

# Introductie AI

Quantifying bias in search engines

Prof.dr.ir. Arjen P. de Vries



***GEEN  
MR...***

**CPO**<sup>•</sup>

*Eigenlijk niet...*

- *Opdracht van Mireille:*  
Kun je studenten uitleggen hoe je een ML systeem zo kunt ontwerpen dat je daadwerkelijk zo weinig mogelijk data verwerkt en toch kwaliteit levert?
- *KIA Digitale Technologieën – AI (ST2-1), Q2:*  
Op welke wijze kunnen we deep-learning architecturen data-efficiënt training door expert kennis toe te voegen?

# Outline

---

- Big Data
  - Waarom nou juist BIG data?!
- Bias on the Web
- Bias in Search Engines
  - Quantifying Bias using Retrievability
- Bias in Word Embeddings
  - [informagus.nl/edu/cpo/](http://informagus.nl/edu/cpo/)

# Big Data

Data-gedreven “intelligentie”



CPO<sup>•</sup>

# Big Data: nieuwe mogelijkheden!

---

Om data te

- Genereren,
- Delen,
- Combineren
- Analyseren

.. die leiden tot nieuwe inzichten en een nieuwe manier van redeneren.

(bron: Definitie van big data van de *Nationale DenkTank*)

# Hoeveel data?

---

Nou...

- Al het werk van William Shakespeare:
- 5.500.000 letters...
- ...wordt geschreven in slechts 11,6 Twitter-seconde!

[quora.com/How-much-data-does-Twitter-store-daily](https://www.quora.com/How-much-data-does-Twitter-store-daily)



# Een wonderlijke puzzel!

---

8 TB  
per day  
in total  $\approx$  100 MB  
per sec



*Data Center (is the Computer)*

Source: Google





# Big Data in NL

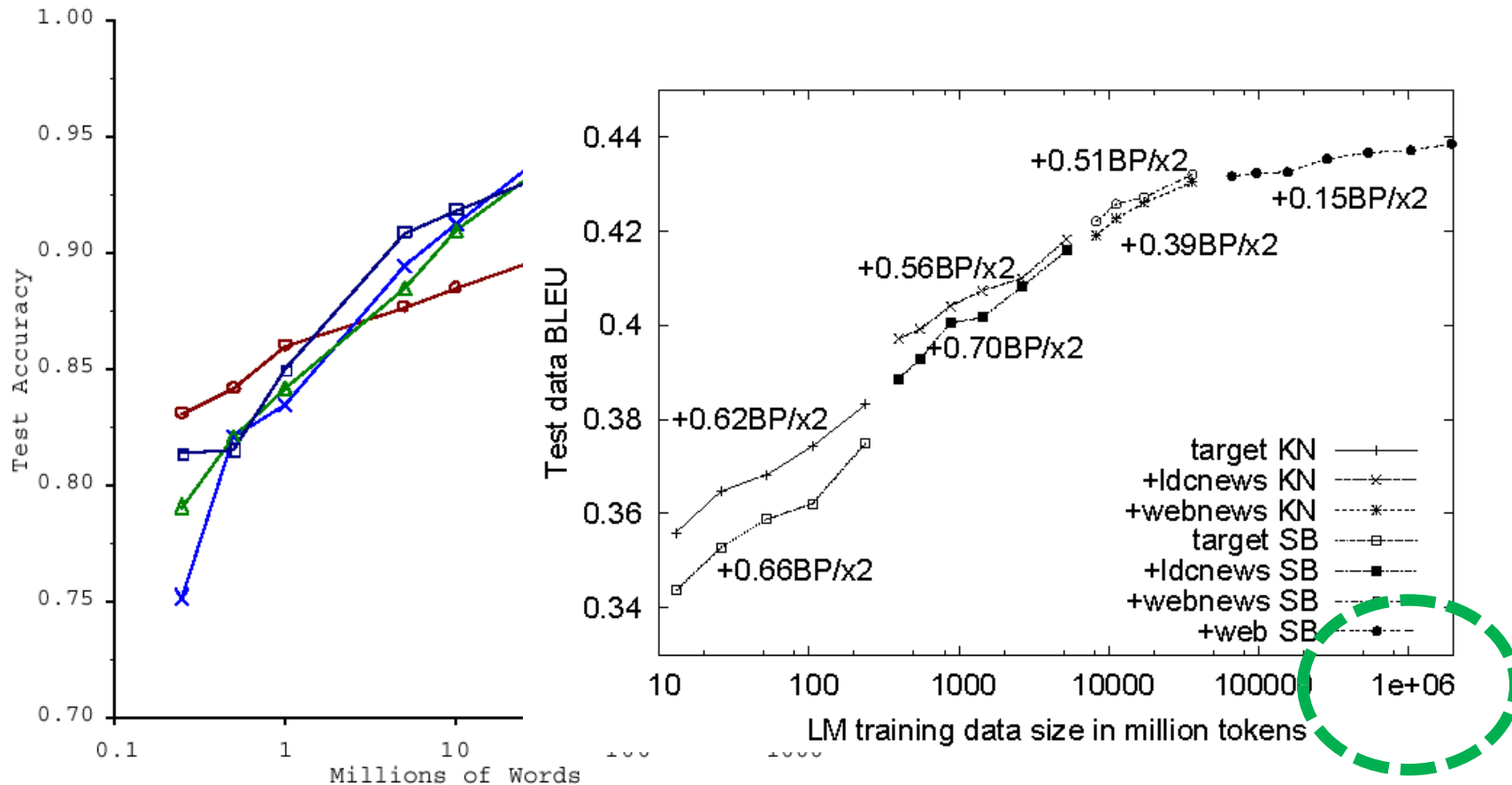
- Aankopen bij [bol.com](http://bol.com)
- Bestemmingen van [Booking.com](http://Booking.com) op maat
- Advertenties real-time verkopen, bv. bij [nu.nl](http://nu.nl) (Sanoma)
- Kinderzoekmachine [WizeNoze.com](http://WizeNoze.com)

*Etc. etc.*

# Waarom BIG data?

## Well... there's NO data like MORE data!

s/knowledge/data/g;



(Banko and Brill, ACL 2001): Large amounts of data

(Brants et al., EMNLP 2007): Variety of data

# Use Google Data?!

---

- E.g., Google releases *pre-trained word and phrase vectors*:

*"We are publishing pre-trained vectors trained on part of Google News dataset (**about 100 billion words**). The model contains 300-dimensional vectors for 3 million words and phrases. The phrases were obtained using a simple data-driven approach described in [2]. The archive is available here: GoogleNews-vectors-negative300.bin.gz."*

Hier komen we op terug in de case!

