



Eelco Herder

Who told you that?

Transparency and control for privacy-aware
personalization

Eelco Herder

27 September 2017

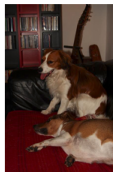


Eelco Herder

Who is Eelco Herder?



de Volkskrant





Eelco Herder

My research in a nutshell

My core research field is User Modeling and Personalization.

The application areas of my research have been extended to various related fields, including social media, technology-enhanced learning and mobile, ubiquitous computing.

In particular, I am interested in how users interact with programs, how they search for information, and how the return to places where they have been before.





Eelco Herder

Looking for a nice restaurant in Tilburg

You would probably recommend me the most popular or well-known restaurants. Or the ones that you like best.

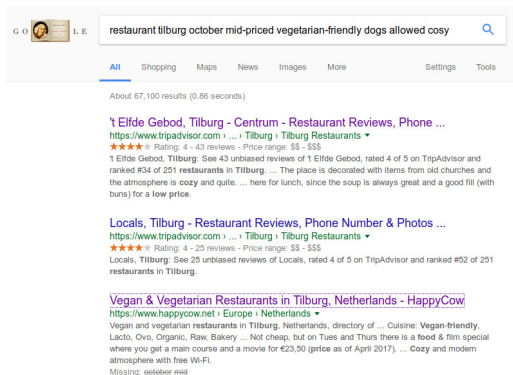
But did you know that I am a vegetarian, that I prefer mid-price restaurants, love Indian and Italian cuisine, find atmosphere more important than quality of food, would like to take the dog with me, and do not have a car at hand?





Eelco Herder

So I might better ask Google



Google search results for the query: "restaurant tilburg october mid-priced vegetarian-friendly dogs allowed cosy".

Search results include:

- 1 Elfde Gebod, Tilburg - Centrum - Restaurant Reviews, Phone ...**
<https://www.tripadvisor.com> ... Tilburg Tilburg Restaurants
★★★★★ Rating: 4 - 43 reviews - Price range: \$\$ - \$\$\$
1 Elfde Gebod, Tilburg: See 43 unbiased reviews of 1 Elfde Gebod, rated 4 of 5 on TripAdvisor and ranked #34 of 251 restaurants in Tilburg. ... The place is decorated with items from old churches and the atmosphere is cozy and quite. ... here for lunch, since the soup is always great and a good fill (with buns) for a low price.
- Locals, Tilburg - Restaurant Reviews, Phone Number & Photos ...**
<https://www.tripadvisor.com> ... Tilburg Tilburg Restaurants
★★★★★ Rating: 4 - 25 reviews - Price range: \$\$ - \$\$\$
Locals, Tilburg: See 25 unbiased reviews of Locals, rated 4 of 5 on TripAdvisor and ranked #52 of 251 restaurants in Tilburg.
- Vegan & Vegetarian Restaurants in Tilburg, Netherlands - HappyCow**
<https://www.happycow.net> Europe Netherlands
Vegan and vegetarian restaurants in Tilburg, Netherlands, directory of ... Cuisine: Vegan-friendly, Lacto, Ovo, Organic, Raw, Bakery ... Not cheap, but on Tues and Thurs there is a food & film special where you get a main course and a movie for €23,50 (price as of April 2017). ... Cozy and modern atmosphere with free Wi-Fi.
Missing: october mid

Hmm...

- ▶ Which aspect (or keyword) is more important?
- ▶ The average query length is 4.29 words



Eelco Herder

We expect interaction to be personalized

In a library, a person looks for some books on China. What will the librarian recommend?¹

- ▶ Is the person a *small child* who saw a TV show about China and wants to learn about this exotic country?
- ▶ Or a *high school student* working on a paper?
- ▶ Perhaps a *prospective tourist*?
- ▶ A scholar interested in *Eastern philosophy*?
- ▶ Someone who can *read Chinese*?

¹Elaine Rich: User Modeling via Stereotypes. Cognitive Science 3, 329-354 (1979)



Eelco Herder

Most likely the librarian will make an educated guess, based on the person's appearance:

- ▶ age, style of clothing, accent, choice of words, ...





Eelco Herder

This initial guess might be confirmed or refuted by observations.

- ▶ It is assumed that a European cannot read Chinese, unless said otherwise
- ▶ Children are generally not (yet) interested in Eastern philosophy, but there are exceptions
- ▶

The educated guess, a *stereotype* can be refined with follow-up questions.

