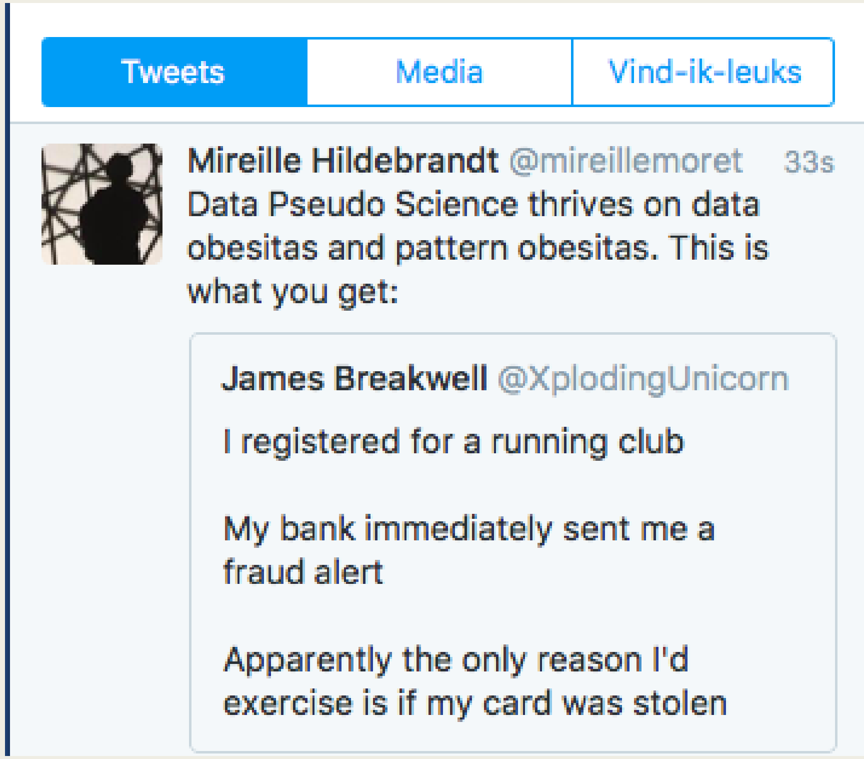
"Go ahead and think that I'm not really thinking. I thought you would think that."

GETTING *DATA PROTECTION* RIGHT

Prof. dr. Mireille Hildebrandt
Interfacing Law & Technology Vrije Universiteit Brussel
Smart Environments, Data Protection & the Rule of Law Radboud University



Mireille Hildebrandt @mireillemoret 33s

Data Pseudo Science thrives on data obesitas and pattern obesitas. This is what you get:

James Breakwell @XplodingUnicorn

I registered for a running club

My bank immediately sent me a fraud alert

Apparently the only reason I'd exercise is if my card was stolen

what's next?

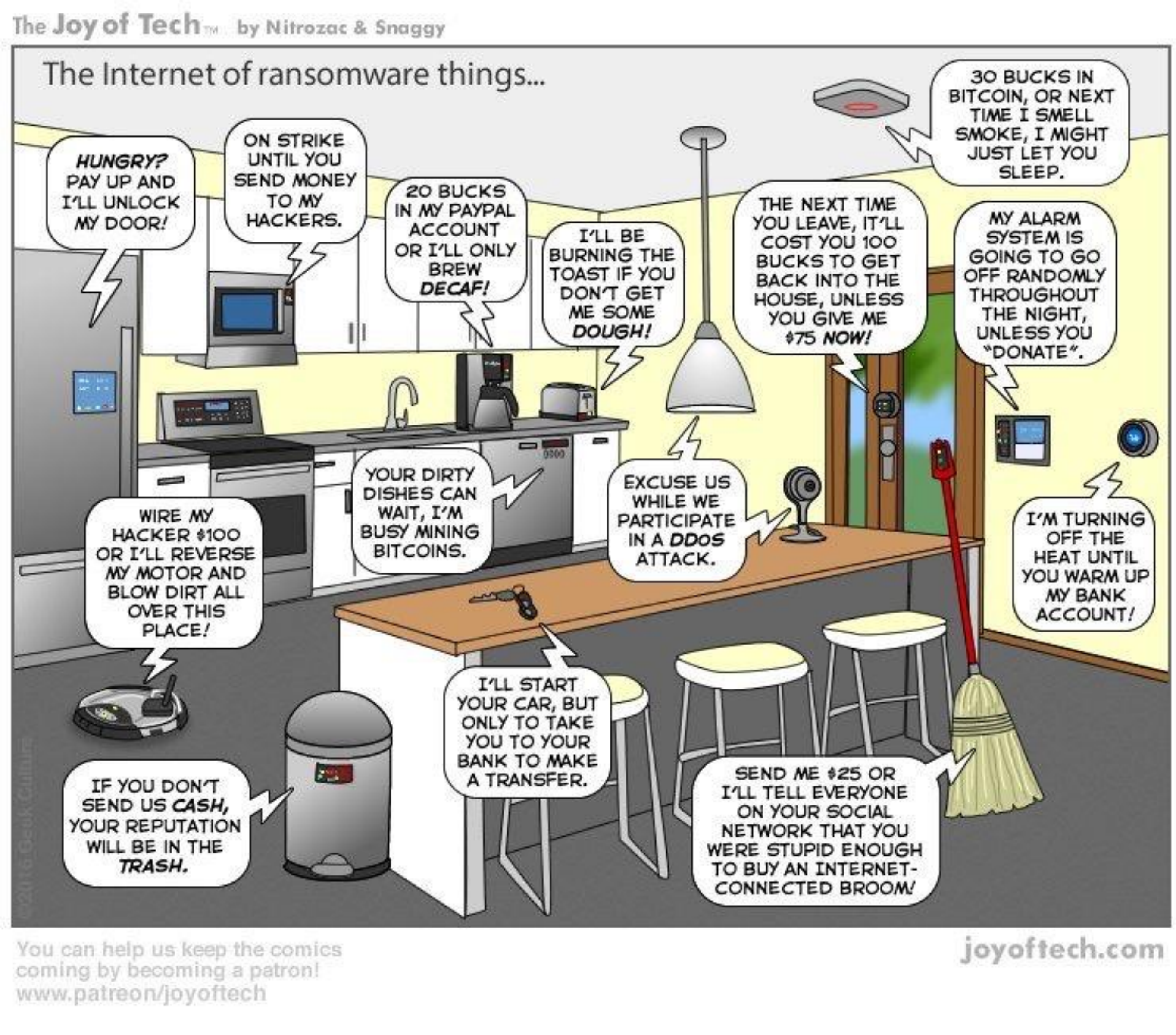
1. From online to onlife
2. Machine Learning
3. Data Protection

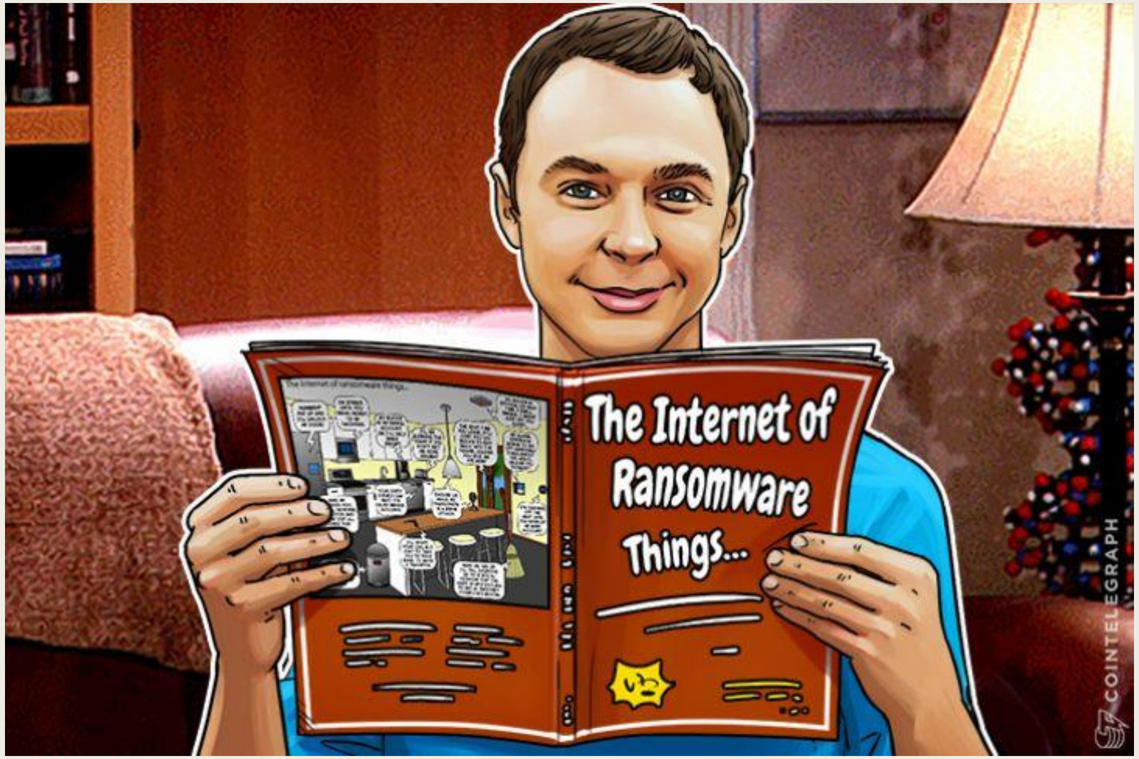
what's next?

