Copyright Laws:
Borrowing from State or British Legislation?

Niels Pepels
Max Planck Institute for European Legal History
pepels@rg.mpg.de

Andreas Wagner
Max Planck Institute for European Legal History
wagner@rg.mpg.de

The migration of the Field Code was a central event in Anglo-American legal history, but no historian has traced the extensive borrowings of the Field text nor recognized the political furor that greeted the code outside New York.[1] Every aspect of a civil justice system, from the rules granting access to courts, to lawyers, to remedies (whether damages, injunctions, or seizure of property) was covered by the code, making its New York-specific rules politically contentious both inside and outside the Empire State. As the Field Code migrated around the country, commentators in each jurisdiction raised the same complaint: how could legislation borrowed from another state represent the popular will and best interests of *this* state?

Understanding the history of the Field Code requires not only attention to its political context but also a detailed examination of the substance of what was borrowed and what was revised in each jurisdiction. Exploring these borrowings is a daunting task, however. Procedure codes were long, technical documents, and although each jurisdiction copied large swaths of text, each also modified the text along the way, sometimes with a simple *Nevada* for *California*, sometimes with more foundational changes to civil remedies. Although Stewart's cut-and-paste code found its way into the archives, most draft legislation did not, and few codifiers explained in detail how they produced their texts. Traditional close reading or textual criticism of some 98,000 distinct sections of law across 20,000 pages comprising 7.7 million words is simply not a feasible research task for a historian who wishes to track these borrowings.

Yet by turning to the digital analysis of texts, we have resolved this difficulty and tracked how states borrowed their codes of civil practice from one another. Within the corpus of legislation, algorithmic analysis of texts can reverse engineer and visualize what the archive

[1] Roscoe Pound, "David Dudley Field: An Appraisal," and Alison Reppy, "The Field Codification Concept," in Alison Reppy, ed., *David Dudley Field: Centenary Essays* (New York University School of Law, 1949); Stephen Subrin, "David Dudley Field and the Field Code: A Historical Analysis of an Earlier Procedural Vision," *Law and History Review* 6 (1988): 311–373; Robert G. Bone, "Mapping the Boundaries of a Dispute: Conceptions of Ideal Lawsuit Structure from the Field Code to the Federal Rules," *Columbia Law Review* 89 (1989): 1–118. See also the literature on procedure and codification cited below.

revealed in figure 1: which texts were borrowed, which modified, and how extensively.[2] Our method works especially well for legal texts, for reasons we will explain.

The computational analysis of texts is a method which historians can use across the discipline to study many topics. We have used text analysis—specifically, methods to detect text reuse—to show how law migrates through borrowings, just as scholars on the *Viral Texts* team have demonstrated how newspaper articles were reprinted in the nineteenth-century United States.[3] While we have developed a method that discovers borrowings of exact words and phrasings, other forms of digital text analysis can track the diffusion of concepts. Making due allowance for the specific historical questions and sources at hand, computational text analysis is broadly applicable to historical problem that involves the spread of words or ideas. The specific method that we outline could be extended beyond codes of civil procedure to legal statutes or treatises. Yet it could also be used, for instance, to track the spread of hymns in collections of hymnbooks in religious history, or the reuse of sections of medical textbooks in the history of science. Since historians by and large work with textual sources and increasingly with digitized texts, computational text analysis should become a part of the historian's toolbox. Lara Putnam has recently described the importance of "digitization and OCR, which make words above all available" for historians to search and read, while observing that "computational tools can discipline our term-searching if we ask them to."[4] Historians have been slow to take up text analysis even as the aid to more traditional reading that Putnam recommends. Yet as we demonstrate in this article, such methods can reveal patterns inaccessible to the traditional historian.[5]

[2]We have released two repositories with all the code used for this project. Lincoln Mullen, "textreuse: Detect Text Reuse and Document Similarity," R package version 0.1.3 (2015–): https://github.com/ropensci/textreuse, includes our implementation of LSH and other algorithms suitable for use by other scholars. (This package was peer-reviewed by rOpenSci, a collective of academic developers who use the R programming language.) A second repository contains all of our code specific to the migration of the Field Code: https://github.com/lmullen/civil-procedure-codes/. These are the most essential software packages that we used, except for those cited elsewhere: R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria (2016): https://www.R-project.org/; Hadley Wickham and Romain Francois, "dplyr: A Grammar of Data Manipulation," R package version 0.4.3 (2016): https://CRAN.R-project.org/package=dplyr; Hadley Wickham and Winston Chang. "ggplot2: An Implementation of the Grammar of Graphics," R package version 2.1.0 (2016): https://CRAN.R-project.org/package=ggplot2; Hadley Wickham, "stringr: Simple, Consistent Wrappers for Common String Operations," R package version 1.0.0 (2016): https://CRAN.R-project.org/package=stringr; Hadley Wickham, <80><9c>tidyr: Easily Tidy Data,<80><9d> R package version 0.4.1 (2016): https://CRAN.R-project.org/package=tidyr; Gabor Csardi and T. Nepusz, "The igraph Software Package for Complex Network Research," *InterJournal, Complex Systems* 1695 (2006): http://igraph.org.

[3]Ryan Cordell and David A. Smith, *Viral Texts: Mapping Networks of Reprinting in 19th-Century Newspapers and Magazines*, NULab for Texts Maps and Networks, Northeastern University (2012–): http://viraltexts.org.

[4]Lara Putnam, "The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast," *American Historical Review* 121, no. 2 (2016): 399–400.

[5]Only a few historians have used computational text analysis, including Cameron Blevins, <80><9c>Space, Nation, and the Triumph of Region: A View of the World from Houston,<80><9d> *Journal of American History* 101, no. 1 (June 1, 2014): 122<80><93>47, doi:10.1093/jahist/jau184; Benjamin M. Schmidt, *Sapping Attention* (blog): http://sappingattention.blogspot.com/; Sharon Block, "Doing More with Digitization: An Introduction to Topic Modeling of Early American Sources," *Common-Place* 6, no. 2 (January 2006): http://www.common-place.org/vol-06/no-02/tales/; Dan Cohen, Frederick Gibbs, Tim Hitchcock, Geoffrey

The first contribution of this article is to demonstrate our methods as applied to a corpus of nineteenth-century civil procedure codes. The second contribution is to integrate what we learned from the text analysis with the more conventional approaches of political and cultural history to explain why the migration of the Field Code mattered. On the national level the extent of legislative borrowing followed a pattern American historians have described as a "Greater Reconstruction" in which the former Confederate South and the Far West showed a remarkable kinship. Scholars have typically described Greater Reconstruction as a federal development, featuring the creation of national citizenship, a national economy, and a larger federal apparatus centered in Washington, D.C. This article shows that Greater Reconstruction had its state-level dimensions as well. The uniform practice of law and adjudication of civil remedies was not structured by Washington mandates, however, but by the anxiety that New York financial capital would follow only New York civil remedies. At the more local level, our digital computations can trace modifications within code traditions, for instance, the ways western and midwestern codifiers altered New York's law to accommodate hardening conceptions of racial competencies in the civil courts.

