14. J. Bollen *et al.*, *PLoS ONE* **4**, e4803 (2009).
15. J. Ginsberg *et al.*, *Nature* **457**, 1012 (2009).
16. J. Pratt, J. Rhine, B. Smith, C. Stuart, J. Greenwood, *Extra-Sensory Perception After Sixty Years: A Critical Appraisal of the Research in Extra-Sensory Perception* (Holt, New York, 1940).
17. G. V. Glass, M. L. Smith, *Educ. Eval. Policy Anal.* **1**, 2 (1979).
18. S. A. Greenberg, *BMJ* **339**, b2680 (2009).
19. J. E. Hunter, F. L. Schmidt, *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (Sage, Thousand Oaks, CA, ed. 2, 2004).
20. Cochrane Collaboration Reviewer's Handbook, www.cochrane.org/training/cochrane-handbook.
21. S. M. Stigler, *The History of Statistics: The Measurement of Uncertainty Before 1900* (Belknap/Harvard Univ. Press, Cambridge, MA, 1986).
22. K. Pearson, *Br. Med. J.* **3**, 1243 (1904).
23. K. O'Rourke, *J. R. Soc. Med.* **100**, 579 (2007).
24. L. V. Hedges, *Stat. Sci.* **7**, 246 (1992).
25. J. P. Ioannidis, *PLoS Med.* **2**, e124 (2005).
26. A. S. Gerber, N. Malhotra, *Sociol. Methods Res.* **37**, 3 (2008).
27. A. Rzhetsky, I. Iossifov, J. M. Loh, K. P. White, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4940 (2006).
28. U. Shwed, P. S. Bearman, *Am. Sociol. Rev.* **75**, 817 (2010).
29. M. Cokol, I. Iossifov, C. Weinreb, A. Rzhetsky, *Nat. Biotechnol.* **23**, 1243 (2005).
30. M. Cokol, R. Rodriguez-Esteban, A. Rzhetsky, *Genome Biol.* **8**, 406 (2007).
31. J. P. Ioannidis, T. A. Trikalinos, *J. Clin. Epidemiol.* **58**, 543 (2005).
32. M. Cokol, I. Iossifov, R. Rodriguez-Esteban, A. Rzhetsky, *EMBO Rep.* **8**, 422 (2007).
33. J. P. Ioannidis, *JAMA* **294**, 218 (2005).
34. T. S. Kuhn, *The Essential Tension: Selected Studies in Scientific Tradition and Change* (Univ. of Chicago Press, Chicago, 1977).
35. A. Tversky, D. Kahneman, *Science* **185**, 1124 (1974).
36. M. Solomon, *Philos. Sci.* **59**, 439 (1992).
37. D. L. Smail, *On Deep History and the Brain* (Univ. of California Press, Berkeley, CA, 2008).
38. J. Henrich, S. J. Heine, A. Norenzayan, *Behav. Brain Sci.* **33**, 61 (2010).
39. M. E. Harris, Ed., *Ways of Knowing: New Approaches in the Anthropology of Experience and Learning* (Berghahn, Oxford, 2007).
40. E. Martin, *Signs* **16**, 485 (1991).
41. R. Barnhardt, *Anthropol. Educ. Q.* **36**, 8 (2005).
42. K. Knorr-Cetina, *Epistemic Cultures: How the Sciences Make Knowledge* (Harvard Univ. Press, Cambridge, MA, 1999).
43. R. Collins, *The Sociology of Philosophies: A Global Theory of Intellectual Change* (Belknap/Harvard Univ. Press, Cambridge, MA, 1998).
44. S. Wuchty, B. F. Jones, B. Uzzi, *Science* **316**, 1036 (2007); 10.1126/science.1136099.
45. M. E. J. Newman, *Phys. Rev. E* **64**, 016131 (2001).
46. M. E. J. Newman, *Phys. Rev. E* **64**, 016132 (2001).
47. M. E. J. Newman, *Proc. Natl. Acad. Sci. U.S.A.* **101** (suppl. 1), 5200 (2004).
48. D. J. Kevles, *The Physicists: The History of a Scientific Community in Modern America* (Knopf, New York, ed. 1, 1978).
49. T. Lenoir, *Instituting Science: The Cultural Production of Scientific Disciplines* (Stanford Univ. Press, Stanford, CA, 1997).
50. P. Coffey, *Cathedrals of Science: The Personalities and Rivalries That Made Modern Chemistry* (Oxford Univ. Press, London, 2008).
51. S. Allesina, http://arxiv.org/abs/0911.0344 (2009).
52. S. Thurner, R. Hanel, http://arxiv.org/abs/1008.4324 (2010).
53. J. A. Evans, *Soc. Stud. Sci.* **40**, 757 (2010).
54. E. M. Armstrong, D. P. Carpenter, M. Hojnacki, *J. Health Polit. Policy Law* **31**, 729 (2006).
55. J. A. Evans, J. Reimer, *Science* **323**, 1025 (2009).
56. myExperiment, www.myexperiment.org.
57. The Polymath Blog, http://polymathprojects.org.
58. Zooniverse, www.zooniverse.org/home.
59. WolframAlpha, www.wolframalpha.com.
60. E. Evangelou, T. A. Trikalinos, J. P. Ioannidis, *FASEB J.* **19**, 1943 (2005).
61. R. Foote, *Science* **318**, 410 (2007).
62. H. M. Collins, *Sci. Stud.* **4**, 165 (1974).
63. T. M. J. Fruchterman, E. M. Reingold, *Softw. Pract. Exper.* **21**, 1129 (1991).

