

2009

Identifiable Neuro Ethics Challenges to the Banking of Neuro Data

Judy Illes

Sofia Lombera

Follow this and additional works at: <https://scholarship.law.umn.edu/mjlst>

Recommended Citation

Judy Illes & Sofia Lombera, *Identifiable Neuro Ethics Challenges to the Banking of Neuro Data*, 10 MINN. J.L. SCI. & TECH. 71 (2009).
Available at: <https://scholarship.law.umn.edu/mjlst/vol10/iss1/6>

The Minnesota Journal of Law, Science & Technology is published by the University of Minnesota Libraries Publishing.



Identifiable Neuro Ethics Challenges to the Banking of Neuro Data

Judy Illes * & Sofia Lombera**

Laboratory and clinical investigations about the brain and behavioral sciences, broadly defined as “neuroscience,” have advanced the understanding of how people think, move, feel, plan and more, both in good health and when suffering from a neurologic or psychiatric disease. Shared databases built on information obtained from neuroscience discoveries hold true promise for advancing the knowledge of brain function by leveraging new possibilities for combining complex and diverse data.¹ Accompanying these opportunities are ethics challenges that, in other domains like the sharing of genetic information, have an impact on all parties involved in the research enterprise. The ethics and policy challenges include regulating the content of, access to, and use of databases; ensuring that data remains confidential and that informed consent

© 2009 Judy Illes and Sofia Lombera.

* Judy Illes, Ph.D., *Corresponding author*, National Core for Neuroethics, The University of British Columbia. The authors are grateful to Dr. Jack Van Horn, Dr. Peter Reiner, Patricia Lau, and Daniel Buchman for valuable discussions on the future of neuro data banks. Parts of this paper were presented at the conference “Emerging Problems in Neurogenomics: Ethical, Legal & Policy Issues at the Intersection of Genomics & Neuroscience,” February 29, 2008, University of Minnesota, Minneapolis, Minnesota by Judy Illes. Full video of that conference is available at <http://www.lifesci.consortium.umn.edu/conference/neuro.php>. Supported by CIHR/INMHA CNE-85117, CFI, BCKDF, and NIH/NIMH # 9R01MH84282-04A1. *All rights reserved.*

** Sofia Lombera, B.Sc., is a Manager for Research & Global Partnerships, National Core for Neuroethics at the University of British Columbia.

1. See generally Shun-Ichi Amari et al., *Neuroinformatics: The Integration of Shared Databases and Tools Towards Integrative Neuroscience*, 1 J. INTEGRATIVE NEUROSCIENCE 117, 117–28 (2002); Jan G. Bjaalie & Sten Grillner, *Global Neuroinformatics: The International Neuroinformatics Coordinating Facility*, 27 J. NEUROSCIENCE, 3613, 3613–15 (2007); International Neuroinformatics Coordinating Facility, <http://incf.org/> (last visited Dec. 11, 2008).

procedures account for future use and commercialization of data; and managing unexpected findings, data anonymization, and recontact procedures. Designing tools to address these challenges in parallel to the technical development of databases is pivotal to their success. While the centralization of neuroscience data in repositories has been met with considerable enthusiasm, this reaction is not uniform throughout the neuroscience community. Policy makers and developers should consider database organization, data sharing, and the obligations and expectations of investigators and accessors as neuroinformatics initiatives move forward. In this article, we identify specific ethics challenges presented by banked collections of “brain data” - genetic, molecular, structural, functional, and behavioral, obtained from human subjects, and we propose directions for research to foster the sharing of data in the future.

NEUROINFORMATICS: A MODEL FOR DATA SHARING IN NEUROSCIENCE

The motivation for the United States’ Human Brain Project (“HBP”) originated during the 1980s from discussions among neuroscientists and program directors at the National Institutes of Health (“NIH”) and the National Science Foundation (“NSF”) who supported the development of neuroinformatics tools that would enable sharing of data among neuroscience investigators.² The tools involve a distributed set of “[w]eb-based databases, analytical tools, and knowledge management systems to foster sharing of data for all domains of neuroscience research.”³ Analytical tools developed in parallel to the databases now allow investigators to study the reliability of methods, to ensure that results are

2. Gordon M. Shepherd et al., *The Human Brain Project: Neuroinformatics Tools for Integrating, Searching and Modeling Multidisciplinary Neuroscience Data*, 21 *TRENDS IN NEUROSCIENCE* 460, 460 (1998); see also Michael F. Huerta & Stephen H. Koslow, *Neuroinformatics: Opportunities Across Disciplinary and National Borders*, 4 *NEUROIMAGE* S4, S4 (1996).

3. NIH GUIDE FOR GRANTS AND CONTRACTS, PROGRAM ANNOUNCEMENT, THE HUMAN BRAIN PROJECT (NEUROINFORMATICS): PHASE I - FEASIBILITY; PHASE II - REFINEMENTS, MAINTENANCE AND INTEGRATION (Dec. 3, 2002) available at <http://grants.nih.gov/grants/guide/pa-files/PAR-03-035.html>; see generally Stephen H. Koslow, *Sharing Primary Data: A Threat or Asset to Discovery?*, 3 *NATURE REVIEWS NEUROSCIENCE* 311 (2002).

reproducible, and carry out meta-analysis not supported by individual data sets. These data sets also allow researchers lacking access to equipment, such as brain scanners, to mine existing data.⁴ The HBP was ultimately created in response to a congressionally mandated initiative in the early 1990s and an Institute of Medicine review of progress in brain mapping.⁵ The momentum in this area was also international. At the Organization of Economic Cooperation and Development (“OECD”) Megascience Forum in 1999,⁶ the creation of a neurosciences database was a highlighted recommendation in the “effort to understand the structure, function, and development of the brain . . . [which] represents one of the great scientific challenges of the 21st century.”⁷ In 2002 Tom Insel, now head of the National Institutes of Mental Health (“NIMH”) of the NIH and colleagues wrote: “[W]e are entering a decade for which data-sharing will be the currency for progress in neuroscience.”⁸ Indeed, in recognition of this, the OECD created a Neuroinformatics Working Group and later the International Neuroinformatics Coordinating Facility (INCF) headquartered at Karolinska Instiutet in Stockholm, Sweden.⁹

The Human Brain Project:¹⁰ Phase I Feasibility Studies Report¹¹ was the first to describe the practical implications of this effort to the scientific community and signaled the beginning of the initiative in the United States. Under the HBP grant program, first phase studies were focused on feasibility and proof of concept; later phase studies focused on refinements, including further testing of the tools across sites,

4. Governing Council of OHBM, *Neuroimaging Databases*, 292 SCIENCE 1673–76 (2001).

5. Shepard, *supra* note 2, at 461.

6. OECD MEGASCIENCE FORUM WORKING GROUP ON BIOLOGICAL INFORMATICS, FINAL REPORT OF THE OECD MEGASCIENCE FORUM WORKING GROUP ON BIOLOGICAL INFORMATICS 50 (1999), *available at* <http://www.oecd.org/dataoecd/24/32/2105199.pdf>.

