Gendered Data in Medieval and Early Modern Sources
The *Gendered Networks* and *Digital Edgeworth Network* Projects

Máirín MacCarron
University College Cork, Ireland

Abstract   The *Gendered Networks* and *Digital Edgeworth Network* projects applied data-driven approaches and digital analysis to material that might not seem entirely conducive to such methodologies. The process of creating viable datasets for both projects was researcher-led due to the complexity of the data. The resulting datasets highlighted aspects of the source material that had hitherto received little attention and led to the development of new research questions. Our data-driven approach ensured that all individuals were included in the data gathering process and allowed us bring to the fore people who might otherwise have been deemed invisible.

Keywords   Data analysis. Gendered data. Social network analysis. Correspondence networks. Bede. Maria Edgeworth. Narrative sources.

Summary   1 Introduction. – 2 *Women, Conflict and Peace: Gendered Networks in Early Medieval Narratives*. – 3 *Digital Edgeworth Network*.

1 Introduction

Data-driven approaches to historical sources have expanded the potential of historical inquiry. The advent of easily accessible computer processing power and the mass digitisation of a whole range of sources have allowed historians to ask questions at a scale that would have been unthinkable for earlier generations. However, the

process of transforming complex historical evidence into data is not straightforward and often requires a different methodological approach which can lead to a researcher requiring a new relationship with their sources.[1] The first step in such endeavours centres on the creation of usable datasets, with an increasing emphasis on not just big but smart data, that is, enriched and structured data (Schöch 2013). While the emphasis on history as 'data' can provoke unease, historians have always managed and categorised their information, and it is not such a major conceptual leap to think of these 'datasets' as enriched and open to additional forms of analysis, whether quantitative or digital.[2] Additionally, in an influential article over a decade ago, Johanna Drucker argued that 'capta' is a better word in the Humanities, because 'data' means 'what is given', while 'capta' means 'what is taken', which is a more accurate reflection of the creation of datasets in the humanities (Drucker 2011; Vitali, Pasqual, *supra*).

Terminology aside, however, the biggest challenge when attempting to create viable 'datasets' from humanities sources concerns the reductive nature of the process. The more complex the source – whether life experience, a painting, or a written text – the more complicated the process of data collection, because not all the complexity can be retained (D'Ignazio, Klein 2020, 10). This is not of itself a negative because meaningful data analysis is not possible without some reduction, and decisions about what to retain is part of a heuristic process that feeds back into the analysis. The challenge for the historian is to determine an acceptable level of loss vis-à-vis the potential gains from data analysis, and to ensure the data that is gathered can address the research questions posed. The affordances of a data-driven approach are many: for example, in analysis of social groups, it is possible to quantify relationship data and visualise social networks in whole new ways. In addition, if all individuals are included in the dataset, irrespective of their perceived or otherwise social importance, we can bring to the fore previously less visible or over-looked figures (Hillner, MacCarron, forthcoming).

The two projects that will be discussed here, *Women, Conflict and Peace: Gendered Networks in Early Medieval Narratives* and the *Digital Edgeworth Network*,[3] both used data-led approaches and digital methods, including social network analysis, in analysing our chosen corpora of sources. In what follows, I will outline the data models that we developed and share some of our findings, including our

---

[1] Hitchcock 2013 has lucidly highlighted the challenges historians face from mass digitisation and increasing reliance on searches run by OCR.

[2] On quantitative approaches, cf. Lemercier, Zalc 2019.

[3] See below for full information including details of funders and composition of project teams.

success in highlighting the presence of individuals or patterns in our datasets that have received little attention in scholarly analysis of these sources. As gender representation was a concern of both projects, though less so in the *Digital Edgeworth Network*, we opted for researcher-led data gathering, rather than computerised methods of data extraction, such as distant reading, which have been shown to be deficient when attending to non-dominant subjects, including women (Klein 2018; Cvetkovich et al. 2010).[4] Although machine learning data gathering processes can ward against potentially subjective data collecting practices, and are not as time-consuming as extracting data through close reading, these methods cannot match the same level of detail as a trained researcher engaged in textual analysis which is what both projects required.

2 *Women, Conflict and Peace: Gendered Networks in Early Medieval Narratives*

Women, Conflict and Peace: Gendered Networks in Early Medieval Narratives (hereafter referred to as *Gendered Networks*), used a data-driven approach to interrogate the role of women in early medieval narrative sources.[5] This was an interdisciplinary project with historians, physicists and a computer scientist as full members of the team.[6] The project's goal was to produce a quantitative large-scale investigation of women's representation in narrative sources, particularly focussing on their roles as connectors of men. This, in part, emerged from the idea that women were often presented as 'peaceweavers' in heroic narratives: political marriages were often arranged in attempts at conflict resolution, whereby a woman would marry the son of a family with which her family were at war and move to her new husband's region in an attempt to weave peace between their

---

4   Mandel 2016 suggests ways to counteract the gender bias in so many digital projects.

5   This project was funded by the Leverhulme Trust RPG 2018-014, and ran from 1 September 2018 to 31 January 2021; the project website is available here: https://sites. google.com/sheffield.ac.uk/genderednetworks/home.

