Utrecht, June 13th 2019

Business Intelligence

Lecture 07 - Managerial, Ethical and Privacy Considerations

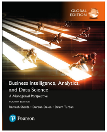
Georg Krempl

Algorithmic Data Analysis
Information and Computing Sciences
Utrecht University, The Netherlands

With particular thanks to

- Armel Lefebvre (tutor, `A.E.J.Lefebvre@uu.nl`)
- Vincent Lemaire (Data Analytics Research Project Manager at Orange Labs)
- Oskar Gstrein and Andrej Zwitter (Big Data Ethics and Law, Data Research Center, University of Groningen)
- Gerhard Kunnert (Constitutional Service, Austrian Ministry of Justice)
- Ana Maldonado and Alex Ridden (Data Scientists, Altius Europe)

Managerial, Ethical and Privacy Considerations



GLOBAL EDITION

Business Intelligence, Analytics, and Data Science

A Managerial Perspective

FOURTH EDITION

Ramesh Sharda • Dursun Delen • Efraim Turban

Pearson

Figure: Textbook [Sharda et al., 2018, Chapter 8] Sharda, Delen, Turban & King (2018). Business Intelligence, Analytics & Data Science: A Managerial Perspective 4th Global Edition, Pearson. ISBN-13: 9781292220567

IEEE Access

Privacy-Preserving Data Mining: Methods, Metrics, and Applications

RICARDO MENDES AND JOÃO P. VILELA

Figure: Paper [Mendes and Vilela, 2017] Mendes, Vilela (2017). Privacy-Preserving Data Mining: Methods, Metrics & Applications IEEE Access vol. 5, 2017

SCIENCE ADVANCES | RESEARCH ARTICLE

SOCIAL NETWORKS

Leaking privacy and shadow profiles in online social networks

David Garcia

Figure: Paper [Garcia, 2017] Garcia (2017). Leaking privacy and shadow profiles in online social networks Science Advances 3(8) 2017

PNAS

Private traits and attributes are predictable from digital records of human behavior

Michal Kosinski, David Stillwell, and Thore Graepel

Figure: Paper [Kosinski et al., 2013] Kosinski, Stillwell, Graepel (2013). Private traits and attributes are predictable from digital records of human behavior PNAS 110(15) 2013

Why 'Anonymous' Data Sometimes Isn't | WIRED https://www.wired.com/2007/12/why-anonymous-data-sometimes-isnt/

WHY 'ANONYMOUS' DATA SOMETIMES ISN'T

LAST YEAR, NETFLIX published 10 million movie rankings by 500,000 customers, as part of a challenge for people to come up with better recommendation systems than the one the company was using. The data was anonymized by removing personal details and replacing names with random numbers, to protect the privacy of the recommenders.

Arvind Narayanan and Vitaly Shmatikov, researchers at the University of Texas at Austin,

1 of 10 11/06/18 22:04

Figure: Paper [Schneier, 2007] Schneier (2007). Why Anonymous Data Sometimes Isn't Wired Security 2007

Outline and Summary[1]

- Introductory Examples
  See the links given in the examples.
- Ethics
  See [Sharda et al., 2018, chapter 7], [High-Level Expert Group on AI, 2019]
- European Ethics Guideline on Trustworthy AI
  See European Commission's website on Trustworthy AI
- Privacy Preserving Data Mining
  See [Mendes and Vilela, 2017] and [Sharda et al., 2018, chapter 7]
- European General Data Protection Regulation (GDPR)
  See European Commission's website on GDPR, Trunomi's GDPR Portal and Official Text of the General Data Protection Regulation
- Summary & Concluding Remarks

▸ Start ▸ Appendix

[1] Note: This lecture integrates knowledge from several further sources (cited where used, except if my own).

Outline and Summary[1]

- Introductory Examples
  See the links given in the examples.

- Ethics
  See [Sharda et al., 2018, chapter 7], [High-Level Expert Group on AI, 2019]

- European Ethics Guideline on Trustworthy AI
  See European Commission's website on Trustworthy AI

- Privacy Preserving Data Mining
  See [Mendes and Vilela, 2017] and [Sharda et al., 2018, chapter 7]

- European General Data Protection Regulation (GDPR)
  See European Commission's website on GDPR, Trunomi's GDPR Portal and Official Text of the General Data Protection Regulation

- Summary & Concluding Remarks

▸ Start ▸ Appendix

[1] Note: This lecture integrates knowledge from several further sources (cited where used, except if my own).

From Shapping Baskets to Pregnancy Prediction

Target Supermarket

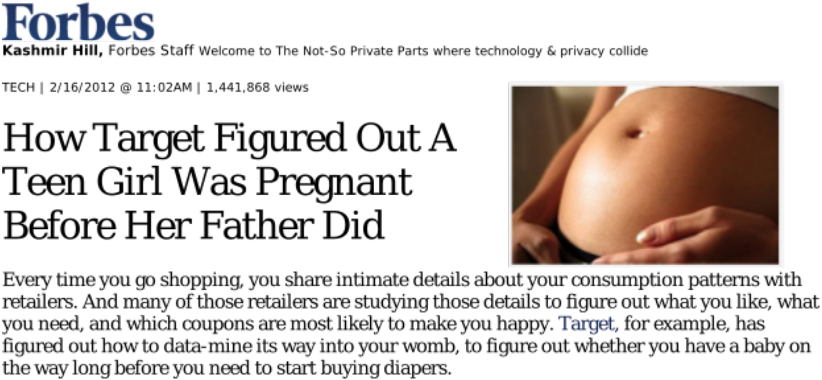
- Use basket analysis to identify prospective parents
- Goal: Send them customised promotional material

> *"Lotions, for example. Lots of people buy lotion, but one of Pole's colleagues noticed that women on the baby registry were buying larger quantities of unscented lotion around the beginning of their second trimester. Another analyst noted that sometime in the first 20 weeks, pregnant women loaded up on supplements like calcium, magnesium and zinc. Many shoppers purchase soap and cotton balls, but when someone suddenly starts buying lots of scent-free soap and extra-big bags of cotton balls, in addition to hand sanitizers and washcloths, it signals they could be getting close to their delivery date."*

Potential Consequences?

