FROM PRIVACY TO ALGORITHMS’ FAIRNESS
Some Hints on the Complex Journey of ICT Ethics

Chiara Sabelli (SISSA)
Mariachiara Tallacchini (UCSC)
Premise 1: :)

Premise 2: Ethics here is more the dialogue between informal and formal normativity, soft law and hard law

Privacy and its meanings, and data protection as the prevailing value/right/answer in ICT

Emergence of limits of data protection: big data and machine learning, IoT, ...

Fairness as the new mantra in EU and US

Algorithms’ fairness and its ambiguity

Whose fairness?
THE RIGHT TO PRIVACY.

"It could be done only on principles of private justice, moral fitness, and public convenience, which, when applied to a new subject, make common law without a precedent; much more when received and approved by usage."

WILLES, J., in Millar v. Taylor, 4 Burr. 2303, 2312.

THAT the individual shall have full protection in person and in property is a principle as old as the common law; but it has been found necessary from time to time to define anew the exact nature and extent of such protection. Political, social, and economic changes entail the recognition of new rights, and the common law, in its eternal youth, grows to meet the demands of society. Thus, in very early times, the law gave a remedy only

Samuel D. Warren,
Louis D. Brandeis.

Boston, December, 1890.
Privacy as data protection in ICT
Privacy as data protection in ICT

Council of Europe Convention for the protection of individuals with regard to automatic processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)
Privacy as data protection in ICT

Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Council of Europe Convention for the protection of individuals with regard to automatic processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)
Privacy as data protection in ICT

- Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data
- Council of Europe Convention for the protection of individuals with regard to automatic processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)
- Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling of 23 November 2010
Privacy as data protection in ICT

Council of Europe Convention for the protection of individuals with regard to the processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)

Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data and on the free movement of such data of profiling of 23 November 2010

REGULATION (EU) 2016/679 ON THE PROTECTION OF NATURAL PERSONS, AND REPEALING DIRECTIVE 95/46/EC (GENERAL DATA PROTECTION REGULATION)
Privacy as data protection in ICT

Council of Europe Convention for the protection of individuals with regard to the processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)

Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling of 23 November 2010

REGULATION (EU) 2016/679 ON THE PROTECTION OF NATURAL PERSONS,..., AND REPEALING DIRECTIVE 95/46/EC (GENERAL DATA PROTECTION REGULATION)

The Health Insurance Portability and Accountability Act (HIPAA) (42 U.S.C. §1301 et seq.)
Privacy as data protection in ICT

Council of Europe Convention for the protection of individuals with regard to the processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)

Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data and profiling of 23 November 2010

REGULATION (EU) 2016/679 ON THE PROTECTION OF NATURAL PERSONS..., AND REPEALING DIRECTIVE 95/46/EC (GENERAL DATA PROTECTION REGULATION)

The Health Insurance Portability and Accountability Act (HIPAA) (42 U.S.C. §1301 et seq.)

The Financial Services Modernization Act (Gramm-Leach-Bliley Act (GLB)) (15 U.S.C. §§6801-6827)
Privacy as data protection in ICT

Council of Europe Convention for the protection of individuals with regard to the processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)

Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling of 23 November 2010

REGULATION (EU) 2016/679 ON THE PROTECTION OF NATURAL PERSONS..., AND REPEALING DIRECTIVE 95/46/EC (GENERAL DATA PROTECTION REGULATION)

The Health Insurance Portability and Accountability Act (HIPAA) (42 U.S.C. §1301 et seq.)

The Financial Services Modernization Act (Gramm-Leach-Bliley Act (GLB)) (15 U.S.C. §§6801-6827)

The Federal Trade Commission Act (15 U.S.C. §§41-58) (FTC Act)
Privacy as data protection in ICT

Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Council of Europe Convention for the protection of individuals with regard to automatic processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)

Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling of 23 November 2010

REGULATION (EU) 2016/679 ON THE PROTECTION OF NATURAL PERSONS..., AND REPEALING DIRECTIVE 95/46/EC (GENERAL DATA PROTECTION REGULATION)

The Health Insurance Portability and Accountability Act (HIPAA) (42 U.S.C. §1301 et seq.)

The Financial Services Modernization Act (Gramm-Leach-Bliley Act (GLB)) (15 U.S.C. §§6801-6827)

The Federal Trade Commission Act (15 U.S.C. §§41-58) (FTC Act)

Children's Online Privacy Protection Act (COPPA) (15 U.S.C. §§6501-6506)
Privacy as data protection in ICT

Council of Europe Convention for the protection of individuals with regard to the processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)

Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling of 23 November 2010

The Health Insurance Portability and Accountability Act (HIPAA) (42 U.S.C. §1301 et seq.)