6   The core team consisted of three historians (Julia Hillner, Máirín MacCarron, Ulriika Vihervalli), three physicists (Silvío Dahmen, Ralph Kenna, Sandra Prado) and one computer scientist (Ana Bazzan); we were assisted by another physicist (Pádraig MacCarron) and two historians (Rob Heffron and El Bailey). The historians prepared the datasets using close reading and interpreted results; the scientists wrote bespoke codes to mathematically analyse the datasets, and produced data visualisations including graphs, charts and tables; both groups worked together in analysing and interpreting our findings. See the team page on the project website: https://sites.google. com/sheffield.ac.uk/genderednetworks/team.

families.[7] We examined narrative sources, specifically histories and hagiographies (spiritual biographies, usually accounts of saints' lives), from the late Roman world, Merovingian Francia and early medieval Britain, that dated from the fourth to the eighth centuries. We wished to determine if the representation of women and their social role was regionally specific and if it changed during our chosen time period. In applying social network analysis to our datasets, we were specifically interested in character networks and dynamic networks (i.e., assessing temporal influence in a social network).[8]

The process of creating datasets for *Gendered Networks* was complicated because we required a sophisticated data model that could be applied to all our sources that would allow us to conduct social network analysis at a deeper level. To do so, we needed to distinguish between several different types of relationship, and ultimately settled on a data model that accounted for twenty-one different categories of relationships. These were grouped under nine broad headings and expanded into sub-headings for greater clarity: for example, the first heading, 'Family', covers a range of relationships including domestic slaves because we wished to indicate the crucial role such people played in the household **[tab. 1]**. Occasionally, the classification of a relationship is not obvious: for example, if a pagan barbarian king murdered a Christian king from a different polity, should that be recorded as military, political or religious hostility? Ultimately, our decision was determined by the context provided by the source: if our author indicated that the interaction was informed by both political and religious motives, we recorded the connection under both headings; if it was purely political, it was only recorded under that heading and so on. Finally, categories 20 and 21 – post mortem and supernatural – are included because we occasionally encountered interactions between the living and the dead, or with angels and demons. We included such relationships to respect the integrity of the sources by recording all individuals and all connections, but these categories can also be removed from our analyses by focussing on categories 1 to 19 only.

---

7 Cf. *Beowulf*, line 1942: the Old English word was 'freothuwebbe'. A classic study of this phenomenon is Rosenthal 1966.

8 On applying social network analysis to these sources, cf. Hillner, MacCarron, forthcoming.

Table 1 *Gendered Networks* Data Model

1	Family	Kinship
2		Marriage & betrothal
3		Concubinage
4		Domestic relations (inc. household slaves/ servants)
5		Fosterage
6	Friendship	Friendship
7	Religious	Church meetings/councils
8		Monastic/Religious *familia*
9		Spiritual kinship (e.g., godparents)
10	Patronage	Patronage
11	Politics	Political connections
12		Physical diplomacy including messengers
13	Hostility	Military violence
14		Political hostility
15		Domestic hostility
16		Religious hostility
17		Gender-based violence
18	Transmission of Information	Letter recipients
19		Sources of information, textual or oral
20	Post Mortem	Post mortem (connections between the living and the dead)
21	Supernatural	Supernatural/Divine interventions

In addition to categorising relationships between characters in our sources, we also recorded attribute data for every character. This included their gender, whether they were named or unnamed in the source, if they were a group or individuals, and their ethnicity, social status, religious status, and geographical location when this information was available. Such information ensures our datasets are rich and allow us to address a range of research questions. Our central research question concerned gender, but the process of data gathering led to the development of additional research inquiries: for example, we detected variations in naming practices which led us to add entries to our datasets concerning whether characters were named when they appeared in our sources. On investigating these practices, we discovered that nearly two-thirds of female characters were left nameless by our authors, whose decisions were influenced by pragmatic, rhetorical or socio-cultural reasons. In addition, our attention to naming practices highlighted the important role played by nameless characters in historical sources and revealed that such individuals can become a scaffold on which plots are structured (Hillner, MacCarron, Vihervalli 2022). This research direction underlined the

importance of carefully managing the presence of the anonymous and unnamed in data-driven research projects.

The complexity of our data model with its attention to relationship categories and character attribute data ensured that our data gathering could only be done by close reading of each text, as explained in the introduction. We treated each unit of a text, whether chapters or sections, as discrete, and we recorded a character anew every time they appeared in a new unit of the text. This method enabled us to weight relationships; i.e., in addition to categorising relationships by type, we also measured the number of times that any two characters were connected in a text, leading to greater depth in our analysis of social networks. Secondly, taking account of the units within texts allowed us to trace progression over time in each source. Our particular concern was to follow plot development in the narrative time of the author, rather than in chronological time, though these often mirrored each other; however, if a narrative moves backwards and forwards in chronological time, our method allows us to follow the narrative's internal chronology. We used the time data to analyse dynamic networks within our texts, i.e., assessing influence over time, as well as engaging in more traditional, static social network analysis (Prado et al. 2020).

