



THE SIXTH WORLD CONGRESS ON EXTENSIVE READING

7-11 August 2023

Bridging the LitERacy Gap

Combining ER and DR(Distant Reading) Using Literary Texts with Computer Coding

- Youngwoo KIM
- Hyunkee AHN
- Yongkook WON
- Kwanghyun PARK
- Junghye Fran CHOI
- Joohae KIM



Reading is important, but ...

Americans Reading Fewer Books, on Average

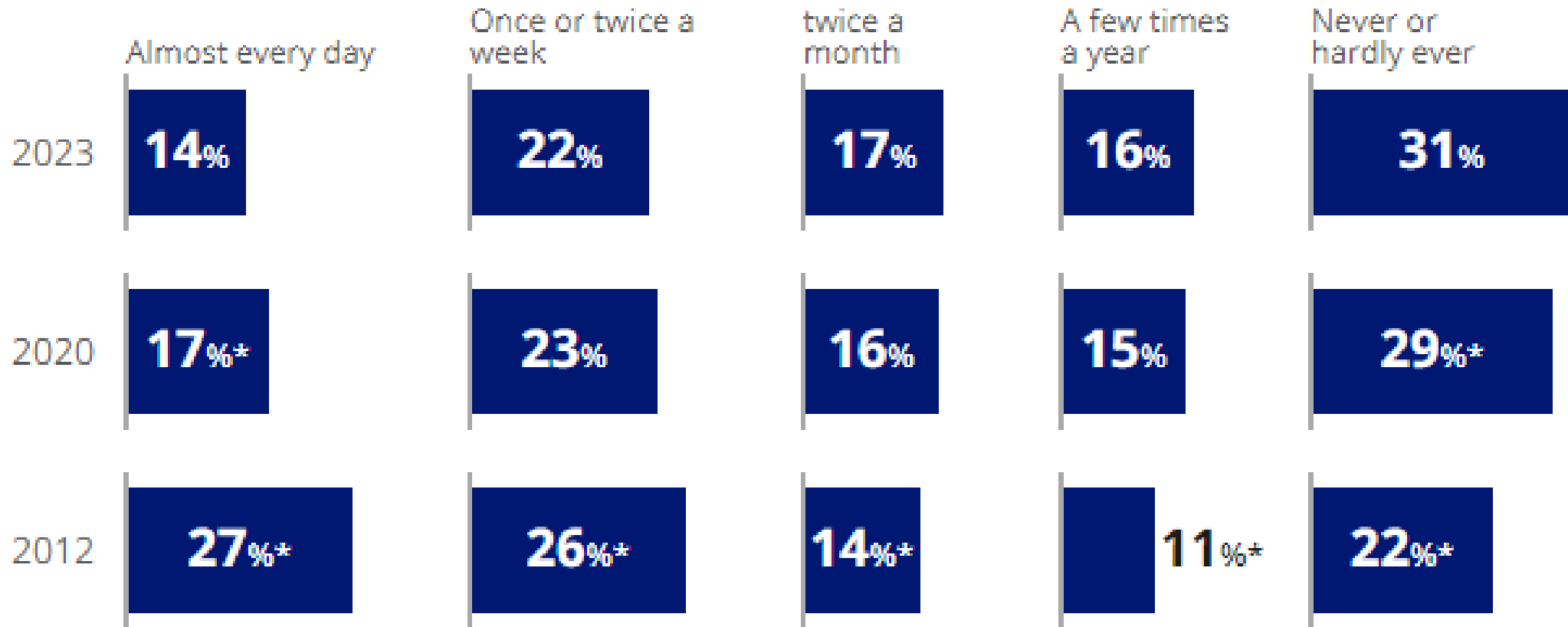
During the past year, about how many books did you read, either all or part of the way through?

— Average number of books



Americans read fewer books.

<https://news.gallup.com/poll/388541/americans-reading-fewer-books-past.aspx>



Among many U.S. children, reading for fun has become less common.

<https://www.nationsreportcard.gov/highlights/ltt/2023/>

Children are reading less than ever before, research reveals

Just a quarter of under-18s read each day, study shows in run-up to World Book Day



📷 Just 53% of children said they enjoyed reading 'very much' or 'quite a lot'. Photograph: Danny Lawson/PA

Children are reading less than ever before in U.K.

<https://www.theguardian.com/education/2020/feb/29/children-reading-less-says-new-research>



What about people in your country?

Reading is good for

- fun,
- experience, and
- knowledge.

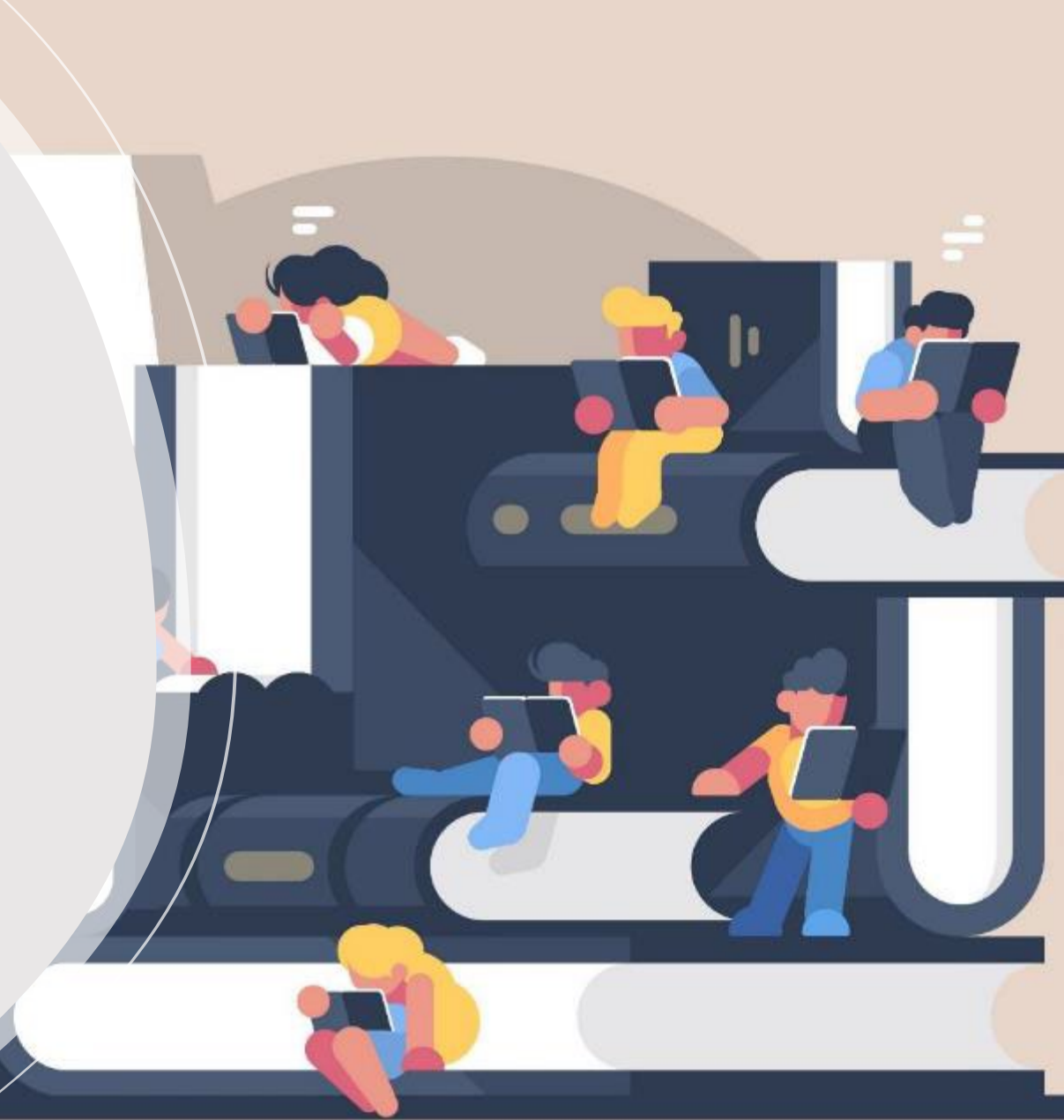
However, other activities such as watching videos, playing games, and so on, can also provide enjoyment, experiences, and knowledge.



