



Authenticating the *John Hancock* of Online Primary Legal Materials

The technical and policy concerns at play

By Anna Russell and Jane Larrington

We often take for granted the authenticity and official status of legal primary source documents. We receive print legal publications from well-known publishers with long-established printing practices and under contract with government entities that provide a reliable chain of custody. We don't have cause to wonder whether the United States Code volumes from the Government Printing Office (GPO) provide an authentic version of the primary source material. We know who created the material, who published it, and who sent it.

Unfortunately, these assurances don't translate easily to the online world. When you download a document from a website that purports to be a government site, how do you know the document is actually coming from that government entity? How do you know that its content hasn't been altered? If there are differing versions of a document, how can you prove which is the official version?

Digital Authentication: What it Means

What does it actually mean to authenticate something in cyberspace? The process of digital authentication is technically complex, but the basic concepts behind the technology are relatively easy to understand. Digital authentication of primary legal materials is concerned with protecting these three security principles:

- **Authenticity:** verifying the source of the material
- **Integrity:** confirming that the material is unaltered
- **Nonrepudiation:** creating evidence that makes it difficult for sources to later deny that they produced the material.

IT professionals and mathematicians have developed standard online protocols using well-known tools to address these three security goals.

Basic Tools of Digital Authentication

The Trusted Third Party

Also known as the certificate authority or CA, the trusted third party can be a commercial company or part of an organization's IT infrastructure. The GPO and several states implementing digital authentication

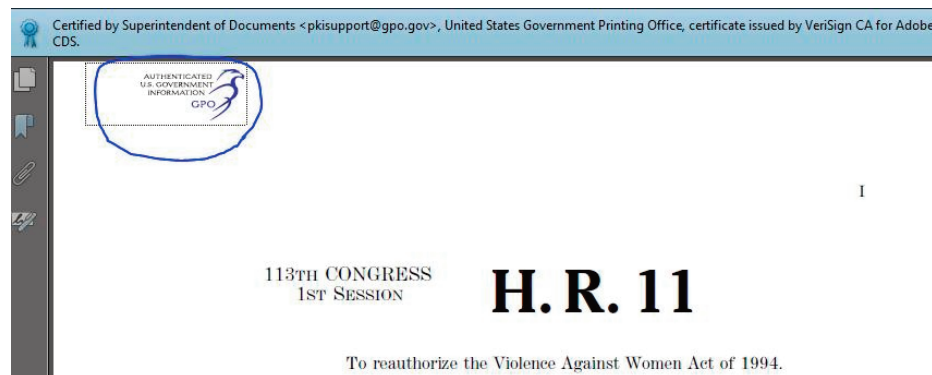
are currently working with commercial CAs like VeriSign, Entrust, or Digicert. The CA's role is crucial to digital authentication because it provides indicia of the original source's authenticity as well as verification to protect against the source's ability to repudiate ownership.

The Certificate

The certificate is similar to the official stamp or seal on an authentic document in the print world. In the online environment, the idea is much the same but includes embedded authentication data (see Image 1).

An encryption technology known as public-key cryptography safeguards the certificate itself by encoding its data. This public-key encryption is a one-way encryption technology with a private-public key pair system. The key pair system works by creating a mathematical

Image 1



VeriSign's certificate for H.R. 11

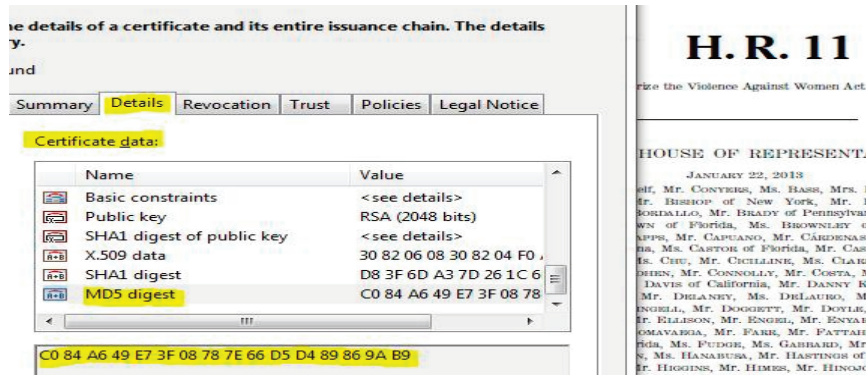
puzzle of particular pieces of data. The CA keeps one key (or encryption code) hidden while making the other key code freely available. Web browsers, in fact, come preinstalled with public keys from many commercial certificate authorities. The CA closely guards the private keys for these key pairs. The public key, when connected to the hidden key, decodes the data to solve the puzzle.

