Windows on Waverley: exploring the effect of variations in the construction of literary social networks

Karen Wade and Siobhán Grayson
Nation, Genre and Gender Project
UCD Humanities Institute, Insight Data Analytics
karen.wade@ucd.ie; siobhan.grayson@ucdconnect.ie

The Nation, Genre and Gender Project:
a summary of our data

Project statistics (August 2016)

- 46 novels by 29 different authors, published between 1800 and 1922
- 1709 total chapters
- 9,630 total unique characters

www.nggprojectucd.ie/

Completed novels: list available at
https://theseaofbooks.com/novels-in-our-collection
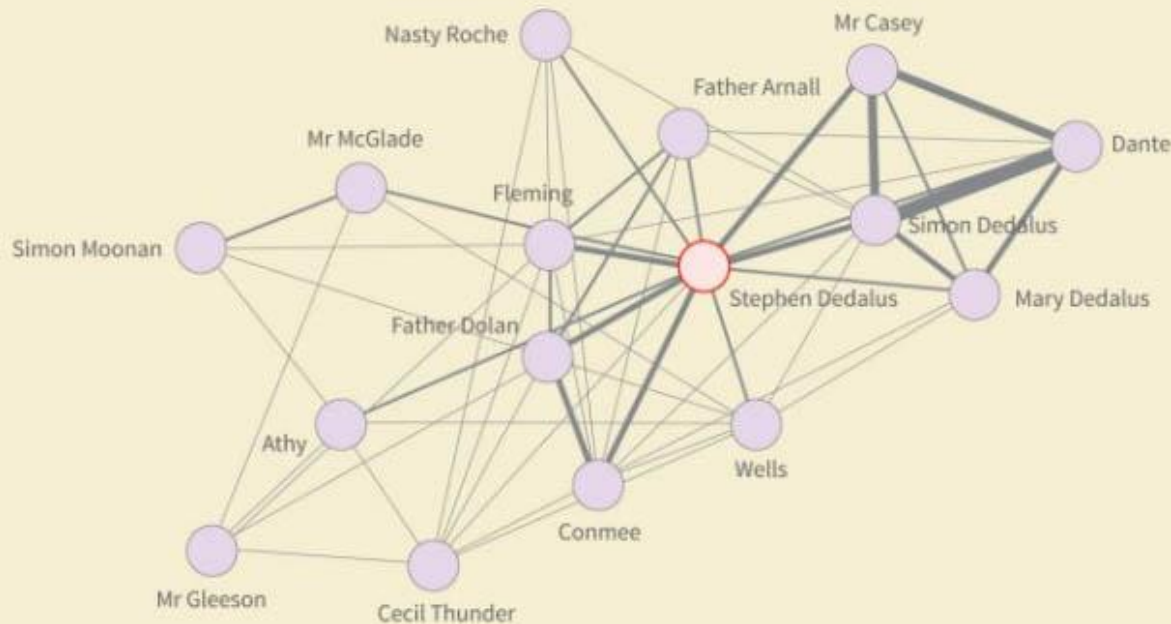
Authors by nationality
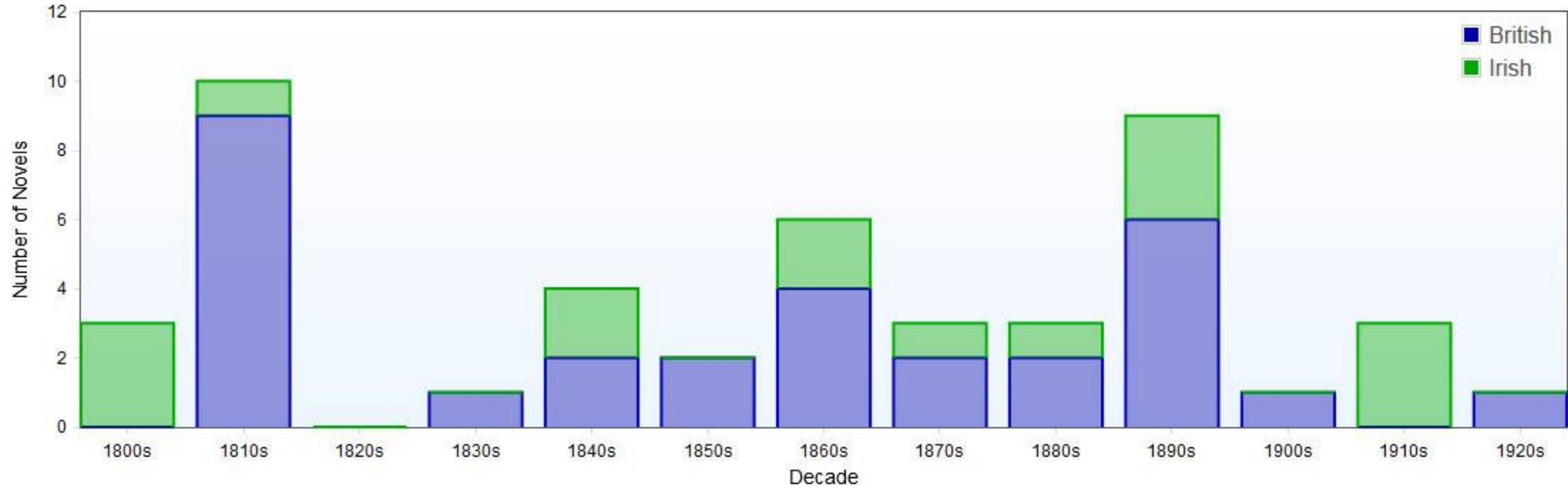
Authors by gender
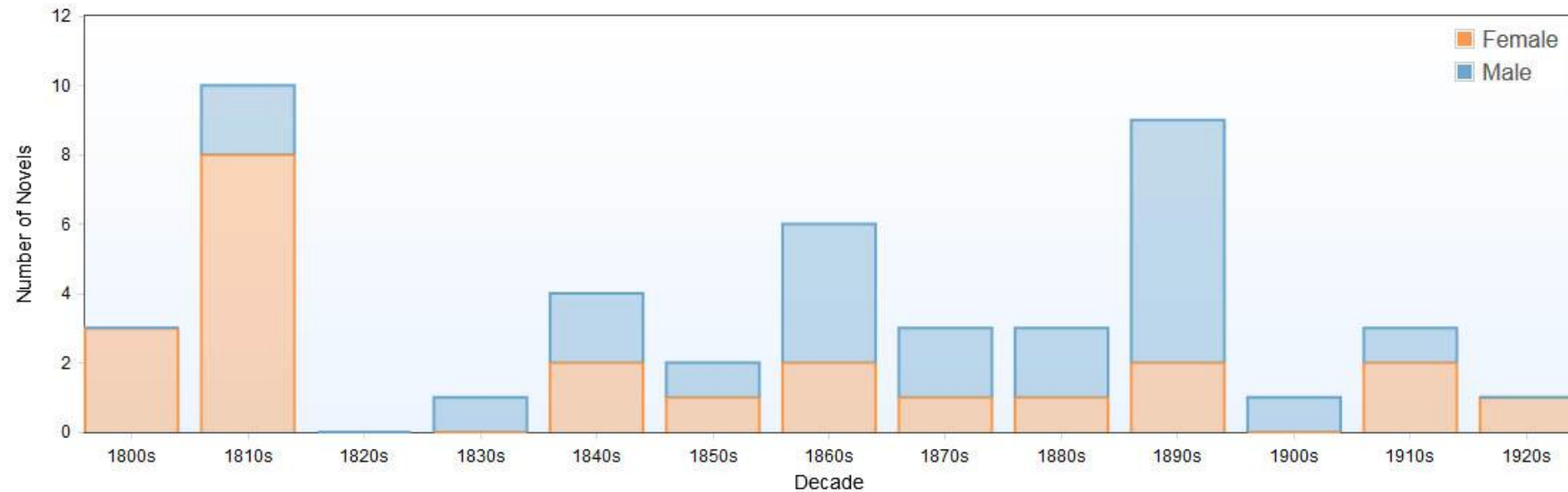
A Portrait of the Artist as a Young Man

CHAPTER 1 – *The social world of Stephen Daedalus as a young boy*

Nation
& Genre
Gender

His strongest relationships are with his parents, but he is also connected to extended family like Dante and Uncle Charles, his school 'fellows', Catholic clergy and his childhood sweetheart, Eileen.

