

# Sagas and genre: A case for application of network analysis to manuscripts preserving Old Norse-Icelandic saga literature

Katarzyna Anna Kapitan <sup>1\*</sup>, Tarrin Wills <sup>2</sup>

<sup>1</sup>Linacre College and Faculty of English, University of Oxford, UK

<sup>2</sup>Dictionary of Old Norse Prose, University of Copenhagen, Denmark

\*Correspondence: Katarzyna Anna Kapitan. E-mail: [katarzyna.kapitan@linacre.ox.ac.uk](mailto:katarzyna.kapitan@linacre.ox.ac.uk)

## Abstract

This study applies statistical approaches to the analysis of the genre relationships of Old Norse-Icelandic literature in order to expand our understanding of the relationships between works, their transmission, and their possible modes of reception, as manifested in the extant manuscripts. This article contributes to the ongoing discussion of the genre boundaries of Old Norse-Icelandic literature and presents an alternative method of engaging with this material in the form of computer-assisted analysis, i.e. data visualization and network analysis. Using data collected from major online databases of Old Norse-Icelandic manuscripts, we present the most complete to date network of co-occurrences in manuscripts of works belonging to a number of literary genres. The present study empirically demonstrates the manifoldness of the connections between the Old Norse-Icelandic works which transcend traditional scholarly genre boundaries. The study identifies two main communities within the network: a community of romances, or works of narrative fiction, which includes mainly legendary sagas (*fornaldarsögur*) and chivalric sagas (*riðdarasögur*), and a community of historicizing narratives, or pseudo-history, which includes mainly sagas of Icelanders (*Íslendingasögur*) and kings' sagas (*konungasögur*).

## 1 Introduction

For decades leading scholars in the field of Old Norse-Icelandic studies have been pointing to the manuscript context as a key for understanding the genre classification of Old Norse-Icelandic literature. The frequent co-occurrence of certain works in manuscripts is interpreted as evidence for close associations of these works in regard to either their production, reception, or both.<sup>1</sup> Not infrequently, however, scholars arrive at contradicting conclusions regarding the genre boundaries of Old Norse-Icelandic literature when analysing manuscripts that preserve this literature. These contradicting conclusions, discussed in detail in Section 2.1, may be a result of the limitation of manual data collection, processing, and analysis. The analysis and interpretation of the manuscript context in which literary works appear require processing of a significant amount of data, which arguably lies outside the ability of a single researcher using traditional hermeneutic methods. According to the recent estimates, around 10,000 Icelandic manuscripts survive until today, and

roughly 900 of them can be dated to the Middle Ages (Jónsson, 2021). While the proportion of manuscripts containing sagas among these codices is unknown, there have been attempts to list all manuscripts containing works belonging to various genres of Old Norse-Icelandic literature, for example, chivalric sagas, *riðdarasögur* (Kalinke and Mitchell, 1985) and legendary sagas, *fornaldarsögur* (Driscoll, 2018), each listing hundreds of codices. For comparison, Stephen Mitchell (1991) exemplifies his argument about manuscript context of the legendary sagas by references to three manuscripts, while Ármann Jakobsson (2012) mentions around twenty medieval manuscripts, but discusses very briefly only a selected handful of the mentioned codices.

Taking a point of departure in the disagreement among scholars regarding the genre boundaries, and considering the great number of extant manuscripts which preserve Old Norse-Icelandic sagas, we have turned towards computational methods of data gathering and visualization to contribute new insights into

Our main hypothesis is that the manuscript context of selected texts should reflect their genre affiliation. Explicitly, we expect that a text belonging to a particular genre should appear most frequently with other representatives of the same genre. So, if a borderline-case saga appears most frequently with certain types of texts in extant manuscripts, we can assume that the community which produced and preserved these manuscripts considered these co-occurring texts as generically related to each other. By application of network analysis we hope to gain insight into relationships between texts in extant manuscripts and problematize the validity of analytical generic distinctions existing in Old Norse-Icelandic studies.

