



15th BILETA Conference: “ELECTRONIC DATASETS AND ACCESS TO LEGAL INFORMATION”.

Friday 14th April 2000.
University of Warwick, Coventry, England.

Tears Shed Over Peer Gynt's Onion: Some Thoughts on the Constitution of Public Legal Information Providers

Thomas R. Bruce[*]
Co-Director, Legal Information Institute
Cornell Law School
New York, USA

-
- [Introduction](#)
 - [The dimensions of the debate](#)
 - [Disclaimers](#)
 - [Roadmap](#)
 - [Demands of the marketplace: what are we trying to achieve?](#)
 - [Comprehensiveness](#)
 - [Currency](#)
 - [Reliability](#)
 - [How new technology shifts the nature of the discussion](#)
 - [Technology affects production as well as distribution](#)
 - [By increasing access, new technology explodes the meaning of "public"](#)
 - [What are the "competing" models?](#)
 - [A pause](#)
 - [Public vs. private considered generally](#)
 - [Legal information serves only lawyers; forcing the public to bear the cost is unfair.](#)
 - [Private-sector publication deprecates service to unprofitable audiences](#)
 - [Free public information does not add the value demanded by niche markets](#)
 - [Extensive public-sector information unfairly competes with private-sector entrepreneurship](#)
 - [Non-competition and deadlock: the US Code](#)
 - [Should government publish its own works?](#)
 - [Factors ostensibly arguing against government as its own publisher](#)
 - [Currency](#)
 - [The Supreme Court](#)
 - [The United States Code, again.](#)
 - [Plant, equipment, and expertise.](#)
 - [Factors arguing for self-publication by government](#)
 - [Publishing as an \(unwitting\) interpreter of law](#)
 - [Punctuation](#)
 - [Whitespace and structure](#)

- [An X-Rated Divertimento](#)
- [Third-party vs. self-publication](#)
 - [Reliability concerns favor self-publication](#)
 - [Gains of third party centralization depend on the level of value added before transfer, and are self-limiting](#)
 - [Centralized publication diminishes incentives for improvement in the process of drafting and creation](#)
 - [What is the best environment for innovation?](#)
 - [Distributed models depend crucially on a painful process of standards development](#)
- [What can we do, practically?](#)

Introduction

The dimensions of the debate

Eight years' experience with Internet delivery of public legal information provides much to draw on as we think about what we might do in the future. It is difficult to codify that experience into prescriptions about how public legal information providers ought to be constituted and organized. There is no question that the systems of the past failed to give the public adequate access to the law that the public is expected to obey. In the United States the barriers have been largely economic; in theory, most law has been accessible to most everyone, at a price[1]. In the United Kingdom and Canada, Crown copyright has acted as a brake on efforts at electronic dissemination[2] <http://www.droit.umontreal.ca/crdp/en/equipes/technologie/conferences/dac/vaver/vaver.htm>[1]. At this point many if not most of us believe that the public is neither well-served by a private legal-information oligopoly nor by existing government-operated systems. The problem, then, is to determine which of a number of alternative arrangements might serve best in the future.

There is general agreement about where public legal information services must fall in the overall market structure[3] of electronic legal information. Most authors[4]2000 (1D) <http://www.law.warwick.ac.uk/jilt/00-1/bruce.htm>[1]1997 (2D) imagine them to be relatively "plain-vanilla"[5]2000 (1D) services that offer information to the public without diminishing potential addition of value by publishers and other public- or private-sector actors. As such they necessarily have two characteristics or constraints: they must be sited somewhere near the headwaters of the information stream, and they must operate without any exclusive rights or practical monopolies that would serve to dam that stream. So far, so good. Is it possible for us to make more detailed prescriptions?

In his editor's preface to the most recent edition of the Journal of Information Law and Technology [6]2000 (1D), Abdul Paliwala characterizes this search for useful prescriptions as a "crucial debate" between the centralized approach represented by the Australasian Legal Information Institute and the distributed model I advocate elsewhere in that journal[7]2000 (1D). A more sophisticated analysis would be at once simpler and more complex than the bipolar nature of a debate would suggest. Like Peer Gynt with the onion, we can peel away complex layers to reveal two distinct, separable halves. There is no wholly convincing *a priori* argument in favor of either. I believe that complex circumstances viewed over the long term favor a distributed model. Different circumstances might favor a different approach in the short term, and ultimately some union of the two as successors in time.

Disclaimers

"Different circumstances" is a code phrase, of course. It is difficult to be sure how well experience in one country will translate into action in another. The United States is an idiosyncratic venue. We

have an uncharacteristically complex court system. Our long experience with private-sector publication is sufficiently diverse to support virtually any position. We have long denied copyright in government works[8]http://www.spatial.maine.edu/tempe/perritt_2.htm[1], but we for many years permitted expansive copyright claims in the apparatus of citation[9]<http://www.law.fsu.edu/journals/lawreview/frames/241/wymafram.htm>[1], and we permit and even promote exclusive, private arrangements with publishers[10] that have an equally restrictive effect. We present a confused and confusing picture, and one from which it may not be useful to generalize. It may well be that for Americans complexity is the deciding factor, leaving distributed solutions as the only ones that will scale properly. But the situation in America may also be exactly the kind of pathological case that produces solutions useful in more tractable settings.

Most of us in the public legal-information game are in a reactive mode. We see ourselves as freeing the law from private-sector oligopoly or from Crown copyright or from bureaucratic inefficacy. But it may be a mistake to let ourselves be led entirely by what we believe has gone before, or to assume that the characteristics our institutions have exhibited in the past are inescapable. There is nothing foreordained that prevents government from publishing law in a timely manner, that prevents the private sector from setting a reasonable price on published law products, that limits the utility of laws to lawyers, or that relegates academics and librarians to the role of passive commentators and consumers. We live in a time when literally everything is up for grabs - including some of the assumptions made about the evils of the *ancien régime*.

Roadmap

Like the layers of the onion, the layers of this analysis tend to be difficult to separate and array neatly into a linear structure. I have somewhat arbitrarily divided the discussion into three brief stage-setting sections that outline the characteristics of the new marketplace for legal information and the nature of the "competing" centralist and decentralized models. These in turn are followed by detailed exploration of three broad discussion points that organize a plethora of closely related questions.

Demands of the marketplace: what are we trying to achieve?

In the sections immediately following I consider the primary demands placed by consumers on legal-information systems, whether public or private. At times it may appear that I am questioning whether or not end users really know what is in their best interest. If I am doing so, I am doing so in a way that can happen only at this particular point in time. Though this is a less troublesome problem than it was a few short years ago, consumers still have trouble seeing the qualities they prize implemented in the new information environment, not because they are not there but because they have taken on new, largely cosmetic externalities[11]<http://lawlibrary.ucdavis.edu/LAWLIB/Dec98/0118.htm>[1]. To some extent, this is an evanescent problem that reflects the youth of the Internet legal-information environment. In larger part it is a problem first of recognizing what was baby and what was bathwater in the old environment, and then of figuring out where the baby has gone in the new milieu. The planner of legal information services must then somehow predict what the baby is likely to be like when it is full-grown. Internet-based legal information services are far from mature.

End users have been conditioned by training, experience, and careful marketing campaigns to value particular aspects of familiar systems like LEXIS and WESTLAW. Those systems are strongly branded, and have a deservedly high reputation among those who have been able to use them. It is not at all surprising (though it is at times dismaying) that experienced users find it difficult to recognize those same virtues when they are produced by new and unfamiliar implementations. At the core of this phenomenon is a bias induced by thirty years' experience with older computer systems and older modes of industrial organization: centrality equals reliability. The Internet approach stands

in sharp contrast as it argues the contrary: decentralization equals reliability, attainability, and scalability. On some profound but subliminal level this is news that shocks and bewilders. New, distributed models of computing that are reflected in distributed information systems and distributed models of business organization must seem inherently anarchic and therefore inherently suspect, no matter their virtues. That suspicion will subside in time, to a degree. But it will never vanish entirely until we become more discerning than we are about what was necessary about older ways of doing things and what was merely incidental.

Comprehensiveness

Legal professionals feel a strong need for comprehensive electronic sources of legal information. They remain oddly naïve about the nature of comprehensiveness in the new, distributed network environment and, to a degree, about their own research practices. It is understandable that lawyers researching caselaw would want to feel that they had searched or seen every single available case on point. What is slightly surprising is their seeming insistence that these results come from a central provider, even when it can be demonstrated to their intellectual satisfaction that an aggregation of providers is providing the same collection of cases searchable from a central point[12].

This is all the more interesting given that their beliefs about what constitutes comprehensiveness are demonstrably subjective and often wildly off-base. What they believe to be a comprehensive search result falls far short of actual comprehensiveness; one study shows a high degree of satisfaction with results that produce surprisingly low levels of recall[13]Dabney, "The Curse of Thamus: An Analysis of Full Text Legal Document Retrieval," 78 Law Library Journal [5] <http://www.yale.edu/lawweb/lawcrs/lasdab1.htm>[1997 (2)]. This suggests that we may be dealing in realms that are not entirely rational, and that have more to do with learned behavior and the need for assurance that one is doing one's job properly. It may also have to do with more than one hundred years' worth of marketing by large commercial providers, each of which is eager to show itself more comprehensive than the others. Indeed, it is difficult to see end-user criteria as separable from vendor claims, in part because the features being trumpeted often do serve market needs and in part because the influence of vendors has been so pervasive both with lawyers and with the librarians who make their information-purchasing decisions.

It was not always that way, though one has to reach back a considerable distance to find a time when it was not. The earliest critiques of the West system took it to task for being *too* comprehensive and hence, non-selective[14]. In 1889, John B. West responded to these complaints in a way that foreshadowed much that was to come:

It is one of the greatest merits of the National Reporter System that it gives all the cases. Some of our critics call it a "Blanket System," and we are disposed to accept the analogy. No policy of insurance is so satisfactory as the blanket policy; and that is the sort of policy we issue for the lawyer seeking insurance against the loss of his case through ignorance of the law set forth in the decisions of the highest courts.[15]

To take Mr. West up on his analogy, concerns about comprehensiveness are exactly as legitimate as concerns about ones' house burning down. Sensible people take sensible precautions, and they pay for insurance policies at a rate they believe appropriate to the value of their goods. They can be induced to pay still more if they can be persuaded that failure to do so is risky.[16] This is not empty analogy; the success of state-oriented CD-ROM collections and online CALR contracts limited in volume and scope shows that practitioners are recognizing that some goods only demand a moderate amount of insurance. Practical lawyers with limited library budgets have probably always done so. [17] It remains important that we recognize that some of our attitudes toward any risk involved have been shaped by the subtle influence over more than a century, reinforced by the adversarial nature of the system and the competitive nature of graduate legal education.

