

From **data** to **decision-making** under conditions of uncertainty

August 26th to 28th, 2025

RISK
Summer
School
2025

Session: New technologies and data governance, security and quality.

Dr Emilie Terrier (CRJ, UGA) & Dr Sara Correia-Hopkins (Swansea University)

August 29th 2025, 1:30-3:30 pm



Risk Institute
Université Grenoble Alpes

DECODE Project

Decode stands for : DEvelop Compences DEvoted to the management of uncertainty with creativity and drama tools



Swansea University
Prifysgol Abertawe

Session Outline

Part I – Data Governance (Emilie) – up to 50 minutes

Part II – Data Quality and Ethics (Sara) – 30 minutes

Part III - Scenario-Based Learning – 40 minutes

- Group work – 20 minutes
- Plenary discussion – 20 minutes

Learning Objectives

- Develop an understanding of data governance issues, with reference to the UK, France and the EU context more widely.
- Identify key debates in data governance.
- Identify key issues of data security, ethics and data quality, where secondary data is used in research.
- Critically reflect on the application of a range of approaches to data governance and ethics.

Part I – Data Governance

DR EMILIE TERRIER (CRJ, UGA)

Part II – Data Quality and Ethics

DR SARA CORREIA-HOPKINS (SWANSEA UNIVERSITY)

What is Secondary Data?

Data collected by others for non-research purposes.

Used by researchers to answer research questions.

Types:
Administrative data, social media data, metadata

Governance Frameworks (GDPR Focus)

GDPR applies in France/EU; assimilated in UK; global influence – the ‘Brussels Effect’:

- (Bradford 2019)

Legal basis (Article 6):

- Consent 6(1)(a) or public interest 6(1)(e).

Special categories require additional justification (Article 9):

- Scientific purposes 9(2)(j) or public data 9(2)(e)
- Hall and Macdonald (forthcoming)

Terms of Service Challenges



INCREASING RESTRICTIONS
ON API ACCESS.

(Parry 2024)



ETHICAL CONCERNS OVER
SELECTIVE ACCESS.



LEGAL RISKS OF VIOLATING
PLATFORM TERMS.

Data Quality Issues



Lack of control over data generation.



Limited metadata access.

E.g., Correia, 2022; Correia-Hopkins, 2024.



Misalignment with research needs.

‘Representation’ and ‘Measurement’ Errors (Zhang 2012)

Measuring and Addressing Quality

Daas et al (2009)

Source: Relevance, governance, privacy.

Metadata: Definitions and clarity.

Data: Accuracy, bias, missing data.

Ethical Issues – Informed Consent

Informed consent - voluntary, uncoerced choice, full cognizance of the purpose and design of the research project, including what participation involves, any associated risks, and one's rights (Israel & Hay, 2012).

- Individualist (Katian) ethics ill-equipped for internet age (Buchanan, 2017)?



OFTEN INFEASIBLE IN
LARGE-SCALE RESEARCH.



ETHICAL WAIVER POSSIBLE
WITH SAFEGUARDS.



EMPHASISE PUBLIC
INTEREST AND
COMMUNITARIAN ETHICS.

(Macdonald & Correia-Hopkins, forthcoming)

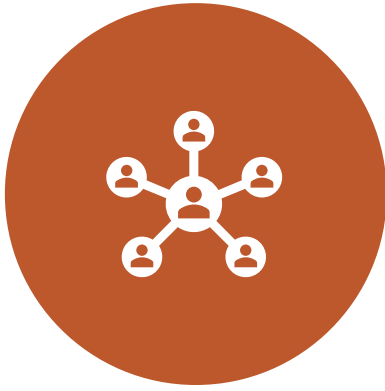
Ethical Issues – Privacy Harms

Risk of re-
identification.



Anonymisation vs.
pseudonymisation
(GDPR's definition
and practices).

Privacy Harms: Mitigation Strategies



USE ORGANISATIONAL
ACCOUNTS OR HIGH-FOLLOWER
PROFILES.



PARAPHRASE POSTS INSTEAD OF
QUOTING.



AVOID PUBLISHING ACCOUNT
NAMES.