Nasty Roche
Mr Casey
Father Arnall
Mr McGlade
Dante
Fleming
Simon Moonan
Simon Dedalus
Stephen Dedalus
Mary Dedalus
Father Dolan
Athy
Wells
Conmee
Mr Gleeson
Cecil Thunder

Strength of association

Weak	1
	2
	3
	4
	5
Strong	6

Data gathering and annotation: a four-stage process

1. Creation of a character dictionary

Definitive Name ▲	Aliases
Edward Waverley	captain butler, captain edward waverley, captain waverley, edward, edward waverley, frank stanley pseudonym, his little nephew, little edward, mr edward waverley, mr wauverley, mr. edward waverley, mr. wauverley, our hero, our young hero, rosy cherub, squoire, waverley, young edward, young waverley
Lady Alice Waverley	lady alice waverley, lady waverley
Miss Rachel Waverley	aunt rachel, miss rachel waverley, mrs rachel, mrs rachel waverley, mrs. rachel, mrs. rachel waverley
Nigel Waverley	nigel, nigel waverley, sir nigel, your grandfather

2. Annotation of complete text to clarify

A difference in political opinions had early separated Sir Everard from his younger brother Richard Waverley, the father of Edward Waverley. Sir Everard had inherited from his sires the whole train of Tory or High-Church predilections and prejudices which had distinguished the house of [Waverley] since the Great Civil War. Richard Waverley, on the contrary, who was ten years younger, beheld himself born to the fortune of a second brother...

3. Accounting for presence of first-person annotator
   - where necessary

I (Victor Frankenstein) am by birth a Genevese, and my family is one of the most distinguished of that republic.

4. Compiling attribute information for each character

Definitive Name ▲	Attributes
Alphonse Frankenstein	father, husband, male, swiss, syndic
Caroline Beaufort	daughter, female, mother, swiss, wife
Ernest Frankenstein	brother, child, male, son, swiss

Breakdown of Characters

- 9,630 total unique characters
- 18,377 total character aliases used to refer to characters
- Gender breakdown: 22.7% female, 55.4% male, 22.0% unspecified

Attribute	Frequency	%
servant	664	8.0%
wife	623	7.5%
husband	462	5.5%
aristocrat	412	4.9%
father	408	4.9%
child	406	4.9%
animal	390	4.7%
mother	346	4.1%
son	336	4.0%
daughter	309	3.7%
catholic	272	3.3%
brother	272	3.3%
sister	267	3.2%
army	257	3.1%
clergy	199	2.4%
politician	186	2.2%
speculative	170	2.0%
cousin	163	2.0%
doctor	149	1.8%
soldier	140	1.7%

Top 20 most common attributes

Attribute	Frequency	%
English	4830	57.9%
Irish	1223	14.7%
Scottish	185	2.2%
French	118	1.4%
American	71	0.9%
Swiss	38	0.5%
Austrian	38	0.5%
Jamaican	34	0.4%
German	34	0.4%
Belgian	32	0.4%
Russian	23	0.3%
Italian	22	0.3%
Spanish	10	0.1%
Hungarian	8	0.1%
Welsh	7	0.1%
Polish	6	0.1%
Dutch	5	0.1%
Greek	5	0.1%
Other	9	0.1%
Unspecified	1640	19.7%

Character nationalities

Our research design principles

Nothing is accidental in the world of a novel:

- every character, no matter how minor, should be included, and

- every character mention affects the nature of the social networks

Radical inclusivity

In the process of data gathering, our annotators identify every possible character for inclusion in the social networks - from

historical figures

to

characters appearing as groups

to

living animals

This can result in enormous communities - *Pride and Prejudice* has a relatively modest cast, at 117 characters, while *Middlemarch*'s population is 333, and *Vanity Fair* contains over 600 characters.