64. This research benefited from NSF grant 0915730 and responses at the U.S. Department of Energy's Institute for Computing in Science (ICiS) workshop "Integrating, Representing, and Reasoning over Human Knowledge: A Computational Grand Challenge for the 21st Century." We thank K. Brown, E. A. Cartmill, M. Cartmill, and two anonymous reviewers for their detailed and constructive comments on this essay.

10.1126/science.1201765

PERSPECTIVE

Access to Stem Cells and Data: Persons, Property Rights, and Scientific Progress

Debra J. H. Mathews,[1]* Gregory D. Graff,[2] Krishanu Saha,[3] David E. Winickoff[4]

Many fields have struggled to develop strategies, policies, or structures to optimally manage data, materials, and intellectual property rights (IPRs). There is growing recognition that the field of stem cell science, in part because of its complex IPRs landscape and the importance of cell line collections, may require collective action to facilitate basic and translational research. Access to pluripotent stem cell lines and the information associated with them is critical to the progress of stem cell science, but simple notions of access are substantially complicated by shifting boundaries between what is considered information versus material, person versus artifact, and private property versus the public domain.

Access to data and materials is critical to the progress of science generally (*1*–*3*), but plays a particularly important role in stem cell science. In the field of pluripotent stem cell research, data and information associated with cell lines are essential to their utility and management. A number of factors currently limit the sharing of data and materials in the field, including strategic behavior of individual scientists, ethical intricacies in using human cell lines, and a complex landscape of intellectual property rights (IPRs). Efforts to manage property relations and develop an accountable system of stewardship to handle the ethical and legal obligations to materials donors could help ameliorate these obstacles, benefiting academic and industry scientists and promoting both basic and translational research.

Difficulties in procuring and managing human cells arise because they are increasingly transcending three distinctions drawn in ethics, law, and common practice—distinctions between information and materials, persons and artifacts, and private property and the public domain (Fig. 1). Some allowance for blurring between and across these conventional distinctions can be useful and productive for science, but it also creates substantial challenges. For example, a cell line is much richer as a research tool by virtue of its connection to an individual human, but that connection also raises privacy concerns if the tool is shared.

Work to Date, Work to Do

Other research communities in the life sciences have experienced problems similar to those outlined above, and some have addressed them through new institutions such as public DNA sequence databases, tissue banks, and mouse repositories (*3*). Within stem cell science, there are important efforts under way to improve access to both cell lines and associated information [e.g., UK Stem Cell Bank, European Human Embryonic Stem Cell Registry (hESCreg)] (*4*). However, these efforts struggle to keep pace with evolving data requirements and the proliferation of new induced pluripotent stem cell lines and do not provide all of the information about

[1]Johns Hopkins Berman Institute of Bioethics, Deering Hall, 208, 1809 Ashland Avenue, Baltimore, MD 21205, USA. [2]Department of Agricultural and Resource Economics, B328 Clark Hall, Colorado State University, Fort Collins, CO 80523, USA. [3]The Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA. [4]Department of Environmental Science, Policy, and Management, 115 Giannini Hall, University of California, Berkeley, CA 94720 USA.

