



# Final Year Project

# Markson's Memory

Conor Kelleher  
13395381

Supervisor: Mark Keane

# Markson's Memory

© 2014 Markson's Memory

# Markson's Memory

- Digital Humanities

# Markson's Memory

- Digital Humanities
- Text Analytics

# Markson's Memory

- Digital Humanities
- Text Analytics
- Natural Language Toolkit (NLTK)

# *Wittgenstein's Mistress*

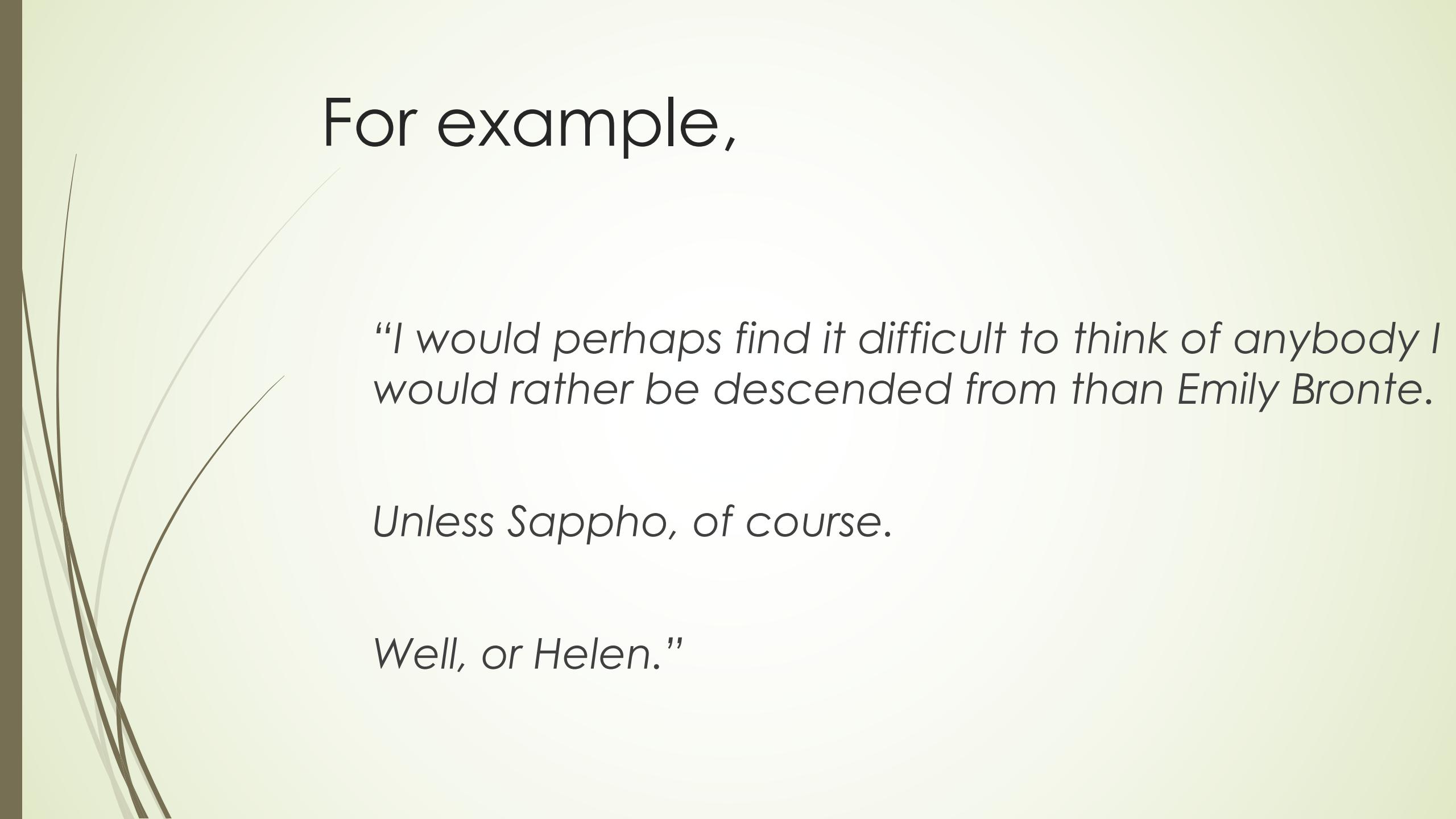
-David Markson

# Wittgenstein's Mistress

In the beginning, sometimes I left messages in the street.

David Markson

"As precise and dazzling as Joyce . . .  
Original, beautiful, and an absolute masterpiece."  
—Ann Beattie  
Copyrighted Material



# For example,

*“I would perhaps find it difficult to think of anybody I would rather be descended from than Emily Bronte.*

*Unless Sappho, of course.*

*Well, or Helen.”*

# Wittgenstein's Mistress

## Close Reading

In the beginning, sometimes I left messages in the street.

David Markson

"As precise and dazzling as Joyce . . .

Original, beautiful, and an absolute masterpiece."

—Ann Beattie  
Copyrighted Material

# Wittgenstein's Mistress

In the beginning,  
sometimes I left  
messages in the  
street.

David Markson

Close Reading

vs.

Distant Reading

"As precise and dazzling as Joyce . . .  
Original, beautiful, and an absolute masterpiece."  
—Ann Beattie  
Copyrighted Material

# Approach

- 
- 
- 
- 

## Wittgenstein's Mistress

In the beginning, sometimes I left messages in the street.

David Markson

"As precise and dazzling as Joyce, . . .  
Original, beautiful, and an absolute masterpiece."  
—Ann Beattie  
Copyrighted Material

# Approach

- Extract Repeated Phrases
- 
- 
- 

## Wittgenstein's Mistress

In the beginning, sometimes I left messages in the street.

David Markson

"As precise and dazzling as Joyce, . . .  
Original, beautiful, and an absolute masterpiece."  
—Ann Beattie  
Copyrighted Material

# Approach

- Extract Repeated Phrases
- Find Named Entities
- 
- 

# Wittgenstein's Mistress

In the beginning, sometimes I left messages in the street.

David Markson

"As precise and dazzling as Joyce, . . .  
Original, beautiful, and an absolute masterpiece."  
—Ann Beattie  
Copyrighted Material

# Approach

- Extract Repeated Phrases
- Find Named Entities
- Construct Networks
- 

# Wittgenstein's Mistress

In the beginning, sometimes I left messages in the street.

David Markson

"As precise and dazzling as Joyce, . . .  
Original, beautiful, and an absolute masterpiece."  
—Ann Beattie  
Copyrighted Material

# Approach

- Extract Repeated Phrases
- Find Named Entities
- Construct Networks
- Track Thought Patterns

# Wittgenstein's Mistress

In the beginning, sometimes I left messages in the street.

David Markson

"As precise and dazzling as Joyce . . .  
Original, beautiful, and an absolute masterpiece."  
—Ann Beattie  
Copyrighted Material

# Finding Phrase Repetition

# Phrase Extraction – N-Grams

*“Why have I written that his name was Adam?*

# Phrase Extraction – N-Grams

*“Why have I written that his name was Adam?”*

3-Grams: {“Why have I”, “have I written”, “I written that”, “that his name”,  
“his name was”, “name was Adam”}

# Phrase Extraction – N-Grams

*“Why have I written that his name was Adam?”*

3-Grams: {“Why have I”, “have I written”, “I written that”, “that his name”,  
“his name was”, “name was Adam”}

...

