



# Modeling Overlapping Textual Data and Describing the Relationship Between the Overlapping Layers.



Leslie Alday, Project Supervisor: Sarah Stanley  
Office of Digital Research and Scholarship, FSU Libraries

## Introduction

- ❖ Digital humanities research.
- ❖ Computer-based representation of documents using a markup language allows better access to the information in a document (DeRose et al. 12).
  - ❖ Extensible Markup Language (XML) using Text Encoding Initiative (TEI) guidelines.
- ❖ How do you model and describe textual data that overlaps on a page?
  - ❖ “If an Identity” document logical structure using TEI guidelines (“TEI Consortium”).
  - ❖ “Algebraic Manipulations” document graphical structure using embedded transcription (“11 representations”).

## Methods

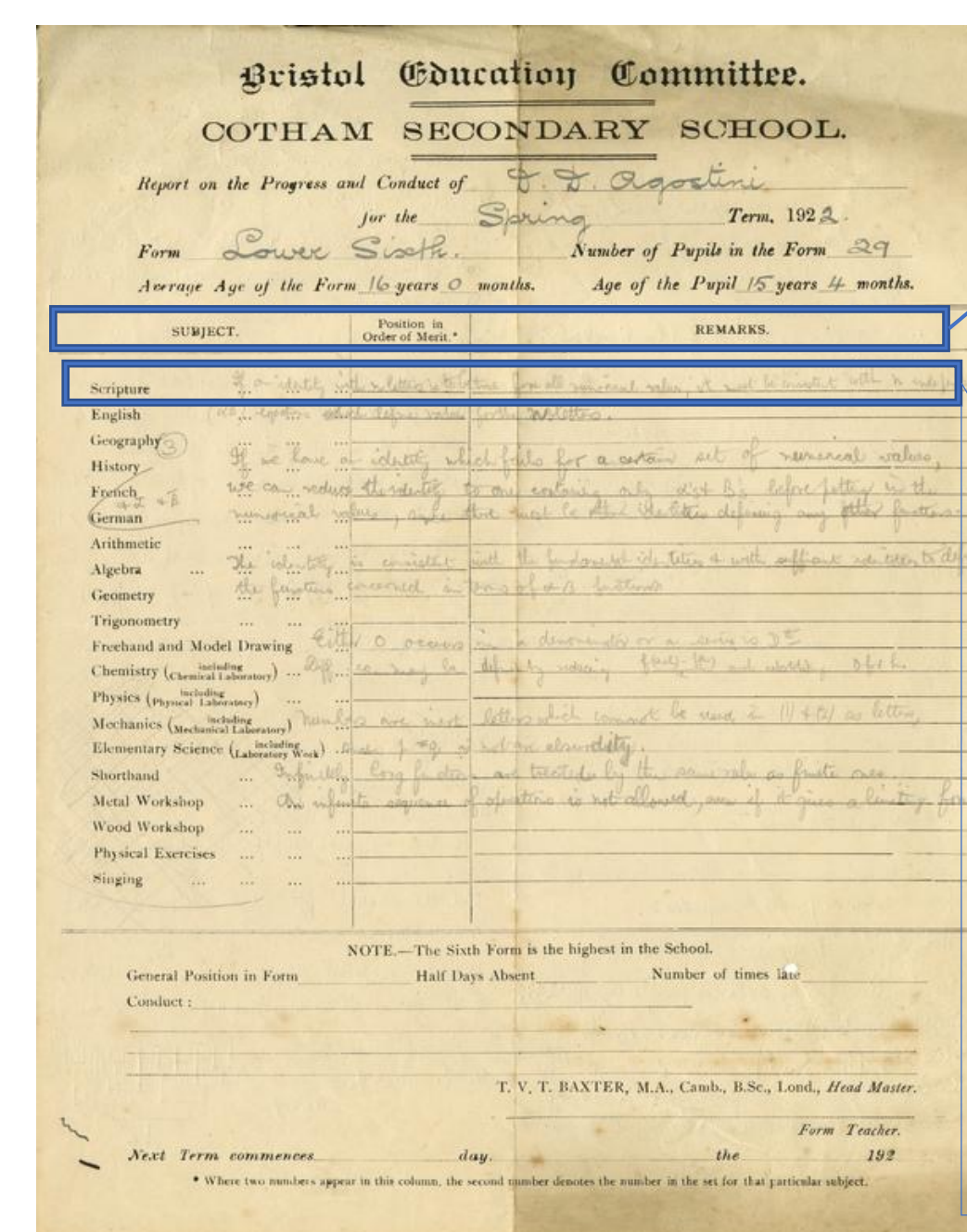
- ❖ The first layer of the “If an Identity” document was encoded considering the logical structure of the text and was treated as a table.
  - ❖ Using elements like `<table>`, `<row>`, and `<cell>` as described in TEI guidelines (“14 Tables”).
- ❖ The second layer of the “If An Identity” document was transcribed using the logical structure of the text and linking it to the rows they are written in on the table.
  - ❖ Using the attributes “corresp” and “xml:id” to point to the row where the line is found.
- ❖ The first and second layers of the “Algebraic Manipulations” were transcribed using the graphical organization of the different objects.
  - ❖ Using elements like `<sourceDoc>`, `<surface>`, and `<zone>` as described in TEI guidelines (“11 Representation”).
- ❖ Both documents were published after being transcribed.

