Societal and Ethical Challenges in the Era of Big Data:

Exploring the emerging issues and opportunities of big data management and analytics

June 28, 2017 from 11.00 to 12.45
ICE/ IEEE Conference, Madeira (PT)

Agenda Item	Speaker
Presentation of e-SIDES	Richard Stevens, Research Director IDC European Government Consulting
Presentation of workshop objectives	
Ethical and legal issues overview	Gabriella Cattaneo, Associate VP IDC European Government Consulting
Interactive session on ethical and legal issues	
Presentation of results and discussion	
Societal and economic issues overview	Daniel Bachlechner, Senior Researcher Fraunhofer ISI
Interactive session on societal and economic issues	
Presentation of results and discussion	
How to get involved with e-SIDES	Stefania Aguzzi, Senior Consultant IDC European Government Consulting
Next steps and conclusions	Richard Stevens, Research Director IDC European Government Consulting

Societal and Ethical Challenges in the Era of Big Data

Exploring the Emerging Issues and Opportunities of Big Data Management and Analytics

---

ICE/IEEE 2017

Madeira, June 28 (11.00-12.45)

e-Sides Ethical and Societal Implications of Data Sciences
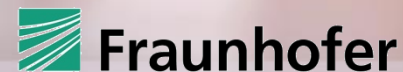
e-SIDES

About e-SIDES

e-SIDES in a nutshell

Involve the complete value chain of big data stakeholders to reach a common **vision for an ethically sound approach** to processing big data

Improve the dialogue between data subjects and big data communities (industry, research, policy makers, regulators) and, thereby, to **improve the confidence of citizens** towards big data technologies and data markets.

e-SIDES Partnership

IDC Analyze the Future

Coordinator

Communication

Community engagement

Fraunhofer

Technical partner for socio-economic research

eLAW LEIDEN

Technical partner for legal and ethics-related research

e-SIDES key objectives

Identify, discuss and validate **ethical and societal implications of privacy-preserving** big data technologies

Liaise with **big data community** (researchers, business leaders, policy makers and society) through events

Provide ethical-legal and societal-economic advice to **facilitate responsible research and innovation** on big data technologies

Provide collective **community position paper** with recommendations for responsible research and innovation on big data
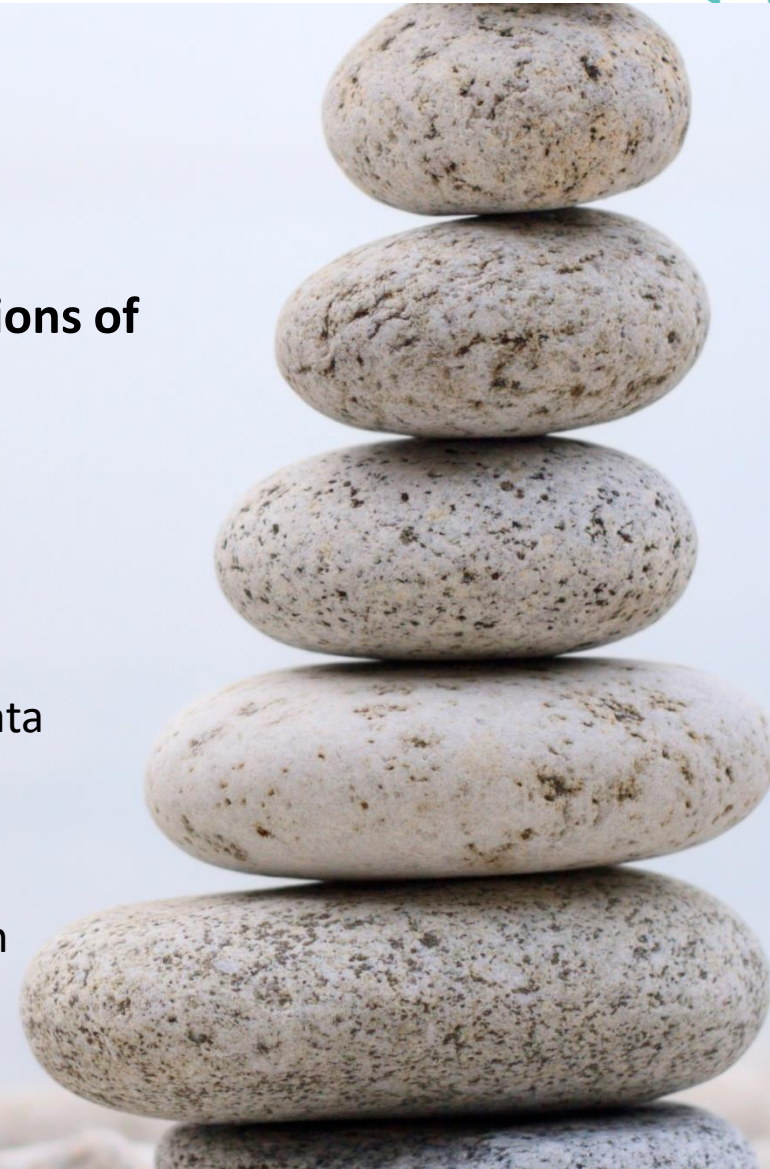
Workshop: Societal and Ethical Challenges in the Era of Big Data