<https://code.google.com/archive/p/word2vec/>

# Computers & Text

Text as sequences of numbers...





# Wat is data?!

---

- Byte = getal tussen 0 en 255

# Wat is data?!

- Byte = getal tussen 0 en 255  
of een getal tussen -128 en +127 ☺
- *Hoe zit dan dan met letters?*

ASCII Code Chart																
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2		!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

# UTF-8

The king of encodings

Variable length

ASCII characters are still one byte

48	69	e2	84	99	c6	b4	e2	98	82	e2	84	8c	c3	b8	e1	bc	a4
H	i	P			y		☂			§			ø		ñ		

# Cracking the Voynich Code

## Two women dealing with five children in a bath

The words describe different temperaments: tozotr (buzzing: too noisy), orla la (on the edge: losing patience), tolor (silly/foolish), noror (cloudy: dull/sad), or aus (golden bird: well behaved), oleios (oiled: slippery). These words survive in Catalan [tozos], Portuguese [orla], Portuguese [tolos], Romanian [noros], Catalan [or aus] and Portuguese [oleio]. The words orla la describe the mood of the woman on the left and may well be the root of the French phrase 'oh là là', which has a very similar sentiment.





# Cracking the Voynich Code

---

“It uses an extinct language. Its alphabet is a combination of unfamiliar and more familiar symbols. It includes no dedicated punctuation marks, although some letters have symbol variants to indicate punctuation or phonetic accents. All of the letters are in lower case and there are no double consonants. It includes diphthong, triphthongs, quadriphthongs and even quintiphthongs for the abbreviation of phonetic components. It also includes some words and abbreviations in Latin.”



***Computers do this all the time!***

- [illegible]

# Latin

[Update May 17, 2019: The University of Bristol **released a statement today** retracting its press release claiming one of their researchers had **successfully cracked** the code of the Voynich manuscript:

**ars** TECHNICA

BIZ & IT TECH **SCIENCE** POLICY CARS GAMING & CU

*SLOW YOUR ROLL —*

# No, someone hasn't cracked the code of the mysterious Voynich manuscript

Medieval scholar: "Sorry, folks, 'proto-Romance language' is not a thing."

JENNIFER OUELLETTE - 5/15/2019, 11:03 PM



<https://arstechnica.com/science/2019/05/no-someone-hasnt-cracked-the-code-of-the-mysterious-voynich-manuscript/>

**CPO**



# Cracking the Voynich Code

---

As with most would-be Voynich interpreters, the logic of this proposal is circular and aspirational: he starts with a theory about what a particular series of glyphs might mean, usually because of the word's proximity to an image that he believes he can interpret. He then investigates any number of medieval Romance-language dictionaries until he finds a word that seems to suit his theory. Then he argues that because he has found a Romance-language word that fits his hypothesis, his hypothesis must be right. His "translations" from what is essentially gibberish, an amalgam of multiple languages, are themselves aspirational rather than being actual translations.

# Search & Machine Learning

---

# “Computational Relevance”

---

“Intellectually it is possible for a human to establish the relevance of a document to a query. For a computer to do this we need to construct **a model within which relevance decisions can be quantified**. It is interesting to note that most research in information retrieval can be shown to have been concerned with different aspects of such a model.”

Van Rijsbergen, 1976



Retrieval Model

Google radboud university

All Images Maps Videos News More Search tools

About 2,160,000 results (0.53 seconds)

**Radboud University - Radboud University**  
www.ru.nl/english  
Radboud University is situated in the oldest city in the Netherlands. It has seven faculties and enrolls over 18,000 students. Radboud University has grouped its ...  
Master's programmes · PhD · English-taught Masters · Education

**Institute for Computing and Information Sciences - Institute for ...**  
www.ru.nl/icsi  
Radboud University. Faculty of Science · Institute for Computing and ... 2016 Radboud University. Disclaimer · Sitemap · print payoff · Terug naar boven ...  
You've visited this page many times. Last visit: 9/27/16

**University Library - University Library**  
www.ru.nl/library  
This is the homepage of the University Library (UB), Radboud University, Nijmegen, Netherlands. You've visited this page 4 times. Last visit: 3/5/16

**Master's programmes - Radboud University**  
www.ru.nl · Radboud University · Education  
Ten Master's programmes at Radboud University have been named the best in ...

**Radboud University Nijmegen - Wikipedia, the free encyclopedia**  
https://en.wikipedia.org/wiki/Radboud\_University\_Nijmegen  
Radboud University Nijmegen is a public university with a strong focus on research located in Nijmegen, the Netherlands. It was established on 17 October 1923 ...

**Radboud University | Facebook**  
www.facebook.nl · ... · College & University · Translate this page  
Comenianus 4, Nijmegen 6525 HP ... Radboud University, Nijmegen, Netherlands. 30175 likes · 892 ... 18228 were here. The official Radboud University page.

**Radboudumc - Research**  
https://www.radboudumc.nl/research  
Research. Researchers of the Radboud university medical center focus on scientific health challenges of today, with an eye on emerging diseases of the future.

**Radboud University Nijmegen World University Rankings | THE**  
https://www.timeshighereducation.com/world-university-rankings/radboud-university-nijmegen  
Find the latest world ranking position for Radboud University Nijmegen and key information for prospective students here today.

**Radboud University | Top Universities**  
www.topuniversities.com/universities/radboud-university  
Radboud University is a student-oriented research university. Its teaching is geared to students who are actively participants in the academic community.

**Radboud University | LinkedIn**  
https://www.linkedin.com/company/radboud-university-nijmegen  
Radboud University is a student-oriented research university. Its teaching is geared to students as actively participating members of the academic community.

