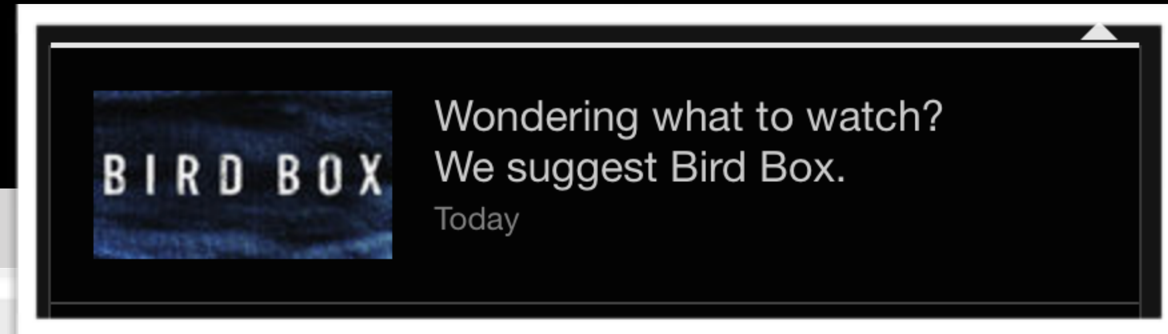
RECSYS 2019

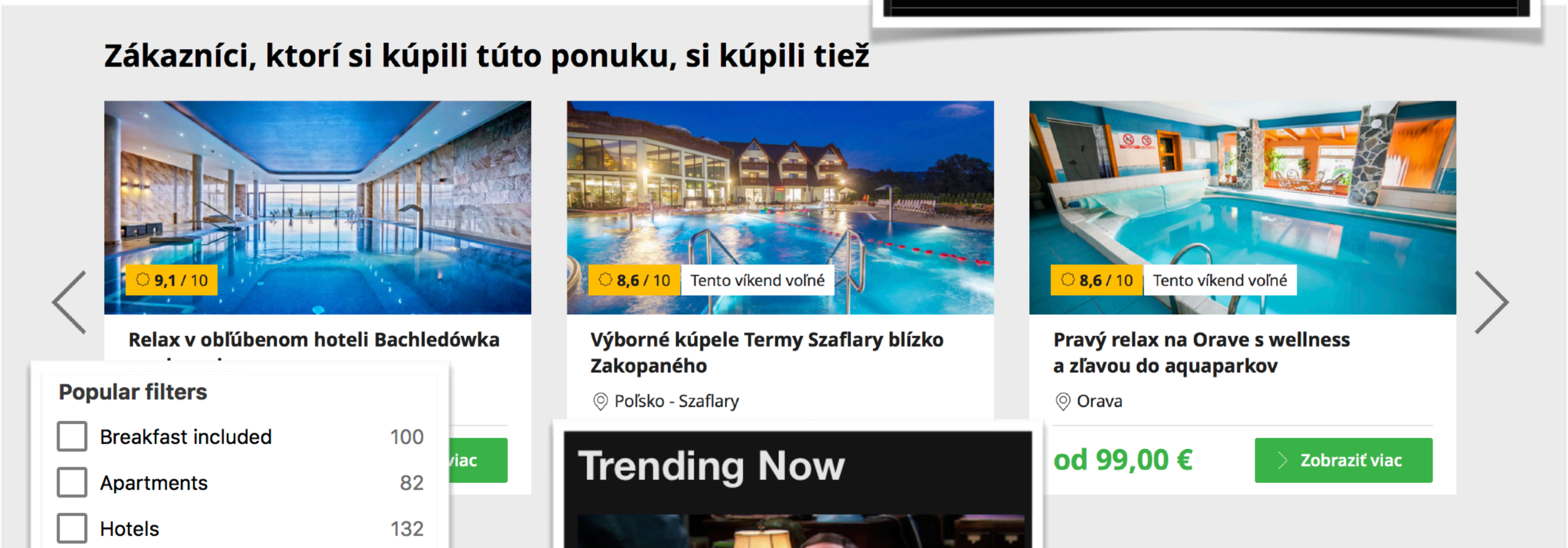
Peter Gašpar, Miroslav Rác, Michal Kompan

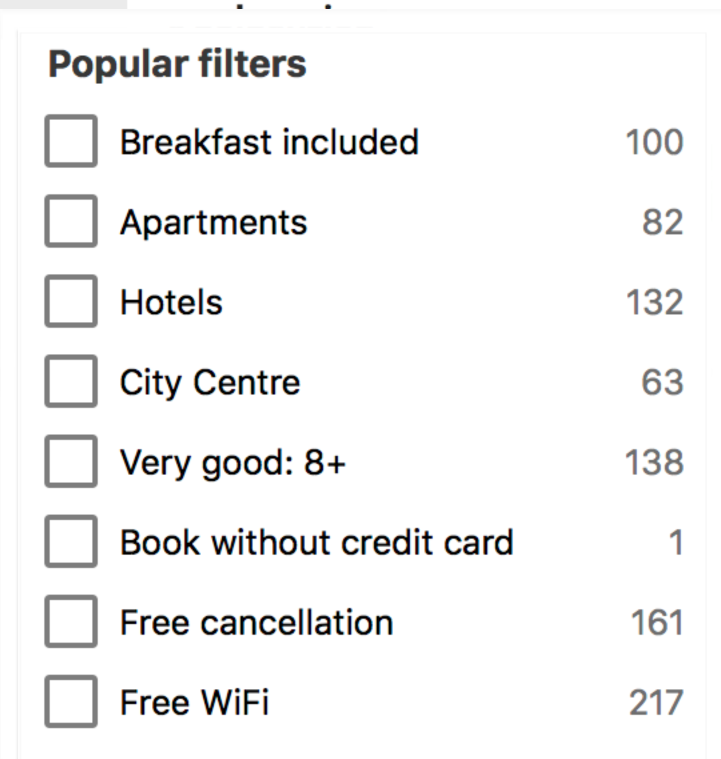
2.10.2019

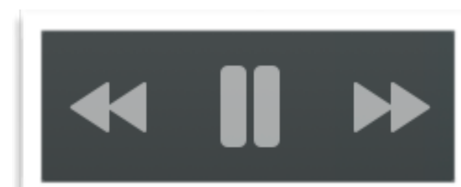
Odporúčanie

Čo je to odporúčanie?









Odporúčanie

Základný úvod

Aké základné odporúčanie poznáme?

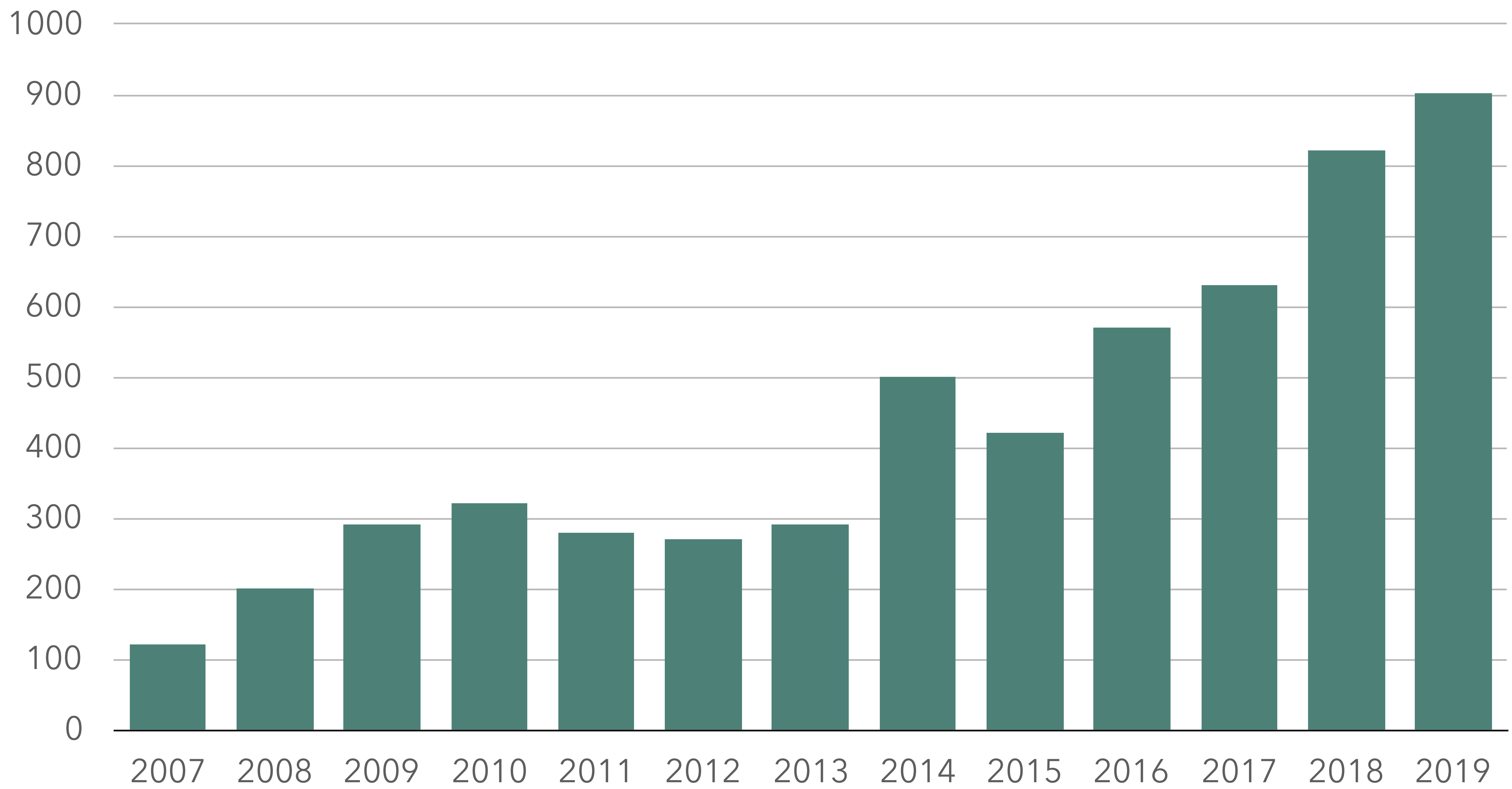
- Kolaboratívne,
- obsahové,
- hybridné.

Skúmame rôzne **metriky**:

