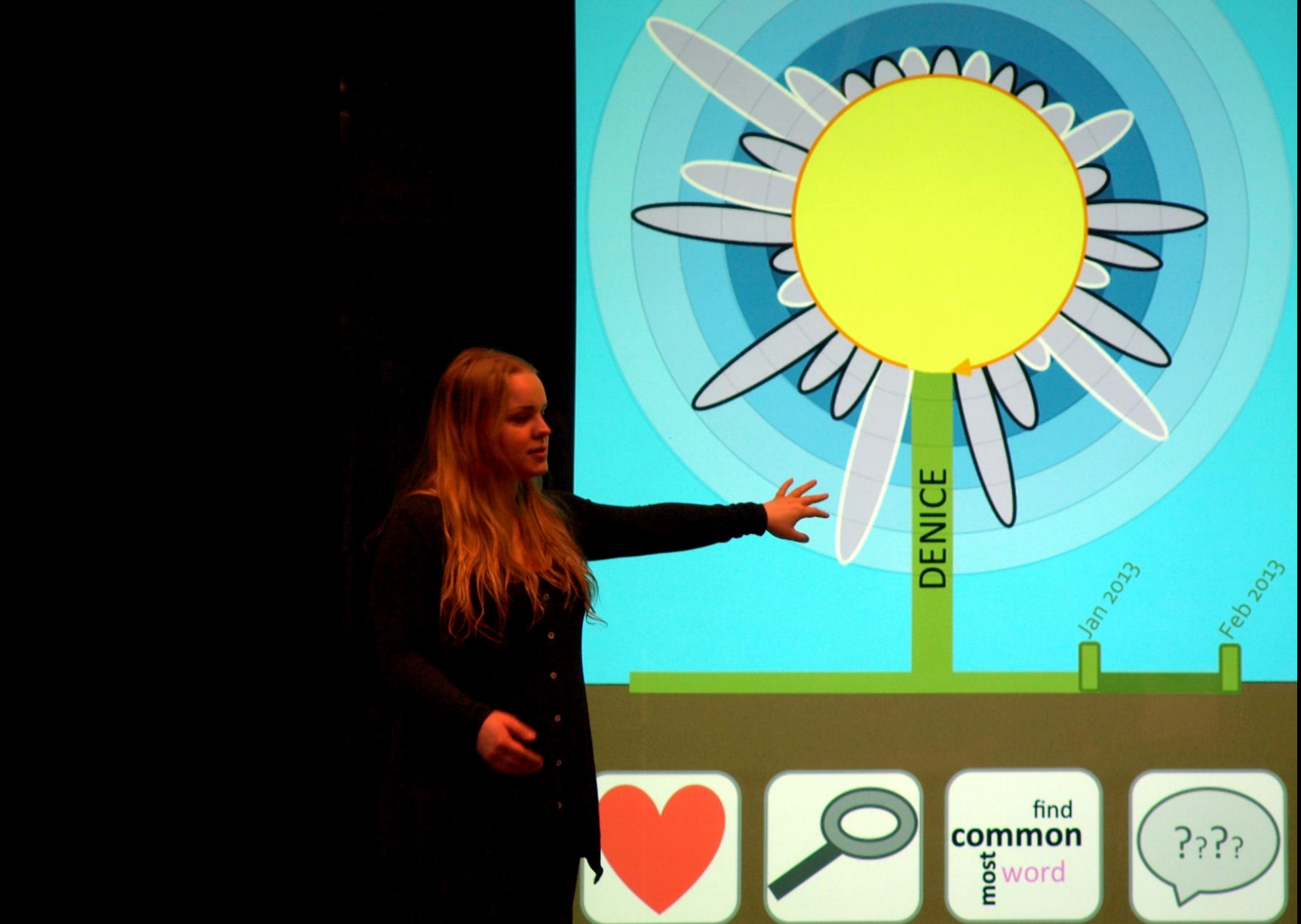
Information Visualization Lectures 2 - 3

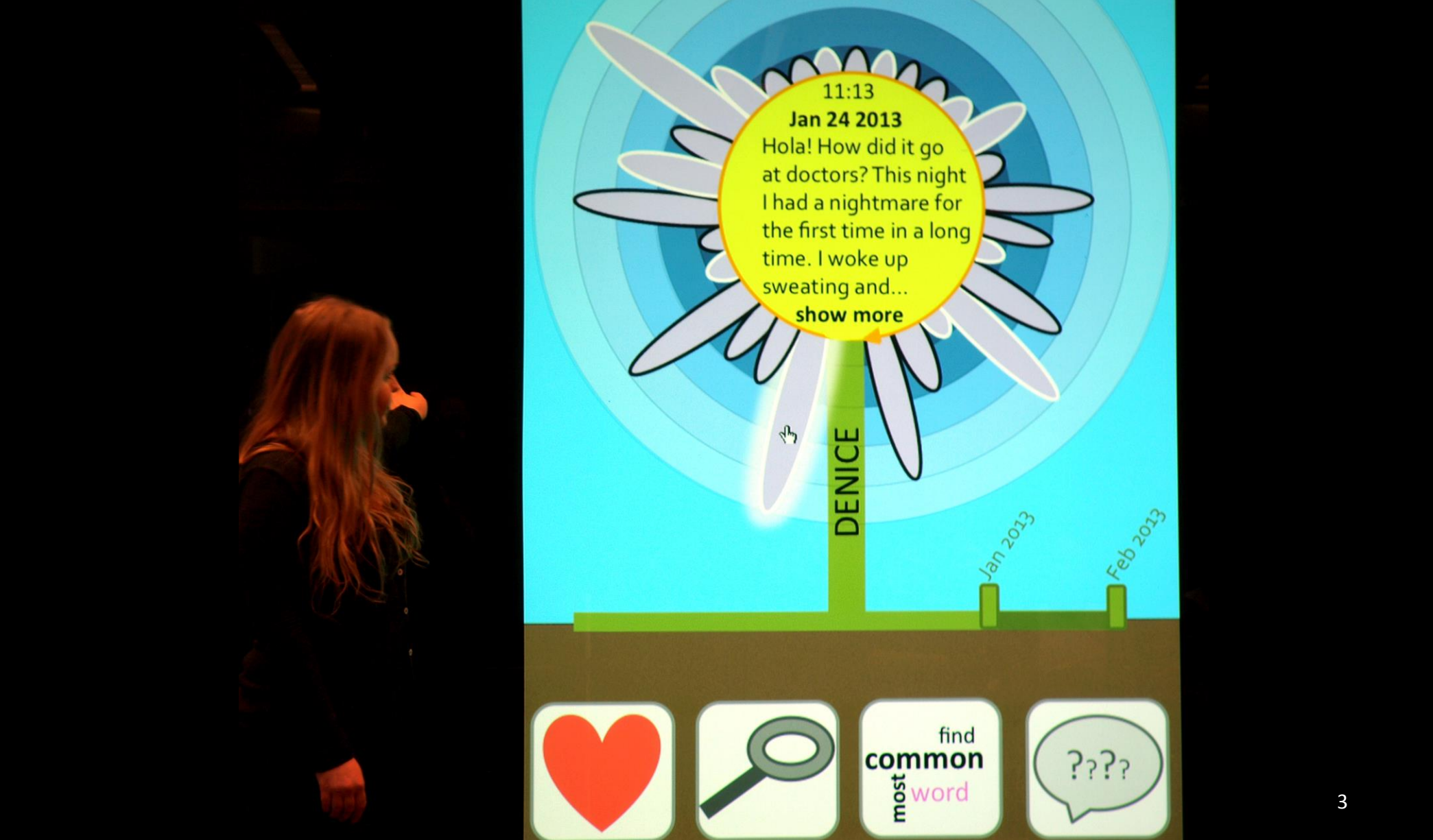
IVIS13 student Stephanie Dawoud presenting the final project Loves Me, Loves Me Not.

KTH VETENSKAP OCH KONST

VICSTHLM
VISUALISATION INTERACTION COLLABORATION

Mario Romero 2014/01/28


DENICE
Jan 2013
Feb 2013
find
common
most
word
????


11:13
Jan 24 2013
Hola! How did it go at doctors? This night I had a nightmare for the first time in a long time. I woke up sweating and...
show more
DENICE
Jan 2013
Feb 2013
find
most common word
???


WENT
SAID
BEFORE
I
SCREAMING
DON'T
OUT
CALL
DADDY
TIME
KNOW
ONE
HEARD
MOMMY
DENICE
Jan 2013
Feb 2013
find
most
common
word
???

Prelude Videos

- Microsoft Vision 2020 – link
- Precision Information Environments Envisioning the future of emergency management – link

Outline

1. Why Information Visualization
2. Discussion on Readings
3. Break: Zap the Bugs
4. Visualization Pipeline
5. Break: Oculus Rift
6. Project 2
7. Case Study: Milo
8. Case Study: Canvas Dance

WHY INFORMATION VISUALIZATION?

http://www.thehumanfaceofbigdata.com/

The trouble with data

https://www.emc.com/infographics/digital-universe-business-infographic.htm

Challenge

- Transform data into information
- Transform information into insight

Human Vision

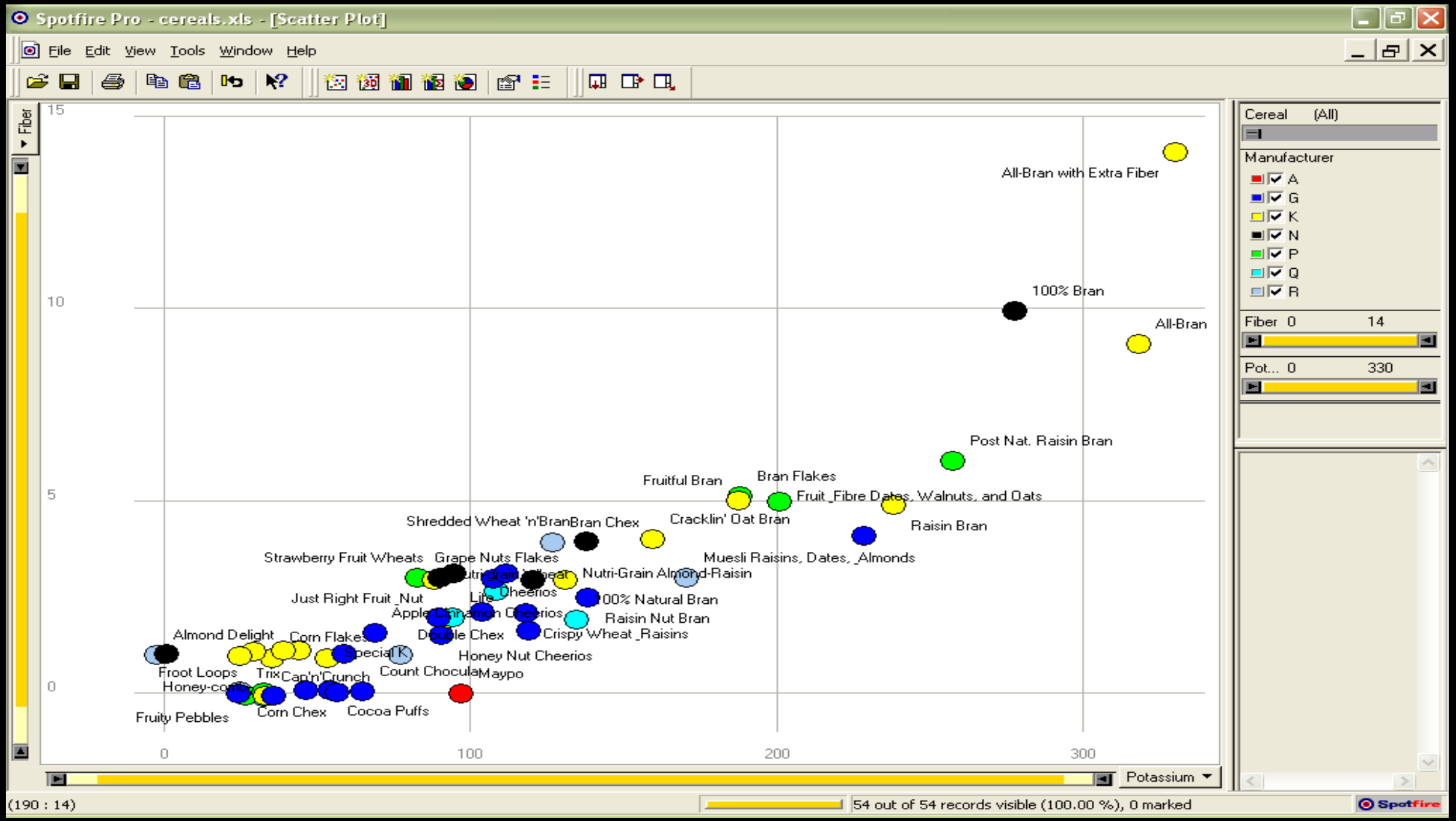
- Highest bandwidth
- Fast, parallel
- Pattern Recognition
- Pre-attentive
- Extends memory and cognitive capacity
- People think visually
- Brain: 30% vision, 8% touch, 3% hearing

Example

- Which cereals have the most and least potasium?
- Is there a relationship between potasium and fiber?