7. *Id.* at 52.

8. Thomas R. Insel et al., *Neuroscience Networks*, 1 PLOS BIOLOGY 9, 10 (2003).

9. See Bjaalie & Grillner, *supra* note 1; International Neuroinformatics Coordinating Facility, <http://incf.org/> (last visited Dec. 11, 2008).

10. Nat'l Inst. of Mental Health, Neuroinformatics: Human Brain Project, www.nimh.nih.gov/research-funding/scientific-meetings/recurring-meetings/human-brain-project/index.shtml (last visited Oct. 26, 2008).

11. NIH GUIDE FOR GRANTS AND CONTRACTS, PROGRAM ANNOUNCEMENT, THE HUMAN BRAIN PROJECT: PHASE I FEASIBILITY STUDIES (Oct. 6, 1995) *available at* <http://grants.nih.gov/grants/guide/pa-files/PA-96-002.html>.

improvements, maintenance, and integration with other related web-based resources.¹² As there is great diversity in the types of data generated by neuroscience research, novel approaches to collecting, manipulating, combining, displaying, retrieving, managing, and disseminating them have been vital to making these data available for scientific collaboration and electronic use. In response, neuroscience data repositories (e.g., the University of California—Los Angeles laboratory on mapping brain structure and function that houses among other data those from the Alzheimer’s Disease Neuroimaging Initiative,¹³ the Biomedical Informatics Research Network,¹⁴ and BrainNet Europe II¹⁵) have been developing at a steady pace. Some contain specialized data, for example, gene expression in the mouse brain¹⁶ (The Allen Brain Atlas¹⁷), single and multi-unit recordings (e.g., CoCoMac,¹⁸ Ear Lab,¹⁹ SenseLab²⁰), and structural magnetic resonance imaging (“MRIs”) (e.g., Surface Management System Database, SumsDB²¹). Others, such as the functional MRI Data Center²² (“fMRIDC”) are repositories for imaging data obtained from functional magnetic resonance imaging (“fMRI”), in combination with other data collected from imaging modalities such as positron emission tomography (“PET”),

12. *Id.* at 5–6.

13. UCLA, Laboratory of Neuro Imaging, www.loni.ucla.edu (last visited Oct. 26, 2008).

14. Biomedical Informatics Research Network, www.nbirn.net (last visited Oct. 26, 2008).

15. BrainNet Europe, www.brainnet-europe.org (last visited Oct. 26, 2008); see also Jeanne E. Bell et al., *Management of a Twenty-First Century Brain Bank: Experience in the BrainNet Europe Consortium*, 115 ACTA NEUROPATHOL 497, 499 (2008).

16. Harry Hochhesier & Judith Yanowitz, *If Only I Had a Brain: Exploring Mouse Brain Images in the Allen Brain Atlas*, 99 BIOLOGY CELL 403 (2007).

17. Allen Inst. for Brain Sci., www.brain-map.org (last visited Oct. 26, 2008).

18. CoCoMac, www.cocomac.org (last visited Oct. 26, 2008).

19. Boston Univ., Ear Lab, <http://earlab.bu.edu/> (last visited Oct. 26, 2008).

20. Yale Univ., Sense Lab, <http://senselab.med.yale.edu/> (last visited Oct. 26, 2008).

21. Wash. Univ. St. Louis, Van Essen Lab, <http://sumsdb.wustl.edu/sums/index.jsp> (last visited Nov. 7, 2008).

22. Univ. of Cal., The The fMRI Data Center Data Center Home Page, <http://www.fmridc.org/fmridc> (last visited Oct. 26, 2008).

electrophysiology (“EEG”) and magnetoencephalography (“MEG”).²³ Tools such as the Neuroimaging Archive Toolkit (XNAT), “a software platform designed to facilitate common management and productivity tasks for neuroimaging and associated data” have also been developed to allow investigators to mine banked data.²⁴ Recently, the first round of the Collaborative Research in Computational Neuroscience²⁵ (“CRCNS”) data sharing program supported the preparation of data sets for both electrophysiology and behavior.²⁶

Beyond statistical power, benefits of sharing brain data include the stability, relationships, integration, and distribution of the structure and function of the brain at both the microscopic and macroscopic levels.²⁷ Cross-modality interoperability, that is the ability to utilize and leverage data from different data acquisition methods, including genome and protein data, is also a key goal.²⁸ Given that not all originally collected data are used in published findings, their availability for others to mine maximizes utility and reduces the cost of neuroscience investigations.

The Society for Neuroscience (“SfN”) Neuroscience Database Gateway²⁹ (“NDG”) was a project released to the SfN

23. See generally Marcia Baringa, *Neuroimaging: Still Debated, Brain Image Archives Are Catching On*, 300 SCIENCE 43 (2003); Daniel Gardner, *Neurodatabase.org: Networking the Microelectrode*, 7 NATURE NEUROSCIENCE 486 (2004); John D. Van Horn et al., *The Functional Magnetic Resonance Imaging Data Center (fMRIDC): The Challenges and Rewards of Large-Scale Databasing of Neuroimaging Studies*, 356 PHIL. TRANSACTIONS ROYAL SOC’Y LOND. B 1323 (2001) (Eng.).

24. Daniel S. Marcus et al., *The Extensible Neuroimaging Archive Toolkit*, 5 NEUROINFORMATICS 11, 11 (2007); Wash. Univ. St. Louis, XNAT, www.xnat.org (last visited Nov. 11, 2008).

25. Univ. of Cal., Collaborative Research in Computational Neuroscience Data Sharing, www.crens.org (last visited Oct. 26, 2008).

26. Jeffery L. Teeters et al., *Data Sharing for Computational Neuroscience*, 6 NEUROINFORMATICS 47 (2008).

27. See generally Daniel H. Geschwind, *Sharing Gene Expression Data: an Array of Options*, 2 NATURE REVIEWS NEUROSCIENCE 435 (2001); Stephen H. Koslow & Steven E. Hyman, *Human Brain Project: A Program for the New Millennium*, 17 EINSTEIN Q. J. BIOLOGY MED. 7 (2000); John Mazziotta et al., *A Probabilistic Atlas and Reference System for the Human Brain: International Consortium for Brain Mapping (ICBM)*, 356 PHIL. TRANSACTIONS ROYAL SOC’Y B 1293 (2001) (Eng.); John C. Mazziotta, *Imaging: Window on the Brain*, 57 ARCHIVES NEUROLOGY 1413 (2000).

28. Koslow & Hyman, *supra* note 27, at 14.

29. SfN Neuroscience Database Gateway, http://www.sfn.org/index.cfm?pagename=NDG_main (last visited Oct. 30, 2008).

community in early May 2004.³⁰ The Brain Information Group³¹ (“BIG”) of SfN was charged with “evaluating the current status of neuroscience databases; assessing future directions of neuroscience data management . . . and promoting enhanced awareness of the potential for databases to benefit the neuroscience community.”³² Through the successful implementation of integrated databases of neuroscience information, the NDG grew rapidly from its initial seventy-six resources in 2004 to over a hundred resources today, including HBP’s own repository.³³ The INCF began to create an inventory of data sources and software resources available to the neuroscience community in 2007 and has made this information available via its website.

