On the Automatic Tracing of Intertextuality by Meaning

Walter J. Scheirer

School of Engineering and Applied Sciences, Department of Molecular and Cellular Biology, Center for Brain Science
Harvard University
What is the connection?

Fie on’t! ah fie! ’tis an unweeded garden,
That grows to seed;
things rank and gross in nature
Possess it merely.

(Hamlet, Act 1, Scene 2,
lines 135-137)

Thorns also and thistles shall it bring forth to thee;
and thou shalt eat the herb of the field;

(Genesis 3:18)
“Thus, by judicious transfers and imitation he brought it about that, in what we read from another source in his work, we prefer his version or we marvel that it sounds better than its source. Let me say that he took from others from a half to about a full line of verse, then full passages slightly changed and rewritten, but with their meaning left intact, so that their source might be obvious…”

— Macrobius on Vergil, *Saturnalia*
Intertextuality

- Kristeva: “Any text is constructed as a mosaic of quotations; any text is the absorption and transformation of another.”

- The Classicist: Two-word pairs are the most basic and common form of text reuse (*loci similes*).

- Hinds: A “poetic of corresponding inexactitude, which draws on but also distances itself from the rigidities of philological and intertextualist fundamentalisms alike.”
Computational approaches

- Lee 2007, Global Linear Models
- Ganascia et al. 2013, Hash Coding
- Smith et al. 2013, Sequence Alignment
- Büchler et al. 2014, Noisy Channel Analysis (eTRACES)
- Coffee et al. 2014, Bi-gram matching (Tessereae Project)
Limitation of lexical matching

Tesserae Benchmark: Lucan’s *Civil War* Book 1 vs. Virgil’s *Aeneid*

Bi-gram matching necessarily misses **33%** of human identified parallels
The process of reading

How did we identify the Shakespeare / Genesis parallel?

1. Examination of semantic content
   - Plant life → Gardens

2. Examination of sentiment
   - Something has gone wrong

3. Search of an internalized corpus
   - Eureka! Garden of Eden
Semantic Analysis

- Popular strategy in natural language processing
  - Designed around the notion of word co-occurrence
    - Deerwester et al. 1990, *Latent Semantic Indexing*
    - Blei et al. 2003, Latent Dirichlet Allocation
  - Commonly deployed over large corpora
    - Jockers 2013, *Macroanalysis*
Semantic analysis of small samples

We don’t always have the luxury of a lot of text:

“Thy brother’s blood the thirsty earth hath drunk.”
(King Henry VI, Part 3.)

“And it came to pass, when they were in the field, that Cain rose up against Abel his brother”
(Genesis 4:8.)
Training set

1. And God said, Let there be light: and there was light.

2. And God said, Let there be lights in the firmament of the heaven

3. And God said, Behold, I have given you every herb bearing seed,

4. And on the seventh day God ended his work which he had made;

5. And the woman said, The serpent beguiled me, and I did eat.

6. And Adam called his wife's name Eve; because she was the mother of all living.

7. And it came to pass, when they were in the field, that Cain rose up against Abel
his brother,

8. And Cain said unto the LORD, My punishment is greater than I can bear.

9. And Cain went out from the presence of the LORD, and dwelt in the land of Nod,

10. And to Seth, to him also there was born a son;
Query: “Thy brother’s blood the thirsty earth hath drunk.”

0. And God said, Let there be light: and there was light.

0. And God said, Let there be lights in the firmament of the heaven

0. And God said, Behold, I have given you every herb bearing seed,

0. And on the seventh day God ended his work which he had made;

0. And the woman said, The serpent beguiled me, and I did eat.

0. And Adam called his wife’s name Eve; because she was the mother of all living.

1. And it came to pass, when they were in the field, that Cain rose up against Abel his brother,

2. And Cain said unto the LORD, My punishment is greater than I can bear.

3. And Cain went out from the presence of the LORD, and dwelt in the land of Nod,

0. And to Seth, to him also there was born a son;
## Bag-of-Words model

7. And it came to pass, when they were in the field, that Cain rose up against Abel his brother.

8. And Cain said unto the LORD, My punishment is greater than I can bear.

9. And Cain went out from the presence of the LORD, and dwelt in the land of Nod.

---

### Documents

<table>
<thead>
<tr>
<th>Terms</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>abel</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>brother</strong></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cain</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>field</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>pass</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rose</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>bear</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>greater</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>lord</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>punishment</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>dwell</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>land</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>nod</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>presence</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

How did the Shakespearean phrase “Thy brother’s blood the thirsty earth hath drunk” match these three bags?

