Peter W. Martin

Meiji University - 19 May 2001

The Legal Information Institute (LII) - Providing Catalysis, Innovation, and Integration in a Complex Legal Information Environment

{slide}

I. The LII - From First (1992) to One Among Many (2001)

{slide}For over eight years, the Cornell Legal Information Institute (LII) has been engaged in the electronic dissemination of legal information – on disk and via the Internet. Over those years, no more than a brief moment in the history of a university, the scale, complexity and ambition of our activities have expanded beyond anything we imagined in 1992 when my collaborator, Thomas Bruce, and I founded the institute. Dramatic though those changes have been, they are small compared to the total transformation of the legal information environment that has taken place in the United States and elsewhere during the same period of time. We have been, I believe, and will continue to be a key player in that transformation. However, we are hardly its most important feature. Consequently, this paper is not only about the Legal Information Institute but also about the larger context. It seeks to trace the understandings we have gained about the importance of broad and effective access to legal information from our distinctive vantage point.

A core founding and sustaining principle of our institute is that a university-based, non-commercial activity has an important role to play both in exploring new modes of education and in extending public access to legal information. Central to that role is a sustained program of applied research on how digital technology can be used to achieve those closely related aims.

At the beginning, we stood alone. Our institute ran the first Net server focusing on a discipline outside the physical sciences (initially a gopher). We created and released as freeware the first Web browser to run under Windows (Cello) – a necessary step in those early days toward providing effective hypertext access to law via the Internet. It is startling to realize how different those times were. In 1993 the LII's original Web server held a hypertext version of the U.S. Constitution, an HTML front-end to one hundred or so Supreme Court decisions at another university's ftp site, the Uniform Commercial Code and a few federal statutes – all created in HTML 1.0 by hand mark-up. More a proof of concept undertaking than a resource for serious researchers the LII site responded to a few hundred data requests a week. At that time our disk-based publications for law students drew far more attention and use. They contained the core codes for a number of important law school courses in a rich hyper-linked and searchable form and were appreciated by computer savvy law students, although only rarely by those who taught them.

Flash forward to today. At present, Cornell's Legal Information Institute runs the most heavily used non-commercial, comprehensive law site in the United States. We operate an array of servers that respond to far more than a million data requests a day, representing tens of thousands, sometimes hundreds of thousands of user sessions. (And

neither figure accounts for the traffic at our mirror site in Europe.) On days when the Supreme Court releases decisions, summaries linked to the opinions in full text are dispatched via e-mail to over 20,000 initial recipients of our free electronic bulletin. Since we encourage redistribution, we have no idea how many individuals are, in the end, reached by this free service. Needless to say, the audience is much larger than and quite different from that reached by the Cornell Law Review and the other two print journals published by our law school. The institute also produces CD-ROMs and downloadable course materials and has, for five years, offered law courses over the Internet to students at a growing number of other U.S. law schools.

Pioneers do not necessarily survive; being first has as many hazards as advantages. A year ago some didn't believe this to be true of the Internet. Today, they know otherwise. I am convinced that the Legal Information Institute continues to thrive and grow because of important strategic decisions made initially and in the years since 1992. As those years have seen enormous changes in the environment surrounding our activity, the key decisions have been subject to frequent revisiting.

{slide}Let me list a few of the more important ones:

That the institute should remain non-commercial and based at Cornell University

Our institute and its principals have faced and resisted numerous opportunities to exchange the commitment to research and non-commercial public access for economic gain. While other Internet projects that began in American universities have, during the Internet explosion, moved to some commercial form, often with large profit to their

founders, we have held to our original non-commercial path. We have also taken pains to avoid individual or institutional partnerships with commercial publishers that posed serious risk of compromising our commercial neutrality. In our setting that meant rejecting special relationships with Westlaw and LEXIS. On the other hand, we have been quite willing to draw revenue from the commercial sector through data and software licensing or consulting. Our institutional setting has given us access to a wide range of expertise, linked our program to deeply held values of discovery and public service, and insulated our work from the direct effects of political and market forces.

That a centralized comprehensive collection was, in the U.S. environment, not an attainable goal

Our founding vision went well beyond a shift in the law school's support for publication (an activity in which it had long engaged) to a new, digital form. We intended for our institute to become itself a center of serious research on how digital technology might be used to improve access to legal information and education. Our research in this area has, from the start, been applied or experimental, rather than purely theoretical. We have built a succession of new products and services designed to be useful to a variety of constituencies, both familiar and new. That led to early confusion about our aims. Commercial publishers imagined us to be a competitor, when instead we were simply providing an advance look at technology applications and forms of information diffusion that were destined to become widespread.

While our research has conspicuously involved several high use, test-bed collections (notably the decisions of the U.S. Supreme Court, decisions of New York's highest court, the procedural rules of the federal courts, and the compilation of federal

legislation known as the U.S. Code), we have never imagined ourselves building or sustaining a comprehensive collection of federal law materials, let alone the legal materials from all fifty states. A portal site - "yes." A comprehensive law data warehouse - "no."