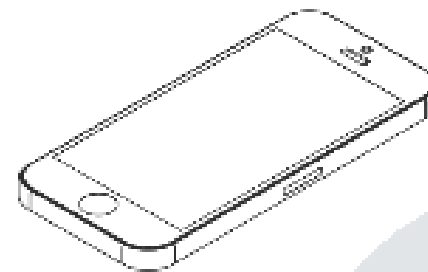# Workshop objectives

Introduce e-SIDES

Present our initial work

Validate results and collect your inputs

How to interact

Go to **www.menti.com** and use the code **74 93 18**

What type of organization do you represent?

1 Grab your phone   **2** Go to www.menti.com   **3** Enter the code
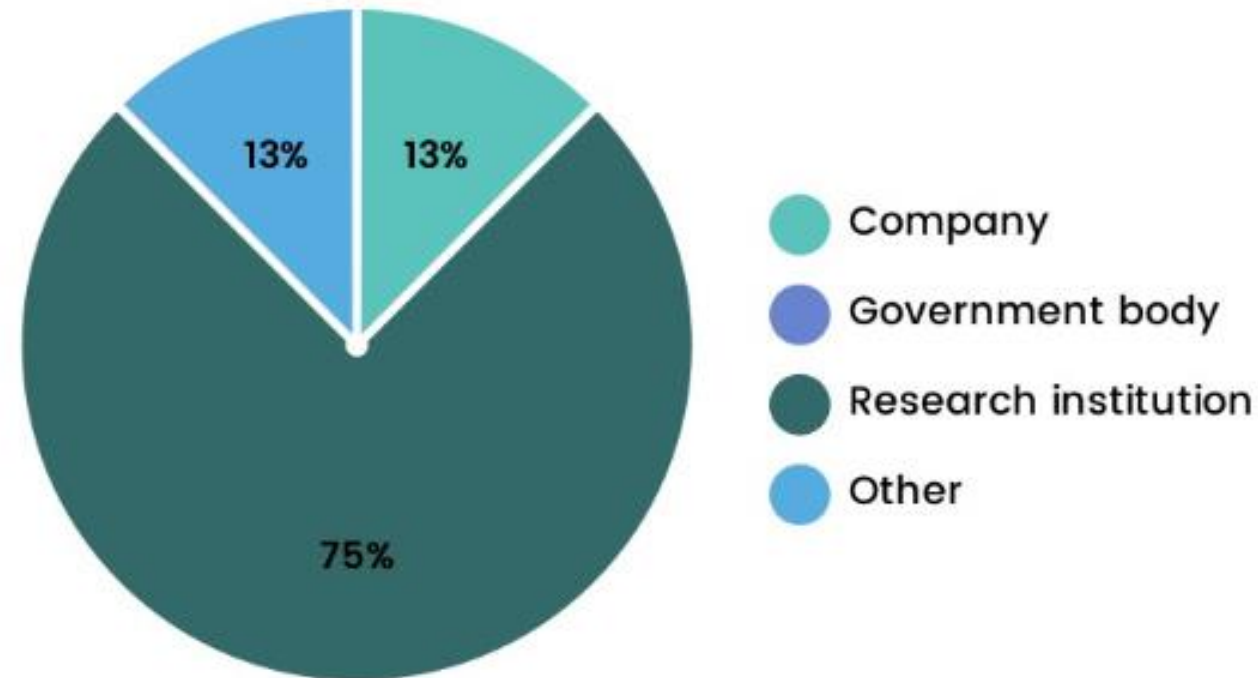74 93 18 and vote!

