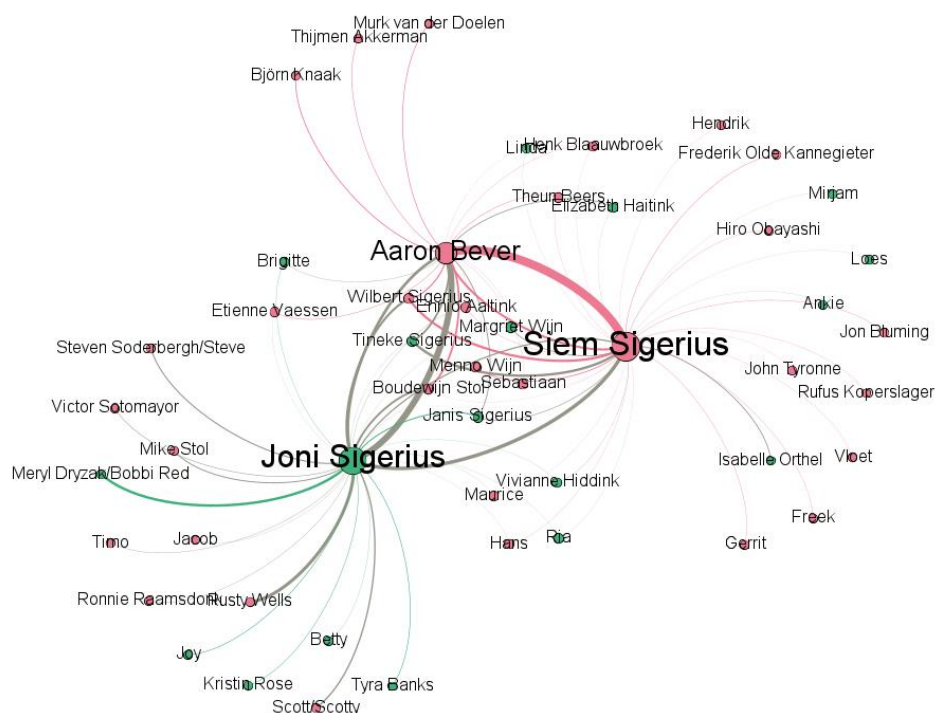# Tools and Methods. Final Assignment

Part A. Character relation

From this graph we can gather information about who the main characters are, which are indicated by the circles. The size of the circle shows us the importance and the weight each of these characters have in the network of the novel. We can identify very quickly which characters are female and male, thanks to the use of different colours for each gender. This kind of plot also gives us a general idea of the novel, regardless of whether we have read it or not. For example: in the story there is a strong relationship between Aaron and Siem, who are both connected with Joni. Siem is one of the main/protagonist characters, and "Why is the protagonist significant here?", as Moretti asks, "Not for what is 'in' it; not for its essence, but for its function in the stability of the network" (2011, p.88). This stability is reflected in the multiple relationships that Siem, and also Joni, has with other characters, and how they build a strong enough network to keep the story on track.

Comparing the example of Bonita Avenue with the example of Hamlet that Moretti mentions in his article, we also see a large difference in interconnectedness between side characters. In a play like Hamlet, all characters go on the same stage and there are many opportunities to interact with each other, building an interconnected clique. In this graph for Bonita Avenue, the central position of the three viewpoint characters becomes clear: Since the edges are based on their mentions of characters, there is no independent connection between non-viewpoint characters. Removing either Aaron, Joni, or Siem would result in a large portion of the network vanishing with them.

But this also leads to the limitations of a network approach such as this one. As Moretti points out, there are important pieces of information that get lost in this network visualisation: We see a small representation of the secondary characters, when perhaps they are a strong part to keep the network working. We can also lose the context of the story. The storyline of a tale is made up of characters as well as maybe additional events, ideas, surroundings, etc. This is reduced to a series of nodes and edges - useful for comprehending the relations between characters in a time-independent format, but useless for other aspects of a novel.