Finally, one must also ask, "Comprehensiveness for whom?" The limitations of print made it difficult if not impossible for John West to publish more than the "decisions of the highest courts" in 1889. More than a century later, the large electronic publishers still lose comprehensiveness as one moves lower and lower in the hierarchy of the courts - and closer and closer to the concerns of the average private citizen. Certainly neither LEXIS nor WESTLAW is a good source of the more localized legal information that impacts private citizens far more directly and frequently than the rulings of distant if important appellate courts. And comprehensiveness is of no concern to anyone who is barred from access to the database, whether by its cost, or by accidents of geography, or by its failure to take notice of the laws and regulations having the most impact for them. Comprehensiveness takes on new meaning - and becomes significantly more challenging - as we try to imagine a system that will provide universal access to the law most directly affecting the public.

Currency

Currency takes on two meanings in the legal-information marketplace. People want up-to-date versions of statutes and regulations. They also want timely reporting of judicial decisions, though here the notion of timeliness is more relaxed. As with other aspects of the legal information market, this criterion has at times taken on inflated importance as competing commercial entities seize on it as a point of competition. For example, a few years ago the currency of Supreme Court decisions was at issue, with each of the American commercial providers claiming to "scoop" the other by a matter of minutes or hours. It is difficult to see what difference so timely a level of timeliness could possibly make to an actual legal process; the information services were acting more like news organizations[18].

As a rule American courts seem to feel more of an obligation to provide currency than comprehensiveness. The US Supreme Court operates a scrupulously up-to-the-minute dissemination service in the form of Project Hermes, but as yet has no Web site offering a fixed archive of opinions. Several courts put opinions online only until they appear in the official reporter or for arbitrarily limited timespans[19]<http://www.alaska.net/~akctlib/sp.ht>[m]<http://www.cobar.org/coappcts/scndx.ht>[m]; others segregate older opinions into separate archives [20]. Motives for removal vary; some courts profess concern for comparative inaccuracies that may creep in as official versions are refined, while others keep their reasons to themselves. A cynic might suspect reluctance to interfere with the revenue streams that flow from exclusive publishing contracts. It may also be that there is a well-meaning but misdirected attempt being made to establish a sharp division between public and private systems using timeliness as the value criterion, much as a free stock ticker might offer quotes that are delayed by some minutes to avoid stepping on a sister for-fee service.

Closely related to the notion of timely dissemination of opinions is the idea of notification or current awareness - that along with the release of a decision or the passage of new legislation something would occur that pushes the event into the view of the user. Usually this also involves some degree of filtering; the customer does not want to be hit over the head by every event, but only some definable subset. This is a need expressed primarily by professionals, but it is worth noting that the need is not restricted to *legal* professionals - it applies equally to anyone affected by law or regulation in some ongoing and vital way. For the general public, this need is often filled by traditional news media, trade associations, specialty newsletters, or simply by word of mouth.[21]<http://www.victimservices.org/fw97eo.htm>[1] Such systems are remarkably easy to build once the infrastructure of speedy information gathering is in place, but of course they rely crucially on a timely stream of information from the source.

Reliability

Concerns about reliability come in several flavors. Some users are concerned about both short- and

long-term availability of cases through public-sector publishers: these are concerns about service availability, on the one hand, and archiving, on the other. Others worry that the texts being published may not be as accurate as those produced by commercial houses.

What is interesting about these assertions is that the virtues attributed to private-sector services are in some sense accidental and reputational. There is nothing about a private-sector service that would, *a priori*, cause it to publish reliable texts that are continuously available and scrupulously archived [22]. There is nothing that would, *a priori*, prevent a public-sector provider from doing so, as many of us are proving. Indeed, digital signatures make it possible for Internet based providers to prove that what is being transmitted is *exactly* what was received from a court or legislature, without alteration of any kind [23]. There is no guarantee, other than time and experience, that texts from West are inherently more reliable than those published by public-sector actors on the Internet. WESTLAW and LEXIS have reliable brands; the public-sector publishers do not as yet, because they have had no time to establish them. In short, concerns about reliability are revealed to be concerns about the new, and they can really only be allayed by experience with new, Internet-based brands.

How new technology shifts the nature of the discussion

Most discussions of the new Internet technology and its effects on legal publication have focussed on the dramatic way in which it lowers entry barriers for would-be legal publishers. Reduced entry barriers create a more competitive marketplace, and enable self-publication by governmental bodies and other kinds of public-sector entities. These are not the only effects of the new technology, and Internet technology is not the only one with which we should be concerned.

Technology affects production as well as distribution

Early commentators on law-on-the-Net [24], and most commentators since, have been quick to point out that Net distribution of legal information is quite inexpensive when compared to print distribution. Given that most creators of electronic legal information now create it using electronic means, the incremental cost of Net distribution is very small indeed. The emphasis on distribution is understandable given that it is the most prominent feature in the landscape of change, but it has also biased the discussion.

Our preoccupation with the pace at which government has or has not made use of new technologies of distribution [25] <http://merlin.law.mercer.edu/elaw/inter2.htm#firs> [t] <http://www.droit.umontreal.ca/crdp/en/equipes/technologie/conferences/dac/bruce/bruce.htm> [1] has distracted us from the technologies of production. It has been almost a decade since RIA, the tax publishing branch of Thomson, built its SGML-based TIGRE system. A book [26] that uses it as a case study has been a standard reference for SGML implementors for more than five years now. Yet SGML has yet to make significant inroads into the Law Revision Counsel's office where the tax code itself is edited and maintained. Some would argue that this demonstrates that government simply cannot shift to new technology as quickly as a better-funded, profit-seeking, competitive private sector can. That may be true, but it is not foreordained. There are many counterexamples [27] 2000 (1) <http://www.law.warwick.ac.uk/jilt/00-1/arnold.htm> [1] http://www.seyboldseminars.com/seybold_report/reports/P2514001.HT [M] <http://www.ncsc.dni.us/NCSC/TIS/JEDDI/Jeddi-1.htm> [m] in which government has kept pace with the private sector in adopting new publishing technologies. The real problem is that few have yet imagined the creation and distribution of legal text as a fully re-engineered process reaching from draftsman through legislature to citizen. In the US, some branches of government have moved quite quickly to develop an Internet presence, but relatively few have revamped the means by which

they produce the content distributed via that apparatus. Many government organizations have been paving over cowpaths, accelerating and widening distribution of a product that is still produced by relatively inefficient means. This is a classic example of incrementalism at work where reengineering is needed[28]. It says nothing about the ultimate capacity of government as a publisher.

By increasing access, new technology explodes the meaning of "public"

In 1995, it was still possible to make the assertion that most use of published law products was made by lawyers[29]. That may no longer be so, if it ever was, and the distinctions that made such a claim possible are breaking down. It was true then, and probably remains true now, that most *law books* are in the hands of lawyers. But there have always been countless ways in which citizens have used legal information encountered in other texts and contexts. In the past, the texts consisted largely of glosses, guides, and handbooks - secondary sources constructed for the benefit of the average citizen, like income-tax self-help guides, or regulatory manuals prepared by trade associations for their members. The Web at once puts all those scattered sources into one space, and offers the possibility of linking those secondary sources to primary material that many people wish to see for themselves. That interest is elastic, and price sensitive, but it is there nonetheless and it is huge. Often, it is not interest in a legal-information product as large and lumpy as a book, but instead in some paragraph or section or smaller granule that is especially important to the searcher. Often, too, it is professional - but the profession is not lawyering. It is, instead, a profession bound up with law in some less direct but nonetheless fundamental way - that of (e.g.) a health-care administrator, or pension plan manager, or police officer, or inventor[30].

Two notions proceed from this. First, our definition of "public access" to law has implicitly but dramatically changed. We must now imagine an expanded public seeking smaller and more relevant granules of information, and seeking it via the Internet. Second, increased granularity, a diversity of specialized interests, and a vastly expanded audience, taken together, strongly imply a need for constellations of niche legal information providers. Let us consider these ideas in turn.

One could outline the changing nature of public access succinctly by saying that government can no longer discharge its public-access responsibility by insisting on the delivery of a few free law books to a few libraries. Current contracts between state government in the US and private legal publishers actually do contain public access provisions, and at one time they probably seemed reasonable. One from New York State, a state with over 18 million residents, reads:

6. Said contract shall require the contractor to furnish the state library with fifty-eight copies of the court of appeals and appellate division reports and three copies of the miscellaneous reports, and also to furnish copies of each of said publications as follows: One of each to the clerk of each county, for the use of the county; one of each to the attorney general, for the use of his office; one of each to the state comptroller, for the use of his office; one of each to the clerk of the court of appeals, for the use of that court, and one of each to each judge or justice of a court of record, for the use of his office; and one of each to the various public law libraries in the state, and the expense of delivery thereof shall be borne by the state. [31]

As the libraries mentioned here are not public libraries, but public *law* libraries; it is not hard to guess who the likely users of this "public" access were thought to be. And with the percentage of American households with Internet access reaching roughly 37%[32] <http://www.altavista.com/av/content/pr091599.htm>[m], and nearly 100% of public schools and libraries connected, it is clear that the Internet is a better venue for public access anyway. One might also adduce other advantages: law on the Internet is more accessible than "book law" intellectually as well as physically, offering a familiar interface and more colloquial forms of organization. One can

only imagine how private legal publishers would react to a requirement that they provide free access to law via a web site.

A legal information product is in fact a *cross* product - the intersection of a particular primary source with a particular audience. This is not just a matter of creating secondary sources for niche audiences, as lawyers do when they produce client newsletters, practice guides, or CLE materials. Marc Galanter writes that law "usually works not by exercise of force but by information transfer, by communication of what's expected, what forbidden, what allowable, what are the consequences of acting in certain ways".^[33] Like any other communication system, primary legal sources can be rendered more or less effective by structuring and presenting them differently for different audiences.

