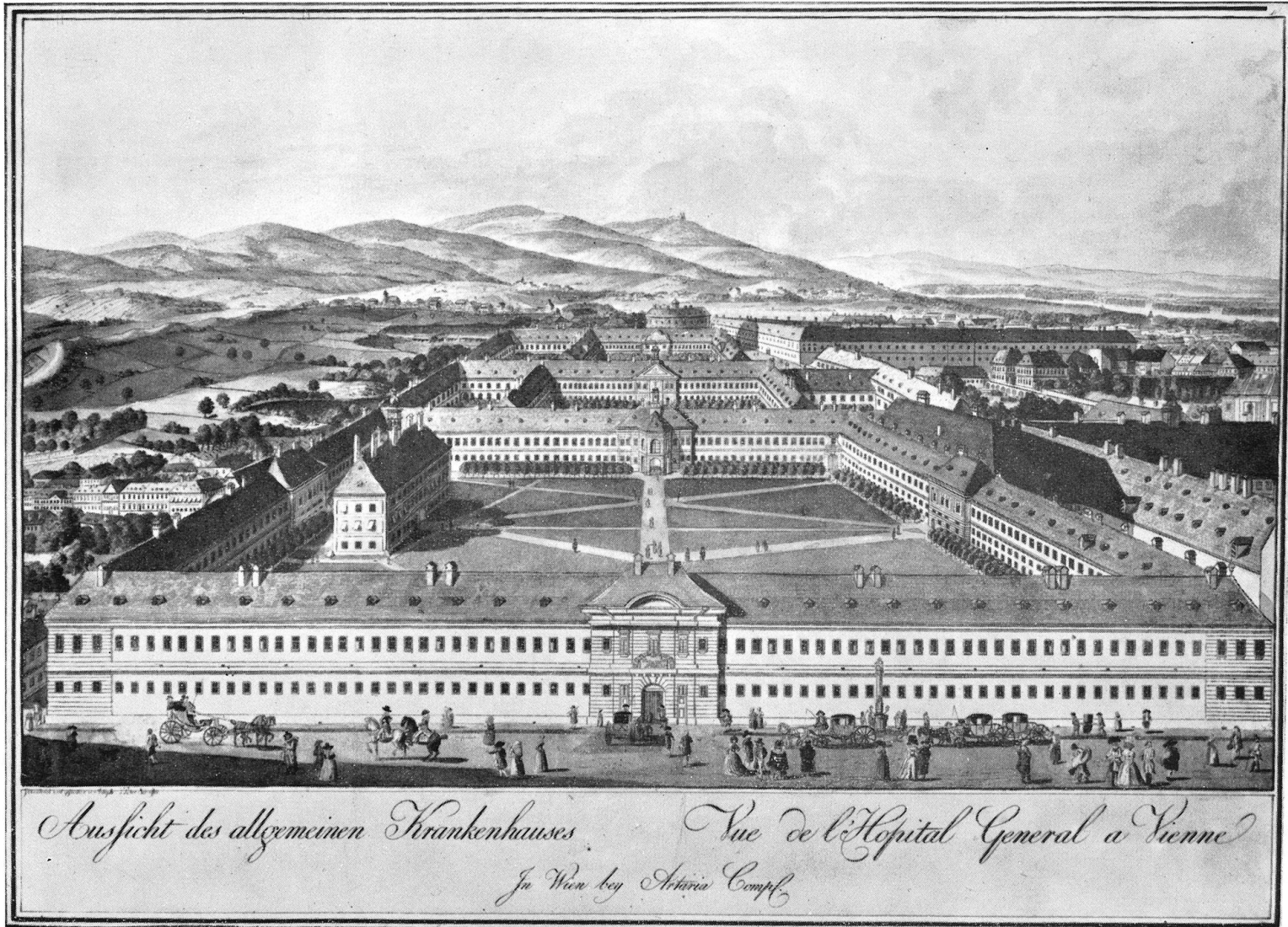




Size is Not Everything!

Bögel György
bogelgy@ceubusiness.org

The world of **Small Data**



The data set

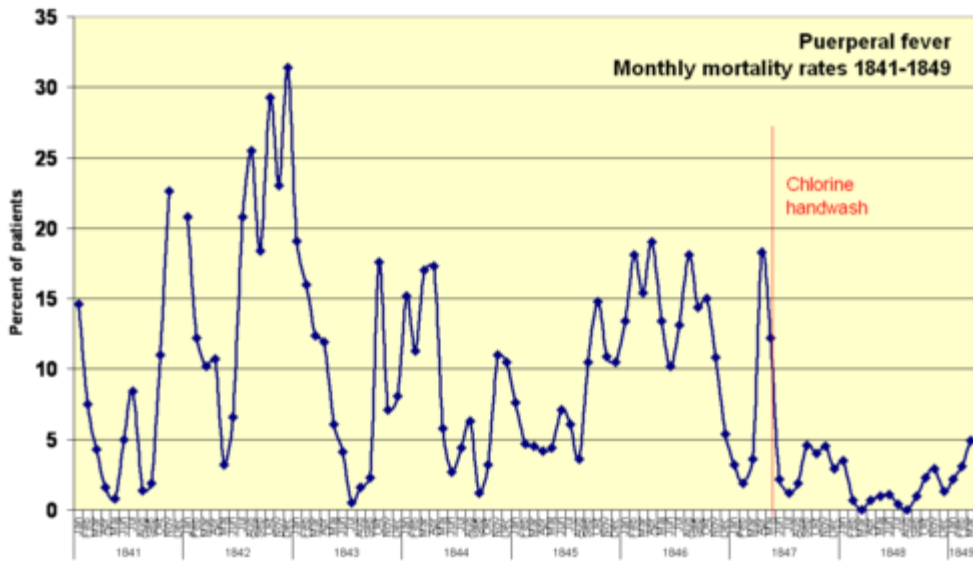
Year	First clinic			Second clinic		
	Births	Deaths	Rate (%)	Births	Deaths	Rate (%)
1841	3,036	237	7.8	2,442	86	3.5
1842	3,287	518	15.8	2,659	202	7.6
1843	3,060	274	9.0	2,739	164	6.0
1844	3,157	260	8.2	2,956	68	2.3
1845	3,492	241	6.9	3,241	66	2.0
1846	4,010	459	11.4	3,754	105	2.8

Data, information, knowledge



- Ignác Semmelweis
- Years before germ theory, Pasteur and Koch
- No acceptable scientific tools, he had to rely on mortality **data**
- Realized a **pattern**
- Looked for **explanations** – and found one!

...and the solution



- Doctors' wards – Midwives wards
- Hand washing with chlorinated lime
- Rejected by the medical community
- Widespread acceptance only years after his death

Data can save lives!

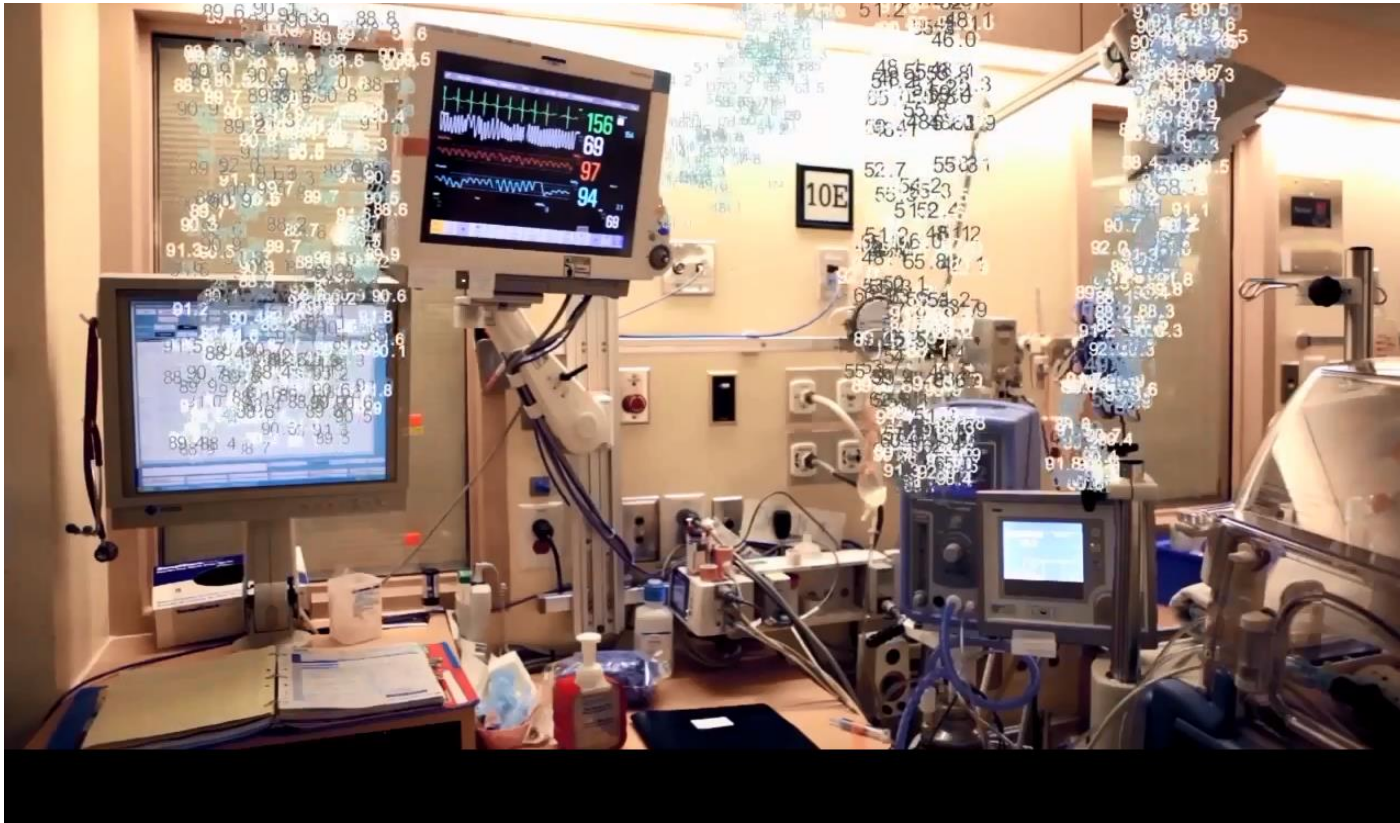
But...