presnosť, pokrytie, zotrvanie zákazníka, spokojnosť, …

RecSys 2019

Počet účastníkov

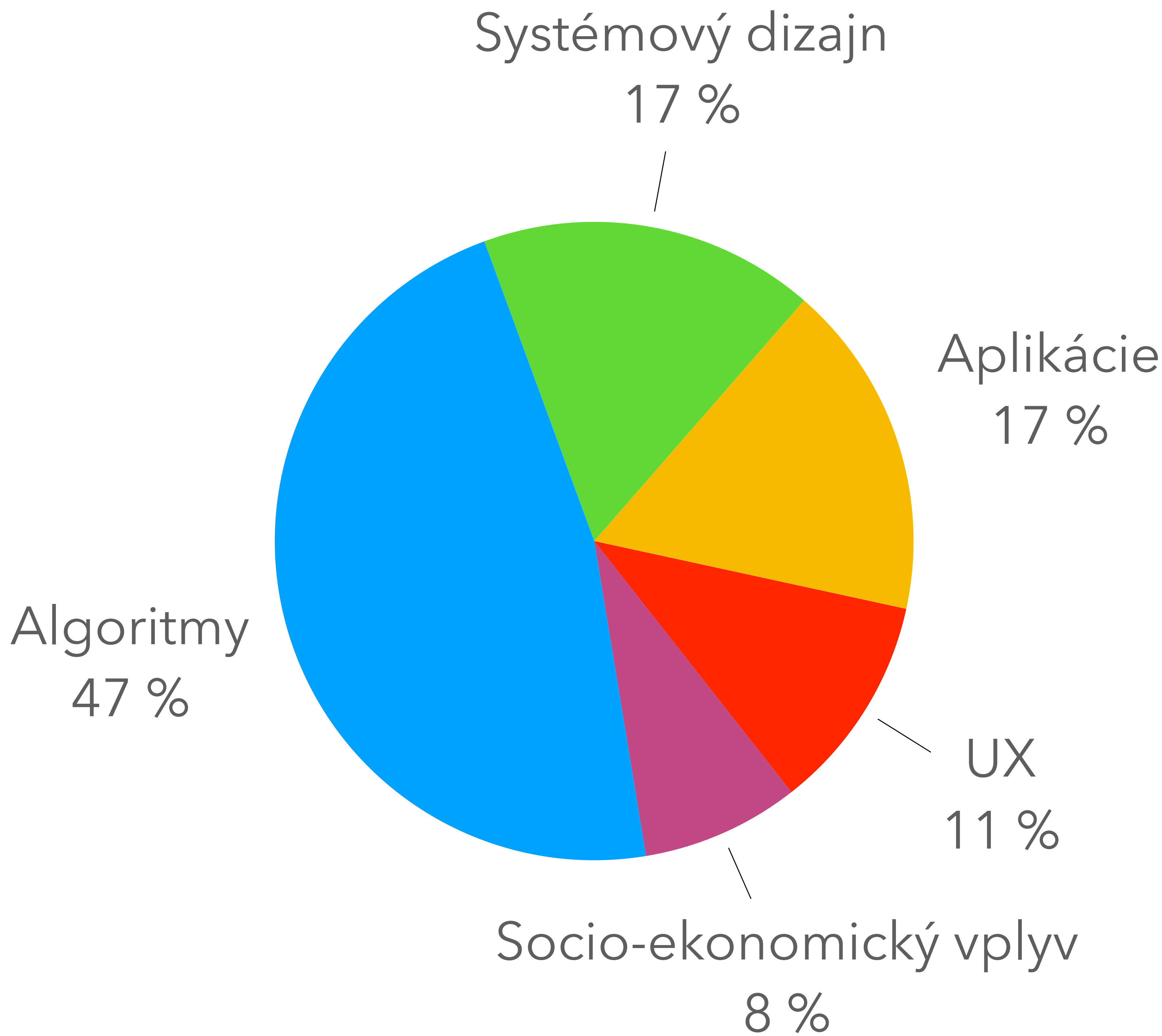


datalys

RECSYS 2019

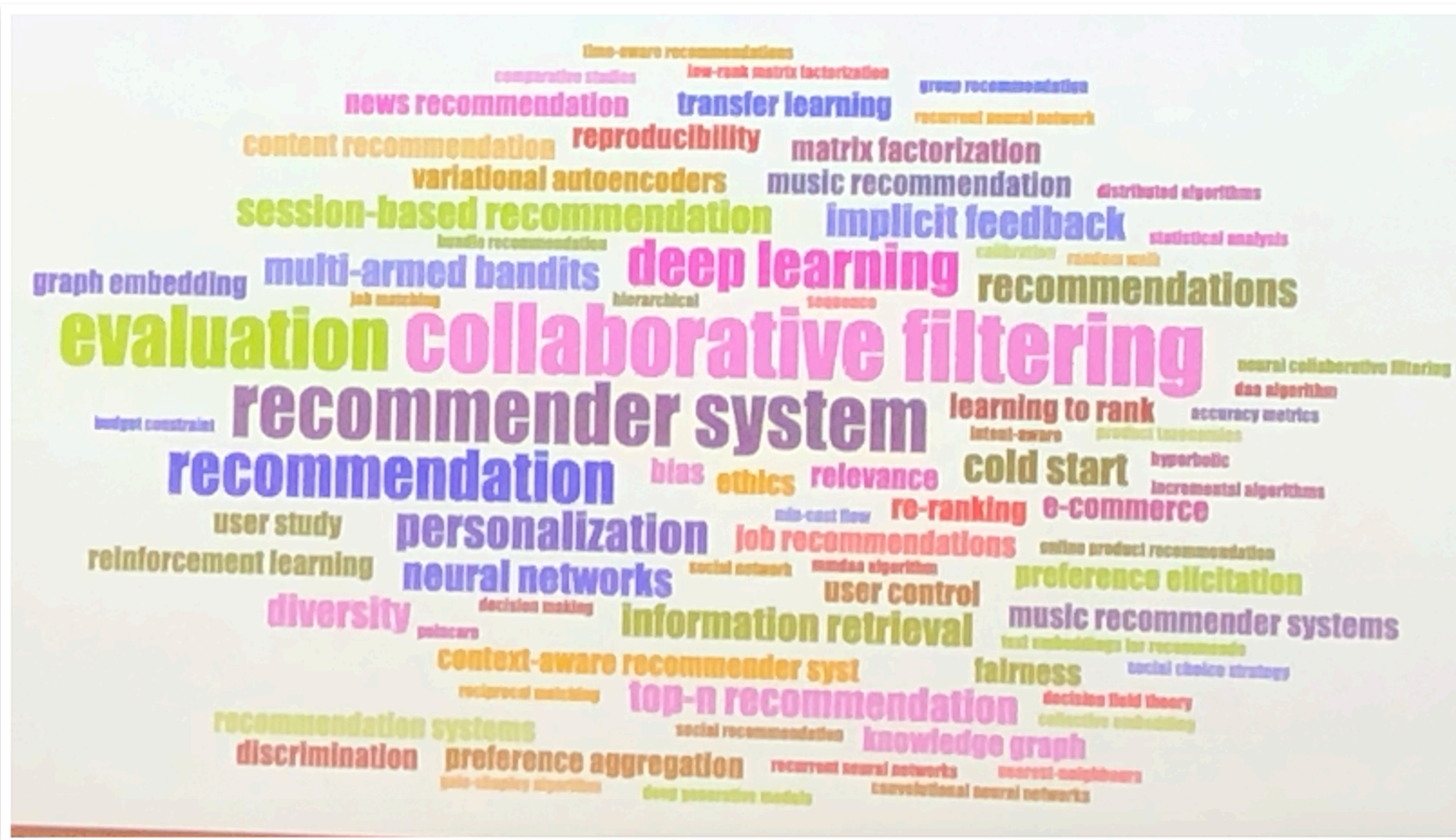
RecSys 2019

Program



RecSys 2019

Kľúčové slová



KEYNOTES

RECSYS 2019

datalys

Keynote 1: Rude Awakenings from Behaviourist Dreams. Methodological Integrity and the GDPR

Mireille Hildebrandt (Vrije Universiteit Brussels, Belgium)

Sú skutočne **personalizované** reklamy účinné?

Často postačuje primerané využitie kontextuálnych informácií namiesto obrovského množstva dát, ktoré dnes zbierajú Facebook a Google.

Keynote 1: Rude Awakenings from Behaviourist Dreams. Methodological Integrity and the GDPR

Mireille Hildebrandt (Vrije Universiteit Brussels, Belgium)

MOBILE

When Procter & Gamble Cut $200 Million in Digital Ad Spend, It Increased Its Reach 10%

Unilever is also reevaluating its budget

By Lauren Johnson | March 1, 2018 PREMIUM

https://www.adweek.com/brand-marketing/when-procter-gamble-cut-200-million-in-digital-ad-spend-its-marketing-became-10-more-effective/

Keynote 1: Rude Awakenings from Behaviourist Dreams. Methodological Integrity and the GDPR

Mireille Hildebrandt (Vrije Universiteit Brussels, Belgium)

After GDPR, The New York Times cut off ad exchanges in Europe — and kept growing ad revenue

JANUARY 16, 2019 by Jessica Davies

When the General Data Protection Regulation arrived last year, The New York Times didn't take any chances.

The publisher blocked all open-exchange ad buying on its European pages, followed swiftly by behavioral targeting. Instead, NYT International focused on contextual and geographical targeting for programmatic guaranteed and private marketplace deals and has not seen ad revenues drop as a result, according to Jean-Christophe Demarta, svp for global advertising at New York Times International.

https://digiday.com/media/gumgumtest-new-york-times-gdpr-cut-off-ad-exchanges-europe-ad-revenue/

Keynote 1: Rude Awakenings from Behaviourist Dreams. Methodological Integrity and the GDPR

Mireille Hildebrandt (Vrije Universiteit Brussels, Belgium)

- Who is paying for recommender systems? Who is profiting from them?
  - Developers?
  - Big tech? Platforms? Publishers? Advertisers?
  - Users?

- What is the goal:
  - Mining preferences? Catering to end-users for their own sake?
  - Increasing ad revenue or sales or influencing voting behaviours?
  - **Attention grabbing, holding, hooking → promoting addiction**

Keynote 1: Rude Awakenings from Behaviourist Dreams. Methodological Integrity and the GDPR

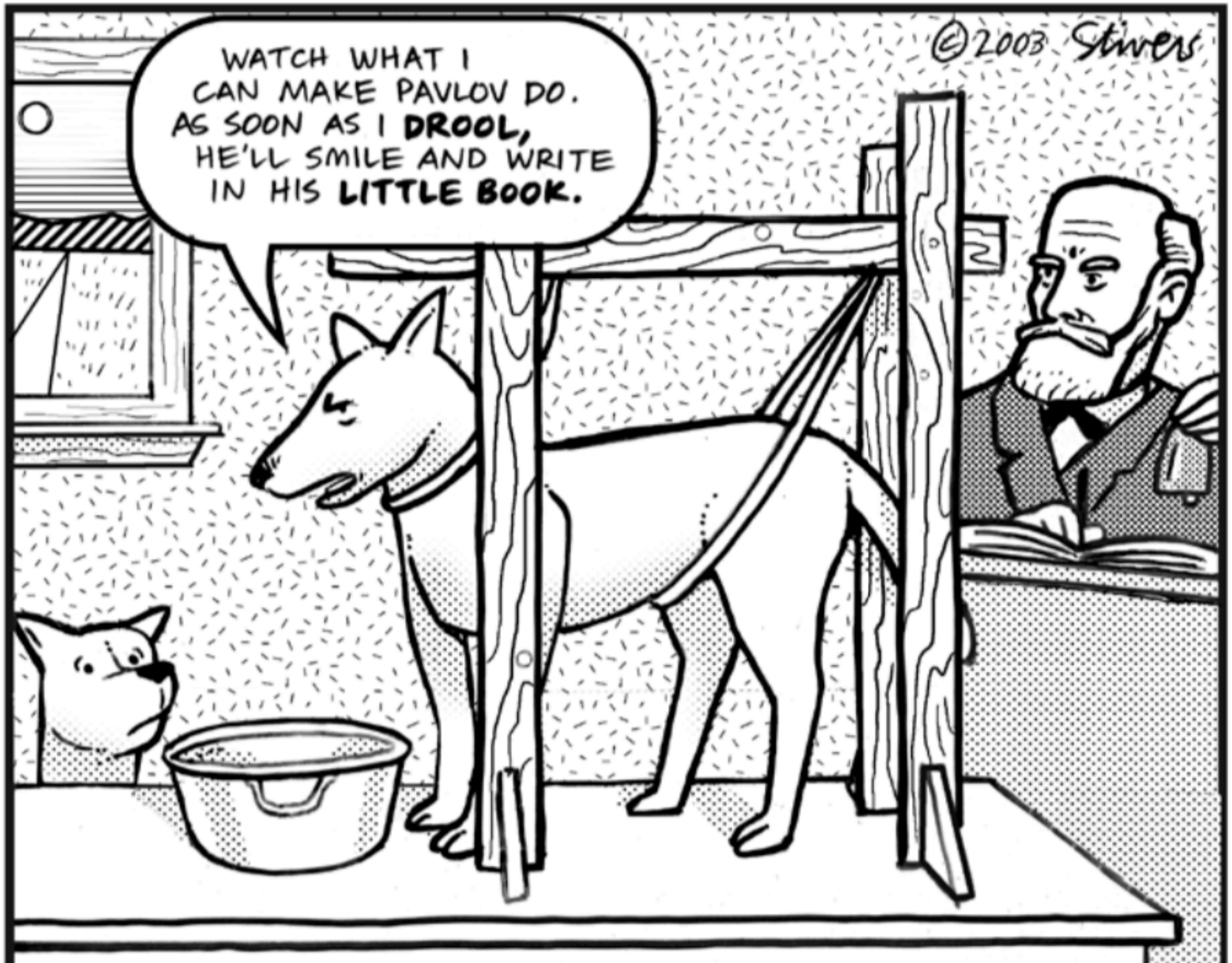
Mireille Hildebrandt (Vrije Universiteit Brussels, Belgium)