This is especially clear when the legendary sagas (*forðaldarsögur*) and chivalric sagas (*riddarasögur*) are discussed. Traditionally, the thematic division establishes the boundaries between the legendary sagas, which are set in the North, and the chivalric sagas are set outside the North on courts of exotic kings and rulers. While this thematic distinction is not unproblematic in itself, for example, because of various exotic travels that appear in legendary sagas, it becomes even more complicated when codicological evidence is taken into account, leading scholars to contradicting conclusions. According to [Mitchell \(1991, p. 21\)](#), ‘the notion that the *forðaldarsögur* represented a particular category of saga in the minds of medieval Icelanders is suggested by the codicological evidence, for the *forðaldarsögur*

On the other hand, the previous scholarship has shown that some other works which are traditionally considered as members of the *fornaldarsögur* group upon close examination appear to have much in common with some *riddarasögur*, so romances or chivalric literature (Kalinke, 1990). This generic fluidity of *riddarasögur* and *fornaldarsögur* in the Middle Ages has been confirmed through application of network analysis to medieval manuscripts catalogued in *Handrit.is* conducted by Blobel. Blobel's study has demonstrated, that in the analysed pre-1540 manuscripts catalogued in *Handrit.is*, *riddarasögur* and *fornaldarsögur* can be grouped together with *Íslendingasögur* into one category of 'secular narratives' (Blobel, 2015, p. 33).

The notion of the instability of manuscripts as artefacts is also an important factor not only in ours, but also in any dataset that analyses manuscripts in their current form. As [Stegmann \(2016\)](#) has recently demonstrated, Icelandic manuscripts did not necessarily keep their original form throughout centuries, as, for example, around 53% of paper manuscripts in the Arnamagnæan Collection (hosted today in Copenhagen and Reykjavik) were rearranged at least once. Therefore our interpretation of the context in which literary texts appear today is more likely to give

omits a very large number of later manuscripts and information about poetic works. It nevertheless provides a valuable supplement to *Handrit* with hundreds of manuscripts from collections not covered by *Handrit*.

Another very reliable source of data was the digital catalogue of the *Stories for all time* project (henceforth *FASNL*), which is a thematic catalogue devoted exclusively to the manuscripts preserving the legendary sagas.<sup>6</sup> At the time of our experiments, it contained XML-based descriptions of 818 manuscripts held at twenty-three libraries in Europe and North America. The descriptions are conformant to the Text Encoding Initiative Guidelines. The catalogue uses the authority list for the legendary sagas, but other works are catalogued following more or less the same principles as *Handrit*. Texts that can be found in the catalogued manuscripts are encoded with `msItem` elements, which contain uniform titles. Each `msItem` has a class attribute that refers to a given text's genre affiliation. Descriptions of manuscripts held at the three partner institutions of *Handrit* were initially based on the *Handrit* descriptions and then adjusted, standardized, and expanded for the purposes of the project. Therefore these files were compatible with the same processing workflow as *Handrit* but needed to be resolved and given priority where overlapping with *Handrit*.

Finally, the database of the *Skaldic Project* (skaldic.org), provided the framework for the initial handling of the data used in the present study. The Skaldic database focuses on Old Norse-Icelandic poetic corpus and its relationship to the prose corpus. Its structure was compatible with the project's aims of aggregating and analysing data from various sources and its flexible web application could be extended to assist in the cleaning of the data, therefore it was introduced as an intermediate step in preparation of the final dataset. Queries written in structured query language (SQL) could build and analyse the hundreds of thousands of potential connections between texts and manuscripts so that the authors could quickly identify issues with the imported data. This process of data cleaning is further discussed in the following section.

## 2.3 Data cleaning

In order to prepare our data for network visualization which would illustrate the genre affiliation of various works as represented in the catalogues together with the co-occurrence of these texts in the manuscripts we needed two tables. The first one was a simple list of works (or ‘uniform titles’), which will be treated as nodes in the network, and attributes that reflects their genre affiliation (‘class’), so the nodes can be sorted accordingly to their genre. The second one was a table of edges, in which each work was connected to other



works it co-occurs with in the manuscripts, and the number of manuscripts that this co-occurrence appears was reflected as weight of the connection. So if two works appear together only in a single manuscript, the edge's weight equals to 1.

