



EDUCATING ABOUT THE PAST IN HOPES OF A MORE EQUITABLE FUTURE:

Identifying, Building, and Using Collections as Data for Social Justice

Amanda Henley, Lorin Bruckner, Sarah Carrier, James Dick,
Hannah Jacobs, and Matt Jansen

*Libraries are using their unique collections and staff expertise to make digital collections accessible for computational research. The Always Already Computational: Collections as Data project provided national forums and resources including the Santa Barbara Statement on Collections as Data¹ to guide institutions in preparing collections for computational research. The Collections as Data: Parts to Whole project² has funded two cohorts of institutions to create their own collections as data. This paper will outline the experience of librarians and scholars at the University Libraries at the University of North Carolina at Chapel Hill (UNC-Chapel Hill) in completing a collections as data project centered on social justice. The project, *On the Books: Jim Crow and Algorithms of Resistance*³ (OTB), was one of the projects funded by Collections as Data Parts to Whole. OTB created a plain text corpus of over 100 years of North Carolina session laws and used machine learning to identify Jim Crow laws. The project seeks to educate about the past, in hopes of a more equitable future.*

PROJECT ORIGIN

OTB came about through a combination of events and initiatives. In recent years, the UNC-Chapel Hill University Libraries has been growing capacity to collaborate on text analysis. Staffing in this area was increased by first sharing staff across departments, and ultimately reworking a couple of positions.

Libraries staff were members on the IMLS-funded project “Digging Deeper, Reaching Further, Libraries Empowering Users to Mine the HathiTrust Digital Library Resources.”⁴ The curriculum for that project was improved as it was taught in multiple iterations, and

many Libraries staff members were introduced to text analysis at one of the four, day-long workshops held between 2016-2018.

Staff members at the University Libraries have made deliberate efforts to work across the Libraries and across campus. A Special Collections Digital Scholarship Working Group encourages collaboration across library divisions. This group was created out of recognition that, “As more archival collections are digitized or born-digital, the work of archivists increasingly overlaps with the work of librarians who are responsible for research data and digital scholarship.”⁵ The group developed closer working relationships while learning from one another and investigating collaboration opportunities, including new uses of digital collections as data. The goal is to create a “collaborative culture that more comprehensively supports library collections and more holistically serves library users.”⁶

Finally, librarians co-founded a Carolina Seminar called, “Transforming Inquiry Through Digital Text Analysis”, consisting of an interdisciplinary group of librarians and humanities scholars who met weekly to discuss text analysis, work on projects, and plan programs to bring in outside experts.

The idea for OTB came from Sarah Carrier, North Carolina Research & Instruction Librarian in Special Collections, who had received a question from a high school social studies teacher seeking a comprehensive listing of North Carolina Jim Crow laws. Ms. Carrier realized that beyond the work of Pauli Murray through 1950, no such inventory exists, and the potential impact of such a resource would be invaluable. Consulting the special collections print volumes of laws passed by the General Assembly of North Carolina, each quite lengthy and in total numbering in the dozens, proved neither efficient nor feasible. While these volumes had been digitized and made available online,⁷ searching each volume’s PDF for keywords had not only its faults and inconsistencies, but without a single text corpus of the laws passed during a specific date range, the potential time commitment was impractical. Ms. Carrier had attended one of the day-long text analysis workshops offered for Library staff, and her interest had been piqued as to whether text analysis could be used to identify the Jim Crow laws. The timing coincided with the hiring of new staff in the libraries to support digital scholarship, so her inquiry was referred to the newly hired Digital Scholarship Specialist, who brought the idea to the Carolina Seminar. The group quickly decided to pilot new methods to investigate feasibility. The idea was discussed more at the Special Collections Digital Scholarship Working Group.

