Aalto University

Behavioral modeling
+ optimization
+ machine learning

A New Type of Intelligence for Intelligent User Interfaces

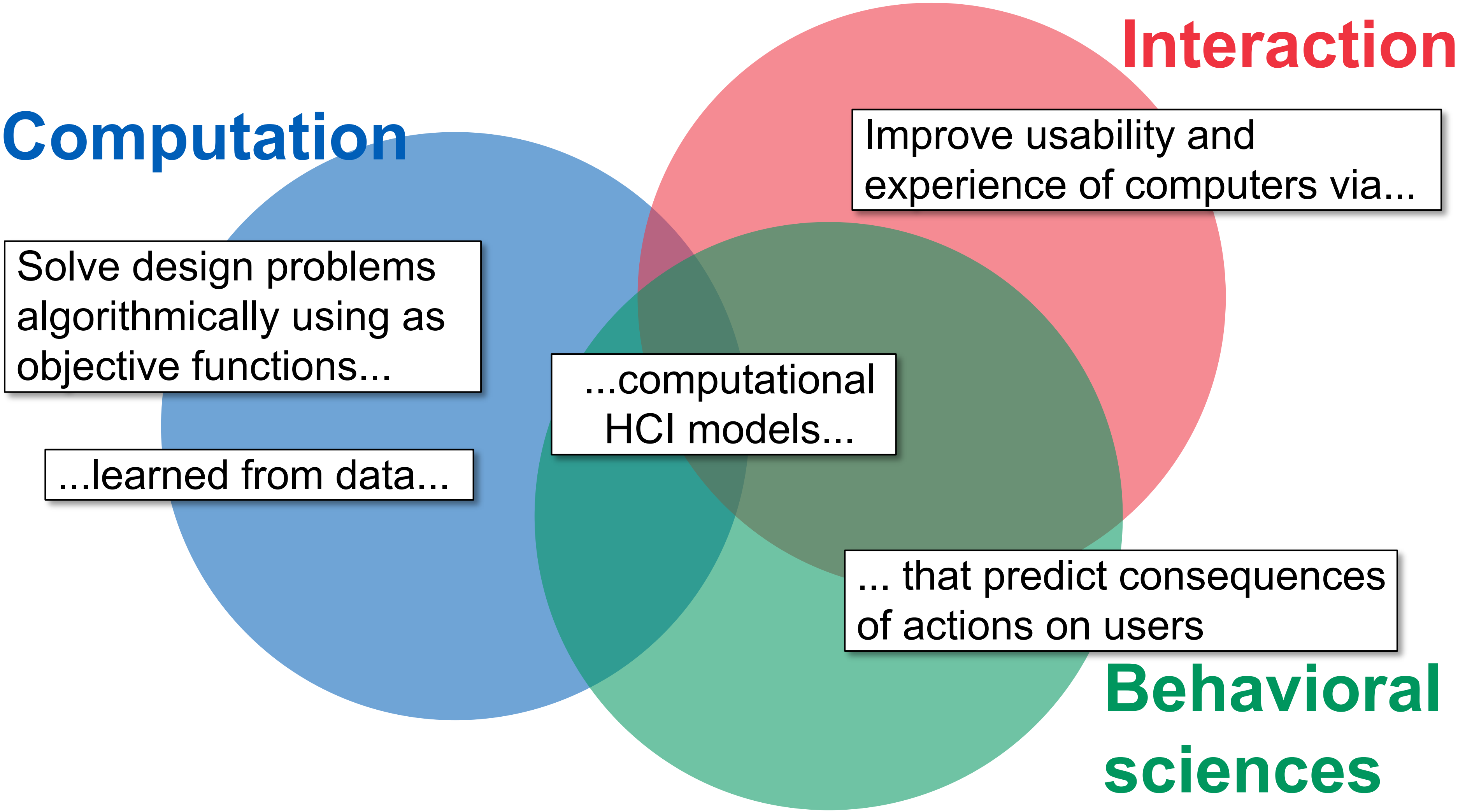
An alternative to the
"deep learning temptation"

Ant
Fin
Aal

userinterfaces.aalto.fi

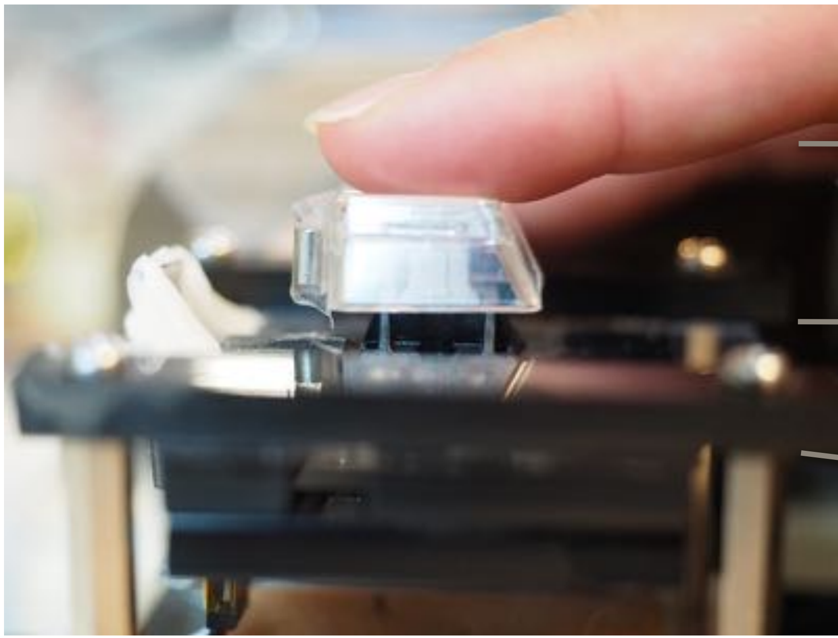
19th
ICWE 2019
19TH INTERNATIONAL CONFERENCE ON WEB ENGINEERING

A new type of intelligence for IUI




CS and EE at
Aalto University

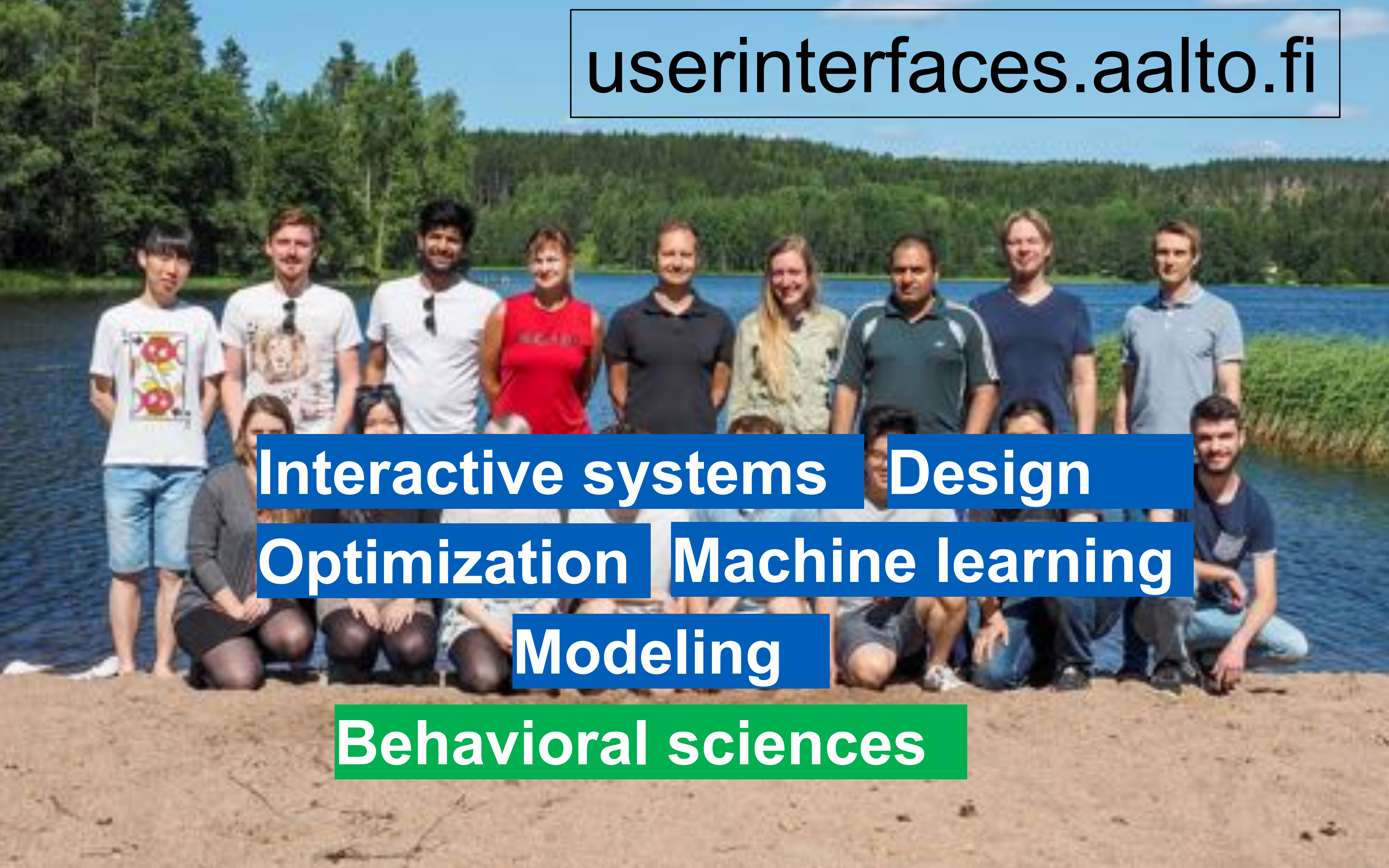
I'm a computational cognitive scientist working on HCI



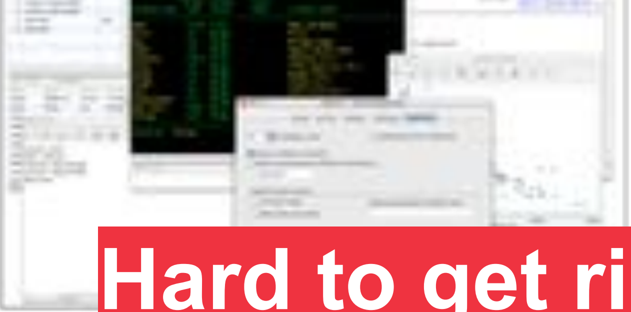
I model human-technology interaction

... and develop **new computational principles** of design and adaptation

...in order to **improve** computing for humans


userinterfaces.aalto.fi
Interactive systems
Design
Optimization
Machine learning
Modeling
Behavioral sciences

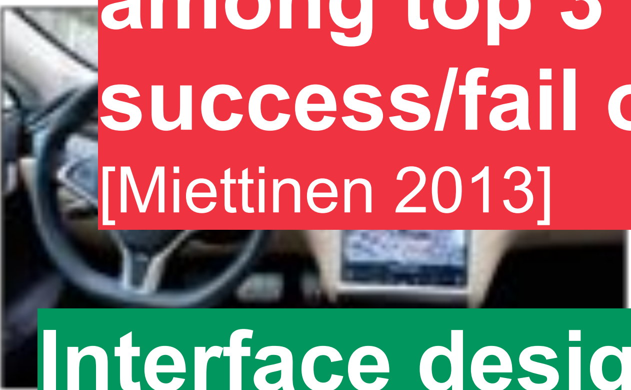
Opportunities for intelligent UIs



G

Touch interfaces

Hard to get right: UID among top 3 reasons for success/fail of ICT projects
[Miettinen 2013]

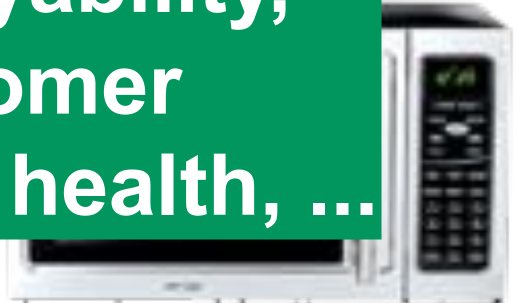


ces

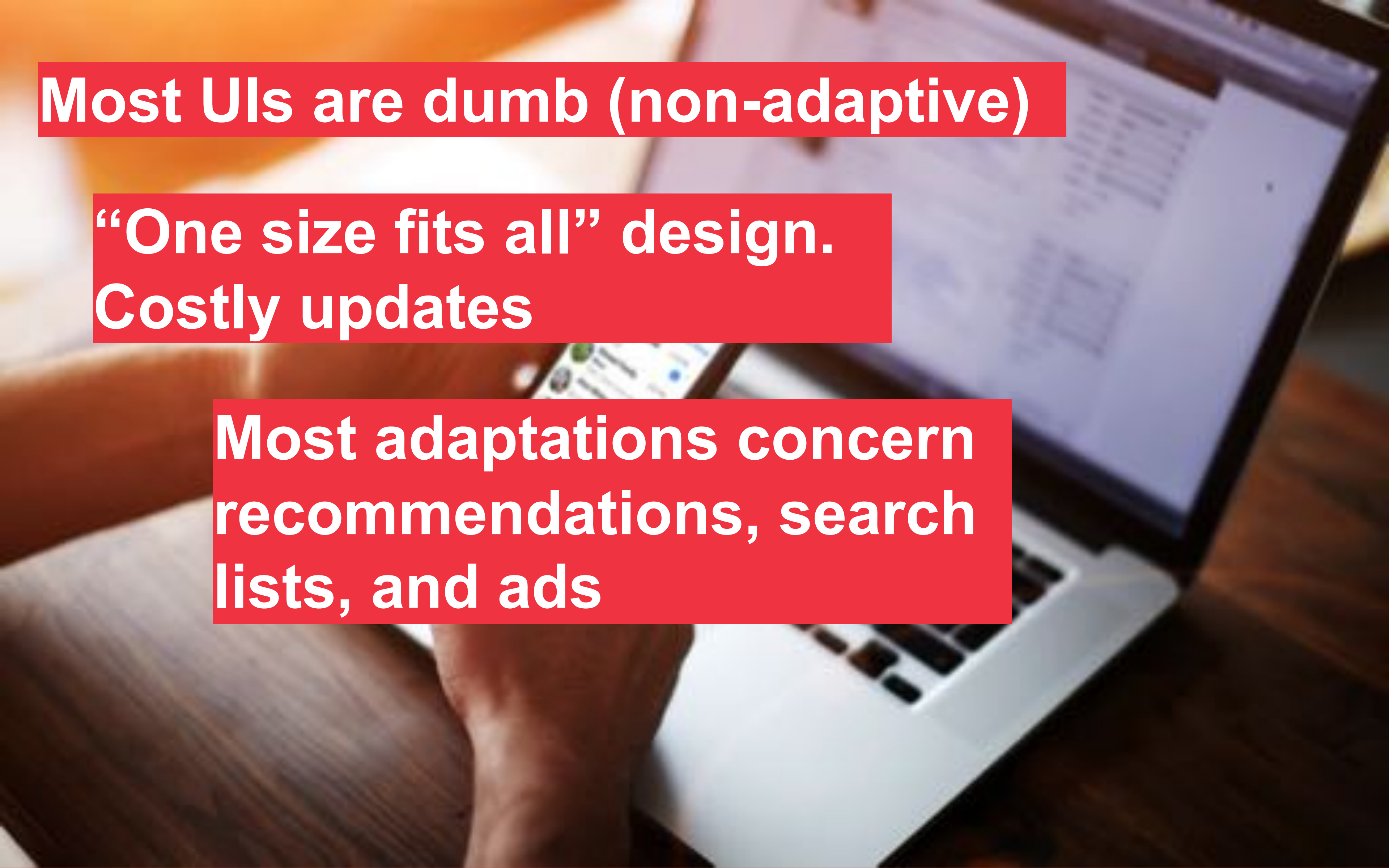
Dialogue interfaces

Interface design affects productivity, enjoyability, satisfaction, customer loyalty, inclusion, health, ...

Discrete input devices



Panel-and-pushbuttons interfaces



Most UIs are dumb (non-adaptive)

“One size fits all” design.
Costly updates

Most adaptations concern recommendations, search lists, and ads

Design creates barriers

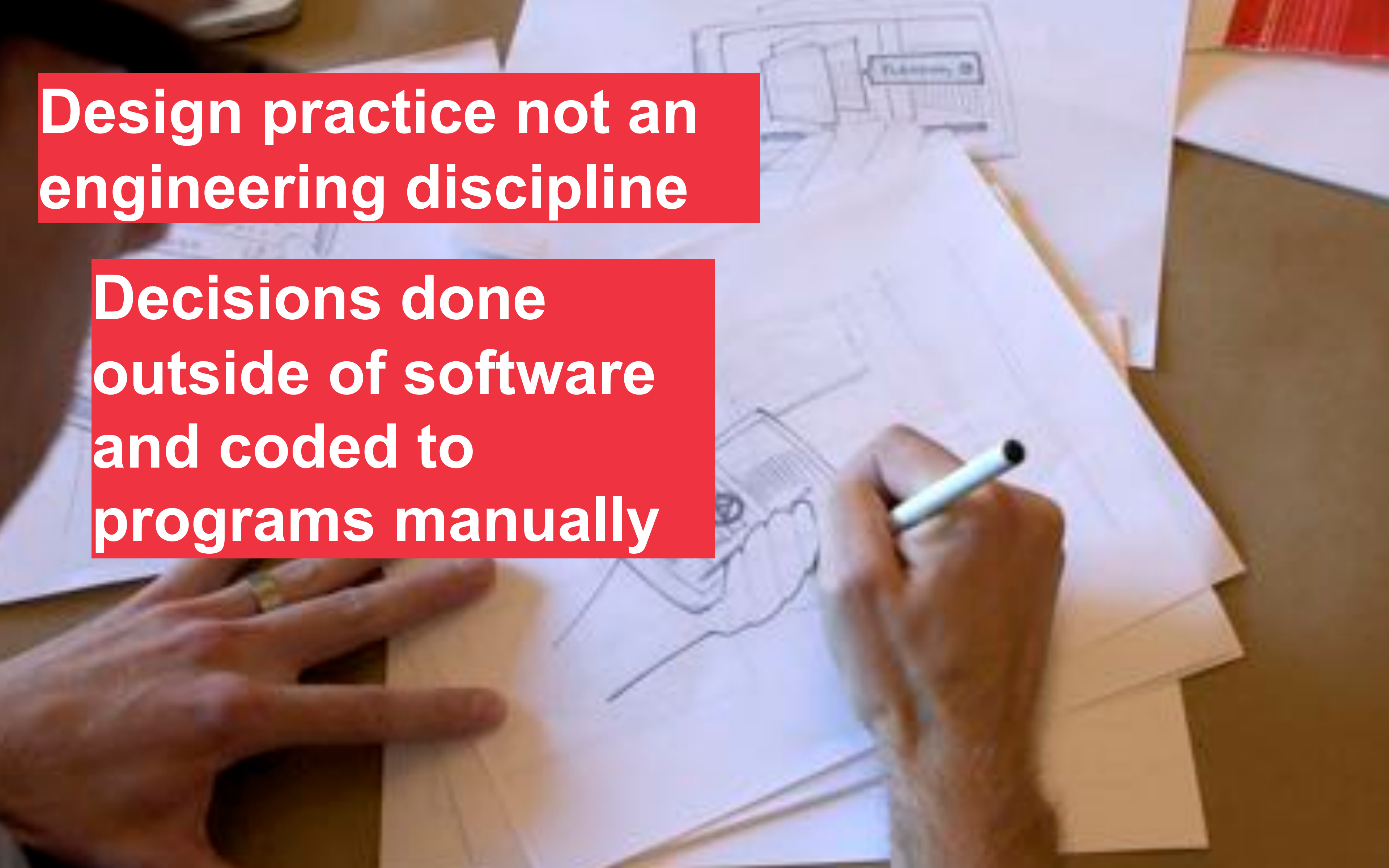
Increasing reliance on e-services widens the digital gap

[OECD 2018]

Novel UIs is a major area of tech investment

Abysmal success rate

We’re bad at transferring knowledge



Design practice not an engineering discipline

Decisions done outside of software and coded to programs manually

Historical backdrop

Methods for intelligent UIs studied since 1970s

Logic

Information

Control

Optimization

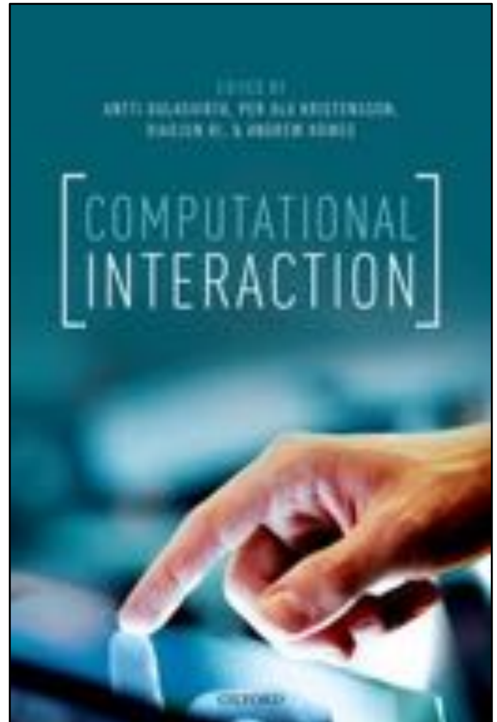
Agents

Cognitive models

Learning

Probability

Neural networks



Computational Interaction (edited)

Oxford University Press
2018

Aalto University

A paradox!

If algorithmic methods are superior, where are they?

Four waves of AI that have hit HCI

Wave 0: Cybernetics

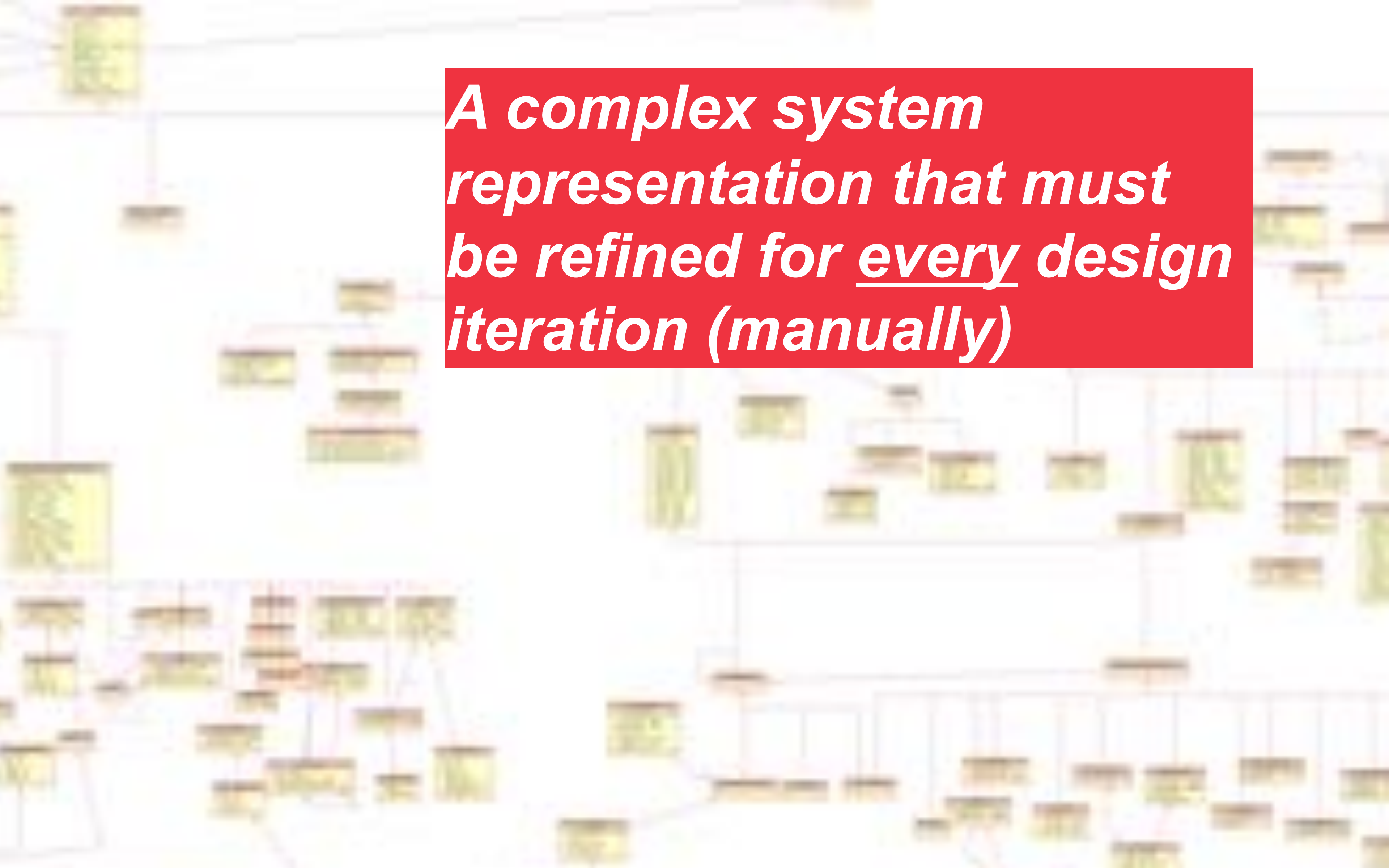
Wave 1: Rules and logic

Wave 2: Cognitive modeling

Wave 3: Pre-DL supervised learning

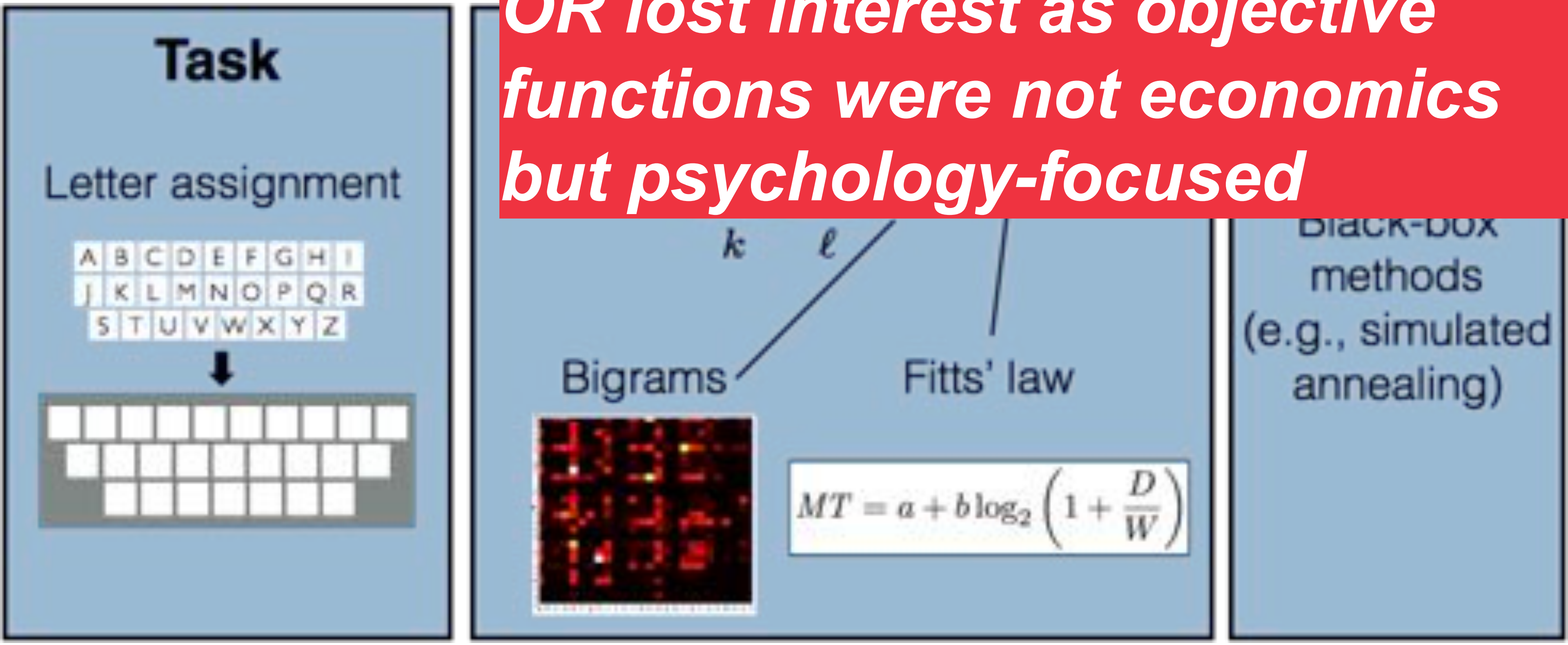
Wave 4: Deep learning

Let’s go through them to learn why they have <u>not</u> revolutionarized the field



