University of California
2019 Ethics, Compliance and Audit Symposium
REACHING NEW HEIGHTS

Regulatory and Ethical Challenges in "Big Data" Research

Kristen Rosati

Coppersmith Brockelman PLC

krosati@cblawyers.com

UNIVERSITY OF CALIFORNIA Ethics, Compliance & Audit Symposium
REACHING NEW HEIGHTS

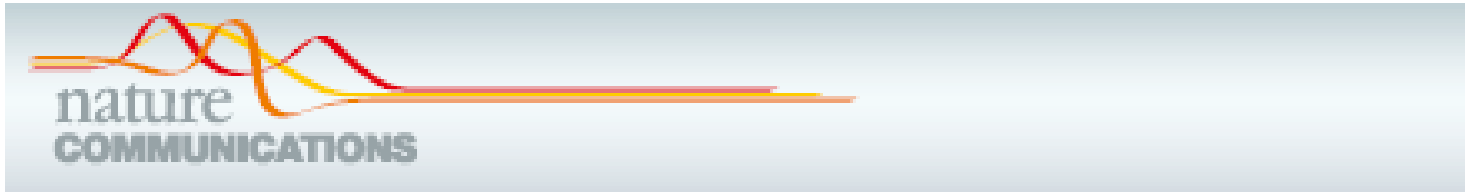

GDPR


CALIFORNIA
CONSUMER
PRIVACY ACT

The New York Times

July 23, 2019, Gina Kolata

Your Data Were 'Anonymized'? These Scientists Can Still Identify You
Computer scientists have developed an algorithm that can pick out almost any American in databases supposedly stripped of personal information.