The world of **Big Data**



- Dr. Carolyn McGregor
- Computer scientist
- Data-based analysis of customer behavior in banking and retail
- **2007: moving to healthcare** after a personal tragedy

What did she find there?



An electrocardiogram alone records 1,000 readings per second

From data to action



- Teamed up with IBM and University of Ontario
- Collect data, find meaning in numbers, look for patterns, build predictive models
- Provide insights into which treatments tend to work and which do not

Insight and prediction

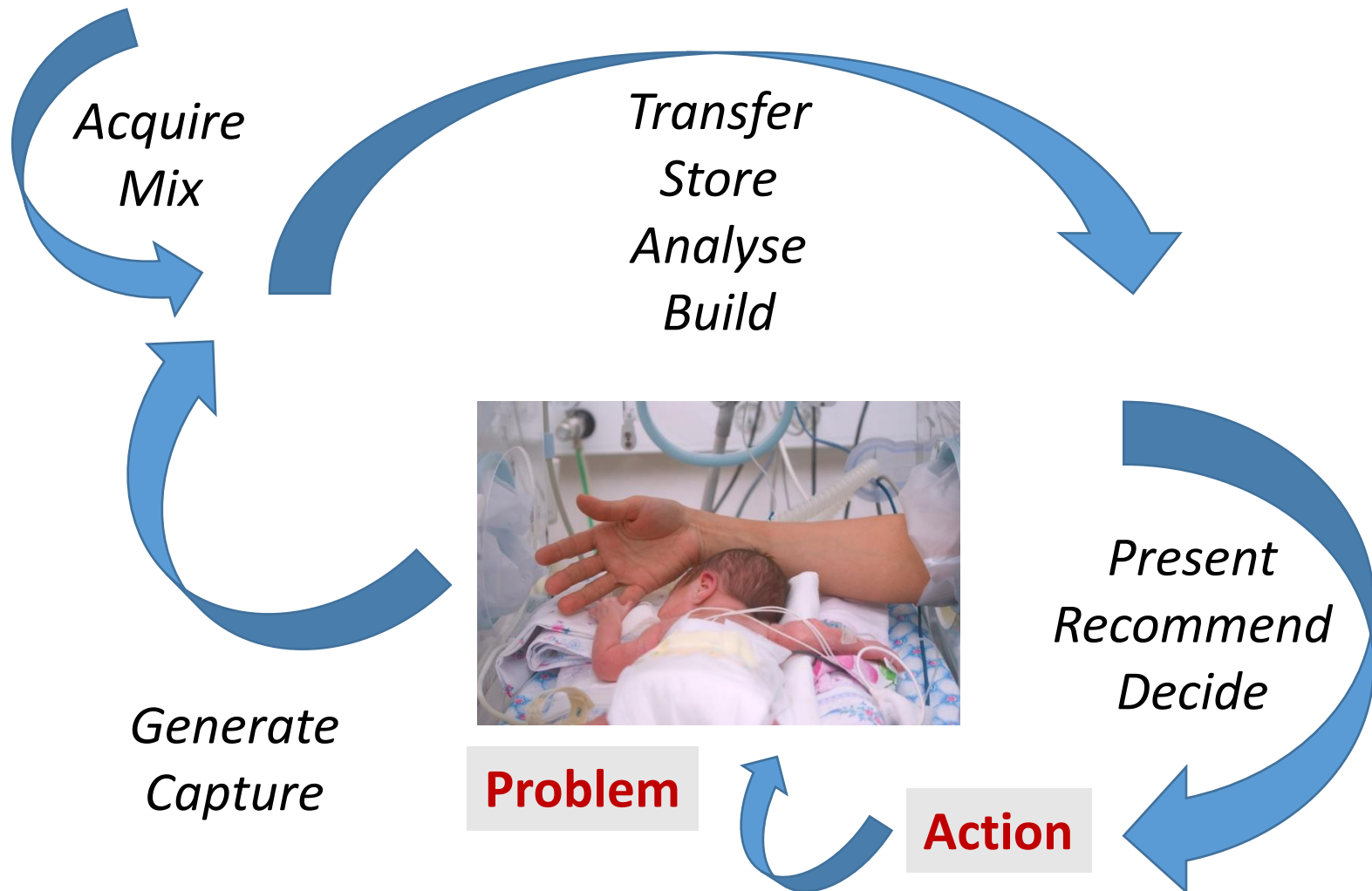


- Premature babies
- Capturing 16 different **data streams**
- Around 1,260 **data points** per second
- **Real-time** processing
- Looking for **correlations**
- **Predicting infections** 24 hours before overt symptoms appear
- E.g. very constant vital signs often are detected prior to a serious infection

Big Data can save
more lives!

But size is not everything...

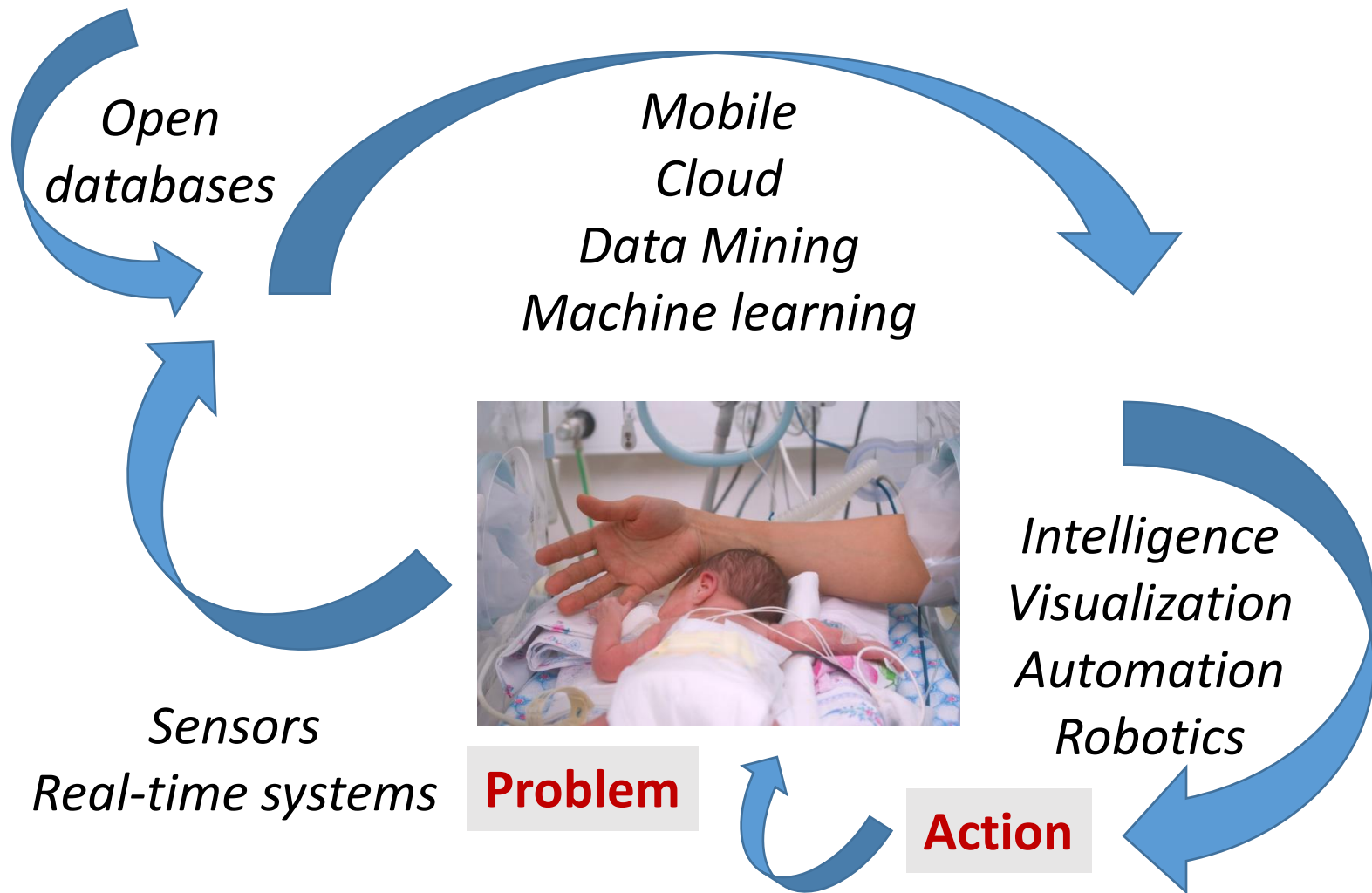
Needed: a smart system
How to predict health problems?



