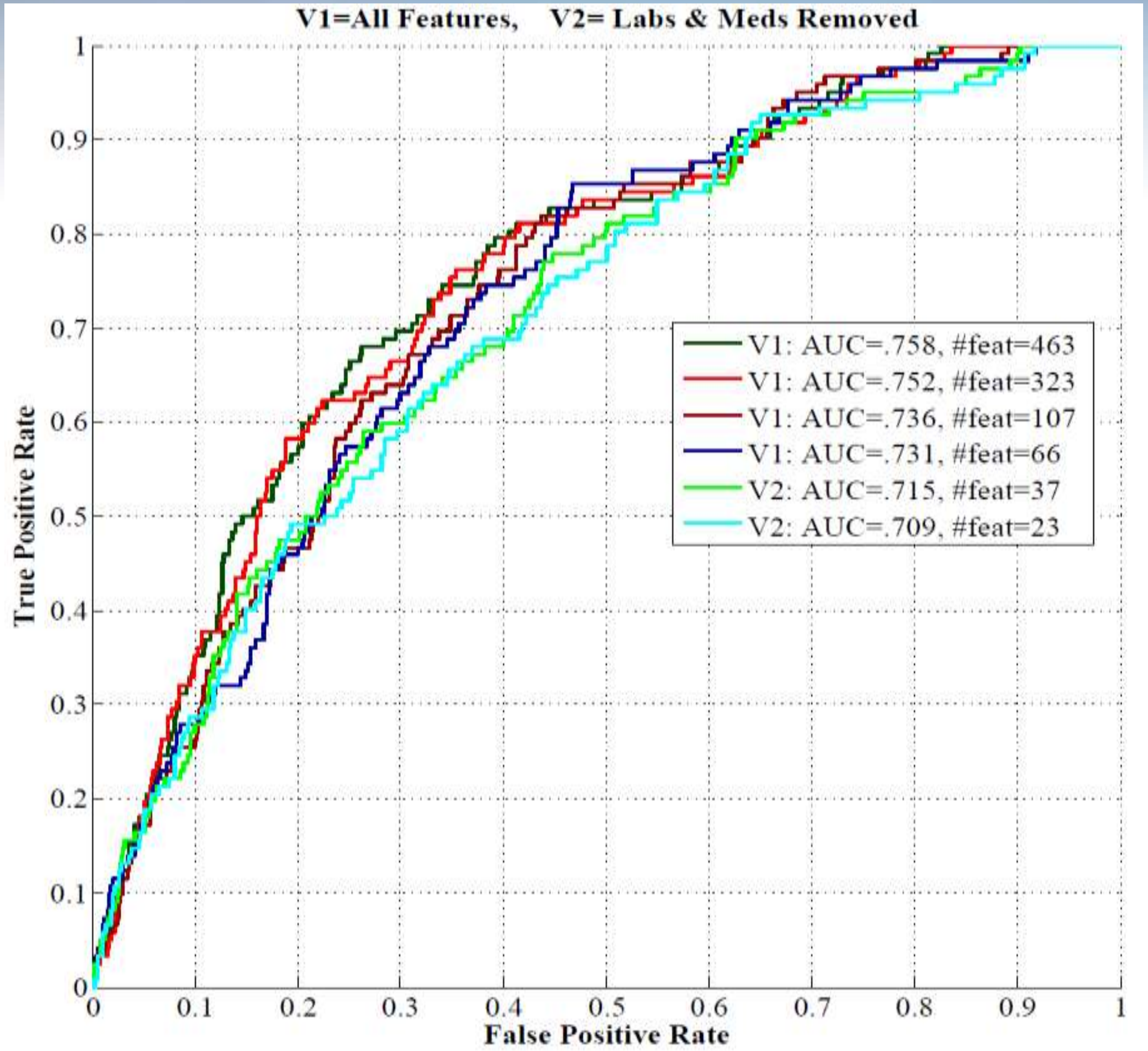
In October 2012, CMS will begin to track readmission and impose financial penalties on hospitals with higher–than–expected readmission rates for certain conditions. Other payers will certainly follow.

It is clear that hospital admissions and readmissions are becoming a critical parameter for tracking care delivery from both a financial and quality perspective.

Readmissions Manager for Microsoft Amalga is an innovative solution to help organizations address this very important business need.

Readmissions Manager Targets Avoidable Hospital Readmissions

Engineering Real-World Solutions



Predictive Platform goes live…

Microsoft Amalga - recazang

US - Sample Hospital

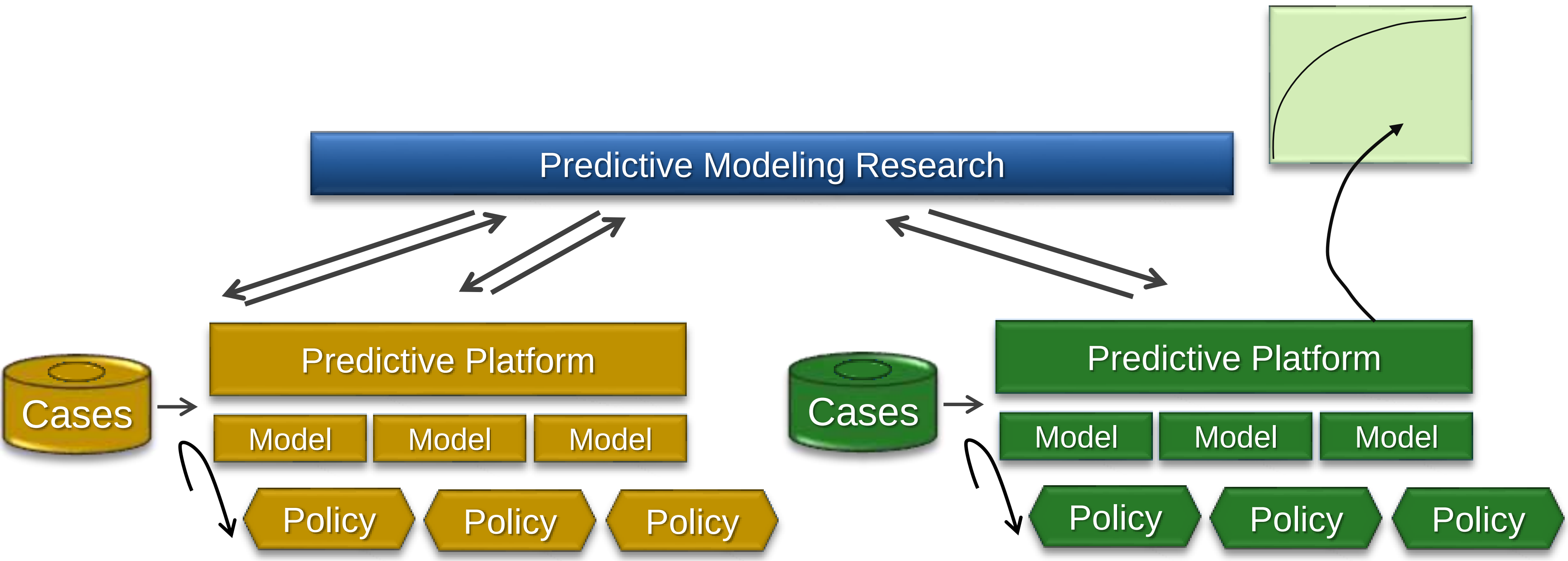
M3L Inp/Inp Readmission Prediction Last... | Filter | Sort | Shortcut | Find | Zoom-in | Refresh | System

None | All ro... | Dev | Data Mining | Info | Input | Forms | Admin | Dashboard | New Task

ACCOUNT	ADMITDTTM	DISCHARGEDTTM	AGE	SEX	PROB_NUM_%	FACTOR
	12/03/2010 14:57	12/08/2010 18:03	62	F	37.9	Num past 6m visits = 6 to 10 / P
	12/08/2010 18:45	12/08/2010 18:45	74	M	32.72	stayed <1 day in the hospital / Pa
	11/16/2010 16:14	12/08/2010 18:50	48	M	30.83	Patient had dx = Chronic renal fa
	12/02/2010 13:49	12/08/2010 18:14	68	M	29.05	Patient had dx = Disorders of flui
	12/01/2010 05:26	12/08/2010 18:55	44	M	28.54	
	12/01/2010 19:08	12/08/2010 18:13	61	M	27.36	Patient had dx = Acute renal failu
	11/30/2010 21:50	12/08/2010 18:52	70	M	18.05	Patient had dx = Other personal
	12/08/2010 08:51	12/08/2010 18:45	68	M	16.57	stayed <1 day in the hospital
	12/03/2010 20:32	12/08/2010 17:50	80	M	16.18	Patient had dx = Disorders of flui
	12/01/2010 01:13	12/08/2010 18:06	79	M	15.52	
	12/08/2010 18:39	12/08/2010 18:39	22	F	14.53	stayed <1 day in the hospital / Av
	12/08/2010 19:01	12/08/2010 19:01	25	F	14.42	stayed <1 day in the hospital / Pa
	12/08/2010 18:05	12/08/2010 18:05	24	M	14.39	stayed <1 day in the hospital
	12/08/2010 18:26	12/08/2010 18:26	53	F	13.59	stayed <1 day in the hospital / 44

Learning from In-World Experiences

- Automation ⟷ expert handholding?
- Data differences → universal schema
- Cross-site learning & sharing



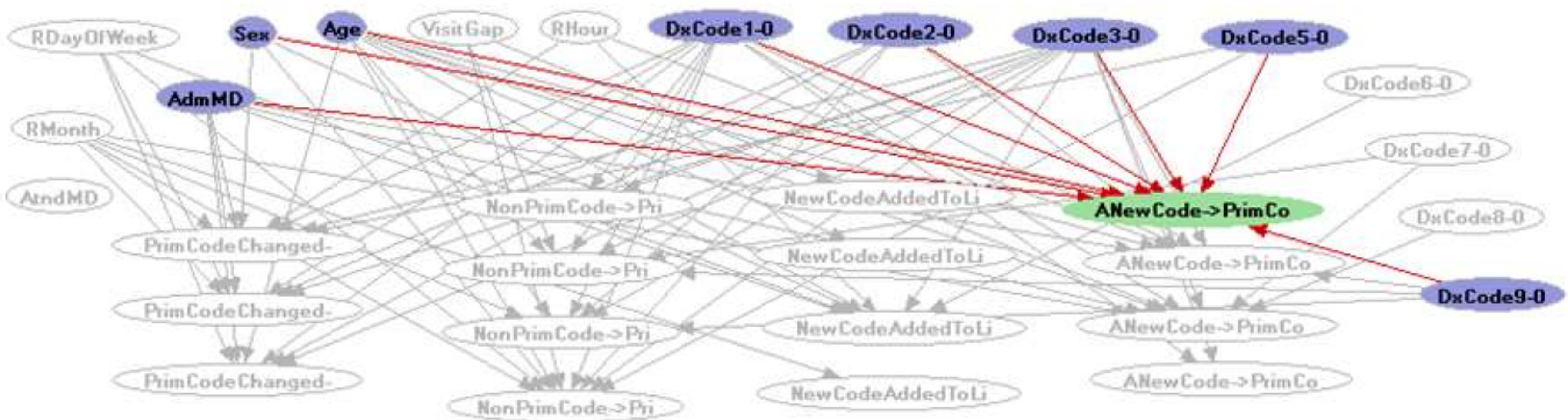
Predictions from Clinical Data

- ED discharge → Inpatient within 72 hrs, 30 days
- Inpatient discharge → Inpatient within 30 days
- CHF discharge → CHF inpatient within 30 days
- Inpatient → infection within 48hrs, 72hrs, stay
- Death within 30 days
- New kinds of models: Surprise

Surprise Model for Clinical Care

- Predict surprising outcomes

“The patient you’re discharging now will likely return within 3 days with a 1^0 dx that is not currently on the chart.”



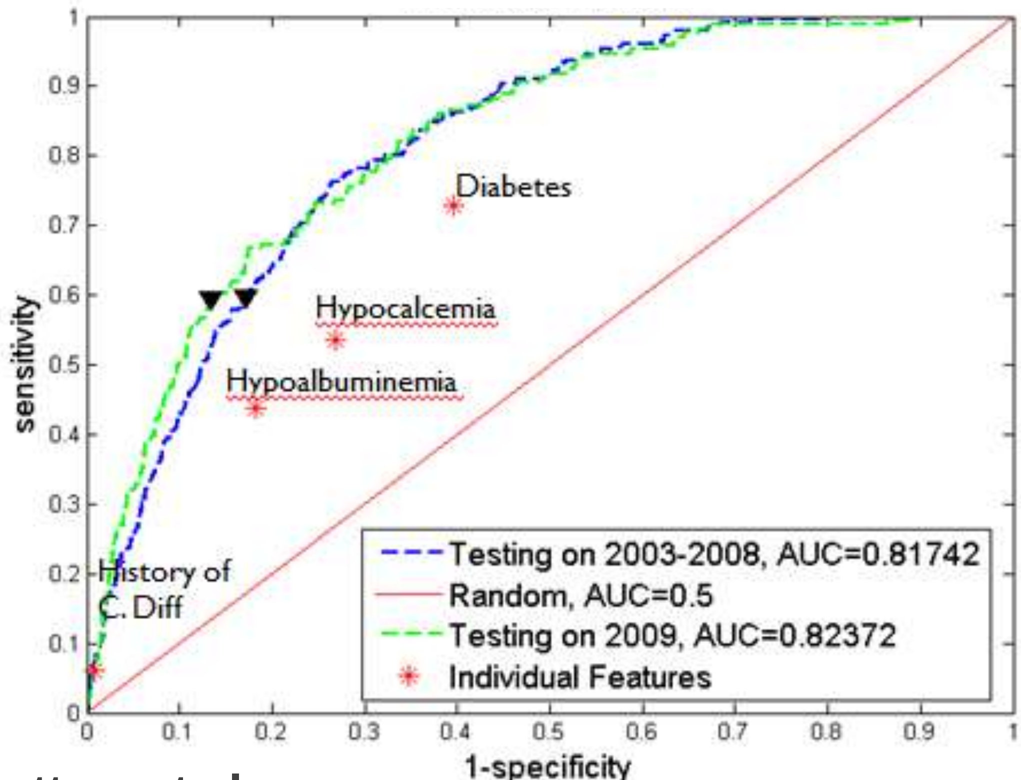
With J. Gatewood, P. Koch, M. Bayati, M. Braverman

Learning about Time and Space

Hospital-Associated Infection

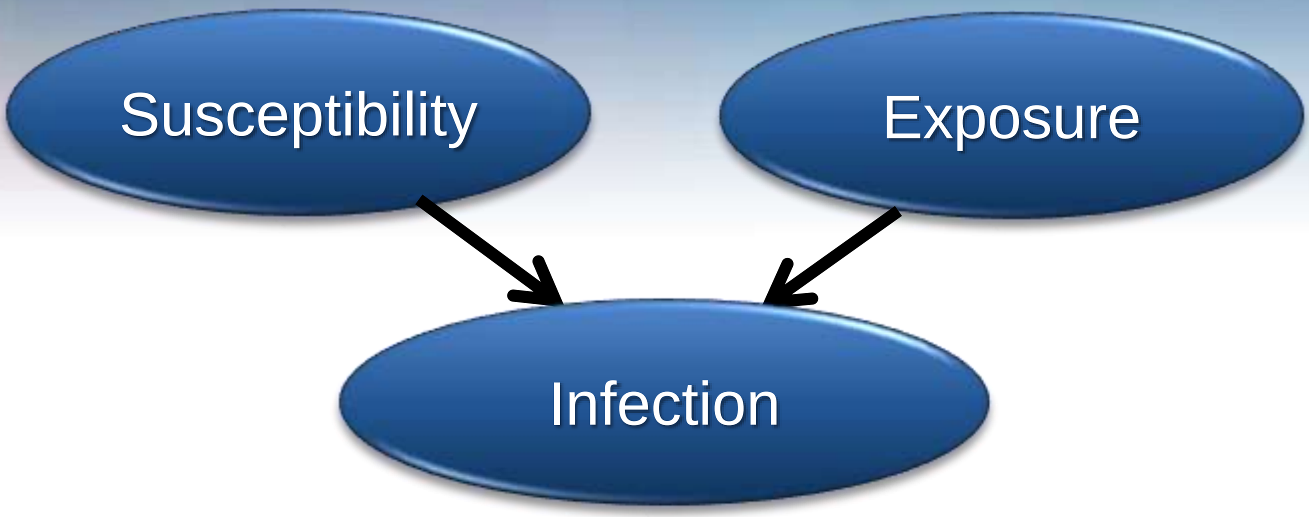
- 1 in 20 hospital visits, ~$20 billion/yr.
- 5% death (top 10 cause of death in US)

Predicting C.Difficile < 48 hrs

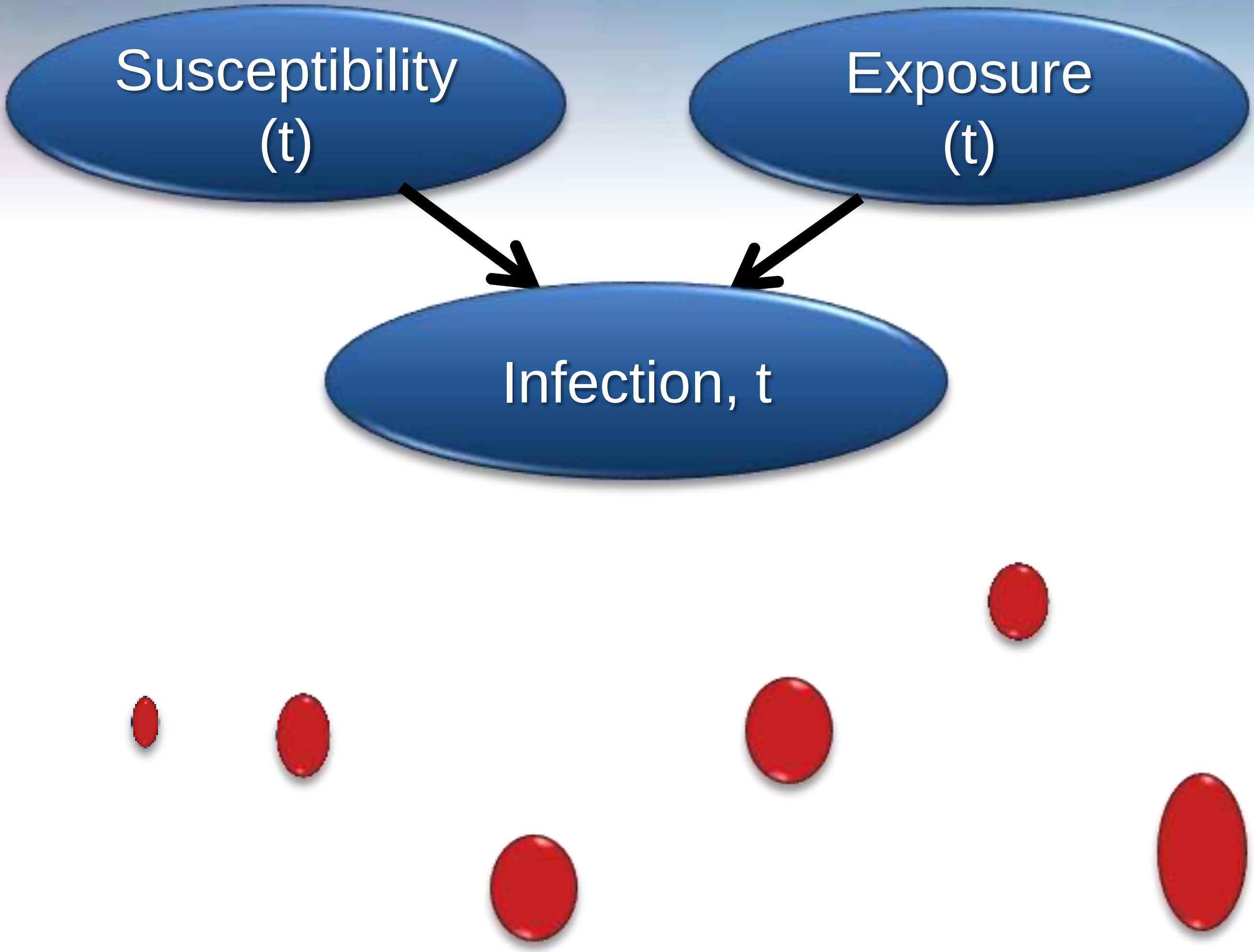


With J. Wiens, J. Guttag, et al.

