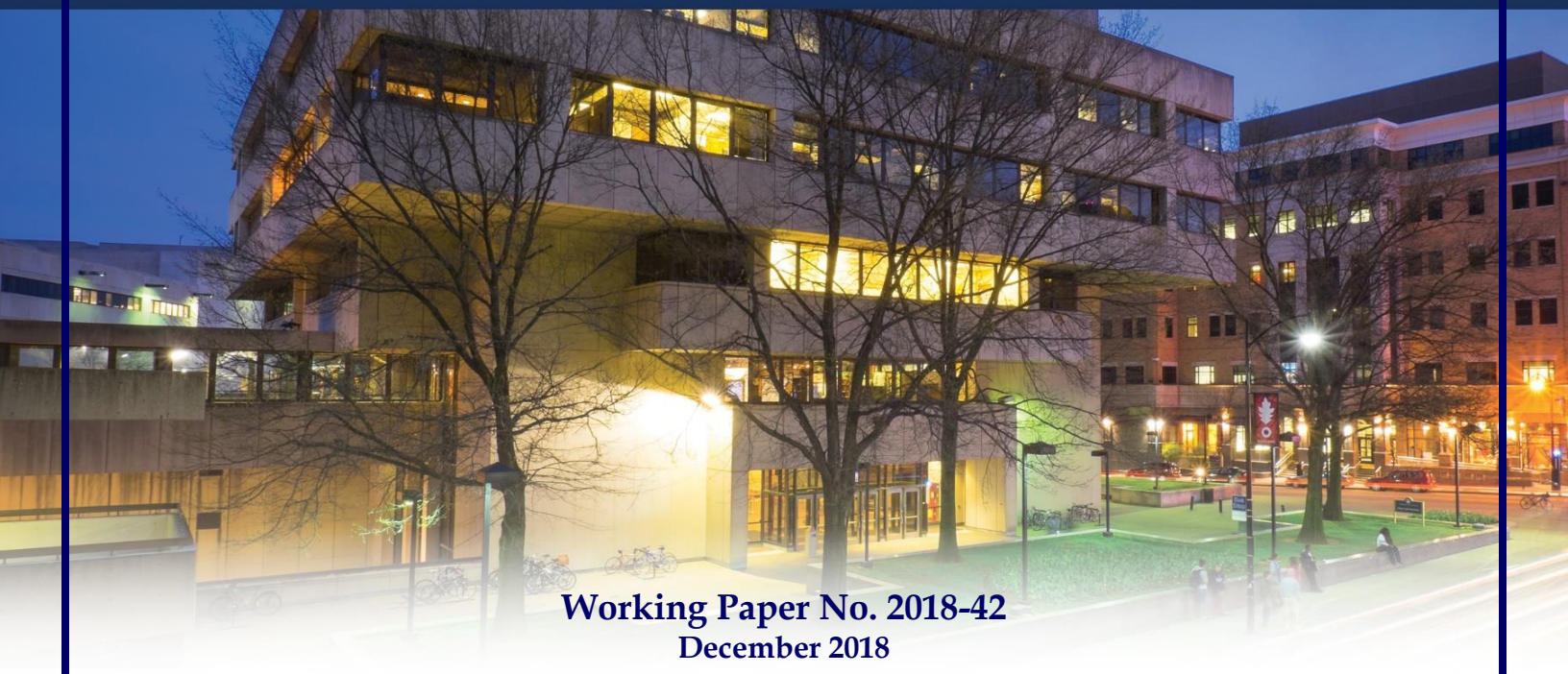
Legal Studies Research Paper Series

PittLaw
Faculty of Law



Working Paper No. 2018-42
December 2018

Biobanks as Knowledge Institutions

Michael J. Madison

Forthcoming *in* Minssen, Timo, Herrmann, Janne Rothmar, and Schovsbo, Jens eds. *Global Genes, Local Concerns: Legal, Ethical and Scientific Challenges in International Biobanking Cheltenham*, UK: Edward Elgar Publishing.

University of Pittsburgh School of Law
3900 Forbes Avenue
Pittsburgh, Pennsylvania 15260-6900

www.law.pitt.edu
412.648.7855
E-mail: madison@pitt.edu

Biobanks as Knowledge Institutions

Michael J. Madison*

Forthcoming *in* Minssen, Timo, Herrmann, Janne Rothmar, and Schovsbo, Jens eds. *Global Genes, Local Concerns: Legal, Ethical and Scientific Challenges in International Biobanking* Cheltenham, UK: Edward Elgar Publishing.

Draft of June 18 2018

Abstract

This chapter describes biobanks as institutions for collection, preservation, curation, and production of knowledge and information, in both material and immaterial forms. That characterization calls for research and comparative analysis of the broad diversity of specific biobanks, using a standardized research framework. Such a framework is identified and described here, as the knowledge commons framework. The chapter describes applications of the framework to biobanks to date and suggests directions for future research.

I. Introduction

Among the earliest references to biobanks in the scholarly literature is a paper titled "BIOBANK, a computerized data storage and processing system for the vascular flora of Iowa," published in 1979 in the Proceedings of the Iowa [US] Academy of Science.[1] While thoughtful and structured collections of biological specimens date back hundreds of years, to early "physic gardens" of plants cultivated for their medicinal properties,[2] that newer usage of "biobank" identifies a pair of contemporary themes. First, how are we to construct systems that effectively integrate material samples with accompanying immaterial information or data, in a usable way? Second, how are we to ensure that multiple systems of that sort are sustainable over time as matters of scientific practice and economic support, and that they are interoperable with each other? These are questions of both law and technology. Biobanks today are both critical

* Professor of Law and Faculty Director of the Innovation Practice Institute at the University of Pittsburgh School of Law, Pittsburgh, PA, USA. Email: madison@pitt.edu.