- Irritated customers
- Privacy issues

From Shapping Baskets to Pregnancy Prediction



Forbes

Kashmir Hill, Forbes Staff Welcome to The Not-So Private Parts where technology & privacy collide

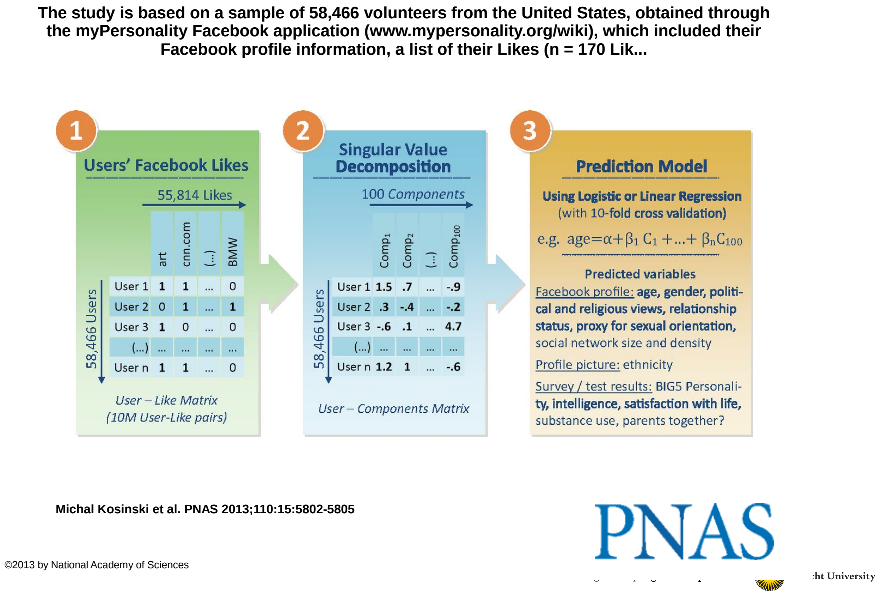
TECH | 2/16/2012 @ 11:02AM | 1,441,868 views

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Figure: How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did
Source: Hill (2012) [Hill, 2012])

Example: Predicting Private Traits and Attributes From Facebook Likes



The study is based on a sample of 58,466 volunteers from the United States, obtained through the myPersonality Facebook application (www.mypersonality.org/wiki), which included their Facebook profile information, a list of their Likes (n = 170 Lik...

1

Users' Facebook Likes

55,814 Likes

58,466 Users

	art	cnn.com	(...)	BMW
User 1	1	1	...	0
User 2	0	1	...	1
User 3	1	0	...	0
(...)	...	...	...	...
User n	1	1	...	0

User – Like Matrix
(10M User-Like pairs)

2

Singular Value Decomposition

100 Components

58,466 Users

	$Comp_1$	$Comp_2$	(...)	$Comp_{100}$
User 1	1.5	.7	...	-.9
User 2	.3	-.4	...	-.2
User 3	-.6	.1	...	4.7
(...)	...	...	...	...
User n	1.2	1	...	-.6

User – Components Matrix

3

Prediction Model

Using Logistic or Linear Regression
(with 10-fold cross validation)

e.g. $age = \alpha + \beta_1 C_1 + ... + \beta_n C_{100}$

Predicted variables

Facebook profile: **age, gender, political and religious views, relationship status, proxy for sexual orientation,** social network size and density

Profile picture: ethnicity

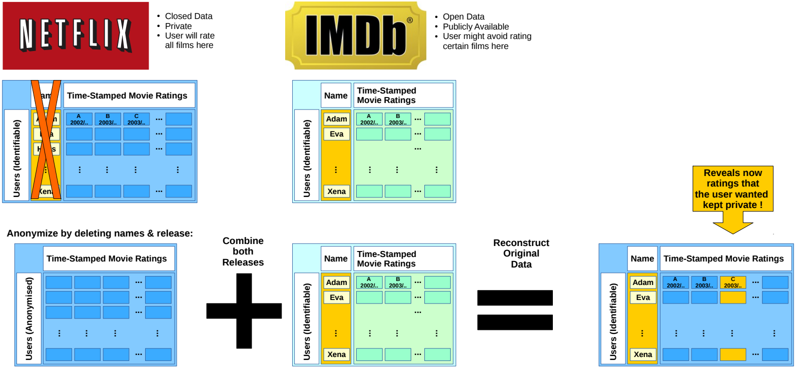
Survey / test results: **BIG5 Personality, intelligence, satisfaction with life,** substance use, parents together?

Michal Kosinski et al. PNAS 2013;110:15:5802-5805

PNAS

Figure: Predicting Private Traits and Attributes From FB Likes (Source: PNAS, [Kosinski et al., 2013])

Data Anonymisation: Why 'Anonymous' Data Sometimes Isn't [2]



- Closed Data
- Private
- User will rate all films here

- Open Data
- Publicly Available
- User might avoid rating certain films here

Anonymize by deleting names & release:

Combine both Releases

Reconstruct Original Data

Reveals now ratings that the user wanted kept private !

[2] See [Schneier, 2007] for further cases: https://www.wired.com/2007/12/why-anonymous-data-sometimes-isnt

From Personal Data to Anonymized Data . . .

Massachusetts Group Insurance Case

- Massachusetts Group Insurance Commission, 1990s
- Goal: Help research by releasing health data
- Released data on state employees that showed every single hospital visit
- Anonymised by deleting obvious identifiers such as name, address, and Social Security number.
- Good idea?

From Personal Data to Anonymized Data . . . and Back!