- Data-driven recsys have trained their algorithm (learner) on behavioural data:
  - Online impression, click and conversion behaviours (RTB, AB testing)
  - Purchasing, watching, location (mobility) behaviours (LBS, Amazon, Netflix)
  - Image and voice recognition (Echo, Alexa)
  - All and any (sentiment analysis)

- Behavioural data =
  - Machine readable data (surf behaviour, sensor data, mobility data, etc.)
  - Data is a trace, representation, or imprint of something else
  - Based on the methodological individualism of behaviourism
  - Shares assumptions with rational choice theory and behavioural economics

Keynote 1: Rude Awakenings from Behaviourist Dreams. Methodological Integrity and the GDPR

Mireille Hildebrandt (Vrije Universiteit Brussels, Belgium)



Keynote 1: Rude Awakenings from Behaviourist Dreams. Methodological Integrity and the GDPR

Mireille Hildebrandt (Vrije Universiteit Brussels, Belgium)

- Personal data processing:
  - Any data relating to an identifiable person (very broad concept)
  - Data subject (natural person)
  - Data processing (from collection to storage to whatever)
  - Behavioural data that can single out = personal data

  - Goal of the GDPR is to enable processing within the Union
    - Based on equivalent protection throughout the Union
    - Mix of public enforcement and private law liability

Keynote 1: Rude Awakenings from Behaviourist Dreams. Methodological Integrity and the GDPR

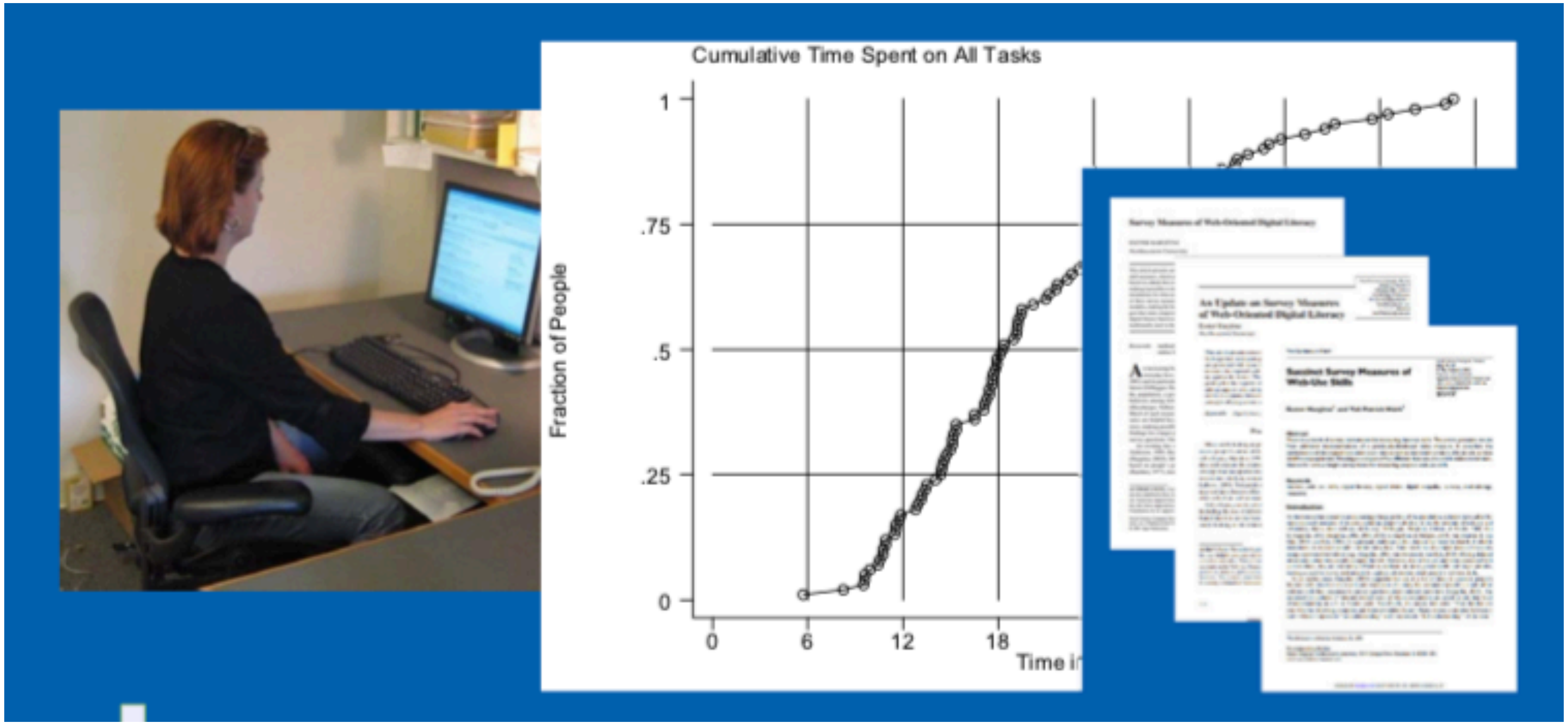
Mireille Hildebrandt (Vrije Universiteit Brussels, Belgium)

Keď príde na trh nový liek, musí prejsť zložitým procesom overovania a registrácie.

Prečo sa to neuplatňuje aj v doméne strojového učenia?

Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)



Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)

Hypotézy:

1. Všetci mladí ľudia sú digitálne zdatní.

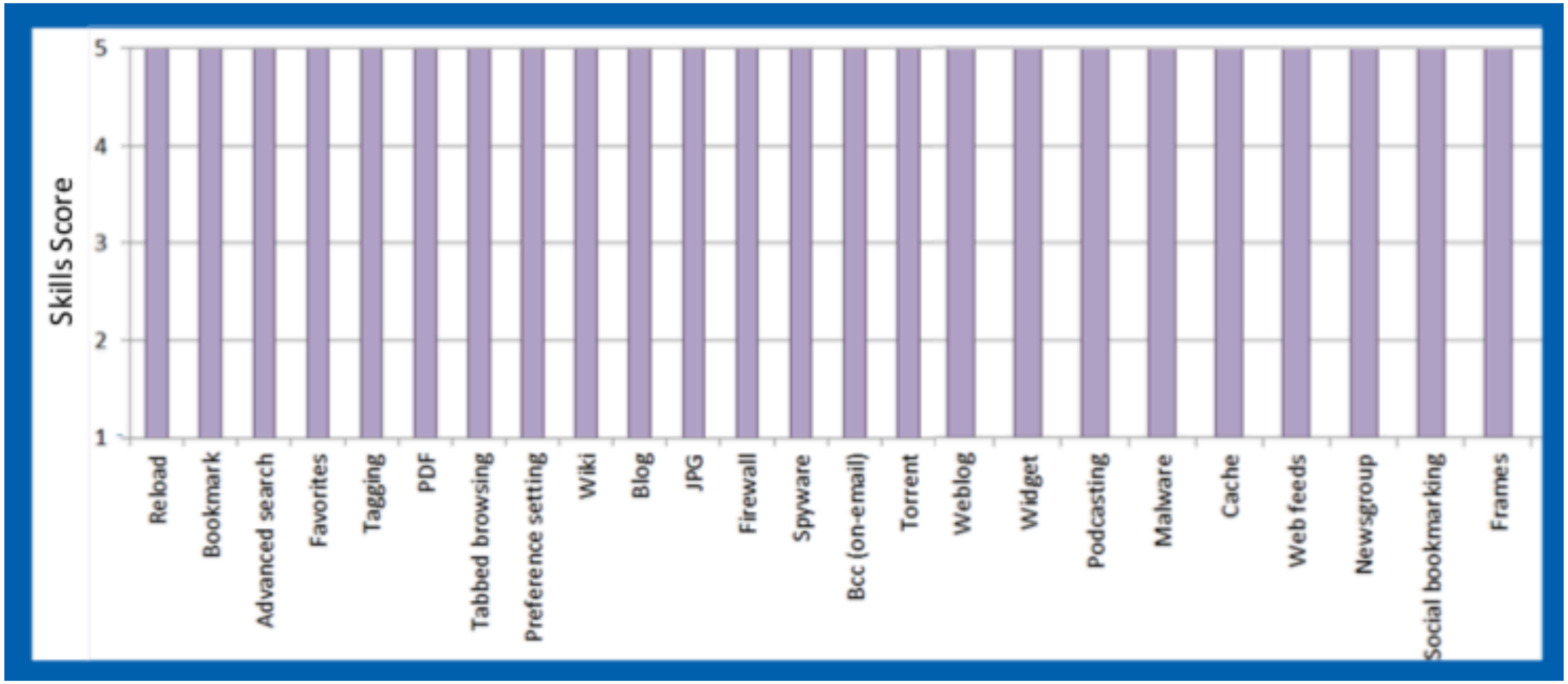
2. Mladí ľudia sú zdatnejší ako starší ľudia.

Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)

1. Všetci mladí ľudia sú digitálne zdatní.

Očakávania:

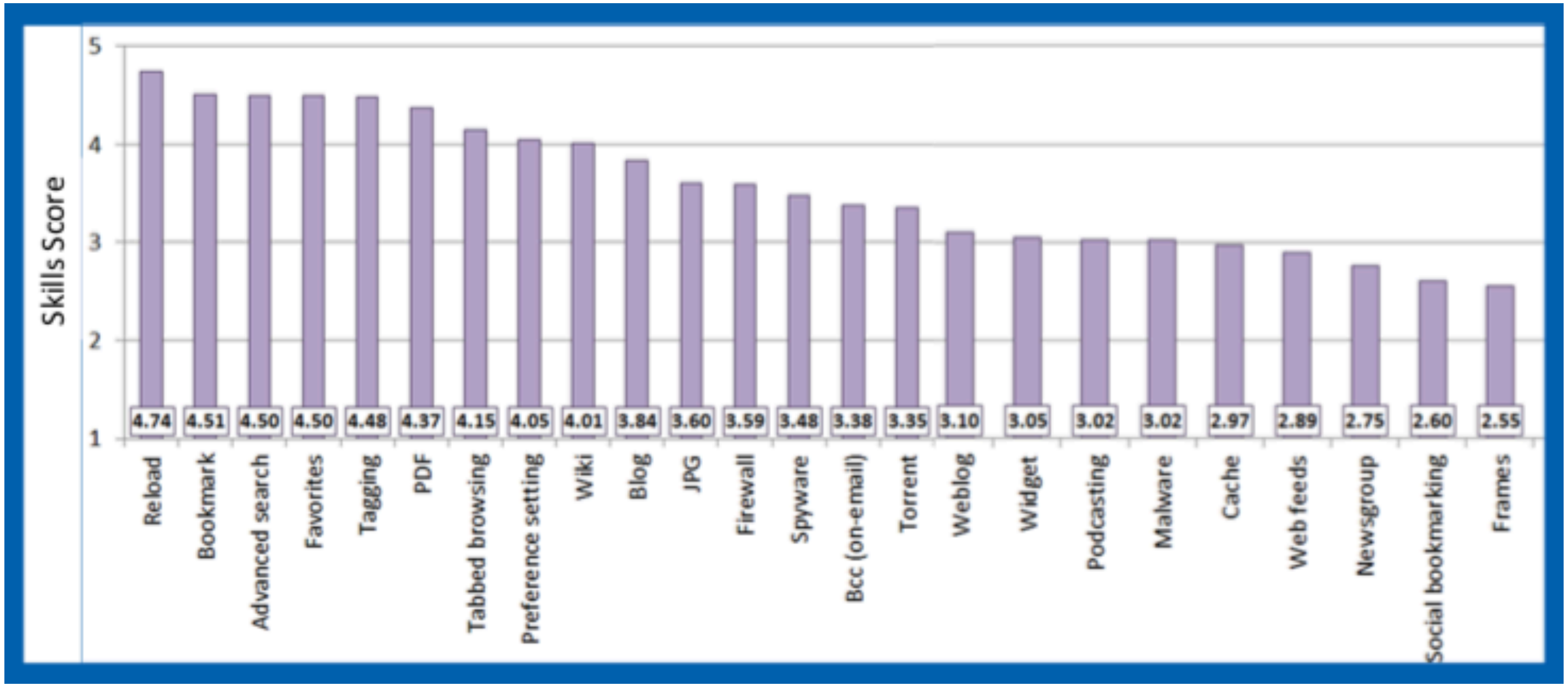


Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)

1. Všetci mladí ľudia sú digitálne zdatní.

Realita:



Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)

Which of the following sites is most likely to be the Web site of a bank called “Bankomat”?