6-Grams: {“Why have I written that his”, “have I written that his name”, “I  
written that his name was”, “written that his name was Adam”}

# Phrase Extraction – N-Grams

*“Why have I written that his name was Adam?”*

3-Grams: {“Why have I”, “have I written”, “I written that”, “that his name”,  
“his name was”, “name was Adam”}

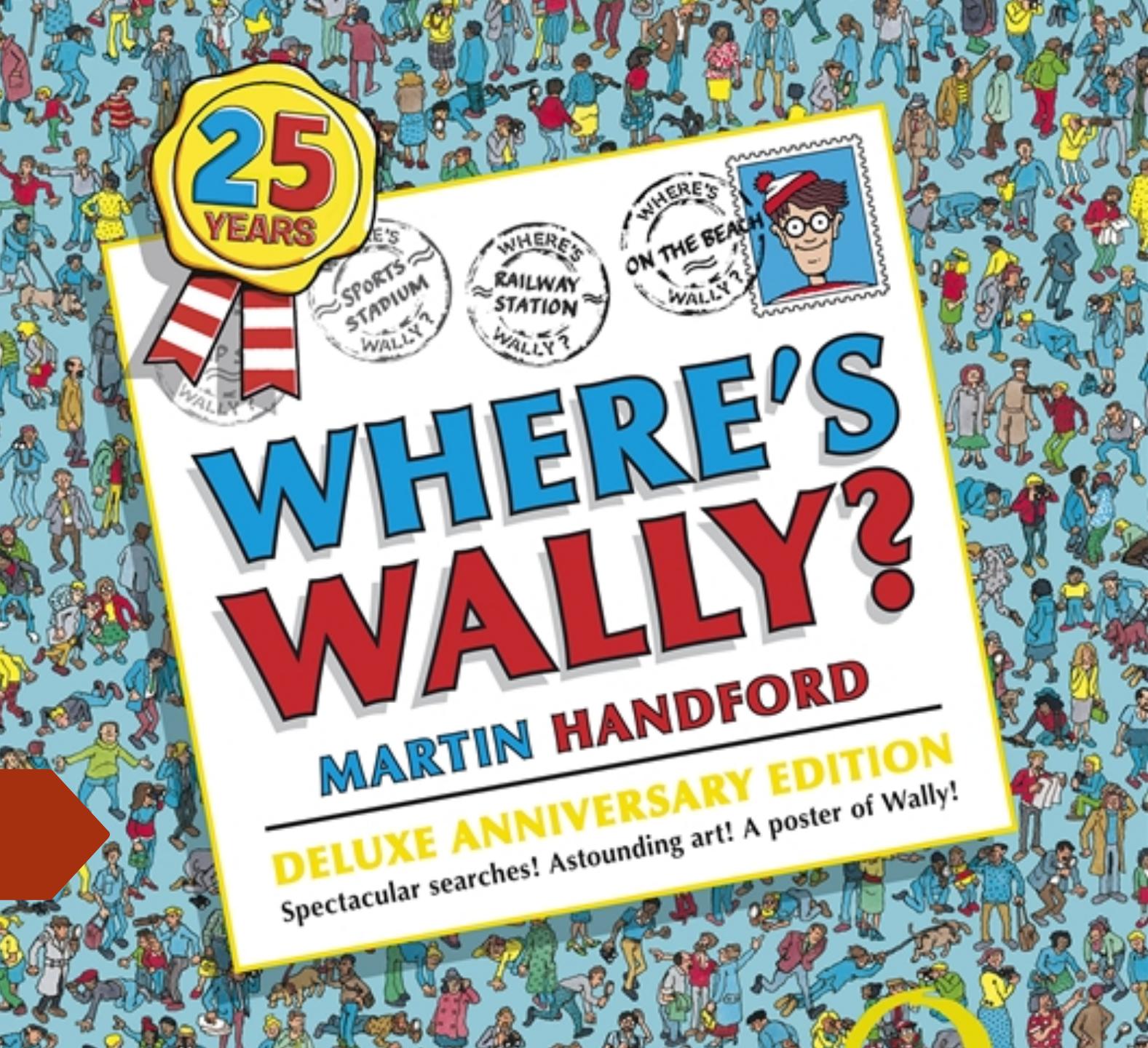
...

6-Grams: {“Why have I written that his”, “have I written that his name”, “I  
written that his name was”, “written that his name was Adam”}

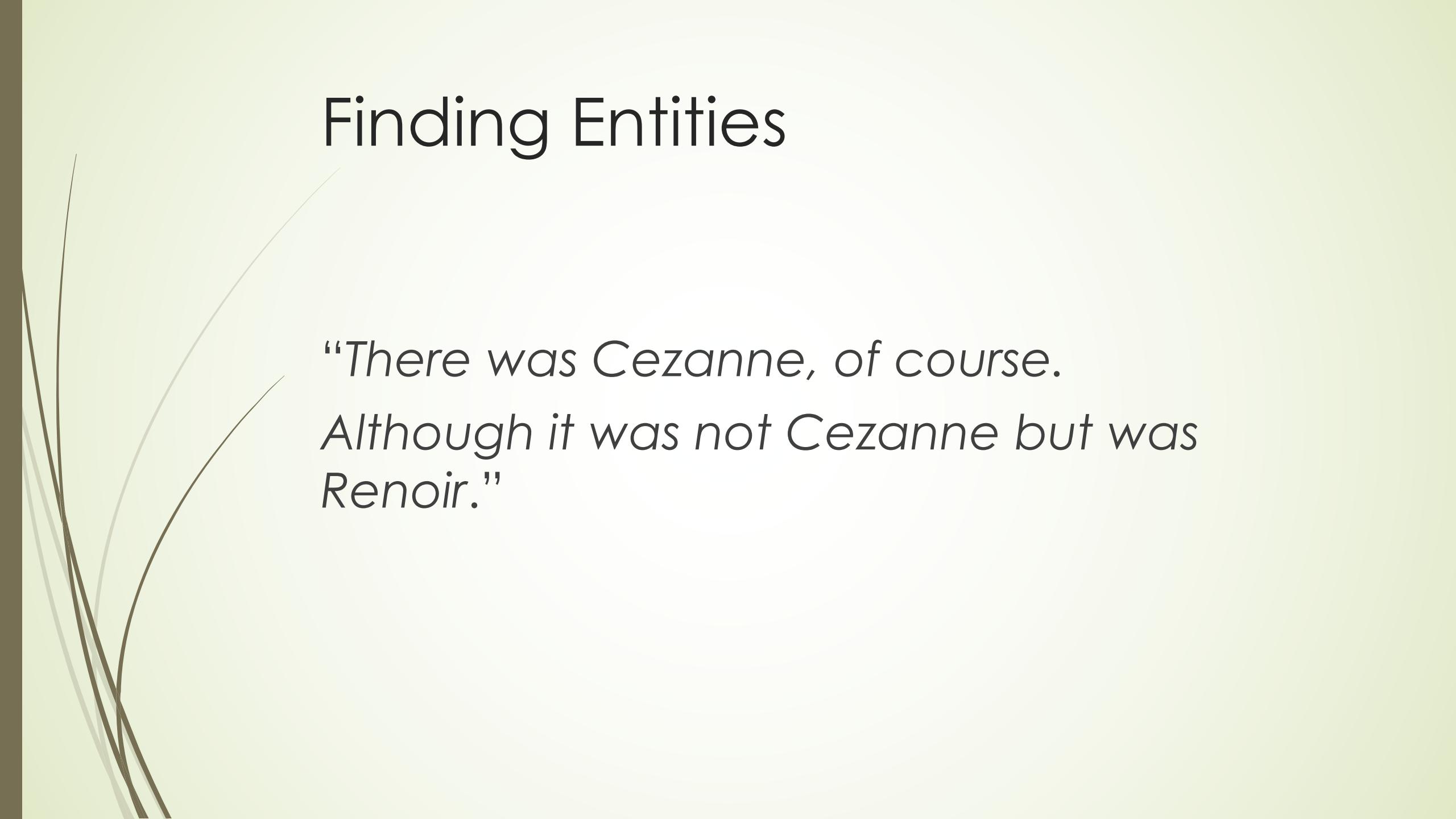
...