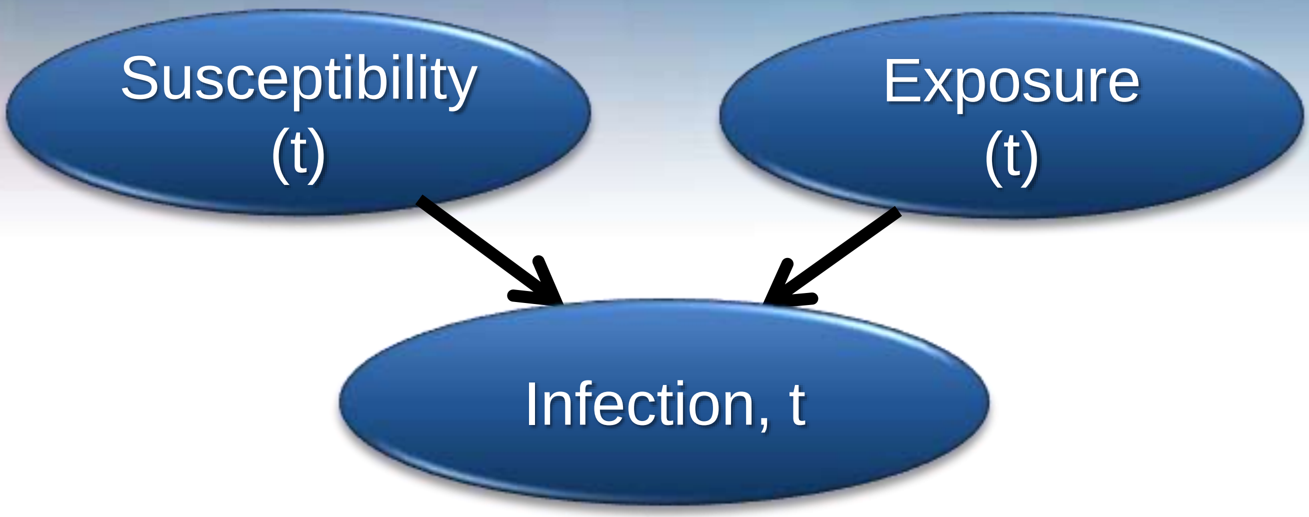
Learning about Time and Space

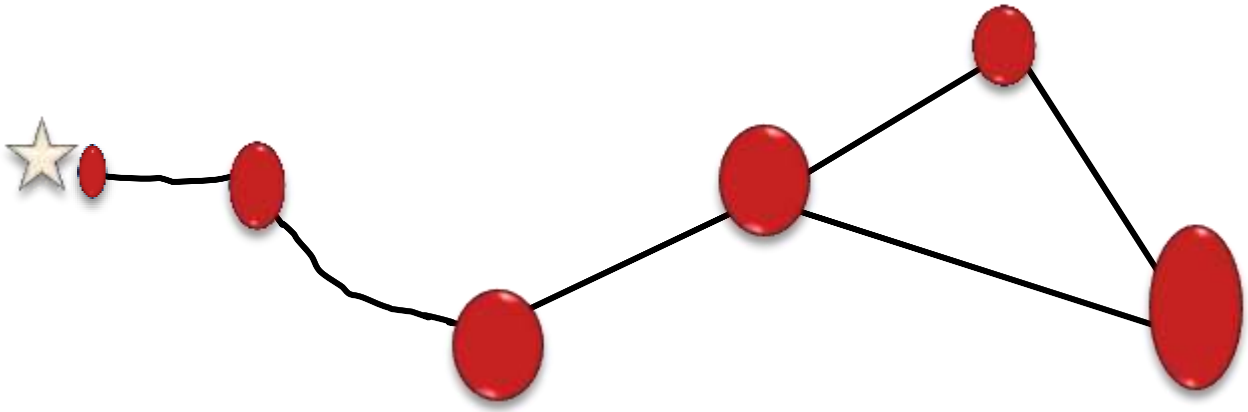


Learning about Time and Space



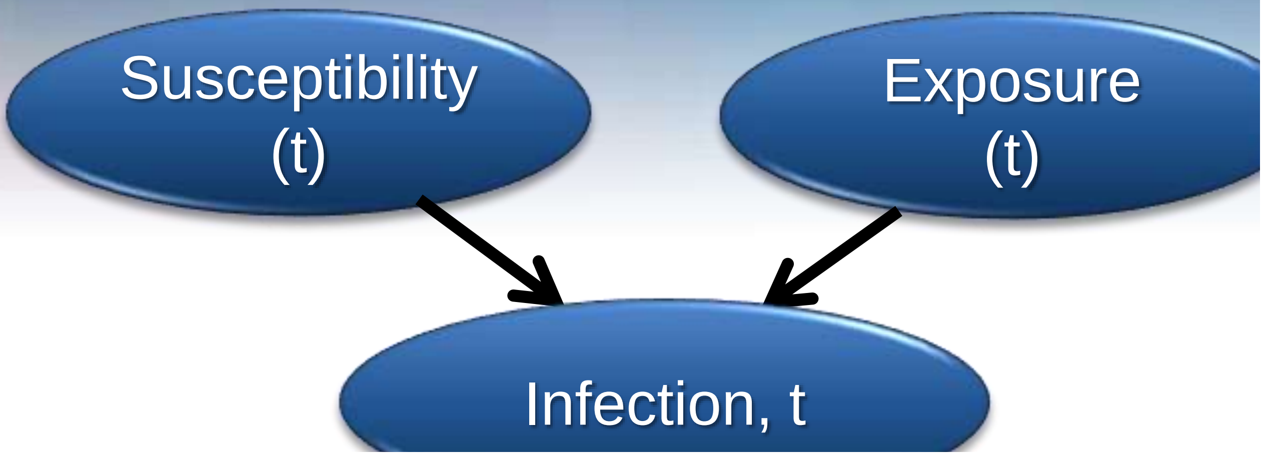
Learning about Time and Space

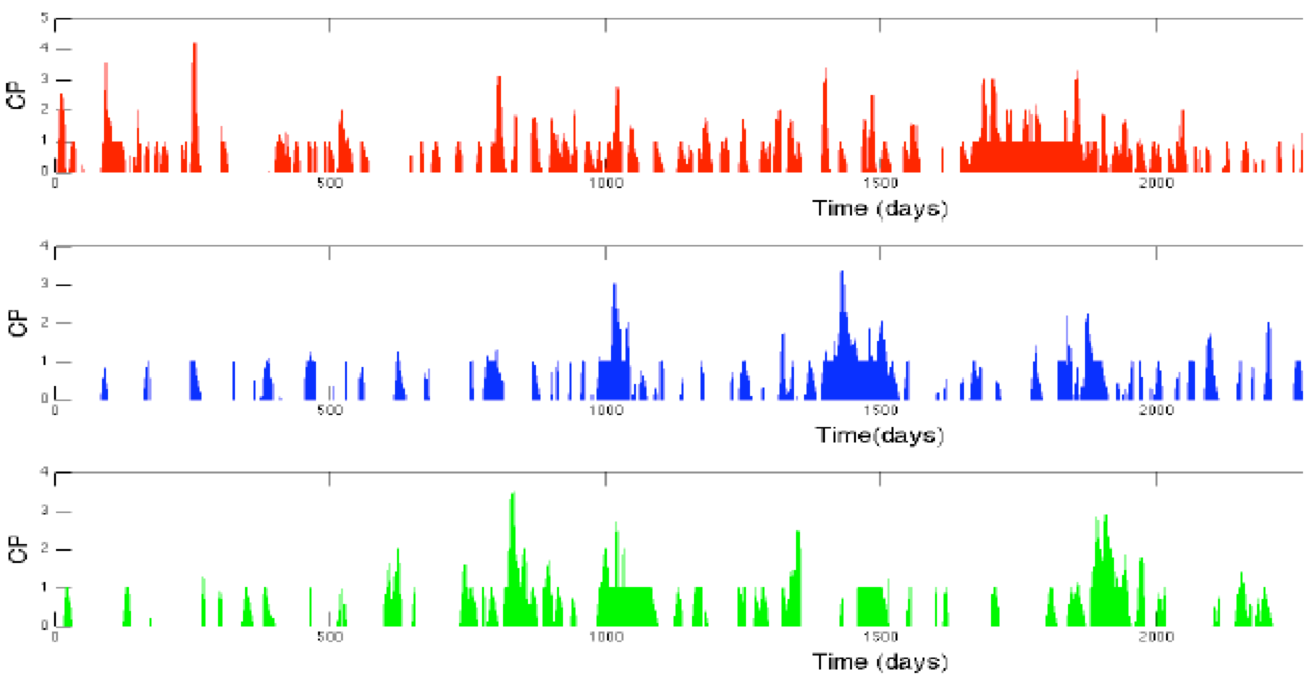




Area under curve 0.69 → .80 [time]

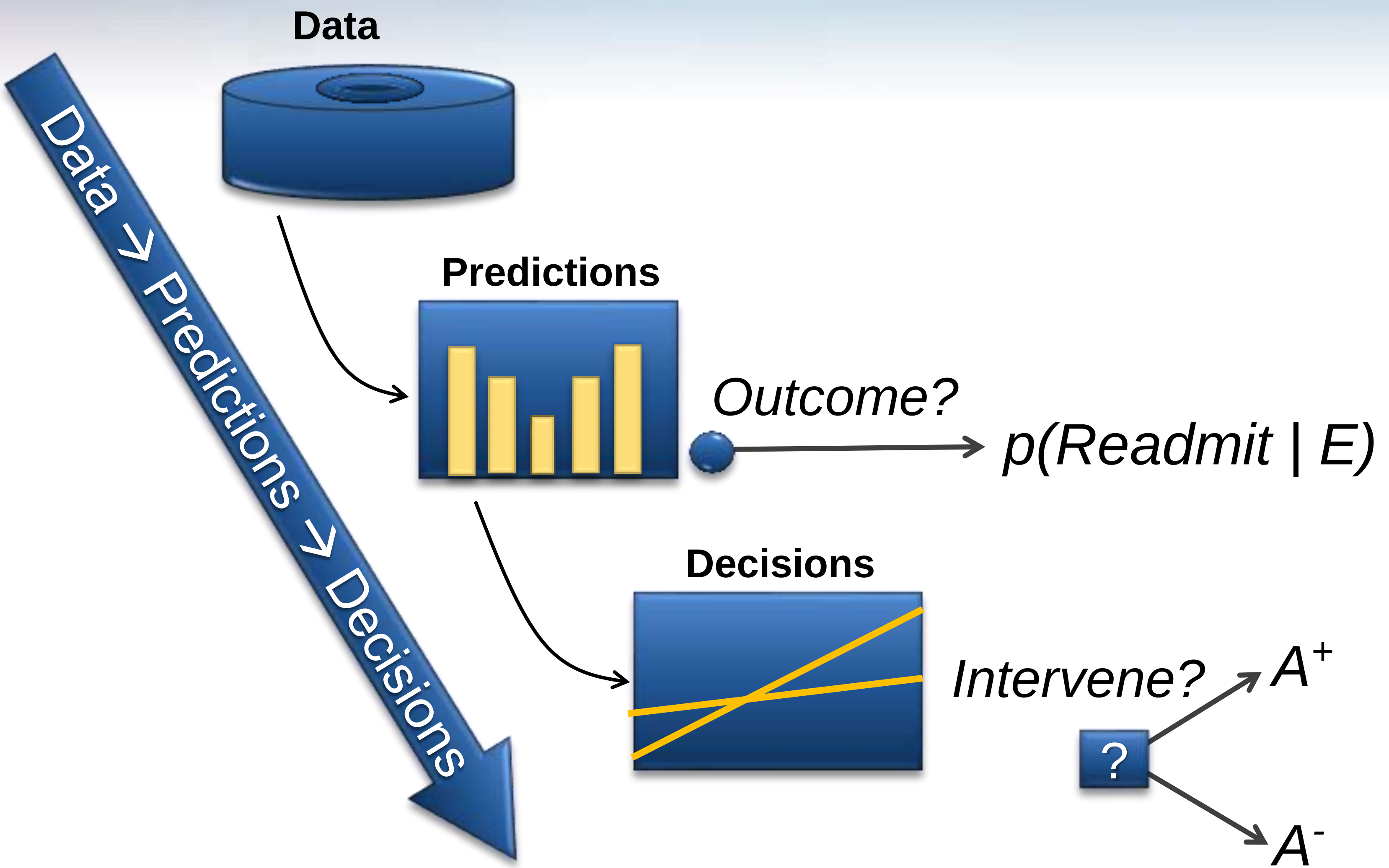
Learning about Time and Space





Location	Description
1C	Patient Care Unit
1E	Patient Care Unit
1G	MedSTAR ICU
1H	Patient Care Unit
2C	Patient Care Unit
2E	Patient Care Unit
2G	Intensive Care Unit (ICU)
2H	Patient Care Unit
2NE	Patient Care Unit
2NW	Patient Care Unit
3C	Patient Care Unit
3D	Patient Care Unit
3E	Patient Care Unit
3F	Patient Care Unit
3G	Intensive Care Unit (ICU)
3NE	Patient Care Unit
4C	Patient Care Unit
4D	Patient Care Unit
4E	Patient Care Unit
4F	Patient Care Unit
4G	Intensive Care Unit (ICU)
4H	Intensive Care Unit (ICU)
4NW	Patient Care Unit
5C	Patient Care Unit
5D	Patient Care Unit
5E	Patient Care Unit
5F	Patient Care Unit
5NE	Patient Care Unit
5NW	Patient Care Unit

Predictions → Decisions



Study: Congestive Heart Failure

- Most frequent dx for hosp. Medicare patients
  - 6–10% of folks over 65
  - $35 billion/yr US

<u>Decision:</u>

Invest in post-discharge program for patient?

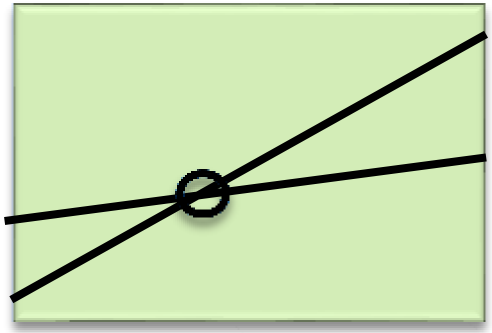
Multiple interventions proposed.