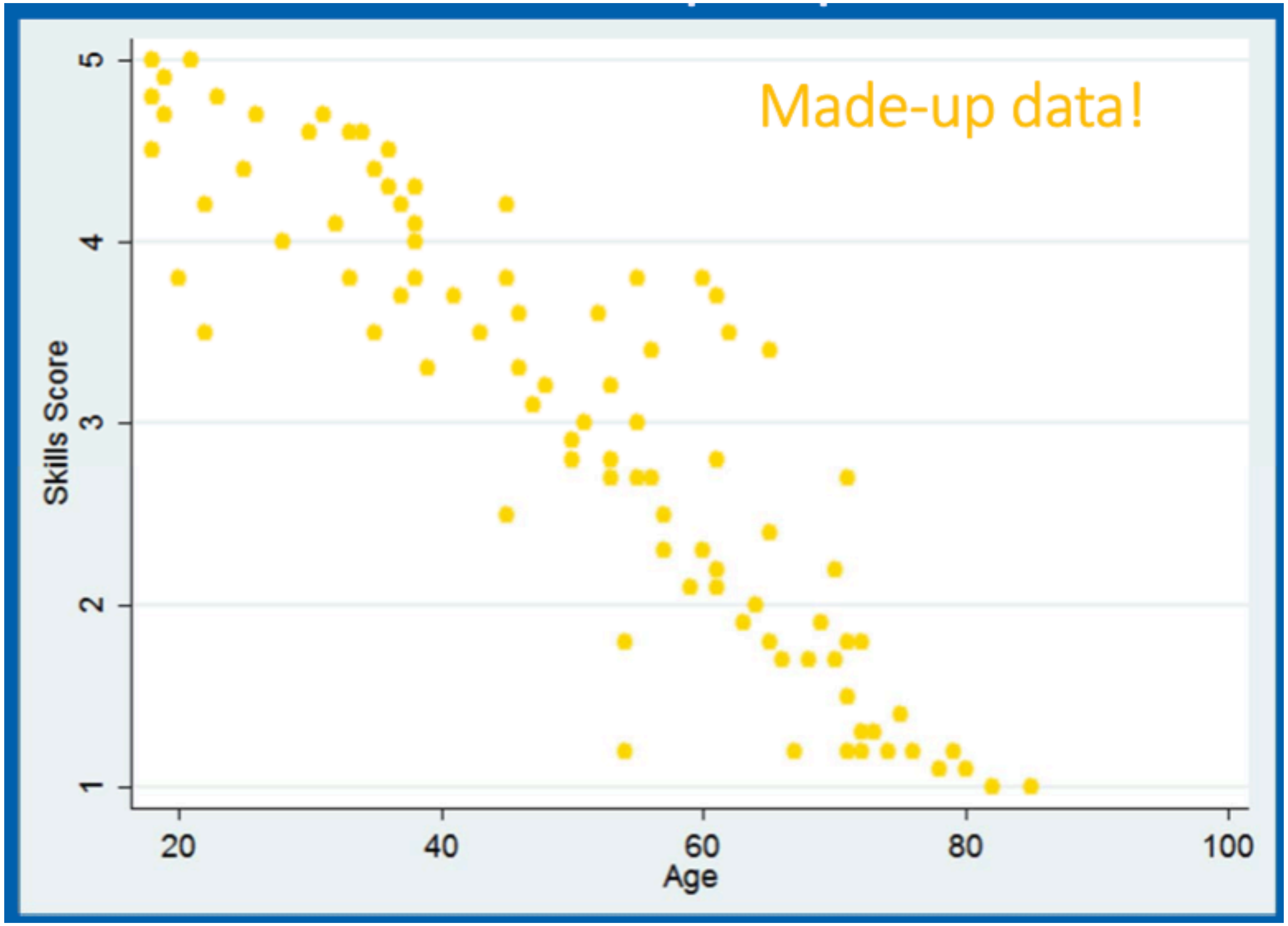
- [x] www3.da-us.bankomat.com/cgi-bin/citifili/portal/I/I.do — 11%
- [ ] www.bankomat.ve/rify.com — 35%
- [ ] www.us.bankomat.businessportal.ru/bankomat/index.php — 50%
- [ ] www.krezmin.nu/bankomat.com — 4%

Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)

2. Mladí ľudia sú zdatnejší ako starší ľudia.

Očakávania:

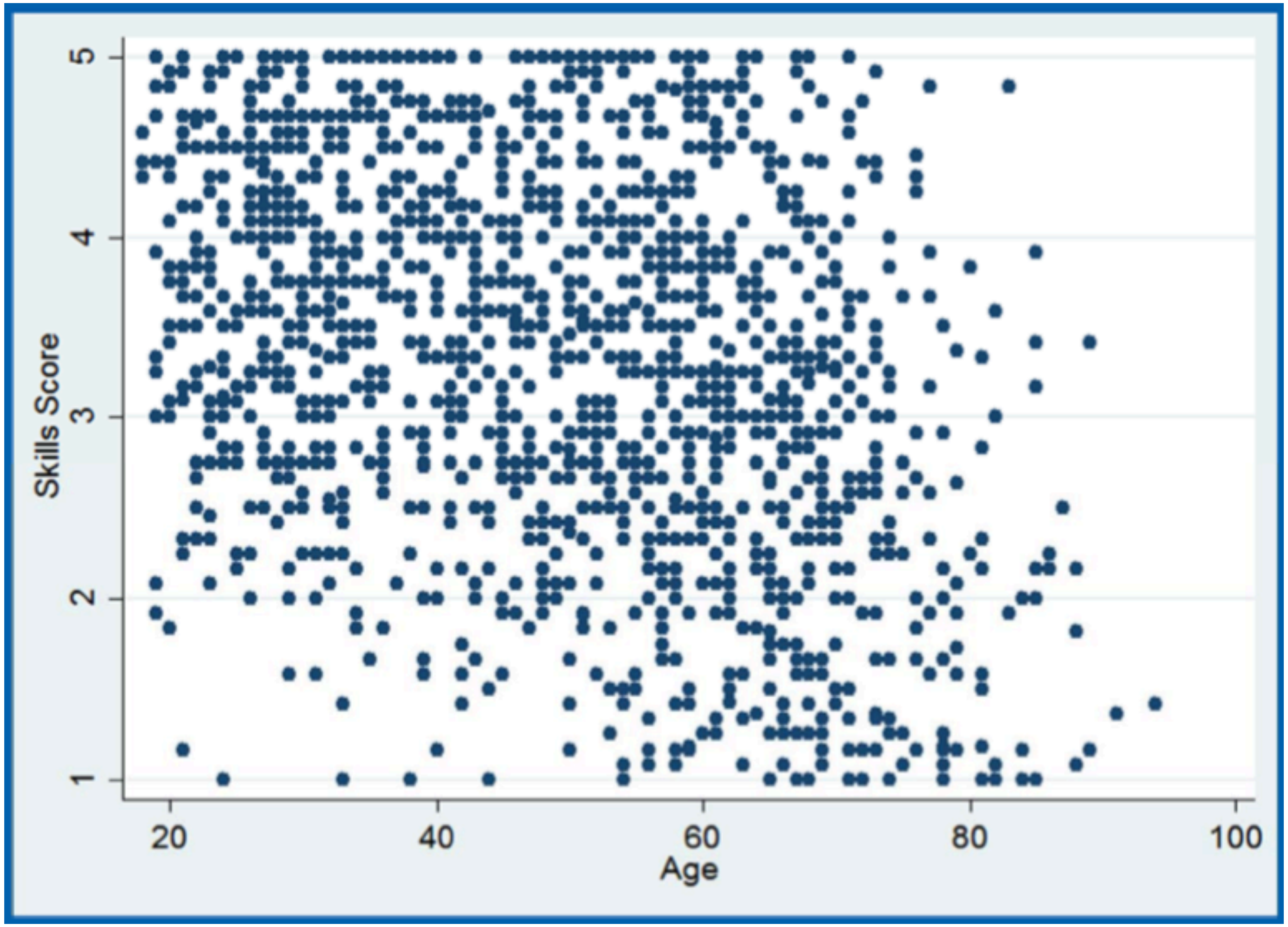


Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)

2. Mladí ľudia sú zdatnejší ako starší ľudia.

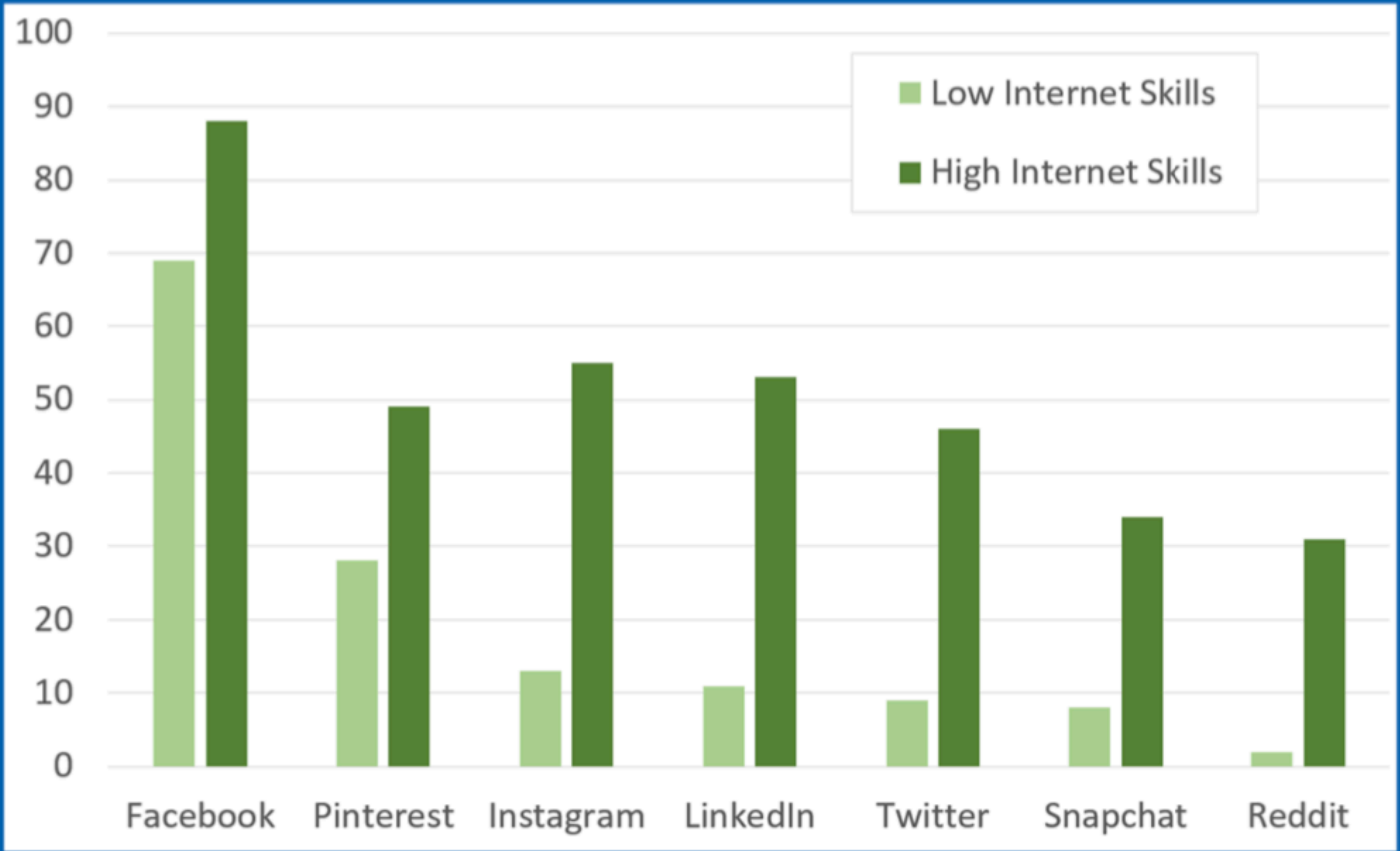
Realita:



Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

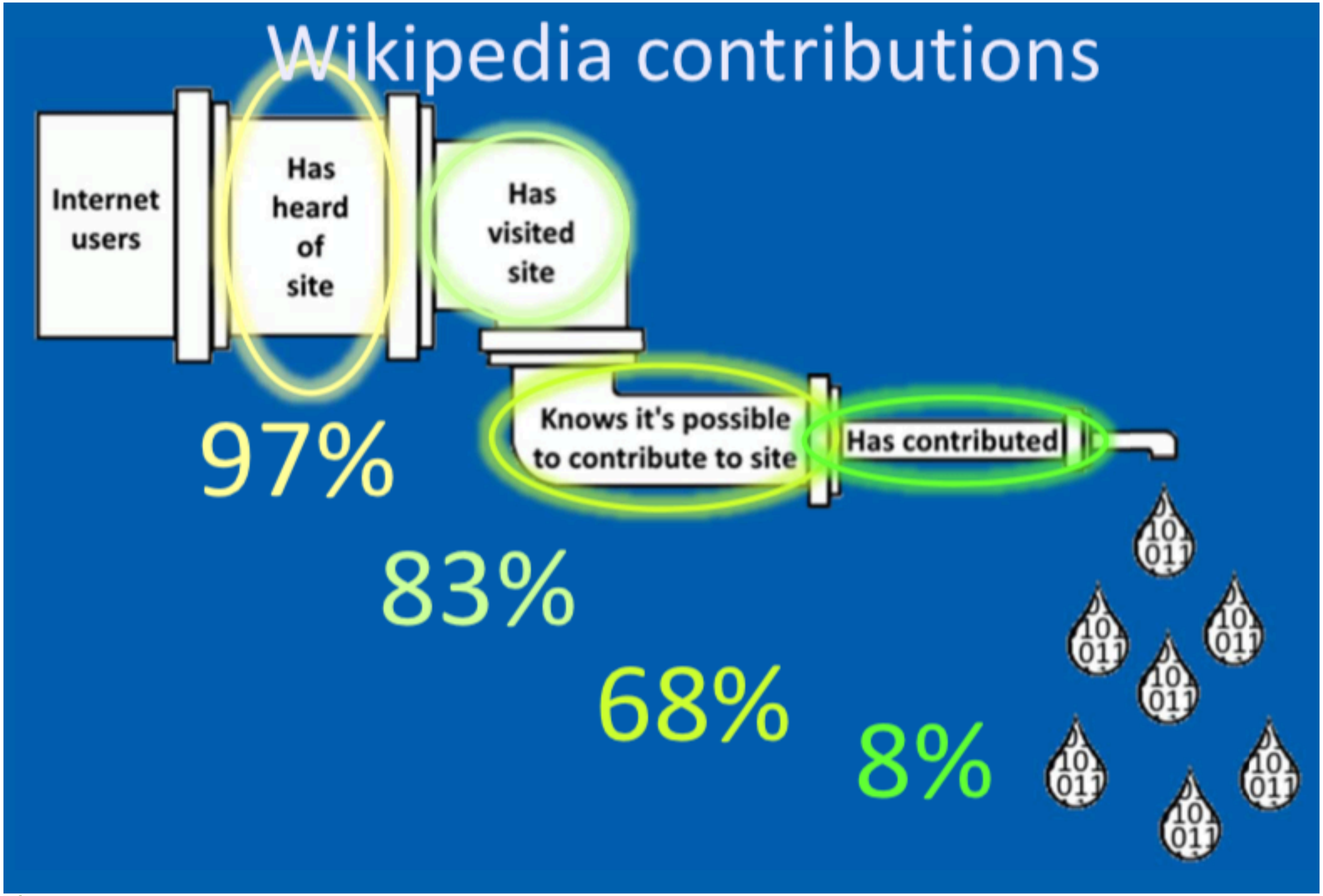
Eszter Hargittai (University of Zurich, Switzerland)

Use of social network sites by Internet skills, US adults



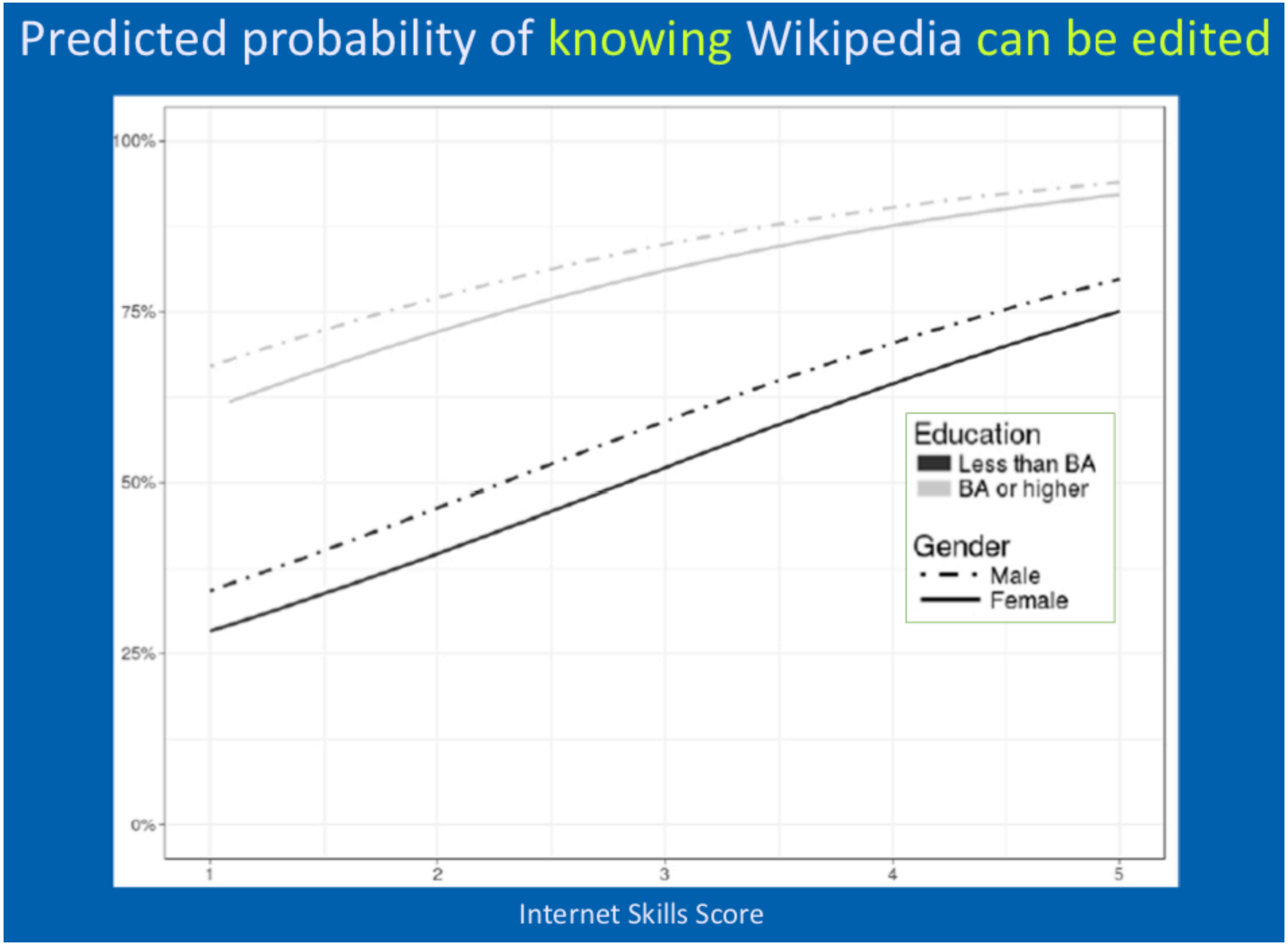
Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)



Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)



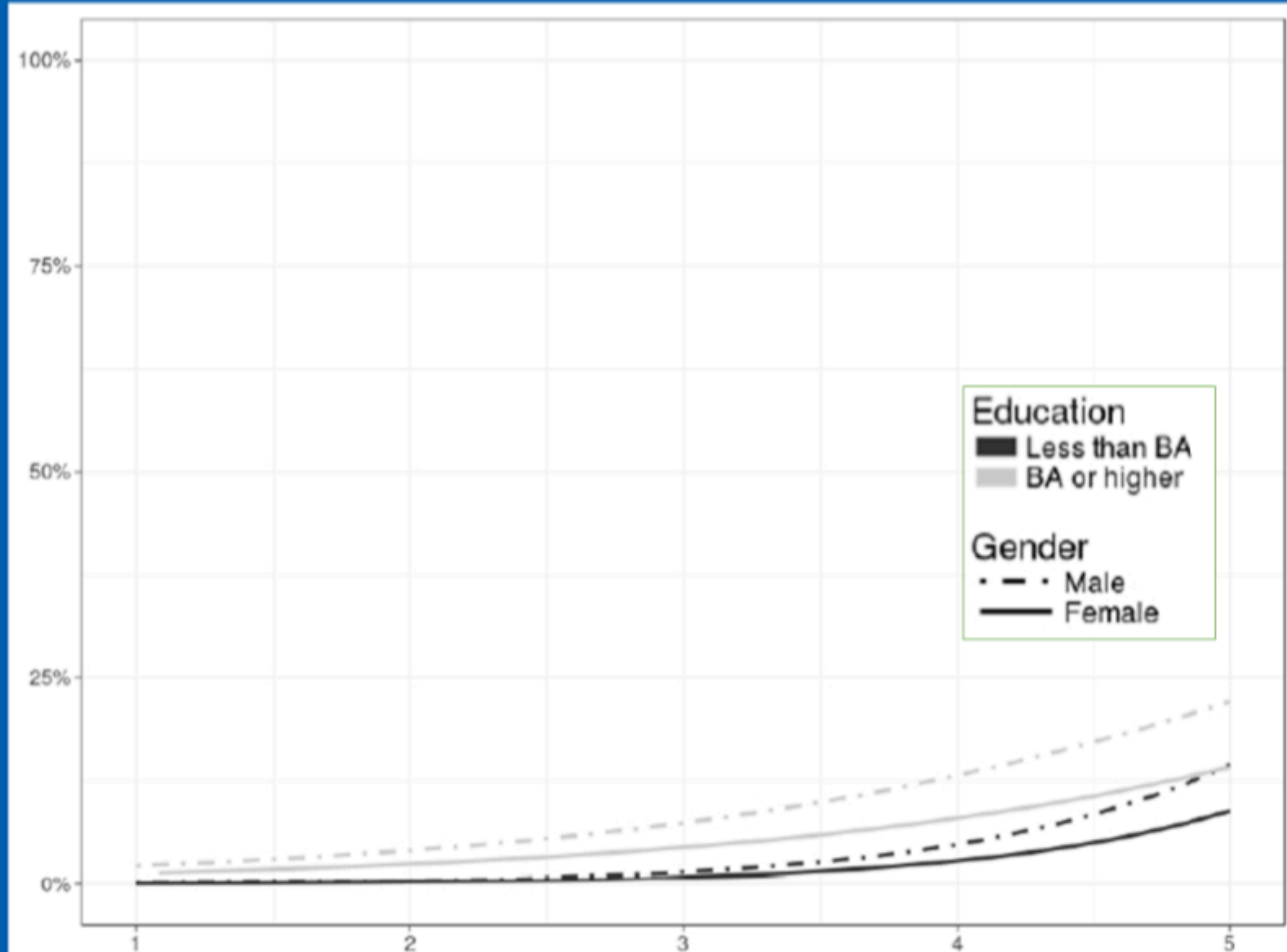
Keynote 2: Inequality in Online Participation and Why it Matters for Recommendation Systems Research

Eszter Hargittai (University of Zurich, Switzerland)