*To whom correspondence should be addressed. E-mail: dmathews@jhu.edu

cell lines that scientists may need. Existing banks and repositories also suffer unnecessary overlap and duplication of effort and lack interoperability, limiting their utility to scientists (*3*, *5*, *6*).

For example, a recently issued statement by the Hinxton Group (*7*) notes that problems with access are twofold, involving access both to cell lines themselves and to critical information about those lines, including technical, provenance, and IPRs characteristics. They note, "… many cell lines are being derived and characterized, though not all lines are being published in the literature, even in the academic sector. Furthermore, useful cell lines created from human materials (especially those created with public funds) and their associated data should be distributed and used widely, constrained only by the wishes of the materials' donors" (*7*). In addition, the statement observes that we currently have "a situation in which even a diligent stem cell researcher or entity that wishes to respect IPRs will face considerable uncertainty and enormous costs if they try to survey the IPRs landscape" (*7*). To address these concerns, the Hinxton Group recommended developing publicly available electronic hubs for accessing a range of relevant data linked to individual stem cell lines. The vision is to coordinate existing resources—including the UK Stem Cell Bank, the NIH Human Embryonic Stem Cell Registry, the International Stem Cell Registry at the University of Massachusetts, the European Human Embryonic Stem Cell Registry, and the International Stem Cell Banking Initiative—together with the patent landscaping efforts of the Japan Patent Office and the UK Intellectual Property Office, and academic efforts such as the Stanford Program on Stem Cells in Society (*7*).

When designing resources to facilitate access to information and materials, careful thought must be given to the kinds of data needed and how coexistence of different categories of data may influence a resource's utility. Public provision of some kinds of data (e.g., identifying patient information, and proprietary information) would create substantial ethical and economic concerns; proper provision will require creative solutions to both legal and technical challenges.

The three crucial boundaries—between information and materials; persons and artifacts; and private property and the public domain—have become highly dynamic, shifting relative to one another as the field advances. If the task of developing common material and data resources is conceptualized as one of facilitating and managing these three dualisms, it may help explain why existing resources have had difficulty gaining traction. Such a perspective also highlights the kind of functions these resources must perform to be really useful. The challenge is to allow productive blurring between these categories (e.g., tight connections between information and material), while maintaining critical boundaries (e.g., protecting donor privacy). Any successful architecture must manage these tensions capably.

Information Versus Materials

Today, a strong and durable cut between what constitutes "data" and what constitutes "materials" cannot be easily made. For example, the recent federal court decision invalidating a patent to Myriad Genetics on a gene associated with breast cancer (*8*) found that the informational nature of DNA, not its "physical embodiment," was more meaningful in considering the validity of Myriad's patent. With DNA, productive use of the material is, today, inextricably linked to the information it embodies. A similar principle can be applied to human cell lines. Technical data (e.g., derivation, culture history, genetic and epigenetic characteristics), donor information (e.g., provenance, medical history, family history), and ownership information [e.g., IPR, Material Transfer Agreements (MTAs), contract conditions] together determine the utility of a cell line. For example, if a donor has, through an informed-consent document, agreed just to a limited set of uses of their cells, that information must always accompany those cells and govern their use. Furthermore, since the cells are alive, what a particular cell line "is" today may be different from what it "is" at a different point in the future. Although a cell line does not change categorically (it's the same cell line), it does change qualitatively (e.g., at the genetic or epigenetic level), and those simultaneous changes in material and data are real, significant, and inextricably linked. Other associated information may also change—such as medical events experienced by the donor or changes in ownership rights relevant to the line—and may affect the utility of that line. For example, noting that a donor responded to a particular small-molecule therapy could make the corresponding cell line useful for genetic and biochemical studies on mechanism of action.

Persons Versus Artifacts

The connection to human donors complicates how we regard human cell lines. The lines are simultaneously human-made research tools and entities derived from individual human beings. The cells containing a donor's complete genome, and information critical to the use of the artifact, or discovered in the course of work with that artifact, will constitute information about that person. Furthermore, for disease-specific pluripotent cell lines, clinical and phenotypic information about the materials donor is increasingly collected. For all cell lines, provenance information is only of interest because informed-consent documents outline researchers' contractual obligations to the donor as a legal person.