[1] Lawrence J. Eilers, 'BIOBANK, a computerized data storage and processing system for the vascular flora of Iowa.' (1979) Proceedings of the Iowa Academy of Science. 86: 15.

[2] American Medical Association, 'A History of Botanic Gardens.' (1915) Journal of the American Medical Association. 65(2): 170; Edward M. Holmes, 'Horticulture in Relation to Medicine.' (1906) Journal of the Royal Horticultural Society. 31: 42.

institutions in their own right and also critical illustrations and examples of broader governance questions concerning shared knowledge resources.

In that regard, this chapter makes and examines three related claims. First, biobanks are knowledge institutions, both in the sense that they store knowledge and information about the world in material and immaterial forms and also in the sense that they enable researchers to produce new knowledge, including useful applications of basic knowledge. Second, the broader questions of design and management that biobanks pose are resource governance questions rather than questions solely of law or of public policy. Third, despite the varied and diverse nature of biobanks today (indeed, precisely because of their diversity), their social and scientific importance dictates the need for a robust program of research of a comparative nature to identify shared features that contribute to their success (where they succeed) and features that likely contribute to problems or even failure. Both the importance of biobanks and associated governance challenges have only grown larger and more complex as biobanks meet the era of data science. In that regard the chapter points to examples of emerging scholarly literature that focuses on governance challenges of material and data in biobank contexts. The chapter suggest directions for future work, building on the emerging knowledge commons governance research framework.

II. A Series of Related Definitions

This section sets out some preliminary definitions.

Biobanks Defined

Biobanks today are defined generally (along with related terms "biorepository" and "Biological Resource Center," or BRC) to include structured collections of biological materials and associated data, stored for purposes of both present and future scientific research. Materials may be collected from humans, animals, cell and bacterial cultures, and plant and other environmental resources. Repositories of human biological materials and accompanying data generally are referred to as biobanks. Institutions that deal with plant and environmental samples generally are referred to as biorepositories or BRCs. Collections vary widely in size (from population-based collections to small collections for clinical or academic study), in purpose (from disease-specific or sample-based research), and in organizational status (non-profit to academic to commercial). Biobanks and biorepositories each raise some distinct governance concerns, and biobanks as such are often analyzed as an institutional mode in

order to isolate legal and public policy concerns that are linked to storage and use of human specimens and related data.

For purposes of this chapter, the similarities among these institutions take precedence over the distinctions. Here, biobanks include all repositories of biological information, stored in any manner and on any media, including information stored primarily as data rather than primarily as tissues or biological specimens or samples. In all cases, biobanks and biorepositories are resource pools, composed of some population of agents contributing specimens and related informational material; some population of agents (perhaps overlapping with the first) having the power to access and perhaps withdraw material; and a character defined by the fact that the social value of the pooled samples likely exceeds the value of each sample considered in isolation.

In sum, biobanks and biorepositories are resource sharing institutions, not solely exchanges or clearinghouses for access to individual samples. They are not merely tissue and specimen banks. Associated information and data may be hand-collected and curated, but increasingly they are apt to be collected and stored in high speed and large scale networked databases and related systems. Only a thin line may distinguish a biobank of material specimens or a network of biobanks focused on related types of specimens, on the one hand, from a large-scale bioinformatics dataset, on the other hand. If that distinction may be somewhat easier to describe today, then it may become more fluid in the near future. Today, biobanks are necessarily intertwined with data science.

The contemporary challenge for researchers is to grapple with the range and diversity of biobanks so that they can be analyzed in some systematic way. Despite such an inclusive beginning, or perhaps because of definitional and classification problems in the broad biobank field,[3] no single, integrated census of biobanks provides their total number.

Knowledge Defined

Biobanks exist to preserve knowledge and information in systematic ways for future generations (the preservation, curatorial, or stewardship function) and also to support generating new scientific and medical knowledge (the production function). Implicit but critical in that framing are definitions of knowledge and information. This chapter does not limit itself to formal or technical definitions of either one. "Information" is often regarded as "raw" or

[3] Gail E. Henderson, et al., 'Characterizing biobank organizations in the U.S.: results from a national survey.' (2013). Genome Medicine. 5: 3.

"unprocessed" descriptions about natural, physical, or social world. "Knowledge" is often regarded as "refined" information, "refined" because it has been treated by human analysis and converted into something that approaches shared scientific or cultural truth. Both knowledge and information are described primarily as immaterial "things," that is, conceptual objects of human engagement and thought.

This chapter adopts a different, more expansive, and more fluid framework for describing (rather than defining) knowledge and information, in order to capture the range of functions that biobanks perform. Knowledge and information are related expressions for the proposition that intellectual, creative, and scientific investigation of human, natural, and physical experience lead to the embodiment of those investigations in shared thoughts and ideas, in conceptual objects, and in material objects.[4] Those embodiments are identified and defined primarily in pragmatic terms, because the definitions align with the functionality of the resulting objects for practical uses, rather than in strict ontological terms. Scientific knowledge remains a source of truth in the sense that it represents the shared understanding of an expert community of practitioners. The development of digital technology during the later 20th century exposed the lack of utility in sharp distinctions between conceptual and material objects (such as ideas and things); between static forms of knowledge and dynamic processes of knowledge production, distribution, and use; and between micro and macro scales of knowledge.