DEVELOPING ENABLING TOOLS

The development of tools for handling ethical and policy issues that complement the development of technical tools for sharing of neuroscience data is vital to the realization of a truly enabling toolbox.³⁴ Ethics and policy tools can be integrated into existing initiatives such as the INCF’s training workshops which are designed to help neuroscientists with technical issues that arise while using neuroinformatics resources. Unlike an ethical response that may be sought only after difficult issues have surfaced, a solution-oriented, ethical-technical partnership can be a powerful force in nurturing the scientific enterprise. This endeavor can be informed by previous work on the ethical, legal, and policy issues surrounding storing tissue and blood samples. In the neuroscience context, the structure of database sharing in terms of content, access, and the culture of ownership; confidentiality, consent and commercialization; and strategies for managing incidental findings and subjects’ privacy are essential to the success of the enterprise. Here we begin to

30. SfN Neuroscience Database Gateway, *supra* note 29.

31. The SfN Brain Information Group, http://www.sfn.org/index.cfm?pagename=ndg_big (last visited Oct. 26, 2008).

32. SfN Brain Information Group, *supra* note 31.

33. SfN Neuroscience Database Gateway, *supra* note 29.

34. See generally Daniel Gardner et al., *Towards Effective and Rewarding Data Sharing*, 1 NEUROINFORMATICS 289 (2003); Thomas R. Insel et al., *Limits to Growth: Why Neuroscience Needs Large Scale Science*, 7 NATURE NEUROSCIENCE 426, 426–27 (2004).

identify some of the ethical and policy challenges with neuroscience databanking.

CONTENT, ACCESS, AND THE CULTURE OF OWNERSHIP

Image-based data are primary drivers for neuroinformatics efforts.³⁵ This is “[b]ecause image-based data are rich in content, large in size and laborious to obtain ...”³⁶ The fast-growing fMRIDC, for example, was introduced to the neuroscience community in June 2000 by the *Journal of Cognitive Neuroscience (JOCN)*.³⁷ Between 2000 and 2006, *JOCN* required that all authors who published in the journal submit their data to the fMRIDC.³⁸ The fMRIDC, funded by NSF and NIH, the W.M. Keck Foundation, and Sun Microsystems Center of Excellence, was created by Professors Michael Gazzaniga and John Van Horn,³⁹ then at Dartmouth College.⁴⁰ The goal was to “speed the progress and the understanding of cognitive processes and the neural substrates that underlie them.”⁴¹ The fMRIDC met these goals as a publicly accessible database of peer-reviewed fMRI studies by storing information that may enable others to re-use data, replicate original studies, and to generate and test new hypotheses.⁴² One of the creators of the database (John Van Horn) reports sending data sets to laboratories at sites around the world, and enabling new collaborations.⁴³ The fMRIDC also provided training opportunities in technology development⁴⁴

35. Maryann E. Martone et al., *E-Neuroscience: Challenges and Triumphs in Integrating Distributed Data from Molecules to Brains*, 7 *NATURE NEUROSCIENCE*, 467, 467 (2004).

36. *Id.* at 468.

37. Van Horn, *supra* note 23, at 1324–25.

38. See Journal of Cognitive Neuroscience Instructions for Authors, J. COGNITIVE NEUROSCIENCE <http://jocn.mitpress.org/misc/ifora.shtml> (“The Journal of Cognitive Neuroscience no longer requires submission of imaging data to the National FMRI Data Center.”).

39. John Darrell Van Horn et al., *Sharing Neuroimaging Studies of Human Cognition*, 7 *NATURE NEUROSCIENCE*, 473–81 (2004).

40. Univ. of Cal. Santa Barbara, The fMRI Data Center, <http://www.fmridc.org/fmridc/aboutus/funding.html> (last visited Nov. 7, 2008).

41. Univ. of Cal. Santa Barbara, The fMRI Data Center, <http://www.fmridc.org/fmridc/aboutus/generalinfo.html> (last visited Nov. 7, 2008).

42. *Id.*

43. Interview with John Van Horn (Nov. 2008).

44. See John Darrell Van Horn & Alomit Ishai, *Mapping the Human Brain: New Insights from fMRI Data Sharing*, 5 *NEUROINFORMATICS* 146, 147

and, like other repositories that draw on policies for data sequence storage in the genetics community, gave anyone the right to publish findings based on mined fMRIDC datasets.⁴⁵ Authors whose papers are based on results from datasets obtained from the Data Center are expected to provide descriptive meta-information for data use, credit original study authors, and acknowledge the fMRIDC and accession number of the data set.⁴⁶

The Organization for Human Brain Mapping⁴⁷ (“OHBM”), an international professional organization dedicated to the progress of neuroimaging research, also favored the concept of brain data sharing for its potential to enable comparison of data across studies, improve reliability and reproducibility, promote meta-analyses, and create access to data for those who cannot afford neuroimaging equipment.⁴⁸ As the *JOCN* data sharing mandate brought the challenges of data sharing to the foreground, the OHBM quickly responded with a task force dedicated to the topic.⁴⁹ The work of the OHBM Neuroinformatics Subcommittee task force culminated in a 2001 *Science* publication framing the critical elements necessary for an informed discussion of the issues.⁵⁰ Among the most pressing were data content, data access, data ownership, database structure, and interaction with the community.⁵¹ The OHBM also highlighted issues of database structure, including whether hybrid structures should be constructed for the specific purpose of storing and maintaining neuroimaging data.⁵²

(2007) (noting that fMRI data from previously published studies can be used to train other neuroscientists).

45. Univ. of Cal. Santa Barbara, The fMRI Data Center, <http://www.fmridc.org/fmridc/help/faq.html#DataSharing> (last visited Nov. 7, 2008) (“ For data housed at the Center, anyone has the right to publish findings based on these datasets. Papers whose results are based on datasets obtained from the Center should credit the authors of the original study and acknowledge the Center and accession number of the dataset.”); Van Horn, *supra* note 23, at 1323.

46. Van Horn, *supra* note 23, at 1333.

47. The Organization for Human Brain Mapping, www.humanbrainmapping.org (last visited Oct. 25, 2008).

48. The Governing Council of the Organization for Human Brain Mapping, *Neuroimaging Databases*, 292 *SCIENCE* 1673, 1673 (2001).

49. *Id.* at 1674.

50. *Id.* at 1673.

51. *Id.* at 1675–76.

52. *Id.* at 1676.