- Only one shared word
- Word Co-occurrence

---
LSI Algorithm: Training

1. Define a window to slide around the texts we want to match against ("documents")

2. Apply stoplist to documents

3. Sample with window to collect a bag-of-words

4. Group bags into a document-term matrix

5. Reduce dimensionality with Singular Value Decomposition (SVD)
   - Free parameter: choose the number of "topics" to retain
   - This produces the model we’ll use for matching
LSI Algorithm: Testing

1. Project a query in the model space, yielding a vector $A$

2. For all of the document windows:
   a. Project the window into the model space, yielding a vector $B$
   b. Compute cosine similarity between $A$ and $B$

3. Rank all of the similarity scores (higher is better)
Why not try LDA?

Vector of Topics

“King”

“War”

“England”

“France”
Why not try LDA?

“Thy brother's blood the thirsty earth hath drunk,
Broach'd with the steely point of Clifford's lance;
And, in the very pangs of death he cried,
Like to a dismal danger…”

500 Character query; 1000 character documents
Gensim (Rehurek and Sojka, 2010) and MALLET (McCallum, 2002)

Run 1: 45:0.981; 1:0.311; 31:0.725
Run 2: 66:0.455; 87:0.221; 45:0.156
... 
Run n: 10:0.940; 27:0.931; 54:0.922
Validation

- Small sample sizes
- Different languages
- Can we find new thematic matches?
- Can this tool help to produce new criticism?
Case Studies
Study #1: Latin Poetry

• Back to our Lucan / Vergil benchmark

  - *Civil War* Book 1: 685 hexameter lines
  
  - *Aeneid*: 9,896 hexameter lines

  - Parallels from four commentaries:
    
    ▸ Heitland and Haskins (1887), Thompson and Bruére (1968), Viansino (1995), and Roche (2009).

35 Missing Parallels (1/3 of total)
Algorithm parameters

- Gensim (https://radimrehurek.com/gensim) LSI implementation
- Query window: 500 characters around and including a target line of text
- Document window: 1,000 characters around and including a line to match against
- Stoplist: 250 most common words
- Bag-of-words: inflected form of each word replaced with the set of all possible stems
- Number of topics: 10, 15, 20
- Count a hit in the top 50 results as a “match”
Top match between a Query and Aeneid 3

qualis, cum turbidus Auster
reppulit a Libycis inmensum Syrtibus aequor
fractaque veliferi sonuerunt pondera mali,
desilit in fluctus deserta puppe magister
nauitaque et nondum sparsa conpage carinae
naufragium sibi quisque facit, sic urbe relieta
in bellum fugitur. nullum iam languidus aeo
evaluit revocare parens coniuxvxe maritum
fletibus, aut patrii, dubiae dum vota salutis
conciperent, tenuere lares; nec limine quisquam
haesit et extremo tunc forsitan urbis amatae
plenus abit visu: ruit inrevocabile volgus.
o faciles dare summa deos cademque tueri
difficiles!

(Civil War 1.498 – 511)

postquam res Asiae Priamique evertere gentem
immeritam visum superis, ceciditque superbunm
Ilium et omnis humo fumat Neptunia Troia,
diversa exsilia et desertas quaerere terras
auguriis agimur divum, classemque sub ipsa
Antandro et Phrygiae molimur montibus Idae,
incerti quo fata ferant, ubi sistere detur,
contrahimusque viros. vix prima inceperat aetas
et pater Anchises dare fatis vela iubebat,
litora cum patriae lacrimans portusque relinquuo
et campos ubi Troia fuit. feror exsul in altum
cum sociis natoque penatibus et magnis dis.

