PROMISCUOUS DATA-SHARING
IN TIMES OF DATA-DRIVEN ANIMISM

Mireille Hildebrandt
professor of Interfacing Law and Technology (Free University Brussels)
professor of Smart Environments, Data Protection and the Rule of Law (Radboud University)

- *It would be nice if all of the data which sociologists require could be enumerated because then we could run them through IBM machines and draw charts as the economists do. However, not everything that can be counted counts, and not everything that counts can be counted.*

William Bruce Cameron (1963)

- *not everything that can be counted counts, and not everything that counts can be counted.*

William Bruce Cameron (1963)

- *Qualculation implies qualifiction.* *Things have to qualify before they can enter a process of qualculation. More work. Though, to press the point again, this can be – this is – done in an endless number of ways. With an endless range of mechanisms and devices.*

- *(…)* *it takes a lot of work to make the rarefactions of (…) non-calculability.*

Michel Callon & John Law 2003

"Go ahead and think that I'm not really thinking. I thought you would think that."

1. purpose limitation

2. enabling big data research

3. the new animism

1. *purpose limitation*

2. enabling big data research

3. the new animism

post-Brexit data protection in the UK

■ 2 year's notice to leave the EU

■ GDPR in force as of May 2018

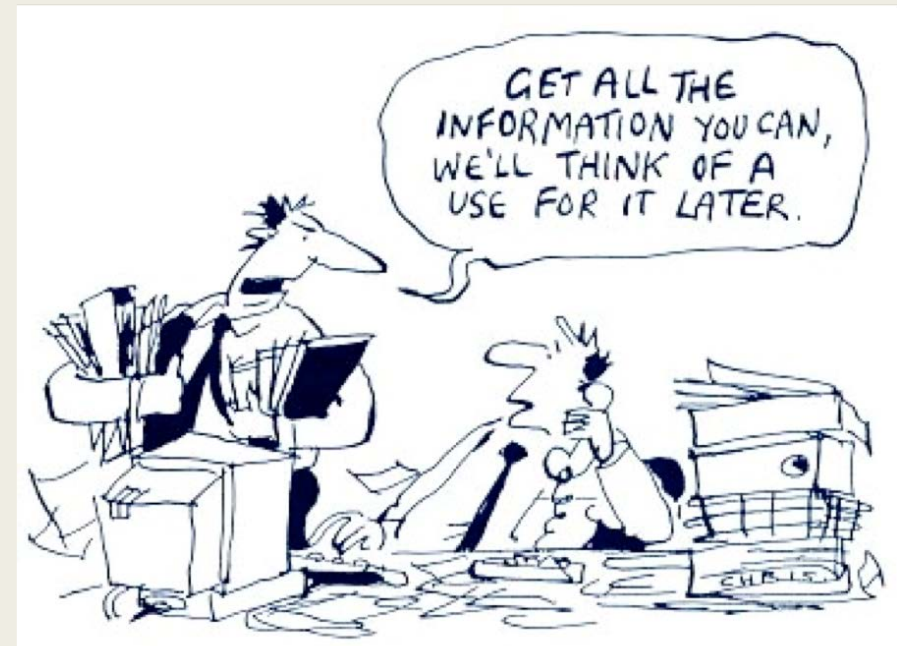After that we may expect that:

1. UK joins the EEA (with Norway, Iceland and Lichtenstein), **OR**

2. UK qualifies for an 'adequacy decision' of the European Commission

GDPR compliance will be crucial in both instances, but interpretation may vary

purpose limitation

■ data protection: purpose limitation (PL)

– *explicit, legitimate, specific purpose*

– *ML mantra: we don't know yet, PL = outdated principle*

data obesitas

purpose limitation

- ■ ML means:
- – *improve performance P at task T based on experience E*
- – *training set, hypothesis space, target function*
- – *off-training set loss, Wolpert's No Free Lunch theorem*

purpose limitation

- A well-defined learning problem requires a *well-specified task, performance metric, and source of training experience.*

