

UIC John Marshall Journal of Information Technology & Privacy Law

Volume 25
Issue 4 *Journal of Computer & Information Law*
- Symposium

Article 9

2008

Information Governance, 25 J. Marshall J. Computer & Info. L. 673 (2008)

Michael J. Madison

Follow this and additional works at: <https://repository.law.uic.edu/jitpl>



Part of the [Computer Law Commons](#), [Internet Law Commons](#), and the [Science and Technology Law Commons](#)

Recommended Citation

Michael J. Madison, Information Governance, 25 J. Marshall J. Computer & Info. L. 673 (2008)

<https://repository.law.uic.edu/jitpl/vol25/iss4/9>

This Symposium is brought to you for free and open access by UIC Law Open Access Repository. It has been accepted for inclusion in UIC John Marshall Journal of Information Technology & Privacy Law by an authorized administrator of UIC Law Open Access Repository. For more information, please contact repository@jmls.edu.

INFORMATION GOVERNANCE

MICHAEL J. MADISON*

I want to talk about, in general terms, why cultural forms and information forms of various sorts exist. In more specific terms, I want to talk about information law, and policy.

First, I want to talk about what information is and what it means. I want to spend some time imagining what the world of law and policy would look like if we did not rely on the premise that there is information out there in the world that simply exists. My basic theoretical point is this: what we think of as information in its various forms, whether we are talking about content, knowledge, data, or the various forms that we apply to intellectual property, are functions of law and other processes as much as they are inputs to law and those processes. Therefore, in terms of trying to figure out some of the complicated law and public policy problems that we are dealing with today and the overlaps among different conventional doctrinal categories, we need to spend some time figuring out where information comes from and what the role of law is in producing information in different contexts. How are information and law managed? How do they manage each other? How do they jointly participate in managing other activities? Once we have a better understanding of those processes, we can figure out where and how regulatory or legal intervention is appropriate and likely to be effective.

Next, a quick aside: why do I bother going to the trouble of working through the sort of synthetic question, "What is information and where does it come from?" The answer is simply that I want to see what information buys me. I do not have a proposal that I can sell that is concretely better than the existing scheme of doctrinal categories that we have been relying on for a long time. But I do think that we have been relying on those categories for a long time, whether we are relying on privacy, security, free speech, copyright, patent, or other related things. We find ourselves in quite a pickle in terms of the various policy problems that we have and the overlaps among different phenomena. I

* Michael Madison is the associate dean for research and associate professor of law at the University of Pittsburgh, School of Law where he teaches and writes about intellectual property and information technology. His J.D. is from Stanford Law School and he has a B.A. from Yale University.

really want to open the door to a conversation about thinking about the world in a different way. I want to see if, by thinking about the world in a different way, we can come up with some intriguing, and possibly useful, ideas.

What would the world look like if we thought about where information comes from and the dynamic relationship between information, law, and other phenomena? What kinds of issues, what kinds of concepts would we confront? And what do I mean when I say that information is a function of law and other processes as much as an input into it?

I will talk through these questions to offer a sense of their nuances, rather than try to wrap up an answer. Step one is to start with the foundational problem that various law and policy structures of all types need to be interrogated in terms of their legitimacy. So whatever information form or information-related practice we are talking about, whether it is law, culture, society, or technology, has to be scrutinized in terms of how it is recognized socially. It also has to be scrutinized in terms of its sustainability or durability. Is the information going to be temporary or long term? In addition to thinking about particular information, we scrutinize both the forms -the static forms, the conceptual forms, and the physical forms- and the practices - the ways people use those forms to gather themselves into groups, networks, and institutions.

To approach these questions I start with the concept that I call governance and apply it to the regulation of information and the production of information. All of the subsidiary questions, for example, that intellectual property owners, consumers, and scholars call the production, distribution, or exploitation of copyrighted information are manifestations and species of what I call governance.

What do I mean by governance? I am borrowing a term that has been developed in a couple of different literatures. It is in American legal scholarship; the idea of governance today is fleshed out in what is called the New Governance scholarship. This is a body of work that really talks about administrative regulation -regulated industries, administrative law, labor and employment law, securities law, banking, financial institutions. It is the idea that there are public regulatory institutions and private firms and other actors in a partnership and in a dynamic relationship with one another. There is also a New Governance literature that talks about public/private partnerships in international development efforts: resource management on a global scale, thinking about water and other environmental resources and how those problems are addressed.

There is also the idea of governance that comes out of the political science literature, associated with Elinor Ostrom and her colleagues at Indiana University. This literature talks about management of common

resources and common pool resources, again primarily in the environmental resource context. All of these pieces of literature share a couple of basic themes, and it is really those thematic elements from which I draw my ideas.