What type of organization do you represent?
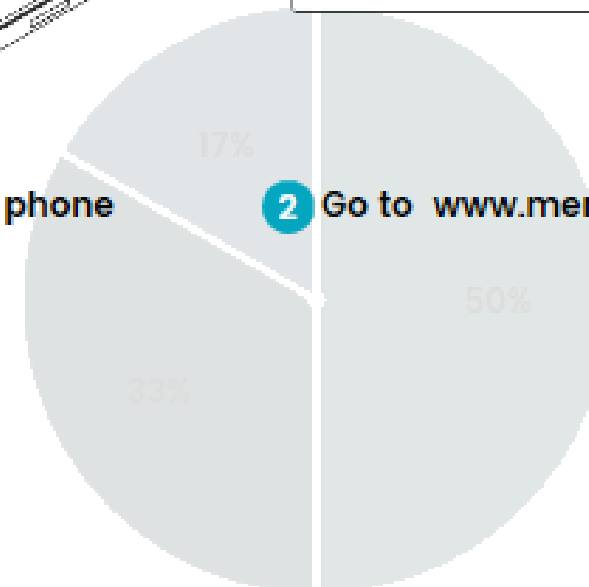
- Company — 13%
- Government body
- Research institution — 75%
- Other — 13%

Ethical and Legal Issues Overview

Gabriella Cattaneo, IDC European Government Consulting

Ethical versus legal issues

- Focus on ethical *and* legal issues
  - Privacy, security
  - Discrimination, stigmatization, polarization
  - Consent, autonomy, self-determination
  - Transparency, integrity, trust

Rapid technological developments may cause ethical issues
  - Violations of moral principles (e.g., human dignity)
  - Conflicting moral principles (e.g. privacy vs. security)
  - New moral principles? (e.g. right to be forgotten, right not to know)

Legislation may not be up to date to address these issues…

… therefore ethical issues may become legal issues

Ethics: how *should* we behave?
Law: how *must* we behave?

Inherent conflicts between Big Data and Personal Data Protection

Personal data protection principles	Big Data risks
Purpose limitation (legitimate and specified before collection)	BD often used for secondary purposes not yet known at collection time
Consent (simple, specific, informed and explicit)	If purpose not clear, cannot ask consensus
Lawfulness (consent obtained and/or processing needed for legitimate purposes)	Without purpose limitation and consent lawfulness is doubtful
Necessity and data minimization	BD needs accumulating data for potential use
Transparency and openness	Individuals cannot keep track of their data

Why do we care?

