The University Of Sheffield.

Genetic and Health Data : ‘identifiability’ and ‘consent’ challenges for tomorrow

MJ Taylor, School of Law, University of Sheffield
m.j.taylor@sheffield.ac.uk

37th International Privacy Conference. 26.10.15.

THE 1ST STUDENT EXPERIENCE SURVEY 2014-15

Context

- Increasingly fluid interpretive contexts – challenges axioms
  - Data persistently relates – or does not relate – to a particular individual
    - ➢ Interpretation and application of data to an individual will rely upon data about others
  - Data is collected for specific purposes that can be prospectively described
    - ➢ Unknowable what information will be generated through future analysis and linkage of data and the potential utility of that information.

- Aim: Want to protect fundamental rights and freedoms, esp. privacy, while facilitating use for purposes that persons have reason to accept and are consistent with reasonable expectation

Fluid interpretive contexts raise (at least) two questions..

1. Can unique (genetic) data be(come) non-identifiable in context?
   - And, if identifiable, can it be pseudonymised?

2. How can we recognise limits of specific, explicit consent, AND ensure use *remains* consistent with reasonable expectations and purposes that persons have reason to accept?

1. GENETIC DATA & IDENTIFIABILITY

Identifiability – 95/46/EC

- 2(a) ‘personal data’ shall mean any information relating to an identified or identifiable natural person (‘data subject’); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;

Identifiability – 95/46/EC

- 2(a) ‘personal data’ shall mean any information relating to an identified or identifiable natural person (‘data subject’); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;

Identifiability – 95/46/EC

- 2(a) ‘personal data’ shall mean any information relating to an identified or identifiable natural person (‘data subject’); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;

Q. Is an individual identified or identifiable through genetic data?

Identifiability – 95/46/EC

Identified if,

"In general terms, a natural person can be considered as 'identified' when, within a group of persons, he or she is 'distinguished' from all other members of the group."

(WP136, p12)

Identifiability – 95/46/EC

Identifiable if,

- *"although the person has not been identified yet, it is possible to do so" (WP136, P12)*
- *"So, the question of whether the individual to whom the information relates is identified or not depends on the circumstances of the case." (WP136, P13)*

Identifiability – 95/46/EC

Identifiable if, in the circumstances,

the person to whom the data relates has not yet been distinguished from others
… by those processing the data.

Can data, identified by 'A', be transferred to 'B', and – in the hands of 'B' – be identifiable in that new context?

Identifiability

Important because – if an individual is identifiable (but not identified), then Recital 26 is relevant.

- “to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person;
- “the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable”

Identifiability

Important because – if an individual is identifiable (but not identified), then Recital 26 is relevant.

- "to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the

If identifiable, then we can take contextual controls into account when assessing whether it is "personal data": any contractual, technical, and organisational efforts to prevent (re)identification of an individual *are* relevant *to assessment of identifiability.*

Scenario : Deposit unique variant in an ‘open access’ research database

- A mutation is identified through molecular genetic testing in clinic.
- To more accurately determine the clinical significance through frequency data and phenotype information, the data are routinely submitted to several databases. The detail on these databases can be quite considerable but the data is restricted to clinical and mutation information.
- The results often consist of information regarding 1 or 2 mutations in the gene of interest. Some of those mutations are apparently unique to that individual and their family.
- **Are those who (operate or) use the database processing personal data?**

Identifiability

IF individual is identifiable from genetic data, but not identified, then:

- Relative uniqueness of genetic data is not as important as the assessment of whether there are “means likely reasonably to be used” to identify the person

Identifiability = data + context

IF individual is identifiable from genetic data, but not identified, then:

- Relative uniqueness of genetic data is not as important as the assessment of whether there are “means likely reasonably to be used” to identify the person

CONSENT, REASONABLE EXPECTATION, AND ACCEPTABLE USE

What is consent? (in Proposed Regulation)

- Article 4(8) 'the data subject's consent' means any freely given specific, informed and explicit indication of his or her wishes by which the data subject, either by a statement or by a clear affirmative action, signifies agreement to personal data relating to them being processed;

What is consent? (in Proposed Regulation)

- Article 4(8) 'the data subject's consent' means any freely given specific, informed and explicit indication of his or her wishes by which the data subject, either by a statement or by a clear affirmative action, signifies agreement to personal data relating to them being processed;

What happens if – due to fluidity of context – it is impracticable to seek specific, informed and explicit consent for particular use?
Option a – It can't happen
Option b – It can happen only on alt. legal basis

Under what conditions will health/genetic data be available for research? Parliament's position

- Article 81(2): "Processing of personal data concerning health which is necessary for historical, statistical or scientific research purposes shall be permitted only with the consent of the data subject, and shall be subject to the conditions and safeguards referred to in Article 83."

If this were the only option, then a lot of research using genetic data could not take place.
(s81(1b) 'broad consent' notwithstanding)

Under what conditions will health/genetic data be available for research? Parliament's position

- Article 81(2): 2a. Member States law **may provide for exceptions** to the requirement of consent for research, as referred to in paragraph 2, with regard to research that serves a high public interests, if that research cannot possibly be carried out otherwise.
- **The data in question shall be anonymised, or if that is not possible for the research purposes, pseudonymised under the highest technical standards, and all necessary measures shall be taken to prevent unwarranted re-identification of the data subjects.**
- However, the data subject shall have the right to object at any time in accordance with Article 19.

Article 4(2a) 'pseudonymous data' means personal data that cannot be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organisational measures to ensure non-attribution

- Article 81(2): 2a. Member States law **may provide for exceptions** to the requirement of consent for research, as referred to in paragraph 2, with regard to research that serves a high public interests, if that research cannot possibly be carried out otherwise.
- **The data in question shall be anonymised, or if that is not possible for the research purposes, pseudonymised under the highest technical standards, and all necessary measures shall be taken to prevent unwarranted re-**

As an aside - note this can only apply to unique genetic variants *if it is the view that even unique variants are only indirectly identifiable: so can be pseudonymised.*

Under what conditions will health/genetic data be available for research? Parliament's position

- Article 81(2): 2a. Member States law **may provide for exceptions** to the requirement of consent for research, as referred to in paragraph 2, with regard to research that serves a **high public interests**, if that research cannot possibly be carried out otherwise.
- The data in question shall be anonymised, or if that is not possible for the research purposes, pseudonymised under the highest technical standards, and all necessary measures shall be taken to prevent unwarranted re-identification of the

High public interest is not defined. Any medical research? Genetic research that has passed particular thresholds?

Unde
availa

- Article
of cons
serves
otherw
- The da
purpos
measu

- What is the guarantee that data will only be used (consistently) under circumstances that individuals have reason to accept and expect?
- No requirement for consultation or public reason(ing) to test acceptability.

High public interest is not defined. Any medical research? Genetic research that has passed particular thresholds?

Summary and Conclusion

- Risks

  1 – (Unique) health data not used for research because consent requirements cannot be met (or cannot be pseudonymised – or de-identified), or

  2- Alt. to consent is relied upon and health data is used for (any) medical research without consulting people because public interest is under-defined.

- Neither alternative supports the ambition of ensuring that data is used only for purposes that persons have reason to accept, consistent with their reasonable expectations.