MSR: M. Bayati, M. Braverman, E. Horvitz
WHC: G. Ruiz, M. Smith, K. Mack

Study

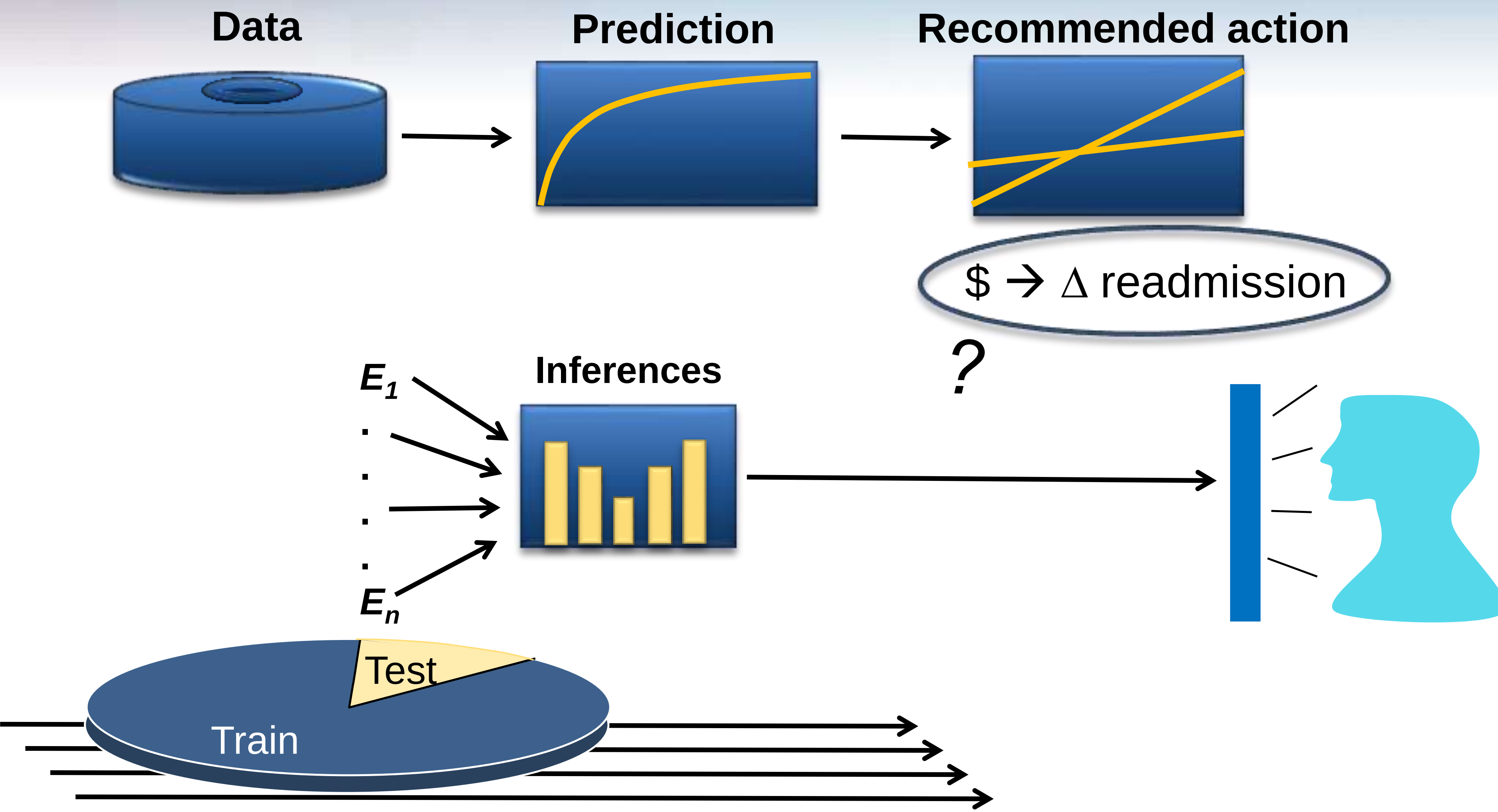
- Train: 4,485 hospitalizations for CHF, 2004-2007
- Testing: 1,319 hospitalizations for CHF, 2008
    - Mean stay: 8.4 days
    - Mean cost: $18,435

Decision model: *Probability threshold on predicted likelihood of readmissions for enlistment in special program.*



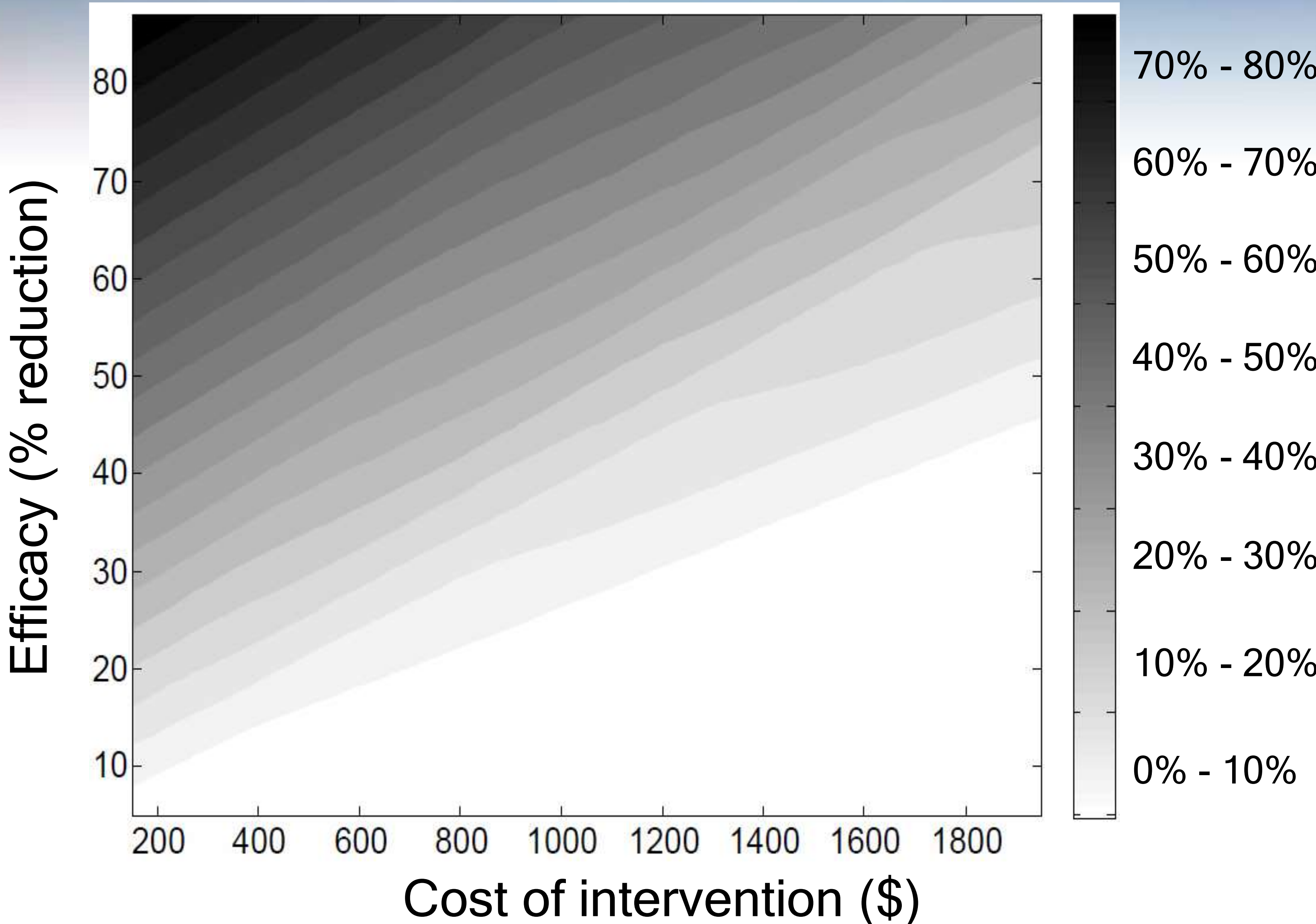
Expected Value of Decision System

- Probe the expected value of fielding a system(!)



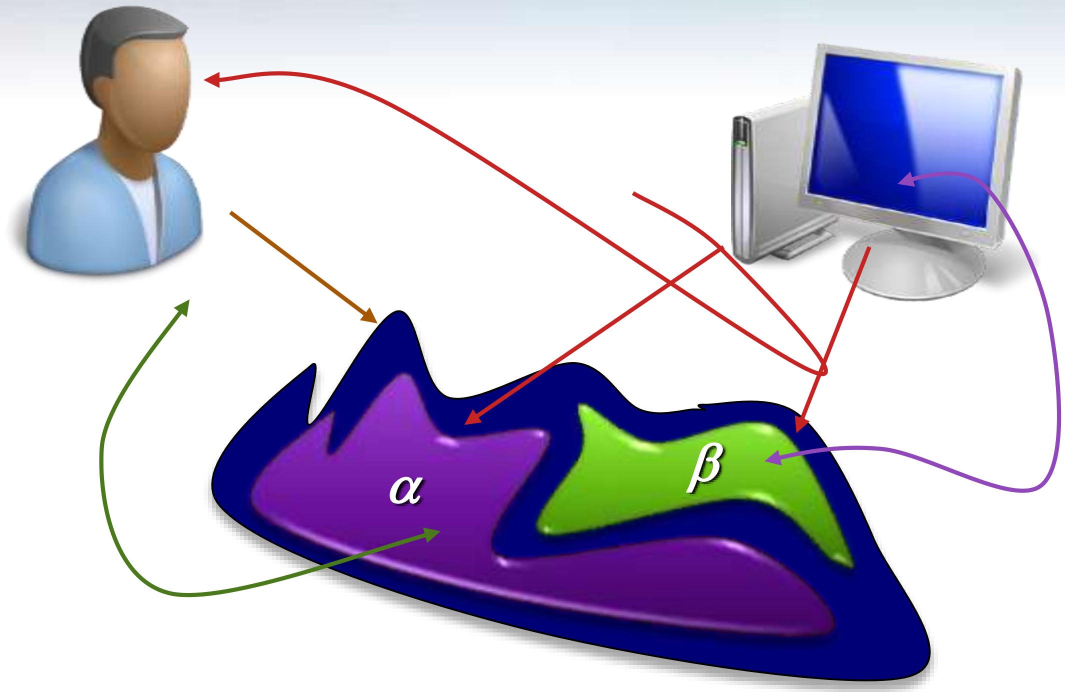
Expected Value of Fielding System

Prediction-centric action: train: 2004-2007, test: 2008



Focus: Complementary Computing

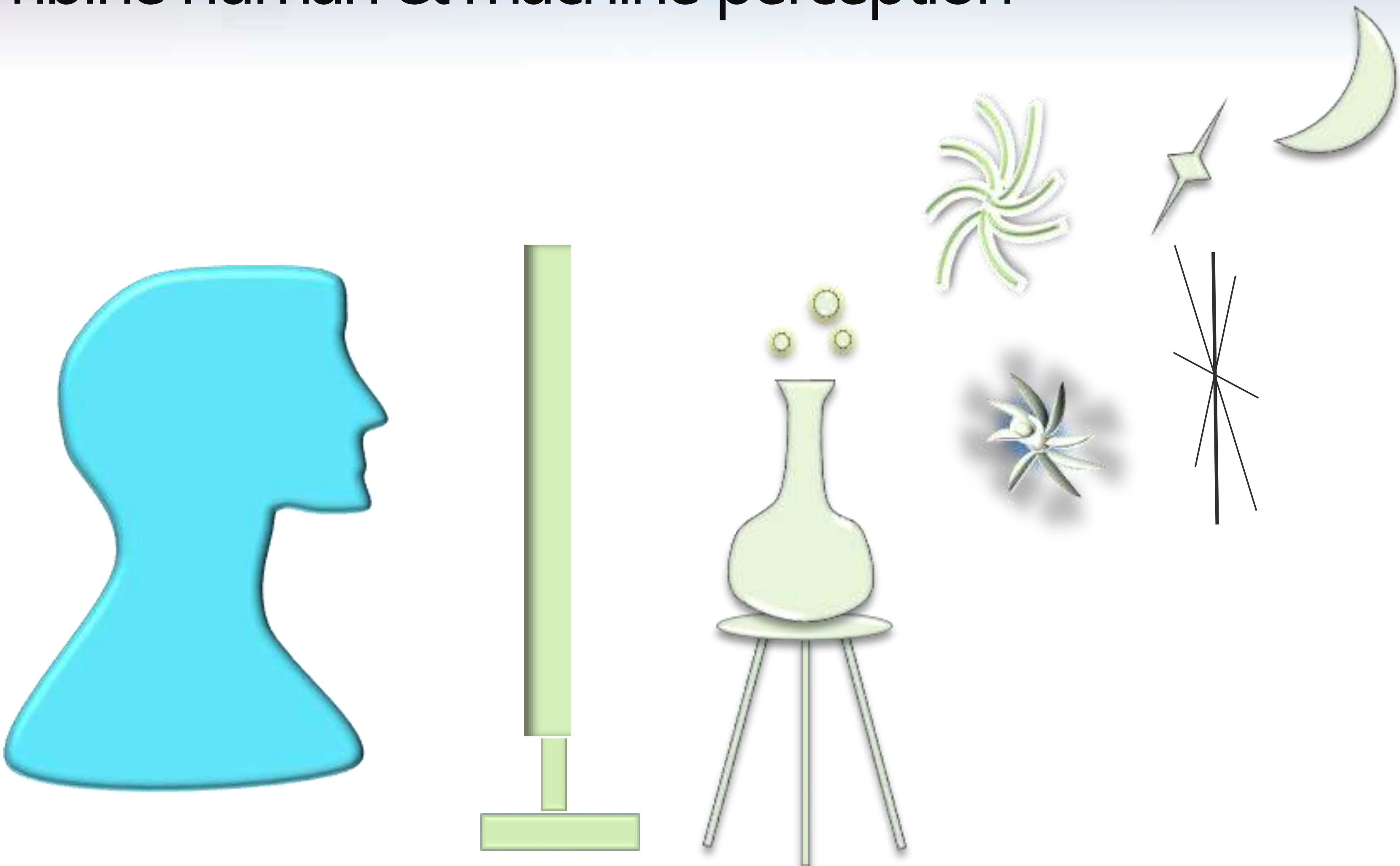
On vision of human-computer symbiosis



Machine learning & inference to leverage contributions from machine & human

Study: Citizen Science

Apply machine learning and decision making to combine human & machine perception



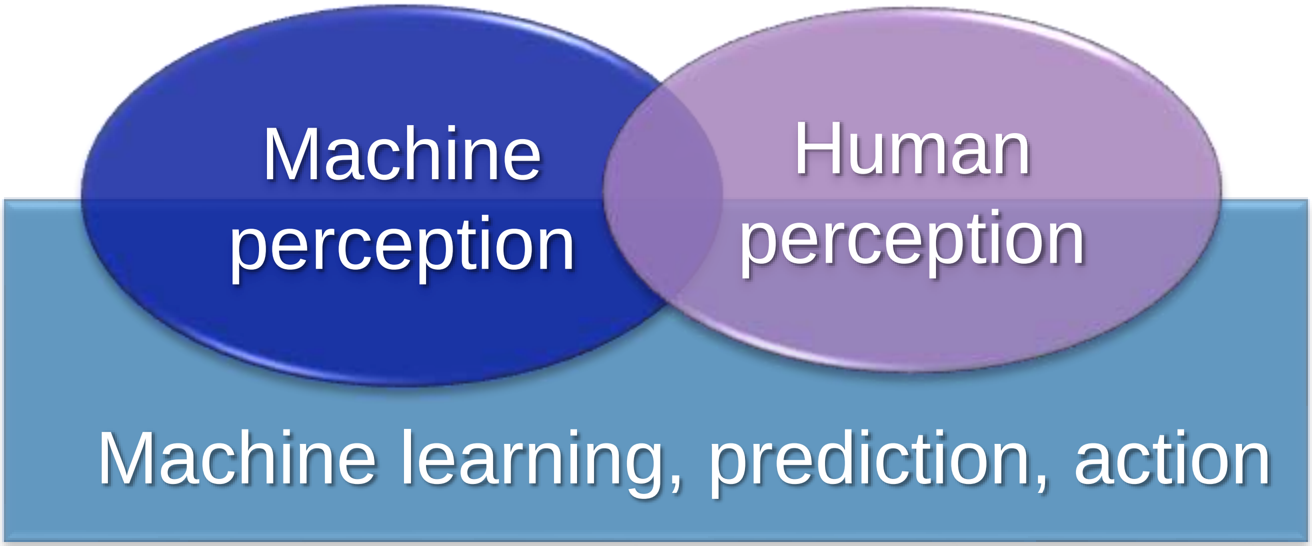
Galaxy Zoo: Tag Galaxies

- Sloan Digital Sky Survey:
  ~10^6 galaxies, ~120k quasars, ~225k stars

Volunteer DB: 886k galaxies, 34m votes, 100k, people

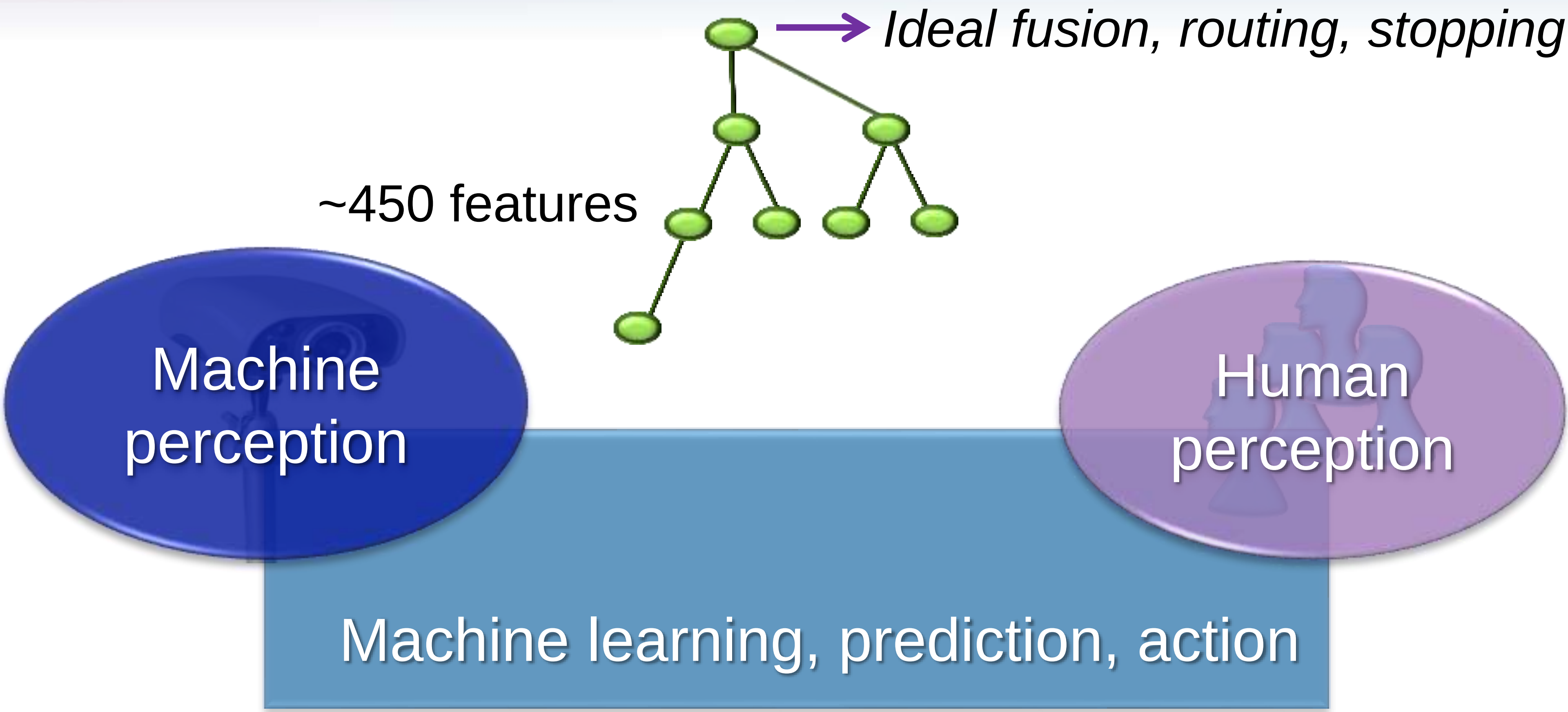
CrowdSynth

- Machine learning for fusion & task routing
- Learn from machine vision & votes



CrowdSynth

- Machine learning for fusion & task routing
- Learn from machine vision & votes



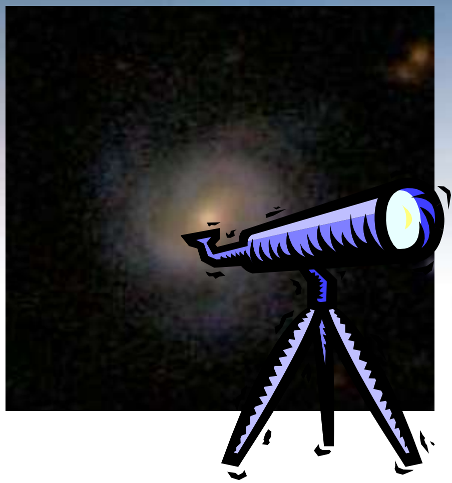
E. Kamar, S. Hacker, P. Koch, C. Lintott, H.