9-Grams: {“Why have I written that his name was Adam”}

# Named Entity Recognition

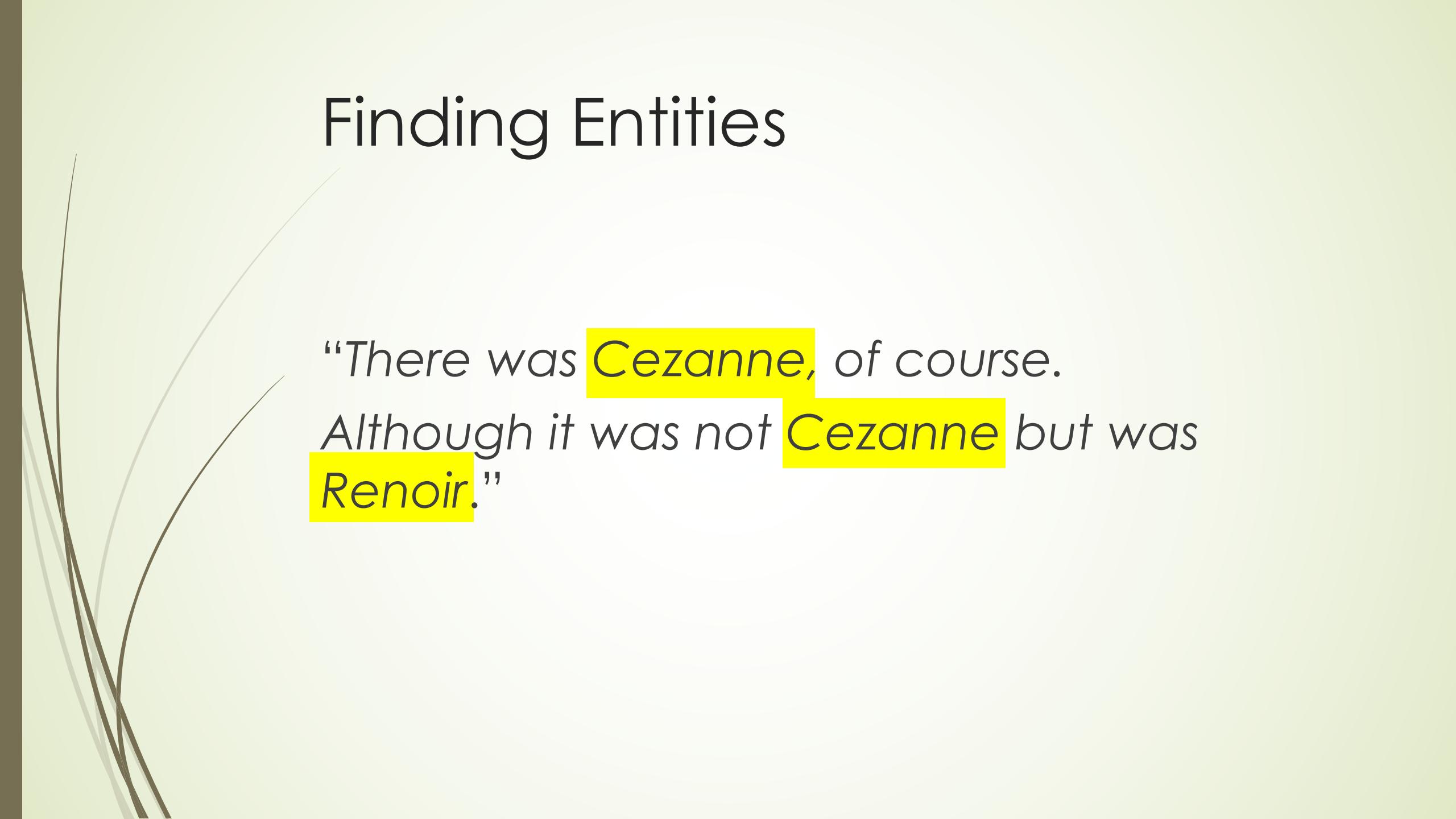


# Finding Entities



“There was Cezanne, of course.  
Although it was not Cezanne but was  
Renoir.”

# Finding Entities



“There was Cezanne, of course.  
Although it was not Cezanne but was  
Renoir.”

# Finding Entities

► What's in a name?

# Finding Entities

- ▶ What's in a name?
- ▶ Knowledge-based (Rules)

# Finding Entities

- ▶ What's in a name?
- ▶ Knowledge-based (Rules)
- ▶ Probability-based (Training Corpus)

# Finding Entities

- ▶ What's in a name?
- ▶ Knowledge-based (Rules)
- ▶ Probability-based (Training Corpus)
- ▶ Stanford NER (Named Entity Recognition)