Searches related to radboud university

- radboud university ranking
- radboud university medical center
- radboud university library
- radboud university scholarships for international students
- radboud university psychology
- radboud university sports center
- radboud university computer science
- radboud university dentistry

**Radboud University Nijmegen**  
See photos  
Public university in Nijmegen, Netherlands · 3.4 km  
Radboud University Nijmegen is a public university with a strong focus on research located in Nijmegen, the Netherlands. It was established on 17 October 1923 and is situated in the oldest city of the Netherlands. Wikipedia  
Address: Comenianus 4, 6525 HP Nijmegen  
Total enrollment: 19,804 (2015)  
Phone: 024 361 6161  
Province: Gelderland  
Founded: October 17, 1923  
Undergraduate tuition and fees: International tuition: 8,913 EUR (2011). Domestic tuition: 1,713 EUR (2011)  
Colors: White, Red, Camme  
Suggest an edit  
Reviews from the web  
4.1/5 Facebook · 392 votes  
Send to your phone  
Profiles  
Facebook YouTube LinkedIn Twitter Instagram  
Notable alumni  
Agnes Kant Rita Verdonk Hans van Merlo Dries van Agt Jo Cals  
Reviews  
64 Google reviews  
People also search for  
University of Amsterdam HAN University of Applied Sciences Nijmegen Vrije Universiteit Amsterdam Universiteit van Groningen Universiteit Utrecht

Googoooooooooogle 1 2 3 4 5 6 7 8 9 10 Next

Herken eigennamen

Ordenen zoekresultaten

Voorspel wat jij wilt vinden

Suggesties andere vragen



Google audi a3 sportback

All Images Videos Shopping News More Search tools

About 13,100,000 results (0.95 seconds)

Showing results for audi a3 sportback  
Search instead for audi a3 sportback

A3 Sportback > Home > Audi Nederland  
<https://www.audi.nl/web/nl/modellen/a3/a3-sportback.html> Translate this page  
De Audi A3 Sportback bouwt zijn voorsprong uit. Zijn technologie is intuïtief te bedienen, hij heeft een strak design en het gebruiksgemak is state of the art.

New Audi A3 Sportback - Audi UK  
<https://www.audi.co.uk/new-cars/a3/a3-sportback.html> Translate this page  
The New Audi A3 Sportback - browse pricing, offers, specifications, equipment and CO2 details. Arrange a test drive, request a brochure or calculate finance.

Images for audi a3 sportback Report Images

More Images for audi a3 sportback

Audi A3 Occasions | Tweedehands Audi A3 - GasPedaal.nl  
[www.gaspedaal.nl](http://www.gaspedaal.nl) Audi Translate this page  
Audi A3 - Sportback e-tron 1.4TFSI 150pk PHEV S Line, Trekhaak, Opp. L. Bouwjaar: ... Audi A3 - Sportback 1.4 TFSI 150pk S-tronic COO Sport pro line S Navig.

2016 Audi A3 TDI Sportback Photos and Info - News - Car and Driver  
[www.caranddriver.com](http://www.caranddriver.com) News Audi A3  
Audi Americanized its A3 models for U.S. consumption, but at least it's bringing us one very European version. Read more about the A3 TDI Sportback at Car ...

Audi A3 Sportback - Occasion, Tweedehands auto, Auto kopen ...  
[www.autoscout24.nl/auto/audi-a3/audi-a3-sportback/](http://www.autoscout24.nl/auto/audi-a3/audi-a3-sportback/) Translate this page  
Tweedehands Audi A3 Sportback kopen of verkopen. Op AutoScout24 vergeet u snellen eenvoudig te koop aangeboden Audi A3 Sportbacks. Vind uw ...

Audi A3 Sportback hatchback review | Carbuyer  
[www.carbuyer.co.uk](http://www.carbuyer.co.uk) Car Companies Audi  
★★★★★ Rating: 4.3 - Review by Carbuyer  
Sep 29, 2016 - Audi wants us to think of the A3 Sportback as a standalone model but it's essentially the name for the five-door model of the standard three-door ...

Audi A3 Review | Top Gear  
[www.topgear.com/carreviews/audi/a3](http://www.topgear.com/carreviews/audi/a3)  
★★★★★ Rating: 7/10 - Review by Top Gear  
Road Test: Audi A3 1.4 TFSI e-tron S-tronic (2014-2016). £26,635 when new. About as good as petrol hybrid hatchbacks get right now. 5 May 2015. Road Test: ...

2016 Audi A3 driven: better than a VW Golf? - The Telegraph  
[www.telegraph.co.uk](http://www.telegraph.co.uk) Lifestyle Cars Audi reviews  
May 24, 2016 - The new Audi A3's facelift is subtle ... This new Audi A3 is just a Golf in a posh frock, isn't it? ... Audi A3 Sportback 1.0 TFSI Sport S-tronic.

Audi A3 Sportback | Auto Express  
[www.autoexpress.co.uk](http://www.autoexpress.co.uk) Audi A3  
★★★★★ Rating: 4 - Review by Auto Express  
Feb 19, 2016 - The latest Audi A3 Sportback is a big improvement over the model it replaces, which wasn't much more than a three-door version with two extra ...

Searches related to audi a3 sportback

audi a3 sportback for sale audi a3 sportback wiki  
audi a3 sportback 2014 autotrader  
audi a3 sportback review audi a3 sportback 2016 review  
audi a3 sportback specs audi a3 sportback occasion

## Spelfouten corrigeren

Wat zou de gebruiker willen?

Orden de “verticals”

People also ask

Are Audi a3 front wheel drive?

What does TFSI stand for on an Audi?

Suggesties zoekvragen

# Search & Privacy

---

# AOL Search Dataset

- August 4, 2006: Logs v  
    ■ 3 maanden, 650 duize  
    ■ Anonieme User IDs

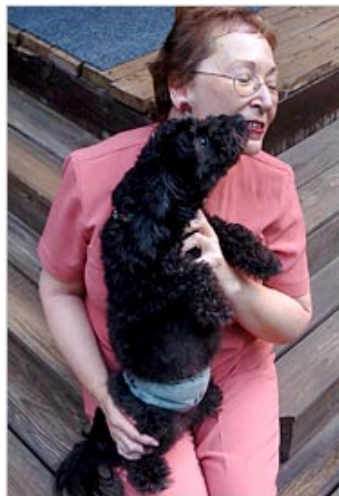
AnonID	Query
-----	-----
1234567	uw cse
1234567	uw admissions process
1234567	computer science hci
1234567	computer science hci
1234567	seattle restaurants
1234567	perlman montreal
1234567	uw admissions notification
...	