Sloan Digital Sky Survey

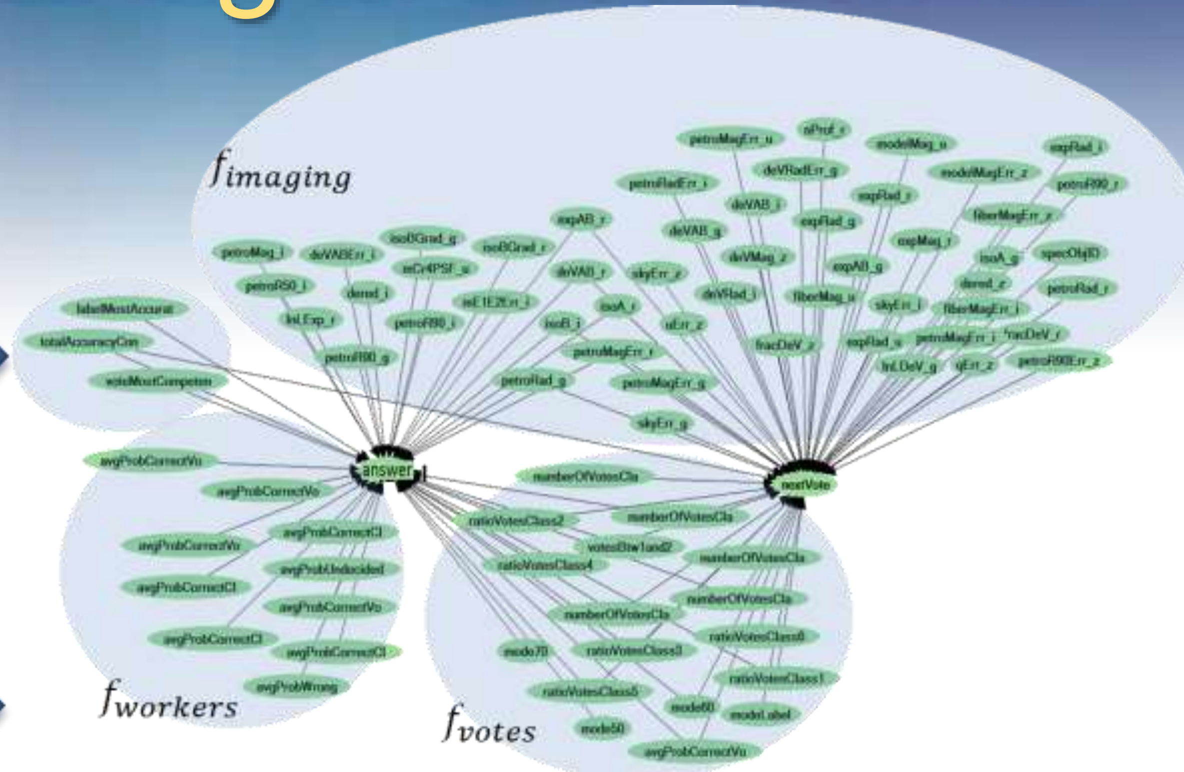
453 features

Attribute	Description
$petroMag_{ug}$	Petrosian magnitude colors. A color was calculated for four independent pairs of bands in SDSS (u, g, r, i, z).
$petroRad_u * z$	Petrosian radius, transformed with redshift to be distance-independent.
$invConIndx_u$	Inverse concentration index. The ratio of the 50% Petrosian magnitude to the 90% Petrosian magnitude.
$isoRowcGrad\_u * z$	Gradient of the isophotal row centroid, transformed with redshift to be distance-independent.
$isoColcGrad_u * z$	Gradient of the isophotal column centroid, transformed with redshift to be distance-independent.
$isoA_u * z$	Isophotal major axis, transformed with redshift to be distance-independent.
$isoB_u * z$	Isophotal minor axis, transformed with redshift to be distance-independent.
$isoAGrad_u * z$	Gradient of the isophotal major axis, transformed with redshift to be distance-independent.
$isoBGrad_u * z$	Gradient of the isophotal minor axis, transformed with redshift to be distance-independent.
$isoPhiGrad_u * z$	Gradient of the isophotal orientation, transformed with redshift to be distance-independent.
$texture_u$	Measurement of surface texture.
$lnLExp_u$	Log-likelihood of exponential profile fit.
$lnLDeV_u$	Log-likelihood of De Vaucouleurs profile fit.
$fracDev_u$	Fraction of the brightness profile explained by the De Vaucouleurs profile.

Machine Learning for Prediction



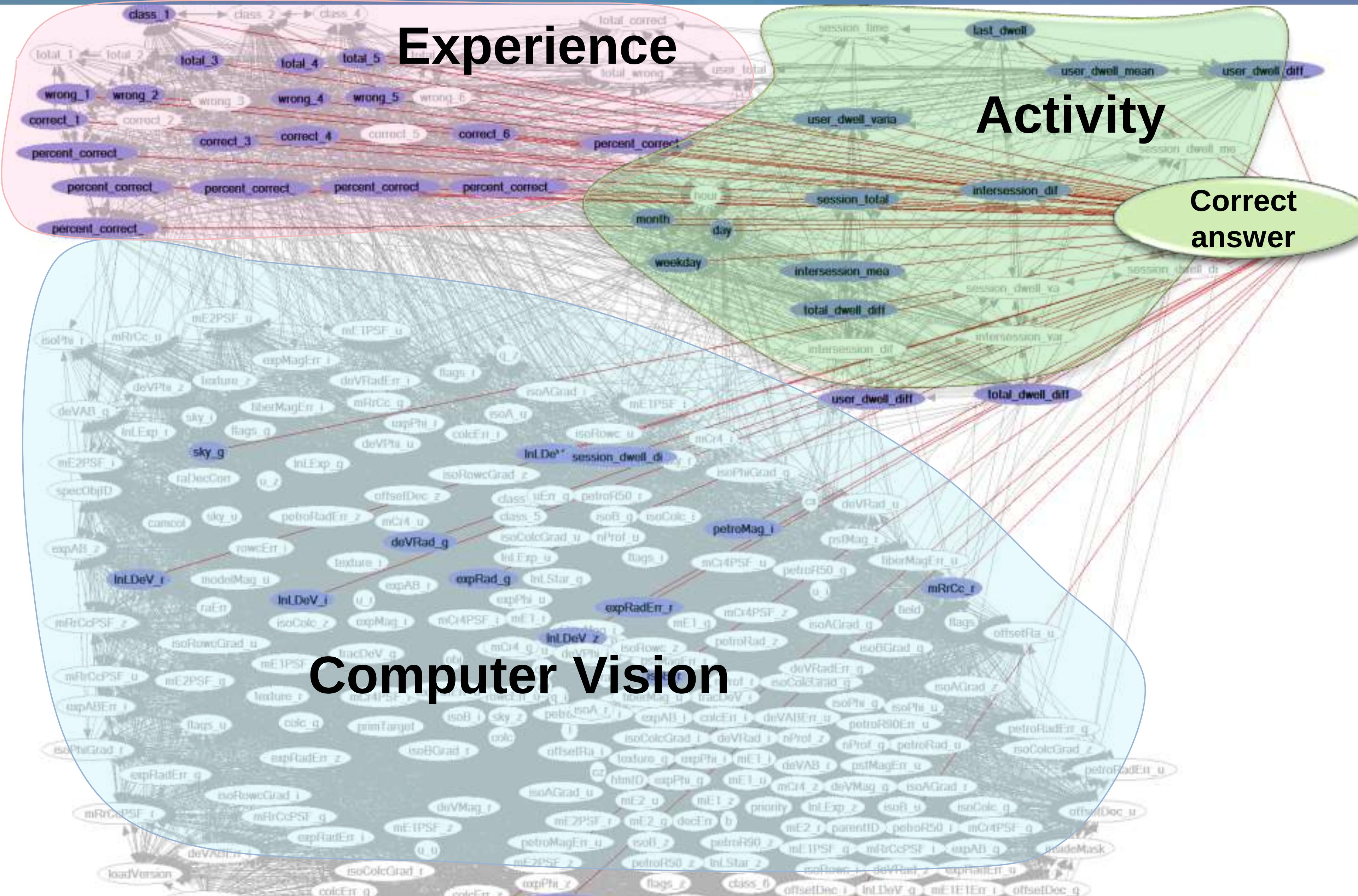
Task features



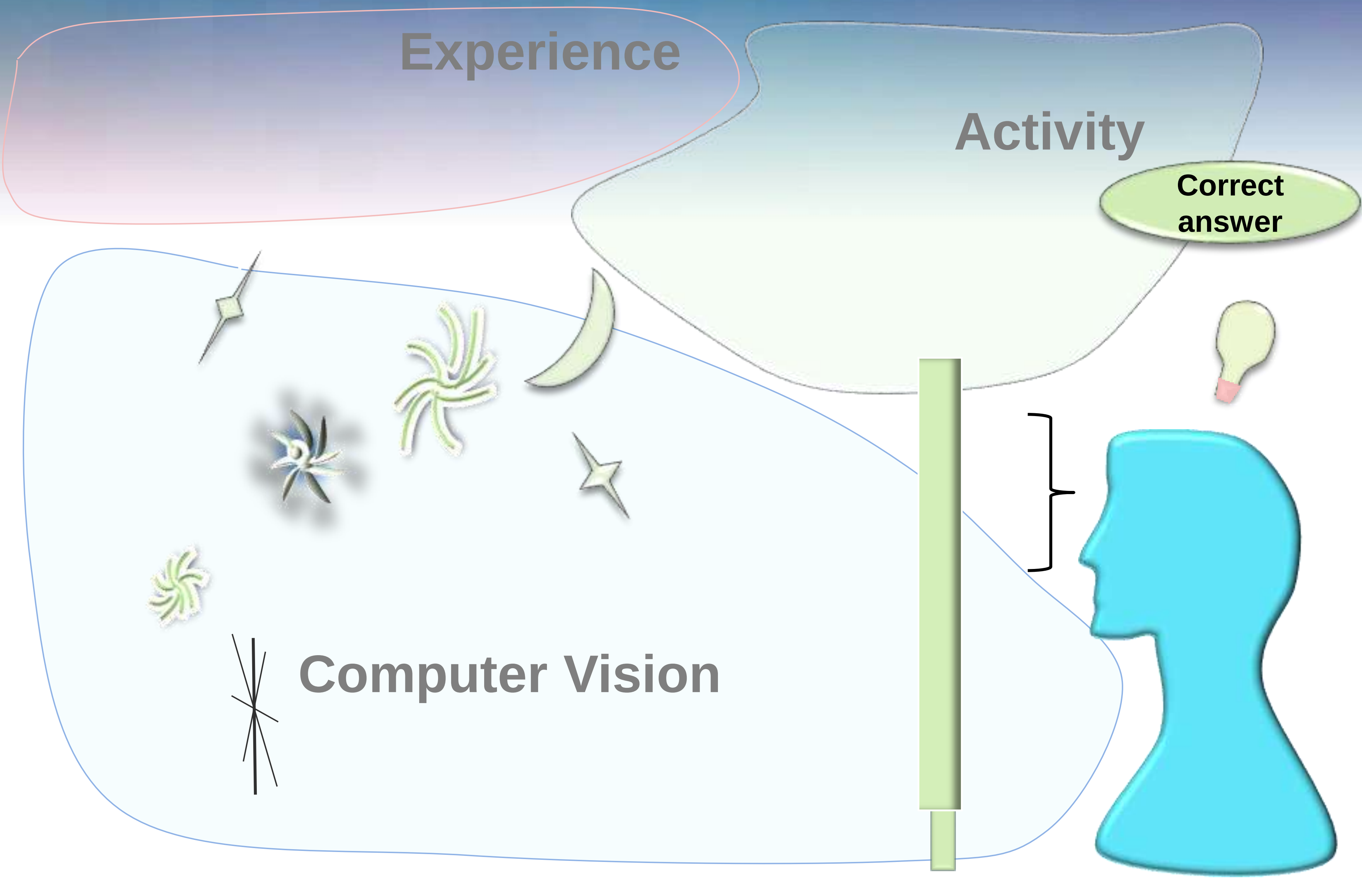
Vote features

- **Object model** assesses likelihood of world state
- **Vote model** predicts worker assessments