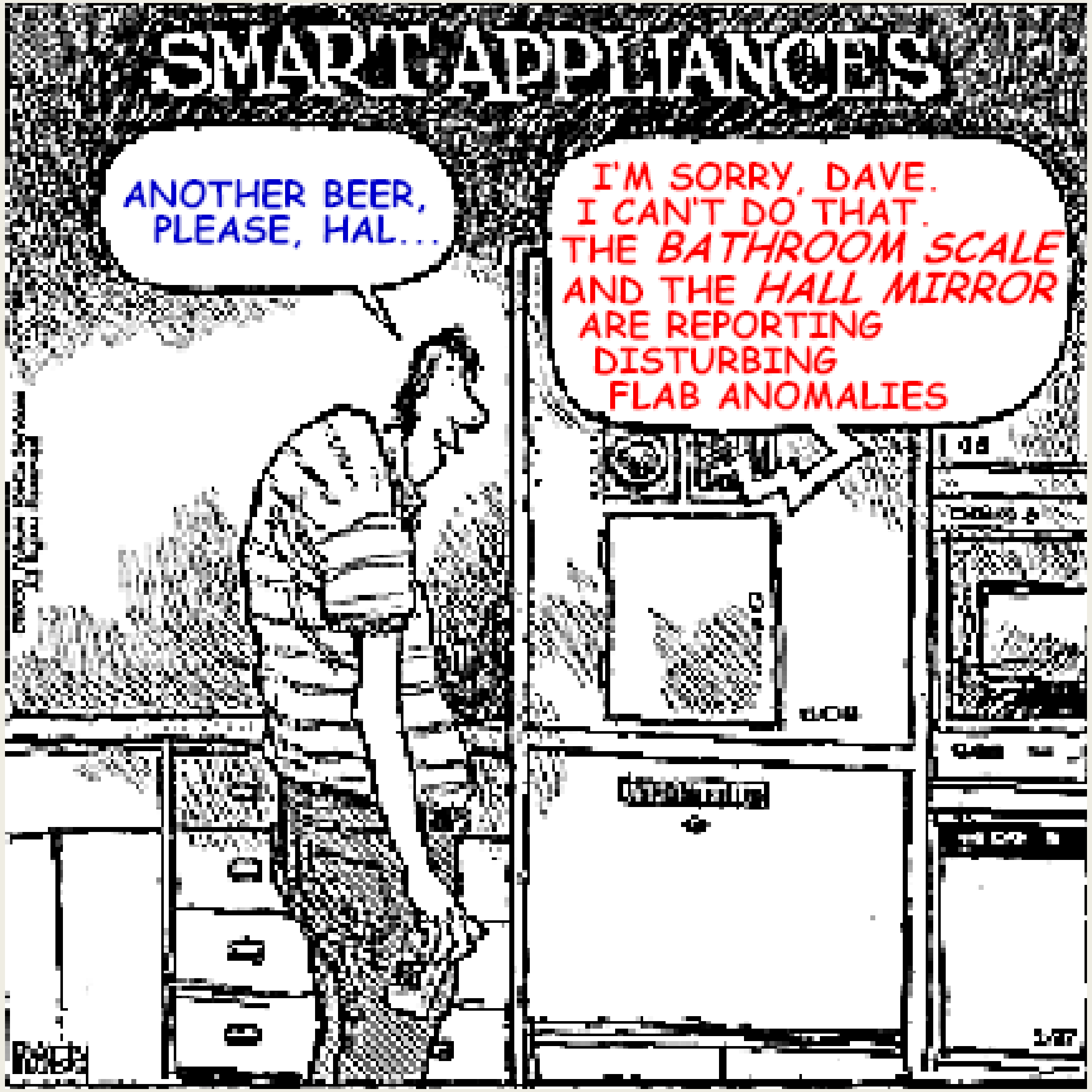
1. From online to onlife

online → onlife

- internet: packet switching & routing, network structure,
- world wide web: hyperlinking
- search engines, blogs, social media, web portals
- web platforms [network effects & filter bubbles; reputation & fake news]
- mobile applications [moving towards IoT, wearables]
- IoT: cyberphysical infrastructures [connected cars, smart energy grids]
- cloud computing, fog computing & edge computing


The Joy of Tech™ by Nitrozac & Snaggy
The Internet of ransomware things...
HUNGRY? PAY UP AND I'LL UNLOCK MY DOOR!
ON STRIKE UNTIL YOU SEND MONEY TO MY HACKERS.
20 BUCKS IN MY PAYPAL ACCOUNT OR I'LL ONLY BREW DECAF!
I'LL BE BURNING THE TOAST IF YOU DON'T GET ME SOME DOUGH!
THE NEXT TIME YOU LEAVE, IT'LL COST YOU 100 BUCKS TO GET BACK INTO THE HOUSE, UNLESS YOU GIVE ME $75 NOW!
30 BUCKS IN BITCOIN, OR NEXT TIME I SMELL SMOKE, I MIGHT JUST LET YOU SLEEP.
MY ALARM SYSTEM IS GOING TO GO OFF RANDOMLY THROUGHOUT THE NIGHT, UNLESS YOU "DONATE".
YOUR DIRTY DISHES CAN WAIT, I'M BUSY MINING BITCOINS.
EXCUSE US WHILE WE PARTICIPATE IN A DDOS ATTACK.
I'M TURNING OFF THE HEAT UNTIL YOU WARM UP MY BANK ACCOUNT!
WIRE MY HACKER $100 OR I'LL REVERSE MY MOTOR AND BLOW DIRT ALL OVER THIS PLACE!
I'LL START YOUR CAR, BUT ONLY TO TAKE YOU TO YOUR BANK TO MAKE A TRANSFER.
IF YOU DON'T SEND US CASH, YOUR REPUTATION WILL BE IN THE TRASH.
SEND ME $25 OR I'LL TELL EVERYONE ON YOUR SOCIAL NETWORK THAT YOU WERE STUPID ENOUGH TO BUY AN INTERNET-CONNECTED BROOM!
©2016 Geek Culture
You can help us keep the comics coming by becoming a patron! www.patreon/joyoftech
joyoftech.com


The Internet of ransomware things...
The Internet of Ransomware Things...
COINTELEGRAPH