Each of these perspectives contributes something valuable to our understanding of the world. In the context of biobanks, each individual specimen is a source and embodiment of information about itself; in the pool of specimens and data that composes the biobank, they are parts of a larger knowledge resource that is a source and embodiment of information about a larger collective of humans, plants, and so on. The specimen is a knowledge component; the biobank is a knowledge system, or a knowledge institution.

Data Science Defined

Biobanks do not exist solely with respect to the material samples collected. The information or data associated with each sample is itself a critical and related knowledge resource, and it should be considered for purposes of analysis jointly with the material to which it relates. The scale of the information resources collected in biobanks, particularly population-based biobanks for

[4] Henry Plotkin, *Darwin Machines and the Nature of Knowledge* (Cambridge: Harvard University Press 1997).

genetic research, suggests the importance of combining modes of analysis relative to specimen collection itself with modes of analysis as understood in the related field of data science, sometimes referred to as data-driven science or data-intensive science. For present purposes, references to data science refer to statistical and other analysis of data sets typically collected, stored, and processed in digital forms. Data models, the analytic frameworks used to define what data is collected and stored and the formats in which it is stored and shared, are as critical to data science as the data themelves. The popular phrase "Big Data" captures a part of what is distinctive about data science in modern biobanking and other scientific contexts, in the sense that the phrase speaks to the scale and speed at which information is collected in extra-large data sets and, in principle, made available to researchers. Data science extends beyond such large collections to include the systematic study of data in any form.[5] The point here is that biobanking governance is necessarily linked to data governance.

Governance Defined

The challenges of biobank classification illustrate the corresponding challenge of analyzing biobank governance. Governance is the object of analysis rather than (or, to be clear, in addition to) law as such. Governance refers to the multiple relationships among various institutional actors, from individuals to governments, and a specified resource or set of resources. Those relationships make consist of regulation, or discipline, or other modes of control or alignment. In most instances of interest here, governance addresses one or more social dilemmas associated with management of the resource, that is, with the fact that the resource, taken in isolation, does not "manage itself" as an autonomous "thing." Those dilemmas arise from the fact of potential conflicts between multiple objectives represented by different actors and institutions with interests in the resource. Knowledge sharing occurs on several dimensions simultaneously: temporal, generational, geographic, disciplinary, and beyond. Not every actor and not every objective can be fully accommodated at all times. When governance succeeds, it mediates and enables knowledge institutions, including biobanks, to thrive. Sometimes, governance does not succeed; sometimes, governance is designed and interposed intentionally but fails to accommodate competing objectives of different actors. Sometimes, governance emerges and evolves, more or less organically, and the processes of adaptation produce less and more thriving.

---

[5] E.g., Christine L. Borgman, *Big Data, Little Data, No Data: Scholarship in the Networked World* (Cambridge, MA: MIT Press 2015); Matthew J. Salganik, *Bit By Bit: Social Research in the Digital Age* (Princeton, NJ: Princeton University Press 2017).

Like virtually all institutions for generating, storing, or otherwise managing knowledge resources, biobanks operate at the intersection of multiple, overlapping regulatory or disciplinary frameworks, which themselves operate at multiple levels. Formal positive law and regulation is only one source of institutional order. Even that formal law may come from multiple legal domains (intellectual property law, privacy law, and antitrust or competition law, for example) and from multiple institutions and sources (legislatures, courts, administrative or regulatory bodies). Informal frameworks (communities of practice, social norms, ideologies) may play important roles in disciplining both individual and collective behavior. Formal and informal systems each may be expressed via well-defined institutions, such as firms, universities, and other legal entities, via markets, and via individuals who internalize relevant expectations. Both formal and informal disciplinary rules, norms, and expectations may be relevant with respect to scientific researchers who deposit, withdraw, and access and use biobank samples and information; with respect to information scientists who design and maintain the biobank itself; and with respect to computer scientists and programmers who maintain the information technology and network infrastructure that is typically associated with a biobank.

Law, policy, design, culture, economics, and ethics are combined, by necessity. What may be called "origin stories" of biobanks often play a central role in determining the narrative framework within which a given biobank exists, thus influencing the choice of relevant legal or disciplinary frameworks for application and analysis. My choice above to invoke the historical example of the small "physic garden" as an early example of a "biobank" itself suggests a historical narrative that distinguishes my summary, in part, from a summary that might have begun, alternatively, with large modern gene banks.

Part of the concern for any governance analysis, in short, is scale. With respect to biobank resources, what sorts of collectives matter for purposes of preserving, accessing, and using the knowledge? Part of the concern is temporal. How should biobanks blend current interests in knowledge with future interests? Part of the concern is hierarchical or sequential. At times, it may be helpful to characterize biobanks as forms of knowledge infrastructure, in the sense that biobank resources offer a broadly distributed and widely shared knowledge resource as an input into a diverse array of potential knowledge outputs. When and how should biobanks support or enable other research? In what respects, if any, should biobanks be concerned with supporting or enabling not only other research but also applications of biobank resources for direct human benefit?

Resolution of conflicts embedded in each of these concerns may not always align. In the specific instance, therefore, the research lens broadens: What helps a biobank thrive? What undermines a biobank? These are not purely legal or public policy questions; as this chapter has emphasized, they are, in a broad sense, governance questions. The chapter next turns to exploring how those questions may be researched in a systematic and comparative way. What commonalities are relevant, and what case-specific circumstances matter?