At the time GIC released the data, **William Weld, then Governor of Massachusetts, assured** *the public that GIC had* **protected patient privacy by deleting identifiers***. In response, then-graduate student Sweeney started hunting for the Governor's hospital records in the GIC data. She knew that Governor Weld resided in Cambridge, Massachusetts, a city of 54,000 residents and seven ZIP codes. For twenty dollars, she purchased the complete voter rolls from the city of Cambridge, a database containing, among other things, the name, address, ZIP code, birth date, and sex of every voter.* **By combining this data with the GIC records, Sweeney found Governor Weld with ease***. Only six people in Cambridge shared his birth date, only three of them men, and of them, only he lived in his ZIP code. In a theatrical flourish,* **Dr. Sweeney sent the Governor's health records (which included diagnoses and prescriptions) to his office***.*

Are Decision Makers Wiser Today?

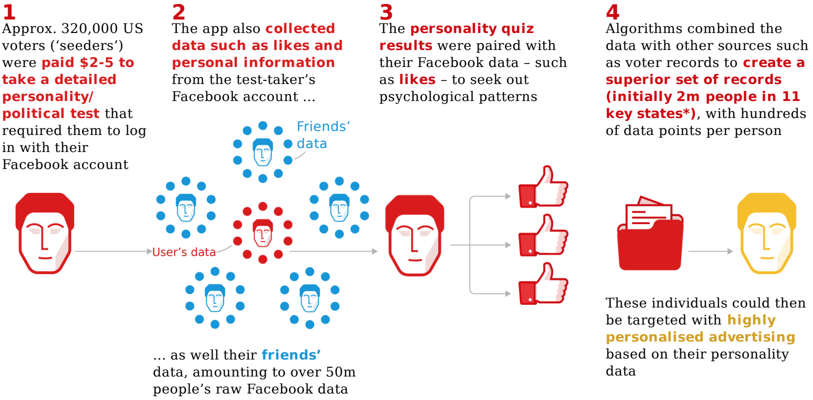
Austria, 2018

- Proposed Data Protection Adjustment Law ("Datenschutz-Anpassungsgesetz"):
- Allow companies and universities access to personal data for research purposes
- Anonymize the data by replacing subjects' names by a pseudo-identifier

> *"Geplant ist konkret, dass persönliche Daten der Österreicherinnen und Österreicher, die der Bund erhoben und abgespeichert hat, für Forschungszwecke abgefragt werden dürfen ('Registerforschung').* **Die Namen der Betroffenen sollen durch eine Kennzahl ersetzt werden, um die namentliche Zuordnung ihrer Daten zu verhindern.**"
>
> *ORF.AT, 11.4.2018, http://orf.at/stories/2433720/2433715/*

- Finally, this paragraph was withdrawn . . .

Cambridge Analytica and Facebook



1 Approx. 320,000 US voters ('seeders') were **paid $2-5 to take a detailed personality/ political test** that required them to log in with their Facebook account

2 The app also **collected data such as likes and personal information** from the test-taker's Facebook account ...

... as well their **friends'** data, amounting to over 50m people's raw Facebook data

3 The **personality quiz results** were paired with their Facebook data – such as **likes** – to seek out psychological patterns

4 Algorithms combined the data with other sources such as voter records to **create a superior set of records (initially 2m people in 11 key states*)**, with hundreds of data points per person

These individuals could then be targeted with **highly personalised advertising** based on their personality data

Guardian graphic. *Arkansas, Colorado, Florida, Iowa, Louisiana, Nevada, New Hampshire, North Carolina, Oregon, South Carolina, West Virginia

Figure: Cambridge Analytica: how 50m Facebook records were hijacked
C. Cadwalladr and E. Graham-Harrison, The Guardian, March 17, 2018
`https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm`

Shadow Profiles: Facebook, Whatsapp, and the Likes[3]

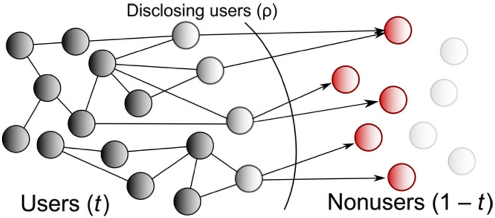


Figure: Shadow Profiles (Source: [Garcia, 2017])

> *The results of this analysis indicate that* **personal information of people outside Friendster could have been predicted with the data shared by Friendster users**, *supporting the* **shadow profile hypothesis**. *This phenomenon can be explained by the mixing patterns of* **sexual orientation and relationship status** *that existed in the network, which* **revealed information of nonusers through the contact lists of users**.
>
> *. . . These results call for new privacy paradigms that take into account the fact that* **individual privacy decisions do not happen in isolation** *and are* **mediated by the decisions of others**.

[3] See e.g. [Garcia, 2017] for Shadow Profile Analysis in Social Networks.

Shadow Profiles: Infering DNA of (Most of) Iceland's Population

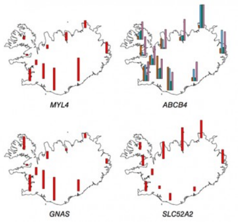


Figure: Risky DNA mutation distribution in Iceland (Source: [Regalado, 2015])

> *The company [. . .] has collected* **full DNA sequences on 10,000 individuals**. *And* **because people on the island are closely related**, *[..] it can now also extrapolate to* **accurately guess** *the DNA makeup of* **nearly all other 320,000 citizens** *of that country, including those* **who never participated in its studies**.
>
> *[. . . The company] has been unable to warn people of the danger because of ethics rules governing DNA research. [Regalado, 2015]*

- Genome study to predict genetic predisposition of cancer risk
- Data collected on 10,000 individuals from Iceland
- But: Close genetic relationship among population allows to infer cancer risk for individuals who did not participate
- Moral dilemma: Inform? Inform relatives? Ignore?

