

# Big Data, Data Science and The Great A.I. Awakening

-Ethical and societal implications-



Prof. Roberto V. Zicari  
Frankfurt Big Data Lab  
[www.bigdata.uni-frankfurt.de](http://www.bigdata.uni-frankfurt.de)

Int`l CAPP Conference  
Berlin, November 16, 2017

# Data as an Economic Asset



*“I think we’re just beginning to grapple with implications of data as an economic asset”*

*-Steve Lohr (The New York Times)*

# Data and Economic Power



- ❧ The companies with big data pools can have great economic power.
- ❧ Today, that shortlist would include Google, Microsoft, Facebook, Amazon, Apple and Baidu.

# What is more important, vast data pools, sophisticated algorithms or deep pockets?



☞ “No one can replicate your data. It’s the defensible barrier, not algorithms.”

-- Andrew Ng, Stanford professor



# Algorithms and Data



“AI is akin to building a rocket ship. You need a huge engine and a lot of fuel. The rocket engine is the learning algorithms but the fuel is the huge amounts of data we can feed to these algorithms.”

-- Andrew Ng

# Interplay and implications of big data and Artificial Intelligence



- ❧ What is the interplay and implications of big data and artificial intelligence?
- ❧ The data revolution has made the recent AI advances possible.

# AI and Big Data for higher-stakes decisions



- ❧ Technology is moving beyond increasing the odds of making a sale, to being used in higher-stakes decisions like medical diagnosis, loan approvals, hiring and crime prevention.
- ❧ *What are the societal implications of this?*

# Marketing and advertising



- œ One stage in the life cycle of an emerging science.  
**Marketing is a low-risk - and, yes, lucrative**

In marketing and advertising, a decision that is better on average is plenty good enough.

You've increased sales and made more money.  
**You don't really have to know why.**

--Steve Lohr



What is the difference between using data for marketing and for other more critical decisions?



☞ *“What happens if my algorithm is wrong? Someone sees the wrong ad. What’s the harm? It’s not a false positive for breast cancer.”*

-- *Claudia Perlich*, Chief scientist at an ad-targeting start-up in New York, Dstillery.

# Ethical Implications



- ❧ But the other decisions are **practically** and **ethically** very different.
- ❧ These are crucial decisions about **individual people's lives**. Better on average isn't good enough.
- ❧ For these kinds of decisions, issues of **accuracy, fairness and discrimination** come into play.

# Automate or Augment humans?



- ❧ You can use AI technologies either to **automate** or to **augment** humans.
- ❧ Are **computer system designers** (i.e. **Software Developers, Software Engineers, Data Scientists**), the ones who will decide what the impact of these technologies are and whether to replace or augment humans in society?

# What are the ethical responsibilities of designers of intelligent systems?



☞ “I think the most important aspect of this question is the simple **acknowledgement** that **intelligent system designers do have ethical responsibilities.**”

--John Markoff (The New York Times)

“We have a **profound ethical responsibility** to design systems that have a positive impact on society, obey the law,  
and adhere to our highest ethical standards.”

-Oren Etzioni

Chief Executive Officer [Allen Institute for Artificial Intelligence.](#)



# Ethics by Design?



## ❧ **Do we need some sort of auditing tool?**

The technology has to be able to explain itself, to explain how a data-driven algorithm came to the decision or recommendation that it did.

How much **Transparency** is desired?

Do we wish “**Human in the loop**” for most of these kinds of decisions for the foreseeable future?

# If humans delegate decisions to machines, who will be responsible for the consequences?



- ❧ Ben Shneiderman (University of Maryland) argues against autonomous systems.
- ❧ His point is that it is essential to **keep a human in the loop**. If not you run the **risk of abdicating ethical responsibility for system design**.
- ❧ **Is it realistic?** If something can be partially automated, will it eventually be fully automated?
- ❧ Do we need to **regulate the development of artificial intelligence?**

# Human Motivations



*“It is absolutely essential that we control the machines, and every indication is that we will be able to do so in the foreseeable future. I do worry about human motivations too. Someone said: I’m not worried about robots deciding to kill people, I’m worried about politicians deciding robots should kill people.”*

-- Oren Etzioni

# Who is responsible?



∞ Data, AI and Intelligent systems are becoming sophisticated tools in the hands of a variety of stakeholders, including political leaders.



# Good Intention vs. Good Data



- ❧ *Good data* reflects reality and thus can help us gain insights into how the world works. That does not make such discovery ethical, even though the discover is correct.
- ❧ *Good intentions* point towards an ethical use of data, which helps protect us against unethical data uses, but does not prevent false big data analysis. This is a long way of saying we need both, albeit for different reasons.

--Viktor Mayer-Schönberger, Professor of Internet Governance and Regulation at Oxford University (UK).

The thing that **motivates** my **actions** will determine the **direction** I am going



# The Individual and Collective Conscience



The **individual and collective conscience** is the existential place where the most significant things happen.

**Research, Change, Decision and Choice** can take two diametrically opposite directions:

can be either “pro or contra” the human person”.

Source: "The good society and the future of jobs: Can solidarity and fraternity be part of business decisions?" MAY 8 -10, 2014 – Vatican City

# Big Data for the Common Good



- ❧ Is there a real potential in using data-driven methods to both help charities, civil society organizations the government, develop better services and products, and understand civil society activity?
- ❧ What are the key lessons and recommendations for future work in this space?
- ❧ We need more researchers team up with decision makers in charities, and more broadly civil society organizations and the government to utilize Big Data to improve our understanding of the key challenges that our society is facing.
- ❧ We also need decision makers and especially policy makers to better understand the power of Big Data.



# “Data for Humanity”



✧ An initiative I started with **Andrej Zwitter** (Professor at the University of Groningen) at the end of 2015, with the goal to:

*bring people and institutions together who share the **motivation to use data for the common good.***

which calls for the use of **five ethical principles for the use of data-** has reached more than 1,000 signatories worldwide.

[www.bigdata.uni-frankfurt.de/dataforhumanity/](http://www.bigdata.uni-frankfurt.de/dataforhumanity/)

# Five ethical principles when using data



- ❧ Do no harm
- ❧ Use data to help create peaceful coexistence
- ❧ Use data to help vulnerable people and people in need
- ❧ Use data to preserve and improve natural environment
- ❧ Use data to help create a fair world without discrimination