Part III – Scenario-Based Learning

Interactive Exercise Overview

1. Discuss and make notes of the following in your groups:

- Key data governance issues;
- Key issues of data quality;
- IP issues;
- Key ethical issues, including the risks of harm and/or actual harms caused by the actions that took place;
- How could/should the potential for harm been identified in advance, mitigated and/or prevented.



Scenarios

- Scenario 1 – Hate speech on social media
- Scenario 2 – Cybercrime police records
- Scenario 3 – Hospital and education data
- Scenario 4 – Online Grooming Vigilantes
- Scenario 5 – Data altruism in mental and brain health

2. After these group discussions you will feedback to the whole room.

Conclusion

Secondary data offers rich opportunities.

Requires careful attention to governance, IP issues, ethics and quality.

GDPR provides a useful framework for global research but concepts like “public interest” need critical examination.

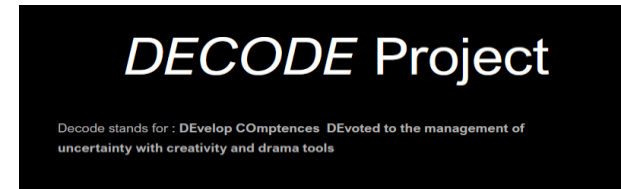
| Merci |
| Thank you |
| Diolch |

Dr Emilie Terrier (CRJ, UGA)
emilie.terrier@univ-grenoble-alpes.fr

Dr Sara Correia-Hopkins
(Swansea University)
s.correia@swansea.ac.uk



Risk Institute
Université Grenoble Alpes



This project is supported by the French National Research Agency (ANR-20-SFRI-0007) under France 2030 investment program.

References

- Bradford, A. (2019). *The Brussels Effect: How the European Union Rules the World*. OUP.
- Buchanan, E. (2017). Considering the ethics of big data research: A case of Twitter and ISIS/ISIL. *PLOS ONE*, 12(12), e0187155. <https://doi.org/10.1371/journal.pone.0187155>
- Correia, S. (2022). Making the most of cybercrime and fraud crime report data: A case study of UK Action Fraud. *International Journal of Population Data Science*, 7(1), 260–272. <https://doi.org/10.23889/ijpds.v7i1.1721>
- Correia-Hopkins, S. (2024). Measuring the Quality of Administrative Cybercrime Data, a UK/USA Comparison. In L. Huey & D. Buil-Gil (Eds.), *The Crime Data Handbook*. Policy Press. <https://doi.org/10.51952/9781529232073.ch018>
- Daas, P. J. H., Ossen, S., Vis-Visschers, R., & Arends-Tóth, J. (2009). *Checklist for the quality evaluation of administrative data sources*. Statistics Netherlands.
- Hall, Jonathan KC and Macdonald, Stuart (forthcoming) ‘Publicising Terrorism in Private: Criminal Law, Online Safety and the Meaning of ‘Public Communications’’
- Israel, M., & Hay, I. (2012). Research Ethics in Criminology. In D. Gadd, S. Karstedt, & S. Messner (Eds.), *The SAGE Handbook of Criminological Research Methods* (pp. 500–516). SAGE Publications Ltd. <https://doi.org/10.4135/9781446268285>
- Kant, I. (1785). *Groundwork for the Metaphysics of Morals* (A. W. Wood, Ed.). Yale University Press.
- Macdonald, S., & Correia-Hopkins, S. (forthcoming). Data Protection, Security and Storage. In S. Lakhani, J. Droogan, S. Macdonald, & L. Khalil (Eds.), *Routledge Handbook of Online Violent Extremism*. Routledge.
- Parry, Douglas (2024) ‘Restrictions on data access impede crucial societal research’ *University World News: Africa Edition*, 9 May. Available at: <https://www.universityworldnews.com/post.php?story=20240509000643443#:~:text=In%20recent%20years%2C%20however%2C%20social,for%20training%20large%20language%20models>
- Zhang, L.-C. (2012). Topics of statistical theory for register-based statistics and data integration. *Statistica Neerlandica*, 66(1), 41–63. <https://doi.org/10.1111/j.1467-9574.2011.00508.x>