22. Have you ever edited a Wikipedia page by fixing a mistake or adding new material?

- No
- Yes

Predicted probability of contributing to Wikipedia



Internet Skills Score

PAPER SESSIONS

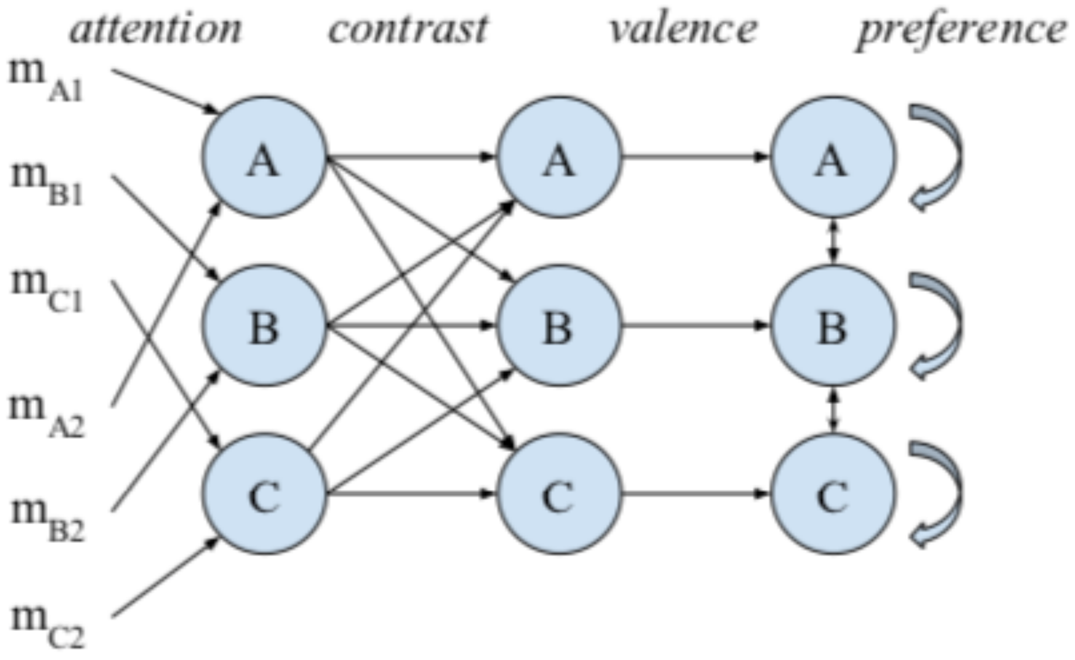
RECSYS 2019

datalys

From preference into decision making: modeling user interactions in recommender systems

Qian Zhao et al. | https://dl.acm.org/citation.cfm?id=3347065

Modelovanie správania používateľa: prehliadanie, akcia, neakcia.



The connectionist network representation of DFT (Decision Field Theory).

A deep learning system for predicting size and fit in fashion e-commerce

Abdul-Saboor Sheikh (Zalando) et al. | https://dl.acm.org/citation.cfm?id=3347006

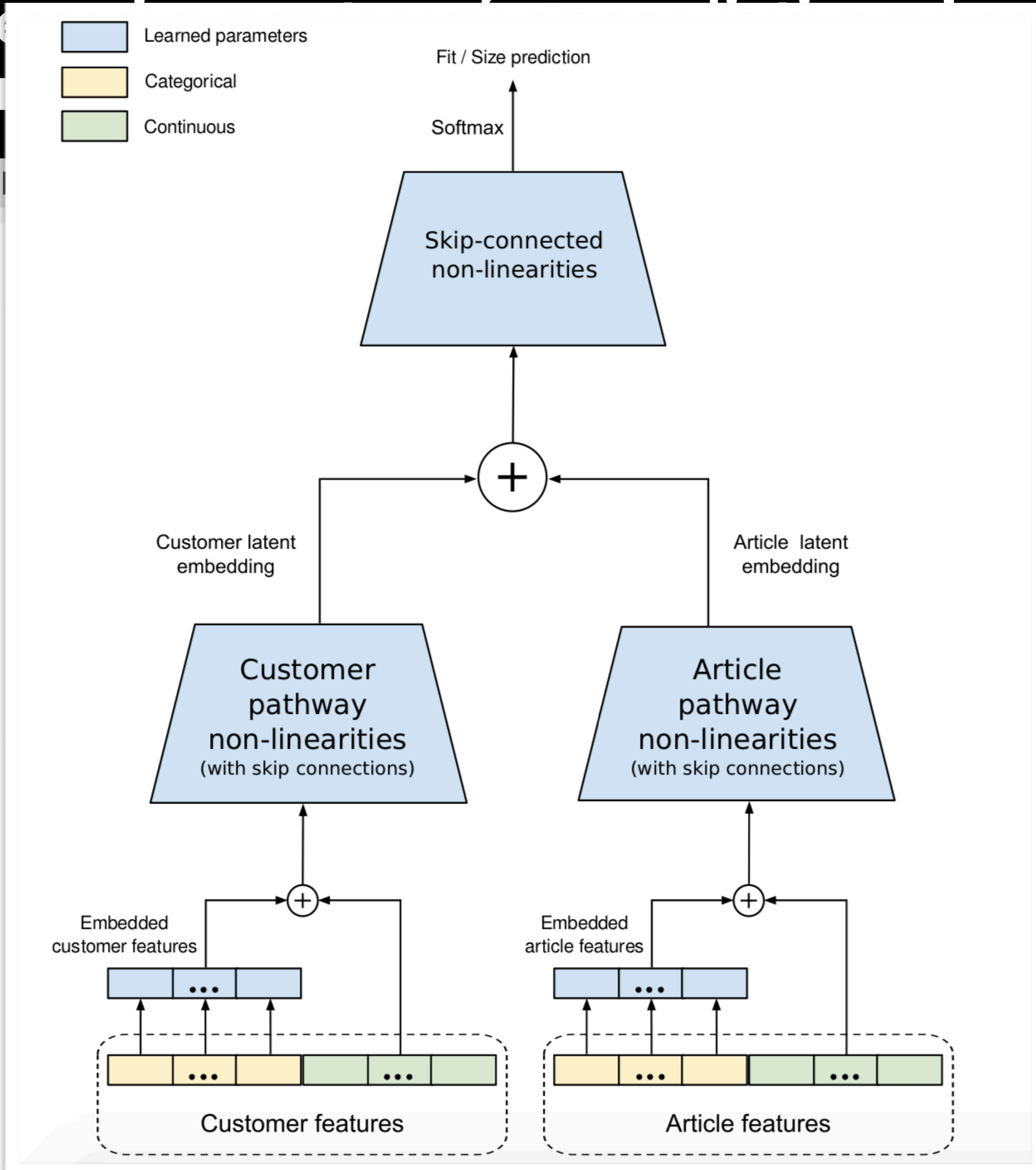
Personalizovaná predpoveď veľkosti oblečenia

- využíva rôzne parametre pre naučenie abstraktnej reprezentácie veľkostí na základe interakcií medzi používateľmi a produktami,
- navyše využili vlastnosti používateľov a produktov.

Features/Dataset	ModCloth	RentTheRunWay
Article	category*, quality, item id*, size	category*, rating, rented for*, item id*, size
Customer	shoe width*, shoe size, waist, bust, cup size, bra size, hips, height, user id*	age, body type*, bust size+, height, weight, user id*

A deep l… and fit in fashio…

Abdul-Saboor Sheil…



datalys

RECSYS 2019

Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches

Maurizio Ferrari Dacrema et al. | https://dl.acm.org/citation.cfm?id=3347058

V doméne **odporúčania s využitím hlbokého učenia** sa autori zamerali na dve hlavné oblasti:

1. **Reprodukovateľnosť výsledkov.** Je výskum zopakovateľný na základe článku?

2. **Prínos.** Dokážu prezentované prístupy prekonať vyladené základné metódy?

datalys

RECSYS 2019

Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches

Maurizio Ferrari Dacrema et al. | https://dl.acm.org/citation.cfm?id=3347058

Reprodukovateľnosť výsledkov:

- dostupná funkčná verzia zdrojových kódov,
- aspoň 1 dataset z článku je dostupný a možné ho rovnako rozdeliť na train/test.

Conference	Rep. ratio	Reproducible
KDD	3/4 (75%)	[17], [23], [48]
RecSys	1/7 (14%)	[53]
SIGIR	1/3 (30%)	[10]
WWW	2/4 (50%)	[14], [24]
Total	7/18 (39%)	

Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches

Maurizio Ferrari Dacrema et al. | https://dl.acm.org/citation.cfm?id=3347058

Prínos:

- porovnanie so základnými baselinami, hyper-parameter tunning.

Conference	Rep. ratio	Reproducible
KDD	3/4 (75%)	[17], [23], [48]
RecSys	1/7 (14%)	[53]
SIGIR	1/3 (30%)	[10]
WWW	2/4 (50%)	[14], [24]
Total	7/18 (39%)	

Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches

Maurizio Ferrari Dacrema et al. | https://dl.acm.org/citation.cfm?id=3347058

Table 2: Experimental results for the CMN method using the metrics and cutoffs reported in the original paper. Numbers are printed in bold when they correspond to the best result or when a baseline outperformed CMN.

	CiteULike-a			
	HR@5	NDCG@5	HR@10	NDCG@10
TopPopular	0.1803	0.1220	0.2783	0.1535
UserKNN	**0.8213**	**0.7033**	**0.8935**	**0.7268**
ItemKNN	**0.8116**	**0.6939**	0.8878	**0.7187**
$P^3\alpha$	**0.8202**	**0.7061**	0.8901	**0.7289**
$RP^3\beta$	**0.8226**	**0.7114**	**0.8941**	**0.7347**
CMN	0.8069	0.6666	0.8910	0.6942
	Pinterest			
	HR@5	NDCG@5	HR@10	NDCG@10
TopPopular	0.1668	0.1066	0.2745	0.1411
UserKNN	**0.6886**	**0.4936**	0.8527	**0.5470**
ItemKNN	**0.6966**	**0.4994**	**0.8647**	**0.5542**
$P^3\alpha$	0.6871	**0.4935**	0.8449	**0.5450**
$RP^3\beta$	**0.7018**	**0.5041**	**0.8644**	**0.5571**
CMN	0.6872	0.4883	0.8549	0.5430
	Epinions			
	HR@5	NDCG@5	HR@10	NDCG@10
TopPopular	**0.5429**	**0.4153**	**0.6644**	**0.4547**
UserKNN	0.3506	0.2983	0.3922	0.3117
ItemKNN	0.3821	0.3165	0.4372	0.3343
$P^3\alpha$	0.3510	0.2989	0.3891	0.3112
$RP^3\beta$	0.3511	0.2980	0.3892	0.3103
CMN	0.4195	0.3346	0.4953	0.3592

Table 5: Experimental results for CDL on the dense *CiteULike-a* dataset.

	REC@50	REC@100	REC@300
TopPopular	0.0038	0.0073	0.0258
UserKNN	**0.0685**	0.1028	0.1710
ItemKNN	**0.0846**	**0.1213**	0.1861
$P^3\alpha$	**0.0718**	**0.1079**	0.1777
$RP^3\beta$	**0.0800**	**0.1167**	0.1815
ItemKNN-CBF	**0.2135**	**0.3038**	**0.4707**
ItemKNN-CFCBF	**0.1945**	**0.2896**	**0.4620**
CDL	0.0543	0.1035	0.2627

Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches

Maurizio Ferrari Dacrema et al. | https://dl.acm.org/citation.cfm?id=3347058

Dagmar Monett @dmonett · 27. 7.

A "worrying analysis":

"18 [#deeplearning] algorithms ... presented at top-level research conferences ... Only 7 of them could be reproduced w/ reasonable effort ... 6 of them can often be outperformed w/ comparably simple heuristic methods."

Paper:
lnkd.in/dTaGCTv

#AI

Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches

Maurizio Ferrari Dacrema
Politecnico di Milano, Italy
maurizio.ferrari@polimi.it

Paolo Cremonesi
Politecnico di Milano, Italy
paolo.cremonesi@polimi.it

Dietmar Jannach
University of Klagenfurt, Austria
dietmar.jannach@aau.at

02v1 [cs.IR] 16 Jul 2019

ABSTRACT

Deep learning techniques have become the method of choice for researchers working on algorithmic aspects of recommender systems. With the strongly increased interest in machine learning in general, it has, as a result, become difficult to keep track of what represents the state-of-the-art at the moment, e.g., for top-n recommendation tasks. At the same time, several recent publications point out problems in today's research practice in applied machine learning, e.g., in terms of the reproducibility of the results or the choice of the baselines when proposing new models.