Cereal Data

	A	B	C	D
1	Cereal	Manufacturer	Fiber	Potassium
2	100% Bran	N	10	280
3	100% Natural Bran	Q	2	135
4	All-Bran	K	9	320
5	All-Bran with Extra Fiber	K	14	330
6	Almond Delight	R	1	0
7	Apple Cinnamon Cheeric	G	1.5	70
8	Bran Chex	R	4	125
9	Bran Flakes	P	5	190
10	Cap'n'Crunch	Q	0	35
11	Cheerios	G	2	105
12	Cocoa Puffs	G	0	55
13	Corn Chex	R	0	25
14	Corn Flakes	K	1	35
15	Count Chocula	G	0	65
16	Cracklin' Oat Bran	K	4	160
17	Cream of Wheat (Quick)	N	1	0
18	Crispy Wheat & Raisins	G	2	120
19	Double Chex	R	1	80
20	Froot Loops	K	1	30
21	Frosted Flakes	K	1	25
22	Fruit & Fibre Dates, Wal	P	5	200
23	Fruitful Bran	K	5	190
24	Fruity Pebbles	P	0	25
25	Golden Grahams	G	0	45
26	Grape Nuts Flakes	P	3	85
27	Honey Nut Cheerios	G	1.5	90
28	Honey-comb	P	0	35
29	Just Right Fruit & Nut	K	2	95
30	Life	Q	2	95
31	Lucky Charms	G	0	55
32	Maypo	A	0	95
33	Muesli Raisins, Dates, &	R	3	170
34	Multi-Grain Cheerios	G	2	90
35	Nutri-Grain Almond-Rais	K	3	130
36	Nutri-grain Wheat	K	3	90
37	Oatmeal Raisin Crisp	G	1.5	120
38	Post Nat. Raisin Bran	P	6	260
39	Product 19	K	1	45
40	Quaker Oatmeal	Q	2.7	110
41	Raisin Bran	K	5	240
42	Raisin Nut Bran	G	2.5	140
43	Rice Krispies	K	0	35
44	Shredded Wheat	N	3	95
45	Shredded Wheat 'n'Bran	N	4	140
46	Shredded Wheat spoon	N	3	120
47	Smacks	K	1	40
48	Special K	K	1	55
49	Strawberry Fruit Wheats	N	3	90
50	Total Corn Flakes	G	0	35
51	Total Raisin Bran	G	4	230
52	Total Whole Grain	G	3	110
53	Trix	G	0	25
54	Wheaties	G	3	110
55	Wheaties Honey Gold	G	1	60


Spotfire Pro - cereals.xls - [Scatter Plot]
File Edit View Tools Window Help
Fiber
15
10
5
0
0
100
200
300
All-Bran with Extra Fiber
100% Bran
All-Bran
Post Nat. Raisin Bran
Fruitful Bran
Bran Flakes
Fruit _Fibre Dates, Walnuts, and Oats
Raisin Bran
Cracklin' Oat Bran
Shredded Wheat 'n'Bran
Bran Chex
Muesli Raisins, Dates, _Almonds
Strawberry Fruit Wheats
Grape Nuts Flakes
Nutri-Grain Almond-Raisin
Just Right Fruit _Nut
Cheerios
100% Natural Bran
Raisin Nut Bran
Almond Delight
Corn Flakes
Double Chex
Crispy Wheat _Raisins
Special K
Honey Nut Cheerios
Froot Loops
Trix
Count Chocula
Maypo
Cap'n'Crunch
Corn Chex
Cocoa Puffs
Fruity Pebbles
Potassium
Cereal (All)
Manufacturer
A
G
K
N
P
Q
R
Fiber 0 14
Pot... 0 330
(190 : 14)
54 out of 54 records visible (100.00 %), 0 marked
Spotfire

Thought

- What if I read the data to you?

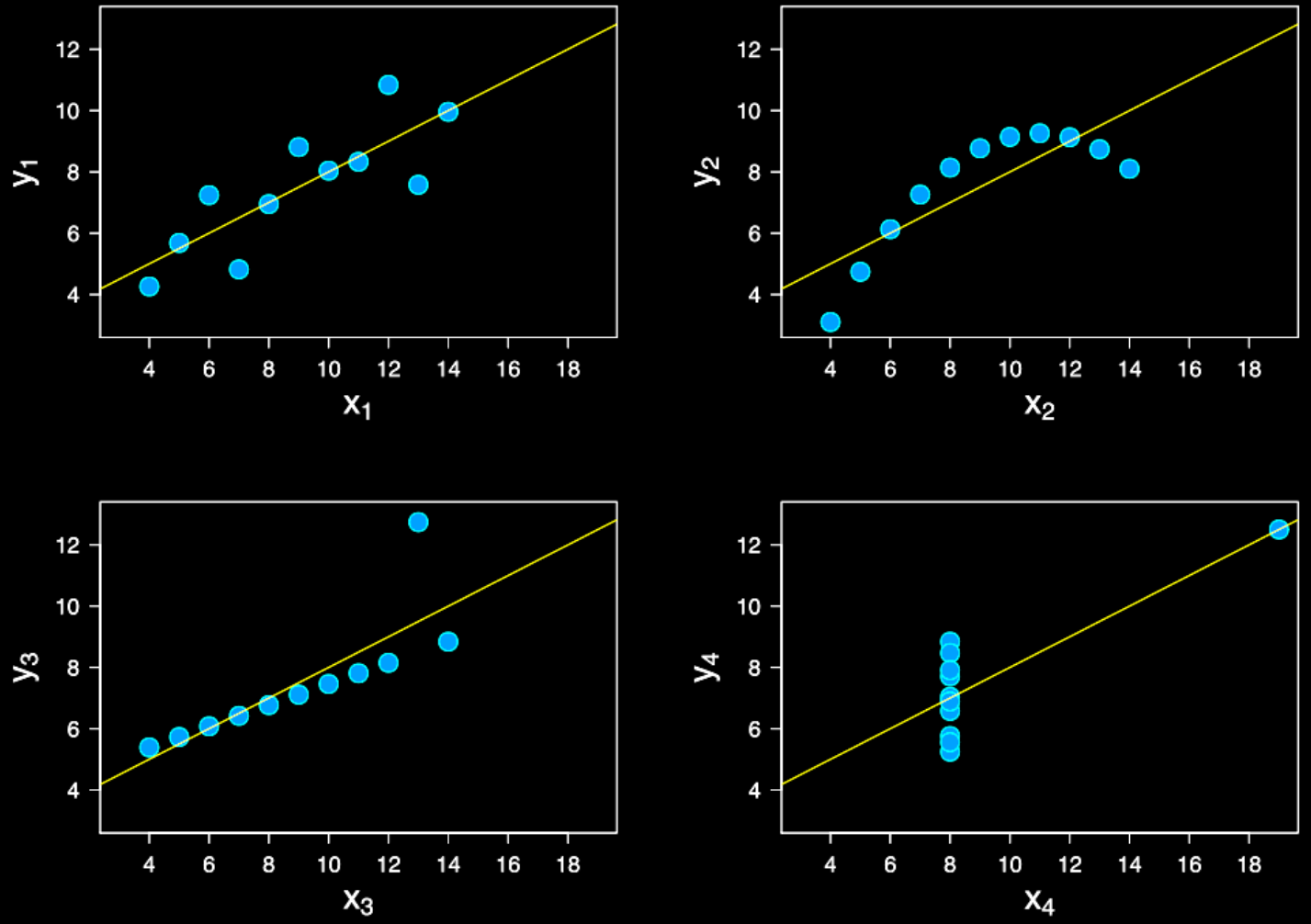
Anscombe's quartet

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
6.00	7.24	6.00	6.13	6.00	6.08	8.00	5.25
4.00	4.26	4.00	3.10	4.00	5.39	19.00	12.50
12.00	10.84	12.00	9.13	12.00	8.15	8.00	5.56
7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91
5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89

Statistics

Property	Value
Mean of *x* in each case	9 (exact)
Variance of *x* in each case	11 (exact)
Mean of *y* in each case	7.50 (to 2 decimal places)
Variance of *y* in each case	4.122 or 4.127 (to 3 decimal places)
Correlation between *x* and *y* in each case	0.816 (to 3 decimal places)
Linear regression line in each case	$y = 3.00 + 0.500x$ (to 2 and 3 decimal places, respectively)

Anscombe's quartet Visualized



Visualization

- “The use of computer-supported, interactive visual representations of data to amplify cognition.” From [Card, Mackinlay Shneiderman ‘98]

Visualization

- Often thought of as process of making a graphic or an image
- Really is a cognitive process
  - Form a mental image of something
  - Internalize an understanding
- “The purpose of visualization is insight, not pictures”
- Insight: discovery, decision making, explanation

Main Idea

- Visuals help us think
  - Provide a frame of reference, a temporary storage area
- Cognition → Perception
- Pattern matching
- External cognition aid
  - Role of external world in thinking and reason

Larkin & Simon ’87 Card, Mackinlay, Shneiderman ‘98

Formally

“Contained within the data of any investigation is information that can yield conclusions to questions not even originally asked. That is, there can be surprises in the data…To regularly miss surprises by failing to probe thoroughly with visualization tools is terribly inefficient because the cost of intensive data analysis is typically very small compared with the cost of data collection.”

W. Cleveland The Elements of Graphing Data

Purpose

- Analysis

Understand your data better and act upon that understanding

1. Presentation

Communicate and inform others more effectively

Analysis

- Find extremes
- Mean, var, std, ste,...
- Relations
- What’s missing
- Identify ambiguity and noise
- Commonality, mode

When to apply InfoVis?

- Other techniques:
  - Statistics
  - Data mining
  - Machine Learning
- InfoVis:
  - Exploratory data analysis
    - Don’t know what you are looking for
    - Don’t have a hypothesis
    - Want to know what question to ask

“A graphic display has many purposes but it achieves its highest value when it forces us to see what we were not expecting.”

H. WAINER

InfoVis Tasks

- Search
  - Find a specific piece of information
    - How many games has Sweden won in the world cup?
    - How many rental apartments are available in Stockholm?
- Browse
  - Look over or inspect something in a more casual manner
    - Learn about nutrition
    - How does the weather affect transportation in Stockholm?

InfoVis Taks (cont.)

- Analysis
  - Compare
  - Contrast
  - Outliers
  - Extremes
  - Patterns
- Assimilation
- Monitoring
- Awareness

Presentation

- Use visualization to communicate
  - Ideas
  - Influence
  - Explain
  - Persuade
- Evidence and support
- Summarize
- Aggregate
- Unite

Two Key Challenges of InfoVis

- Scale
  - Large datasets
  - Datasets with largely varying scales
    - Seconds
    - Days
    - Years
    - Centuries
- Diversity
  - Data types
  - Forms
  - Sizes

Ben Shneiderman

Bob Amar and John Stasko

Instructions

1. If you did not read the papers, please, go to the hallway and come back when you are done reading them.
2. Else, form groups of 3 (or 4) students
3. Discuss the main points of the papers
4. Summarize the 3 main contributions from each paper
5. Post this both on the Facebook wall and the KTH social wall
6. Be ready to present to the entire group
7. You have 10 minutes

Readings' Contributions

Harry Schröter, Ivo van Bon, Wouter Jansen

MaGer MaGer Rafel Saad Shaun Basil Mendonsa

The main idea of Ben Shneiderman paper is: Overview first, zoom and filter, then details-on-demand.
Overview: gain an overview of the entire collection.
Zoom: zoom in on items of interest.
Filter: filter out uninteresting items.
Details-on-demand: select an item or group and get details when needed.

He categorizes the tasks into: 1-dimensional, 2-dimensional, 3-dimensional, temporal, multi-dimensional, tree and network.

The paper of Amar and Stasko concerns the Rationale Gap, the gap between perceiving a relationship and actually being able to explain confidence in that relationship and the usefulness of that relationship and the Worldview Gap, the gap between what is being shown and what actually needs to be shown to draw a straightforward representational conclusion for making a decision.
In this paper, the author gives ways to avoid these gaps. For example, a way to avoid the Worldview Gap is to a system can help bridge the Worldview Gap by providing support for the formulation and verification of user hypotheses.

Readings' Contributions

Tommy Feldt
Karenina Gunnarson
Daniel Molin
Daniel Månsson

Schneiderman - The eyes have it:

1. Visual information seeking mantra: A design guideline for all types of information visualization: "Overview first, zoom and filter, then details on demand". It is basically how the human mind searches for information and it is therefore important for that any visualization to follow it. It could be used both as a design principle and as a heuristic for evaluating visualizations.
2. Data type: The author introduces 7 different data types that a visualization may involve: 1-dimensional, 2-dimensional, 3-dimensional, temporal, multidimensional, tree, network.
3. Visualization tasks: The author presents 7 visualization tasks: overview, zoom, filter, details-on-demand, relate, history, extract. Any general visualization tool should support all of these.