The evident scale and decentralization of the U.S. legal system, combined with the firmly established market presence of commercial legal information vendors saved us from any delusion that we might be a non-commercial LEXIS or Westlaw. We have been extremely careful not to undertake more than we could maintain and continue to develop.

Our aim has been to influence not own or control. Consequently, we find gratifying evidence of our success in the numerous legal Web sites, of all kinds, that embody elements of format and functionality that we originated. Since the available technologies and the reachable user base have been changing at unprecedented speeds, our efforts to work effectively with law content at their intersection have been stimulating, influential, and some days overwhelming.

That the explosion of legal information sources of all types on the Internet represented fresh opportunity rather than a diminished role

We have held to the view that there is an important role for academically-based activities like ours, even as the Internet has become the dominant delivery path for all commercial legal information providers in the U.S., old and new, and as public bodies have begun in growing numbers to use the Internet to provide free public access to the

law for which they are responsible. While Westlaw and LEXIS have brought their comprehensive and integrated collections to the Net, where they compete with LOISLAW, recently acquired by Wolters Kluwer, all are surrounded by fee or other barriers that cut off large and important segments of the public and severely limit innovation in both information delivery and education.

The proliferation of public sites – hosted by or working with courts, legislatures, administrative agencies, state and city governments – has at the same time created the potential for a truly open, distributed, public information system. But this remains a potential, not an actuality. It is an essential but not a sufficient condition for free and widely accessible legal information.

Although we now operate in a crowded field, that means more to do not less and more difficult choices about priorities than when we stood alone.

That the distinct contribution an activity like the LII can make in the complex U.S. legal information environment is as catalyst (innovating, leading through example) and integrator

Several years ago, a public spirited group of American law school librarians, technology people, and others gathered at the Georgetown Law Center to explore ways of bringing the decisions of the U.S. Court of Appeals to the Internet. This court, which is divided into thirteen different units, called "circuits," resolves all appeals that arise in the American federal court system, that are not subsequently dealt with by the Supreme Court. As the Supreme Court takes very few cases a year, the final interpretation on many important points of law falls to the Court of Appeals.

At the time of the Georgetown meeting, two schools, the University of Texas and Emory University, had already begun to distribute the decisions of the circuits for their regions on the Internet. Other schools at this meeting quickly volunteered to distribute the rest. Our institute was not tempted either by the entire project or any of its obvious pieces. First, we were certain the scale exceeded anything we needed for research. We were already working with the 75-80 decisions a year of the U.S. Supreme Court and the 200 or so of New York's highest court. Indeed, the scale and lack of data consistency across the thirteen circuits placed any such ambition beyond our reach. The annual output of the entire court exceeds 25,000 decisions and while all of the circuits must interpret and apply the same national law, each jealously guards its autonomy on such matters as data systems, decision format, court procedures, and schedule.

Observing, however, that this distributed federal law collection would need integration the LII built a cross-site full text index – to enable users to search for decisions dealing with particular topics of federal law without having to visit multiple sites and master the idiosyncrasies of diverse search engines.

The good news is that in the years since we undertook this project all but one of these federal courts have established their own servers (leading a number of the original law school intermediaries to drop this service). Regrettably, though predictably, these public sites have not been designed to facilitate cross-site linking or indexing and in that respect they are less useful than their academic precursors. These units of the same court, which cannot coordinate their schedules for recruiting law clerks or any of numerous other details of carrying out their parallel tasks, have each contracted for decision database services with little regard for the interests of those seeking to access and read

their decisions, let alone those seeking to integrate their work product with that of other circuits. As a result, the LII's role as example and integrator has become more important rather than less.

There is in the U.S. no public body with the responsibility of coordinating the distribution of judgments from the full range of federal courts. The judges themselves and their clerks all use commercially distributed legal information and so have little personal stake or insight into the serious limitations in the manner in which their respective courts have implemented the public access ideal in this new digital environment. Our institute's search engine, rebuilt only a month ago to deal with the idiosyncrasies of some of the new public sites, is the only means by which these separate collections which collectively reach back over 5 years, exists as a single resource on such key legal topics as copyright, civil rights, labor law and federal securities and banking law. Integrated with the decisions of the Supreme Court on one side and the U.S. Code on the other, both resources we maintain, they become part of a strong federal law library.

And finally, that key to future leadership in these ways is the collection of human and information resources assembled at our university and our deep experience with education

We established our institute in 1992 with the conviction that digital technology should facilitate a quantum shift in the distribution of legal information and also make it possible for a university law school to become a serious electronic publisher of its own research. To explain the venture to colleagues and alumni we analogized its aims to those that prompted Cornell to establish its first law journal in 1915 and two additional ones in

later years. Journals like these, we pointed out, were costly. In light of the school's purposes for producing them they would be free if they could be free. All whose work they contain seek the widest possible readership and expect no financial return. But with print, the incremental costs of production and distribution prevent "giving copies away" without limit. The Net, we argued, removed that frustrating constraint.

