On the Books: Jim Crow and Algorithms of Resistance

CNI Spring Meeting
May 4, 2020

Lorin Bruckner, Amanda Henley, Kimber Thomas
Jim Crow

• Codified system of racial apartheid

• This project uncovers some of the many laws that were put in place to maintain racial segregation in the U.S. South between 1865-1968.

• The laws uncovered here demonstrate how, for African Americans, especially, Jim Crow affected nearly every aspect of their daily lives, including attacks on their freedom and dignity.
Figure 2. Outside Looking In, Mobile, Alabama. Adapted from the Gordon Parks Foundation website, retrieved from www.gordonparksfoundation.org. Copyright 1956.
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
Identifying Laws

• Reviewed 900+ laws

• Laws were separated into two categories: yes or no

• Laws identified as "yes," or Jim Crow laws, presented evidence of legalizing or enforcing racial segregation

• Laws identified as "no," or non-Jim Crow laws, presented no evidence of legalizing or enforcing racial segregation
Sec. 2. That the board of commissioners of Cherokee county shall set apart from the taxes collected for the year one thousand eight hundred and eighty-one the sum of six hundred dollars for the payment of jurors for said county for the year one thousand eight hundred and eighty-two.
those doing a temporary business; to establish, regulate and control cemeteries; to provide for the manner in which bodies may be interred therein or removed therefrom, and for beautifying, ornamenting, and keeping the same in condition; to provide suitable grounds for the enlargement, extension, or establishment of new cemeteries, providing separate cemeteries for white and black;
Expertise Needed

• Scholarly expertise
• Deep knowledge of the collection
• Project management
• Technical/data skills
  • Programming skills
  • Version control
  • OCR
  • Text analysis
  • Metadata
  • Software development

• Cross-departmental collaboration fostered through a Special Collections Digital Scholarship Working Group
Project Team

Neil Byers, Graduate Assistant – Documentation and Content Developer

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

Kimber Thomas, African American Studies Scholar

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

Student Workers

Montana Eck, Julia Long, Ashley Mullikin, Siri Nallaparaju, Tim Oyeleke, and Jenna Patton

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
Workflow and Processes
Compiled comprehensive listing of volumes for our corpus


Gathered OPAC record unique IDs > linked to Internet Archive Objects > generated download links with Internet Archive API.
1917—Chapter 401

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 super-
tendent 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 invested 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. 5. 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 indus-

tious. 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 counter-signed by the county treasurer, and the coupons of said bonds shall bear the engraved or lithographed signature of the treas-

urer 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 fif-
teen 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 author-
ized, 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 25th 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 section be stricken out and the following inserted in lieu thereof:

"Sec. 2. 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
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 semi-annually, 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.
OCR

OCR’d over 80,000 pages for the corpus
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 280,000 sections** split.
Text Analysis
Supervised and Unsupervised Machine Learning
Supervised

• Training the model with laws labeled by experts:
  • Pauli Murray
  • Richard Paschal
  • William Sturkey
  • Kimber Thomas
• Our training currently set contains 900 laws
  • 100 "Yes", 800 "No"
  • Needs to be better developed
Models

- Experimenting with Naïve Bayes and Gradient Booster
- Naïve Bayes performs better with 90% accuracy
  - High accuracy comes from its ability to correctly predict laws that are NOT Jim Crow
  - Need to continue to develop training set so more Jim Crow laws can be identified
Unsupervised

• Topic Modeling
  • No annotation or training
  • Algorithm detects differences and similarities between the law texts
  • Laws placed into groups or "topics" based on those differences
Topic Models
(Using Partial Data, 5% Sample)
Deliverables and Next Steps
Deliverables

• Outreach and Education
  • Presented to several audiences: Librarians, Digital Humanists, K-12 Teachers
  • White paper

• Website
  • Corpus of laws available to download as text
  • Planned assessment date: 3 years
  • Corpus will be retained through Carolina Digital Repository.

• GitHub code repository
  • Jupyter Notebooks for explaining code
The example images below demonstrate the "banding" process. It is important to note that the images below show the bands superimposed on the original images rather than their diff versions. In reality, `get_bands` works with `diff` and not with `orig`.

```python
# hide input
bandimages = [bandimage(orig1, ang1, band_dict1, bheight, orig_bbox1),
               bandimage(orig2, ang2, band_dict2, bheight, orig_bbox2),
               bandimage(orig3, ang3, band_dict3, bheight, orig_bbox3)]
display(bandsdisplay(bandimages))
```
Next Steps

- Carolina K-12 to create a curriculum using our deliverables.
- Seeking additional funding to further our work:
  - Improve law splitting
  - Create search functionality for website
  - Improve machine learning models, expand analysis
    - Can we classify the Jim Crow laws?
    - Exploratory analysis – are there temporal or geographic trends?
- Create OER for use in academic classrooms
- Investigate potential of using our methods to ID Jim Crow laws from other states.
Why is this happening at the Library?
UNC Data Science Initiative

• Will include a new school for data science
• Libraries will be tasked with supporting data focused education and research
• Collections are becoming data and data are becoming collections
Thank you

References