Then, why
we still need
to read?



There are
many reasons.
Among them,





Reading ensures time for thinking.

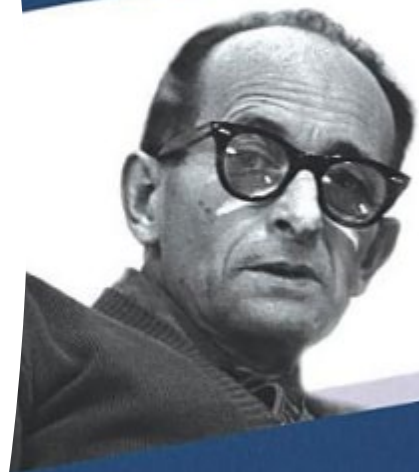


And Extensive Reading is good for thinking.



Because ER
provides a
variety of
inputs.

She [Arendt] concluded that Eichmann's inability to speak coherently in court was connected with his incapacity to think, or to think from another person's point of view. His shallowness was by no means identical with stupidity (p.14).



PENGUIN  CLASSICS

HANNAH ARENDT

EICHMANN IN JERUSALEM

Evil comes from a failure to think. It defies thought for as soon as thought tries to engage itself with evil and examine the premises and principles from which it originates, it is frustrated because it finds nothing there. That is the banality of evil (p.15).



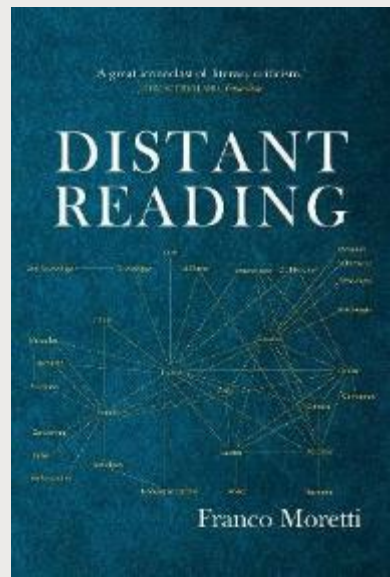
However, ER requires a significant time commitment that is precious in this era.



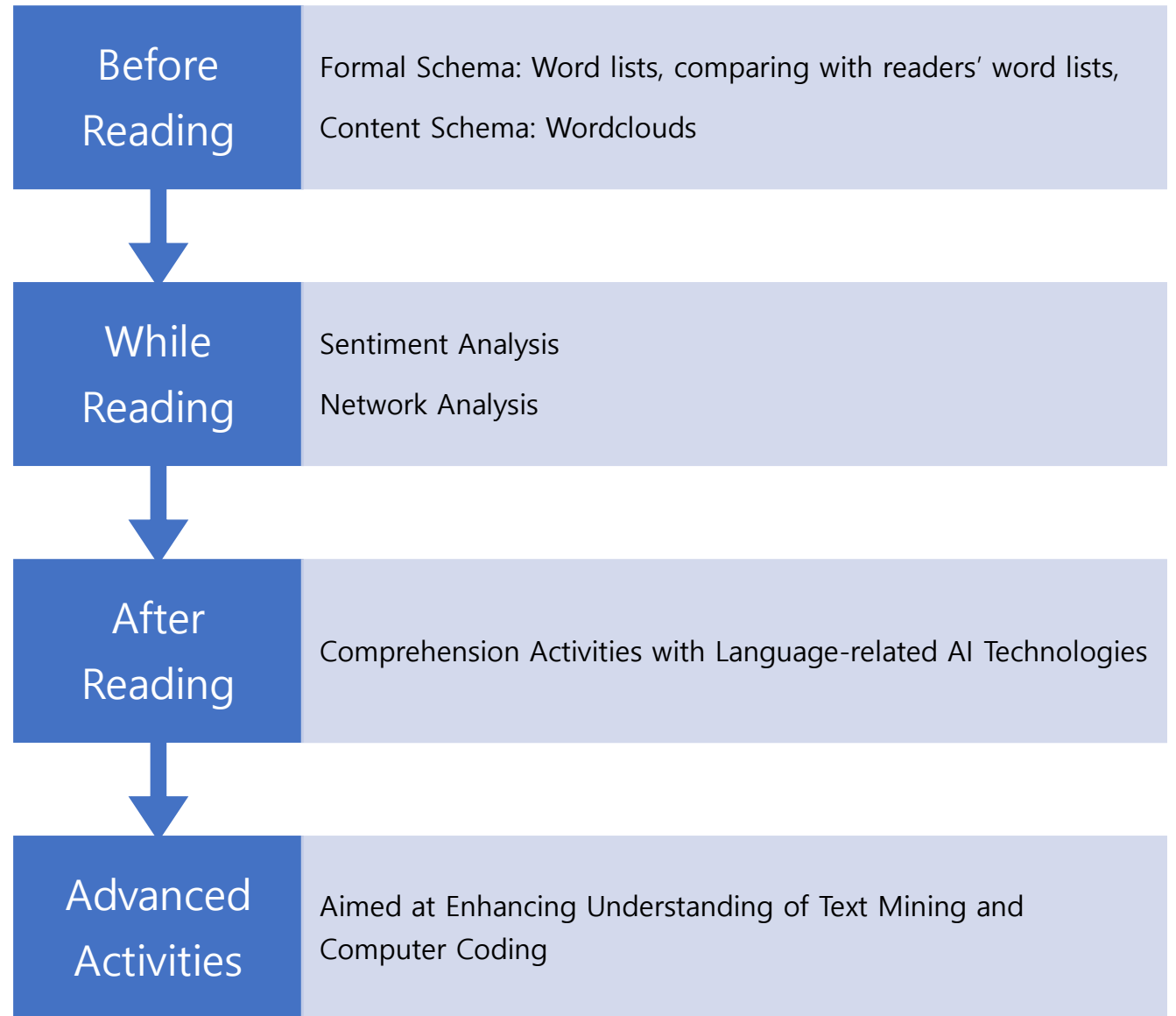
Text data analysis is gaining a lot of attention under the name of text mining.



Distant reading
with text mining
can be
combined with
ER.



Here are
some text
mining
methods
for ER.



Frequency analysis enables the creation of a list of words

that appear frequently in literary texts.

Such list of words can contribute to the formation of a formal schema, which assists readers in understanding literary texts.

	word	n
1	gatsby	197
2	tom	177
3	daisy	150
4	house	95
5	eyes	89
6	looked	85
7	time	83
8	car	73
9	night	73
10	moment	70

11	door	68
12	gatsby's	67
13	jordan	64
14	hand	63
15	people	62
16	wilson	62
17	voice	58
18	girl	47
19	half	47
20	sport	47

Words from literary texts can be compared with

words from readers' word lists.

Readers can improve their vocabulary skills cumulatively by comparing the words of literary texts with the words from their own word lists.

