

# Rechtsgeschichte Legal History

[www.rg.mpg.de](http://www.rg.mpg.de)

<http://www.rg-rechtsgeschichte.de/rg24>  
Zitiervorschlag: Rechtsgeschichte – Legal History Rg 24 (2016)  
<http://dx.doi.org/10.12946/rg24/341-343>

Rg **24** 2016 341 – 343

**Kellen Funk**  
**Lincoln A. Mullen**

## A Servile Copy: Text Reuse and Medium Data in American Civil Procedure

**Kellen Funk,  
Lincoln A. Mullen**

## A Servile Copy: Text Reuse and Medium Data in American Civil Procedure

At the opening of the first Nevada legislature in 1861, Territorial Governor James W. Nye, a former New York lawyer, warned the assembly that it had the burden to erect a functioning government within a short legislative session. »Happily for us, a neighboring State whose interests are similar to ours, has established a code of laws« that Nye argued could »be made applicable« to Nevada. That neighboring state was California, and the California mining lawyer William Morris Stewart

followed Nye’s instructions to the letter. Stewart literally cut and pasted the California Practice Act into the session bill, crossing out *state* and *California* and substituting *territory* and *Nevada* where necessary. (Figure 1.) Stewart copied not just California’s code but also its method of codification. California too had borrowed its code in a similar if not quite as extensive manner from New York.

Although, as scholars are aware, codification was not nearly so extensive in America as in