Outline and Summary[1]

- Introductory Examples
  See the links given in the examples.
- Ethics
  See [Sharda et al., 2018, chapter 7], [High-Level Expert Group on AI, 2019]
- European Ethics Guideline on Trustworthy AI
  See European Commission's website on Trustworthy AI
- Privacy Preserving Data Mining
  See [Mendes and Vilela, 2017] and [Sharda et al., 2018, chapter 7]
- European General Data Protection Regulation (GDPR)
  See European Commission's website on GDPR, Trunomi's GDPR Portal and Official Text of the General Data Protection Regulation
- Summary & Concluding Remarks

▸ Start ▸ Appendix

[1] Note: This lecture integrates knowledge from several further sources (cited where used, except if my own).

Example: Trolley Problem[4]

You / your algo

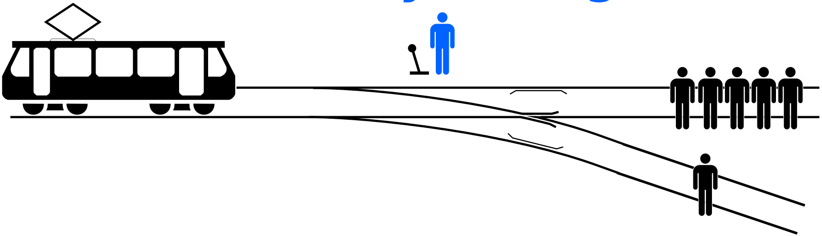


Figure: The Trolley Problem (Source: Wikimedia Commons / Zapyon)

- **Consequentialism:** “The greatest good for the greatest number” (utilitarianism, Bentham 1748–1832)
- **Categorical Moral Reasoning:** “Act only according to that maxim whereby you can, at the same time, will that it should become a universal law.” (Kant, 1724–1804)
- **Question: How is this related to, e.g., quantifying classification performance?**

[4] For a more detailed discussion of this and related moral dilemma, see Michael Sandel’s Harvard lecture *Justice: What’s The Right Thing To Do? Episode 01 “THE MORAL SIDE OF MURDER”* `https://www.youtube.com/watch?v=kBdfcR-8hEY`

Example: Trolley Problem

Considerations of a Car Producer

> *If you know you can save at least one person, at least save that one.* **Save the one in the car.** *If all you know for sure is that one death can be prevented, then that's your first priority.*
>
> *Christoph von Hugo, October 2016*
> *Head of Active Safety at Daimler AG*

AI/ML/BI Decisions

- Why bother about transparency and ethics?
- For example, should regulate algorithms?
- Market solution? Insurance? Regulation?

Discussion: Show Algorithms be Regulated?[5]

Markus Ehrenmann, Swisscom

“People have a right to an explanation about the decisions that affect them. And they have a right not to be discriminated against. This is why we have to be in a position to comprehend the decision-making processes of algorithms and, where necessary, to correct them.”

Mouloud Dey, SAS

“Creativity can’t be stifled nor research placed under an extra burden. Our hand must be measured and not premature. Creative individuals must be allowed the freedom to work, and not assigned bad intentions a priori. Likewise, before any action is taken, the actual use of an algorithm must be considered, as it is generally not the computer program at fault but the way it is used.”

[5] From M. Kassner (2017). Biases in algorithms: The case for and against government regulation. https://www.techrepublic.com/article/biases-in-algorithms-the-case-for-and-against-government-regulation

Discussion: Show Algorithms be Regulated? (2)[7]

Kenneth Cukier, Viktor Mayer-Schönberger

“The use, abuse, and misuse of data by the U.S. military during the Vietnam War is a troubling lesson about the limitations of information as the world hurls toward the big-data era. The underlying data can be of poor quality. It can be biased. It can be misanalyzed or used misleadingly. And even more damning, data can fail to capture what it purports to quantify.
We are more susceptible than we may think to the ’dictatorship of data’—that is, to letting the data govern us in ways that may do as much harm as good. The threat is that we will let ourselves be mindlessly bound by the output of our analyses even when we have reasonable grounds for suspecting that something is amiss.”

Alex Howard

“Our world, awash in data, will require new techniques to ensure algorithmic accountability, leading the next-generation of computational journalists to file Freedom of Information requests for code, not just data, enabling them to reverse engineer how decisions and policies are being made by programs in the public and private sectors. To do otherwise would allow data-driven decision making to live inside of a black box, ruled by secret codes, hidden from the public eye or traditional methods of accountability. Given that such a condition could prove toxic to democratic governance and perhaps democracy itself, we can only hope that they succeed.[6]

[6] TechRepublic (2014). Data-driven policy and commerce requires algorithmic transparency.” https://www.techrepublic.com/article/data-driven-policy-and-commerce-requires-algorithmic-transparency

[7] From M. Kassner (2017). Biases in algorithms: The case for and against government regulation. https://www.techrepublic.com/article/biases-in-algorithms-the-case-for-and-against-government-regulation

Lessons From Those Illustrative Cases

- Often conflict between privacy/data protection and usability of data/knowledge

- Not only **direct identifiers** like names, or IDs, are critical,
- but also **pseudo- or quasi-identifiers** like date of birth, postal code, etc.

- Not only **characteristics from explicitly provided data** might be exposed,
- but also **further characteristics** that are inferred from this data or from **combining it with other data sources**

- Not only data about **participating individuals** might be exposed,
- but also about **non-participating individuals (shadow profiles)**

- Involves legal, privacy, and ethical/moral aspects

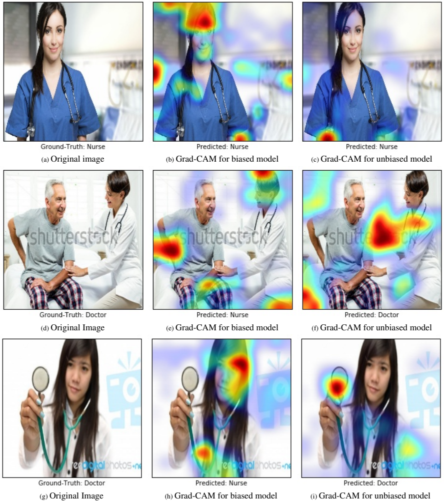
Fairness and Transparency: Motivating Examples: Wolf vs. Dogs