Persons expect a *personalized* advice, even though the librarian does not know them.



Eelco Herder

And the same seems to yield for Web stores.



Jeff Bezos, amazon.com

If I have 3 million customers on the Web, I should have 3 million stores on the Web

When is personalization useful?

My supermarket is not personalized. Still, I can find all products that I need. Probably just because my needs are similar to everyone else's needs.



Personalization is deemed useful when:

- ▶ there are so many things to choose from that there is a need for guidance or recommendations
- ▶ the system is used by people with different goals and backgrounds



Eelco Herder

The ideal recommender

Your partner, your best friend or your mother probably knows a lot about you:

- ▶ the food you like, the books you read, the movies you watch
- ▶ things that interest you or that upset you
- ▶ your current needs, aspirations and goals
- ▶ dates of your birthday, your kids' birthdays, and holidays
- ▶ secret desires and phantasies





Eelco Herder

Still, this does not guarantee that your mother will buy you a present that you like.



It can be something that

- ▶ you already have
- ▶ you hate for some reason only known to you
- ▶ she bought to surprize you (sometimes this works out perfectly fine, though)

Eelco Herder

Personalization techniques

Most information-oriented websites are not personalized. They provide the same kind of content to anyone.

The screenshot shows the BBC News website interface. At the top, there is a navigation bar with categories like News, Sport, Weather, Travel, Culture, Autos, TV, Radio, and a search box. The main headline is "Snowden 'has not entered Russia'", accompanied by a photo of Edward Snowden and a sub-headline "Russia says tightens US intelligence leader Edward Snowden has not entered the country, as the US requests demands that he should be handed over". Below this, there are other news items: "Taliban assault in Kabul secure zone" and "Nelson Mandela's condition 'unchanged'". On the right side, there is a "Magazine" section with articles like "Loud and clear" and "Weighty debate". An advertisement placeholder is visible at the bottom right.

In many cases, this is fine. Users find the information they need - as long as the navigation structure is understandable.

Semi-personalized systems

Semi-personalized systems do not make use of a user profile. They try to adapt the content to the (estimated) needs *of everyone*.

For example:

- ▶ trending topics on Twitter
- ▶ most read items on a news site

Trending on Fox News

TRENDING HEADLINES	TRENDING VIDEOS
<p>All U.S. World Politics Entertainment Leisure Health SciTech Opinion Sports Travel</p> <p>Journalist Michael Hastings sent chilling email to colleagues before death 239</p> <p>Obama planning to sidestep Congress for next phase in climate change agenda 215</p> <p>Russian official says Snowden hasn't crossed into Russian territory, reject... 163</p> <p>Will anyone stand by Paula Deen, or is she toast? 29</p>	<p>Sen. Cruz: Immigration bill not designed to fix problems</p> <p>Snowden case a test for resetting US-Russia relations?</p>




Eelco Herder

A personal touch and personal functionality

Particularly transaction sites (online stores) require users to log in or to create an account.

It is customary that these sites 'welcome' the user, which creates some basic 'mutual recognition'.

Hello **Eelco**



[Show unread posts since last visit.](#)
[Show new replies to your posts.](#)
There are [19 members](#) awaiting approval.
Total time logged in: 115 days, 8 hours and 41 minutes.



Eelco Herder

Why bother to log in

A personal greeting is not a very convincing reason for asking users to register.

More convincing - and practical:

- ▶ Easier, quicker checkout (less information to fill out when buying an item)
- ▶ Access to previous orders, perhaps with the probability to re-order
- ▶ Saving of default location, default language, default values



Eelco Herder

I personally really appreciate this sort of functionality - most users expect (ecommerce) sites to offer such features.

Meine Übersicht

Hallo, Herr Eelco Herder!
Von Ihrer Benutzerkonto-Übersicht aus haben Sie die Möglichkeit, Ihre letzten Vorgänge einzusehen und Ihre Benutzerkonto-Daten zu bearbeiten. Wählen Sie dazu einen der untenstehenden Links, um Informationen anzusehen oder zu bearbeiten.