Personal data protection principles	Big Data risks
Individual rights (to access, rectify, erase/be forgotten)	Lacking transparency, individuals have difficulty to exercise their rights
Information security	Collection of large quantities of data increases risks of violations and abuse
Accountability (compliance with principles)	Compliance does not hold and hence cannot be demonstrated
Data protection by design and by default	Anonymise data! But: Too much anonymization = no use for BD Too little anonymization = possible re-identification

Legal issues in the implementation of DP principles

Ineffective purpose limitation

Lack of fully informed consent

The volume, variety and velocity of Big Data sets may have unintended negative consequences even when processing neutral data with legitimate processes

Blurring of sensitive data concept

Harmful consequences from data processing

Ethical issues: discrimination, stigmatization, polarization

BD profiling, categorising, classifying of data about individuals may...

Weaken solidarity and social cohesion

Weaken individual autonomy by manipulating choices

Create bias about some groups of people

Ethical issues: trust, autonomy, self-determination

Information asymmetry, dependency on BD holders, BDA-driven algorithms making automated choices (selection, pricing) may lead to...

Loss of Trust

Lack of Moral Responsibility

Interactive Session on Ethical and Legal Issues

Gabriella Cattaneo, IDC European Government Consulting

Which of the following issues should be prioritized in the development of big data technologies in order to minimise negative impacts?

Lack of fully informed consent for use of personal data — 7

Ineffective purpose limitation of the exploitation of personal data — 6.3

Blurring of the concept of sensitive data — 6.3

Harm done through processing not clearly understood — 6.5

Solidarity –undermining social cohesion by using personal data profiling — 5.8

Loss of trust because of dependency on Big data providers — 4.5

Risk of reducing autonomy through the manipulation of individual choices — 5

Profiling, categorisation and correlation create Bias in society — 5.8

Avoidance of moral responsibility when decision making is machine generated. — 5

Least relevant / Most relevant

Societal and Economic Issues Overview

Daniel Bachlechner, Fraunhofer ISI

Sources

FP7 and H2020 projects assessing impacts of different (big data related) technologies

- BIG
- CANVAS
- CAPITAL
- CLARUS
- Coco Cloud
- CONSENT
- CRISP
- DwB

- ENDORSE
- ENFORCE
- Inter-Trust
- SIAM
- Socialising Big Data
- SURVEILLE
- SysSec

Unequal access

Unequal access to data and technology, and information asymmetry leading to unequal chances

- Not everybody or every organization is in the same **starting position** with respect to big data

- The **digital divide**, for instance, refers to inequalities between those who have computers and online access, and those who don't

- **Access** to contact data, a privacy policy or information about data collection, processing and sharing depends on capabilities

- Relevant inequalities also exist between **organizations** of different industries, sizes and regional contexts

Example

Online policies on websites typically require a through legal and technological understanding to be fully understood, if they are found at all

Normalization

The classification of people and organizations based on broad categories leading to restricted access to information and services

- People are put into broad **categories** whose characteristics are determined by what is most common and thus expected to be most likely

- **Filter bubbles** result when an algorithm selectively guesses what information somebody wants to see based on information about the individual as well as other similar individuals

- The **breadth of choices** is restricted and **pluralism** pushed back

- Normalization also happens on an **organizational level** but seems to be less critical

Example

Recommendations for products in online shops such as Amazon

Discrimination

Unfair treatment of people and organizations based on certain characteristics leading to immediate disadvantages and unequal chances

- People or groups are **treated differently** depending on certain characteristics including age, disability, ethnicity or gender

- Big data technologies to some extent allow concluding initially **unknown characteristics** from others in the same or other datasets

- Discriminating people or groups might make **economic sense** and is **difficult to be detected**

- Data or algorithms upon which people are discriminated may be **incorrect** or **unreliable**

Example

Predictive policing, no-fly lists, or personalized pricing are examples where discrimination in the context of big data becomes visible

Dependency

The dependency of people and organizations from organizations and technology leading to limitation of flexibility

- People and organizations **depend on others** collecting or processing data, or providing access to data
- **Switching** from one organization to another is often linked to high costs, if it is possible at all
- For many types of data or data-related services, there is a **limited number** of providers and a considerable share of them is based **outside the EU**
- **Business practices** as well as **security measures** can usually not be affected by externals
- Organizations are also highly dependent on the **data** as well as the **big data technologies** they use