- Designing a machine learning approach involves a number of design choices, including choosing the type of training experience, the target function to be learned, a representation for this target function,

Tom Mitchell

purpose limitation

- A well-defined learning problem requires a *well-specified task, performance metric, and source of training experience.*

- Designing a machine learning approach involves a *number of design choices*, including choosing the type of training experience, the target function to be learned, a representation for this target function,
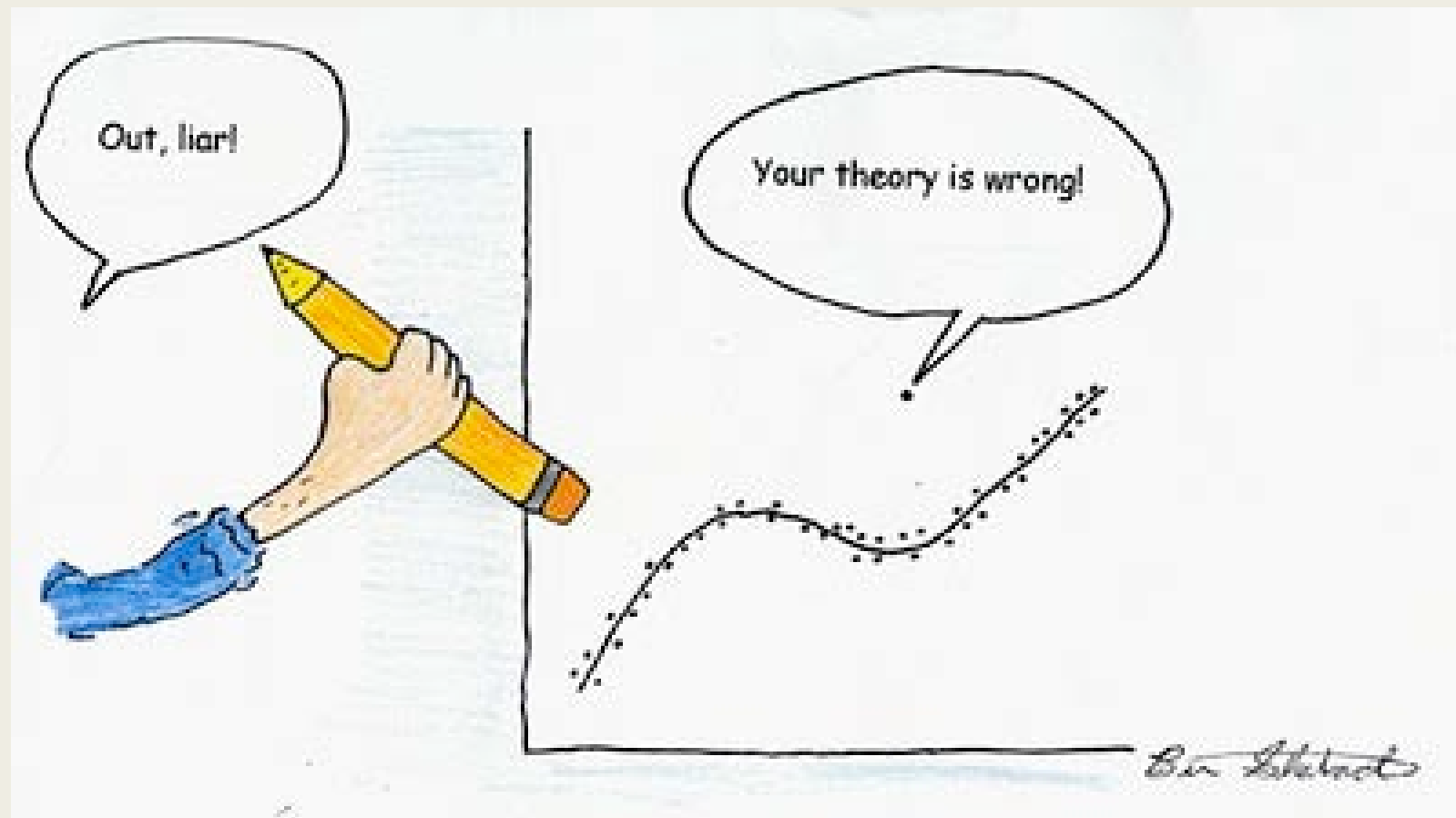
<div align="right">Tom Mitchell</div>