Letzte Bestellungen
[Alle ansehen](#)

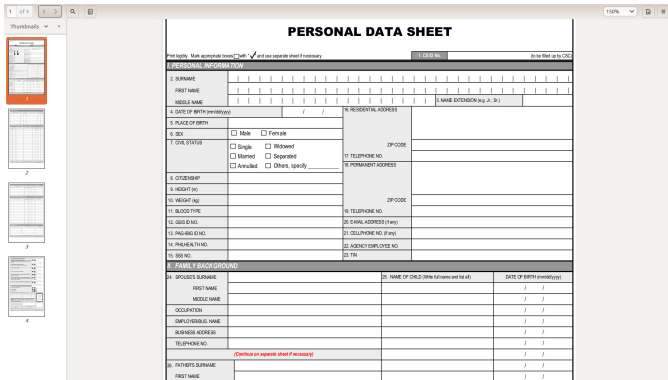
Bestellung	Datum	Senden an	Bestellsumme	Status	
200156704	02.06.13	Herr Eelco Herder	40,86 €	Bestellung komplett	Bestellung ansehen Nachbestellen
200144927	22.04.13	Herr Eelco Herder	40,94 €	Bestellung komplett	Bestellung ansehen
200134329	14.03.13	Herr Eelco Herder	73,44 €	Bestellung komplett	Bestellung ansehen
200122541	30.01.13	Herr Eelco Herder	49,96 €	Bestellung komplett	Bestellung ansehen
200109249	07.12.12	Herr Eelco Herder	47,38 €	Bestellung komplett	Bestellung ansehen

Personalization goes beyond this 'basic' functionality.



Eelco Herder

It's good that the government has some data about me
I wouldn't want to fill out the same personal data over and over
again. Would you?



PERSONAL DATA SHEET

Print legibly. Mark appropriate boxes and use separate sheet if necessary. 3. CDS No. Date filed up by CDS

I. PERSONAL INFORMATION

2. SURNAME			
FIRST NAME			
MIDDLE NAME			
3. NAME EXTENSION (e.g. Jr., Sr.)			
4. DATE OF BIRTH (mm/dd/yyyy)	/	/	
5. PLACE OF BIRTH			
6. SEX	<input type="checkbox"/> Male <input type="checkbox"/> Female		
7. CIVIL STATUS	<input type="checkbox"/> Single <input type="checkbox"/> Widowed <input type="checkbox"/> Married <input type="checkbox"/> Separated <input type="checkbox"/> Annulled <input type="checkbox"/> Other, specify _____		
8. CITIZENSHIP			
9. HEIGHT (in)			
10. WEIGHT (kg)			
11. BLOOD TYPE			
12. SSN ID NO.			
13. PANG-BBS ID NO.			
14. PHARMAHEALTH NO.			
15. BSN NO.			
16. RESIDENTIAL ADDRESS			
ZIP CODE			
17. TELEPHONE NO.			
18. PERMANENT ADDRESS			
ZIP CODE			
19. TELEPHONE NO.			
20. CAREGIVER ADDRESS (if any)			
21. CELLPHONE NO. (if any)			
22. AGENCY EMPLOYEE NO.			
23. TIN			

II. FAMILY BACKGROUND

19. SPOUSAL SURNAME	20. NAME OF CHILD (Only full names and full ad.)	DATE OF BIRTH (mm/dd/yyyy)
FIRST NAME		/ /
MIDDLE NAME		/ /
OCCUPATION		/ /
EMPLOYER/BUS. NAME		/ /
BUSINESS ADDRESS		/ /
TELEPHONE NO.		/ /
<i>(Check on separate sheet if necessary)</i>		
24. PATIENT'S SURNAME		/ /
FIRST NAME		/ /

Understanding the user

Tools such as Google Analytics show general trends:

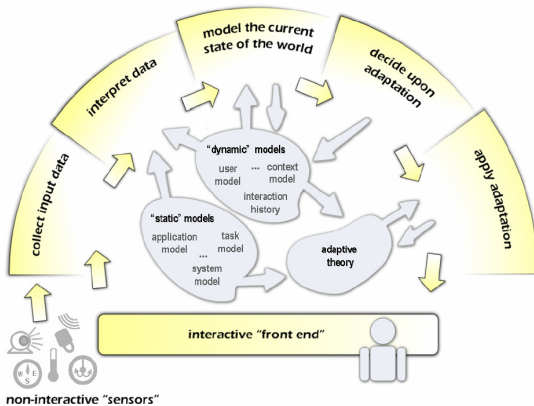
- ▶ number of visits and users
- ▶ where do users come from, which systems do they use
- ▶ popular pages and keywords



User modeling is about getting to know the *individual* user.