(Aeneid 3.1-12)
Just as when the swirling south wind drives the vast
sea back from the Libyan Syrtes, and the shattered
mass of the mast, with its sail, groans, the helmsman
abandons the stern and leaps into the waves; and
though the fittings of the hull are not yet strewn
apart, each sailor fashions his own personal
shipwreck; so too they desert the city and flee into
war. Parents, frail with age, cannot call back their
sons, nor wives, by their tears, their husbands; nor
the ancestral homes, so long as they place their
hopes on an unlikely salvation. No one hesitated on
his threshold, to depart, perhaps, with a final look,
filled with the love of his city. The crowd rushed on,
heedless. How easily the gods give everything, how
little they care to preserve it.

(Aeneid 3.1-12)

*Translation by C. Forstall & N. Coffee*
Analysis

• Shared themes:
  - Abandonment (diversa exsilia et desertas quaerere terras, litora cum patria lacrimans portusque relinquo)
  - Naval imagery (classem, vela, portus)

• Roche: BC 1.504-7 and AEN Book 2; **AEN 3.1-12**
  - Contrasting imagery (Aeneas’s concern, Roman disregard)
Bag-of-Words Control

• Does dimensionality reduction via SVD actually add anything?
  - Cosine similarity between bag-of-words representations
  - If LSI is better, we should see better match scores for relevant parallels
**Recovered Civil War 1 (BC) – Aeneid (AEN) commentator parallels**

<table>
<thead>
<tr>
<th>BC Line</th>
<th>AEN Line</th>
<th>Shared Context</th>
<th>Topics</th>
<th>LSI Rank</th>
<th>LSI Prec.</th>
<th>BoW Rank</th>
<th>BoW Prec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.60</td>
<td>1.291</td>
<td>Destiny of Caesar; peace</td>
<td>10</td>
<td>3</td>
<td>0.18</td>
<td>86</td>
<td>0.00</td>
</tr>
<tr>
<td>1.139</td>
<td>4.441</td>
<td>The blowing wind; tree</td>
<td>20</td>
<td>1</td>
<td>0.22</td>
<td>1</td>
<td>0.28</td>
</tr>
<tr>
<td>1.141</td>
<td>2.626</td>
<td>The blowing wind; tree</td>
<td>15</td>
<td>1</td>
<td>0.32</td>
<td>45</td>
<td>0.00</td>
</tr>
<tr>
<td>1.193</td>
<td>2.774</td>
<td>An apparition</td>
<td>20</td>
<td>33</td>
<td>0.08</td>
<td>47</td>
<td>0.00</td>
</tr>
<tr>
<td>1.193</td>
<td>3.47</td>
<td>An apparition</td>
<td>15</td>
<td>42</td>
<td>0.06</td>
<td>145</td>
<td>0.00</td>
</tr>
<tr>
<td>1.291</td>
<td>11.492</td>
<td>Horses</td>
<td>20</td>
<td>26</td>
<td>0.02</td>
<td>212</td>
<td>0.00</td>
</tr>
<tr>
<td>1.490</td>
<td>11.142</td>
<td>Flight</td>
<td>15</td>
<td>17</td>
<td>0.22</td>
<td>23</td>
<td>0.08</td>
</tr>
<tr>
<td>1.504</td>
<td>2.634</td>
<td>Abandonment</td>
<td>15</td>
<td>3*</td>
<td>0.52</td>
<td>70</td>
<td>0.00</td>
</tr>
<tr>
<td>1.504</td>
<td>3.11</td>
<td>Abandonment; Navy</td>
<td>15</td>
<td>4</td>
<td>0.16</td>
<td>215</td>
<td>0.00</td>
</tr>
<tr>
<td>1.673</td>
<td>2.199</td>
<td>Omens; terror</td>
<td>15</td>
<td>31</td>
<td>0.02</td>
<td>162</td>
<td>0.00</td>
</tr>
<tr>
<td>1.676</td>
<td>4.68</td>
<td>Dido as Bacchant</td>
<td>15</td>
<td>1</td>
<td>0.40</td>
<td>2</td>
<td>0.08</td>
</tr>
<tr>
<td>1.676</td>
<td>6.48</td>
<td>Prophecy</td>
<td>15</td>
<td>39</td>
<td>0.44</td>
<td>148</td>
<td>0.00</td>
</tr>
<tr>
<td>1.695</td>
<td>6.102</td>
<td>Frenzied discussion</td>
<td>20</td>
<td>22</td>
<td>0.20</td>
<td>31</td>
<td>0.20</td>
</tr>
</tbody>
</table>