The efficacy of our approach can be seen in the following case study concerning Osthryth (d. 697), queen of the kingdom of Mercia in early medieval Britain. Osthyrth was born into the royal house of Northumbria in the mid-seventh century, became queen of Mercia during the 670s (precise date is unknown), and was subsequently murdered by her own nobles in the late-seventh century.[9] Her presence in the source record is minimal: she appears three times in Bede's *Historia ecclesiastica gentis anglorum* (completed c. 731) and once in the *Life* of Wilfrid of York (c. 634-709/10) which was written within a few years of Wilfrid's death by Stephen, a member of Wilfrid's community (Bede 1969; Stephen 1927).[10] Osthryth appears in the narrative of Bede's *Historia* due to her role as queen of Mercia, and hers was undoubtedly a politically-motivated marriage intended to bring peace between the warring kingdoms of Mercia and Northumbria (Bede 1969, 3.11, 4.21). Indeed, the marriage of Osthryth and Æthelred was the third such marriage between the royal families of Mercia and Northumbria within a generation, none of which

---

9   The early medieval kingdom of Mercia roughly corresponds with the modern-day English midlands; the kingdom of Northumbria spanned present-day northern England and southern Scotland. Mercia and Northumbria shared a border, which was often in flux in this period.

10   On Bede, cf. Darby, MacCarron 2023. On Stephen and the *Vita Wilfridi*, cf. Thacker 2013; Stancliffe 2013.

succeeded in bringing about peace.[11] That Osthryth was murdered by her Mercian nobles underlines the difficulties of her role, although the specific reasons for her murder are unknown; Bede recorded her death in a short recapitulation of events included in the final chapter of the *Historia* (5.24) but not in the main narrative.

Our other source for Osthryth, the *Life of Wilfrid*, refers to her only once. Wilfrid of York was a contentious character who was exiled from his bishopric of Northumbria three times, the first two at the hands of King Ecgfrith of Northumbria, Osthryth's brother. On the occasion of his second exile, Stephen related that the bishop fled first to Mercia, then to the kingdom of the West Saxons, further south, and he finally found refuge in the kingdom of the South Saxons. Wilfrid's hagiographer claimed that his hero had been driven from the kingdoms of the Mercians and the West Saxons because of royal kinship networks, noting that the Northumbrian king's sister, i.e., Osthryth, was queen in Mercia and his sister-in-law was queen of the West Saxons. Stephen suggested that these royal couples expelled Wilfrid to ingratiate themselves with Ecgfrith (Stephen 1927, 40).[12] This chapter is Osthryth's only appearance in Stephen's account and she is unnamed, as was the queen of the West Saxons whose name is unknown.

Osthryth is clearly something of a shadowy figure in the narratives of the only two sources that mention her. Consequently, she seems somewhat peripheral in the networks of both sources. However, when we consider her place in these networks using metrics such as degree and betweenness centrality, her significance in the network structure is greater than one might expect. Degree is the number of connections that any character, or node, has in a network: for example, a character who is connected to five other nodes would have a degree of five. Betweenness centrality represents the role that an individual node plays in connecting others within the network: that is, an individual character, A, could have a low degree, but high betweenness if the characters that A is connected to are connected to many others, meaning that A is important in connecting those that A is between. The dataset we created from Bede's *Historia ecclesiastica* contains 594 characters, and just over 12% of those are women: 72 women, 509 men, and 13 non-gender.[13] Of these 594 characters, Osthryth is ranked seventy-fourth for degree, and thirty-first for betweenness. We see a similar pattern in Stephen's *Life of Wilfrid*,

---

11   Cf. Bede 1969, 3.21, for the earlier marriages of Osthryth's half-sister, Alhflæd, and half-brother, Alchfrith, to Æthelred's brother, Peada, and sister, Cyniburg, respectively. Indeed, many years before Osthryth's murder by her Mercian nobles, Alhflæd was accused of involvement in the murder of her husband, Peada: Bede 1969, 3.24.

12   For discussion of royal networks and Wilfrid, cf. Hillner, MacCarron 2021, 31-40.

13   Non-gender is used for those whose gender is unknown including angels and demons which are of indeterminate gender in our sources.

for which our dataset has 156 characters, and just over 15% are women: 24 women, 140 men, and 3 non-gender. Osthryth is placed sixty-fourth for degree, and fifty-sixth for betweenness.

There are two intriguing elements here; firstly, for a relatively inconsequential character, Osthryth punches above her weight in terms of network metrics in both works. This is especially notable in the *Life of Wilfrid* where she appears only once, and nameless, but is in the top half for both metrics. This result is, no doubt, largely due to her connections to royal houses, but it is nonetheless a significant result. Secondly, in both sources, Osthryth ranked higher for betweenness centrality than for degree. Essentially, she is not distinguished by the number of her direct connections in either source; however, she is of greater importance in both networks for connecting others. This is especially apparent in Bede where, as shown before, her betweenness centrality is surprisingly high, considering the size of the network and that she only appears three times in the text. It also highlights the nature of her role within this society: queens were connectors, of men and women, a reality that Bede's account reflects when examined from the perspective of relationship connections. Indeed, this reality may also underlie Stephen's account, because Osthryth is relatively highly placed for betweenness in the *Life of Wilfrid* despite only appearing once in the text.