Eelco Herder

The adaptation process: a process-oriented view²



² Alexandros Paramythis, Stephan Weibelzahl, Judith Masthoff. Layered evaluation of interactive adaptive systems: framework and formative methods. *User Model. User-Adapt. Interact.* 20(5): 383-453 (2010)



Eelco Herder

Explicitly provided user data on the Web

Personal data has always been published and shared in the World Wide Web.

The Web 2.0 has promoted this even more by offering a variety of services where users can publish data without requiring specific technical skills on:

- ▶ social networking sites
- ▶ photo and document sharing sites
- ▶ collaborative work environments
- ▶ blogs
- ▶ and many other sites.

Implicitly collected user data

In addition, many of our activities on our computers and on the Web are logged in some way.

- ▶ **Ecommerce sites** register which items we browsed for and which we bought.
- ▶ **Social networking sites** keep track of the messages, pictures and likes that we broadcast and send to our friends.
- ▶ **Browsers** maintain a history of sites that we visited and store cookies.
- ▶ **Desktop search engines** index all programs and files that we used.





Eelco Herder

Who or what can access the data?

Much personal data is only accessible by ourselves and is used for **private purposes**, such as reflection, archival and refinding.

Many other traces are **shared with others** or broadcast to the world - voluntarily or involuntarily³.

These public or semi-public traces define our online presence: the way we are seen by the outside world, based on which others judge who we are.

And even personal data that is not disclosed is often **used for inferring** our interests or for providing recommendations.

³Kaweh Djafari Naini, Ismail Sengor Altingovde, Ricardo Kawase, Eelco Herder, Claudia Niederee. Analyzing and Predicting Privacy Settings in the Social Web. Proc. UMAP 2015.



Eelco Herder

Information asymmetry⁴

A fundamental privacy concern is the *information asymmetry* between

- ▶ website providers as data collectors and
- ▶ users as data providers

due to the absence of adequate control mechanisms of how user data are collected and whether they are disseminated.



⁴See for instance Tsai, J. Y., Egelman, S., Cranor, L., Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, 22(2), 254-268.

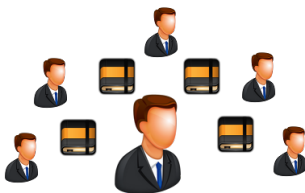


Eelco Herder

Limited user awareness

In most environments it is hard to figure out what exactly is being logged or to inspect and regulate which traces are stored, used by other applications or published to the outside world.

This raises serious privacy issues, of which the average user is often not aware.



Eelco Herder

The consequences can be real⁵

The screenshot shows a news article from a German website. The article title is "Fire Me! Who wants to get fired?" and the sub-headline is "Diese Website outet meckernde Job- und Chef-Hasser". The article features a photo of a woman with her arms raised in a cheering gesture. To the right of the article are three tweets:

- Tweet 1: "My boss is a fucking cunt"
- Tweet 2: "My boss is a fucking idiot for opening on thanksgiving. You're so fucking selfish"
- Tweet 3: "I hate creating new business forms. My boss is a fucking jerk!"

In 2008, 8% of U.S. companies with 1000 workers or more reported firing an employee based on information released on social networks.

⁵Ricardo Kawase, Bernardo Pereira Nunes, Eelco Herder, Wolfgang Nejdl, Marco Antonio Casanova. Who Wants To Get Fired. Proc. Web Science 2013



Eelco Herder

The end of privacy via obscurity⁶

Through the powerful reach of search engines, obscure pieces of personal information - such as email messages sent a decade ago to niche forums or newsgroups - are increasingly retrievable by a simple keyword search.

As a result, any 'privacy via obscurity' that generally kept such information from public view has been diminished.

Hello Google+. Long time no see. I just checked in to add a colleague ...



<https://plus.google.com/113584688972177896510/posts/Xnd4pUNCgST>

Eelco Herder

Nov 30, 2011 - Hello Google+. Long time no see. I just checked in to add a colleague (who added me) to my circles. Let's check what I missed thus far outside the Facebook world

What's new? My first message in Google+, that's what's new. Hi ...



<https://plus.google.com/113584688972177896510/posts/CSihY4MkAk>

Eelco Herder

Jul 1, 2011 - What's new? My first message in Google+, that's what's new. Hi everyone and thanks to Ricardo for the invite

⁶Zimmer, M. (2008). The externalities of search 2.0: The emerging privacy threats when the drive for the perfect search engine meets Web 2.0. *First Monday*, 13(3).