Amar & Stasko - A knowledge task-based framework for design and evaluation of information visualizations:

1. Why visualizations fail: Three reasons why many information visualizations currently fail to support decision making: limited affordances of the visualization systems (the systems can not do what the users need them to do), predetermined representations (the representations are not agile/adaptable) and a decline of determinism in decision-making (visualizations fail to deal with uncertainty in the data).
2. Identifying gaps: The authors identify 2 analytic gaps that limit current systems: the rationale gap (i.e. a problem with the relationship between the perceived and what is represented), and the worldview gap (i.e. the gap between what is presented and what the user needed to be presented)
3. Bridging the gaps: The authors present a number of concrete suggestions for strategies to bridge the gaps. For instance, "expose uncertainty" "providing support for discovery" and "concretize relationship"

The Eyes Have It

The Visual Information Seeking Mantra
(Overview first, zoom and filter, then detail-on-demand)

Task by Data Type Taxonomy
Match data with relevant/useful visualization methods to successfully visualize the data.

Advanced Filtering
It's difficult to design an interface with high usability when implementing Boolean filtering.

A knowledge task-based framework for Design and Evaluation of Information Visualization
Presents two issues: The Rationale Gap and The Worldview Gap

The Rationale Gap – The difference between perceived correlations and causality and being able to explain confidence and usefulness of a relationship.

The Worldview Gap – the difference between what is shown and what needs to be shown.

For each gap, three tasks are proposed in order to "bridge" the gaps.

Rationale Gap:
- Expose Uncertainty
- Concretize Relationships
- Formulate Cause and Effect

Worldview gap
- Determination of Domain Parameters
- Multivariate Explanation
- Confirm Hypotheses

These tasks can be used both for design and evaluation existing designs.

Johan Blomgren, Johan Wikström, Rasmus Ansin

Readings' Contributions

Key points from group Jacob Håkansson, Ida Renström, Fiorella Grados and Hanna Hasselqvist.

The Eyes Have It:
- **The mantra: Overview first, zoom and filter, then details-on-demand**
- **"History" and "Extract" are said to be rare in information visualizations (however nowdays these are more common functions)**
- **Research typically deals with one type of data and the real world requires many**

A Knowledge Task-Based Framework:
Division into tasks (based on common gaps in information visualisations):
- **Rationale: Expose Uncertainty, Concretize Relationships, Formulate Cause and Effect**
- **Worldwide: Determination of Domain Parameters, Multivariate Explanation, Confirm Hypotheses**
- **Using the tasks for design and evaluation**

The eyes have it:

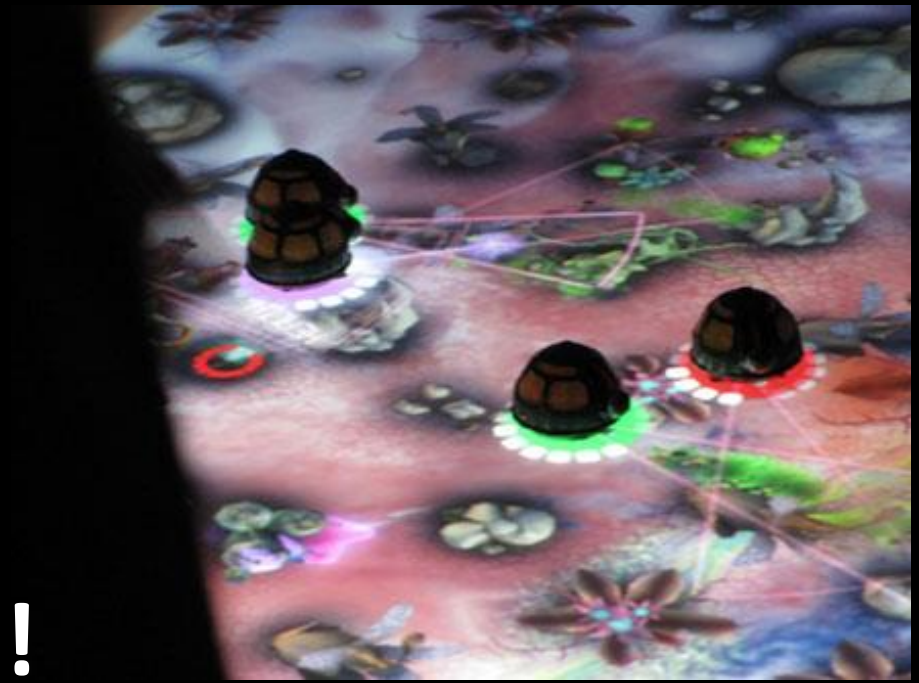
1. The visual information seeking mantra: "Overview first ,zoom and filter, then details-on-demand".
2. In the article Shneiderman lists different data types and task.
3. The computer is potentially the magic lens for finding, sorting, filtering, and presenting the relevant items. It's good with visual interfaces.

A Knowledge Task-Based Framework for Design and Evaluation of Information Visualization

1. There could be gaps between what the visualization show and the insights you can draw. The authors call them rationale gaps and world gaps. This way the limitations of the visualization systems are classified.
2. 2. The authors present tasks that could help bridging these gaps. 3. These tasks could be used to evaluate a visualization.
3. //Anna, Moa and Terese

Break (Bonus: Dispersion)

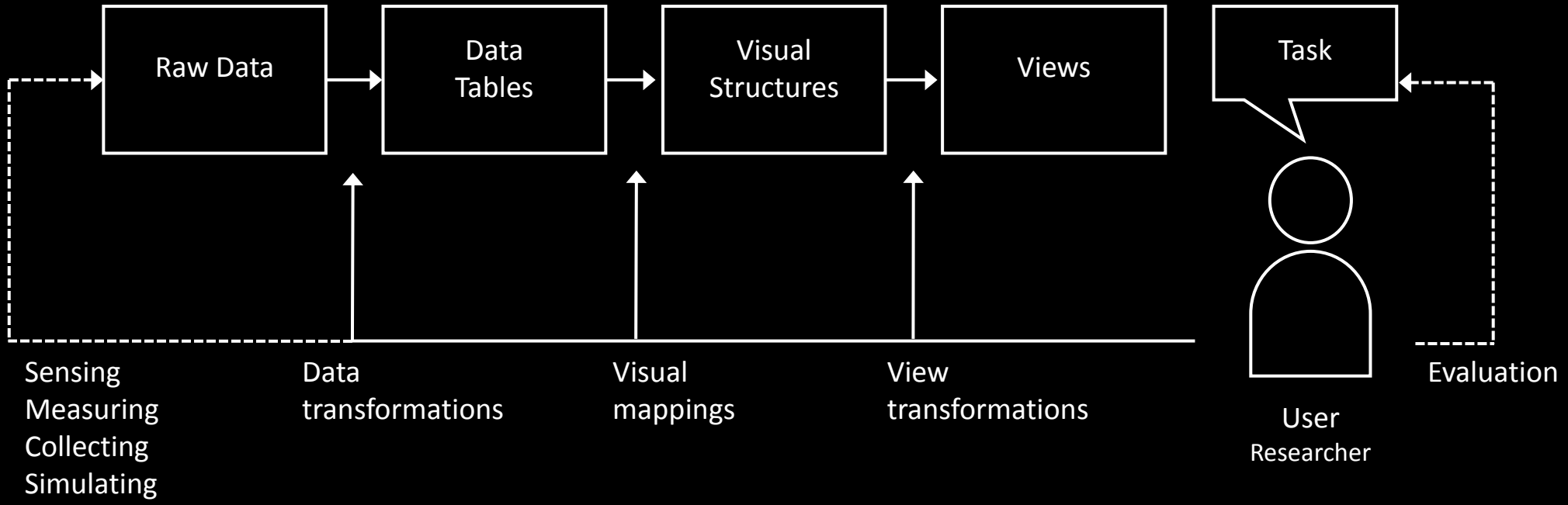
ZAP THE BUGS!



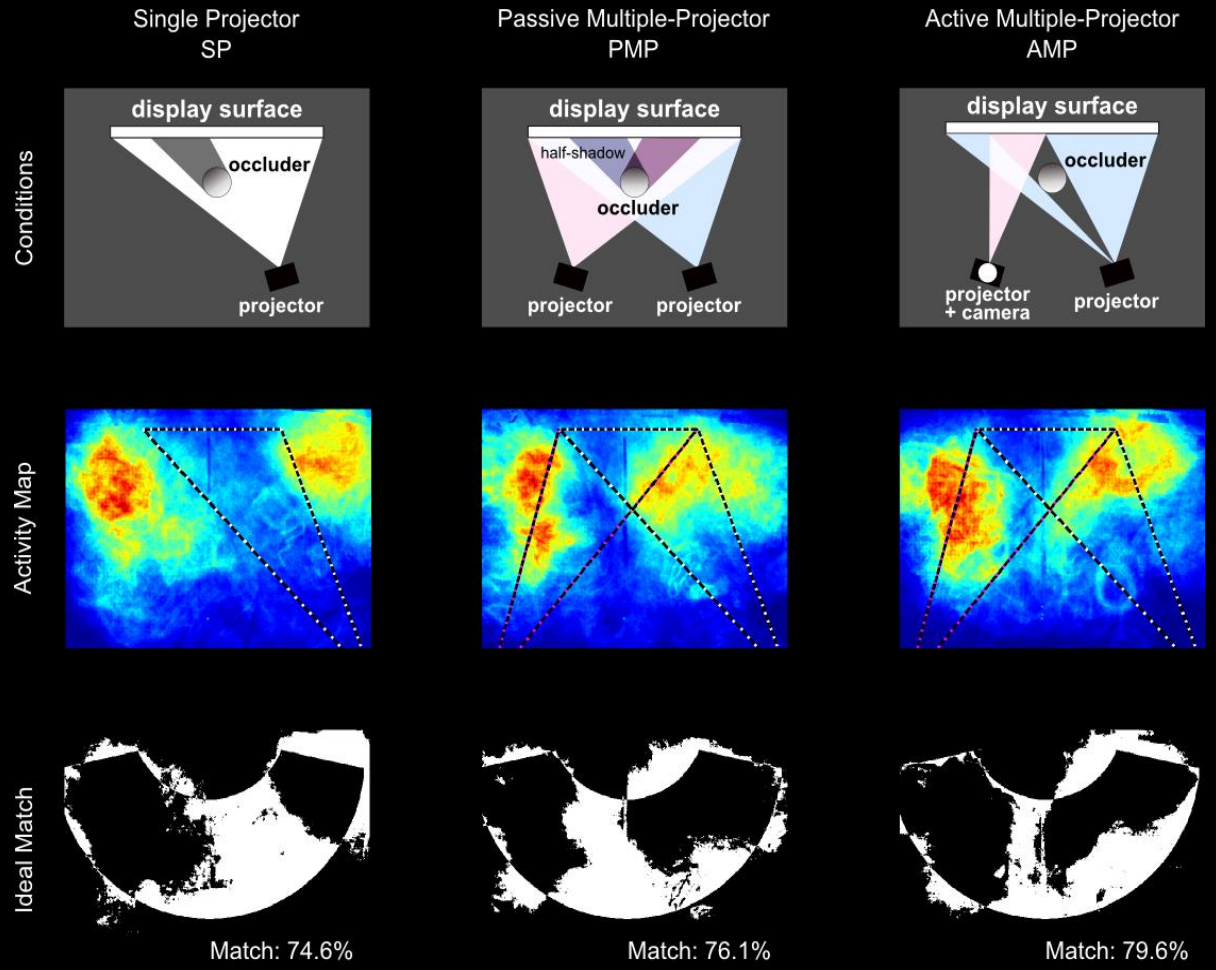
INFORMATION VISUALIZATION PIPELINE

Information Visualization Pipeline

expanded from **Readings in Information Visualization: Using Vision to Think**
By Stuart K. Card, Jock D. Mackinlay, Ben Shneiderman, 1999