- Idea: Learn a classifier to discriminate between wolfs and dogs
- Approach: Acquire some images of wolfs and some of dogs,
  Learn a classifier (here a deep neural network (DNN))
- Works seemingly fine - but did the classifier learn **the right thing?**

(a) Husky classified as wolf (b) Expanation (snowy background)

Figure: Misclassification due to training data (Source: [Ribeiro et al., 2016, Fig 11])

Fairness and Transparency: Motivating Examples: Doctors vs. Nurses[8]



(a) Original image (b) Grad-CAM for biased model (c) Grad-CAM for unbiased model

(d) Original Image (e) Grad-CAM for biased model (f) Grad-CAM for unbiased model

(g) Original Image (h) Grad-CAM for biased model (i) Grad-CAM for unbiased model

Figure A4: Grad-CAM explanations for model1 and model2. In (a-c) we can see that even though both models made the right decision, the biased model (model1) was looking at the face of the person to decide if the person was a nurse (b), whereas the unbiased model, was looking at the short sleeves to make the decision (c). For example image (d) and example (g) the biased model made the wrong prediction (misclassifying a doctor as a nurse) by looking at the face and the hairstyle (e, h), where as the unbiased model made the right prediction looking at the white coat, and the stethoscope (f, i).

Figure: Source: [Selvaraju et al., 2016, Figure A4]

[8] Thanks to Vincent Lemaire for the suggestion of this example!

Non-Discriminatory Classification, Fairness and Transparency

- Key questions:
  - How to ensure that prediction models are fair?
    e.g., do not use any (ethically/legally) discriminatory information
  - How to verify this / make them transparent?

  See part 2 “Fairness and GDPR” of Vincent Lemaie’s talk (cmp. [Wachter et al., 2018, Selvaraju et al., 2016])
- Not discusses due to lack of time...
- Some Literature References:
  - “Why should I trust you?: Explaining the predictions of any classifier” (dog-vs-wolf example above) [Ribeiro et al., 2016]
  - “Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization” (nurs-vs-doctor example above) [Selvaraju et al., 2016]
  - “A multi-dimensional approach to disinformation” [De Cock Buning and et al., 2018], lead by Madeleine de Cock Buning (Universiteit Utrecht)
  - “It’s not privacy, and it’s not fair” [Dwork and Mulli, 2013]
  - “Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data” [Veale and Binns, 2017]
  - “Big Data Ethics” [Zwitter, 2014]
  - **See also examples and links in the appendix!**

Outline and Summary[1]

- Introductory Examples
  See the links given in the examples.
- Ethics
  See [Sharda et al., 2018, chapter 7], [High-Level Expert Group on AI, 2019]
- European Ethics Guideline on Trustworthy AI
  See European Commission's website on Trustworthy AI
- Privacy Preserving Data Mining
  See [Mendes and Vilela, 2017] and [Sharda et al., 2018, chapter 7]
- European General Data Protection Regulation (GDPR)
  See European Commission's website on GDPR, Trunomi's GDPR Portal and Official Text of the General Data Protection Regulation
- Summary & Concluding Remarks

▸ Start ▸ Appendix

[1] Note: This lecture integrates knowledge from several further sources (cited where used, except if my own).

European Ethics Guidelines for Trustworthy AI

Trustworthy AI

- European Ethics Guidelines for Trustworthy AI
- published this year: [High-Level Expert Group on AI, 2019]

Components of Trustworthy AI:

1. lawful: complying with all applicable laws and regulations
2. ethical: ensuring adherence to ethical principles and values
3. robust: from both technical and social perspective
   Avoid to cause even unintentional harm

Trustworthy AI: Requirements

7 Key Requirements

1. Human agency and oversight
   Empower humans to make informed decisions, foster their fundamental rights

2. Technical Robustness and safety
   Resilient and secure, reliable and reproducible

3. Privacy and data governance
   Respect privacy and data protection, adequate data governance mechanisms

4. Transparency
   Explain decisions, inform human that interaction partner is AI

5. Diversity, non-discrimination and fairness
   AI systems should be accessible to all, avoid unfair biases

6. Societal and environmental well-being
   Benefit all human beings, including future generations
   Consider environmental, social and societal impact

7. Accountability
   Ensure responsibility and accountability for AI systems and their outcomes

Outline and Summary[1]

- Introductory Examples
  See the links given in the examples.
- Ethics
  See [Sharda et al., 2018, chapter 7], [High-Level Expert Group on AI, 2019]
- European Ethics Guideline on Trustworthy AI
  See European Commission's website on Trustworthy AI
- Privacy Preserving Data Mining
  See [Mendes and Vilela, 2017] and [Sharda et al., 2018, chapter 7]
- European General Data Protection Regulation (GDPR)
  See European Commission's website on GDPR, Trunomi's GDPR Portal and Official Text of the General Data Protection Regulation
- Summary & Concluding Remarks

▸ Start ▸ Appendix

[1] Note: This lecture integrates knowledge from several further sources (cited where used, except if my own).

Privacy Models [9]

k-Anonymity

- Entities are grouped into sets of *k* undistinguishable records (so-called equivalence classes)
- Does not consider that entities might share the same sensitive attribute value (e.g., a positive test result)

l-Diversity

- Expands k-Anonymity by requiring at least *l* “well-represented” values for sensitive attributes
- Does not consider difference in the distribution of sensitive attributes between whole population and an equivalence class

t-Closeness

- Expands l-Diversity by requiring closeness in the distribution of sensitive attributes between whole population and an equivalence class

[9] See [Mendes and Vilela, 2017, Table 2] for details.

Privacy Models (2) [10]

ϵ-Differential Privacy

- Requires that a single record does not considerably affect the outcome of the analysis of the data set
- Recent concept, ensures participation of (single) individual has not a considerably larger effect than non-participation
- No experimental guide on setting ϵ, and might require heavy data perturbation

[10] See [Mendes and Vilela, 2017, Table 2] for details.