We can further develop our understanding of Osthryth's role in these sources, and better understand the sources themselves, when we examine the representation of relationship categories in these networks. It has long been noted that Stephen and Bede treat women very differently in their works, though the reasons for these differences have been disputed. In an influential publication, Stephanie Hollis argued that women have more agency in Stephen's *Life of Wilfrid* than Bede allows for them in his works (Hollis 1992). However, a close analysis of the women in Stephen's account indicates that many are engaged in hostile relationships with Wilfrid **[fig. 1]**.[14] The graph represents hostile relationships in the *Life of Wilfrid* as defined by the *Gendered Networks* data model. The four categories in use are military, political, religious, and domestic hostility. The characters' gender is indicated by the colour and trajectory of the triangles: male nodes are teal-coloured upward-pointing triangles, female nodes are magenta-coloured downward-pointing triangles. The size of the triangles is determined by the number of connections that each character has, i.e., their degree. Wilfrid is clearly at the centre and is the largest triangle, indicating he has the highest degree in this network.

---

14   The networks in figures 1 and 2 were drawn by Ana Bazzan, Sílvio Dahmen and Sandra Prado using a bespoke programme written for the *Gendered Networks* project. For further discussion of what follows, cf. Hillner, MacCarron 2021, 37-40.
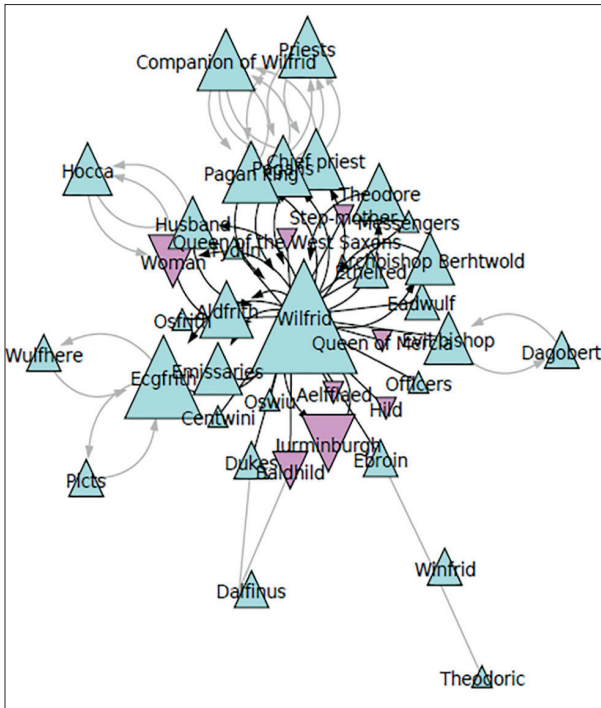
Osthryth, here presented as 'Queen of Mercia' because she is name-less in the text, is to the right of Wilfrid and is one of several women in close proximity to him. The number of nodes in the graph visually represents the high level of hostility that Wilfrid generated during his life, and highlights Stephen's view that women were central to this; the significance of women in this graph becomes more apparent when we compare figure 1 with hostility graphs for Bede's *Historia*.