Publishing for such niche markets is an activity that is both profitable for private sector concerns and a matter of necessity for public bodies that have an explicit or implicit mandate to serve audiences with special needs and perspectives. Unsubtle examples include language translation and specialized presentation for the physically disabled. Other less obvious needs exist, as well. Agencies routinely rearrange statutory material into structures and sequences more useful to field workers or program administrators^[34]<http://www.swrcb.ca.gov/html/wtrlaws.htm>^[1]
<http://www.upenn.edu/researchservices/spmanual/agency.htm>^[1]. Professional associations, trade groups, and lobbying organizations build legal guides intended for niche markets in a particular industry or demographic slice, such as the elderly. In the private sector, the fact that publishers must provide products for a range of media and audiences - law in different packages - is often cited as a major competitive reason to adopt markup technologies such as SGML and XML^[35] that facilitate reaggregation of legal information into new publications. They are not the only organizations with reason to do this.

The major points to be made here are two. First, the vastly expanded notion of "public" that is implied when the means of public access is the Internet is not a simple expansion of numbers of readers. It is a combinatorial explosion of niche audiences for legal information. It is difficult if not impossible to service the needs of such a diversity of diversities without reliance on the private sector. They cannot be cut out of the picture ^[36]. At the same time, it would be a mistake to assume that servicing of niche legal-information markets is an activity that is the *exclusive* concern of the private sector. Public creators of legal information often service niches directly, and must be as mindful of the diversity of audiences that they serve^[37]<http://www.census.gov>^[v]
http://www.ssa.gov/SSA_Home.htm^[1].

What are the "competing" models?

The competing centralist and distributed models represented by the Australasian Legal Information Institute and its American elder brother have arisen from very different circumstances that more than anything else have determined their characteristics. The critical differences between the two countries are differences of scale and history. America has a good many more jurisdictions and a baroque system of courts and legislatures, and has experienced almost total domination of law publishing by the private sector. Australia's legal and government communities are much smaller, and the private sector has not had a chokehold on publishing. For these reasons, Australia is a place where it is much more practical to talk about a centralized public legal information system operating outside of government, and to take the political and institutional steps necessary to set it up. In the United States, the notion of a centralized operation is not merely daunting but probably impossible outside the private sector.

As a result, two different visions of the public legal-information regime have arisen. AustLII acts as a kind of comprehensive, centralized, sophisticated service bureau for official bodies interested in publishing legal information, and enjoys close, quasi-official relationships with those whose

information it publishes[38]http://elj.warwick.ac.uk/jilt/LegInfo/97_2gree/contents.htm[m]
<http://www.law.warwick.ac.uk/jilt/00-1/austin.htm>[1]. By contrast, the LII is non-comprehensive, has no official relationship with any government body or agency, and receives no public funding except indirectly through our parent school. We operate three major flagship collections (the decisions of the US Supreme Court, the United States Code, and the decisions of the New York Court of Appeals) as examples of what it is possible to do with legal information, but we do not aim for comprehensiveness of any kind. Instead, we aspire to be a center for applied research that will serve as an example to a wide variety of self-publication efforts in the United States and elsewhere. We imagine that a comprehensive public legal information regime will be an aggregation of different low- or no-cost providers acting under a variety of arrangements, principally self-publication[39]
 2000 (1D)<http://www.law.warwick.ac.uk/jilt/00-1/bruce.htm>[1], and achieving interoperability through common standards and practices. If pressed, we would probably advocate the formation of something like a W3C[40]<http://www.w3c.org/>[1] for law, a consortium that could develop and promulgate interoperability standards, but we would not imagine that it would take on responsibility for comprehensive publication at any level, including service as a comprehensive portal.

At this point in the proceedings, it would be tempting to save a great deal of trouble and effort for both author and reader by pointing out that a global distributed model can contain any number of localized centers of concentration. We could then declare that synthesis has occurred at a higher level of abstraction, and adjourn. Unfortunately that would do even less to illuminate the public policy issues than any further scribblings by the author are likely to. Nevertheless, as the discussion unfolds, it may be useful to bear in mind that there are ways for the two visions to coexist, either at different scales or as successors in time, an approach I shall say more about later.

A pause

With background behind us, let us now do a little stocktaking. We need to consider our central problem - the ideal structure for public legal information services - along a series of dimensions rather than as a simple choice between centralized and distributed models. For in reality we have a series of binary choices to make, leading us down a variety of potential paths. A useful way to categorize different sets of alternatives might be

- * Public-sector versus private sector
- * Publication by government versus publication by others, be they private- or public-sector actors.
- * Publication by centralized publisher versus self-publication by creators, with no consideration of whether the centralized publisher is public, academic, or private.

Obviously one can produce some odd and contradictory results by combining these three sets of alternatives willy-nilly - there is no such thing as a private-sector, government-run, self-publication operation, for instance[41]. But a look at some prominent examples suggests that in fact we should look at all three dimensions if we are to accurately describe many existing operations. For instance, the US Government Printing Office is a centralized, public-sector operation run by government; AustLII is a centralized, public-sector operation that is not. The loose consortium that publishes the opinions of the US Courts of Appeal is both public- and private-sector, with elements of self-publication that are within government, elements of self-publication undertaken by arrangement with academic institutions outside government, and elements like FindLaw that are private-sector actors espousing public-sector ideology, all linked in a distributed system.

One can imagine other dualities having to do with funding sources or technology. One might also make a practically infinite number of distinctions based on price point, though this would be a tortuous analysis unless we were to restrict ourselves to three points that we could define as "wholly

free", "cost-recovery", and "what the market will bear"[42]. Let us, however, abolish such complexity at a stroke by declaring that we will be concerned with services that are priced at or below cost, as they would likely be in a public-access regime, and that we will limit ourselves to Internet based technologies[43]. For the moment, it is enough to consider the three major choice-points listed above.

Public vs. private considered generally

Most of the discussion that treats the question of public versus private responsibility for legal information publishing ends up recommending some form of coexistence. Neither public nor private means can by themselves answer every demand that exists for legal information. The debate is not really about whether public-sector information ought to exist to the exclusion of private services, or whether private-sector services can answer all public needs, but where the dividing line should be set between the two. Curiously, most of the arguments on this point end up being normative; there is little as yet in the way of empirical evidence or economic theory that would favor placing the dividing line in one particular place or another. Here are some arguments that have been made in the past:

Legal information serves only lawyers; forcing the public to bear the cost is unfair.

Several years' experience with legal information on the Internet permits us to more readily see this argument as the nonsense that it always was. To be fair, it seemed a lot more reasonable when we had only print and commercial online services to consider. In 1995, it was quite possible to get hold of the wrong end of the binoculars and argue, as one commentator did, that because nobody but lawyers made use of law books or of Westlaw, it was unfair to expect the public to pay for such a service:

...Any government effort to distribute free or low-cost legal information will benefit lawyers more than anyone else. By far, the greatest use of legal information is made by lawyers in various guises, largely to make money. If the government does expend significant resources to create an entirely new legal database, the greatest beneficiaries will be the for-profit legal community...One would have to be incredibly naïve to think that such a program will end up really being designed for the average citizen. Subsidizing the practice of law is hardly an attractive goal for public spending at this time[44].

On the whole one would prefer naivete to the disingenuousness of this argument, which willfully ignores two aspects of the market for public legal information. First, the non-lawyer market was and is extremely price-sensitive, and flew under the radar of WESTLAW and LEXIS largely because it was unwilling to pay LEXIS and WESTLAW prices. Before the existence of less costly published law sources, it was a difficult market to detect or evaluate (and there is still much to be done on that score). Second, it was a market that chose not to employ the professional means and methods of legal research either because it had no access to them or because it did not know how to apply them. Instead, it found other means to pursue its informational needs, largely through secondary sources and through public libraries. It is unusual to see (e.g.) someone with a question about the changes in the practical limits on their activities as a police officer that flow from a recent Supreme Court decision wandering the stacks of a law library[45]. It is almost inconceivable that such a person would use a traditional commercial electronic service at full price. From the perspective of conventional research systems encountered in law libraries these people were invisible[46] <http://www.law.uh.edu/swall/schedule2000.htm>[1]. However, to deny that they exist is to deny that anyone ever bought a tax-preparation aid or a book on making a will, consulted the local building code, or read about an appellate court decision affecting (say) the limits on their behavior in the

workplace.

Private-sector publication deprecates service to unprofitable audiences

The idea here is a simple one: if it is profit that motivates legal publishers to publish, then law that turns a profit will be published first. A look at existing services will show that for the most part this proposition is true. The law that is available from commercial publishers is largely the law employed in the service of deep pockets. By way of illustration one might contrast the level of publishing activity surrounding the law governing mergers and acquisitions with that of the law of veteran's benefits. The first is complicated but highly profitable; the latter is insanely complicated and distinctly unprofitable, since lawyer fees are capped.

In this sense law publishing as a communications system resembles others in which there is a public-policy interest in universal service, such as postal or telephone service. The market will not act in such a way as to provide universal access to law; it has to be legislated. Drafting practical legislation on this point would be a daunting task; the notion of universal or public legal information service begs the question, "What level of value is to be provided for free?"[\[47\]](#) At the moment we have little idea of where we should set the level of value-added in a universal, public system.

Free public information does not add the value demanded by niche markets

Information provided by a universal, public service can never add value in all the different ways that niche markets might expect. This has always been true to a degree, as witness the variety of specialty legal publications serving different varieties of legal practice. It is even more the case now that the dramatically larger audience available on the Internet has revealed niche markets that were previously unreachable as such, and as notions of "mass customization" increasingly dominate consumer expectations generally[\[48\]](#). In practical terms this plays out as an issue of scalability. It is made more acute by the fact that niche legal information products crucially depend on added editorial value that is always expensive to provide and often the product not of general legal expertise but of specialized knowledge in a field identified with the target audience, such as accounting, finance, or engineering. It is not to be expected that freely available public sources could supply nearly the quality of niche product that the private sector could.

Extensive public-sector information unfairly competes with private-sector entrepreneurship

The "unfair competition" card is usually the first one played in any debate over the relative value of public and private regimes. I have saved it for last because, like some other propositions we discuss here, it turns on the level of value being added by the public source. A dump truck full of free legal information published on paper cannot compete with even the most primitive electronic source offered for sale at any remotely rational price. At the other end of the spectrum, a private source that adds costly features offering only marginal improvements in functionality will not last long in an open market[\[49\]](#). These are examples at the extremes; things become a lot less clear as we choose examples of public and private systems that are closer to one another in the value they add to the same dataset. As yet, there is no country that has grappled with detailed questions of how much added value the public has the right to expect from a free system. AustLII, the LII, CRDP, and others who are focussed on technical innovation are at least implicitly asserting their right to add as much value as innovation permits in the hope of forcing commercial systems to still higher levels[\[50\]](#). This is good insofar as it provides leadership, but it falls far short of being a public-policy determination

that would establish useful guidelines for content creators.