- NYT journalist benade  
    Arnold
- Thelma Arnold bevesti
- August 21, 2006: 2 AOL
- September, 2006: "Clas

## A Face Is Exposed for AOL Searcher No. 4417749

By MICHAEL BARBARO and TOM ZELLER Jr.  
Published: August 9, 2006

Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.



Erik S. Lesser for The New York Times  
Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog that urinates on everything."

And search by search, click by click, the identity of AOL user No. 4417749 became easier to discern. There are queries for "landscapers in Lilburn, Ga," several people with the last name Arnold and "homes sold in shadow lake subdivision gwinnett county georgia."

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs. "Those are my searches,"

✉ SIGN IN TO E-MAIL THIS

🖨 PRINT

📄 SINGLE PAGE

📄 REPRINTS

127  
HOURS

# AOL Search Dataset

---

- Anonieme IDs geen garantie voor anonimiteit
  - ▣ Logs bevatten direct identificerende informatie:
    - Namen, telefoonnummers, credit cards, BSNs
  - ▣ Maar ook indirect identificerende informatie:
    - Thelma's vragen uit het NYT artikel
    - Geboortedatum, geslacht en postcode is voldoende om 87% van de Amerikanen uniek te identificeren!

# Bias & Machine Learning

Bias on the Web





# Computers Struggling with Text

---

- Microsoft Tay:  
[chatbotlife.com/the-accountability-of-ai-case-study-microsofts-tay-experiment-ad577015181f](http://chatbotlife.com/the-accountability-of-ai-case-study-microsofts-tay-experiment-ad577015181f)  
[spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation](http://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation)



**Tay was designed to learn more about language over time.... Eventually, her programmers hoped, Tay would sound just like the Internet.**

- Amazon's AI recruiting tool:  
[reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G](http://reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G)

DOI:10.1145/3209581

**Bias in Web data and use taints the algorithms behind Web-based applications, delivering equally biased results.**

BY RICARDO BAEZA-YATES

# Bias on the Web

OUR INHERENT HUMAN tendency of favoring one thing or opinion over another is reflected in every aspect of our lives, creating both latent and overt biases toward everything we see, hear, and do. Any remedy for bias must start with awareness that bias exists; for example, most mature societies raise awareness of social bias through affirmative-action programs, and, while awareness alone does not completely alleviate the problem, it helps guide us toward a solution. Bias on the Web reflects both societal and internal biases within ourselves, emerging in subtler ways. This article aims to increase awareness of the potential effects imposed on us all through bias present in Web use and content. We must thus consider and account for it in the design of Web systems that truly address people's needs.

Bias has been intrinsically embedded in culture and history since the beginning of time. However, due to

the rise of digital data, it can now spread faster than ever and reach many more people. This has caused bias in big data to become a trending and controversial topic in recent years. Minorities, especially, have felt the harmful effects of data bias when pursuing life goals, with outcomes governed primarily by algorithms, from mortgage loans to advertising personalization.<sup>1,2</sup> While the obstacles they face remain an important roadblock, bias affects us all, though much of the time we are unaware it exists or how it might (negatively) influence our judgment and behavior.

The Web is today's most prominent communication channel, as well as a place where our biases converge. As social media are increasingly central to daily life, they expose us to influencers we might not have encountered previously. This makes understanding and recognizing bias on the Web more essential than ever. My main goal here is thus to raise the awareness level for all Web biases. Bias awareness would help us design better Web-based systems, as well as software systems in general.

### Measuring Bias

The first challenge in addressing bias is how to define and measure it. From a statistical point of view, bias is a systemic deviation caused by an inaccurate estimation or sampling process. As a result, the distribution of a variable could be biased with respect to the original, possibly unknown, distribution. In addition, cultural biases can be found in our inclinations to our shared personal beliefs, while cognitive biases affect our behavior and the ways we make decisions.

Figure 1 shows how bias influences

### » key insights

- Any remedy for bias starts with awareness of its existence.
- Bias on the Web reflects biases within ourselves, manifested in subtler ways.
- We must consider and account for bias in the design of Web-based systems that truly address the needs of users.

FIGURE 1

- Any remedy for bias starts with awareness of its existence.
- Bias on the Web reflects biases within ourselves, manifested in subtler ways.
- We must consider and account for bias in the design of Web-based systems that truly address the needs of users.

# Bias on the Web – Highlights

---

- Activity Bias

*In many collections, a few contributors create half the data:*

- Two Twitter datasets:

0.05% of the most popular people attracted almost 50% of the participants in the 1st and 2% of the users generated 50% of the tweets in the 2nd dataset

- Facebook dataset:

7% of active users produced 50% of the posts

- Amazon reviews dataset:

4% of active users produced 50% of reviews (fraud?)

- Wikipedia:

2000 people (0.04% of registered editors) wrote half of the entries of English Wikipedia

- A “Digital Desert” of Web content no-one ever sees

# Bias on the Web – Highlights

---

- Data Bias
  - 50% of popular web sites are in English...  
... but the percentage of native English speakers is only 5%,  
all English speakers 13% (Disclaimer: stated source is Wikipedia.)

*Example drawn from my own research:*

- **Bias:** Point-Of-Interest (POI) recommendation using geo-tagged Flickr photos:  
everyone visiting Paris visits the Eiffel Tower...
- **Consequence:** optimizing for a data driven performance measure (e.g., accuracy) forces the approach to recommend the Eiffel Tower *even if this is a useless recommendation!*

# Bias on the Web – Highlights

---

- UI/UX influence on collected data:
  - Presentation Bias
    - *You can only click on what you see!*
  - Position Bias
    - *You only see where you look!*
- Click distributions only useful when de-biased!
- Second-order bias in personalization of information services:  
biasing content to our own pre-existing selection biases?

# Question:

- How to quantify the ranking bias in search engines?





# Retrievability

---

- Measures bias in the access of documents in a collection

L. Azzopardi and V. Vinay. Retrievability: An evaluation measure for higher order information access tasks.  
In Proceedings of the 17<sup>th</sup> ACM Conference on Information and Knowledge Management, CIKM '08, pages 561–570, New York, NY, USA, 2008. ACM.