Propeller Health: asthma control

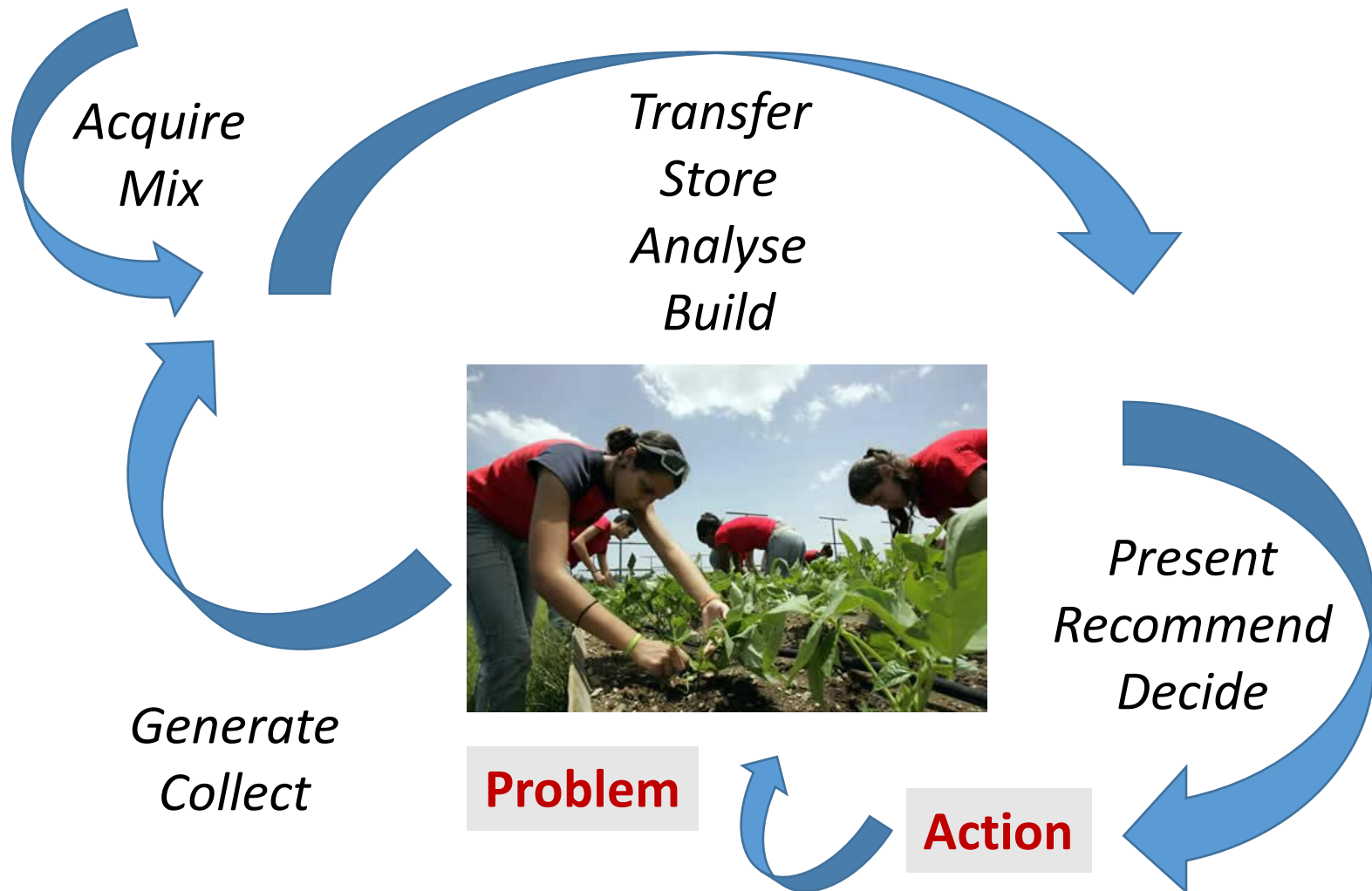


Development everywhere (examples only)



Smart **agriculture**

How to help Mother Nature?

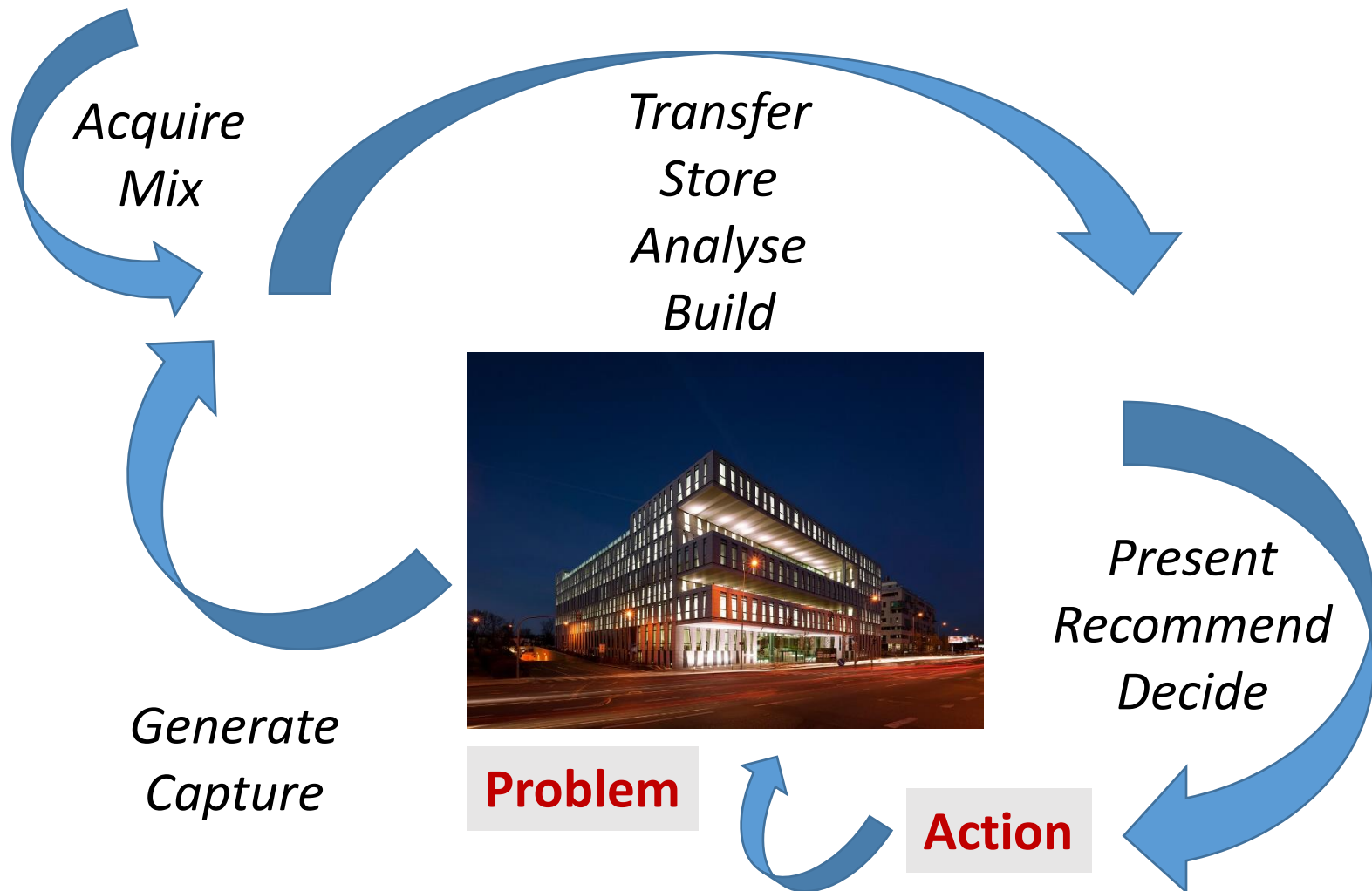


Smart thinning (Blue River Technology)



Smart **enterprise**

(Marketing and **sales**, production, maintenance, finances, HR...)