As Perritt points out in the context of GIS data[51], this debate does not end with asking whether government will itself be a competitor or not. Once government realizes that it has the ability to *control* competition by making decisions about what resources it will offer for free, and by choosing who it will allow to add value to its creations, other considerations take over. There is a temptation to restrict access in order to increase revenue flow to government; value-adders will pay dearly for exclusivity (and, by contrast, nothing if they feel that a competing public source is pre-empting their market). At another extreme, government can eliminate all competition and thus effectively limit knowledge of its activities. Thus, issues of exclusivity and censorship flow directly from competitive considerations.

Non-competition and deadlock: the US Code

Often, too, our perception of whether competition is "fair" or "unfair" is accidental: it depends on whether it was the private or public sector that first made a service available. Nowhere is this better illustrated than in the current systems surrounding the publication of the US Code. There are at least four electronic versions available. The United States Government Printing Office offers one that is anywhere from 12 to 24 months out of date, depending on the title[52]. The version offered by the LII is derived from the GPO version, and hence is equally out of date but more readily navigable because of better hypertext functionality. By contrast, electronic editions offered by private publishers are completely current. They are kept so by large teams of editors. This is not an inexpensive effort and one imagines that the cost is, ultimately, passed back to the consumer.

This situation presents a number of issues. First, one might ask whether or not citizens are entitled to the same level of service as lawyers and government officials - particularly as we contemplate a service where currency is important and large numbers of the governed are affected. Second, there is the issue of duplicated effort and efficiency. Would it not make sense to have government provide an up-to-date version, and save the private sector acquisition and updating costs that are incurred by every downstream publisher? Yet I am quite sure that if government were to begin providing an up-to-date version this would be seen as undercutting the market for private versions in an undesirable way, and would be quite controversial. Government initially offered too low a level of added value in the public version; perhaps such snapshots or consolidations were all that was possible at the time. It will, however, be very difficult to change matters because the private sector has developed products that go the next logical step in adding value, and can complain that competitive processes would be unfairly undercut by the release of a more capable public product. Most would argue that government should go ahead and do what it should have done in the first place by making an up-to-date version available and thus placing the burden on the private sector to find new levels of value to add. Such an outcome is desirable, but it would take considerable time and effort to achieve. Resistance to change would be high and it would take time to resolve the issue of competition. For practical purposes the situation is one of deadlock.

Again, there is nothing to argue conclusively for public information services over private ones or vice versa. The need for universal service provides a *raison d'etre* for public versions, since it is unlikely that the private sector will adequately service unprofitable niches. The need for relatively deep layers of added value to serve professional and specialized markets, as well as some arguments having to do with government control of information, argue that private-sector services are needed too. What is more difficult is to decide how the two ought to coexist. If we make mistakes in setting the appropriate level of value in public versions those mistakes will be hard to undo. If the level of value is set too high, it will stifle private-sector services at birth; if too low, it will be hard to attain more reasonable levels of public service without controversy over unfair competition. It might not be all that difficult to answer questions about what the public is entitled to were we actually to consider the needs of the public. Unfortunately, to the extent that there is any real public-policy debate in this area at all, it is taking place between lawyers, legislators, and specialists most of whom have an

outdated, incomplete, and artificially polarized picture of what the public does with legal information.

Should government publish its own works?

My own opinion is that government has a responsibility to electronically publish its own creations, be they judicial opinions, legislation, or regulations. It is simply an obligation to inform citizens of the rules by which they will be governed. Many commentators maintain that there is nothing that constrains government to carry out that function itself, that government is not a particularly skilful or conscientious publisher, and that there is much in the historical record to demonstrate this lack of competence. While the historical record speaks for itself, it does not determine what will happen in the future, and there is much that argues for government self-publication. Let us look at some of the arguments arrayed on both sides.

Factors ostensibly arguing against government as its own publisher

Currency

Most of those who judge the creators of legal information incompetent to publish their own work base their attacks on problems with currency of information. Their central claim is that any process of dissemination undertaken by government is inevitably slower than one accelerated by market forces. Historically, this may well have been true^[53], but at this point one has the impression that these arguments are being made by rote and with only cursory examination of the present state of affairs. Government has been slow in the past, but in some places it is not now, and there is no guarantee that private industry is any faster. Indeed, recent experience in the United States shows that government operations with universal-service mandates can provide highly efficient and effective service^[54][http://epinions.lycos.com/srvc-Postal-Brand-USPS/PriorityMail\[1\]](http://epinions.lycos.com/srvc-Postal-Brand-USPS/PriorityMail[1]), especially when goaded by similar operations in the private sector.

The Supreme Court

Consider the following:

Competition for profits has produced quality. Anyone wanting a simple example of what happens when the market is not at work need only compare the United States Reports, the officially produced version of United States Supreme Court decisions, with privately published reporters covering the same cases. The United States Reports is compiled by a Court-appointed reporter and published by the Government Printing Office. It is a well-produced set of reports that provides readers with Supreme Court decisions one or two years after they have been handed down by the Court and long after they have been published in the private reporters. Whereas a government-subsidized system can live with that sort of egregious delay, private publishers responding to market demands and seeking profits cannot. Who would subscribe to such a set unless forced? ^[55]

Berring tries here to sketch as inevitable something that is merely accidental, as an examination of the Court's current electronic-publication practice reveals. In fact the *most current* source of decisions from the Court is the Court itself, which releases its opinions through the electronic service known as Project Hermes simultaneously with their delivery from the bench. To be sure, these are bench opinions, subject to editing before their appearance in the official reports, but they are good enough for most people who have a real need for currency. If they were anything less then it would simply be unconscionable for the Court to release them. Indeed the electronic avatars of the private-sector entities that Berring admires so much obtain their electronic versions of the opinions from the

Court by this means - and are two to four hours slower in placing them online than the LII, a non-profit provider. Creators of legal information are not ineluctably slow in placing their information online. They can lead the pack. And in some ways it makes little sense to use past performance in print media as an indicator of future performance in the electronic venue[56].

The United States Code, again.

Legislation presents disseminators with a more difficult problem than judicial opinions do. It is far more expensive to keep current; it undergoes continual revision and periodic consolidation, whereas judicial opinions stabilize in official form relatively quickly. And usually more is at stake. As a rule, legislatures speak to more parties more powerfully than courts do; judicial decisions that directly affect more than those immediately involved are comparatively rare, whereas with legislation the situation is reversed. Legislation has the larger audience by far.

Our earlier discussion of the US Code mentioned four means of electronic access, two of which are freely available and poorly updated, standing in sharp contrast to a set of for-profit systems that are well-updated and expensive. In the context of timeliness, it is worth noting that governments elsewhere are doing much better than the American Federal Government, both at the state level in the United States[57] and elsewhere[58]2000 (1)]. Untimely delivery by government is not some inevitable result of self-publication-by-bureaucracy. Instead, it is a process-engineering problem demanding solution in its own right[59]. To "solve" the problem by handing it to a third party, public or private, is simply an incrementalist, Band-Aid solution that removes responsibility from a place where it belongs: with the people creating the information. Government can provide timely information if it wants to, and if it is made clear that abdicating responsibility to third parties is no longer an option.

Plant, equipment, and expertise.

Part of the problem is that there was a long period where such an abdication of responsibility was the only practical alternative, given the constraints of print. Early legal-publishing operations were almost always turned over to private printers. In the UK this was at least partly a matter of patronage; in the United States, government could not afford to operate the printing plant and distribution facilities needed even in Colonial times. Often a contract for printing the laws was not seen as a boon; printers complained that the activity was simply unprofitable, and had to be coerced:

It was left to individual enterprise to carry this important object into effect; and as the undertaking would be attended with considerable expense, and interruption of other business, without any prospect of private advantage; no professional gentleman, for a period of a few years, appeared willing to make the requisite sacrifice[60]

In time, of course, legal publishing came to be hugely profitable[61], and given the industrial technology of print it remained efficient to place publishing operations outside of government in a well-capitalized private sector. Until recently, computerized research systems were no different - the creation of centralized mainframe systems and of the access systems that would permit lawyers to reach them were if anything more capital-intensive than the creation of a printing plant. In this respect the early electronic systems held no seeds of change - it still made sense to develop capital-intensive electronic publishing systems in the private sector. Thus in matters of industrial organization centralized electronic systems were little different from their print predecessors. As a distributed system, the Internet presents the first real challenge to the print-oriented mindset that says that legal publishing can only take place in a centralized system. Legal publishing is now something that government can practically do on its own[62].

Factors arguing for self-publication by government

Publishing as an (unwitting) interpreter of law

Some consumer concerns about the reliability of legal information are closely related to questions about whether or not the artifacts of publishing can act in ways that affect interpretation of the law. Librarians writing about headnoting and indexing systems have long realized that systems that organize caselaw, such as the West key numbering system, also serve to pigeonhole decisions in ways that may obscure their meaning or usefulness.[63] In retrospect it seems fairly obvious that these editorially-constructed secondary sources, like any system of organization, have always had the power to miscategorize, or to categorize in ways that lose accuracy and relevance with time and with an evolving understanding of the meaning (or more properly meanings) of a particular case or statute. Insofar as full-text search bypasses such categorizations these arguments need not detain us. The points I want to pursue here are tied more closely to publication of primary sources, and to our overarching question of whether it is better for the creators of legal information to do their own publishing, or not. We can take it as given that courts and legislatures are in a better position to clarify what they themselves were trying to say than publishers are. The question, then, is whether the act of publication creates any interpretive effects, and if so what the consequences might be.

Punctuation

The strongly-worded disclaimers placed at the beginning of electronic legal documents provide compelling evidence that courts see the potential for difficult consequences to flow from inaccuracies. Probing into the history of cases revolving on concrete punctuation errors yields a less clear picture; at times courts have considered punctuation an "official" part of decisions, and at times they have not.[64] Nevertheless, interpretation can sometimes turn on a comma[65]. The aim of a publication system should be simply to avoid the need for courts to consider such questions; we are not concerned so much about what the law has said about the issue of simple publishing errors as we are with avoiding occasions on which we might find out. Such frolics are costly, and perhaps unnecessary. The question, then, is who is more vested in avoiding errors - a court or legislature publishing its own work, or a third party?