The OTB project served as a catalyst for the University Libraries to address the organizational challenges introduced by collections as data and clarify the methods, roles, and services surrounding the creation and maintenance of these collections. “A successful turn towards collections as data development requires inclusive organizational experimentation spanning archivists, technologists, subject experts, catalogers, and more.”⁸

COLLECTIONS AS DATA PRINCIPLES

Goals for OTB were largely informed by principles from The Santa Barbara Statement on Collections as Data.⁹ The project aimed to lower barriers to use by providing materials scoped to users with a variety of technical expertise, including K-12 students and teachers, those interested in doing general legal research, and those interested in creating collections as data or doing text analysis. Thorough documentation and data transparency were key project goals. “Libraries do not often provide access to the scripts that generate collection derivatives, access to processes for cleaning or subsetting data, access to custom schema that have been used, indications of how representative digital holdings are relative to overall holdings, nor is the quality of data typically indicated.”¹⁰ All scripts that were written to create the OTB corpus are accessible on a GitHub site,¹¹ along with documentation. The white paper describes processes and workflows, lists all volumes included in the corpus, and provides an OCR accuracy assessment.¹² By using open-source software and providing documented code, OTB offers a free and customizable framework so the approach can be implemented in other contexts.

Lastly, the project was guided by ethical commitments. “These commitments work against historic and contemporary inequities represented in collection scope, description, access, and use.”¹³ OTB attempts to document the under-represented injustices that are part of the South’s legacy. “In order to fully understand the shadow that Jim Crow continues to cast over us today, it is necessary to know how ostensibly democratic government at all levels and in all places used law to advance white interest while disadvantaging the interests of African Americans and other minorities.”¹⁴

The project used an algorithmic approach to discover racist laws during the Jim Crow Era. The project team acknowledges the limits of this approach. “The number and existence of statutes does not create the full measure of the de jure discrimination by government in North Carolina. It was the operation of those laws, the implementation, the interpretation of those laws which almost matter more... than the number of statutes themselves.”¹⁵ The project team has also thought critically about algorithms.

ALGORITHMS OF RESISTANCE

All computer interaction is mediated by algorithms, the human-defined processes used by machines to complete tasks. Algorithms determine search results, social media feeds, GPS routes, targeted advertisements, and so on. They do this by following “a set of instructions, rules, and calculations designed to solve problems.”¹⁶ If programmers are careless—or worse yet, malicious—when writing algorithms, the results can be damaging for minority populations. Safiya Noble calls these “algorithms of oppression” because they support existing negative views of communities of color and sow further division through “digital redlining.”¹⁷ Noble’s work focuses on the racialized ways that women in particular are represented in online search results: searches for “black girls” and “black beauty” yield racist, sexist, derogatory material while “white girls” and “beauty” yield stereotypical, no less racist or sexist, ideals of beauty as both white and overtly feminine in appearance.

Algorithms, as Noble has shown and Zeynep Tufekci has argued, have very real consequences: they are “computational agents who are not alive, but who act with agency in the world.”¹⁸ These “computational agents” act out the intentions (and unchecked biases) of their most-often white, male programmers. The lack of diversity in computer programmers leads to the oversight of algorithmic biases in code. Hiring more women and people of color in technology roles is the best solution, but there are also other ways to resist algorithmic oppression.

Joy Buolamwini and the Algorithmic Justice League (AJL), including design justice theorist and activist Sasha Costanza-Chock, are carrying out algorithmic resistance through their focus on exposing and eliminating biased and harmful AI technologies.¹⁹ They work against what Ruha Benjamin has termed “discriminatory design.” We draw from their examples to outline a set of guidelines in use by OTB for algorithmic resistance:

- First, we *acknowledge that we live in a society historically built on racist ideologies and practices* and while we do not *intend* to systematically exclude marginalized people, we recognize that *any algorithmic platform, analysis, or tool we create or use may still impose harm, whether small or significant.*²⁰
- Despite these challenges, we seek to *use algorithms to resist discriminatory policies and to enact positive change* by:
 - *Centering those whose knowledge and experiences directly connect to our work;*²¹ and *Enabling community agency and control.*²² We are engaging instructors at the K-12 and college levels with the aim of making them aware of these laws and working with them to identify potential pedagogical uses. Through education and community engagement, we make students, scholars, and the general public aware of our work and our uses of algorithms. In addition to promoting awareness, we seek feedback on how our work with algorithms may be impacting others and how we can address that impact.
 - *Considering our own identities, assumptions, and privileges.*²³ We recognize that the technical part of our team is overwhelmingly white, so we rely on the advice of experts in history and African American studies to inform our work.
 - *Ensuring transparency in our uses and creation of algorithms.*²⁴ We have posted our documented code on GitHub and are creating step-by-step tutorials and examples.
 - *Keeping humans in the loop.*²⁵ We used a combination of automated and manual processes to prepare the corpus to the best of our ability. We trained the algorithms we use to identify Jim Crow Laws with training sets identified and refined by scholars in multiple disciplines.
 - *Prioritizing process over product.*²⁶ Our team spent a year developing a thorough process to reduce as many corpus errors as possible.