purpose limitation

- We shall see that most current theory of machine learning rests on the *crucial assumption* that the distribution of training examples is identical to the distribution of test examples. Despite our need to make this assumption in order to obtain theoretical results, it is important to keep in mind that this assumption must often be violated in practice.

Tom Mitchell

'overfitting' or 'black swan'

purpose limitation

No Bias no Direction no Learning (the map is the territory):

– *The bias that is necessary to mine the data will co-determine the results*

– *This relates to the fact that the data used to train an algorithm is finite*

– *'Reality', whatever that is, escapes the inherent reduction*

– *and: Data is not the same as what it refers to or what it is a trace of*

– *a healthy return to Hume's scepticism and Gadamer's Wahrheid und Methode*

purpose limitation

Conclusion from a data science perspective:

– any type of ML makes assumptions
– involves productive bias
– 'just searching for any pattern' makes no sense: *pattern obesitas*

– pruning: select *before* you collect, *while* you collect and *after* you collect
– lean and agile computing!
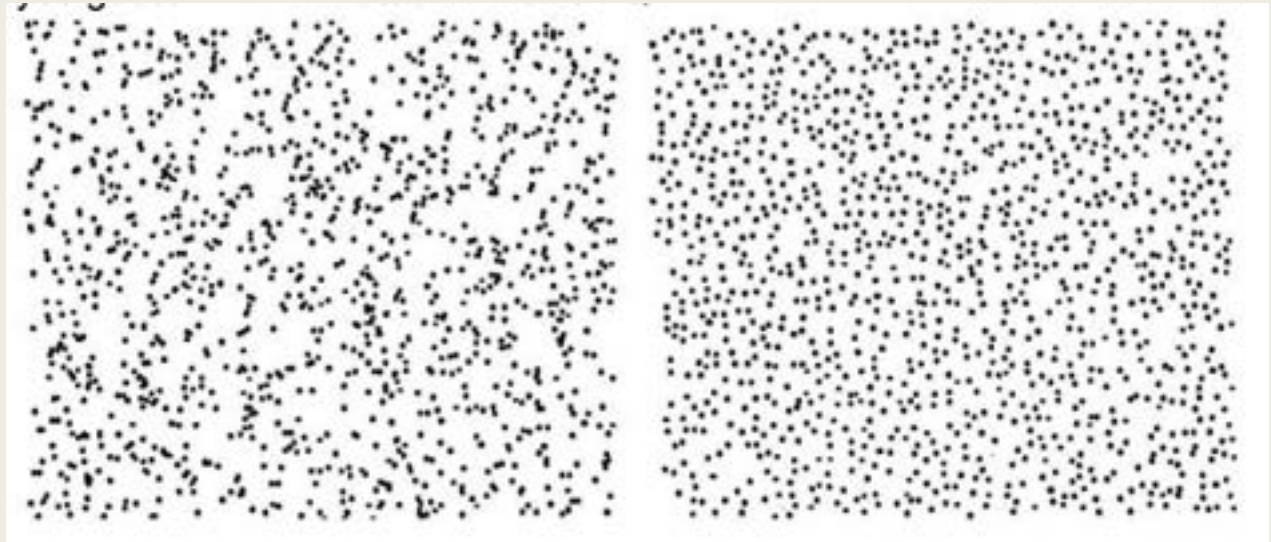– developing and testing a hypothesis space requires direction (purpose)

pattern obesitas

WIRED

AATISH BHATIA   SCIENCE   12.21.12   4:48 PM

WHAT DOES RANDOMNESS LOOK LIKE?

pattern obesitas

- One of the patterns is randomly generated. The other imitates a pattern from nature. Can you tell which is which?

pattern obesitas

1. The one on the left, with the clumps, strands, voids, and filaments (and perhaps, depending on your obsessions, animals, nudes, or Virgin Marys) is the array that was plotted at random, like stars.