III. The Knowledge Commons Research Framework

Connecting the conceptual and practical definitions of the last section – biobanks, knowledge, data science, and governance – yields the intermediate proposition that biobanks are knowledge commons institutions. The term 'commons' should not mislead. As used here, commons refers to institutional arrangements for managing shared access to a pooled or collected resource. In adopting the knowledge commons characterization for biobanks, the chapter follows usage by other scholars.[6] The purpose of the identifying biobanks as knowledge commons is that doing so sets a foundation for the claim that understanding biobanks governance is a matter for empirical research. That research should be systematic and comparative across multiple biobank cases studies, to identify commonalities among the diversity of biobank forms.

This section describes one such approach to the relevant research: comparative institutional analysis using the knowledge commons research framework. The framework supplies an analytic basis for systematic comparative analysis of knowledge sharing institutions of all kinds. The framework is set out in brief in this section. Biobanks are one leading illustration. The question for researchers is not "is a biobank a commons?"; rather, the question is, "are biobanks governed as commons?" The next section summarizes relevant research to date on biobanks as knowledge commons.

Background

[6] Andrea Boggio, 'Population Biobanks' Governance: A Case Study of Knowledge Commons,' in Katherine J. Strandburg, Brett M. Frischmann, and Michael J. Madison (eds.), *Governing Medical Knowledge Commons* (Cambridge: Cambridge University Press 2017); Barbara J. Evans, 'Genomic Data Commons,' in Katherine J. Strandburg, Brett M. Frischmann, and Michael J. Madison (eds.), *Governing Medical Knowledge Commons* (Cambridge: Cambridge University Press 2017).

Following Elinor Ostrom's groundbreaking work on institutions for resource management in the natural resource and environmental contexts,[7] Frischmann, Madison, and Strandburg describe knowledge commons generally as governance solutions for shared resources subject to social dilemmas.[8] In the first place a resource is identified or created; next, use of that resource is purposefully shared by some population of producers and/or consumers. In the second place, a number of possible social dilemmas exist that are associated with the shared production and/or use of that resource, deriving generally from interests in social collaboration and cooperation. Commons address one or more of those dilemmas. Commons are forms of governance, or management, of shared resources.

With respect to natural resources, Ostrom and her colleagues and collaborators demonstrated the viability of a range of sustainable, durable commons governance strategies that preserve the resource over time, implemented by local groups and communities using well-structured convention and custom. Commons are collectively managed governance systems, often marked by the absence of formal, market-based property law systems. In their re-purposing of Ostrom's work, Frischmann, Madison, and Strandburg set out a research framework to investigate the viability of equivalent commons governance strategies with respect to knowledge, scientific, and cultural resources.[9] Knowledge commons governance may differ from natural resource commons governance in key respects, beginning with the fact that knowledge commons resources, unlike forests or fisheries, are naturally nonrival or nondepletable and therefore naturally or inherently shareable. The case for sustainable commons governance is neither inherently stronger nor weaker as a result. Instead, cases of knowledge commons must be researched from the beginning, rather than analyzed solely by analogy to natural resource commons. The point of the research framework presented here, like any research framework, is to permit research and data collection to proceed under a common set of assumptions and questions, even if specific research methods and disciplinary foundations may vary from researcher to researcher or field to field. The framework is neither theory nor model. Strong theorizing and modelling may follow the research but only light and tentative theorizing, if any, should precede it.

---

[7] Elinor Ostrom, *Understanding Institutional Diversity* (Princeton, NJ: Princeton University Press 2005); Elinor Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge: Cambridge University Press 1990).

[8] Brett M. Frischmann, Michael J. Madison, and Katherine J. Strandburg (eds.), *Governing Knowledge Commons* (Oxford: Oxford University Press 2014).

[9] *Id.*

The framework as described is borrowed from a 2014 book by Frischmann, Madison, and Strandburg titled *Governing Knowledge Commons*.[10] That book presents the framework and applies it to a set of case studies of institutions defined in part by knowledge sharing practices with respect to one or more knowledge resources. While the framework is designed for application at the institutional level, which are referred to as commons, the intuitions and preliminary investigation that animated its development are applicable more broadly. Knowledge resources come in many forms; governance comes in many forms. Information and knowledge are principally immaterial, intangible resources, but they may be embodied in material forms, in flows of knowledge as well as in forms, and in labor and skill and time as well as in embodied creation and other materials. The balance of this section gives a fuller account of the framework.

The knowledge commons framework builds on a series of related intuitions. Commons governance means knowledge and information management characterized by domains of managed openness and sharing of relevant resources. The first intuition is that commons governance is in broad use in day-to-day practice in a variety of domains and across a variety of scales. Documenting evidence to justify that intuition is a primary goal of the framework.

The second intuition is that such structured openness in the management of both natural and cultural resources is likely to lead to socially beneficial and/or socially productive outcomes. Salient among the class of cases where commons governance is successful and sustainable are contexts where social interest in positive spillovers from bilateral market transactions is high. Commons may sustain the production of spillovers when the market otherwise may not. Describing the commons framework in terms of spillovers from bilateral market transactions runs a substantial risk of characterizing an information or knowledge context exclusively in "scarce resource" terms rather than in "abundant resource" terms. Care must be taken in applying the framework to understand the nature of the resources in question.

The final intuition is that a standard framework for identifying and assessing commons across a variety of domains can support the development of more sophisticated tools for realizing the potential for commons solutions in new institutional settings. It can also help to distinguish commons solutions from other solutions, such as an approach grounded solely in formal IP law, or in formal privacy law or contract law, might be preferred. Applying the knowledge

[10] *Id.*

commons research framework is an exercise in analyzing colloquial commons institutions, such as "scientific research" taken in the aggregate, in a nuanced way and in application to concrete examples.