One theme is that governance reflects a dynamic interaction among a variety of regulatory agents out there in the universe. These are things that affect our behavior in different ways. I include individuals, groups, networks, firms, institutions, and governments all in one cluster of agents. These are conceptual agents or conceptual influences on our behavior. A second cluster of conceptual agents includes narrative, metaphor, history, and other forms and uses of language.

The third cluster of governance agents are the material elements of our world, that is, physical objects and boundaries in terms of where we find ourselves in space and in place. Therefore, the fact that we are all in this room having this conference and this conversation today, is an element of a governance framework that guides how we engage with one another here today. These clusters and clusters of them interact with one another in a dynamic way under the overall governance framework.

Information regulation is both an output and input of that interaction. Information is both an element of the material environment and an element of the conceptual environment. The materiality of information is something we construct using law and public policy tools. We construct narratives about information that give us guidance about how we should go about behaving vis-à-vis information.

For example, the idea of a work of authorship in copyright is a construct that law gives us. It is not something that authors of copyrighted work sit around consciously processing. The same is true with the idea of an invention in patent law or a mark in trademark law. What is an invention? What is a work of authorship? What is a mark for legal purposes? These are all things that are significantly given by the way the legal system interacts with authors and inventors and businesses that use them.

This is a dynamic thing, or a feedback loop, rather than a process that produces fixed outputs. Here is a quick illustration in patent law: the idea of the art that is the basis for assessing what is a patentable invention is a cyclical dynamic thing that is constantly refreshed and re-informed by the inventions that are recognized in the Patent Office and the courts. What is the market, even in a straight free-market analysis, is often something that is refreshed by the way the intellectual property system treats what is protected in the market and what is not protected.

What governance does not do is automatically privilege the public authority or the governmental authority in a given system and assume that what it does is necessarily the most important in determining what

is legitimate and what is likely to be sustainable or durable in any given information environment. What government or public authority does is clearly important, but it is important to think in the context of other material, conceptual and social actors.

Let me give you an example of what this might do. What I want to do now is apply this to a particular case or particular situation, to see how it might play out in contrast to the more conventional way of thinking about information-related law and policy. This is an intellectual property ("IP") law example, although I plan as I go forward with this project to get into other non-IP related things.

My IP law example is the Google Library Project. Basically, the Google Library Project works with a library. Google signs a contract with a library, in most cases a university library, and agrees to digitize the entirety of the library's collection. A copy of that digital archive is then delivered back to the library that supplied the content, and the content is also added to Google's overall inventory of this material. Google then takes the Google Library contents and makes them available to all of us on the Internet. This information becomes searchable and accessible to everybody through the Google Library interface. For copyrighted material in which there has been no objection registered by the copyright owner, it is searchable. However, only portions, or "snippets," of the works are available and are relatively limited segments of the content available to searchers. Publishers of copyrighted material do have the ability to notify Google on their interest of opting out entirely, in which case bibliographic data is available via Google Library, but not the content. Public domain material is fully available.

Now, a whole collection of doctrinal problems are associated with this project, so it is very controversial. There is a competitor project called the Open Content Alliance ("OCA"). The OCA is engaged in a project that has similar goals and a similar technical aspiration. However, slightly different licensing terms are associated with what they do with the content at OCA. Basically, the OCA does not archive copyrighted material without the copyright owner's consent.

I want to focus on Google Library precisely because of the somewhat controversial nature of it. One is simply the copyright question. Is Google itself a copyright infringer or is Google itself engaging in a form of fair use? There is also the question of whether the users of the Google Library service are fair users of the material. The question becomes whether someone searching the Google archive and pulling up search results, even if he or she is simply relying on the "snippet" version, is a fair user of this material. In addition, there is a question of access to material in the public domain. That is, Google digitizes the entirety of a library's collection, and then redelivers a copy of that archive to the library itself. When this occurs, there is a license restriction built in so

that that particular library agrees by virtue of its contract with Google not to supply that digital archive to any other search engine. Google is in effect the exclusive supplier of that content to the world at large, including that library's collection of public domain material.

That raises a question about contractual restrictions on access to public domain content. By virtue of Google's size and position as a market leader in the search industry, there is a question about the monopolization of knowledge—more of a conventional antitrust type of question. There is a significant question raised about the security and privacy of patron data. If someone is going to use the Google Library archive, what expectations of privacy do you have vis-à-vis the material they use? What obligation has Google undertaken to protect user-related data?

There are also international issues that arise in this example using Google. Most of the debate about Google Library has been framed in the context of U.S. intellectual property law and U.S. privacy law. But there are concerns of various sorts about access to knowledge in both the developed world in Europe and Asia and also in the developing world, where access to the Internet and knowledge generally is a much broader and more complex question. Google and the OCA in a sense bring the contents of the world's libraries to the developed world. They bring the information to the folks who already have access to the hardcopy material. But there is not a lot being done to bring the contents of these libraries to developing countries, where access to Internet resources generally is more limited. There are a whole host of conventional doctrinal and policy questions.