The issue with linked data and mash-ups⁷

Mash-ups, clever combinations of existing applications, provide interesting new applications and user-friendly one-stop services.

Further, many social networking applications have methods for synchronizing content and contacts, as well as for integration in other sites.

As a result, users receive better, more integrated services, but they are also exposed to serious threats to informational privacy.



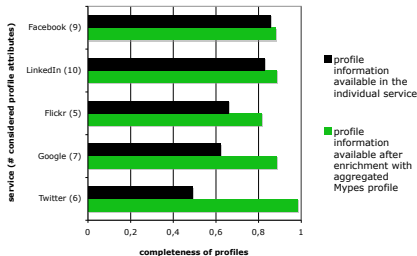
1,497 562

⁷Bizer, C., Heath, T., Berners-Lee, T. (2009). Linked data-the story so far. Semantic Services, Interoperability and Web Applications: Emerging Concepts, 205-227.

Completeness of user profiles

Users often do not fill out their profiles completely. For example, Twitter only asks 6 attributes, but these profiles are only completed up to 49%⁸.

Aggregating data from different sources leads to more complete user profiles.



⁸ Fabian Abel, Eelco Herder, Geert-Jan Houben, Nicola Henze, Daniel Krause. Cross-system User Modeling and Personalization on the Social Web. *UMUAI Journal on User Modeling and User-Adapted Interaction* 23 (2-3), 2013, pp 169-209



Eelco Herder

The issue is often not the data itself, but the interpretation⁹

Mr. Iwanyk suspected that his TiVo thought he was gay, since it kept recording gay programs. He tried to tame TiVo by recording war movies and other “guy stuff.”

“I overcompensated. It started giving me documentaries on Goebbels and Eichmann. It stopped thinking I was gay and decided I was a crazy guy.”



⁹Zaslow Jeffrey. 2002. If TiVo Thinks You Are Gay, Heres How to Set It Straight. Wall Street Journal, November 26.



Eelco Herder

Some interpretations may seem quite far-fetched at first...¹¹



Apart from the fact that people who like curly fries are more intelligent, it is also possible to find out whether a teen girl is pregnant¹⁰

¹⁰ <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#77184eaf6668>

¹¹ Jennifer Golbeck: The curly fry conundrum: Why social media likes say more than you might think. TEDxMidAtlantic 2013. https://www.ted.com/talks/jennifer_golbeck_the_curly_fry_conundrum_why_social_media_likes_say_more_than_you_might_think


Eelco Herder

Ultimately the question is what is being done with your user profile

The image shows a composite of two web pages. The top page is a Mediamarkt product page for a Whirlpool F5C70410 MIT washing machine, priced at 389,-. The bottom page is a Criteo retargeting advertisement. The Criteo ad includes a header with the Mediamarkt logo and a navigation bar for 'Vegatopia' and 'horecagrootkeukenshop.nl'. The main content of the Criteo ad is titled 'WHY AM I BEING SHOWN THIS BANNER?' and explains that it shows products viewed on the Mediamarkt website. It lists two products: 'ASUS Notebook E1420NA-SMART' and 'HP Notebook 14 EP501ND'. A privacy notice at the bottom states: 'CRITEO TAKES YOUR PRIVACY VERY SERIOUSLY. Criteo collects browsing data such as product viewed and no identifying data. For more information on Criteo's privacy policy, visit Criteo.com. Criteo employs personalized banners that are more likely to be of interest to you than non-targeted advertising. We personalize our banners based on your recent browsing behavior.'

Being shown an advertisement for washing machines after having searched (and bought one) at Mediamarkt might be a bit creepy. But it doesn't have real consequences. Apart from that, I can even check why this advertisement has been shown to me.


But what if the price I have to pay is heavily influenced by my user profile?¹²



Contents lists available at [ScienceDirect](#)

International Journal of Hospitality Management

journal homepage: www.elsevier.com/locate/ijhosman



Tourists' willingness to pay for an accommodation: The effect of eWOM and internal reference price

Marta Nieto-García*, Pablo A. Muñoz-Gallego, Óscar González-Benito

Instituto Multidisciplinar de Empresa, Universidad de Salamanca, Campus Miguel de Unamuno, Edificio FES, 37007, Salamanca, Spain

ARTICLE INFO


Article history:
Received 23 August 2016
Received in revised form 3 November 2016
Accepted 12 December 2016
Available online 27 December 2016