The *Historia* is a large text presented in five books and covers the history of Britain from the invasions of Julius Caesar in 55 and 54 BC up to the first decades of the eighth century, although most of the work is concerned with events in the seventh century. As the work is so large, it can be difficult to analyse visually, so figure 2 presents the hostility graphs for books 4 and 5 separately [fig. 2]. These books cover the periods of Wilfrid's enforced exiles in the late seventh and early eighth centuries. As with the *Life of Wilfrid*, these graphs represent hostile relationships defined by the *Gendered Networks* data model, and include the same four categories of military, political, religious, and domestic hostility. The colour-coding, trajectory and size of triangles are the same as in figure 1.
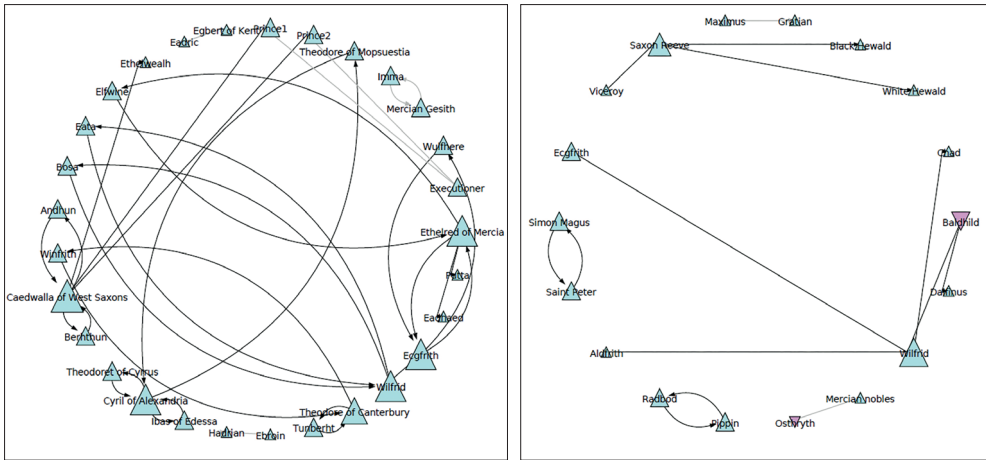
It can be difficult to perform detailed visual analysis of these graphs, which is why network metrics (as outlined above) are essential when analysing complex networks. The key visual finding from the graphs in figure 2 is the small number of women who appear, especially when compared with the hostility graph from the *Life of Wilfrid*. Indeed, there are no women in the hostility connections for book 4, despite the fact that there are many women in this book; moreover, book 4 is sometimes referred to as the 'book of abbesses' because Bede devoted so much attention to female religious houses.[15] There are two women in the hostility connections from book 5, but only one is connected to Wilfrid: Baldhild, a Frankish queen, who murdered the archbishop of Lyon, Wilfrid's early patron (Bede 1969, 5.19; Stephen 1927, 6). Osthryth is represented in a hostile relationship with her Mercian nobles, to represent her murder, but she is not connected to Wilfrid. A comparison of these hostility graphs reveals the extent to which Bede omitted the female network hostile to Wilfrid in his account, despite this being a significant feature of Stephen's account, and Stephen was one of Bede's key sources for the *Historia*. These

---

[15] Cf., for example, Abbess Æthelburh and the female community at Barking, Bede 1969, 4.6-10; Abbess Æthelthryth and Ely, Bede 1969, 4.19-20; Abbess Hild and the double-monastery of Whitby, Bede 1969, 4.23-4; Abbess Æbbe and the double-monastery at Coldingham, Bede 1969, 4.25.

findings suggest that Bede may have suppressed evidence of women hostile to Wilfrid, rather than women in general.[16]

This analysis reveals the ways in which our understanding of these sources can be enhanced when taking a data-led approach. A quantitative and digital approach combined with qualitative analysis has added another dimension to our interpretation of the presentation of women in the structure and development of these narratives. Our second project concerns a different type of source material, specifically correspondence collections, but also conducive to a data-driven approach.

3 *Digital Edgeworth Network*

The *Digital Edgeworth Network* (hereafter referred to as *DEN*) was developed to explore and analyse the archive of the celebrated author, Maria Edgeworth (1768-1849) and the Edgeworth family.[17] The Edgeworths were an Anglo-Irish family who owned extensive properties in the midlands of Ireland when Ireland was under British rule. Their archive contains over 40,000 items and the holdings are primarily held in two institutions in different countries: the Bodleian Library in Oxford, UK, and the National Library of Ireland in Dublin.[18] *DEN* was devised as a small-scale proof-of-concept project which would explore the potential of data-driven methods for approaching a complex and divided archive. The project was jointly based in the University of Oxford and University College Cork.[19]

Maria Edgeworth is the primary reason for interest in the Edgeworth archive, but we wished to examine the collection in the broadest possible terms rather than focussing on the family's most famous member. As the archive is extensive, we focussed on the Edgeworth

---

16 Stephen's pejorative attitude towards queens is further evidenced in calling two of these women Jezebels: the Frankish queen, Baldhild (Stephen 1927, 6); and Queen Iurminburgh of Northumbria, who was Osthryth's sister-in-law (Stephen 1927, 24). Jezebel was Ahab's queen in 1 Kings 16-9, she was accused of killing the prophets and became an archetypal evil queen (Nelson 1986). It is notable that Bede did not resort to this *topos*, despite having Stephen's *Life of Wilfrid* as a source. For further discussion of Bede's nuanced approach to female characters in the *Historia*, cf. MacCarron 2017; Prado et al. 2020.

17 For an introduction to Maria Edgeworth's life and works, cf. Ó Gallchóir 2021.

18 There are also small collections of Edgeworth papers elsewhere, for example, Archives at Yale.

19 *DEN* was jointly funded by the Irish Research Council and the Arts and Humanities Research Council in the UK as part of the UK-Ireland Collaboration in Digital Humanities Networking programme: https://research.ie/2020/07/27/12-new-uk-ire-land-digital-humanities-collaborations-announced/. The team was Clíona Ó Gallchóir and Máirín MacCarron (Cork), Ros Ballaster and Anna Senkiw (Oxford). The project ran from 1 August 2020 to 30 November 2021.