The Standards for Privacy of Individually Identifiable Health Information (HIPAA Privacy Rule) (45 C.F.R. Parts 160 and 164) apply to the collection and use of protected health information (PHI) (HIPAA rules revised 2013)

The Financial Services Modernization Act (Gramm-Leach-Bliley Act (GLB)) (15 U.S.C. §6801-6827)

The Federal Trade Commission Act (15 U.S.C. §§41-58) (FTC Act)

Children’s Online Privacy Protection Act (COPPA) (15 U.S.C. §§6501-6506)

The Standards for Privacy of Individually Identifiable Health Information (HIPAA Privacy Rule) (45 C.F.R. Parts 160 and 164) apply to the collection and use of protected health information (PHI) (HIPAA rules revised 2013)

REGULATION (EU) 2016/679 ON THE PROTECTION OF NATURAL PERSONS...., AND REPEALING DIRECTIVE 95/46/EC (GENERAL DATA PROTECTION REGULATION)
Privacy as data protection in ICT

Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Council of Europe Convention for the protection of individuals with regard to automatic processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)

Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling of 23 November 2010

Regulation (EU) 2016/679 on the protection of natural persons and repealing Directive 95/46/EC (General Data Protection Regulation)

The Health Insurance Portability and Accountability Act (HIPAA) (42 U.S.C. §1301 et seq.)

The Standards for Privacy of Individually Identifiable Health Information (HIPAA Privacy Rule) (45 C.F.R. Parts 160 and 164) apply to the collection and use of protected health information (PHI) (HIPAA rules revised 2013)

The Health Insurance Portability and Accountability Act (HIPAA Omnibus Rule revision of Security and Accountability Act (HBreach Notification Rule (45 C.F.R. Part 164))

The Financial Services Modernization Act (Gramm-Leach-Bliley Act (GLB)) (15 U.S.C. §§6801-6827)

The Standards for Privacy of Individually Identifiable Health Information (HIPAA Privacy Rule) (45 C.F.R. Parts 160 and 164) apply to the collection and use of protected health information (PHI) (HIPAA rules revised 2013)

The Federal Trade Commission Act (15 U.S.C. §§41-58) (FTC Act)

Children’s Online Privacy Protection Act (COPPA) (15 U.S.C. §§6501-6506)
Privacy as data protection in ICT

- Council of Europe Convention for the protection of individuals with regard to the processing of personal data of 28 January 1981 (ETS No 108) and its Additional Protocol of 8 November 2001 (ETS No 181)
- Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data
- Recommendation CM/Rec(2010)13 of the Committee of Ministers of the Council of Europe to Member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling of 23 November 2010
- REGULATION (EU) 2016/679 ON THE PROTECTION OF NATURAL PERSONS...., AND REPEALING DIRECTIVE 95/46/EC (GENERAL DATA PROTECTION REGULATION)

The Health Insurance Portability and Accountability Act (HIPAA) (42 U.S.C. §1301 et seq.)
- The Federal Trade Commission Act (15 U.S.C. §§41-58) (FTC Act)
- The Standards for Privacy of Individually Identifiable Health Information (HIPAA Privacy Rule) (45 C.F.R. Parts 160 and 164) apply to the collection and use of protected health information (PHI) (HIPAA rules revised 2013)
- The Health Insurance Portability and Accountability Act (HIPAA Omnibus Rule revision of Security and Accountability Act (HBreach Notification Rule (45 C.F.R. Part 164) (45 C.F.R. Part 164)
- The Fair Credit Reporting Act (15 U.S.C. §1681) (and the Fair and Accurate Credit Transactions Act (Pub. L. No. 108-159)
- The Children's Online Privacy Protection Act (COPPA) (15 U.S.C. §§6501-6506)
Privacy differences US/EU

**US**
- Born as a **negative right** to be let alone
- Used to cover issues of **autonomy and integrity**: abortion, reproductive technologies → non-interference, integrity, autonomy
- Evolved toward: **consumer right** data protection
- Protected primarily through TOS and self-regulations

---

**EU**
- Born as a **positive right** protection of integrity/dignity
- Used to cover some situations: all ICT issues, right to information
- Evolved toward: **fundamental human rights** → privacy + data protection
- Protected through mandatory general legislation
(The construction of) privacy meanings in the US and COE/EU

U.S. Constitution

Amendment 14 (Ratified 07/09/1868)
...nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.

Amendment 4 (Ratified 12/15/1791)
The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches

European Convention on Human Rights
Article 8 – Right to respect for private and family life (and right to information)
1. Everyone has the right to respect for his private and family life, his home and his correspondence.

Charter of the Fundamental Rights of the European Union
Article 7 - Respect for private and family life
Everyone has the right to respect for his or her private and family life, home and communications.

Article 8 - Protection of personal data
1. Everyone has the right to the protection of personal data concerning him or her.
2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law.
What privacy can do ...

**US**

Roe v. Wade (USSC 1973)

This right of privacy, ... founded in the Fourteenth Amendment’s concept of personal liberty (as a fundamental right).... is broad enough to encompass a woman’s decision whether or not to terminate her pregnancy.

**EU**

ECHR (1998) - Case of Guerra and Others v. Italy

Article 8

Failure to provide local population with information about risk factor and how to proceed in event of an accident at nearby chemical factory
"The right to be let alone is indeed the beginning of all freedom"

(EDPS Meeting the challenges of big data 7/2015)

HARVARD
LAW REVIEW.