Part B. Textual relations
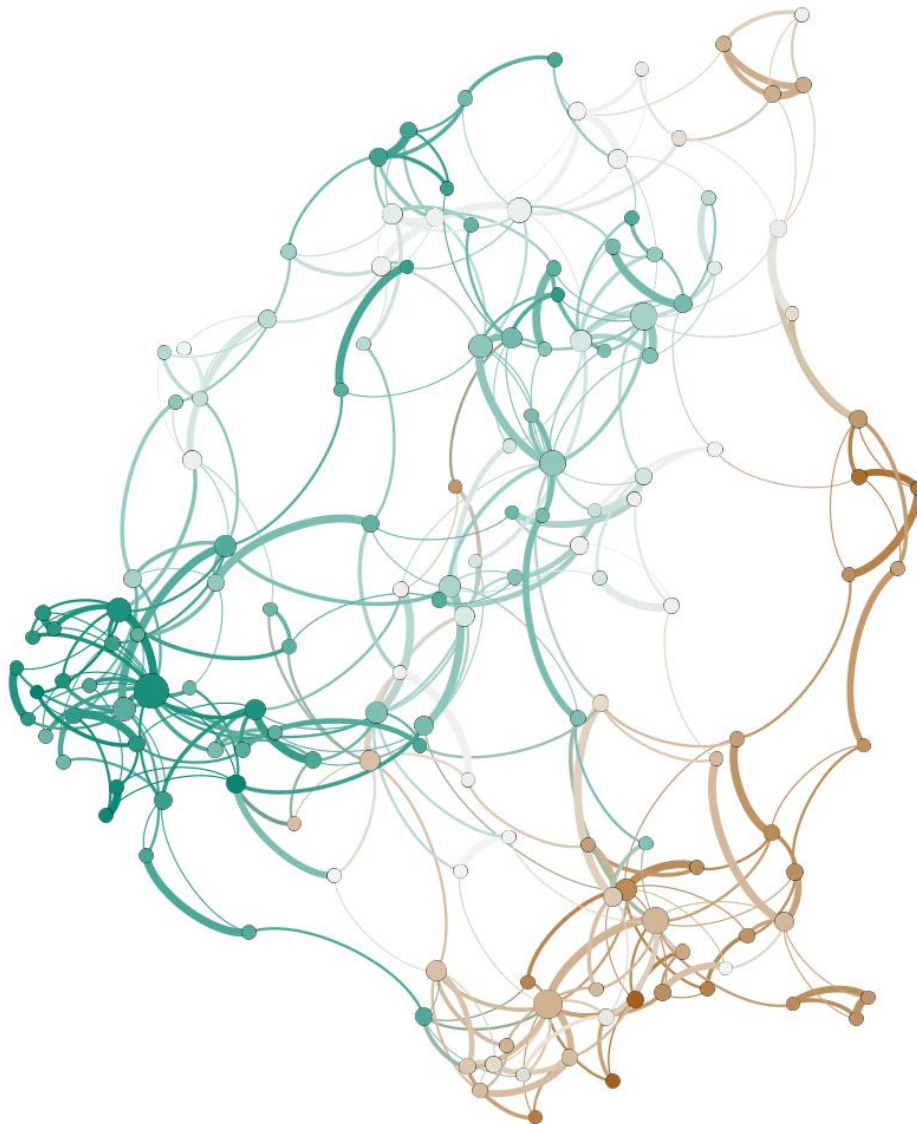
The network visualisation above uses the Yifan-Hu algorithm. Nodes are sized according to degree and coloured according to the year their book was released in, while edges are sized and coloured according to their weight.

A piece of information that I chose not to represent was the gender of the authors, as this does not seem to influence a novel's style as much as other factors. On this visualisation, female authors are broadly clustered to the right and male authors to the left. Male authors also have the books with the largest degrees, though a more thorough statistical analysis would be needed for a real interpretation of this. However, there are many stylistic connections between male and female authors. There is a potential analysis that is hard to do with the current data, but could be potentially interesting: Transforming the current undirected stylistic similarity edges into directed ones based on the publication year of the novels. This could maybe show that some authors were more influenced by others or influenced others stronger.

Coming back to the analysis that can be done with the current data, there are two obvious clusters forming, seemingly according to their books' publication years: A 1790 to 1830 cluster on the bottom right and a 1900 to 1930 cluster on the left. Smaller clusters appear to be related to authors: Jane

Austen clusters mostly by herself on the top right, while Maria Edgeworth and Fanny Burney cluster to the right. A potential interpretation of this is that novels first tend to cluster if they have the same author and then form to larger clusters when they were published in a period with strong stylistic tendencies. This would explain why the novels of individual authors always tend to be connected, but there are then only two identifiable clusters according to period.

Apart from these generalisations, the visualisation also allows us to identify outliers. Why, as mentioned above, are Austen, Edgeworth, and Burney so unconnected to other novels – apart from Edgeworth's *Castle Rackrent*, which is completely separated from her other novels? Why is Mary Noailles Murfree's *The Frontiersman* located at the edge of the 1790-1830 cluster, even though it was published in 1904? Why are John Galt's *Annals of the Parish* and James Hogg's *Three Perils of Man* located at the edge of the 1900 to 1930 cluster, even though they were published in 1821 and 1822, respectively? Why does F. Scott Fitzgerald's *The Beautiful and the Damned* seem to have by far the largest degree of interconnectedness in the network? These questions cannot be answered through this network visualisation alone: They call for a potential further close reading investigation.

This is a visualisation of the connections between the world's 200 largest airports. Nodes represent airports, with their size based on the total number of seats in flights connected to them (also to airports not represented in the visualisation) and their colour based on their degree in the network. Edges are routes, with their weight based on the number of routes connecting two airports. The data comes from the World Bank and openflights.org.

The visualisation seems to neatly show the current "hub" model of international air travel, where airlines mostly target a few large hub airports with their international connections, from where short-range flights then connect to regional airports. This results in the clustering into continents seen in the visualisation, with the hub airports at the "inner" sides of the continental clusters. Dubai provides connections to most of the Arab world, while London Heathrow, Frankfurt, Paris Charles de Gaulle, and Amsterdam Schiphol provide them to Europe. Incheon, Hong Kong, Bangkok, and Shanghai Changi are the hubs for Asia.

America is an interesting case, as the airports are much smaller based on total seat counts, but have heavily weighted connection flights between them (not well visible in this export of the network). This could be interpreted to mean that the hub model is not so strong for America, as travel seems to

happen on a more distributed national level even between the smaller airports. John F. Kennedy in New York, Toronto and Los Angeles are still large international hubs, but they do not come close to those of Europe or Asia. This would make sense taking into account the denser rail networks in parts of Europe and Asia, which probably handle more national level travel than in the American case, but further investigation would be needed to confirm such a hypothesis.

The visualisation can also be interpreted from an economic point of view, as hub airports are heavily associated with certain airlines: Frankfurt with Lufthansa, Heathrow with British Airways, Dubai with Emirates, etc. We can assume that these airlines are responsible for large parts of the degrees of their respective home hubs, which therefore to a certain degree also reflect these airlines' economic influence over international air travel.