In this work, we report the results of a systematic analysis of algorithmic proposals for top-n recommendation tasks. Specifically, we considered 18 algorithms that were presented at top-level research conferences in the last years. Only 7 of them could be reproduced with reasonable effort. For these methods, it however turned out that 6 of them can often be outperformed with comparably simple heuristic methods, e.g., based on nearest-neighbor or graph-based techniques. The remaining one clearly outperformed the baselines but did not consistently outperform a well-tuned non-neural linear ranking method. Overall, our work sheds light on a number of potential problems in today's machine learning schol-

1 INTRODUCTION

Within only a few years, deep learning techniques have started to dominate the landscape of algorithmic research in recommender systems. Novel methods were proposed for a variety of settings and algorithmic tasks, including top-n recommendation based on long-term preference profiles or for session-based recommendation scenarios [36]. Given the increased interest in machine learning in general, the corresponding number of recent research publications, and the success of deep learning techniques in other fields like vision or language processing, one could expect that substantial progress resulted from these works also in the field of recommender systems. However, indications exist in other application areas of machine learning that the achieved progress—measured in terms of accuracy improvements over existing models—is not always as strong as expected.

Lin [25], for example, discusses two recent neural approaches in the field of information retrieval that were published at top-level conferences. His analysis reveals that the new methods do *not* significantly outperform existing baseline methods when these are carefully tuned. In the context of recommender systems, an in-depth analysis presented in [29] shows that even a very recent neu-

47 · 1,3 tis. · 2,6 tis.

https://twitter.com/dmonett/status/1154876433564655616

Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches

Maurizio Ferrari Dacrema et al. | https://dl.acm.org/citation.cfm?id=3347058

Fantómový progres môže byť spôsobený:

- **výber baselineov** – napr. zvolená neurónová sieť ako baseline (hoci sa ukázalo, že nie je baseline),
- nedostatok informácií o **optimalizácii baselinov**,
- **chyby pri delení** dát a pri implementácií **metrík**,
- množstvo datasetov, metrík, protokolov na delenie dát, ktoré sťažujú jednoduchšie a priame porovnanie.

Ghosting: Contextualized Inline Query Completion in Large Scale Retail Search

Lakshmi Ramachandran et al. | https://dl.acm.org/citation.cfm?id=3346995

datalys

RECSYS 2019

Ghosting: Contextualized Inline Query Completion in Large Scale Retail Search

Lakshmi Ramachandran et al. | https://dl.acm.org/citation.cfm?id=3346995

Recently searched query	bose bluetooth headphones	dslr camera	wireless charger
Frequency	jvc headphones	nike socks men	iphone charger
Session context	jvc bluetooth headphones ✓	nikon dslr camera ✓	iphone wireless charger ✓
Recently searched query	dance clothes for women	bike shorts	adidas trail running shoes
Frequency	zumba fitness dvd	bike seat	trash bags
Session context	zumba clothes for women ✓	bike shorts women ✓	trail running shoes men ✓

Ghosting: Contextualized Inline Query Completion in Large Scale Retail Search

Lakshmi Ramachandran et al. | https://dl.acm.org/citation.cfm?id=3346995

Recently searched query	bose bluetooth headphones	dslr camera	wireless charger
Frequency	jvc headphones	nike socks men	iphone charger
Session context	jvc bluetooth headphones ✓	nikon dslr camera ✓	iphone wireless charger ✓
Recently searched query	dance clothes for women	bike shorts	adidas trail running shoes
Frequency	zumba fitness dvd	bike seat	trash bags
Session context	zumba clothes for women ✓	bike shorts women ✓	trail running shoes men ✓

Table 3: Evaluating ghosting on a production system. The better model has a larger positive impact on acceptance, cart-add rate and net sales and larger negative impact on spell-correction rate and avg. prefix length. Statistically significant results are marked with "*" (t-test, p-value < 0.05). The session context model shows improvements on all metrics.

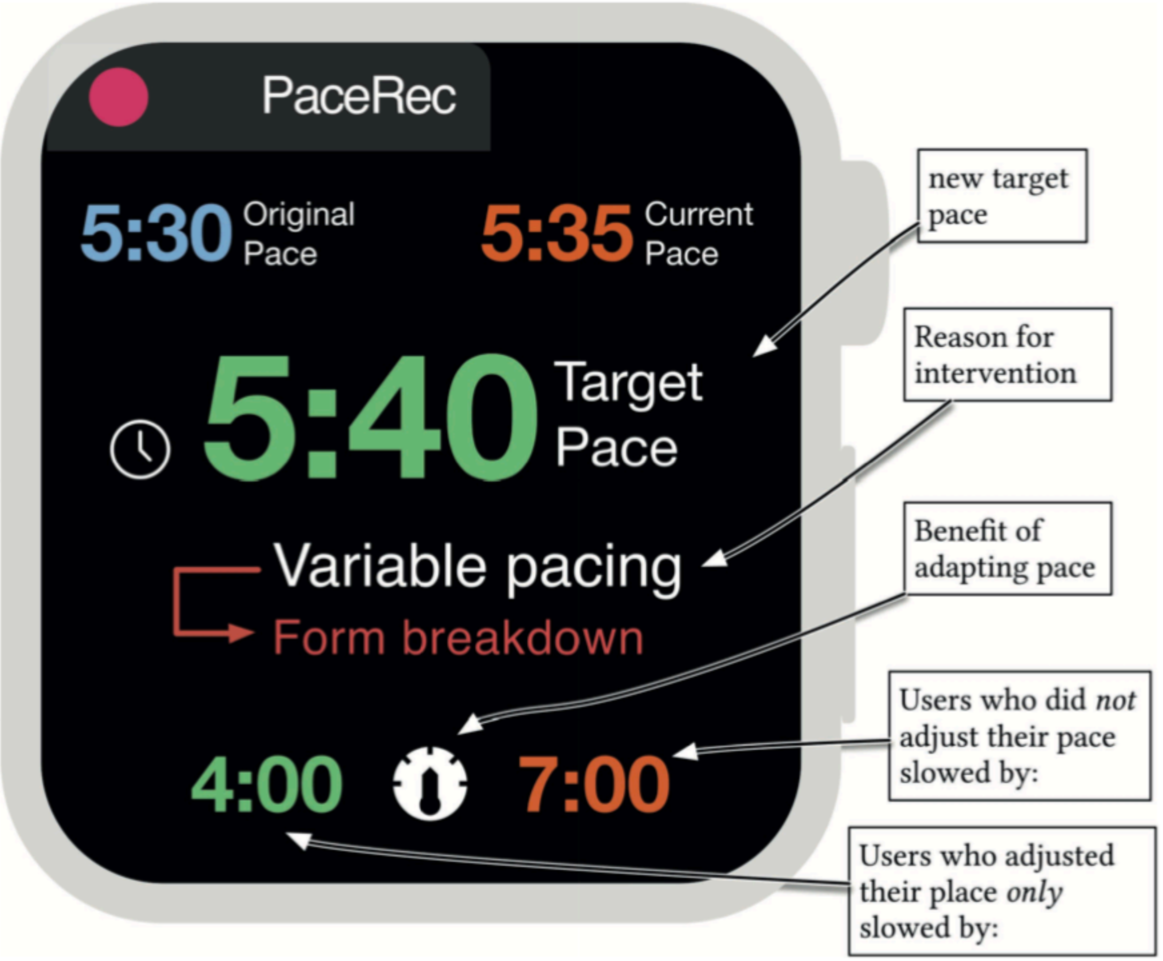
Ghosting model	acceptance	spell-correction rate	average prefix length	cart-add rate	net sales
Frequency	**+9.01***	**-9.52***	**-7.51***	-0.03	-0.15
Session context	+6.18*	-4.42*	-7.12*	**+0.04***	**+0.14***

Pace My Race: Recommendations for Marathon Running

Jakim Berndsen et al. | https://dl.acm.org/citation.cfm?id=3346991

Predikcia času ukončenia behu použitá na odporúčanie.

- Odporúčanie bežeckej stratégie v rôznych častiach behu.



INDUSTRY SESSIONS

RECSYS 2019

datalys

IKEA: Designer-Driven Add-to-Cart Algorithms

Sandhya Sachidanandan

Odporúčanie produktov s využitím obrázkov.

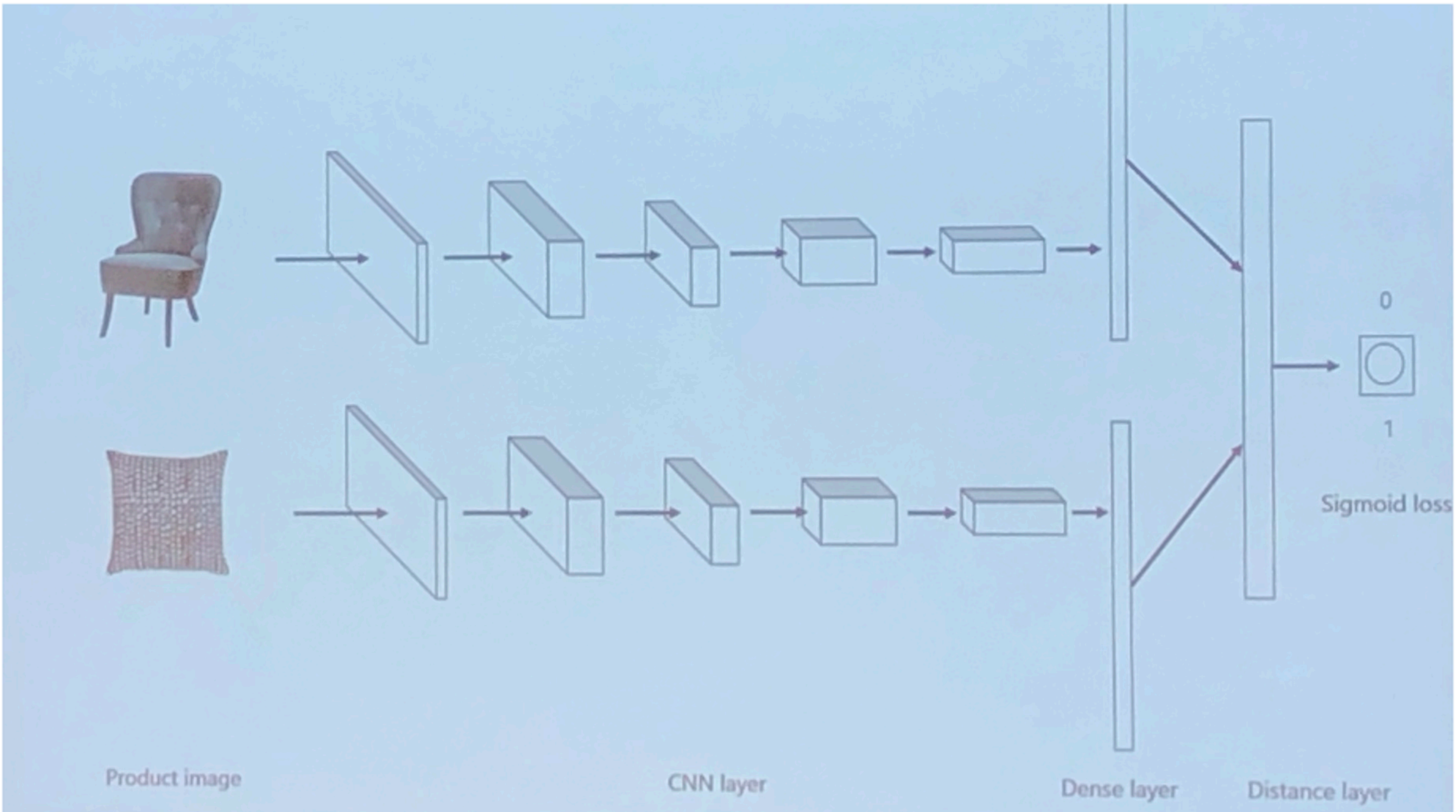
Siamské neurónové siete:

- pozitívne príklady: spolunákupy + expertné pravidlá
- negatívne príklady: rôzne farby, materiály, bez spolunákupov

Cieľ predpovedať: Hodí sa to k sebe?