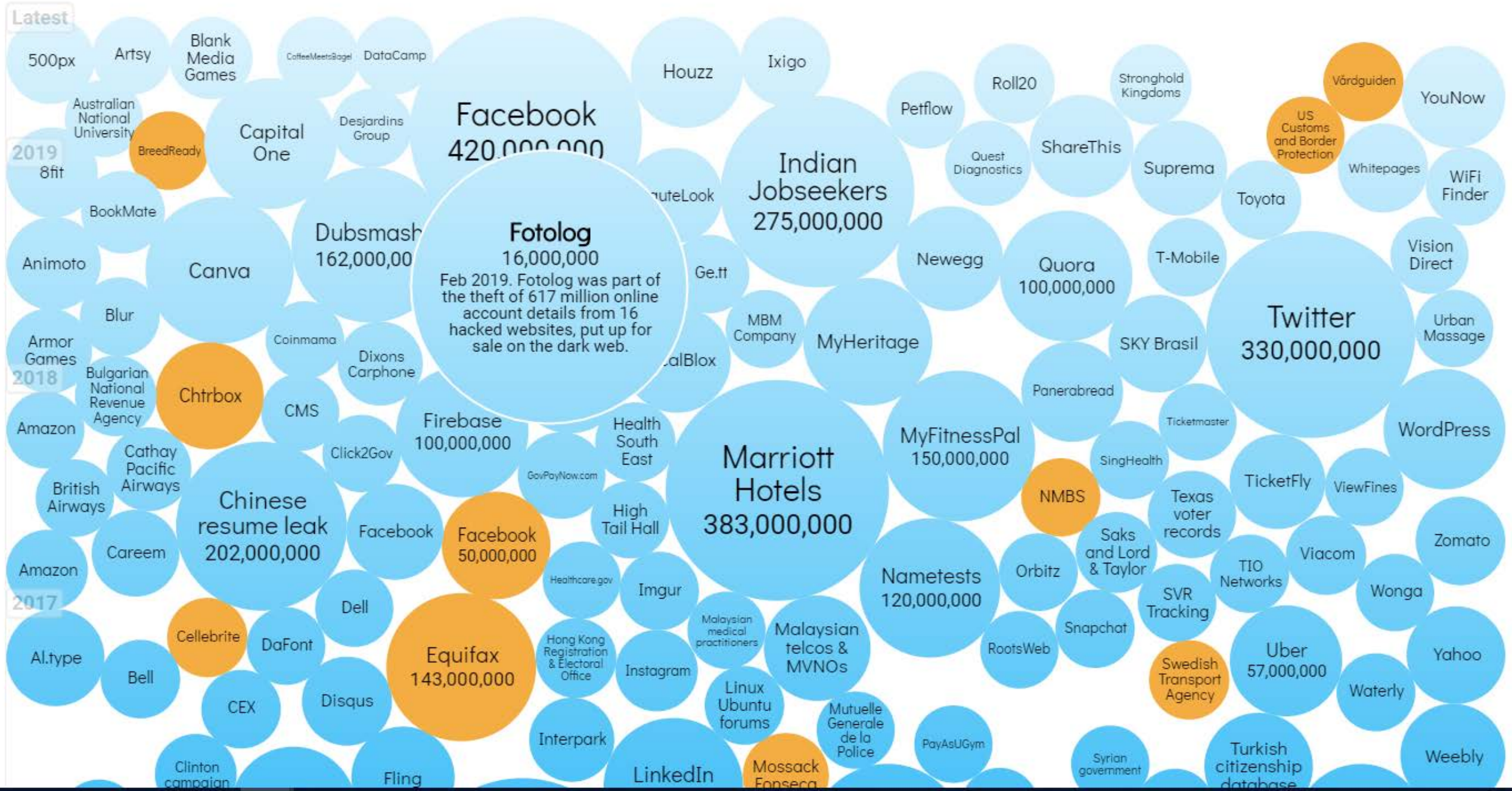


ARTICLE

https://doi.org/10.1038/s41467-019-10933-3 OPEN

Estimating the success of re-identifications in incomplete datasets using generative models

Luc Rocher[1,2,3], Julien M. Hendrickx[1] & Yves-Alexandre de Montjoye[2,3]


Latest
2019
2018
2017
500px
Artsy
Blank Media Games
CoffeeMeetsBagel
DataCamp
Houzz
Ixigo
Australian National University
BreedReady
Capital One
Desjardins Group
Facebook
420,000,000
Petflow
Roll20
Stronghold Kingdoms
Vårdguiden
YouNow
US Customs and Border Protection
Whitepages
WiFi Finder
8fit
BookMate
Quest Diagnostics
ShareThis
Suprema
Toyota
Indian Jobseekers
275,000,000
Animoto
Canva
Dubsmash
162,000,00
Fotolog
16,000,000
Feb 2019. Fotolog was part of the theft of 617 million online account details from 16 hacked websites, put up for sale on the dark web.
Ge.tt
Newegg
Quora
100,000,000
T-Mobile
Vision Direct
Blur
MBM Company
MyHeritage
Twitter
330,000,000
Urban Massage
Armor Games
Coinmama
Dixons Carphone
SKY Brasil
Bulgarian National Revenue Agency
Chtrbox
CMS
Panerabread
Ticketmaster
WordPress
Amazon
Cathay Pacific Airways
Click2Gov
Firebase
100,000,000
Health South East
GovPayNow.com
Marriott Hotels
383,000,000
MyFitnessPal
150,000,000
SingHealth
TicketFly
ViewFines
British Airways
Chinese resume leak
202,000,000
Facebook
Facebook
50,000,000
High Tail Hall
NMBS
Texas voter records
Careem
Saks and Lord & Taylor
Zomato
Amazon
Healthcare.gov
Imgur
Nametests
120,000,000
Orbitz
TIO Networks
Viacom
Wonga
Dell
SVR Tracking
Al.type
Cellebrite
DaFont
Equifax
143,000,000
Malaysian medical practitioners
Malaysian telcos & MVNOs
RootsWeb
Snapchat
Uber
57,000,000
Yahoo
Bell
Hong Kong Registration & Electoral Office
Instagram
Linux Ubuntu forums
Swedish Transport Agency
Waterly
CEX
Disqus
Interpark
Mutuelle Generale de la Police
PayAsUGym
Syrian government
Turkish citizenship database
Weebly
Clinton campaign
Fling
LinkedIn
Mossack Fonseca

Agenda

- Use of data without consent
  - HIPAA
  - Common Rule
  - 42 CFR Part 2
  - California Consumer Privacy Act (CCPA)
  - EU General Data Protection Regulation (GDPR)
- Use of data with consent
- Key elements of an ethical data governance process

HIPAA

- De-identified data
  - Is genetic information Protected Health Information (PHI)?
    - Genetic information is "health information"
    - Health information is PHI if it is "individually identifiable information": identifies the individual or "there is a reasonable basis to believe the information can be used to identify the individual"
    - Office for Civil Rights (OCR) has concluded that not all genetic information is "individually identifiable," but has not provided guidance on when genetic information is individually identifiable
    - Common interpretation: genetic information is not PHI unless it is accompanied by HIPAA identifiers or unless you know recipient has the ability to link the genetic information to a person's identity