Learn Rich Models of Individuals' Abilities

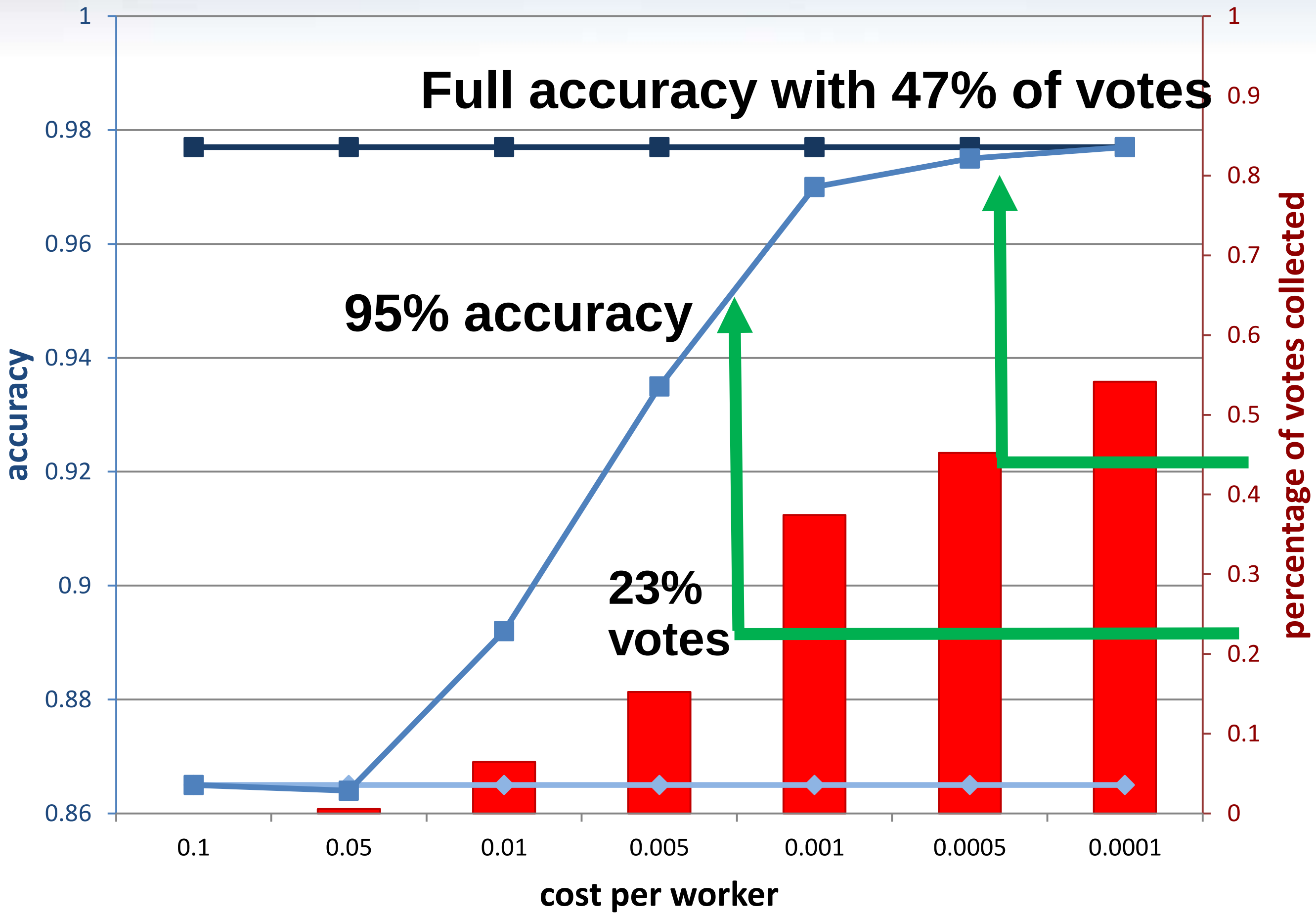


Learn Rich Models of Individuals' Abilities

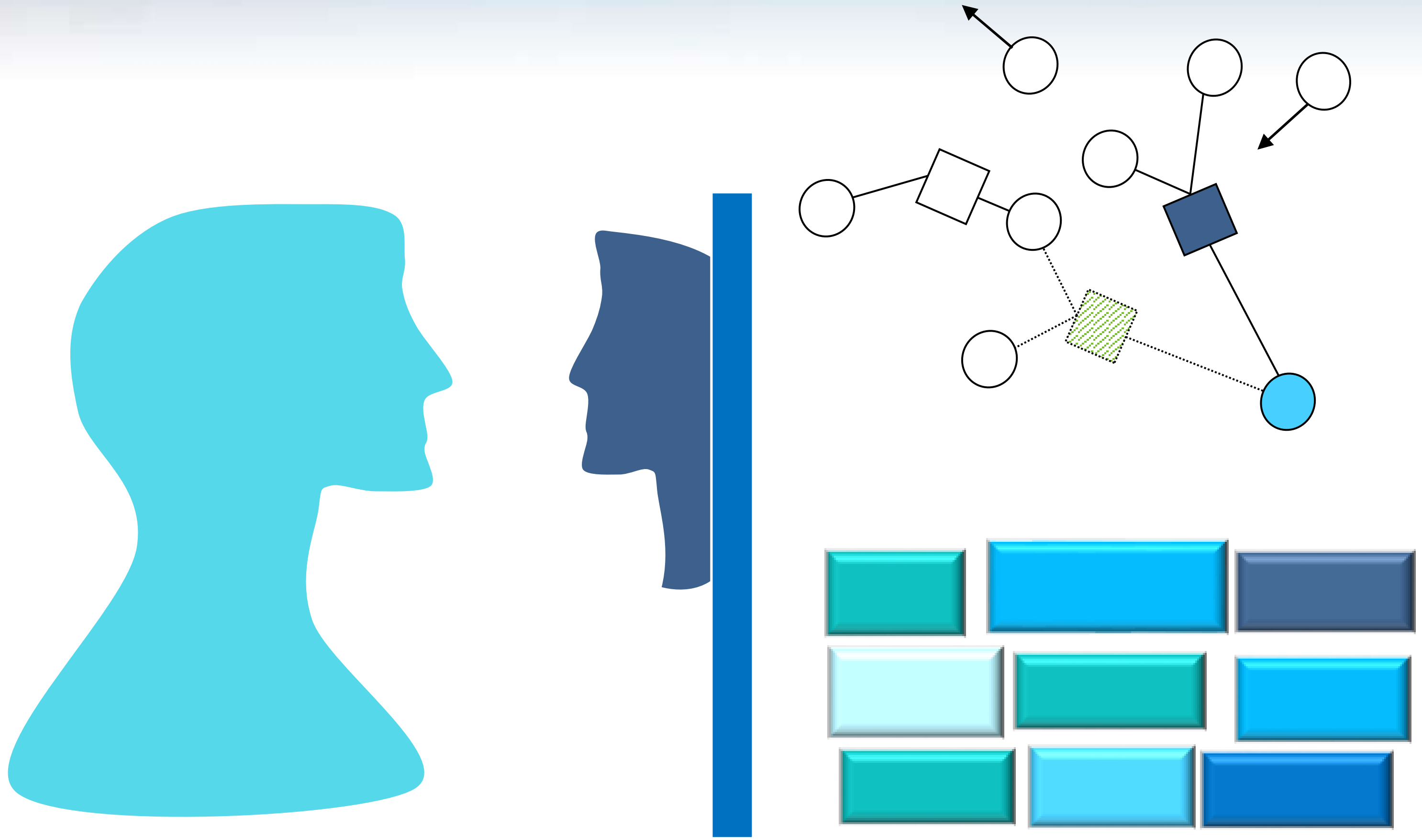


Power of Complementary Computing

- New efficiencies, stopping criteria



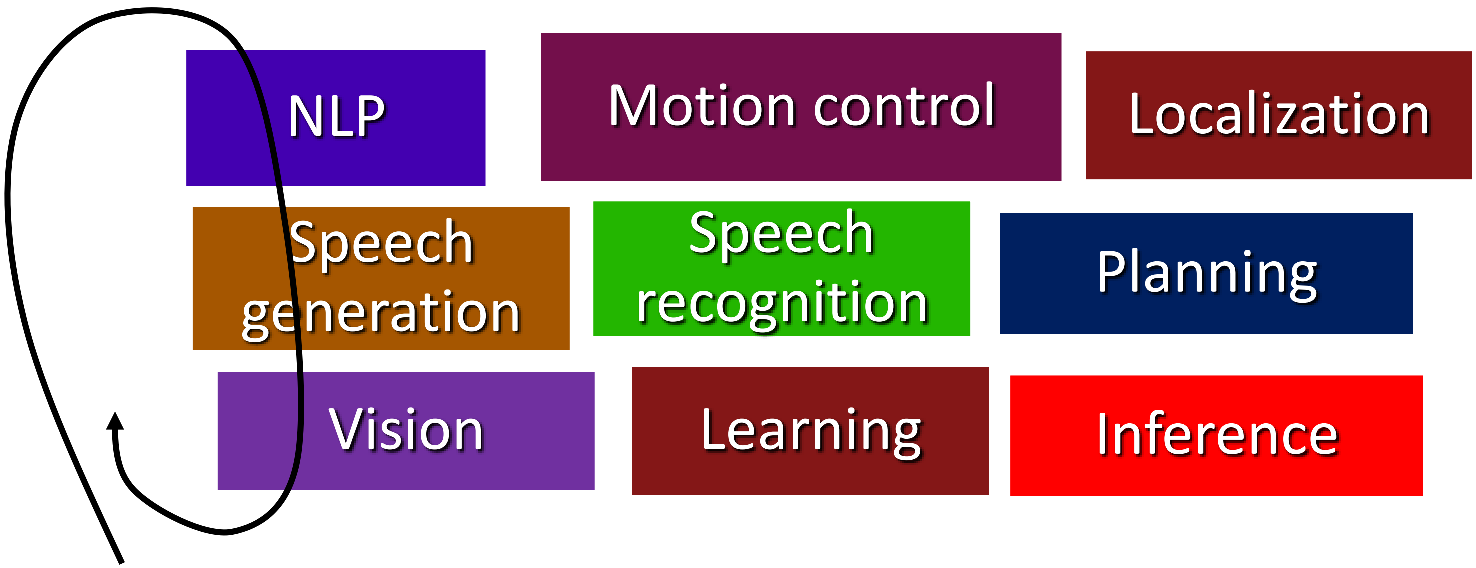
Focus: Integrative Intelligence



With Dan Bohus, Ece Kamar, Paul Koch, Anne Loomis Thompson

Intelligence via Composition

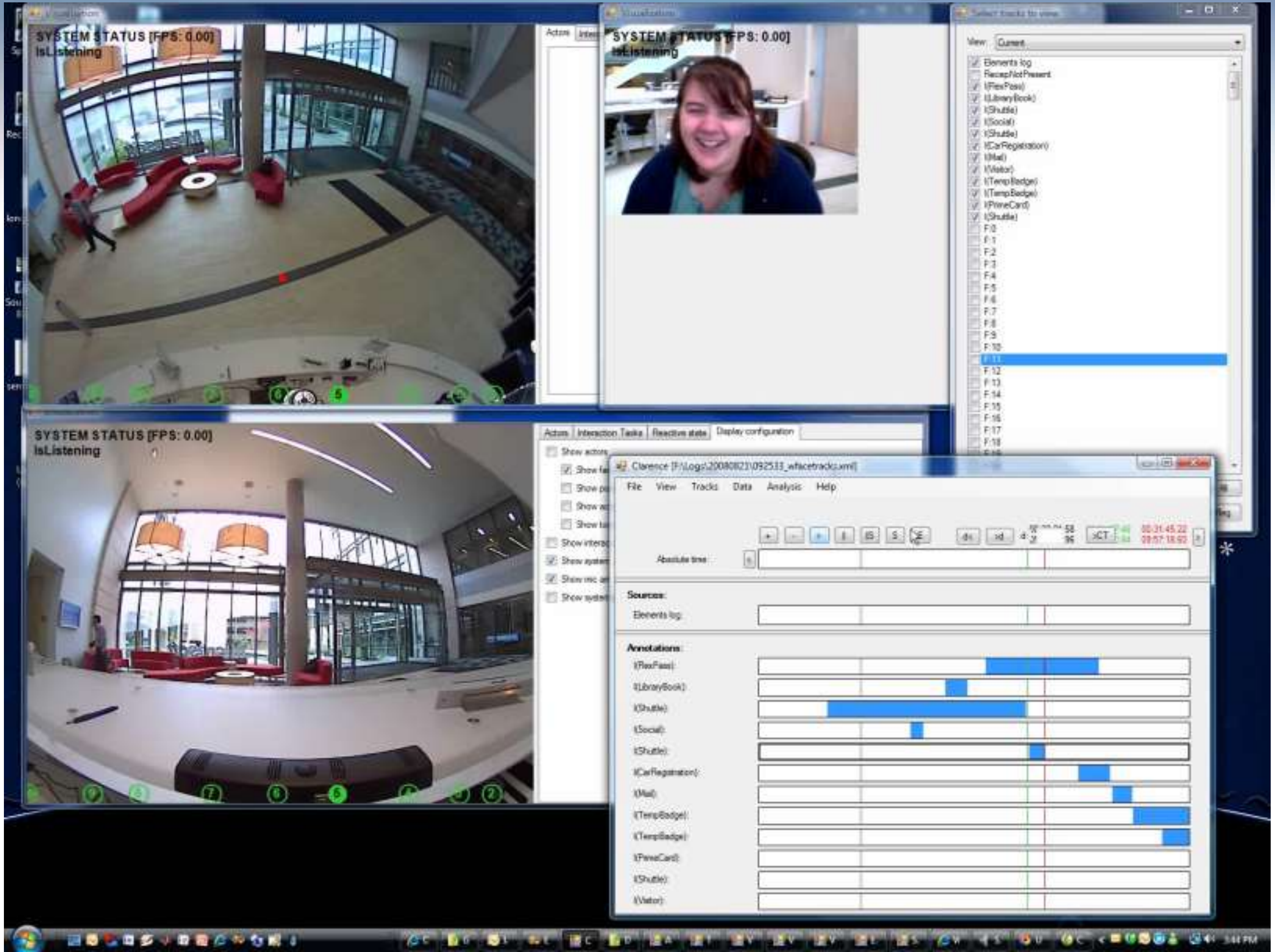
- Leveraging tapestry of components
- Understanding synergies & dependencies
- Whole more than sum?



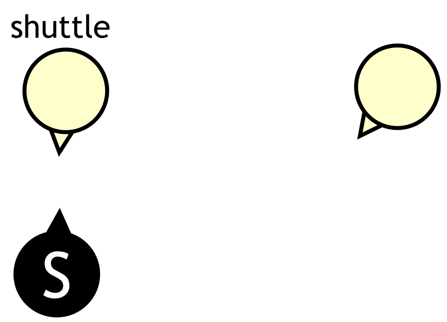
$$\text{Whole} \gg \textstyle\sum_i \text{part}_i\ ?$$