Keywords:
Online reviews
Willingness to pay
Internal reference price
eWOM valence
eWOM volume
Dynamic pricing

ABSTRACT

Understanding the determinants of consumers' willingness to pay (WTP) is an important challenge especially for practitioners. This study evaluates the effect of external information (eWOM valence and volume) and internal information (internal reference price) on consumers' WTP for an accommodation. The results of an online experiment (n = 766) show a direct effect of valence on WTP which is strengthened by both volume and the internal reference price. Consumers with high reference prices are more sensitive to the effect of an increase in valence. Moreover, internal reference price has a non-linear influence on WTP. The findings suggest the relevant role of eWOM as well as internal reference price in determining consumers' WTP. The inclusion of these two variables in dynamic pricing strategies could lead to greater benefits for hospitality managers.

© 2016 Elsevier Ltd. All rights reserved.



¹² Nieto-García, M., Muñoz-Gallego, P. A., González-Benito, Ó. (2017). Tourists willingness to pay for an accommodation: The effect of eWOM and internal reference price. *International Journal of Hospitality Management*, 62, 67-77.

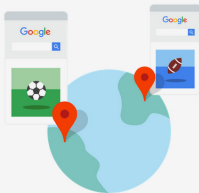
Personalized search results

Useful when the personalization is done right. Annoying when factors are taken into account that are wrong or not relevant - and one cannot turn it off.

Considering context

Information such as your location, past search history and Search settings all help us to tailor your results to what is most useful and relevant for you in that moment.

We use your country and location to deliver content relevant for your area. For instance, if you're in Chicago and you search "football", Google will most likely show you results about American football and the Chicago Bears first. Whereas if you search "football" in London, Google will rank results about soccer and the Premier League higher. [Search settings](#) are also an important indicator of which results you're likely to find useful, such as if you set a preferred language or opted in to [SafeSearch](#) (a tool that helps filter out explicit results).



Google's personalization is like my mother buying the wrong present



We cannot predict or analyze 100% accurately your current needs and context. Even your mother cannot.

This is also the reason why a filter bubble feels uncomfortable: it is not self-chosen or self-made.

Plus there are other factors at work, like commercial interests. And that's why transparency and control is needed.



Eelco Herder

Judy Kay: Scrutable Adaptation¹³

Capable of being understood through study and observation;
comprehensible and understandable upon close examination

- ▶ What information does the system collect about me to drive this personalisation?
- ▶ What does this system do with that information?
- ▶ With whom does it share the information?
- ▶ What information is in this part of my environment?
- ▶ How is it combined with other information about me?
- ▶ What is the meaning of "prefers low cost travel"?
- ▶ How did the system conclude that I prefer to travel at low cost?
- ▶ How can I get a big picture of the models related to my travel?
- ▶ How did the system choose to put this information here and now?
- ▶ What did it present a week ago?
- ▶ What would it present if I were a rich man?
- ▶ How do I change any of the above processes?

¹³Judy Kay. Scrutable Adaptation: Because We Can and Must. Proc. Adaptive Hypermedia 2006.



Eelco Herder

A scrutable hypertext¹⁴

Home	Contents	Your Profile	Make Notes	Change Topic	Help
<ul style="list-style-type: none"> Removed by personalisation Added by personalisation Key: <ul style="list-style-type: none"> No colour - content added for everyone 					Personalisation Highlighted 2 items removed 4 items added Return to normal page view
Morning TV Guide Your morning viewing schedule:					
4:30am-5am					
Added because profile has: You are interested in Current Affairs programs					because your profile has: <ul style="list-style-type: none"> You are NOT a member of SIG1 You are interested in Current Affairs programs You are interested in Lifestyle programs You are interested in Childrens programs
<ul style="list-style-type: none"> SBS - Filipino News 					
Added because profile has: You are interested in Childrens programs					
<ul style="list-style-type: none"> Ch.9 - Round the Twist 					
Removed because profile has: You are NOT a member of SIG1					
<ul style="list-style-type: none"> Ch.10 - Enjoying Everyday Life with Joyce Meyer 					

- ▶ Users were able to scrutinise simple forms of personalisation, but mixed results for more complex forms.
- ▶ The profile tool was effective and users found it easy to use.
- ▶ *“Surprised and upset that the system made them a member of the Special Interest Group for religious TV programs.”*

¹⁴ A Scrutable Adaptive Hypertext. Marek Czarkowski. PhD thesis, University of Sydney, 2006.