	A	B	C
	word	n	source
5	sport	47	gls2000
0	afternoon	42	gls2000
0	suddenly	40	gls2000
8	baker	38	gls2000
1	garage	28	gls2000
8	inquired	26	gls2000
4	feet	23	gls2000
6	hair	22	gls2000
3	yellow	22	gls2000
4	chair	19	gls2000
3	inside	19	gls2000
7	telephone	19	gls2000
6	dozen	18	gls2000
9	grey	18	gls2000
2	island	18	gls2000
9	slowly	18	gls2000
9	rain	17	gls2000

Wordclouds can be created

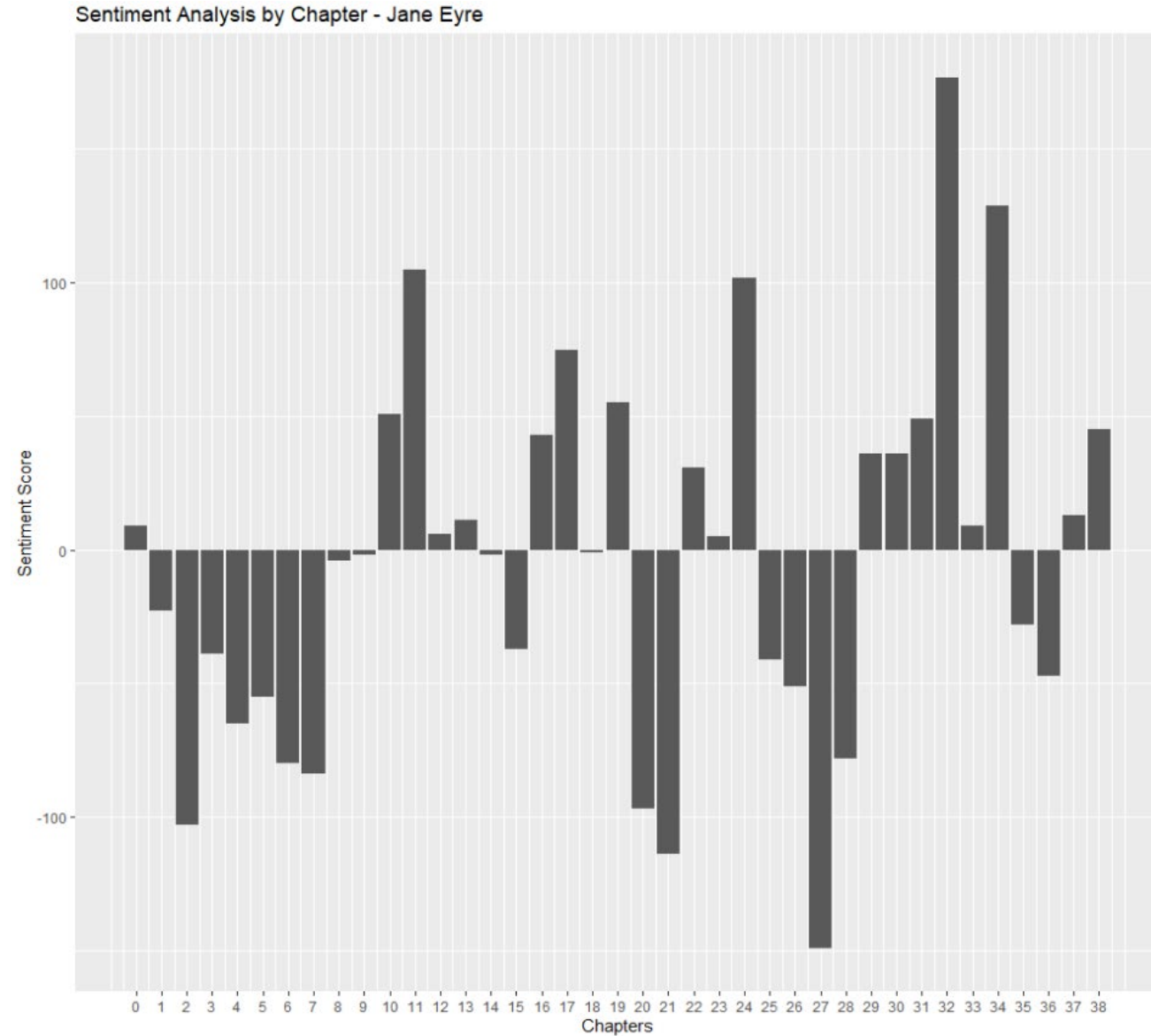
using words from literary texts.

Wordclouds can assist in forming content schemas, which aid readers in understanding literary texts.



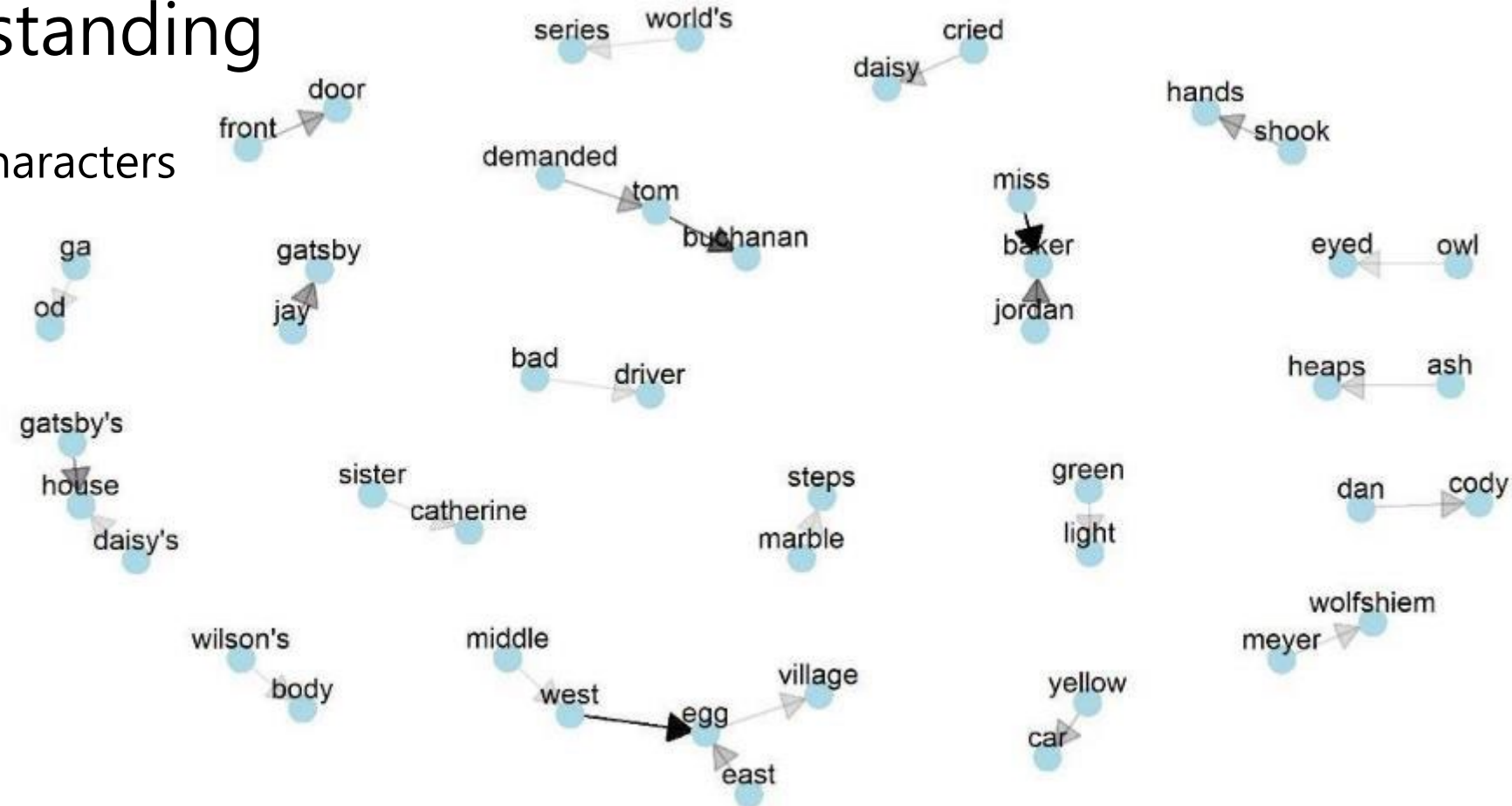
By conducting sentiment analysis according to the chapter,

it is possible to gain an understanding of the development process of a literary work.