Management of violations and a multitude of issues surrounding interactions with the community were further identified, pointing to the diversity of challenges associated with banked neuroscience data.⁵³

CONFIDENTIALITY, CONSENT, AND COMMERCIALIZATION

Internet-accessibility of databases has heightened concerns about consent and the confidentiality of research participant information. The federal human subject protection law mandates that all identifying information be removed from data prior to submission for sharing.⁵⁴ Latanya Sweeney of Carnegie Mellon University suggested that true de-identification of medical information—written records, images—may be inherently flawed since it is possible to match data to other databases and identify the individual.⁵⁵ By contrast, Amanda Bischoff-Grethe from the University of California—San Diego and colleagues have described a new technique for de-identifying images from magnetic resonance, that appears to be robust.⁵⁶ Nonetheless, new possibilities for reconstructing facial and cranial features from a brain image make old confidentiality rules about identifying information a particularly vexing problem today.⁵⁷ Moreover, while institutional ethical review, safety, and quality assurance are fundamental, prospective secondary data uses expand the horizon of these considerations.

Seminal work by Ellen Wright Clayton et al.⁵⁸ and others underscore the complexity of the underlying ethical, legal, and social problems surrounding the status, storage, and current and future use of human materials. The focus of these scholars in the mid-1990s was on organs, gametes, embryos, tissue, blood, and cells. Attention now is also on neuroscience data.

53. *Id.*

54. 45 C.F.R. § 46.101(4) (2008); John D. Van Horn & Michael S. Gazzaniga, *Databasing fMRI Studies—Towards a “Discovery Science” of Brain Function*, 3 NATURE REV. NEUROSCIENCE 314, 314 (2002).

55. See generally Latanya Sweeney, *Maintaining Patient Confidentiality When Sharing Medical Data Requires a Symbiotic Relationship Between Technology and Policy* (Mass. Inst. Tech. Artificial Intelligence Lab., Working Paper No. AIWP-WP344b, 1997).

56. Amanda Bischoff-Grethe et al., *A Technique for the Deidentification of Structural Brain MR Images*, 28 HUMAN BRAIN MAPPING 892, 902 (2007).

57. *Id.* at 893.

58. See generally Ellen Wright Clayton, *Ethical, Legal, and Social Implications of Genomic Medicine*, 349 NEW ENG. J. MED. 562 (2003).

As donors, patients, and research participants everywhere have become “sources,” consent, choice, contact, and controls are topics of ongoing interest and identified priorities for databanks.⁵⁹

Institutional Review Boards (“IRB”) are “administrative bodies established to protect the rights and welfare of human research subjects.”⁶⁰ Members of these boards are charged with reviewing research protocols and ensuring federal guidelines for protection of privacy and informed consent, among others are followed. The establishment of data banks added future uses of data to the topics under consideration. Leslie Wolf and Bernard Lo wrote specifically on the IRB issues in the control of future uses of data and disclosure of results to donors in research involving stored biologic materials.⁶¹ They found that IRBs address many significant issues but concluded that challenges remain.⁶² The authors identified best practices within institutions as those that embodied a rationale and examples in protocols, provided a checklist to walk investigators through pertinent issues, and highlighted particular issues that investigators might not anticipate.⁶³ Wolf and Lo further emphasized the need for scrupulous protection of the rights and welfare of subjects, especially children and those without decisional capacity.⁶⁴ Current IRB policies state that protocols using de-identified data do not need to go

59. See A. Cambon-Thomsen et al., *Trends in Ethical and Legal Frameworks for the Use of Human Biobanks*, 30 EUR. RESPIRATORY J. 373, 376-378 (2007) (Eur.).

60. Institutional Review Board Guidebook, Chapter I Institutional Administration, http://www.hhs.gov/ohrp/irb/irb_chapter1.htm (last visited Dec. 11, 2008).

61. Leslie E. Wolf & Bernard Lo, *Untapped Potential: IRB Guidance for the Ethical Research Use of Stored Biological Materials*, 26 IRB: ETHICS & HUMAN RES. 1,1 (2004).

62. *Id.* at 5–6.

63. *Id.* at 6–7.

64. See *id.* at 2 (noting a lack of consensus on the consent requirements for storing a child’s biological material); see also Jeffrey R. Botkin, *Preventing Exploitation in Pediatric Research*, 3 AM. J. BIOETHICS 31 (2003) (arguing that the inclusion of children in research should be guided more by the goal of protecting children from exploitation, and less by assumptions about formal decision making capacities); Leonard H. Glantz, *Conducting Research with Children: Legal and Ethical Issues*, 34 J. AM. ACAD. CHILD ADOLESCENT PSYCHIATRY 1283, 1285 (1996) (noting that being a child or being disabled has been historically associated with being questionably treated in a research setting).

through ethics review. However, given the possibility to reconstruct identifying features from brain scans, further exploration of whether mining neuroscientific data requires IRB approval is warranted.

Possible novel banking issues, such as group harm and group consent, have been addressed by commentators such as Michael J. Malinowski from Louisiana State University⁶⁵ and Henry T. Greely of Stanford University.⁶⁶ Possible group effects, in particular regarding confidentiality and consent, have resulted in opposition to some publicized projects.⁶⁷ In opposition, Clayton argued for more detailed content, scope, and transparency of consent, especially as withdrawal of data or material after it has been collected and banked is a key unresolved area.⁶⁸ A general blanket consent to all future research has not been considered sufficient to meet standards of consent⁶⁹—a reality faced by one of three partners in the HBP consortium whose research was held up for several years because the local IRB objected to the blanket consent that subjects were asked to provide.⁷⁰ If people must be given adequate information on which to base a decision, a consensus on permissible secondary uses should be developed if the subjects did not expressly consent to those uses.⁷¹ Indeed, a major continuing goal is how best to align practices of repositories with requirements of ethics committees.

Commercialization raises further ethical issues, including preventing exploitation of vulnerable populations, balancing

65. Michael J. Malinowski, *Technology Transfer in BioBanking: Credits, Debits, and Population Health Futures*, 33 J. L. MED. & ETHICS 54, 59 (2005) (applying an intellectual property approach to biobanking issues).

66. See Henry T. Greely, *Human Genomics Research: New Challenges for Research Ethics*, 44 PERSP. BIOLOGY & MED. 221, 221 (2001) (arguing that current recombinant DNA research fails to protect the interests of human subjects and imperil future research by violating the public's trust).

67. Melissa A. Austin, et al., *Genebanks: A Comparison of Eight Proposed International Genetic Databases*, 6 COMMUNITY GENETICS 37, 43 (2003).

68. Ellen Wright Clayton, *Informed Consent and Biobanks*, 33 J.L. MED. & ETHICS 15, 19 (2005).

69. *Id.* at 19–20 (noting that many are opposed to the use of blanket consent because it does not allow individuals to make an informed choice).

70. Anne Beaulieu, *Research Woes and New Data Flows: A Case Study of Data Sharing at the fMRI Data Center, Dartmouth College, USA*, in *THE PUBLIC DOMAIN OF DIGITAL RESEARCH DATA* 65, 85 (Paul Wouters & Peter Schröder eds., 2003) (Neth.).