In the years since that insight has moved along several related paths. We have worked with several other U.S. law schools to create a new distributed system for the digital distribution of formal legal scholarship produced by faculty and students. Perhaps, I should remind you that in the United States every law school publishes at least one law journal, many like Cornell produce several. More remarkable still, those publications are edited by students. Our institute has succeeded in redirecting some of that student energy and talent on which the print journals depend to the production of shorter legal commentary of greater immediate value to lawyers and judges. Working under faculty supervision, the Legal Information Institute student editors produce an electronic bulletin reporting on the important decisions of New York's highest court within days not months of those rulings. Students and faculty members are also deeply involved in the production and review of the editorial content of our Web and disc publications. Finally, as I shall explain in greater detail toward the end of this paper, our faculty's experience in teaching law figures prominently in the LII's future plans.

{slide}

II. Compelling Reasons for Legal Information to be Free, Accessible, and Interoperable

Supporting and informing the Legal Information Institute's activities and strategic decisions through this period has been a steadily growing recognition of the tight, enduring connection between free and effective access to legal information and justice, between its unchecked flow and effective, transparent government.

From earliest times "communication" has been central to law. As the technology of communication has changed, the impact of those changes on law and the central actors in the law process (law-makers, law-appliers, lawyers, and citizens) has been profound. The introduction of the technology of writing, then the printing press, then widespread literacy and the growth of organized libraries each transformed the law activity.

Better access to and improved communication of law have been consistent goals for reformers throughout recent history. In the early 19th century statutes were passed in several American states that required judges to write out their decisions rather than simply speak them so that accurate copies might be distributed in print. America's late 19th century codification and restatement movements were premised significantly on a view that law derived from the mosaic of judicial opinions was too inaccessible. In the 20th century enactment of a federal Administrative Procedure Act and subsequent mandates that governmental regulations be written in non-technical language illustrate the same reformist thrust toward improving the legal system through better access, including better understanding of law.

Since many legal norms do not operate through citizen self application, the quality of communication within the structure of government is equally important to the law's performance. In areas like tax and social security, law operates through vast government agencies, which intersect the lives and activities of large numbers of citizens. Key qualities of government performance such as accuracy, timeliness, consistency, efficiency, and equity (like cases treated like, different cases, with appropriate difference) are strongly influenced by how communication of governing legal norms is accomplished within these agency structures. In areas of the law where judges or judges and law enforcement officials are essential elements of the law application process, the concerns are quite similar even as the means of communication have traditionally been different. Public bodies and those who do their work are among the most important users of legal information.

In some instances, concern that people and enterprises be able to know the grounds of their accountability, "ignorance of the law being no excuse," captures the rationale for these pre-digital reforms, but in others the aims are better understood affirmatively. That is to say whatever goals the law is pursuing and through whatever intermediate means, the prime instrument is communication. Efforts to make law more accessible, more understandable, more clearly expressed are ultimately efforts to make law more effective and in a democracy, more accountable and responsive.

In New York and some other states legislation provides for publication and placement of reported appellate court decisions in county and public law libraries – as a means of providing free access to the state's law. A similar provision for free distribution of statutes exists in nearly all states that publish their own.

Liberated by digital technology from the marginal costs of printing, shipping, and storing which force hard choices about how many copies to print, where to place them, and for whom, law-making bodies might be expected to embrace free distribution of their output by entities like ours and indeed to undertake it themselves. Our experience teaches that there are many reasons they may fail to do so, at least in any way that effectively promotes accessibility and interoperability.

In this failure public bodies are often aided and abetted by others who benefit from controlled access to law. Where there are legal information "haves" and legal information "have nots", significant power resides with the "haves". To the extent that direct access to legal information at the source is difficult or costly, those who can acquire it and can control its subsequent distribution can reap a large profit. Add in such elements as inertia, the force of existing working patterns and relationships, limited resources, and preoccupation with other demanding tasks and what is surprising is not how uneven progress has been toward broad, free access to legal information in the U.S. but rather how widespread and steady it has been.

Where the conditions have been particularly favorable the results have demonstrated in very clear ways what social gains can be realized from free and uninhibited public access to important legal information. I should like to provide examples from four different sources of American law making and administration.

{slide}The most conspicuous examples of public access to law at the federal level in the U.S. are those government agencies whose responsibilities require that they interact with large numbers of the public directly within a complicated legal framework. The two most heavily used government Web sites are not those maintained by the White House,

the U.S. Congress, the Supreme Court, or the Attorney General. ***{slide}***They are, in order of use, the sites of the Internal Revenue Service and the Social Security Administration. Both these agencies have assembled quite comprehensive collections of the legal information within their respective areas of programmatic responsibility – the first dealing with federal tax liability, the second with entitlement to retirement, disability and death benefits. Both collections are relatively easy to use and include all relevant statutes, regulations, and less formal agency manuals and guides. Both also provide a full set of forms, along with complete instructions for their completion. ***{slide}***Indeed, the Social Security Administration now offers an on-line benefit application process as well as a deep parallel collection of information in America's second language, Spanish. As recently as three years ago a commercial publisher was marketing a comparable collection of Social Security legal materials for over USD 1,000.