SMART APPLIANCES
ANOTHER BEER, PLEASE, HAL...
I'M SORRY, DAVE. I CAN'T DO THAT. THE BATHROOM SCALE AND THE HALL MIRROR ARE REPORTING DISTURBING FLAB ANOMALIES

onlife: data driven agency

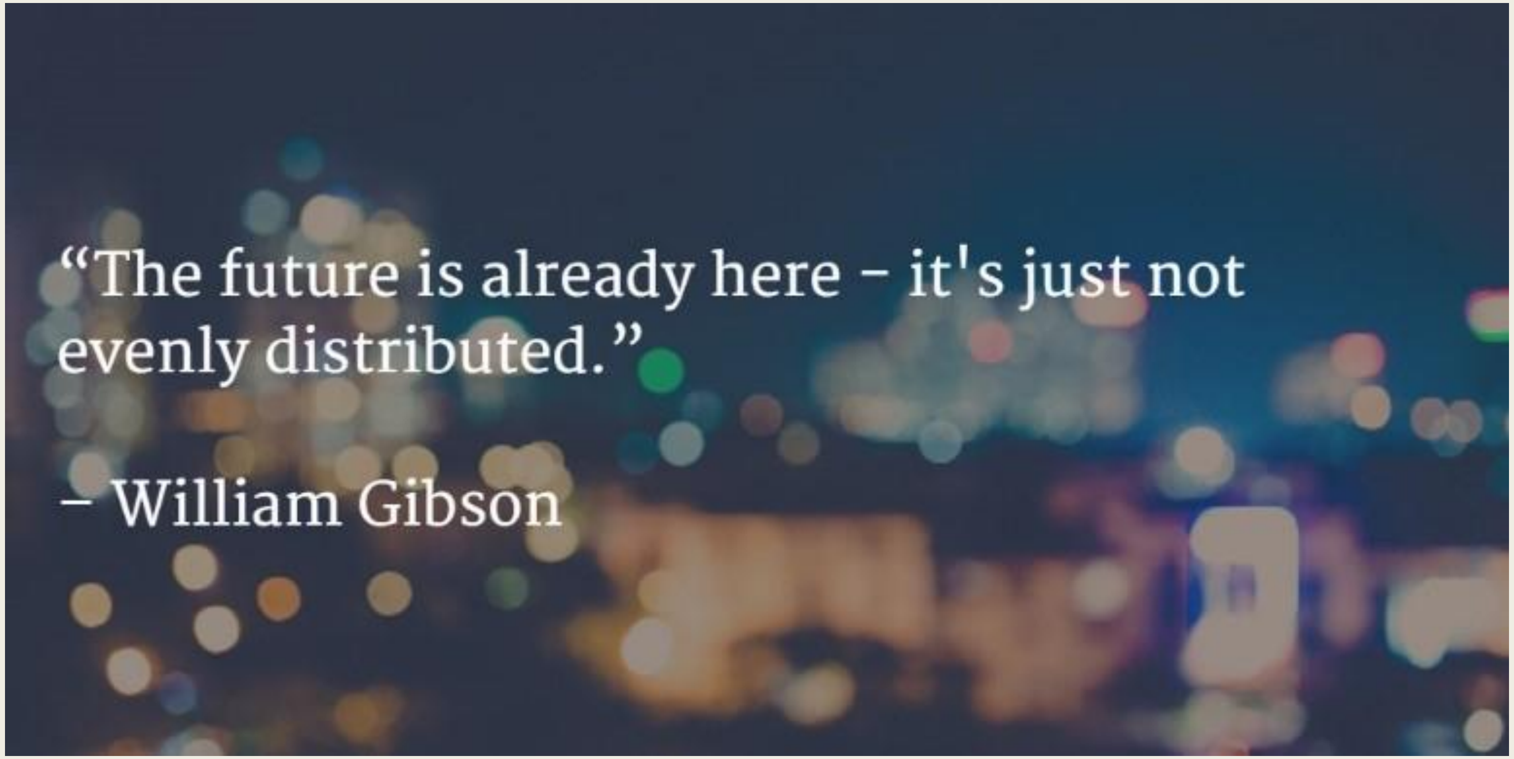
- creating added value from big data or small data
- predicting behaviours
- pre-empting behaviours

- interplay of backend & frontend of computing systems
- interfaces *enable* but they also *hide*, nudge and force [AB testing, ‘by design’ paradigms]

onlife: digital unconscious

Big Data Space:

- accumulation of behavioural and other data
- mobile and polymorphous data & hypothesis spaces
- distributed storage [once data has been shared, control becomes a challenge]
- distributed access [access to data or to the inferences, to training set & algos]



“The future is already here – it's just not evenly distributed.”

– William Gibson

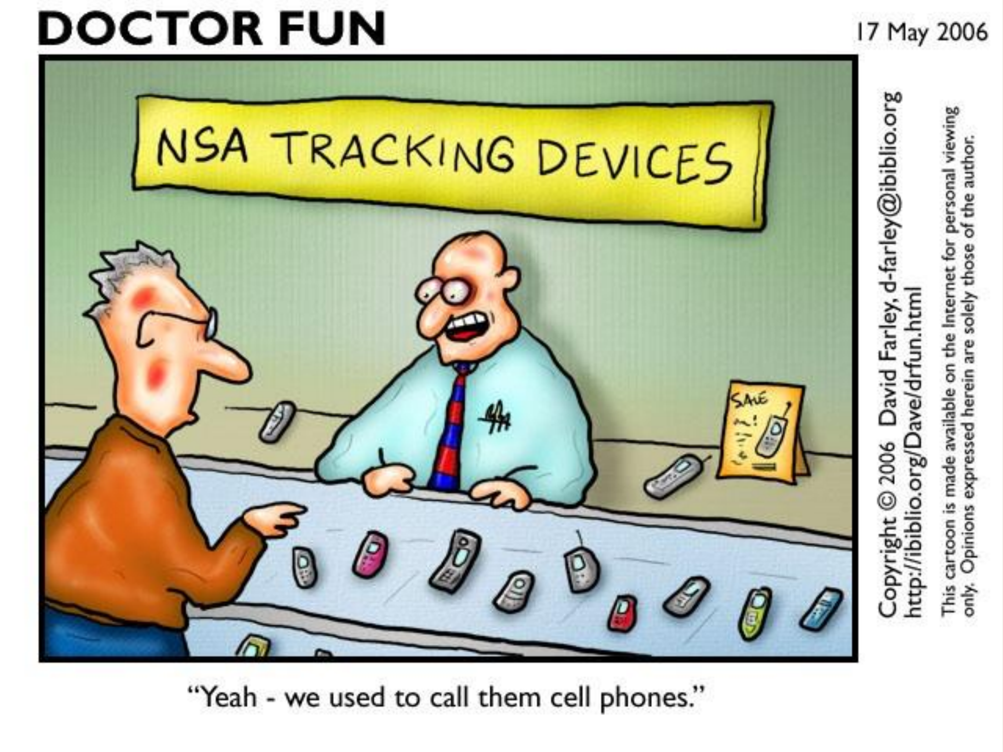
onlife: digital unconscious

Big Data Space:

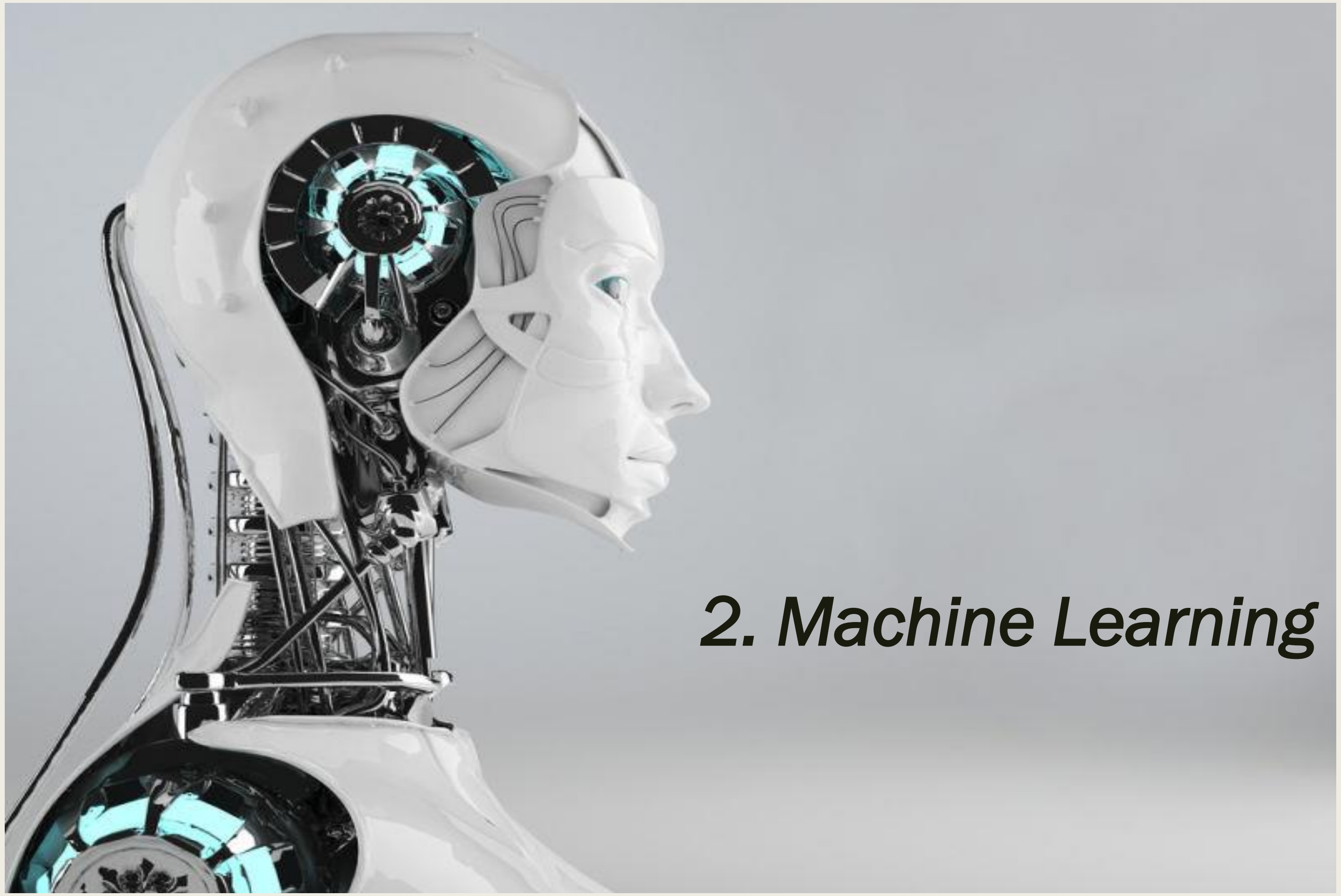
the envelop of big data space drives human agency, providing convenience & resilience

Weiser's ***calm computing***, IBM's ***autonomic computing***:

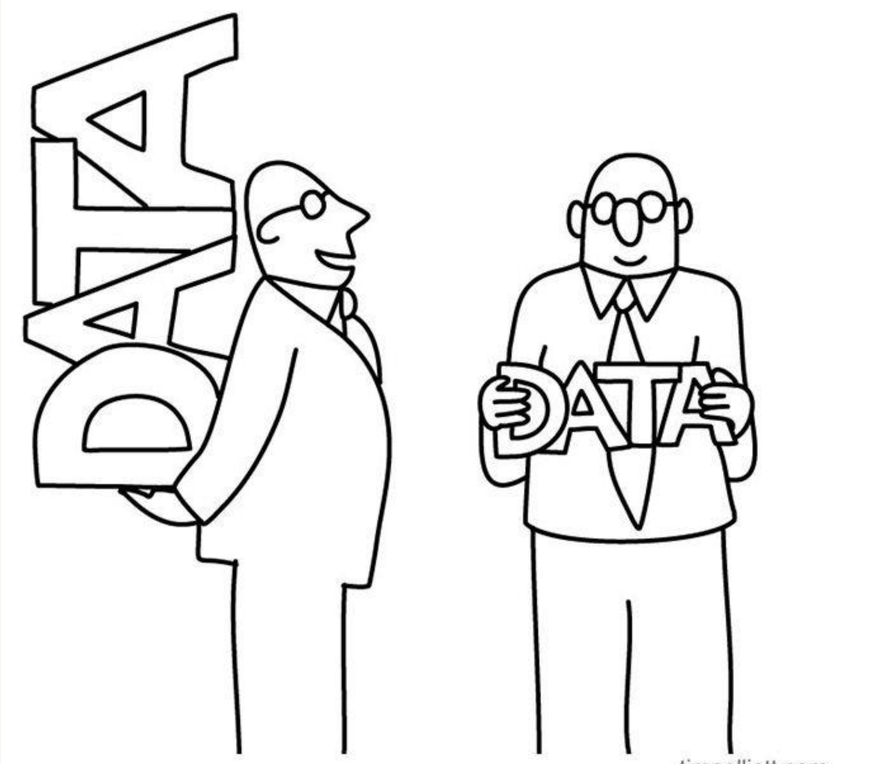
- increasing dependence on the dynamics of interacting data driven cyberphysical systems


DOCTOR FUN
17 May 2006
NSA TRACKING DEVICES
SALE
Copyright © 2006 David Farley, d-farley@ibiblio.org
http://ibiblio.org/Dave/drfun.html
This cartoon is made available on the Internet for personal viewing only. Opinions expressed herein are solely those of the author.
"Yeah - we used to call them cell phones."

what's next?



2. Machine Learning



"I think you'll find that mine is bigger..."

big data, open data, personal data

- BIG
- – *volume (but, n=all is nonsense)*
- – *variety (unstructured in sense of different formats)*
- – *velocity (real time, streaming)*

- OPEN as opposed to proprietary? reuse? repurposing? public-private?
- – *creating added value is hard work, not evident, no guarantees for return on investment*

- PERSONAL data: IoT will contribute to a further explosion of personal data
- – *high risk high gain (think DPIA)? anonymisation will mostly be pseudonymisation!*

machine learning (ML)

“we say that a machine learns:

- with respect to a particular task **T**,
- performance metric **P**, and
- type of experience **E**,

if

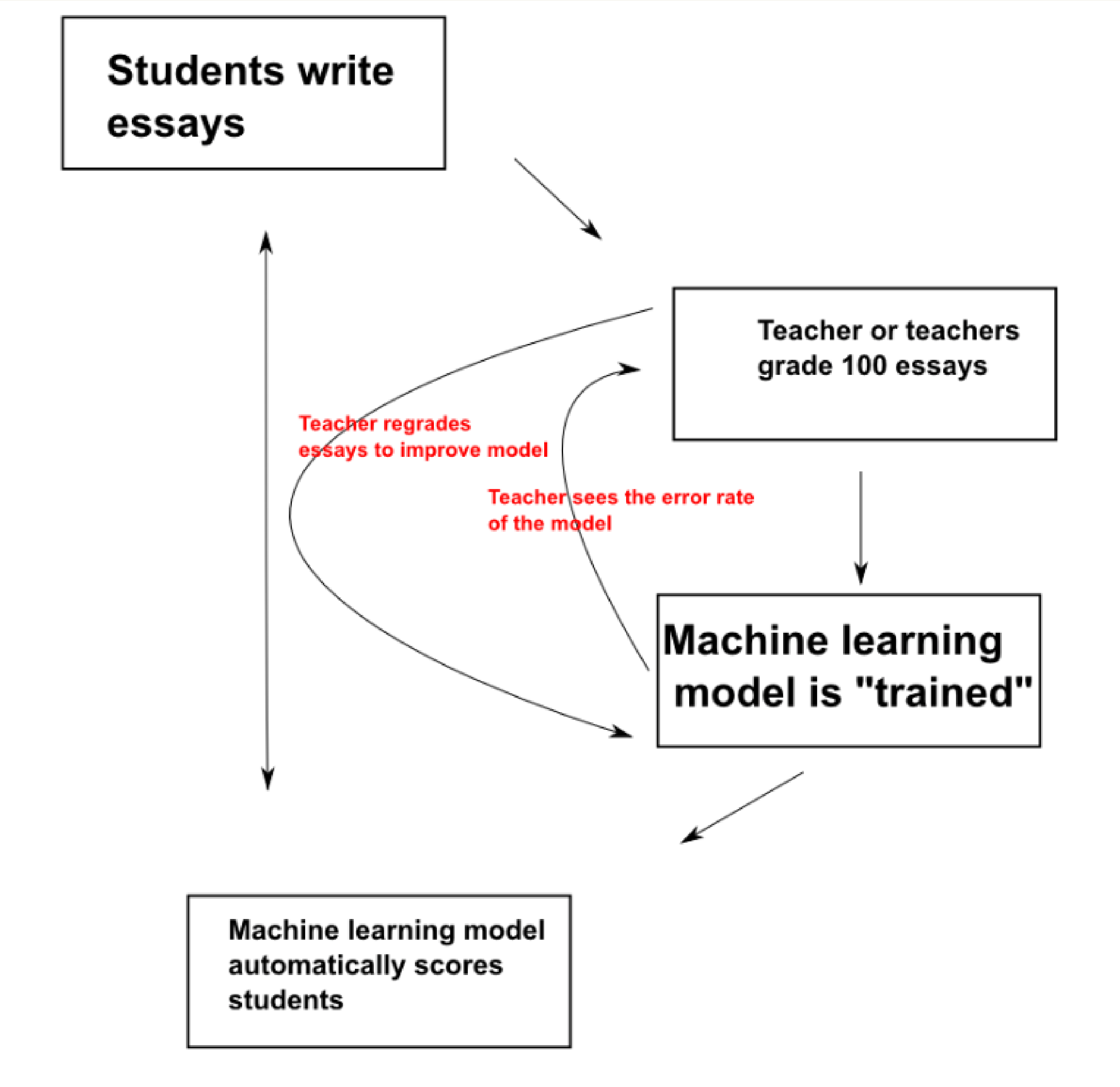
- the system reliably improves its performance **P**
- at task **T**,
- following experience **E**.”