A complex system representation that must be refined for every design iteration (manually)

Operations research started and stopped with keyboards



OR lost interest as objective functions were not economics but psychology-focused

Burkard et al. 1977; Light & Anderson 1993; Zhai et al. 2000

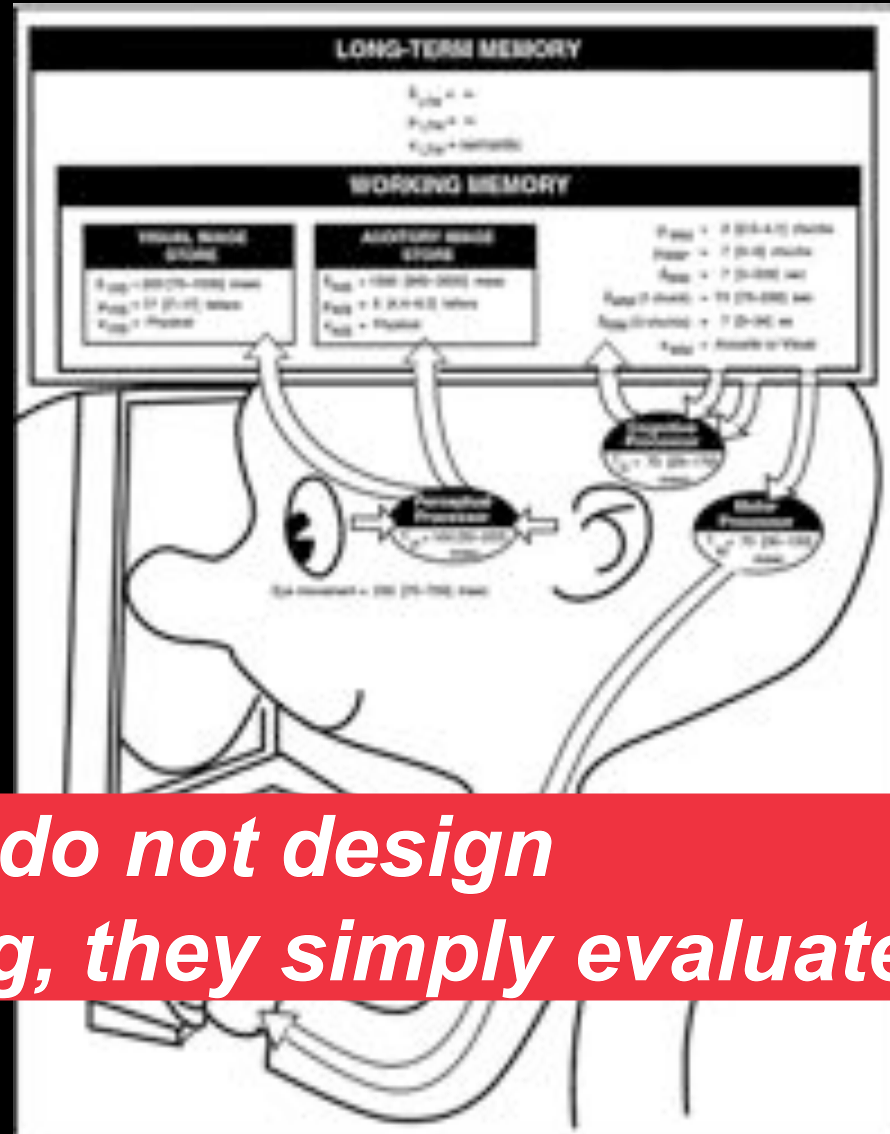
Psychology as the science of design

August Dvorak

Herbert Simon

Stuart Card

Predictive cognitive models (1980s)



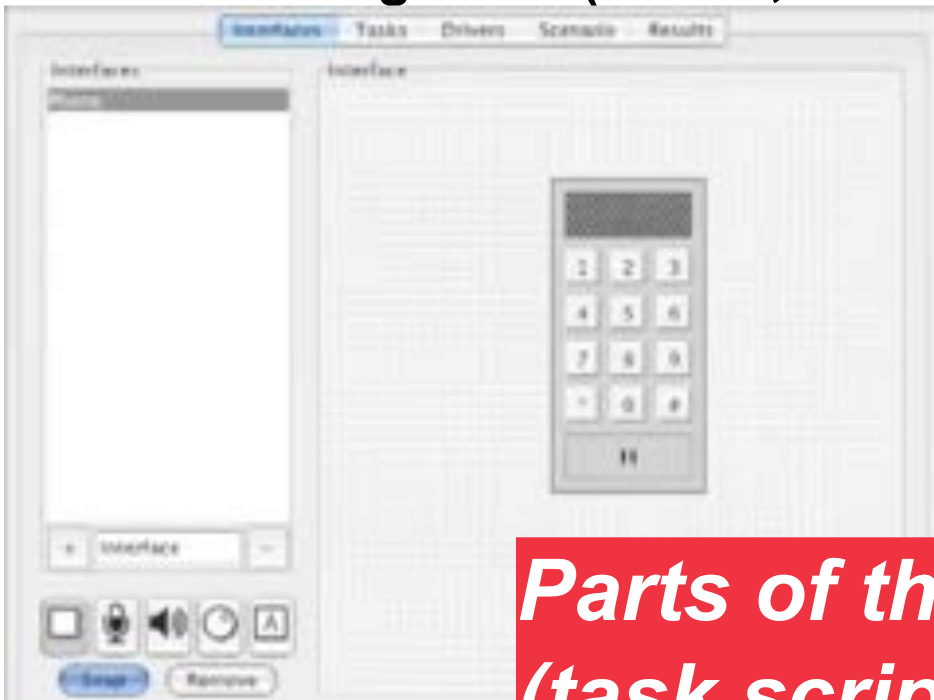
Card et

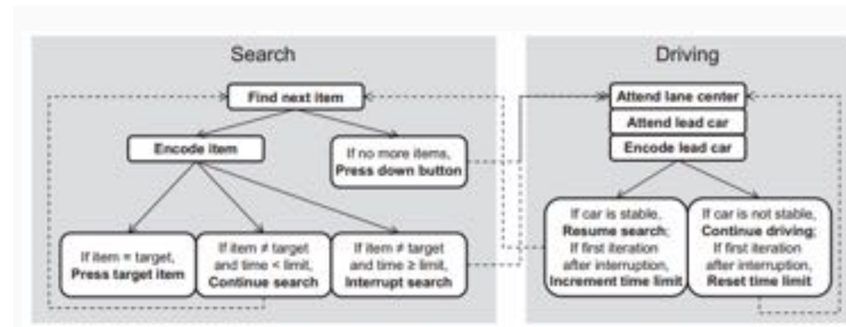
Models do not design anything, they simply evaluate

Oulasvirta, Antti
ICWE 2019

Generative cognitive models

Threaded Cognition (ACT-R, a driving simulator)

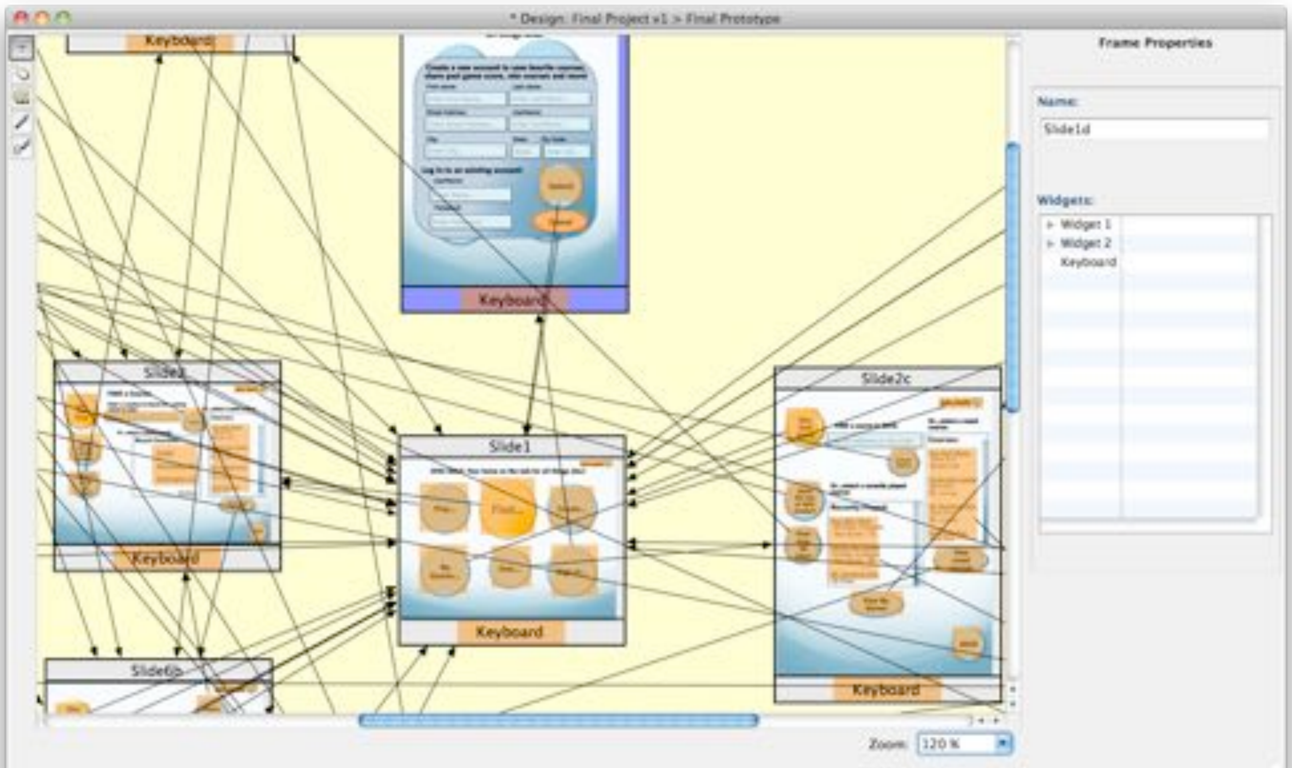




Parts of the cognitive model (task script) must be redefined every time design changes

Tools for using cognitive models in evaluation

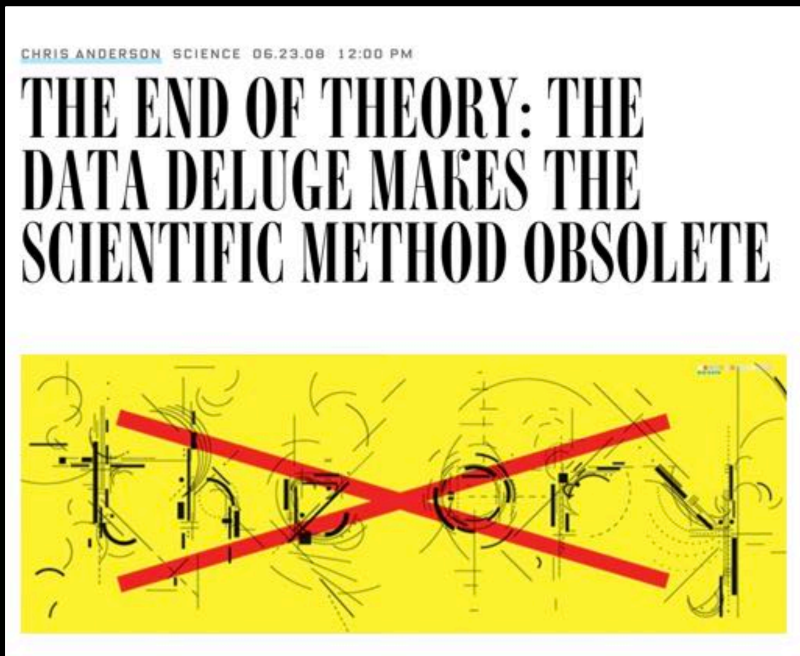
CogTool



Work of Bonnie John / IBM and CMU

Big data and machine learning

"Correlationalism and the data deluge"


CHRIS ANDERSON SCIENCE 06.23.08 12:00 PM

THE END OF THEORY: THE DATA DELUGE MAKES THE SCIENTIFIC METHOD OBSOLETE

"Theory is dead. Data is the new king!"

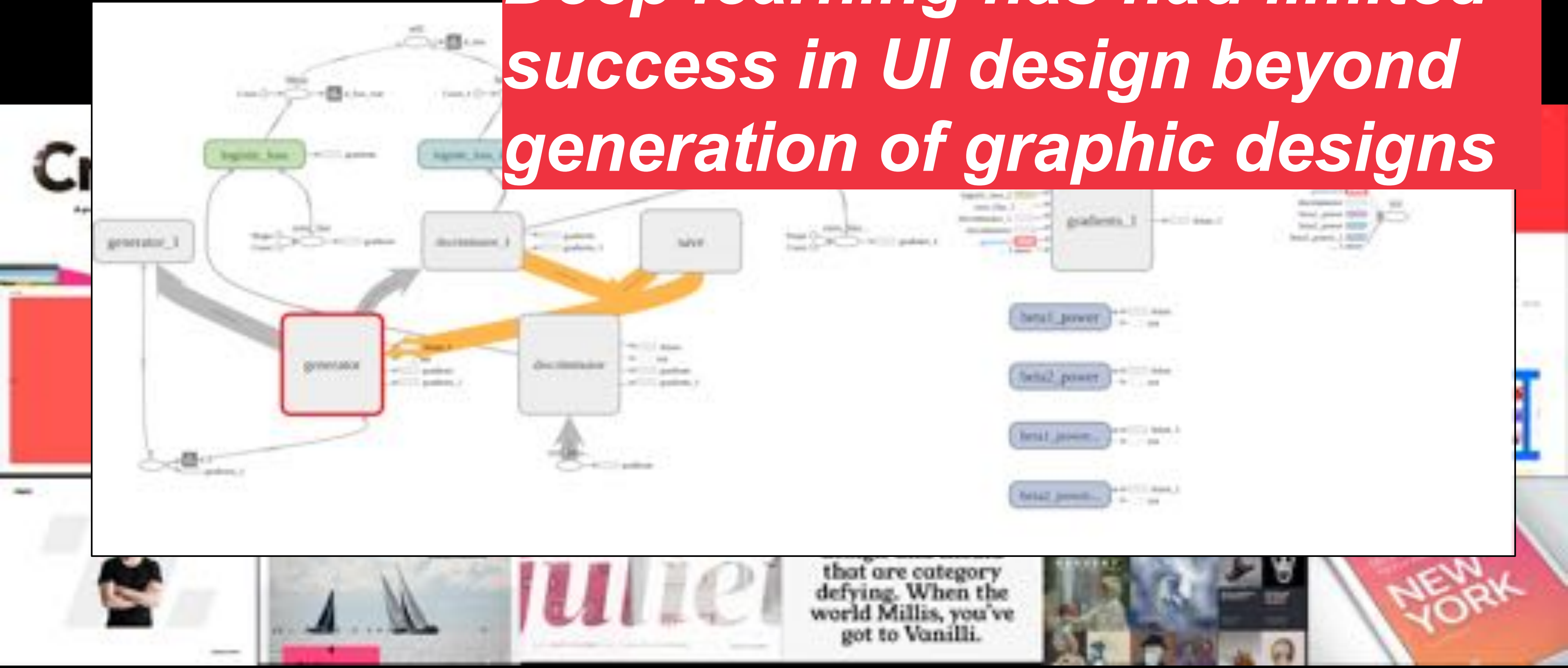
Oulasvirta, Antti
ICWE 2019

Generative art by deep learning

Combines known patterns. Not easily controlled toward desirable properties

Expero

Deep learning has had limited success in UI design beyond generation of graphic designs



Oulasvirta, Antti
ICWE 2019

Autodesk

Combination of DL and <u>physics models</u> has been a success in computational design. But is insufficient for HCI

Wix promises AI website design, but fails to deliver true artificial intelligence

Wix's new automated design functionality is helpful, but it doesn't match the definition of artificial intelligence

Lots of hot air and broken promises...

4 Hard to change designs

3 Design is ill-specified and combines creative and problem-solving activities

User interface design is hard

2 Humans are hard to predict

1 Very large search spaces

Design is "speculation"

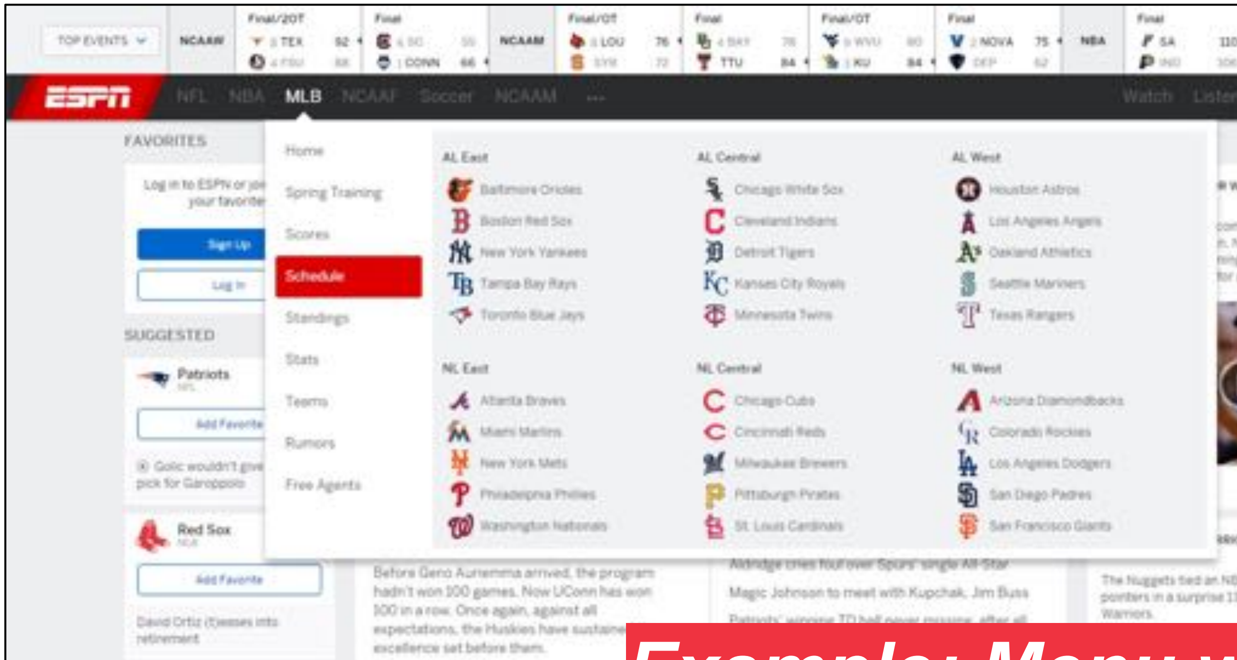
A useful theory contains <u>counterfactual</u> information about what will happen <u>if</u> a design decision is taken:

If design was <this>

then

interaction *would* be <this>

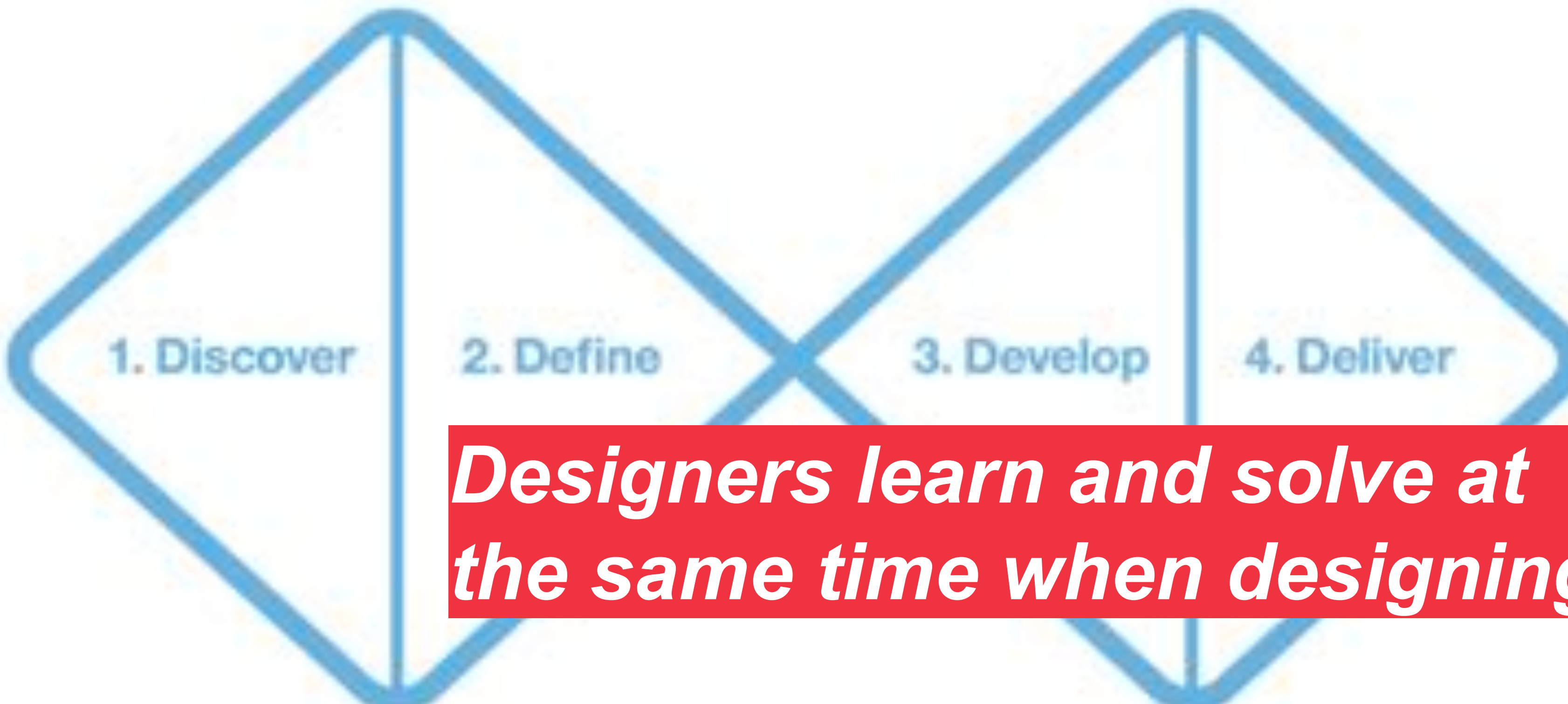
Design involves very large search spaces