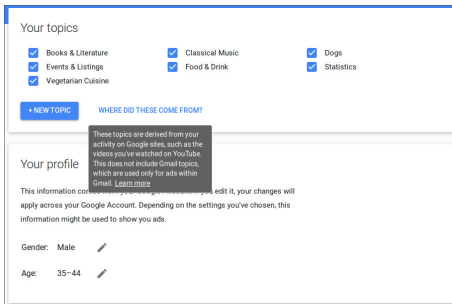


Eelco Herder

The usability of privacy controls is improving quickly

I wanted to show the arcane privacy controls for Google Ads. Apparently, Google (like Facebook) recently put a lot of effort in the usability.

Still, how much time would users want to spend on settings for *advertisements*?



Your topics

- Books & Literature
- Events & Listings
- Vegetarian Cuisine
- Classical Music
- Food & Drink
- Dogs
- Statistics

[+ NEW TOPIC](#) [WHERE DID THESE COME FROM?](#)

Your profile

These topics are derived from your activity on Google sites, such as the videos you've watched on YouTube. This does not include Gmail topics, which are used only for ads within Gmail. [Learn more](#)

This information controls the ads you see. If you edit it, your changes will apply across your Google Account. Depending on the settings you've chosen, this information might be used to show you ads.

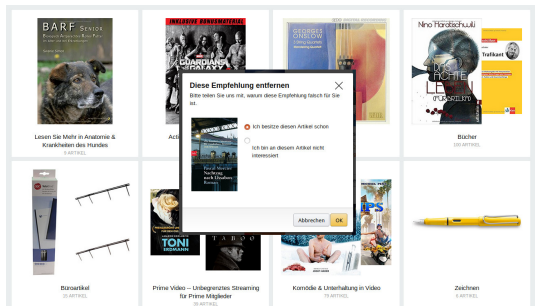
Gender: Male

Age: 35-44



Eelco Herder

Amazon offers methods for controlling my recommendations

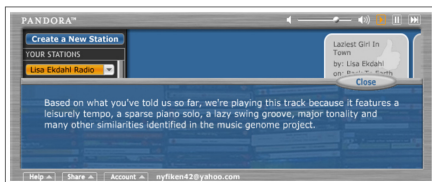


- ▶ On an item-by-item basis (not scalable)
- ▶ It is not explained what else is in my user model

Why not explanations for any type of personalization?

This is an active research field in recommender systems¹⁵.

Aim	Definition
Transparency (Tra.)	Explain how the system works
Scrutability (Scr.)	Allow users to tell the system it is wrong
Trust	Increase users' confidence in the system
Effectiveness (Efk.)	Help users make good decisions
Persuasiveness (Pers.)	Convince users to try or buy
Efficiency (Efc.)	Help users make decisions faster
Satisfaction (Sat.)	Increase the ease of use or enjoyment



¹⁵ Tintarev, N., Masthoff, J. (2011). Designing and evaluating explanations for recommender systems. Recommender Systems Handbook, 479-510.



Eelco Herder

Data security and legislation are only part of the solution

Sometimes scrutability is deemed unnecessary or cumbersome, sometimes simply as not wanted.

Why should a company explain why you have to pay a higher price?

Regulations and fines might be a solution. But ultimately you want the user to be in control, or at least to be able to be in control.



FOK! Welkom op FOK!

Gamy dat we je zo bruta lastgelegd. De overheid wil echter graag dat we je melden dat we op FOK! niet als eke andere site doet, cookies gebruiken. Onze website voor het plaatsen van je aanvragen. Maar het stikken op de privacy leeg of ook. Immers dat je ekele melding en je akkoord met het plaatsen van cookies en het plaatsen van FOK! is. Het is belangrijk om het niet meer over te laten. Het geldt op FOK!

Het gebruik van cookies is veilig. Er wordt geen directe persoonlijke informatie, zoals je naam, adres, telefoonnummer, mobiele of internetnummer in cookies opgeslagen. Daarvoor kunnen cookies ook niet worden gebruikt voor je mail en internetverkeer. Maar cookies kan nodig voor als je een keer terug gaat. Daar gebruik te blijven maken van deze site je je aanmelding afkondt over je wachtwoord.