(Tom Mitchell)

http://www.cs.cmu.edu/~tom/mlbook.html

types of machine learning

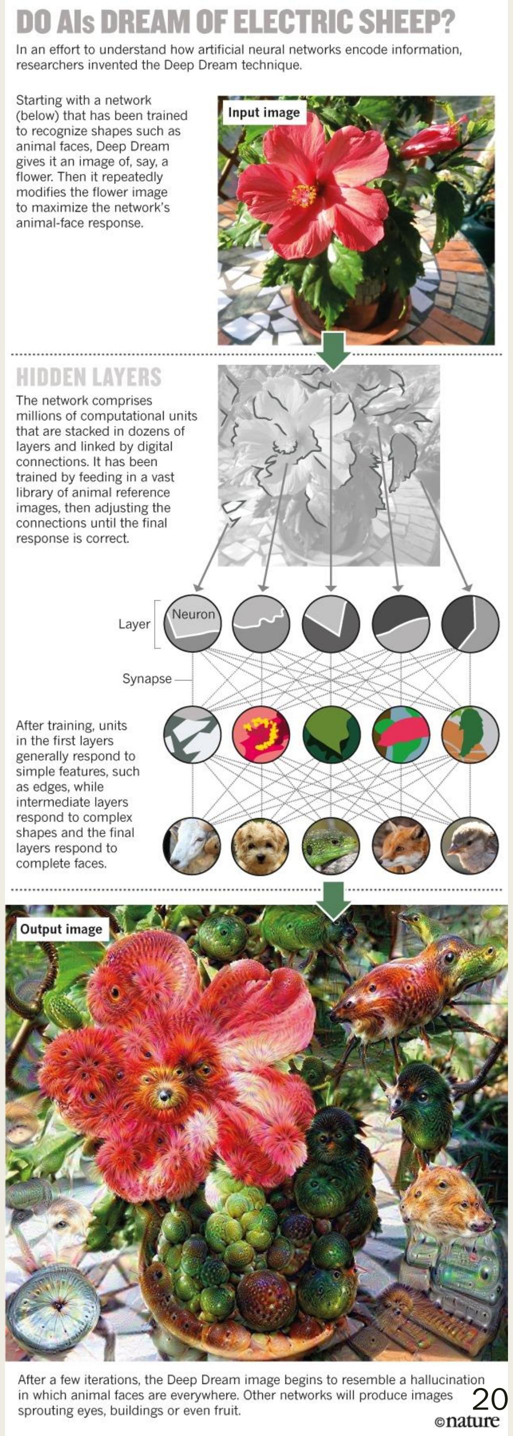
- *supervised* (learning from examples – requires labelling, domain expertise)
- *reinforcement* (learning by correction - requires prior domain expertise)
- *unsupervised* (bottum up, inductive – danger of overfitting)


Students write essays
Teacher or teachers grade 100 essays
Teacher regrades essays to improve model
Teacher sees the error rate of the model
Machine learning model is "trained"
Machine learning model automatically scores students

bias
optimisation
spurious correlations

- 2. have a network trained to recognize animal faces
- 1. present it with a picture of a flower
- 2. run the algorithms
- 3. check the output (see what it sees)

http://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731



DO AIs DREAM OF ELECTRIC SHEEP?

In an effort to understand how artificial neural networks encode information, researchers invented the Deep Dream technique.

Starting with a network (below) that has been trained to recognize shapes such as animal faces, Deep Dream gives it an image of, say, a flower. Then it repeatedly modifies the flower image to maximize the network's animal-face response.

HIDDEN LAYERS

The network comprises millions of computational units that are stacked in dozens of layers and linked by digital connections. It has been trained by feeding in a vast library of animal reference images, then adjusting the connections until the final response is correct.

After training, units in the first layers generally respond to simple features, such as edges, while intermediate layers respond to complex shapes and the final layers respond to complete faces.

After a few iterations, the Deep Dream image begins to resemble a hallucination in which animal faces are everywhere. Other networks will produce images sprouting eyes, buildings or even fruit.

Wolpert: no free lunch theorem

Where

d = training set;

f = ‘target’ input-output relationships;

h = hypothesis (the algorithm's guess for f made in response to d); and

C = off-training-set ‘loss’ associated with f and h (‘generalization error’)

How well you do

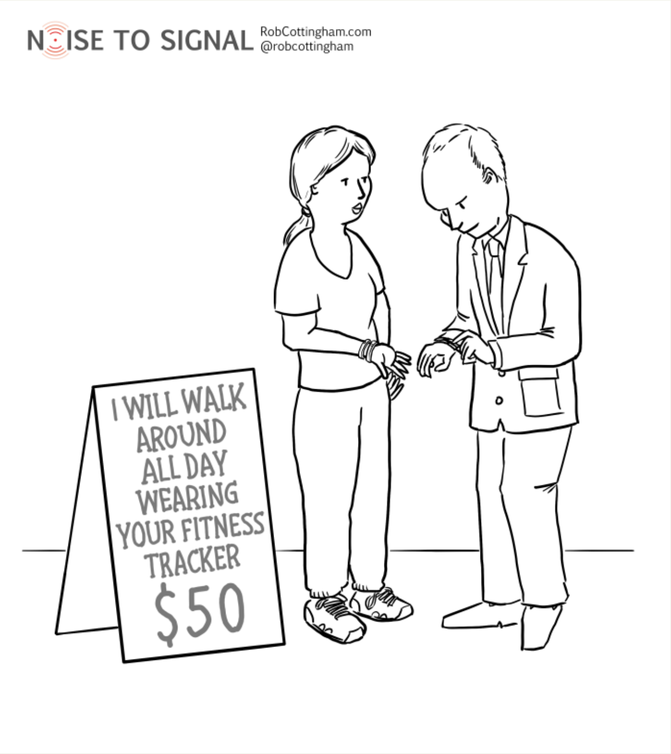
is determined by how ‘aligned’ your learning algorithm P(h|d) is with the actual posterior, P(f|d).