Example: Menu with 50 items can be organized in 10^{80} ways


~~Physics~~
Behavioral and cognitive sciences

Design is ill-specified



Designers learn and solve at the same time when designing

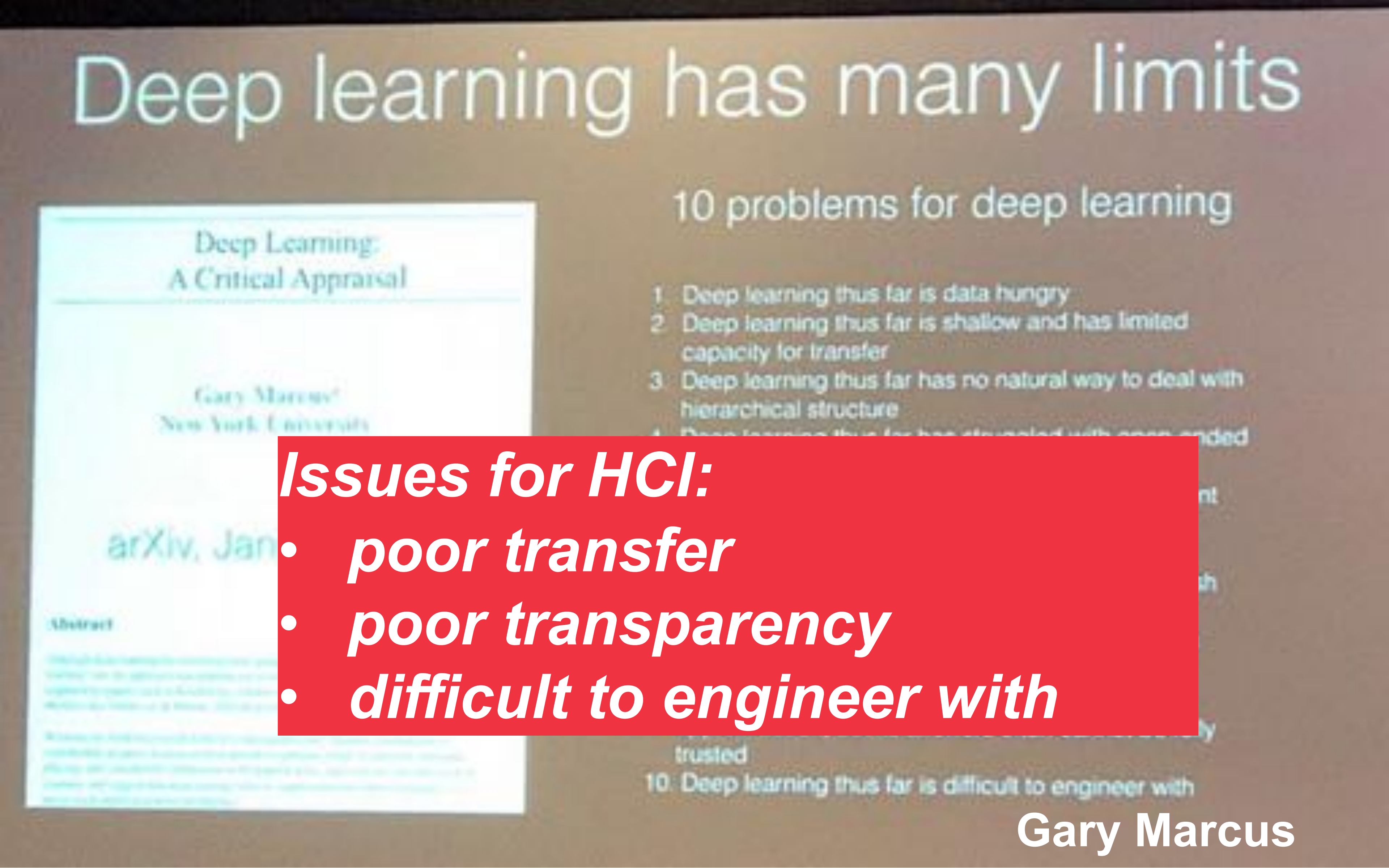
"Concept" **"Artefact"**

UK Design Council

Designs are "best achievable compromises"

Hard to know what the best compromise is without a human in the loop

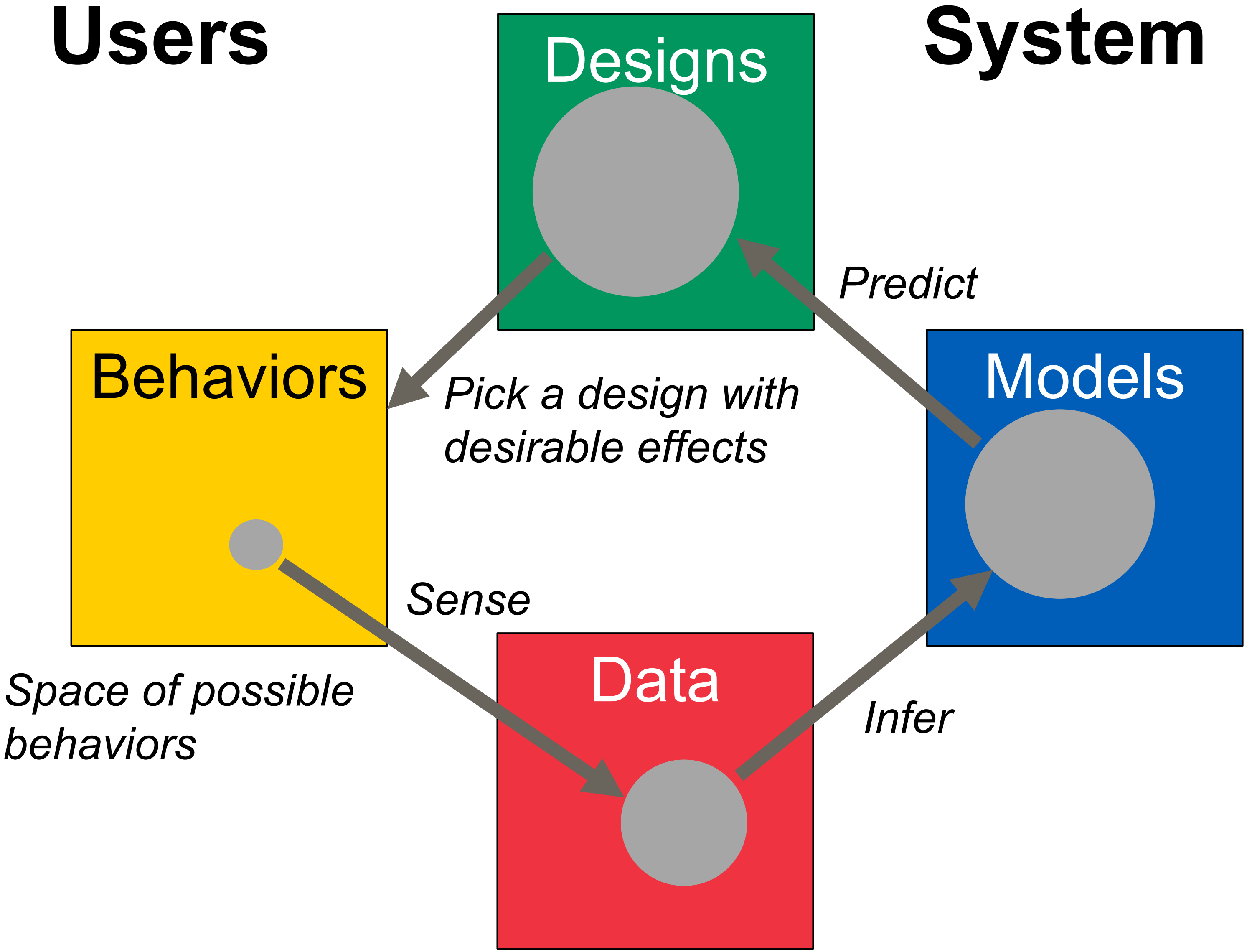
Deep learning has had limited success in UI generation & adaptation



Issues for HCI:

- ***poor transfer***
- ***poor transparency***
- ***difficult to engineer with***

Gary Marcus


Users
System
Designs
Predict
Behaviors
Pick a design with
desirable effects
Models
Sense
Data
Space of possible
behaviors
Infer

This talk

1 Approach

2 Applications

3 Summary

Lots of examples coming up

Approach

Vision: A new type of MI for IUI

(1) Able to anticipate the consequences of its actions

(2) Represents its rationale in human-relatable terms

(3) Chooses designs counterfactually

(4) Learns and updates itself in the light of user data

(5) Operates conservatively, admitting the inherent fallibility of its input data and model

The design problem

Find the design (*d*) out of candidate set (*D*) that maximizes <u>goodness</u> (*g*) in given conditions (*θ*):

$$\max_{d \in D} g(d, \theta)$$

Expanded...

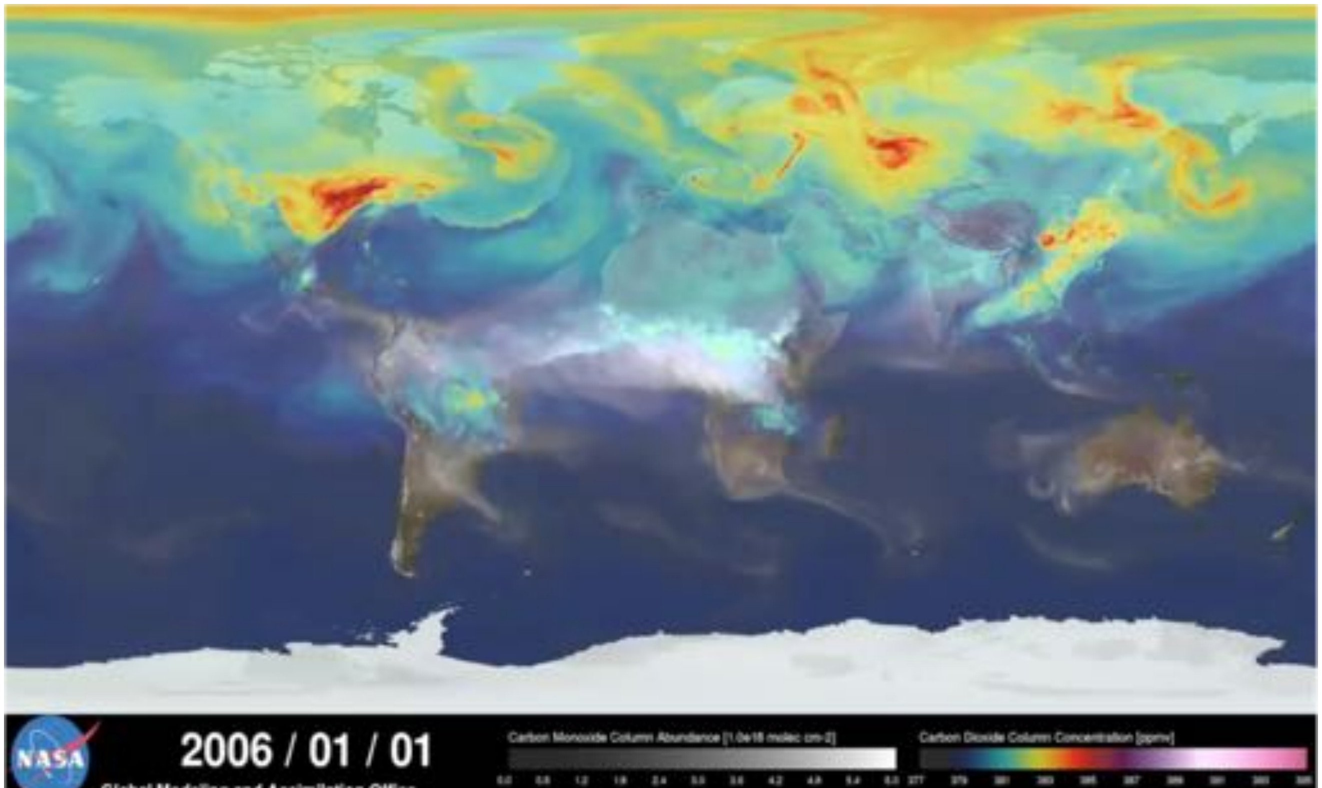
"Choose the design that maximizes expected goodness to users, given observations"

$$\max_{d \in D} E(g(d)|o)$$

"Choose the design that maximizes expected goodness predicted by a model inferred from observations"

$$= \max_{d \in D} \int g\big(M(\theta, d)\big) p(\theta|o) d\theta$$

Inspired by simulation models in science and engineerings



Why we need causal models

1. Very hard inference problems

- Many possible explanations to human behavior

2. The is-ought problem

- Data alone does not prescribe a desirable future

3. High costs of design error

- Very few chances to “try one’s luck” in interaction

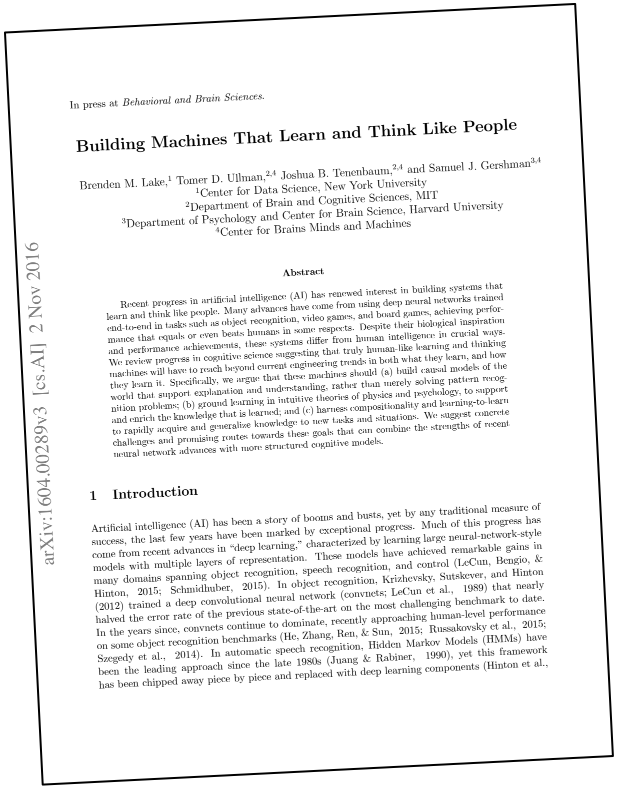
4. Deferral of reward

- The value of an action to humans is rarely directly observable

We cannot solve these with a correlationalist approach

FCAI

“Machines that think like humans”


In press at *Behavioral and Brain Sciences.*

Building Machines That Learn and Think Like People

Brenden M. Lake,[1] Tomer D. Ullman,[2,4] Joshua B. Tenenbaum,[2,4] and Samuel J. Gershman[3,4]
[1]Center for Data Science, New York University
[2]Department of Brain and Cognitive Sciences, MIT
[3]Department of Psychology and Center for Brain Science, Harvard University
[4]Center for Brains Minds and Machines

arXiv:1604.00289v3 [cs.AI] 2 Nov 2016

Abstract

Recent progress in artificial intelligence (AI) has renewed interest in building systems that learn and think like people. Many advances have come from using deep neural networks trained end-to-end in tasks such as object recognition, video games, and board games, achieving performance that equals or even beats humans in some respects. Despite their biological inspiration and performance achievements, these systems differ from human intelligence in crucial ways. We review progress in cognitive science suggesting that truly human-like learning and thinking machines will have to reach beyond current engineering trends in both what they learn, and how they learn it. Specifically, we argue that these machines should (a) build causal models of the world that support explanation and understanding, rather than merely solving pattern recognition problems; (b) ground learning in intuitive theories of physics and psychology, to support and enrich the knowledge that is learned; and (c) harness compositionality and learning-to-learn to rapidly acquire and generalize knowledge to new tasks and situations. We suggest concrete challenges and promising routes towards these goals that can combine the strengths of recent neural network advances with more structured cognitive models.

1 Introduction

Artificial intelligence (AI) has been a story of booms and busts, yet by any traditional measure of success, the last few years have been marked by exceptional progress. Much of this progress has come from recent advances in “deep learning,” characterized by learning large neural-network-style models with multiple layers of representation. These models have achieved remarkable gains in many domains spanning object recognition, speech recognition, and control (LeCun, Bengio, & Hinton, 2015; Schmidhuber, 2015). In object recognition, Krizhevsky, Sutskever, and Hinton (2012) trained a deep convolutional neural network (convnets; LeCun et al., 1989) that nearly halved the error rate of the previous state-of-the-art on the most challenging benchmark to date. In the years since, convnets continue to dominate, recently approaching human-level performance on some object recognition benchmarks (He, Zhang, Ren, & Sun, 2015; Russakovsky et al., 2015; Szegedy et al., 2014). In automatic speech recognition, Hidden Markov Models (HMMs) have been the leading approach since the late 1980s (Juang & Rabiner, 1990), yet this framework has been chipped away piece by piece and replaced with deep learning components (Hinton et al.,

Lake et al. 2016

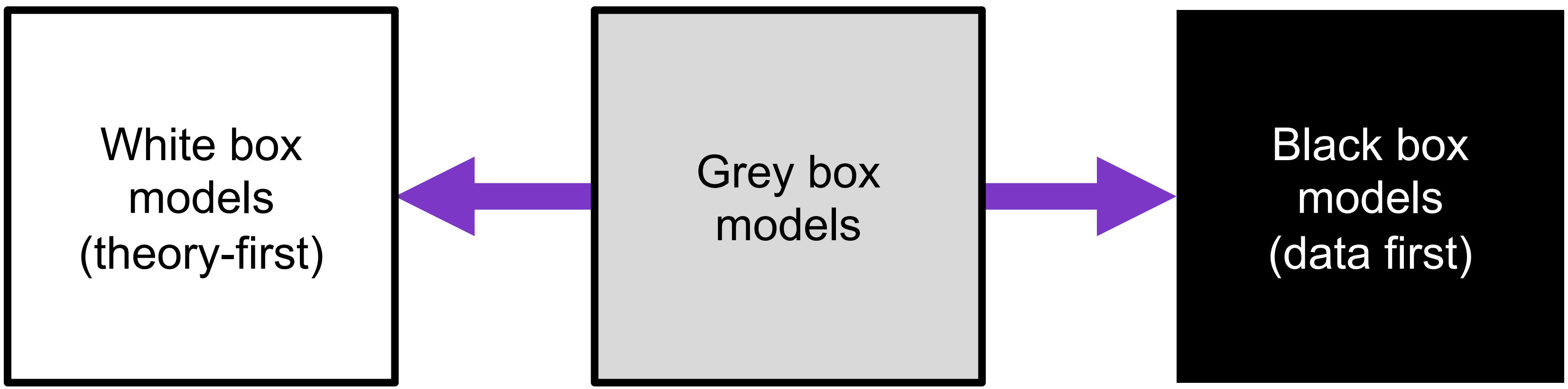
(a) build causal models of the world that support explanation and understanding, rather than merely solving pattern recognition problems;

(b) ground learning in intuitive theories of physics and psychology, to support and enrich the knowledge that is learned;

(c) harness compositionality and learning-to-learn to rapidly acquire and generalize knowledge to new tasks and situations.

"Grey boxing"

We combine models of human behavior with optimization and probabilistic inference

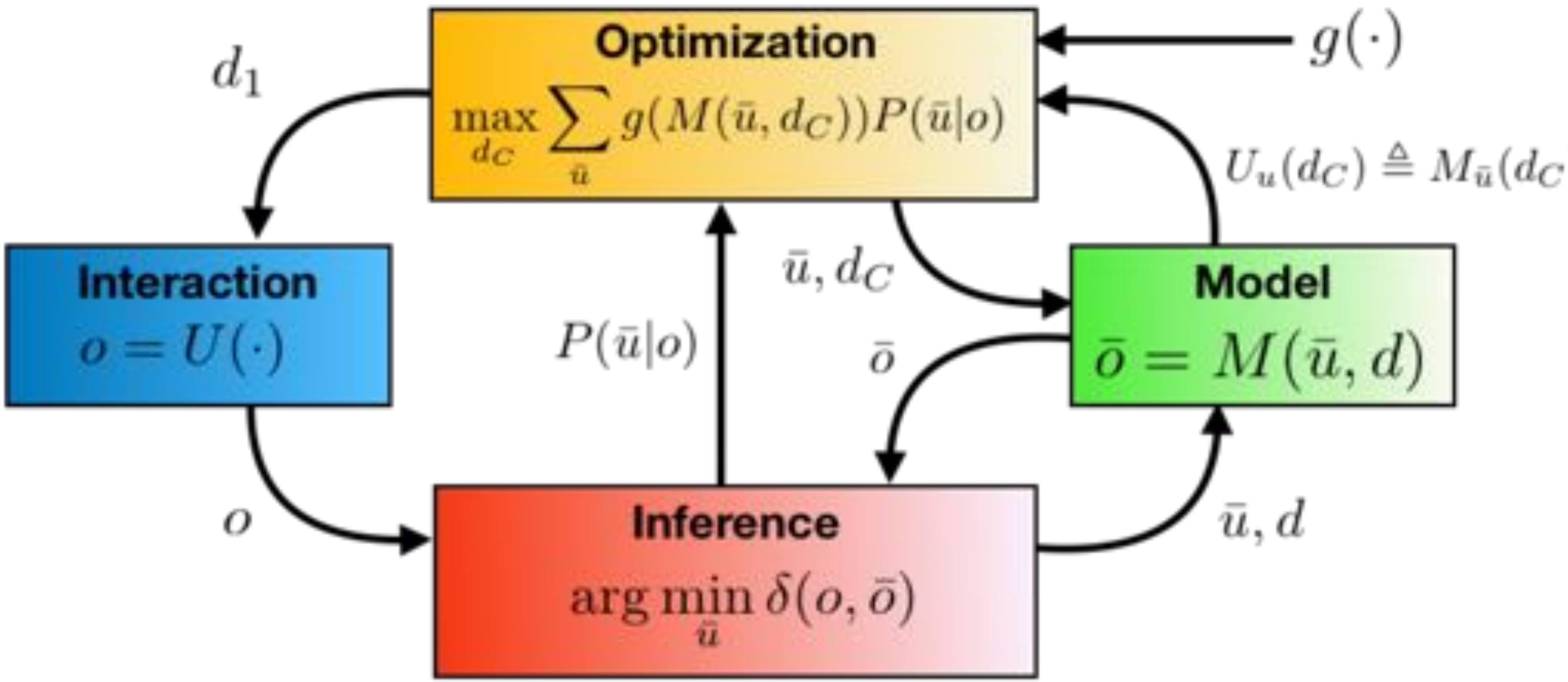


Fully controllable
High generality
Less data hungry
Hard to construct

High rep. power
Data hungry
Poor controllability
Low generality

“The stool”

Model-driven human-computer interaction



Achievements 2010-2019: Overview

Solve design problems using optimization

- Many HCI problems defined by now
- Significant improvements in usability in hard problems

Support designers' creativity and problem-solving

- Provide guarantees for result quality

Assist in the interpretation of complex behavioral data

Drive adaptive interfaces

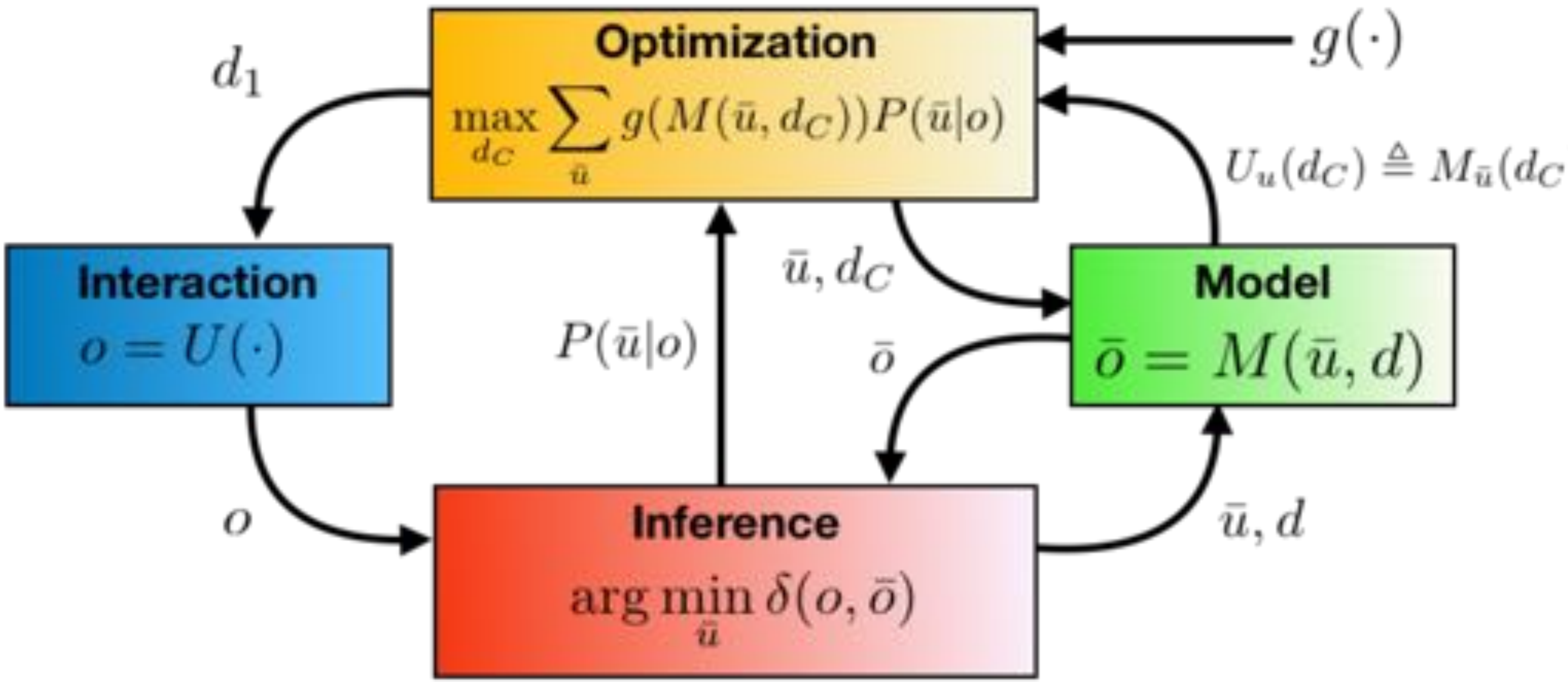
- Improvements to an individual 5-25% in usability