{slide}My second example comes from one of America's smaller states. I remind you that my country's fifty states all have their own law and legal institutions – legislatures, administrative agencies, and courts. It is the law that they make and apply that bears most directly on the key areas of domestic life – employment, education, family responsibilities, crime, death transfers, and even commercial transactions. As is often true at points of dramatic change, those least well served by the old regime can more readily see and seize the full advantages of the new. This is such a case. A strong example of what can be accomplished by a publicly run judicial Web site has first arisen in a state with fewer than a million residents, North Dakota.

North Dakota is one of many states too small to warrant their own set of printed law reports; it is not a commercially attractive legal information market. For most of the

last century, decisions of its Supreme Court and an annual handful of selected decisions of the state's intermediate court of appeals were published by the major commercial law publisher in the U.S. in a regional compilation with those of six other states.

{slide}In 1997, the North Dakota Supreme Court adopted its own "media-neutral" case citation system (in full conformance with the recommendations of several national bodies). The day an opinion is released it is assigned its permanent official citation. Fox versus Fox decided on May 4, 2001 appeared on the court's Web site that day as 2001 ND 88. During the period the court rules allow the parties to seek a rehearing, it will carry a warning of that possibility in red. When the period has passed, that will be removed. References to specific portions of that opinion need not await its appearance in print for the system includes paragraph numbers. ***{slide}***Here you see paragraph 17. Under court's 1997 citation rule, a citing reference to a particular passage in a 1997 decision is as you see it Zuger v. Zuger, 1997 ND 97, ¶ 13 or Zuger v. Zuger, 1997 ND 97, ¶ 13, 563 N.W. 2d 804. (Zuger v. Zuger being the 97th decision of the North Dakota Supreme court in 1997 and the referenced passage being in paragraph 13.)

The Court also established an official web site to which decisions are released in final, official, citable form – released and archived. The following year, 1998, the site added decisions of the North Dakota Court of Appeals in the same final and official form. Today, lawyers, judges, businesses and citizens of the state have unprecedented access to judicial opinions.

{slide}My third example comes from a neighboring state, Minnesota. Among state legislature sites, Minnesota's currently sets the standard. In countless ways it is superior to either of the federal government sites offering the U.S. Code. Like other top

legislative sites it offers an up-to-date version of the full state code. The database is structured so that users can find relevant provisions by following the logical structure of the compilation (selecting first the relevant title, then chapter, then sections). And the data architecture allows anyone else placing legal material on the Net to create direct links to individual sections as well as to larger units. This legal database shares an important trait with North Dakota's judicial site. Both were constructed to serve government workers, as well as the public. They were designed for serious use, rather than mere public relations.

{slide}My final example is drawn from the level of law-making and application that may well have the greatest day-to-day impact on small businesses and the lives of citizens – namely municipal codes. In countless communities, including surprisingly large ones, the collected laws are poorly maintained and inefficiently distributed. Here is an account prompted by a young lawyer's recent attempt to secure the dog ordinances of the City of Binghamton, New York – a city quite close to Cornell. The lawyer told me: "Believe it or not, the city clerk told me that no complete copy of the Binghamton Code is available to the public anywhere, even the public library. The only way to get an up-to-date version is to go to (or call) the clerk's office."

In a growing number of U.S. cities digital technology and the Internet have enabled officials to do what Binghamton has not. In such places as Rochester, New York; Cincinnati, Ohio; Boone County, Kentucky; Fitchburg, Massachusetts; and Yuma City, Arizona, citizens troubled by barking dogs, interested in establishing an ambulance service or restaurant, or curious about how close the road they can build a garage or place a sign can find the pertinent law on-line.

As these diverse public bodies have discovered and now demonstrate to others, law data in this form is cheap to produce, transmit, and store. Users don't need to own or be close to a dedicated "library" space. Because the cost threshold is so low, many public bodies – courts, legislatures, and administrative agencies – are discovering that they need not depend on commercial intermediaries for dissemination of their work product. Those gathering law in this form, around a particular problem or issue can readily separate out, transport, file and work with the material they judge most relevant. And when done right – like good computer code – it is interoperable, that is, capable of being link to or combined in other ways with related information from other sources.

These examples, when compared with others, also demonstrate that digital distribution alone is not enough.

My early enthusiasm over the growing number of public bodies releasing law in digital form, thrust me into a public exchange with Vance Opperman, then President of the West Publishing Company. He dismissed these sources as offering only "raw data," uttering the phrase in a pejorative tone that suggested sewage. It was a deft rhetorical move, and suggested an important truth: "not all data are of equal value."

{slide}Let us begin with the now obvious difference between furnishing data in print and offering it in digital format. Moving content from print to digital format is costly, running currently at two to three dollars per page for printed English language legal documents. This is a burden we have not shouldered. Everything our institute has done has begun with digital material – in most cases digital material acquired from a public source. The Supreme Court of the United States began releasing its decisions in electronic format in May of 1990, a full decade before it established its own Web site.