Je ontvangt deze melding maar één keer per browser (mits je het gebruikende cookie niet verwijdert en je browser toelate cookies van FOK! op te slaan - zie de tekst). In geval van problemen kun je contact voor de klantendienst. Het is niet mogelijk FOK! te blokkeren zonder cookies te accepteren. Je kunt cookies wel zelf kiezen in je browser blokkeren retour.

Ja, ik wil een goed werkende site...
...en mijn gegevens plaatsen

[Natuurlijk accepteer ik de cookies](#)

© 2014 FOK! - Alle rechten voorbehouden



Eelco Herder

Requiring companies to do so without telling them which rules exactly to adhere to doesn't work - they will find a workaround.

Our job is to provide metaphors, workflows and templates that companies have to adhere to, to ensure that:

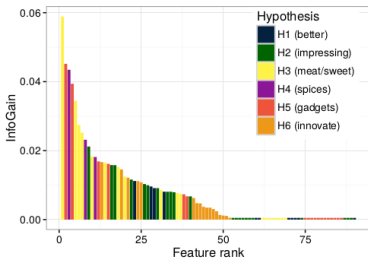
- ▶ it is transparent how search results are adjusted
- ▶ why your Facebook feed contains which items (and which not)
- ▶ why you are given a loan (or not)
- ▶ ...



Eelco Herder

Counter argument: our models are too hard to explain

Clustering methods, support vector machines, regression methods, ... they all can provide insight in which factors are most important (i.e. have the highest loadings). Often, a small number of small features provides sufficient insight¹⁶.



¹⁶ Figure taken from: Markus Rokicki, Eelco Herder, Tomasz Kusmierczyk and Christoph Trattner. Plate and Prejudice: Gender Differences in Online Recipes. Proc. UMAP 2016.



Eelco Herder

Counter argument: how can you tell whether a price or feed is personalised or not?

Some companies might not be willing to provide the full details on how they personalize.

This is why more and more studies focus on reverse-engineering the inner workings of personalized search engines and the like¹⁷.

Room type	Sleeps	Today's price	Your choices	Select room	Confirm your reservation
<p>Double Room</p> <p>2 people and 2 children 12</p> <p>No high demand - only 3 rooms left on our site!</p> <p>1 double bed</p> <p>Private bathroom • Flat-screen TV</p> <p>Concierge</p> <p>Shower • Safety Deposit Box • TV • Telephone</p> <p>Hairdryer • Iron • Radio • Desk</p> <p>Household • Fan • Hair • Heating</p> <p>Satellite Channels • Bath or Shower</p> <p>Wake-up service • Electric kettle</p> <p>Wardrobe/Closet • Terrace • Linen</p> <p>Coffee/tea • Toilet paper</p> <p>Prices are per room, including 21 % VAT, Breakfast</p>	2	<p>US\$192</p> <p>US\$192</p>	<ul style="list-style-type: none"> Continental breakfast included FREE cancellation before 18 Sept 2017 NO PREPAYMENT NEEDED - pay at the property 	<p>US\$192</p>	<p>Reserve</p> <p>Confirmation is immediate</p> <p>30 other people looking now, according to our Booking.com travel scientists</p>
<p>Jackpot! This is the cheapest price you've seen in London for your dates</p> <p>1 night (Thu 21 Sept - Fri 22 Sept)</p>					
<p>Triple Room</p> <p>1 single bed and 1 double bed</p> <p>Private bathroom • Flat-screen TV</p> <p>Concierge</p> <p>Prices are per room, including 21 % VAT, Breakfast</p>	3	<p>US\$182</p> <p>US\$182</p>	<ul style="list-style-type: none"> Continental breakfast included FREE cancellation before 18 Sept 2017 NO PREPAYMENT NEEDED - pay at the property 	<p>US\$182</p>	<p>Reserve</p> <p>Last changed only 1 week left on our site!</p>

¹⁷ e.g. Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., Wilson, C. (2013, May). Measuring personalization of web search. In Proceedings of the 22nd international conference on World Wide Web (pp. 527-538). ACM.

High-level research agenda

- ▶ Understand the impact of personal data on information and services served to the user
- ▶ Understand user choices and user understanding regarding online privacy
- ▶ Develop generic metaphors and approaches for scrutability (transparency and control)
- ▶ Evaluate user understanding, acceptance and adoption of these tools





Eelco Herder

Questions?



Privacy Engineering, User Modeling, Personalization,
Recommendation, Web Usage Mining,
Data Analysis and Visualization,
Usability, Evaluation