Example

Data-intensive organizations such as NHS hospitals in the UK had to stop operating after being attacked with ransomware

EU Commission has led several competition cases against Google's dominance in several markets (e.g. online advertising)

Many online services can only be used after providing requested data (take-it-or-leave-it)

Intrusiveness

The intrusion into the peoples' privacy and organizations' business practices leading to reduction of freedom

- Big data has integrated itself into nearly every part of people's **online** and to some extent also in their **offline experience**

- Data is **stored for long periods** of time and the potential to analyze the data or to integrate it with other data grows

- General suspicion of public authorities and an insatiable appetite of organizations for ever more data infringe people's **freedom**

- The **behavior of people** including how they live, work and interact is affected by intrusive big data applications

Example

The impact of the integration of big data and video surveillance is considered to have particular potential for being intrusive

CCTV, body cameras and drones are increasingly used without the consent of the people observed

Super markets use video data with face recognition to classify and "guide" customers

Non-transparency

The lack of transparency of organizational algorithms and business practices leading to control loss

- Algorithms are often like **black boxes**, they are not only opaque but also mostly unregulated and thus perceived as uncontestable

- People and organizations cannot be sure who is **collecting, processing or sharing** which data

- There are limited means to **check** if an organization has taken suitable measures to protect sensitive data

- Law **enforcement** is often constrained by a lack of resources of public authorities

- There is a lack of **practical experience** with respect to audits including privacy impact assessments

Example

Data subjects' right to information often impossible to exercise

Right to information limited to data storage

Abusiveness

The potential for abuse of data and technologies leading to control loss and deep mistrust

- Data as well as big data technologies may be used for **illegal purposes** or for purposes that fall into a legal grey zone

- It is difficult to **check** the validity of results of data analyses if they look plausible

- Data or algorithms can be **manipulated** in order to reach desired results

- The border between data use and abuse is **blurry** at times

Example

Data collected to remove security flaws may be used by criminals to take over vulnerable systems

Summary

Normalization
The classification of people and organizations based on broad categories leading to restricted access to information and services

Dependency
The dependency of people and organizations from organizations and technology leading to limitation of flexibility

Unequal access
Unequal access to data and technology, and information asymmetry leading to unequal chances

Abusiveness
The potential for abuse of data and technologies leading to control loss and deep mistrust

Non-transparency
The lack of transparency of organizational algorithms and business practices leading to control loss