The same algorithm can be used for oppression or resistance. How we design and use algorithms, and how they impact those vulnerable to discriminatory policies, determines whether they *enact* oppression or resistance. OTB is an attempt to work, however imperfectly, for that resistance.

WORKFLOW & PROCESSES

A summary of the OTB workflow and processes follow. Additional information is available in the OTB white paper.²⁷

Creating an inventory of North Carolina Jim Crow laws required first and foremost a collection of digitized pages from North Carolina law books. The collection used for OTB was digitized between 2009-2011 under the IMLS grant Ensuring Democracy through Digital Access,²⁸ which digitized state documents and made them accessible through the Internet Archive. OTB focused on session laws passed during the Jim Crow era (1865-1967), a total of 96 volumes.

While the digitized collection was crucial to the project's success, image quality varied and presented OCR challenges. Images included indexes, title pages, tables, etc. Likewise, some text was skewed on the images, and pages contained marginalia and headers that could interfere with analysis. To improve OCR results, each image was adjusted or rotated, as needed, using the Python Imaging Library, Pillow. Prior to OCR, each image was trimmed to isolate the main text.

Once image adjustment was complete, Python-tesseract was used for OCR and the OCR quality was assessed. North Carolina session laws are organized by chapter and section. To identify Jim Crow laws, the individual laws (or sections) needed to be separated. Much of the time and effort required to build the corpus was spent isolating individual sections for analysis. This lengthy process involved an initial split of the volumes into chapters and sections, numerous iterations of corrections, generation of the final split files, and an error assessment.

The initial splits used pattern matching with regular expressions to identify the beginning of new chapters and sections. After these splits were made, likely problems were identified, such as pattern matching errors, OCR errors, image scan errors, etc. Chapter splits that were not identified because of text missing from the digital images or lost in the OCR process were transcribed. Some volumes could not be split automatically, including those that identify chapters with Roman numerals, those with chapter headings in the margins, and, lastly, some volumes that simply did not split well. Those volumes were split manually. Due to the extremely time-intensive nature of this work, some split errors remained in the first version of the corpus. The corpus of all session laws contains 53,515 chapters and 297,790 sections. Initially, 27,327 chapter/section split errors were identified. 89.7% of the errors were corrected for version 1 of the corpus. Additional work was done to correct the remaining known errors during phase 2 of the project. Version 2 of the corpus is forthcoming.

The team investigated both unsupervised machine learning (topic modelling) and supervised machine learning methods to identify Jim Crow laws. Supervised machine learning was found to be most effective. Supervised machine learning requires a training set, which is used to “teach” the computer how to identify Jim Crow laws using labeled examples of both Jim Crow laws and laws that are not Jim Crow. The training set was comprised of Jim Crow laws identified by legal scholars Pauli Murray²⁹ and Richard Paschal,³⁰ as well as laws selected through random and targeted sampling that were classified as Jim Crow or not Jim Crow by William Sturkey and Kimber Thomas, disciplinary scholars on the project team. An XGBoost model was selected to predict Jim Crow laws.³¹ The model was calibrated using 80% of the training set, and 20% was used to assess performance; the final model used all labeled laws to make predictions. Laws with a calibrated probability of 90% or more were classified as Jim Crow laws, identifying 775 Jim Crow laws (in addition to those in the training set). Analysis using version 2 of the corpus and the Jim Crow laws identified by the first version, then confirmed or corrected by experts, is underway.

