

BIG DATA

CHAPTER 9: CONTROL

A BIT OF HISTORY

- * Before Gutenberg designed the first printer, books were controlled by the Catholic Church.
- * The Bible became a best seller.
- * Governments started to establish censorship and authors requested copyright.

A BIT OF HISTORY

- * In the XIX century, intellectuals for speech of freedom that brings responsibility.
- * Nowadays, this is regulated by laws controlling the misuse of information.
- * With the big data, new professions could emerge to interpret the algorithms under big data and warn people of the dangers.

BIG DATA GOVERNANCE

- * PRIVACY

- * PROPENSITY

- * ALGORITHM AUDITING

FROM PRIVACY TO ACCOUNTABILITY

- * Each individual person has the right of controlling his information.
- * This is not useful in big data because the information can be reused in different ways that were not specified when they were collected.
- * The authors suggest treat data based in control and avoiding harm persons who ceded the data.

PROPENSITY

- * Authors mention that Big data cannot be used in justice.
- * Justice cannot condemn a person based on Big Data.



BREAKING THE BLACK BOX ALGORITHM AUDITING

- * Behavior of the systems can be predicted and analyzed by the right persons.
- * The answer for one question might use different algorithms and data sets so the traceability tends to be extremely difficult.
- * Big data requires a group of people to control the companies and create a set of procedures

BREAKING THE BLACK BOX ALGORITHM AUDITING

- * New professions related to computer science, mathematics and statistics will appear.



BREAKING THE BLACK BOX ALGORITHM AUDITING

- * External algorithmist will offer an impartial vision of the accuracy and validity of the predictions made by using Big data.
- * They will be under a code of conduct.
- * In case of breaking one of the statements, the algorithmists could face legal problems.

BREAKING THE BLACK BOX ALGORITHM AUDITING

- * Internal algorithmist work inside the company observing the activities and avoiding possible harms to the owners of the data.
- * They will require freedom and impartiality.
- * They will protect people of misusing of data.

What can we avoid using this
methods?



selena gomez is a |

selena gomez is a **mexican**

selena gomez is a **good role model**

selena gomez is a **cancer**

Press Enter to search.



politicians are |

politicians are **like diapers**

politicians are **corrupt**

politicians are **scum**

politicians are **lizards**

Press Enter to search.



obama is a

obama is a **robot**

obama is a **cactus**

obama is a **trader**

obama is a **gdi tank**

Press Enter to search.

NOW THE QUESTION