Lots of advances last 10 years

A bit more detail…

Mathematical definition of design tasks

Model-driven human-computer interaction

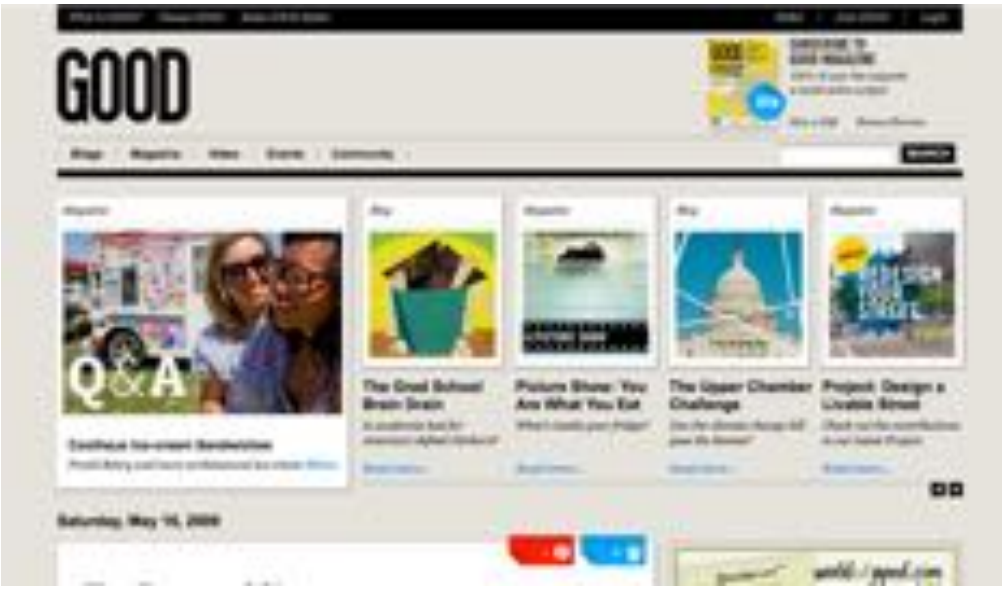


Defining design problems
Example: IP definition of label selection

$$
\begin{aligned}
\min \quad & \left(\sum_{\ell \in L} t_\ell y_\ell, \sum_{u \in U} \sum_{\ell \in L_u} m_{u\ell} x_{u\ell}, \sum_{k,\ell \in L} d_{k\ell} y_k y_\ell \right) & & \\
\text{subject to} \quad & \sum_{\ell \in L_u} x_{u\ell} = 1 & \forall u \in U & \\
& y_\ell \geq \sum_{u \in U_\ell} x_{u\ell} & \forall \ell \in L & \qquad (20) \\
& y_\ell \leq 1 & & \\
& x_{u\ell} \in \{0, 1\} & \forall u \in U, \ell \in L & \\
& y_\ell \geq 0 & \forall \ell \in L &
\end{aligned}
$$

Definition of a problem allows (1) analysis of problem complexity and (2) choice of best solver

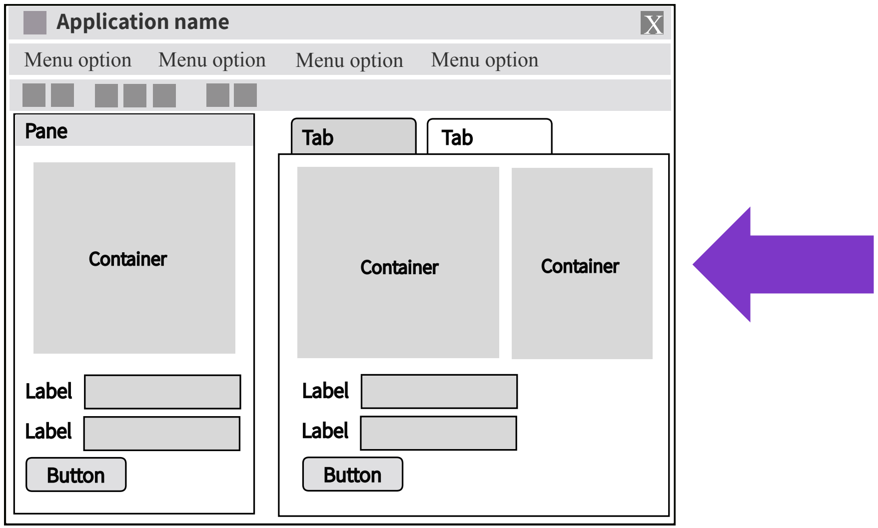
Defining textbook-level GUI design problems as optimization problems



Widget selection
Dialogue and form design
Command sets
Navigation structures
Windowing
Notification scheduling
Task allocation
Assortment design
…

Advances in 2014-2019

Formulation of an existing scope of UI design problems as *known* combinatorial problems



Assignment

Selection

Packing

Ordering

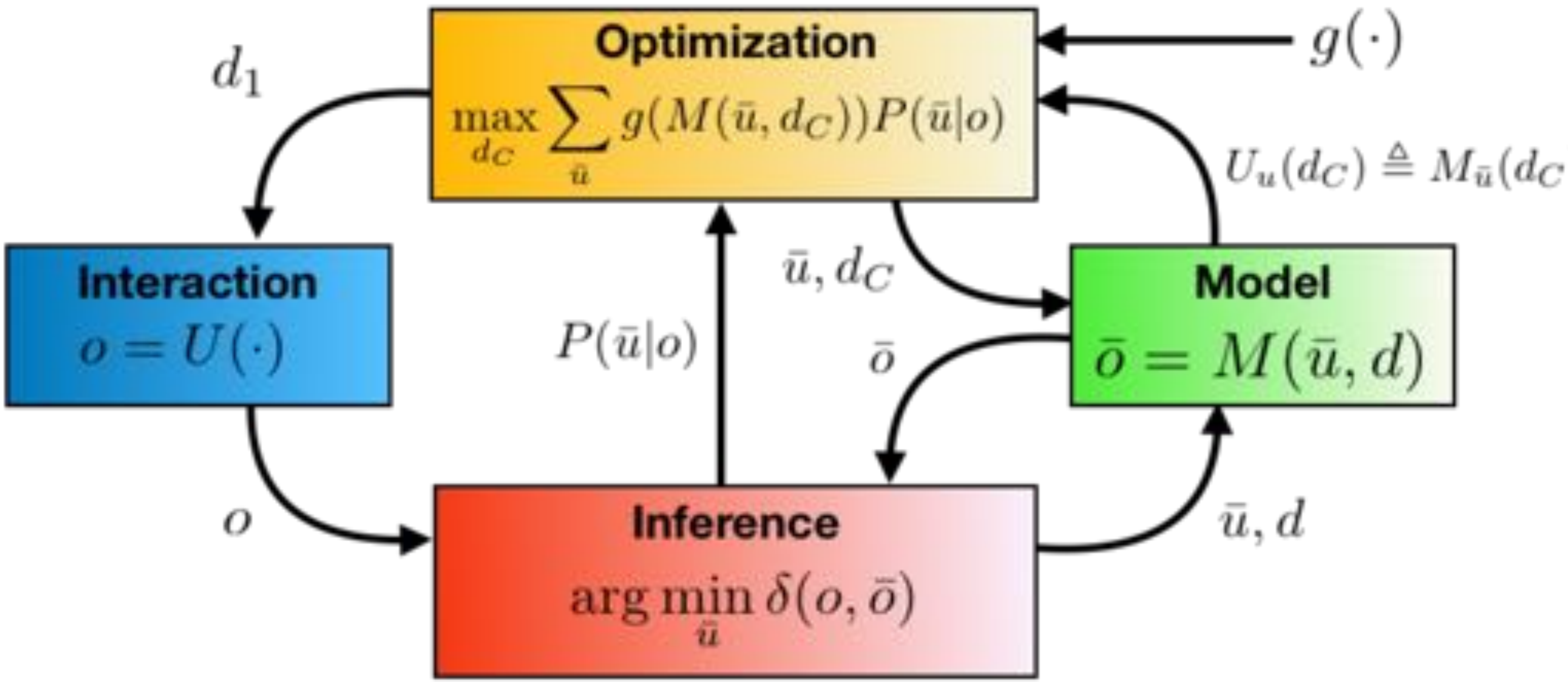
Layouts

[Oulasvirta et al. *submitted*]

A bit more detail…

Predictive models of human behavior

Model-driven human-computer interaction



Many modeling families

Behavioral heuristics

if-then rules

Mathematical models

$$y = f(\mathbf{x}, \boldsymbol{\beta}) + \epsilon$$

Generative models

E.g., symbolic, neural, or biomechanical

→ A multi-objective function for design

$$g(\boldsymbol{x}) = \omega_1 g_1(\boldsymbol{x}) + \cdots + \omega_q g_q(\boldsymbol{x})$$

Example: Model of menu selection

$$T(\theta) = \frac{1}{R} \sum_{i=0}^{i=l} G_s(i,\theta) + G_d(i,\theta) + G_p(i,\theta)$$

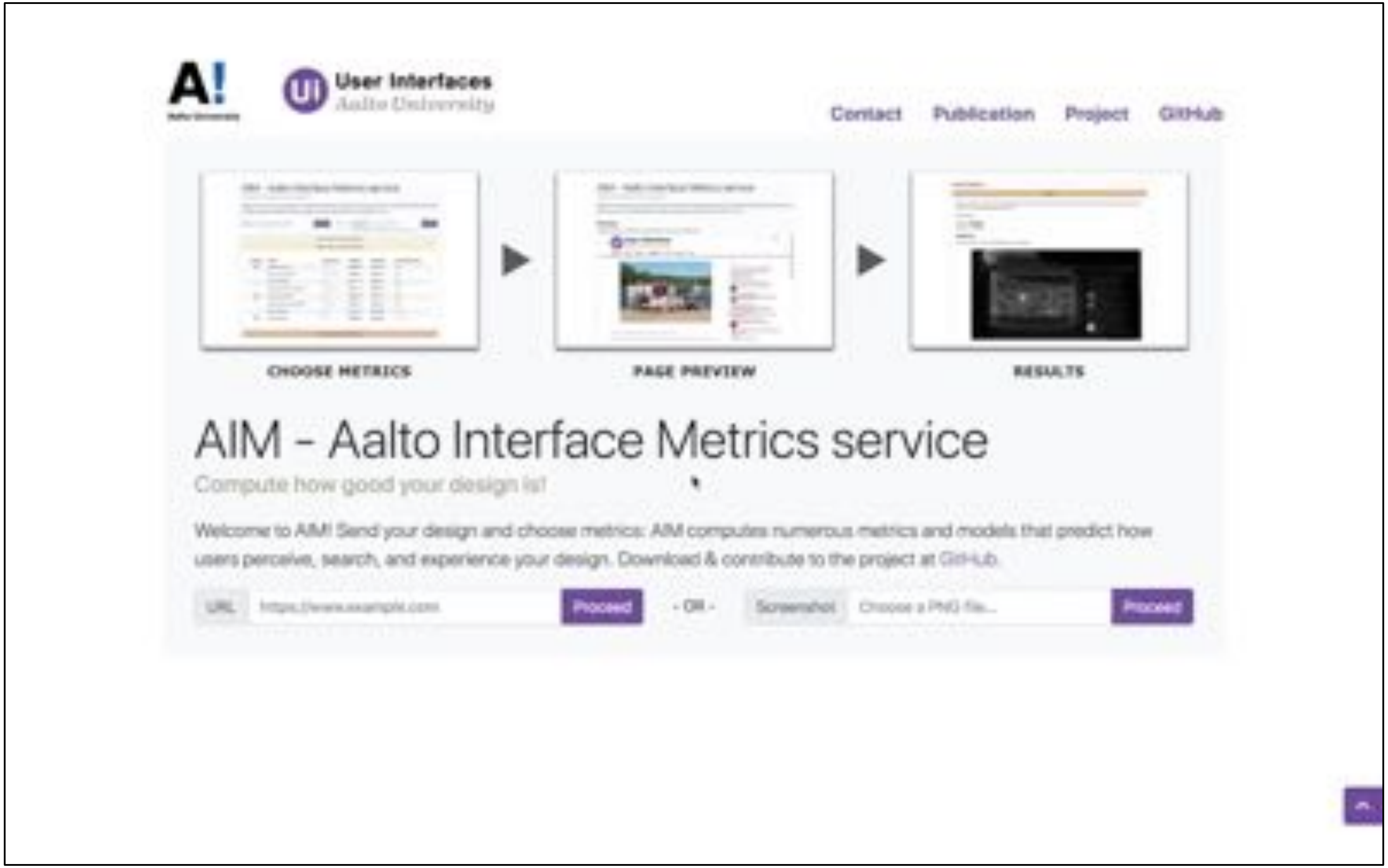
$$G_s(i,\theta) = \begin{cases} a_s \times \exp(-b_s \times p_i) + c_s & \text{if } i \leq t \\ 0 & \text{otherwise} \end{cases}$$

$$G_d(i,\theta) = \psi_{t,\sigma_d^2}(i) \times (a_d \times \exp(-b_d \times p_i) + c_d)$$

$$G_p(i,\theta) = a_p + b_p \times \log(1 + \alpha \times t) \times \psi_{t,1.1}(i)$$

Bailly, Oulasvirta, Brumby, Howes CHI 2014

interfacemetrics.aalto.fi



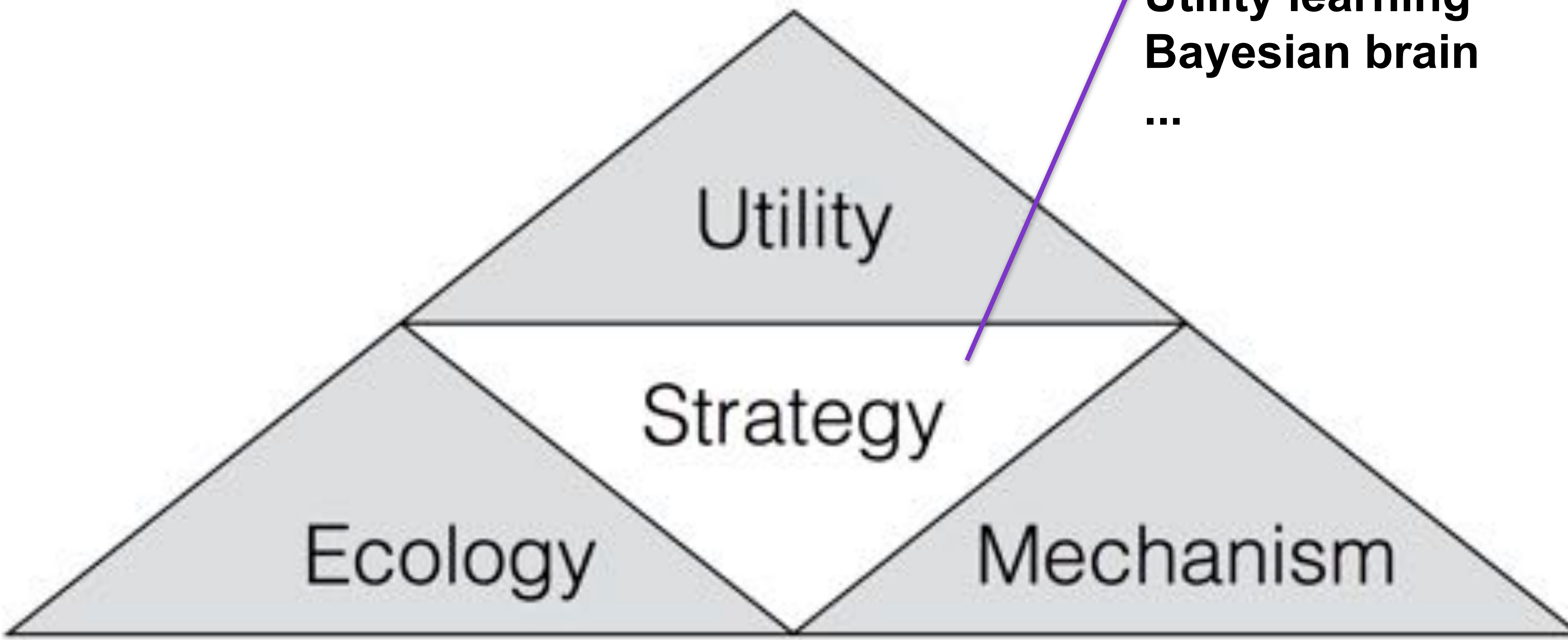
Aalto University

Machine learning is revolutionarizing cognitive modeling

How should I search for “politics”?

ML can predict users' behavioral strategies (how people adapt)

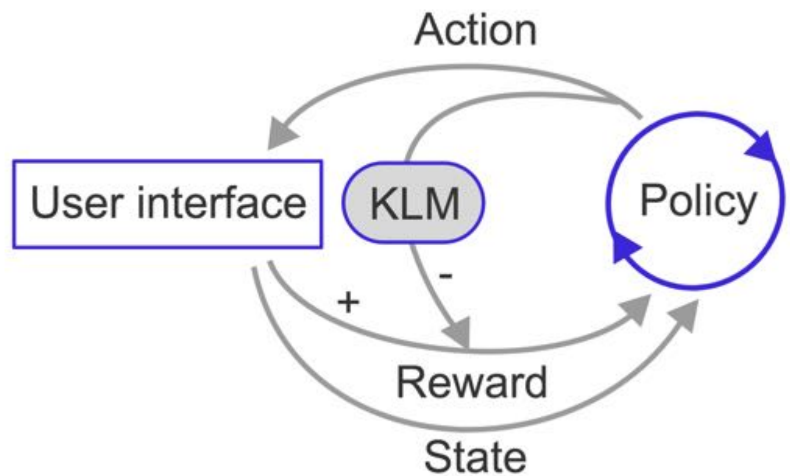
Policy learning
Utility learning
Bayesian brain
...



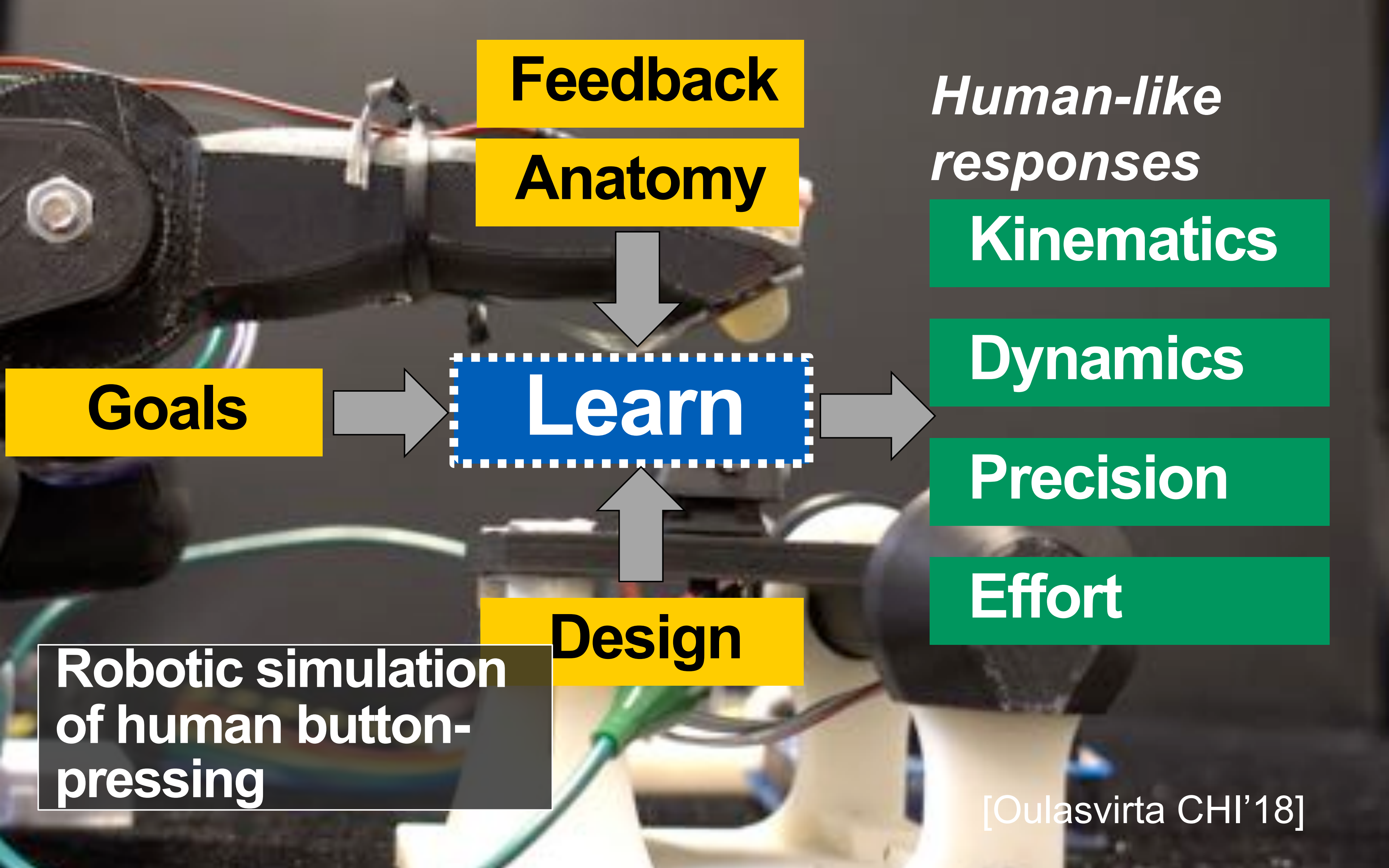
Payne & Howes 2013; Chen et al. 2015

RL-KLM

Automated assessment of task completion time on point-and-click UIs with reinforcement learning & KLM

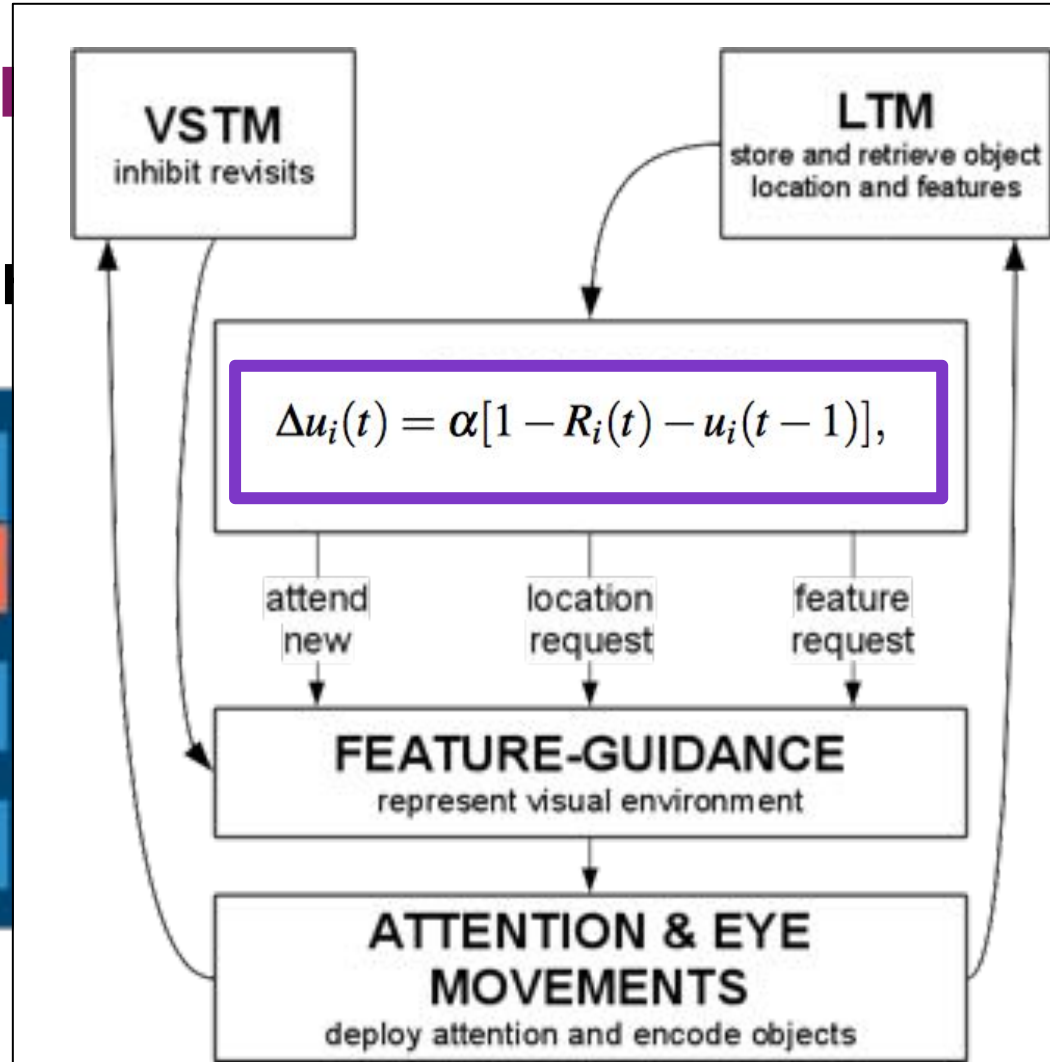


Leino et al. Proc. IUI'19


Feedback
Anatomy
Human-like responses
Goals
Learn
Kinematics
Dynamics
Precision
Effort
Design
Robotic simulation of human button-pressing
[Oulasvirta CHI'18]