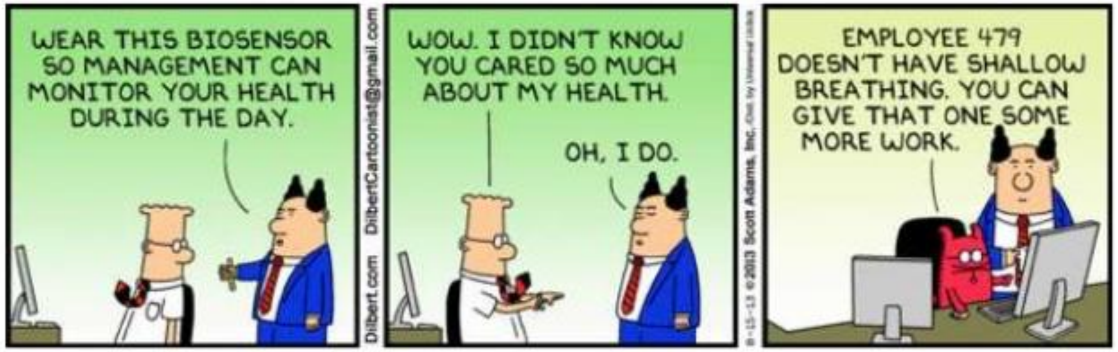
Check http://www.no-free-lunch.org

Wolpert: no free lunch theorem

Summary:

- *The bias that is necessary to mine the data will co-determine the results*
- *This relates to the fact that the data used to train an algorithm is finite*
- *‘Reality’, whatever that is, escapes the inherent reduction*
- *Data is not the same as what it refers to or what it is a trace of*


NOISE TO SIGNAL
RobCottingham.com
@robcottingham
I WILL WALK AROUND ALL DAY WEARING YOUR FITNESS TRACKER
$50


WEAR THIS BIOSENSOR SO MANAGEMENT CAN MONITOR YOUR HEALTH DURING THE DAY.
WOW. I DIDN'T KNOW YOU CARED SO MUCH ABOUT MY HEALTH.
OH, I DO.
EMPLOYEE 479 DOESN'T HAVE SHALLOW BREATHING. YOU CAN GIVE THAT ONE SOME MORE WORK.
Dilbert.com DilbertCartoonist@gmail.com
8-15-13 ©2013 Scott Adams, Inc./Dist. by Universal Uclick

trade-offs

- NFL theorem
  - *overfitting, overgeneralization*
- training set, domain knowledge, hypotheses space, test set
  - *accuracy, precision, speed, iteration*
- low hanging fruit
  - *may be cheap and/or available but not very helpfull*
- data nor algorithms are objective
  - *bias in the data, bias of the algos, guess what: bias in the output*
- the more data, the larger the hypotheses space, the more patterns
  - *spurious correlations, computational artefacts*

data hoarding & obesitas

- data obesitas: lots of data, but often incorrect, incomplete, irrelevant (low hanging fruit)
  - *any personal data stored presents security and other **risks** (need for DPIA, DPbD)*
  - ***purpose limitation** is crucial: **select before you collect** (and while, and after)*

- pattern obesitas: trained algorithms can see patterns anywhere, added value?
  - *training set and algorithms **necessarily** contain bias, this **may** be problematic (need for DPIA, DPbD)*
  - ***purpose limitation** is crucial: to prevent spurious correlations, to **test relevance***

agile and lean computing

- agile software development:
  - *iteration instead of waterfall*
  - *collaboration domain experts, data scientists, whoever invests*
  - *initial purpose (prediction of behaviour, example: tax office, car insurance)*
  - *granular purposing (testing specific patterns, AB testing to nudge specific behaviour)*

- lean computing:
  - *less data = more effective & more efficient*

- methodological integrity:
  - *make your software testable and contestable: mathematical & empirical software verification*
  - *secure logging, open source*

what's next?

4. Data Protection Law

privacy and autonomy

- ***the implications of pre-emptive computing:***
  - *AB testing & nudging*
  - *pre-emption of our intent, playing with our autonomy*
  - *we become subject to decisions of data-driven agents*
  - *this choice architecture may generate manipulability*

non-discrimination

- three types of *bias:*
  - *bias inherent in any action-perception-system (APS)*
  - *bias that some would qualify as unfair*
  - *bias that discriminates on the basis of prohibited legal grounds*

the opacity argument in ML:

1. intentional concealment

- trade secrets, IP rights, public security

2. we have learned to read and write, not to code or do machine learning

- monopoly of the new ‘clerks’, the end of democracy

3. mismatch between mathematical optimization and human semantics

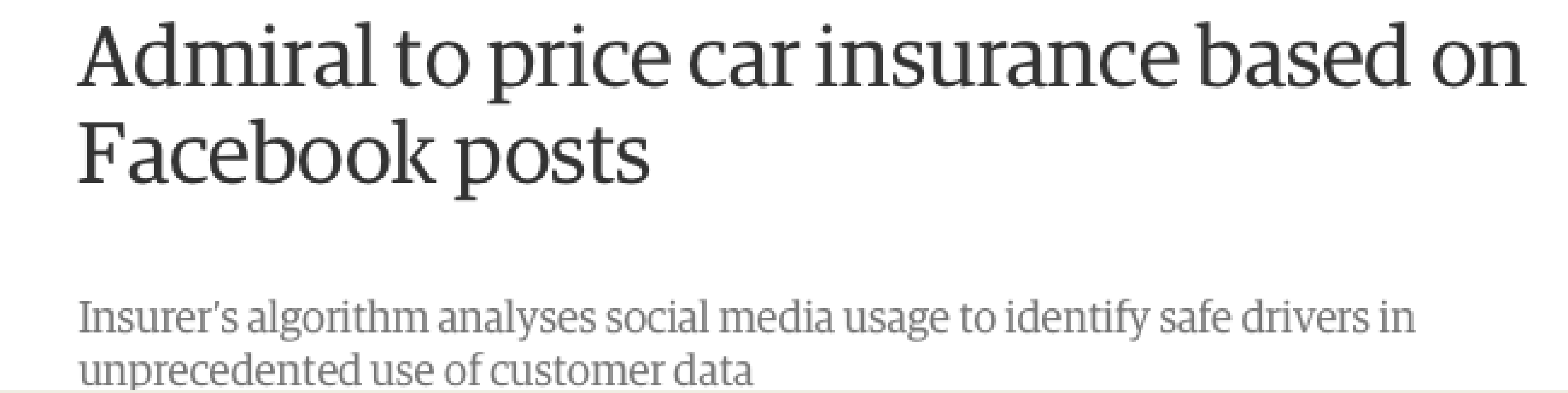
- when it comes to law and justice we cannot settle for ‘computer says no’

- *inspired by: Jenna Burrell, How the machine ‘thinks’: Understanding opacity in machine learning algorithms’, in* **Big Data & Society***, January-June 2016, 1-12*

due process

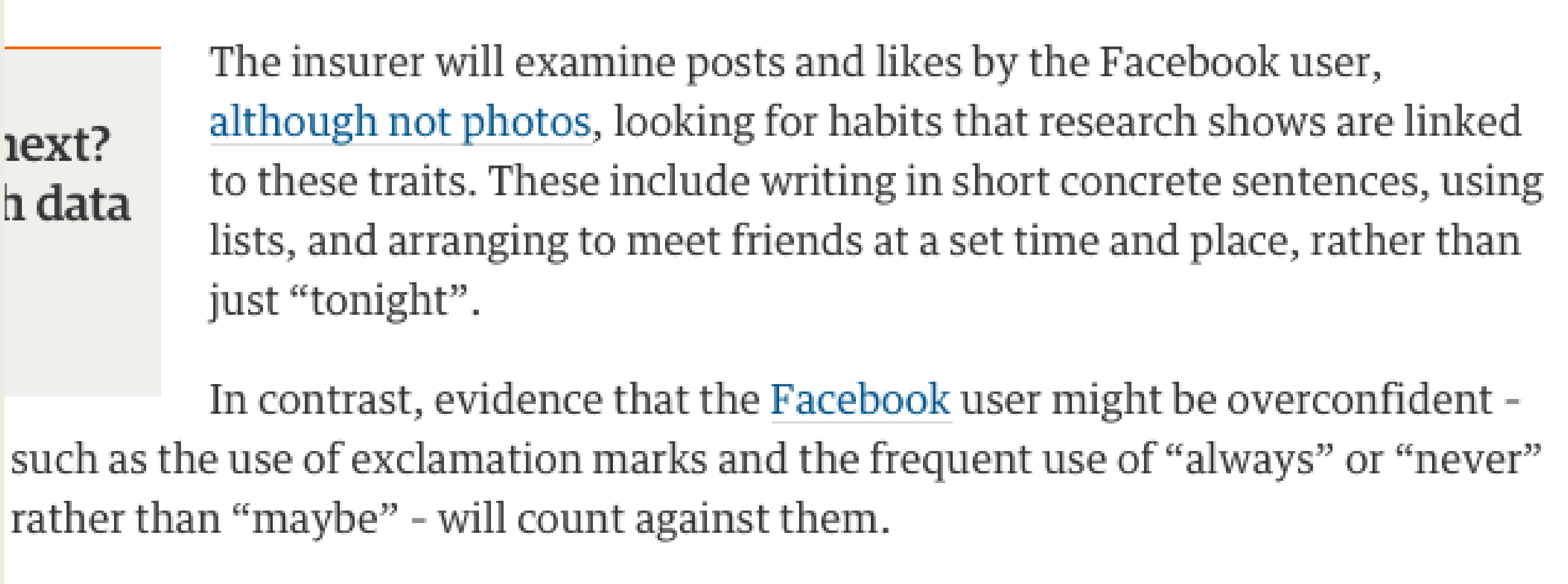
- in the case of automated decisions taken by AI systems we need:

1. *to know **that** ML or other algorithms determined the decision*
2. *to know **which data points** inform the decision and how they are weighted*
3. *which are the **envisaged consequences** of the employment of the algorithms*



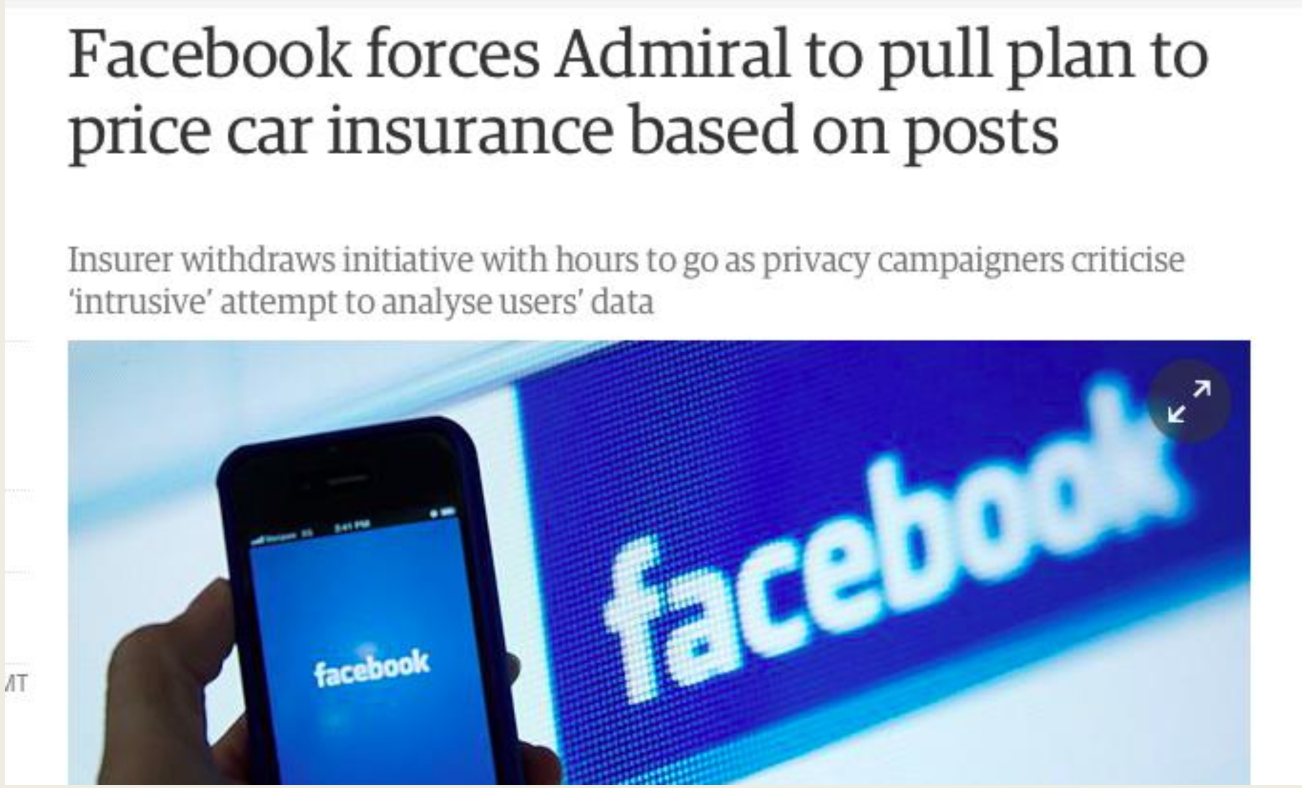
Admiral to price car insurance based on Facebook posts

Insurer's algorithm analyses social media usage to identify safe drivers in unprecedented use of customer data



The insurer will examine posts and likes by the Facebook user, although not photos, looking for habits that research shows are linked to these traits. These include writing in short concrete sentences, using lists, and arranging to meet friends at a set time and place, rather than just "tonight".

In contrast, evidence that the Facebook user might be overconfident - such as the use of exclamation marks and the frequent use of "always" or "never" rather than "maybe" - will count against them.



Facebook forces Admiral to pull plan to price car insurance based on posts

Insurer withdraws initiative with hours to go as privacy campaigners criticise 'intrusive' attempt to analyse users' data

Nature editorial 22 september 2016

- "To avoid bias and improve transparency, algorithm designers must make data sources and profiles public."
- "People should have the right to see their own data, how profiles are derived and have the right to *challenge* them."
- "Some proposed remedies are technical, such as developing new computational techniques that better address and correct discrimination both in training data sets and in the algorithms — a sort of affirmative algorithmic action."

the Onlife's Choice Architecture

- nudge theory, cognitive psychology, behavioural economics
  - what options does an environment give its inhabitants?

- what options does a data-driven environment give its 'users'?
- which are the defaults? e.g. withdrawal of consent in GDPR
- architecture is politics

Data Protection Law's Choice Architecture

- how does DP law constrain and reconfigure the Onl*ife*'s choice architectures?

1. what choice architecture does DP law provide data subjects?
2. what choice architecture does DP law provide data controllers?

data minimisation

= a choice architecture for data controllers:

- think ‘training sets’: select before you collect
- think of how to avoid ‘low hanging fruit’
- think of how to ensure accuracy, relevance, pertinence
- data minimisation, if done well, should avoid both data and pattern obesitas

– *detect productive bias, while also detecting unfair or prohibited bias*

– *make data sets available for inspection and contestation*

purpose limitation

= a choice architecture for data controllers

- think ‘training sets’: select before you collect (*and while you collect and after*)
- think of how to avoid ‘low hanging fruit’ (*GIGA*)
- think of how to ensure accuracy, relevance, pertinence (*depending on purpose*)
  - *purpose specification, if done well, should avoid both data and pattern obesitas*
  - *purpose should direct the development and employment of data-driven applications*
  - *experimentation can be a purpose, but not in itself*
- the choice of algorithms should be informed by the purpose

automated decision rights

- current choice architecture of AI:
- ML, IoT is meant to pre-empt our intent
- to run smoothly under the radar of everyday life
- it is all about continuous surreptitious automated decisions

automated decision rights

= choice architecture for data subjects (EU legislation)

1. the right not to be subject to automated decisions that have a significant impact
2. the right to a notification, an explanation and anticipation if exception applies

automated decision rights

= choice architecture for data subjects:

1. the right not to be subject to automated decisions that have a significant impact, unless

a. *necessary for contract*

b. *authorised by EU or MS law*

c. *explicit consent*

under a and c: right to human intervention, possibility to contest

prohibition to make such decisions based on sensitive data

automated decision rights

= choice architecture for data subjects:

2. the right to a notification, an explanation and anticipation if exception applies

- *existence of decisions based on profiling*
- *meaningful information about the logic involved (= explanation?)*
- *significance and envisaged consequences of such processing*

legal protection by design

- Data Protection by Design:
  - *JASP (open source) v SPSS (proprietary)*
  - *can give you output based on Bayesian and classical statistics*
  - *on same data set*
  - *imagine training same algorithms on different data sets*