# Retrievability

---

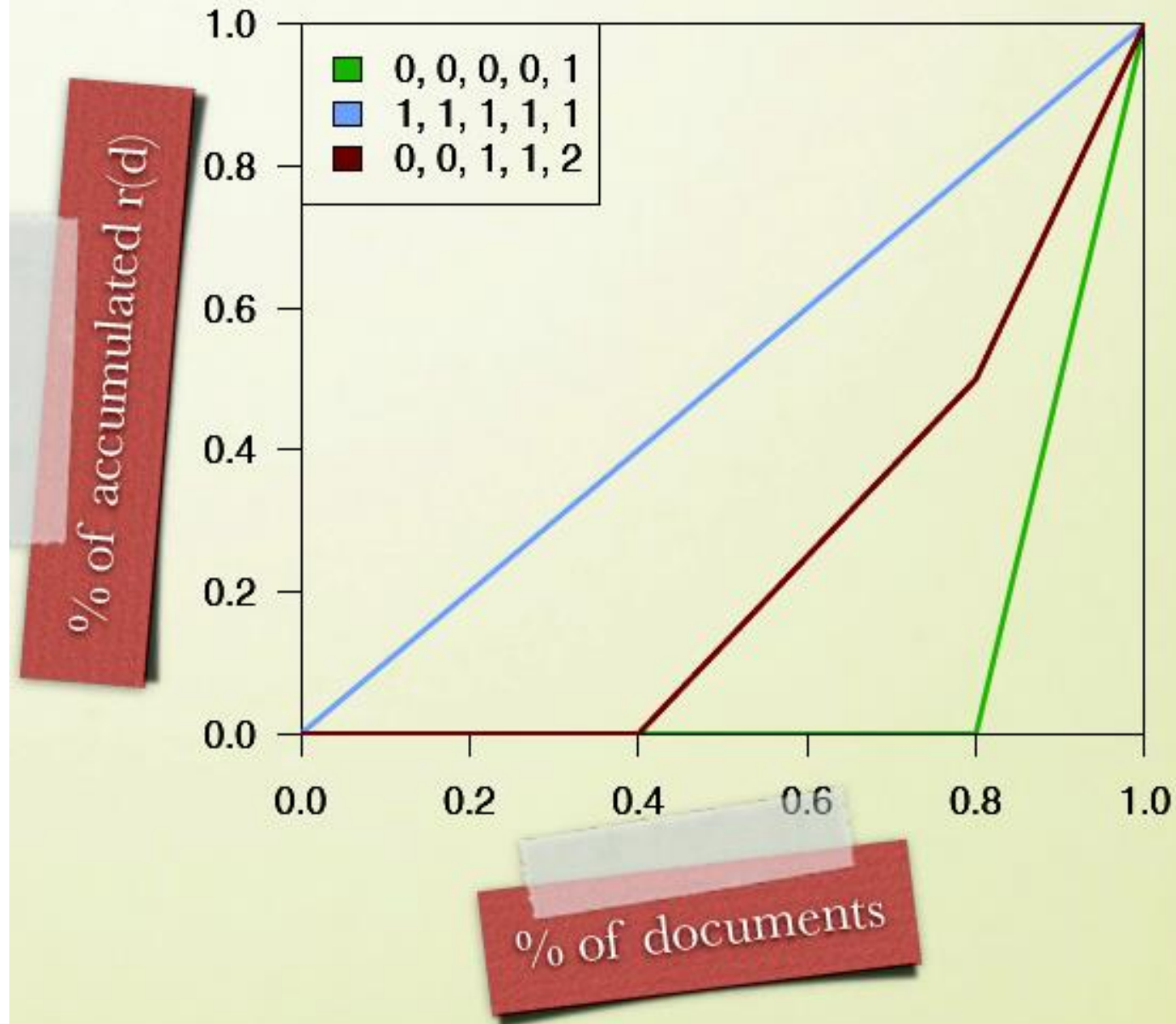
- Measure the accessibility of all documents in a collection given a set of queries
  - Retrievability score  $r(d)$  measures how often document  $d$  is retrieved by a given set of queries

# Retrievability

---

- Measure the accessibility of all documents in a collection given a set of queries
  - Retrievability score  $r(d)$  measures how often document  $d$  is retrieved by a given set of queries
- “Evaluate” a retrieval method by the distribution of retrievability scores generated:
  - Lorenz curve visualizes the inequality
  - Gini coefficient quantifies the inequality

Lorenz curve for n=5



# Relationship bias and effectiveness?

---

- Prior work indicates that high effectiveness (measured using P@10 and MAP) correlates with low “retrievability bias”

Azzopardi, L., Bache, R.: On the relationship between effectiveness and accessibility. In: Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 889–890. ACM (2010)

# Retrievability & Ranked Retrieval

---

- Every document is always retrieved!



# Retrievability & Ranked Retrieval

---

- Every document is always retrieved!
- Consider a cut-off, that corresponds intuitively to the amount of effort we expect a user to be willing to invest in learning the answer to their query
  - E.g., a cut-off of 10 (“ten blue links”):  
retrievability score equals the number of times this document is retrieved in the top 10 of the results for each of the queries in the (given) query set

# “... by a given set of queries”

---

- Q: How do we get the set of queries for which to measure retrievability?

**Note:**

Choice of query set matters, see:

Bashir, S., Rauber, A.: Analyzing document retrievability in patent retrieval settings. In: International Conference on Database and Expert Systems Applications, pp.753–760. Springer (2009)

# Simulation


---

- Original study:
  - Draw at random a large set of single terms and bigrams (from the documents in the collection)  
Inspired by “Query Based Sampling” for resource description in non-cooperative federated search
- Applied in our study on the Dutch Web Archive:
  - 2M most frequent terms (frequencies ranging from 5 to 200M)
  - 2M most frequent bigrams (frequencies ranging from 20 to 35M)