Network analysis can assist in understanding

the characteristics of characters and the relationships within literary texts.



Improving the
comprehension
of literary texts
can be achieved

through various language-related
AI technologies.



VS



VS

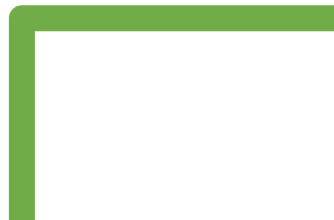




Through advanced activities, readers can acquire the skills

to analyze the language data of literary texts using computer coding techniques.

```
er - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
code_LaskarPelangi_RainbowTroops.R*
# Activity 2: Collecting Text Data
# In order to use text data in the Collecting Text Data activity of this
# project, the text of 'Laskar Pelangi' and 'Rainbow Troops' are secured.
# Loading packages
library(tidyverse)
# Setting a file folder
folder <- "literaryworks/"
# Setting text file folders
l_file <- paste0(folder, "LaskarPelangi.txt")
t_file <- paste0(folder, "TheRainbowTroops.txt")
# Getting text from text files
l_text <- readLines(l_file)
t_text <- readLines(t_file)
```





Here are sample literary works

- Jane Eyre
- The Great Gatsby
- **The Red Badge of Courage**
- Through the Arc of the Rain Forest
- Rainbow Troops / Laskar Pelangi

The image features a blue background with a repeating geometric pattern of small triangles. A white, torn-paper-like shape is cut out from the center, containing the word "DEMO" in a bold, dark blue, sans-serif font. The bottom edge of the image is also torn, revealing a white background.

DEMO

Demonstration

Extensive Reading and Language Data Activities



1. CHOOSE A BOOK AND READ IT

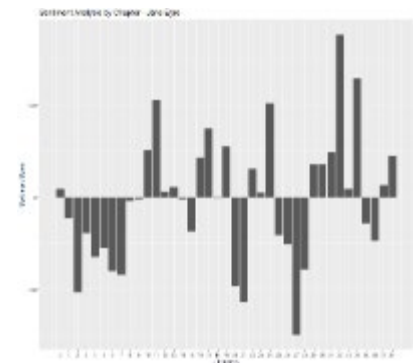


2. UTILIZE THE BOOK'S TEXT DATA

A	B	C	D
no	chapter	text	word_count
1	1	JANE EYRE	2
3	2	AN AUTOBIOGRAPHY	2
4	3	by Charlotte Brontë	3
5	4	ILLUSTRATED BY F. H. TOWNSEND	5
6	6	London	1
7	7	SERVICE & PATON	3
8	8	5 HENRIETTA STREET	3
9	9	1897	1
10	11	The illustrations	2
11	12	in this Volume are the copyright of	7
12	13	SERVICE & PATON, London	4
13	15	TO	1
14	16	W. M. THACKERAY, ESQ.	4
15	18	this Work	2
16	19	IS RESPECTFULLY INSCRIBED	3
17	21	BY	1
18	22	THE AUTHOR	2
19	24	PREFACE	1
20	25	A preface to the first edition of 'Jane Eyre' being unnecessary, I give none; this second edition demands a few words both of acknowledgment and miscellaneous remark.	27



3. MINING THE TEXT



ER and Language Data Activities

1. Choose a book and read it
 - 1) Choose a book to read. An out-of-copyright book is preferable.
Sample: *Jane Eyre: An Autobiography* by Charlotte Brontë
 - 2) Obtain the book to read. For ebooks, get the file.
Sample: *Jane Eyre* at Project Gutenberg (<https://www.gutenberg.org/ebooks/1260>)
 - 3) Read the book.

For Kindle: <https://www.amazon.com/gp/help/customer/display.html?nodeId=G5WYD9SAF7PGXRNA>

Consider using a reading tracking app to record your book reading activity.

Ex: Goodreads, StoryGraph, Basmo, Italic Type, Bookshelf, Bookly

Jane Eyre: An Autobiography

Charlotte Brontë



Jane Eyre at Project Gutenberg

Project Gutenberg [About](#) [Search and Browse](#) [Help](#)


Quick search [Donation](#)

[Project Gutenberg](#) > [71,255 free eBooks](#) > [13 by Charlotte Brontë](#)

Jane Eyre: An Autobiography by Charlotte Brontë


Jane Eyre: An Autobiography

Charlotte Brontë



Project Gutenberg

[Twitter](#) [Facebook](#) [Print](#)



Download This eBook

Format	Size			
Read this book online: HTML5	1.1 MB			
Read this book online: HTML (as submitted)	1.1 MB			
EPUB3 (E-readers incl. Send-to-Kindle)	2.8 MB			
EPUB (older E-readers)	2.8 MB			
EPUB (no images)	537 kB			
Kindle	3.1 MB			
older Kindles	3.0 MB			
Plain Text UTF-8	1.0 MB			
More Files...				

Similar Books

[Readers also downloaded...](#)

<https://www.gutenberg.org/ebooks/1260>

Amazon Kindle

<https://www.amazon.com/gp/help/customer/display.html?nodeId=G5WYD9SAF7PGXRNA>

The screenshot shows the Amazon website's help page for Kindle content. At the top, the Amazon logo is on the left, and the delivery location 'Deliver to KIM, Seoul 134' is in the center. A search bar is on the right. Below the navigation bar, the page title is 'Help & Customer Service'. A left sidebar contains a list of help topics under 'Kindle Content Help', including 'Return a Kindle Book Order', 'Identify Which Kindle Books Have Been Read', 'Kindle Content isn't Showing in Your Library', 'Check Out a Public Library Book', 'Transfer Your Kindle Content to Another Country or Region', 'Update Your Kindle Book Version', 'Enable Whispersync for Kindle Books', 'Redeem a Kindle Book Gift', 'Learn About Kindle Vella', 'Buy Tokens for Kindle Vella', 'Return a Kindle Vella Episode for Refund', 'Send Ebooks to Your Kindle E-Reader Manually', 'Learn About Kindle Unlimited', 'Update Your Kindle Unlimited Membership Payment Information', 'Cancel Your Kindle Unlimited Subscription', 'Share Kindle Unlimited Content with Amazon Household Members', and 'Buy Kindle Books for Others'. The main content area is titled 'Find more solutions' and has a search input field. Below that, the breadcrumb path is 'Digital Services and Device Support > Kindle Content Help >'. The main heading is 'Learn About Sending Documents to Your Kindle Library'. The text explains that 'Send to Kindle' is a free service for sending documents to Kindle devices and apps. A note states that starting in early 2023, MOBI files will no longer be supported. The section 'Send to Kindle for Web' describes uploading documents directly from a computer. A list of supported file types includes Microsoft Word (.DOC, .DOCX), HTML (.HTML, .HTM), RTF (.RTF), Text (.TXT), JPEG (.JPEG, .JPG), GIF (.GIF), PNG (.PNG), BMP (.BMP), PDF (.PDF), and EPUB (.EPUB).

amazon Deliver to KIM, Seoul 134 All Search Amazon

All Today's Deals Buy Again Registry Customer Service Young... 's Amazon.com Gift Cards Browsing History Sell

Help & Customer Service

< All Help Topics

Kindle Content Help

- Return a Kindle Book Order
- Identify Which Kindle Books Have Been Read
- Kindle Content isn't Showing in Your Library
- Check Out a Public Library Book
- Transfer Your Kindle Content to Another Country or Region
- Update Your Kindle Book Version
- Enable Whispersync for Kindle Books
- Redeem a Kindle Book Gift
- Learn About Kindle Vella
- Buy Tokens for Kindle Vella
- Return a Kindle Vella Episode for Refund
- Send Ebooks to Your Kindle E-Reader Manually
- Learn About Kindle Unlimited
- Update Your Kindle Unlimited Membership Payment Information
- Cancel Your Kindle Unlimited Subscription
- Share Kindle Unlimited Content with Amazon Household Members
- Buy Kindle Books for Others

Find more solutions

Digital Services and Device Support > Kindle Content Help >

Learn About Sending Documents to Your Kindle Library

Send to Kindle is a free service that allows you to send documents to your Kindle library on your Kindle devices and Kindle app. Upload documents through the web, an authorized email address, the Kindle app for iOS and Android devices, or a Send to Kindle application if available. All documents sent through Send to Kindle are protected with end to end encryption.