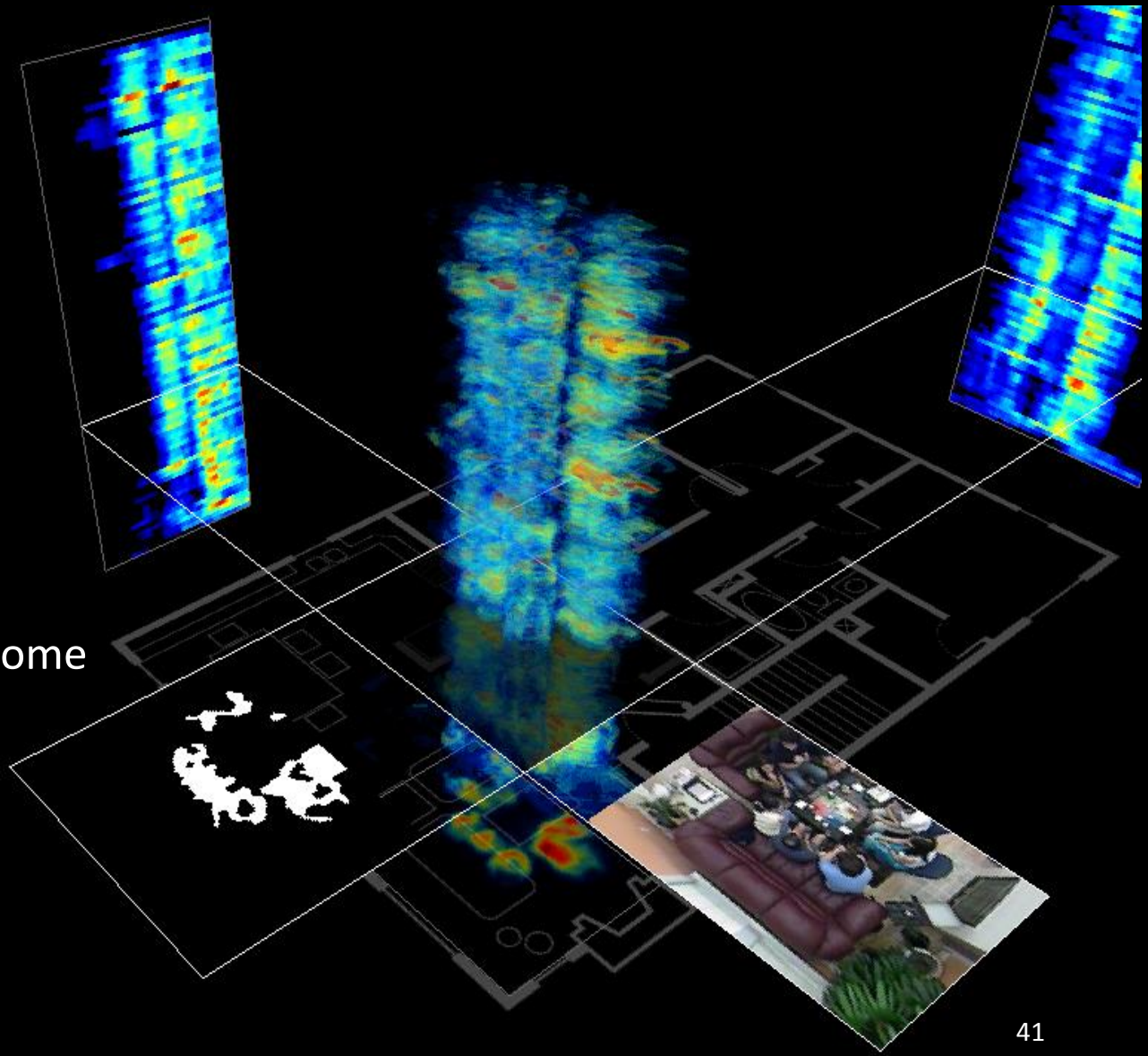


Virtual Rear Projection Study

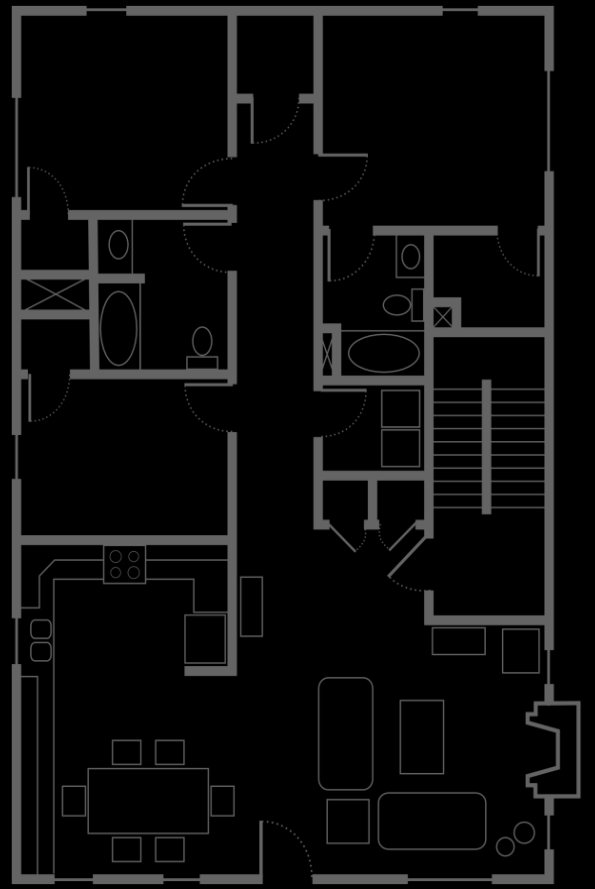


Visualizing Behavior in the Home

VIZ-A-VIS



Capturing and Visualizing Behavior



Aware Home

Compute and Aggregate Motion

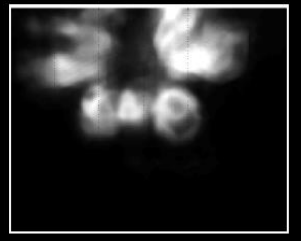
10 cameras in public areas

- **Image ←→ Space**
- **Fixed background**
- **High Resolution**