# Network Analysis



# Same-paragraph mentions

# Same-paragraph mentions

Achilles

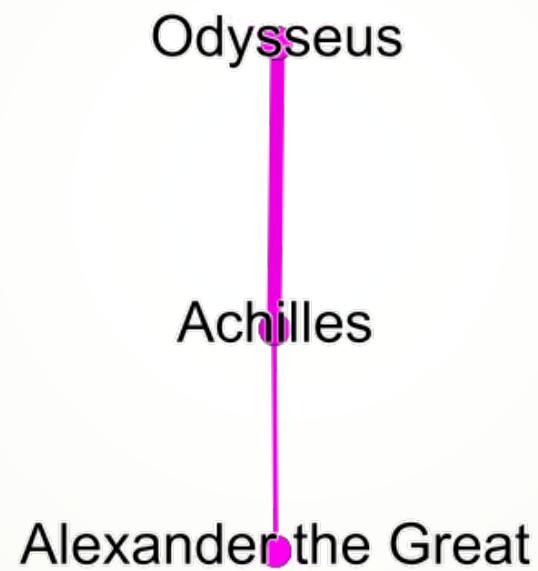
# Same-paragraph mentions

Odyssesus

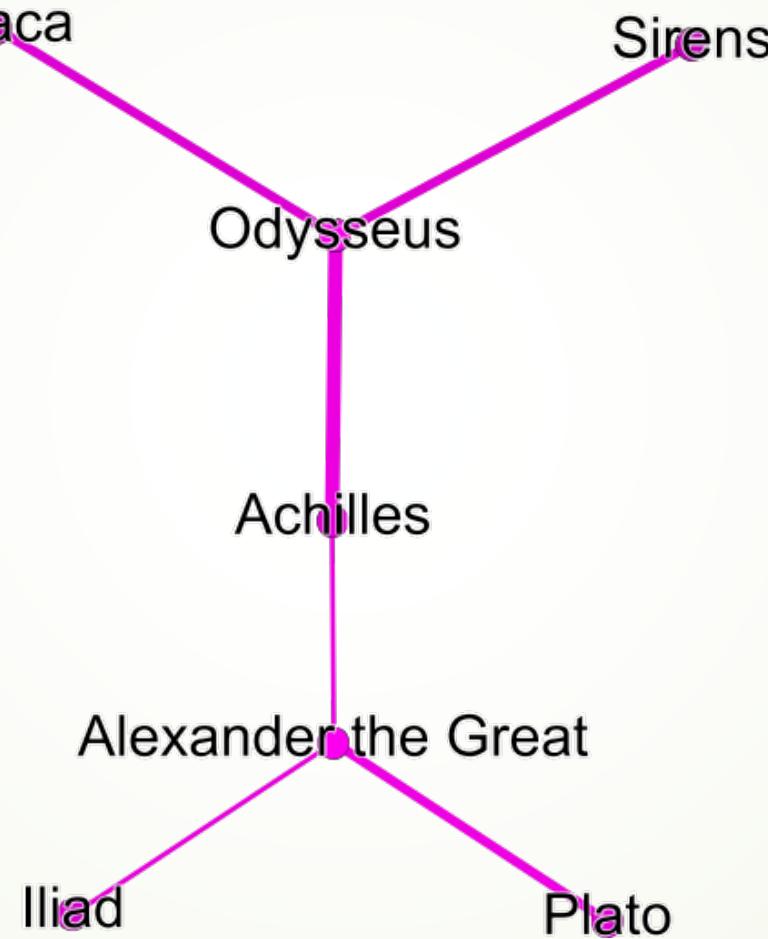
Achilles



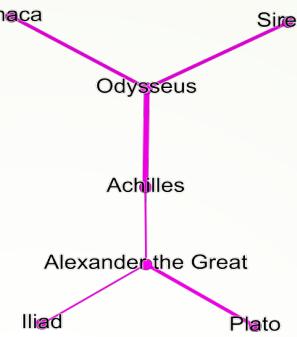
# Same-paragraph mentions



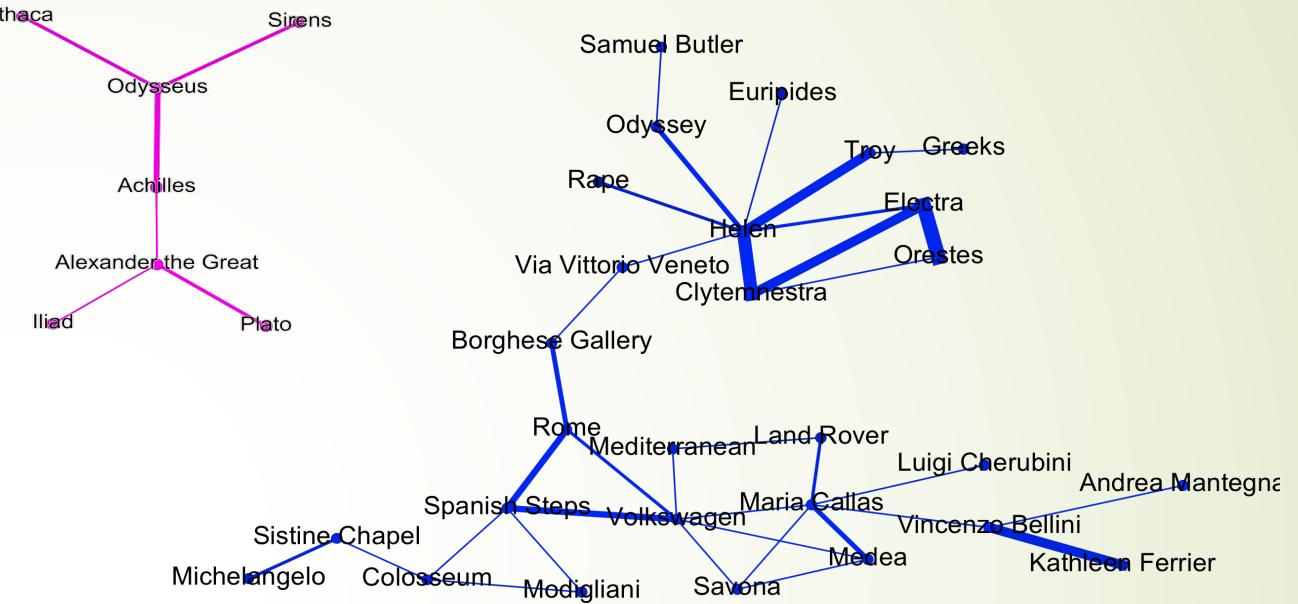