Note: Beginning in early 2023, you'll no longer be able to send MOBI (.AZW, .MOBI) files to your Kindle library using Send to Kindle. This change won't affect any MOBI files already in your Kindle library. You can still read them with Kindle. MOBI is an older file format and won't support the newest Kindle features for documents.

Send to Kindle for Web

Upload documents (200 MB or smaller) directly from your computer.

Supported file types include:

- Microsoft Word (.DOC, .DOCX)
- HTML (.HTML, .HTM)
- RTF (.RTF)
- Text (.TXT)
- JPEG (.JPEG, .JPG)
- GIF (.GIF)
- PNG (.PNG)
- BMP (.BMP)
- PDF (.PDF)
- EPUB (.EPUB)

Amazon Kindle

<https://www.amazon.com/gp/help/customer/display.html?nodeId=G5WYD9SAF7PGXRNA>

Send to Kindle for Web

Upload documents (200 MB or smaller) directly from your computer.

Supported file types include:

- Microsoft Word (.DOC, .DOCX)
- HTML (.HTML, .HTM)
- RTF (.RTF)
- Text (.TXT)
- JPEG (.JPEG, .JPG)
- GIF (.GIF)
- PNG (.PNG)
- BMP (.BMP)
- PDF (.PDF)
- EPUB (.EPUB)

If you share content using Send to Kindle for Web, make sure to enable **Add to your library** so that documents in your library can be downloaded to any device linked to your account.

*Take
Action*

Use Send to Kindle for Web to upload documents to your device.

Send to Kindle

Amazon Kindle



To import and read personal files on any device linked to your account, add them to your library using [Send to Kindle](#).

The screenshot shows the Amazon Kindle desktop application window titled "Multicom". The interface includes a menu bar (File, View, Go, Tools, Help), a search bar for titles and authors, and a "Kindle Store" link. The main area displays a library of documents with the following details:

Document Type	Progress	Additional Info
PDF	9%	
PDF		NEW
PDF	95%	
PDF	38%	
DOC	0%	

Each document card features a "Send to Kindle" icon (a pair of scissors) and a checkmark at the bottom, indicating the document's status.

Amazon Kindle

The screenshot shows the Amazon 'Send to Kindle' page. At the top, the Amazon logo is on the left, followed by the delivery location 'Deliver to KIM, Seoul 134', a search bar with 'All' and 'Search Amazon', and the user's account information 'Hello, Youngwoo Account & Lists'. Below the navigation bar, the main heading is 'Send to Kindle' with the tagline 'Send once, read anywhere'. The central area is divided into two sections: 'File Upload' and 'Other ways to send'. The 'File Upload' section features a dashed box for file placement, three document icons, the text 'Drag and drop files here', a yellow 'Select files from device' button, and a link to the terms of service. It also lists supported file types (PDF, DOC, DOCX, TXT, RTF, HTML, HTML, PNG, GIF, JPG, JPEG, BMP, EPUB) and a maximum file size of 200 MB. The 'Other ways to send' section lists six options: Web, iOS, Android, Microsoft Word, Email, and Chrome Extension, each with a corresponding icon.

amazon Deliver to KIM, Seoul 134 All Search Amazon Hello, Youngwoo Account & Lists

All Today's Deals Buy Again Registry Customer Service Young... 's Amazon.com Gift Cards Browsing History Sell

Send to Kindle

Send once, read anywhere

File Upload

Drag and drop files here

Select files from device

By using Send to Kindle, you agree to the terms [here](#).

Supported File Types: PDF, DOC, DOCX, TXT, RTF, HTML, HTML, PNG, GIF, JPG, JPEG, BMP, EPUB

Max File Size: 200 MB

Other ways to send

- Web
- iOS
- Android
- Microsoft Word
- Email
- Chrome Extension

StoryGraph

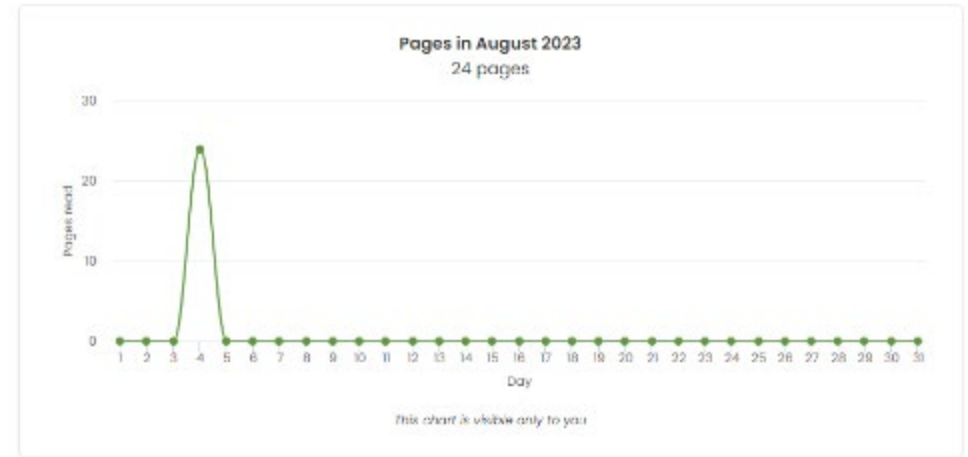


The StoryGraph

Because life's too short for a book you're not in the mood for.

Sign up

Sign in



Home Stats Reading Challenges Community Giveaways Search all books...

Current reads

2 books Reading Journal

A Course of Pure Mathematics
Cambridge Mathematical Library
G.H. Hardy
522 pages • first pub. 1953 (editions)

nonfiction mathematics reference science
informative medium-paced

3%
Started Aug 3, 2023
currently reading
mark as "finished"
owned buy

Eichmann in Jerusalem: A Report on the Banality of Evil
Hannah Arendt
312 pages • first pub. 1963 (editions)

nonfiction history philosophy politics
challenging informative reflective
slow-paced

2%
Started Aug 4, 2023
currently reading
mark as "finished"
owned buy

ER and Language Data Activities

2. Utilize the book's text data
 - 1) Obtain an Excel file with the book's text data.

Sample: [JaneEyre_Results_Text.xlsx](#)

- 2) Handle the Excel file in a computer.

Or use Google Sheets.

- 3) Utilize Excel's features:
Searching, sorting, tabulation, etc.