The importance of low-frequency characters

Grayson, Siobhán et al, ‘**Discovering Structure in Social Networks of 19th Century Fiction**’, in *Proceedings of the 8th ACM Conference on Web Science* (presented at the Web Science 2016, Hanover, Germany, 2016), pp. 325-26

Full text available at www.nggprojectucd.ie/about/ (bottom of page)

Summary: Jane Austen and Charles Dickens make different narrative use of minor or "low-frequency" characters in their respective nove
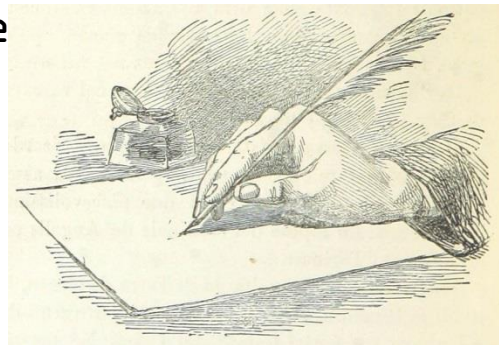
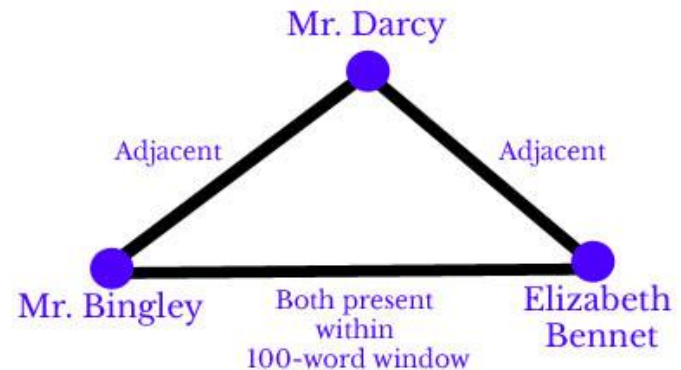# Two network strategies

Collinear

Adjacent   Adjacent

Mr. Bingley followed his advice. Mr. Darcy walked off; and Elizabeth Bennet remained with no very cordial feelings towards him.