Visual sampli

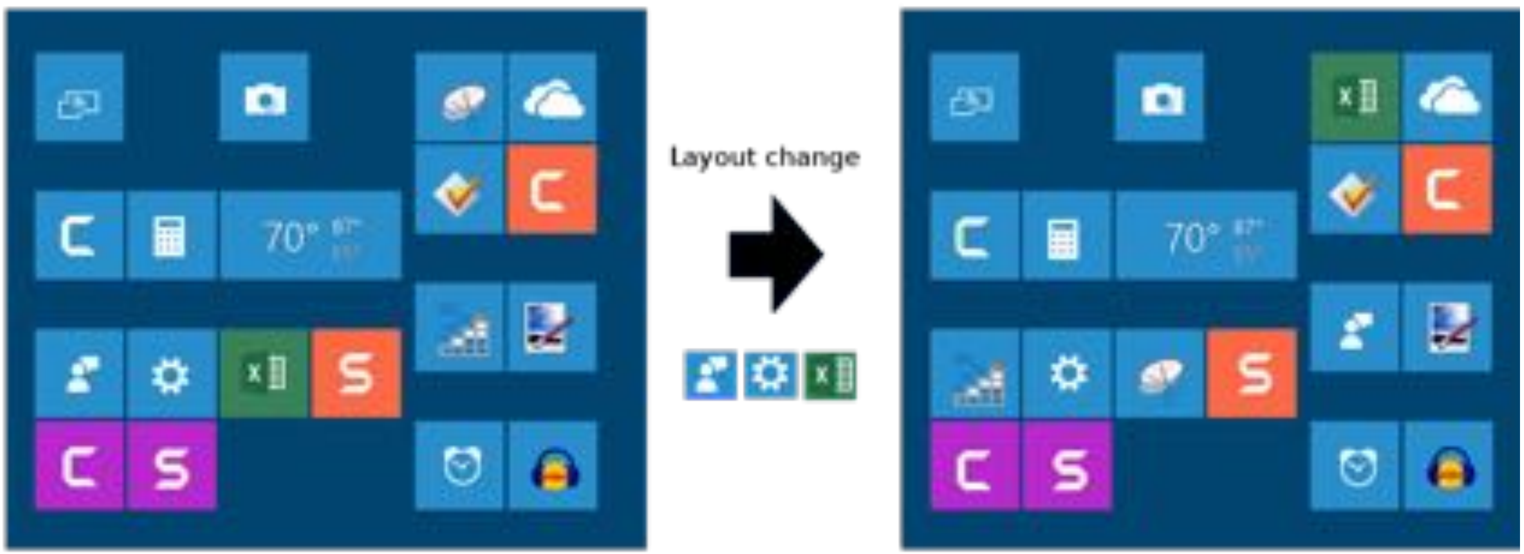
Modeling visual searc



Jokinen et al. Proc. CHI 2017

Visual sampling after layout change

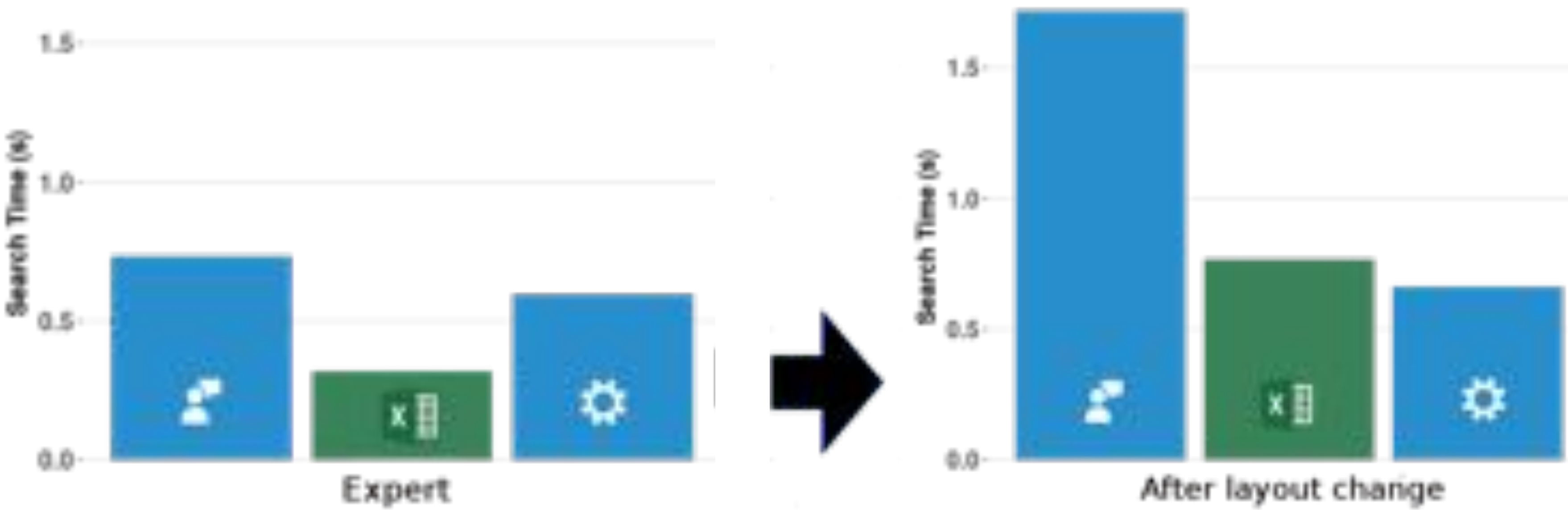
Reorganization of visual search behavior after a layout has changed



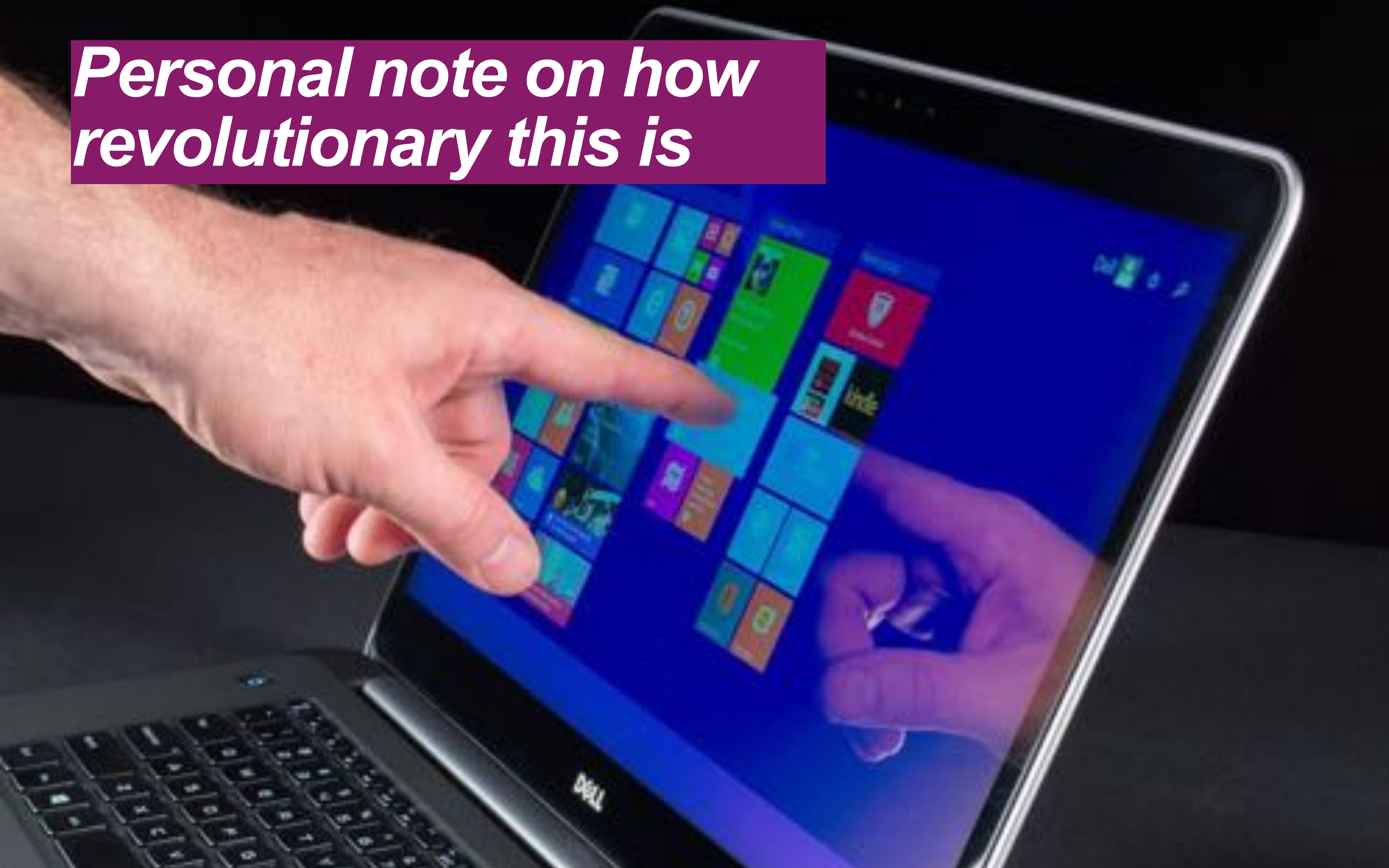
Jokinen et al. Proc. CHI 2017

Results: example

Effects of layout change on visual sampling strategy and therefore search costs



Personal note on how revolutionary this is

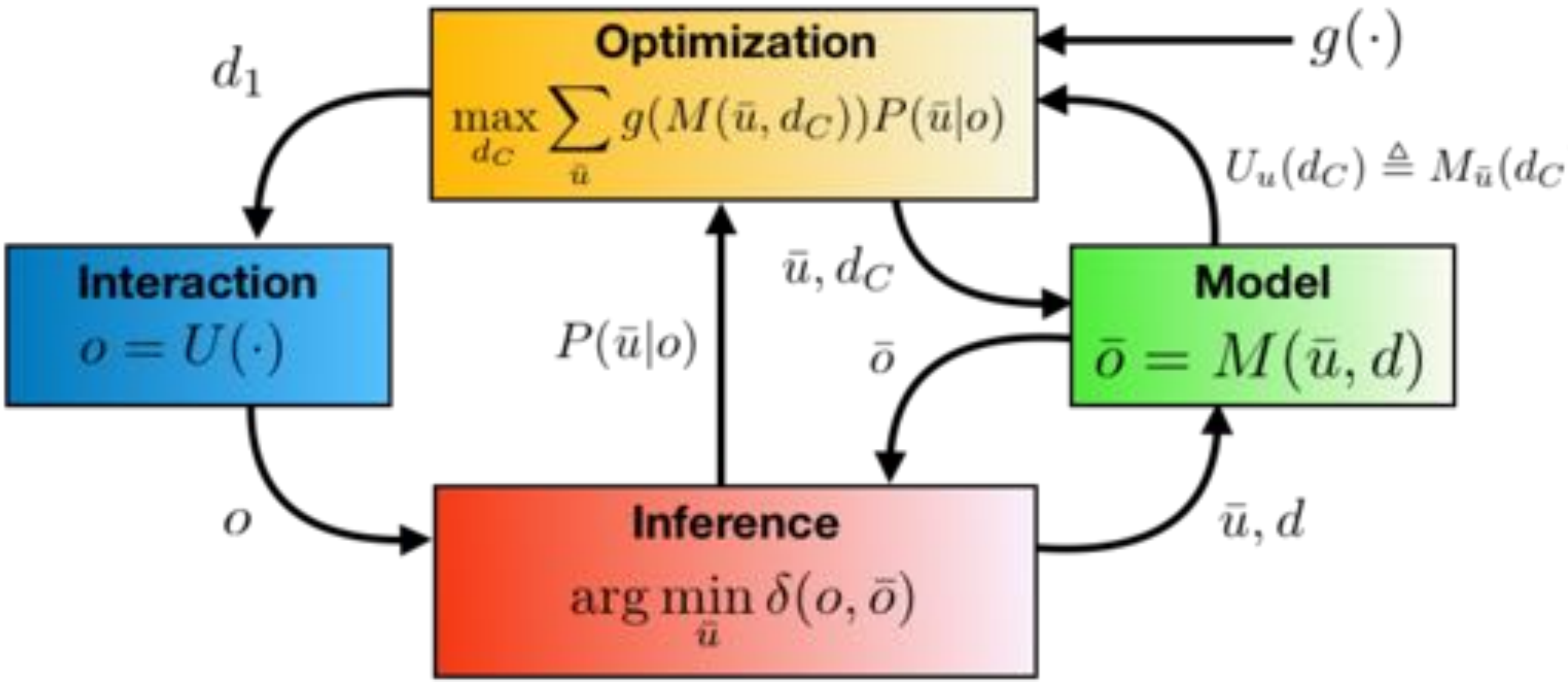


Likelihood-free inference allows learning model parameters from data

A bit more detail…

Inference

Model-driven human-computer interaction

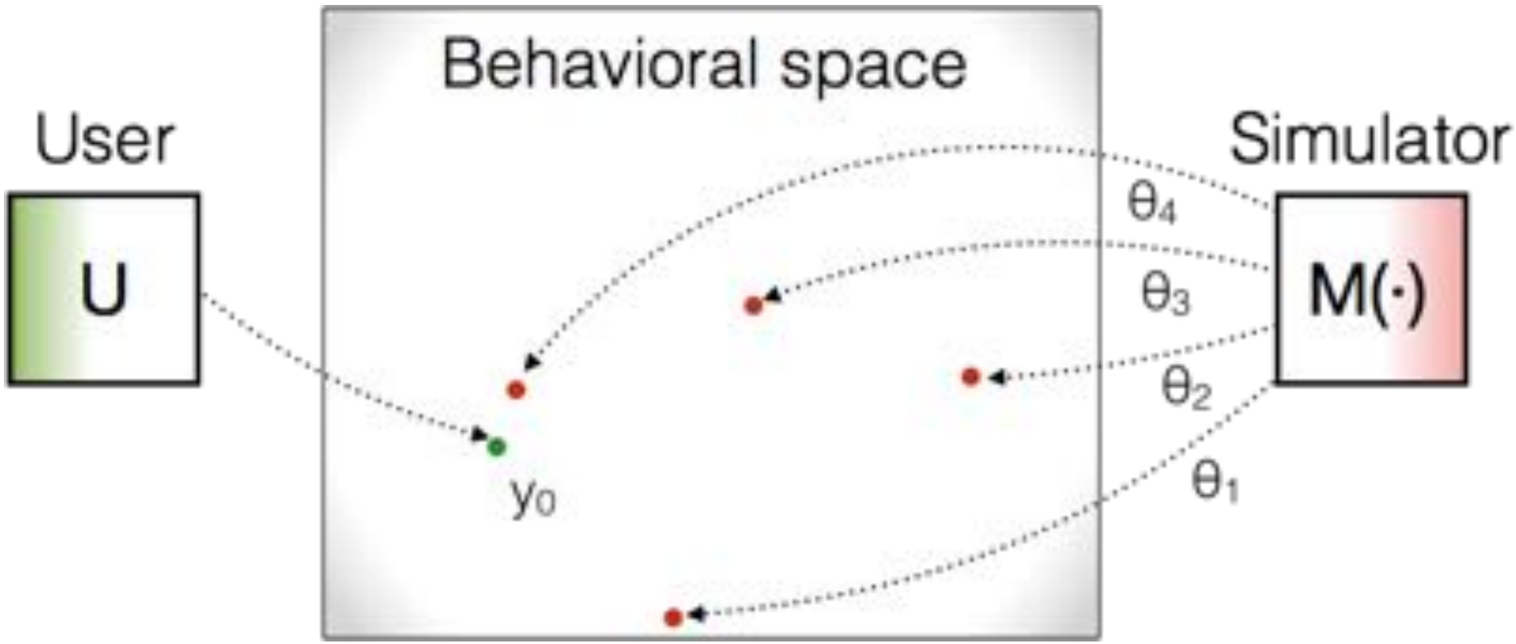


What are we allowed to infer from a click?

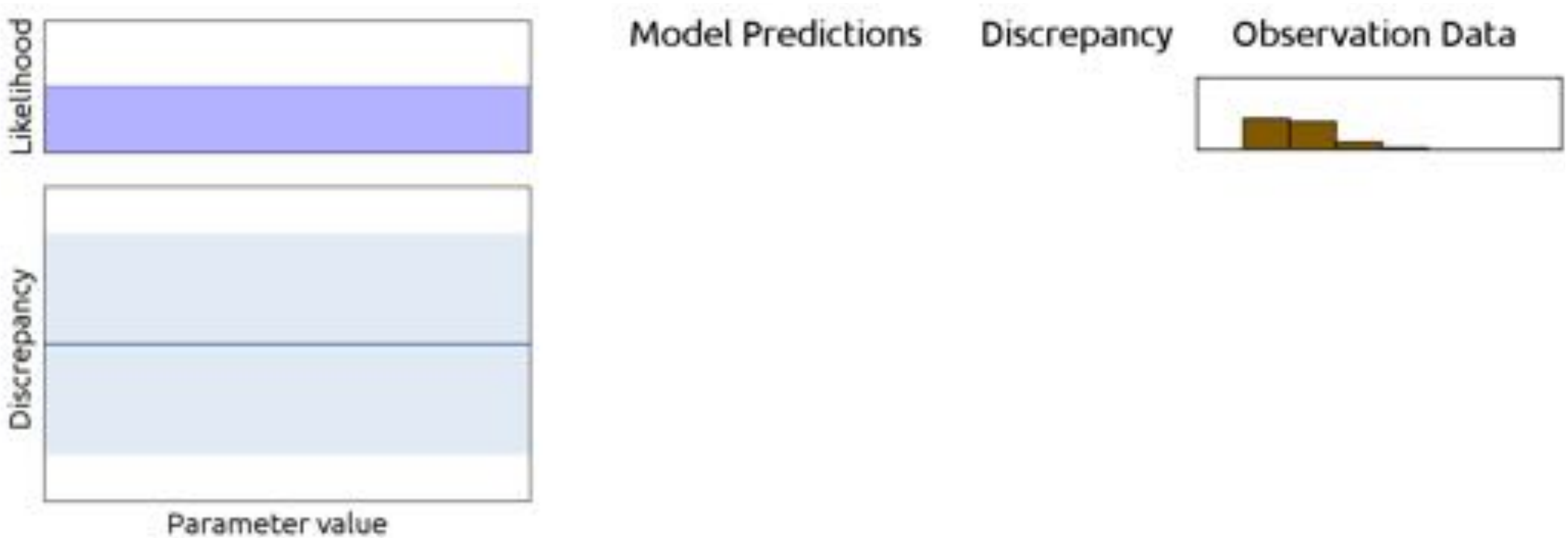
Algorithmic Sherlock Holmes

Fitting simulator parameters to log data with Bayesian methods

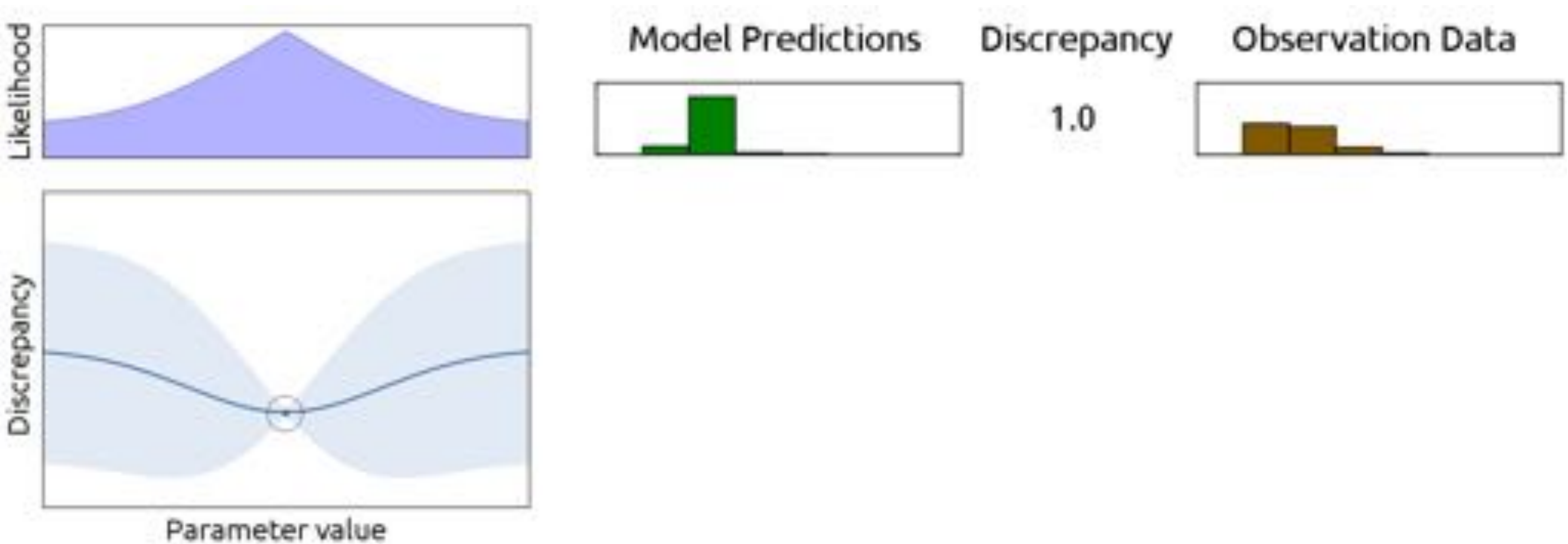
Approximate Bayesian Computation (ABC)