Privacy Preservation Along Data Lifecycle Phases [11]

Data Collection Privacy

- Randomisation of data at the collection time,
- additive or multiplicative noise

Data Publishing Privacy

- Randomisation before releasing data
- Generalisation: replacing a sensitive value with a more general one
- Suppression: removal of sensitive attributes
- Anatomisation: de-associate pseudo-identifiers and sensitive attributes in two separate tables
- Perturbation: replace original data by synthetic values with identical statistical information

[11] See [Mendes and Vilela, 2017] for details.

Privacy Preservation Along Data Lifecycle Phases (2) [12]

Data Mining Output Privacy

- Filtering of data mining algorithms' output
- Association Rule Hiding
- Downgrading Classifier Effectiveness
- Query Auditing and Inference Control

Distributed Privacy

- Privacy preservation when data mining is distributed among multiple entities

[12] See [Mendes and Vilela, 2017] for details.

Outline and Summary[1]

- Introductory Examples
  See the links given in the examples.
- Ethics
  See [Sharda et al., 2018, chapter 7], [High-Level Expert Group on AI, 2019]
- European Ethics Guideline on Trustworthy AI
  See European Commission's website on Trustworthy AI
- Privacy Preserving Data Mining
  See [Mendes and Vilela, 2017] and [Sharda et al., 2018, chapter 7]
- European General Data Protection Regulation (GDPR)
  See European Commission's website on GDPR, Trunomi's GDPR Portal and Official Text of the General Data Protection Regulation
- Summary & Concluding Remarks

▸ Start ▸ Appendix

[1] Note: This lecture integrates knowledge from several further sources (cited where used, except if my own).

EU General Data Protection Regulation (GDPR)[13]

- The GDPR codifies and unifies data privacy laws across all European Union member countries.
- Enforcement as of May 25, 2018, after a 2-year grace period.
- Potentially devastating penalties for non-compliance with the provisions of the GDPR regarding collection/use of personal data
- Applicable to any business collecting personal data from a citizen of the EU
- Any data related to a natural person that can be used to directly or indirectly identify this person is considered personal data.

[13] Official law: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN

GDPR: Key Principles in a Nutshell

- Explicit consent, and easy withdrawal thereof
- Breach notification within 72 hrs of discovery
- Right to access one own's personal data:
  Upon data subject's request, companies must provide information about
  - whether personal data pertaining to them is being processed
  - where it is being processed,
  - for what purpose,
  - a copy of the personal data being processed in an electronic format (free of charge).
- Right to be forgotten:
  When asked by data subject, all personal data must be erased, any further dissemination and processing of the data must be stopped!
- Data portability:
  Companies must provide mechanisms for providing a subject with their personal data in a commonly used and machine-readable format (free of charge)

GDPR: Key Principles in a Nutshell (2)

- Privacy by Design:
  - Collect and process only data that is absolutely necessary for the completion of companies' business
  - Access to this data must be restricted to those employees who need it for completing the process (the subject has consented to)
- Record Keeping and Data Protection Officers:
  - Maintain detailed records pertaining to the collection, processing, and storage of personal data
  - For companies that are large (e.g., $\geq$ 250 employees) or process data of many natural persons (e.g., $\geq$ 5000 per annum)[14]: Requirment to establish Data Protection Officers, who oversee personal data use

[14] For the precise regulation, refer to the EU regulations.

GDPR: Consent

- The GDPR requires a business to ask for **explicit permission from the (data) subject before** processing any personal data.
- The request must use clear language (no hidding of permissions in, e.g., Terms and Conditions)
- The consent must be given for a specific purpose and must be requested separately from other documents / policy statements.
- It must be as easy to withdraw consent as it is to give it.

GDPR's Key Principles: Personal Data

Identified, Identifiable, and Anonymous

The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. To determine whether a natural person is **identifiable**, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments. **The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.** This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.

GDPR's Key Principles: Proportionality, Transparency, Personal Rights

Principle of Proportionality

The processing of personal data should be designed to serve mankind. The right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality

Transparency: Requesting information

Information provided under Articles 13 and 14 and any communication and any actions taken under Articles 15 to 22 and 34 **shall be provided free of charge**. [Exceptions: unfounded or excessive requests]

Rights of the Data Subject w.r.t. the Controller

- access to and rectification or erasure of personal data
- restriction of / objection to processing of personal data
- right to data portability

GDPR's Key Principles: Requesting One Own's Personal Information

Art. 13: Information to be provided where personal data are collected from the data subject

The controller shall provide

- the identity and the contact details of the controller
- the contact details of the data protection officer,
- the purposes and legal basis of the data processing
- the recipients or categories of recipients of the personal data
- if transfered to a 3rd country: reference to the appropriate or suitable safeguards
- the period for which the personal data will be stored
- the existence of **automated decision-making, including profiling** [..] meaningful **information about the logic involved**, as well as the significance and the envisaged **consequences** of such processing for the data subject

GDPR's Key Principles: Requesting One Own's Personal Information (2)

Art. 14: Information to be provided where personal data have **not been obtained** from the data subject

The controller shall provide

- as above, but the categories of personal data concerned (rather than the recipients / categories of recipients of p.d.)
- the legitimate interests pursued by the controller or by a 3rd party
- information about the source for the data

Outline and Summary[1]

- Introductory Examples
  See the links given in the examples.
- Ethics
  See [Sharda et al., 2018, chapter 7], [High-Level Expert Group on AI, 2019]
- European Ethics Guideline on Trustworthy AI
  See European Commission's website on Trustworthy AI
- Privacy Preserving Data Mining
  See [Mendes and Vilela, 2017] and [Sharda et al., 2018, chapter 7]
- European General Data Protection Regulation (GDPR)
  See European Commission's website on GDPR, Trunomi's GDPR Portal and Official Text of the General Data Protection Regulation
- Summary & Concluding Remarks

▸ Start ▸ Appendix

[1] Note: This lecture integrates knowledge from several further sources (cited where used, except if my own).