The code would encounter similar difficulties in each jurisdiction that adopted it. Even the shortest version of the Field Code was significantly longer than any other state statute before the Progressive legislation of the twentieth century. Unlike statutory compilations that sometimes took the name of a "code" but made no changes to existing law, the Field Code opened by abolishing the hallmarks of prior practice and instituting "hereafter" a new form of action with substantial revisions to basic matters of civil remedies.[6] In the states where it was imported, there was no getting around the fact that the code introduced much new law, yet legislators were unable to read, critique, and amend the code within the brief period of a legislative session. "It is folly to undertake to pass a code in a sixty day session," wrote the *Montana Post*, "and the best way would be for the Assembly to select one from a State or Territory which would come near meeting our wants, and slide

Rockwell, et al., "Data Mining with Criminal Intent," white paper, 31 August 2011, http://criminalintent.org; Robert K. Nelson, "Mining the Dispatch, website, Digital Scholarship Lab, University of Richmond, http://dsl.richmond.edu/dispatch/; Dan Cohen,"Searching for the Victorians," 4 October 2010: http://www.dancohen.org/2010/10/04/searching-for-the-victorians/; Micki Kaufmann, *"Everything on Paper Will Be Used Against Me:" Quantifying Kissinger*, digital project (2012–16): http://blog.quantifyingkissinger.com/; E. Thomas Ewing, Samah Gad, Bernice L. Hausman, Kathleen Kerr, Bruce Pencek, and Naren Ramakrishnan, "An Epidemiology of Information: Datamining the 1918 Flu Pandemic," project research report (2 April 2014): http://vtechworks.lib.vt.edu/bitstream/handle/10919/46991/An%20Epidemiology%20of%20Information%20Project%20Research%20Report_Final.pdf?sequence=1; Michelle Moravec, " 'Under this name she is fitly described': A Digital History of Gender in the History of Woman Suffrage" (March 2015): http://womhist.alexanderstreet.com/moravec-full.html.

[6]Final Report of the Commissioners, 225–226, 554. For examples of "codes" that did not alter previously enacted statutes, see, for instance, Report, Appendix to the Journals of the Senate and Assembly of the State of Tennessee (1857), 191 ("The digest presents the law substantially as it now exists in the State. I have neither felt at liberty nor deemed it advisable to innovate largely upon the established system."); 1897 New Mexico Compiled Laws 9 ("The commissioners were given no authority to revise."); 1866 Illinois Compiled Laws v ("We cannot change the text, but we can arrange and systematize the entire legislation of the state upon any given subject."); 1849 Wisconsin Revised Statutes, "Advertisement" (commission "directed the subscriber to arrange the chapters into parts and titles as he thought proper, re-arranging the order of the sections or transposing them from one chapter to another, whenever it would not alter the meaning of the law.").

it through with the fewest changes possible."[7] Sliding the code through eased the problem of time but exacerbated the problem of local sovereignty. "To be governed by a foreign law, especially when that law is not preknown to the people whose conduct is to be regulated thereby ... is something repugnant to the idea of Democratic Republican government," complained the *The Miner's Express* in Iowa.[8]

How, then, did states and territories achieve a politically acceptable balance between efficiency and sovereignty, borrowing law for sake of time but endowing it with popular legitimacy in each locale? Quite apart from the technicalities of legal practice, the American federation of civil government into (depending on the year) more than thirty or forty separate jurisdictions makes it hard to describe a phenomenon that was truly national despite its state-centered enactments. It requires a sense of how much law was borrowed in each location and to what degree innovations were introduced. But precisely because these questions concern codes—texts that comprehensively and systematically cover a given subject—they are ideal sources for the techniques of digital history.

2. Detecting borrowings among procedure codes

To discover how the Field Code migrated to other jurisdictions, we compiled a corpus of potentially relevant laws, including separately bound codes of civil procedure as well as codes or statutes appearing within session laws and statutory compilations from around the Atlantic world. The corpus comprises 135 statutes from the nineteenth century, which amounts to 7.7 million words organized into 98,000 regulations. It includes the initial enactment of every U.S. code of civil procedure, as well as procedure statutes and re-enacted codes from jurisdictions reputed to have been legally influential, including French and British codes. The corpus does not include every nineteenth-century statute of procedural law. While a comprehensive project may be illuminating in its own ways, our specific question of how New York legislation influenced other American jurisdictions permits a more curated corpus.[9]

Curating a corpus to answer a specific question is one of two ways in which digital history can proceed. To quote Jason Heppler, digital history–like all historical work—can begin either "with a corpus looking for a question, or a question looking for a corpus." Computational text analysis in digital history is often conceived of as beginning with sources, particularly with large datasets such as the Google Books or Hathi Trust corpora. These large corpora are sometimes called "big data"—though it must be emphasized almost never by digital historians who actually work with them—on which "distant reading" can be practiced. While it is salutary for historians to have their research questions shaped by the broadest possible contexts, it is not apparent that digital historians can readily move from these omnibus corpora to answering the specific research questions that animate various historical fields. In this article we demonstrate an alternative approach,

[7] *Montana Post*, January 21, 1865.

[8] *The Miner's Express* (Dubuque, IA), February 26, 1851.

[9] For full citations to all of the codes that we used, plus links to electronic versions at the Hathi Trust, Google Books or other sources when available, see Kellen Funk, "American Civil Procedure: Law on the Books" (2015–16): http://kellenfunk.org/civil-procedure/procedure-law/.

which we might unimaginatively label "medium data." The amount of legislation that governed American civil practice is impressive, since every state amended and re-enacted procedure statutes nearly every decade. But while a corpus of procedural legislation requires some computational sophistication, the techniques are far less complex than those derived for truly big data. We have gathered a large but narrowly constrained corpus centered on solving a well-defined research question.[10] This corpus is large enough that digital historical methods provide results that a scholar could not obtain through traditional methods, but sufficiently circumscribed so as to directly address a discipline- and field-specific question.[11]

[10]Our approach draws on an earlier generation of digital history which collected sources, as exemplified in *The Valley of the Shadow* project: William G. Thomas III and Edward L. Ayers, "The Differences Slavery Made: A Close Analysis of Two American Communities," *The American Historical Review* 108, no. 5 (December 1, 2003): 1299<80><93>1307, doi:10.1086/587017. See also Dan Cohen and Roy Rosenzweig, *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web* (Philadelphia: University of Pennsylvania Press, 2005), ch. 6, digital edition hosted at Roy Rosenzweig Center for History and New Media, George Mason University: http://chnm.gmu.edu/digitalhistory/.