Once the CA has vetted the identity of a particular source or originator, it embeds data for a source's document into

or "hash digest" for the document. This hash value functions like a fingerprint of the file, sensitive to even minor punctuation changes, and remains unchanged for the life of the unaltered document. To maintain the integrity of the hash value, it is encrypted using the CA's private key and embedded in the document's certificate. Each time the authenticated document is opened by an end-user online, the hash value is recalculated and checked with the original, encoded hash value. See Image 2 for an

The processing department for the state legislature would: (1) save Government Code Section 100.1 in a file format publicly accessible and compatible with digital authentication technologies, such as a PDF; (2) run the security software purchased from a commercial CA to e-sign, create the hash value, and produce a certificate from the CA; and (3) upload the certified and e-signed PDF to the state's online repository. When opened from a web browser, the PDF will be stamped and signed by the embedded data contained in the digital certificate. A screenshot of a GPO certificate for H.R. 11, shown in Image 4, provides some of the embedded data included in the certificate, such as the name of the CA, the validity period, and the originator's information.

Image 2



The encrypted hash value attached to H.R. 11

a certificate. You can view the embedded authentication data by double-clicking or right-clicking the certificate area. Embedded data includes: (1) the name of the CA, (2) verification of the originator's e-signature, (3) the public key code, (4) an encoded "hash value" of the document itself, and (5) the certificate's validity period. Certificates are typically good for one to five years; the CA should issue new certificates upon expiration of the old. When a certificate stamps a digital file during the valid time period and the CA uses time-stamping technology, the certificate will remain valid indefinitely (as long as the document is not tampered with). Without the time stamp, the certificate will become invalid once it expires. For a short discussion on digital certificates, visit Microsoft's "What are Digital Certificates?" page at [msdn.microsoft.com/en-us/library/office/aa190113\(v=office.10\).aspx](http://msdn.microsoft.com/en-us/library/office/aa190113(v=office.10).aspx).

The Signature

The signature is the protected kernel at the heart of digital authentication for primary law documents. An e-signature is created when the primary law document (in PDF, DOC, or XML format) is processed through a CA's digital authentication software. The software runs a hash algorithm or computation to produce a "hash value"

example of the encrypted hash value attached to House Bill 113, H.R. 11.

Additionally, see Image 3 for the signature verification that an end-user sees upon downloading and viewing the bill. (It's marked with the exact date and time the document was downloaded and the hash value was verified.)

Putting the Process into Practice

Let's walk through the online authentication process with a hypothetical primary legal document, "State X, Government Code Section 100.1." The state legislature for State X wants to place its Government Code section 100.1 online and allow the public to access, download, and use the documents but prevent tampering or corruption of the document.

Image 3



The signature verification that an end-user sees upon downloading and viewing H.R. 11

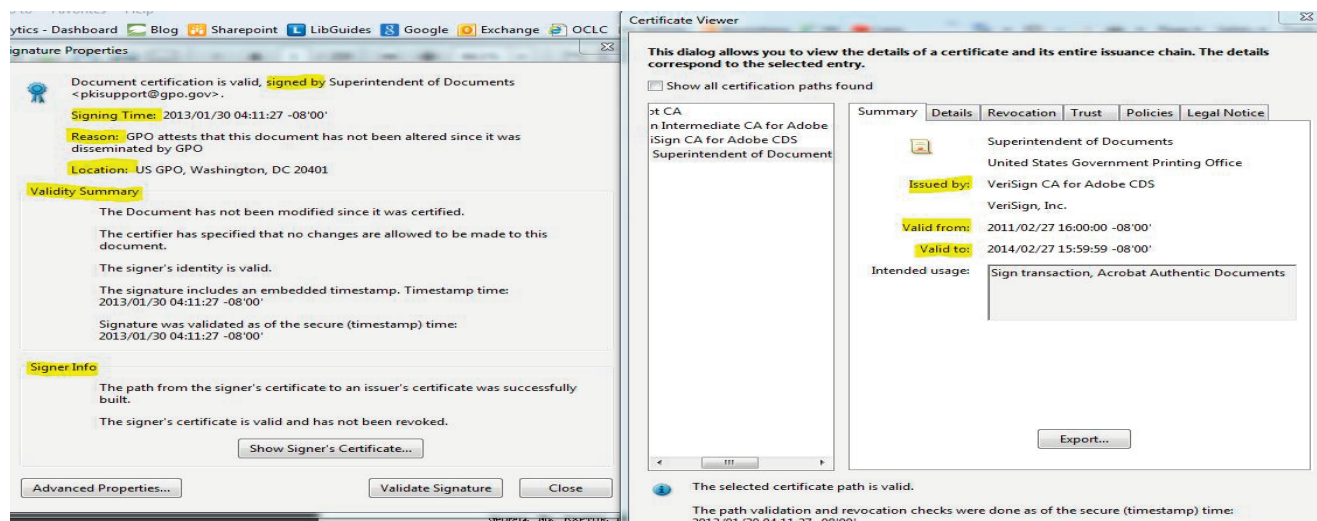
Policy and Logistical Issues with Digital Authentication

The GPO has been working on authentication issues for more than a decade, but most state governments have lagged behind in authenticating their online primary legal materials. Although all states provide electronic access to some primary legal material, only a handful currently authenticate any of that material, and none of them authenticate all of their electronically available primary legal material.