71. See Edward J. Janger, *Genetic Information, Privacy and Insolvency*, 33 L. MED. & ETHICS 79, 79 (2005) (noting, particularly, the difficulty of addressing issues of secondary uses when a biobank faces insolvency).

costs and benefits, and avoiding conflicts of interest.⁷² Concerns about commercialization of information led to the cessation of gene banking in at least one country, Tonga.⁷³ International documents, in particular, suggest discomfort with the idea of gain from the transfer or exchange of human genetic material and information.⁷⁴ Commercial involvement in the development of useful products from tissue is generally not discouraged so long as there is attention to scientific and social norms. Mary Anderlik from Baylor College of Medicine points out that “although many uncertainties remain, consensus seems to be forming on a number of issues.... [and] a few countries have enacted general legislation providing for comprehensive regulation of biobanks through licensure.”⁷⁵ One example of intellectual property privileges and commercialization in neuroscience is represented by the Brain Resource Company, whose promotional material offers “large quality controlled database of normative subjects and with a range of clinical disorders,”⁷⁶ and provides fee-for-service analysis reports to clients.⁷⁷ In 2002 the OECD Working Group noted that although the short-term impact of proprietary databases on open neuroscience appeared to be small, long-term and larger

72. See Mary R. Anderlik, *Commercial Biobanks and Genetic Research: Ethical and Legal Issues*, 3 AM. J. PHARMACOGENOMICS 203, 204 (2003) (“[A] commercial orientation may be at odds with the values that should inform custodianship of this special resource, or lead to neglect of standards that uphold public trust in the research enterprise and ensure respect for the rights and interests of research subjects.”); Shun-Ichi Amari et al., *Neuroinformatics: The Integration of Shared Databases and Tools Towards Integrative Neuroscience*, 1 J. INTEGRATIVE NEUROSCIENCE 117, 123 (2002) (noting the necessity of guidelines governing rights of researchers and costs of maintaining databases); Mark A. Rothstein, *Expanding the Ethical Analysis of Biobanks*, 33 LAW, MED. & ETHICS 89, 90 (2005) (discussing potential harmful results for individuals and groups when data is not stripped of identifying information); BUYING IN OR SELLING OUT: THE COMMERCIALIZATION OF THE AMERICAN RESEARCH UNIVERSITY 133 (Donald G. Stein ed., 2004).

73. Bob Burton, *Proposed Genetic Database on Tongans Opposed*, 324 BRIT. MED. J. 443, 443 (2002) (discussing a failed proposal by Autogen to establish a database of genetic information on the population in Tonga in the South Pacific).

74. See Anderlik, *supra* note 72, at 206–207.

75. *Id.* at 203–04.

76. Peter Eckersley & Gary Egan, *Legal and Policy Questions for International Collaboration in Neuroscience*, 1 PROC. 9TH INT’L CONF. ON NEURAL INFO. PROCESSING 202, 202 (2002).

77. *Id.*

effects should be anticipated.⁷⁸ Following the 9th International Conference on Neural Information Processing, Peter Eckersley and Gary Egan of the University of Melbourne urged prospective consideration of issues arising from relationships between public and private contributions to neuroinformatics resources and the construction of a policy framework.⁷⁹

INCIDENTAL FINDINGS, DATA ANONYMIZATION, AND RECONTACT

Research with identifiable samples involves risk of discovery of unexpected and potentially unknown clinical significance, missed incidence, violation of the donor's privacy through discovery, and disclosure of sensitive information (intrinsic harm), or risk of discrimination by disclosure of information to third parties (consequential harms).⁸⁰ Participants must be told that when samples (data) are used anonymously, the participants cannot be given specific information about findings related to their samples. Bartha Knoppers from the University of Montreal, in Montreal, Canada has favored a coded model (or double-coded via a third party or "tissue trustee model") for biobank samples because it gives subjects an option to opt-out upfront or to recontact later.⁸¹ Mary Anderlik Majumder at Baylor College of Medicine has described an initiative funded by National Institute on Science and Technology to create a secure web-based consent mechanism for patients to communicate with researchers in a dynamic and anonymous fashion.⁸² But, as Clayton points out, recontact can be a real "wild card"⁸³: what investigators do when they are faced with undesired information from a research participant with whom they have had no prior contact is an open question. Since neuroinformatics resources will be populated by data and mined by investigators on a global scale, standardization of protocols for managing incidental findings, data anonymization and subject recontact are necessary.

78. See Amari, *supra* note 1.

79. Peter Eckersley & Gary Egan, *Legal and Policy Questions for International Collaboration in Neuroscience*, 1 PROCEEDINGS OF THE 9TH INTERNATIONAL CONFERENCE ON NEURAL INFORMATION PROCESSING (ICONIP'02) 202-05 (Lipo Wang et al., eds 2002).

80. Rothstein, *supra* note 72.

81. Bartha Maria Knoppers, *Biobanking: International Norms*, 33 L. MED. & ETHICS 7, 12 (2005).

82. Mary A. Majumder, *Cyberbanks and Other Virtual Research Repositories*, 33 L. MED. & ETHICS 31, 32 (2005).

83. Clayton, *supra* note 68, at 20.

The National Bioethics Advisory Commission⁸⁴ (“NBAC”) recommended that IRBs should develop general guidelines for disclosure of results in current or future research when (a) the results are scientifically valid and confirmed, (b) the results have implications for subjects’ health concerns, and (c) a course of action to ameliorate or treat the concern is readily available.⁸⁵ Although these guidelines provide a strong basis for framing approaches in neuroscience, they do not readily apply to brain incidental findings today. Discoveries about frequency and clinical significance, including false positives are ongoing, and treatment, especially in the case of certain neurodegenerative diseases, is still elusive. In the case of shared data, the Office for Human Research Protections⁸⁶ (“OHRP”) suggests that the Common Rule⁸⁷ does not apply to investigators who receive coded information as long as they do not have access to the code key.⁸⁸ However, conflicting regulations between OHRP, the Food and Drug Administration and the Health Information Portability and Accountability Act in the United States have “led to chaos” in that they fail to provide clear guidance or instruction.⁸⁹ Moreover, in light of the dynamic pace of scientific progress, refinement of ethical norms and changes in public opinion, approaches, and protocols may require adjustments that were not foreseeable at the outset.⁹⁰

84. The President’s Council on Bioethics, www.bioethics.gov (last visited Oct. 29, 2008); The President’s Council on Bioethics, http://www.bioethics.gov/reports/past_commissions/index.html (last visited Nov. 11, 2008) (NBAC expired in 2001 and was replaced by The President’s Council on Bioethics).

85. NAT’L BIOETHICS ADVISORY COMM’N, RESEARCH INVOLVING HUMAN BIOLOGICAL MATERIALS: ETHICAL ISSUES AND POLICY GUIDANCE. 72 (1999), *available at* http://www.bioethics.gov/reports/past_commissions/nbac_biological1.pdf.

86. Office for Human Research Protections, www.hhs.gov/ohrp (last visited Oct. 29, 2008).

87. *See* Protection of Human Subjects, 45 C.F.R. § 46 (2003). Part A is referred to as the “Common Rule” and is the federal policy governing human subjects. *Id.*

88. Clayton, *supra* note 68, at 16; OHRP, Guidance on Research Involving Coded Private Information or Biological Specimens, <http://www.hhs.gov/ohrp/humansubjects/guidance/cdebiol.htm> (last visited Dec. 11, 2008).