How ABC works

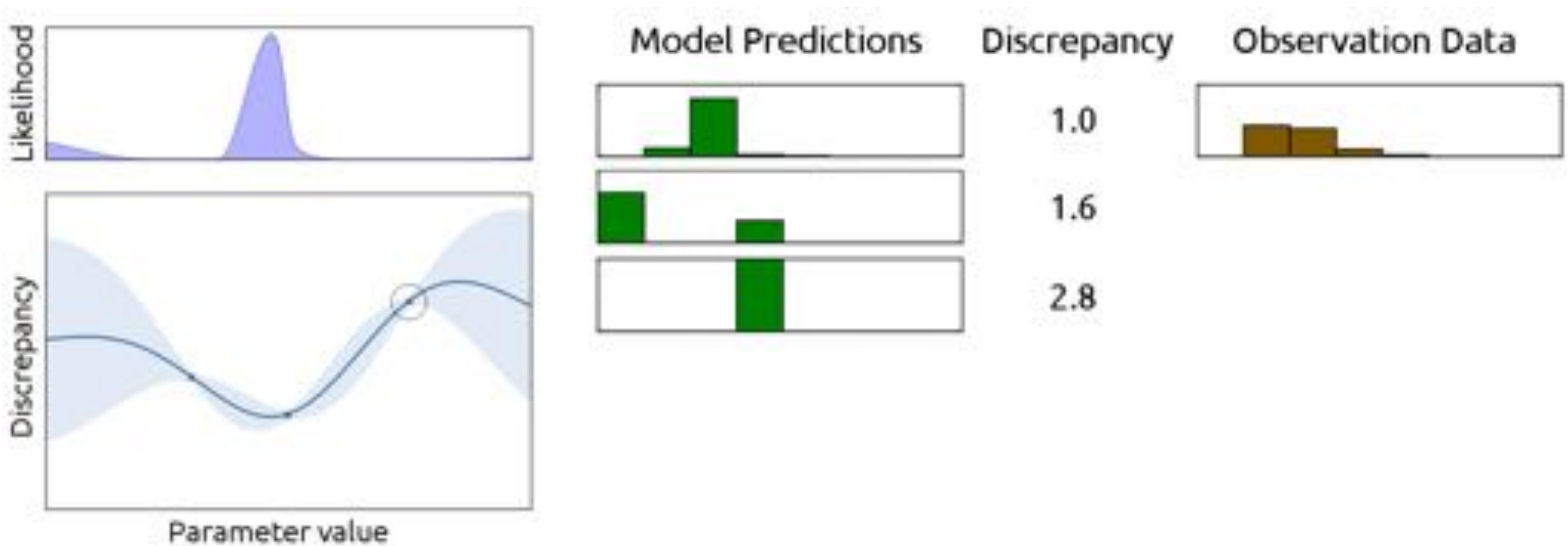


How ABC works

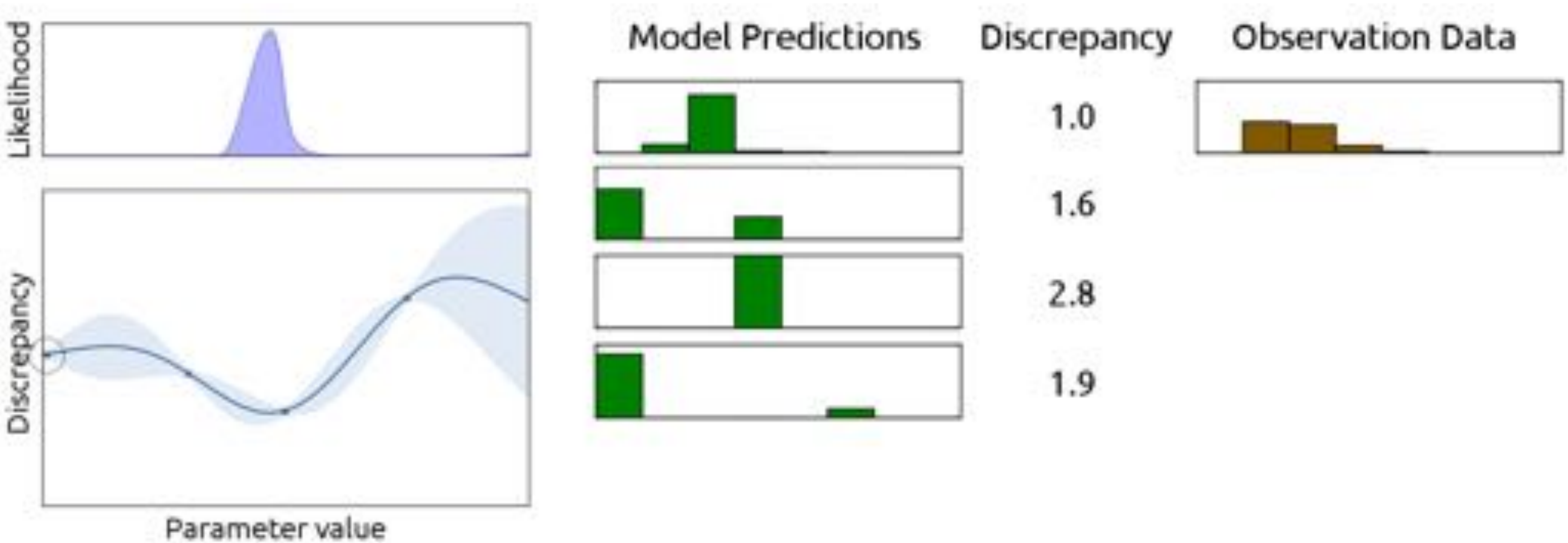


How ABC works

How ABC works

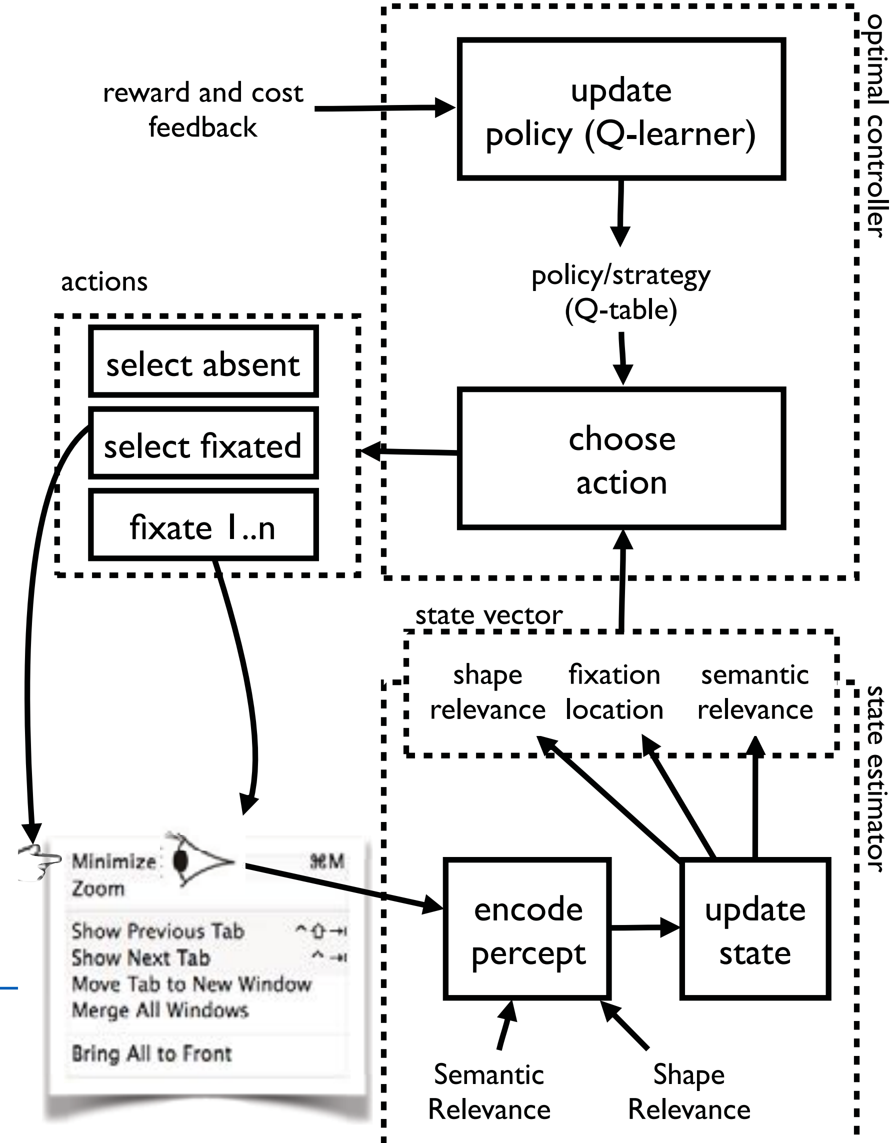


How ABC works



Example: A model of menu search

Finds optimal gaze pattern given menu design and parameters of the visual and cognitive system



Aalto University

[Chen et al. CHI'15]

ABC helps us infer model parameters from data logs

Given click times, predict model (HVS) parameters

Click times

Parameter	Description
f_{dur}	Fixation duration
d_{sel}	Time cost for selecting an item (added to the duration of the last fixation of the episode if the user made a selection)
p_{rec}	Probability of recalling the semantic relevances of all of the menu items during the first fixation of the episode
p_{sem}	Probability of perceiving the semantic relevance of ... and below of the ...

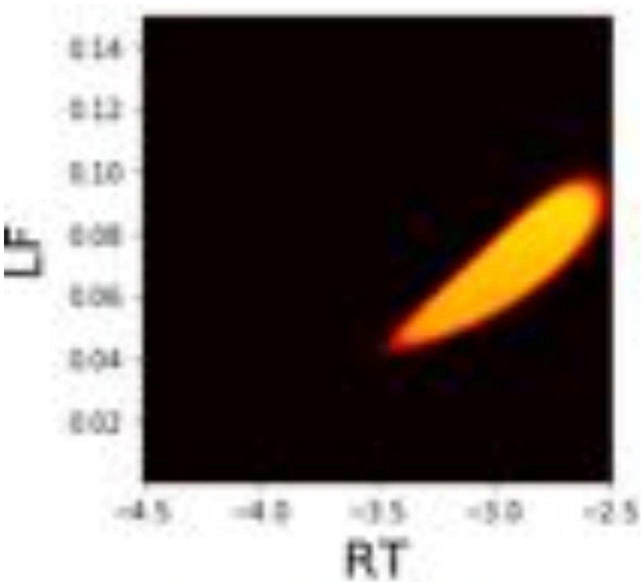
Table 1. Parameters inferred wit

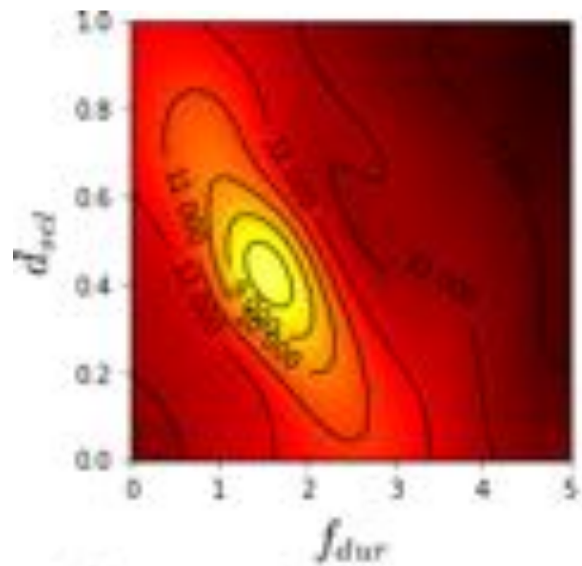
Generate user-like behavior for new designs

Kangasrääsiö et al. CHI 2107, Cognitive Science

ABC yields posterior estimations

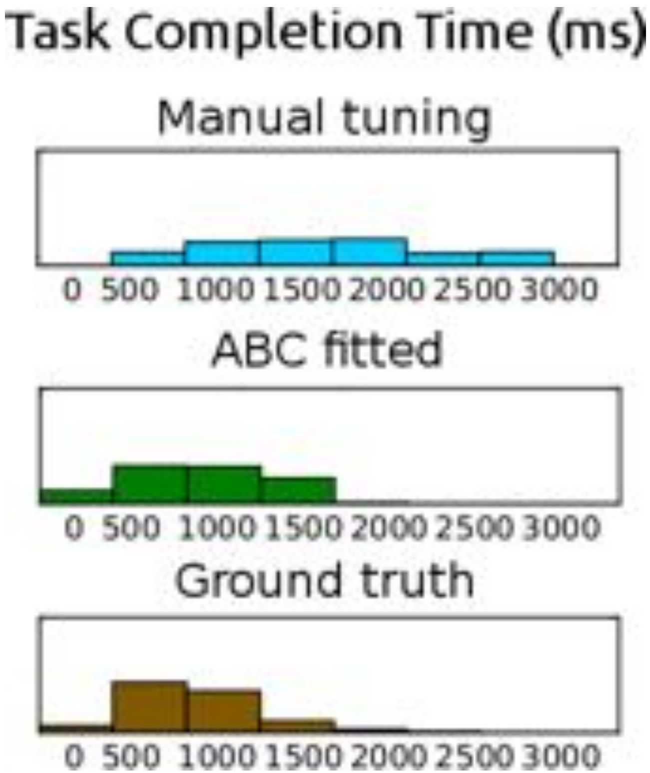
Bayesian inference yields a posterior distribution for model parameters





"Possible explanations for the data"

ABC improves fit over manual tuning



Mean TCT 1.49 s

Mean TCT 0.93 s

Mean TCT 0.92s

Kangasrääsiö et al. CHI 2107

Examples of model-driven inferences

Visual attention: The items the user is looking for are hard to find

Motor control: The interactions are too hard or cumbersome to execute

Navigation: User does not understand site structure

Aesthetics: The page is perceived ugly and confusing

Skills: User does not have sufficient motor or conceptual skills to use the page

Intentions and preferences: What items and which types of services or interaction users like

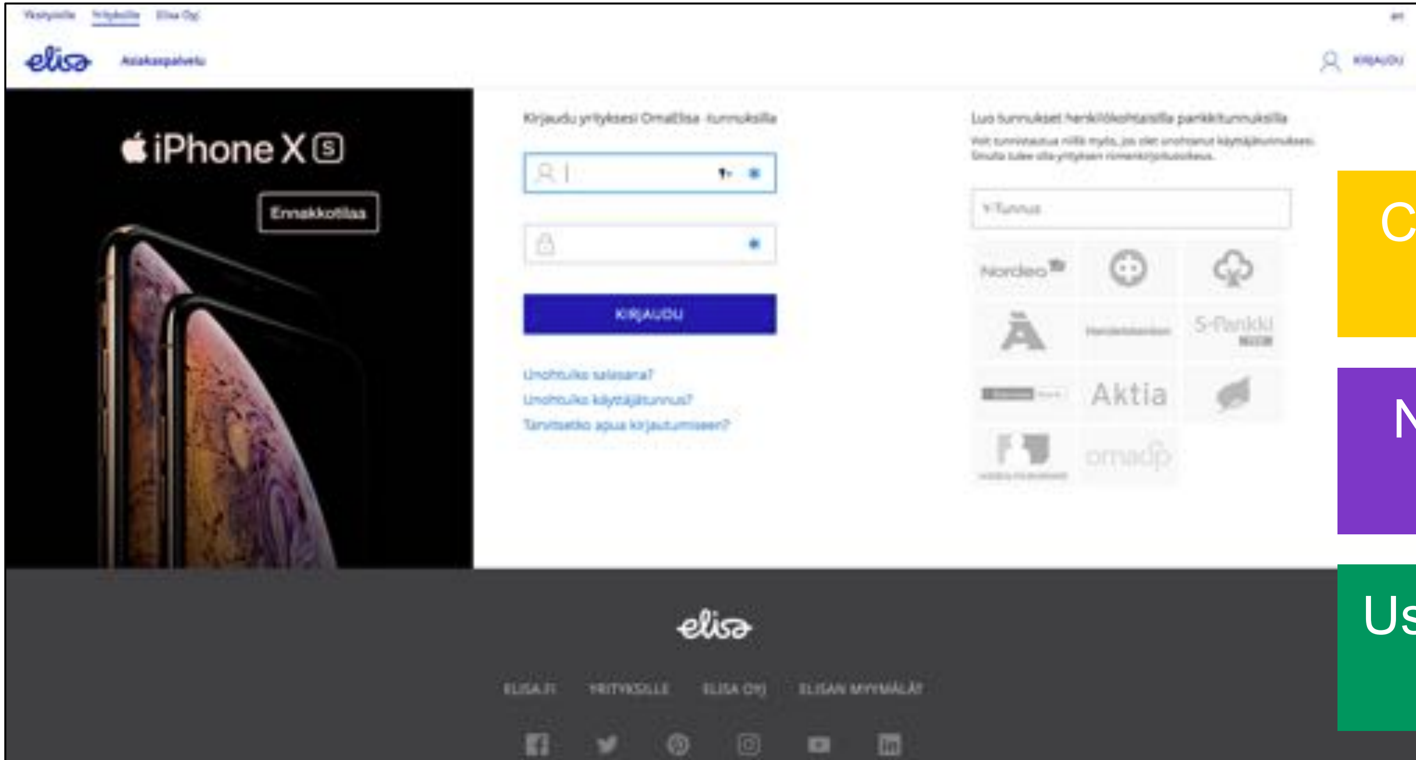
Errors and mistakes: Users inadvertently do something they would not have liked to

Cultural background: Language and other cultural explanations to style of use

Decision-making: Users' strategies and goals in decision-making

Permission from Esko Kurvinen / Elisa

Example: Getting lost



Clickstream data

Navigation models

Users getting lost?

Dashboard

Applications

Computational design of keyboard layouts

The new standard for the French AZERTY keyboard

Solved a very hard combinatorial problem: over 10^{213} possible designs

Anna Feit, Mathieu Nancel et al.

Example: Hierarchical menus

Niraj Dayama

Ordering commands and assigning them into tabs and groups.
two objectives: selection time and associativity:

Tab 1
- Full-Screen-Mode
- Read-Mode
- Read-Aloud
- Print
- Properties
- Save-As
- Send-File
- Open
- Revert
- Close
- Exit
- Save
- Save-As-Other

Tab 2
- Copy
- Take-A-Snapshot
- Select-All
- Deselect-All
- Cut
- Delete
- Copy-file-to-clipboard
- Undo
- Paste
- Redo

Tab 3
- Cascade
- Tile
- Zoom
- Page-Navigation
- New-Window
- Minimize-All-Windows
- Rotate-View
- Page-Display

Tab 4
- Look-up-selected-wor
- Check-Spelling
- Find
- Advanced-Search
- Analysis

Reader app (Adobe)

26% faster to use than commercial designs

Browser (Firefox)

Tab 1
- New
- Open-File
- Open-Location
- Select-All
- Find-Again
- Save-Page-As
- Print
- Page-Setup
- Clear-Recent-History
- User-History-Item
- Start-Page
- Show-All-History
- Restore-Previous-Session

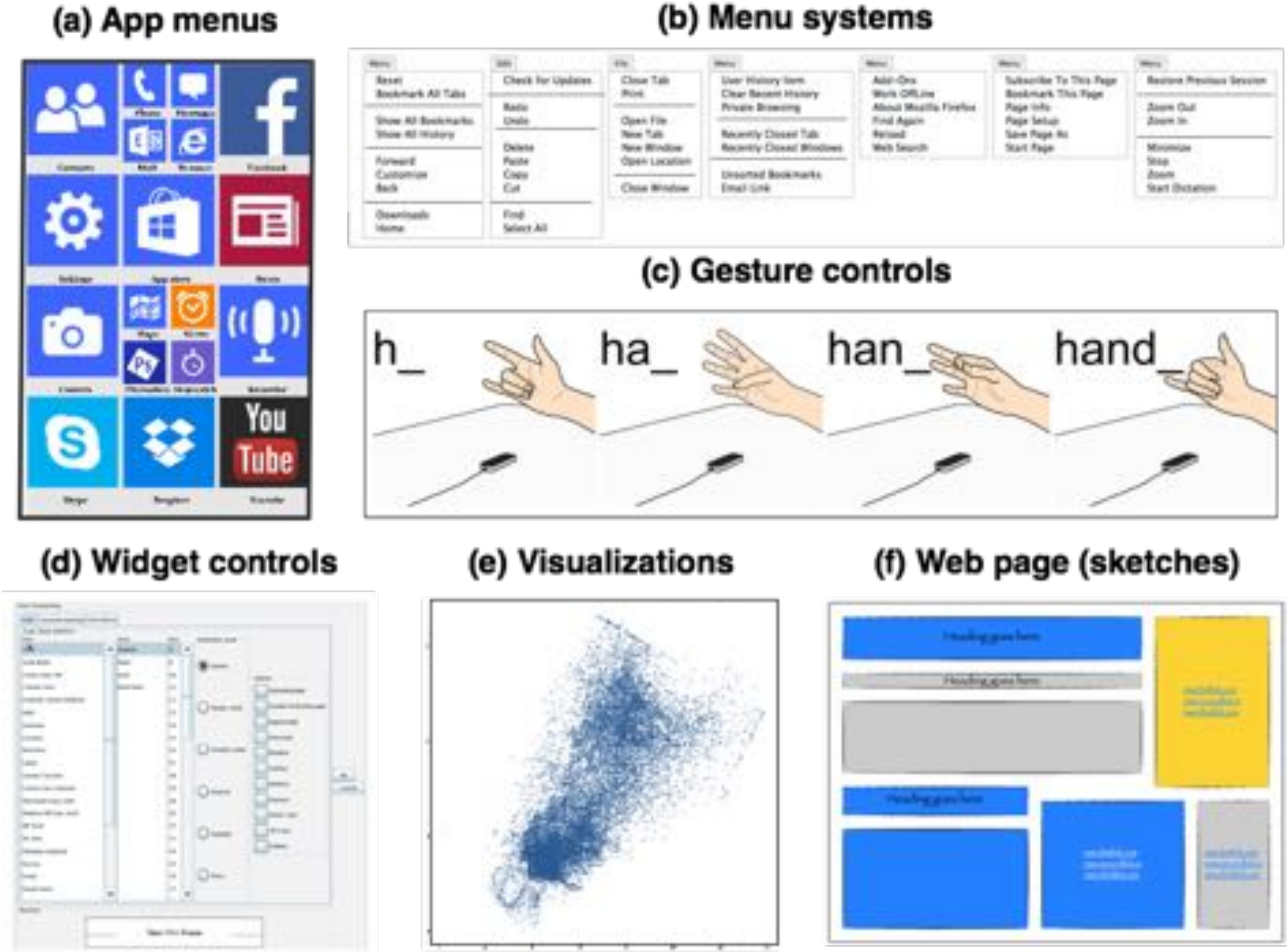
Tab 2
- Downloads
- About-Mozzila-Firefox
- Subscribe-To-This-Page
- Check-for-Updates
- Unsorted-Bookmarks
- Web-Search
- Bookmark-This-Page
- Bookmark-All-Pages
- Show-All-Bookmarks
- Add-Ons
- Work-Offline

Tab 3
- Private-Browsing
- New-Window
- Close-Window
- Email-Link
- Start-Dictation
- RecentlyClosed-Tab
- Recently-Closed-Windows
- Close-Tab
- Customize
- Zoom-In
- Minimize
- Zoom
- Zoom-Out

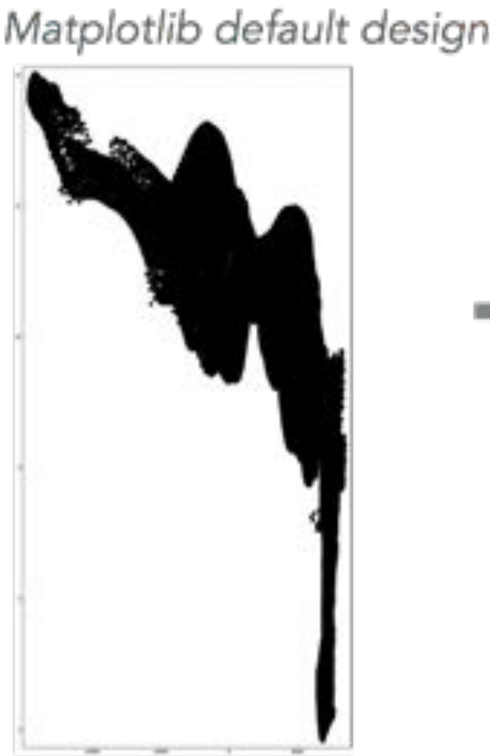
Tab 4
- Find
- Home
- Back
- Forward
- Reset
- Page-Info
- Reload-3
- Stop
- Undo
- Redo

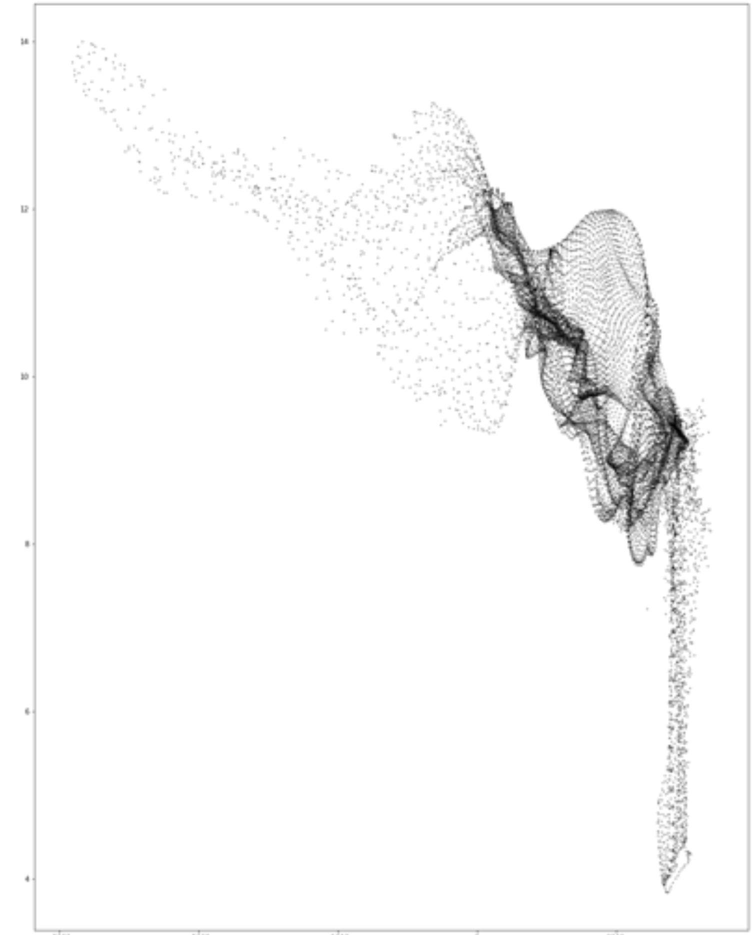
Tab 5
- Copy
- Paste
- Delete
- Cut


Antti, you're right.
But ... You're not dreaming BIG enough
Donald Norman
... at my NordiCHI keynote in 2014


(a) App menus
(b) Menu systems
(c) Gesture controls
h_
ha_
han_
hand_
(d) Widget controls
(e) Visualizations
(f) Web page (sketches)