Examining commons in knowledge and information contexts builds on the framework pioneered by Ostrom and her colleagues known as the Institutional Analysis and Development (IAD) framework, but it adds some important modifications. The IAD framework has been used principally to structure analysis of solutions to collective action problems in natural resource contexts (so-called action arenas, or action situations) such as forests, fisheries, and irrigation systems.

One, the knowledge commons framework differs from the IAD framework in certain key respects. Unlike the IAD framework, it does not assume the agency of rational, choice-selecting, self-interested individuals, as the IAD framework tends to do. The social dilemmas subject to analysis in the knowledge commons context arise not only because of conflicts among rational, self-directed actors seeking to maximize their own benefits from resource access and use. Relevant social dilemmas arise because of conflicts among actors with diverse backgrounds, interests, and motivations. The knowledge commons framework accepts the role of historical contingency and of both inward-and outward-directed (selfless or other-oriented) agents in the evolution of collective or commons institutions.

Two, unlike natural resource commons, which largely take the existence of their resources for granted: fish, trees, water, and the like, knowledge commons identify resource design and creation as variables to be described and analyzed. As intellectual resources (that is, as forms of knowledge and information), patents, copyrights, and underlying inventions, creations, and data, and related material objects, are shaped by a variety of institutional forces rather than by nature.

Three, critically, the knowledge commons framework does not assume that the relevant resources are rival and depletable. The knowledge commons framework generally assumes precisely the contrary: that intangible information and knowledge resources are nonrival, nondepletable public goods. They may be closely linked, however, to depletable material resources. The dilemma to be solved is not primarily a classic "tragic commons" overconsumption problem. Instead, it is more likely (in part) an underproduction problem (how to produce the knowledge resource?) and (in part) a coordination problem. How may different actors coordinate their activities and interests in order to make appropriate and productive use of a shared resource? That is not an exclusive

list, however. In applying the framework to any particular case, care must be given to describing the authentic character of the social dilemmas present.

Finally, Ostrom's work largely ignored or discounted the role of the state in commons governance. The knowledge commons framework necessarily accepts the possibility that the state may have one or more key roles to play in managing shared knowledge resources.

The Framework

Against that background, the knowledge commons framework proposes to undertake comparative institutional analysis by evaluating cases of commons resources via a series of questions, or clusters, to be applied in each instance.

Each case study investigation begins with a general description of the history and character of the problem that is being addressed by governance in the specific case or context. This may be an explanation that is internal to the governed institution(s) (problems and explanations may emerge from stories told by participants, either today or historically, or both), or an explanation that is external to the governed institution (such as the public goods account of the rise of IP law).

A researcher should ask whether the relevant resource or case is characterized from the outset by patent rights or other proprietary rights, as in the case of a patent pool, or by a legal regime of formal or informal openness, as in the case of public domain data or information collected in a government archive. That characterization influences the description of the social dilemmas that the governance institution addresses. A particular regime might involve securing the benefits of sharing data and information, or sharing rights in information, or sharing both. The character of the commons solution might involve encouraging the production of new resources, or coordinating holders of different property interests or holders of different public domain knowledge resources, for example.

Answering those initial questions sets a baseline against which a commons governance regime has been constructed. Within that regime, one next asks definitional questions. What are the relevant resources, taking into account both intangible and tangible resources and their individual or social character? What are the relationships among these resources, the baseline, and any relevant legal regime (for example, what a scientist considers to be an invention, what patent law considers to be an invention, and the boundaries of the patent itself are three related but distinct things)? What are the boundaries

and constitution (membership) of the collective, community, or communities that manage access to and use of those resources? How is membership acquired (this may be informal, formal, or a blend of the two), and how is membership governed? What is good behavior within the group, and what is bad behavior? Who polices that boundary, and how?

Next are questions concerning explicit and implicit goals and objectives of commons governance, if any such goals and objectives exist. It is possible that commons governance regimes emerge from historical contingency rather than via planning. Is there an identified resource development or management dilemma that commons governance is intended to address, and what commons strategies are used to address that dilemma?

How "open" are the knowledge and information resources and the community of participants that create, use, and manage them? *Governing Knowledge Commons* argues that commons governance regimes involve significant measures of resource and community sharing and openness. The details of this openness should be specified in both absolute and relative terms, along with their contributions to the effectiveness of commons. Some commons and commons resources have precise and fixed definitions of both resources and community membership. Either resources or membership or both may be more fluid, with boundaries defined by flexible standards rather than by rules.

A large and critical cluster of questions concerns the dynamics of commons governance, or what Ostrom refers to as the "rules-in-use" of commons: the interactions of commons participants and resources. Included in this cluster of questions are: (1) details of stories of the origins, histories, and operations of commons; (2) formal and informal (norm-based) rules and practices regarding distribution and coordination of commons resources among participants, including rules for appropriation and replenishment of commons resources; (3) the institutional setting(s), including the character of the regime's possibly being "nested" in larger scale institutions and being dependent on other, adjacent institutions; (4) relevant legal regimes, including but not limited to property law; (5) the structure of interactions between commons resources and participants and institutions adjacent to and outside the regime; and (6) dispute resolution and other disciplinary mechanisms by which commons rules, norms, and participants are policed.