DEFINING JIM CROW LAWS

A central challenge for this project has been arriving at a working definition of a Jim Crow law. OTB has drawn heavily from Rev. Dr. Pauli Murray's seminal work on race and law in the United States, *States' Laws on Race and*

Color, published in 1950. In this work, Murray uses the term “Jim Crow” with relative infrequency, generally preferring the term “segregation law”. Murray cites the US Supreme Court case *Henderson v. US* (1949) as “The Legal Case against Segregation”, accepting the case as an authoritative text on the subject.³² While Murray does not define segregation or Jim Crow, *Henderson* provides definitions for both, and Murray broadly follows these definitions throughout her work. The *Henderson* case defines “segregation” by stating that the type of segregation being discussed is “Racial segregation under compulsion of law” or “according to their color”³³ and that such practices “brand[s] the Negro with the mark of inferiority and assert[s] that he is not fit to associate with white people.”³⁴ *Henderson* defines Jim Crow laws as those that create conditions which “strengthen and enhance” “racial prejudice.”³⁵ Combining these definitions of segregation laws and Jim Crow laws, OTB has come to a working definition of a Jim Crow law as *any law designed to encourage or enforce separate treatment of any group on the basis of race*. This working definition of a Jim Crow law is necessarily broad and as such is subject to different interpretation by experts. As described by disciplinary scholar William Sturkey, “In many of the laws, there’s no question about the intent—the law segregating schools is clearly a Jim Crow law. Other laws might be up for interpretation. We were looking for anything that required racial segregation or stratification in any way.”³⁶

Building the training set required interpretation of the laws. Interpretation depends on the reviewer’s historical knowledge and understanding. Due to the subjective nature of interpretation, there was some disagreement among the reviewers’ assessments. Team member and attorney James Dick reviewed the training set during the second phase of the project to ensure consistency across the training set. This work led to a more nuanced understanding of the different types of Jim Crow laws.

These laws created a “complex system of racial apartheid designed to create a racial hierarchy throughout the American South”, featuring “complex and creative ways that people came up with to keep races separated.”³⁷ These laws built on one another to create a structure for apartheid. Laws initiating segregation necessitate the creation of additional laws—the *structural progeny* of Jim Crow. Structural progeny, in this context, refers to laws that did not, in and of themselves, establish a structure for segregation or discrimination but were necessary for the continued enforcement or administration of that structure. Such laws are extremely common relative to laws establishing new segregated institutions and illustrate the tremendous complexity and cost of maintaining the Jim Crow system. Structural progeny frequently included laws calling for payment to be rendered or funds to be allocated to existing (segregated) institutions, or to the staff of those institutions, such as teachers at individual (segregated) schools.

Close review of the Jim Crow laws revealed that they could be categorized. Three categories were identified: *explicit*, *implicit*, and *extrinsic*. Explicit Jim Crow laws are those that clearly called for or enforced segregation or discrimination based on race. These laws cover a wide variety of archetypical Jim Crow topics. Examples include laws requiring separate restrooms, railway cars, and schools for members of different races, and laws preventing interracial marriages. These laws generally included explicitly racialized language such as “white”, “colored”, and “Indian” and were thus most readily identified. Also classified as explicit were laws that did not call for additional racial segregation but were the structural progeny of Jim Crow laws. A good example of how this structural progeny-type Jim Crow law appears is 1903 Session Law Chapter 364, Section 1, which provides in relevant part that:

[T]he Treasurer of the county school fund of Person County be and he is hereby authorized to pay to Narcissa V. Mason, a colored school-teacher of School District No. 5, for [the] colored race, in Cunningham Township, Person County, the sum of eighty dollars for services as teacher of a public school in said district from October 30, 1899, to March 13, 1900.³⁸

This law, while not establishing segregation, only existed because it was necessary to manage an existing racially segregated institution, namely a racially segregated public school. This type of law was categorized as explicit, as it uses racially explicit language and is the legal progeny of previous, explicit school segregation laws and can thereby be interpreted as enforcing previous segregation laws.

Implicit Jim Crow Laws are those that contain no clear racialized language to indicate an intent to segregate or discriminate, but are broadly understood, due to their subject matter, as including an intent to discriminate against and/or segregate on the basis of race. Such laws were categorized as implicit Jim Crow laws. A good ex-

ample is 1929 Session Law Chapter 195, Section 11,³⁹ which provides for the establishment of a board to determine eligibility of prospective students for a public school. Taken on its face value, the text of Section 11 would not indicate that it is a Jim Crow law. However, when taken in context with the chapter heading of Chapter 195, which states explicitly that its intent is “to codify and prescribe the racial qualification of those seeking admission into Cherokee Indian Normal School”, it is clear that Section 11 is in fact establishing an entity to enforce racial segregation and is, by implication, a Jim Crow law.