Legal, Privacy and Ethical Considerations: Legal Issues[15]

Legal Issues to Consider

- What is the value of an expert opinion in court when the expertise is encoded in a computer?
- Who is liable for wrong advice (or information) provided by an intelligent application?
- What happens if a manager enters an incorrect judgment value into an analytic application?
- Who owns the knowledge in a knowledge base?
- Can management force experts to contribute their expertise?

[15] Overview from [Sharda et al., 2018, Chapter 8.5].

Legal, Privacy and Ethical Considerations: Privacy Considerations[16]

Definitions of Privacy

“The claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others” [Westin, 1968]

“The right of an individual to be secure from unauthorised disclosure of information about oneself that is contained in an electronic repository” [Bertino et al., 2008]

“The right to be left alone and the right to be free from unreasonable personal intrusions” [Sharda et al., 2018]

Considerations

- Collecting information about individuals: How much is too much?
- Mobile User Privacy: Location-based analysis/profiling
- Homeland Security and Individual Privacy
- Recent Issues in Privacy and Analytics
- Who owns our private data?

[16] Overview from [Sharda et al., 2018, Chapter 8.5].

Legal, Privacy and Ethical Considerations: Ethical Considerations[17]

Model for Ethical Reasoning [Mason and Culnan, 1995]

Fundamental questions:

- Who is the agent?
- What action was actually taken or is being contemplated?
- What are the results or consequences of the act?
- Is the result fair or just for all stakeholders?

Hierarchy of ethical reasoning:

- Each ethical judgment or action is based on rules and codes of ethics
- These rules and codes themselves are based on principles
- Principles in turn are grounded in ethical theory

[17] Overview from [Sharda et al., 2018, Chapter 8.5].

Business Intelligence and Data Science: Outlook to Further Courses

Analytics Tasks

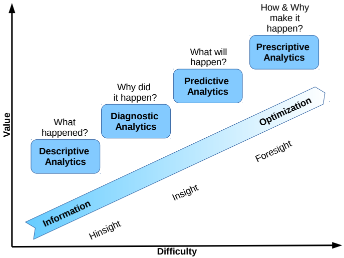
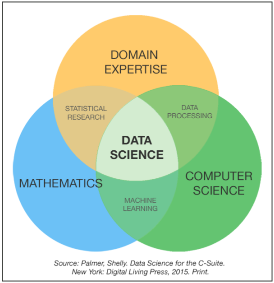


Figure: Analytic Ascendancy Model
(based on Gartner's model [Laney, 2012])

Data Science Competences



,

Figure: Data Science Venn Diagram
(Interpretation and Source: [Palmer, 2015]

Any More Questions?

Thank you!

Appendix

China's Social Credit Score

> *"Part financial credibility indicator and part compliance mechanism, the social credit system aims to generate a score for individuals and institutions in China based on data like tax filings and driving demerits. And while consumers may reap rewards, the score also functions as a signal mechanism for authorities about whom or what deserves to be penalized." CNBC 2017/03/16* [18]

- China is developing a "**Social Credit Score**"
- Nation-wide program is scheduled to begin in 2020
- Benefits for "good behaviour" include, e.g., waivers of security deposits for various rentals
- Problematic for freedom of speech, e.g.,

> *"My social-media account has been canceled many times, so the government can say I am a dishonest person," he said. "Then I can't go abroad and can't take the train." (Chinese novelist, in Washington Post, 2016/10/20)*

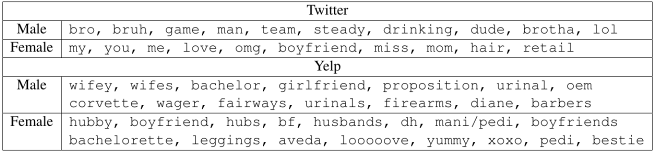
[18] See `https://www.scu.edu/ethics/focus-areas/internet-ethics/resources/chinas-social-credit-score/` and `http://www.cnbc.com/2017/03/16/china-social-credit-system-ant-financials-sesame-credit-and-others-give-scores-that-go-beyond-fico.html`

DM/ML for Revealing and Hiding Sensitive Information[20]

Inferring Personal Information from Text[19]

- Given two tweets, is the writer female/male?
  - "OMG I'm sooooo excited!!!"
  - "dude I'm so stoked."

- "Obfuscating Gender in Social Media Writing" [Reddy and Knight, 2016]:



	Twitter
Male	bro, bruh, game, man, team, steady, drinking, dude, brotha, lol
Female	my, you, me, love, omg, boyfriend, miss, mom, hair, retail
	Yelp
Male	wifey, wifes, bachelor, girlfriend, proposition, urinal, oem corvette, wager, fairways, urinals, firearms, diane, barbers
Female	hubby, boyfriend, hubs, bf, husbands, dh, mani/pedi, boyfriends bachelorette, leggings, aveda, looooove, yummy, xoxo, pedi, bestie

Figure: Words with highest gender associations (Source: [Reddy and Knight, 2016, Table 1]

- Machine Learning-based Classifier:
  - multiple exclamation marks, *omg*, lengthening intensifier ⇒ female
  - *dude*, *stoked*, no lengthening/exclamation ⇒ male

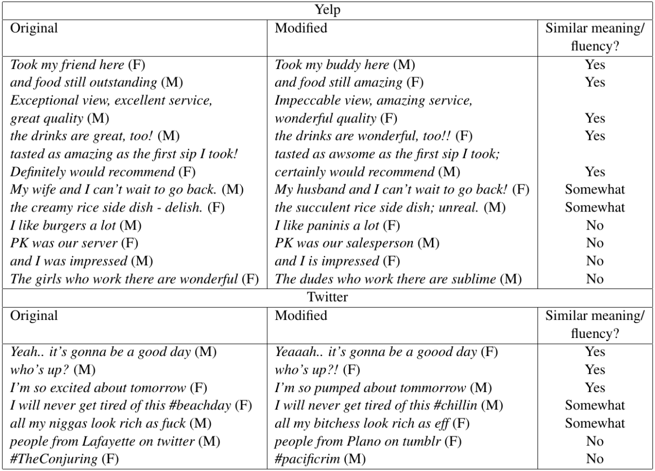
- Such rules can be learned to **reveal sensitive information**, e.g., age, sex, (political) views, . . .