Situated Interaction

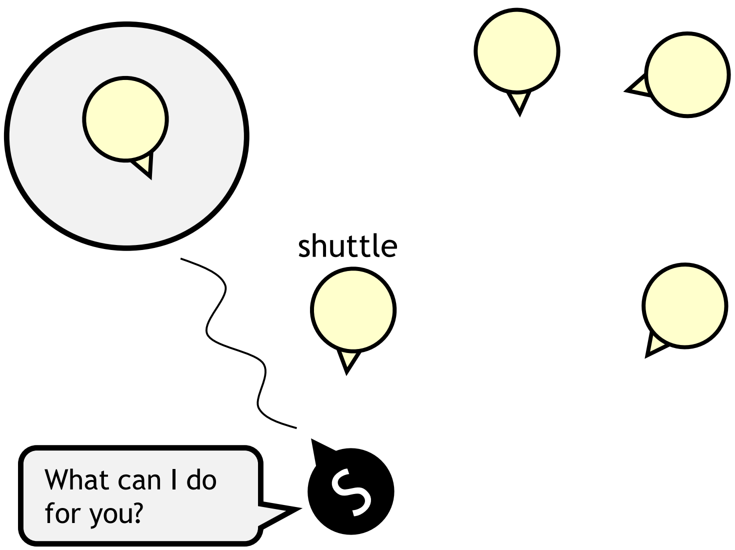
Situated Interaction Project



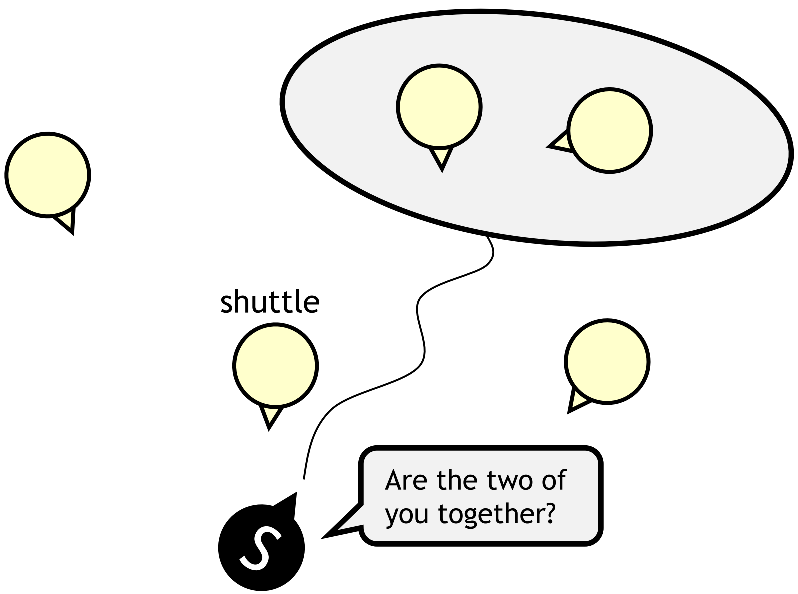
Models of Multiparty Collaboration



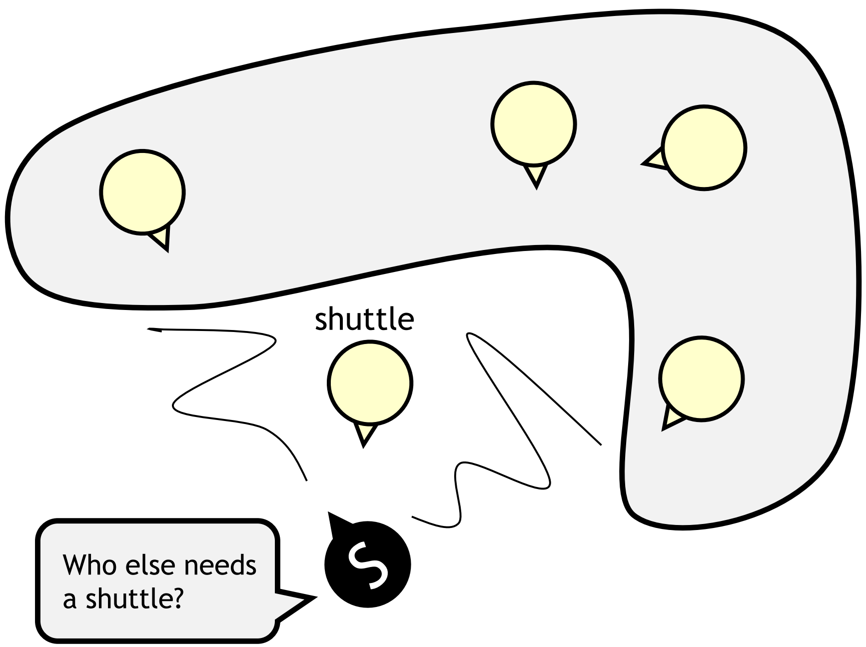
Models of Multiparty Collaboration



Models of Multiparty Collaboration



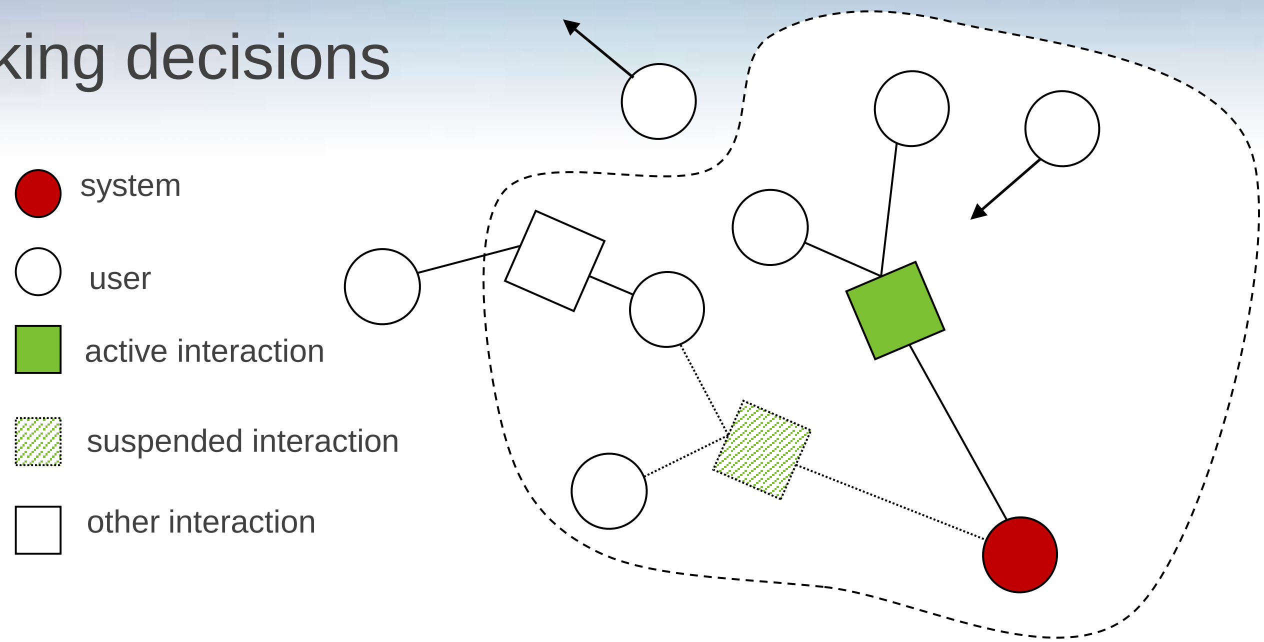
Models of Multiparty Collaboration

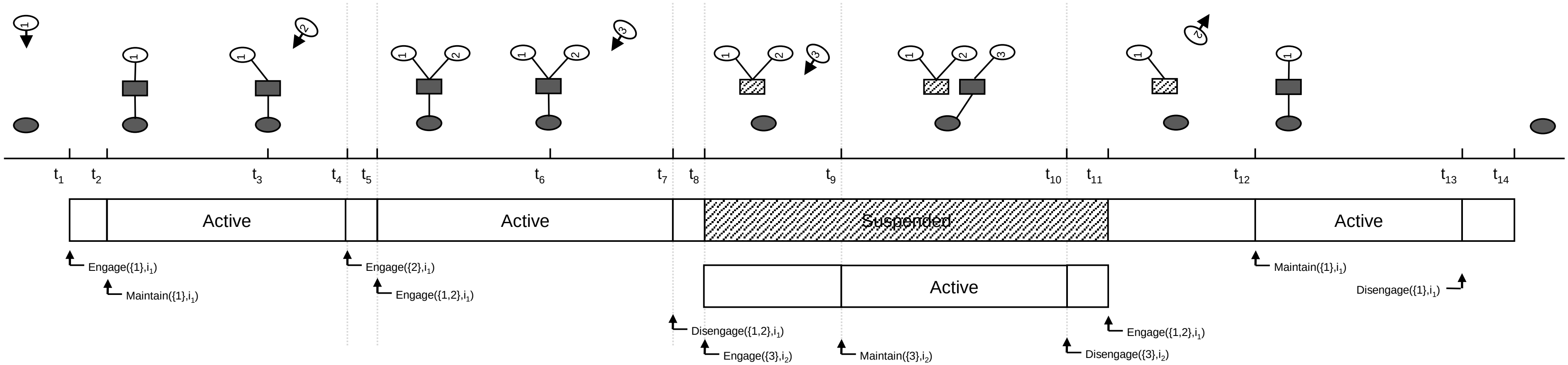


Contributions & Turns in the Open World

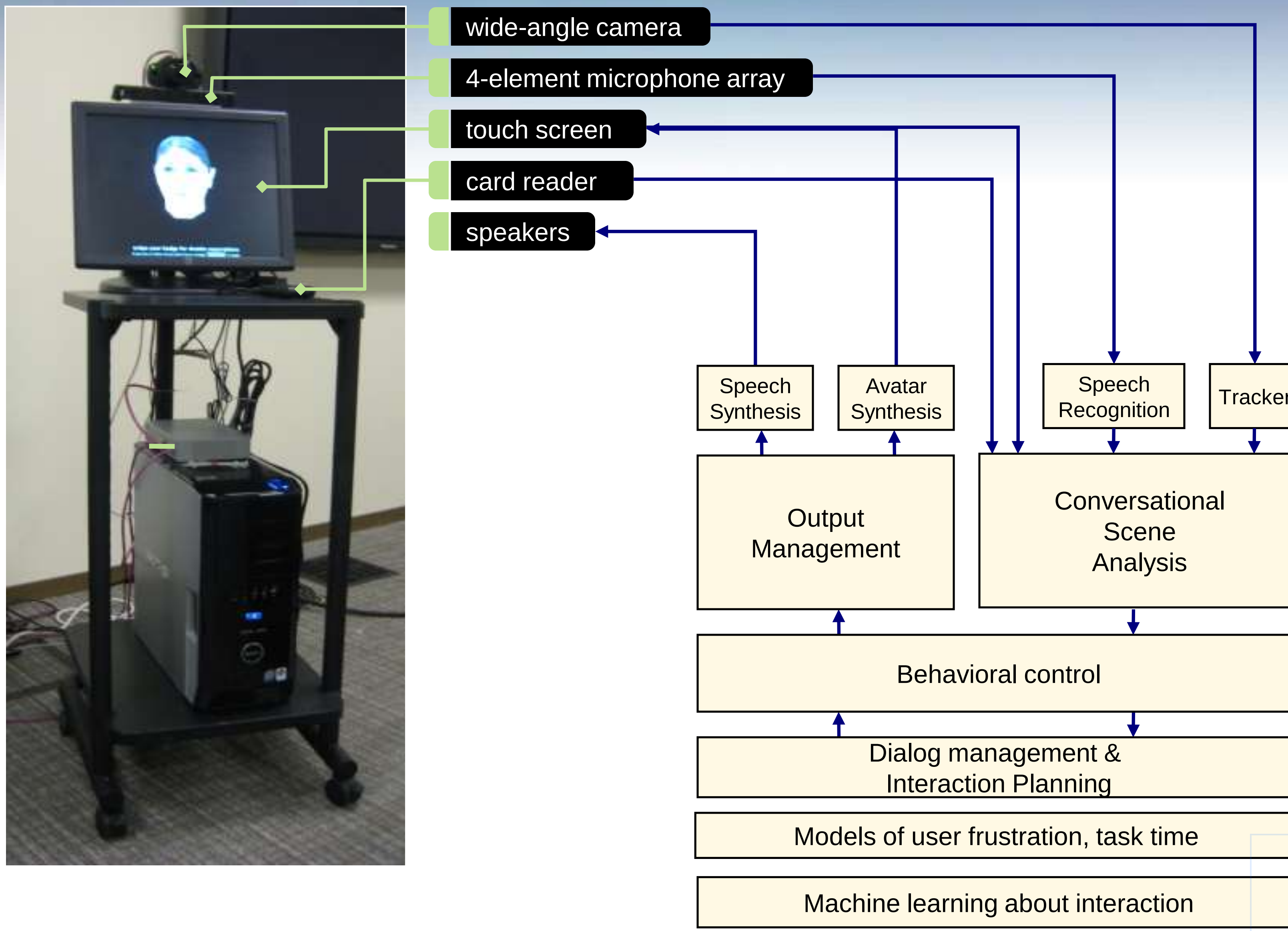
Track conversational dynamics

Make turn-taking decisions





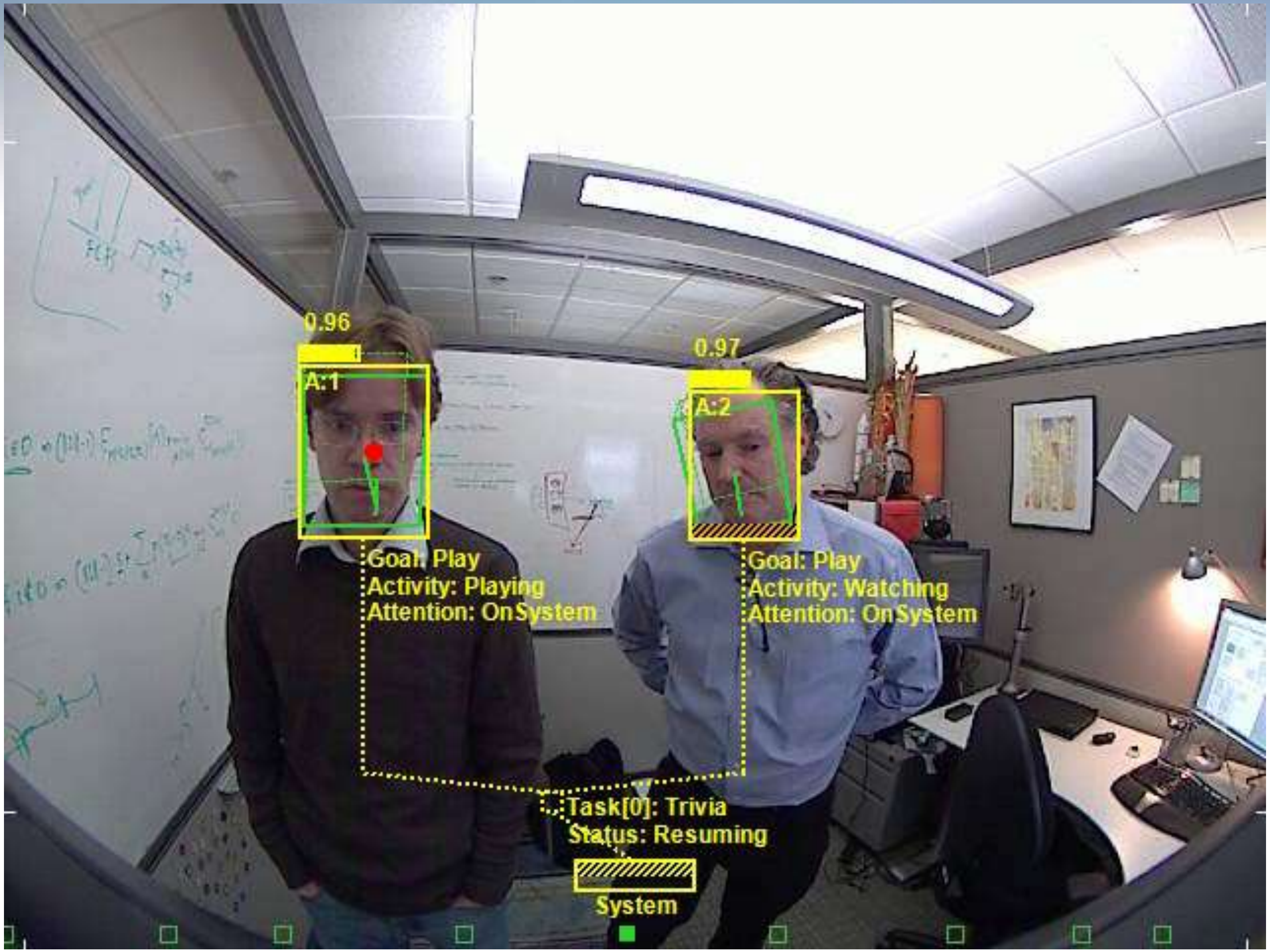
Composing a Platform



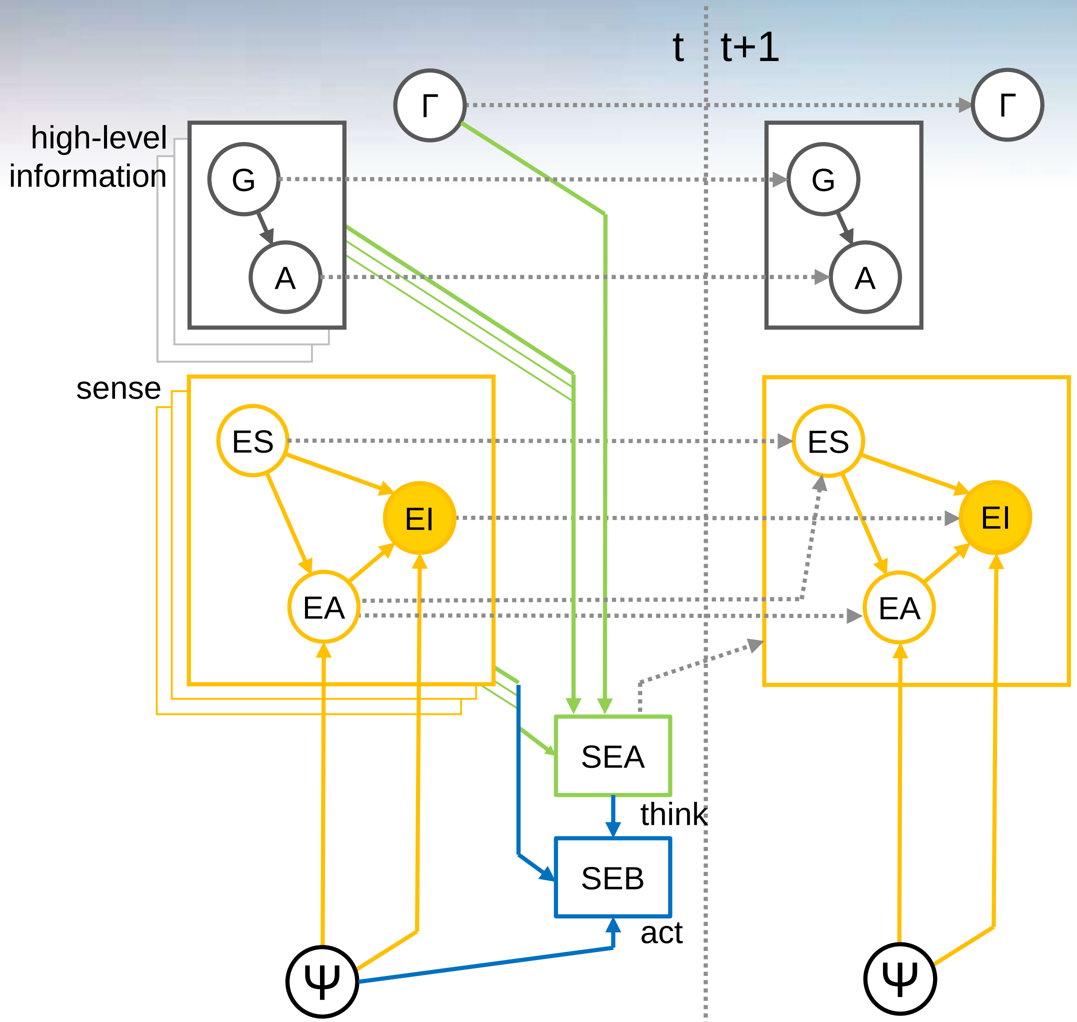
Multiple Tasks and Participants

Multiple Tasks and Participants

Multiparty Collaboration & Turn Taking



Representation and Inference



Ψ = all sensory evidence

ES = engagement state
{ engaged, not-engaged }

EA = engagement action
{ maintain, disengage, engage, no-action }

EI = engagement intention
{ engaged, not-engaged }

SEA = system engagement action
{ maintain, disengage, engage, no-action }

SEA = π(ES, EA, EI, G, A, Γ)

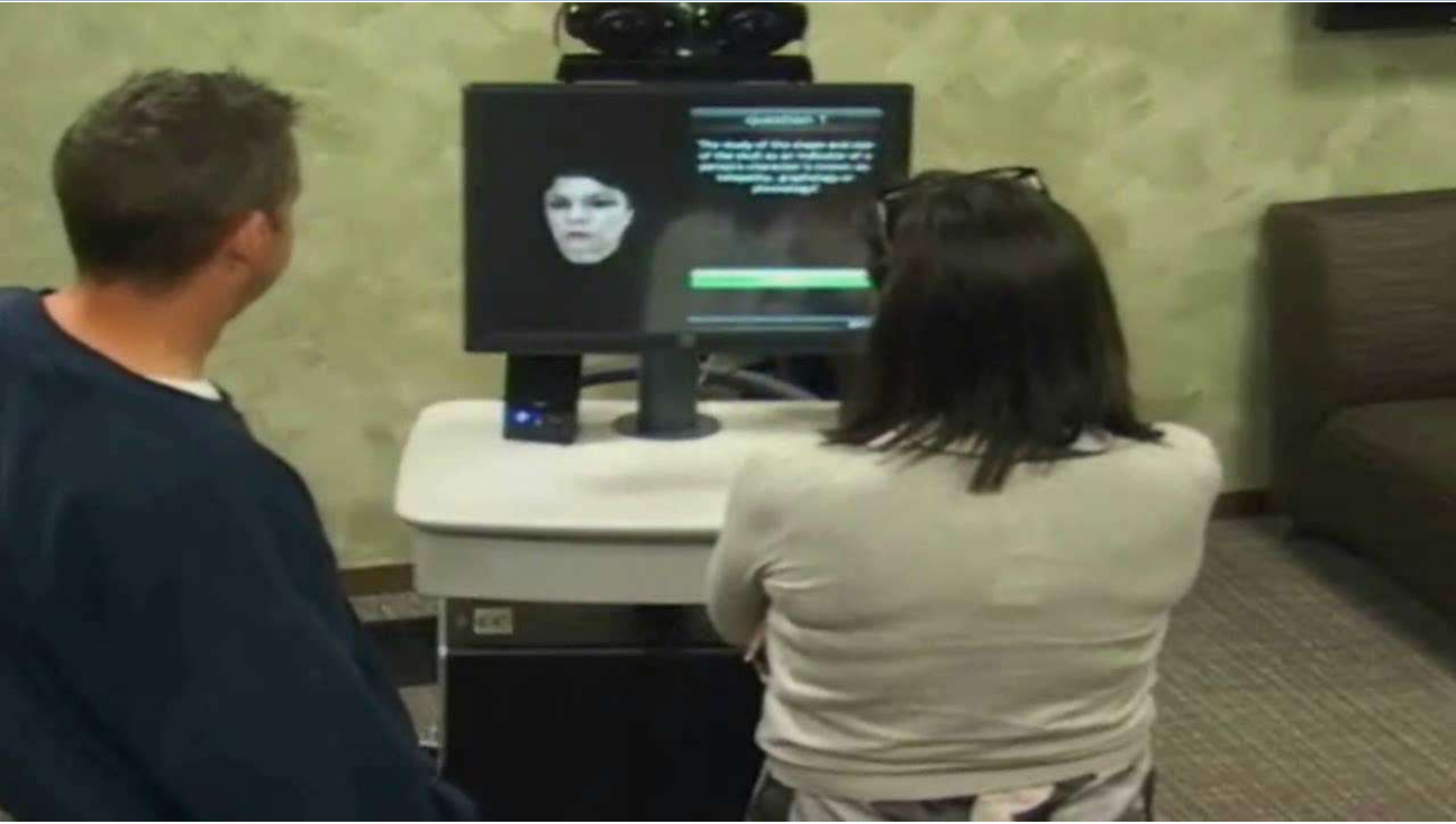
G = high-level goal
{ shuttle, register, other }

A = high-level activity
{ interacting, waiting-for-receptionist waiting-for-other, passing-by }