Finally, Extrinsic Jim Crow Laws are such laws that neither contain language clearly calling for or enforcing segregation or discrimination based on race nor are broadly understood to have an intent to discriminate against and/or segregate on the basis of race but which our experts have marked as Jim Crow based on knowledge of specific extrinsic factors surrounding the passage of a law. A good example is 1925 Session Law Chapter 295 Section 10,⁴⁰ which provides funding for the construction of a new courthouse. The text of this section of the law contains no explicit language calling for racial segregation, nor does it evince a clear intent to segregate, but our disciplinary scholar was aware that the funds allocated for the construction of a new courthouse were used in part to fund the construction of a confederate monument. With this extrinsic knowledge as context, 1925 Session Law Ch. 295 section 10 can then be classified as a Jim Crow law.

Data analysis can be misleading or useless without contextualization through disciplinary knowledge. The scholarly interpretation of Jim Crow underpins the consistency of our training set and analysis; classification of the types of Jim Crow laws may be used to expand this analysis in the future.

PROJECT OUTREACH

Ensuring that project outcomes included clear connections and possibilities for pedagogical use has been a central goal since the beginning, particularly since the idea for OTB came from a practical need for teaching the history of Jim Crow to high school students. The OTB website includes a *Teach* section where educators can search laws and find ways to incorporate the subject into their curriculum. The website includes lesson plans, a diverse list of primary and secondary sources, an interactive historical timeline of Jim Crow in North Carolina, and a glossary of common and necessary terms.

To ensure that an inventory of laws could be made available for those teaching Jim Crow, particularly K-12 instructors, and that the website could become a portal for materials that support teaching history in the K-12 classroom, OTB teamed up with Carolina K-12,⁴¹ part of the Carolina Public Humanities initiative at UNC-Chapel Hill. Carolina K-12 staff are curriculum experts with the background and resources to ensure that any content on the site is appropriate and adaptable to K-12 teaching, and that the website is easy to understand and use. This group of K-12 experts created a new lesson plan specifically for teachers interested in using the site and recommended lesson plans already in their database to list on the OTB website.

Furthermore, Carolina K-12 regularly offers workshops for public school teachers in North Carolina and, through the OTB partnership, planned a workshop for teachers on the subject of Jim Crow, with the website being a centerpiece. This first workshop was held January 25, 2021 via Zoom, with over 200 attendees. Other workshops and outreach will be conducted by OTB team members alongside Carolina K-12, including work with school media specialists in the state.

Following the project’s ethical commitments and guiding concept of algorithms of resistance, the team developed educational modules for learners wanting to engage with OTB methods. These modules are designed to empower learners to build on the work begun by OTB and to apply the project’s methods to their own work. They introduce concepts of algorithms, discuss algorithms of resistance, and demonstrate using Python to perform OCR and exploratory analysis of historical texts.⁴² Created in Jupyter Notebooks, published on GitHub, and hosted on Binder (mybinder.org), these modules can be used in individual learning as well as in classroom instruction. When completed, the modules will be linked to from the OTB website. Planning is also under way for their possible uses in an undergraduate course at UNC-Chapel Hill, a graduate digital humanities course at Duke University, and in a workshop series hosted by the Digital Humanities Collaborative of North Carolina in 2021-22. Jupyter Notebooks afford educators and learners the ability to intersperse code with text and image, allowing learners to read about and test technical concepts in the same space.

ORGANIZATIONAL CONSIDERATIONS FOR CREATING COLLECTIONS AS DATA

OTB required a range of skills and expertise, including subject expertise, technical skills, and data skills. This breadth of expertise and diversity was required to be successful. Each person brought to the team their specialized knowledge that everyone relied on. Ensuring successful collaboration among such a large group of individuals with different specializations required intentionality. Although the team did not begin the project ascribing to the concept of radical collaboration, this was how it operated. Nancy McGovern describes radical collaboration as, “coming together across disparate, but engaged, domains in ways that are often unfamiliar or possibly uncomfortable to member organizations and individuals in order to identify and solve problems together, to achieve more together than we could separately.”⁴³