	A	B	C	D
1	no	chapter	text	word_count
2	1	0	JANE EYRE	2
3	2	0	AN AUTOBIOGRAPHY	2
4	3	0	by Charlotte Brontë	3
5	4	0	ILLUSTRATED BY F. H. TOWNSEND	5
6	6	0	London	1
7	7	0	SERVICE & PATON	3
8	8	0	5 HENRIETTA STREET	3
9	9	0	1897	1
10	11	0	The Illustrations	2
11	12	0	in this Volume are the copyright of	7
12	13	0	SERVICE & PATON, London	4
13	15	0	TO	1
14	16	0	W. M. THACKERAY, ESQ.,	4
15	18	0	This Work	2
16	19	0	IS RESPECTFULLY INSCRIBED	3
17	21	0	BY	1
18	22	0	THE AUTHOR	2
19	24	0	PREFACE	1
20	25	0	A preface to the first edition of "Jane Eyre" being unnecessary, I gave none: this second edition demands a few words both of acknowledgment and miscellaneous remark.	27

Excel Search

	A	B	C	D
1	no	chapter	text	word_count
818			waiter who answered the summons.	
819	1601	11	"Thornfield? I don't know, ma'am; I'll inquire at the bar." He vanished, but reappeared instantly—	15
820	1603	11	"Is your name Eyre, Miss?"	5
821	1605	11	"Yes."	1
822	1607	11	"Person here waiting for you."	5
823	1609	11	I jumped up, took my muff and u passage: a man was standing by t dimly saw a one-horse conveyanc	
824	1611	11	"This will be your luggage, I suppo he saw me, pointing to my trunk	
825	1613	11	"Yes." He hoisted it on to the veh got in; before he shut me up, I as	
826	1615	11	"A matter of six miles."	
827	1617	11	"How long shall we be before we	
828	1619	11	"Happen an hour and a half."	
			He fastened the car door, climbed to his own seat outside, and we set off	

찾기 및 바꾸기

찾기(D) 바꾸기(P)

찾을 내용(N): Eyre 설정된 서식 없음 서식(M)...

범위(H): 시트 대/소문자 구분(C)

검색(S): 행 전체 셀 내용 일치(O)

찾는 위치(L): 수식 전자/반자 구분(B) 옵션(O) <<

모두 찾기(F) 다음 찾기(N) 닫기

j_text | j_text_words | j_text_df_search_Rochester | j_text_df_search_Jane_Bessie

Excel Sort

	A	B	C
1	word	n	
2	the	7856	
3	i	7169	
4	and	6632	
5	to	5238	
6	a	4470	
7	of	4368	
8	you	2970	
9	in	2769	
10	was	2526	
11	it	2406	
12	my	2236	
13	me	2056	
14	he	1901	
15	her	1716	
16	that	1659	
17	as	1593	

Navigation: j_text | j_text_words | j_text_df_search_Roch

	A	B	C
1	word	n	length
2	characteristically	1	18
3	conventionalities	2	=LEN(A3)
4	conscientiousness	1	17
5	disadvantageously	1	17
6	inconsiderateness	1	17
7	misunderstandings	1	17
8	incomprehensible	3	16
9	disproportionate	2	16
10	responsibilities	2	16
11	indiscriminately	1	16
12	misunderstanding	1	16
13	personifications	1	16
14	superciliousness	1	16
15	superintendent's	1	16
16	unimpressible	1	16
17	unsubstantiality	1	16

Navigation: j_text | j_text_words | j_text_df_search_Roch

Excel Making Tables

The screenshot displays an Excel spreadsheet with a PivotTable. The PivotTable is located in columns B and C, with 'Chapters' as the row label and 'Texts' as the column label. The data is as follows:

Chapters	Texts
11	14
12	4
13	20
14	9
15	8
16	23
17	29
18	21
19	16
20	24
21	7
22	12
23	13
24	24
25	8
26	27
27	15
28	3
31	1
32	1

The PivotTable Fields task pane is open on the right side of the screen. It is titled '피벗 테이블 필드' (PivotTable Fields). The fields 'chapter' and 'text' are selected and listed in the task pane. Below the field list, there are sections for '필터' (Filter), '열' (Columns), '행' (Rows), and '값' (Values). The '행' (Rows) section is set to 'chapter' and the '값' (Values) section is set to 'Texts'. At the bottom of the task pane, there is a checkbox for '나중에 레이아웃 업데이트...' (Update layout later...) and an '업데이트' (Update) button.

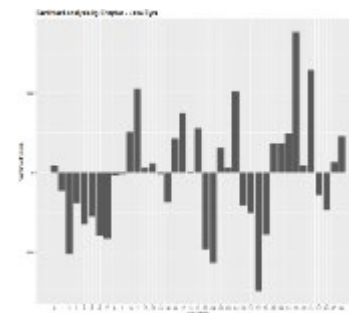
Excel

Copy Text

	A	B	C
1	no	chapter	text
6			the impression made on me by her nursery tales. She was pretty too, if my recollections of her face and person are correct. I remember her as a slim young woman, with black hair, dark eyes, very nice features, and good, clear complexion; but she had a capricious and hasty temper, and indifferent ideas of principle or justice: still, such as she was, I preferred her to any one else at Gateshead Hall.
3982		21	"Missis had been out of health herself for some time: she had got very stout, but was not strong with it; and the loss of money and fear of poverty were quite breaking her down. The information about Mr. John's death and the manner of it came too suddenly: it brought on a stroke. She was three days without speaking; but last Tuesday she seemed rather better: she appeared as if she wanted to say something, and kept making signs to my wife and mumbling. It was only yesterday morning, however, that Bessie understood she was pronouncing your name; and at last she made out the words, 'Bring Jane—fetch Jane Eyre: I want to speak to her.' Bessie is not sure whether she is in her right mind, or means anything by the words; but she told Miss Reed and Miss Georgiana, and advised them to send for you. The young ladies put it off at first; but their mother grew so restless, and said, 'Jane, Jane,' so many times, that at last they consented. I left

ER and Language Data Activities

3. Mining the text (to be applied in additional/advanced stages)
 - 1) Obtain the text data of the book as a text file. An excel file will do. Sample: [JaneEyre.txt](#)
 - 2) Prepare a coding environment.
For R, use Posit Cloud (<https://posit.cloud/>),
For Python, use Google Colab (<https://colab.google/>).
 - 3) Obtain text mining code written in R or Python.
Sample: <https://posit.cloud/content/6256020>
 - 4) Mine text of the book with codes provided.
 - 5) Utilize the text mining results (excel file, visuals, etc.) for your reading activity.
[JaneEyre_Results_Text.xlsx](#), [JaneEyre_Results_Visual.zip](#)



Posit Cloud

 **posit** Cloud

[Log In](#) [Sign Up](#)

Friction free data science

Posit Cloud (formerly RStudio Cloud) lets you access Posit's powerful set of data science tools right in your browser – no installation or complex configuration required.

[GET STARTED](#)

[ALREADY A USER? LOG IN](#)

If you already have a shinyapps.io account, you can log in using your existing credentials.



Posit Cloud Project

The screenshot displays the Posit Cloud web interface. The browser address bar shows the URL `https://positcloud/content/6256020`. The page title is "Your Workspace / LanguageDataAnalysis_ER" by user "Youngsoo Kim".

