



CREATE

script

Reproducible computing and the Law

Burkhard Schafer
University of Edinburgh



Arts & Humanities
Research Council



CREATE

Overview

- Legal incentives
- Legal Obstacles
 - Intellectual property
 - Data protection
 - Liability
- Legal Facilitators
 - Intellectual property
 - Sustainability and legal form



Arts & Humanities
Research Council

script

Why Reproducible Research?

- To enhance benefits from research by facilitating uptake
- To improve quality of research through collective testing



Arts & Humanities
Research Council



CREATE

script

Why Reproducible Research?

- To enhance benefits from research by facilitating uptake
- Legal ramifications:
 - in UK and increasingly throughout Europe, open access promoted as condition for public funding



Arts & Humanities
Research Council



CREATE

script



CREATE

- EPSRC POLICY FRAMEWORK ON RESEARCH DATA

- EPSRC-funded research data is a public good produced in the public interest and should be made freely and openly available with as few restrictions as possible in a timely and responsible manner.
- EPSRC recognises that there are legal, ethical and commercial constraints on release of research data. To ensure that the research process (including the collaborative research process) is not damaged by inappropriate release of data, research organisation policies and practices should ensure that these constraints are considered at all stages in the research process



Arts & Humanities
Research Council

script

In the US (stop press...)

- 22.2.2013: Office of Science and Technology Policy OSTP memorandum requesting federal research funders
- 9.5.2013: Presidential Executive Order 78 FR 28111: "Making open and machine readable the new default for government information"



Arts & Humanities
Research Council



CREATE

script

A generic problem

- Legal rulings also depend on how far you are from the North Pole; a meridian can make the difference between truth and falsehood. What is true on one side of the Pyrénées is wrong on the other side.



Blaise Pascal



Arts & Humanities
Research Council



CREATE

script

Other relevant “promoting” legislation

- Freedom of Information Act (2000)
- Freedom of Information (Scotland) Act (2002)
- Environmental Information Regulations 2004
- Environmental Information (Scotland) Regulations 2004).



Arts & Humanities
Research Council



CREATE

script

Why Recomputable Research?

- To improve quality of research through collective testing



Arts & Humanities
Research Council



CREATE

script

Reasons for RR



Arts & Humanities
Research Council



CREATE

script

- RC creates the type of knowledge increasingly required when science is used in court



Arts & Humanities
Research Council



CREATE

script

US: Daubert Standard

- Whether the theory can and has been tested.
- Whether it has been subject to peer review.
- The known or expected rate of error.
- Whether the theory or methodology employed is generally accepted in relevant scientific community



Arts & Humanities
Research Council



CREATE

script

UK

- English law reform proposal mirrors Daubert
- Some cases claim that we already apply Daubery equivalent standards in UK
- R v T



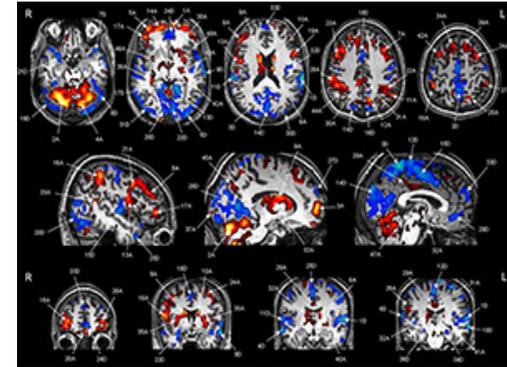
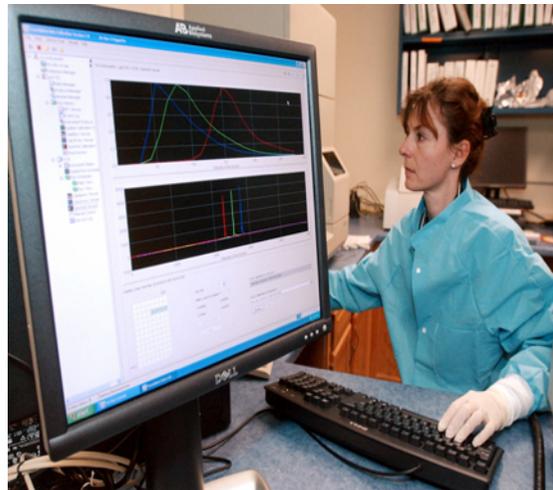
Arts & Humanities
Research Council



CREATE

script

Computer enabled forensics



Arts & Humanities
Research Council



CREATE

script

Problem?

- “reasonable doubt” threshold very low
- Right to reply first?
- Wider issues of public acceptance of science



Arts & Humanities
Research Council



CREATE

script

Other interaction with forensics?

- “Integrity of crime scene”?
- Re-investigating past investigations or historical crimes?



Arts & Humanities
Research Council



CREATE

script

Legal Obstacles

- Intellectual property
- Data protection
- Liability



Arts & Humanities
Research Council



CREATE

script

Legal Obstacles

- Intellectual property



Arts & Humanities
Research Council



CREATE

script

Types of research outputs

- Articles
- Raw data
- Structured and processed data
- Software code



Arts & Humanities
Research Council



CREATE

script

3 Questions

- Downstream: Do I have the rights to distribute everything needed for replication?
- Upstream 1: How can I avoid to accidentally prevent others from using/replicating my results?
- Upstream 2: How can I protect myself from inappropriate use of material I only wanted distributed for some purposes?