# Same-paragraph mentions



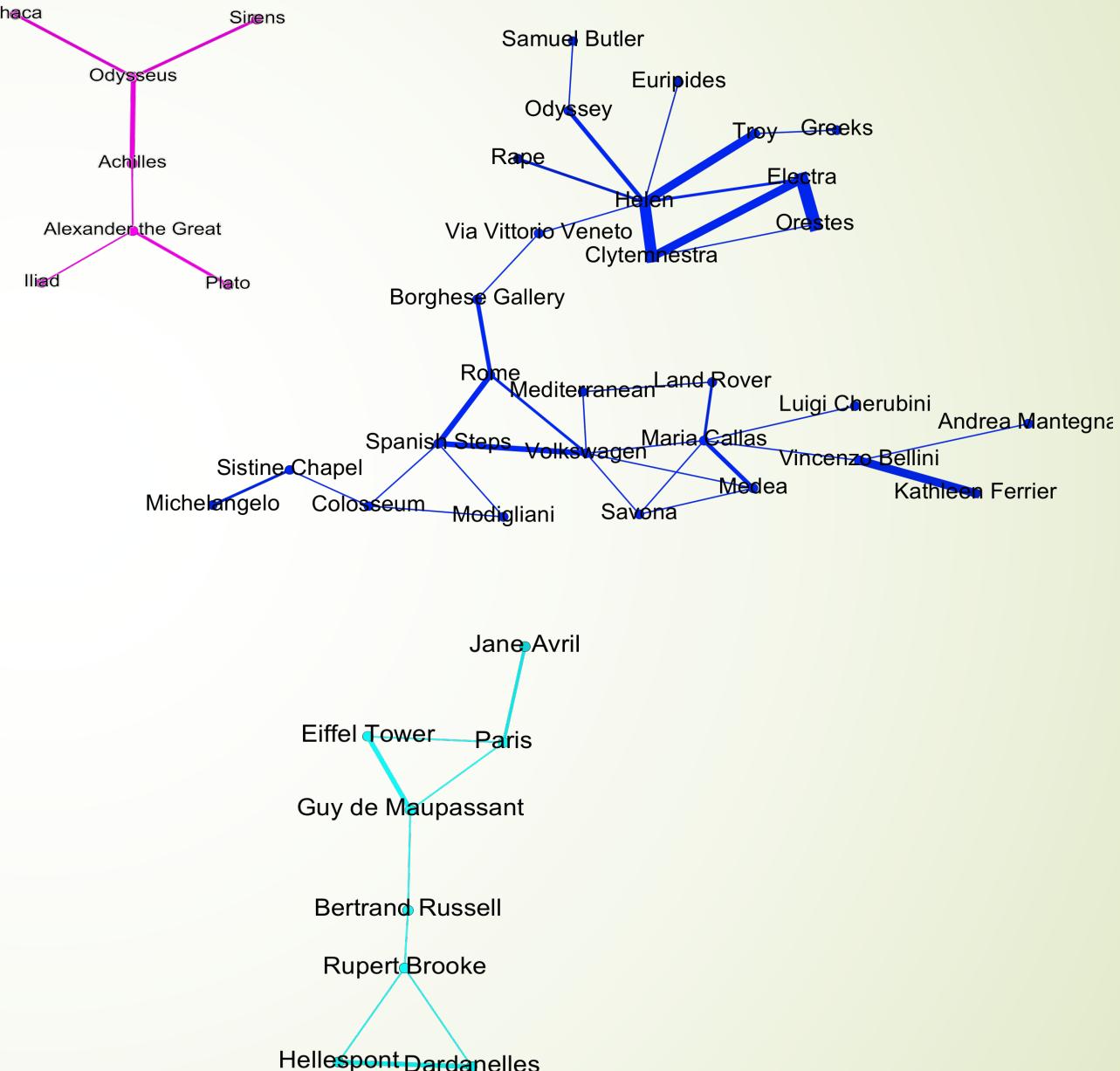
# Same-paragraph mentions



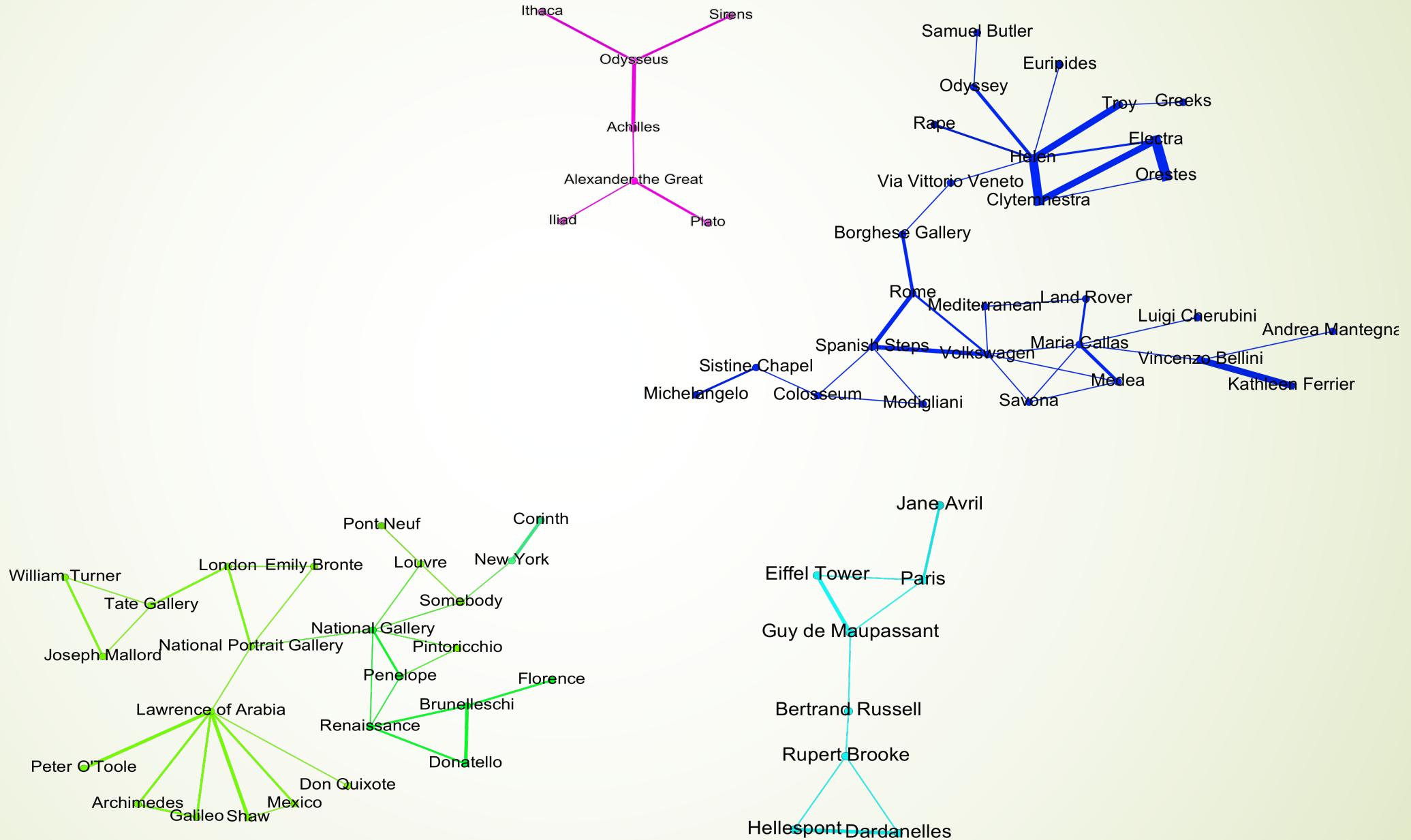
# Same-paragraph mentions



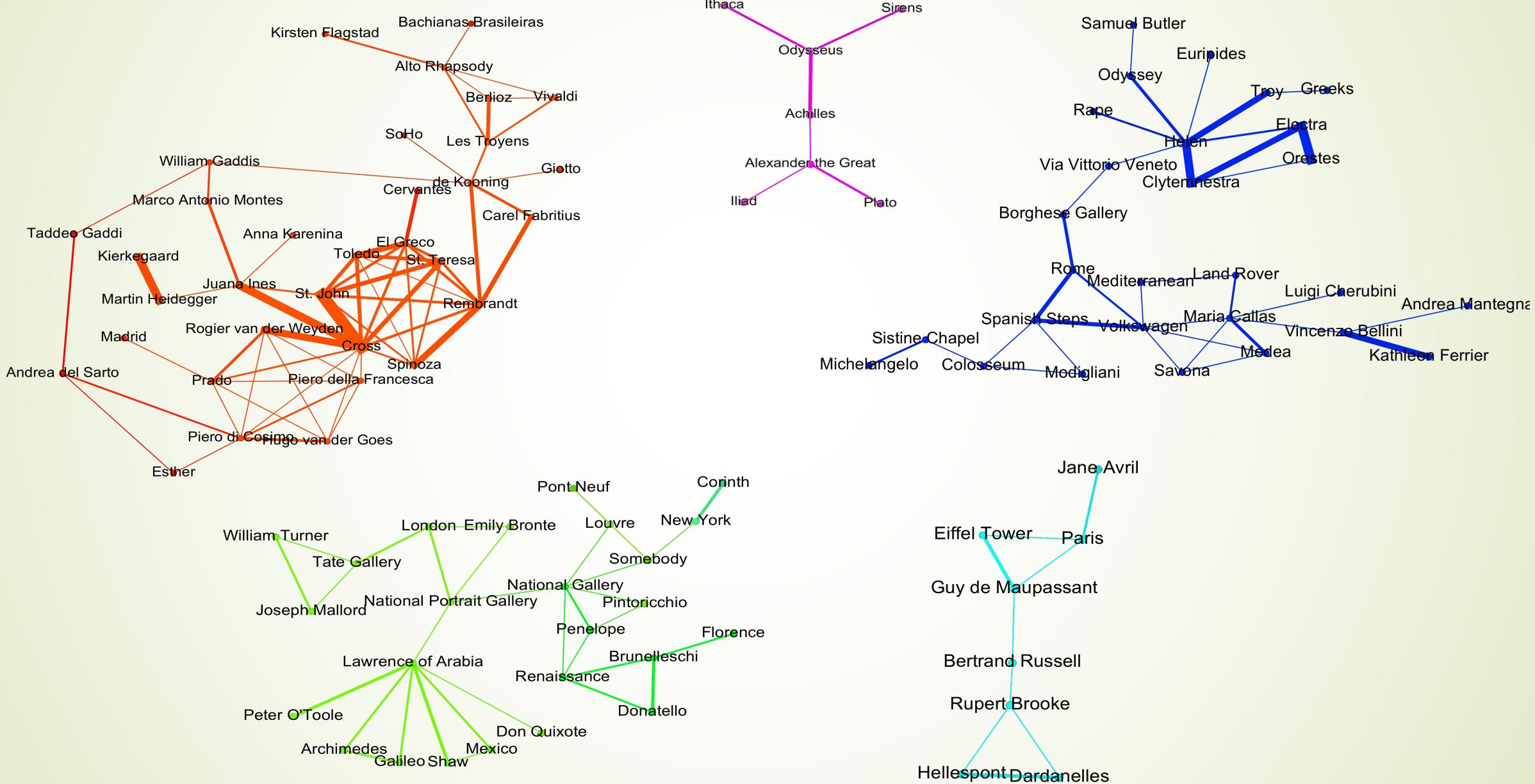
# Same-paragraph mentions



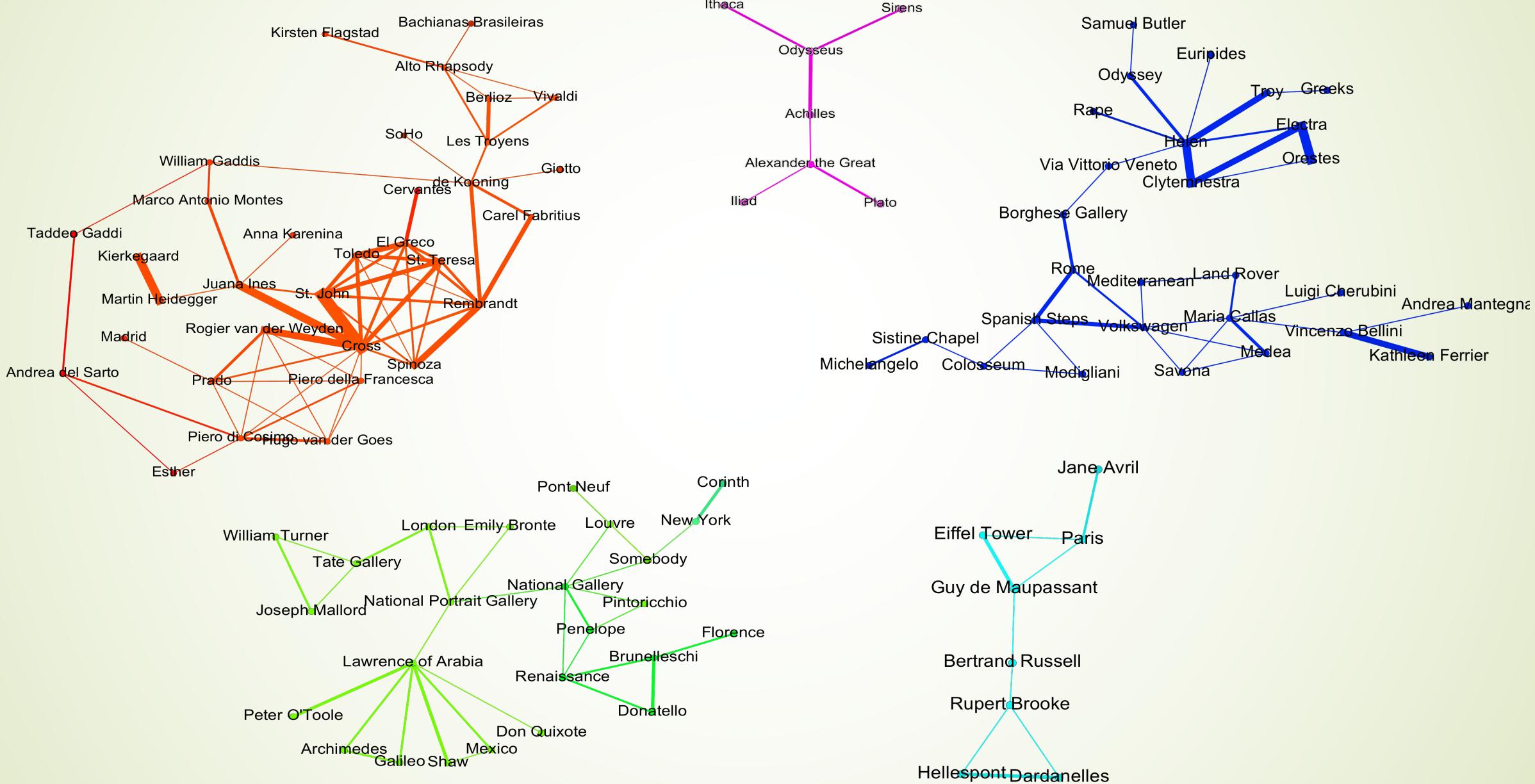