Adjacent   Mr. Darcy   Adjacent

Mr. Bingley   Elizabeth Bennet
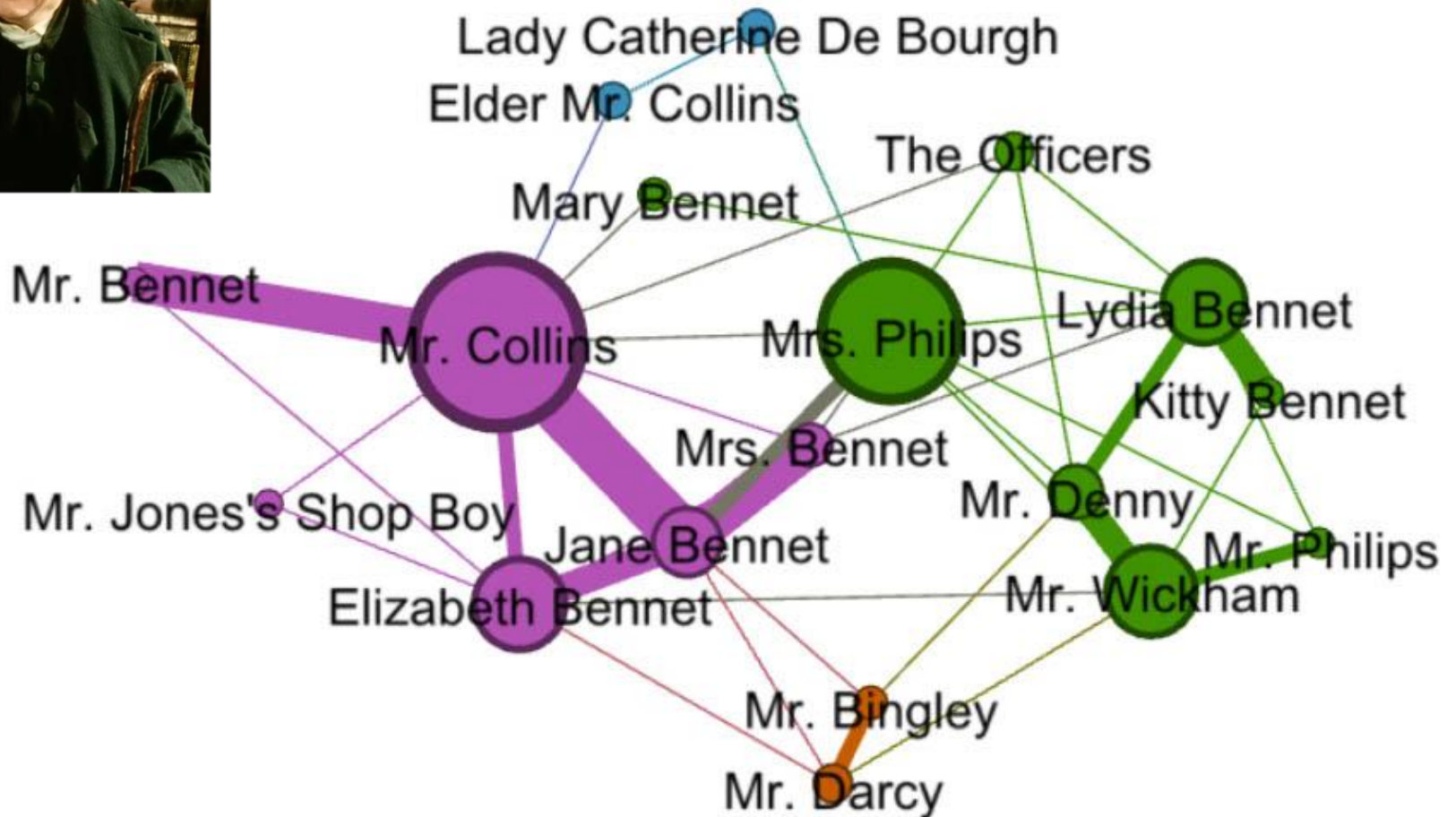
Two network strategies

Collinear

Adjacent   Adjacent

Mr. Bingley followed his advice. Mr. Darcy walked off; and Elizabeth Bennet remained with no very cordial feelings towards him.

Adjacent   Adjacent

Mr. Darcy

Mr. Bingley   Elizabeth Bennet

Combination

Adjacent   Adjacent

Mr. Bingley followed his advice. Mr. Darcy walked off; and Elizabeth Bennet remained with no very cordial feelings towards him.

Both present within 100-word window

Mr. Darcy

Adjacent   Adjacent

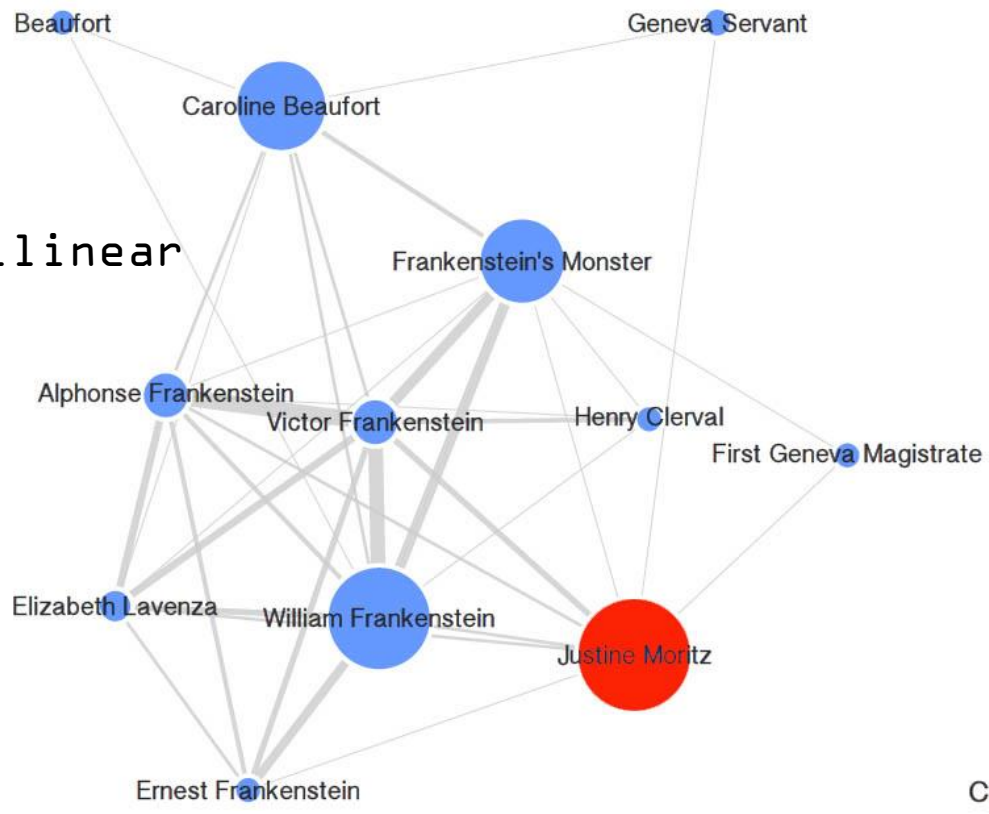Mr. Bingley   Both present within 100-word window   Elizabeth Bennet