F_{t-1}

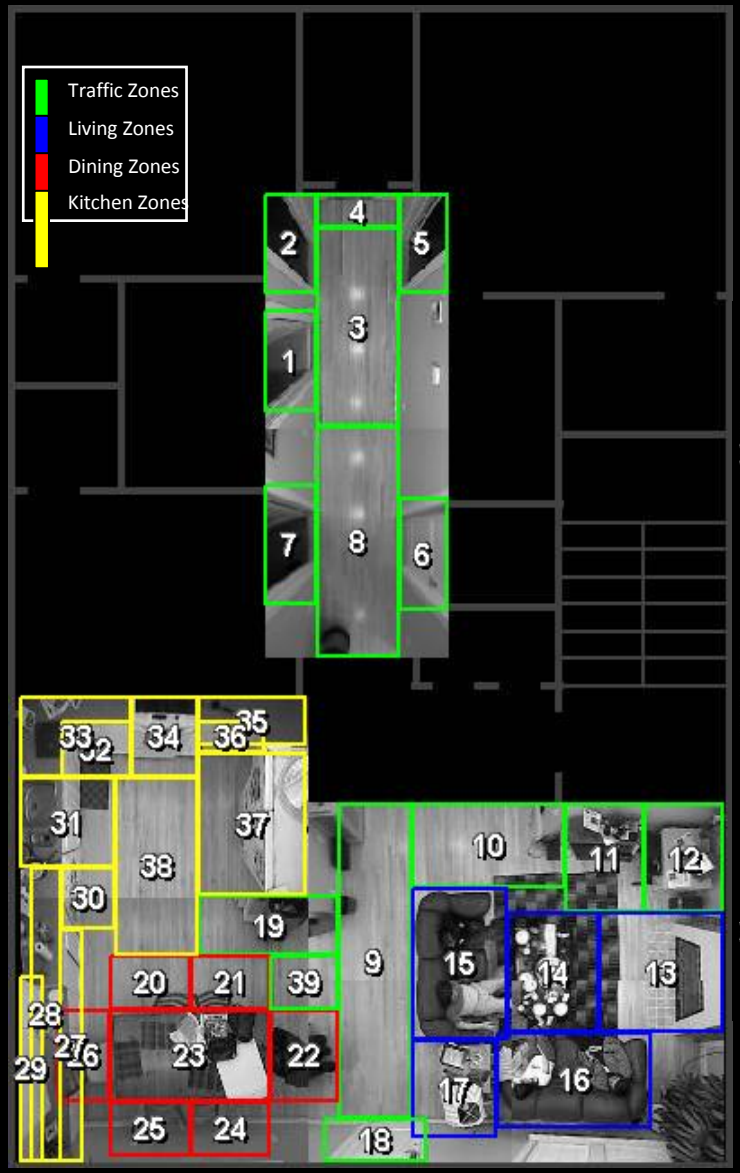
F_t

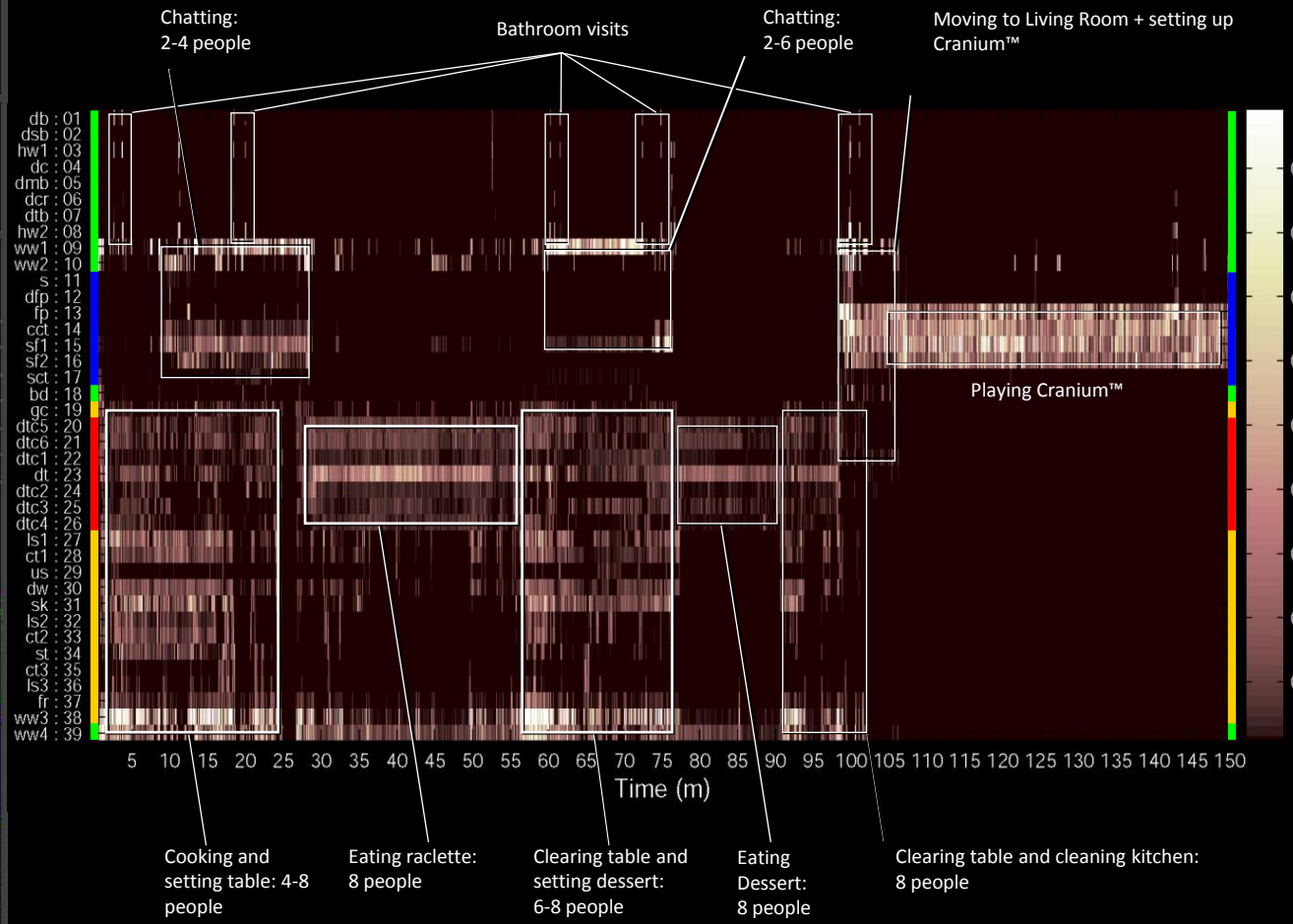
$AFDt = |F_t - F_{t-1}| > \Theta$



$\sum_t AFDt$

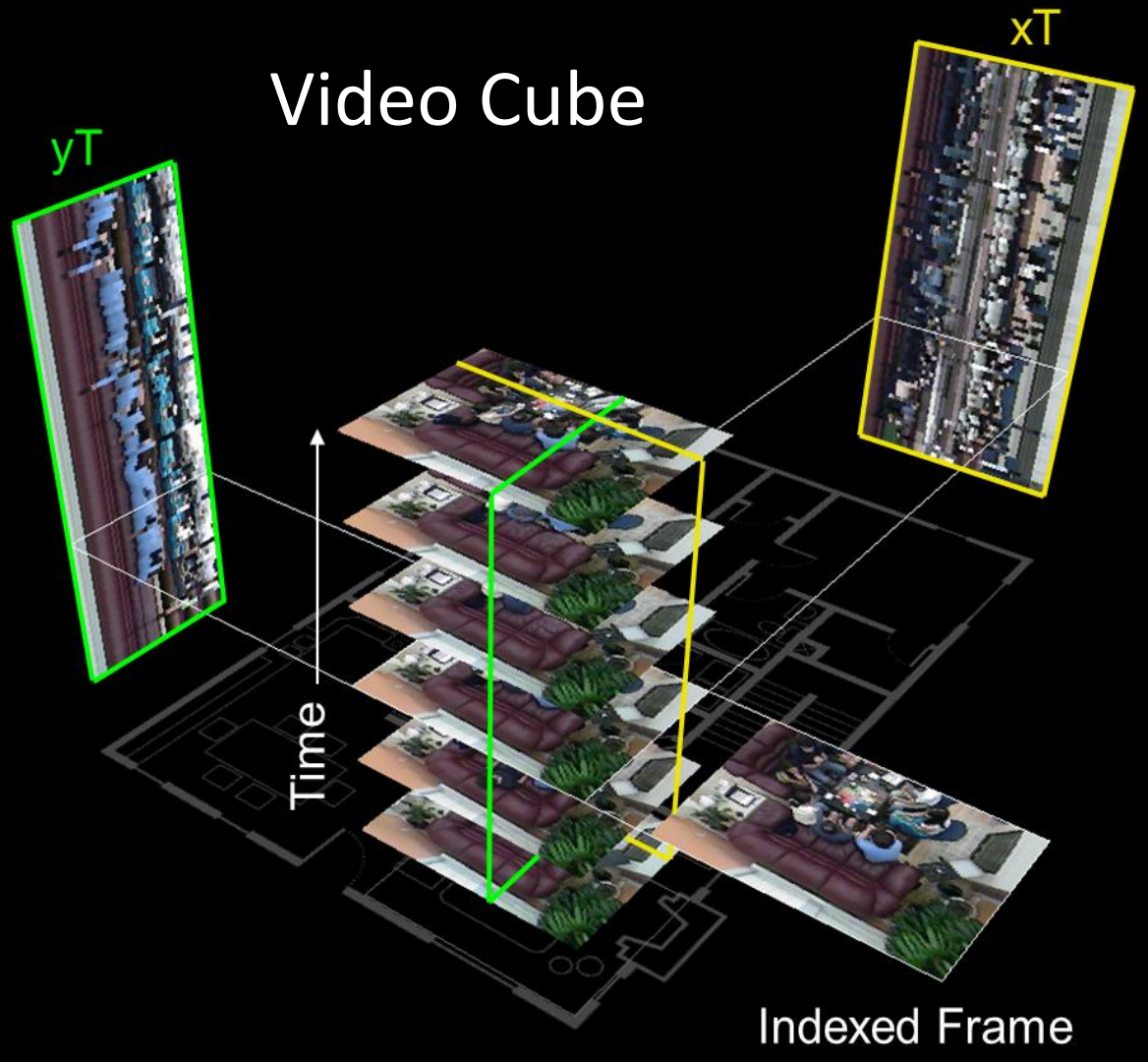
Activity Table

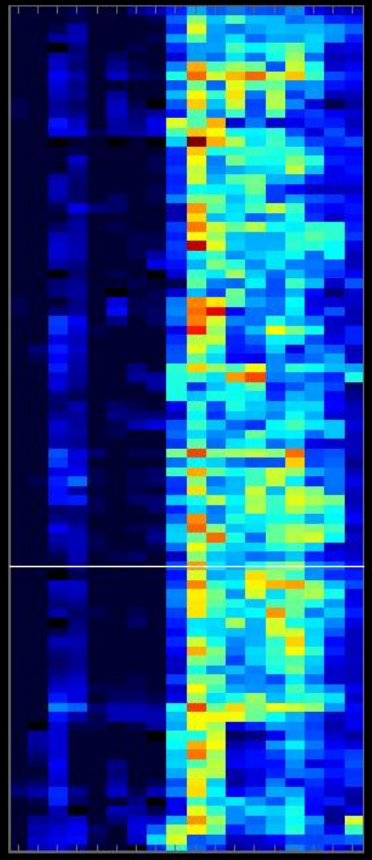


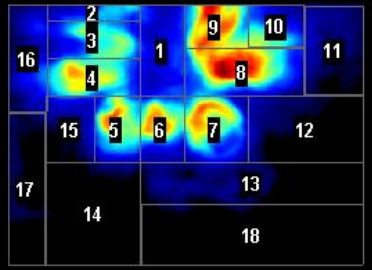


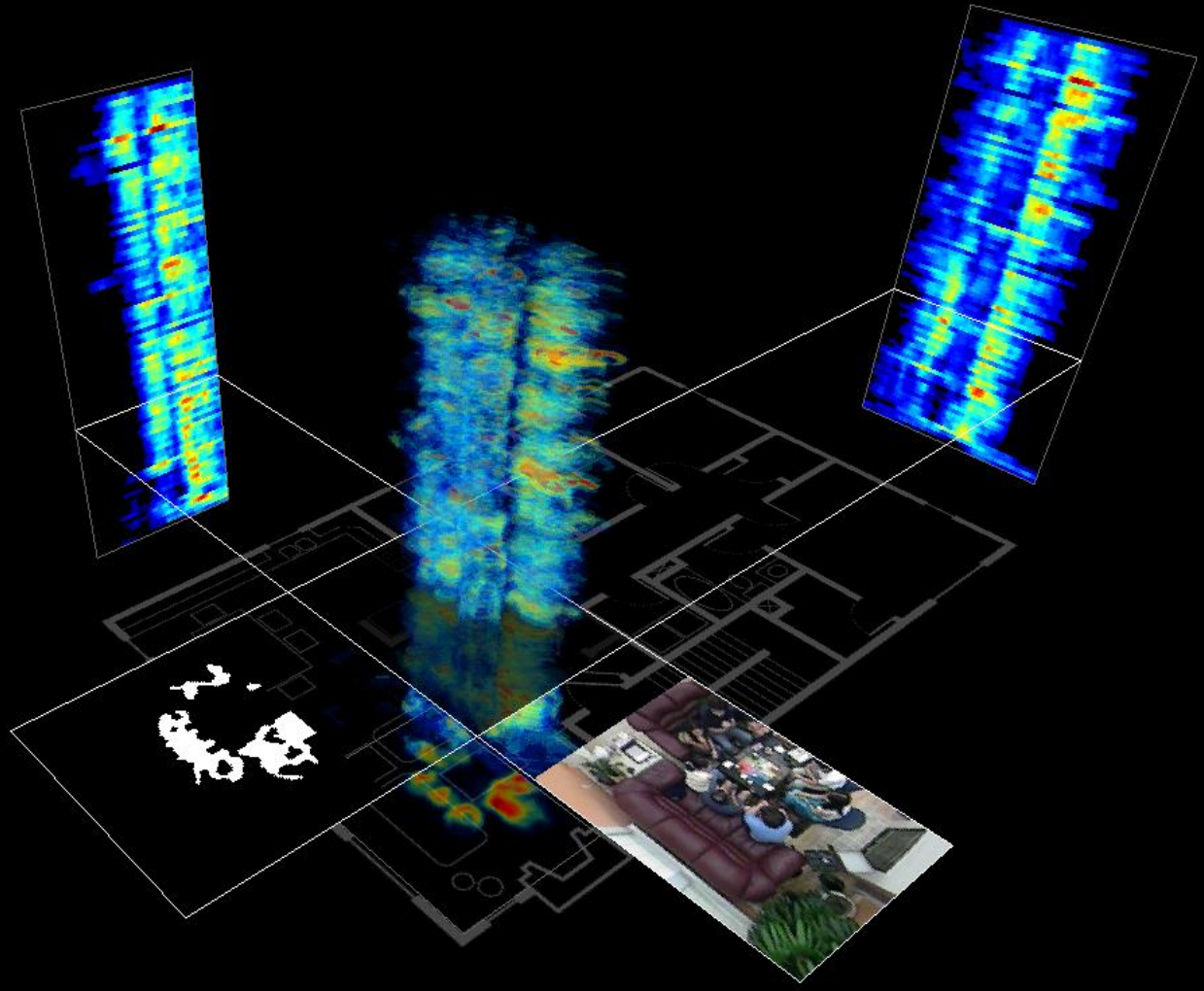
Video Cube

Playback Console