2. The points on the right records the positions of glowworms on the ceiling of the Waitomo cave in New Zealand. These glowworms aren't sitting around at random, they're competing for food, and nudging themselves away from each other. They have a vested interest against clumping together.

purpose limitation

Business, Government, Research: Explain your purpose!

- if you can't explain to a person why you need her data

- to figure out what business case it may help to discover

- don't expect her to share the data for that – unknown – purpose

Business, Government, Research: Do not take for granted that you have a right to process my data!

- you may want to employ my house, bag, kidneys or brain capacity to figure out a new business case

- I am, however, under no obligation to share these 'things' with you

- why should it be different concerning my personal data?

1. purpose limitation

2. *enabling big data research*

3. the new animism

enabling big data research

■ *enabling re-use & re-purposing for ML applications*

– *what legal ground?*

(a) consent: for a specific purpose

(f) legitimate interest: balancing against rights and freedoms data subjects

enabling big data research

- **enabling re-use & re-purposing for ML applications**

  – *what specific and legitimate purpose?*

- experimentation for medical, educational, marketing, scientific research purposes

  – *science, sensitive data, data life-cycle management, open source, secure logging*

  – *note that the inferences may be used to target citizens, users, consumers*

  – *information must be provided; safeguards and derogations apply*

enabling big data research

- **provision of information**

  - also if no direct link between data controller and data subject

  - the identity and the contact details of the controller; the purposes as well as the legal basis; the categories of personal data concerned; the recipients or categories of recipients of the personal data; intention to transfer personal data to a recipient in a third country.

  - storage period, legitimate interest (if applicable), the existence of the right to request access, rectification, erasure, of the right to object as well as the right to data portability; the existence of the right to withdraw consent at any time, the right to lodge a complaint with a supervisory authority; source and whether it came from publicly accessible sources; the existence of automated decision-making, including profiling, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

enabling big data research

■ **provision of information**

art 14.5

(b) the provision of such information *proves impossible or would involve a disproportionate effort*, (...). *In such cases the controller shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available*;

enabling big data research

- **provision of information**

Art. 89 **Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes**

1.Processing for archiving purposes in the public interest, *scientific* or historical *research* purposes or *statistical purposes*, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that *technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation*. Those measures may include *pseudonymisation* provided that those purposes can be fulfilled in that manner. (…).

2.Where personal data are processed for *scientific or historical research purposes or statistical purposes*, Union or Member State law may provide for *derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article* in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.

enabling big data research

- **provision of information**

Art. 89 Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes

3.Where personal data are processed for *archiving purposes in the public interest*, (...).