What if we think about the Google Library Project in different ways? I am not going to say that I think Google Library is a good or a bad thing. But if we think about it in different ways, can we figure out ways to enhance the likelihood that it would be accepted and perceived as a legitimate enterprise or not? Or, if you want to disrupt it and marginalize it in some way, could you shift the paradigm in different ways? I want to think about that question in terms of my overall governance framework, and then break that down into conceptual questions, material questions, and institutional/social questions.

Governance as a framework means that they are not three distinct questions: material, conceptual, and social. These questions overlap. But in using those questions, how do we conceptualize the Google Library Project? Here are some quick notes. When the Project got underway two years ago, it was not called the Google Library Project, and the search function was not called Google Book Search. Those are the terms that are attached to it now. When the project was first announced, it was called Google Print. The first round of controversy over this whole project was a controversy over Google Print. Google realized, in my opinion, that the linguistic and metaphoric associations with Google Print

were very different and more problematic from a variety of points of view than those associated with Google Library. The association with libraries and librarians and librarianship and books is a much more benign and presumptively legitimate framework linguistically than Google Print. Google Print makes it sound like Google is in the publishing business, and that it is publishing books without the permission of the copyright holder, and we know that is bad. This is metaphoric governance, but also institutional. Google wants to align itself with a nonthreatening discipline.

A second controversy is that the product of Google Book Search (if you search copyrighted books) is "snippets." You do not get excerpts, you do not get limited copies, you get small portions of the actual book. "Snippets" makes it sound like you get a sentence or two at a time. But you do not get a sentence or two at a time, in fact you get pages at a time. You get a limited number of pages, with search terms highlighted, but you get pages. You get a meaningful amount of text, but Google calls them "snippets." The more courts and the public and the user community can be persuaded that snippets are really what is going on, it minimizes the competitive impact of what Google is doing. Snippets are small, meaningless things. "Pages" sounds more substantial than "snippets."

A third controversy is if you look at the homepage for Google Library, it calls the overall enterprise an enhanced or an annotated catalog of the world's books. This is Google's goal: to create an overall annotated catalog. Cataloging is fine, we know that cataloging is fine; it is part of being a library. Libraries supply catalogs. Catalogs are presumptively legitimate.

What about the material aspects of this? One interesting thing about the Google Library Project is it takes all of these individual hardcopy books, things that we have all gotten accustomed to (in our case, for decades; in the world's case, for centuries) and presents them to us in this seamless database. It is all just one enormous searchable database. Culturally, I think we are skeptical about that. Data as "data" in this giant sense makes us nervous. To the extent that the Google Library Project manages to reproduce the books in a book-like format, that is reassuring. The more it looks like we are simply accessing a virtual version of something that we are already comfortable with, a library or a catalog or both, the more plausible the whole thing is. The more this content simply gets dumped into an undifferentiated database, the more problematic it feels. I have a long paper that I published a couple of years ago that is conventionally referred to as the "things pa-

per,” published in *Case Western Reserve Law Review*.¹ In that article, I explain at great length how things and objects are very, very important in our conceptual apparatus even as we move into a fully virtual and digitized environment. We still like objects and things. The point is both law and private actors can create and manipulate materiality to influence how the world of information is governed.

The last element of this that I want to mention is the social and institutional aspects of it. The fight over Google Book Search was initiated by the Authors Guild, the union representing authors, but there are other institutional interests as well. Universities have an interest in this, and it is not an accident that Google bought access to university collections. We think of universities as repositories of knowledge and distributors and developers of new knowledge. The librarians and librarianship community is very active in the debate over this. The Google Library archive is of obvious interest and value to scholars and researchers, particularly regarding older material. Then there is the public, and consumers that are in the mix. So, we are left to sort out the various group-related interests in this process and how they participate in governing the creation and use of Google Library content. Some of that can be done conceptually, at a linguistic or metaphoric or narrative level, but it is also important to think of the governance impact of the actual practices that scholars, librarians, authors, or other general-purpose consumers bring to Google Library. As a resource primarily for scholars or librarians, Google Library is much less significant in market terms, and therefore should be less offensive in copyright terms, than if it is primarily a consumer resource.

In conclusion, I have three quick notes. First, this governance framework is a way of thinking about problems; it is not necessarily a way of developing specific solutions. Secondly, some people who have described the project say it is a radical reinvention that I am engaged in. I think of it as a very conservative kind of thing, that is, I am really thinking about conserving and stabilizing an otherwise very controversial debate. Third, I think of this as especially useful in planning and assessing on a prospective basis. How do we assess what is likely to work, and what is likely to be stable and legitimate going forward? I think in the future that these issues are going to be more important than being able to simply resolve this lawsuit or that one.

1. Michael J. Madison, *Law as Design: Objects, Concepts, and Digital Things*, 56 *CASE W. RES. L. REV.* 381 (2005).

