

On the Books: Jim Crow and Algorithms of Resistance

Collections as Data Summative Forum
Las Vegas, Nevada, January 17, 2020

Maria Estorino, Amanda Henley, William Sturkey



Dr. Reverend Pauli Murray (1910-1985)

- Lawyer
- Priest
- Civil rights activist
- Author, *States Laws on Race and Color* (1951)



Figure 1. Murray, Pauli, 1910-1985. Adapted from The Carolina Story: A Virtual Museum of University History digital collection, retrieved from The University of North Carolina at Chapel Hill University Libraries website.



Figure 2. Outside Looking In, Mobile, Alabama. Adapted from the Gordon Parks Foundation website, retrieved from www.gordonparksfoundation.org. Copyright 1956.



Figure 1. Murray, Pauli, 1910-1985. Adapted from The Carolina Story: A Virtual Museum of University History digital collection, retrieved from The University of North Carolina at Chapel Hill University Libraries website.

On the Books

- Motivated by a reference question:

Where do I find a list of NC Jim Crow laws?

About On the Books



Project to make North Carolina legal history accessible as a text corpus.



100+ years of North Carolina public, private, and local session laws



Project Goals:

- Create corpus of NC Session Laws from 1865/66-Identify discoverable NC segregation statutes during the Jim Crow era using text analysis
- Create project website listing statutes discovered
- Produce white paper
- Git repository

Project Team

Lorin Bruckner, Data Visualization Services Librarian - Text Analysis and Visualization Expert

Sarah Carrier, North Carolina Research and Instructional Librarian - Special Collections Expert

Rucha Dalwadi, Research Assistant - Documentation and Content Developer

María R. Estorino, AUL for Special Collections & Director of the Wilson Library - Executive Sponsor and Liaison to the Library Leadership Team

Amanda Henley, Head of Digital Research Services - Principal Investigator and Project Lead

Matt Jansen, Data Analyst - Text Analysis Expert and Statistician

William Sturkey, Faculty Member of History - Disciplinary Scholar

Nathan Kelber, Ithaka – Collaborator (former PI and Project Lead)



Additional Project Consultants and Collaborators

Daniel Anderson, Professor for Pilot Project: [NC Jim Crow Laws, 1899-1919](#)

Ryan Cordell, OCR Specialist and author of [A Research Agenda for Historical and Multilingual Optical Character Recognition](#)

Kristen Foote, Research Assistant and Lead for Pilot Project: [NC Jim Crow Laws, 1899-1919](#)

Anna Goslen, Metadata Librarian

Aaron S. Kirschenfeld, NC Legal Information Expert

Richard Paschal, NC Legal Scholar

Steve Segedy, Software Developer

Ryan Shaw, Information Science Expert

Kimber Thomas, African American Studies Scholar

Completed work

- Process
- Progress on Deliverables
- Challenges

1865/66 <https://archive.org/details/privatelawsofstal186566nor>
1865/66 <https://archive.org/details/publiclawsofstat186566nor>
1866/<https://archive.org/details/privatelawsofstal186666nor>
1866/<https://archive.org/details/publiclawsofstat186666nor>
1868
1868/<https://archive.org/details/privatelawsofstal186868nor>
1868/69 <https://archive.org/details/publiclawsofstat186869nor>
1869/70 <https://archive.org/details/privatelawsofstal186970nor>
1869/<https://archive.org/details/publiclawsofstat186969nor>
1870/<https://archive.org/details/privatelawsofstal187070nor>
1870/<https://archive.org/details/publiclawsofstat187070nor>
1871/<https://archive.org/details/privatelawsofstal187171nor>
1871/<https://archive.org/details/publiclawsofstat187171nor>
1872/<https://archive.org/details/privatelawsofstal187272nor>
1873/<https://archive.org/details/publiclawsofstat187373nor>
1874/75 <https://archive.org/details/lawsresolutionso187475nor>
1876/77 <https://archive.org/details/lawsresolutionso187677nor>
1879 <https://archive.org/details/lawsresolutionso1879nort>
1880 <https://archive.org/details/lawsresolutionso1880nort>

Compiled comprehensive listing of volumes for our corpus

96 volumes in our corpus. Volumes span from 1865/66 – 1967.