Vol. IV. DECEMBER 15, 1890. No. 5.

THE RIGHT TO PRIVACY.

"It could be done only on principles of private justice, moral fitness, and public convenience, which, when applied to a new subject, make common law without a precedent; much more when received and approved by usage."

WILLIS, J., in Millar v. Taylor, 4 Burr. 2303, 2313.

THAT the individual shall have full protection in person and in property is a principle as old as the common law; but it has been found necessary from time to time to define anew the exact nature and extent of such protection. Political, social, and economic changes entail the recognition of new rights, and the common law, in its eternal youth, grows to meet the demands of society. Thus, in very early times, the law gave a remedy only

Samuel D. Warren,
Louis D. Brandeis.

Boston, December, 1890.
When two notions of privacy diverge

The two interpretation of the right to privacy diverge
• emotion: the Facebook case on PNAS
• information: filter bubble / echo-chamber (Facebook)
Emotional contagion experiment

Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock

Abstract

Emotional states can be transmitted from one individual to another, even when this interaction occurs without direct communication (i.e., via social networks). However, the extent to which people are affected by the emotional states of others is currently unknown. This study used Facebook to experimentally manipulate exposure to emotional content and measured the extent to which this exposure affected the emotional valence of posts made by Facebook users. The results demonstrate that emotional contagion is widespread and robust, occurring across large social networks, even when emotions are not expressed directly. The findings suggest that emotional contagion may be an important mechanism for the spread of spoken language, and that these effects may occur on a massive scale. The implications of these findings for our understanding of social influence and communication are discussed.

Results

The experiments demonstrated that emotional contagion occurs through social networks, even when emotions are not expressed directly. The results showed that exposure to emotional content led people to post content that was consistent with the exposure. The findings suggest that emotional contagion may be an important mechanism for the spread of spoken language, and that these effects may occur on a massive scale. The implications of these findings for our understanding of social influence and communication are discussed.
Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer\textsuperscript{a,1}, Jamie E. Guillory\textsuperscript{b,2}, and Jeffrey T. Hancock\textsuperscript{b,c}

\textsuperscript{a}Core Data Science Team, Facebook, Inc., Menlo Park, CA 94025; and Departments of \textsuperscript{b}Communication and \textsuperscript{c}Information Science, Cornell University, Ithaca, NY 14853

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

criticism

Study should have been retracted by PNAS because of lack on IC

defence

posts were analysed with a text mining algorithm […] such that no text was seen by the researchers

“As such, it was consistent with Facebook’s Data Use Policy, to which all users agree prior to creating an account on Facebook, constituting informed consent for this research.”
Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer\textsuperscript{a,1}, Jamie E. Guillory\textsuperscript{b,2}, and Jeffrey T. Hancock\textsuperscript{b,c}

\textsuperscript{a}Core Data Science Team, Facebook, Inc., Menlo Park, CA 94025; and Departments of \textsuperscript{b}Communication and \textsuperscript{c}Information Science, Cornell University, Ithaca, NY 14853

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

Study should have been retracted by PNAS because of lack on IC

even if privacy in the sense of data protection was not breached, the autonomy/integrity of the users was undermined

defence

posts were analysed with a text mining algorithm [...] such that no text was seen by the researchers

“As such, it was consistent with Facebook’s Data Use Policy, to which all users agree prior to creating an account on Facebook, constituting informed consent for this research.”
Obtaining informed consent and allowing participants to opt out are best practices.

As a private company Facebook was not under the rules for research on human subjects (TOS, not IC).

It is nevertheless a matter of concern that the collection of the data by Facebook may have involved practices that were not fully consistent with the principles of obtaining informed consent and allowing participants to opt out.
The Filter Bubble

Facebook learns what is relevant for you studying your actions (likes, comments, shares) and gives you a personalised “information diet”. Facebook algorithms act as “editors”. [Pariser 2011]

Credit: Eli Pariser, TED Talk, march 2011
Measuring Filter Bubbles

Anatomy of news consumption on Facebook

Ana Lucia Schmidt\textsuperscript{a}, Fabiana Zollo\textsuperscript{a,1}, Michela Del Vicario\textsuperscript{a}, Alessandro Bessi\textsuperscript{b}, Antonio Scala\textsuperscript{a,c}, Guido Caldarelli\textsuperscript{a,c}, H. Eugene Stanley\textsuperscript{d}, and Walter Quattrociocchi\textsuperscript{a,2}

\textsuperscript{a}Laboratory of Computational Social Science, Networks Department, IMT Alti Studi Lucca, 55100 Lucca, Italy; \textsuperscript{b}IUSS Institute for Advanced Study, 27100 Pavia, Italy; \textsuperscript{c}IS-CNR Uos “Sapienza,” 00185 Rome, Italy; and \textsuperscript{d}Department of Physics, Boston University, Boston, MA 02115