[11]We have preferred to use the term "digital history" when referring to our own work, in part because the term "digital humanities" has largely come to refer to the work of digital literary and digital media scholars, but primarily because we wish to see digital scholars make disciplinary, rather than interdisciplinary contributions. We are working primarily in the field formerly known as humanities computing, but there are other forms of digital history such as digital public history or spatial history. On the role of disciplines and the importance of field specific argumentation, see, Stephen Robertson, "The Differences between Digital Humanities and Digital History," 289–307, and Cameron Blevins, "Digital History's Perpetual Future Tense," 308–324, both in *Debates in the Digital Humanities 2016*, ed. Matthew K. Gold and Lauren F. Klein (Minneapolis: University of Minnesota Press, 2016); William G. Thomas III, "The Promise of the Digital Humanities and the Contested Nature of Digital Scholarship," in *A New Companion to the Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth (Malden, MA: Wiley Blackwell, 2016), 524–37. Nevertheless, digital text analysis in the humanities has mostly been published by literary scholars, including Stephen Ramsay, *Reading Machines: Toward an Algorithmic Criticism* (Urbana: University of Illinois Press, 2011); Franco Moretti, *Distant Reading* (New York: Verso, 2013); Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History* (Urbana: University of Illinois Press, 2013); Ted Underwood, *Why Literary Periods Mattered: Historical Contrast and the Prestige of English Studies* (Stanford: Stanford University Press, 2013). Related but less well regarded by humanities scholars is work in "culturomics": see Jean-Baptiste Michel et al., "Quantitative Analysis of Culture Using Millions of Digitized Books", *Science* 331, no. 6014 (2011), 176–182; doi:10.1126/science.1199644. On the uselessness of the term "big data," see Ted Underwood, "Against (Talking About) 'Big Data,' " *The Stone and the Shell*, blog post, 10 May 2013: https://tedunderwood.com/2013/05/10/why-it-matters-that-we-dont-know-what-we-mean-by-big-data/. For an overview of digital history projects involving text analysi, see the Roberston essay cited above. For an example of curating a corpus aimed at research questions, see Ted Underwood, Boris Capitanu, Peter Organisciak, Sayan Bhattacharyya, Loretta Auvil, Colleen Fallaw, J. Stephen Downie, "Word Frequencies in English-Language Literature, 1700–1922," dataset, v0.2 (HathiTrust Research Center, 2015) doi:10.13012/J8JW8BSJ.

Only a few scholars have turned their attention to the computer analysis of legal texts for historical purposes, including Paul Craven, "Deteccin automtica y visualizacin de dominios especficos similares en documentos: anlisis DWIC y su aplicacin en el Proyecto Master & Servant [Automatic Detection and Visualization of Domain-Specific Similarities in Documents: DWIC Analysis and its Application in the Master & Servant Project]," published on CD-ROM in F. J. A Perez et al., eds., *La Historia en una nueva frontera [History in a New Frontier]* (Digibis: Ediciones de la Universidad de Castilla-La Mancha, 1998); Paul Craven and Douglas Hay, "Computer Applications in Comparative Historical Research: The Master & Servant Project at York University, Canada," *History and Computing* 7, no. 2 (1995); Paul Craven and W. Traves, "A General-Purpose Hierarchical Coding Engine and Its Application to Comparative Analysis of Statutes," *Literary and Linguistic Computing* 8, no. 1 (1993): 27–32, doi:10.1093/llc/8.1.27; Eric C. Nystrom

Because most codes were public statutes, they were widely printed and distributed and therefore found their way into libraries digitized by Google Books. We drew primarily from the Google Books, filling in gaps from other databases as necessary. We used optical character recognition software (OCR) to create plain-text versions of the codes, which we edited lightly, correcting section markers by hand as necessary and writing a script to fix the most obvious OCR errors.[12]

The most important step we took in processing the files was to split each section of the code into its own text file. Codes varied in how they were organized, but they all divided specific regulations into *sections* (or, on occasion, *articles*). Not only does the discursive form of these texts provide a handy organizational scheme for digital methods, but historically sections were also the way legislators borrowed their texts. Codifiers took their sources apart by sections, rearranging here, editing, drafting, and then re-combining there. Despite the fact that states differed widely on what topics they included in "civil procedure," sectioning the codes allowed us to assess similarity even among codes of quite different lengths and coverage. For instance, we know that California's 1851 code was derived from New York's 1850 code. (Stephen J. Field, David Dudley's brother, was the lawyer who imported New York's code into California.[13]) But the New York code is over 150,000 words long, whereas California's code was just over 50,000 words long. Those disparate lengths mean that comparing all of the California code to all of the New York code is less meaningful than comparing each section in the California code to each section in the New York code, where matching sections will have a similar length.

Having divided the texts according to a historically justified pattern, our next step was to compare each section to every other section and measure the similarity between them. To continue the New York-to-California example, consider the following pairs of sections. The first pair is from the final draft of the New York Field Code. These sections completely abolished prior practice and began to rebuild the procedure system from the ground up (figure 3).

In the theory of Euro-American lawyers, California had no prior practice to abolish, so the code began more simply (figure 4).

---

and David S. Tanenhaus, "The Future of Digital Legal History: No Magic, No Silver Bullets," *American Journal of Legal History* 56, no. 1 (2016): 150–67, doi:10.1093/ajlh/njv017; Dan Cohen, Frederick Gibbs, Tim Hitchcock, Geoffrey Rockwell, et al., "Data Mining with Criminal Intent," white paper, 31 August 2011, http://criminalintent.org.

[12]In each instance we downloaded an entire volume of sessions laws, statutory compilations, or single-volume codes of procedure and then cropped out irrelevant pages, marginalia and footnoted commentary, leaving only the statutory text. After several trials of various implementations of Tesseract (open source) and I.R.I.S. (proprietary) OCR programs, we determined that Nitro Pro PDF, which relies on I.R.I.S. software, offered the best OCR tool for this project. I.R.I.S. provides slightly more accurate readings of nineteenth-century typefaces than Tesseract, and Nitro Pro's implementation makes words, not characters, the fundamental unit of output. The latter feature made cropping between marginalia and the statute more reliable. We removed hyphenated line breaks and standardized spelling for common terms that evolved over the nineteenth century (e.g., *indorsement*).

[13]William Wirt Blume, "Adoption in California of the Field Code of Civil Procedure: A Chapter in American Legal History," *Hastings Law Journal* 17 (1966): 701; Stephen J. Field, *Personal Reminiscences of Early Days in California* (1893), 75–78.

§ 554. The distinction between actions at law and suits in equity, and the forms of all such actions and suits, heretofore existing, are abolished ; and, there shall be in this state, hereafter, but one form of action, for the enforcement or protection of private rights and the redress or prevention of private wrongs, which shall be denominated a civil action.

§ 555. In such action, the party complaining is known as the plaintiff, and the adverse party as the defendant.

Figure 1: Final Report of the Commissioners on Practice and Pleadings (New York, 1850), 225–25, 554–555.

§ 1. There shall be in this State but one form of Civil Action, for the enforcement or protection of private right, and the redress or prevention of private wrongs.

§ 2. In such action the party complaining shall be known as the plaintiff, and the adverse party as the defendant.

Figure 2: 1851 California Laws 51 1–2.

The pairs are obviously related to one another, both in terms of their legal force and in terms of the actual words used.

A common method for measuring the similarity of two documents involves dividing texts up into tokens of consecutive words (called n-grams) and calculating a Jaccard similarity score, defined as ratio between the number of tokens that the two document have in common to the total number of tokens that appear in both documents. We used five-word tokens and shingled them, meaning that for the New York sections above, the first token was "the distinction between actions at," the second token was "distinction between actions at law," and so on. These tokens each contain more meaning than a single word, yet because they are shingled they are robust to changes in the text or noisy OCR. A Jaccard similarity score will always be in a range between 0 (complete dissimilarity) and 1 (complete similarity).[14]

Comparing the section pairs above produces the matrix of similarity scores in table 1.