**Gathered OPAC record unique IDs > linked to Internet Archive Objects >
generated download links with Internet Archive API.**

Image Challenges

Marginalia

Differences between volumes

Skew

Superintendent and assistants.	boys, and a building for girls with separate accommodations for white and colored girls. The board of commissioners may employ a superintendent and such other assistants as may be deemed necessary to manage and control the institution under such rules as shall be established by such board: <i>Provided</i> , that the superintendent employed by the said board shall have the right and is hereby authorized to require obedience from all of the inmates of said reformatory, and is hereby intrusted with the authority to correct and punish any inmate thereof to the same extent as a parent may, under the law, impose upon his own child.
Proviso: Superintendent to require obedience.	
Authority to correct and punish.	
Work required.	SEC. 3. That any or all of the inmates may be required to work, from time to time, at light work on the farm, and in such shop or shops as may be established on the property, the object being to provide regular, light, suitable and humane employment, to teach the inmates good moral character, how to work, and to be industrious.
Object.	
Bond issue authorized.	SEC. 4. That in order to provide funds for the purchase of property and the erection of buildings and equipping the same, as provided in section two of this act, the board of commissioners are hereby authorized and empowered to issue coupon bonds bearing interest payable annually or semiannually, at a rate not exceeding six per cent per annum, to such an amount as the board of commissioners may determine, not exceeding the sum of thirty thousand dollars, in such denominations as the board of commissioners may determine, payable not more than twenty years from the date of issue. Said bonds shall be signed by the chairman of the board of county commissioners and countersigned by the county treasurer, and the coupons of said bonds shall bear the engraved or lithographed signature of the treasurer and shall be made payable, both principal and interest, at such place or places as the board may determine. Said bonds shall be sold at public sale, after due notice, and the proceeds thereof shall be applied for the purposes herein set out, and no other.
Interest.	
Amount.	
Denominations.	
Maturity.	
Authentication.	
Sale of bonds.	
Specific appropriation.	
Special tax.	SEC. 5. That the board of county commissioners, in order to provide for the payment of said bonds and coupons, are hereby authorized and empowered to levy a special tax of not more than five cents on the one hundred dollars worth of property and fifteen cents on the poll, observing the constitutional equation in the levy thereof. And the said board of commissioners are hereby authorized and empowered to purchase and retire one or more of said bonds annually, as the proceeds of the special tax may be found sufficient for that purpose, and without regard to the date of maturity of the particular bond or bonds; such bond or bonds, however, to be purchased and retired in the order of the date of their maturity.
Limit of rate.	
Constitutional equation.	
Annual purchase and retirement of bonds.	

S. B. 128

CHAPTER 15

AN ACT TO AUTHORIZE, EMPOWER, AND DIRECT THE SHERIFF AND/OR TAX COLLECTOR OF CUMBERLAND COUNTY TO ACCEPT COUNTY VOUCHERS IN PAYMENT OF TAXES.

The General Assembly of North Carolina do enact:

SECTION 1. That the sheriff and/or tax collector of current or back taxes of Cumberland County be and he is hereby authorized, empowered, and directed to accept in payment of all taxes due said county vouchers issued by the county of Cumberland under the authority of the board of commissioners of said county, when said vouchers are so issued in satisfaction and settlement of the obligations of said county for services rendered to or purchases made by it, and the said vouchers so accepted by said sheriff and/or tax collector shall be included in and accepted by the said board of commissioners of Cumberland County in any settlement of taxes submitted and made by said sheriff and/or tax collector.

County vouchers acceptable in paying Cumberland County taxes.

SEC. 2. All laws and clauses of laws in conflict with this act are hereby repealed.

Disposition of vouchers.

SEC. 3. This act shall be in force and effect from and after its ratification.

Conflicting laws repealed.

Ratified this the 30th day of January, A.D. 1933.

S. B. 146

CHAPTER 16

AN ACT TO AMEND CHAPTER TWENTY-FIVE OF PUBLIC-LOCAL LAWS OF ONE THOUSAND NINE HUNDRED AND TWENTY-NINE RELATING TO THE OFFICE OF AUDITOR OF RUTHERFORD COUNTY.

The General Assembly of North Carolina do enact:

SECTION 1. That section one, chapter twenty-five, Public-Local Laws one thousand nine hundred and twenty-nine, be amended by striking out, in lines three and four, the words, "together with the Board of Education" and by striking out the word "boards" in line nine and inserting in lieu thereof the word "board."