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved January 31, 2017 (received for review October 14, 2016)
Anatomy of news consumption on Facebook

Ana Lucía Schmidt\textsuperscript{a}, Fabiana Zollo\textsuperscript{a,1}, Michela Del Vicario\textsuperscript{a}, Alessandro Bessi\textsuperscript{b}, Antonio Scala\textsuperscript{a,c}, Guido Caldarelli\textsuperscript{a,c}, H. Eugene Stanley\textsuperscript{d}, and Walter Quattrociocchi\textsuperscript{a,2}

\textsuperscript{a}Laboratory of Computational Social Science, Networks Department, IMT Alti Studi Lucca, 55100 Lucca, Italy; \textsuperscript{b}IUSS Institute for Advanced Study, 27100 Pavia, Italy; \textsuperscript{c}ISC-CNR Uos “Sapienza,” 00185 Rome, Italy; and \textsuperscript{d}Department of Physics, Boston University, Boston, MA 02115

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved January 31, 2017 (received for review October 14, 2016)

---

Fig. 2. Community structure. (Left) Backbone of the projections on pages of the users’ likes ($G_l^p$). (Right) Comments ($G_c^p$). The color of the nodes indicate the Fast Greedy community. Nodes in $G_l^p$ are ordered according to the detected communities, whereas in $G_c^p$, the nodes follow the same order as in $G_l^p$. 

---
Privacy in Big Data and Machine Learning
Privacy in Big Data and Machine Learning

Big Data and Machine Learning pose new privacy issues.
Privacy in Big Data and Machine Learning

Big Data and Machine Learning pose new privacy issues.

A private database (medical records, loan applications, ...) provides useful information for the population represented by the sample.
Privacy in Big Data and Machine Learning

Big Data and Machine Learning pose new privacy issues.

A private database (medical records, loan applications, …) contains useful information for the population represented by the sample.

Simple solutions could be:
Big Data and Machine Learning pose new privacy issues.

**private database** (medical records, loan applications, …)

useful information for the **population** represented by the sample

simple solutions could be:

1. **ANONIMYSING DATA** (Netflix [NS2008])
Big Data and Machine Learning pose new privacy issues.

_Anonymising data_ (Netflix [NS2008]) is _unsafe_.

Simple solutions could be:

1. **Anonymising data** (Netflix [NS2008])
Privacy in Big Data and Machine Learning

Big Data and Machine Learning pose new privacy issues.

**private database** (medical records, loan applications, …)

useful information for the population represented by the sample

simple solutions could be:

1. **ANONIMYSING DATA** (Netflix [NS2008])

2. **RELEASE A LOT OF STATISTICS ON RAW DATA** (Genome Wide Association Study [G2013])
Privacy in Big Data and Machine Learning

Big Data and Machine Learning pose new privacy issues.

![Diagram: Database with arrow pointing to useful information for the population represented by the sample]

private database (medical records, loan applications, ...)

simple solutions could be:

1. **ANONIMYSING DATA** (Netflix [NS2008])

2. **RELEASE A LOT OF STATISTICS ON RAW DATA** (Genome Wide Association Study [G2013])
Differential privacy

private database → sanitized information → data analysts

SANITIZER (algorithm)

[Dwork 2006]
Differential privacy

It has the property that if a single individual participates, the outcome "does not change"
Differential privacy

private database \rightarrow \text{sanitized information} \rightarrow \text{data analysts}

SANITIZER (algorithm)

it has the property that if a single individual participates, the outcome “does not change”

[Dwork 2006]
Differential privacy

SANITIZER (algorithm)

- it has the property that if a single individual participates, the outcome "does not change"

- data analysts can learn something new about the population as a whole
Differential privacy

SANITIZER (algorithm)

it has the property that if a single individual participates, the outcome "does not change"

data analysts can learn something new about the population as a whole

INDIVIDUALS MAY STILL BE AFFECTED BY THESE LEARNINGS
Differential privacy

[Dwork 2006]

private database

sanitized information

data analysts

it’s not privacy!
it’s not about humans,
it’s about algorithms

INDIVIDUALS MAY STILL BE AFFECTED BY THESE LEARNINGS
FAIRNESS
FAIRNESS

the new mantra
in the EU and the US
REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
of 27 April 2016

on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

Principles

Article 5

Principles relating to processing of personal data

1. Personal data shall be:

(a) processed lawfully, fairly and in a transparent manner in relation to the data subject ("lawfulness, fairness and transparency");

(b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ("purpose limitation");
Recommendation CM/Rec(2010)13
of the Committee of Ministers to member states
on the protection of individuals with regard to automatic processing
of personal data in the context of profiling

(Adopted by the Committee of Ministers on 23 November 2010
at the 1099th meeting of the Ministers’ Deputies)
Statement on Statement of the WP29 on the impact of the development of big data on the protection of individuals with regard to the processing of their personal data in the EU