The interface is divided into several sections:

- Left Sidebar:** Contains navigation options such as "Your Workspace", "SharedSpace", "New Space", "Learn" (with sub-items like Guide, What's New, Printers, Cheat Sheets), "Help" (with sub-items like Current System Status, Post Community), and "Info" (with sub-items like Plans & Pricing, Terms and Conditions).
- Top Bar:** Includes a "Save a Permanent Copy" button, which is highlighted by a hand cursor. Other icons include a temporary copy warning, a user profile, and a refresh button.
- Console:** Shows the R version 4.3.1 (2023-06-16) with the "Beagle Scouts" logo. It displays the R license text: "R is free software and comes with ABSOLUTELY NO WARRANTY. You are welcome to redistribute it under certain conditions. Type 'license()' or 'licence()' for distribution details. R is a collaborative project with many contributors. Type 'contributors()' for more information and 'citation()' on how to cite R or R packages in publications. Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help. Type 'q()' to quit R."
- Environment Panel:** Shows the current environment as "Global Environment" and indicates that the environment is empty.
- File Explorer:** Displays a file tree for the "Cloud - project" directory. The files listed are:

Name	Size	Modified
..	0 B	Jul 29, 2023, 10:14 AM
..history		
1_SameGoalAndObtainingTextData.R	2.4 KB	Aug 4, 2023, 11:03 AM
2_ExploringPreprocessingTextData.R	3 KB	Aug 4, 2023, 11:07 AM
3_AnalyzingTextDataAndUtilizingResults.R	6.9 KB	Aug 4, 2023, 11:03 AM
libraryworks		
outputs		
project.Rproj	205 B	Aug 4, 2023, 9:21 PM

Posit Cloud Running R Script

The screenshot displays the Posit Cloud interface for a workspace named "LanguageDataAnalysis_ER". The main editor shows an R script with the following content:

```
17
18 # The following literary work will be used in this project:
19 # Jane Eyre by Charlotte Brontë.
20
21 # It can be obtained from Project Gutenberg:
22 # https://www.gutenberg.org/ebooks/1260
23
24 # In the shared Posit Cloud project, the text file has already been downloaded and
25
26 # loading packages
27 # if prompted to install the packages, install them.
28 library(tidyverse)
29 library(openxlsx2)
30
31 # Setting a file folder for text files
32 input_folder <- "literaryworks/"
33
34 # Setting a file folder for output files
35 output_folder <- "outputs/"
36
37
38
```

The console output shows the R session starting with R 4.3.1. The user enters the commands from the script, and the console displays the following output:

```
R 4.3.1 . /cloud/project/ >
Type "q()" to quit R.

> # loading packages
> # if prompted to install the packages, install them.
> library(tidyverse)
— Attaching core tidyverse packages — tidyverse 2.0.0 —
✓ dplyr 1.1.2 ✓ readr 2.1.4
✓ forcats 1.0.0 ✓ stringr 1.5.0
✓ ggplot2 3.4.2 ✓ tibble 3.2.1
✓ lubridate 1.9.2 ✓ tidyr 1.3.0
✓ purrr 1.0.1
— Conflicts — tidyverse_conflicts() —
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag() masks stats::lag()
! Use the conflicted package to force all conflicts to become errors
>
```

The interface also shows a file explorer on the right with a table of files:

Name	Size	Modified
..		
.Rhistory	0 B	Jul 29, 2023, 10:14 AM
1_SettingGoalsAndObtainingTextData.R	2.4 KB	Aug 4, 2023, 11:03 AM
2_ExploringPreprocessingTextData.R	3 KB	Aug 4, 2023, 11:07 AM
3_AnalyzingTextDataAndUtilizingResults.R	6.9 KB	Aug 4, 2023, 11:03 AM
literaryworks		
outputs		
project.Rproj	205 B	Aug 4, 2023, 9:21 PM

Posit Cloud Word Cloud

The image shows a screenshot of the Posit Cloud interface. On the left, a browser window displays a word cloud with various words in different colors and orientations. The words include: hand, heart, door, left, miss, rochester, day, house, love, st, heard, night, sir, time, life, john, eyes, looked, and don't. On the right, the R console shows the following code and output:

```
> j_text_df_words_withoutstopwords_top_20_angle_wordcloud
Warning message:
In wordcloud_boxes(data_points = points_valid_first, boxes = boxes, :
Some words could not fit on page. They have been placed at their original positions.
> # Saving the word cloud as an image file
Warning messages:
1: In wordcloud_boxes(data_points = points_valid_first, boxes = boxes, :
Some words could not fit on page. They have been placed at their original positions.
2: In wordcloud_boxes(data_points = points_valid_first, boxes = boxes, :
Some words could not fit on page. They have been placed at their original positions.
> ggsave(paste0(output_folder, "j_top20_wordcloud.jpg"), j_text_df_words_withoutstopwords_top_20_angle_wordcloud)
Saving 4.91 x 3.05 in image
>
```

Below the console, there are links for "Info", "Plans & Pricing", and "Terms and Conditions".

Posit Cloud Bar Chart

The screenshot displays the Posit Cloud interface for a workspace named "LanguageDataAnalysis_ER". The main editor shows R code for sentiment analysis and visualization. The code includes steps for joining data, grouping by sentiment, and creating a faceted bar chart. The console shows the execution of these commands, and the environment pane lists the data objects used. The resulting bar chart, titled "Sentiment Analysis - Jane Eyre", shows the frequency of words associated with negative and positive sentiments.

```
73  
74 # Saving the bar graph of sentiment analysis results  
75 j_text_df_words_counts_sentiment_bargraph <- j_text_df_words_counts %>%  
76   inner_join(bing_data) %>%  
77   group_by(sentiment) %>%  
78   slice_max(n, n = 10) %>%  
79   ungroup() %>%  
80   mutate(word = reorder(word, n)) %>%  
81   ggplot(aes(word, n, fill = sentiment)) +  
82   geom_col(show.legend = FALSE) +  
83   facet_wrap(~sentiment, scales = "free_y") +  
84   labs(y = "Sentiment Analysis - Jane Eyre", x = NULL) +  
85   coord_flip()  
86  
87 ggsave(paste0(output_folder, "j_sentiment_bargraph.png"), j_text_df_words_counts_sen  
88  
89 # Displaying a bar graph of sentiment analysis results - by chapter  
90 j_text_df_words %>%  
91   inner_join(get_sentiments("bing")) %>%  
92   count(chapter, sentiment) %>%  
93
```

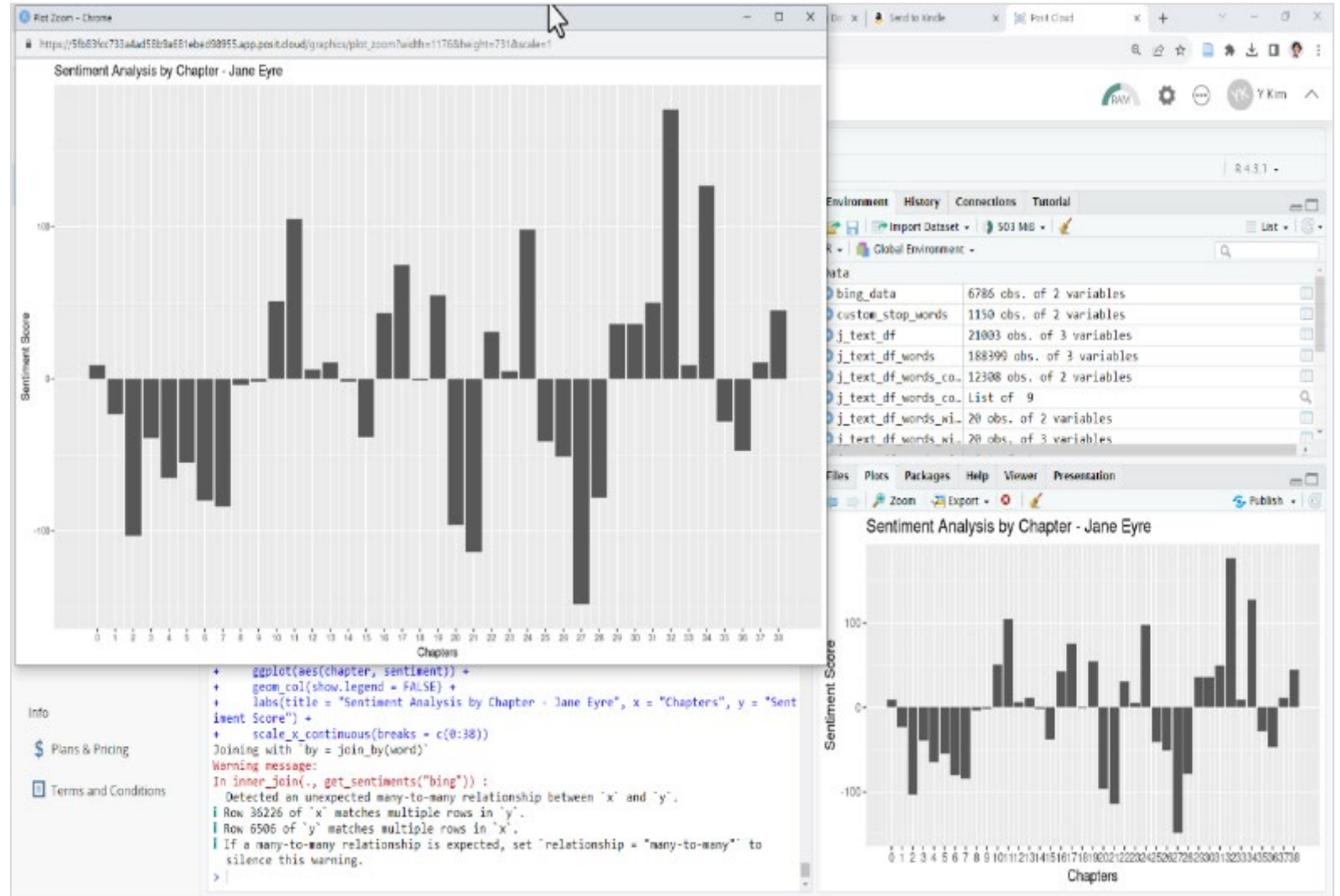
Environment pane data:

Object	Description
bing_data	6786 obs. of 2 variables
custom_stop_words	1150 obs. of 2 variables
j_text_df	21803 obs. of 3 variables
j_text_df_words	188399 obs. of 3 variables
j_text_df_words_co.	12308 obs. of 2 variables
j_text_df_words_co.	List of 9
j_text_df_words_wi.	20 obs. of 2 variables
j_text_df_words_wi.	20 obs. of 3 variables

Bar Chart Data (Estimated):

Sentiment	Word	Count
negative	strange	100
	dark	90
	cold	75
	poor	65
	fear	60
	doubt	55
	wild	55
	fall	50
	death	50
	poke	50
positive	love	150
	master	100
	pleasure	80
	quiet	60
	fine	55
	strong	55
	happy	50
	grace	50
	smile	50
	sweet	50

Posit Cloud Sentiment Analysis



Posit Cloud Word Network

The screenshot displays the Posit Cloud web interface for a workspace named "LanguageDataAnalysis_ER". The main area contains an R script with the following code:

```
146 library(igraph)
147 library(ggraph)
148
149 j_text_df_bigrams_filtered_counts_graph <- j_text_df_bigrams_filtered_counts %>%
150   filter(n > 5) %>%
151   graph_from_data_frame()
152
153 set.seed(2017)
154 a <- grid::arrow(type = "closed", length = unit(.15, "inches"))
155
156 ggraph(j_text_df_bigrams_filtered_counts_graph, layout = "fr") +
157   geom_edge_link(aes(edge_alpha = n), show.legend = FALSE, arrow = a, end_cap = c("circle", "circle"), size = 1) +
158   geom_node_point(color = "lightblue", size = 5) +
159   geom_node_text(aes(label = name), vjust = -0.5, hjust = 0.5) +
160   theme_void()
161
162 ggsave(paste0(output_folder, "j_network_bigrams.jpg"), height=4.5, width=8.5)
163
164 ## Looking for details in the novel
165
166
```

The console output shows the execution of the script, including the creation of the graph and the saving of the plot as "j_network_bigrams.jpg". A warning message is also visible: "Warning message: Using the 'size' aesthetic in this geom was deprecated in ggplot2 3.4.0. Please use 'linewidth' in the 'default_aes' field and elsewhere instead. This warning is displayed once every 8 hours. Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was generated."

On the right side of the interface, the "Environment" panel shows a list of data objects:

Object	Description
a	List of 1
bing_data	6786 obs. of 2 variables
custom_stop_words	1150 obs. of 2 variables
j_text_df	21003 obs. of 3 variables
j_text_df_bigrams	171698 obs. of 3 variables
j_text_df_bigrams_	85645 obs. of 2 variables
j_text_df_bigrams_	14396 obs. of 4 variables
j_text_df_bigrams_	12916 obs. of 3 variables

Below the environment panel, a word network plot is displayed. The nodes represent words, and the edges represent their relationships. The plot shows a complex network of words, with some words appearing more central than others. The words are colored light blue and are connected by lines of varying thickness, representing the frequency of co-occurrence.

Posit Cloud Outcomes

The screenshot displays the Posit Cloud interface for a workspace named "LanguageDataAnalysis_ER". The main editor shows R code for text analysis, including filtering for the word "Rochester" and saving the results to an Excel file. The console shows the output of the R code, displaying a list of text segments containing the word "Rochester".

```
176  
177 # Searching where Rochester appears  
178 j_text_df %>%  
179   filter(str_detect(text, "Rochester"))  
180  
181 # Saving an Excel the parts where Rochester appears  
182 j_text_df %>%  
183   filter(str_detect(text, "Rochester")) %>%  
184   write_xlsx(paste0(output_folder, "j_text_df_search_Rochester.xlsx"), "Rochester")  
185  
186  
187 ## After Reading Activity  
188  
189 # Using Language-related AI Technologies such as ChatGPT  
190  
191 # Additional/Advanced Activities:  
192 # Project Title: Language Data Analysis with Literacy Texts  
193 # Additional/Advanced Activities:  
194  
195 # These activities are aimed at enhancing the understanding of  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000
```

The file explorer on the right shows the output folder containing several files, including Excel spreadsheets and image files. A mouse cursor is pointing to the "j_text_df_search_Rochester.xlsx" file.

이름	수업일 날짜	유형	크기
bing_data.xlsx	2023-08-04 오후 9:24	Microsoft Excel ...	92KB
network_bigrams.jpg	2023-08-04 오후 9:26	JPG 그림	410KB
sentiment_bargraph.png	2023-08-04 오후 9:25	PNG 그림	70KB
sentiment_bargraph_chapter.png	2023-08-04 오후 9:25	PNG 그림	56KB
text.xlsx	2023-08-04 오후 9:22	Microsoft Excel ...	827KB
text_df_search_jane_besse.xlsx	2023-08-04 오후 9:26	Microsoft Excel ...	0KB
text_df_search_rochester.xlsx	2023-08-04 오후 9:26	Microsoft Excel ...	22KB
text_words.xlsx	2023-08-04 오후 9:23	Microsoft Excel ...	171KB
top20_bargraph.jpg	2023-08-04 오후 9:24	JPG 그림	254KB
top20_wordcloud.jpg	2023-08-04 오후 9:24	JPG 그림	271KB
stopwords.xlsx	2023-08-04 오후 9:23	Microsoft Excel ...	20KB



THE SIXTH WORLD CONGRESS ON EXTENSIVE READING

7-11 August 2023

Bridging the LitERacy Gap

Combining ER and DR(Distant Reading) Using Literary Texts with Computer Coding

Contact: youngwoolearning@gmail.com

Youngwoo KIM

Yongkook WON

Junghye Fran CHOI

Hyunkee AHN

Kwanghyun PARK

Joohae KIM

Terima **Kasih**

Matur Suksma