Table 1: A subset of the section-to-section similarity matrix.

	NY1850-554	NY1850-555	CA1851-001	CA1851-002
NY1850-554		0	0.14	0
NY1850-555			0	0.41
CA1851-001				0
CA1851-002				

As we would expect the first sections (New York 554 and California 1) have a score of 0.14, which indicates that they are similar but have significant differences, while the second sections (New York 555 to California 2) have a much higher similarity score of 0.41 since only a few words were changed. Just as important, when we compare the first section in New York to the second section in California, we get a score of 0; the two sections are nothing like each other.

The aim, then, was to create a triangular matrix like the one above, but with approximately 98,000 rows and 98,000 columns, containing the similarity scores for each possible pair of sections. While this is easy enough to conceptualize, such a matrix is actually quite large, containing about 4.8 billion comparisons. This would take an unreasonable amount of computation time, and most of these comparisons would be unnecessary since each section has no relationship to most other sections. We therefore implemented the minhash/locality sensitive hashing algorithm to detect pairs of possible matches quickly. Instead of comparing all tokens to one another, this algorithm samples tokens from each document to find probable matches, and then Jaccard scores can be calculated for only those probable matches (that is, many of the needless calculations that produce scores of zero get cut out).[15]

[14]The formal definition of the Jaccard similarity for two sets, A and B, is $J(A,B) = \frac{|A \cap B|}{|A \cup B|}$.

[15]We implemented LSH as described in Jure Leskovec, Anand Rajaraman, and Jeff Ullman, *Mining of*

The result was a matrix of similarity scores, with over 45,000 genuine matches. For each section in the corpus, after making some adjustments to remove anachronisms and spurious matches, we were able to identify the section from which it was most likely borrowed.[16] In other words, we had traced the work of the commissioners' scissors and paste-pots through the course of their codes.

3. Patterns of borrowing among Field Code jurisdictions

The computational evidence that we assembled revealed patterns in how law migrated at several different scales of analysis.[17] We used the similarity matrix as the input to three different digital history techniques: network analysis, visualizations, and clustering.

At the broadest scale of analysis, we aggregated the section-to-section borrowings into a summary of how many sections each code borrowed from each other code. We therefore

Massive Datasets, 2nd ed. (Cambridge University Press, 2014), ch. 3, http://www.mmds.org; the algorithm was first described in Andrei Z. Broder, "On the Resemblance and Containment of Documents," in *Compression and Complexity of Sequences 1997: Proceedings*, (IEEE, 1997): 21–29, http://gatekeeper.dec.com/ftp/pub/dec/SRC/publications/broder/positano-final-wpnums.pdf. Other digital humanities projects, most notably *Viral Texts*, have used other means for detecting text reuse. The most prominent of these are algorithms for sequence alignment. (Our "textreuse" package for R also implements the Smith-Waterman local sequencing algorithm, derived from gene sequencing.) Yet the older and simpler LSH algorithm sufficed for our purposes because legal sources are easily divided into discrete sections which can be treated as independent documents. For other approaches, see David Bamman and Gregory Crane, "Discovering Multilingual Text Reuse in Literary Texts," white Paper, Perseus Digital Library (2009): http://www.perseus.tufts.edu/publications/2009-Bamman.pdf; Timothy Allen, Charles Cooney, Stphane Douard, et al., "Plundering Philosophers: Identifying Sources of the Encyclopdie," *Journal of the Association for History and Computing* 13, no. 1 (2010): http://hdl.handle.net/2027/spo.3310410.0013.107; Glenn Roe, Russell Horton, and Mark Olsen, "Something Borrowed: Sequence Alignment and the Identification of Similar Passages in Large Text Collections," *Digital Studies / Le Champ numrique* 2, no. 1 (2010): http://www.digitalstudies.org/ojs/index.php/digital_studies/article/view/190/235; David A. Smith, Ryan Cordell, and Elizabeth Maddock Dillon, "Infectious Texts: Modeling Text Reuse in Nineteenth-Century Newspapers," in *2013 IEEE International Conference on Big Data*, 2013, 86–94, doi:10.1109/BigData.2013.6691675; David A. Smith, Ryan Cordell, Elizabeth Maddock Dillon, et al., "Detecting and Modeling Local Text Reuse," *Proceedings of IEEE/ACM Joint Conference on Digital Libraries* (IEEE Computer Society Press, 2014); Christopher Forstall, Neil Coffee, Thomas Buck, Katherine Roache, and Sarah Jacobson, "Modeling the scholars: Detecting Intertextuality through Enhanced Word-level N-gram Matching" *Digital Scholarship in the Humanities* 30, no. 4 (2015): 503–515; Douglas Ernest Duhaime, "Textual Reuse in the Eighteenth Century: Mining Eliza Haywood's Quotations," *Digital Humanities Quarterly* 10, no. 1 (2016): http://www.digitalhumanities.org/dhq/vol/10/1/000229/000229.html.

[16] We filtered this matrix based on what we knew about the process of borrowing. We removed any match below a threshold that we determined by checking a sample of matches. Because Jaccard similarity scores are symmetric, we also removed anachronistic matches. For instance, a code from 1851 obviously did not borrow from a code from 1877. Furthermore, in chains of borrowing (e.g., NY1850 to CA1851 to CA1868 to CA1872 to MT1895) the latest section might have a high similarity to all of its ancestors, but it was in fact borrowed only from the most recent parent. We therefore filtered the similarity matrix to remove matches within the same code, anachronistic matches, and spurious matches beneath a certain threshold. Then if a section had multiple matches, we kept the match from the chronologically closest code, giving preference to codes from the same state, unless there was a substantially closer match from a different code.

[17] Attention to big and small scales is described in Shawn Graham, Ian Milligan, and Scott Weingart, *Exploring Big Historical Data: The Historian<80><99>s Macroscope* (Imperial College Press, 2015).