Adopted on 16 September 2014
ARTICLE 29 DATA PROTECTION WORKING PARTY

Opinion 7/2015

Meeting the challenges of big data

A call for transparency, user control, data protection by design and accountability

on the impact of the development of big data with regard to the processing of their personal data in the EU

Adopted on 16 September 2014
Fairness in the EU

Opinion 7/2015
Meeting the challenges of big data
A call for transparency, user control, data protection by design and accountability
Adopted on 16 September 2015

Opinion 8/2016
EDPS Opinion on coherent enforcement of fundamental rights in the age of big data
Fairness in the EU

• Fairness: fair prices and fair processing of data

• The challenges and risks of big data therefore call for more effective data protection. The question is not whether to apply data protection law to big data, but rather how to apply it innovatively in new environments

• Transparency, user-control, user-friendly design, accountability,
EP on fundamental rights

Fundamental rights implications of big data

European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI))

Article 6 – Right to liberty
Article 10 – freedom of thought
Article 11 – freedom of expression and information
**EP on fundamental rights**

**Fundamental rights implications of big data**

European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI))

| Article 6 – Right to liberty |
| Article 10 – freedom of thought |
| Article 11 – freedom of expression and information |

Union law on .....the **right to equality and non-discrimination**, as well as the **right of individuals to receive information** about the logic involved in automated decision-making and profiling
EP on fundamental rights

**Fundamental rights implications of big data**

European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI))

Article 6 – Right to liberty
Article 10 – freedom of thought
Article 11 – freedom of expression and information

Union law on .....the **right to equality and non-discrimination**, as well as the **right of individuals to receive information** about the logic involved in automated decision-making and profiling

whereas big data ... also entails significant risks with regard to fundamental rights, such as ... **freedom of expression and non-discrimination**
Fundamental rights implications of big data

European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI))

Article 6 – Right to liberty
Article 10 – freedom of thought
Article 11 – freedom of expression and information

Union law on …..the right to equality and non-discrimination, as well as the right of individuals to receive information about the logic involved in automated decision-making and profiling

whereas big data … also entails significant risks with regard to fundamental rights, such as … freedom of expression and non-discrimination

… to ensure that data-driven technologies do not limit or discriminate access to a pluralistic media environment…
Fairness in the US

Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights

Executive Office of the President
May 2016
**Source of unfairness:** choice of inputs and choice of how algorithms work

1) Challenges relating to data used as inputs to an algorithm:

- Poorly selected data (certain data are important to the decision but not others)
- Incomplete, incorrect, or outdated data
- Selection bias (data inputs not representative of a population)
- Unintentional perpetuation and promotion of historical biases

2) Challenges related to how algorithms work (algorithms as proprietary black boxes, unknowable by the user):

- Poorly designed matching systems
- Personalization and recommendation services that narrow instead of expand user options
- Decision-making systems that assume correlation necessarily implies causation
- Data sets that lack information or disproportionately represent certain populations

Machine learning—the “science of getting computers to act without being explicitly programmed.”
“In data we trust”, but ... [Dwork 2013]

“Data fundamentalism,” the notion that correlation always indicates causation, and that massive data sets and predictive analytics always reflect objective truth.
“In data we trust”, but …

“Data fundamentalism,” the notion that correlation always indicates causation, and that massive data sets and predictive analytics always reflect objective truth.

BUT

Classification systems are neither neutral nor objective, but are biased toward their purposes.

While automated decision making systems “may reduce the impact of biased individuals, they may also normalise the far more massive impacts of system-level biases and blind spots.” [GJr2010]

Datasets and algorithms reflect choices about data, connections, inferences, interpretation, and thresholds for inclusion that advance a specific purpose.
Data as social mirror

TRAINING

historical data

TESTING

CLASSIFICATION

new data
"Weapons of math destruction"

2016 book by a former Wall Street analyst, now turned into an algorithmic auditor

"Models are opinions embedded in mathematics"

SCALE / OPACITY / DAMAGE MINORITIES / NEGATIVE FEED-BACK LOOP / DO NOT UNDERGO DEEP SCRUTINY

predictive policing
criminal justice

teachers evaluation
higher education

finding
a job

access to credit

[O’Neil 2016]
**Predictive policing**

**Problem:** reduction of local police agents.

The head of the local police wants to **optimise the patrolling strategy**
He invests in a predictive policing software from the company **PredPol**

**Based on the geography** of past crimes the algorithm predicts where it is most probable that the future crimes will happen.

But: **Geography is an excellent proxy for race!**
Predictive policing

**Problem:** reduction of local police agents.

The head of the local police wants to **optimise the patrolling strategy**
He invests in a predictive policing software from the company **PredPol**

**Based on the geography** of past crimes the algorithm predicts where it is most probable that the future crimes will happen.

But: **Geography is an excellent proxy for race!**
Arrested people fill a questionnaire and a **Risk Recidivism Model** is applied to the results to assess the criminal defendant’s likelihood of becoming a recidivist and consequently decide how long to sentence him/her to prison.