Clayton, *supra* note 68, at 16.

89. Rothstein, *supra* note 72, at 93.

90. *See* Mylene Deschenes & Clementine Sallee, *Accountability in Population Biobanking: Comparative Approaches*, 33 L. MED. & ETHICS 40, 41

With increasing demands comes the need for ongoing reform of regulations for protecting human research participants. Inadequate resources for IRBs and median costs to academic medical centers for the system of protecting participants of nearly \$750,000 per year per institution⁹¹ make essential the proactive embodiment of ethics principles that could enhance coherence and efficiency.

CULTURE OF DATA SHARING

While increased statistical power and cost efficiency are commonly noted as benefits of data sharing, proponents are not without opposition. In the genetics literature, researchers have reported intentionally withholding data for reasons related to the sheer workload associated with sharing, as well as to protect publication opportunities for themselves and other faculty, especially junior faculty and fellows.⁹² For brain imaging, for example, Arthur Toga argued that in order for data to be appropriately understood and used, the data must be considered in the context of the sample, methodology, and analysis with which they were collected and generated.⁹³ Patient confidentiality and the relinquishment of personal benefit constitute another central theme in resistance to the principles of brain data sharing.⁹⁴ Moreover, in the absence of a standard paradigm for collecting data, comparison across studies may be more difficult than expected. This issue also raises questions about who will be responsible for converting data into a standard format and how this procedure might take place.

The backing of the *JOCN* mission for data sharing was not unanimous.⁹⁵ While some journals remained uncertain and others favored the approach of informal encouragement rather than formal guidelines,⁹⁶ journal-by-journal disclosure of

(2005).

91. Jeremy Sugarman et al., *The Cost of Institutional Review Boards in Academic Medical Centers*, 352 *NEW ENG. J. MED.* 1825, 1826 (2005).

92. Eric G. Campbell et al., *Data Withholding in Academic Genetics: Evidence From a National Survey*, 287 *JAMA* 473, 478 (2002).

93. Arthur W. Toga, *Imaging Databases and Neuroscience*, 8 *THE NEUROSCIENTIST* 423, 423–24 (2002).

94. Peter Aldhous, *Prospect of Data Sharing Gives Brain Mappers a Headache*, 406 *NATURE* 445, 445–46 (2000).

95. *Id.*

96. See Gordon M. Shepherd, *Supporting Databases for Neuroscience Research*, 22 *J. NEUROSCIENCE* 1497 (2002) (encouraging authors of fMRI

policies for data sharing gained importance.⁹⁷ The five core “UPSIDE” principles (Uniform Principle for Sharing Integral Data and Materials Expeditiously) for sharing data and materials of *The Proceedings of the National Academy of Sciences*⁹⁸ (“PNAS”) state: (1) data that would enable other investigators to verify or replicate claims should be included in published reports; (2) if that information is too cumbersome for inclusion in a publication, it should be made freely available through other means (e.g., online); (3) by the time of publication, the data should be deposited in a publicly accessible repository that has been agreed upon by the authors; (4) means of accessing data should be anticipated by authors and addressed in the methods sections of publications; and (5) patented material should be made available under a license for research use.⁹⁹ The *Journal of Neuroscience* made banking genetic sequencing data mandatory, but fMRI data submission, for example, is only encouraged.¹⁰⁰ *Science* initially required its contributors in the field of protein data banking to deposit their genomic sequences and crystallographic coordinates in public databases like the Protein Databank and to wait a year before meeting the contributing requirement, but today requires that the data be deposited before publication.¹⁰¹ *Nature*¹⁰² and *PNAS*¹⁰³ also have this specific requirement now, and data-hold

articles accepted for publication in the *Journal of Neuroscience* to submit their data).

97. See Aldhous, *supra* note 94.

98. Proceedings of the National Academy of Sciences, www.pnas.org (last visited Oct. 29, 2008).

99. Nicholas R. Cozzarelli, *UPSIDE: Uniform Principle for Sharing Integral Data and Materials Expeditiously*, 101 PROC. NAT’L ACAD. SCI. 3721, 3721 (2004).

100. JOURNAL OF NEUROSCIENCE, POLICIES, http://www.jneurosci.org/misc/ifa_policies.shtml (last visited Nov. 8, 2008).

101. SCIENCE, DATABASE DEPOSITION POLICY, http://www.sciencemag.org/about/authors/prep/gen_info.dtl#datadep (last visited Nov. 8, 2008).

102. NATURE, AVAILABILITY OF DATA AND MATERIALS, http://www.nature.com/authors/editorial_policies/availability.html (last visited Nov. 8, 2008); Nature Publishing Group Editorial Policies: Availability of Data and Materials, http://www.nature.com/authors/editorial_policies/availability.html (last visited Dec. 11, 2008).

103. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, JOURNAL POLICIES, <http://www.pnas.org/site/misc/iforc.shtml#submission> (last visited Nov. 8, 2008); Proceedings of the National Academy of Sciences of the United States, Information for Authors Page section vii: Materials and Data

policies have been shortened or have disappeared altogether. Both of these journals now require that investigators submit data sets to either large database and supply accession numbers before publication.

In studying the issue of trust in data sharing practices and policies, Beaulieu found that sociological hurdles were profound even though the coupling of sharing and publication was designed to be a trust-building mechanism.¹⁰⁴ Even before Beaulieu's 2003 work was published, Ari Patrinos, Director of Biological and Environmental Research at the Department of Energy, was quoted as saying that "[i]t would be 'a mistake' ... to adopt a simple rule forcing authors to choose between releasing control of all their data at publication or not publishing."¹⁰⁵ Today's new requirements by the NIH in the United States and the Canadian Institutes of Health Research¹⁰⁶ ("CIHR") in Canada to make all scientific articles, whether deriving fully or partially from NIH-sponsored projects, electronically available through the Internet adds another layer of complexity to the discussion.¹⁰⁷ The CIHR requires that all research papers generated through funded projects be freely accessible through the publisher's website or an online repository within six months of publication.¹⁰⁸ NIH requires data sharing in several areas, such as DNA sequences, mapping information, and crystallographic coordinates, and expects "the timely release and sharing' [of data] to be no later than the acceptance for publication of the main findings from the final data set."¹⁰⁹ These policies apply to all data from

Availability, <http://www.pnas.org/site/misc/iforc.shtml#viii> (last visited Dec. 11, 2008).

104. Beaulieu, *supra* note 70, at 72; see also Jeremy P. Birnholtz & Matthew J. Bietz, *Data at Work: Supporting Sharing in Science and Engineering*, in PROCEEDINGS OF THE 2003 INTERNATIONAL ACM SIGGROUP CONFERENCE ON SUPPORTING GROUP WORK 339 (Mark Pendergast et al. eds., 2003) (suggesting that data sharing contributes to the establishment and maintenance of communities).

