Abstract
“African American Arts Coverage in the Black Press” is a text and data analysis project focused on understanding how the Black Press (periodicals, newspapers, magazines, etc.) covered African American visual artistic production (painting, sculpture, photography, etc.) during the twentieth century.

Our principal research questions are aimed at understanding how the Black press was reporting on developments in the African American visual arts scene: who, what, when.

Which visual artists were being mentioned and in what context?
What types of visual art received coverage?
Are there chronological trends in arts coverage?

Does this information match up with the scholarship on twentieth century African American artistic production?

Data Used
A major hurdle this project faced was gaining access to the data since it all falls under copyright. The well timed receipt of a PROQUEST TDM Studio Research Grant enabled the team to utilize copyrighted materials in the ProQuest created environment.

ProQuest Historical Newspapers: Black Periodicals Database. Created chronological slices of data (1919-1940 and 1960-1970) to test methods on materials when African American visual production was high. Using ProQuest TDM Studio environment limited the data set size to under 2,000,000 documents.

- Afro-American
- Chicago Defender

Methods
In order to answer the research questions we had posed, our aim was to look through the text within the historical newspapers and find the articles where African American Artists were being mentioned.

There were two major limitations in this project:
1. The articles were read into the system through Optical Character Recognition (OCR) and therefore the quality of the text being parsed was questionable.
2. The workspace environment was heavily limited due to limited connection to the internet which meant all major text processing tools (Spacy, NLTK, etc.) had to be manually downloaded in order to be used.

In order to assess the quality of OCR, we searched through the articles and created a readability metric. This metric was essentially a measure of the proportion of recognizable English words that existed with each article.

Search Algorithm
We created a list of artist names, the variations they could have in their mention in it and whether that mention was truly something we can trust.

With the search algorithm we looked into (among other questions): 1. How many artist mentions were happening each year 2. Who was being mentioned the most. Who was being mentioned the least. 3. The disparity in mentions based on the gender of the artist. 4. What were the co-occurrences of mentions of artists (who was being mentioned together)

Initial results indicate a correlation between the spikes in visual arts mention and major periods in African American art history (e.g., Harlem Renaissance of the 1920s)

Results and Analysis
With the search algorithm we looked into (among other questions):
1. How many artist mentions were happening each year
2. Who was being mentioned the most. Who was being mentioned the least.
3. The disparity in mentions based on the gender of the artist.
4. What were the co-occurrences of mentions of artists (who was being mentioned together)

Below are some figures representing what we looked into.

Next Steps
Some next steps involve incorporation of both the search mechanisms and the readability score to develop a unified approach. This unified approach would also take into account the number of art related terms being mentioned in order to increase the quality of matches we are getting (i.e. is the article being captured really an art related article. In the end we should get a score associated with how likely it is that a given article has a given mention in it and whether that mention was truly something we can trust.

Acknowledgements
This work was funded by an Andrew W. Mellon Fellowship for the Digital Humanities and made possible in part due to a ProQuest TDM Studio Research Grant.

Professor VanDiver would like to thank Sarah Swanz and Ali Yaqoob for their dedication to shaping and helping launch this research project.