HIPAA

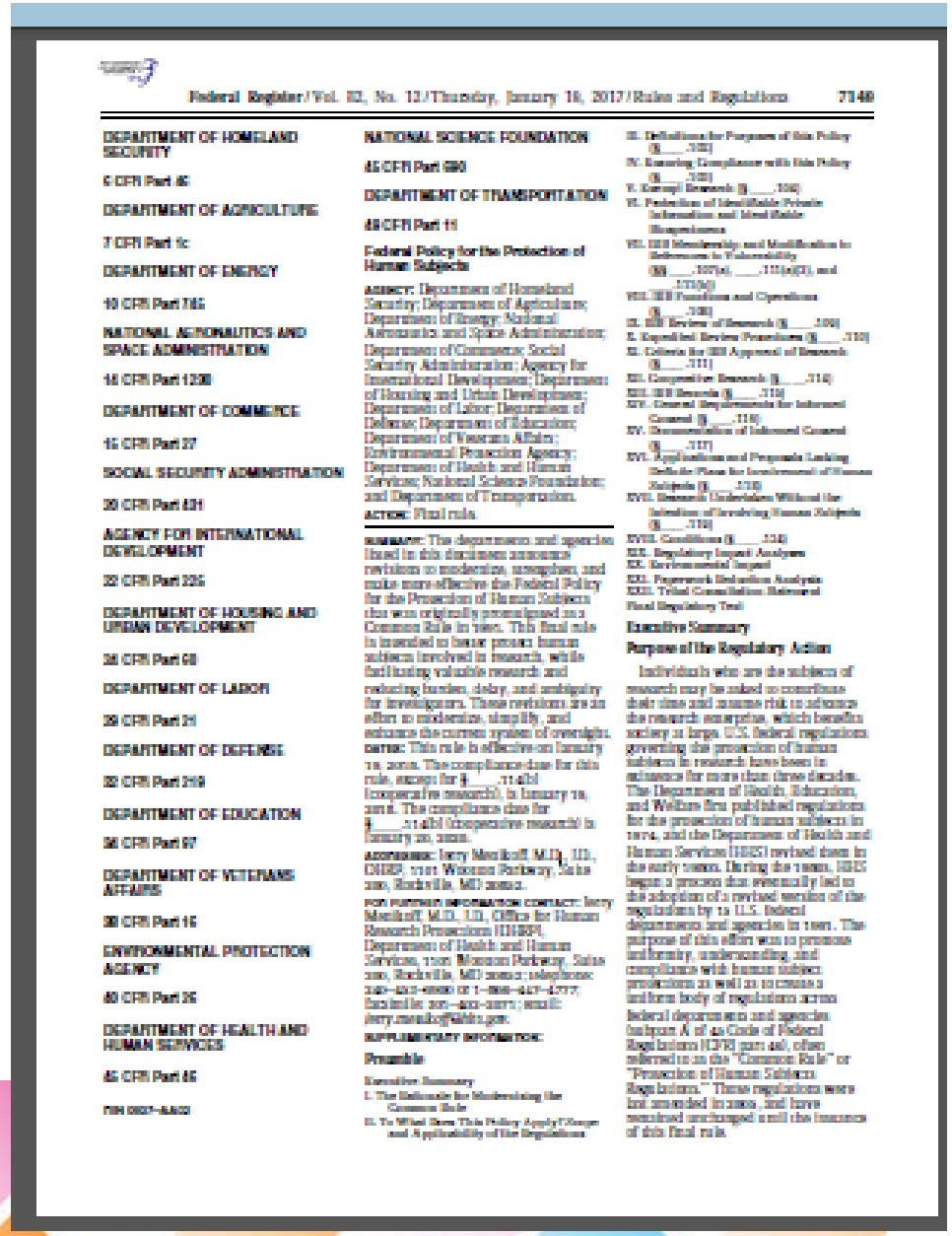
- De-identified data: data use agreement recommended
  - Prohibits re-identifying or contacting individuals*
  - Restricts use/disclosure to research project, prohibits any other use/disclosure unless required by law*
  - Restricts access to defined personnel/agents/subcontractors working on the research project (with agreement to comply with DUA terms)*
  - Expressly prohibits downstream disclosures without permission
  - Prohibits combining with other datasets without permission
  - Requires reporting of unauthorized use or disclosure*
  - Requires reasonable safeguards*
  - Data destruction when project is done (unless necessary to support research results or if validly retained for future research)
- Limited data sets: data use agreement required terms *