At the beginning of the project, María Estorino, senior administrator on the team, led the group in some practices to facilitate work across reporting lines. These practices included forming the Special Collections Digital Scholarship working group with membership across different library units and getting to know one another. The instinct to get started with the work right away was strong, but “A new community thrives by devoting time to getting acquainted.”⁴⁴ The team took the time for each member to share with the group the skills they brought to the project and what they hoped to gain. This helped everyone understand their roles and how to best organize the work. Working and meeting schedules were established to ensure that the project stayed on track. Ultimately, these practices began the process of building trust. “Being able to rely upon others results from accrued trust based on the perceived reliability of partners.”⁴⁵ Before team members knew one another well enough to trust, good intent was assumed; everyone operated under good faith that each team member was doing their best. As the project progressed and additional outside expertise was needed, others were eager to join. “When you engage in radical collaboration, you participate in an interaction of two or more people allowing the group to achieve and sustain outcomes that members could not individually, the resulting community flourishes—successes are visible and measurable, and people want to join.”⁴⁶

EXTENDING THIS WORK

The second phase of OTB, funded by the Association for Research Libraries’ Venture Fund, ends in May 2021. This phase involved improving the corpus through additional data cleaning efforts, additional analysis work, the creation of educational modules, improving the project website, and improving and finalizing the training set for release.

The release of the training set is the final piece needed to allow for the expansion of this work—the identification of Jim Crow laws in states beyond North Carolina. The scripts and workflows on the project’s GitHub site are available for the preparation of law corpora from other states, and the training set can be used to train models to identify Jim Crow laws. States’ legislatures can be influenced by legislation from neighboring states. A future study might investigate how Jim Crow laws were propagated throughout the states, and how they differed between regions.

BIBLIOGRAPHY

- Algorithmic Justice League. “Learn More—The Algorithmic Justice League.” Accessed April 12, 2021. <https://www.ajl.org/learn-more>.
- Always Already Computational—Collections as Data. “The Santa Barbara Statement on Collections as Data.” Accessed February 2, 2021. <https://collectionsasdata.github.io/statement/>.
- Bruckner, Lorin, Sarah Carrier, Amanda Henley, Matt Jansen, and Sturkey, William. “On the Books: Jim Crow and Algorithms of Resistance.” Presented at the Association of Southeastern Research Libraries Webinar Series, June 30, 2020. <https://vimeo.com/434504369>.
- Benjamin, Ruha. *Race after Technology: Abolitionist Tools for the New Jim Code*. Medford, MA: Polity, 2019.
- Chen, Tianqi, and Carlos Guestrin. “XGBoost.” In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. <https://arxiv.org/abs/1603.02754>.
- Collections as Data—Part to Whole. “Part to Whole.” Accessed March 5, 2021. <https://collectionsasdata.github.io/part2whole/>.
- Costanza-Chock, Sasha. *Design Justice*. Cambridge, MA: MIT Press, 2020. <https://design-justice.pubpub.org/>