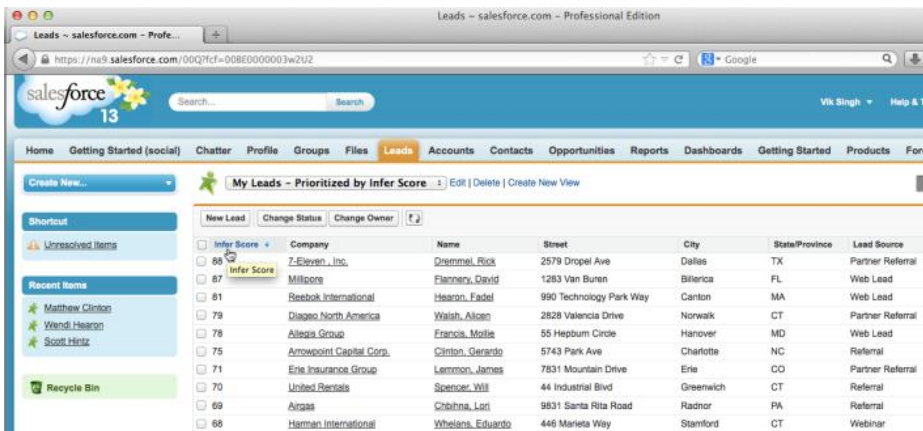
Infer: predictive customer scoring



- Identifying the **most promising leads** based on a user's historical deal-making tendencies and external data about potential leads

- Relies on **150 signals**, including dozens of online data feeds

- Investor: **Andreessen Horowitz**



Leads - salesforce.com - Professional Edition

https://na9.salesforce.com/00Q2fcf-008E0000003w2U2

salesforce 13

Home Getting Started (social) Chatter Profile Groups Files Leads Accounts Contacts Opportunities Reports Dashboards Getting Started Products Fore

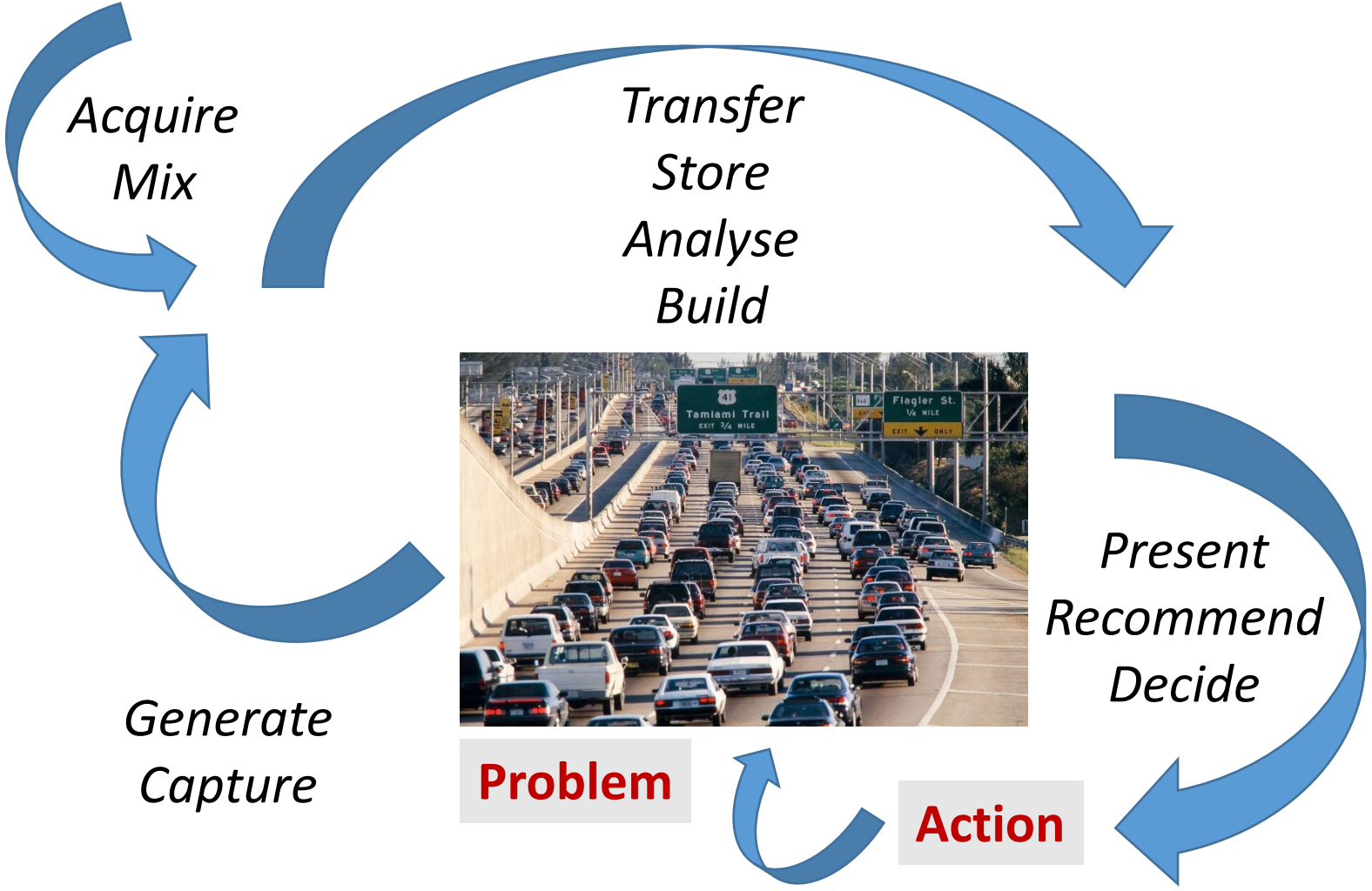
Create New...

My Leads - Prioritized by Infer Score | Edit | Delete | Create New View

Infer Score	Company	Name	Street	City	State/Province	Lead Source
86	7-Eleven, Inc.	Dremmel, Rick	2579 Dropel Ave	Dallas	TX	Partner Referral
87	Millipore	Flannery, David	1283 Van Buren	Billerica	FL	Web Lead
81	Reebok International	Hearon, Fadel	990 Technology Park Way	Carlton	MA	Web Lead
79	Diageo North America	Walsh, Aileen	2828 Valencia Drive	Norwalk	CT	Partner Referral
78	Allegis Group	Francis, Mollie	55 Heppburn Circle	Hanover	MD	Web Lead
75	Arrowpoint Capital Corp.	Clinton, Gerardo	5743 Park Ave	Charlotte	NC	Referral
71	Erie Insurance Group	Lemmon, James	7831 Mountain Drive	Erie	CO	Partner Referral
70	United Rentals	Spencer, Will	44 Industrial Blvd	Greenswich	CT	Referral
69	Aygas	Cisshina, Lori	9831 Santa Rita Road	Radnor	PA	Referral
66	Harman International	Whelans, Eduardo	446 Marieta Way	Stamford	CT	Webinar

Smart traffic

How to avoid jams?



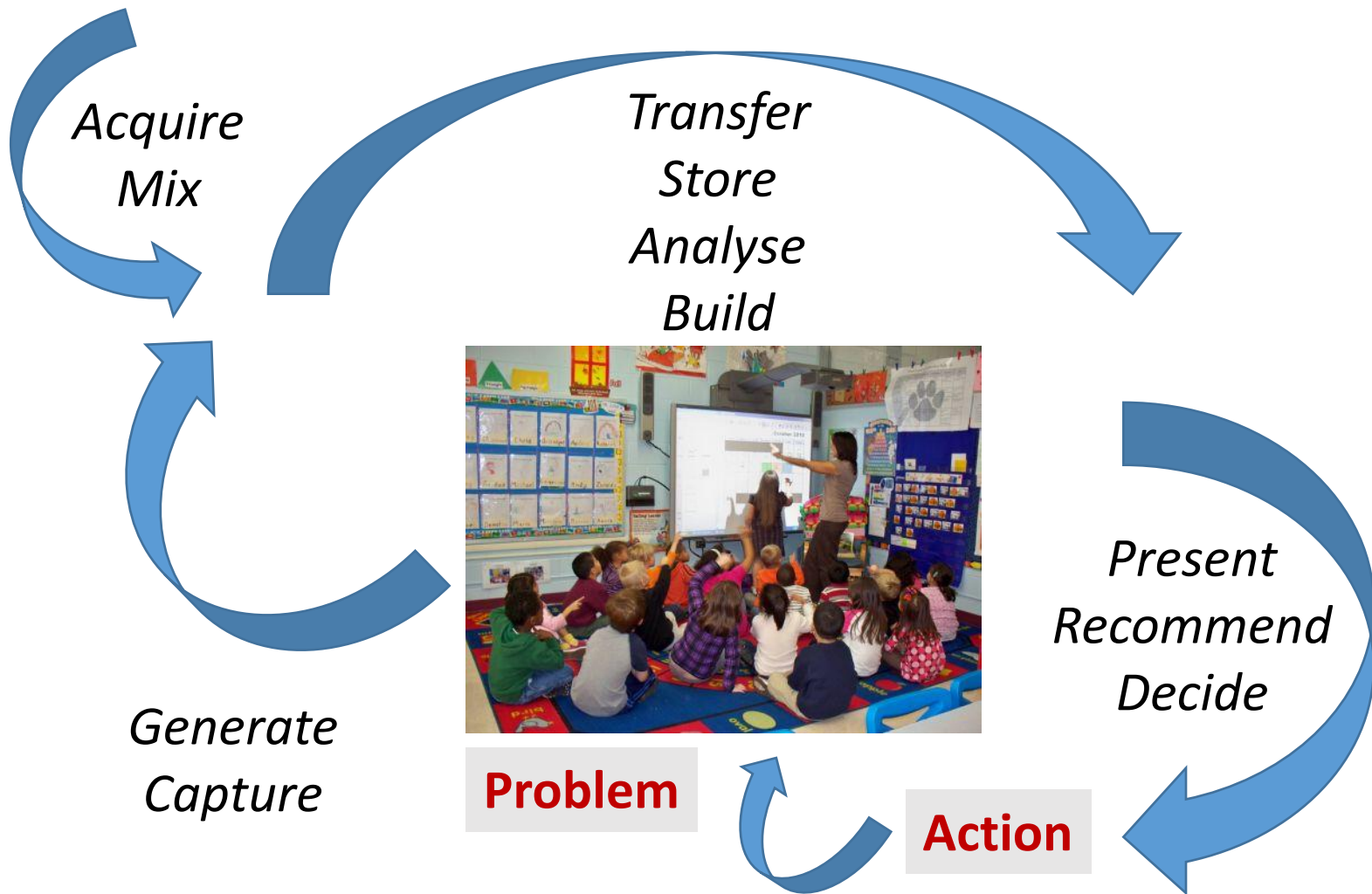
Waze: community based traffic and navigation app