## Results

- ❖ “If an Identity” published with encoding describing logical text structure.
  - ❖ Reading the document from the left to right.
  - ❖ Objects in the order that they are encountered.
- ❖ “Algebraic Manipulations” published with encoding based on the graphical structure.
  - ❖ Different objects are transcribed with tags that mark their location in an XY coordinate plane.

## Results

- ❖ If an Identity with n Letters Is to Be True...



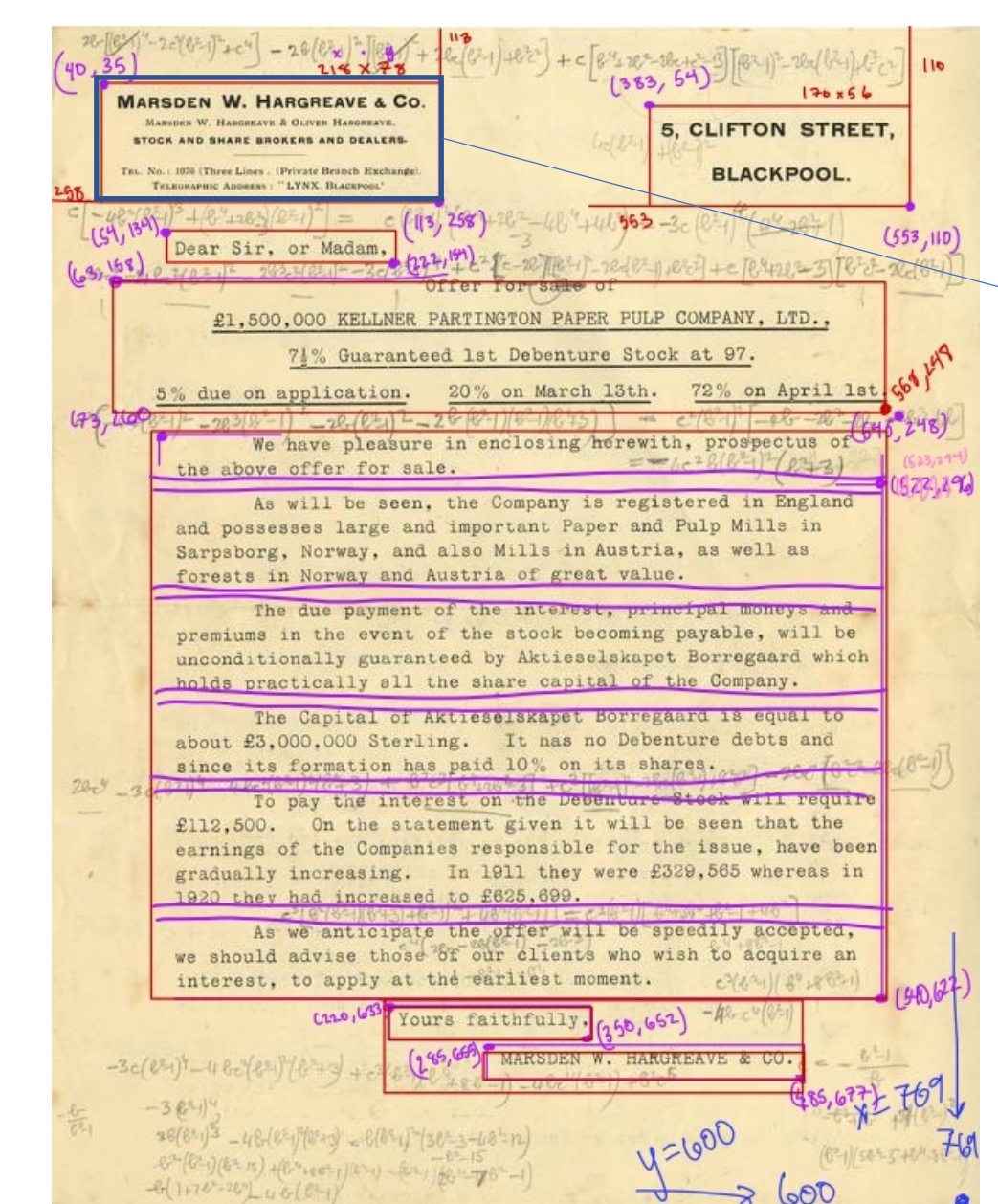
❖ Examples of how the first layer was transcribed using the logical structure.

```
<table rows="21" col="3">
  <row role="label">
    <cell role="label">SUBJECT</cell>
    <cell role="label">Position in Order of Merit.
    <hi rend="sup">*</hi><note place="bottom"><hi rend="sup">*</hi>Where two numbers
    appear in this column,
    the second number denotes the number in the set for that particular subject.</note></cell>
    <cell role="label">REMARKS.</cell>
  </row>
  <row role="data" xml:id="scripture01" corresp="#line01">
    <cell role="label">Scripture ... ..</cell>
    <cell role="data"> </cell>
    <cell role="data"> </cell>
  </row>
```