HIPAA

- IRB waiver of HIPAA authorization to use/disclose fully-identifiable PHI
  - Common documentation problem -- IRB grant of HIPAA authorization waiver does not include all required documentation requirements:
    - Date of action
    - Statement that IRB has determined that waiver satisfies the following criteria: (1) no more than minimal risk to privacy, based on an adequate plan to protect identifiers, adequate plan to destroy the identifiers at the "earliest opportunity consistent with conduct of the research," adequate written assurance that no PHI will be reused or disclosed (except as required by law, for HIPAA-compliant research, or for authorized oversight of study); (2) research could not be conducted without the waiver; (3) research could not be conducted without the PHI
    - Description of PHI needed for the research
    - Statement that has been reviewed under normal or expedited procedures
- Preparatory to research activities
  - Cannot "remove" PHI from the covered entity

Common Rule

- Non-identifiable data
  - Upcoming guidance on what technologies will generate identifiable data – whole genome sequencing up first for evaluation
  - Potential disconnect with HIPAA
- The new exemption for secondary use of data that is regulated by HIPAA
  - Applies to research within and between covered entities – watch university "hybrid entities"
  - De-identified information



Federal Register / Vol. 82, No. 12 / Thursday, January 19, 2017 / Rules and Regulations 7149

DEPARTMENT OF HOMELAND SECURITY

6 CFR Part 46

DEPARTMENT OF AGRICULTURE

7 CFR Part 1c

DEPARTMENT OF ENERGY

10 CFR Part 745

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

14 CFR Part 1230

DEPARTMENT OF COMMERCE

15 CFR Part 27

SOCIAL SECURITY ADMINISTRATION

20 CFR Part 431

AGENCY FOR INTERNATIONAL DEVELOPMENT

22 CFR Part 225

DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT

24 CFR Part 60

DEPARTMENT OF LABOR

29 CFR Part 21

DEPARTMENT OF DEFENSE

32 CFR Part 219

DEPARTMENT OF EDUCATION

34 CFR Part 97

DEPARTMENT OF VETERANS AFFAIRS

38 CFR Part 16

ENVIRONMENTAL PROTECTION AGENCY

40 CFR Part 26

DEPARTMENT OF HEALTH AND HUMAN SERVICES

45 CFR Part 46

NATIONAL SCIENCE FOUNDATION

45 CFR Part 690

DEPARTMENT OF TRANSPORTATION

49 CFR Part 11

Federal Policy for the Protection of Human Subjects

Executive Summary

Purpose of the Regulatory Action

42 C.F.R. Part 2

- 42 C.F.R. Part 2 amended on 1/17/11 (effective 3/21/17) and on 1/3/18 (effective 2/2/18) -- new regulations expected soon!
- Part 2 regulations apply to (1) "federally assisted" substance use disorder "programs"; and (2) "lawful holders" that receive Part 2-protected data under the regulations (providers with consent and a re-disclosure notice, health plans with consent and researchers without consent)
- Part 2 –protected data
  - Identifies a patient as having (or having had) a substance use disorder
  - Was obtained by a "federally assisted" Part 2 "program"

42 C.F.R. Part 2

- 2017 amendments changed the old rule requiring approval by the program director for research
- Now may use or disclose Part 2 data if determination that the recipient:
  - Is a HIPAA covered entity or business associate and has HIPAA authorization or waiver of authorization;
  - Is subject to the Common Rule and has informed consent or waiver of informed consent or is exempt;
  - If both HIPAA covered entity and subject to Common Rule, complies with <u>both</u>

42 C.F.R. Part 2

- If not a HIPAA covered entity or business associate, and not subject to the Common Rule, requires patient consent
- Part 2 consent requirements are problematic for research
  - Consent form could permit disclosure to: (1) a research institution with a treating relationship with the patient; (2) to a research institution without a treating relationship if re-discloses only to treating providers; or (3) to specific named individuals
  - Other requirements not consistent with HIPAA or the Common Rule

42 C.F.R. Part 2

- An individual or entity that receives Part 2 for research:
  - Is fully bound by the Part 2 Regulations and must resist in judicial proceedings any efforts to obtain access to patient records except as permitted by the Part 2 Regulations;
  - Must not re-disclose patient identifying information except back to the individual or entity from whom that patient identifying information was obtained or as permitted under the data linkage provisions;
  - May include Part 2 data in research reports only in aggregate form in which patient identifying information has been rendered non-identifiable such that the information cannot be re-identified and serve as an unauthorized means to identify a patient, directly or indirectly, as having or having had a substance use disorder;
  - Must maintain and destroy patient identifying information in accordance with the security policies and procedures; and
  - Must retain records in compliance with applicable federal, state, and local record retention laws