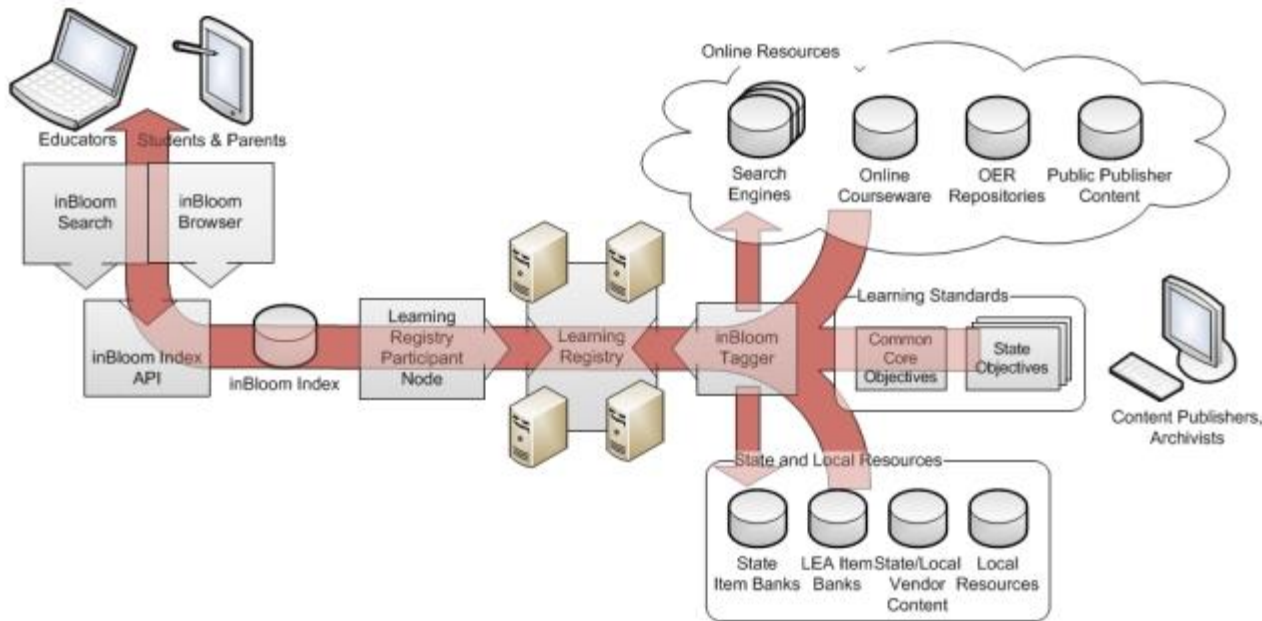
- Drivers **share** real-time traffic and road info
- **Report** jams, accidents, roadworks, speed traps, petrol prices, etc.
- Active users are **tracked** automatically
- 2013: **50m users** in 100+ countries

Smart **education**

How to customize teaching and learning?

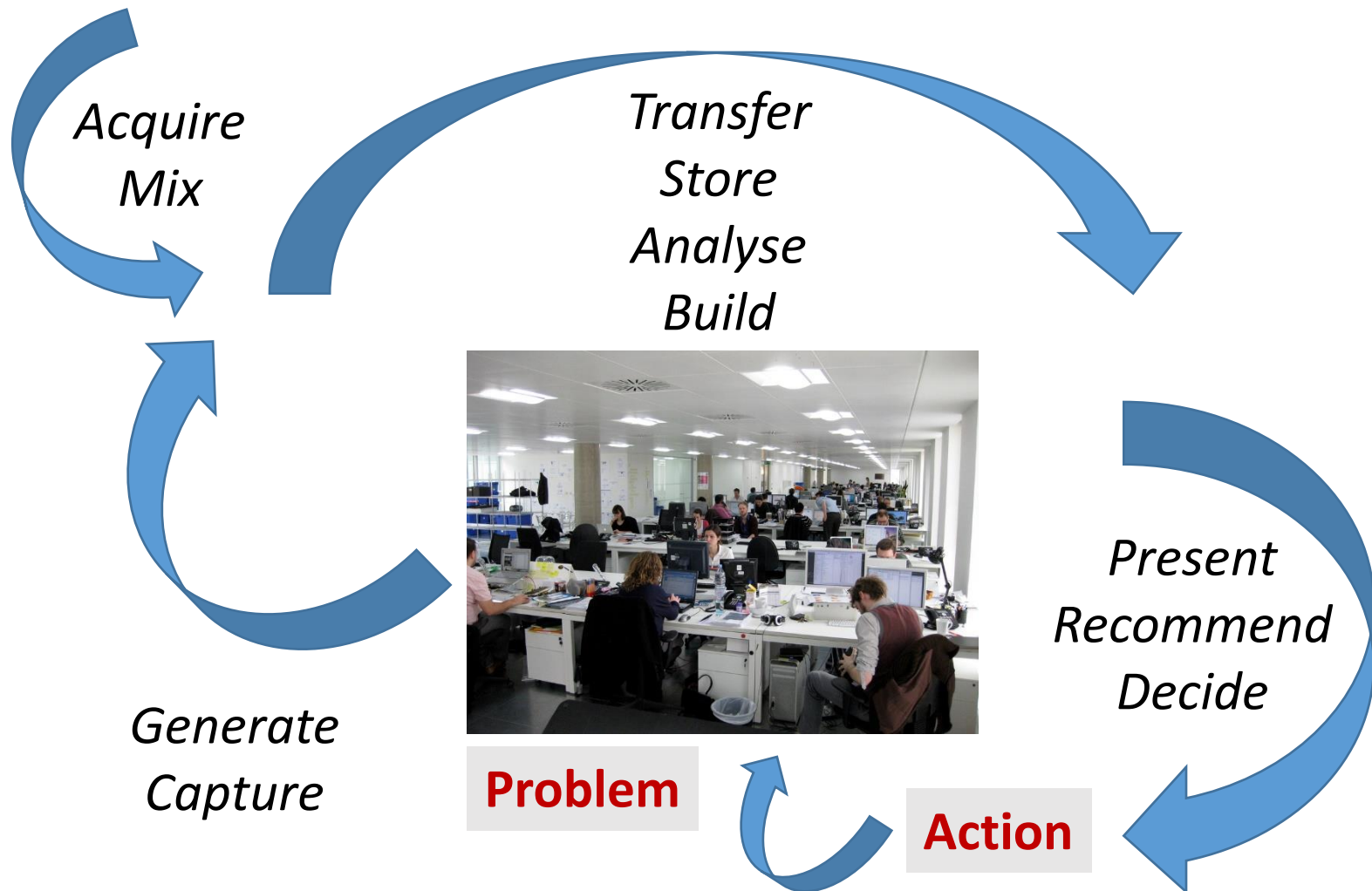


inBloom: customized education



Smart **organization**

How to improve cooperation and teamwork?