# Same-paragraph mentions



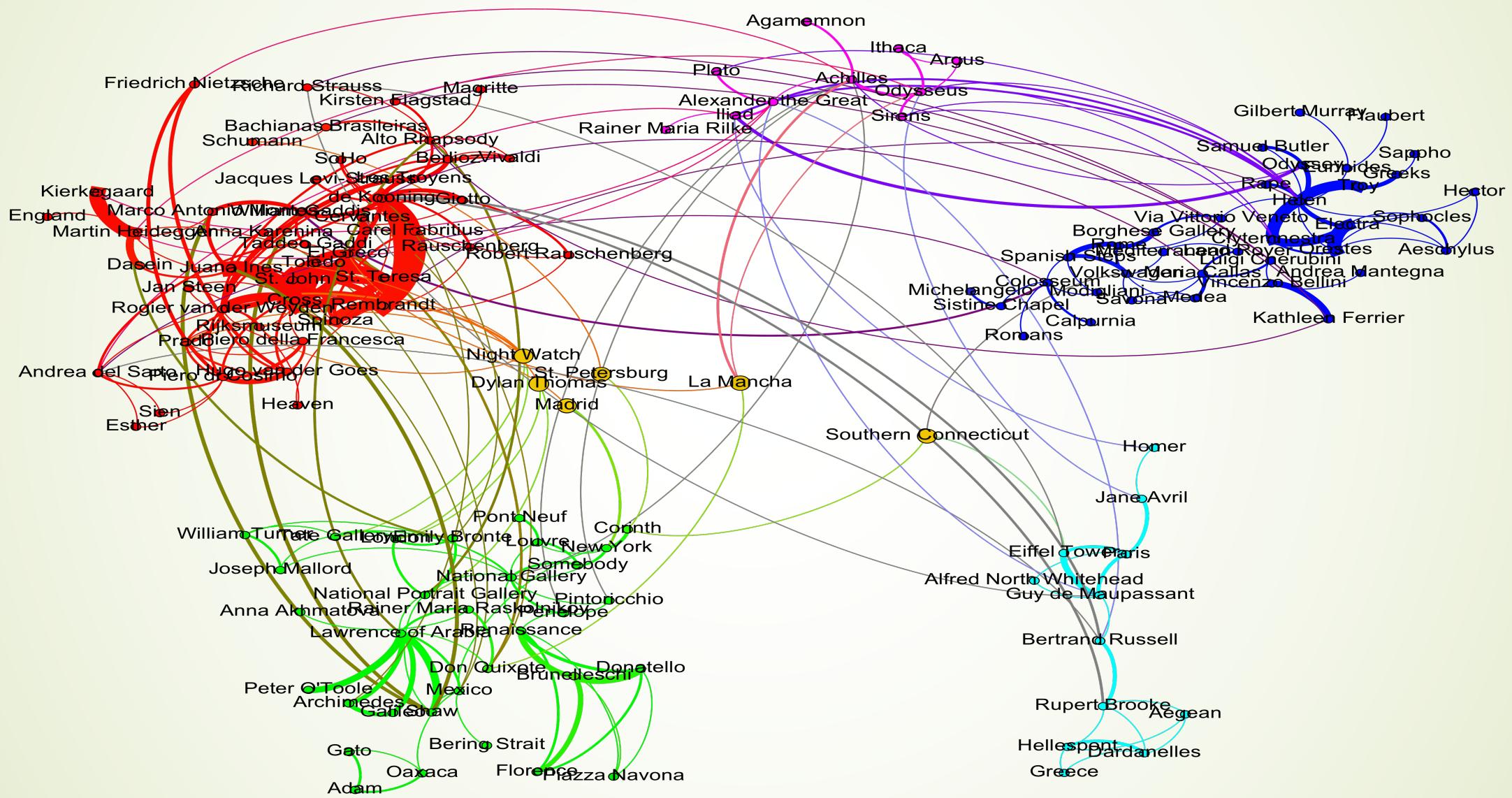
# Same-paragraph mentions



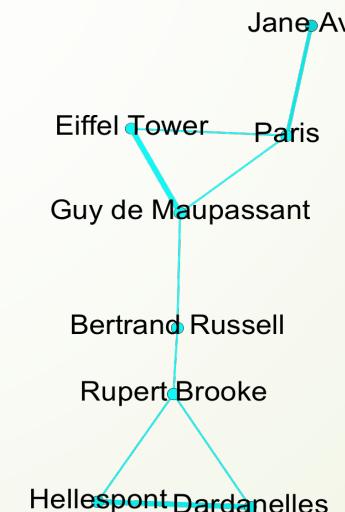
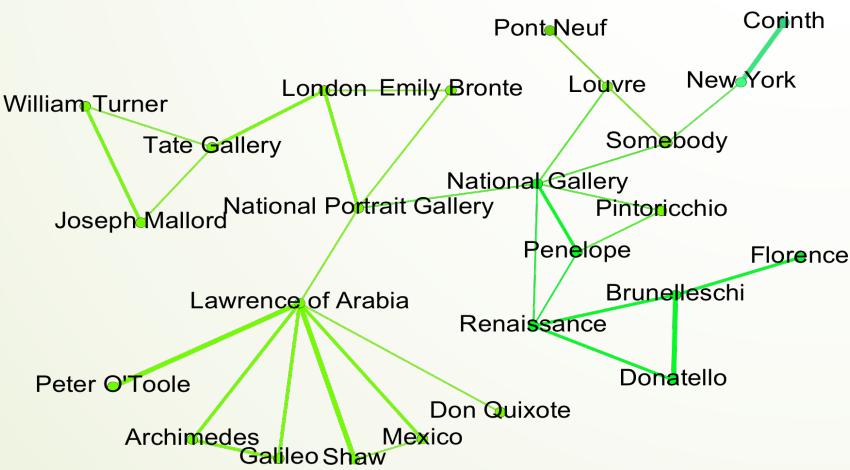
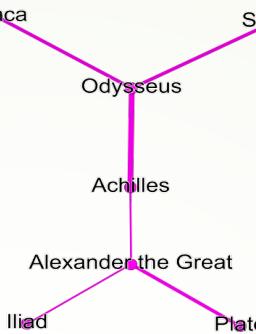
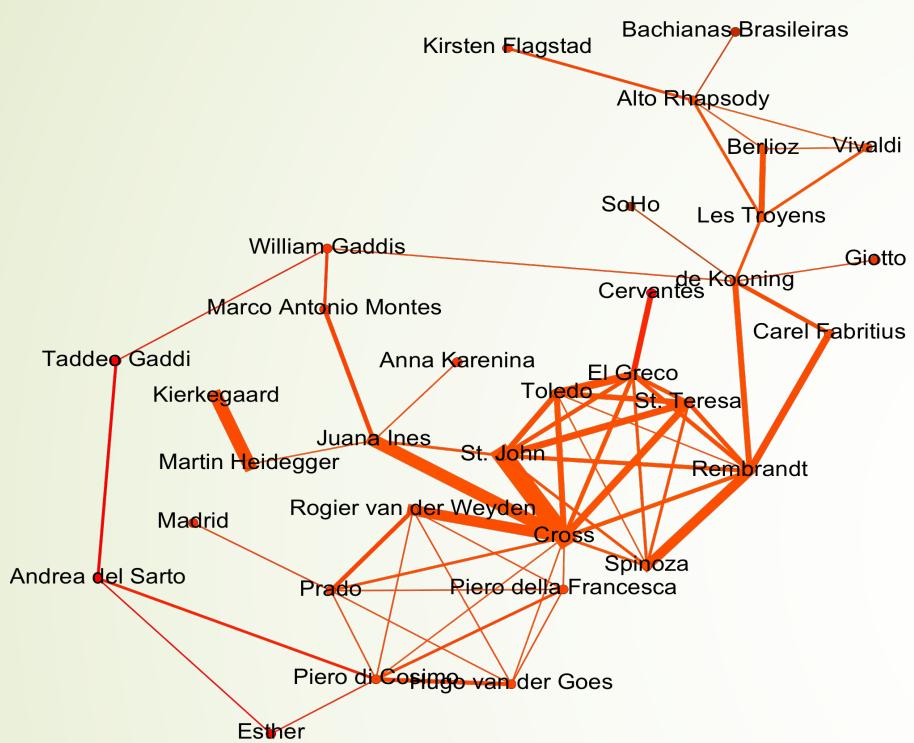
# One paragraph further...