The New York Court of Appeals established a dial-up bulletin board at around the same time. By the nineties courts, legislative bodies, and agencies were preparing their output with computers. While print was still their formal or official distribution medium, digital release posed minimal incremental costs. Both the U.S. Supreme Court and New York Court of Appeals financed these incremental costs by charging subscription fees. The former set up a system limited to information brokers or resellers and priced it accordingly. The New York Court set a much lower annual fee of $30. Even with the added long distances charges for those outside the Albany area this put the resource directly in the hands of lawyers and small newspapers.

In working with digital data from these two sources over the years we have learned that while less costly than conversion from print, digital data can carry its own considerable costs. These courts like many public bodies in the U.S. have not yet recognized that digital data can be delivered "free" but configured in ways that severely reduce accessibility, resulting in heavy burdens, both for re-distributors whether non-commercial like our institute or profit-driven entities like West Group or LEXIS, and for ultimate users.

In January 1997, when the Legal Information Institute first undertook programmatic conversion of U.S. Supreme Court decisions to HTML, the Court was releasing its decisions in word-processing format – Wordperfect 5.1. In the summer of that same year, the Court shifted internally to Microsoft Word. Rather than release opinions as Word documents, the Court began with the October 1997 term to release its decisions in the proprietary PDF format. The change came with little warning and insufficient time to allow us to build and fully test what had to be totally new conversion

software. West, LEXIS, and the New York Times also had to contend with this same inattention to the needs of subscribers to the Court's electronic distribution service, though with far greater resources. Fortunately, somewhat later in the same term the Court added an SGML-like format – a hybrid of structural and presentational markup. Unfortunately, this came too late to save the LII the effort of creating software to convert PDF.

{slide}Why go into these technical details? It is precisely in such technical details lies the difference between effective, free and costly, limited public access. Too many public law-making bodies that have undertaken digital distribution of law data have done so without any thought to facilitating redistribution with added value. Distributing only in PDF is a telltale sign. PDF is not friendly to subsequent machine processing. Those who want a court opinion to "look like a court opinion" on the screen or upon being sent to a laser printer are very fond of the format. But for those who would link the references within a document to the cited material, add key words and other metadata, create sophisticated full-text indices, and integrate a document's content with other related law materials PDF is a major barrier.

Subtler barriers lie in format changes and inconsistencies produced by simple inattention. Bodies that exercise great care to assure the quality and consistency of their output in print can wreak havoc on the data systems of others that build on their opinions, enactments, or rules because they will release data that will print handsomely on a page but be utterly confusing to text processing software or search engines. Our work with the opinions of the New York Court of Appeals has given us repeated painful lessons in the many different ways that a majority opinion can be joined with a dissent, the variety of

ways to set off main headings within an opinion or the date of the decision – all the while printing quite handsomely. Until public bodies take digital distribution as seriously as they do print, this will remain a problem.

There is a related way in which too many public bodies in the U.S. undercut the value of their digital distribution of legal information. By declaration and reinforcing practices they withhold full recognition from this version of their law so that both those who know and care enough to be risk averse and those who are easily persuaded by official warnings are pushed toward other final and official (and expensive) versions.

{slide}This is what the U.S. Supreme Court says about the decisions it releases in digital format.

> **Caution:** These electronic opinions may contain computer-generated errors or other deviations from the official printed slip opinion pamphlets. Moreover, a slip opinion is replaced within a few months by a paginated version of the case in the preliminary print, and–one year after the issuance of that print–by the final version of the case in a U. S. Reports bound volume. In case of discrepancies between the print and electronic versions of a slip opinion, the print version controls. In case of discrepancies between the slip opinion and any later official version of the opinion, the later version controls.

{slide}Finally, too few public legal information sites are built with an open architecture. Large numbers reinforce jurisdictional boundaries with data system barriers that frustrate efforts to connect with closely related legal material held on other sites, public or private. This public site distributing decisions of one of the U.S. Court of Appeals circuits is surrounded by several such barriers. It archives decisions in zip files

by date. There are no tools for search or retrieval of individual opinions. And the site's structure blocks others from adding such value directly on top of its archive.

Primary legal texts are peculiarly fragmentary or recombinant. At least that is true of the American legal system. Although units of the U.S. Code are called chapters, they are not like the chapters in a novel, written to be read from start to finish, one after another. Those working with the law must gather relevant provisions around a problem or issue, following cross references in one section that link it to others that sharpen or qualify its effect, tracing back to determine if any of the operative words or phrases are defined elsewhere. Individual appellate decisions rarely can be understood without reference to numerous others, including later ones. And since decisions cannot themselves refer to later opinions that overrule, disapprove or qualify their holdings data systems must do that work. This high degree of textual interconnection is why such large gains can be realized by placing legal materials in a searchable, hypertext environment. Much of our institute's research has concerned techniques, both automated and editorial, that aid the gathering of related legal materials from multiple sources.