* also found by Tesserae lexical matching

**Recovered 12 missing parallels**
Additional recovered Civil War 1 (BC) – Aeneid (AEN) commentator parallels with Bag-of-Words approach

<table>
<thead>
<tr>
<th>BC Line</th>
<th>AEN Line</th>
<th>Shared Context</th>
<th>BoW Rank</th>
<th>BoW Prec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>4.628</td>
<td>War</td>
<td>48</td>
<td>0.02</td>
</tr>
<tr>
<td>1.8</td>
<td>12.313</td>
<td>Hostility</td>
<td>23</td>
<td>0.06</td>
</tr>
<tr>
<td>1.226</td>
<td>4.624</td>
<td>Broken treaty</td>
<td>13</td>
<td>0.10</td>
</tr>
<tr>
<td>1.226</td>
<td>12.435</td>
<td>Fortune</td>
<td>38</td>
<td>0.04</td>
</tr>
<tr>
<td>1.674</td>
<td>4.300</td>
<td>Dido as Bacchant</td>
<td>28</td>
<td>0.02</td>
</tr>
<tr>
<td>1.678</td>
<td>10.670</td>
<td>Questioning destination</td>
<td>17</td>
<td>0.16</td>
</tr>
<tr>
<td>1.685</td>
<td>2.554</td>
<td>Decapitation; shore</td>
<td>45</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Recovered 7 additional missing parallels
Exploratory work

<table>
<thead>
<tr>
<th>BC Line</th>
<th>AEN Line</th>
<th>Shared Context</th>
<th>Topics</th>
<th>LSI Rank</th>
<th>BoW Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>4.98</td>
<td>War</td>
<td>15</td>
<td>1*</td>
<td>1</td>
</tr>
<tr>
<td>1.141</td>
<td>2.252</td>
<td>The blowing wind</td>
<td>15</td>
<td>6</td>
<td>128</td>
</tr>
<tr>
<td>1.291</td>
<td>11.291</td>
<td>Conquest</td>
<td>20</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1.353</td>
<td>1.647</td>
<td>City; nation</td>
<td>15</td>
<td>1*</td>
<td>26</td>
</tr>
<tr>
<td>1.504</td>
<td>3.639</td>
<td>Abandonment; nautical imagery</td>
<td>20</td>
<td>2</td>
<td>129</td>
</tr>
<tr>
<td>1.676</td>
<td>6.809</td>
<td>Bacchus</td>
<td>15</td>
<td>1</td>
<td>81</td>
</tr>
<tr>
<td>1.676</td>
<td>4.304</td>
<td>Bacchus</td>
<td>10</td>
<td>36</td>
<td>295</td>
</tr>
</tbody>
</table>

* also found by Tesserae lexical matching
Civil War 1.504 also matched:

sed fugite, o miseri, fugite atque ab litore funem rumpite.

nam qualis quantusque cavo Polyphemus in antro lanigeras claudit pecudes atque ubera pressat,

centum alii curva haec habitant ad litora vulgo infandi Cyclopes et altis montibus errant.

*But flee, you wretches, flee and slash the cables from the shore. For as great and tall as Polyphemus is who lives in his hollow cave, keeps wooly flocks, and milks their udders, a hundred such other monstrous Cyclopes live together along the curved shore, and wander the steep mountains.*

*(Aeneid 3.639-44)*
Bacchus re-contextualized

nam, qualis vertice Pindi
Edonis Ogygio decurrit plena Lyaeo . . .