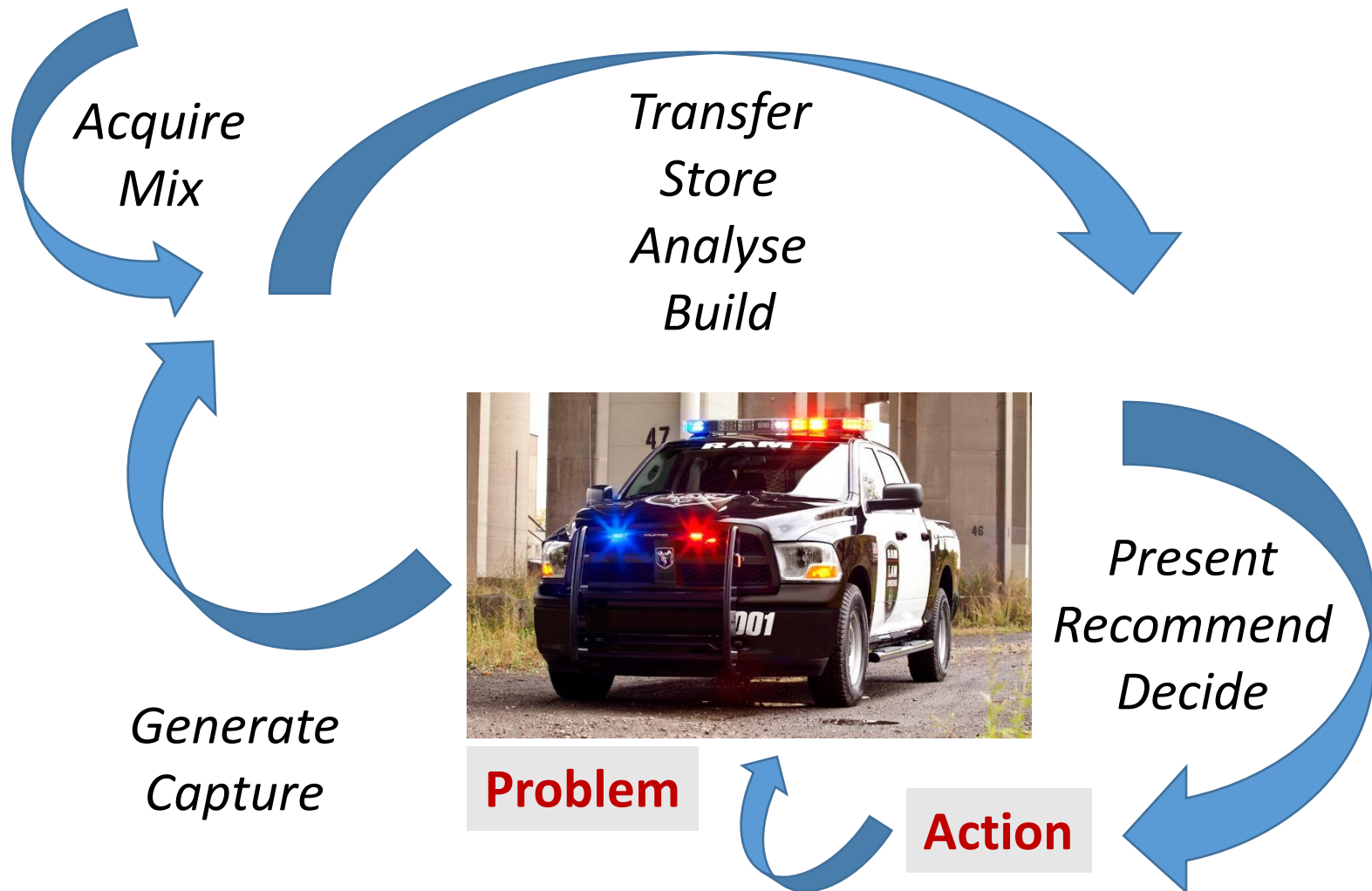
Sociometric Solutions: measuring how people collaborate



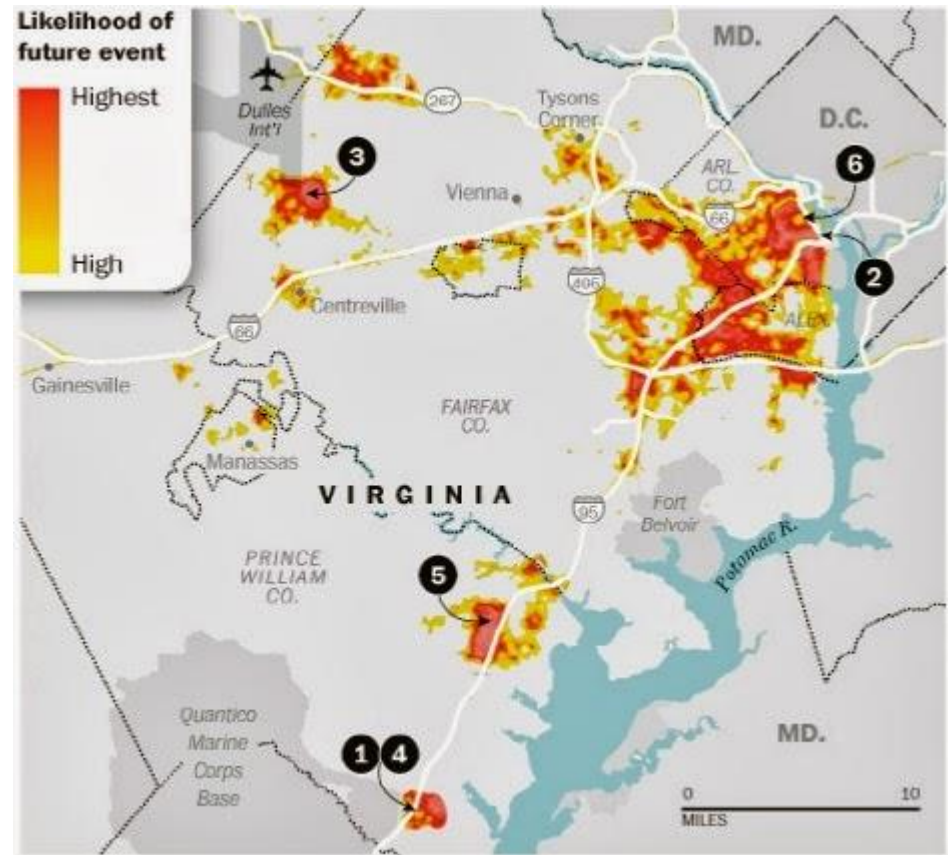
- Analyzing communication patterns with **social sensing technology**
- **Sensor ID badges** for collecting tons of data
- To drive innovative organizational **transformations**

Smart **police**

How to prevent / fight crime?