Errors that omit, replace, or add a character are clumsy and obvious - more so than errors of paragraphing and indentation, errors that affect the way in which we perceive the text to be structured, or errors based on the application of logical meaning through typography or tagging. These latter are particularly important to consider as we undergo a fundamental shift in markup technologies. In a sense, the fuzzy way in which essentially presentational markup schemes like traditional typography and HTML bind logical meaning to the appearance of a text is a kind of circuit breaker. It makes it possible to deny well after the fact of publication that the *appearance* of the text has any inherent relationship to its *meaning*[66]. On the other hand, technologies that bind logical meaning to markup more tightly, such as XML and SGML, raise real questions about who is to be entrusted to do the markup. To be sure most markup schemes will be fairly general[67] and most likely confined to fairly incontestable metadata like the name of the author of an opinion or the date of enactment of a statute. No matter how much law students might wish for it, it is not likely that we will ever see a judicial opinion containing tags like <PAYATTENTION> or <DICTUM class="IMPORTANT" duration="ETERNAL">, even from a court wanting to lend weight to its own statements. It is possible to get into more subtle difficulties, however.

Whitespace and structure

Even casual reading of the United States Code (and many other statutes) reveals numerous examples in which the layout of text on a page is closely related to the meaning of a statute. Typically these take the form of an introductory catchline, followed a list of specifics that pertain to it, followed by a paragraph providing further detail that modifies all or part of the preceding list of specifics. A look at 27 USC 203 as it appears on the LII web site[68] shows the following (sic):

(c) It shall be unlawful, except pursuant to a basic permit issued under this subchapter by the Secretary of the Treasury -

(1) to engage in the business of purchasing for resale at wholesale distilled spirits, wine, or malt beverages; or

(2) for any person so engaged to receive or to sell, offer or deliver for sale, contract to sell, or ship, in interstate or foreign commerce, directly or indirectly or through an affiliate, distilled spirits, wine, or malt beverages so purchased. This subsection shall take effect July 1, 1936.

This section shall not apply to any agency of a State or political subdivision thereof or any officer or employee of any such agency, and no such agency or officer or employee shall be required to obtain a basic permit under this subchapter.

As the text appears on our Web site, it is not instantly clear whether the paragraph break that is implied between "1936" and the words "this section" separates a paragraph that is applicable to item 2 only, as would be implied by the fact that it is indented to the same level as (2), or whether it applies to both (1) and (2), as would appear to be the case if it were indented at the same level as the leading catchline for (c). And while the word "section" would appear to tie it to (c), it is not at all clear that that is so, particularly if the indentation is off. In fact, this entire discussion could be rendered completely incomprehensible to the reader if the final, printed format of the block quotation above is not as I am seeing it now (struggles with my word processor as I created it leave me thinking that that is a strong possibility). Consultation with the original text before it passed through our formatting software indicates that the current layout is, in fact, an error.

An X-Rated Divertimento

Errors like this produce real effects. When a producer of pornographic films starring a then-underage Traci Lords was prosecuted under 18 USC 2252 - the item on the LII Web site that is most frequently linked to, never mind by whom[69] -- punctuation and layout were important factors. The majority opinion of the Supreme Court in *United States v. X-Citement Video*, 513 US 64 (1994), states:

The critical determination which we must make is whether the term "knowingly" in subsections (1) and (2) modifies the phrase "the use of a minor" in subsections (1)(A) and (2)(A). The most natural grammatical reading, adopted by the Ninth Circuit, suggests that the term "knowingly" modifies only the surrounding verbs: transports, ships, receives, distributes, or reproduces. Under this construction, the word "knowingly" would not modify the elements of the minority of the performers, or the sexually explicit nature of the material, because they are set forth in independent clauses separated by interruptive punctuation. But we do not think this is the end of the matter...[70]

What, we wonder, renders those clauses independent if not whitespace and punctuation, and who would be responsible for errors in interruptive punctuation, if not the publisher? Our concern here is not so much with the particular legal interpretation of publishing artifacts -- indeed, the opinion goes on to imply that concerns about punctuation and layout are trumped by considerations of meaning and effect [71]. What is important is that we realize that legal questions can turn on relatively subtle questions of textual markup and layout, so that mistakes by publishers creates the potential for costly litigation. It is the cost of determination rather than the outcome that is most critical to us as we think about future systems. Courts, legislatures, and agencies themselves have better reason to be wary of potential mistakes than third parties do, no matter whether those third parties are public or private. Creators are simply more averse to the consequences of error. As we pointed out in the section on reliability above, this is not a question of who is more careless, but of who has better incentive to correct those errors that will inevitably occur and the power to do so directly and quickly.

Third-party vs. self-publication

Subtly different from the issue of whether government is competent to publish at all is the question of whether each agency within government ought to publish its own works, or whether publication activities ought to be concentrated in a single entity.

Reliability concerns favor self-publication

Third-party commercial publishers are often held up as models of reliability, especially when they are compared to third-party operations on the Internet. As we have already learned, this is often more a matter of strong brand identification than anything else. In retrospect, it is unsurprising that legal professionals would want some proof that third-party operations having no official relationship with the creators of legal content and no existing brand name would provide accurate content. Over time these suspicions have eased. It is important to remember that they probably never applied to operations like GPO Access that are, in fact, directly operated by government. An informal and hasty survey of public-library web sites, for instance, shows that most favor "official" government sites in their lists of public information resources, even when third-party sites offer greater functionality and hugely improved response. And the 'horse's mouth' argument is persuasive: who is a more reliable provider of government information than the government itself?

Also, direct feedback from information consumers offers an efficient means for creators to improve information quality at source. Inaccuracies and errors can and do creep in at all stages of the publication process. The more quickly these can be corrected, and the closer to the ultimate source the correction occurs, the lower the ultimate cost of the error[72]. Experience teaches us that a user community acts in powerful ways through the immediate feedback mechanisms provided by the Net; that power can be efficiently harnessed to correct errors near the source, but only if the source is reachable.

Gains of third party centralization depend on the level of value added before transfer, and are self-limiting

Arguments in favor of a centralized publishing regime rest on questions of standardization, reliability, and efficiency. These three strands of argument share the presumption that a central publishing entity could use a set of common, internal tools and standards for information handling and storage, ultimately resulting in presentation of the information to the public in some uniform, standardized way. Thus, when we talk about centralization, we imagine that the centralized, third-party publisher can wrestle a diversity of incoming information streams, each presumably with its own idiosyncratic format, into a standardized product that is available to an audience that wants the benefits of common functionality. This activity is imagined to be efficient because the group involved is able to spread the costs of technology and of developing technological expertise across information streams produced by many creators, whereas without them each creator would have to develop publishing expertise on its own. Other presumed virtues include standardization, which speaks for itself, and reliability, which we suppose in this context to be a matter of placing all of one's eggs in one basket, the better to watch it.

"Standardization" is a simple label for what is really a complex bundle of practical problems for publishers generally and for electronic publishers in particular[73]2000 (1D). Different information creators may use different electronic file formats for their information, such as different word-processing programs, plain ASCII, HTML, or a database format. Some formats are less readily converted than others; the most familiar example of this is the Adobe Acrobat/PDF format, which is extremely difficult to convert to other formats but remains quite faithful in appearance to printed versions.[74] But transformation of a creator's work into something that conforms to a set of

standards is not merely a matter of electronically transforming one file format into another. There are issues of editorial conformance involved that go beyond the typographic to the structure of the documents themselves. To reliably convert a judicial opinion, for example, we must be able to reliably (and preferably automatically) determine what the names of the parties are, what the date of decision is, who the author is, where headings, major and minor sectional divisions fall, and so on. We do this not only for typographic purposes but so that we may appropriately tag metadata in our standardized version, extract text features for special treatment by a search engines and indexing software, cite it, and so on. Because courts and legislatures vary in the way they format materials, so must our text-conversion-and-conformance software vary from court to court, legislature to legislature, and agency to agency. Because the techniques on which such software depends are fundamentally matters of sophisticated pattern recognition, rather like a souped-up version of a word-processing search and replace function, conversion software tends to be very difficult to generalize beyond a single corpus[75].

Experience at the LII illustrates this point, though at times we rather wish it didn't. We are now on our third generation of software that does conversion for each of three flagship collections we maintain, including the opinions of the US Supreme Court, the United States Code, and the New York Court of Appeals - three times three or nine packages in all, written in the last five to seven years. Part of the reason that we change the software from time to time is that we want added functionality. But it has been uncontrollable change in the formats used by creators that has most urgently compelled us to develop successive versions. Courts and legislatures make improvements and changes too, and we must adapt our mechanisms to them when they do. Typically this involves extensive recoding and an awkward period of revision and debugging as we discover hidden variations in the new formats and new complexities that we did not encounter in whatever test data we had when we began coding. Retrospective projects are even more difficult. Like Schliemann digging in the ruins of Troy, we typically find that past history comes in layers. We see not one but several document structures that change from time to time to suit the whims of reporters of decisions, clerks of court, legislative publishing and printing systems, and other ephemeral actors. My aim here is not to catalog the gripes of a third-party publisher so much as it is to point out that, as with all encoding projects, there is much that needs to be encoded whose appearance and structure varies both from creator to creator and from year to year. In the world of conversion and standardization, this kind of diversity equals cost.[76]
<http://www.sims.berkeley.edu/~hal/Papers/publish.htm>[1]

Cost, in turn, affects the scale at which a centralized third party can operate. Each new, non-standardized corpus that the publisher takes on represents significant short- and long-term conversion costs. How tempting, then, to insist that all newcomers meet some sort of input data standard that will make the conversion job both easier and less expensive. In fact, this is what organizations like AustLII do[77], and I envy them their ability to do it. In the United States, where there are many jurisdictions and a tradition of judicial independence that at times could be better described as judicial contrariness[78], we do not enjoy such a happy state of cooperation with creators.

There are two morals to the story. First, there are intensely practical reasons why a centralized, third party approach may not scale well; it is expensive to convert diverse income streams into a common format and then into equally diverse products. Obviously, that suggests that the approach works better when the number of creators to be serviced is more limited, where an only-a-few-sizes-fit-all approach is appropriate, or where the inevitably-better-funded private sector is at work. More importantly, centralized, public-sector, third-party publishing operations with resource limits can only scale up to a certain point before they must compel a structure that is fundamentally indistinguishable from self-publication by creators. This is so because third-party publishers must insist on increasing levels of standardization in the data they are given in order to maintain operations as the scale and complexity of their input streams increase out of proportion to their resources. Ultimately the extent to which the input standard being sent by the creators so closely

reflects the practical needs of the output product that the creators might just as well be publishing the material directly.