Γ = grouping information

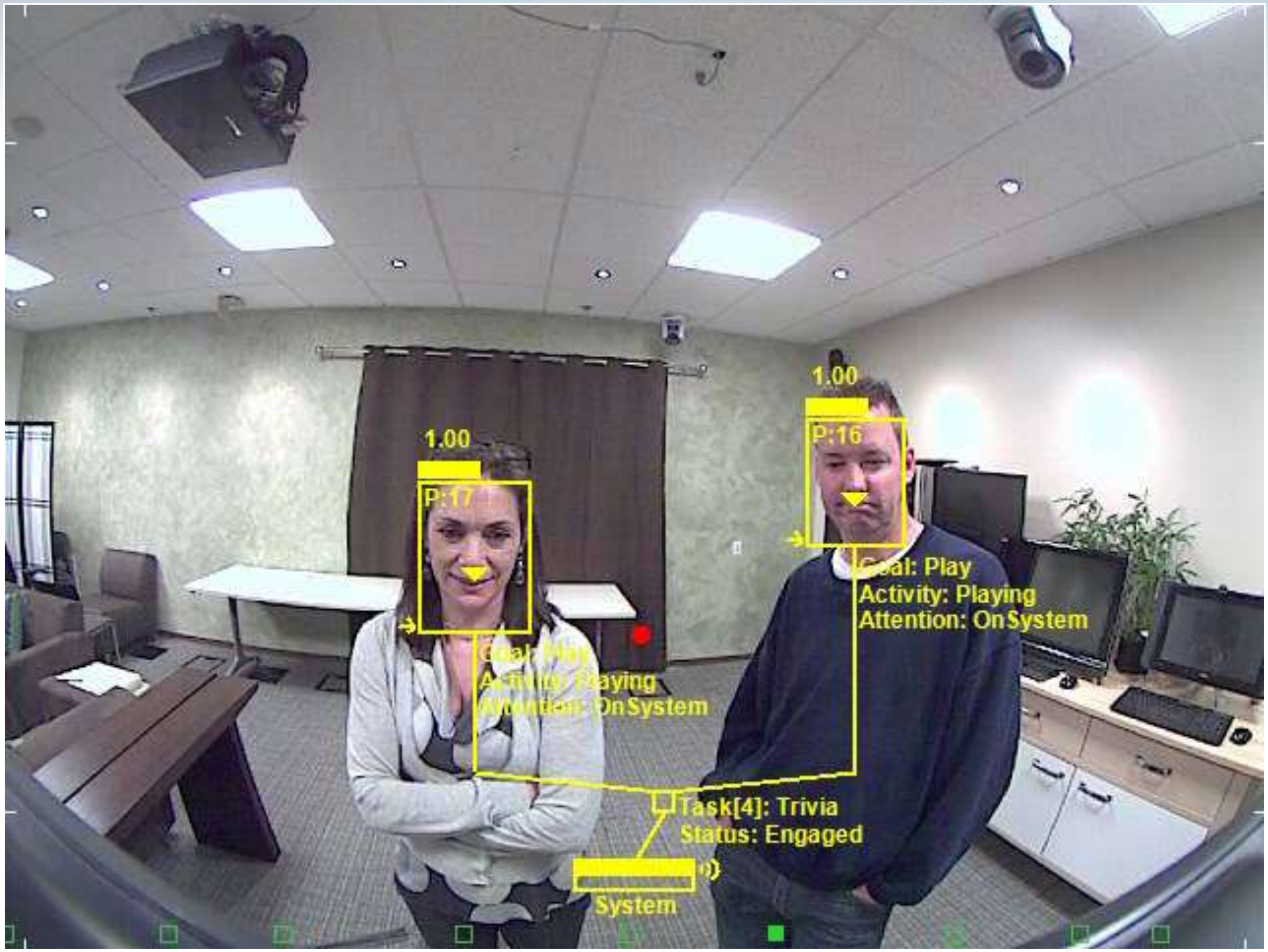
SEB = system engagement behavior
{ glance, greet, excuse-me, etc. }

Experiments



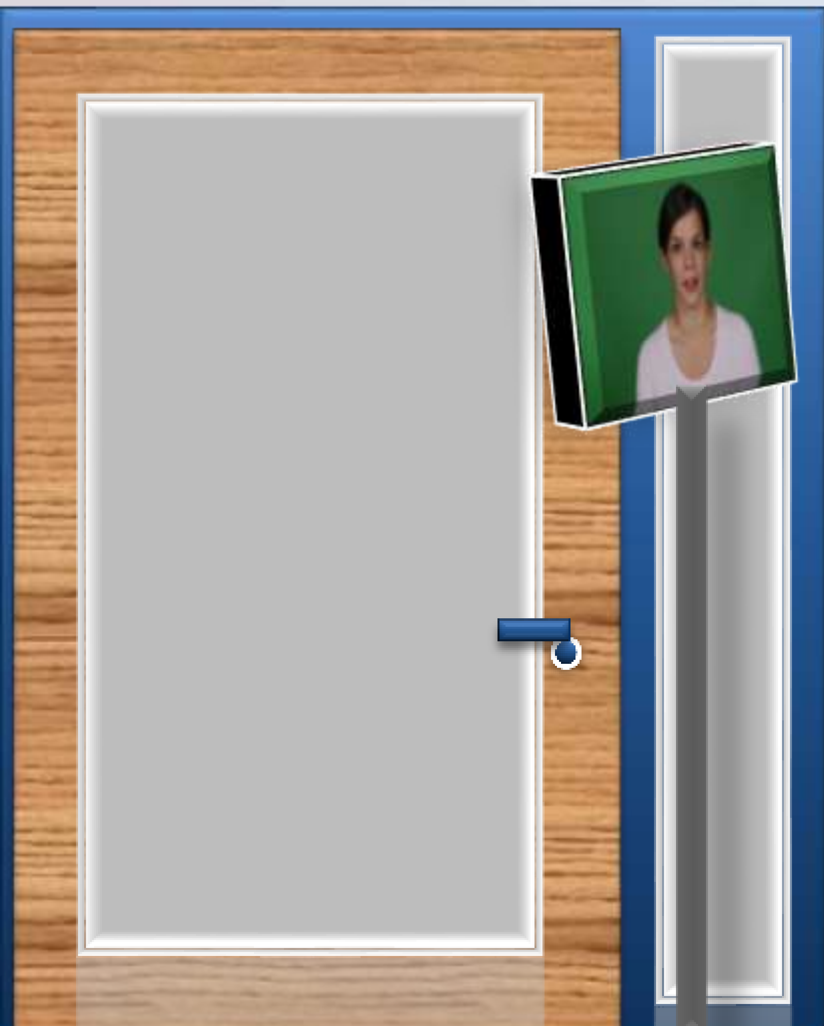
Experiments

- P: arrow indicates direction of attention
- P: P has floor
- P: P is the target of the floor release
- P: P is releasing the floor
- P: P is trying to take the floor (performs TAKE action)
- P: P is speaking
- P: P is an addressee
- indicates system's gaze direction