4.Where processing referred to in paragraphs 2 and 3 serves *at the same time another purpose, the derogations shall apply only to processing for the purposes referred to in those paragraphs*.
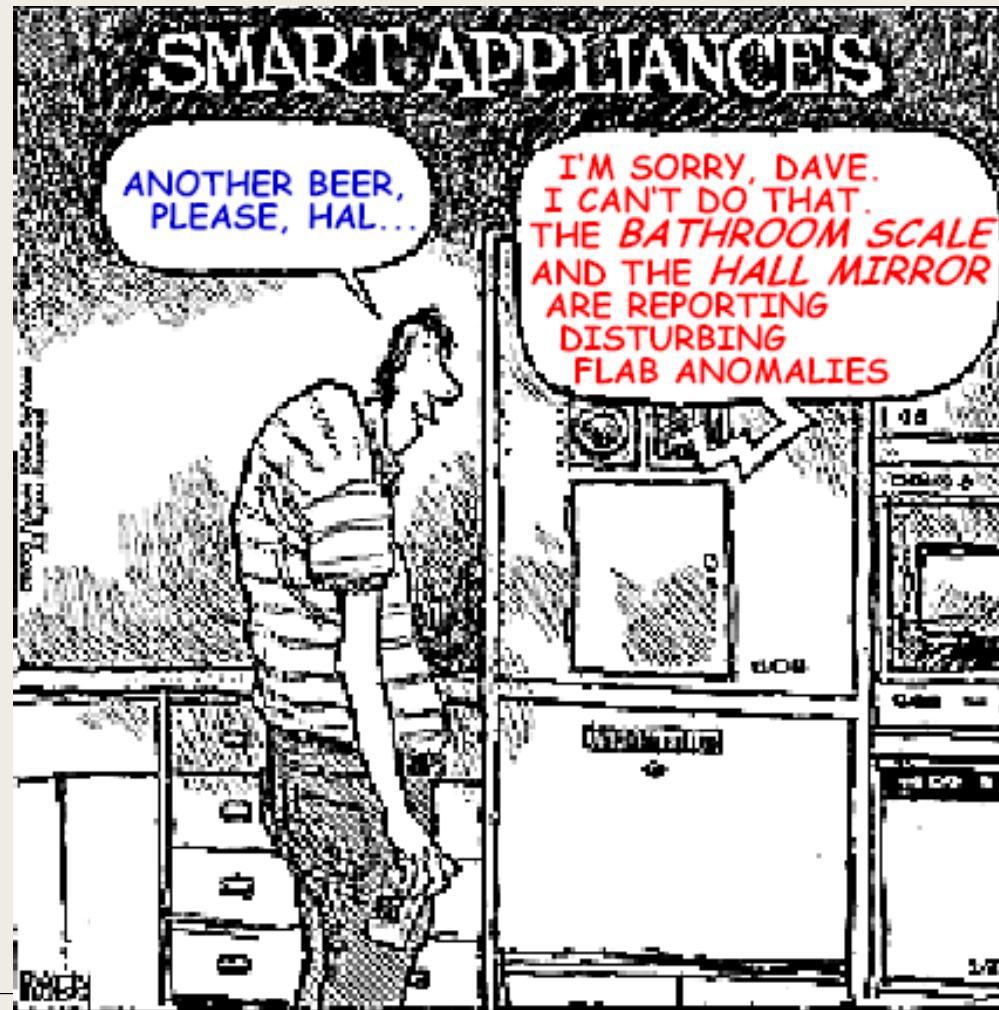
1. purpose limitation

2. enabling big data research

3. *the new animism*

on the internet, nobody can see you're a bot

Lalith Polepeddi, a (human) teaching assistant and researcher on the Jill Watson project at the Georgia Institute of Technology. PHOTO: LALITH POLEPEDDI

- "I have been accused of being a computer," says TA Lalith Polepeddi, a computer-science master's student who was needled for responding to messages with lightning speed. "I don't take it personally."

THE WALL STREET JOURNAL.

http://www.wsj.com/articles/if-your-teacher-sounds-like-a-robot-you-might-be-on-to-something-1462546621

A-HED

Imagine Discovering That Your Teaching Assistant Really Is a Robot

Students mostly couldn't tell 'Jill Watson' wasn't human; 'Yep!'

data-driven agency

- our ethics is based on things being passive and mind being active

- the distinction no longer holds

- things now perceive their environment, 'behave' and learn from the feedback

- data-driven applications are capable of anticipating our machine-readable behaviours

data-driven agency

■ Merleau Ponty, Don Ihde, Bruno Latour, Andy Clark (*bien étonnés de se trouver ensemble*):

– *the extended mind – outside the 'skinbag'*

– *the women with the feather on her hat*

– *the monk with glasses*

– *delegating cognitive tasks to technological artefacts*

– *teenagers with their mobile phone*

– *our cognitive resources are part of the extended mind*

– *now, what about cognitive machines (AIs), what's new?*

data-driven agency

- our behaviours form the cognitive resources for data-driven application

- we become part of the extended mind of online and onlife applications

- these cognitive machines (e.g. IBM's Watson, or Google's self-driving car) can pre-empt us

- this is a mutation in our relationship with the world we inhabit, it is now being animated

- **we need a new animism:** recognition that our world is saturated with mindless agents

- **we need a new intentional stance**: anticipating how 'they' anticipate us

Profile transparency

Information to be provided: Art. 13.2.f and 14.2.g

■ *the existence of* automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

Profile transparency

Information to be provided: Art. 13.2.f and 14.2.g

- the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, *meaningful information about the logic involved*, as well as the significance and the envisaged consequences of such processing for the data subject.