In order to generate such tables automatically from the Skaldic Project database, we needed to resolve number of issues with the heterogenous dataset we have created. Data from the *Dictionary of Old Norse Prose*, which had previously been imported some 10 years earlier, were updated with the latest information from the Index. This provided a consistent reference point for the cleaning of further data. The *Handrit* and *FASNL* data were first imported as raw XML, prioritizing the *FASNL* data when overlaps occurred. They were then processed using the libxml library to find the relevant information: siglum/shelfmark, dating, and location. The manuscript shelfmarks were normalized automatically to the same format as those of the existing dataset. The resulting shelfmark was used to search for the same manuscript in the existing dataset and update the information accordingly, or where not found, to create a new record for that manuscript.

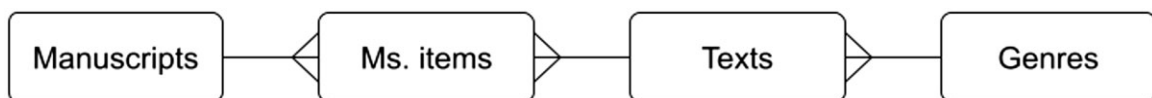
A major challenge with the data imported from *Handrit* was that the manuscript contents described in the XML files do not refer to an external authority list, so we had to rely on the ‘uniform title’ encoded in the XML files to find different manuscripts of the same work. The uniform titles, however, are quite often far from uniform, with various spelling variations and versions of the same title. These were imported into a separate table and a series of automated normalizations applied. Manual normalization was performed through the web interface, which involved updating and checking the links to the skaldic database’s table of text titles (Fig. 2). The resulting table of normalizations can be used for future updates from these and other sources.

Finally, the genre classification of the works catalogued in *Handrit* manuscripts, which is represented as a class attribute on `msItem` elements, are highly heterogenous. Not infrequently, a single work can be assigned different classes in different manuscript descriptions, or contain more than one value. This means that the same work could be classified as a *for-naldarsaga* in one manuscript description, but as an *Íslendingasaga* in another, or as both in yet another. We have made an editorial decision that each work can

belong only to a single genre (have only one value of the ‘class’ attribute). This allowed us to assign the most frequently assigned genre (class value) for a work consistently across the corpus. The genre classifications in the FASNL files had been manually cleaned up and standardized by Timothy Rowbotham, who checked all the works that had more than one class attributed in the initial dataset. Clean classes were then applied to the table of texts, resolving any remaining inconsistencies deriving from varying classification within the datasets originating from *Handrit* and FASNL. The cleaned data were then imported into the original structure of the skaldic project, resulting in a data model as shown in Fig. 1, indicating one-to-many relationships. From this structure we needed to extract a table of nodes and a table of edges in order to import into the Gephi software (see section 2.4).

An SQL query first extracted the text nodes into a table, limiting the genre designation to the types which lie at the interest of the present study. These are sagas of Icelanders and related short stories, kings' sagas and related short stories, legendary sagas and related short stories, and romances. Eddic poems, *rímur*, and legal works played a role of a control group.<sup>7</sup> Only works classified according to these categories were included in the resulting dataset. The edges table was generated by linking each row in the text table—via the manuscript item table—to all manuscripts containing the text. Then each of these manuscripts was linked to all of the other texts found in the manuscripts table. The results were grouped according to pairs of nodes, with the number of manuscripts containing both texts forming the weighting for each edge.

The fast results allowed us to identify and resolve some issues with the result set. Some duplicate results were found due to incorrect resolution from the different data sources and instances where one work appeared more than once because it was broken up by other works. In cases when two texts were catalogued as a single work (or single msItem), for example, the two closely related legendary sagas, *Völsunga saga* and *Ragnars saga loðbrokar*, they were excluded from the dataset, but they were included when catalogued separately. Skaldic poems were excluded, even when appearing in the texts that lie within the scope of this project, as these are mainly recorded within the other works and not as individual items. Keeping our research questions in mind, we excluded compilations of