Arts & Humanities
Research Council



CREATE

script

Types of IP Rights

- Patents
- Registered Trademarks
- Unregistered trademarks
- Copyright
- Database right
- Integrated circuit layout design protection
- Publication rights
- Intellectual rights to magic methods?



Arts & Humanities
Research Council



CREATE

script

Require action

- Patents
- Registered Trademarks



Arts & Humanities
Research Council



CREATE

script

Patents

- Need (costly) registration
- File first/publish first principle
- Violating “by accident” easy



Arts & Humanities
Research Council



CREATE

script

Trademark

- „Linux“



Arts & Humanities
Research Council



CREATE

script

Copyright

- Exists automatically
- Protects expression of an idea, not idea itself
- Relevant here: academic articles and documentation/explanation/metadata for database and software code
- Default: rests with author,
 - Unless “during the course of employment”



Arts & Humanities
Research Council



CREATE

script

Copyright I Commercial Exploitation



- reproduce the work, prepare derivative works based upon the original
- life of the author plus 70 years.



Arts & Humanities
Research Council



CREATE

script

- Moral rights



Arts & Humanities
Research Council



CREATE

script

Moral rights II

- "I know you will laugh at me," he replied, "but I really can't exhibit it. I have put too much of myself into it."
- Lord Henry stretched himself out on the divan and laughed.



Arts & Humanities
Research Council



CREATE

script

Moral Rights III

- Right to be identified as author or director
- Protection against False attribution of work
- Right to object to derogatory treatment of work
 - But: contribution to periodicals excepted
- Right to privacy of certain photographs and films



Arts & Humanities
Research Council



CREATE

script

Managing copyright

- Licensing
- Creative Commons
 - Varying ranges of permissiveness
 - BY: Licensees may copy, distribute, display and perform the work and make derivative works based on it only if they give the author or licensor the credits in the manner specified by these
 - SA:Share alike – viral license



Arts & Humanities
Research Council



CREATE

script

- CC-BY discouraged for software.
- Several alternatives available, e.g.



Arts & Humanities
Research Council



CREATE

script

Berkeley Software Distribution (BSD)

Copyright (c) <YEAR>, <OWNER>

All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met: Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

Neither the name of the <ORGANIZATION> nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission



Arts & Humanities
Research Council



CREATE

script

Other examples

- GNU Lesser General Public License (typical for code library that other modules link to.)
- MIT License
- Apache License (which also grants patent license if needed to run a product)



Arts & Humanities
Research Council



CREATE

script

Database

- - Raw data unprotected in US
 - But specific arrangement can attract copyright
 - In EU, sui generis data base right.
 - CC-BY 4.0 first CC license to include databases
 - But alternatives exist, e.g. Open Data Commons Attribution License (ODC-By) v1.0 -



Arts & Humanities
Research Council



CREATE

script

Problem

- upstream: license conflict, incompatibility and `most restrictive wins`
- Solutions?
 - Analyzing Open Source License Compatibility Issues with Carneades, Thomas F. Gordon
 - Governatori, Guido, et al. "One License to Compose Them All." *The Semantic Web-ISWC 2013*. Springer Berlin Heidelberg, 2013. 151-166.



Arts & Humanities
Research Council



CREATE

script

Legal Obstacles II: Liability

- Can be managed through disclaimers
- ...to a degree at least
- Can become void if commercial service is provided



Arts & Humanities
Research Council



CREATE

script

Legal Obstacles III: Data Protection

- Unproblematic:



Don't loo

I juss tooked da curlers out an I
hazent styled dem yet!

ICANHASCHEEZBURGER.COM



Arts & Humanities
Research Council



CREATE

script

Instead...

- "personal data" covers any data that can be used to identify a living individual.
- Anonymised or aggregated data is not regulated by the Act,
- providing the anonymisation or aggregation has not been done in a reversible way.



Arts & Humanities
Research Council



CREATE

script

- *Sayers v Smithkline Beecham Plc*
- 'Provided that the reports are sufficiently anonymised so as to prevent any of the confidential medical information in the reports from being identified as relating to any of the claimants or the controls, no question of an infringement arises.