(Civil War 1.674-5)

nec qui pampineis victor iuga flectit habenis
Liber, agens celso Nysae de vertice tigris

(Aeneid 6.804-5)

For just as a Thracian bacchant, filled
with Theban Bacchus, rushes down from
the summit of Mount Pindus . . .

Nor did Bacchus, who in victory guides his chariot
with reins of vine, leading his tigers from the
summit of lofty Nysa, [traverse as much land as
Augustus will rule].

And

BC 1.676 and AEN 4.300-3
Discussion

• Recovered 19/35 of the missing parallels

• What can we do better?
  - Algorithms that exploit context more like human readers
  - Revisit LDA? Need a solver that can handle small inputs
Study #2: Fan fiction on the web

wattpad.com:
- 2 million writers
- 100,000 new pieces per day
- 20 million readers
Song, dialogue, and the novel

“Some novelists draw on their own experience; others borrow from history, mythology or classic literary tropes. Anna Todd, a 25-year-old debut novelist in Texas, found inspiration in Harry Styles, the tousle-haired heartthrob from the British boy band One Direction.”

Why fan fiction?

• Emerging literatures

• Direct participation in a popular narrative

• Interactive engagement with a text

• Which new texts will be successful? Publishers want to know
After

- Atypical style for a romance novel
  - Conversational first person narrative reminiscent of social media

- Obvious intertexts signaled via direct citation

Does the novel share something more with a broader literary culture?
“Today will be our last day on Pride and Prejudice, I hope you all have enjoyed it, for today’s discussion we will be talking about Austen’s use of foreshadowing. As a reader, did you expect her [Elizabeth] and Darcy to end up together in the end?” Professor asks and I raise my hand as always. Liam and I are usually the first to answer, and usually the only.

“Miss Young” he calls on me.

“The first time I read the novel I was on the edge of my seat to see whether they would end up together. Even now, as I have read it at least ten times I still feel anxious during the beginning of their relationship. Mr. Darcy is so cruel and says hateful things about Elizabeth and her family so I didn’t know if she would forgive him, let alone love him.” I answer and smile.

(“After,” Chpt. 24, Page 2)
A Silent Prayer

- “Supernatural” Romance
- Also a statement on the first generation immigrant experience
  - Conflict between religious tradition and contemporary culture
- Clever use of allusion and a plot framed by the Qu’ran
“No, there are many, more than the human population. They have a world within our world, but we can’t see them unless they want to be seen. There are males and females. There are good ones and bad ones, as well. You know what the interesting part is? Each human has been assigned a Jinni, a demon. We don’t see it but it is always there. We recognize it when our soul is not strong enough to protect our body and our demon takes over, which makes us do all of the possible sins.”

(“A Silent Prayer,” Chpt. 16, Page 4)

Yet the foolish among us hath spoken that which is extremely false of GOD; but we verily thought that neither man nor genius would by any means have uttered a lie concerning GOD. And there are certain men who fly for refuge unto certain of the genii; but they increase their folly and transgression: and they also thought, as ye thought, that GOD would not raise any one to life. And we formerly attempted to pry into what was transacting in heaven; but we found the same filled with a strong guard of angels, and with flaming darts: and we sat on some of the seats thereof to hear the discourse of its inhabitants; but whoever listeneth now, findeth a flame laid in ambush for him, to guard the celestial confines.

(The Qu’ran, Chapter 72, 4-9)
A Silent Prayer and the Bible

“Humans are made of clay. God collected seven different kinds of clay from the earth to create Adam. Clay supports and gives life, whereas, fire destroys and burns.” She paused for a moment and continues. “But like Satan, humans also carry pride, envy and arrogance, which sometimes work as a fire. There is a very thin line between humans and Jinn. Though they are invisible to humans, they do exist.

