

**Processing pseudonymised data for  
research, statistics and archiving:  
*Korea in the shoes of an EU Member State***

Graham Greenleaf, Professor of Law &  
Information Systems, UNSW Australia

**AI, Ethics & Privacy Workshop  
(AI 윤리와 프라이버시 워크숍)**

January 20, 2020, Korea University Seoul, Korea

# Approach: Korea in EU shoes

- New PIPA Art. 28-2 attempts to emulate GDPR
  - “data processors may process pseudonymized data for statistics, scientific research, and public interest archiving, etc ., without the consent of data subjects.”
- Others are expert on what revised PIPA means
- I will consider what an EU Member State can do
- We cannot be sure what Art. 28-2 might mean to the EU’s assessment of adequacy of Korean law
- Asked to address six questions ...

# Main Q: Presumption of compatibility?

1. Data controllers want to use personal data collected for purpose A for purpose B
  - without having to obtain *consent* for purpose B
2. GDPR art. 6(4) requires purpose B to be *compatible* with purpose A, for non-consensual use to be made
  - Must *consider 5 factors* in art. 6(4): risky if wrong
3. But art. 5(1)(b) provides a *presumption of compatibility* for '**ARS uses**'
  - ‘**archiving** purposes in the public interest, scientific or historical **research** purposes or **statistical** purposes”
4. If this presumption applies, *avoids* risks of art. 6(4)

# 1. Pseudonymous data is still personal data

- *Pseudonymous* data is still personal data
  - Art. 4(5) definition requires: (i) data is processed so it cannot be linked to an individual without further info.; + (ii) that info. is held separately
- *Anonymous* data is not personal data
  - ‘data subject is no longer identifiable’ taking into account ‘all of the means reasonably likely to be used’ (Rec. 26)
- Anonymity is an objective test, and hazardous
  - Increasingly difficult to achieve; dangerous to assume
  - So difficult, this increases the attraction of art. 5(2) and the ‘presumption of compatibility
- BUT pseudonymous data, by itself, is not within 5(2)

## 2. Only 4 categories of purposes are presumed compatible uses

- Art. 5(2): ‘**archiving** purposes in the public interest, scientific or historical **research** purposes or **statistical** purposes’ (‘ARS uses’)
  - Korea omits ‘historical research’ in art. 28-2
- GDPR uses only these terms, 11 times
  - There is no ‘etc’; only ARS uses qualify
- Any uses of pseudonymised data other than ARS uses are not exempt from art. 6(4)
  - Pseudonymisation does not allow big data uses

## 3 Rest of GDPR still applies to ARS uses

- ‘Presumption of compatibility’ for ARS uses does not make the GDPR inapplicable
- Where GDPR rights can affect ARS processing:
  1. Rights of access & correction
  2. Rights to restrict or suspend processing
  3. Right to erasure (incl. right to be forgotten)
  4. Right to object (unless necessary in public interest)
  5. Limits on automated processing (+ ARS processing cannot include individual decision-making)

## 3 Rest of GDPR still applies to ARS uses (cont.)

- 3 exceptions where GDPR does not apply:
  1. A separate 'lawful basis' (art. 6(1)) for the ARS use is not needed, according to Recital 50.
  2. Personal data can be stored 'for longer periods' (than completion of processing) if processing is solely for ARS use (art. 5(1)(e))
  3. Derogations from 6 GDPR articles (only) are possible, by Member State laws (= Korean law)
    - Only if exercise of rights would 'render impossible or seriously impair' the ARS processing (art. 89(2)-(4))
- Presumed compatibility ≠ Get out of Jail Free

## 4. Is commercial 'scientific research' OK?

- Probably no bar on profit-seeking companies
- Is the processing 'scientific research?'
  - Is research the purpose of the processing?
  - Not enough that it uses 'scientific method'
  - Is it a research project complying with sectoral methodological and ethical standards? (EDPS & Art 29 WP)
- Is the processing primarily for research, or is it primarily for some other purpose? (EDPS)
  - Processing for routine commercial purposes not OK

## 5. Safeguards based on data minimisation are required

- Data minimisation is the main aim of safeguards
  - ‘...with technical and organisational measures aimed at minimising the processing of personal data in pursuance of the proportionality and necessity principles’ (Recital 156)
- Art. 89(1)(1) says safeguards ‘may include pseudonymisation’, but also implies that if *anonymisation* can be used, it should be

## 6. Disclosure to 3<sup>rd</sup> Ps, and publication

- **Disclosure** of pseudonymised data to 3rd parties
  - Irrelevant that disclosure is in definition of processing
  - Disclosure increases risk of misuse / data breaches, so is contrary to data minimisation safeguards
  - For whose purpose is disclosure?
- **Publication** of pseud. data is too dangerous
  - subsequent uses are uncontrollable
  - risks of re-identification are not preventable

## 7. Use of pseudonyms to link databases for ARS uses

- Can multiple databases be linked for ARS uses?
  - Rec 157 *supports* ‘coupling information from registries’, but does not clarify its exact role
  - How can an art. 5(2) presumption of compatibility operate when the different databases are collected for different purposes?
  - Linking one individual = profiling, but not illegal
  - Art. 6(4) tests of compatibility seem appropriate
  - Alternatively: new art. 6(1) lawful processing
- Perhaps linking registries is *supported but not presumed* to be compatible.

# 8. Sensitive data in ARS uses

- Art. 9(2)(j) allows processing of any type of sensitive data for ARS uses, under conditions:
  1. **'necessary'** for ARS purposes
  2. art. 89(1) observed – safeguards (data minimisation)
  3. 'shall be **proportionate** to the aim pursued';
  4. shall **'respect the essence** of the right to data protection'; and
  5. shall 'provide for suitable and specific **measures to safeguard** the fundamental rights and the interests of the data subject'
- 3-5 must be provided by Member State (=Korean) laws
- Including sensitive data in ARS uses cannot be presumed
- Choice: (i) separate procedures for sensitive data; OR (ii) non-sensitive data requires the same higher standard procedures.

# 9. Fines & compensation

- Failure to comply may be costly
  - Higher levels of administrative fines for infringing art. 89, art. 9(2)(j) etc
  - Compensation actions (including group actions) by data subjects whose data has been wrongly used

# Art 5(2) ≠ 'Get Out of GDPR Free' card