Recent cases make it clear that what happens to human research materials, especially "immortalized cell lines," can be deeply meaningful to those from whom the materials have come. A recent book about Henrietta Lacks (*9*) and the creation of the HeLa cell line has spurred new discussions about informed consent, commercial use, and, importantly, the continuity between human research subjects and the cells derived from their bodies. Similar issues were raised in a dispute between the Havasupai tribe and Arizona State University, in which tribal members alleged that blood samples were used in genetic studies for which they had not given permission (*10*).

Researchers currently protect confidentiality of donors' data by one of two methods: anonymization, where identity is irreversibly severed from the material to prevent any future re-identification; or de-identification, where coded and linked identifying information is retained separately from the material. As the paradigm of single-center cell banking is increasingly replaced by multisite arrangements involving institutions across many jurisdictions, there are concerns that these confidentiality practices may be insufficient (*5*), and that the melding of persons and artifacts may lead to unprecedented violations of privacy and consent (*11*).

Blurred Boundaries In Human Stem Cell Research

Person ↔ Artifact		Information ↔ Materials		Private ↔ Public	
Deserving of autonomy, privacy, and respect	Instrument of research community	Biological facts about the cells and materials donor	Manipulatable research objects	Private asset for commercial translation	Common resource for academic and commercial use

Fig. 1. Blurred boundaries in human stem cell research.

Private Property Versus Public Domain

Following trends seen elsewhere in the sciences, stem cell researchers—and the companies or universities for which they work—are increasingly taking private ownership of early-stage technologies, including cell lines, genes, and associated data (Fig. 2). Simultaneously, researchers in the field draw upon a common repository of knowledge and technologies considered to be in the public domain. However, the boundary demarcating what is public from what is private has become fluid and ill defined.

In theory, that which is in the public domain is defined by the absence of private legal claims of ownership and control. Thus, to the extent that private property is not clearly demarcated, the public domain is left ill defined. In stem cell science, the landscape of IPRs is complex, and its boundaries fuzzy. IPRs are not uniformly recognized, registered, or enforced globally: A technology may be closely held by a private owner in one country and effectively left in the public domain in others. Different, inseparable aspects of a technology may be subject to separate property claims; for example, patents could cover both the process to create pluripotent cells and the reagents necessary to do so. Moreover, multiple, narrow claims over interdependent aspects (or uses) of given technologies can create dense thickets of ownership claims that are costly to negotiate and transact.

This can be particularly problematic when information and materials are inseparable, as for a given technology, they and their uses may be protected separately—even by different types of property right. Materials may be owned as physical property. Methods of derivation or propagation may be patented. Associated information may simply be held as confidential or maintained in a private database.

The complexity of IPRs claims can also be compounded by the person-artifact dualism that characterizes human cell lines. For, in addition to any third-party IPRs claims over a cell line, methods associated with its derivation, or uses of it as an artifact, the donor may have legitimate personal property claims over the cells or their use, as well as privacy rights over associated information.

Full information and well-defined property rights are necessary conditions for markets to function efficiently. Without these conditions met in the market for stem cell technologies, search costs, transaction costs, and risk are imposed and detract from the incentives that IPRs are intended to provide. What is needed is as reliable information as possible on where stem cell–related IPRs have and have not been claimed, held valid, and remain in force. This will, moreover, enhance the reliability of the public domain.

A Promising Start

The proposed information and materials hubs for stem cell research (*7*) may be a solution. Such hubs could improve access to data and materials generally and serve a gate-keeping function for access to various stakeholders, being mindful of ethics and IPRs concerns and providing a solution for mediating the complex and blurred distinctions described above.

The construction of such a resource could begin with developing a centralized portal for access to existing resources, one that aggregates key data characterizing their available materials. Additional features, such as IPRs information, and provenance and consent characteristics, could be added as funding becomes available to support the necessary programming and database research. The integration of these kinds of information would produce the greatest value-added to the community (*12*). Critical to the success of such a resource are commitments from the members of the scientific community to contribute to and curate the information in the resource, and from funding agencies to support the work.