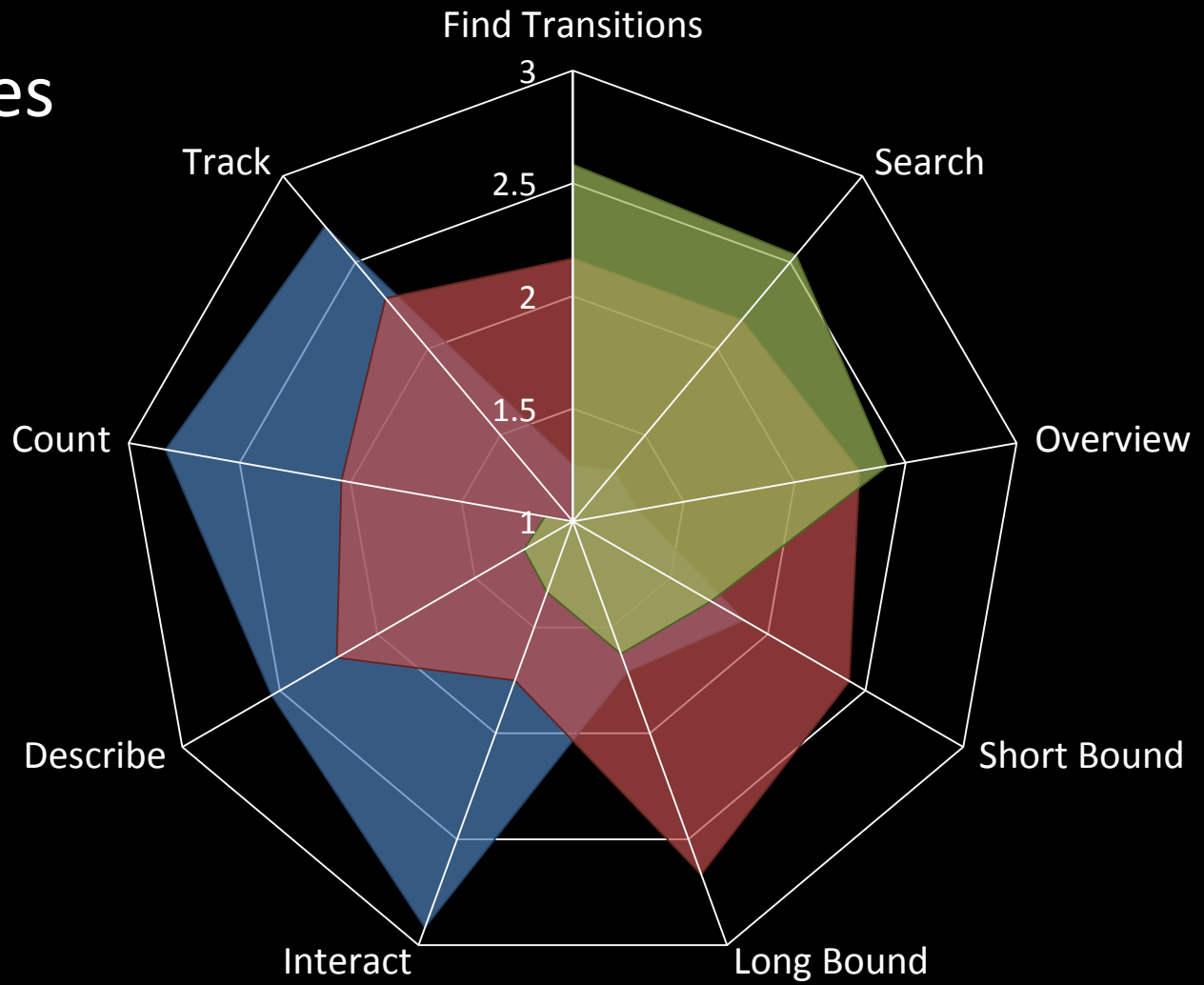


Activity Cube

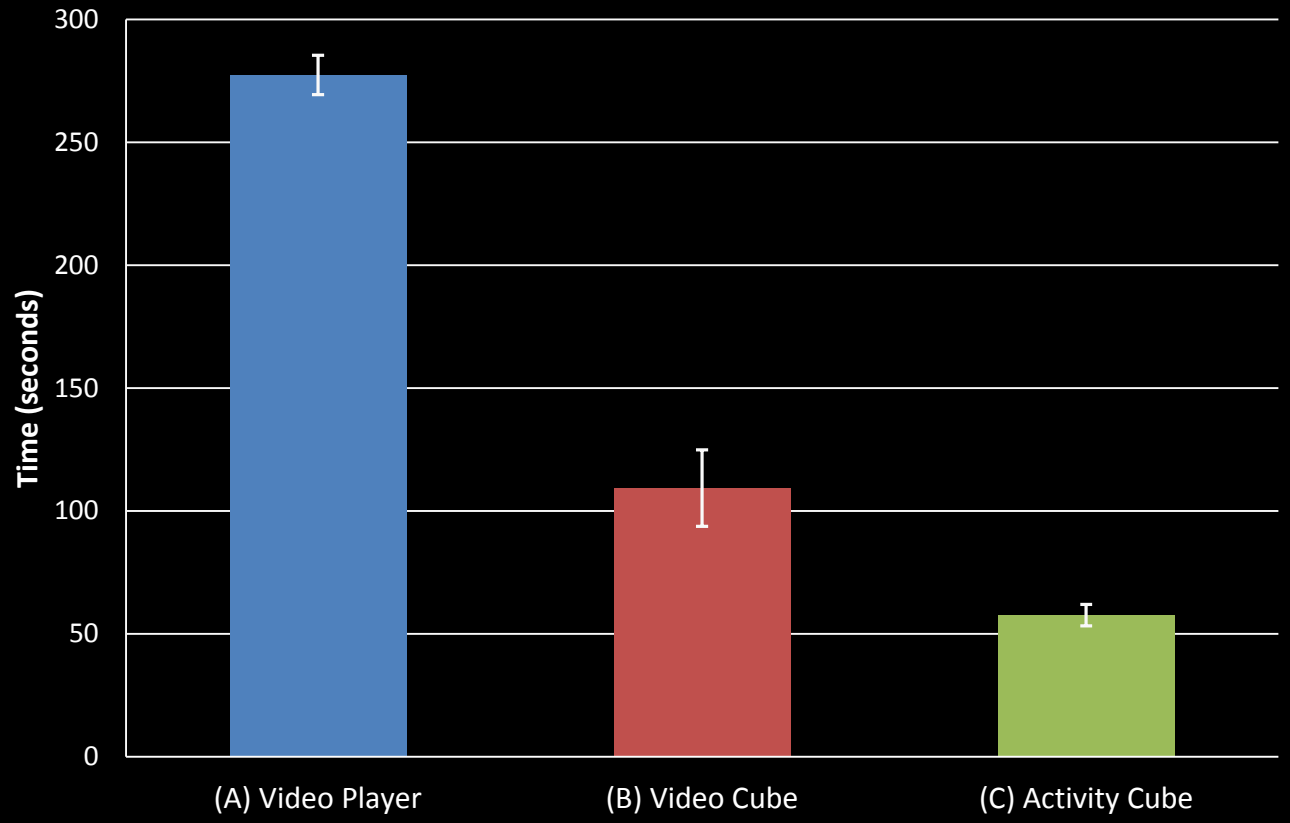
Evaluating Video Visualizations of Human Behavior

Mario Romero
Alice Vialard
John Peponis
John Stasko
Gregory Abowd

User Studies



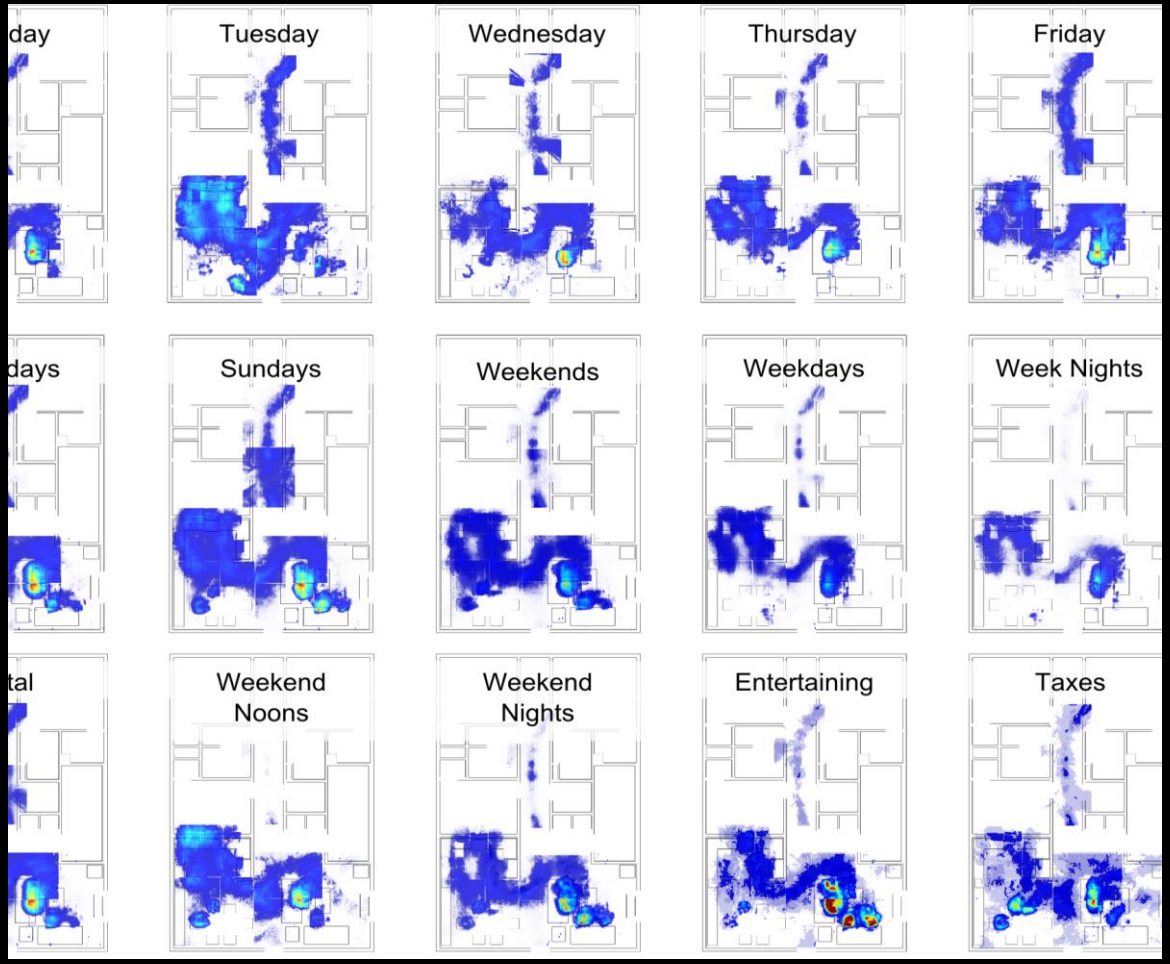
Search Time


300
250
200
150
100
50
0
Time (seconds)
(A) Video Player
(B) Video Cube
(C) Activity Cube