Ch. 25, Public-Local Laws 1929, amended, as to Rutherford County Auditor.

SEC. 2. That section three of said chapter be stricken out and the following inserted in lieu thereof:

"SEC. 3. That the Auditor thus appointed shall be paid an annual salary of two thousand dollars (\$2,000.00) payable monthly out of the general county fund of said county, the same

Salary of Auditor, \$2,000.00.

Marginalia

Special tax.

Limit of rate.

Constitutional
equation.

Annual purchase
and retirement of
bonds.

SEC. 5. That the board of county commissioners, in order to provide for the payment of said bonds and coupons, are hereby authorized and empowered to levy a special tax of not more than five cents on the one hundred dollars worth of property and fifteen cents on the poll, observing the constitutional equation in the levy thereof. And the said board of commissioners are hereby authorized and empowered to purchase and retire one or more of said bonds annually, as the proceeds of the special tax may be found sufficient for that purpose, and without regard to the date of maturity of the particular bond or bonds; such bond or bonds, however, to be purchased and retired in the order of the date of their maturity.

Image
preprocessing

Marginalia removed with
custom Python Script

Images rotated

blank, color-balanced
margins added back to
the images to improve
OCR

boys, and a building for girls with separate accommodations for white and colored girls. The board of commissioners may employ a superintendent and such other assistants as may be deemed necessary to manage and control the institution under such rules as shall be established by such board: *Provided*, that the superintendent employed by the said board shall have the right and is hereby authorized to require obedience from all of the inmates of said reformatory, and is hereby intrusted with the authority to correct and punish any inmate thereof to the same extent as a parent may, under the law, impose upon his own child.

SEC. 3. That any or all of the inmates may be required to work, from time to time, at light work on the farm, and in such shop or shops as may be established on the property, the object being to provide regular, light, suitable and humane employment, to teach the inmates good moral character, how to work, and to be industrious.

SEC. 4. That in order to provide funds for the purchase of property and the erection of buildings and equipping the same, as provided in section two of this act, the board of commissioners are hereby authorized and empowered to issue coupon bonds bearing interest payable annually or semiannually, at a rate not exceeding six per cent per annum, to such an amount as the board of commissioners may determine, not exceeding the sum of thirty thousand dollars, in such denominations as the board of commissioners may determine, payable not more than twenty years from the date of issue. Said bonds shall be signed by the chairman of the board of county commissioners and countersigned by the county treasurer, and the coupons of said bonds shall bear the engraved or lithographed signature of the treasurer and shall be made payable, both principal and interest, at such place or places as the board may determine. Said bonds shall be sold at public sale, after due notice, and the proceeds thereof shall be applied for the purposes herein set out, and no other.

SEC. 5. That the board of county commissioners, in order to provide for the payment of said bonds and coupons, are hereby authorized and empowered to levy a special tax of not more than five cents on the one hundred dollars worth of property and fifteen cents on the poll, observing the constitutional equation in the levy thereof. And the said board of commissioners are hereby authorized and empowered to purchase and retire one or more of said bonds annually, as the proceeds of the special tax may be found sufficient for that purpose, and without regard to the date of maturity of the particular bond or bonds; such bond or bonds, however, to be purchased and retired in the order of the date of their maturity.

S. B. 128

CHAPTER 15

AN ACT TO AUTHORIZE, EMPOWER, AND DIRECT THE SHERIFF AND/OR TAX COLLECTOR OF CUMBERLAND COUNTY TO ACCEPT COUNTY VOUCHERS IN PAYMENT OF TAXES.

The General Assembly of North Carolina do enact:

SECTION 1. That the sheriff and/or tax collector of current or back taxes of Cumberland County be and he is hereby authorized, empowered, and directed to accept in payment of all taxes due said county vouchers issued by the county of Cumberland under the authority of the board of commissioners of said county, when said vouchers are so issued in satisfaction and settlement of the obligations of said county for services rendered to or purchases made by it, and the said vouchers so accepted by said sheriff and/or tax collector shall be included in and accepted by the said board of commissioners of Cumberland County in any settlement of taxes submitted and made by said sheriff and/or tax collector.

SEC. 2. All laws and clauses of laws in conflict with this act are hereby repealed.

SEC. 3. This act shall be in force and effect from and after its ratification.