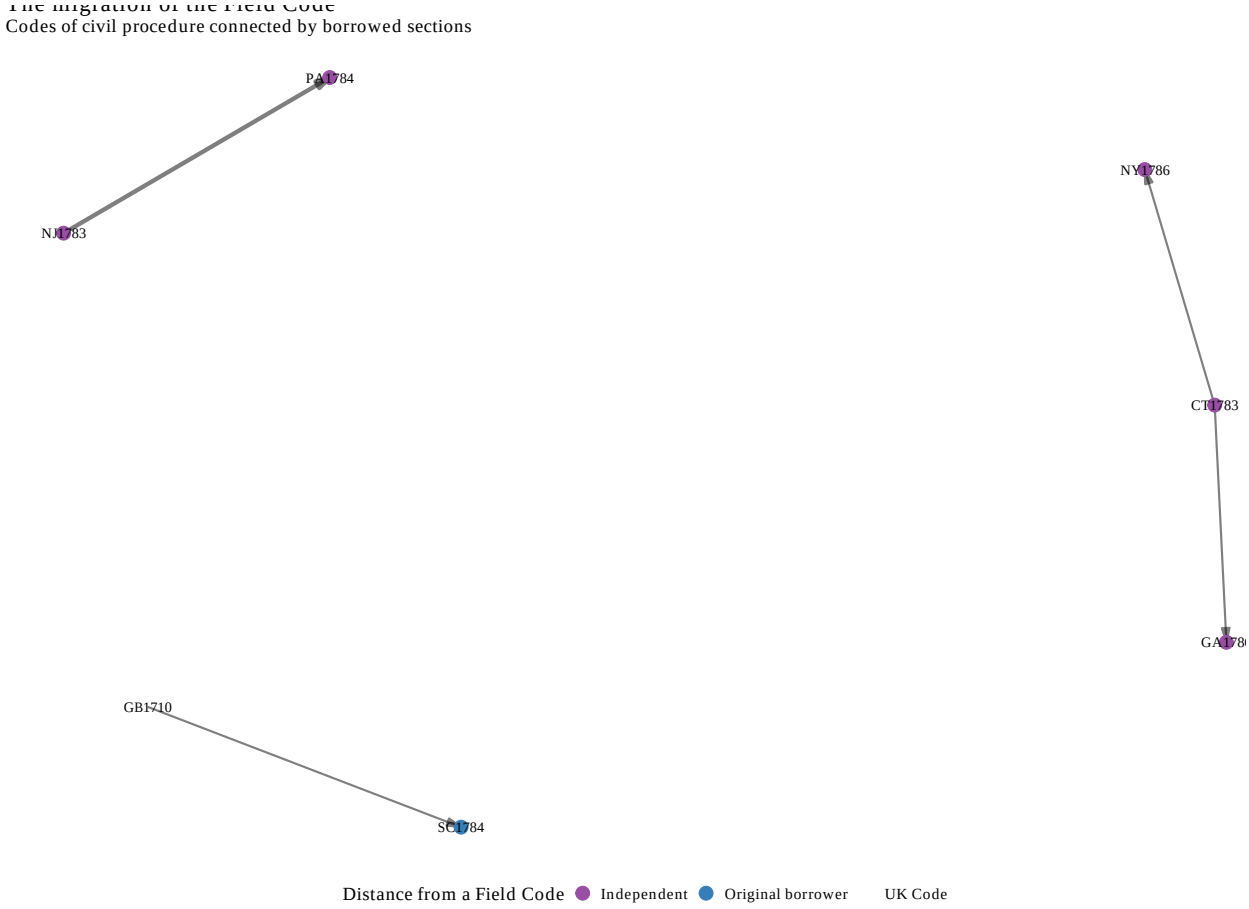

Figure 3: The structure of borrowings among nineteenth-century codes of civil procedure. Note that several versions of New York's Field Code were at the center of the network, while other states such as California and Ohio became centers of regional variations on the Field Code. States that adopted any of the variations on the Field Code became part of a network centered on New York capital.

can show the connections from one code to another. The resulting network graph reveals the genealogy of civil procedure in the United States.

The New York Field Codes, especially the finished draft of 1850, were central to the entire network.[18] New York gave rise to different regional traditions within the procedural

[18]That New York codes are central is obvious from the visualization, but we also confirmed this by formal measures of centrality used in network analysis. A network is simply a list of edges (in our case the number of sections borrowed) between nodes (in our case, the codes). Because even our efforts at determining the best match for each section sometimes attributed a section to an incorrect code, we pruned the edges of the graph so that each code was connected to another code only if it borrowed at least fifty sections or twenty percent of its sections.

Within New York there was a definite chronological progression from the 1848, 1849, 1850, 1851, and 1853 versions of the code, but the development was not chronologically linear. The state legislature enacted the 1848, 1849, and 1851 codes, and these show strong similarities in their relationships. The 1850 and 1853 versions were David Dudley Field's ideal drafts of the code which were never enacted. They were, however,

network. Variations in the Field drafts meant that different states could borrow different versions of the Field Code. Field<80><99>s 1850 draft—never actually enacted in New York—was the primary progenitor of several families of codes in California, Kentucky, Iowa, and Ohio, each of which in turn became major contributors to the law of neighboring states. The 1851 New York code—a small revision to the original 1848 code—became the progenitor of codes for Wisconsin, Florida, North Carolina, and South Carolina. While Field considered the 1850 version to be the definitive, ideal version of the code, all of the New York codes from 1848 to 1853 became models for other jurisdictions. In many cases, the commissions likely used whatever version of the code they had at hand. The Field Code was not a single volume on the shelf, but a series of drafts, any of which might be more accessible in different regions and in different years.

Even later New York codes can be considered a separate family. In 1876 a New York commission produced a new code attempting to consolidate all the case law and statutory amendments subsequent to the 1851 Field Code. David Dudley Field was upset by the changes introduced in this revision. A count by a "friend" of Field's found that only three sentences of the Field Code had carried over word-for-word into the latest edition. With respect to Mr. Field or his "friend," we found that the connection between the codes somewhat stronger than he thought, although his conclusion that, textually, the 1876 code did "not appear to be the same thing as before," remains sound.[19]

Finally, our corpus included a number of statutes which stood outside the Field Code tradition, such as Virginia and West Virginia regulations, statutes from Massachusetts and Maine, and southern codes from Georgia to Louisiana. These statutes show that the dominance of the Field Code was not total, and a number of older jurisdictions remained outside of its ambit.[20] But nearly every jurisdiction established or reconstructed after 1850 became a part of the Field Code network, and no other tradition achieved anywhere near the same coherence across state lines.

In addition to the overview of the relationship between codes, we can also see more detail by visualizing the pattern of borrowings within each code. To illustrate this, we will follow one branch of the Field Code network, beginning with the family started by California's 1850 and 1851 codes.

California's 1850 code, enacted in the period when California was entering the Union as a state, was borrowed almost entirely from New York's 1849 Field Code. The compiler Elisha Crosby did lift one portion from the mixed civilian/common law code of Louisiana, the rules for ordering a new trial to revisit an earlier jury<80><99>s verdict. New trials

printed with wide margins, quality typesetting, and—in the 1850 draft—extensive explanatory notes, all with an eye towards other jurisdictions copying them as a model. Those two codes show stronger similarity to one another than to the enacted drafts.

[19]David Dudley Field, The Latest Edition of the New York Code of Civil Procedure (1878), 21.

[20]Non-Field jurisdictions occasionally exhibited a borrowing relationship within a state or across two states. In in a few unusual instances they contributed to codes which were derived from the Field Code. For example, Alabama's 1852 Code provided a few sections to Tennessee's 1858 code, and some states like Wisconsin copied, along with the Field Code, large passages of pre-code legislation from earlier in the state's history.

were not provided for in the New York Code until the finished draft in 1850. Most of the sections that were not borrowed, as with many of the codes, have to do with parts that describe the system of courts or provide sample forms of pleading or sheriff<80><99>s writs that were peculiar to each state.

When Stephen Field revised California<80><99>s code in 1851, he largely redrafted it from the updated code his brother David Dudley had completed for New York in 1850. This includes the portion of the code on new trials previously borrowed from Louisiana.[21] The remainder of the code was borrowed from the 1850 California code. (Many of the non-matching sections are tables of contents.) Thus California based the majority of its law of civil remedies entirely on New York's code not once, but twice. California made few to no innovations to the Field Code beyond a rearrangement of its provisions and their application to the new state<80><99>s particular system of courts.