Starting with **COMPAS** scores for 10,000 criminal defendants in Broward County, Florida, they looked at the difference between who was predicted to get rearrested by COMPAS versus who actually did.

[ProPublica, May 2016]
Arrested people fill a questionnaire and a **Risk Recidivism Model** is applied to the results to assess the criminal defendant’s likelihood of becoming a recidivist and consequently decide how long to sentence him/her to prison.

Starting with **COMPAS** scores for 10,000 criminal defendants in Broward County, Florida, they looked at the difference between who was predicted to get rearrested by COMPAS versus who actually did.

[ProPublica, May 2016]
Arrested people fill a questionnaire and a **Risk Recidivism Model** is applied to the results to assess the criminal defendant’s likelihood of becoming a recidivist and consequently decide how long to sentence him/her to prison.

**Prediction Fails Differently for Black Defendants**

|                                | White | African American |
|--------------------------------|-------|------------------|
| Labeled Higher Risk, But Didn’t Re-Offend | 23.5% | 44.9%            |
| Labeled Lower Risk, Yet Did Re-Offend      | 47.7% | 28.0%            |

[ProPublica, May 2016]
Arrested people fill a questionnaire and a **Risk Recidivism Model** is applied to the results to assess the criminal defendant’s likelihood of becoming a recidivist and consequently decide how long to sentence him/her to prison.

### Prediction Fails Differently for Black Defendants

|                                    | WHITE  | AFRICAN AMERICAN |
|------------------------------------|--------|------------------|
| Labeled Higher Risk, But Didn’t Re-Offend | 23.5%  | 44.9%            |
| Labeled Lower Risk, Yet Did Re-Offend       | 47.7%  | 28.0%            |

[ProPublica, May 2016]
Evaluating public schools teachers

IMPACT:
evaluation program of DC public schools

Value Added Model:
from the data analysis company Mathematica Policy Research

Sarah Wisocki was fired at the end of her second year as a middle school teacher in DC, even if students and parents were happy about her.

she never had the chance to look at the algorithm!
Evaluating public schools teachers

IMPACT:
evaluation program of DC public schools

Value Added Model: from the data analysis company Mathematica Policy Research

Sarah Wisocki was fired at the of her second year as a middle school teacher in DC, even if students and parents were happy about her

she never had the chance to look at the algorithm!
Higher Education

US News Ranking of 1800 colleges and universities in the US: the ranking in short was destiny.

For-profit colleges: exploit ‘predatory advertising’ to target students eligible for federal education funds.

Corinthian College was a giant in the industry. California Attorney General complaint that it targeted ‘isolated, impatient individuals with low self-esteem’.
Governing algorithms

- Support research into mitigating algorithmic discrimination, building systems that support fairness and accountability, and developing strong data ethics frameworks.
- Encourage market participants to design the best algorithmic systems, including transparency and accountability mechanisms such as the ability for subjects to correct inaccurate data and appeal algorithmic-based decisions.
- Promote academic research and industry development of algorithmic auditing and external testing of big data systems to ensure that people are being treated fairly.
- Broaden participation in computer science and data science, including opportunities to improve basic fluencies and capabilities of all Americans.
- Consider the roles of the government and private sector in setting the rules of the road for how data is used.
While data systems should remove inappropriate human bias....

Careful attention should be paid:
- to ensure that the use of big data does not contribute to systematically disadvantaging certain groups;
- to avoid exacerbating biases by encoding them into technological systems.

Need to develop:
- a principle of “equal opportunity by design” designing data systems that promote fairness and safeguard against discrimination from the first step of the engineering process and continuing throughout their lifespan.
Fairness through awareness
Fairness through awareness
Fairness through awareness
Fairness through awareness
Fairness through awareness

CLASSIFIER

TRAINING
Fairness through awareness

TRAINING optimisation problem
Fairness through awareness

TRAINING
optimisation problem

RUNNING
Fairness through awareness

TRAINING
optimisation problem

RUNNING
Fairness through awareness

TRAINING
optimisation problem

RUNNING
Fairness through awareness

TRAINING optimisation problem

RUNNING
Fairness through awareness

implementing fairness = putting constraints on the optimisation problem
Fairness through awareness

implementing fairness = putting constraints on the optimisation problem

treat similar individual similarly (Lipschitz classifier) [Dwork 2013]
Fairness through awareness

implementing fairness = putting constraints on the optimisation problem
treat similar individual similarly (Lipschitz classifier) [Dwork 2013]
Fairness through awareness

Optimisation problem

Implementing fairness = putting constraints on the optimisation problem

Treat similar individuals similarly (Lipschitz classifier) [Dwork 2013]

Probability to belong to the blue category

Feature space
Fairness through awareness

implementing fairness = putting constraints on the optimisation problem

probability to belong to the blue category

TREAT similar individual similarly (Lipschitz classifier) [Dwork 2013]
Fairness through awareness

implementing fairness = putting constraints on the optimisation problem
treat similar individual similarly (Lipschitz classifier) [Dwork 2013]
Fairness through awareness