Ratified this the 30th day of January, A.D. 1933.

S. B. 146

CHAPTER 16

AN ACT TO AMEND CHAPTER TWENTY-FIVE OF PUBLIC-LOCAL LAWS OF ONE THOUSAND NINE HUNDRED AND TWENTY-NINE RELATING TO THE OFFICE OF AUDITOR OF RUTHERFORD COUNTY.

The General Assembly of North Carolina do enact:

SECTION 1. That section one, chapter twenty-five, Public-Local Laws one thousand nine hundred and twenty-nine, be amended by striking out, in lines three and four, the words, "together with the Board of Education" and by striking out the word "boards" in line nine and inserting in lieu thereof the word "board."

SEC. 2. That section three of said chapter be stricken out and the following inserted in lieu thereof:

"SEC. 3. That the Auditor thus appointed shall be paid an annual salary of two thousand dollars (\$2,000.00) payable monthly out of the general county fund of said county, the same

OCR

OCR'd over 80,000 pages for the corpus

OCR Quality Assessment

	A	B	D	E	F	
			is the word deliniation correct (red box)	is the info in the red box OCR'd correctly?	Unsure?	
1	text	wordsampleImageBuffer				
2	or	or i	y	y		
3	in	in :	y	y		
4	succession,	succession, v	y	y		
5	it	it i	y	y		
6	town,	town, :	y	y		
7	machinery,	machinery, n	y	y		
8	over	over t	y	y		
9	provided	provid	n	n	y	
10	the	the	y	y		
11	said	said o	y	y		
12	fer-	fer-	y	y		
13	real	real e	y	y		
14	seventy-three,	seventy-three, b	y	y		
15	shall	shall c	y	y		
16	enact,	enact, '	y	y		
17	McKey's	McKey's I	y	n	y	
18	immediately	immediately a	y	y		
19	discretion	discretion is	y	y		
20	may	ority m	n	n	y	
21	and	and c	y	y		


```

<root>
  <volume title="" date_range="1865/1866">
    <type lawtype="(ie public, private, constitution)">
      <law chapter="" start_page="" start_image="">
        <section section_number="">
          "Text goes here"
        </section>
        <section section_number="">
          "Text goes here"
        </section>
      </law>
      <law chapter="" start_page="" start_image="">
        <section section_number="">
          "Text goes here"
        </section>
        <section section_number="">
          "Text goes here"
        </section>
      </law>
    </type>
  </volume>
</root>

```

Devised Metadata Schema

- XML
- Consulted metadata librarian to create our own schema

Corpus creation + Text Analysis

Can we determine which laws are Jim Crow?

Parse/Annotate Laws

Challenges

- Errors in original volumes.
- Chapter breaks in the margins for some volumes
(These were split by hand)
- A few volumes have Roman Numerals (did not OCR well)
- Numbers OCR poorly

Using regular expressions to fix common errors.

Likely to be the biggest limitation in our corpus.

Currently we have **over 286,000 sections** split.