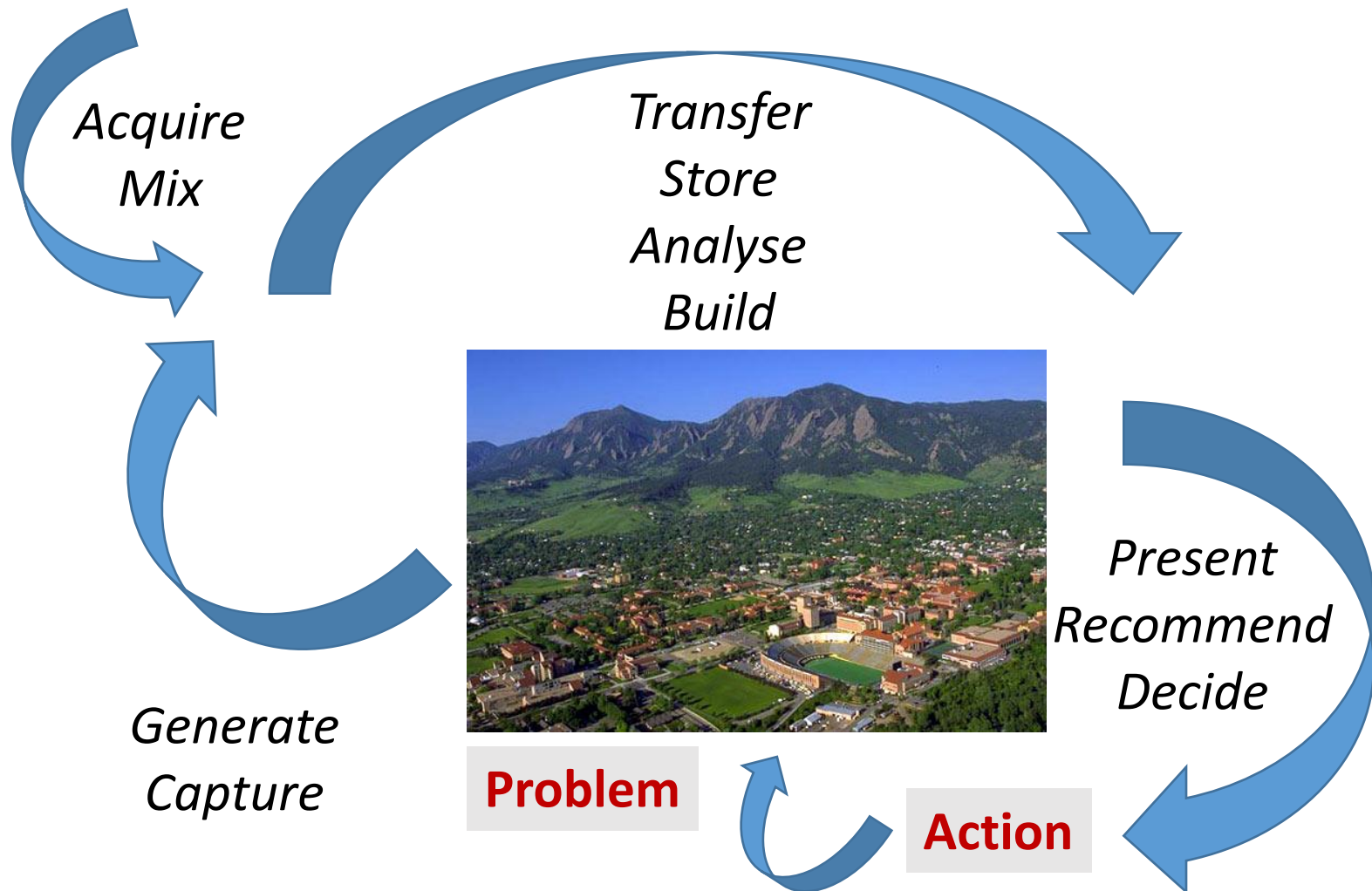


Predicting crime

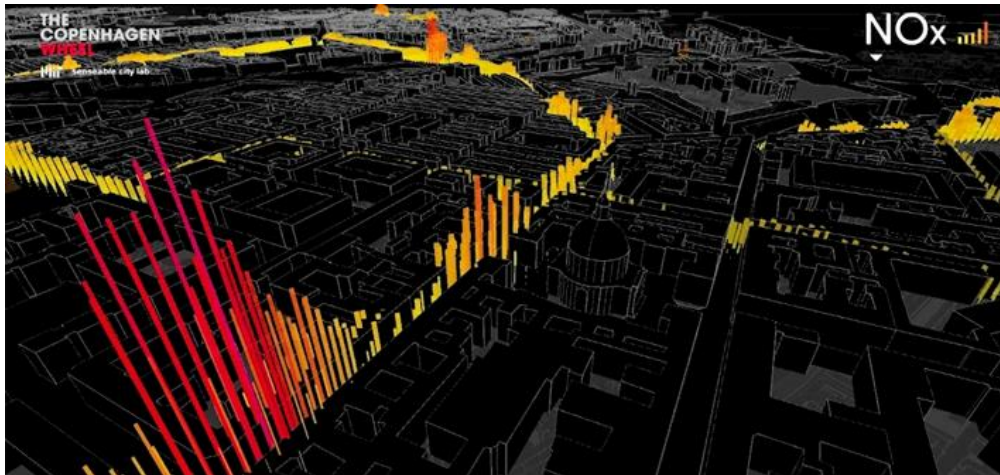


Smart town

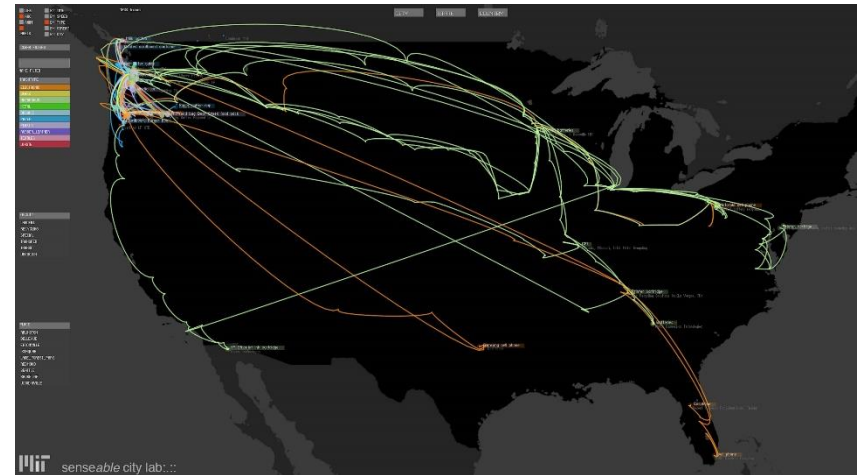
How to make it efficient, green, cheap, creative...



MIT SENSEable City Laboratory



Pollution project



Garbage project

Opportunity matrix – the Big Data ecosystem

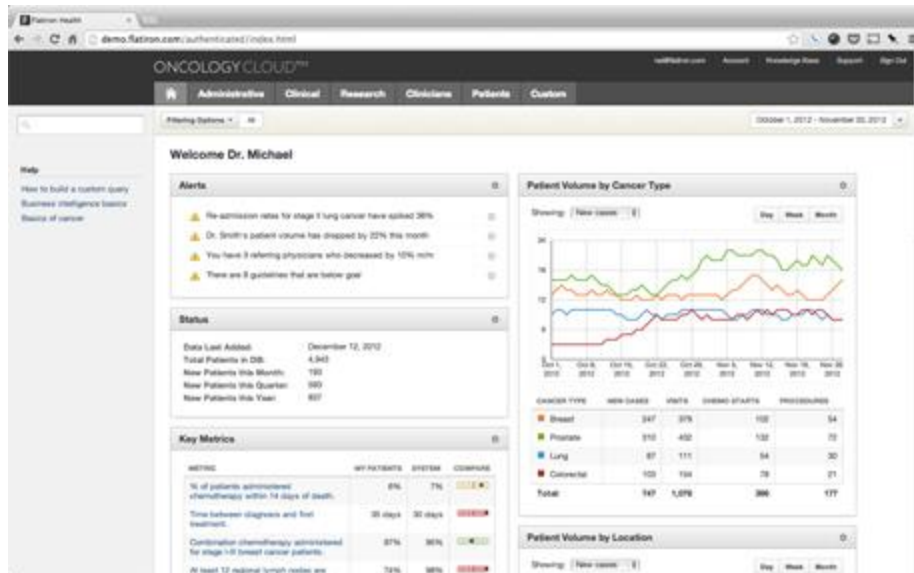
	Business	Healthcare	Education	Energy
Generating	<input type="checkbox"/>			
Capturing				
Mixing				
Transferring				
Storing				
Analyzing/Mining				
Modelling				
Presenting				
Recommending				
Acting				

Detailed description: The image shows a 10x5 opportunity matrix. The columns represent sectors: Business, Healthcare, Education, and Energy. The rows represent data lifecycle stages: Generating, Capturing, Mixing, Transferring, Storing, Analyzing/Mining, Modelling, Presenting, Recommending, and Acting. A green square is in the 'Generating' row under 'Business'. A brown rectangle highlights the 'Capturing' row under 'Healthcare'. A red rectangle highlights the 'Analyzing/Mining' row across all sectors. A blue border highlights the 'Energy' column across all sectors. Dashed arrows point right from the top of the 'Energy' column and down from the left of the 'Business' column.