At this point the attributes of the system have been specified, and it becomes possible to identify and assess outcomes as the system operates in practice. In Ostrom's IAD framework, outcomes are typically assessed in terms of the resources themselves. Has a fishery been managed in a way that sustains

fish stocks over time? Do commons participants, such as the members of a fishing community, earn returns in the commons context that match or exceed returns from participation in an alternative governance context, such as private market transactions? In knowledge commons settings, resource-based outcome measures may be difficult to identify and assess. Sustaining the products of the resource, individually or in combination, may be the point. A patent pool may serve as knowledge commons governance for a particular industry, but the success of the pool is measured not only by the fact that the pool itself survives. Equally important, the value of the pool may be shown by the production of valuable complex products that could not be produced but for the pooling arrangement. Outcomes take different forms. It may be the case that social patterns of participant interaction constitute relevant outcomes as well as relevant inputs. What may matter is that the community itself thrives, in addition to the knowledge resource that it manages. Levels and types of interaction and combination matter. Participant interaction in the context of a shared resource pool or group may give rise to (or preserve, or modify) an industrial field or a technical discipline. In that specific case, such spillovers may be treated as relevant outcomes.

Having identified relevant outcomes, it becomes possible to look back at the dilemmas that defined commons governance in the first place. Has the regime solved those problems, and if not, then what gaps remain? How do the outcomes produced by commons governance differ from outcomes that might have been available if alternative governance had been employed? Has commons governance created costs or risks that should give policy makers and/or institution designers pause? Costs of administration might be needlessly high; costs of participation might be high. A collection of industrial firms that pool related patents in order to produce complex products may engage in anticompetitive, collusive behavior. Commons governance may facilitate innovation. It may also facilitate stagnation.

In sum, the knowledge commons framework provides a useful method of blending standardization and local adaptation in a systematic way in researching governance of knowledge institutions, such as biobanks, using comparative institutional analysis. From that foundation, the framework then provides the means to undertake more focused queries in order to define the relevant opportunity set for legal/regulatory analysis (that is, specifically with regard to law or policy approaches within the overall governance context). Finally, and most optimistically, results from using the framework may permit specifying a useful set of guidelines and recommendations for further development and design of knowledge institutions themselves. It is important to recognize the study of biobank governance as an opportunity to innovate regarding modes of

governance beyond formal, public or positive law, beyond publicly-enacted regulation, and even beyond formalized "public/private partnerships," as the default or primary modes of encouraging and sustaining these institutions.

Applications: Biobanking Governance as Knowledge Commons

The claim presented in this chapter, that biobank governance should be analyzed as knowledge commons governance, is not novel. Since the initial publication of the knowledge commons research framework in 2010,[11] several researchers have adopted that framework in order to study biobanks and related institutions for collecting and managing biological specimens and associated data. This section reviews the relevant literature briefly, to show in part the viability of this approach and to show in part (by contrast) how other biobank governance literature might be adapted and brought within its scope. This is not an exhaustive literature review or bibliography; it advances the claim by illustrating it.

The most direct applications of the knowledge commons framework appear in case studies published in the edited collection by Strandburg, Frischmann, and Madison titled *Governing Medical Knowledge Commons*.[12] Four contributors to that volume have written case studies that address biobanks or features of biobanks using the knowledge commons governance framework. Peter Lee contributed "Centralization, Fragmentation, and Replication in the Genomic Data Commons," a treatment of methods and technologies used to manage information already contributed to large-scale genomic databases.[13] Barbara Evans wrote a detailed case study of management of genomic databases composed of the results of consumer-driven genomic testing, "Genomic Data Commons."[14] Andrea Boggio offered "Population Biobanks' Governance: a Case Study of Knowledge Commons," whose title directly describes its content.[15] Jorge Contreras contributed "Leviathan in the Commons: Biomedical Data and the State," which directs attention to the key role that state support or coordination

[11] Michael J. Madison, Brett M. Frischmann, and Katherine J. Strandburg, 'Constructing Commons in the Cultural Environment.' (2010) Cornell Law Review. 95:657.
[12] Strandburg, Katherine J., Brett M. Frischmann, and Michael J. Madison (eds.), *Governing Medical Knowledge Commons* (Cambridge: Cambridge University Press 2017).
[13] Peter Lee, 'Centralization, Fragmentation, and Replication in the Genomic Data Commons,' in Katherine J. Strandburg, Brett M. Frischmann, and Michael J. Madison (eds.), *Governing Medical Knowledge Commons* (Cambridge: Cambridge University Press 2017).
[14] Evans, *supra* note 6.
[15] Boggio, *supra* note 6.

may play in commons governance contexts.[16] Jorge Contreras has also published a number of other studies of governance of genomic commons institutions, typically within the knowledge commons framework.[17]

Particular attention should be paid to the impressive recent work, *Governing Digitally Integrated Genetic Resources, Data, and Literature: Global Intellectual Property Strategies for a Redesigned Microbial Research Commons*, by Jerome Reichman, Paul Uhlir, and Tom Dedeurwaerdere.[18] The authors provide a comprehensive account of a central set of biobanking institutions as knowledge commons governance, addressed to the material and data-related results of microbial research. The work is especially valuable in the biobank context because its attention to plant genetic information complements the commons-based study of institutions to manage human biological material. A useful volume that anticipates many of the themes of biobank governance as knowledge commons is *Gene Patents and Collaborative Licensing Models: Patent Pools, Clearinghouses, Open Source Models and Liability*, edited by Geertrui van Overwalle.[19] That collection offers a number of useful perspectives on knowledge and information sharing with respect to biological data, though it does not consolidate them in a systematic way. Recent, important work on the challenges of Big Data and data science, with attention given to biobanking and to the utility of Ostrom's work on commons governance in understanding data science for researchers in the 21st century, is *Big Data, Little Data, No Data: Scholarship in the Networked World*, by Christine Borgman.[20]

The foregoing research may be contrasted with research on biobanks and biobank governance that adopts specific frameworks for analysis rather than a wholesale comparative institutional framework. The work may be broken down generally into several categories, each of which has yielded important valuable contributions but which has not produced a comprehensive, systematic vision of

---

[16] Jorge Contreras, 'Leviathan in the Commons: Biomedical Data and the State,' in Katherine J. Strandburg, Brett M. Frischmann, and Michael J. Madison (eds.), *Governing Medical Knowledge Commons* (Cambridge: Cambridge University Press 2014).