105. Eliot Marshall, *Clear-Cut Publication Rules Prove Elusive*, 295 SCIENCE 1625 (2002).

106. Canadian Institutes of Health Research, www.cihr-irsc.gc.ca (last visited Oct. 29, 2008).

107. See Giorgio A. Ascoli, *Looking Forward to Open Access*, 3 NEUROINFORMATICS 1, 2 (2005) (discussing the benefits and possible harmful side effects of open access to NIH publications).

108. OPEN ACCESS TO HEALTH RESEARCH PUBLICATIONS: CIHR UNVEILS NEW POLICY, Sept. 4, 2007, <http://www.cihr-irsc.gc.ca/e/34851.html>.

109. FINAL NIH STATEMENT ON SHARING RESEARCH DATA, Feb. 26, 2003, <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>.

funded research, not just published data. All grant proposals with direct costs greater than \$500,000 in any single year are expected to have a section on data-sharing in their application.¹¹⁰ Since January 1, 2008, the CIHR requires grantees to deposit all bioinformatics, atomic, and molecular coordinate data into a public database immediately upon publication of results.¹¹¹

The lack of clear funding-agency policies in the face of competing interests, “often far removed from academic research,” has been reported to pose problems for scientists,¹¹² just as perilously as unstable funding has done.¹¹³ Investigators have argued that administrative and organizational management and diversity in science may necessitate a variety of institutional data management approaches, and that establishing and aligning this infrastructure will require proactive, ongoing and dedicated budgetary planning.¹¹⁴ Maximizing effectiveness through the involvement of researchers is critical, since many are unaware of existing policies and opportunities even within their institutions and organizations.¹¹⁵

Existing heterogeneity in international policies makes data sharing across borders potentially even more difficult.¹¹⁶ In the United States, federal government databases are not copyright protected.¹¹⁷ In the European Union, government databases are eligible for protections under law.¹¹⁸ Even within some countries, practices vary with major funding agencies subscribing to different principles.¹¹⁹ Peter Arzberger,¹²⁰ Daniel

110. *Id.*

111. POLICY ON ACCESS TO RESEARCH OUTPUTS, Sept. 2007, <http://www.cihr-irsc.gc.ca/e/34846.html>.

112. Peter Arzberger et al., *An International Framework to Promote Access to Data*, 303 SCIENCE 1777, 1777 (2004).

113. Zeeya Merali & Jim Giles, *Databases in Peril*, 435 NATURE 1010, 1010 (2005) (discussing difficulty of obtaining funding for already existing research projects and databases).

114. Arzberger, *supra* note 112, at 1777.

115. *Id.*

116. *Id.*

117. *Id.*

118. *Id.*

119. *Id.*

120. Arzberger, *supra* note 112, at 1778.

Gardner,¹²¹ among others called for an empirical analysis of views from researchers, funders, and policy-makers, and for solutions to barriers through guidelines for best practices— an analysis that, to the knowledge of these authors, has yet to be conducted formally and reported.¹²² An investment in large scale clinical research and sharing practices such as that promoted by NIH’s roadmap for “Re-engineering the Clinical Research Enterprise”¹²³ will only be returned if it enables both progress from a broad community of scientists and proper coordination and integration of resources.¹²⁴ Finally, in this new era of banking brain data, a range of legal actors—including prosecutors, national security authorities, other governmental agencies, and litigants such as individuals, insurers, and employers — may seek information from banked sources. Individuals who are data sources may also try to access information about themselves for legal use, for example, to prove excuse or mitigating circumstances. Because data can be banked indefinitely, these issues can persist throughout the lifetime of a source individual. Indeed, because some of these data are genomic, they may have implications for offspring and other family members into the indefinite future.¹²⁵

To map the terrain of data banking issues that may challenge us, Eric Racine, formerly from Stanford University and now Director of Neuroethics at the Institute of Clinical Research and Medicine in Montreal Canada, and one of the authors (Judy Illes) conducted an exploratory study of databases to provide empirical evidence for which information is readily accessible on the Internet. Our goal was to determine the preliminary nature of information available on the data sharing websites, the extent of the information, and consistency. We conducted an analysis of a sample of 58 biobanks from top tier returns (10-25) of a search for gene (N=11), blood (N=11), tissue (N=15) and twin study (N=21) data sharing, with duplicates deleted on a finite set of variables. Our results are shown in Table 1.

121. Gardner, *supra* note 34.

122. Arzberger, *supra* note 112, at 1778.

123. NIH ROADMAP FOR MEDICAL RESEARCH, www.nihroadmap.nih.gov/clinicalresearch (last visited Oct. 30, 2008).

124. Insel, *supra* note 34, at 427.

125. Susan M. Wolf et al., *Managing Incidental Findings in Human Subjects Research: Analysis and Recommendations*, 36 *J.L.MED. & ETHICS*, 219, 219–48 (2008).

	Location	Host	Year	Information about privacy	Type of host (public, private)	Information about access	Rationale or background information	References or links to relevant laws or regulations
Biobanks (N=58)	90% (52)	88% (51)	81% (47)	79% (46)	74% (43)	66% (38)	41% (24)	3% (2)
Gene Banks (N=11)	100% (11)	100% (11)	91% (10)	82% (9)	100% (11)	91% (10)	82% (9)	19% (2)
Tissue Banks (N=15)	73% (11)	100% (15)	73% (11)	87% (13)	100% (15)	30% (5)	67% (10)	0
Blood Banks (N=11)	100% (11)	73% (8)	91% (10)	82% (9)	82% (9)	91% (10)	45% (5)	0
Twin Data Banks (N=21)			76% (16)	90% (19)	30% (8)	62% (13)	71% (15)	0

Table 1. Extent and consistency of information for a sample of biobanks (N=58). Percentages (and numbers in parentheses) reflect the presence of data or information for that cell.

Overall, we found that the sites provide a rich resource of accessible information online that largely describes the nature and infrastructure of databases. However, the availability of detailed information decreases as one queries for privacy, host type, access, rationale, and regulations. Of our small sample, gene banks were the most complete of the biobanks surveyed.

A NEUROETHICS DATABASE SHARING APPROACH

Data banking and data sharing, especially in relation to the open-endedness of use and circulation of data, raise novel ethical issues. Open-endedness can be both a challenge and a desire in the eyes of developers and users. It can be problematic for bodies charged with the regulation of research ethics. Scholars have observed that in the field of genetics biobanks, such as GenBank, have “revolutionized the research fields that depend on DNA sequences;”¹²⁶ some neuroscientists expect the power of brain databases to similarly catapult the

126. Baringa, *supra* note 23, at 44.

field to new levels of research and discovery.¹²⁷ Continued innovation and expansion in brain data sharing is supported by efficiency and economy in terms of both experimental power and real dollars.¹²⁸

With steadily growing sharing practices in the neurosciences, this is an opportune time to investigate the attendant ethical issues, and what we are still missing. Rivka Ravid, at the Netherlands Institute for Neurosciences, places “[a]dequate funding for research on [biobanks] ...Standard evaluation protocols for audit of [biobank] performance.... Internationally accepted [standard operating practices] which will facilitate exchange and sharing of specimens and data with the scientific community. . .[and an i]nternationally accepted code of conduct” as top priorities.¹²⁹ These goals should not be met with tools that are prescriptive. Rather, results are needed that enable a broad approach to systems development¹³⁰ and that have an empowering and streamlining effect on existing and newly evolving database practice standards. Guidelines for negotiating scientific collaboration are best formulated as logical, hypothetical guides to ethical judgment about the task at hand, rather than categorical or absolute rules. From a pragmatic perspective, ethical guidance is appropriately subject to reconsideration as discoveries are made and as they are relevant to the time, place, and purpose of inquiry.