family correspondence, which encapsulates both the potential and challenges posed by the entire collection. The Edgeworth correspondence extends from the 1760s to the 1850s and comprises about 10% of the holdings in both libraries. It includes letters to and from multiple members of the Edgeworth family including cousins and in-laws. It is generally accessed through two different avenues. The letters in the Bodleian are listed in a typescript document dating from the 1970s which catalogues the correspondence by sender and is known as the 'Colvin Calendar'. It was created by Christina Colvin, a member of the family, and covers 2,754 items.[20] The letters in the National Library of Ireland (NLI) are catalogued in 'NLI Collection list 40'.[21] The NLI holds just over one-third of the total correspondence with 1,409 items. 'Collection list 40' organises the letters by date irrespective of sender. Consequently, the researcher who wishes to access the Edgeworth correspondence is faced with two distinct collections, which are organised differently: one is catalogued by the sender's name, and the other by the date that letters were sent. In addition to the archive being divided, this also means that it is not possible to examine the entire collection either chronologically or by sender, and both catalogues have gaps. A further challenge is that, because comparatively little of the correspondence has been digitised, a researcher who wants to understand the correspondence has to start with these catalogues. Consequently, both the logistical challenge and time commitment are significant.

In *DEN*, we attempted to address these problems by working with both libraries in an effort to harmonise their catalogues, and by engaging in digital analysis of both catalogues together. Our intended purpose was twofold: firstly, we wanted to demonstrate that the collection could be treated as a whole, though working on the catalogues from both libraries rather than the actual holdings; and secondly, we applied social network analysis to the dataset we created to determine the efficacy of applying digital approaches to a divided archive. As this was a correspondence network, the data model was simpler than that developed for *Gendered Networks*. We created a master spreadsheet of the correspondence which contained data on sender, recipient, date, location of sender, location of recipient, and recorded the gender of sender and recipient. The data creation was a laborious process, but we successfully harmonised the data, in so far as was possible, and highlighted gaps in the catalogues: for

---

20 The 'Colvin Calendar' has been digitised and can be accessed on the Bodleian Libraries website: https://archives.bodleian.ox.ac.uk/repositories/2/archival_objects/64160.

21 'NLI Collection list 40' is available as a PDF on the NLI website: https://www.nli.ie/sites/default/files/2022-12/040_edgeworth.pdf.

example, information relating to senders or recipients is occasionally missing, and location data was frequently omitted. A further complication was that some letters had two or more recipients, and others had two or more senders; although the existence of multiple recipients and senders is not very surprising when assessing a family's interactions, it posed a further challenge when attempting to reduce the complexities of the correspondence to usable data.

The complete spreadsheet contains 4,544 rows, and we recorded far more data than we have, as yet, been able to use. For example, in addition to the obvious sender, recipient, and location information, we also recorded information on other people named in the letters, additional locations, literary works sent with letters (often the case with Maria Edgeworth), and literary works mentioned in letters (an even more common occurrence). Although the process of data gathering, carried out by Anna Senkiw, was painstaking and time consuming, it demonstrated the feasibility of turning this divided archive into a usable dataset that allowed us to engage with and visualise the Edgeworth correspondence in an entirely new way.

Figure 3 shows everyone in the network who sent or received at least one letter [fig. 3]. There are 359 nodes and 545 connections, and the size of each node indicates the number of letters sent or received.[22] The biggest nodes are Maria Edgeworth, and her father, Richard Lovell Edgeworth (both nodes are pink), who dominate the correspondence and is what one would expect as the most prominent members of the family. However, the next most connected members of the correspondence network (green nodes) joined the family at different times: Frances Ann Edgeworth née Beaufort was Richard Lovell's fourth wife, with whom she had six children, and she was a year younger than her stepdaughter, Maria; Charles Sneyd was Richard Lovell's son and twelfth child. The importance of these two members in the network, especially Frances Ann, is somewhat surprising and would have been more difficult to assess using traditional approaches.

Figure 4 shows a subset of the same network, but only represents everyone who appears at least twice in the collection [fig. 4]. Consequently, it is more visually accessible. The same four figures dominate: Maria, Richard Lovell, Frances Ann, and Charles Sneyd. Maria and Richard Lovell are green nodes; Frances Ann and Charles Sneyd are grey-blue; the remaining nodes are dark blue.

In the next stage of our analysis, we restricted the correspondence network to that of the immediate Edgeworth family: specifically Richard Lovell, his wives and children (he had four wives and twenty-two children), his siblings and their families, creating a network