Personal Assistant

Perception, learning, reasoning components



Face ID | Vmail | Calendar

Room acoust. | Email | Location

Multiparty Engagement & Dialog

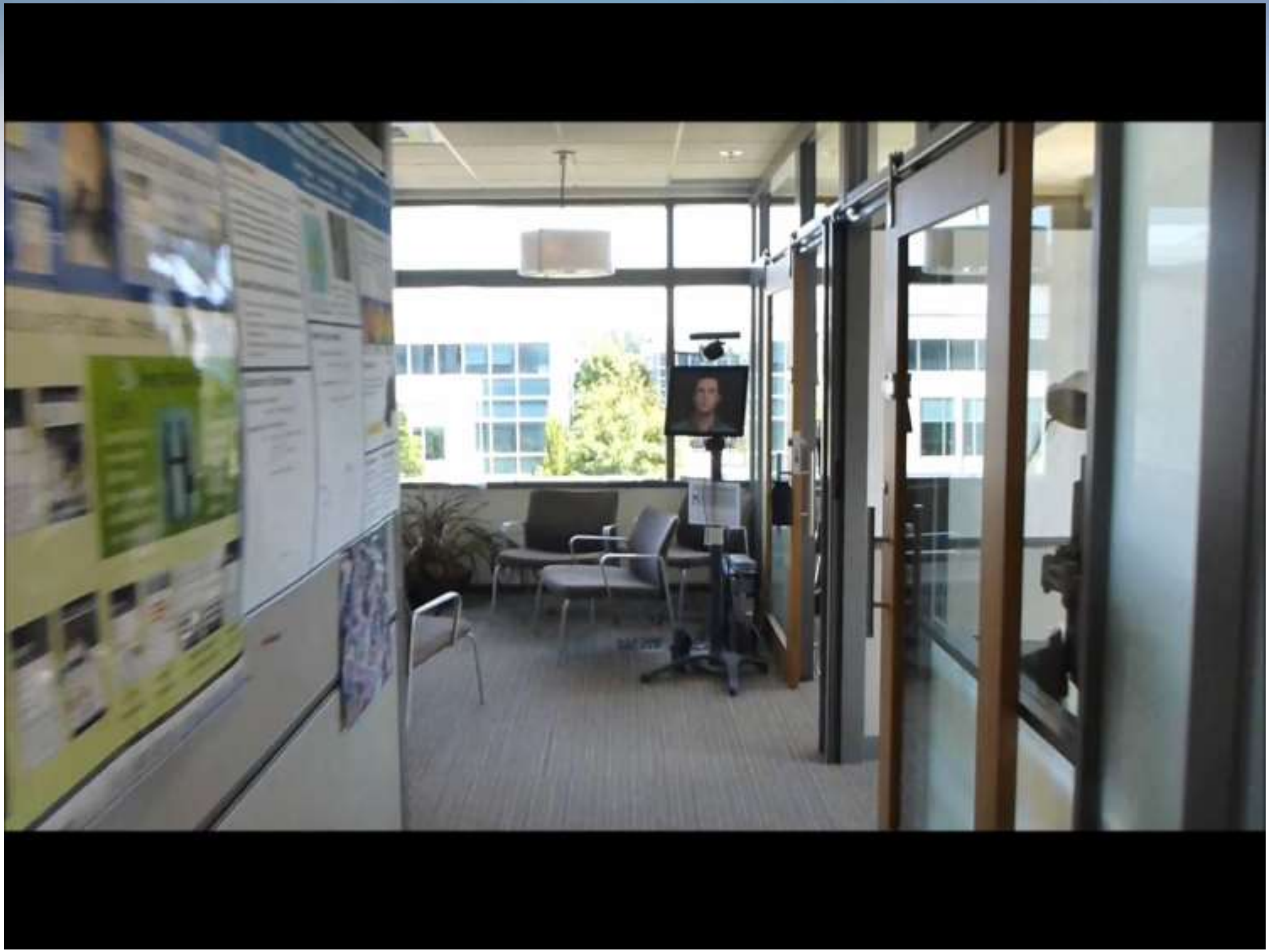
Prediction about *presence*

Prediction of *cost of interruption*

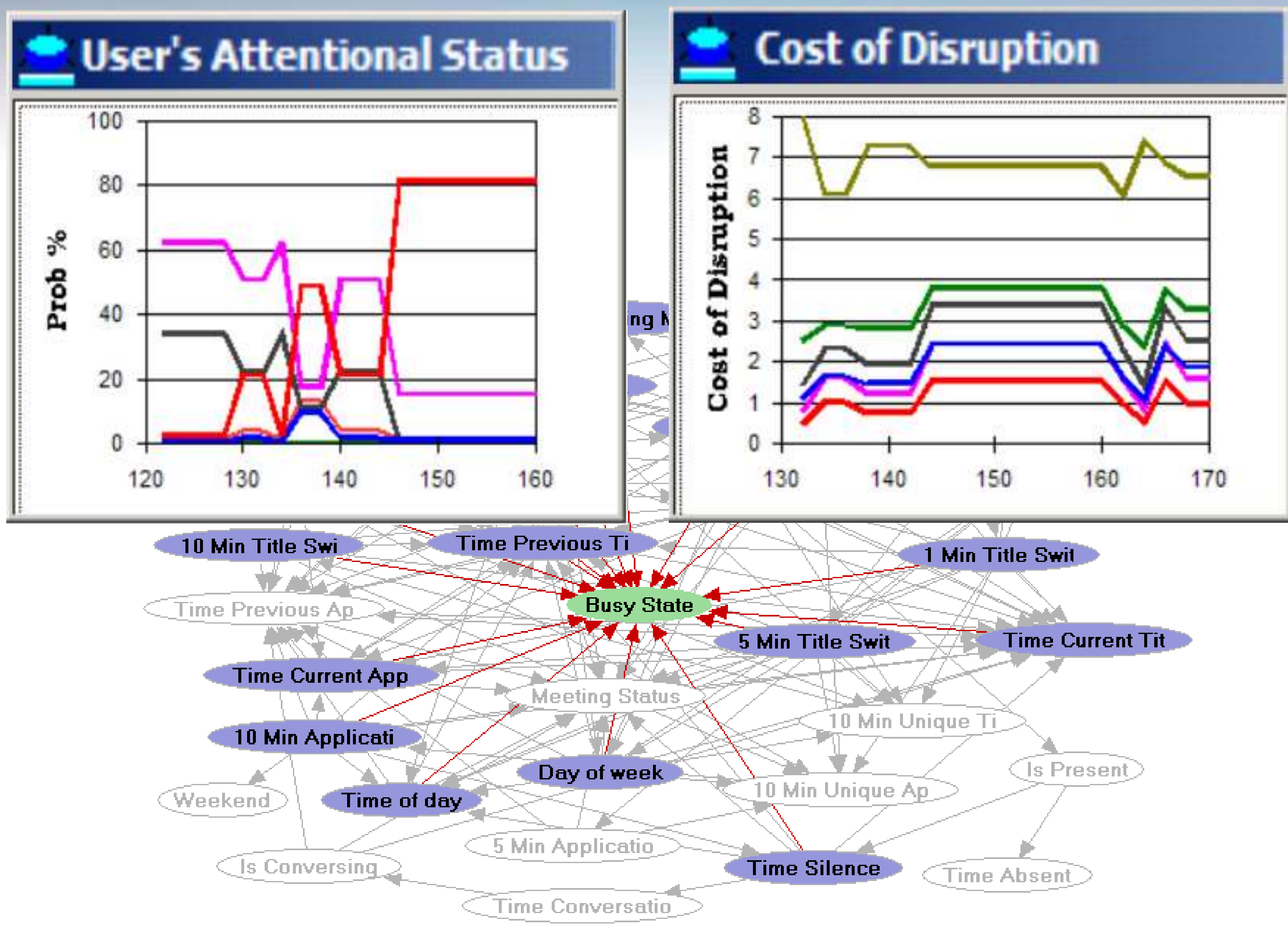
Prediction about *forgetting*

Prediction of *message urgency*

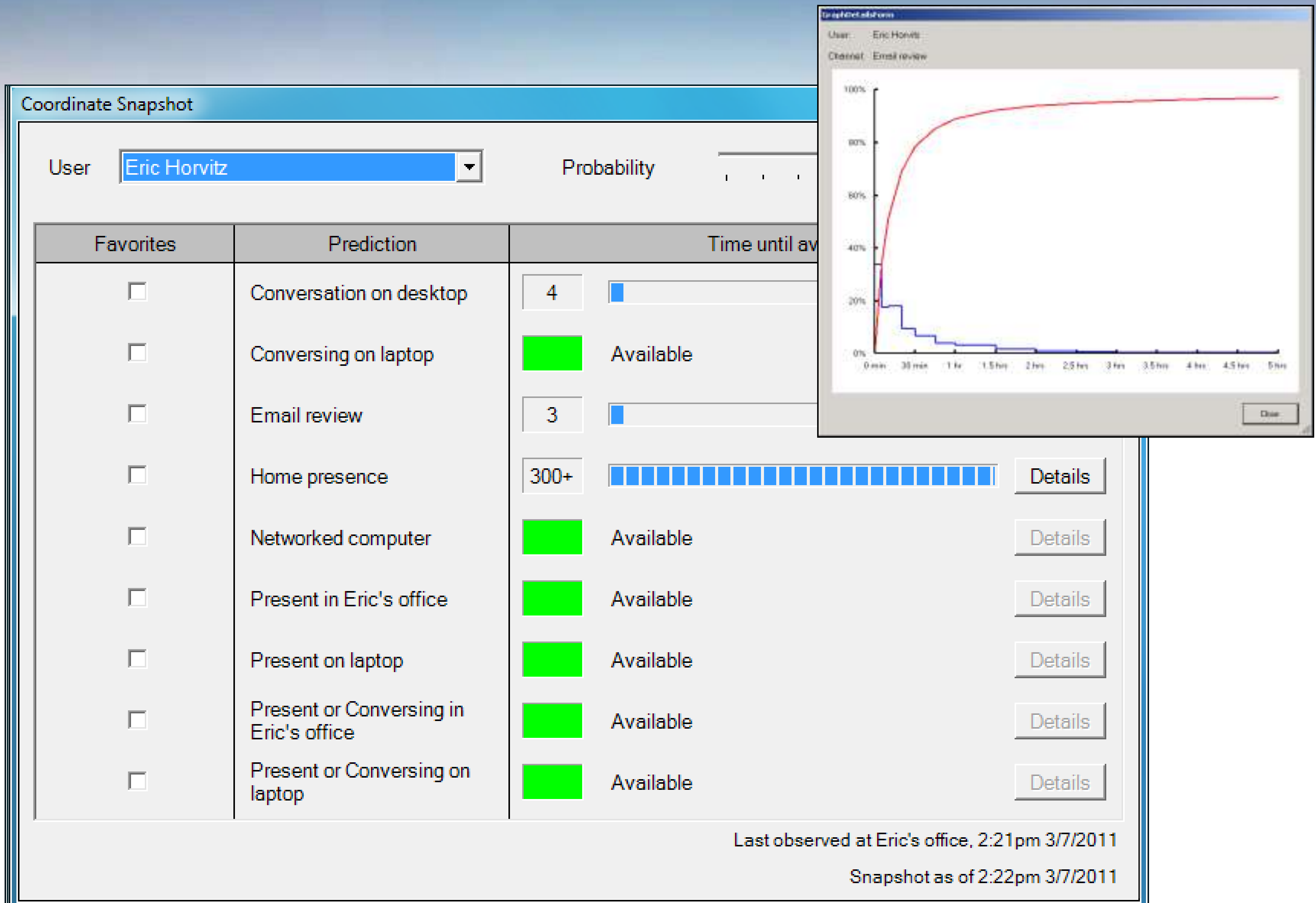
Personal Assistant



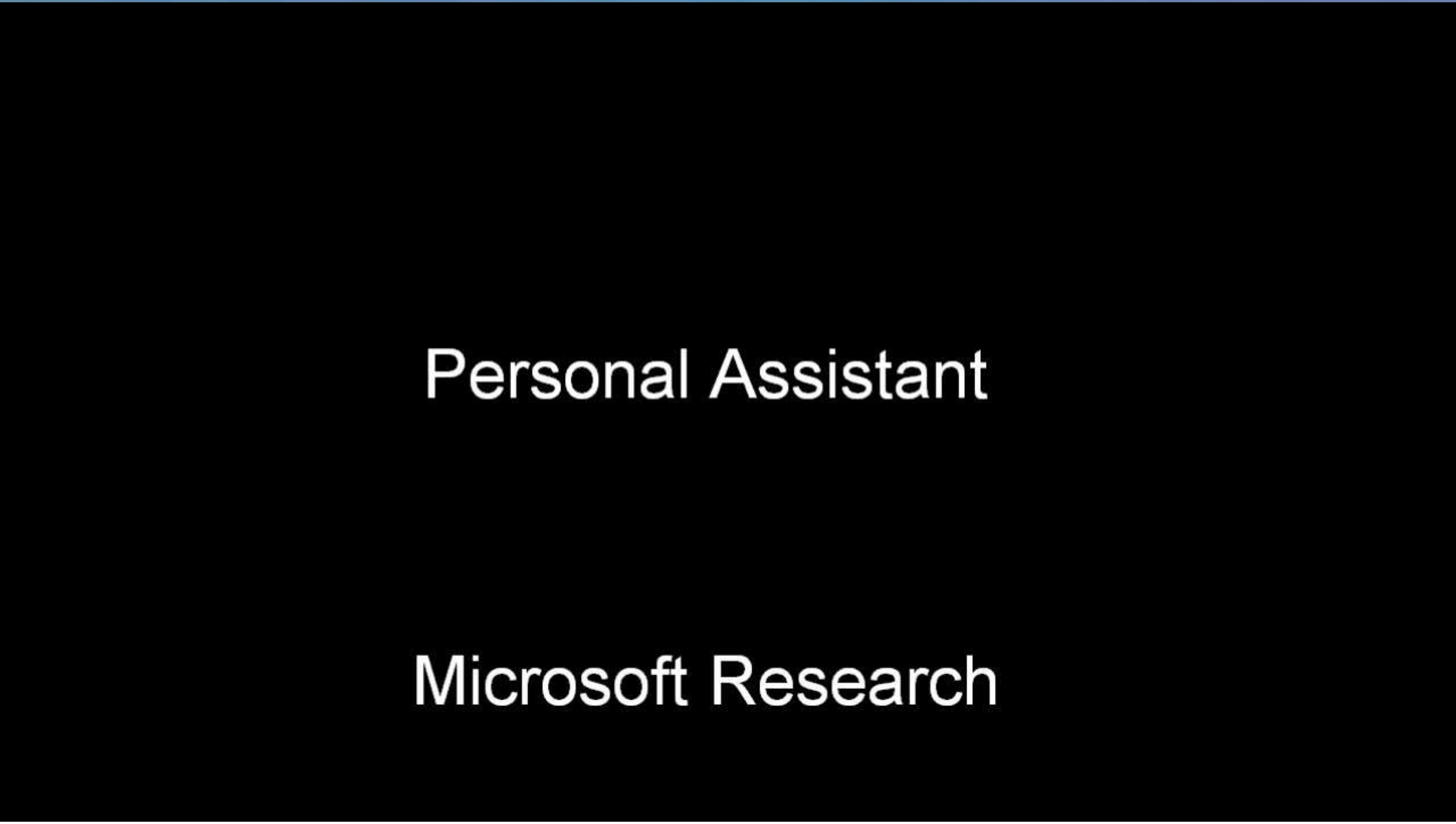
Learned Models: Cost of Interruption



Learned Models of Presence



Personal Assistant

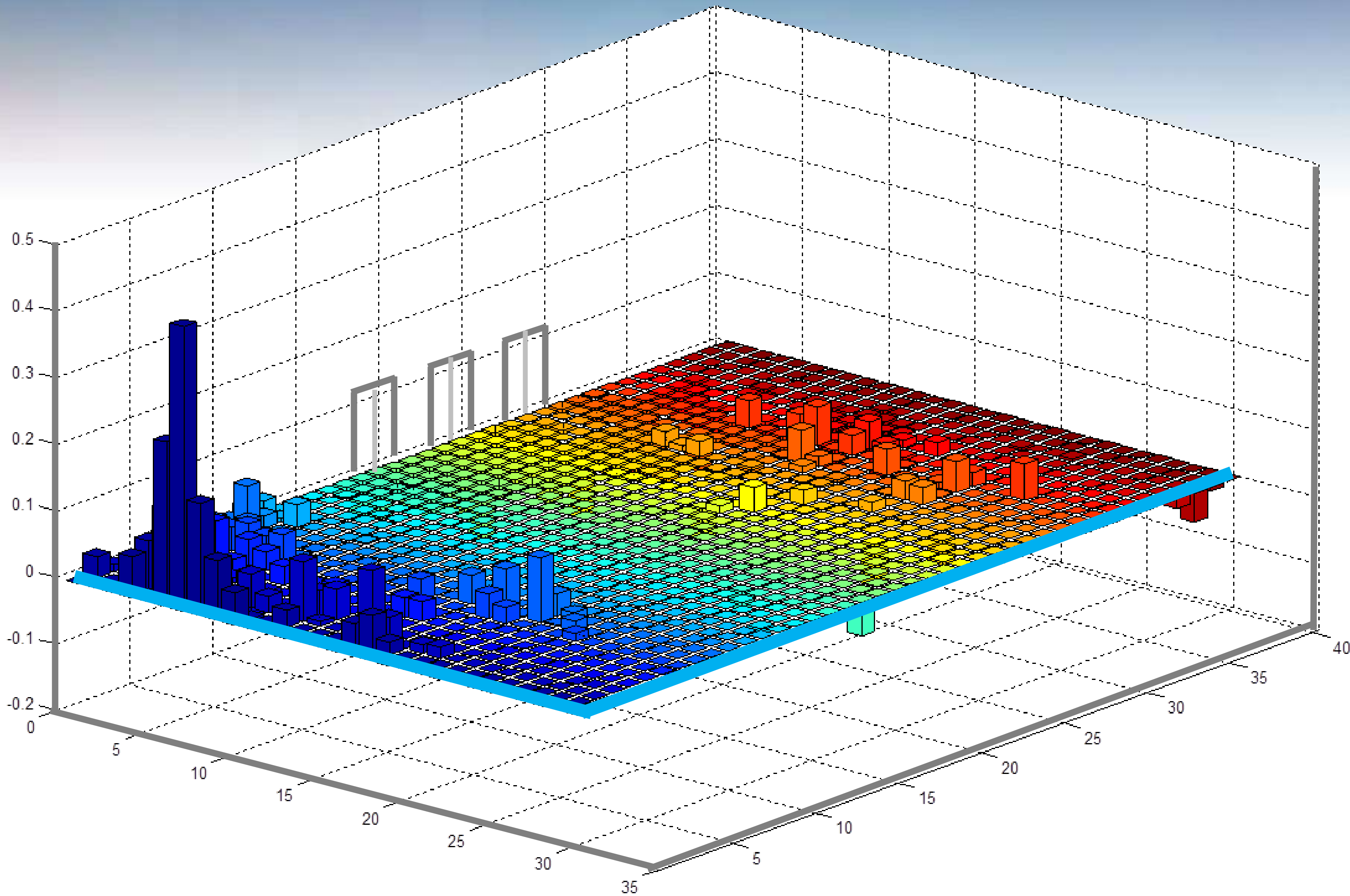


Personal Assistant

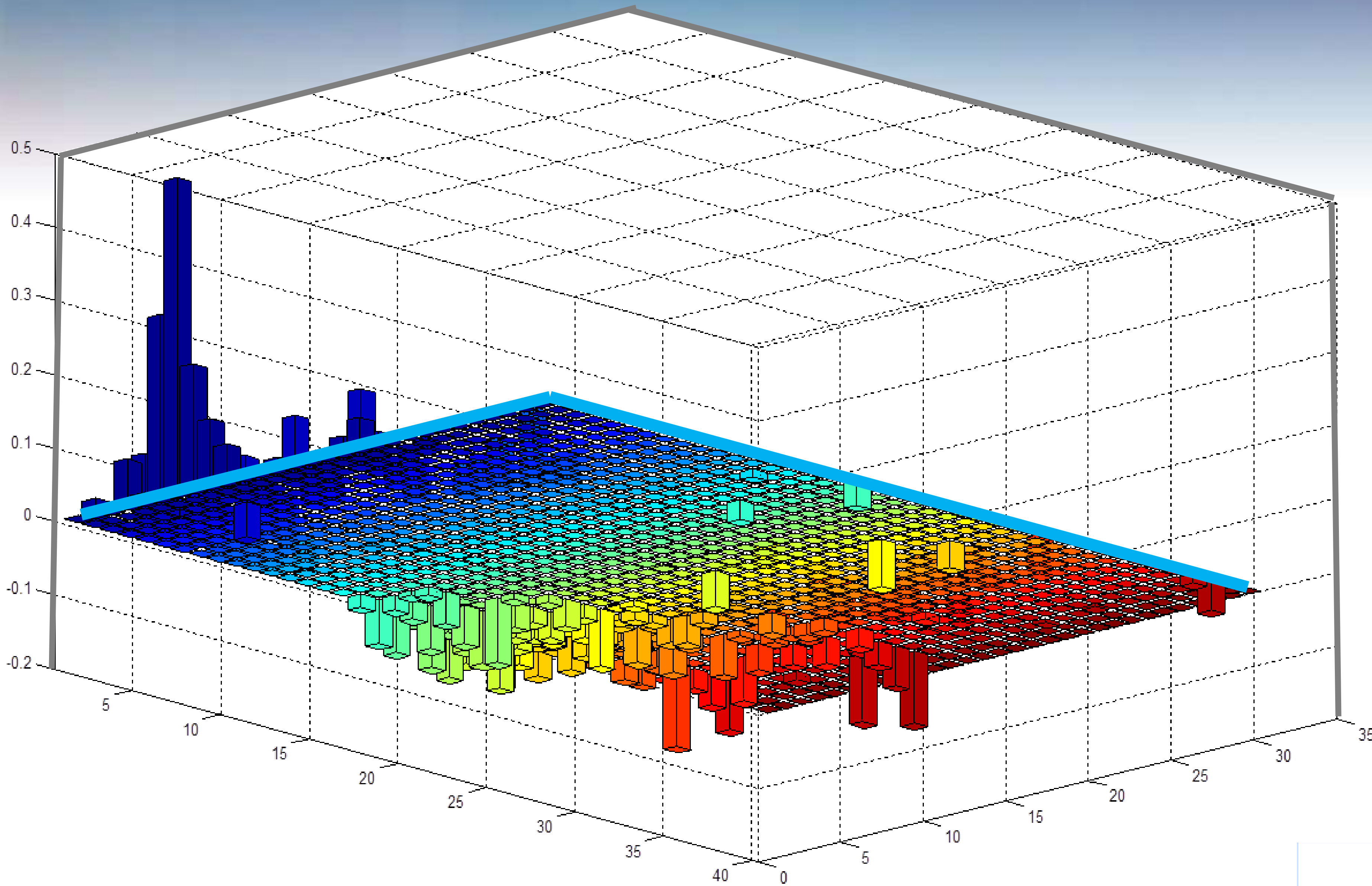
Broader Applications of Platform

Project 3E

Identifying Critical Regions

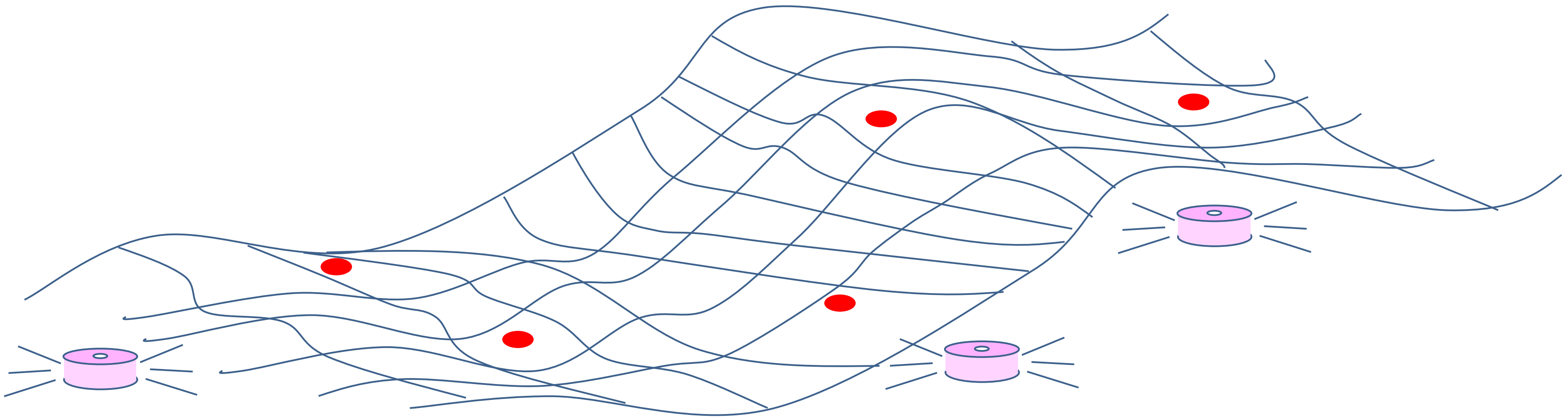


Identifying Critical Regions



Summary

- Applications of sensing, learning, and reasoning still in infancy
- Unprecedented value to people and society
- Principles → Applications → Principles …



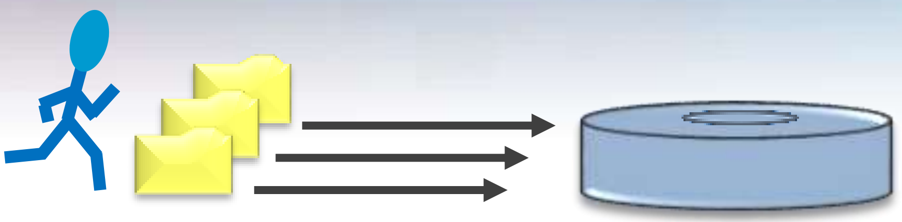
Microsoft Research Asia
Faculty Summit 2012

Privacy, Data, and Machine Learning

Urgency …and optimism

- Clarity, preferences, and handles
- Decision-theoretic mediation
- Differential privacy
- Protected sensing & personalization

Example: Personalized Search (Psearch)

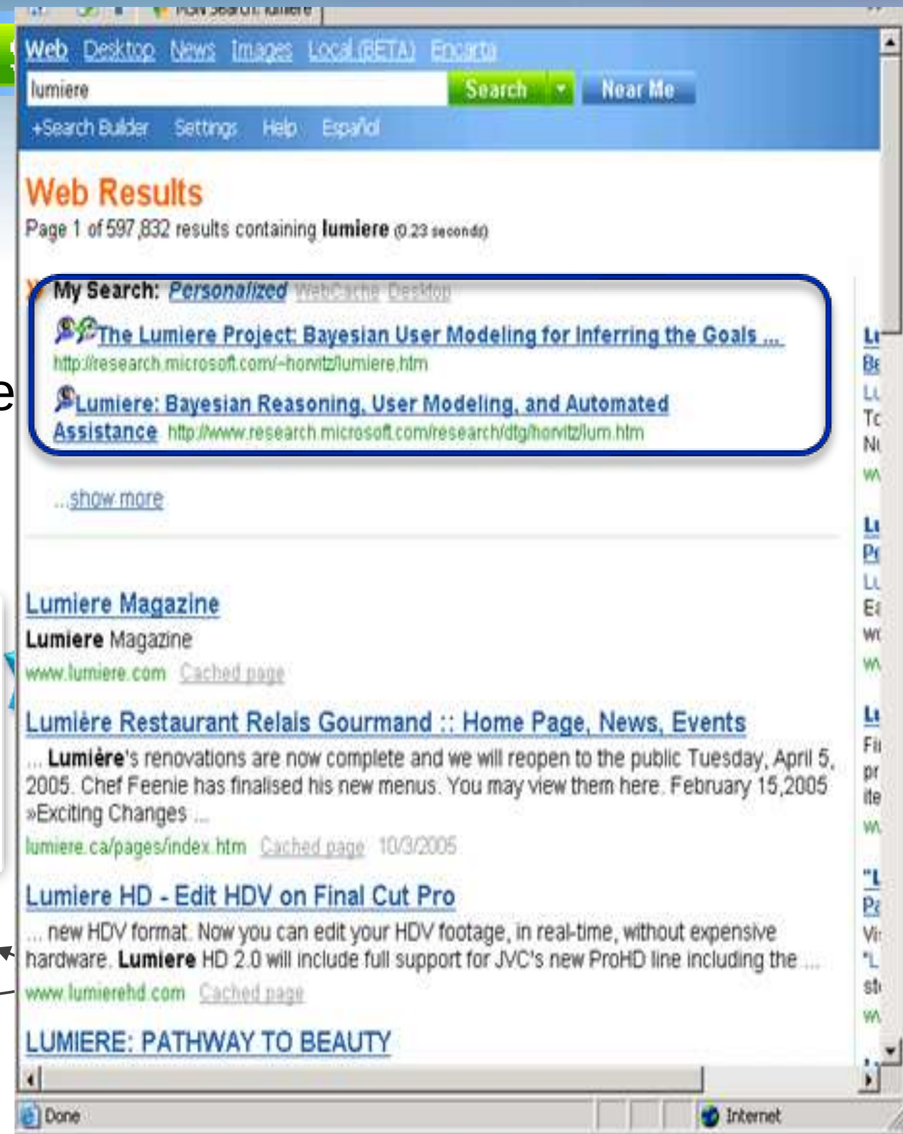


- Email
- Documents
- Web activity
- GPS, wifi

.
.
.

Content & activities store

Personalized ranker



Results from web search engine

With J. Teevan and S. Dumais

Example: Lifebrowser

Images
& videos

Desktop
& search activity

Appts &
events

Locations

Whiteboard
capture