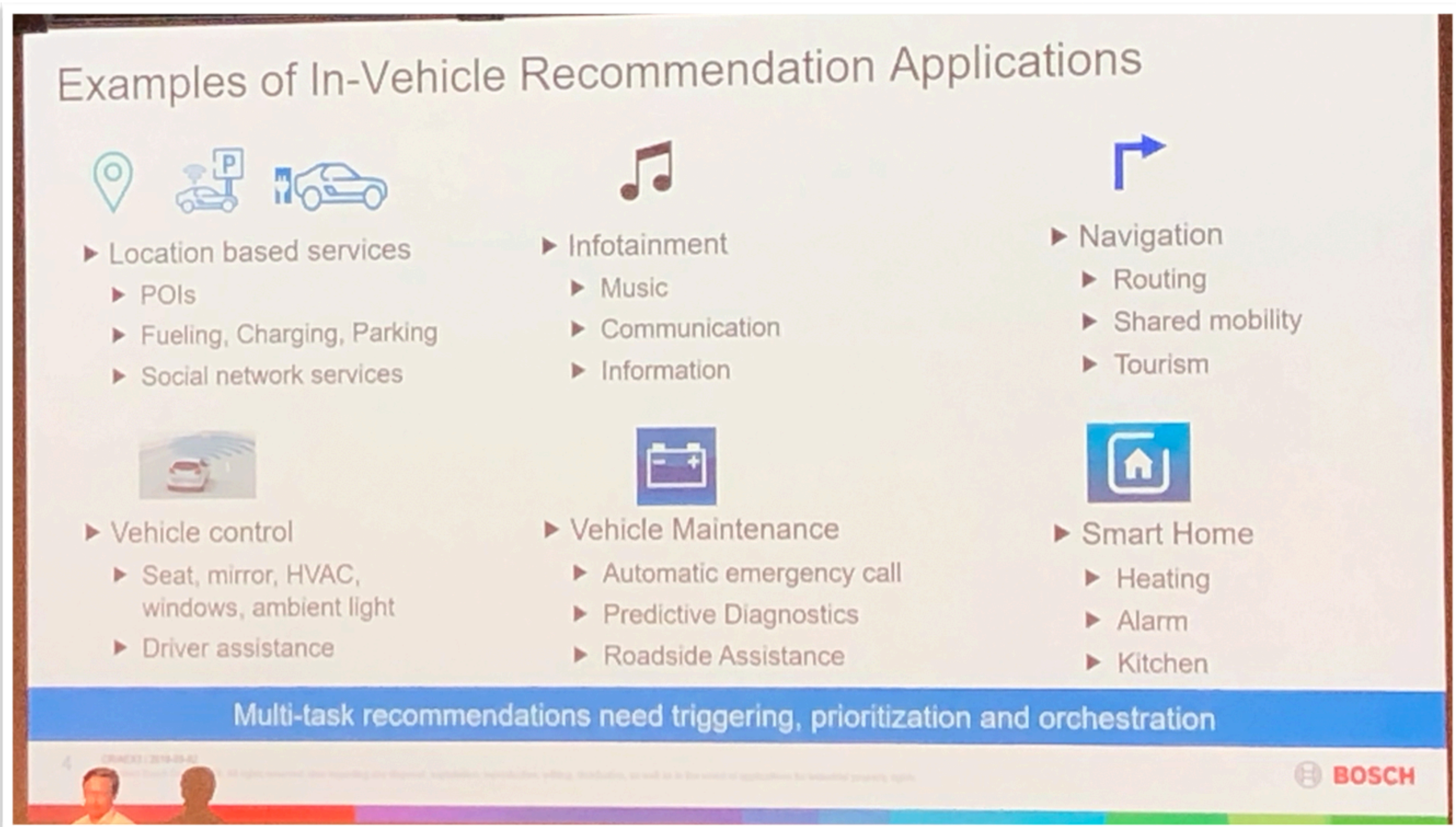
IKEA: Designer-Driven Add-to-Cart Algorithms

Sandhya Sachidanandan



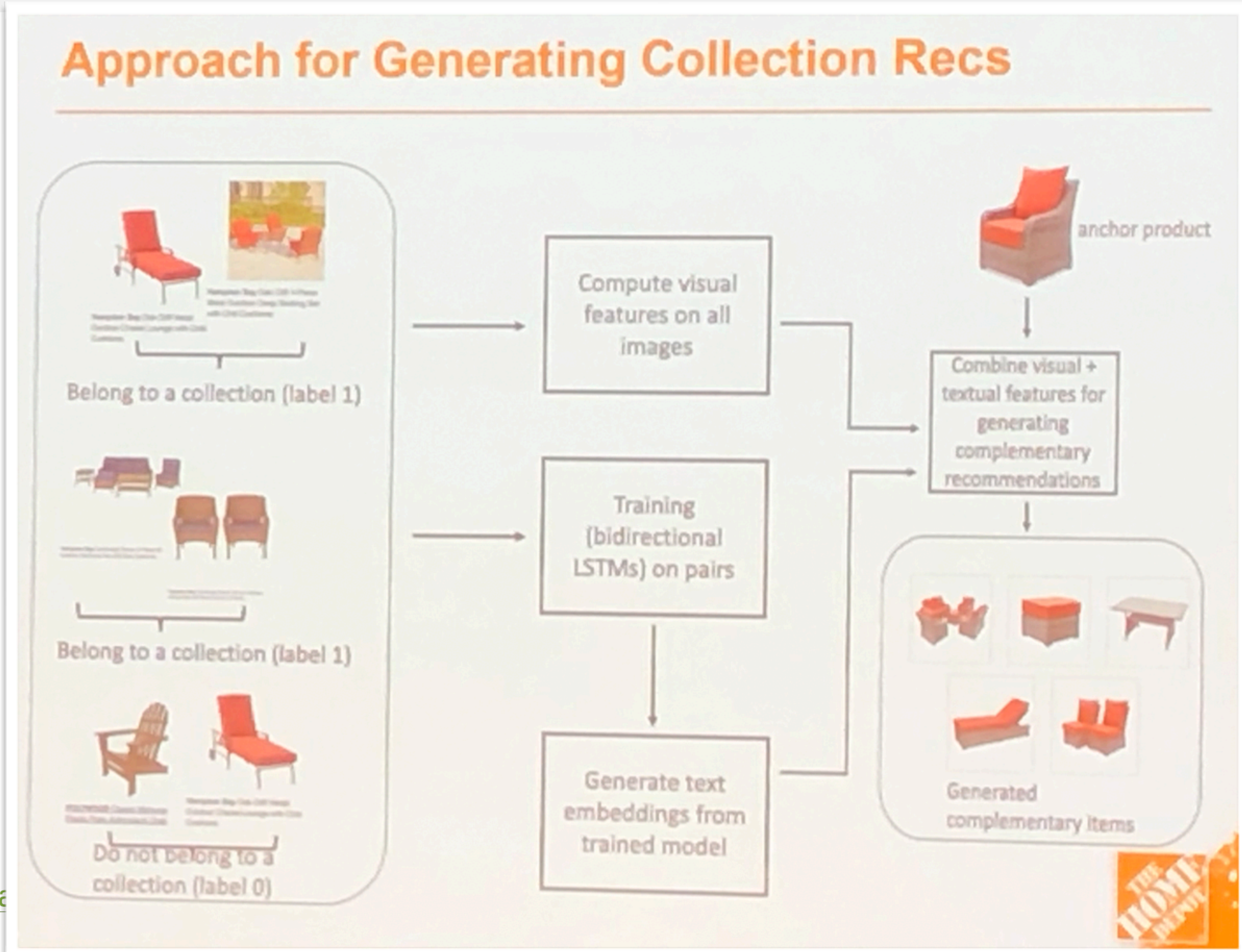
Bosch: Future of In-Vehicle Recommendation Systems

Juergen Luettin, Susanne Rothermel and Mark Andrew



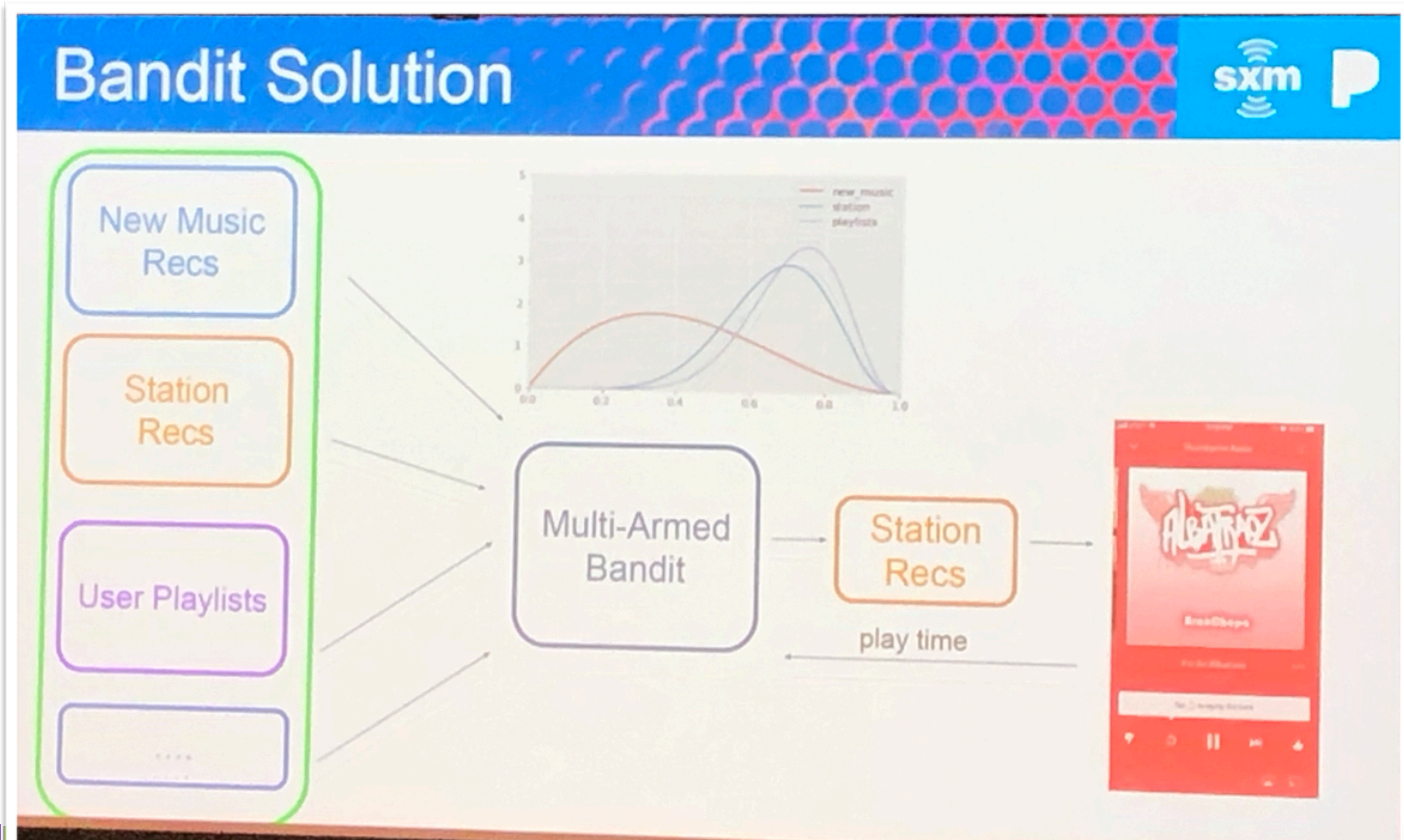
Home Depot: Recommendation in Home Improvement Industry, Challenges and Opportunities

Khalifeh Aljadda



Pandora: Just Play Something Awesome: The Personalization Powering Voice Interactions at Pandora

Vito Ostuni



KODAŇ


Booking.com
centipede
expedia group
Spotify
HUAWEI
NETFLIX
pandora
zalando
amazon
hulu
OpenTable
Medium
wayfair
Adevinta
Google
trivago
THANKS TO OUR
SPONSORS!
RECSYS 201


LIVA

RIO DE JANEIRO, BRAZIL, 22.-26. SEPTEMBER 2020