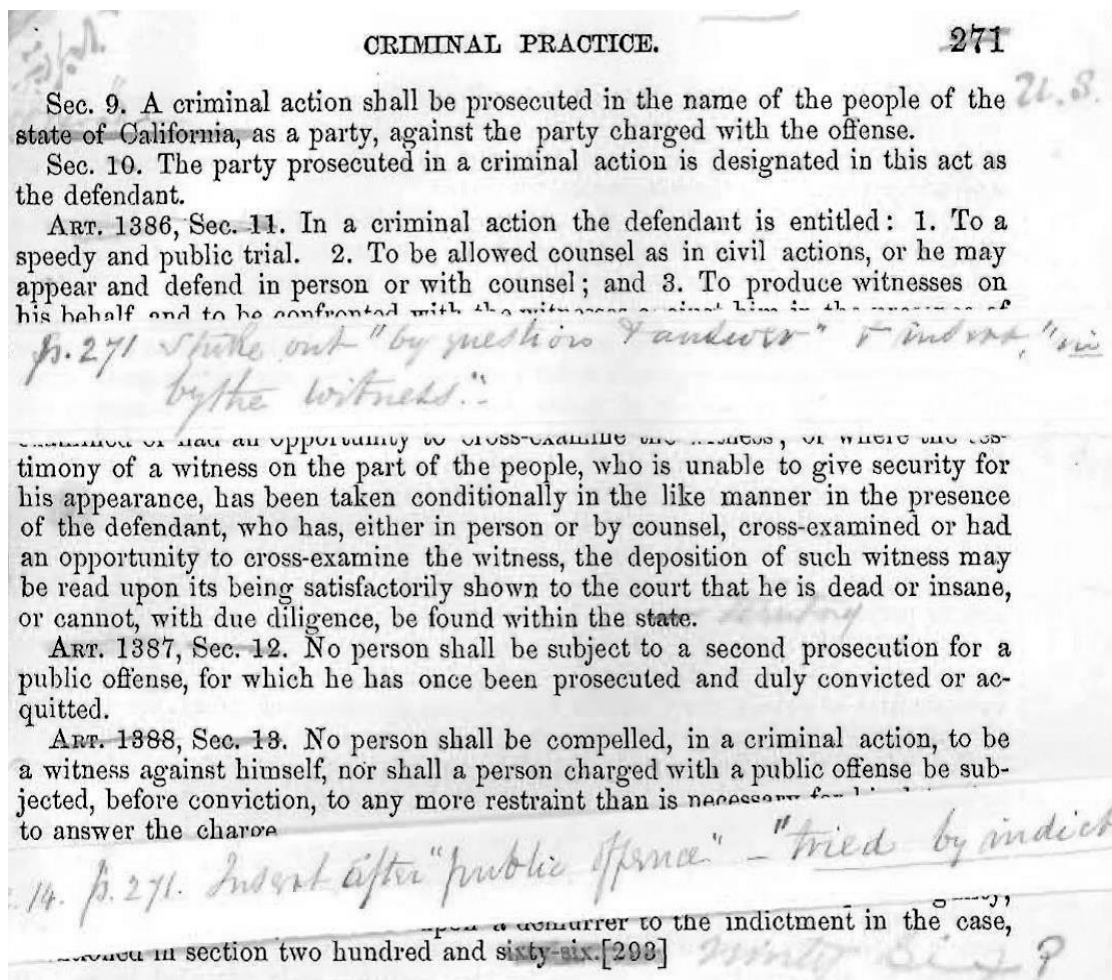


Figure 1. Detail from Council Bill 21, First Territorial Legislative Session (1861), Nevada State Library, Archives and Public Records.

Europe, the field of civil procedure was an exception. In the second half of the 19<sup>th</sup> century, over 30 American common law jurisdictions enacted a code of civil procedure, joining a half-dozen southern states that had inherited codified systems from French or Spanish rule. Proceduralists in the United States are aware that those 30 jurisdictions adapted their codes from a model drafted by the New York trial lawyer David Dudley Field in 1850. However, to date, the extent to which these states and territories »borrowed« Field's actual text, and the political significance of »borrowing« legislation, have gone unexamined. In addition to being numerous, American procedural codes were long documents, often more than 1,000 sections, and they covered more topics than their European counterparts. Altogether, the original Field family of codes comprises some 52,000 laws. A close reading of these rules using the conventional practices of cultural history would not just be daunting – it would be impossible.

We have applied computational methods to detect the reuse of text across these procedure codes. That is, we have reverse engineered the process of borrowing revealed in the archival version of Stewart's Nevada Code but largely hidden in the history of the other codes. We have asked to what extent did one jurisdiction reuse the text of another, taking into account minor amendments in wording or substantive modifications to the law. The course of this project has led us to develop two ideas. First, law is particularly well suited to certain digital methods such as detecting text reuse. But second, those methods are most successful when combined with traditional methods of social and cultural history.

At the level of individual texts, law's discursive organization of its rules simplifies a fundamental problem in detecting the reuse of texts: determining the unit of comparison. If one wants to trace textual influences on *Moby Dick* or *War and Peace*, should one look for borrowings of individual words, phrases, lines, or pages? A legal code inherently suggests what should be compared. Modern codes were almost universally organized into *sections* (sometimes called *articles*), with each section providing one concise regulation. (Although legislation offers particular clarity in its organization, other legal sources, such as treatises or cases also tend to have easily demarcated sections, such as articles or rulings.) More than a handy organizational scheme, sections were historically the way

legislators imported their texts, as a close look at the Nevada Code reveals. Codifiers took their sources apart by sections, rearranging here, editing, drafting, and then re-combining there.

Thus the actual development of a legal text aids the digital deformation and analysis of its rules. At least for 19<sup>th</sup>-century typesetting, legislative section breaks are programmatically detectable, allowing us quickly to dismantle our corpus of codes into its component sections to search for borrowings. Standard methods of scoring text similarity then allow us to retrace legislators' steps reliably over an entire corpus, a task which could not reasonably be completed by hand.

The discursive structure of law lends itself to digital analysis not only at the level of the text but at the higher level of a corpus. Digital text analysts often assume their methods are useful only for »big data« corpora, such as multi-million volume archives of books or newspapers. The jurisdictional boundaries of law allow for and invite a more curated corpus, however. Whether jurisdiction is treated as a function of geography or sovereignty (Germany, the European Union, the common law world) or subject matter (criminal law, private law, procedure, or the ancient persons, things, and actions) the western legal enterprise has from antiquity sought to mark precise boundaries of genre. One could, of course, attempt to trace textual influence through all legal treatises ever written, but these jurisdictional boundaries permit a more targeted examination of bodies of law in which interconnections might be more analytically interesting, such as between corpora of canon law and Roman law treatments of the same topic.

We propose to refer to such projects as »medium data,« defined as projects small enough to fit into memory on a single standard computer but large enough to make naïve approaches to the text implausible. Although our project required the collaboration of a proceduralist and a digital historian, we have accomplished our analysis without the large teams and funding sources that feature in many arresting digital humanities projects. This is largely because 1. our method of selection was provided by the sources themselves – we sought a corpus of *codes of procedure* in the *United States*, and 2. because of the economic interests and demands of the legal profession, those sources were already largely *printed, digitized, and available*.