This is not bad in itself. There is a useful time dimension at work. Even if things reach the state I have just described the creator will have bought itself some time and possibly some technological development and transfer by using the third party publisher as a kind of buffer against change. As we shall see later this sounds (and can work) much more like a process in which the third party educates the creating legislature, court, or agency and leads it through a process of technology transfer. This may well be a result that is less efficient in the short run (there is duplication of technological effort of the kind we remarked earlier in this section). On the other hand, the long term cost is lower because of other efficiencies inherent in self-publication, and because quality is increased.

But suppose that we artificially limit either the complexity of inputs or the diversity of products to be created in such a way that resources are not an issue, and we are never forced to compel standardization by creators. We will make resources available to a public third-party publisher at some reasonable level that meets a fixed level of need on the part of both the upstream creators and the downstream consumers. This sort of steady-state approach has a certain appeal, but it is balanced on a knife edge, and it closely resembles what we know already as government printing operations - externalized into the academic sector, perhaps, but still operating with a relatively fixed budget provided (we imagine) by government. First, it is unlikely that such a steady-state operation would be funded over the long term if it remained outside of government. Government would argue that it needed better control over the operation in order to guarantee its quality, and would then set it up as it were in-house. But even assuming that some solution sited outside government *were* adopted, much would depend on where our "reasonable level" of added value is set. If it is too low, the public standard will be increasingly irrelevant as the private sector moves to fill needs that the public standard does not meet. As a result, the central publisher will be de-funded and will spiral into nothingness as it more and more fails to meet expectations. Or it will protect itself from defunding by becoming more and more a service bureau designed to meet the needs of funding sources - for example, the legislature [79] <http://www.access.gpo.gov/public-affairs/5-99facts.htm> [1]. If the "reasonable level" is set too high, the operation will ultimately be accused of interfering with the private sector or of acting exactly like it - and it will then be defunded as a sop to the private sector [80]. I believe that publishing operations that are incorporated into legislatures, courts, and agencies would stand a better chance of avoiding these scenarios, if only because they are strongly identified with creators and can point at a public mission.

Centralized publication diminishes incentives for improvement in the process of drafting and creation

Let us return for a moment to the question of technological efficiency and centralization. We have already pointed out that a good part of the appeal of a centralized operation is that it concentrates the need for technological savvy and equipment in one place. This presumably means that individual courts and legislatures need not underwrite the continual expense of developing and maintaining technological expertise; the work at the cutting edge will be undertaken by the central operation and shared by everyone. Remember, however, that the third-party, centralized operation is only concerned with a part of the process: the part that distributes finalized (or mostly-finalized) rulings, statutes, or regulations from the creator to the public. They are not concerned with the processes of drafting or deliberation, and in fact handing over publication responsibilities to a third party at least partially disconnects those processes from that of publication. At worst, casting of an outside party in the role of innovator removes all incentive for innovation within the creating body. This is substantially the situation we have now in the United States, where reliance on private sector publishers for all aspects of publishing for many years removed most of the incentive to improve the electronic or print systems run by government itself [81].

What is the best environment for innovation?

Some would argue that this overreliance on the private sector happens not because government and other creators of legal information have abdicated responsibility for innovation but because large, bureaucratic organizations are incapable of it. There is a perception that the private sector is nimbler and better able to seize technological advantage[82]. Both theoretical and empirical work in industrial organization would paint a rather different picture:

A bit of monopoly power in the form of structural concentration is conducive to invention and innovation, particularly when advances in the relevant knowledge base occur slowly. But very high concentration has a favorable effect only in rare cases, and more often it is likely to retard progress by restricting the number of independent sources of initiative...Likewise, it seems important that barriers to new entry be kept at modest levels, and that established industry members be exposed continually to the threat of entry by technically audacious newcomers...What is needed for rapid technical progress is a subtle blend of competition and monopoly, with more emphasis in general on the former than on the latter, and with the role of monopolistic elements diminishing when rich technological opportunities exist[83].

While words like "industrial" and "competition" would seem to flag this as an argument that applies only to private-sector actors, it can equally well be applied to public-sector entities in a very commonsensical way. There is nothing that constrains "independent sources of initiative" to the private sector. And while it may at first seem odd to think of government agencies as being in competition, there is reason to believe that forces strongly resembling competition can be at work as government bodies strive to communicate whatever they imagine their message to be. Legislatures like voters to see what they are up to. Agencies want to be seen as relevant and customer-oriented. Courts, particularly courts of first impression, want their work to be seen and to be respected. All are involved to some degree in competition for public resources. And local governments compete with one another: consider, for example, the jurisdiction that wants to use relaxed regulation as an incentive for private industry to locate there, as opposed to a neighboring jurisdiction. Individual American states often do this using relaxed tax or environmental regimes as their incentives; some have made minor industries out of particular parts of their law, as Delaware has with its requirements for incorporation. Very often this is reflected in their web sites, and held up as an example of red-tape elimination or government streamlining[84]<http://corporate-law.widener.edu/ctofchan.htm>[m]<http://www.state.la.us/state/business.htm>[m]<http://www.state.ri.us/bus/frststp.htm>.

Distributed models depend crucially on a painful process of standards development

Thus far, the arguments arrayed on both sides seem to favor a distributed model - if there were but infinite time to develop it. In order to have a distributed model that supports common interfaces and capabilities we need to formulate workable standards. This is challenging to say the least, and the author yields to no-one in saying that it is a difficult process riddled with extraneous political concerns, vulnerable to the manipulations of vendors and others seeking advantage, and often characterized by pointless hairsplitting. Unfortunately, it is also necessary, and it is important that it include a broad international community.

One might argue, as many have over the years in other contexts, that a *de facto* standard provided by one or another industry leader would do as well as one arrived at through a relatively open and representative process[85]http://www.zdnet.com/anchordesk/story/story_4376.htm[1]. To be sure, there must always be leaders and to the extent that there are leaders any open process will take on some of their biases - witness the dominance of English as the language of the Internet and as an underlying assumption of most search-engine technology, or of the United States in the domain-naming scheme. But it is a key realization of organizations like the W3C that organizations that

formulate broad standards must also be as inclusive as possible[86]. Inclusive processes are necessarily slow ones, and that in turn becomes the disadvantage of a standards-based distributed system. It is hard to get a given level of functionality from it as quickly as one might with a centrally administered one. Acceptance of the standard ultimately depends on the availability of standards-based applications that users actually want, and these may be slow in coming.[87] In a world where technology changes rapidly, and where public bodies are already making decisions about what to do, this may be a critical disadvantage.

What can we do, practically?

The multifaceted discussion that has led us to this point is far from dispositive. Here are the two halves of the onion:

* On the one hand, there is good reason to believe that a standards-based distributed system would be superior to a centralized one. In complicated venues like the United States, it may be the only one that will answer the problem. Indeed, as more and more localized creators of legal information (like municipalities) come online, it may be the only one that would answer the problem in *any* venue, as almost all are complex once local law is taken into consideration.

* On the other hand, such a system is undesirable because it is impossibly slow to implement and because a centralized system can perhaps for a time be more efficient and less costly. In the rapidly-shifting world of the Internet, slow-to-implement may mean never-to-be-implemented.

But these are not strict alternatives, and thinking about time points the way toward our conclusion. Most of the disadvantages of a centralized mechanism show up in the longer term and in an environment of complexity. For that reason it is well adapted to be a short-term solution. It can provide both proof-of-concept and a rallying point that will create the political will necessary for longer-term efforts. It can buy time for standards development. Longer-term efforts should be aimed at creating a decentralized model, because it enjoys better scalability and certain quality advantages, as well as holding government to an important responsibility it owes the public. Thus, the two models are not competitors but successors-in-time.

What do short-term and long-term mean, practically speaking, and how would we know when to make the change from one model to another? The critical factors involved in making that decision are, I think, set out in the discussion of self-limiting technical processes above, conditioned somewhat by the availability of technical expertise to creators. At the point where the standardization requirements that the centralized public system imposes on a creator become as onerous as self-publication, it is probably past time for that creator to stand on its own two feet as a publisher. How far it is past time is a good question. The ideal point for abandonment by the central operation will probably differ in every case as a product of workloads, available expertise, size of the corpus being transferred, and so on. It is also the case that the centralized operation can accelerate its abandonment of creators to their own devices as the independent creator community grows and becomes a source of mutual support. It will not be easy to determine the point at which independence should be granted, but it must be done, and it must be planned for from the beginning.

[*] The author is co-director of the Legal Information Institute at the Cornell Law School. He wishes to thank Genie Tyburski, Daniel Poulin, Abdul Paliwala, and as always Peter Martin for feedback and discussion at formative stages.

[1] Apologists for the system consistently point to its comprehensiveness while conveniently ignoring the cost factor. See, eg., Robert C. Berring, "On Not Throwing the Baby: Planning the Future of Legal Information", 83 *Cal. L. Rev.* 615, at 618 (1995).

[2] See generally Vaver, David. "Copyright and the State in Canada and the United States." *Crown Copyright in Cyberspace Conference*. Montréal: Université de Montréal, Faculté de Droit, Centre de Recherche en Droit Public, 1995, online at [

3] The market structure is well-described in Henry Perritt, Jr., "Market Structures for Electronic Publishing and Electronic Contracting on a National Research and Education Network: Defining Added Value", in *Building Information Infrastructure*, Brian Kahin, ed., McGraw-Hill Primis 1992, 344-402.

[4] See Perritt, *supra* note 3; Bruce, T., "Public Legal Information: Focus and Future", [] *The Journal of Information, Law and Technology (JILT)*, online at [, and Greenleaf G et al, "[The AustLII Papers - New Directions in Law via the Internet](#)",] *The Journal of Information, Law and Technology (JILT)*, online at http://elj.warwick.ac.uk/jilt/leginfo/97_2gree/ .Greenleaf is somewhat more aggressive than either Perritt or myself, wanting the public providers to establish standards of value which the private sector must then exceed in order to attract a market for their products.

[5] See, eg., Chung P et al, "A Defence of Plain HTML for Law:AustLII's Approach to Standards", [.] *The Journal of Information, Law and Technology (JILT)*. Online at <http://www.law.warwick.ac.uk/jilt/00-1/chung.html> .