---

22   The graphs in figures 3 and 4 were drawn by Máirín MacCarron using Gephi 0.9.2.
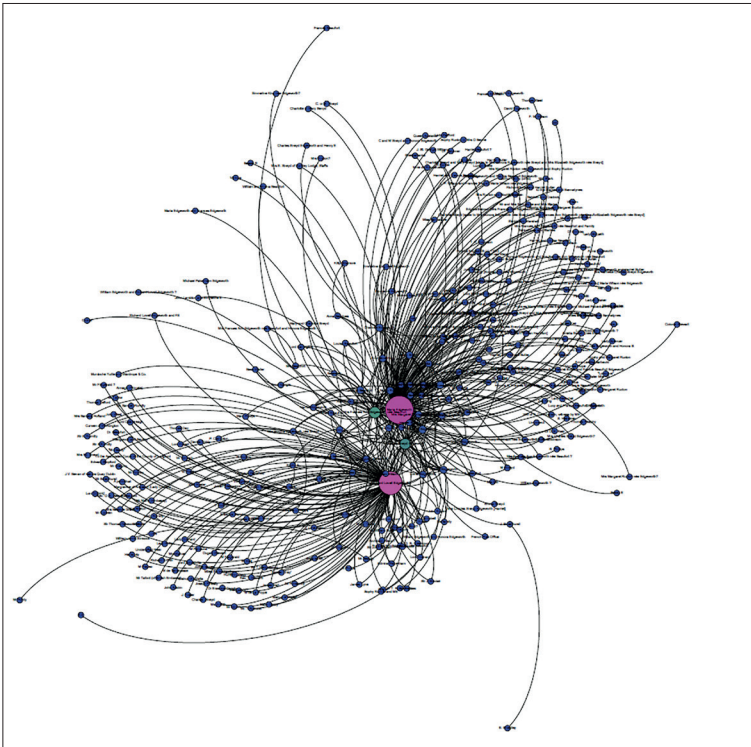
Figure 3   Network of the Edgeworth family correspondence drawn from the 'Colvin Calendar' and 'NLI Collection List 40'

of 38 nodes (Senkiw 2024). This analysis revealed the predominance of women's voices in the network; indeed, women outnumber men by just over two to one (there are 26 women and 12 men). Such an insight can be missed when focussing on the most important figures in a network, rather than a network as a whole. Further analysis also revealed that half of the correspondents in this network were children when they began writing to family members and their letters were preserved alongside those of the adults. Several of these children also appear later in the network as adults.

These networks represent the affordances of applying a data-driven approach to a complex archive such as the Edgeworth papers. Our findings, such as the predominance of children, would have been difficult to identify using other methods because of the limitations of the catalogues. Indeed, we did not set out to look for children, but as our analysis developed their presence became apparent. In attempting to put order on what we were finding, we devised new categories
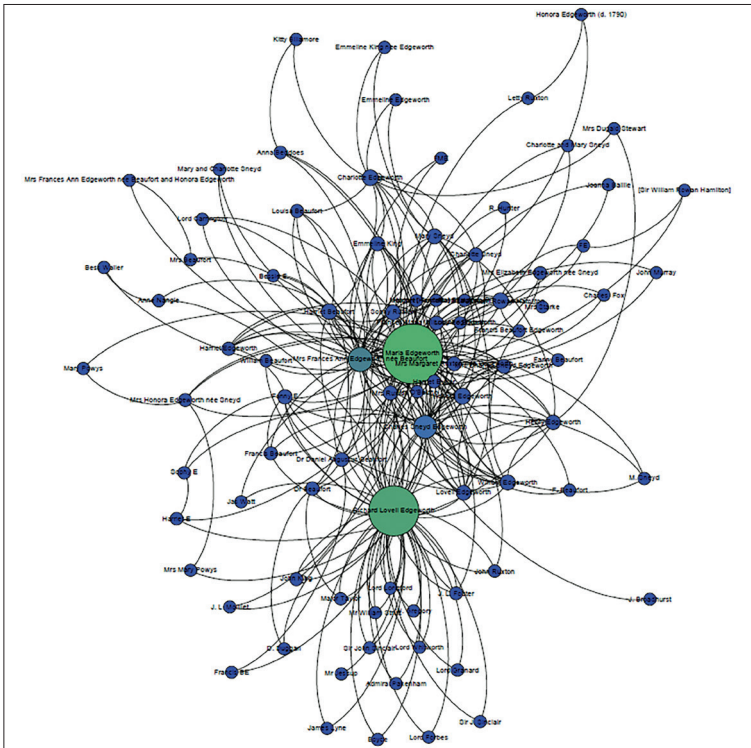
Figure 4 Network of everyone who appears at least twice in the Edgeworth family correspondence as represented in the 'Colvin Calendar' and 'NLI Collection List 40'

to account for the variety in the dataset. We have shared our findings with our library partners, and they are exploring the possibility of including a category for children in subsequent metadata. Our experience indicates, among other things, that when one begins extracting data from historical sources, the very process of extraction reveals types of data that one was not expecting to find. Occasionally, this is data that might previously have been considered invisible.

The *Gendered Networks* and *DEN* projects applied data-driven approaches and digital analysis to material that might not at first seem entirely suitable for such methodologies. Both projects engaged in researcher-led close reading to create their complex and multi-faceted datasets. This was a time-heavy commitment, but the process of data creation allowed us to see aspects of the source material that had hitherto received little attention, which led to the development of new research questions, including the role of nameless characters in narrative texts, and the presence of children in the Edgeworth

correspondence collection. The data-driven approach also ensured that all individuals were included in our data gathering, regardless of their perceived importance, and allowed us to bring to the fore people who have often been overlooked.

Bibliography

Bede (1969). *Bede's Ecclesiastical History of the English People*. Ed. and transl. by B. Colgrave; R.A.B. Mynors. Oxford: Clarendon press. Transl. of: *Historia ecclesiastica gentis anglorum*.