**TRAINING**

optimisation problem

**RUNNING**

implementing fairness = putting constraints on the optimisation problem

treat similar individual similarly (Lipschitz classifier) [Dwork 2013]
Fairness through awareness

implementing fairness = putting constraints on the optimisation problem
treat similar individual similarly (Lipschitz classifier) [Dwork 2013]

probability to belong to the blue category
Fairness through awareness

implementing fairness = putting constraints on the optimisation problem
treat similar individual similarly (Lipschitz classifier) [Dwork 2013]

probability to belong to the blue category
Fairness through awareness

- Implementing fairness = putting constraints on the optimisation problem
- Treat similar individuals similarly (Lipschitz classifier) [Dwork 2013]

Feature space

\[ P(x) \quad P(y) \]

Probability to belong to the blue category

Individual Fairness
And what about group fairness?
And what about group fairness?

1. **sample size disparity:**
   minorities are under-represented in data, and sample size is inversely correlated with the accuracy of the classifier.
And what about group fairness?

1. sample size disparity:
   minorities are under-represented in data, and sample size is inversely correlated with the accuracy of the classifier.

2. cultural differences
   The lesson is that statistical patterns that apply to the majority might be invalid within a minority group.
Algorithmic approaches to fairness do not translate all concepts of fairness (theories of justice)

Algorithmic approaches to fairness have their own biases
1. privileging those that can be translated into mathematical terms (especially distributive justice: e.g. Rawls 1971; Binmore 1994)
2. privileging efficiency as the main value (O’Reilly 2013)

“Who is going to choose the relevant axioms?” a longstanding question in moral and political philosophy (Sen 2009).
The Idea of Justice, Amartya Sen (2009)
THREE CHILDREN AND A FLUTE:
AN ILLUSTRATION

illustrate the problem with an example in which you have to decide which of three children – Anne, Bob and Carla – should get a flute about which they are quarrelling. **Anne claims the flute on the ground that she is the only one of the three who knows how to play it (the others do not deny this), and that it would be quite unjust to deny the flute to the only one who can actually play it. If that is all you knew, the case for giving the flute to the first child would be strong.**

In an alternative scenario, it is **Bob** who speaks up, and defends his case for having the flute by pointing out that he is the only one among the three who is so poor that he has no toys of his own. **The flute would give him something to play with (the other two concede that they are richer and well supplied with engaging amenities). If you had heard only Bob and none of the others, the case for giving it to him would be strong.**

In another alternative scenario, it is **Carla** who speaks up and points out that she has been working diligently for many months to make the flute with her own labour (the others confirm this), and just when she had finished her work, ‘just then’, she complains, ‘these expropriators came along to try to grab the flute away from me’. **If Carla’s statement is all you had heard, you might be inclined to give the flute to her in recognition of her understandable claim to something she has made herself.**
• **Principle of justice are plural**

• **Rights depend on “capabilities”**

• **Criteria need democratic discussion**

• **Assessment through the worlds that are generated**

In an alternative scenario, it is Bob who speaks up, and defends his case for having the flute by pointing out that he is the only one among the three who knows how to play it. If you had heard only Bob and none of the others, the case for giving it to him would be strong. But if you had heard Carla’s statement, and had also heard the others, you would give him something to play with (the other two concede that they are richer and well supplied with engaging amenities). If you had heard only Carla and none of the others, the case for giving it to her would be strong. And if you had heard only the others and none of the others, the case for giving it to the first child would be strong.

When she had finished her work, ‘just then’, she complains, ‘these expropriators came along to try to grab the flute away from me’. If Carla’s statement is all you had heard, you might be inclined to give the flute to her in recognition of her understandable claim to something she has made herself.
thank you!
References

[KGH2013] Adam D.I. Kramer, Jamie E. Guillory, Jeffrey T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks, PNAS 111(24), 8788–8790 (2013).

[Pariser 2011] Eli Pariser, The Filter Bubble: What the Internet is hiding from you, Penguin (2011). [TED Talk: https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=it#t-514256]

[S2016] Ana Lucía Schmidt, Fabiana Zollo, Michela Del Vicario, Alessandro Bessi, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, Walter Quattrociocchi, Anatomy of news consumption on Facebook, PNAS 114(12), 3035–3039 (2016).

[Dwork 2006] Cynthia Dwork, Frank McSherry, Kobi Nissim, Adam Smith, Calibrating noise to sensitivity in private analysis, Proceedings of the Third Theory of Cryptography Conference (2006). See also: https://www.youtube.com/watch?v=lg-VhHlztqo.

[Dwork 2013] Cynthia Dwork, Deirdre K. Mulligan, It’s not privacy and it’s not fair, 66 STAN. L. REV. ONLINE 35 (2013).

[GJr2010] Oscar H. Gandy Jr., Engaging Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems, 12 ETHICS & INFO. TECH. 29, 37-39 (2010).
References

[NS2008] Arvind Narayanan, Vitaly Shmatikov, Robust De-anonymization of Large Sparse Datasets, IEEE Symposium on Security and Privacy (2008).

