



Attorneys at Law

Consortium Standards Bulletin

A ConsortiumInfo.org publication

JUNE 2004

Vol III, No. 6

From the Standards Blog

☐ ☐ June 21, 2004

#17 AVOIDING A NEW DARK AGE Everyone knows that the period following the fall of the Roman Empire is traditionally referred to as the "Dark Ages" of Western culture. And if someone knows only one more datum about the Dark Ages, it's likely to be that we know very little about this period of time (hence, the adjective "dark").

The reason for the darkness is that European society rapidly crumbled into a multitude of illiterate tribal societies following the withdrawal of Roman military protection, as Europe and even Rome itself came under increasing pressure from barbarian attacks. It was the better part of a millennium before these societies reordered themselves into sufficiently complex polities for literacy training once again to become an attractive investment for political, and not just religious, institutions.

Once that point was reached, history, music, story telling, and the other aspects of everyday life that make the memory of one age available to the next once again began to be both recorded and preserved. But the many centuries that had elapsed between the fall of Rome and the rise of the European kingdoms left little to inform us of what life was like during that long period of neglect.

Are we unwittingly consigning much of our own past to a new Dark Ages, with the result that much of the history and lore of the centuries that have elapsed since the end of the last age of illiteracy will fade into obscurity as well?

Consider this: when was the last time that you did research in a library, as compared to online? Today, we are raising a new generation of students that will be increasingly willing to settle for whatever information is available through a browser, in lieu of seeking out the less convenient, but authoritative sources that may exist in hard copy alone. Because Web-based information is so easy to obtain, we unconsciously (or even consciously) adjust our output to what we can find most easily on line. And worse, we are increasingly tempted to employ data from non-official sources, with no assurance as to whether that data is accurate, or up to date.

Even if we are scrupulous about our sources, if we limit our research to Web-available matter, we are still limited to such information as someone has chosen to place upon the Web – and this is where the threat of a new Dark Ages begins to emerge.

Today, there are a variety of incentives for diverse parties to place most types of current information on the Web. But there are comparatively few incentives to make many types of pre-existing data available on line. Consider, for example, this comparison of Google search results on historical, as compared to current, social phenomena:

"Hugh O'Brien" + Boston + Mayor:

61 accurate results. The longest is a single paragraph. Most are passing references

"Britney Spears" + poptart:

About 11,800 (remove the "poptart" from the search toaster and the results jump to "about 4,690,000")

Presumably, there will be more serious researchers fifty years from now that will be interested in the social changes attendant on the election of the first Irish-born mayor of one of America's most patrician cities than will be interested in the social impact of BS.

Even where valuable data from the past is assembled, it is usually “silo” data – narrow, topical collections of data that are not informed by, or assembled with reference to, other information or learning.

Worse yet, much of the useful information that has been made available on line on-line is not public, in the sense of being maintained by an organization that is likely to have a long existence. Rather, persons that derive no economic benefit from their efforts have placed it on the Web. The result is that they are free to expend as much, or as little, future effort into maintaining that data as they wish, and there is no long-term custodian to ensure its ongoing availability.

But worst of all is the fact that all of this information is ephemeral, since few sites bother to maintain comprehensive printed files of their own content. If I wished, I could unplug this site tomorrow, and all of the content would disappear forever, except for what may be archived at sites such as the Internet Archive’s “Wayback Machine” www.archive.org/web/web.php Of course, the Internet Archive is itself a non-profit organization, whose longevity will be determined by the generosity of its supporters and the enthusiasm of its staff. In comparison, there are libraries full of books that are hundreds of years old that simply need to be properly stored in multiple locations to ensure their survival. Presumably, there is no hard copy library being created by the Internet Archive to match it’s own virtual library.

What this means is that if we are to continue to move to a Web-based information society, as assuredly we should and will (and particularly so for the benefit of the third World), we need to create the following:

- Backup hard storage in multiple locations of all information on the Web, so that a war or other disaster would not literally wipe out a substantial part of all new knowledge and history as it is created.
- Multiple revenue models to create the type of economic incentives that are needed to create and preserve trusted, properly organized, easily found, and reliably archived information.

Call it the ***Alexandria Project***, for the great library in which virtually all of the knowledge of the ancient world was supposedly once stored.

Which finally brings us to the standards part.

The first step in creating (as compared to funding) proper archival storage is a pure standards task: defining what type of storage is needed, how data must be formatted for importation, how often and by what means that data is transferred, and so on. As to funding: this should be a national priority, and part of the national budget.

The more interesting question has to do with providing the incentives for compiling old data comprehensively and usefully, and adding new data reliably and appropriately. And, while we’re at it, let’s make it easier to find trusted data as well.

The most obvious solution is to combine the following three elements:

- a way to accredit data
- a way to search for (only) that data, and
- a way to pay for that data

This is not, in fact, as hard to do as it may seem, at least in so far as describing a workable methodology is concerned. Consider each step individually:

- Accreditation is a classic standards exercise: first, create the requirements for trusted material (definition of scope of hosted information, comprehensiveness, reliability, attribution to sources, updating obligations, and so on). Next, create the infrastructure for applying for, maintaining (and rescinding) accreditation credentials.
- Search is an even easier challenge within a certified system, and could be addressed in a variety of ways, from central hosting to unique identifiers supplied by the accrediting body.
- Revenue models, of course, are the hard part. But it seems more likely that the initial Web model of free access will give way to viable payment models as the quality of information becomes

higher, if the cost of access is low. After many early false starts, there are a number of new micro-payment and subscription models that have recently been launched to address the possibility of selling low-cost content and other goods and services, each of which is informed by knowledge of the factors that led to the demise of the earlier, similar efforts. It is my belief that at least some of these models will take hold, perhaps sooner rather than later. In the end, it may be eBay and PayPal that will sufficiently habituate the public to on line impulse payments to enable a conversion to online content shopping.

As with so many other aspects of the Web, it is time to transition from the frontier era to a mature build out of the full content preservation capacity of this revolutionary resource. It is true that the frontier era provided us all with a once in a lifetime, heady opportunity to participate in a riot of innovation. But it is now time to engage in careful infrastructural investment in order to fully realize the potential of the Internet and the Web to assemble a globally-accessible, trusted library of all human knowledge – and thereby protect us all from the catastrophic loss of that same priceless treasure trove as well.

Comments? updegrove@consortiuminfo.org

Copyright 2004 Andrew Updegrove

The opinions expressed in the Standards Blog are those of the author alone, and not necessarily those of Gesmer Updegrove LLP

###

Useful Links and Information:

About the Internet Archive:

www.archive.org/about/about.php

Mayors of Boston (since 1884):

[www.wordiq.com/definition/Boston#Mayors of Boston.2C past and present](http://www.wordiq.com/definition/Boston#Mayors%20of%20Boston.2C%20past%20and%20present)

Typically Spartan online summary of the life, achievements and impact of Hugh O'Brien:

[www.wordiq.com/definition/Hugh O%27Brien](http://www.wordiq.com/definition/Hugh%20O'Brien)

The following site has, needless to say, a unique perspective. Click here for not only a complete list of Boston mayors – but the location of each worthy's internment as well:

politicalgraveyard.com/geo/MA/ofc/boston.html

Postings are made to the Standards Blog on a regular basis. Bookmark:

www.consoriuminfo.org/blog/