[6] Paliwala A, "Access to Legal Information - A welcome return?" , [] *The Journal of Information, Law and Technology (JILT)*. Online at <http://www.law.warwick.ac.uk/jilt/00-1/paliwala.html>

[7] Bruce T, "Public Legal Information: Focus and Future" , [] *The Journal of Information, Law and Technology (JILT)*. Online at <http://www.law.warwick.ac.uk/jilt/00-1/bruce.html>

[8] A particularly succinct and useful description of the history of this proposition is to be found in Perritt, Henry H., "Should Local Governments Sell Local Spatial Databases Through State Monopolies?", online at [. [There are many interesting analogies to be drawn between the publication and sale of GIS data and that of black-letter law. Consider this passage:](#)]

Geographic information is an especially valuable type of public information. Land developers, those concerned with siting environmentally sensitive facilities, and public utilities pay substantial sums for integrated spatial databases. In many cases, profit-seeking entities collect much of the data necessary to an integrated spatial database. At the same time, government agencies, including those charged with collecting and maintaining geographic information like land records and survey information, face stringent budget constraints. As a result, it is natural for public agencies to suppose that they can ease their budget pressures and serve their publics better by appropriating some of the potential revenue stream; they can sell their information. Beyond that, it is natural for them to suppose that the quality of results and perhaps also the size of revenue stream can be increased by "partnerships" with private entities.

Unfortunately, this is but a short step away from imposing restrictions on what other vendors and distribution channels can do. Most public agencies responsible for geographic information have either a natural or de jure monopoly on the information. Monopolists perceive that they can increase their total revenue stream by setting prices higher than they would be in a competitive market. Monopolists also are tempted to extend their monopolies into downstream markets. Thus, public agency decisionmakers, behaving like rational monopolists in the private sector, implement their partnership aspiration by prohibiting private sector competition with their chosen partners. The result is a state monopoly that limits economic and technological benefits to a broad range of potential distributors of the public information. And, as the monopolies are extended downstream by exclusive "partnerships," they block competition in a variety of rapidly changing and diverse markets for value-added information products.

There may be theoretical circumstances of economies of scale and other determinants of natural monopoly in which such arrangements would be economically efficient. However, such situations are rare and difficult to identify because the pace of technology and the fragmentation of the market tend to result in many vendors specializing in products or particular market segments. Moreover, there is no reason to suppose that public decisionmakers are better than consumers and entrepreneurs in picking technologies and product design; yet, that is exactly what they do to set up exclusive arrangements. The best market structure is one in which everyone is free to follow his or her instincts in commercializing new technologies and developing markets. The best information policy is one with a diversity of channels and sources for geographic information. Happily, that information policy is one that coincides with legal entitlements to access.

[9] See the useful summary in James Wyman, "Freeing the Law: Case Reporter Copyright and the Universal Citation System", 24 *Fla. St. U. L. Rev.* 217, online at [

[10] N.Y. Judiciary Law §434 governs these contracts in New York State; the exact terms of the contracts themselves are confidential.

[11] See, for example, the rather confusing list of critiques at [[A less confused assessment that fails equally as prophecy can be found in Cary Griffith, "Internet Yields Inadequate Waves for In-House Surfers"](#)], *Corporate Legal Times*, August 1995.

[12] Indeed, neither Westlaw nor Lexis are themselves monoliths. They are branded federations of distributed collections, and this cellular, federated structure is perpetuating itself through merger and consolidation in the legal publishing industry.

[13] See [[available on the Web at](#)], and Paliwala A et al, 'User Needs in Electronic Law Reporting: A Research Study of The Law Reports', [*The Journal of Information, Law and Technology (JILT)*], online at http://elj.warwick.ac.uk/jilt/leginfo/97_2pal/, summarizing and discussing an earlier study by Blair and Maron.

[14] Woxland, Thomas A. and Patti J. Ogden, *Landmarks In American Legal Publishing*. Exhibit Catalog. Minneapolis: University of Minnesota Law Library (ca. 1996) p.38. As with most materials related to the history of American legal publishing, the catalog and its exhibit received support from a legal publisher (in this case the West Publishing Company).

[15] John B. West, "A Symposium of Law Publishers", 23 *Am. L. Rev.* 396, 406-407 (1889).

[16] Computer manufacturers in the United States have long been accused of using so-called "FUD" tactics (practices preying on Fear, Uncertainty, and Doubt) to market computer products and standards to hapless professionals. They did not invent the practice. At this point it would probably be impossible to come to any sort of agreement about what a reasonable level of comprehensiveness (or conversely, selectivity) might be - though interestingly publishers seem to be finding discrete subsets of legal information they can market on CDROM.

[17] A comprehensive and detailed how-to on this point is available in Kendall F. Svengalis, *Legal Information Buyer's Guide and Reference Manual 1997-98*, Barrington, R.I. : Rhode Island LawPress (1996).

[18] It was and is difficult to imagine a lawyer running into a crowded courtroom in Oklahoma and shouting, "Wait! The Supreme Court has just ruled that..." after the fashion of last-minute television-drama denouement. Apparently the commercial online vendors felt that if a need for that level of timeliness did not exist, it could be created.

[19] See, eg., recent opinions of the Alaska Supreme Court, online at [[or the Colorado Supreme](#)

[Court at .](#)

[20] The Supreme Court of California provides an example; see <http://www.courtinfo.ca.gov/opinions/> .

[21] Advocates for the poor and other purveyors of personal legal services rely on this sort of informal network quite extensively, and there are many among them who advocate the extension of these networks via the Internet. One wonders if this is because they have been so poorly served by other legal-information services in the past. See, eg., *Victim Services Report: The Newsletter of Victim Services*, 8 No. 2 (Fall/Winter 1997), online at [L](#) .

[22] Indeed, there has been remarkably little objective assessment of quality factors among the online commercial publishers. The only attempt that I know of created such rancor and bad behavior from the vendors themselves that the publication dropped it.

[23] Bruce Schneier, *Applied Cryptography: Protocols, Algorithms, and Source Code in C*, New York: John Wiley and Sons(1994). 34-39

[24] See for example Bruce, "The Internet and Legal Information: Projects and Prospects", in *Les Autoroutes Electroniques: usage, droit, et promesses*. Quebec, Canada:Editions Yvon Blais (1995)

[25] See, e.g., Stephen Johnson, "The Internet Changes Everything: Revolutionizing Public Participation and Access to Government Information Through the Internet", preprint online at [L](#) , or Bruce, "[Legal Information, Open Models, and Current Practice](#)", online at [L](#) .

[26] Travis, Brian, and Dale Waladt, *The SGML Implementation Guide, A Blueprint for SGML Migration*. New York: Springer-Verlag (1995). The case study of the TIGRE system appears at pp. 371-384.

[27] Notable efforts with legal information have taken place in Tasmania, and in Canada. State level efforts are underway in the US but thus far the US remains remarkably innocent of SGML in the production of Federal statutes, regulations, and judicial opinions despite the fact that SGML is a Federal Procurement standard. See in particular Arnold-Moore T et al, "Connected to the Law: Tasmanian Legislation Using EnAct", [L](#) . *The Journal of Information, Law and Technology (JILT)*. [\[An early report of Canadian efforts is given in the Seybold Report at . Finally, some state-level projects in the US are reported by the National Center for State Courts at . Thus far most SGML and XML efforts at the state level have been concerned with filing and other matters of court administration rather than with the dissemination of statutes, regulations, and judicial opinions. But of course the most extensive government use of SGML in the United States is in the arena of defense procurement and systems, where it is ubiquitous.](#)

[28] See Hammer, Michael, and James Champy, *Reengineering the Corporation: A Manifesto for Business Revolution*. New York: HarperCollins (1993), pp. 36-39 and 202.

[29] Berring, *supra* note 1, at 621.

[30] It is notoriously difficult to wrestle Web server log files into accurate profiles of an audience. Some rather unscientific tallying of our own, along with the rather better information we get from subscribers to our current-awareness services, leads us to believe that the non-lawyer professional audience accounts for better than half of the more than a million hits the LII site receives daily.

[31] New York State Judiciary Law §434, online at <http://assembly.state.ny.us/cgi-bin/claws?law=53&art=27>

[32] This figure is derived from an AltaVista survey made in September of 1999, online at [\[\]](#); [projections would indicate that the total could now be as high as 45%](#).

[33] Marc Galanter, "The Legal Malaise, or Justice Observed," 19 *Law and Society Review* 537, 545 (1985).

[34] See, eg., [\[\]](#), [a consolidation of state and Federal water laws by the California Water Control Board, or the regulations covering Federal grants to research institutions consolidated at](#) .

[35] Travis and Waldt, *supra* note 26, at 372

[36] Perritt (*supra* note 8) remarks that

"...if the government gives or sells public information below cost, it may undermine market opportunities for private vendors. This is undesirable, because any system of dissemination of public information depends to some degree on the private sector. If the government eliminates competing sources of public information, it raises the price and creates the possibility that the government will control knowledge of its operations."

Perritt assumes (as he has elsewhere) that everything that is not done by the government is done in the private sector. Thus he either ignores the existence of public actors like AustLII or the LII or implies that they closely resemble private-sector activity with a very low price point - a proposition I would agree with.

[37] Government bodies that produce descriptive information like labor statistics seem to have less trouble with this idea. See, e.g., the diversity of reports and alternate information structures available from the US Census Bureau at [\[\]](#). [For an example that is closer to our concerns here, see the Social Security Administration site at , which includes what is essentially regulatory information in Spanish as well as a specialized presentation for children.](#)

[38] Having seen the LII characterized many times in print in ways that conspicuously fail to capture all of its phases and aspects I am suspicious of my own ability to characterize another, equally complex operation. The notions put forward here are gleaned from the examination of materials at the AustLII site (<http://www.austlii.edu.au/austlii/>) and those published by AustLII principals in various venues over the years, including *The AustLII Papers* (online at [\[\]](#)) and [the AustLII technical roadmap, most recently published at](#) .

[39] A convoluted and detailed vision of such a regime is put forward in Bruce, T, "Public Legal Information: Focus and Future", [\[\]](#)*The Journal of Information, Law and Technology (JILT)*, online at [\[\]](#).

[40] The W3C is the WorldWideWeb Consortium housed at the Massachusetts Institute of Technology and administered cooperatively by a troika of universities. See [\[\]](#).

[41] I say this in complete ignorance of how something like a state-owned enterprise might function in this context, a notion that might well be worth pursuing. Recent experience with the US Postal Service would indicate that such a structure might well provide a good level of universal service while holding its own against private competition in niche markets.