Arts & Humanities
Research Council

script

Non-reversible

- "Linkage Attacks"
 - Differential privacy



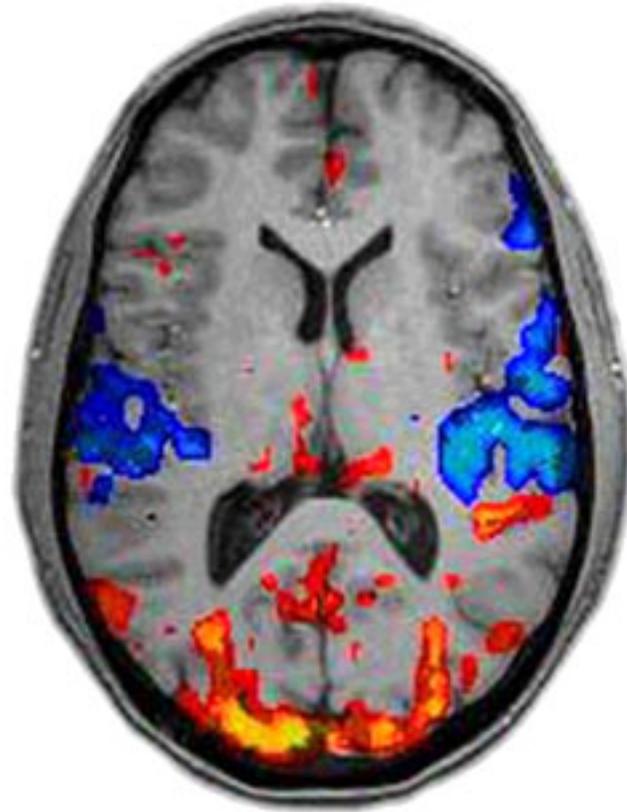
Arts & Humanities
Research Council



CREATE

script

Non-reversible II



Arts & Humanities
Research Council



CREATE

script

Sensitive personal data

- race,
- ethnicity,
- politics,
- religion,
- trade union status,
- health,
- sex life
- criminal record



Arts & Humanities
Research Council



CREATE

script

“Sword and Shield of DP law”



Arts & Humanities
Research Council



CREATE

script

Shield

-
- Anonymisation
- Consent
- Research, historical or statistical analysis (on this more later)



Arts & Humanities
Research Council



CREATE

script

sword

- Personal data shall be obtained only **for one or more specified and lawful purposes**, and shall not be further processed in any manner incompatible with that purpose or those purposes.
- **Personal data shall be adequate, relevant** and not excessive in relation to the purpose or purposes for which they are processed.
- Personal data shall be accurate and, where necessary, **kept up to date**.
- Personal data processed for any purpose or purposes shall **not be kept for longer than is necessary** for that purpose or those purposes.
- Appropriate **technical and organisational measures** shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.
- Personal data **shall not be transferred to a country or territory outside the European Economic Area** unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data



Arts & Humanities
Research Council



CREATE

script

Sensitive personal data

- the individual who the sensitive personal data is about has given explicit consent to the processing.
- The processing is necessary to protect the vital interests of:
 - the individual (in a case where the individual's consent cannot be given or reasonably obtained), or
 - another person (in a case where the individual's consent has been unreasonably withheld).



Arts & Humanities
Research Council



CREATE

script

problems

- Registration duty and triggers for data subject rights
- Withdrawing consent and time limits
- International collaboration



Arts & Humanities
Research Council



CREATE

script

Shield 2: Research Exemption

- Further processing which would otherwise breach the second data protection principle, is not to be regarded as incompatible with the purposes for which the data were obtained, but only where this additional processing is for the research purposes and complies with the relevant conditions



Arts & Humanities
Research Council



CREATE

script

Examples 2

- exempt from the section 7 right of subject access where they are processed for research purposes, in compliance with the relevant conditions, and where the results of the research or any resulting statistics are not made available in a form which identifies data subjects or any of them
- This ensures that, where an individual consents to participate in medical research, for example, the research work will not be hampered or constrained by the possibility that the data subject may seek access to his personal data.



Arts & Humanities
Research Council



CREATE

script

Examples

- by being processed for an additional purpose which had not been specified by the data controller
- keeping indefinitely if they are processed for the research conditions and in compliance with the relevant conditions



Arts & Humanities
Research Council



CREATE

script

Solutions?

- Academic only?
- Simulators or gold standard DBs?



Arts & Humanities
Research Council



CREATE

script

Sustainable infrastructure

- Remember Google Data Service?
- Neither do I....
- Universities: In for the long run, but dependent on individuals
- Incorporated entities: survive individuals
 - But: liability?
 - Funding
 - Commercial – risks liability
 - Trust?

Eligibility for research funding?



Arts & Humanities
Research Council



CREATE

script

Relevant Literature and resources

- Victoria Stodden's blog <http://blog.stodden.net>
- Victoria Stodden, Friedrich Leisch, Roger D. Pen Implementing Reproducible Research, Chapman & Hall 2013 <https://osf.io/s9tya/wiki/home/>
- Victoria Stodden The Legal Framework for Reproducible Scientific Research Licensing and Copyright <http://web.stanford.edu/~vcs/papers/LFRSR12012008.pdf>
- Jerome H. Reichman† and Ruth L. Okediji When Copyright Law and Science Collide: Empowering Digitally Integrated Research Methods on a Global Scale http://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=5351&context=faculty_scholarship
- Nayha Sethi, Graeme Laurie 'Delivering proportionate governance in the era of eHealth: Making linkage and privacy work together' (2013) [*Medical Law International* 13](#)
Abstract:
- Graeme Laurie, Emily Postan 'Rhetoric or Reality: What is the Legal Status of the Consent Form in Health-related Research?' (2012) [*Medical Law Review*](#)



Arts & Humanities
Research Council



CREATE

script