❖ Examples of how the second layer was transcribed using the logical structure.

```
<b corresp="#scripture01" xml:id="line01">If an identity with n letters is to <unclear>be
true</unclear> for all numerical values, n must be consistent with n independent
```

- ❖ Algebraic Manipulations, Doodled on an Old Stock Offer.



❖ Examples of how the first layer was transcribed using the graphical structure.

```
<surface ulx="0" uly="0" lrx="600" lry="769">
  <zone ulx="40" uly="35" lrx="113" lry="258">
    <line>MARSDEN W. HARDGREAVE & CO. </line>
    <line>MARSDEN W. HARDGREAVE & CO. OLIVER HARDGREAVE</line>
    <line>STOCK AND SHARE BROKERS AND DEALERS.</line>
    <line>Tel. No. : 1076 (Three Lines) . (Private Branch Exchange).</line>
    <line>Telegraphic Address: "LYNX. Blackpool"</line>
  </zone>
```

❖ Examples of how the second layer was transcribed using the logical structure.

```
<surface ulx="0" uly="0" lrx="600" lry="769">
  <zone ulx="18" uly="0" lrx="556" lry="43">
    <line>
      2b[<del rend="strikethrough">(b<hi rend="sup">2</hi>-1)<hi rend="sup">4</hi></del>
      -2c<hi rend="sup">2</hi>(b<hi rend="sup">2</hi>-1)<hi rend="sup">2</hi>+c<hi rend="sup">4</hi>]
      -2b(b<hi rend="sup">2</hi>-1)<hi rend="sup">2</hi>[<del rend="strikethrough">(b<hi rend="sup">2</hi>-1)<hi
      rend="sup">2</hi></del>
      +<hi rend="underline">2bc(b<hi rend="sup">2</hi>-1</hi>)+b<hi rend="sup">2</hi>c<hi rend="sup">2</hi>]
      +c<hi rend="underline">b<hi rend="sup">4</hi>+2b<hi rend="sup">2</hi></hi>-2bc+c<hi rend="sup">2</hi>-<hi
      rend="underline">3</hi>]
      [<hi rend="underline">(b<hi rend="sup">2</hi>-1)<hi rend="sup">2</hi></hi>-2bc(b<hi rend="sup">2</hi>-1)+b<hi
      rend="sup">2</hi>c<hi rend="sup">2</hi>]
    </line>
  </zone>
```

## Conclusion

- ❖ Using the logical structure of the text highlights the positions of the text on the table.
  - ❖ The second layer was written using the rows of the table in the first layer as lines.
  - ❖ Transcribing the document alone would not show this relationship.
- ❖ Using the graphical structure of the text shows where exactly on the page the equations were written in relation to the text of the first layer.

## Works Cited

- “11 Representation of Primary Sources - The TEI Guidelines.” *Text Encoding Initiative*, <https://tei-c.org/release/doc/tei-p5-doc/en/html/PH.html>. Accessed 12 Feb. 2022.
- “14 Tables, Formulæ, Graphics, and Notated Music - The TEI Guidelines.” *Text Encoding Initiative*, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/FT.html>. Accessed 23 Jan. 2022.
- Paul M. Dirac. *Algebraic Manipulations, Doodled on an Old Stock Offer*. [http://purl.flvc.org/fsu/fd/FSUDirac\\_12\\_1b\\_0005](http://purl.flvc.org/fsu/fd/FSUDirac_12_1b_0005). FSU Digital Library, Paul A.M. Dirac Papers, 1788-1999.
- . *If an Identity with n Letters Is to Be True...* [http://purl.flvc.org/fsu/fd/FSUDirac\\_12\\_3\\_1\\_0017](http://purl.flvc.org/fsu/fd/FSUDirac_12_3_1_0017). FSU Digital Library, Paul A.M. Dirac Papers, 1788-1999.

Steven J. DeRose, et al. “What Is Text, Really?” *Journal of Computing in Higher Education*, vol. 1, no. 2, Dec. 1990, pp. 3–26.

“TEI Consortium, Eds. TEI P5: Guidelines for Electronic Text Encoding and Interchange.” *Text Encoding Initiative*, <http://www.tei-c.org/Guidelines/P5/>.

## Acknowledgements

- ❖ Special thanks to Sarah Stanley for mentoring and supporting this project
- ❖ Florida State University Special Collections and Archives, Digital Library Center, and the DigiNole Repository Managers and Developers.