The pattern of borrowings in the Washington 1855 code was a rather different case. The Washington code was definitely in the lineage of the 1851 California code, since it borrowed sections from both California directly as well as from Oregon (which was also derived from California). Indiana's 1852 code and Oregon's 1854 code provide the majority of the borrowings. The contiguous bands of borrowings correspond to regulations on judgement borrowed from Oregon and enforcement provisions borrowed from Indiana. This pattern likely came about because one of the Washington code commissioners, Edward Lander, was an Indiana appellate judge from 1850 to 1853, while another commissioner, William Strong, was a justice of the Oregon Supreme Court in the same years. While working on the Washington code, they must have each used the law (and the law books) that they knew best. As a second generation variation on a regional code, the Washington code drew from a variety of sources, even though all these sources basically agreed on the substance of the law.

Finally, we can examine one of the outermost leaves on the family tree of the Field Code in Washington<80><99>s revised code of 1873. The bulk of this code was taken from the early Washington code with only small amendments. The main exception was lengthy set of sections on probate drawn from California's 1872 code. Like many of the last generation codes, the text of the procedure code had stabilized as a local manifestation of a regional tradition. The code was still genuinely a Field Code, with a great deal of similarity to the original New York Field Codes, but its specific form depended on the many edits and rearrangements that code commissioners from several states had made to the text.

So far we have retained the context of the surrounding sections within a particular code. But since our fundamental unit of comparison is section to section, we can use a technique

[21] Final Report (New York, 1850), 804–809, compared to 1851 California Laws 260, 439–441.

called clustering to group sections based on their similarity to one another, regardless of which code they come from. There are innumerable clustering algorithms, but we used the affinity propagation clustering algorithm because its assumptions aligned with the characteristics of our problem. That algorithm finds an "exemplar" item which is most characteristic of the other items in the cluster. That assumption fits nicely with borrowings from the Field Code, where a single section (likely from a Field Code) had many borrowings, but where there could also be innovative sections from other states that might be more influential.[22]

The result was a set of approximately 2,900 clusters which contained at least five sections, though this probably overstates the number of innovative, *ur*-sections in the corpus. The biggest cluster, which concerned the use of affidavits in pleading, contained 103 sections. Within each cluster, we organized the sections chronologically. We were thus able to see the development of the law from jurisdiction to jurisdiction over time. This method provides historians with a way of noticing small changes in the wording and substance of the law. Most discussions of algorithmic reading have focused on "distant reading," or have balanced the claims of distant reading by using it as a means to enable close reading. This method of clustering, however, is a kind of algorithmic close reading. By deforming the texts—taking them out of the context of the codes and putting them into the context of their particular variations—we are able to pay attention to those variations.[23]

Take provisions regulating witness testimony as an example. At common law, parties and interested witnesses were not permitted to testify in their own causes. Field's Code reversed this rule, expanding witness competency as widely as possible: any person "having organs of sense" was to be admitted as a witness in New York, with only the insane and very young children possibly exempted. As the code migrated West, however, legislators added racial exclusions to Field's list. The cluster of sections in the appendix documents how California's codifiers grafted earlier prohibitions from midwestern states into Field's Code. Many other codes then evidenced a remarkable uniformity with California's text (which later changed only to add "Mongolian" to the list of races). Iowa's Code had a more minor influence, and Wyoming developed the only truly original section which made it explicit that the exclusion was based on the racist assumption that non-white peoples were infantilized, a connection only implicit in other jurisdictions.

Uniformity in the law, as shown through these clusters, is just as instructive as variation. The most significant clusters that we investigated related to the collection of debts. These were clusters which went against the typical pattern we observed. While most clusters exhibited regional variation as they grew more distant from the Field Code, clusters having

[22]Brendan J. Frey and Delbert Dueck, "Clustering by Passing Messages Between Data Points," *Science* 315 (2007): 972–976, doi: 10.1126/science.1136800; Ulrich Bodenhofer, Andreas Kothmeier, and Sepp Hochreiter, "APCluster: An R package for Affinity Propagation Clustering," *Bioinformatics* 27 (2011): 2463–2464, doi: 10.1093/bioinformatics/btr406. Even though the affinity propagation algorithm did not fully converge with our dataset, it did an adequate job clustering the documents. Because there was an exemplar section for each cluster, we were able to merge clusters whose exemplars had a high Jaccard similarity score.

[23]Lisa Samuels and Jerome McGann, "Deformance and Interpretation" *New Literary History* 30, no. 1 (1999): 25–56; Mark Sample, "Notes toward a Deformed Humanities," blog post, 2 May 2012: http://www.samplereality.com/2012/05/02/notes-towards-a-deformed-humanities/; Ramsay, *Reading Machines*, 32–57.

to do with creditors' remedies were almost completely uniform across the American West and South. No single section of the code announced its preference for creditors' rights; rather, the acceleration of creditors' remedies resulted from the combination of several sections. In New York's original enacted code from 1848, 107 required a defendant to answer the complaint within twenty days, instead of at the next court session (which in some cases could have been as far as three months away); 202 provided for default judgment as a matter of course, issued by a clerk without a judicial order if no adequate answer was received within the twenty days; 128–133 abolished fictitious pleadings and required answers to state true facts verified by a defendant's oath, all so that no trial would delay the enforcement of uncontestable obligations; finally, the code abolished a traditional thirty-day waiting period between issue of judgment and commencement of execution. These provisions dealt with what merchants and capitalists perceived as an abuse of the common law system, where defendants in cases of debt could stretch out enforcement of debt collection for as long as two years. The Field Code's summary judgment brought down the time of debt collection to a matter of weeks. The code thus traded the rhythms of agriculture for the rhythms of merchant finance.[24]

Clustering each of these sections reveals that western states along with the former Confederate states of South Carolina, North Carolina, and Florida copied each provision almost exactly. Midwestern and Upper South states that had already developed and maintained commercial ties to Chicago and New Orleans by 1850 varied the New York rules, sometimes by requiring answers only in term time, or permitting only a judge to decree default judgment rather than a clerk, in either case effectively stretching out enforcement and making a formal trial more likely. But in the Reconstruction South, and in the West over the same period, regardless of whether a jurisdiction abolished chancery or not, regardless of the racial exclusions it may have placed on witness testimony, the provisions on debt collection remained unchanged. When it came to creditors' remedies, the law of New York became the law of the land.

Conclusion

By addressing our historical questions to a sufficiently large but narrowly defined corpus of sources, we benefited from a useful symbiosis of traditional and digital historical methods. Our computational methods produced useful historical knowledge because they were carefully tailored to what we knew about the data from traditional historical work. We knew that code commissioners worked with "the scissors and paste-pot," as critics complained, and we examined codes in the archives which showed how commissions literally marked up the legislation of other states. While we think that one of the most useful things about digital history is its ability to start with large corpora and then figure out what was interesting from the past, we have shown how digital history can also operate by starting with specific historical questions rather than particular sources. We have shown how a collection of methods from computer science, including minhash/locality-sensitive hashing, affinity propagation clustering, and network analysis, along with the concept of text deformance from literary studies, can be used to good effect in tracking the changes

[24]First Report (New York, 1848), 197.

in the law, as well as any other discursive field whose texts can be readily divided into sections. Finally we have shown how it is possible to work on different scales, using network analysis, visualization, and algorithmic close reading, and thus to gain both a broad overview of the law's migration, as well as a highly detailed view of the changes in the law.