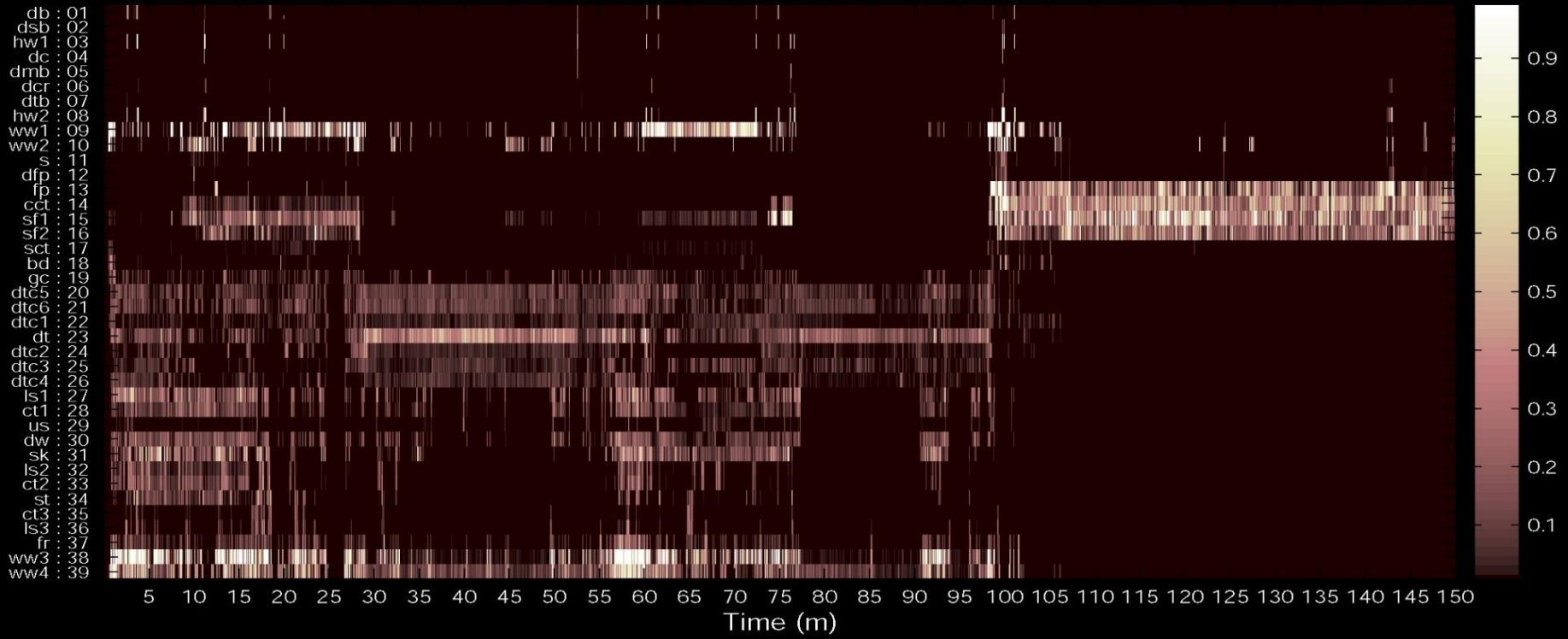
Architecture Study

Data

Behavioral Pattern Discovery



Occupancy Patterns



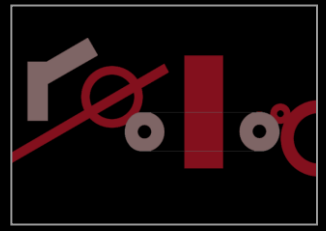
An Art Installation for the Home

TABLEAU MACHINE

Tableau Machine

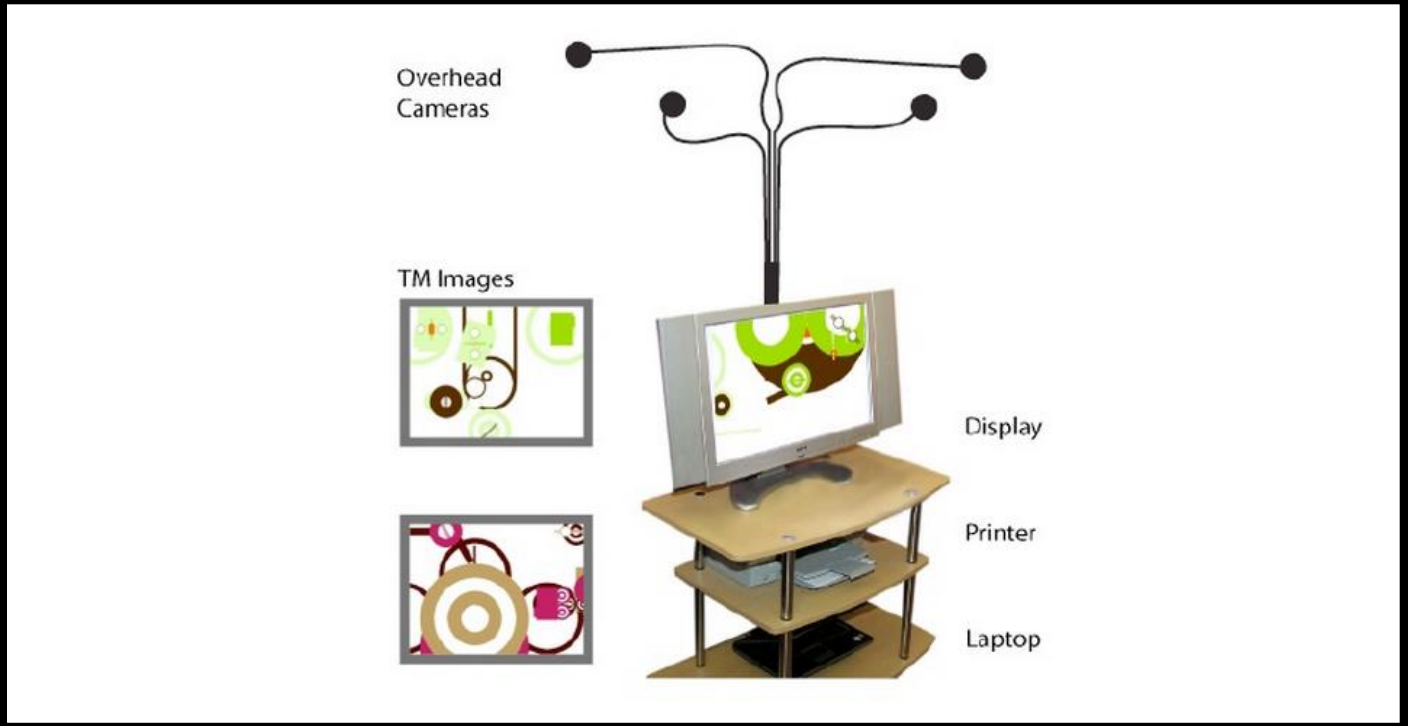
everyday activity

sensing and recognizing



creative visualization

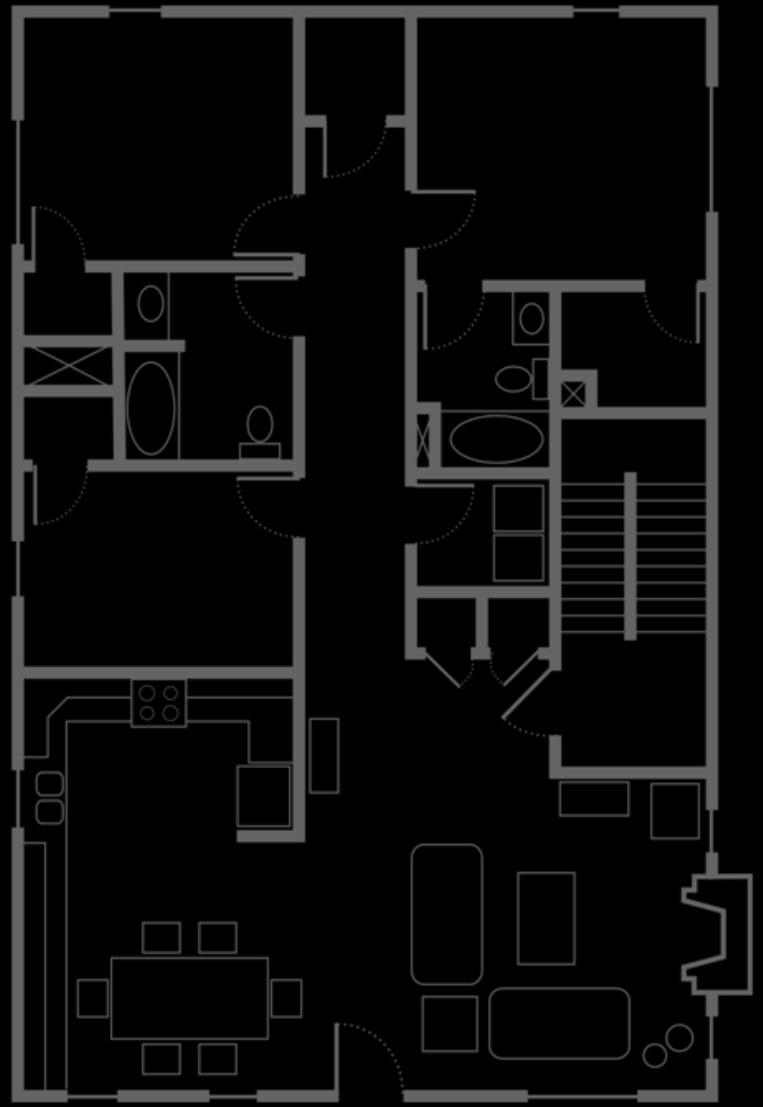
Tableau Machine



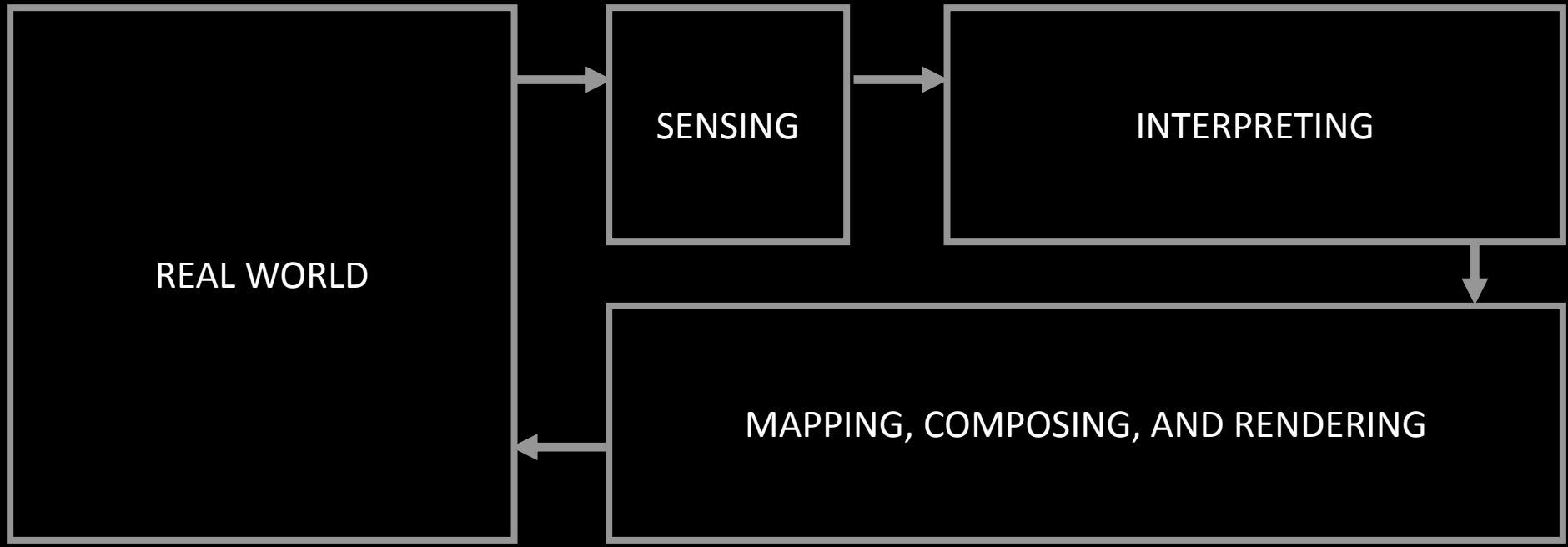
AWARE HOME RESEARCH INITIATIVE

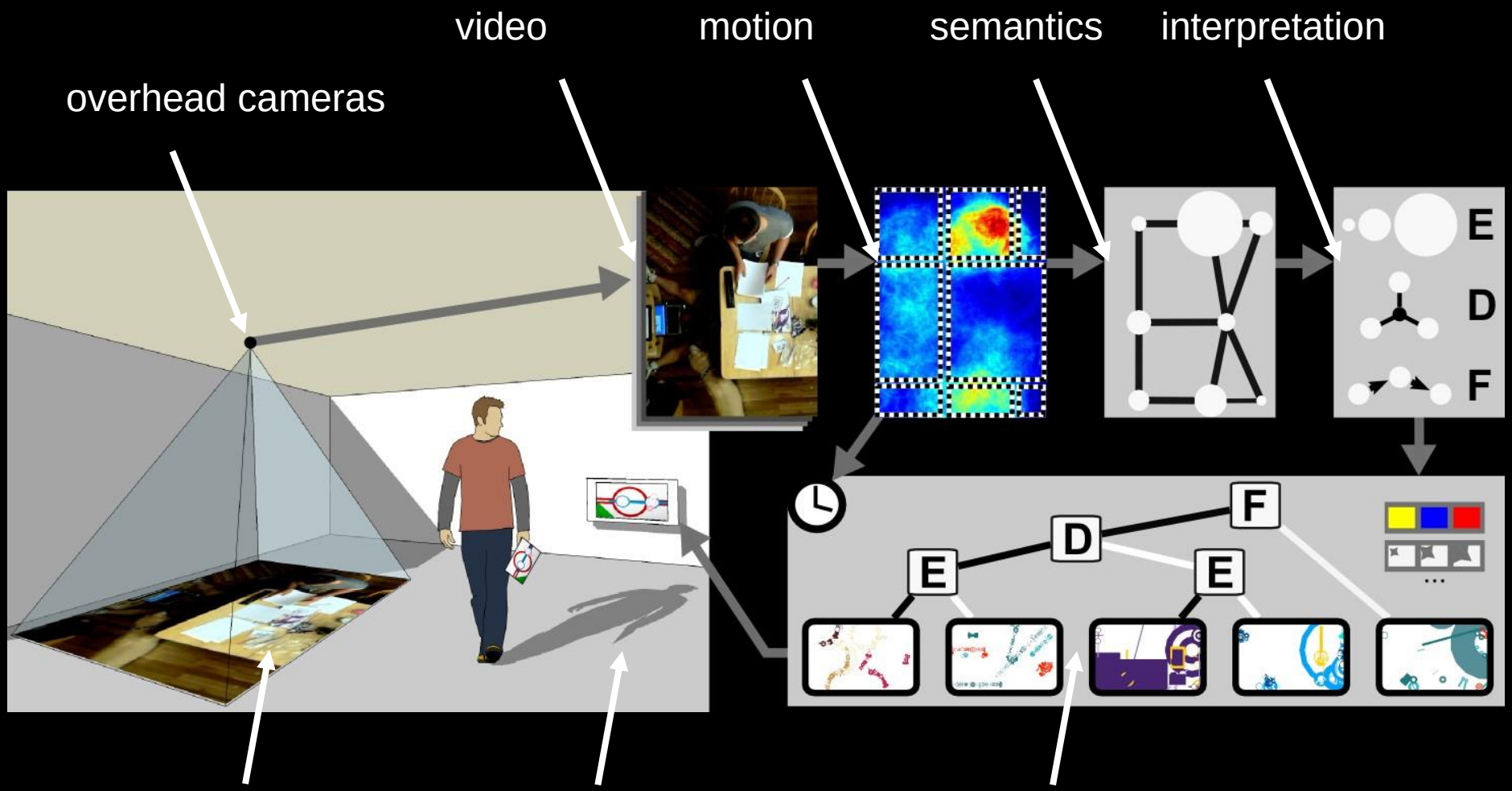
“The Home is a machine for living in….”
Le Corbusier (1923)

“Space is the machine.”
Hillier (1996)

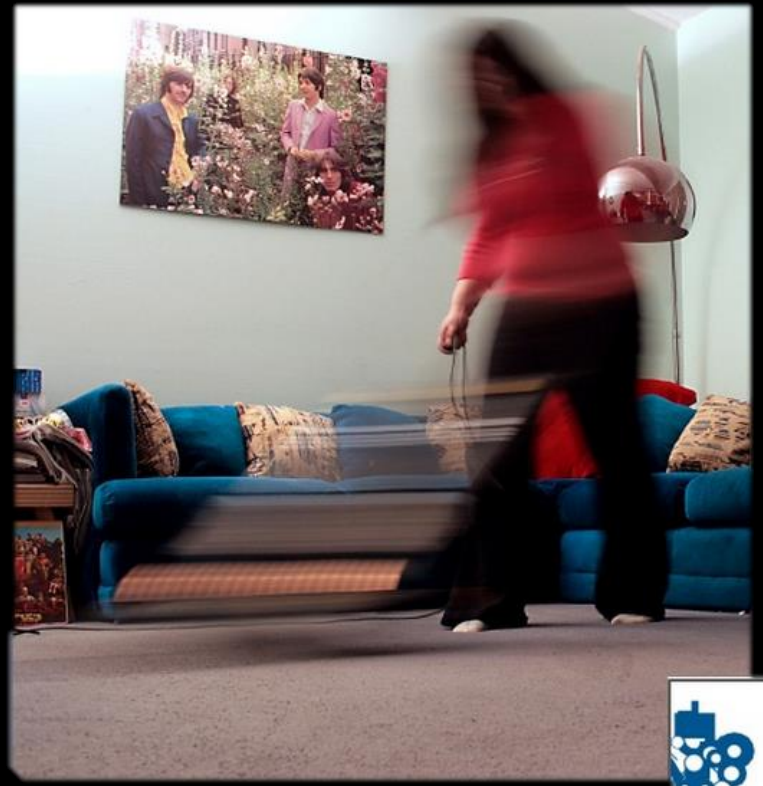


Architecture

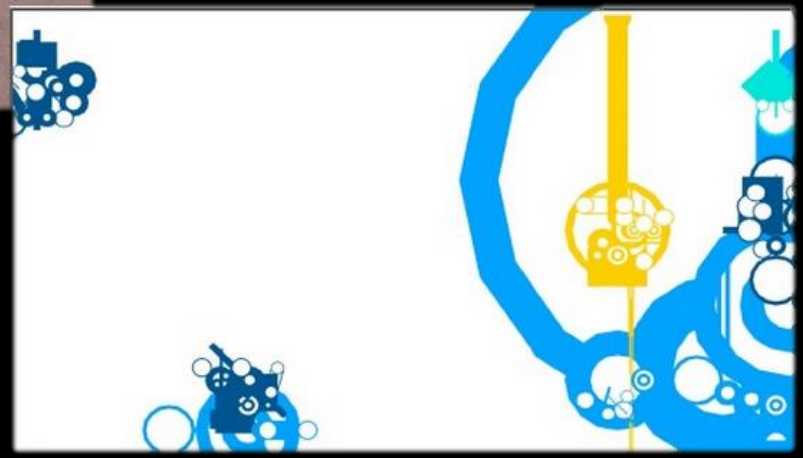



video
motion
semantics
interpretation
overhead cameras
E
D
F
E
D
F
E
interesting places
display and printout
generative architecture

Qiuet Times
(Watching TV, Reading, ...)



Active Times Alone
(cleaning, exercising, ...)