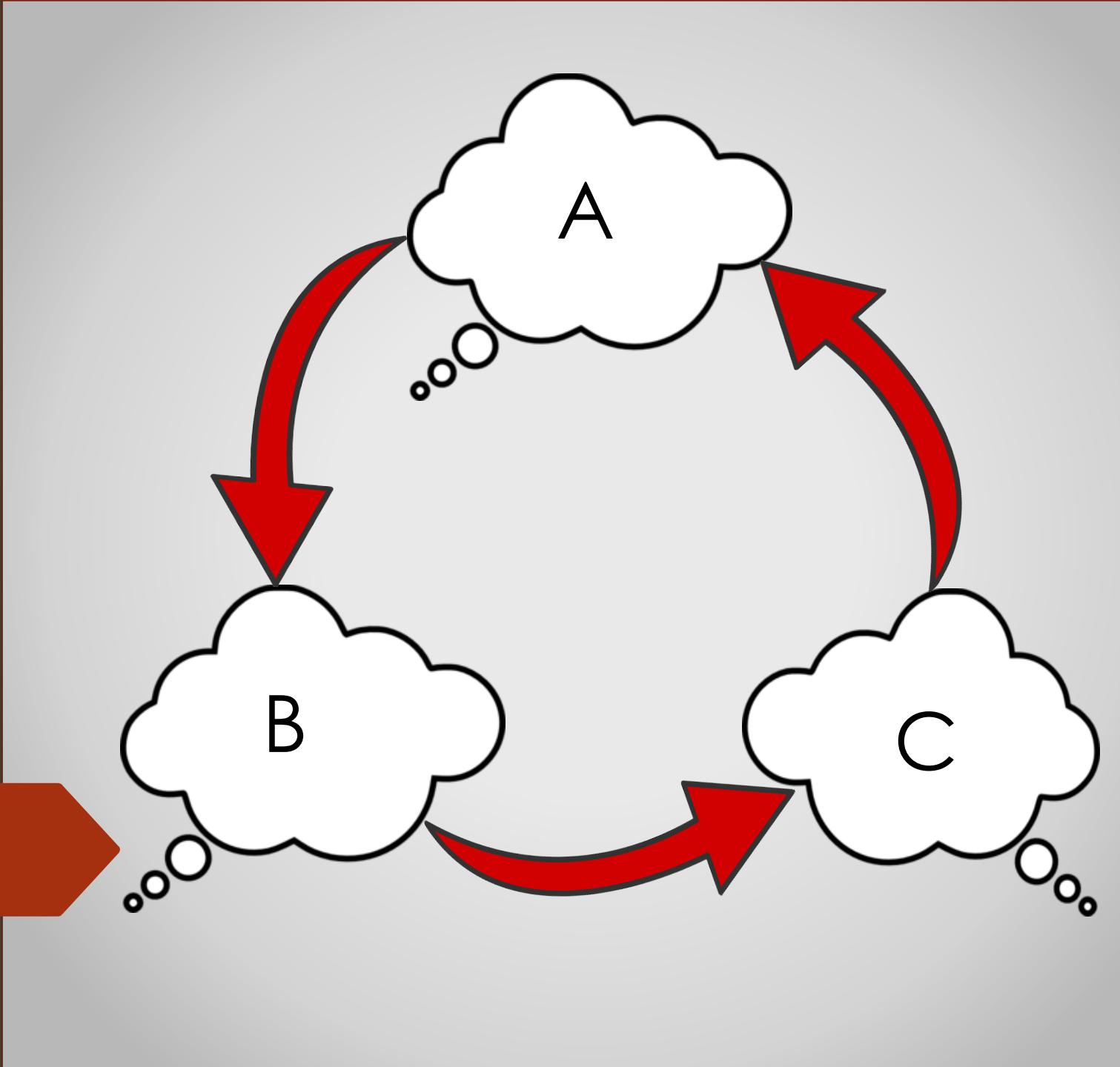
# One paragraph further...



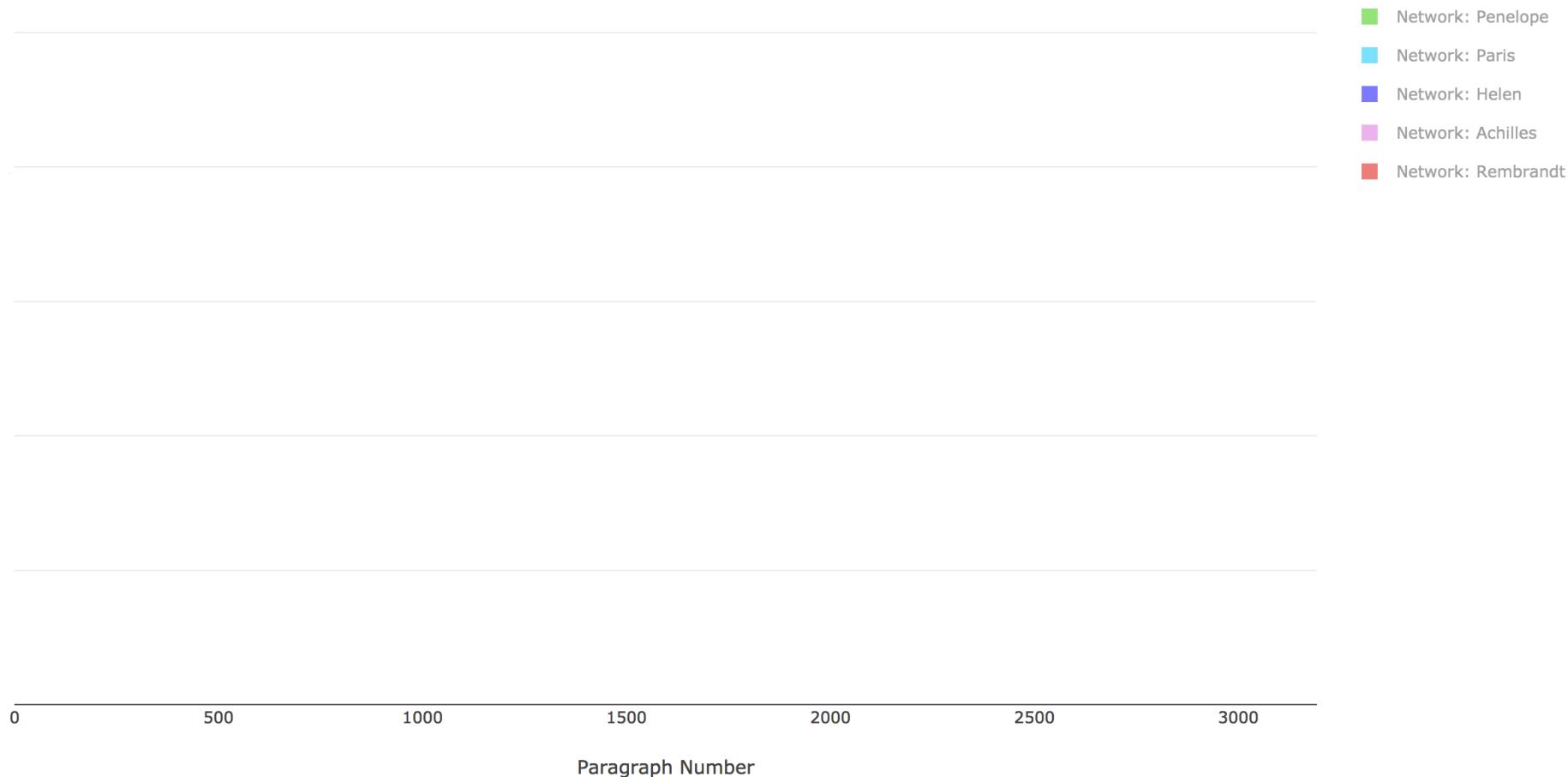
# Order of Relationships / Associative Structure ?