**Figure 1.** Relational model of texts and manuscripts used by the network analysis









literature. There seems to be scholarly disagreement regarding genre affiliation of this text. In the *Lexikon der altnordischen Literatur*, the entry for *Þorsteins saga Geirnefufóstra* reads as follows: ‘Þorsteins saga Geirnefufóstra ist eine vermutl. erst im 18. Jh. verfaßte Saga im Stil der Fornaldarsögur’ (Þorsteins saga Geirnefufóstra is a saga presumably composed in the style of a fornaldarsaga at the earliest in the eighteen century, *Simek and Pálsson*, 2007, p. 391). This suggests that the story was most likely composed no earlier than in the eighteenth century and that it is in style of legendary sagas. Conversely, Guðni Jónsson writes in his introduction to the edition of *Íslendingasögur* that the saga was written in the nineteenth century by a known author: ‘Þorsteins saga Geirnefufóstra gerist á 10. öld að miklum hluta í Noregi, Grænlandi og víðar, en Fljótamenn í Skagafirði koma þar við sögu. [...] sagan, sem nú er til og hér prentuð í fyrsta sinn, er samin af Gísla Konráðssyni snemma á 19. Öld’ (The events of Þorsteins saga Geirnefufóstra take place mostly in Norway, Greenland and in other places in the tenth century, but the men of Fljót in Skagafjörður are also mentioned in the saga [...] the saga that exists today and is printed here for the first time was composed by Gísli Konráðsson early in the nineteenth century). So the contents of the saga fit the bill for *Íslendingasögur*, but its style is more *fornaldarsaga*-like. The position of the story in the network indicates that the communities producing and circulating manuscripts of *Þorsteins saga Geirnefufóstra* considered it more similar to *Íslendingasögur* and *konungasögur*-related materials, than to *fornaldarsögur*, and therefore put it together with these type of texts in manuscripts. This also speaks in favour of Guðni Jónsson’s classification of this text as a *Íslendingasaga*.

This section illustrated that already simple visual observation of the network, in which pre-existing attributes determined colouring of the nodes, proves itself useful in exploration of relationships between texts and allows us to ask further questions. However, the strength of network analysis does not lie in hermeneutic observations of the network, but in its potential for statistical explorations, which are the subject of the following section.

### 3.2 Network and statistics

In the opening paragraph of the previous section, we have suggested that the multiple cross-genre edges visible with a bare eye reveal that the analytical categories applied in the modern catalogues, such as late legendary sagas or late chivalric sagas, do not correspond to the ethnic genre classifications, as represented by the extant manuscripts. In order to evaluate this statement, it is necessary to apply computational methods of

network assessment, which allows us to understand the topography of the network. One of the methods is the application of modularity statistic, ‘which attempts to assess the number of distinct groupings within a network’ (Cherven, 2015, p. 189). Using Gephi’s in-built function ‘Modularity’ (with the following parameters: ‘Randomize: On, Use edge weights: On, Resolution: 1.5’), we identified fourteen communities within the network.<sup>11</sup> Figure 4a–c presents three visualizations of the network, generated with different layouts. Here, unlike in the previous visualizations on Fig. 3, nodes are coloured by their modularity class.<sup>12</sup>

Even though fourteen communities have been distinguished in the network, modularity Classes 2, 3 (bright green nodes visible on Fig. 4b and c), 5, and 13 are all small communities of legal texts, while modularity Classes 6–12 are all small communities of *rimur*. Thus, there are three large communities in our network:

- 1) Modularity Class 0 (magenta nodes on Fig. 4a–c) consisting of texts classified mainly as *fornaldarsögur* (and *fornaldarsögur síðari tíma*) and *riddarasögur* (and *riddarasögur síðari tíma*), but also two *konungasögur*, two *konungapættir*, one set of *rímur*, and one *Íslendingasaga*;
- 2) Modularity Class 1 (cyan nodes on Fig. 4a–c) consisting of Eddic texts and one legendary saga;
- 3) Modularity Class 4 (white nodes on Fig. 4a–c) consisting of mainly *Íslendingasögur* and *pættir*, *konungasögur* and *pættir*, but also some *fornaldarsögur* and *pættir*, three sets of *rímur*, and one *riddarasaga*.