Example: Perceptual optimization of scatterplots



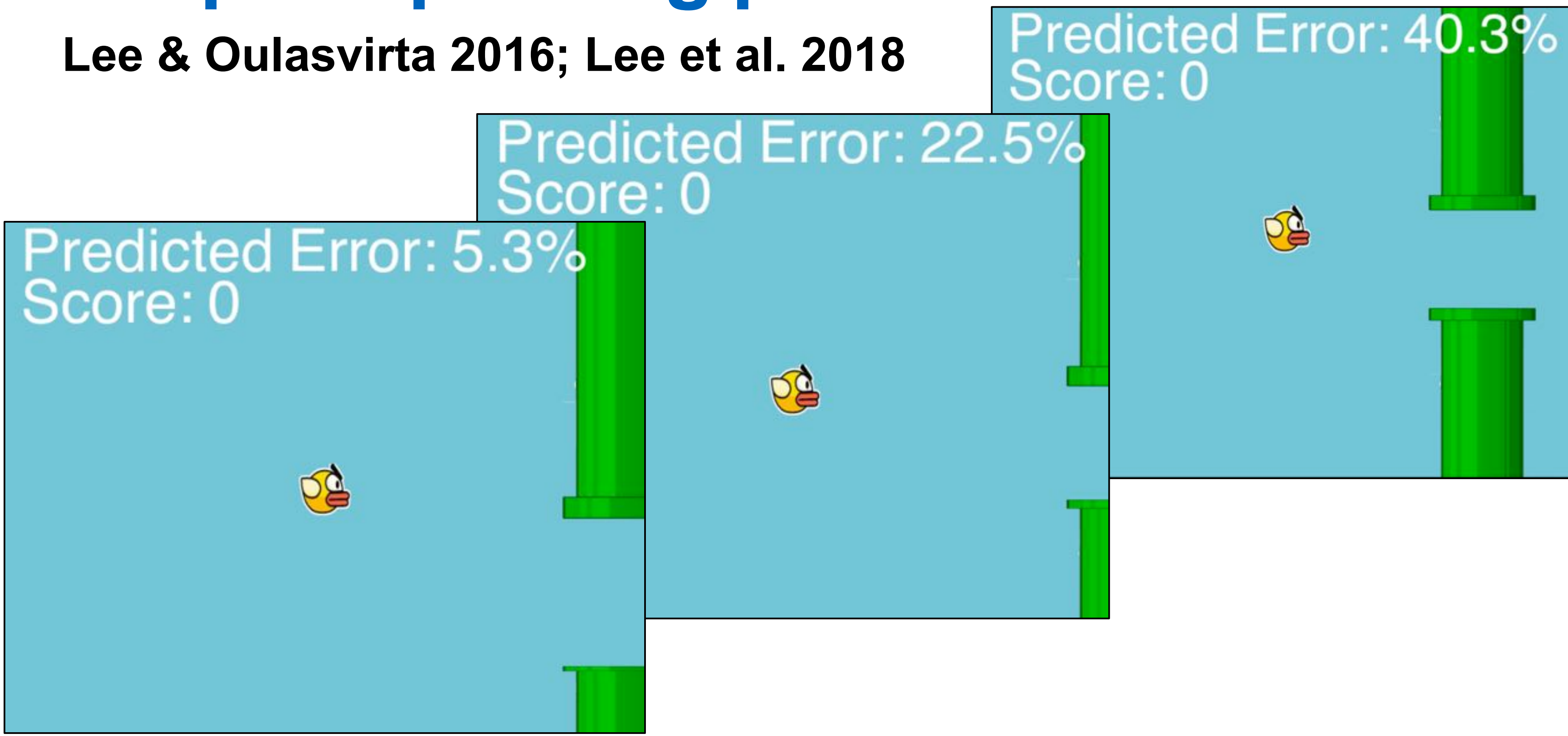


Micallef et al. IEEE TGCV 2017

Optimizing skim reading

Optimization of game levels for temporal pointing performance

Lee & Oulasvirta 2016; Lee et al. 2018



CHI 2018
Engage with CHI

AdaM

Adapting Multi-User Interfaces for Collaborative Environments in Real-Time

Seonwook Park, Christoph Gebhardt, Roman Rädle, Anna Feit, Hana Vrzakova, Niraj Dayama, Hui-Shyong Yeo, Clemens Klokmose, Aaron Quigley, Antti Oulasvirta, Otmar Hilliges

Advanced Interactive Technologies ETH zürich Aarhus University Aalto University University of Eastern Finland University of St Andrews

Tool concepts that "empower not just automate" designers

Interactive optimization and learning systems

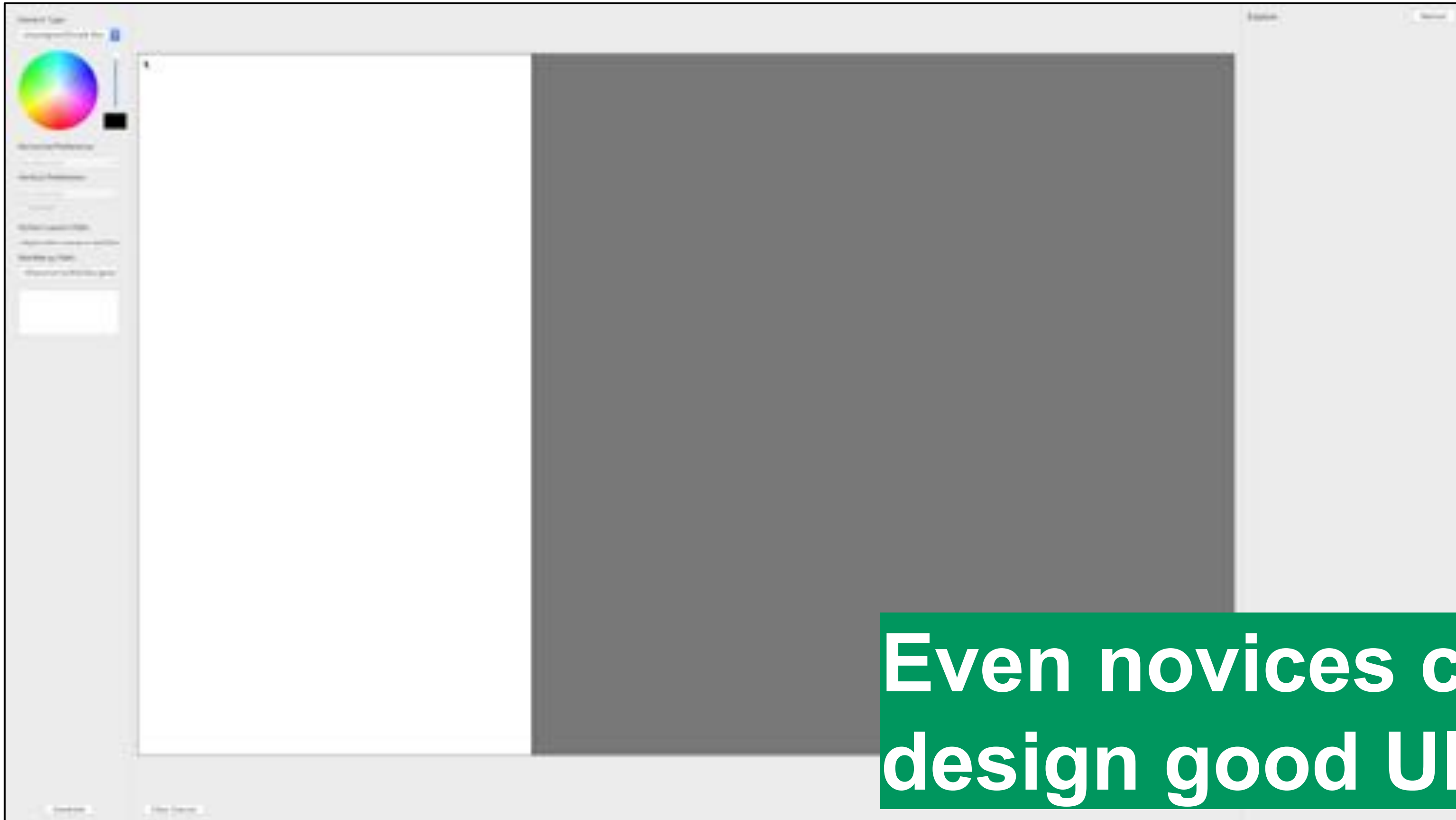
Interactive support for design

MenuOptimizer: Interactive Optimization of Menu Systems

Gilles Bailly
Antti Oulasvirta
Timo Kötzing
Sabrina Hoppe

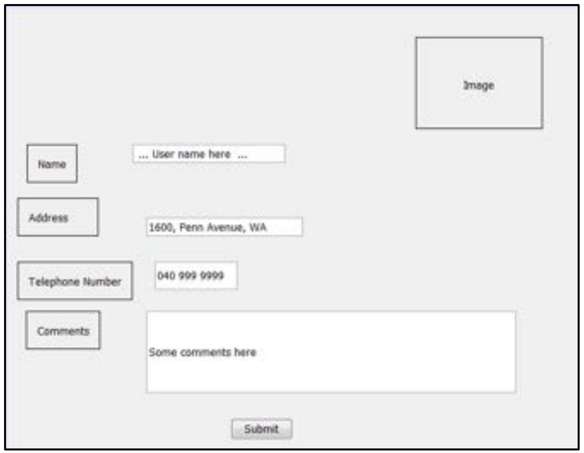
Interactive example galleries

[Ramesh et al. submitted]



Even novices can design good UIs

Style transfer



Initial sketch

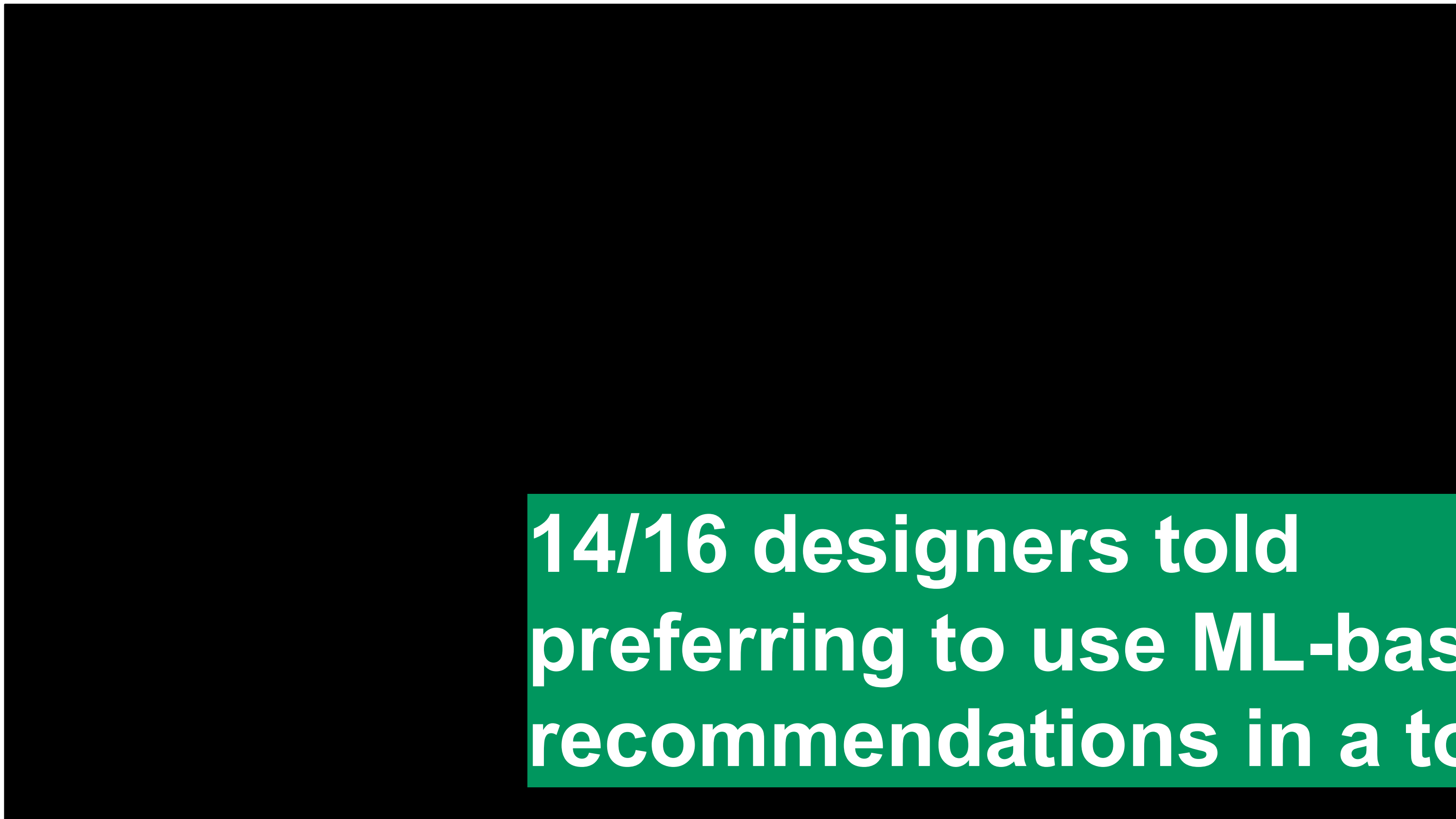
+

Closest Match in Library

=

Transferred design

Using bandits to adapt suggestions to designer's style

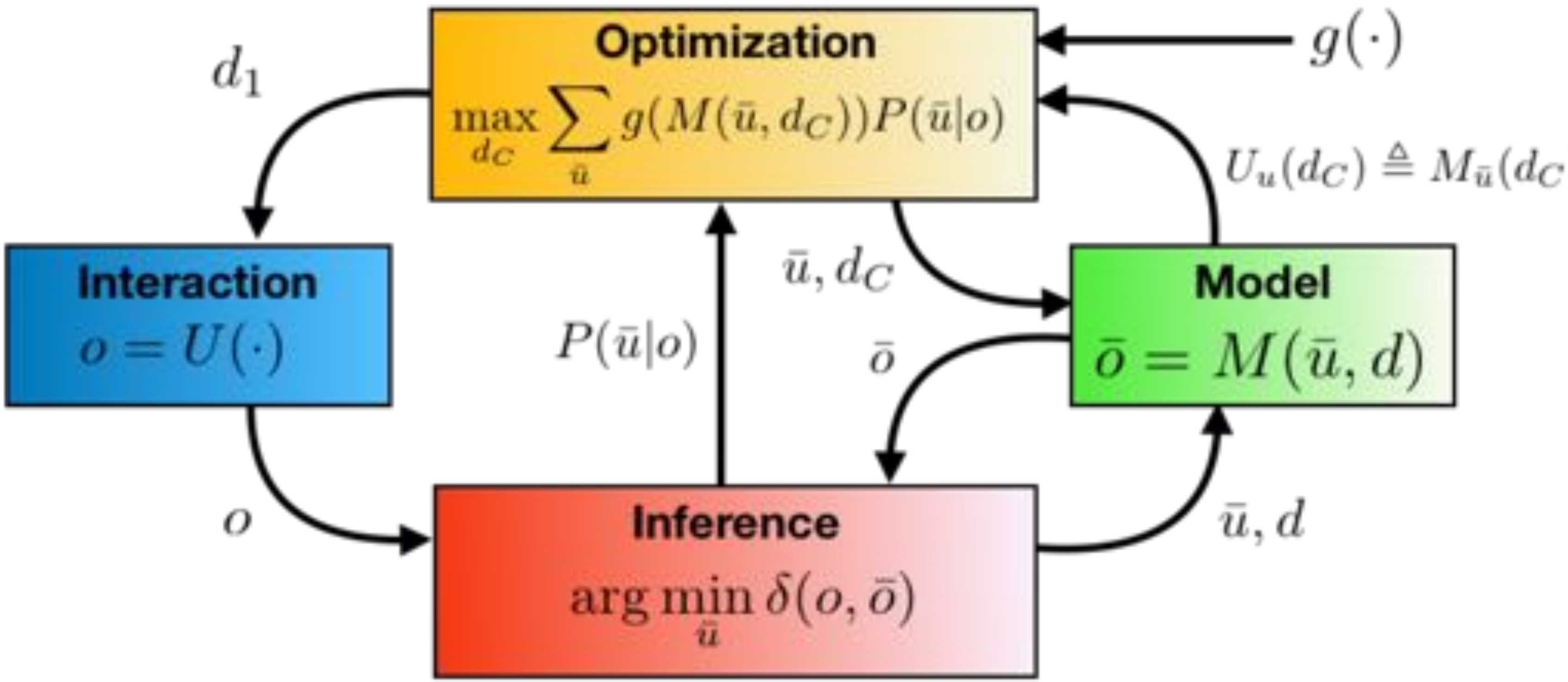


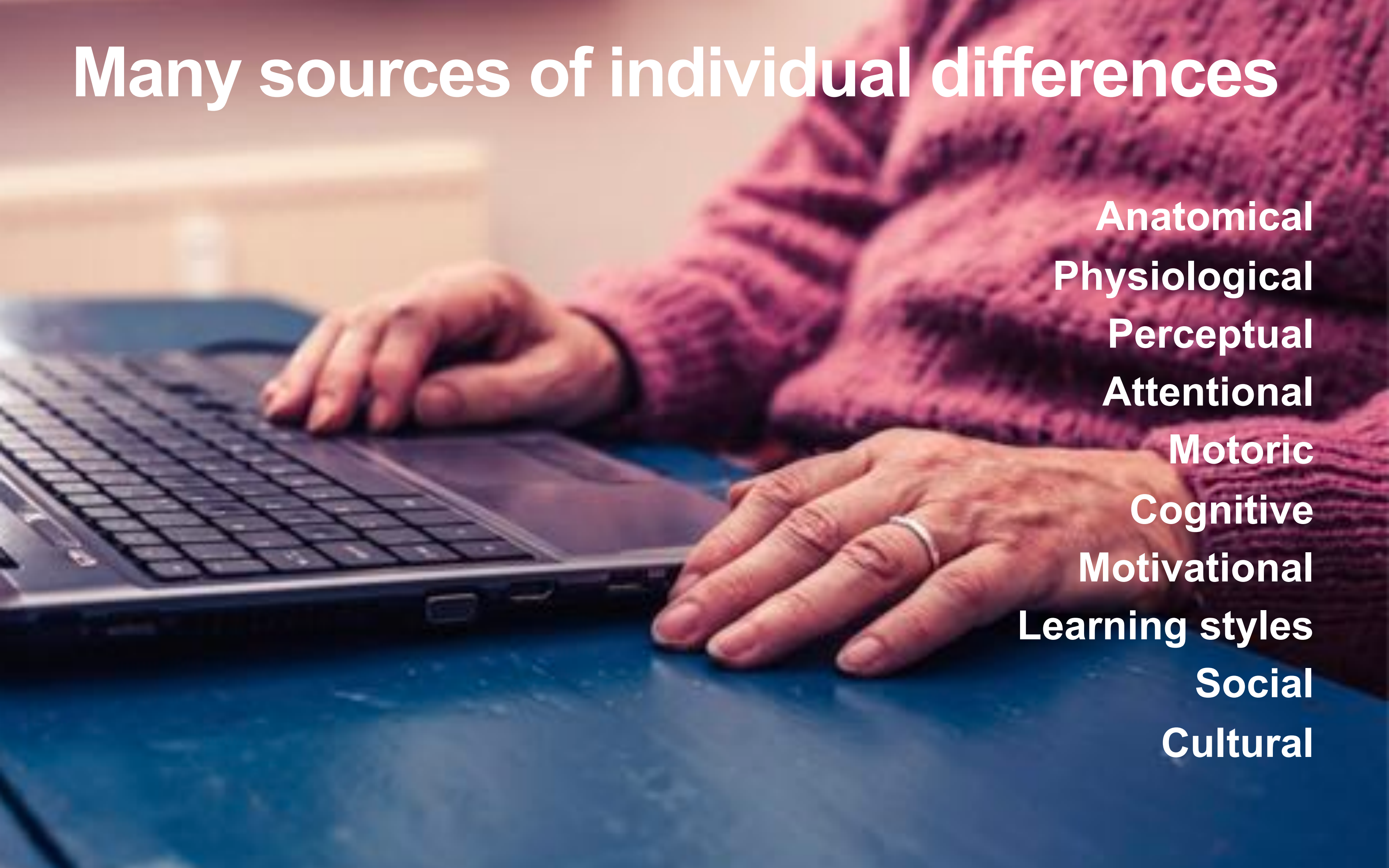
14/16 designers told preferring to use ML-based recommendations in a tool

[Koch et al. CHI'19]

Ability-based adaptation

Model-driven human-computer interaction





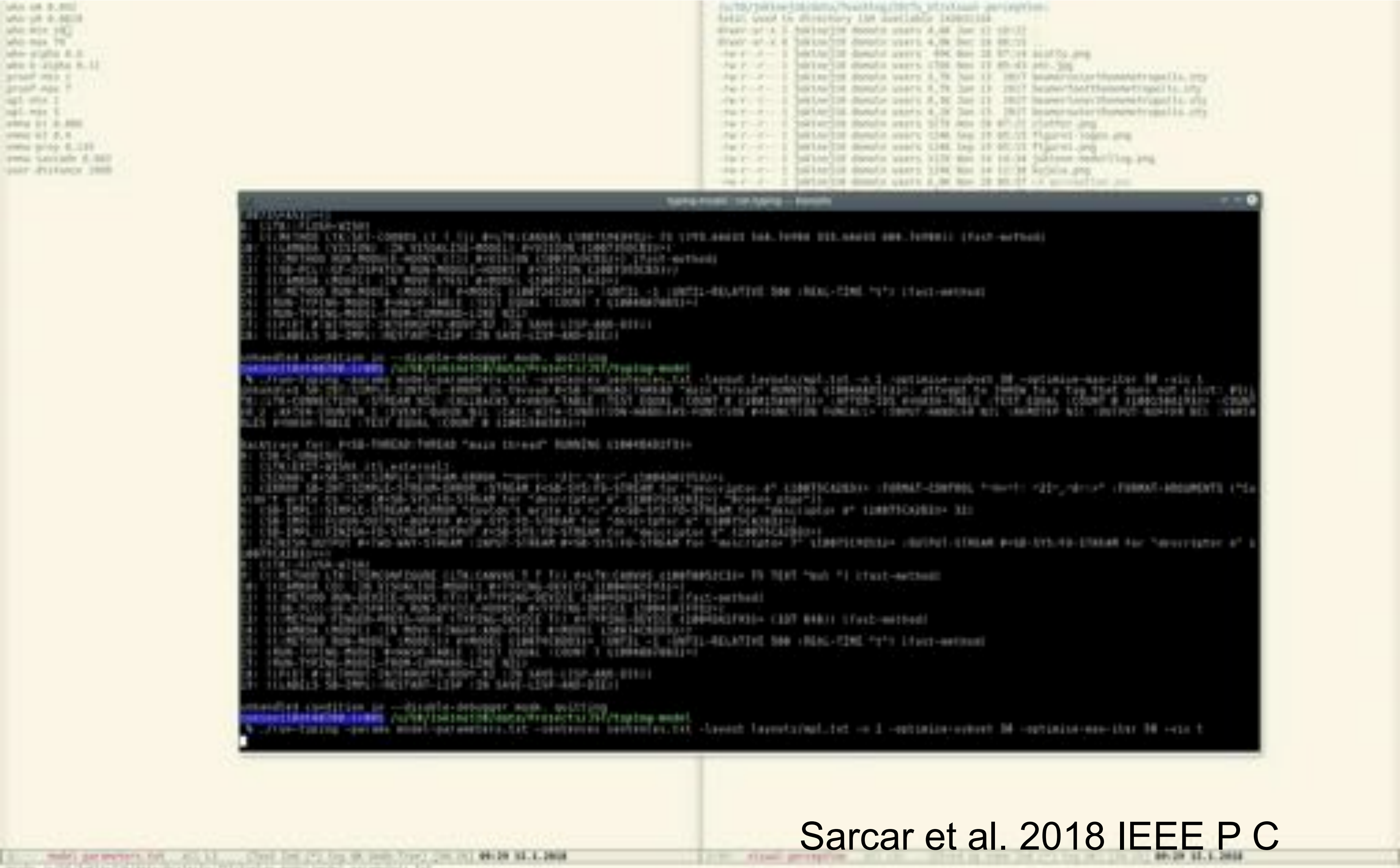
Many sources of individual differences

Anatomical
Physiological
Perceptual
Attentional
Motoric
Cognitive
Motivational
Learning styles
Social
Cultural

Individual differences are represented as model parameters

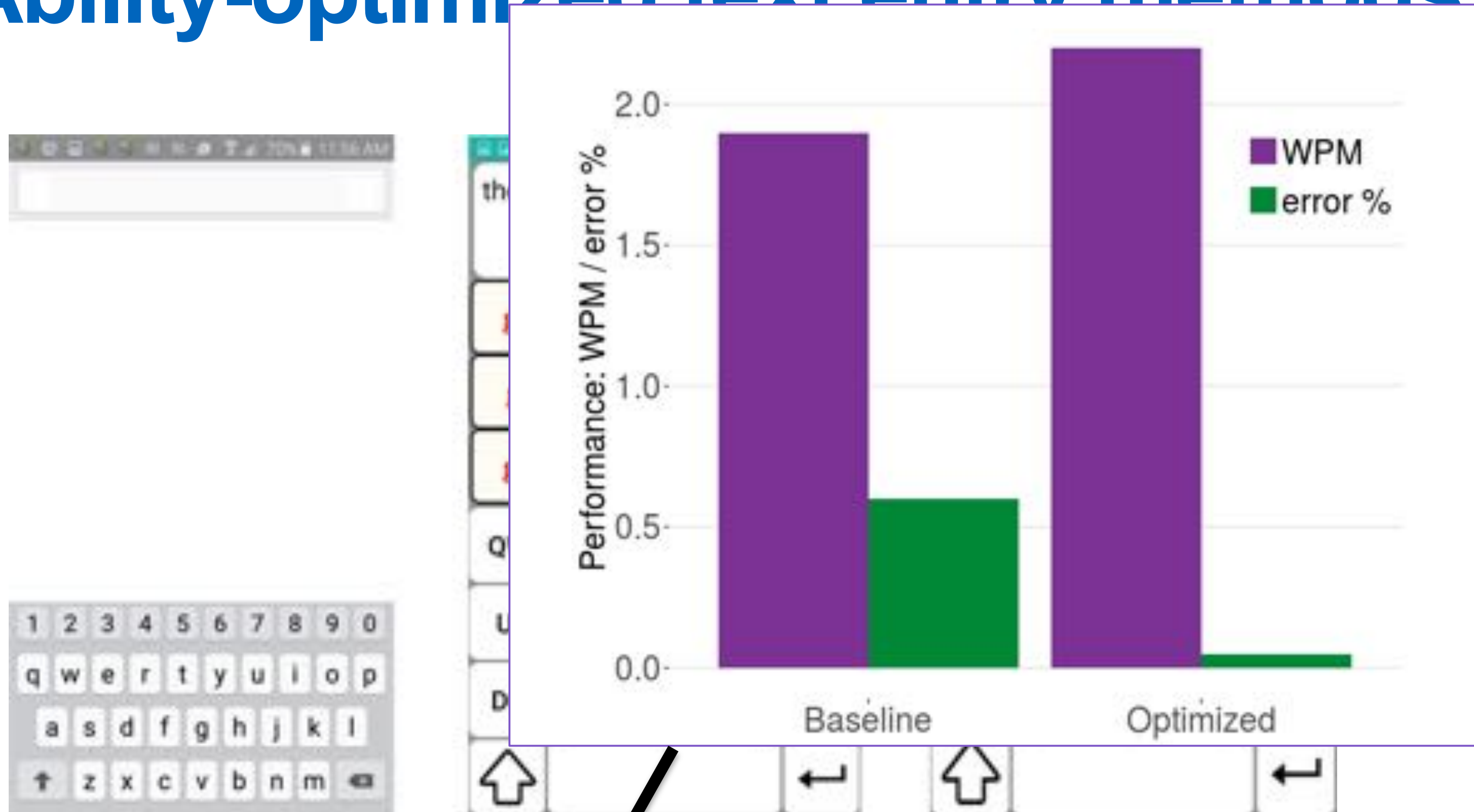
Table 1. Individual abilities modeled by *Touch-WLM*