# Thought Patterns



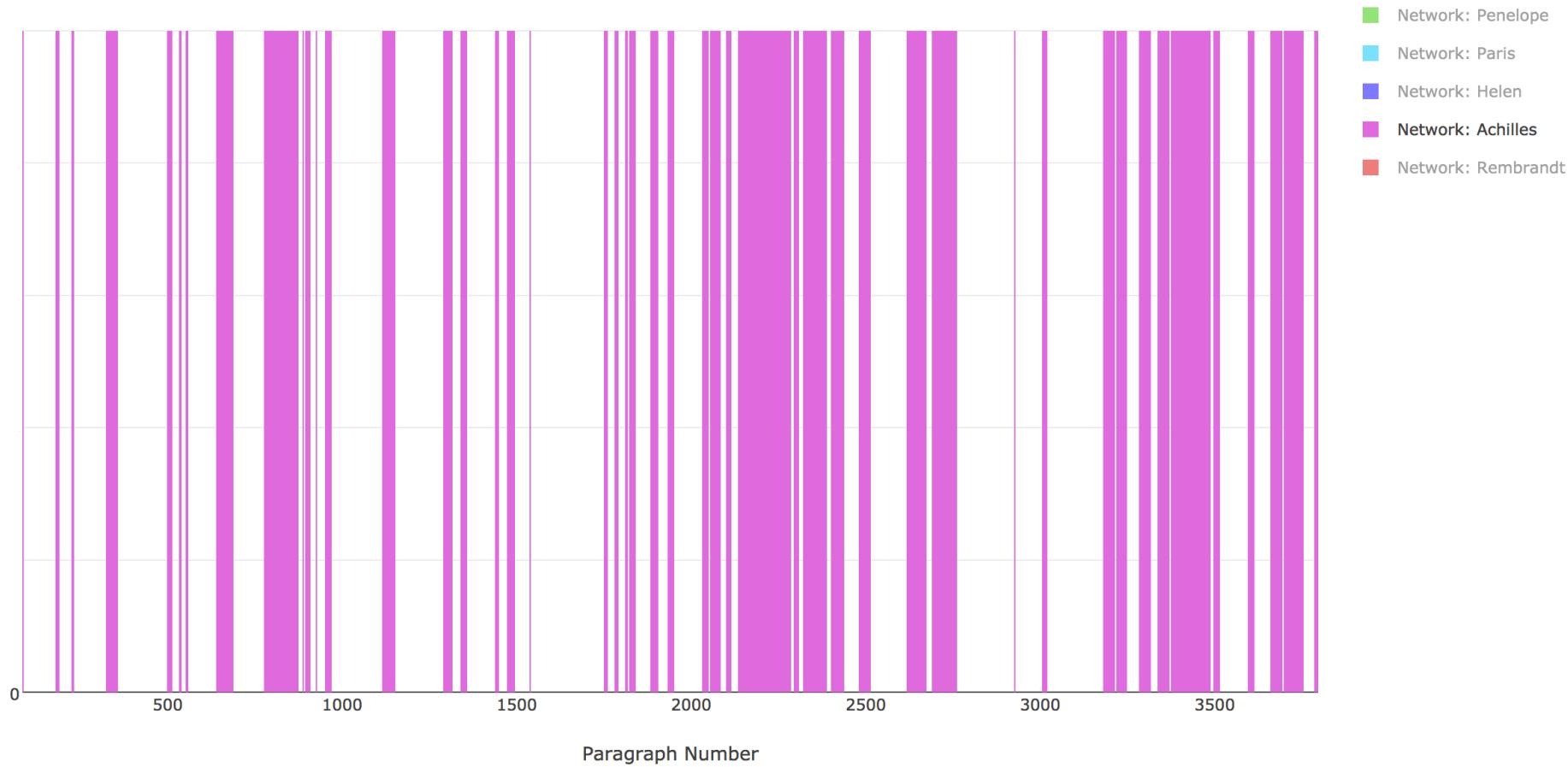
# Adding a Timeline



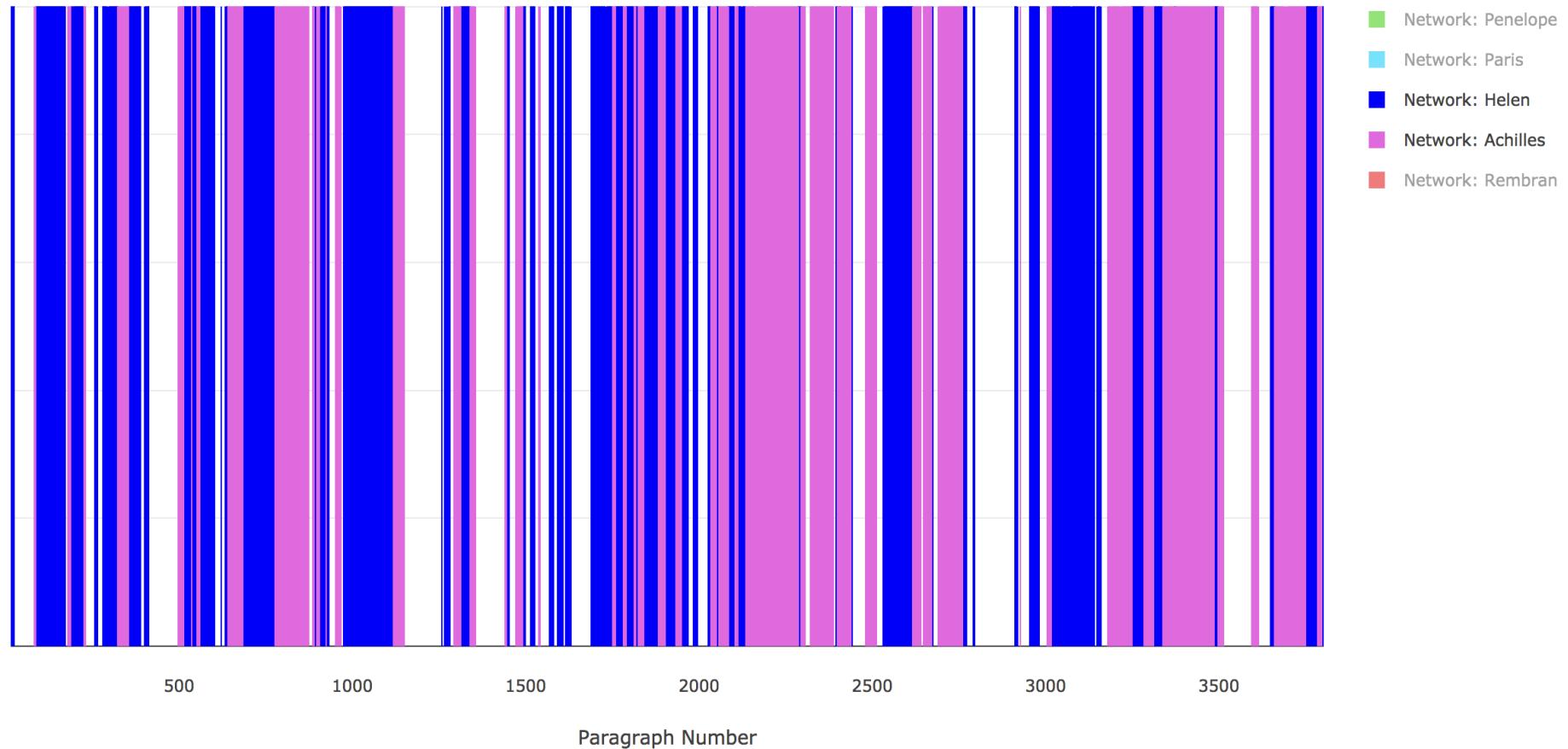
# Adding a Timeline



# Adding a Timeline



# Adding a Timeline



# Adding a Timeline

## Networks

0: Penelope

2: Helen

3: Achilles

Pattern	Count
< 3, 2 >	39
< 2, 3 >	35
< 2, 3, 2 >	24
< 3, 2, 3 >	21
...	
< 2, 3, 2, 3, 2, 3, 2, 3, 2 >	4
< 3, 2, 3, 2, 3, 2, 3, 2, 3 >	3
< 2, 3, 2, 3, 2, 3, 2, 3, 2, 3 >	2
< 2, 3, 2, 3, 2, 3, 2, 3, 2, 0 >	2

# Conclusions



# Conclusions

► Distant Reading applicable to this text

# Conclusions

- ▶ Distant Reading applicable to this text
  - ▶ Significant Phrase-Repetition

# Conclusions

- ▶ Distant Reading applicable to this text
  - ▶ Significant Phrase-Repetition
  - ▶ Clear entity-association networks

# Conclusions

- Distant Reading applicable to this text
  - Significant Phrase-Repetition
  - Clear entity-association networks
  - Repeated thought-patterns

# Conclusions

- ▶ Distant Reading applicable to this text
  - ▶ Significant Phrase-Repetition
  - ▶ Clear entity-association networks
  - ▶ Repeated thought-patterns
- ▶ Possible use in similar studies

# Conclusions

- ▶ Distant Reading applicable to this text
  - ▶ Significant Phrase-Repetition
  - ▶ Clear entity-association networks
  - ▶ Repeated thought-patterns
- ▶ Possible use in similar studies
  - ▶ Generic algorithm

# Conclusions

- ▶ Distant Reading applicable to this text
  - ▶ Significant Phrase-Repetition
  - ▶ Clear entity-association networks
  - ▶ Repeated thought-patterns
- ▶ Possible use in similar studies
  - ▶ Generic algorithm
  - ▶ Clear, quickly generated visuals

# Questions?

[conor.kelleher.1@ucdconnect.ie](mailto:conor.kelleher.1@ucdconnect.ie)

Conor Kelleher  
BSc. Computer Science

# Sample Paragraph Structure

*I have no idea how long ago it was when I was doing that. If I were forced to guess, I believe I would guess ten years.*

*Possibly it was several years longer ago than that, however.*

*And of course I was quite out of my mind for a certain period too, back then.*

# Phrase Extraction – Significant Phrases Only

“Why have I written that”

“Why have I written”

“have I written”

“Why have I”

“have I”

# Phrase Extraction – Significant Phrases Only

“Why have I written that”

~~“Why have I written”~~

~~“have I written”~~

~~“Why have I”~~

~~“have I”~~

# Phrase Extraction – Significant Phrases Only



“Why have I written that”