The history of codification on the American periphery challenges foundational assumptions about American federalism. Scholars commonly speak of regulating at <80><9c>the state level<80><9d> imagining an equality between state sovereignties that exists in tension only with <80><9c>the federal level.<80><9d> But the history of legal practice and civil remedies is one in which the localism fostered by common law practice rapidly gave way to uniform regulations promulgated by New York trial lawyers without the slightest interference of the federal government.[25] The history of the Code also has important implications for recent scholarship seeking to unearth a long tradition of <80><9c>administrative law<80><9d> among the states before the twentieth century. These accounts have largely focused on administrative adjudication or discretionary regulation within a narrow domain, such as customs houses, but have so far neglected the most widespread and significant instance of nineteenth-century administrative lawmaking in America—the spread of remedial codes through extra-legislative commissions.[26] While these histories have sought to demonstrate that nineteenth-century Americans could be quite comfortable with administrative law, accepting it as a normal part of the constitutional order, this chapter has shown how lawmaking by commission generated significant political controversy and raised grave questions about popular sovereignty that over time were merely dodged rather than answered.

This study thus gets at the heart of lawmaking in U.S. history. Lawyers and judges, politicians and newspaper editors warred over whether codes that were drafted by commissioners and borrowed wholesale from beyond a jurisdiction's borders were actually democratic. Codifiers responded by transmuting democratic theory into support for a remedial code that elected legislators had neither the time nor inclination to read. Popular support for commercial development was taken to indicate popular support for New York's civil remedies, especially the cheapened and accelerated collection of debts. In

[25]The equality of the states is a foundational assumption in the much-criticized idea of the states as laboratories for regulatory experimentation. The states-as-laboratories idea emerged from *New State Ice Co. v. Liebmann*, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting). See James A. Gardner, <80><9c>The <80><98>States-as-Laboratories<80><99> Metaphor in State Constitutional Law,<80><9d> *Valparaiso University Law Review* 30 (1996), 475. For a collection of refutations, see Brian Galle & Joseph Leahy, <80><9c>Laboratories of Democracy? Policy Innovation in Decentralized Governments,<80><9d> *Emory Law Journal* 58 (2009), 1333. Even as federalism scholars vigorously refute the idea of states as <80><9c>laboratories<80><9d> for regulative experimentation, they continually pose <80><9c>the federal<80><9d> to <80><9c>the state<80><9d> level, with an assumed equality among the numerous sovereignties in the latter category. See, for instance, James E. Fleming & Jacob T. Levy, eds., *Federalism and Subsidiarity* (NYU 2014); Heather Gerken, *Beyond Sovereignty, Beyond Autonomy: A Nationalist<80><99>s View of Federalism<80><99>s Future* (forthcoming).

[26]See Daniel Ernst, *Tocqueville's Nightmare: The Administrative State Emerges in America, 1900–1940* (Oxford: Oxford University Press, 2014); Jerry L. Mashaw, *Creating the Administrative Constitution: The Lost One Hundred Years of American Administrative Law* (Yale Law Library, 2012); Gautham Rao, *National Duties: Custom Houses and the Making of the American State* (University of Chicago, 2016).

many jurisdictions the exact language of the Field Code remains on the books, and its basic provisions for civil procedure are in force throughout the United States. Without too much exaggeration we might say that our method has revealed the spine of modern American legal practice.

Appendix: Clustering sections involving witness exclusions

This sample cluster brings together sections that involve race-based exclusions from witness testimony in civil trials, with commentary on the history of the variations.

Disqualifying the testimony of non-white parties and witnesses was not new to the procedure codes. The following laws from midwestern states with large free black populations would be echoed in adaptations of the Field Code.

Code	Section
OH 1807	That no black or mulatto person or persons, shall hereafter be permitted to be sworn or give evidence in any court of record, or elsewhere in this state, in any cause depending, or matter of controversy, where either party to the same is a white person, or in any prosecution, which shall be instituted in behalf of this state, against any white person.
IA 1839	A negro, mulatto, or Indian, shall not be a witness in any court or in any case against a white person.
IN 1843	No negro, mulatto or Indian, shall be a witness, except in pleas of the state against negroes, mulattoes, or Indians, and in civil causes where negroes, mulattoes, or Indians alone are parties: every person other than a negro having one-fourth part of negro blood or more, or any one of whose grandfathers or grandmothers shall have been a negro, shall be deemed an incompetent witness, within the provisions of this article.

Racial disqualifications were introduced to the Field Code tradition in California, first in Elisha Crosby<80><99>s draft of 1850, then in Stephen J. Field<80><99>s draft of 1851. Many western states copied Stephen Field<80><99>s provision.

Code	Section
CA 1850	306. No black, or mulatto person, or Indian, shall be permitted to give evidence in any action to which a white person is a party, in any Court of this State. Every person who shall have one eighth part or more of negro blood, shall be deemed a mulatto; and every person who shall have one half Indian blood, shall be deemed an Indian.

Code	Section
CA 1851	394. The following persons shall not be witnesses: lst. Those who are of unsound mind at the time of their production for examination; 2d. Children under ten years of age, who appear incapable of receiving just impressions of the facts respecting which they are examined, or of relating them truly: and; 3d. Indians, or persons having one fourth or more of Indian blood, in an action or proceeding to which a white person is a party: 4th. Negroes, or persons having one half or more Negro blood, in an action or proceeding to which a white person is a party.
OR 1854	6. The following persons shall not be competent to testify: 1. Those who are of unsound mind, or intoxicated at the time of their production for examination; 2. Children under ten years of age, who appear incapable of receiving just impressions of the facts respecting which they are examined, or of relating them truly; 3. Negroes, mulattoes and Indians, or persons one half or more of Indian blood, in an action or proceeding to which a white person is a party.
WA 1855	293. The following persons shall not be competent to testify: 1st. Those who are of unsound mind, or intoxicated at the time of their production for examination. 2d. Children under ten years of age, who appear incapable of receiving just impressions of the facts, respecting which they are examined, or of relating -them truly. 3d. Indians, or persons having more than one half Indian blood, in an action or proceeding to which a white person is a party.
UT 1859	215. The following persons shall not be competent to testify: 1. Those who are of unsound mind or intoxicated at the time of their production for examination. 2. Children under ten years of age, who appear to be incapable of receiving just impressions of the facts respecting which they are examined or of relating them truly. Negroes, mulattos, and Indians, or persons having one fourth of negro or Indian blood, in an action or proceeding to which a white person is a party, but shall not be disqualified from testifying against another.