Variable	Explanation	Domain
Eye movements		
e_K	Encoding time	Foveal encoding
e_k	Eccentricity factor	Parafoveal encoding
t_{prep}	Saccade preparation	Oculomotor command
t_{exec}	Saccade execution	Oculomotor command
t_{sacc}	Saccade velocity	Oculomotor performance
Motor performance		
m_k	Total resource	Motor performance
m_α	Speed–accuracy bias	Motor performance
Strategy		
m_a	Finger accuracy	Motor strategy
l	Letters before proofing	Cognitive strategy
Constants		
s_{key}	Search time for key	Visual search
$t_{confirm}$	Backspace confirmation	Thinking



Sarcar et al. 2018 IEEE P C

Ability-optimized text entry methods



Baseline **Tremor** **Dyslexia**

Sarcar et al. 2018 IEEE P C

Web interaction

Toward self-optimizing web services…

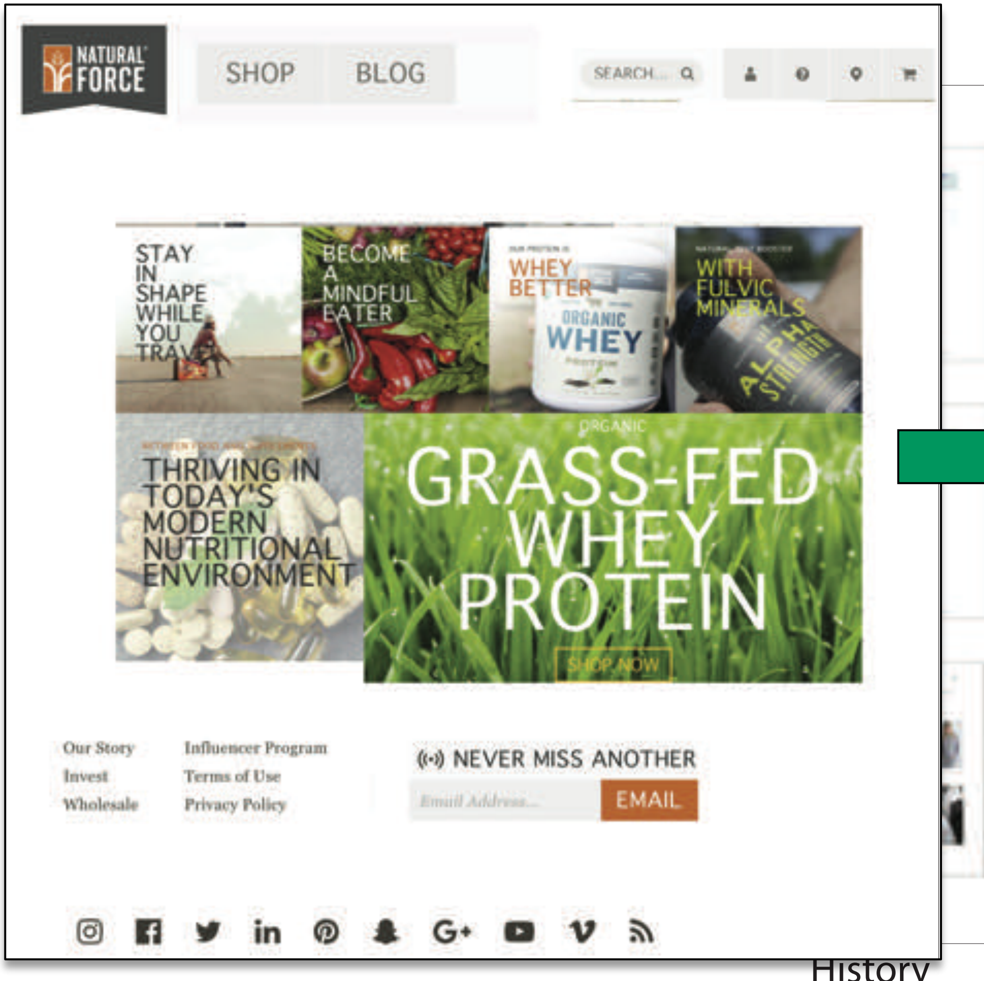
Browser-based menu adaptation

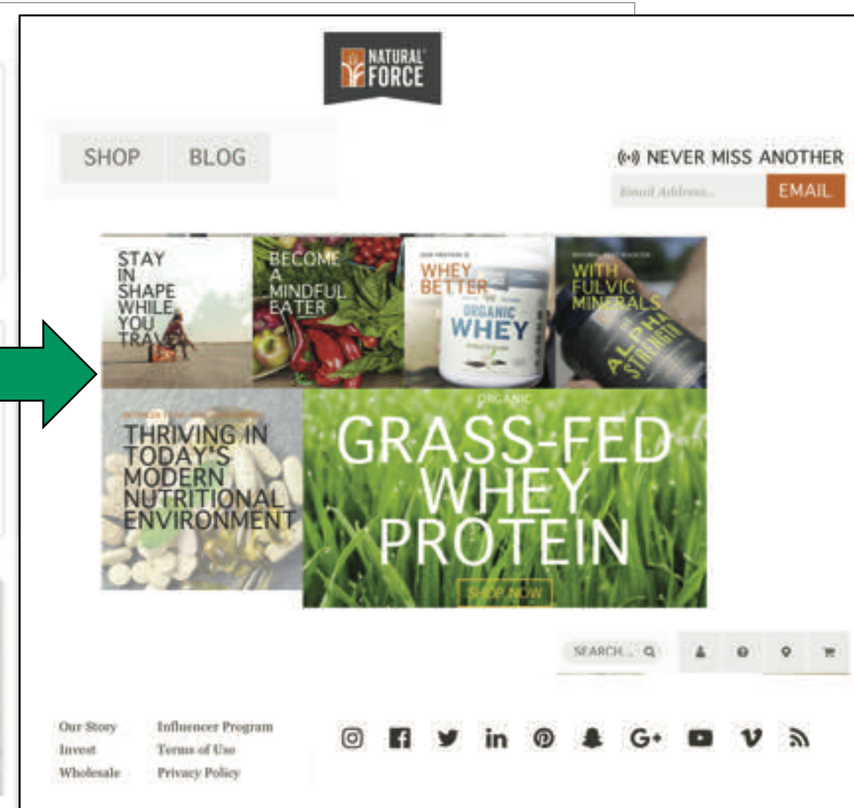
Accelerated

[Gobert et al. IUI’19]

Individualizing web pages

[Todi et al. IUI 2018]





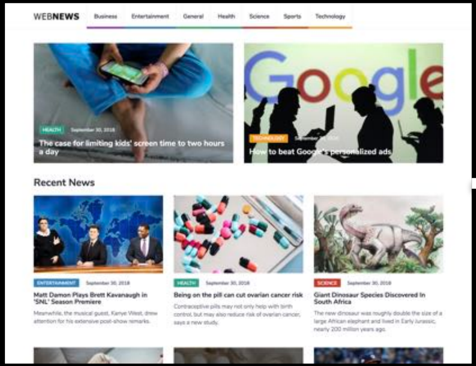
25% faster to find items

“Layout as a service”

On-going work led by Dr. Markku Laine
Team: Dr. Ai Nakajima, Dr. Niraj Ramesh, Kseniia Palin, Dr. Kashyap Todi, Samuel de Pascale

Layout as a Service (concept)

Bootstrap supported atm



Original
web page

Layout
as a service

Easy to set up

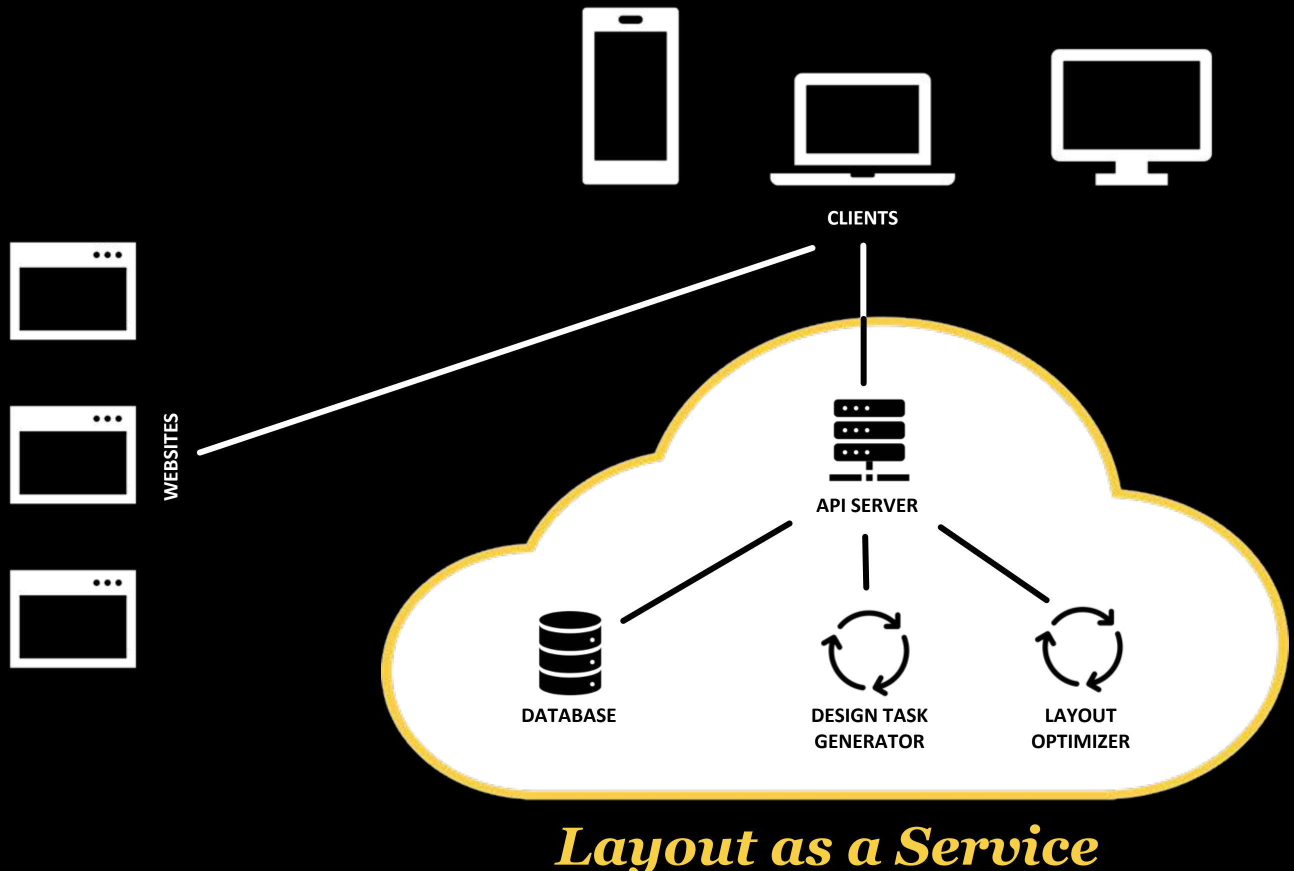
Plug and play

Individual and
group-level
adaptation

Optimized
web page

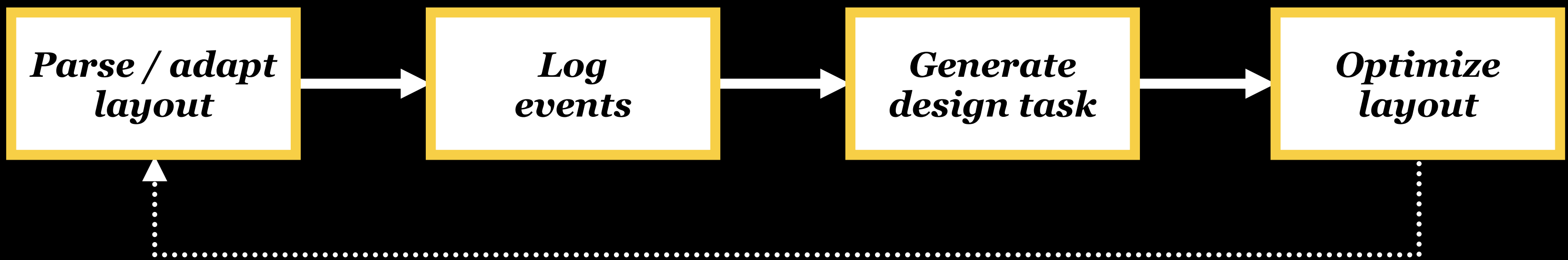
Runtime architecture

Overview



Adaptation lifecycle

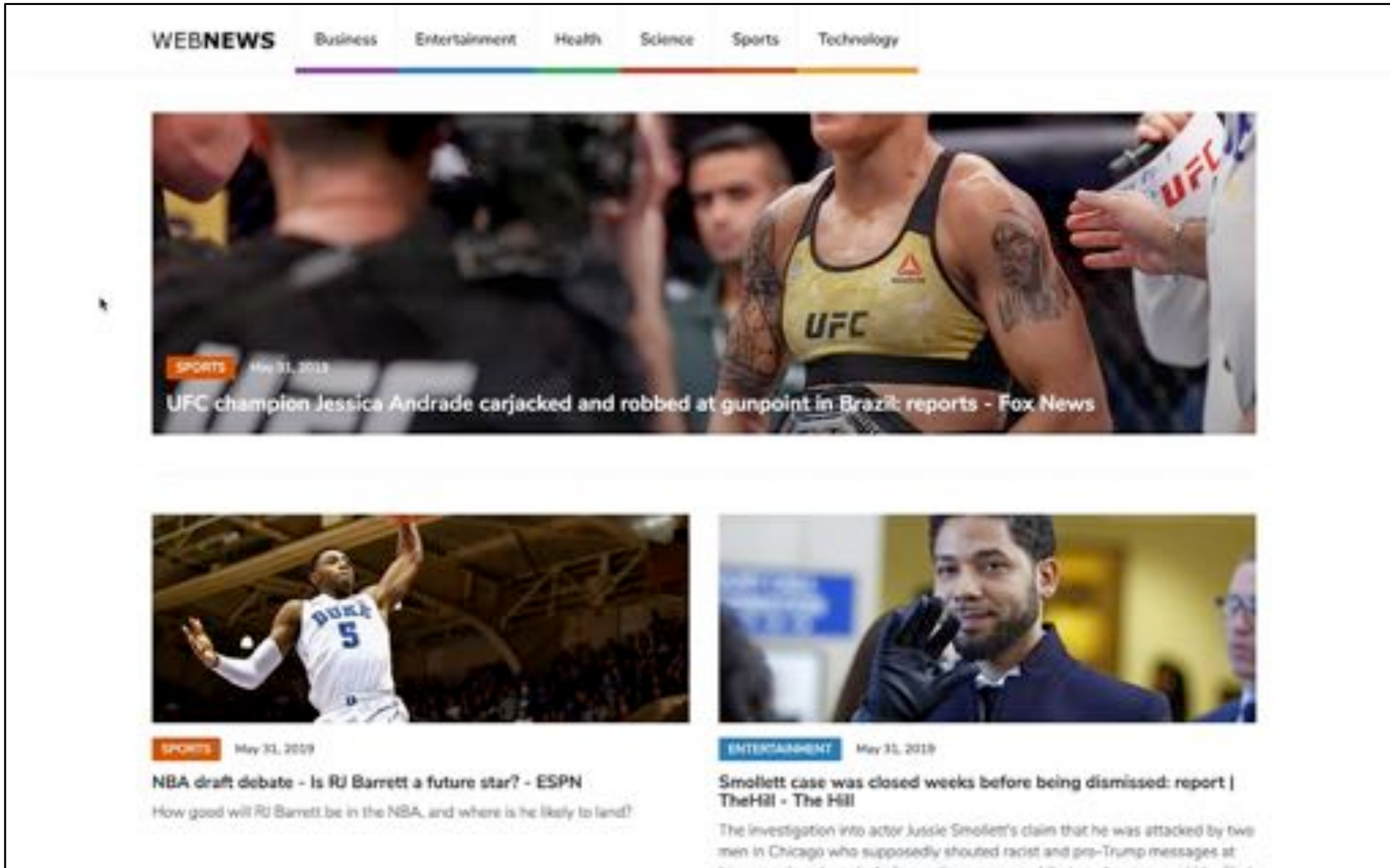
Enables self-optimizing web pages



Demo 1/2

Optimizes for: Perceptual fluency, saliency, selection time

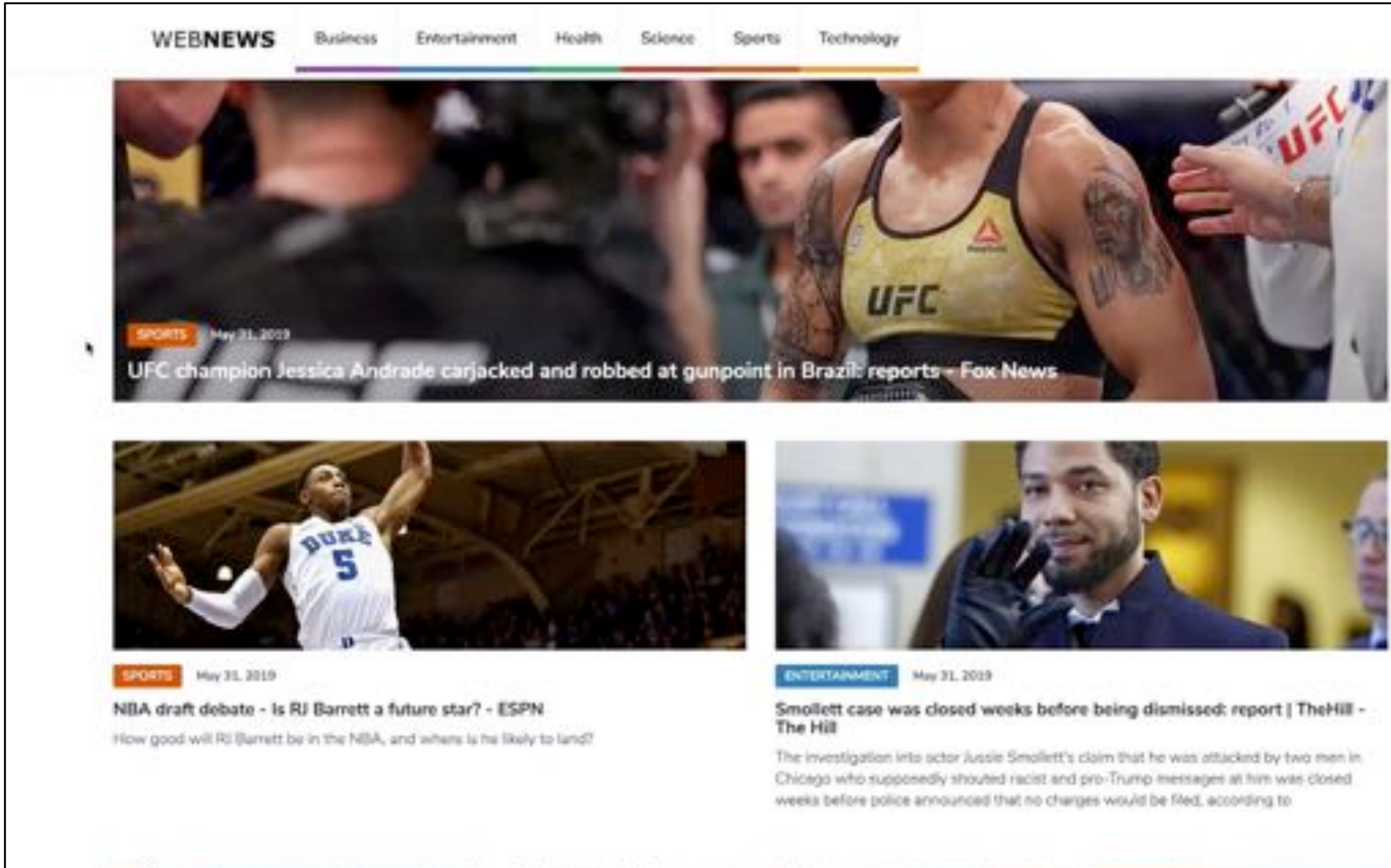
User clicking: 60% sports, 20% entertainment, 20% business



Demo 2/2

Optimizes for: Perceptual fluency, saliency, selection time

Optimized layout



Summary

A powerful way to solve HCI problems

Increasing scope of problems successfully addressed

Plenty of work remains to go beyond proof-of-concepts

Levels of intelligence in design

Level 0: Design by intuition
Level 1: Design by models
Level 2: Design by optimization
Level 3: Design by learning

Level 4: Combine Levels 0 - 3

Future: Combine model-based methods with deep learning

Benefit from their high representational power while retaining causal mechanisms that enable counterfactual decisions and controllability

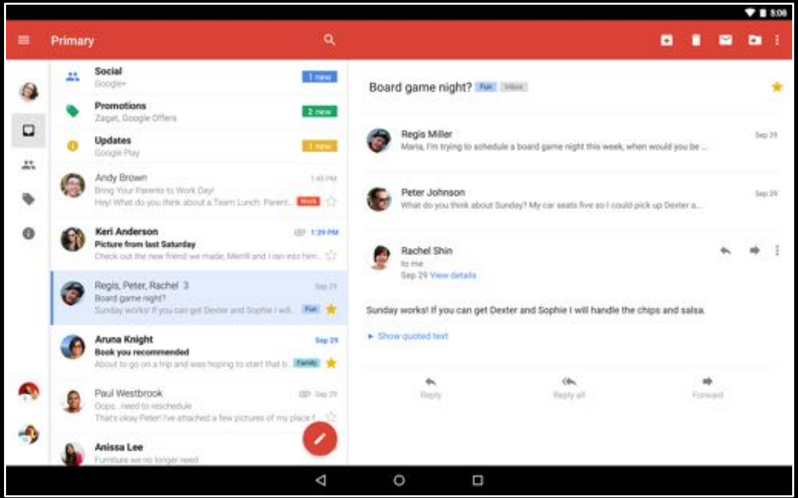
Two *very* hard challenges

The Winograd & Flores' argument

Design depends on linguistic intelligence from the user

The Dreyfus argument

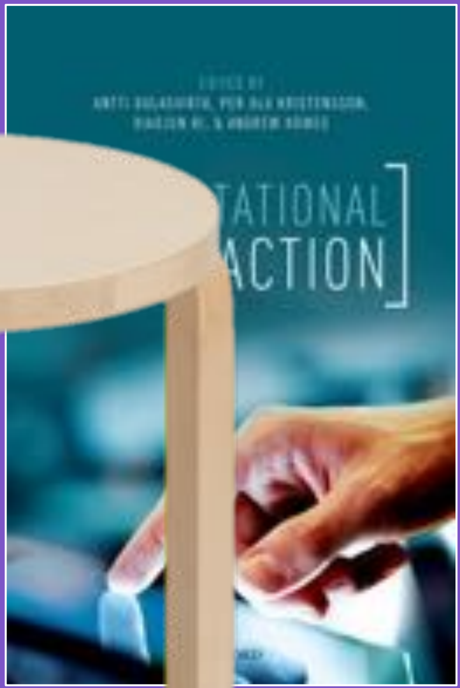
Human-like being-in-the-world and social acculturation required for real intelligence



Visuo-spatial-motor aspects of design are within reach, but social and linguistic not yet?

Aalto University

FCAI



$$\max_{d \in D} \int g\big(M(\theta, d)\big)p(\theta|o)d\theta$$

Thank you

antti.oulasvirta@aalto.fi

Finnish Center for Artificial Intelligence FCAI

Aalto University

userinterfaces.aalto.fi

Daejeon, Korea 19th

ICWE 2019

19th INTERNATIONAL CONFERENCE ON WEB ENGINEERING