[17] E.g., Jorge Contreras, 'Constructing the Genome Commons,' in Brett M. Frischmann, Michael J. Madison, and Katherine J. Strandburg (eds.), *Governing Knowledge Commons* (Oxford: Oxford University Press 2014).

[18] Jerome H. Reichman, Paul F. Uhlir, and Tom Dedeurwaerdere, *Governing Digitally Integrated Genetic Resources, Data, and Literature: Global Intellectual Property Strategies for a Redesigned Microbial Research Commons* (Cambridge: Cambridge University Press 2016).

[19] van Overwalle, Geertrui (ed.), *Gene Patents and Collaborative Licensing Models: Patent Pools, Clearinghouses, Open Source Models and Liability Regimes* (Cambridge: Cambridge University Press 2009).

[20] Christine Borgman, *Big Data, Little Data, No Data: Scholarship in the Networked World* (Cambridge, MA: MIT Press 2015).

the domain. The point to emphasize is that none of these modes has yet been able to integrate multiple perspectives on the social dilemmas associated with biobanks in order to produce a complete and integrated legal and public policy analysis. The work neither describes the biobank landscape in full and nor offers an overall guide to interpreting strengths and weaknesses of specific existing biobanks and to developing new ones.

One mode of research examines biobanks and biobank resources primarily as questions of ownership, thus raising important questions about public accountability and access, proprietary right and incentives, and moral and ethical claims. Ownership interests and claims may be assessed both with respect to patients and consumers and also with respect to scientific researchers and the owners or managers of the relevant biobank enterprise. Claims may be analyzed with respect to material specimens and also with respect to information derived from them. In specific legal terms, the questions include chattel property in the material specimens and patent, copyright, and data ownership with respect to the information resources. The vocabulary of commons may appear, though often in opposition to a term and concept borrowed from the law of property, "anticommons," which speaks to excessively fragmented property interests in a complex market context.[21]

A second mode of research examines biobanks from the standpoint of personal autonomy, privacy and security, and ethics, primarily with regard to patients, consumers, and other subjects of clinical trials but also with respect to scientific researchers and even with respect to the ethical status of biological specimens and information. These are often not framed as property claims. Rather, in legal and public policy terms, questions may be posed in terms of transparency and consent. This literature is unlikely to highlight commons governance or knowledge sharing as a priority. Rather, the research question is typically how to address or accommodate individual and personal privacy interests within an institutional environment that is set up to pool and share information.[22]

---

[21] Examples include the contributions in Giovanni Pascuzzi, Umberto Izzo, and Matteo Macilotti, (eds.), *Comparative Issues in the Governance of Research Biobanks: Property, Privacy, Intellectual Property, and the Role of Technology* (Springer 2013); Peter Lee, 'Toward a Distributive Commons in Patent Law' (2009) *Wisconsin Law Review.* 2009: 917.

[22] Examples include Kris Dierickx and Pascal Borry (eds.), *New Challenges for Biobanks: Ethics, Law and Governance* (Intersentia 2009); Henry T. Greely, 'The Uneasy Ethical and Legal Underpinnings of Large-Scale Genomic Biobanks' (2007) Annual Review Genomics and Human Genetics. 8:343; Mark A. Rothstein and B. M. Knoppers (eds.), 'Regulation of Biobanks' (2005) Journal of Law, Medicine and Ethics. 33(1): 1; Mark Stranger and Jane Kaye, *Principles and Practice in Biobank Governance* (Routledge 2009).

A third mode of research focuses primarily on social norms and technical resources concerning information and knowledge production and sharing within the scientific community. The literature tends to situate proprietary claims and autonomy claims within the broader information collection and production environment, highlighting the normative value of scientific collaboration and the production of new scientific and medical knowledge. Commons governance may be discussed in this context, though often as a normative claim rather than as an analytic framework.[23]

A final mode of analysis of biobanking governance applies traditional styles of governance thinking, focusing on positive law within national legal systems, to biobanks and to biobank enterprises.[24]

The knowledge commons framework, developed through case studies over time, offers the prospect of integrating these results and analyses in a systematic way.

IV. Research and Governance Challenges: Lessons to Date, and the Future

This section sets out critical areas of future inquiry with respect to biobank research, both within the knowledge commons research framework and for possible use by policy analysts and institutional designers. To date, case studies of knowledge commons governance for biobanks are too few in number to conclude with certainty that any key legal reforms are necessary or that developers or managers of biobanks should follow any mandatory guidance. The research so far does suggest some key areas of focus in both respects.[25] Those include the following:

First, it is perhaps most important to understand the various goals and purposes associated with a given knowledge institution, such as a biobank, and how those goals and purposes have evolved and how they relate to one another.

---

[23] A recent example is Helen Yu, 'Redefining responsible research and innovation for the advancement of biobanking and biomedical research.' (2016) Journal of Law and the Biosciences. 3(3): 1.