- Henley, Amanda, Matt Jansen, Lorin Bruckner, Neil Byers, and Rucha Dalwadi. "On the Books: Jim Crow and Algorithms of Resistance White Paper." White Paper, 2020. <https://doi.org/10.17615/hvz4-sr14>.
- Internet Archive. "North Carolina Government Documents—Ensuring Democracy through Digital Access." Accessed March 5, 2021. <https://archive.org/details/ncgovdocs?tab=about>.
- McGovern, Nancy. "Radical Collaboration and Research Data Management: An Introduction." *Research Library Issues*, no. 296 (2018): 6–22. <https://doi.org/10.29242/rli.296.2>.
- Murray, Pauli. *States' Laws on Race and Color; and Appendices Containing International Documents, Federal Laws and Regulations, Local Ordinances and Charters*. Cincinnati: Methodist Church, 1950.
- Noble, Safiya Umoja. *Algorithms of oppression : How Search Engines Reinforce Racism*. New York: NYU Press, 2018.
- North Carolina. Private Laws of the State of North Carolina Passed by the General Assembly at Its Session of 1903. Raleigh : E.M. Uzzell. <http://archive.org/details/privatelawsofsta1903nort>.
- North Carolina. Public Laws and Resolutions Passed by the General Assembly at Its Session of 1929. Raleigh : Mitchell Print. Co. <http://archive.org/details/publiclawsresolu1929nort>.
- North Carolina. Public-Local Laws Passed by the General Assembly ; Private Laws Passed by the General Assembly 1925. Raleigh : Edwards & Broughton Print. Co., state printers. <http://archive.org/details/publiclocallawsp1925nort>.
- Padilla, Thomas. *On a Collections as Data Imperative*. Conference Report. Collections as Data: Stewardship and Use Models to Enhance Access. Washington, D.C.: The Library of Congress, 2016. http://digitalpreservation.gov/meetings/dcs16/tpadilla_OnaCollectionsasDataImperative_final.pdf.
- Padilla, Thomas, Laurie Allen, Hannah Frost, Sarah Potvin, Elizabeth Russey Roke, and Stewart Varner. *Final Report—Always Already Computational: Collections as Data*, 2019. <https://doi.org/10.5281/zenodo.3152935>.
- Paschal, Richard. "Jim Crow in North Carolina A Carolina K-12 Virtual Curriculum Module." Presented at the Carolina Public Humanities, February 10, 2021. <https://www.youtube.com/watch?v=WF1Xtx63Nas&feature=youtu.be>.
- Paschal, Richard. *Jim Crow in North Carolina: The Legislative Program from 1865 to 1920*. Carolina Academic Press, 2020.
- Shanks, Justin, Sara Mannheimer, and Jason Clark. "Radical Collaboration: Making the Computational Turn in Special Collections and Archives." *ResearchDataQ*, November 20, 2019. <https://researchdataq.org/editorials/radical-collaboration-making-the-computational-turn-in-special-collections-and-archives/>.
- Tufekci, Zeynep. "Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency symposium essays." *Colorado Technology Law* 13, 2015: 203– 18.
- UNC-Libraries-Data/OnTheBooks. GitHub Repository, accessed March 5, 2021. <https://github.com/UNC-Libraries-data/OnTheBooks>
- UNC University Libraries. "Algorithms of Resistance Reveal Extent of North Carolina's Jim Crow Laws." Accessed April 6, 2021. <https://library.unc.edu/2021/02/algorithms-of-resistance/>.
- University of Illinois. "HTRC Digging Deeper, Reaching Further—Libraries Empowering Users to Mine the HathiTrust Digital Library Resources." Accessed March 5, 2021. <https://teach.htrc.illinois.edu/>.
- University of North Carolina at Chapel Hill. "Carolina K-12." Carolina Public Humanities. Accessed March 22, 2021. <https://humanities.unc.edu/ck12/>.

NOTES

1. "The Santa Barbara Statement on Collections as Data," Always Already Computational—Collections as Data, accessed February 2, 2021, <https://collectionsasdata.github.io/statement/>.
2. "Part to Whole," Collections as Data—Part to Whole, accessed March 5, 2021, <https://collectionsasdata.github.io/part2whole/>.
3. "On The Books—Jim Crow and Algorithms of Resistance," accessed March 5, 2021, <https://onthebooks.lib.unc.edu/>.
4. "HTRC Digging Deeper, Reaching Further—Libraries Empowering Users to Mine the HathiTrust Digital Library Resources," University of Illinois, accessed March 5, 2021, <https://teach.htrc.illinois.edu/>.
5. Justin Shanks et al., "Radical Collaboration: Making the Computational Turn in Special Collections and Archives," *ResearchDataQ*, November 20, 2019, <https://researchdataq.org/editorials/radical-collaboration-making-the-computational-turn-in-special-collections-and-archives/>.
6. See note 5 above.
7. "North Carolina Government Documents—Ensuring Democracy through Digital Access." Internet Archive, 2010. <https://archive.org/details/ncgovdocs?tab=about>.
8. Thomas Padilla et al., "Final Report—Always Already Computational: Collections as Data," May 2019: 15, <https://doi.org/10.5281/zenodo.3152935>.
9. "The Santa Barbara Statement on Collections as Data."
10. Thomas Padilla, "On a Collections as Data Imperative," Conference Report, Collections as Data: Stewardship and Use Models to Enhance Access (Washington, D.C.: The Library of Congress, September 27, 2016). http://digitalpreservation.gov/meetings/dcs16/tpadilla_OnaCollectionsasDataImperative_final.pdf].
11. "UNC-Libraries-Data/OnTheBooks," accessed March 5, 2021, <https://github.com/UNC-Libraries-data/OnTheBooks>.
12. Amanda Henley et al., "On the Books: Jim Crow and Algorithms of Resistance White Paper," White Paper, 2020, <https://doi.org/10.17615/hvz4-sr14>.