There are substantial challenges in developing such a hub, however, including who will fund it, who will do the work, what the resource will look like, where it will reside administratively, and how the various blurred distinctions will be facilitated and managed in practice. We need to think critically about the design of data architectures to provide gate-keeping functions for access by various stakeholders, such as where restrictions on use or existing IPRs require formal negotiations and legal agreements. Expertise needs to be developed across these domains, with fresh thinking about the dualisms and how to manage them. Further research may be needed to learn how potential donors view their tissues, their relationship to those tissues following donation, and their own rights relative to third-party IPRs claims over those tissues.

These issues are increasingly prevalent across biomedical science, including biobanking, genetics and genomics, and personalized medicine. Community resources of this sort are emerging as necessary infrastructure of the scientific enterprise, no longer an aberration or exception to the rule, but rather the way research communities must function to move forward.

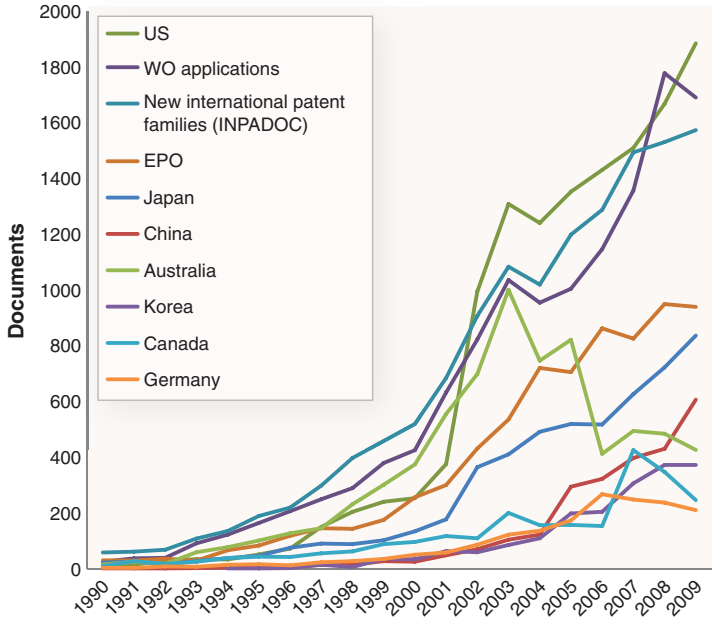

Fig. 2. Stem cell patents and patent applications published by various patent offices. Data source: Thomson Innovation (2010), queried using methods of Bergman and Graff (*13*).

References

1. Editor, *Nature* **461**, 145 (2009).
2. D. Field *et al.*, *Science* **326**, 234 (2009).
3. P. N. Schofield *et al.*, CASIMIR Rome Meeting participants, *Nature* **461**, 171 (2009).
4. G. Stacey, C. J. Hunt, *Regen. Med.* **1**, 139 (2006).
5. S. M. Fullerton, N. R. Anderson, G. Guzauskas, D. Freeman, K. Fryer-Edwards, *Sci. Transl. Med.* **2**, 15cm3 (2010).
6. B. Nelson, *Nature* **461**, 160 (2009).
7. The Hinxton Group, "Policies and Practices Governing Data and Materials Sharing and Intellectual Property in Stem Cell Science" (2011); see www.hinxtongroup.org.
8. Association for Molecular Pathology v. U.S. Patent & Trademark Office, No. 09 Civ. 4515 (Southern District of New York, 29 March 2010).
9. R. Skloot, *The Immortal Life of Henrietta Lacks* (Crown, New York, 2010).
10. M. M. Mello, L. E. Wolf, *N. Engl. J. Med.* **363**, 204 (2010).
11. I. S. Kohane, P. L. Taylor, *Sci. Transl. Med.* **2**, 37cm19 (2010).
12. Winickoff *et al.*, *Yale J. Health Policy Law Ethics* **9**, 52 (2009).
13. K. Bergman, G. D. Graff, *Nat. Biotechnol.* **25**, 419 (2007).

10.1126/science.1201382

Access to Stem Cells and Data: Persons, Property Rights, and Scientific Progress

Debra J. H. Mathews, Gregory D. Graff, Krishanu Saha and David E. Winickoff

Science **331** (6018), 725-727.
DOI: 10.1126/science.1201382