Code	Section
NV 1861	342. The following persons shall not be witnesses: First. Those who are of unsound mind at the time of their production for examination. Second. Children under ten years of age, who, in the opinion of the court, appear incapable of receiving just impressions of the facts respecting which they are examined, or of relating them truly. Third. Indians, or persons having one half or more of Indian blood, and negroes, or persons having one half or more of negro blood, in an action or proceeding to which a white person is a party. Fourth. Persons against whom judgment as been rendered upon a conviction for a felony, unless pardoned by the governor, or such judgment has been reversed on appeal.
ID 1864	352. The following persons shall not be witnesses: First. Those who are of unsound mind at the time of their production for examination. Second. Children under ten years of age, who, in the opinion of the court, appear incapable of receiving just impressions of the facts respecting which they are examined, or of relating them truly. Third. Chinamen or persons having one-half or more of China blood; Indians, or persons having one-half or more of Indian blood, and negroes, or persons having one-half or more of negro blood, in an action or proceeding to which a white person is a party. Fourth. Persons against whom judgment has been rendered upon a conviction or felony, unless pardoned by the governor, or such judgment has been reversed on appeal.
AZ 1865	396. The following persons shall not be witnesses: 1. Those who are of unsound mind at the time of their production for examination. 2. Children under ten years of age, who appear incapable of receiving just impressions of the facts respecting which they are examined, or of relating them truly; and, 3. Indiana or persons having one-half or more of Indian blood, in an action or proceeding to which a white person is a party. 4. Negroes, or persons having one-half or more negro blood, in an action or, proceeding to which a white person is a party.

Code	Section
CA 1868	394. The following persons shall not he witnesses: First. Those who are of unsound mind at the time of their production for examination. Second. Children under ten years of age, who, in the opinion of the court, appear incapable of receiving just impressions of the facts respecting which they are examined, or of relating them truly. Third. Mongolians, Chinese, or Indians, or persons having one-half or more of Indian blood, in an action or proceeding wherein a white person is a party. Fourth. Persons against whom judgment has been rendered upon a conviction for a felony, unless pardoned by the governor, or such judgment has been reversed on appeal.

Kentucky differed from other code states by making no distinction between incompetency (an absolute bar) and privilege (which might be waived). Kentucky also maintained a strict disqualification of parties and interested witnesses while other Field Code states made parties at least partially competent to stand examination.

Code	Section
KY 1851	568. The following persons shall be incompetent to testify: 1. Persons convicted of a capital offense, or of perjury, subornation of perjury; burglary, robbery, larceny, receiving stolen goods, forgery, or counterfeiting. 2. Infants under the age of ten years, and over that age, if incapable of understanding the obligation of an oath. 3. Persons who are of unsound mind at the time of being produced as witnesses. 4. Husband and wife, for or against each other, or concerning any communication made by one to the other, during the marriage, whether called as a witness while that relation subsisted or afterwards. 5. An attorney, concerning any communication made to him by his client in that relation, or his advice thereon, without the client<80><99>s consent. 6. Persons interested in an issue, in behalf of themselves, and parties to an issue, in behalf of themselves or those united with them in the issue. 7. Negroes, mulattoes, or Indians, in any action or proceeding where a white person, in his own right, or as representative of a white person, is a party, except in actions brought to recover a penalty or forfeiture for a violation of law, against a negro, mulatto, or Indian.

Code	Section
MT 1865	320. The following persons shall be incompetent to testify: First, Persons who are of an unsound mind at the time of their production for examination. Second, Children under ten years of age who appear incapable of receiving just impressions of the facts respecting which they are examined or of relating them truly, but the court in its discretion may allow such children to testify, and the facts herein enumerated shall go to their credibility. Third, Husband or wife for or against each other, or concerning any communication made by one to the other during the marriage, whether called as a witness while that relation existed or afterwards. Fourth, An attorney concerning any communication made to him by his client in that relation, or his advice thereon, without the client<80><99>s consent. Fifth, A clergyman or priest concerning any confession made to him, in his professional character, in the course of discipline enjoined by the church to which he belongs, without the consent of the person making the confession. Sixth, A negro, Indian, or Chinaman, where the parties to the action are white persons, but if the parties to an action or either of the parties is an Indian, negro, or Chinaman, a negro may be introduced as a witness against such negro, an Indian against such Indian, or a Chinaman against such Chinaman. A negro within the meaning of this act is a person having one-eighth or more of negro blood, an Indian is a person having one-half or more of Indian blood, and a Chinaman is a person having one-half or more Chinese blood.

Iowa's code began by defining competency purely in terms of understanding the legal oath, and in the same section it barred non-white testimony (even if non-white actors could understand the oath). Wyoming followed the same tack, but softened the racial bar by adopting the same standard used for children: only those adjudged incapable of perceiving and relating facts were barred from testifying.

Code	Section
IA 1851	2388. Every human being of sufficient capacity to understand the obligation of an oath is a competent witness in all cases both civil and criminal except as herein otherwise declared. But an indian, a negro, a mulatto or black person shall not be allowed to give testimony in any cause wherein a white person is a party.

Code	Section
NE 1855	2388. Every human being of sufficient capacity to understand the obligation of an oath is a competent witness in all cases both civil and criminal except as herein otherwise declared. But an indian, a negro, a mulatto or black person shall not be allowed to give testimony in any cause wherein a white person is a party.
WY 1870	325. Every human being of sufficient capacity to understand the obligations of an oath, is a competent witness in all cases, civil and criminal, except as otherwise herein declared. The following persons shall be incompetent to testify: First, Persons of unsound mind at the time of their production. Second, Indians and negroes who appear incapable of receiving just impressions of the facts respecting which they are examined, or of relating them intelligently and truly. Third Husband and wife, concerning any communication made by one to the other during the marriage, whether called as a witness while that relation exists or afterwards. Fourth, An attorney, concerning any communication made to him by his client in that relation or his advice thereon, without the client<80><99>s consent in open court or in writing produced in court. Fifth, A clergyman or priest, concerning any confession made to him in his professional character in the course of discipline enjoined by the church to which he belongs, without the consent of the person making the confession.

States that did not borrow Field's evidence code nevertheless borrowed prohibitions of non-white testimony. Two distinct strands emerge in legislation from the Deep South as well as the Upper South and Midwest.

Code	Section
MS 1848	All negroes, mulattoes, Indians, and all persons of mixed blood, descended from negro or Indian ancestors, to the third generation, inclusive, though one ancestor of each generation may have been a white person, shall be incapable in law, to be witnesses in any case whatsoever, except for and against each other.
AL 1852	2276. Negroes, mulattoes, Indians, and all persons of mixed blood, descended from negro or Indian ancestors, to the third generation inclusive, though one ancestor of each generation may have been a white person, whether bond or e, must not be Witnesses in any cause, civil or criminal, except for or against each other.

Code	Section
TN 1858	3808. A negro, mulatto, Indian, or person of mixed blood, descended from negro or Indian ancestors, to the third generation inclusive, though one ancestor of each generation may have been a white person, whether bond or free, is incapable of being a witness in any cause, civil or criminal, except for or against each other.
DC 1857	A negro shall be a competent witness for or against a negro in any criminal proceeding, and shall be a competent witness in any civil case to which only negroes are parties, but not in any other case.
VA 1860	A negro or indian shall he a competent witness in a case of the commonwealth for or against a negro or indian, or in a civil ease to which only negroes or indians are parties, but not in any other case.
IL 1866	A negro, mulatto or Indian shall not be a witness in any court, or in any case, against a white person. Any person having one-fourth part negro blood shall be adjudged a mulatto.