The Remaking of Reading: Data Mining and the Digital Humanities

Matthew Kirschenbaum
University of Maryland
Google Books

Interesting

Classics
The Internet Archive
Dreams of the Burning Child: Sacrificial Sons and the Father's Witness (Hardcover)
by David Lee Miller (Author)
Key Phrases: The Winter's Tale, Ben Jonson, Things Fall Apart (more...)
No customer reviews yet. Be the first.

List Price: $41.00
Price: $30.11 & this item ships for FREE with Super Saver Shipping. Details
You Save: $10.89 (27%)

Only 1 left in stock--order soon (more on the way).

Want it delivered Thursday, October 4? Order it in the next 21 hours and 51 minutes, and choose One-Day Shipping at checkout. See details

5 used & new available from $30.11

Keep connected to what's happening in the world of books by signing up for Amazon.com Books Delivers, our monthly subscription e-mail newsletters. Discover new releases in your favorite categories, popular pre-orders and bestsellers, exclusive author
# Amazon.com Text Stats

**Readability**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>Easier Percentage</th>
<th>Harder Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fog Index</td>
<td>17.8</td>
<td>89% are easier</td>
<td>11% are harder</td>
</tr>
<tr>
<td>Flesch Index</td>
<td>37.7</td>
<td>70% are easier</td>
<td>30% are harder</td>
</tr>
<tr>
<td>Flesch-Kincaid Index</td>
<td>15.0</td>
<td>90% are easier</td>
<td>10% are harder</td>
</tr>
</tbody>
</table>

**Complexity**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>Fewer Percentage</th>
<th>More Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex Words</td>
<td>17%</td>
<td>60% have fewer</td>
<td>40% have more</td>
</tr>
<tr>
<td>Syllables per Word</td>
<td>1.7</td>
<td>59% have fewer</td>
<td>41% have more</td>
</tr>
<tr>
<td>Words per Sentence</td>
<td>28.0</td>
<td>94% have fewer</td>
<td>6% have more</td>
</tr>
</tbody>
</table>

**Number of**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>Fewer Percentage</th>
<th>More Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characters</td>
<td>555,709</td>
<td>69% have fewer</td>
<td>31% have more</td>
</tr>
<tr>
<td>Words</td>
<td>92,507</td>
<td>70% have fewer</td>
<td>30% have more</td>
</tr>
<tr>
<td>Sentences</td>
<td>3,302</td>
<td>48% have fewer</td>
<td>52% have more</td>
</tr>
</tbody>
</table>

**Fun stats**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words per Dollar</td>
<td>3,072</td>
</tr>
<tr>
<td>Words per Ounce</td>
<td>5,256</td>
</tr>
</tbody>
</table>
act aeneas aeneid again already between body book boy burning calls
cannot child children come culture dead death desire dickens does
dombey dream economy even fantasy father fatherhood figure fillial
first form freud god hamlet hand himself history human image ing itself
jonson king leontes line little loss lost love makes man may might moment
mother must name narrative new now once own passage place play poem
point power priam question re relation ritual sacrifice sacrificial scene
see seems sense shakespeare social something son story subject takes
tale things thought time tion turns virgil whose witness words work world yet
“Humankind writes more than it can read.”
–Gabriel Zaid
But is Reading “at Risk”?

- NEA, 2004

“For the first time in modern history, less than half of the adult population now reads literature, and these trends reflect a larger decline in other sorts of reading.”

<table>
<thead>
<tr>
<th>Table 17</th>
<th>Literary Reading: U.S. Adults, 1982, 1992, and 2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number reading literature (in millions)</td>
<td>95.6</td>
</tr>
<tr>
<td>American population 18 or older (in millions)</td>
<td>168.0</td>
</tr>
<tr>
<td>% reading literature</td>
<td>56.9%</td>
</tr>
</tbody>
</table>
But What is Reading?

- “Cover to cover”
- Excerpts, selections
- Grazing, browsing
- Skimming, scanning
- Speed reading
- Reference/look-up
Thomas Jefferson’s Revolving Bookstand
“Not Reading”

“As long as there have been books there have been more books than you could read. In the life of a professional or scholar, reading in the strong sense of "close reading" almost certainly takes a back-seat to finding out what is in a book without actually reading all or even any of it. There are age-old techniques for doing this, some more respectable than others, and they include skimming or eyeballing the text, reading a bibliography or following what somebody else says or writes about it. Knowing how to "not-read" is just as important as knowing how to read.” –Martin Mueller
A Christmas Gift to a Dear Child in Memory

Chapter 1

Alice was beginning to get very tired of sitting by her sister on the sofa, and of having nothing to do. Once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, 'and she could read it herself sometimes, when her sister wasn’t looking. There was nothing it was any good to think about, she was sure of that.

It was all a mistake about the rabbit, she thought, with a yawn and rub of eyes, that perhaps it had said to itself “Dear, dear! I shall be too late!” (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural); but when the rabbit actually took a watch out of its waistcoat-pocket, looked at it, and then hurried on, Alice started to her feet, for
The adventures of Tom Sawyer (1876)

Author: Twain, Mark, 1835-1910
Digitizing Sponsor: Yahoo!
Book Contributor: University of California Libraries
Language: English
Keywords: Adventure and adventurers; Mississippi River - Juvenile literature; Missouri - Juvenile literature

Special Collections copy 1 = Blue cloth over boards blocked in gold and black; upper board partially detached. Gift of Wilbur Jordan Smith. Printed and illustrated label, Paul Jordan Smith, on upper paste-down. 2 other printed labels on upper paste-down, indicating copy 1 is "second issue...with the frontispiece on verso of the half-title, and the first page of Contents printed on the verso of the first of Preface."