III. Some Salient Forces of Resistance Within the U.S. Legal Information Ecology

{slide}Despite the apparent promise in the number of public law sites, our experience has taught us not to be surprised when government agencies, courts and legislatures fail to embrace or aid free distribution of their output, let alone implement effective digital distribution themselves. Here are some of the reasons for such response.

The first is the power of settled practice – the inertial resistance flowing from patterns of work and strong relationships formed during the long history of print distribution. Often these forces work through or express themselves in attitudes about control, important constituencies, or responsibility.

In many European countries, including those unencumbered by doctrines of government copyright in law, free distribution has nonetheless been frustrated by tight control on the terms of access to official systems of digital distribution. Comfortable with uncontrolled private sector print publication and conditioned by Westlaw and LEXIS to view digital law as no less suitable for competitive, multiple source redistribution, U.S. courts and legislatures have been far quicker to release digital take-offs from their law-making activities than their counterparts in some other countries and to do so without attempting to impose conditions. But that does not mean that the U.S. is not troubled by what I might call the "it is our law and critically important to us and our prime constituencies with whom we already have appropriate arrangements" mindset. Government bodies that have a tight affiliation with a particular business sector may not welcome the transparency and consequent reduction in control that free distribution of their documents could bring.

Courts are susceptible to a very different mindset limitation. I think of it as the "that is not our responsibility as judges" posture. It amounts to a view that the tasks of making law or ruling on cases are separate from dissemination. Individually, judges find it quite easy to see their dominant or even exclusive responsibility as deciding cases. Unless the distribution of those judgments in a useful, official format, is clearly lodged in

a well organized judiciary, it will be left to others – those others being commercial publishers in the U.S. setting.

A distribution process that includes substantial time between official act and final official publication may allow some measure of revision during that period. Many appellate courts, for example, have grown comfortable with, indeed, reliant upon the lag between initial release of their opinions and their appearance in "official law reports," using that time for reference checking and editorial review.

In some jurisdictions those functions are actually performed by a separate office, the office of court reporter. Judges write opinions that are released in "slip form" but then readied by a court reporter for publication in archival form. When reporters add summaries and key words to decisions that commonly occurs after rather than before initial release. Nearly all courts delay the attachment of full citation information to decisions until their appearance in print.

All of the above features are reflected in the current practice of New York's highest court. Decisions handed down (and placed on the Internet) by the New York Court of Appeals are not published in "official form" for several months.

It is an overstatement to say that the version of a decision the court releases in digital format is a draft, but each file at the court's site carries the warning: "This opinion is uncorrected and subject to revision before publication in the New York Reports." Having worked with the court's decisions for six years, we can assure you that is not just a formality. If that is so, why doesn't the court subsequently release the final version at its Web site? The reason lies in yet another factor cutting against free and uninhibited access.

Courts (and legislatures) in large market jurisdictions like New York are able to and therefore tempted to reap some return from their output. Since these bodies are not only a source of law but also heavy users of legal information the contractual arrangements surrounding the production of "official" court reports or an "official" state code can provide a way to finance government operations. The commercial entity undertaking the responsibility of doing official publication in print and now electronic format commonly contracts to furnish the issuing public body and other designated recipients with significant quantities of its information products and services.

The addition of editorial content by a state court reporter or legislative staff creates a composite that is copyrightable. That allows the public body to assure a measure of exclusivity to any potential private sector partner, or to secure a revenue stream from any competitor, or both. Court rules requiring attorneys to cite to the official reports reinforce the exclusive arrangement.

This recipe has worked in New York and California, though not in small population states like North Dakota. Indeed, historically large states have been able to generate competition over these contracts. The current New York contract, let to West Group last year, runs for a term of five years. Its provisions are constrained by both established practice and statute. The contract requires the commercial publisher to provide numerous copies of the published reports to state offices ranging from the state library, through all the state judiciary, to each county and public library in the state. The publisher's price for the sale of the reports to the public – both print volumes and other media or formats – is controlled. Finally, the contractor agrees to provide the hardware, software, and training

necessary to enable the staff of the reporter's office to enter decisions into the contractor's data system driving both print and electronic publication.

Several cycles ago local printers vied for this contract. In view of the scale of the undertaking and the current shape of the legal information marketplace that no longer occurs. My principal point in opening up this entrenched practice is to reveal how one state's judicial system trades the legal information it produces for a wide range of legal information and technology services. That exchange would collapse were decisions released unrestricted and free in final and official form, complete with necessary citation information. Since the New York courts have a direct stake in the value received by the "official publisher" the free versions of decisions of decisions at Court of Appeals site or the site run for the New York State Law Reporter by the "official publisher" are offered only temporarily. The Reporter's site does subsequently offer the official version, but for a fee without rights to redistribute – through a transaction directly with the “official publisher”.

A similar pattern exists in California where the publisher of official reports is also West Group. As in New York, the judiciary has a Web site. It holds only "slip opinions." Initially it held them for only 100 days. It has now begun to archive beyond that period. However, the court site instructs users both that the archive collection is not "provided for purposes of legal research" and that:

> Cases beyond the Web site's retention period are available at Westlaw.com in the CA-ORCS database or individually in WestDoc. Westlaw.com is a fee-based online research service of the publisher of the California Official Reports.