Therefore, to effectively usher in a next generation of brain

127. Kalyani Narasimhan, *Scaling Up Neuroscience*, 7 NATURE NEUROSCIENCE 425, 425 (2004) (noting that gathering vast amounts of data, along with effectively mining and interpreting this information is crucial to further developing brain research).

128. See Xinhau Cao et al., *A Web-Based Federated Neuroinformatics Model for Surgical Planning and Clinical Research Applications in Epilepsy*, 2 NEUROINFORMATICS 101, 101 (2004) (emphasizing the need for efficient data sharing among brain researchers); Rex M. Jakobovits et al., *WIRM: An Open Source Toolkit for Building Biomedical Web Applications*, 9 J. AM. MED. INFORMATICS ASS'N 557, 557 (2002) (emphasizing a need for biomedical systems that facilitate data sharing); Yi-Ren Ng et al., *BrainImageJ: A Java-Based Framework for Interoperability in Neuroscience, with Specific Application to Neuroimaging*, 8 J. AM. MED. INFORMATICS ASS'N 431 (2001) (describing a computer-based infrastructure designed to share data analysis tools).

129. Rivka Ravid, *Standard Operating Procedures, Ethical and Legal Regulations in BTB (Brain/Tissue/Bio) Banking: What Is Still Missing?*, 9 CELL TISSUE BANKING 151, 151 (2008).

130. See Jeffrey Jones & Hugh Preston, *Big Issues, Small Systems: Managing with Information in Medical Research*, TOPICS IN HEALTH INFO. MGMT. Aug. 2000, at 45, 46.

data sharing, which is both necessary and inevitable, empirically-generated data are needed in the following three major categories: data organization; data sharing; and the roles, rights and responsibilities for sharers and bankers.

DATA ORGANIZATION

Research is needed to examine both successful and failed practices of the past five, ten and twenty years, and how they inform the next five to ten years. What features characterize those that have met good receptivity? Why have others encountered significant obstacles or failed to endure? For example, uniformity of requirements between journals or research sponsors has obvious positive implications for streamlined data processing. The downside to such homogeneity, including the time involvement for preparing datasets for sharing is unknown.

Surely better and richer organization of meta-data is one response. X-batch, a software package that provides analysis automation and data management for fMRI neuroimaging laboratories, instantiated an ontology with a detailed record of all fMRI analyses performed and can be part of larger system for neuroimaging data management, sharing, and visualization.¹³¹ The cost-benefit analysis of different approaches to data organization requirements or options has not been studied rigorously. Other types of trade-offs relate to whether standardizing data acquisition to maximize sharing efficiency has an impact on innovation and scientific creativity and whether this has had a positive impact or a stifling one.

A structured organization for disaster recovery and the implications of re-resourcing the same data, effectively giving us an archaeology of our brains in fifty years, are a few of the remaining, but not by any means the least, important challenges for data organization for the next generation.

GROUND RULES FOR DATA SHARING

The bioethics literature, and more recently the neuroethics literature, is replete with discussions about the examination of guidelines, policies, and regulations for the governance of scientific processes. Promotion of professional self-regulation

131. Andrew V. Poliakov et al., *Unobtrusive Integration of Data Management with fMRI Analysis*, 5 NEUROINFORMATICS 3, 3–10 (2007).

over the development of regulations by external bodies has been a particularly common theme in the neuro-focused literature. It is highly relevant to the question of who should establish and update rules for the banking and sharing of neuro data, and ensure an evenness of requirements among investigators. To this end, clear definition of responsible use and re-use of data from banked but not yet published, as well as published, sources is needed.

Once government gets databases started, where does the responsibility for maintenance and sustainability lie? If the commercial sector is to have an increasing role, what financial investments in the organization of data sharing and practices are likely to yield the highest return and what guidelines are needed to ensure ethically sound fiscal benefit? In this regard, disparities in access to banked data that may in fact mimic disparities in the population between not-for-profit and for-profit sharers and accessors (not to mention poorly-funded versus well-funded investigators) are foreseeable and may well give rise to biases in who gets to do business in the future and how. Guidance is needed to protect against discriminatory practices for this eventuality, as well as for issues surrounding ownership and intellectual property of secondary data sets, results, and follow-on innovation.

ROLES, RIGHTS, AND RESPONSIBILITIES

Success in the next generations of data sharing and banking will be driven by the growing motivation by all relevant actors to participate. Remaining reservations to do so are justified by unknowns discussed above for data organization and ground rules, and by this next category of discovery that concerns rights, roles, and responsibilities for sharers and bankers. For example, where are the boundaries of responsibility to respond to investigators with heavy needs or to provide data to investigators in competing groups? The impact of sharing practices on review and readership, selection of journals and sponsors, and on advancement and promotion are related, unsolved meta-academic challenges. With whom does responsibility for re-contact lie in cases of unexpected clinical findings or the discovery of erroneous analyses? These challenges will no doubt gain ever more attention as imaging genomics, the complex bridge between neuroscience and genetics and plausibly the most powerful emerging tool for neurology and psychiatry, will change the neuroscience practice

of the future.

CONTOURS OF AN ONGOING DEBATE AND CONCLUSIONS

The professional community is steadily moving away from post-event policing to proactive engagement of ethics. The best support for this momentum will come from educating the next generation of neuroscientists in critical thinking about the ethical implications of their research, including the ethics of data sharing practice. Luis Marengo from Yale University points out that “[i]t is crucial to orchestrate technologies such as database mediators, metadata repositories, semantic metadata annotations, and ontological services.”¹³² There is much in this regard to learn from past lessons in genomics and international collaborations.¹³³ Interoperability between studies will yield results that maximize the leverage envisioned for data organized for sharing and accessing. Results and recommendations drawn from in-depth comparison studies between neuroscience, genomics, and other domains rich with data will elucidate solutions to difficult organizational challenges, break down cultural barriers involving relevant ground rules, and lift the ethical cloudiness that still surrounds roles and responsibility.

132. Luis Marengo et al., *Interoperability Across Neuroscience Databases*, in *NEUROINFORMATICS* 23, 23 (Chiquito Joaquim Crasto ed., 2007).

133. See B.M. Knoppers et al., *The Public Population Project in Genomics (P³G): A Proof of Concept?*, 16 *EUR. J. HUM. GENETICS* 664 (2008) (discussing building a worldwide collaborative infrastructure to foster interoperability between studies in human population genomics).