Reviews

Average Rating: 5/5

Reviewer: goodjob! - August 2, 2007
Subject: Always good
Always good

Selected Metadata

Title: The adventures of Tom Sawyer
Creator: Twain, Mark, 1835-1910
Did you know that you can help us produce ebooks by proof-reading just one page a day? Go to: Distributed Proofreaders

Online Book Catalog - Overview

Note: we also have offline book catalogs to download and use at home.

Browse by Author, Title, Language or Recently Posted

Our browse pages are ideal to view what's in the collection if you are yet undecided on what you want to read.

The recently posted pages list what new books got added or updated most recently. There is also an RSS Feed (You'll need a feed reader software to read this.)

Freshness: updated nightly.

Authors: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z other

Titles: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z other

Languages with more than 50 books: Chinese Dutch English Finish French German Italian Portuguese Spanish Tagalog

Languages with up to 50 books: Afrikaans Aleut Bulgarian Catalan Cebuano Czech Danish Esperanto Frisian Friulano Galician Gamara Ray Gascon Greek Hebrew Hungarian Icelandic Edo Edo Interlingua Irish Iroquoiian Japanese Khasi Khowar Lithuanian Medium Maltese Mapudungun Miskito Modena Navajo Nias Nde Mende Mya Na Ngiemboon Ojibwe Osage Papiamento Papiamentu Peruvian Quechua Pidgin Polish Punjabi Pulaar Quechua Romani Rongorongo Sranan Samoan Somali Sotho Southern Tswana Spanish Yiddish Yoruba Zulu

Project Gutenberg needs your donation! More Info
With over a third of a million full-text works of poetry, prose and drama in English, together with the definitive online criticism and reference library, Literature Online is the world's largest cross-searchable database of literature and criticism.

Use our interactive demo to find out more about the unique content and features of Literature Online.

LOG IN:
Username: 
Password: 
Log in
and hands are so full of our own things that we have no time for tasks for Him. We need never fear that our hands will be any less full at the end because we have slackened our pace a little now and then to do some slight wayside service for Christ.

There is a story of one who began to run in a race, and ran well—was foremost of all. But by-and-by he stopped to lift up a fallen child and place it out of danger, thus losing something of what he had gained. Farther on, a fainting comrade appealed to his sympathy, and he turned aside to help him.
The idealization that assimilates Hamlet’s violence to the morality of a later age shows itself capable of even greater miracles when it assimilates Betterton at age seventy to an image of youthful expectation. It is almost as if the actor were making good on Hamlet’s taunt to Polonius: the truth about old men’s bodies should not be bluntly set down, “for yourself, sir, shall grow old as I am, if like a crab you could go backward.” Later in the same scene Hamlet marvels at the transformative powers of make-believe:

Is it not monstrous that this player here,
But in a fiction, in a dream of passion,
Could force his soul so to his own conceit
That from her working all his visage wann’d,
Tears in his eyes, distraction in his aspect,
A broken voice, an’ his whole function suit’d
With forms to his conceit? And all for nothing.
(2.2.551–57)

Together these passages form a kind of Shakespearean matrix for Mr. Greenhalgh’s account of the admired Betterton: is it not monstrous that this player here can, like a crab, go backward, forcing his soul so to his own conceit that forty years drop away from his visage? He works this miracle
A Poem on a Page Exists as a Unified Visual Event

so much depends
upon

a red wheel
barrow

glazed with rain
water

beside the white
chickens.
But the Codex Atomizes a Text’s Visual Field
Word Cloud

alone around away bed
behind believe better bit
course creep dear
door enough ever eyes fancy
feel getting head house
jennie lie myself nervous night nothing
paper pattern really
quiet quite room says seems shall
sleep something things
sometimes think tired used
wall-paper windows
wish woman wonder yellow
Alice's Adventures In Wonderland
Nora: “Her breast is fit for pearls . . .”
Nora: “A full fed Rose . . .”
“Everyone then sometime is a whole one to me, everyone then sometime is a whole one in me, some of these do not for longs times make a whole one to me inside me. Some of them are a whole one in me and then they go to pieces again inside me, repeating comes out of them as pieces to me, pieces of a whole one that only sometimes is a whole one in me.”
The Remaking of the Reading of *The Making of Americans*

Image courtesy of Tanya Clement.
Mind the Gaps (Chapters 1 and 4 Especially)

Detail of narrative gaps, chapters 1-6

*Image courtesy of Tanya Clement.*
Longest Repetition (Highest on Y-Axis) Exactly in Middle

Repetitions charted for length per location

Image courtesy of Tanya Clement.
Reading the Z-Axis

Above, repetitions charted by ID, length and count (A); Right, repetitions charted by ID, length and count (B)

Images courtesy of Tanya Clement.
The Text’s Ring Structure

![Diagram of the Text’s Ring Structure]

- **Untitled**
  - 2: Mr. David Hersland and Knowledge (ontological)
  - 3: Martha Hersland’s failed marriage

- **“Martha Hersland”**
  - 4: Contradictions

- **“History of a Family’s Progress”**
  - 5: Contradictions
  - 6: Alfred Hersland’s failed marriage
  - 8: David Hersland and Knowledge (epistemological)
  - 7: “David Hersland
  - 9: “Alfred Hersland And Julia Dehning”

*Image courtesy of Tanya Clement.*
Conclusion: “Digitization”
Questions, Comments?