IV. Forms of Leadership and Leverage Uniquely Possible with an Academically-Based Center

{slide}Relationships and settled patterns of work and thought like these are not easily escaped. Having no direct stake in the benefits received by either party, centers like the Legal Information Institute are able to demonstrate by example the public value lost as a consequence. We continue to distribute and archive the decisions of the New York Court of Appeals to which we add official citation information in hopes that that may speed the day when the state court reporter is charged with doing so.

Although the LII's on-line U.S. Code was once a Net "exclusive" it has long since become one of many. The House of Representatives itself offers a searchable version. Nonetheless, this LII resource continues to draw over 3 million hits a week. The explanation lies not in unique content but distinctive features of format and functionality. While this collection's content is drawn from the government, it has been reformatted and given navigation and finding aids not available elsewhere on the Net.

We continue to add new features that have increase the value of this resource and significantly several of them draw together information services provided by different offices of the federal government. We have, for example, created links between the Code and related portions of the Code of Federal Regulations, and built an updating feature that integrates separate services offered by the LII, the House of Representatives, and the Library of Congress.

Even public law-making bodies that recognize their obligation to provide effective public access to their law still need a lot of help in coming to understand that a handsome,

free, up-to-date collection of PDF files can fail to deliver on that obligation and can actually frustrate it by making it difficult for other public bodies and independent value-adders like the LII from integrating their work with other relevant material. "Open," "modular," and "interoperable" are qualities as important to the value of legal data as they are with computer code. The on-line opinions of the North Dakota Supreme Court can and do link to cited earlier decisions of the court, but references to the North Dakota Century Code, also on-line, are not linked because the legislature's site, built from a database used for bill drafting has not been structured with such use in mind. The LII's on-line U.S. Code, by contrast, has from the start been set up to welcome links – whether from Supreme Court decisions at our own site or the sites of thousands of others, ranging from U.S. government agencies to numerous special interest newsletters.

{slide}For free law content on the Internet to approach its potential value new analogs must be developed for some very old devices that make particular texts locatable – devices for organizing, finding, and sorting whose print predecessors have become so ubiquitous and familiar as to be invisible. The recombinant nature of law data and very public and decentralized nature of the Net underscore the need for interoperability between collections. Interoperability calls for a set of common approaches permit cross-referencing between documents in separate collections and that act to create integrated functionality among them. Our full text index to the decisions of the thirteen circuits of the U.S. Court of Appeals integrates a distributed collection. In doing so it puts pressure on the respective courts to improve the quality and consistency of their digital distribution of decisions.

{slide}The Legal Information Institute aims to be more than a non-commercial distributor of law content. Through example, white papers, workshops, and technical exchanges with peers we have worked to set and spread standards for interoperability, markup, and resource location. Last July we sponsored an international invitational workshop on these technical matters which can have such important consequences. Participants came from all of the major English-speaking jurisdictions, including importantly our colleagues from AustLII, from important U.S. Government web publishers, from the highest quality state sites offering legal information in the United States, as well as from important sites in Norway, South Africa, and elsewhere. We firmly believe that these discussions which have taken place in multiple venues, including today Meiji, represent an important means of improving the cooperative relationships and interoperable technologies shared between non-commercial legal information centers worldwide.

{slide}Like these peers and others putting law content on the Net, the LII has encountered a vastly larger and more diverse audience for legal materials than the commercial publishers and on-line providers previously perceived or dealt with. Often, it is an audience that is highly sophisticated in its needs even though it is not an audience of lawyers. Professionals of all kinds in many countries make use of the legal information we host and organize. This new and important audience is largely ignorant of the idiosyncrasies of legal research and is, in effect, asking why legal research can't be done in ways that are closer to other forms of on-line research. It is a good question. While there are doubtless sound reasons why legal research *must* be different there is also little

doubt that commercial publishers serving specialist audiences have little reason to innovate or to make things easier for non-specialists.

An important target of the LII's research has involved designing and building systems that seek to serve these nontraditional audiences more effectively. We do so in the belief that finding and organizing legal information is not all that easy for lawyers either, and that improvements in the information environment for a broader audience will improve things for legal professionals as well.

Our present and planned future work in this area concerns: mark-up standards and document structuring, metadata and metadata description, and the coordination of this standards work with other public legal information providers. We shall continue to maintain and further develop key collections of primary material as test-beds for this work, with the twin goals of determining that contemplated standards actually work in practice and of demonstrating that the work involved in conforming pre-existing collections can result in worthwhile improvements in functionality.

I have largely described our work in relation to public bodies. Let me turn now to the other side. Neither our current work nor long-term strategy imagines the withering away of private sector legal information vendors. That will not happen within any future I can foresee. The evidence is strongly to the contrary.