[24] Examples include Herbert Gottwies and Alan Petersen (eds.), *Biobanks: Governance in Comparative Perspective* (New York: Routledge 2008); Jane Kaye, Susan M. C. Gibbons, Catherine Heeney, Michael Parker, and Andrew Smart, *Governing Biobanks: Understanding the Interplay between Law and Practice* (Hart Publishing 2012).

[25] The following is adapted from Katherine J. Strandburg, Brett M. Frischmann, and Michael J. Madison, 'Governing Knowledge Commons: An Appraisal,' in Katherine J. Strandburg, Brett M. Frischmann, and Michael J. Madison (eds.), *Governing Medical Knowledge Commons* (Cambridge: Cambridge University Press 2017).

Researchers may be tempted to make assumptions about those goals and purposes and to move directly into questions regarding property rights or privacy and ethics or the role of the state. That is almost certainly an error. The knowledge commons framework suggests that explicit goals need not be taken at face value; it is worth exploring implicit conflicts and dilemmas that are addressed by institutional structures.

Second, the relevant actors may be more numerous and their roles more complex than they first appear. In turn, that means that the relevant goals of the biobank and their interdependencies may also be complex. Interdependencies may yield additional goals and opportunities, or may yield barriers and limitations. Contributors, users, other researchers, managers, patients, subjects, host institutions, funders, external reviewers, and their respective institutional linkages, past, present, and anticipated, should be described. Formal, informal, and normative social and cultural structures should be mapped, changes over time should be detailed, and the different potential of planned and emergent behavior should be considered. Social hierarchies may matter more or less. Shared or distinctive cultural values may play important roles in organizational function. Intrinsic motivations for action may play important roles along with extrinsic ones. Detailed mappings may be necessary to ensure that the full portrait of the institution is described. For example, future access and use may be as important or more important than present access and use. Research use and clinical use may more important than stewardship and preservation or heritage concerns, or may co-exist, though with different values.

Third, the character of the resources may be complex as well. This chapter has highlighted the interdependencies of biobank resource pools that consist primarily of physical or material specimens and pools that consist primarily of intangible information or knowledge resources. Those conceptual interdependencies may correlate with interdependencies regarding technical systems to be developed to collect, store, manage, and access them. They may be constituted as separate but related resource pools, or as integrated resource pools. They may be characterized with differing levels of item-specific identifiability and access/use parameters, and with interdependencies regarding governance strategies concerning access and use.

Many of those interdependencies arise from the multi-faceted aspects of the problems of designing and managing a biobank. A key element of knowledge commons resources is their constructed character. That means that law, policy, and practice may play important roles in determining the form and identity of the resource itself and/or of resource units within it. Law, in this sense, may be definitional as well as regulatory in determining what "counts," in multiple

senses, for purposes of inclusion, extraction, and use of resources within the biobank. That fact may or may not distinguish the character of a resource pool that consists primarily of material specimens from a resource pool that consists primarily of research data. "Things" as resources are designed by human engagement (law, policy, science), along with systems to govern them and those who use them.[26]

Fourth, the knowledge commons research framework, like Ostrom's work on natural resource commons, emphasizes close study of the interactions among relevant actors with respect to the resource, under conditions specified by formal and informal rules. These interactions take place in "action arenas," which may be physical, virtual, or conceptual. Any given knowledge institution may consist of and support multiple action arenas. Action arenas may be centralized or consolidated, and they may be distributed geographically or virtually. Understanding the action arenas of a given case may require research of an almost ethnographic character. At this point, the empirical character of the knowledge commons perspective comes into sharpest focus, because the behavior of various actors in actual practice cannot be assumed to correspond to conceptual frameworks associated with law or any other research discipline, such as economics. As we know from ordinary experience but as we often do not believe as researchers, people do not necessarily act in rational self-interest.

Fifth, assessing the success or failure or other standard of viability of a knowledge commons institution is critical, but methods and standards for doing so in the knowledge and information context are badly underdeveloped. This is a part of the field of commons research particularly, and of comparative institutional analysis in general, that needs the most effort.

V. Conclusion

The primary points of this chapter are the following:

Whether at local, national, regional, or global levels, biobanks are knowledge institutions. They collect, curate, and steward biological materials and associated knowledge and information for the benefit of future generations as well as for present scientific researchers. They house knowledge resources, and they provide important knowledge infrastructure for the production of new knowledge.

---

[26] Michael J. Madison, 'IP Things as Boundary Objects: The Case of the Copyright Work,' (2017) LAWS. 6(3): 13, doi:10.3390/laws6030013; Michael J. Madison, 'Law as Design: Objects, Concepts, and Digital Things,' (2005) Case Western Reserve Law Review. 56: 381.

Biobank resources consist of more than the physical specimens they collect. Increasingly, those specimens are accompanied by critical information and data and/or are media that express critical information and data. Biobank governance is intertwined with data and information governance, and with data science. Law, policy, design, culture, economics, and ethics are combined, by necessity.

Biobank governance consists principally of various modes of knowledge and information sharing with respect to the resource that constitutes any particular biobank. Knowledge sharing occurs on multiple dimensions simultaneously: temporal, generational, geographic, disciplinary, and beyond. As a result, governance is required to enable biobanks to thrive amid multiple possible social dilemmas, meaning potential conflicts between multiple goals and interests represented in the resource.

Systematic study of knowledge sharing institutions, such as biobanks, is most effectively conducted using a standard analytic framework for comparative institutional analysis that is tailored to the dynamics of knowledge. Here, the chapter proposes use of the knowledge commons research framework. Examples and illustrations are given of the framework as applied to biobanks. Preliminary research results are described, and recommendations for future research are suggested.