- science and open society fit well together:
  - *make systems testable (science), make systems contestable (Rule of Law)*
  - *if you can’t test it you can’t contest it*

DP & Privacy Law: Choice Architecture

- individual citizens need:
  - *the capability to reinvent themselves,*
  - *to segregate their data-driven audiences,*
  - *have their human dignity respected by the data-driven infrastructures*
  - *make sure their robotic social companions don't tell on them beyond necessary*
  - *the capability to detect and contest bias in their data-driven environments*

DP & Privacy Law: Choice Architecture

- the architects of our new data-driven world need:
  - *integrity of method: rigorously sound and contestable methodologies (bias)*
  - *accountabiity: (con)testability of both data sets and algorithms*
  - *fairness: testing bias in the training set, testing bias in the learning algorithm*
  - *privacy & data protection: reduce manipulability, go for participation and respect*

‘by design’ paradigm

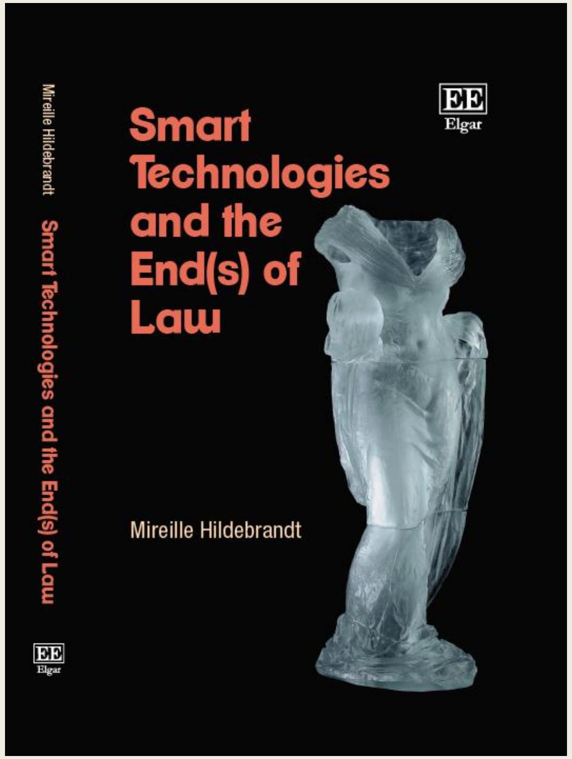
- *architecture is politics*
- *translate fairness, methodological integrity, fundamental rights into the architecture*
- *Data Protection by Default: engineer data minimisation as a requirement*
- *Data Protection by Design: engineer state of the art DP tools as a requirement*

‘by design’ paradigm

- *architecture is politics*
  - *we cannot always be saved by design*
  - *resilience will depend on: testability & contestability*
  - *ultimately this is in the interest of data subjects and controllers*

- *ML as a utility*

THE END IS NEAR
THE END
Roy Delgado



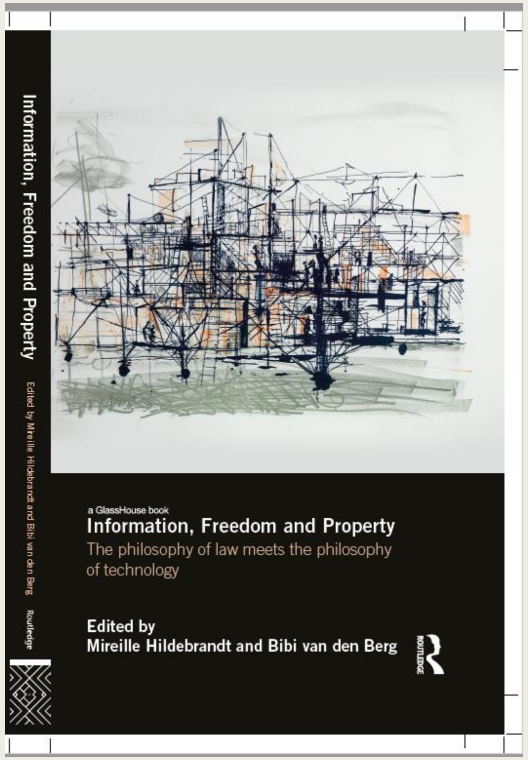
Mireille Hildebrandt Smart Technologies and the End(s) of Law

EE Elgar

Smart Technologies and the End(s) of Law

Mireille Hildebrandt

EE Elgar



Information, Freedom and Property

Edited by Mireille Hildebrandt and Bibi van den Berg

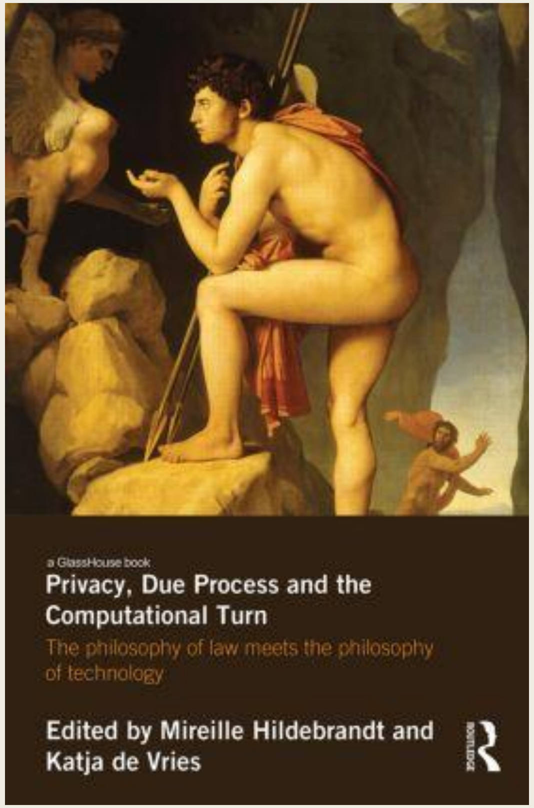
Routledge

a GlassHouse book

Information, Freedom and Property

The philosophy of law meets the philosophy of technology

Edited by
Mireille Hildebrandt and Bibi van den Berg



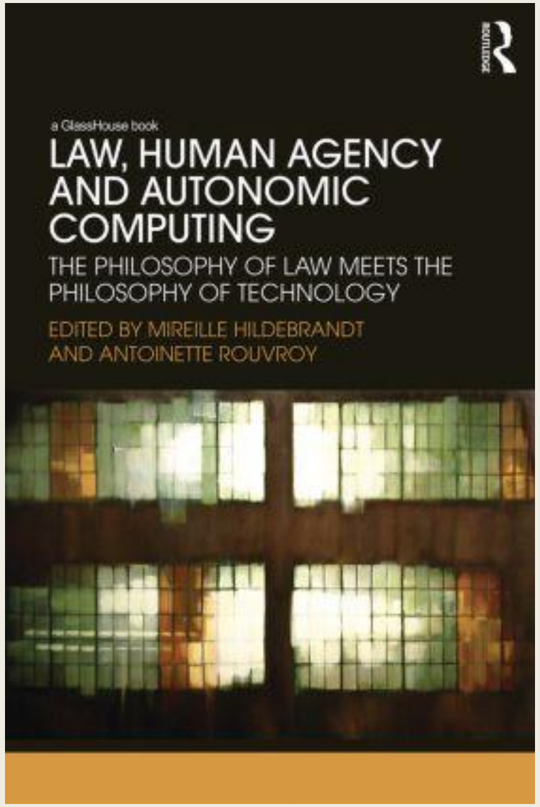
a GlassHouse book

Privacy, Due Process and the Computational Turn

The philosophy of law meets the philosophy of technology

Edited by Mireille Hildebrandt and Katja de Vries

ROUTLEDGE



ROUTLEDGE

a GlassHouse book

LAW, HUMAN AGENCY AND AUTONOMIC COMPUTING

THE PHILOSOPHY OF LAW MEETS THE PHILOSOPHY OF TECHNOLOGY

EDITED BY MIREILLE HILDEBRANDT AND ANTOINETTE ROUVROY