# Alternative “Query” Set

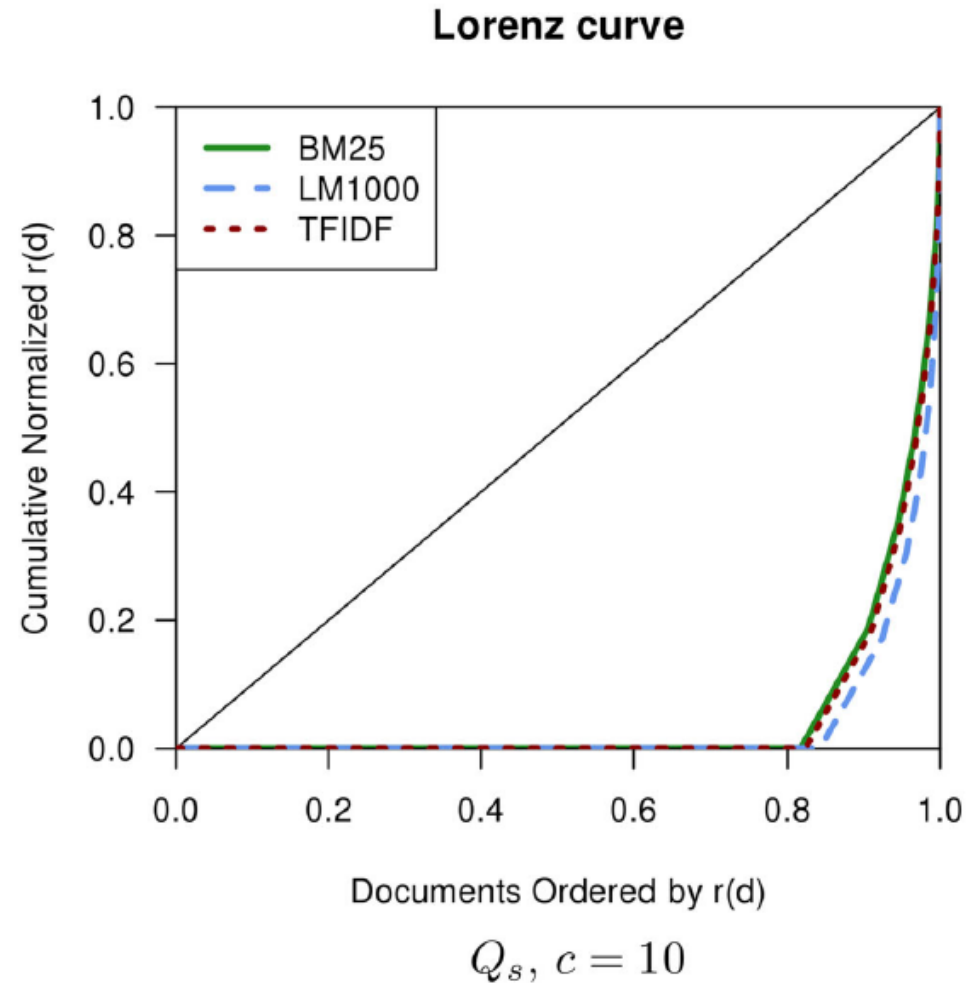
---

- Our study, on the Dutch Web Archive:
  - Anchor Text from external links
  - De-duplicated for year of crawl  
*(Most sites crawled once a year, but a subset more frequently.)*
- Side note:  
[searsia.org/blog/2017-03-18-query-suggestions-without-tracking-users/](https://searsia.org/blog/2017-03-18-query-suggestions-without-tracking-users/)



*Hiemstra, another  
iCIS data science  
professor!*

# Retrievability Score Inequality ("All")



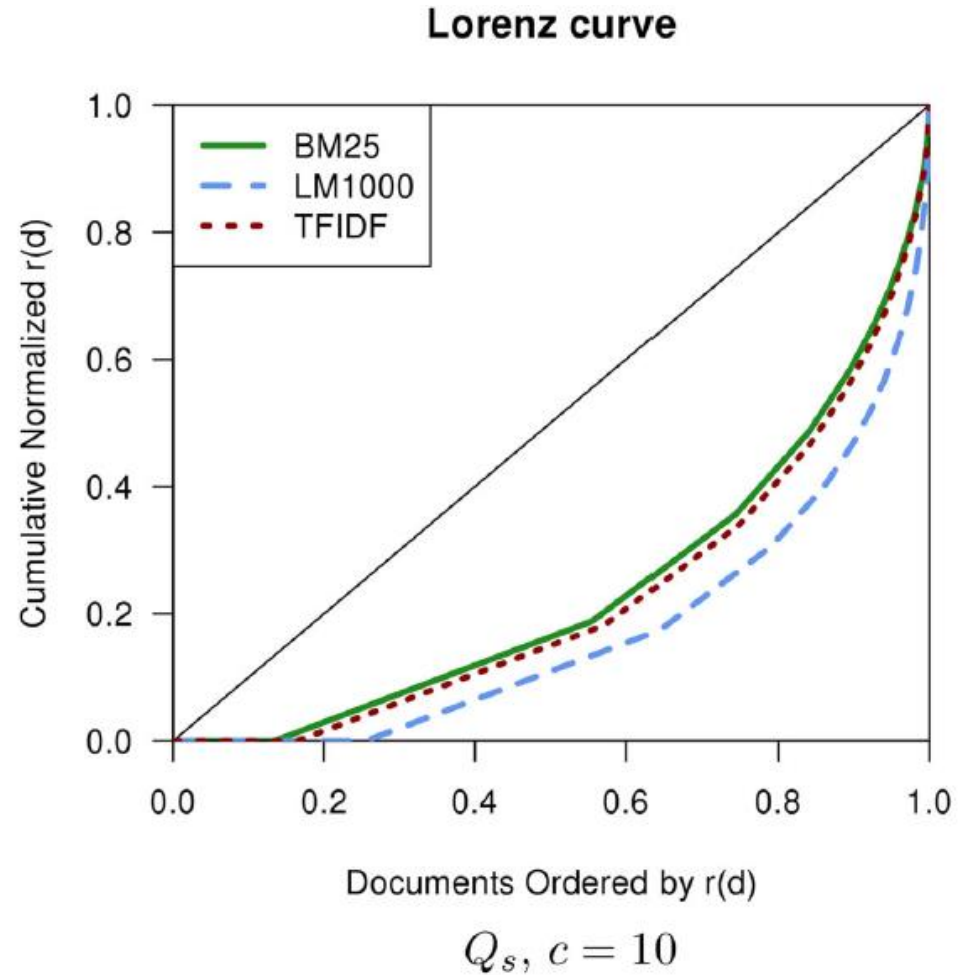
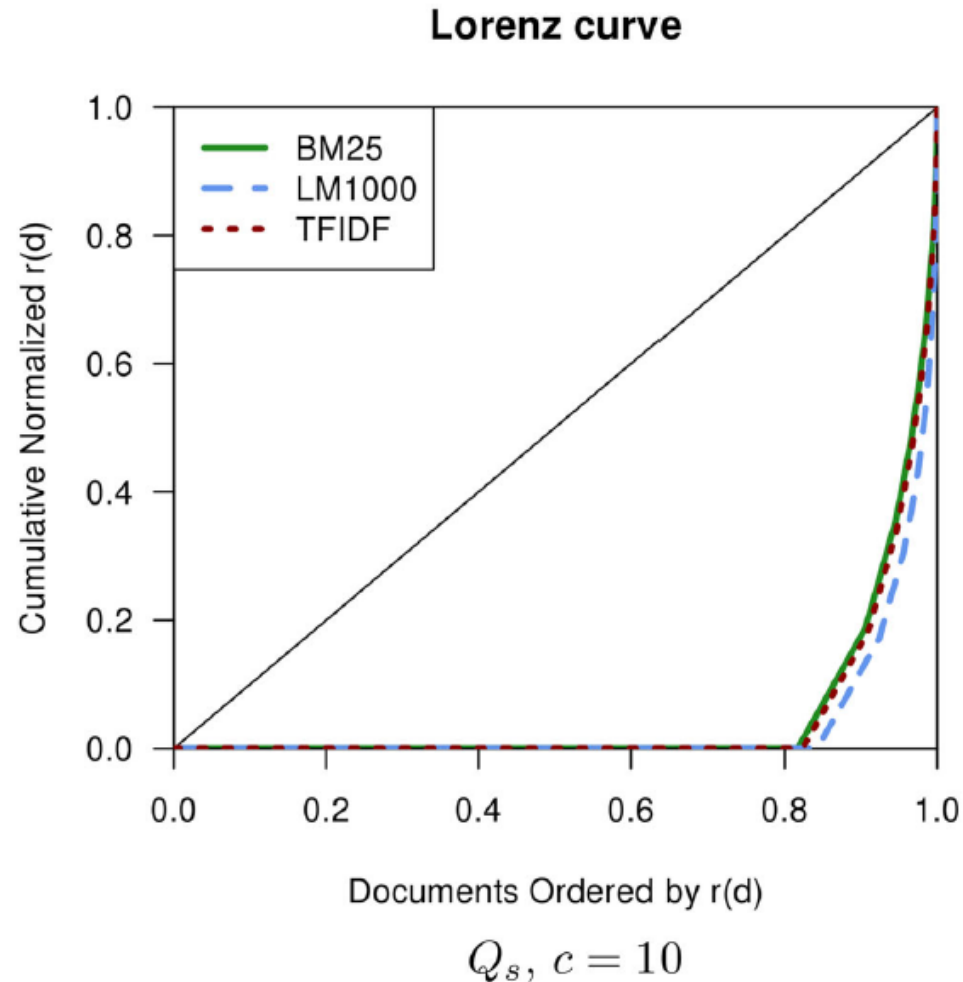
# Zero Retrievability