42 C.F.R. Part 2

- Researchers may request linkages to data sets from data repositories:
  - Obtain IRB review by OHRP-registered IRB to ensure patient privacy is considered and the need for identifiable data is justified;
  - Upon request, provide evidence of the IRB approval of the research project that contains the data linkage component; and
  - Ensure that patient identifying information obtained is not provided to law enforcement agencies or officials
- Data repository that receives Part 2 data is fully bound by the Part 2 regulations and:
  - After providing the researcher with the linked data, must destroy or delete the linked data from its records, including sanitizing any associated hard copy or electronic media, to render the patient identifying information non-retrievable; and
  - Ensure that the patient identifying information is not provided to law enforcement agencies or officials

The California Consumer Privacy Act

- Cal. Civil Code 1798.100-1798.199
- New proposed regulations published October 11, 2019 at https://www.oag.ca.gov/privacy/ccpa
- Applies only to for profit entities
- Will affect any for-profit collaborators:
  - CCPA's definition of de-identified data is not harmonized with the HIPAA standards
  - CCPA's existing clinical trial exemption does not apply non-interventional data-based research

The EU General Data Protection Regulation

- See UCOP Research Policy Analysis and Coordination page : https://www.ucop.edu/research-policy-analysis-coordination/policies-guidance/general-data-protection-regulation/index.html
- Jurisdictional reach:
  - Applies to organizations "established" (with a physician location) within the European Economic Area (EEA)
  - Applies to organizations outside the EEA that offer goods or services to data subjects within the EEA (clinical trial recruitment) or monitor the behavior of data subjects within the EEA (collection of research data from research participants)
- Applies to EEA research collaborator transfer of "Personal Data" to the United States

What is "Personal Data" under the GDPR?

- Any data that directly or indirectly identifies a living person (not just patients)
  - Name, identification number, location data, online identifiers, factors specific to the physical, psychological, genetic, mental, economic, cultural or social identity
- More sensitive data have special protection
  - Genetic data, biometric data for the purpose of creating unique identification, data concerning health, data regarding race, religion, politics, sex
- Treatment of de-identified data
  - No de-identification "safe harbor"– data is "anonymized" if under a "facts and circumstances" test, the data cannot be identified by any means "reasonably likely to be used … either by the controller or by another person"
  - "Pseudonymised" (coded) still personal data

When is Consent Required under the GDPR?

- GDPR requires a legal basis for "processing" data
  - Consent;
  - Necessary for compliance with a legal obligation of "controller";
  - Necessary for purposes of the "legitimate interests" of the controller (which includes research); or
  - Other provisions not generally relevant in the healthcare setting
- GDPR requires additional legal basis for processing special categories of sensitive data
  - Explicit consent;
  - Necessary for preventive or occupational medicine, medical diagnosis, the provision of health or social care or treatment;
  - Necessary for public health;
  - Necessary for scientific research; or
  - Other provisions not generally relevant in the healthcare setting

If Consent Is Sought under the GDPR?

- Guidance from the "Working Party" requires consent ("explicit" consent for sensitive Personal Data) to be:
  - Freely given
  - Specific
  - Informed
  - An unambiguous indication by a statement or a clear affirmative action
- May also need consent to transfer Personal Data to the US (see next slide)
- Right to withdraw consent

Transfer of Personal Data to US under GDPR

- Requirements for transfer of personal data from the EEA to the US may apply to the sender:
  - Consent (and advising data subjects of the risks of transfer to the US);
  - Contract that contains model contractual clauses approved by the European Commission (which impose some GDPR requirements on receiving entity);
  - To US for-profit entities that have been certified under the EU-US "Privacy Shield"; or
  - Pursuant to codes of conduct by associations

Key Elements of Good Data Governance

- Minimize amount of data, and use de-identified information when possible
- Control use of data through data use agreement (even if the data is de-identified and there is no regulatory requirement for a DUA)
  - See previous slide for recommendations
- Placement on UC-controlled data resources if possible, with carefully controlled access
- Oversight mechanism
  - Ensure regulatory compliance
  - Evaluate high risk or politically sensitive data requests

Questions?

Kristen Rosati

Coppersmith Brockelman PLC

krosati@cblawyers.com

602-381-5464