Colgrave, B. (1927). *The Life of Bishop Wilfrid by Eddius Stephanus*. Cambridge: Cambridge University Press.

Cvetkovich, A. et al. (2010). "Women as the Sponsoring Category: A Forum on Academic Feminism and British Women's Writing". *Partial Answers*, 8(2), 235-54.
https://muse-jhu-edu.ucc.idm.oclc.org/article/382600

Darby, P.; MacCarron, M. (eds) (2023). *Bede the Scholar*. Manchester: Manchester University Press.

D'Ignazio, C.; Klein, L.F. (2020). *Data Feminism*. Cambridge (MA): MIT Press.
https://direct.mit.edu/books/oa-monograph/4660/Data-Feminism

Drucker, J. (2011). "Humanities Approaches to Graphical Display". *Digital Humanities Quarterly*, 5(1).
https://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html

Hillner, J.; MacCarron, M. (2021). "Female Networks and Exiled Bishops between Late Antiquity and the Early Middle Ages: The Cases of Liberius of Rome and Wilfrid of York". Bérat, E.O.; Hardie, R.; Dumitrescu, I. (eds), *Relations of Power: Women's Networks in the Middle Ages*. Bonn: Bonn University Press, 19-44.

Hillner, J.; MacCarron, M. (forthcoming). "Social Network Analysis". Flower, R. et al. (eds), *A Companion to Roman Prosopography*. Leiden: Brill.

Hillner, J.; MacCarron, M.; Vihervalli, U. (2022). "The Politics of Female Namelessness between Late Antiquity and the Early Middle Ages, circa 300 to 750". *Journal of Late Antiquity*, 15(2), 367-401.
https://doi.org/10.1353/jla.2022.0021

Hitchcock, T. (2013). "Confronting the Digital: Or How Academic History Writing Lost the Plot". *Cultural and Social History*, 10(1), 9-23.
https://doi.org/10.2752/147800413X13515292098070

Hollis, S. (1992). *Anglo-Saxon Women and the Church: Sharing a Common Fate*. Woodbridge: Boydell.

Klein, L. (2018). "Distant Reading after Moretti". *Arcade: Literature, the Humanities and the World Blog*, 29 January.
https://arcade.stanford.edu/blogs/distant-reading-after-moretti

Lemercier, C.; Zalc, C. (2019). *Quantitative Methods in the Humanities: An Introduction*. Transl. by A. Goldhammer. Charlottesville : University of Virginia Press. Transl. of: *Méthodes quantitatives pour l'historien*. Paris: La Découverte, 2008.

MacCarron, M. (2017). "Royal Marriage and Conversion in Bede's *Historia ecclesiastica gentis anglorum*". *Journal of Theological Studies*, 68(2), 650-70.
https://doi.org/10.1093/jts/flx126

Mandel, L. (2016). "Gendering Digital Literary History: What Counts for Digital Humanities". Schreibman, S.; Siemens, R.; Unsworth, J. (eds), *A New Companion to Digital Humanities*. Oxford: Blackwell, 511-23.
http://library.ucc.ie/record=b2238281

Nelson, J. (1986). "Queens as Jezebels: the Careers of Brunhild and Baldhild in Merovingian History". Nelson, J. (ed.), *Politics and Ritual in Early Medieval Europe*. London: Hambledon, 1-48.

Ó Gallchóir, C. (2021). *Maria Edgeworth*. Brighton: Edward Everett Root.

Prado, S. et al. (2020). "Gendered Networks and Communicability in Medieval Historical Narratives". *Advances in Complex Systems*, 23(3).
https://doi.org/10.1142/S021952592050006X

Rosenthal, J.T. (1966). "Marriage and the Blood Feud in 'Heroic' Europe". *The British Journal of Sociology*, 17(1), 133-44.

Schöch, Ch. (2013). "Big? Smart? Clean? Messy? Data in the Humanities". *Journal of Digital Humanities*, 2(3).
https://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/

Senkiw, A. (2024). "'Much Material of 1821 Not Listed': Troubling Colvin's Calendar of Maria Edgeworth's Correspondence with Digital Analysis". *Digital Studies/Le champ numérique*, 14(1), 1-26.
https://doi.org/10.16995/dscn.9865

Stancliffe, C. (2013). "Dating Wilfrid's Death and Stephen's Life". Higham, N.J. (ed.), *Wilfrid: Abbot, Bishop, Saint: Papers from the 1300th Anniversary Conferences*. Donington: Shaun Tyas, 17-26.

Stephen (1927). *The Life of Bishop Wilfrid by Eddius Stephanus*. Ed. and transl. by B. Colgrave. Cambridge: Cambridge University Press. Transl. of: *Vita Wilfridi*.

Thacker, A. (2013). "Wilfrid: His Cult and His Biographer". Higham, N.J. (ed.), *Wilfrid: Abbot, Bishop, Saint: Papers from the 1300th Anniversary Conferences*. Donington: Shaun Tyas, 1-16.