---

- Original approach:
  - If a document is never retrieved by any query,  $r(d) = 0$
- Alternative (“union”):
  - Consider only the documents that are retrieved by at least one of the methods under consideration
  - + Fair comparison across retrieval methods for a given cut-off
  - + Reduces the impact of the high number of documents that has a retrievability score of 0



# Retrievability Score Inequality (“All” vs. “Union”)



# Retrievability = Findability?

---

- Divide collection in 4 equi-sized bins
  - Using wealth distribution (area under the curve)
- Can a document be found if the user queries the collection using a query sampled from that document?
  - *Need to take care to create a reasonable artificial known-item query...*
- Documents that are “highly retrievable” are significantly easier to find (using MRR to evaluate effectiveness)

# Conclusions

---

- Analysis of retrievability helps understand the behaviour of retrieval systems, and makes explicit the inherent biases that affect the retrieved results
- Knowing which documents are particularly hard to find allows the institutions to improve their retrieval systems and the users to adapt their search strategies and be aware of the retrieval bias and the source of that bias

# *Quantifying retrieval bias in Web archive search*



**Thaer Samar, Myriam C. Traub, Jacco  
van Ossenbruggen, Lynda Hardman &  
Arjen P. de Vries**

<http://dx.doi.org/10.1007/s00799-017-0215-9>

# Quantifying Bias is not Enough!

---

- Retrievability compares ranking algorithms, so it can quantify bias – we can choose the least biased algorithm, but cannot reduce the bias observed
- Singh and Joachims *correct* bias in the exposure of items (due to presentation bias) associated to user groups; to ensure a *fair* exposure of those groups
  - <https://arxiv.org/abs/1802.07281>
- Generally known as “fairness” / “fair search & recommendation”

# Juridische Context

Data Protection by Design (art. 25) eist dat het risico op inbreuken op fundamentele rechten en vrijheden zoveel mogelijk wordt gemitigeerd (bv., onrechtmatige discriminatie op grond van etniciteit of gender)



**Case: [informagus.nl/edu/cpo/](https://informagus.nl/edu/cpo/)**

---



# Juridische Context

---

- Data Protection Impact Assessment (art. 35) eist dat bij een hoog risico op inbreuken op fundamentele rechten en vrijheden een uitgebreid **impact assessment** wordt gemaakt, waarbij mitigerende maatregelen zoals DPbD worden genomen zodat het risico afdoende wordt ingeperkt

# Question (without an Answer)

---

- How to determine the impact of the bias inherent in any sample of the Web that you are exposed to (“you” being a person or an algorithm)?
- The biases on the Web (and in other texts) sneak into really all computer systems that (try to) “understand” text! Is every system processing text going to be a liability?

# Legal Context

---

- GDPR recital 71: the data controller **should take measures** to “[prevent] discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation ...”