- However, we can also use such a machine learning classifier to learn **How to conceal/obfuscate sensitive information!**

[19] See [Reddy and Knight, 2016]

[20] See [Mäder, 2018] (in German).

Hiding Sensitive Information[21]



Yelp

Original	Modified	Similar meaning/ fluency?
Took my friend here (F)	*Took my buddy here* (M)	Yes
and food still outstanding (M)	*and food still amazing* (F)	Yes
Exceptional view, excellent service, great quality (M)	*Impeccable view, amazing service, wonderful quality* (F)	Yes
the drinks are great, too! (M)	*the drinks are wonderful, too!!* (F)	Yes
tasted as amazing as the first sip I took! Definitely would recommend (F)	*tasted as awsome as the first sip I took; certainly would recommend* (M)	Yes
My wife and I can't wait to go back. (M)	*My husband and I can't wait to go back!* (F)	Somewhat
the creamy rice side dish - delish. (F)	*the succulent rice side dish; unreal.* (M)	Somewhat
I like burgers a lot (M)	*I like paninis a lot* (F)	No
PK was our server (F)	*PK was our salesperson* (M)	No
and I was impressed (M)	*and I is impressed* (F)	No
The girls who work there are wonderful (F)	*The dudes who work there are sublime* (M)	No

Twitter

Original	Modified	Similar meaning/ fluency?
Yeah.. it's gonna be a good day (M)	*Yeaaah.. it's gonna be a goood day* (F)	Yes
who's up? (M)	*who's up?!* (F)	Yes
I'm so excited about tomorrow (F)	*I'm so pumped about tommorrow* (M)	Yes
I will never get tired of this #beachday (F)	*I will never get tired of this #chillin* (M)	Somewhat
all my niggas look rich as fuck (M)	*all my bitchess look rich as eff* (F)	Somewhat
people from Lafayette on twitter (M)	*people from Plano on tumblr* (F)	No
#TheConjuring (F)	*#pacificrim* (M)	No

Figure: Examples for Original and Obfuscated Text (Source: [Reddy and Knight, 2016, Table 3]

[21] See [Reddy and Knight, 2016]

Bibliography I

Bertino, E., Lin, D., and Jiang, W. (2008).
A survey of quantification of privacy preserving data mining algorithms, page 183–205.
Springer.

De Cock Buning, M. and et al. (2018).
A multi-dimensional approach to disinformation.
Technical report, Independent High level Group on fake news and online disinformation.

Dwork, C. and Mulli, D. K. (2013).
It's not privacy, and it's not fair.
66 Stanford Law Review, 35.

Garcia, D. (2017).
Leaking privacy and shadow profiles in online social networks.
Science Advances, 3(8).

Hand, D. J., Mannila, H., and Smyth, P. (2001).
Principles of Data Mining.
Adaptive Computation and Machine Learning. The MIT Press.

High-Level Expert Group on AI (2019).
Ethics guidelines for trustworthy AI.

Bibliography II

Hill, K. (2012).
How target figured out a teen girl was pregnant before her father did.
Forbes.

Kosinski, M., Stillwell, D., and Graepel, T. (2013).
Private traits and attributes are predictable from digital records of human behavior.
Proc. of the Nat. Ac. of Sciences (PNAS), 110:5802–5805.

Laney, D. (2012).
Information economics, big data and the art of the possible with analytics.
Presentation copyrighted by Gartner.

Mäder, A. (2018).
Computer lesen auch zwischen den zeilen.
Spektrum der Wissenschaften, Mäders Moralfragen, 4.6.2018.

Mason, R. O. and Culnan, M. J. (1995).
Ethics of Information Management.
Sage.

Mendes, R. and Vilela, J. P. (2017).
Privacy-preserving data mining: Methods, metrics, and applications.
IEEE Access, 5:10562–10582.

Bibliography III

Palmer, S. (2015).
Are you ready for data science?
Blog Entry for Huffpost, posted on 03/10/2015 09:34 pm ET, updated Dec 06, 2017.

Reddy, S. and Knight, K. (2016).
Obfuscating gender in social media writing.
In *Proceedings of the First Workshop on NLP and Computational Social Science*, page 17–26. Association for Computational Linguistics.

Regalado, A. (2015).
Genome study predicts dna of the whole of iceland.
Technology Review, March 2015.

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016).
Why should i trust you?: Explaining the predictions of any classifier.
In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 1135–1144. ACM.

Schneier, B. (2007).
Why anonymous data sometimes isn't.
Wired Security.

Bibliography IV

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2016).
Grad-cam: Visual explanations from deep networks via gradient-based localization.
arXiv preprint arXiv:1610.02391.

Sharda, R., Delen, D., and Turban, E. (2018).
Business Intelligence, Analytics, and Data Science: A Managerial Perspective.
Pearson, 4 edition.

Sherman, R. (2015).
Business Intelligence Guidebook: From Data Integration to Analytics.
Morgan Kaufmann.

Veale, M. and Binns, R. (2017).
Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data.
Big Data and Society, 4(2).

Wachter, S., Mittelstadt, B., and Russell, C. (2018).
Counterfactual explanations without opening the black box: Automated decisions and the gdpr.
Harvard Journal of Law and Technology, 31(2).

Bibliography V

Westin, A. F. (1968).
Privacy and freedom.
Washington Lee Law Rev., 25.

Winston, W. L. (1997).
Operations Research: Applications and Algorithms.
Wadsworth Publishing Company, 3rd edition edition.

Zwitter, A. (2014).
Big data ethics.
Big Data and Society, 1(2).