Profile transparency

Information to be provided: Art. 13.2.f and 14.2.g

■ the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as *the significance and the envisaged consequences* of such processing for the data subject.

Profile transparency

In the case of re-use and re-purposing: Art. 13.3 and 14.4

- Where the controller intends to further process the personal data for *a purpose other than that for which the personal data were obtained*, the controller shall provide the data subject prior to that further processing with information on that other purpose and with any relevant further information as referred to in paragraph 2.

Profile transparency

In the case of re-use and re-purposing: Art. 13.3 and 14.4

- Where the controller intends to further process the personal data for *a purpose other than that for which the personal data were obtained*, the controller shall provide the data subject *prior* to that further processing with information on that other purpose and with any relevant further information as referred to in paragraph 2.

Profile transparency

In the case of re-use and re-purposing: Art. 13.3 and 14.4

- Where the controller intends to further process the personal data for *a purpose other than that for which the personal data were obtained*, the controller shall provide the data subject *prior* to that further processing with *information on that other purpose and with any relevant further information as referred to in paragraph 2.*

Profile transparency

1. The data subject shall have the *right to object*, on grounds relating to his or her particular situation, at any time to processing of personal data concerning him or her which is *based on point (e) or (f) of Article 6(1), including profiling based on those provisions.* The controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims.

2. Where personal data are processed for direct marketing purposes, the data subject shall have the right to object at any time to processing of personal data concerning him or her for such marketing, which includes profiling to the extent that it is related to such direct marketing.

3. Where the data subject objects to processing for direct marketing purposes, the personal data shall no longer be processed for such purposes.

Profile transparency

1. The data subject shall have the right to object, on grounds relating to his or her particular situation, at any time to processing of personal data concerning him or her which is based on point (e) or (f) of Article 6(1), including profiling based on those provisions. The controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims.

2. Where personal data are processed for *direct marketing purposes*, the data subject shall have the right to object at any time to processing of personal data concerning him or her for such marketing, which includes profiling to the extent that it is related to such direct marketing.

3. Where the data subject objects to processing for direct marketing purposes, the personal *data shall no longer be processed for such purposes*.

Profile transparency

4. At the latest at the time of the first communication with the data subject, *the right referred to in paragraphs 1 and 2 shall be explicitly brought to the attention of the data subject and shall be presented clearly and separately from any other information.*

5. In the context of the use of information society services, and notwithstanding Directive 2002/58/EC, the data subject may exercise his or her right to object by automated means using technical specifications.

6. Where personal data are processed for scientific or historical research purposes or statistical purposes pursuant to Article 89(1), the data subject, on grounds relating to his or her particular situation, shall have the right to object to processing of personal data concerning him or her, unless the processing is necessary for the performance of a task carried out for reasons of public interest.

Profile transparency

4. At the latest at the time of the first communication with the data subject, the right referred to in paragraphs 1 and 2 shall be explicitly brought to the attention of the data subject and shall be presented clearly and separately from any other information.

5. In the context of the use of information society services, and notwithstanding Directive 2002/58/EC, the data subject may exercise his or her right to object *by automated means using technical specifications.*

6. Where personal data are processed for scientific or historical research purposes or statistical purposes pursuant to Article 89(1), the data subject, on grounds relating to his or her particular situation, shall have the right to object to processing of personal data concerning him or her, unless the processing is necessary for the performance of a task carried out for reasons of public interest.

The End