Betweenness centrality

Collinear
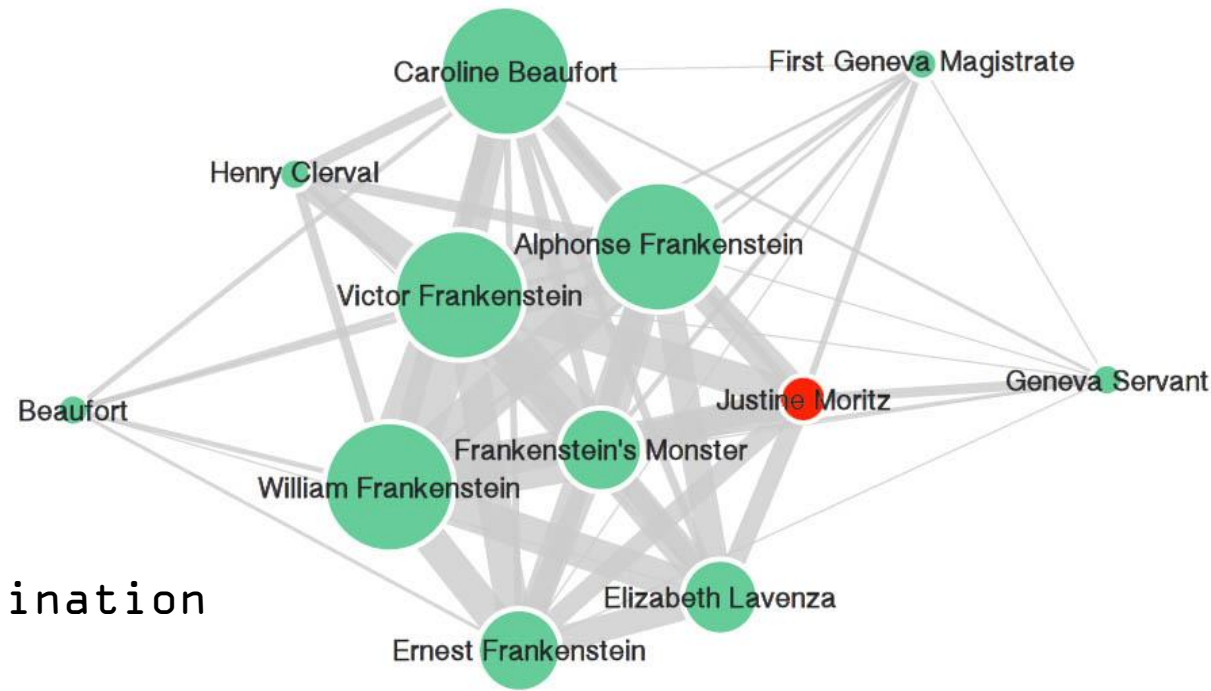
Mary Shelley's
Frankenstein,
chapter 7
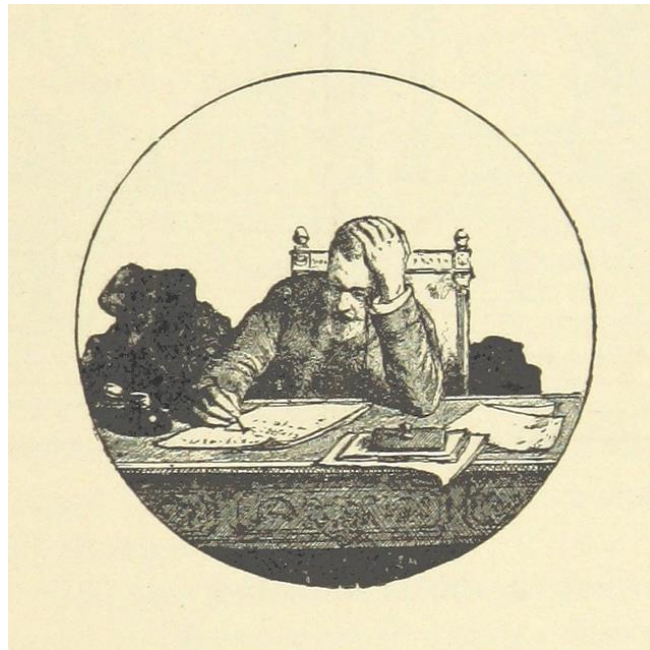
Combination