{slide}In January Wolters Kluwer, the multinational information services company based in the Netherlands acquired a U.S. legal information start-up called Loislaw for USD 95 million, which it combined with a previous acquisition, Aspen Publishers a source of legal commentary in diverse formats. The following month Thomson, owner of what is now called the West Group, paid USD 37 million for FindLaw.com – a

commercial site that had explored a non-fee business model, drawing revenue instead from advertising aimed at the audience collected by free legal information. Findlaw will become a magnet for Westlaw similar to the Lexis-One site now run by Reed Elsevier. Thomson's Legal and Regulatory Group posted a 12% increase in revenue during 2000 for a total of USD 2.6 billion. The Westlaw piece of this group experienced 14% growth. Throughout, Thomson has aggressively shifting from its old print business to electronic products and services. And only last week, Reed Elsevier announced that it was consolidating its worldwide legal information products under a single master brand - LexisNexis.

These three enormous enterprises exhibit several important characteristics which they have confidence, backed by huge investment, will assure a strong presence in a growing market. To begin, in the U.S. alone they have the reach and resources, as no single governmental body has, to assemble and configure a fully functional federal and state legal information collection, reaching back in time before the 1990s. Assuming an ever more complete and consistent implementation of the public responsibility for free and effective release of law data, integration of that data across jurisdictions and back across time, packaged with a single interface, format, and search engine will hold large value for those with comprehensive information needs. In addition, all these commercial law data distributors have assembled deep and broad commentary collections – treatises, journals, specialty update services. Finally, I need not tell this audience that these information companies are transnational. They all see a global market for legal information linked to a global market for business, investment, and trade information.

In information markets where one of these major competitors has an advantage it will, understandably, seek a special relationship with any or all of the public bodies that generate law – offering expertise, attractive prices on information services, and trusted band names. During this critical period of transformation, public bodies and their constituencies need the strong persistent pressure of counter examples, examples that demonstrate the value to both of the release of free, accessible, and interoperable law data, in final, official, citable form. At minimum this will promote robust competition in the commercial sector. But it should do more by enabling smaller entities including non-profit research centers like ours to create integrated collections of public resources, specially focused clusters of commentary and primary law, and education services.

V. The Blurring of the Boundary Separating Information and Education

{slide}From the very start, the Legal Information Institute has woven educational themes and activities with the provision of legal information. We have continually prepared core documents for important law school courses and provided guidance to law school faculty members and others interested in incorporating elements of the LII collection in teaching materials. We realized very quickly that important education about law occurred outside U.S. law schools and we, therefore, prepared a CD-ROM collection of historic Supreme Court decisions that we offer to high schools and colleges. Because many users of our Internet-based resources were not U.S. lawyers and judges we added commentary to our Web site that provides basic overview to over one hundred topics of U.S. law linked both to relevant primary material and to other commentary sources

providing greater depth. And for the past five years we have used the Internet as a virtual classroom – offering courses to law students at scattered law school sites.

{slide}During the academic year just finished, the Legal Information Institute conducted two on-line law courses for students enrolled at seven other American law schools. Both courses employ distance learning methodologies that break loose of fixed schedules, time zones, and expensive fixed facilities – namely, streaming audio linked to Web-based multi-media content, interactive exercises, on-line submission of student written work, faculty-student exchange carried on by means of asynchronous conferencing software, plus administrative systems supporting and managing all of the above.

Inescapably, technology shapes the categories we use to discuss and think about human activity. The set of activities people associated with the word "education" and those they refer to as "research or information gathering" will likely grow far less distinct they converge on the same set of digital technologies.

The successful providers of continuing professional legal education in America have increasingly become publishers of print materials, audio and video tapes to the point that most provide more "education" in this form than through live programs. These materials share the characteristic that they allow the learner to choose the time, place, and topic. Long term, we think it probable that the LII web publications and LII-developed distance learning approaches will interweave. We envision integrating introductory "learning" modules with the LII's overview pages and its deeper faculty-organized libraries (the American Legal Ethics Library and Social Security Library). At the topmost level these learning modules would involve no teacher-student interaction or

evaluation. They would also, however, provide a pathway to richer levels of content and interactivity – distance learning options, if you will – available to those with a need or the desire to go further.

VI. What We Can and Cannot Learn from One Another

{slide}I have learned from my colleagues who work in the field of comparative law and from numerous Japanese graduate students at Cornell how remarkably different our legal systems are. Despite superficial similarities, the institutions, practices, culture, professional and governmental categories of our respective countries cause law to operate in ways that frustrate any straightforward one to one translation of important doctrines, procedures, or programs of legal education. Due to differences in techniques of writing and therefore the process of converting legal documents to digital format and indexing or otherwise manipulating them, comparative legal informatics confronts additional challenges.

At a higher level of generality, however, we have solid common ground and exciting prospects for future collaboration. Despite differences of doctrine and detail all legal systems run on information and communication and perform more effectively if that exchange is free and open. No matter how the public and private sectors have handled the transition to digital exchange, the active involvement of academic centers like those represented here can be a tremendously important force – both within the national setting and as collaborators facilitating international information exchange and education. Drawing upon the very differences of our respective environments, we can help each other gain better perspective on the large and exciting challenges we share.