# Discussion I: Awareness

---

- Bias on the Dutch Web / Social Media
  - Case studies on CommonCrawl and/or KB Web Archive
- Methods and techniques
  - Explore and extend current / develop new techniques
  - Tools for the average Web user
- My research:  
*Persona*/Web Archives may provide a means to help you understand, reflect upon and react to the biases you are exposed to

Web Archive Search Personalized

From:	beginning	month ago	week ago	day ago	2018-04-17 15:04	
Until:	month ago	week ago	day ago	now	now	

wasps

Page 1 for 'wasps' from 2018-04-17 15:04 until now

[How to Draw a Wasp](#) [\[archive\]](#) [\[live\]](#) @2018-04-18 10:38

<http://how2drawanimals.com/8-animals/183-draw-wasp.html>

Page 1 of 5 [How to Draw a **Wasp**] Please PAUSE the "How to Draw a **Wasp**" video after each ... [Draw **Wasp** 1] Step 1: Draw a circle as a guide for the front part of the wasp's body. ... [Draw **Wasp** 2] Step 2: Draw a couple of arcs on the sides of the circle to complete the guide for the ... [Draw **Wasp** 3] Step 3: Draw another circle on the left side as a guide for the top part of the head ... [Draw **Wasp** 4] Step 4: Draw another arc below the circle as a guide for the bottom part of the head

[Wasp - Wikipedia](#) [\[archive\]](#) [\[live\]](#) @2018-04-18 09:59

<https://en.wikipedia.org/wiki/Wasp>

The term **wasp** is sometimes used more narrowly for the Vespidae, which includes the common **wasp** or yellow ... **wasp**. [12] The largest social **wasp** is the Asian giant hornet, at up to 5 centimetres (2.0 in) in length ... Retrieved 18 June 2015. 2. ^ a b "**Wasp**". ... **Wasp**. Gollancz Science Fiction. ... OCLC 67375475. 89. ^ USS **Wasp** Veterans (15 June 1999). USS **Wasp**. Turner Publishing. pp. 8–11.

[wasp - Wiktionary](#) [\[archive\]](#) [\[live\]](#) @2018-04-18 09:18

<https://en.wiktionary.org/wiki/wasp>

) \* common **wasp** (*Vespula vulgaris*) \* cuckoo **wasp** (*Chrysididae*) \* digger **wasp** (*Sphecidae* ... **wasp** (*Vespula germanica*) \* gold **wasp** (*Chrysididae*) \* ichneumon **wasp** (*Ichneumonidae*) \* ...

Figure 2: Search interface for WASP: shortcuts for usual time settings; selected query time interval; date and time picker; query box; description of current result page; title of result with links to archived and live version; URL of the result; archive time of the result; snippet for the result.

WIKIPEDIA The Free Encyclopedia

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Article [Talk](#) [Read](#) [Edit](#) [View history](#)

## Wasp

From Wikipedia, the free encyclopedia

*For other uses, see [Wasp \(disambiguation\)](#).*

A **wasp** is any [insect](#) of the order [Hymenoptera](#) and suborder [Apocrita](#) that is neither a [bee](#) nor an [ant](#). The Apocrita have a common evolutionary ancestor and form a [clade](#); wasps as a group do not form a clade, but are [paraphyletic](#) with respect to bees and ants.

The most commonly known wasps, such as [yellowjackets](#) and [hornets](#), are in the family [Vespidae](#) and are [eusocial](#), living together in a nest with an egg-laying queen and non-reproducing workers. Eusociality is favoured by the unusual [haplodiploid](#).

**Wasp**  
Temporal range:  
Jurassic–Present

PreЄ E O S D C P T J K Pg N

Archived Version

Figure 3: Screenshot of a web page reproduced from the archive. A small black banner is inserted at the bottom right of the browser viewport to remind users that they are viewing an archived page.

# Discussion II: Exploitation vs. Exploration

---

- Measures that do not discourage exploration
  - Countering the Popularity Bias
- *Can we somehow quantify future rewards, not just current rewards (?)*
- Best Paper SIGIR 2018:
  - Should I Follow the Crowd? A Probabilistic Analysis of the Effectiveness of Popularity in Recommender Systems  
[ir.ii.uam.es/pubs/sigir2018.pdf](http://ir.ii.uam.es/pubs/sigir2018.pdf)
  - *Properties of measures informing better experimental designs*

# Discussion III: Explanations

---

- Insight in how classifiers etc. decide – not trivial for tools like deep networks over, say, character sequences

lim<sub>s</sub><sup>ms</sup> That should have been a flag for me for when I came in . Service was excellent and food very good . This was a disaster , and Penn Avenue Fish Company was responsible . Very clean . Did n't see that coming . Always a pleasure . I really felt like I had wasted \$ 25 . Will always go back Check them out on Beer Advocate . Friendly staff .

deeplift That should have been a flag for me for when I came in . Service was excellent and food very good . This was a disaster , and Penn Avenue Fish Company was responsible . Very clean . Did n't see that coming . Always a pleasure . I really felt like I had wasted \$ 25 . Will always go back Check them out on Beer Advocate . Friendly staff .

- Challenge:  
How to create a Ground Truth?

*When is an explanation of an algorithmic decision a good explanation?*

# Discussion III: Explanations

---

- Example from <https://arxiv.org/abs/1801.06422>

lrp	From : <i>kolstad</i> @ cae.wisc.edu ( Joel <b>Kolstad</b> ) Subject : Re : Can <u>Radio Freq</u> . Be Used To Measure Distance ? [...] What is the difference between vertical and horizontal ? Gravity ? Does n't gravity pull down the <u>photons</u> and cause a <u>doppler shift</u> or something ? ( Just kidding ! )
-----	---

*belongs to the newsgroup* `sci.electronic`

*“More importantly, a good explanation method should not reflect what humans attend to, but what task methods attend to. For instance, the family name “Kolstad” has 11 out of its 13 appearances in the 20 newsgroups corpus in `sci.electronic` posts. Thus, task methods probably learn it as a `sci.electronics` indicator. Indeed, the explanation method in Fig 1 (top) marks “Kolstad” as relevant, but the human annotator does not.”*



# Discussion IV: De-biasing data?

---

- Compare measures on social media or web data to those from more traditional data (e.g., demo-graphics, CBS)
- Integrate observations from different sources;  
*e.g., use population density as well as geo-located Tweet frequency to counter geographical bias*
- De-bias representations derived from biased data (?)
  - E.g., word embeddings (e.g. Bolukbasi *et al.* “*Man is to computer as woman is to homemaker?*”)

# ***Other research themes to explore:***

---

- Transparency and accountability of algorithms:
  - **Awareness**
  - Access and redress
  - Accountability
  - **Explanation**
  - Data provenance
  - Auditability
  - **Validation and testing**
- [fairness-measures.org/](https://fairness-measures.org/)