In a collinear network, the **recipient** is the highest ranked character for betweenness

In a combination network, the **sender** is the highest ranked character for betweenr

Pride and Prejudice chapter 35: Darcy's letter to Elizabeth

Collinear network rankings

(1) elizabeth bennet (recipient)
(2) mr. bennet
(3) mr. darcy
(4) lydia bennet
(5) jane bennet
(6) mr. wickham
(7) kitty bennet
(8) mary bennet
(9) mr. bingley
(10) elder mr. darcy

(11) miss bingley
(12) miss darcy
(13) wickham's father
(14) wickham's mother
(15) colonel fitzwilliam

Combination network rankings

mr. darcy (sender) (1)
jane bennet (2)
elizabeth bennet (3)
mr. wickham (4)
elder mr. darcy (5)
mr. bingley (6)
colonel fitzwilliam (7)
mr. bennet (8)
lydia bennet (9)
kitty bennet (10)
mary bennet (11)
miss bingley (12)

Rule holds true for letters in our dataset over 700 words long

Novel	Chapter	Letter word count	Chapter word count	Letter is % of chapter	Letter Recipient	Follows rule?	Letter Author	Follows rule?
Pride and Prejudice	35	3092	3111	99.39	Elizabeth Bennet	Yes	Mr. Darcy	Yes
Frankenstein	0	4163	5681	73.28	Margaret Walton	Yes	Robert Walton	Yes
Frankenstein	24	5249	8449	62.13	Margaret Walton	No*	Robert Walton	Yes
Pride and Prejudice	52	1796	3150	57.02	Elizabeth Bennet	Yes, 2nd place	Mrs. Gardiner	Yes
Frankenstein	6	1342	2806	47.83	Victor Frankenstein	Yes	Elizabeth Lavenza	Yes
Waverley	66	806	2373	33.97	Edward Waverley	Yes	Colonel Talbot	Yes, 2nd place
Waverley	28	713	2601	27.41	Edward Waverley	Yes	Rose Bradwardine	Yes, 2nd place
Pride and Prejudice	46	673	3160	21.30	Elizabeth Bennet	No	Jane Bennet	No
Frankenstein	22	660	3507	18.82	Victor Frankenstein	No	Elizabeth Lavenza	No
Frankenstein	7	687	3694	18.59	Victor Frankenstein	No	Alphonse Frankenstein	No
Pride and Prejudice	26	490	2432	20.15	Elizabeth Bennet	Yes	Jane Bennet	No
Waverley	55	281	1490	18.86	Colonel Talbot	No	Lucy Talbot	No