Discrimination
Unfair treatment of people and organizations based on certain characteristics leading to immediate disadvantages and unequal chances
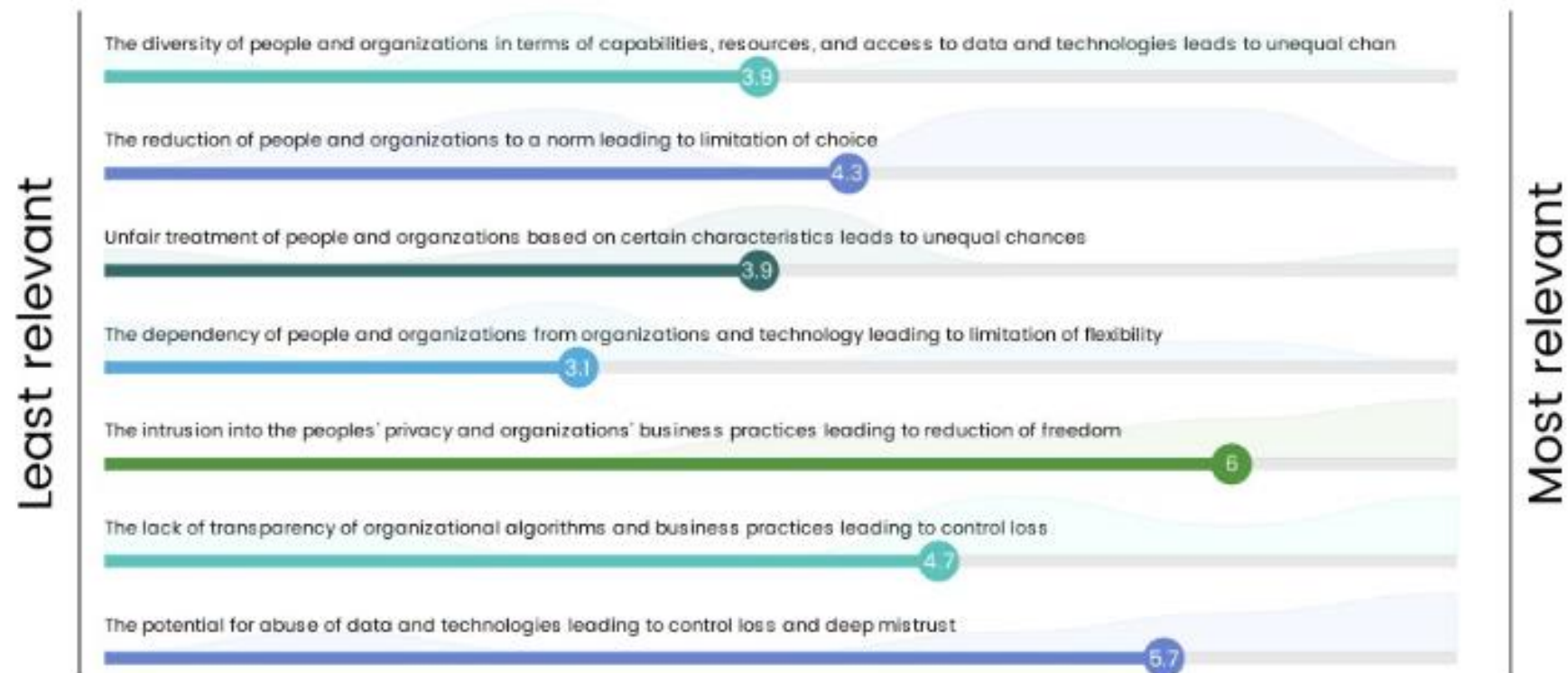
Intrusiveness
The intrusion into the peoples' privacy and organizations' business practices leading to reduction of freedom

Interactive Session on Societal and Economic Issues

Daniel Bachlechner, Fraunhofer ISI

In your opinion, which of the following issues are most urgent and should be prioritized in the development of big data technologies?

The diversity of people and organizations in terms of capabilities, resources, and access to data and technologies leads to unequal chan **3.9**

The reduction of people and organizations to a norm leading to limitation of choice **4.3**

Unfair treatment of people and organizations based on certain characteristics leads to unequal chances **3.9**

The dependency of people and organizations from organizations and technology leading to limitation of flexibility **3.1**

The intrusion into the peoples' privacy and organizations' business practices leading to reduction of freedom **6**

The lack of transparency of organizational algorithms and business practices leading to control loss **4.7**

The potential for abuse of data and technologies leading to control loss and deep mistrust **5.7**

Least relevant — Most relevant

Next Steps and Conclusions

Attend our following conferences, workshops and webinars

Next e-SIDES event: November 2017, TBD

Get access to relevant, novel and continually updated content

Next e-SIDES Report presenting the issues in August 2017

Contribute to our online community paper on website at www.e-sides.eu

Online community initiatives: e-sides.eu

e-SIDES

Ethical and Societal Implications of Data Science

WHAT IS E-SIDES?

Thank you
www.e-sides.eu
info@e-sides.eu

@eSIDES_eu

eSIDES_eu

e-SIDES

e-Sides Ethical and Societal Implications of Data Sciences