Text Analysis

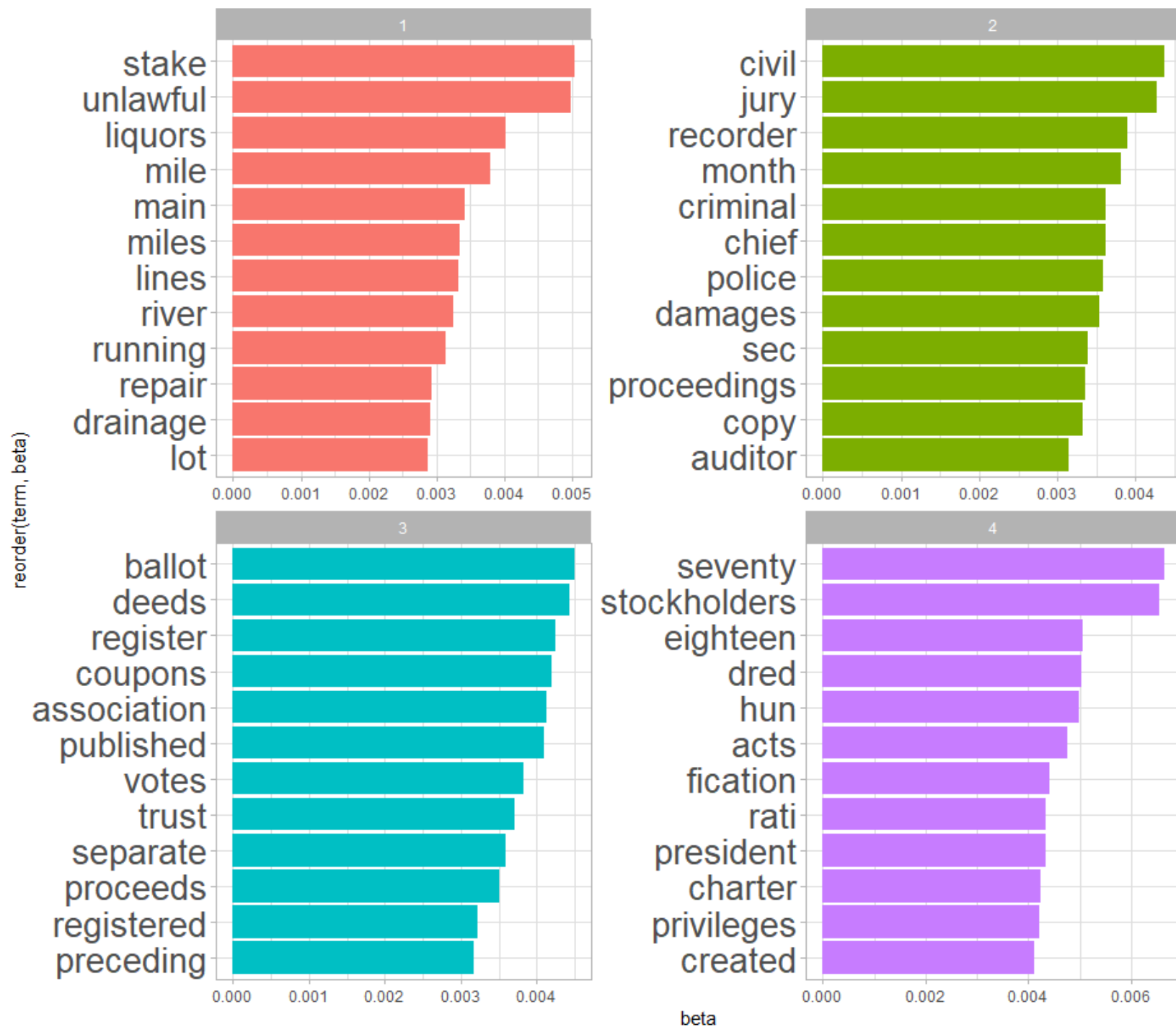
Machine Learning Techniques

Supervised

- Training the model with laws labeled by experts:
 - Pauli Murray
 - Richard Paschal
 - William Sturkey
 - Kimber Thomas
- Our training set is 800 laws (and growing)
 - 16% of laws in the training set are Jim Crow
- Preliminary supervised classification uses a Naïve Bayes model

Topic Models

(Using Partial Data, 5% Sample)



Challenges

- Work at scale it is not going to be perfect.
- New territory for librarians and it makes us uneasy.
- Balance between completing the product and perfecting it.
- Realizing that our corpus will not be perfect, we will strive for transparency, documentation, and reproducibility.

Thank you

References

Parks, G. [ca. 1956]. Outside Looking In, Mobile, Alabama [Digital image]. Retrieved from:
https://s3.amazonaws.com/files.collageplatform.com/prod/image_cache/enlarge/548b3beaaa921a8b5be4cbcc/21bbea5eeebb2b2c0a59e5fdbfc27eb6.jpeg.
Copyright 1956 by Gordon Parks.

North Carolina Collection Photographic Archives, University of North Carolina at Chapel Hill University Libraries [Digital image]. Retrieved from:
https://dc.lib.unc.edu/cdm/ref/collection/vir_museum/id/431

Communication, Education & Outreach

- Presented to several audiences
 - Librarians, Digital Humanists, K-12 Teachers
- White paper
- Website
 - Wordpress platform.
 - **Planned expiration date for the website: 3 years.**
 - Corpora to be retained - Carolina Digital Repository.
- GitHub code repository
 - Jupiter Notebooks for explaining code

Planned activities post grant funding

Carolina K-12 to create a curriculum using our work aimed at K-12.

Continue to improve the corpus

Seek additional funding to continue this work.