[42] It is a great shame that there is no hard data to be had on price sensitivity and demand elasticity in non-lawyer professional markets for legal information. Indeed, there is precious little to be had on the lawyer market for legal information, beyond what one might deduce from the practices of the large commercial vendors and their lower-priced, Internet competitors. Legal information has for so long been treated as a commodity for which there is inelastic demand at the high end of the market

that we have little to go by. One might hope that someone with sufficient curiosity and perseverance to read the footnotes in obscure tracts on the structuring of public legal information systems would be interested in doing some empirical research.

[43] In some settings Internet distribution might better be considered as the foundation of a print-on-demand system. The needs of areas like sub-Saharan Africa, where the information pipes leading in and out of the country are thin, and local telecommunications infrastructure is poor, would argue for Internet delivery of information to a central point from which wider distribution takes place in print. The same may be true of local communities or even neighborhoods where access to bandwidth is low but where logical agents for print distribution exist.

[44] Berring, *supra* note 1, at 627.

[45] Katsh presents a portrait of law libraries as exclusionary institutions separating the body of legal knowledge from other areas of human endeavor. See Ethan Katsh, *The Electronic Media and the Transformation of Law*, Oxford : Oxford University Press (1989) at 215.

[46] It is equally intriguing to consider the degree to which most law librarians see the category "non-lawyers doing legal research" as equivalent to "pro se patron representing herself in a formal proceeding". The author has somewhat cynically remarked that a third category, "homeless derelicts shambling about the reading room", is often seen to be the equivalent of the first two. Note, for example, the perception implied by the inclusion of law library security as a topic in a session on handling the needs of pro se patrons conducted by the Southwest Association of Law Libraries (at D

[47] So-called "plain English" legislation serves a similar purpose in providing the public with intellectual access to the law. Interestingly, the quieter, self-imposed initiatives undertaken by agencies such as the Social Security Administration in rewriting their own regulations have probably been more effective.

[48] See Cox, W. Michael and Richard Alm, "The Right Stuff: America's Move to Mass Customization", in the Federal Reserve Bank of Dallas Annual Report (1998), online at <http://www.dallasfed.org/htm/pubs/annual/arpt98.html>.

[49] It may be that the commercial services in the United States are getting something of an education in this respect. For years they have added features that had little value outside the attainment of momentary advantage over a sole competitor. As the market has opened up to more sellers and as various circumstances conspire to create a more price-sensitive buyer, this kind of duopolistic shot-trading has become less and less useful.

[50] Greenleaf, *supra* note 4.

[51] Perritt, *supra* note 26.

[52] The GPO updates the Code on a continuous, rolling basis that takes about a year to run the cycle of all 50 titles. At any given moment the most out-of-date title will be nearly two years old; the least about a year old. Updates to the public version occur in batches of 3 to 6 titles; so far as we know the schedule is neither regular nor published in advance.

[53] It is John B. West's invention of the reporter system as a more rapid alternative to official reports that is most often held out in support of this claim, and indeed West did provide a valuable innovation that responded to deficiencies in the official reporting system. But this does not mean that government has always been slower, or private publishers faster. In the 1840s, for instance, Congress found it necessary to introduce amendments that would force newspapers into timely reporting of legislation (5 *Stat. L.* 527); much earlier, it had been necessary to subsidize private printers in order

to get them to take on the printing of laws at all (1 *Stat. L.* 724). Perhaps the most disturbing thing about the historical arguments brought to bear on this point is a distinct feeling that history is being written by the victors; it is difficult to find an historical study of American legal publishing in which the West Publishing Company is not somehow involved as sponsor or publisher. This is not to diminish John West's innovation, but it would be a mistake to believe that the conditions that produced it are an unalterable and eternal reality.

[54] I speak here of the United States Postal Service, which has improved enormously in the face of competition from private parcel services and is in fact more than holding its own. Just how radical an improvement this is can be seen by comparing the customer-satisfaction survey at [\[with the opinions expressed in the Berring article quoted below, which holds out the Postal Service as an undesirable example of government inefficiency.](#)

[55] Berring, *supra* note 1, at 621 (footnotes omitted).

[56] Print can be quite fast, too, as the speedy release of the Starr Report on the independent counsel's investigation of President Clinton amply demonstrates. See Christopher Yeich, "Print's wild week; fast release of major news in print and electronic news formats" in *Graphic Arts Monthly*, October, 1998.

[57] See, eg., the Minnesota state statutes, Web-published at <http://www.revisor.leg.state.mn.us/stats/> , or any of the other resources mentioned in the so-called "LII Honor Roll of State Law on the Internet" at <http://wwwsecure.law.cornell.edu/background/states/> .

[58] The system in place in Tasmania is arguably the best available in the world today. See Arnold-Moore T et al, "Connected to the Law: Tasmanian Legislation Using EnAct", [\[.\]The Journal of Information, Law and Technology \(JILT\)](#), online at <http://www.law.warwick.ac.uk/jilt/00-1/arnold.html> .

[59] Indeed the history of the EnAct project in Tasmania as described by Arnold-Moore (see note *supra*) resembles nothing so much as a re-engineering case study, and it would be interesting to see it fully treated in that manner.

[60] 1 Conn. Rep. Xxviii.

[61] West Group is currently a \$1.3 billion business. Simba, Inc. *Electronic Publishing Report*, issue of November 13, 1998.

[62] Of course, fiscal capital is not the only form of capital, and it might remain desirable to concentrate intellectual capital in the form of expertise about electronic systems in a centralized operation or in some other locus outside government. We discuss this issue later.

[63] See Richard Delgado and Jean Stefancic, "Why Do We Tell the Same Stories?: Law Reform, Critical Librarianship, and the Triple Helix Dilemma", 42 *Stanford L. Rev.* 207 (1989), and M. Ethan Katsh, *The Electronic Media and the Transformation of Law* (1989), pp. 218-226.

[64] 2A Norman J. Singer, *Sutherland Statutory Construction* § 47.14 (5th ed. 1992).

[65] As it did in *United States v. Ron Pair Enters., Inc.*, 489 U.S. 235 (1989).

[66] See Singer, *supra* note 64.

[67] For some indicators of what these are likely to be like, see the DTDs published by LexUM at

###

[68] See <http://www4.law.cornell.edu/uscode/27/203.htm>

[69] The author is quick to assert that he comes by this information through the reading of statistical reports on HTTP referrer activity, and not by direct examination of primary sources.

[70] *United States v. X-Citement Video*, 513 US 64 (1994)

[71] This is ordinarily the case when statutory interpretation turns on questions of punctuation or arrangement. See Henry Campbell Black, *Construction And Interpretation Of The Laws* (2d ed. 1911), §§ 86-88.

[72] I suspect that there is an analogy to be made to software-quality engineering, in which the cost of repairing defects increases exponentially as the program moves from testing into production.

[73] For a rather complete statement of AustLII's process, see Austin D et al, "Scalability of Web Resources for Law: AustLII's Technical Roadmap: Past, Present and Future", [[The Journal of Information, Law and Technology \(JILT\)](http://www.law.warwick.ac.uk/jilt/00-1/austin.html)]. <http://www.law.warwick.ac.uk/jilt/00-1/austin.html>.

[74] This faithfulness to print won PDF a foothold with agencies like the IRS that use it primarily to distribute tax forms and other fixed-format materials. This is something that PDF is very good at. Unfortunately, this has been taken to mean that it is generally virtuous as an electronic format where in fact it is far from it.

[75] See generally Travis and Waldt, *supra* note 26, pp. 151-182.

[76] See Hal Varian, "The Future of Electronic Journals", online at [[for comments based on experience with journals that accept multiple electronic submission formats - a context that is in some ways highly suggestive of the changeable nature of judicial output over time.](#)]

[77] See Austin, *supra* note 72.

[78] A recent project the LII undertook for the New York State Court of Claims revealed seventeen different formats for opinions written by seventeen judges, each of whom was a passionate advocate for his or her own method.

[79] This is more or less the situation with the US Government Printing Office at present. See their mission statement, online at [[. It speaks of their mission to serve the Federal community, and says not one word about the public.](#)]

[80] This entire set of arguments is currently being played out in controversy surrounding the potential decommissioning of the National Technical Information Service (NTIS) in the United States. *Hearings of the Technology Subcommittee of the House Science Committee*, Sept. 14, 1999.

[81] Well-documented proof of this assertion is hard to come by, primarily because comparative analyses of public and private resources are rare except in the complaints of law librarians and information-retrieval researchers, neither of whom have been fond of the official systems. Perhaps the best evidence that can be marshalled on this point comes from the abandonment of citation to official state reporters in the last edition of the Bluebook, based on complaints that they were unavailable and generally inferior. If the reader is beginning to suspect that there is a certain circularity about all this - official reports are bad because more resources go to privately published reports, which then gain more resources by asserting that the official reports are bad - then the reader is beginning to get the picture.

[82] See, e.g., Berring, *supra* note 1, at 620. Contrasting sharply with this rosy picture is substantial if undocumented dissatisfaction on the part of information-retrieval researchers. They complain that in reality commercial online publishers have been uninterested in research and have devoted far more substantial resources to the expansion of collections than to improving information quality and the technology underlying their services.

[83] F.M. Scherer, *Industrial Market Structure and Economic Performance*, 2nd. Ed. Chicago: Rand McNally (1980) p. 438.

[84] See, eg. the Delaware Court of Chancery site, [[interestingly enough operated by an academic law library](#)], [Louisiana Business and License Information at](#) , [the Rhode Island "First Stop Business Center" at](#)

[85] Unsurprisingly, it tends to be industry leaders and other quasi-monopolists who argue this. Things are even more fun when two industry leaders or quasi-monopolists dispute territory. See Jesse Berst, "Why Sun and Microsoft Won't Let You Have A Wired Home", available via ZDNet at [[.](#)

[86] An interesting view of how W3C actually thinks about these matters is provided at <http://www.w3.org/Consortium/Process/Process-19991111/>.

[87] There is no better illustration of this proposition than the GOSIP standards process, which produced very few implementations during its 14-year history. An entertaining exposition of its failures and those of standards processes generally is in Marshall T. Rose, *The Little Black Book: Mail Bonding with OSI Directory Services*. Englewood Cliffs: Prentice-Hall (1992). Discussion of the standards process is at pp. 290-301, under the heading " A Long Dark Night for Open Systems".