We prefer to think of this as a »medium data« project for the further reason that digital analysis,

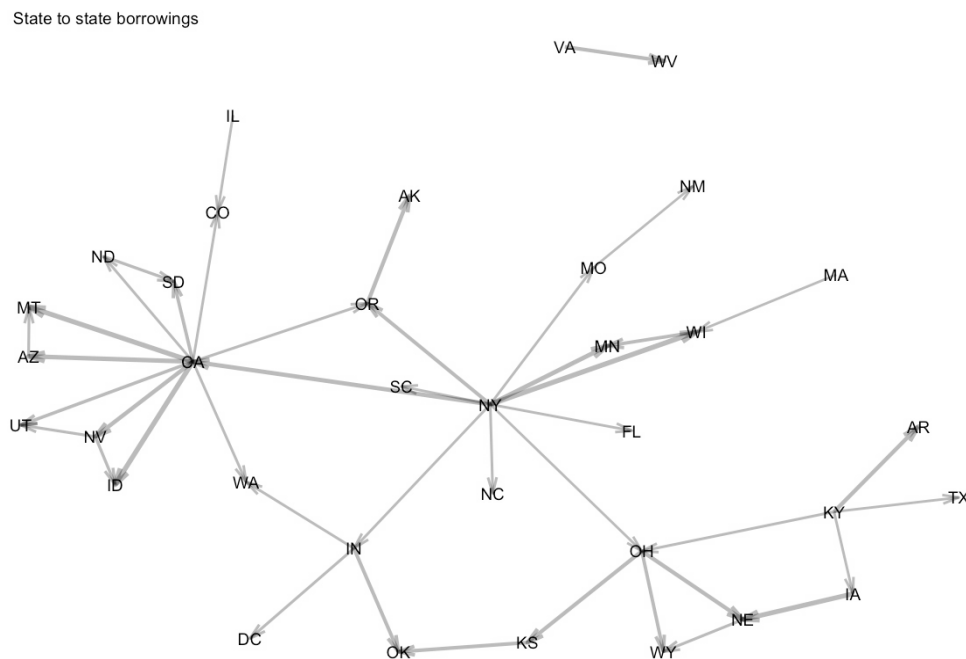


Figure 2. Influence of Field Code jurisdictions upon one another based on number and strength of textual borrowings.

in itself, does not provide sufficient or satisfactory results apart from conventional practices in the humanities. To be sure, deforming codes into sections and rearranging them into the pattern of borrowings allows us a more refined analysis than would be possible merely performing close readings. At a microscopic level, we can »map« the evolution of a single regulation, for instance, the definition of witness competency. As the original Field Code migrated around the nation, some states copied the text verbatim, while others modified the text to disqualify racialized witnesses. The racial disqualification section then generated its own branch on a family tree, as other jurisdictions opted to copy it instead of the original New York text. At a macroscopic level, we can aggregate all these borrowings to reveal a national map of regional code families based on the number of sections borrowed and the degree of similarity among the borrowings. (Figure 2.)

But what is the *significance* of these regional families? A medium data approach uses these digital methods to support further traditional scholarship. Our code enables us to observe patterns of interest by deforming the text to aggregate and analyze the parts. Then we put those parts back together and search out their contexts in conven-

tional sources of social, political, and cultural history.

Take, for instance, the network graph in Figure 2. The codes of western America cluster around California, those of midwestern America around Ohio. But the southern states divide, with a significant portion clustered around New York. Why? The codes that cluster around New York were adopted between 1868 and 1870, that is, when the southern states were being »reconstructed« immediately following the American Civil War. Legislative and lawyerly archives in North Carolina preserve abundant commentary about the Code as a northern imposition, »a servile copy of New York« dictated by northern capitalists and their legal counsel. Indeed, the codifier responsible for North Carolina's Code was the Radical Republican Albion Tourgée, a New York-trained lawyer. The political press condemned Tourgée's Code as »one of the greatest curses inflicted upon North Carolina« and agreed to accept racial political equality if only the new code of practice were repealed. Digital methods allow us to trace Tourgée's pen as he adapted the New York Code, but we then follow those adaptations into a stormy scene of American political and legal culture. ■