(“A Silent Prayer,” Chpt. 11, Page 1)

• Creation story is retold above
• The character of Adam shares his name with the biblical Adam
  - Emphasizes the genre’s trope of the primal man
• Adam’s sister is named “Eva”
  - Brother and sister are symbolic of temptation
A Silent Prayer and Ismael ibn Kathir

Even after my rudeness, not calling him and thanking him for his kindness, he is still worried about me. After so many years, someone is worried about me. No one would believe that I’ve seen another side of Adam, which no one knows. One who is not thankful to a person, is not thankful to God. The prophet’s quote buzzes in my mind and I decide I will call him to say thanks during my journey.

(“A Silent Prayer,” Chpt. 11, Page 1)

“This means respond to the poor with mercy and gentleness.” And proclaim the grace of your Lord. meaning, just as you were poor and needy, and Allah made you wealthy, then tell about Allah’s favor upon you. Abu Dawud recorded from Abu Hurayrah that the Prophet said, Whoever is not thankful to the people, then he is not thankful to Allah.

(Tafsir, Book 93, 216-225)
Corpus

- Primary works: complete texts of *After* and *A Silent Prayer* from wattpad.com

- Comparison works:
  - The Book of Genesis (Authorized Version, digital text from Project Gutenberg)
  - The Qu’ran (translated by George Sale, digital text from Project Gutenberg)
  - *Tafsir* of ibn Kathir (digital text from the Internet Archive)
Algorithm Parameters

- R “lsa” package
- Query window: 100 words
- Document window: 250 words
- Stoplist: 650 common words (Jockers 2014)
- Thresholds over cosine similarities
Qu’ran Chapter 72 and *A Silent Prayer* Chapter 16

- Three query samples from Qu’ran Chpt. 72
- All samples from *A Silent Prayer*
- Threshold: 0.9; 20 topics
  - **Query 1**: Chapter 16, Sample 4 (0.906)
  - **Query 2**: Chapter 16, Sample 3 (0.908)
  - **Query 3**: Chapter 16, Sample 3 (0.906) 100% Precision
- Control: all samples from *After*
  - No matches
Tafsir Chapter 93 and A Silent Prayer
Chapter 11

- One query sample for Tafsir
- All samples from A Silent Prayer
- Threshold: 0.75; 10 topics

- Chapter 6, Sample 4 (0.906)
- Chapter 7, Sample 3 (0.785)
- Chapter 7, Sample 6 (0.802)
- Chapter 7, Sample 8 (0.772)
- Chapter 11, Sample 1 (0.753)
- Chapter 26, Sample 2 (0.817)
- Chapter 26, Sample 6 (0.809)
- Chapter 26, Sample 7 (0.766)
- Chapter 28, Sample 3 (0.777)
- Chapter 28, Sample 4 (0.795)

10% Precision
Matches to Genesis

τ = 0.85
Discussion

- *After* lacks literary merit compared to *A Silent Prayer*, but is massively popular
- Large scale automated study to understand readership trends
- How can we model an overarching relationship between a novel and its framing material?
Wrapping up...
Resources


• Already in Tesserae
  - http://tesserae.caset.buffalo.edu/cgi-bin/lsa.pl

• R implementation will be on GitHub in the coming month
Acknowledgements

Chris Forstall (U. Geneva)

Neil Coffee (U. Buffalo)

Start-Up Grant Award #HD-51570-12
Quantitative Intertextuality

Authors:
Walter Scheirer, Harvard University
Christopher Forstall, University at Buffalo

A remarkable amount of information crosses our eyes and ears each day, yet we adeptly identify what is familiar with seemingly no effort at all. In many cases, what we see or hear has been shaped by recognizable prior sources. Such instances of intertextuality reveal a wealth of data about authorship, influence, and style, making them attractive targets for automatic identification. This volume introduces quantitative intertextuality, a new approach for the algorithmic study of information reuse in text, sound and images. Using a variety of tools drawn from machine learning, natural language processing, and computer vision, readers will learn to trace patterns of reuse across diverse sources for scholarly work and practical applications.

This is the first book to offer a rigorous, quantitative approach to the study of intertextuality. The authors are recognized experts in their respective fields, and have planted the seeds for this work by co-developing popular tools for textual and visual analysis. From practitioners specializing in forensics to students of cultural studies, readers from diverse backgrounds will find something of interest.
Questions?