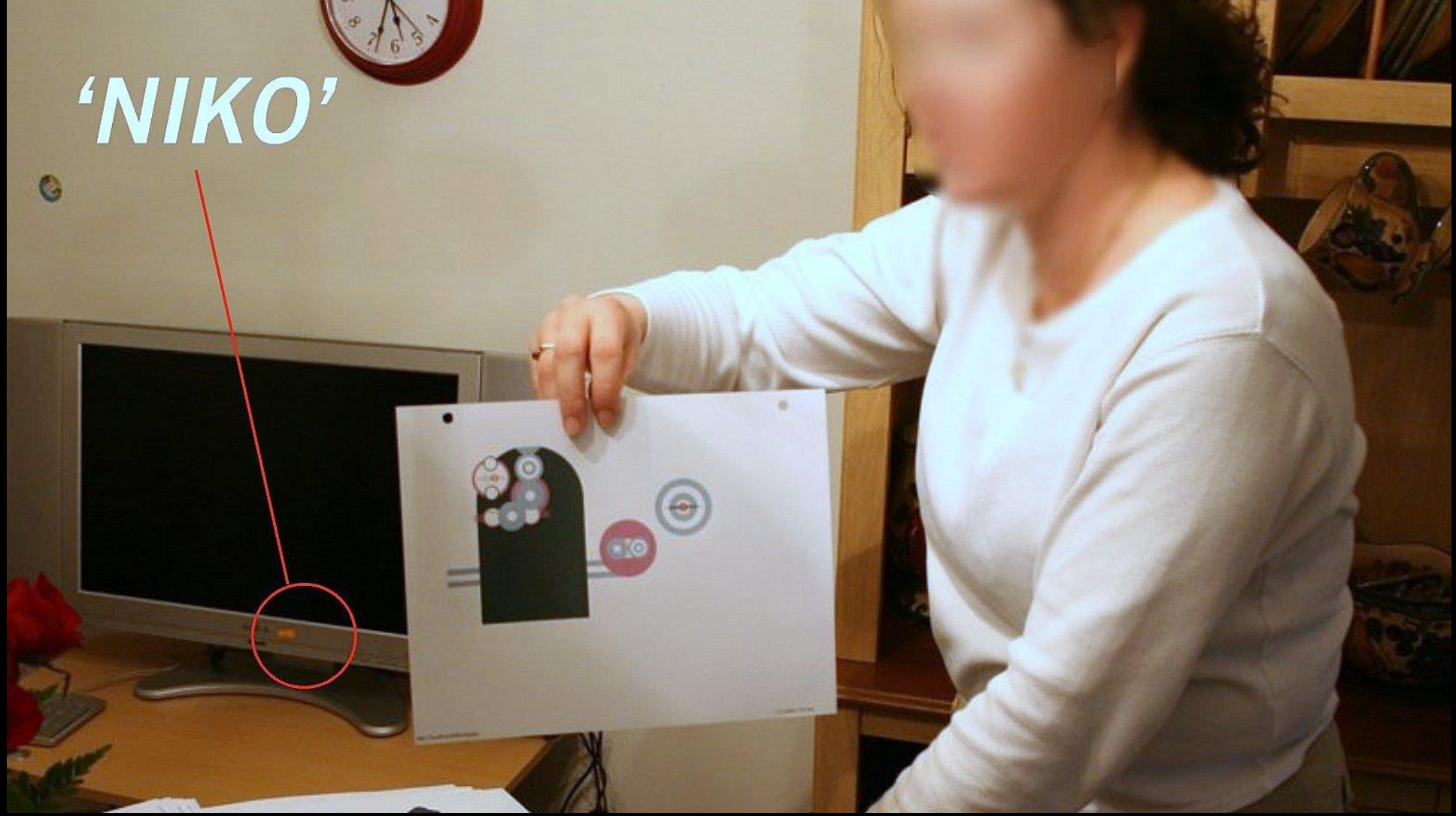


Social Times
(meals, board games, ...)

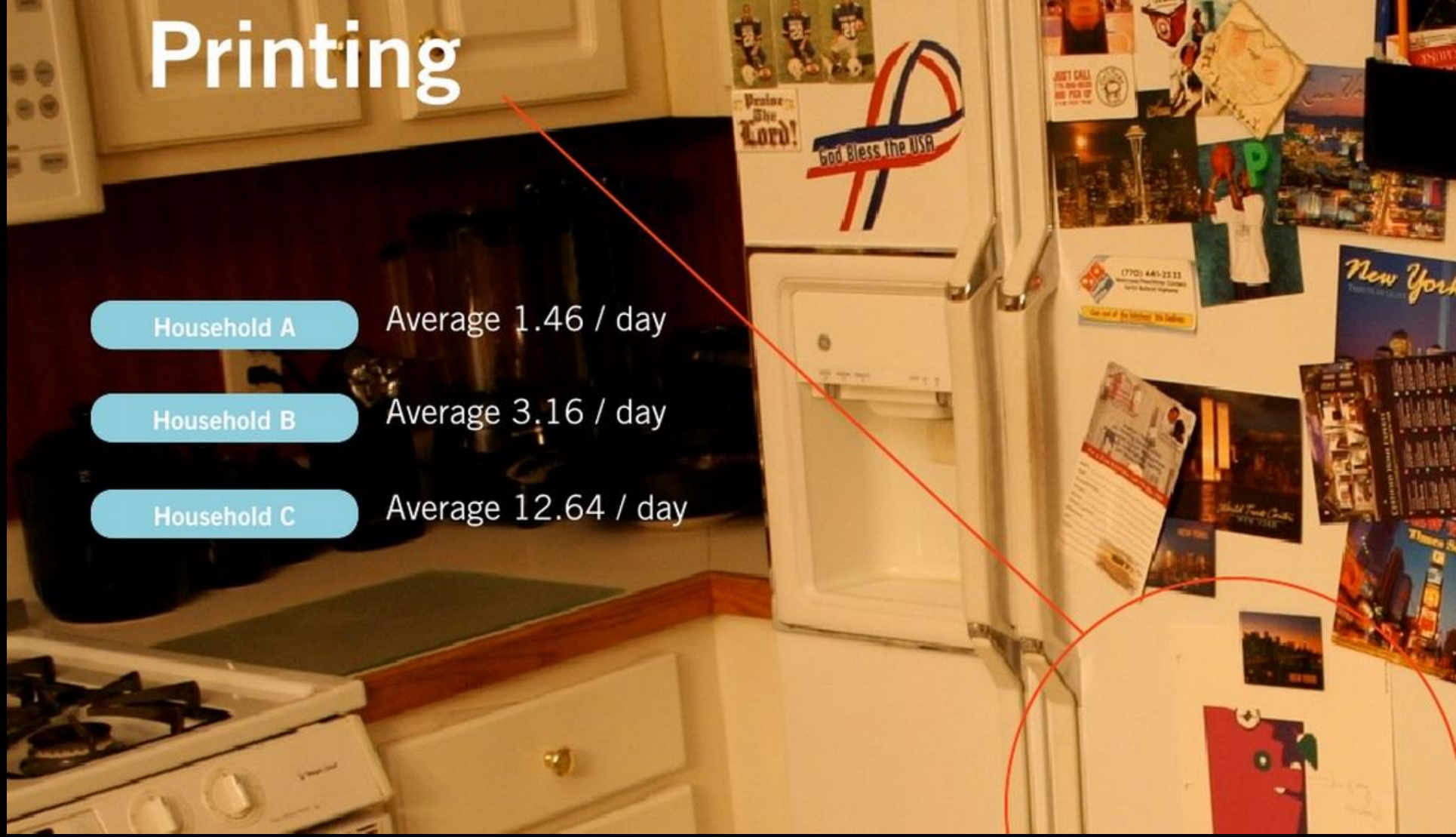
Wild Times
(playing, fighting, ...)

Longitudinal Study

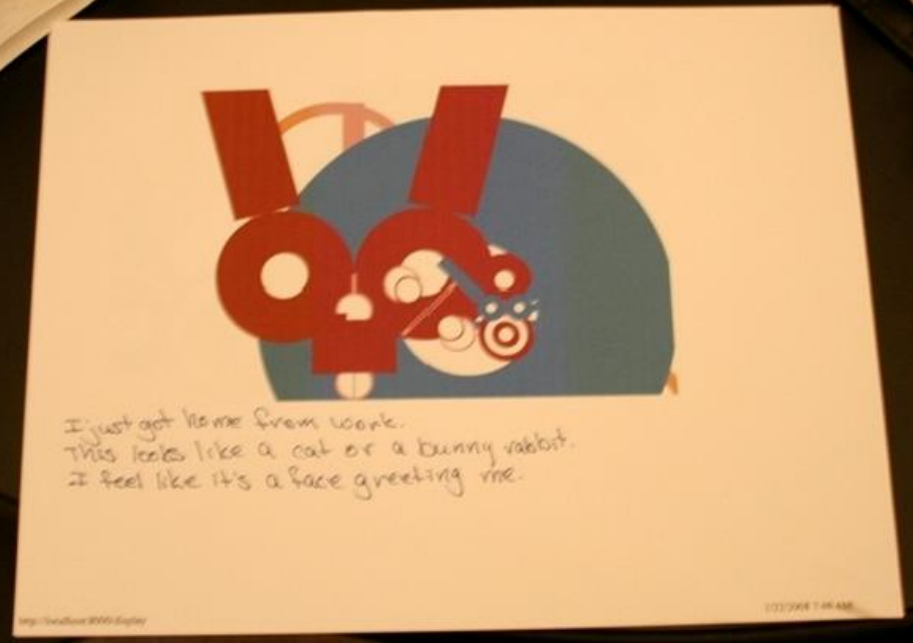
- 3 Homes
- 8 weeks
- Instruments:
  - Interviews
  - Prints
  - Word games
  - Felt boards


‘NIKO’

Printing


Household A
Average 1.46 / day
Household B
Average 3.16 / day
Household C
Average 12.64 / day
Praise the Lord!
God Bless the USA

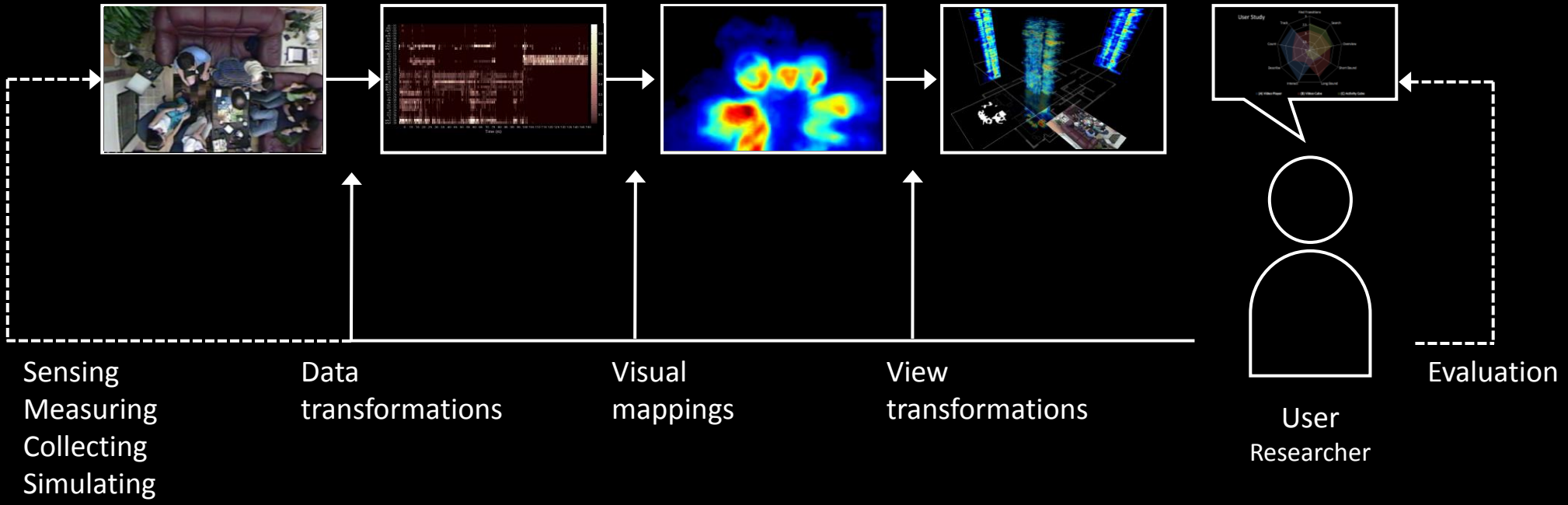
(Hints of) Personality



Household C

Information Visualization Pipeline

expanded from **Readings in Information Visualization: Using Vision to Think**
By Stuart K. Card, Jock D. Mackinlay, Ben Shneiderman, 1999



Reading Assignments

- Introduction to Information Visualization
  - Mazza
  - pdf
- MULTIDIMENSIONAL DETECTIVE
  - Inselberg
  - pdf
- Watch this video http://datajournalism.stanford.edu/

Project 2

1. You have been hired by the World Values Survey
2. Read about it. What is it? What type of data do they show?
3. Learn to use their visualization tools
   - Create and document an analytical trail
4. Use http://d3js.org/ to significantly improve the visualization tools of WVS
   - Create and document a new analytical trail that goes deeper into your research question
5. Create a URL with the code, text, and images describing your results

Thank you!

marior@kth.se