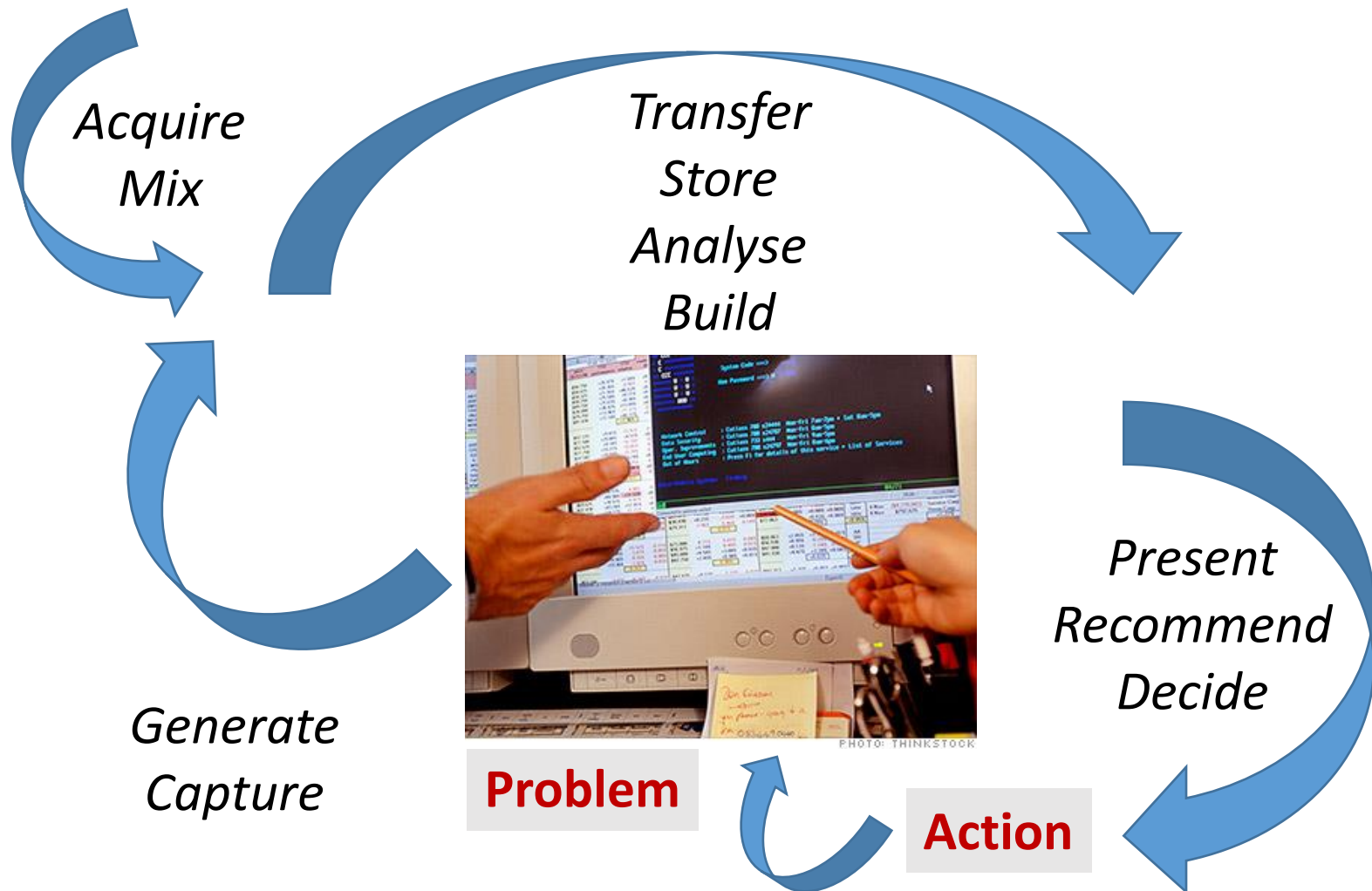
A niche player: Flatiron Health



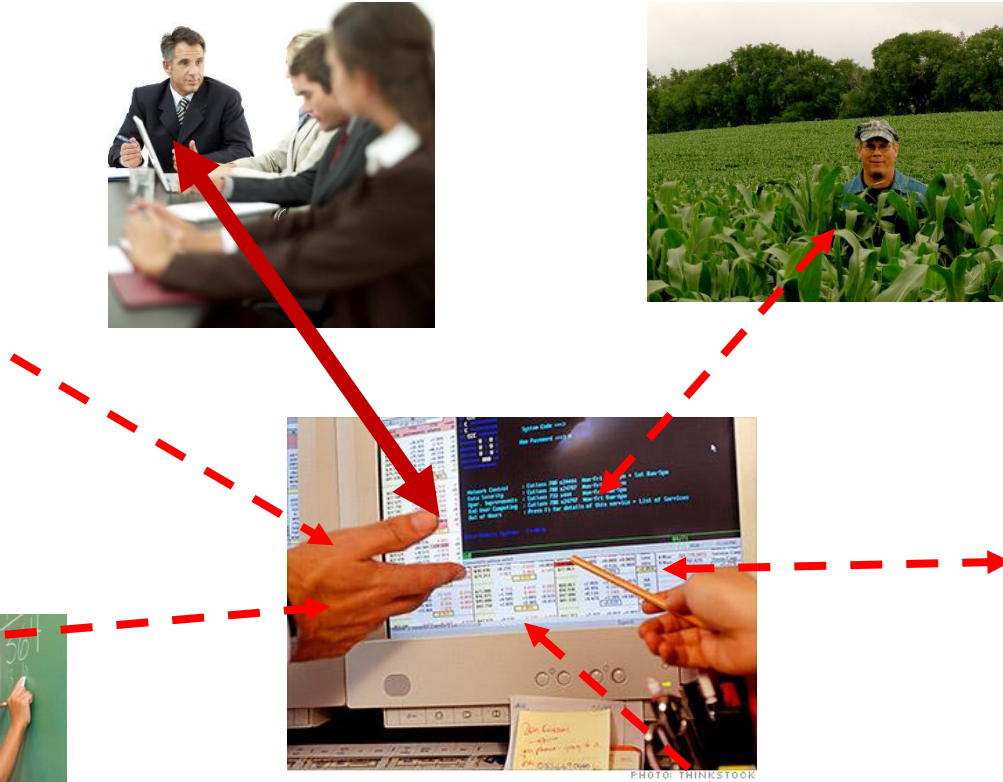
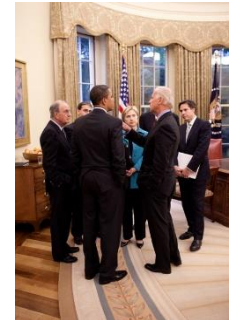
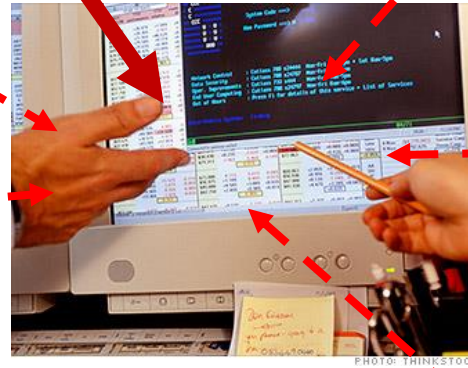
- Organize the world's **oncology data** and make it useful
- Aggregates and transforms **clinical** and **financial** data from EMR and billing systems in real-time
- A data **platform** that provides the most comprehensive view of a patient's experience



Data Scientist – what is this profession about?



Needed: understanding, insight and communication



Back to the world of **Small Data**

CARTE FIGURATIVE des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.

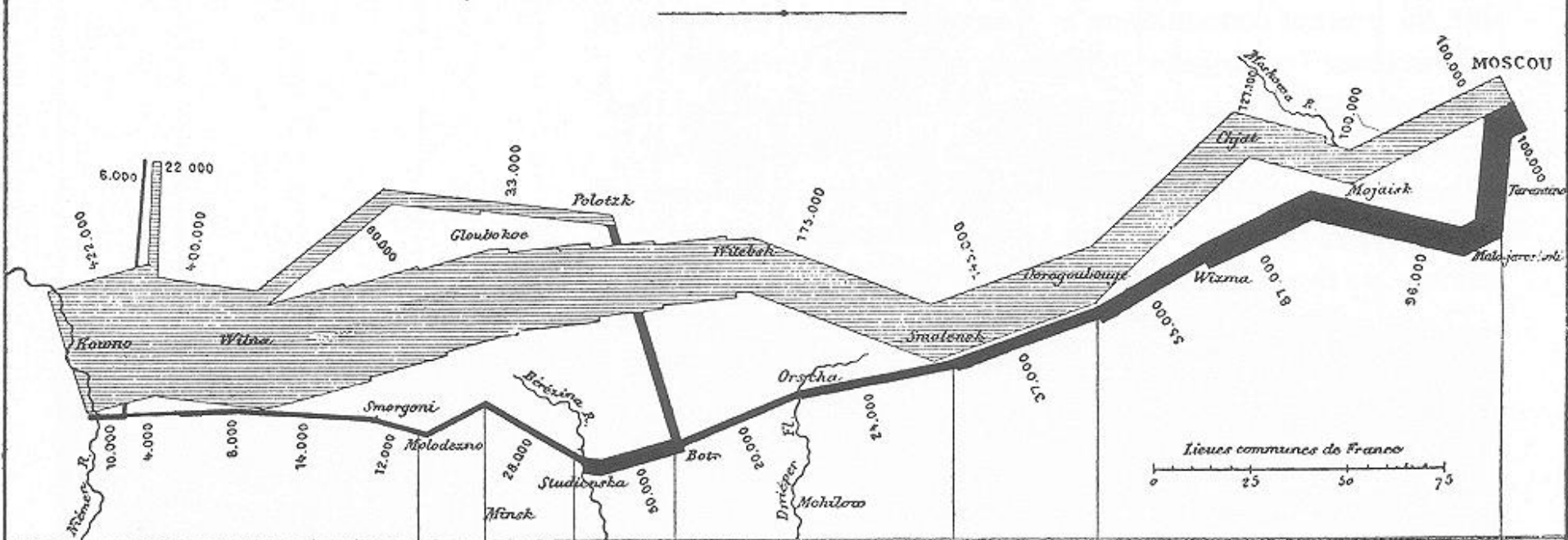
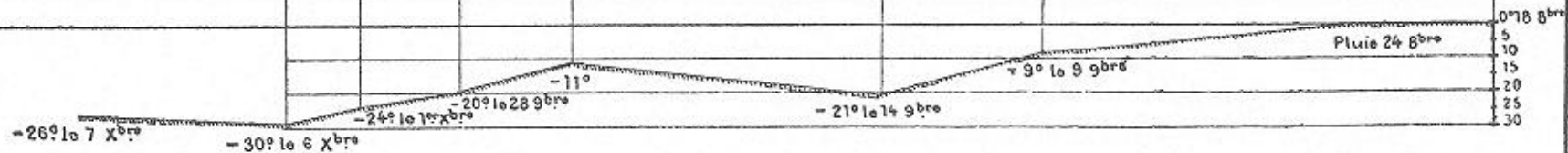


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro





Bögel György
bogelgy@ceubusiness.org