[G2013] Melissa Gymrek, Amy L. McGuire, David Golan, Eran Halperin, Yaniv Erlich, Identifying Personal Genomes by Surname Inference, Science 339, 321-324 (2013).

[T2012] Bill Turque, ‘Creative ... motivated’ and fired, Washington Post, March 6, 2012.

[K2017] Will Knigth, The Dark Secret at the Heart of AI, MIT Technology Review, April 11, 2017.

[ON2016] Cathy O’Neil, Weapons of math destruction, Crown Publishing (2016).

[Propublica 2016] Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica, ‘Machine Bias. There’s software used across the country to predict future criminals. And it’s biased against blacks.’, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[Dwork 2013] Cynthia Dwork, Moritz Hardty, Toniann Pitassiz, Omer Reingoldx, Richard Zemel, Fairness Through Awareness, (2011)
ADDITIONAL SLIDES
“For centuries logicians have been able to neglect the problem of the justification of one’s choice of axioms, by considering the latter either as self-evident or as arbitrary. In the first case, since we must bow to the evidence, we have no choice and therefore no need to justify our acceptance. In the second case, since all choices are considered equally arbitrary, it is impossible to justify any one by showing it to be preferable to any other. When we reject both of these extremes, so reminiscent of realism and nominalism, when we admit that a choice of axioms is possible and that it is not entirely arbitrary, then the justification of choice ceases being a negligible problem.”
Algorithmic Regulation

Tim O’Reilly in “Beyond transparency”, 2013.

“Laws should specify goals, rights, outcomes, authorities, and limits. If specified broadly, those laws can stand the test of time.

Regulations, which specify how to execute those laws in much more detail, should be regarded in much the same way that programmers regard their code and algorithms, that is, as a constantly updated toolset to achieve the outcomes specified in the laws....”
“Laws should specify goals, rights, outcomes, authorities, and limits. If specified broadly, those laws can stand the test of time.

Regulations, which specify how to execute those laws in much more detail, should be regarded in much the same way that programmers regard their code and algorithms, that is, as a constantly updated toolset to achieve the outcomes specified in the laws....”

As outlined in the introduction, a successful algorithmic regulation system has the following characteristics:

1. A deep understanding of the desired outcome
2. Real-time measurement to determine if that outcome is being achieved
3. Algorithms (i.e. a set of rules) that make adjustments based on new data
4. Periodic, deeper analysis of whether the algorithms themselves are correct and performing as expected.
Algorithmic Regulation

Tim O’Reilly in “Beyond transparency”, 2013.

Algorithmic regulation v. de-regulation (markets regulate)

Private corporations as new models for regulation (private sector regulates)

Algorithms as efficiency test for the law (economic criterion)

Effectiveness/enforcement through reputation (users’ aggregate judgments)

The inner morality of algorithms: “We need a moral revolution in business which isn’t about do-goodism; it’s about the right way to do things”
Algorithmic Regulation

Tim O’Reilly in “Beyond transparency”, 2013.

Algorithmic regulation (markets regulate)

Private corporations as new models for regulation (private sector regulates)

Algorithms as efficiency test for the law (economic criterion)

Effectiveness/enforcement through reputation (users’ aggregate judgments)

The inner morality of algorithms: “We need a moral revolution in business which isn’t about do-goodism; it’s about the right way to do things”
Algorithmic Regulation

Tim O’Reilly in “Beyond transparency”, 2013.

Algorithmic regulation v. de-regulation (markets regulate)

Private corporations as new models for regulation (private sector regulates)

Algorithms as efficiency test for the law (economic criterion)

Effectiveness/enforcement through reputation (users’ aggregate judgments)

AGENCY

The inner morality of algorithms: “We need a moral revolution in business which isn’t about do-goodism; it’s about the right way to do things”
Algorithmic Regulation

Tim O’Reilly in “Beyond transparency”, 2013.

Algorithmic regulation v. de-regulation (markets regulate)

Private corporations as new models for regulation (private sector regulates)

Algorithms as efficiency test for the law (economic criterion)

Effectiveness/enforcement through reputation (users’ aggregate judgments)

The inner morality of algorithms: “We need a moral revolution in business about the right way to do things”

We need a moral revolution in business about the right way to do things”

AGENCY

INSCRUTABILITY

NORMATIVITY
It’s neither privacy nor transparency

“Whether the information used for classification is obtained with or without permission is unrelated to the production of disadvantage or discrimination.” [DM2013]

“Privacy solutions can hinder efforts to identify classifications that unintentionally produce objectionable outcomes” [DM2013]

“Exposing the datasets and algorithms of big data analysis to scrutiny - transparency solutions - may improve individual comprehension, but given the independent (sometimes intended) complexity of algorithms, it is unreasonable to expect transparency alone to root out bias.” [DM2013]

Difficult to understand

Some machine learning algorithms, i.e. deep learning, work as “black bloxes” even for those who have programmed them [The Dark Secret at the Heart of AI, MIT Technology Review]