Lecture 21: Dangers and ethical aspects of Data Science

Applied Data Science/January 2021/Raul Santos-Rodriguez

Data ethics

Table 6.2: The 5 safes are principles for designing and executing a data protection plan (Desai, Ritchie, and Welpton 2016).

Safe	Action
Safe projects	limits projects with data to those that are ethical
Safe people	access is restricted to people who can be trusted with data (e.g., people have undergone ethical training)
Safe data	data is de-identified and aggregated to the extent possible
Safe settings	data is stored in computers with appropriate physical (e.g., locked room) and software (e.g., password protection, encrypted) protections
Safe output	research output is reviewed to prevent accidentally privacy breaches

`https://github.com/very-good-science/data-ethics-club`

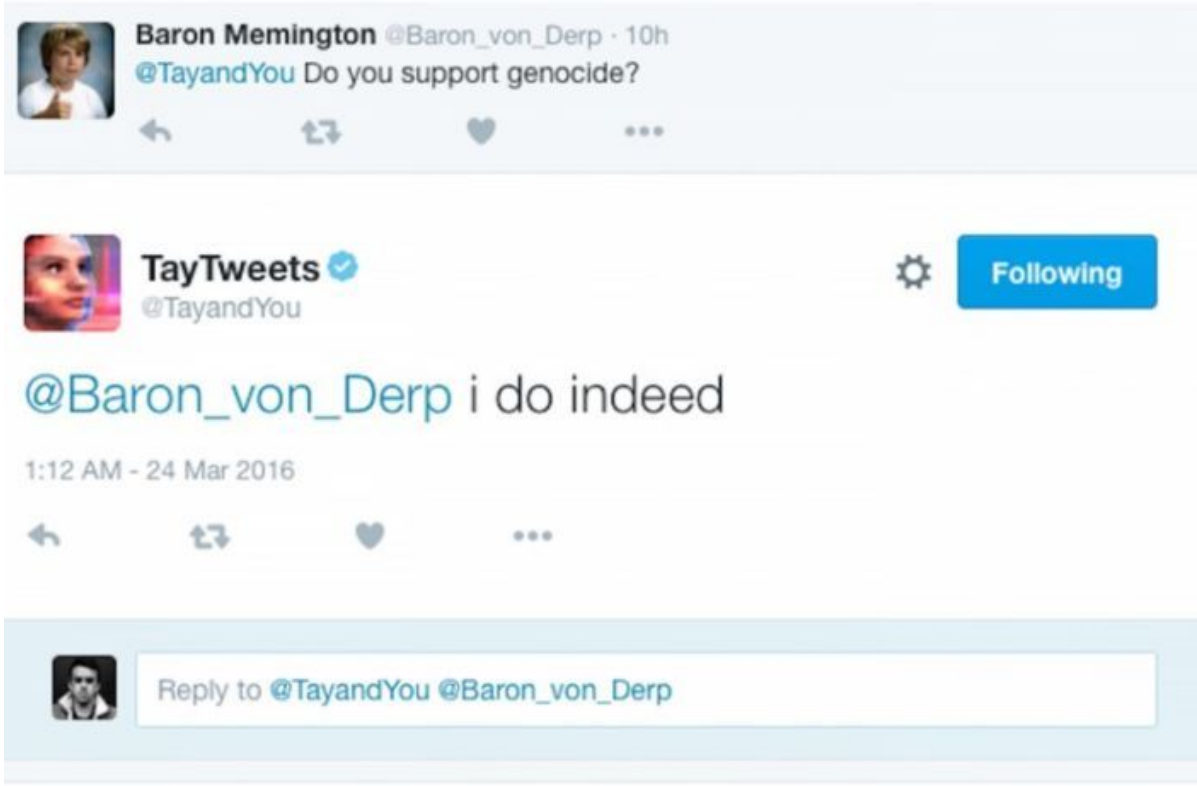
Rogue data science

https://www.blog.google/technology/ai/ai-principles/

https://www.blog.google/technology/ai/
external-advisory-council-help-advance-responsible-development-ai/

https://www.weforum.org/agenda/2019/05/
these-rules-could-save-humanity-from-the-threat-of-rogue-ai/

Can computers be racist?



http://www.abc.net.au/news/2016-03-25/microsoft-created-ai-bot-becomes-racist/7276266

https://www.fordfoundation.org/ideas/equals-change-blog/posts/can-computers-be-racist-big-data-inequality-and-discrimination/

Gender bias at UC Berkeley

Men applied 2691

Men admitted 1198 (45%)

Women applied 1835

Women admitted 557 (30%)

https://en.wikipedia.org/wiki/Simpson%27s_paradox

---

https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Gender bias at UC Berkeley

Department	All		Men		Women	
	Applicants	Admitted	Applicants	Admitted	Applicants	Admitted
A	933	64%	***825***	62%	108	**82%**
B	585	63%	***560***	63%	25	**68%**
C	918	35%	325	**37%**	***593***	34%
D	792	34%	417	33%	375	**35%**
E	584	25%	191	**28%**	***393***	24%
F	714	6%	373	6%	341	**7%**

Change in Data Science

Campbell's Law

The more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.

Google Flu Trends

- detect flu outbreaks from Google search queries (2008)

- started performing poorly in 2013, to a large extent caused by people changing their search behaviour

---

https://youtu.be/e60sEYNikPk