The *Uniform Electronic Legal Material Act (UELMA)* has brought these issues to the fore for many states. *UELMA* requires that online primary legal materials that are designated as official be authenticated, preserved, and permanently accessible to the public. The act leaves the specifics of implementing digital authentication to the individual states. As states adopt *UELMA*, they face a number of technical and logistical challenges and policy decision points.

Among the first decisions for an implementing state will be to select a digital authentication technology to use. A few states did consider authentication of their online primary legal materials in the early 2000s but abandoned efforts

Image 4



A screenshot of an example GPO certificate for H.R. 11

because the technology was too unwieldy and expensive. The technology is streamlined and less costly now, but states may still find that costs are out of reach in this era of ever-shrinking budgets. One of the primary motivating factors for states to consider *UELMA* is to *reduce* publication costs by eliminating the need for an official print publication. Multi-state partnerships may help spread costs, at least those initial-stage costs for investigating options and setting up efficient systems.

UELMA calls for both authentication and preservation. Both priorities can be achieved by implementing a well-planned technology infrastructure. A 2008-2012 preservation project funded by the Library of Congress' National Digital Information Infrastructure and Preservation Program expanded its scope to address authentication. The project brought together a multi-state, multidisciplinary group including partners from historical societies, libraries, archives, and state entities responsible for drafting, revising, and publishing state codes. The project generated a white paper on authentication methods and their associated costs as well as an authentication prototype. These and other valuable resources are posted on the project website: www.mnhs.org/preserve/records/legislativerecords/carol/index.htm.

The logistical and policy challenges each state faces may prove even more formidable than the technology. Legal publishing in many states is not centralized; different entities are responsible for publication of statutory codes, court decisions, regulatory codes, agency decisions, and other primary legal materials. Print versions may

	Legislative materials			Judicial materials				Executive materials		
	Constitution	Session laws	Statutory codes	Court rules	Supreme court	Intermediate app. court	Trial court	Admin. regs/rules	Agency decisions	Register/other
CA	X	X	X							
CO	X	X	X					X		
CT	X		X		X	X	X	X		
HI	X	X	X	X	X	X		X		
IL	X	X	X	X	X	X		X	X	
MA	X	X	X	X	X	X	X	X	X	X
MN	X	X	X					X		
MO	X	X	X	X	X	X	X	X		
NV	X	X	X					X		
ND	X	X	X					X		
OR	X	X	X					X		
RI	X	X	X	X	X	X	X	X	X	

The combinations of legal materials each introducing state includes in its *UELMA* bills

be designated official, but unofficial, online versions may be provided by a commercial vendor, by a centralized state office, or by individual courthouses and agencies. There is a complete lack of uniformity, not just among states, but *within* states, as to how primary legal materials are published both in print and online. Systematic digital authentication of all state primary legal materials is likely impossible until more uniformity and centralization is achieved.

However, states may face legal issues as they try to change the way primary legal materials are published. Some states

statutorily assign publication responsibilities to private publishers. These statutes may require amendment before digital authentication projects can begin. Constitutional separation of power issues may be implicated. The statutorily prescribed centralized publishing of all federal materials hasn't been without controversy; several federal agencies have balked at what they see as legislative encroachment on executive branch sovereignty. State legislative attempts at centralizing publishing may be similarly challenged.

(continued on page 31)

Fortunately, the Uniform Law Commission structured *UELMA* such that individual states can include or exclude particular types of primary legal material. For example, California has chosen to enact *UELMA* only for legislative materials. See the table on page 19 for the combinations of legal materials each introducing state includes in its *UELMA* bills.

Additionally, a brief from the Center for Technology in Government provides a good starting point for states considering *UELMA*: www.ctg.albany.edu/publications/reports/legal_materials/legal_materials.pdf.

Law librarians played a vital role in the creation of *UELMA* and its introduction and passage in a number of states. Our information science and legal expertise, facility with technology,

and understanding of end-user concerns position us as ideal partners on digital authentication and preservation projects. ■

Anna Russell (russell@sandiego.edu) is the electronic resources librarian and **Jane Larrington** (jlarrington@sandiego.edu) is a reference librarian at the University of San Diego, Pardee Legal Research Center.