13. Padilla et al., 20.
14. Richard Paschal, *Jim Crow in North Carolina: The Legislative Program from 1865 to 1920* (Carolina Academic Press, 2020), 194.
15. Richard Paschal, “Jim Crow in North Carolina A Carolina K-12 Virtual Curriculum Module,” <https://www.youtube.com/watch?v=WF1Xtx63Nas&feature=youtu.be>
16. Benjamin, Ruha. “Introduction.” in *Race After Technology: Abolitionist Tools for the New Jim Code*. Newark: Polity Press, 2019. Accessed March 5, 2021. ProQuest Ebook Central. 6.
17. Noble, Safiya Umoja. “Introduction.” In *Algorithms of Oppression : How Search Engines Reinforce Racism*. (New York: NYU Press, 2018), 10.
18. Zeynep Tufekci, “Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency symposium essays,” *Colorado Technology Law* 13 (2015): 207.
19. “Overview.” Algorithmic Justice League, accessed April 12, 2021. <https://www.ajl.org/learn-more>.
20. Benjamin, Ruha. “Engineered Inequity.” In *Race After Technology: Abolitionist Tools for the New Jim Code*. Newark: Polity Press, 2019. Accessed March 5, 2021. ProQuest Ebook Central. 4, 42.; Costanza-Chock, Sasha. “Design Values: Hard-Coding Liberation?” In *Design Justice: Community-Led Practices to Build the Worlds We Need*. Cambridge, MA: MIT Press, 2020. <https://design-justice.pubpub.org/pub/3h2zq86d/release/1>.
21. Costanza-Chock.
22. Costanza-Chock.; Algorithmic Justice League, “Overview”
23. Costanza-Chock.
24. Algorithmic Justice League, “Overview.”
25. Algorithmic Justice League, “Overview.”; Costanza-Chock.
26. Ruha.; Costanza-Chock.
27. Amanda Henley et al., “On the Books: Jim Crow and Algorithms of Resistance White Paper,” White Paper, 2020, <https://doi.org/10.17615/hvz4-sr14>.”
28. Internet Archive, “North Carolina Government Documents—Ensuring Democracy through Digital Access”.
29. Pauli Murray, *States’ Laws on Race*.
30. Richard Paschal, *Jim Crow in North Carolina: The Legislative Program from 1865 to 1920*.
31. Tianqi Chen et al., “XGBoost.” In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. <https://arxiv.org/abs/1603.02754>.
32. Pauli Murray, *States’ Laws on Race*, 676.
33. Pauli Murray, *States’ Laws on Race*, 677.
34. Pauli Murray, *States’ Laws on Race*, 679.
35. Pauli Murray, *States’ Laws on Race*, 685.
36. “Algorithms of Resistance Reveal Extent of North Carolina’s Jim Crow Laws—UNC Chapel Hill Libraries,” UNC University Libraries, accessed April 6, 2021, <https://library.unc.edu/2021/02/algorithms-of-resistance/>.
37. William Sturkey, quoted from Lorin Bruckner et al, “On the Books: Jim Crow and Algorithms of Resistance,” <https://vimeo.com/434504369>.
38. North Carolina, *Private Laws of the State of North Carolina Passed by the General Assembly at Its Session of 1903* (Raleigh), 863, Ch. 364 § 1.
39. North Carolina, *Public Laws and Resolutions Passed by the General Assembly at Its Session of 1929* (Raleigh), 252, Ch. 195 § 11.
40. North Carolina, *Public-Local Laws Passed by the General Assembly ; Private Laws Passed by the General Assembly 1925* (Raleigh), 332, Ch. 295 § 10.
41. “Carolina K-12” Carolina Public Humanities, University of North Carolina at Chapel Hill, accessed March 22, 2021, <https://humanities.unc.edu/ck12/>.
42. “UNC-Libraries-Data/OnTheBooks,” [Github.com](https://github.com/UNC-Libraries-data/OnTheBooks), accessed April 5, 2021, <https://github.com/UNC-Libraries-data/OnTheBooks>.
43. Nancy Y. McGovern. “Radical Collaboration and Research Data Management: An Introduction.” *Research Library Issues*, no. 296 (2018): 6. <https://doi.org/10.29242/rli.296.2>.
44. McGovern, “Radical,” 9.
45. McGovern, “Radical,” 8.
46. McGovern, “Radical,” 11.