The first community (modularity Class 0) is an intriguing one, consisting of 120 works across different genres of Old Norse literature, presented in [Supplementary Materials](#) (nodes\_modularity). While *fornaldarsögur* and *fornaldarsöguir síðari tíma* together with *riddarasögur* and *riddarasöguir síðari tíma* dominate in this group, there are couple of exceptions. One text classified as *Íslendingasaga* appears in this group, *Ármanns saga og Dalmanns*. Even though, *Ármanns saga* is included in Guðni Jónsson's edition of *Íslendingasögur*, it is worth to emphasize, that it is a nineteenth century creation. Just as in the aforementioned case of *Þorsteins saga Greinefufóstra*, the transmission context of *Ármanns saga og Dalmanns* indicates that it was treated in a similar way as most of legendary sagas and romances. In a similar manner, we should interpret the appearance of *konungasögur* and *konungabættir* in this group (*Jómsvíkinga saga*, *Ynglinga saga*, *Hákonar þáttur Hárekssonar*, and *Hrólfs þáttur*). Finally, the single set of *rímur* that appears in this group, *Gríms rímur og Hjálmars*, is the













- Jónsson, M. (2021). Hversu mörg islensk handrit eru varðveitt í útlöndum? Vísindavefurinn blog post. <https://www.visindaufur.is/svar.php?id=81724> (accessed 1 February 2022).
- Kalinke, M. (1990). *Bridal-Quest Romance in Medieval Iceland*. Ithaca and London: Cornell University Press.
- Kalinke, M. and Mitchell, P. M (1985). *Bibliography of Old Norse-Icelandic Romances*. Ithaca, NY: Cornell University Press.
- Kapitan, K. A. (2022). When a king of Norway became a king of Russia: Danish historiography and early transmission and reception of Hrómundar saga Greipssonar. *Scandinavian Studies*, 94(3): 316–351.
- Kapitan, K. A. (2021a). Perspectives on digital catalogs and textual networks of old norse literature. *Manuscript Studies: A Journal of the Schoenberg Institute for Manuscript Studies*, 6(1): 74–97.
- Kapitan, K. A. (2021b). Manuscripts derived from printed editions in the transmission history of Hrómundar saga Greipssonar. In Driscoll M. J. and Mac Cathmhaoil N. (eds), *Hidden Harmonies: Manuscript and Print on the North Atlantic Fringe, 1500–1900*. Copenhagen: Museum Tusulanum, pp. 79–114.
- Kapitan, K. A., Rowbotham, T., and Wills, T. (2017). Visualising genre relationships in Icelandic manuscripts. In *Digital Humanities in the Nordic Countries, Second Conference: Conference Abstracts*. Gothenburg: The University of Gothenburg. [https://web.archive.org/web/20180613002357/http://dhn2017.eu/abstracts/#\\_Toc475332531](https://web.archive.org/web/20180613002357/http://dhn2017.eu/abstracts/#_Toc475332531) (accessed 1 February 2022).
- Lavender, P. (2015). The secret prehistory of the Fornaldarsögur. *The Journal of English and Germanic Philology*, 114(4): 526–51.
- Love, J. (2013). *The Reception of Hervarar saga ok Heiðreks from the Middle Ages to the Seventeenth Century*. Munich: Herbert Utz Verlag.
- Lönnroth, L. (1975). The concept of genre in saga literature. *Scandinavian Studies*, 47(4): 419–26.
- Martin, S., Brown, W. M., Klawns, R., and Boyack, K. (2011). OpenOrd: an open-source toolbox for large graph layout. In *Proceedings SPIE 7868, Visualization and Data Analysis 2011*. <https://doi.org/10.1117/12.871402> (accessed 2 February 2022)
- McTurk, R. (ed) (2005). *A Companion to Old Norse-Icelandic Literature and Culture*. Oxford: Blackwell.
- Mitchell, S. (1991). *Heroic Sagas and Ballads*. Ithaca, NY and London: Cornell University Press.
- Rafn, C. C. (ed) (1829–30). *Fornaldar sögur Nordrlanda eptir gömlum handritum*, 3 vols. Copenhagen: Popp.
- Riva, G. F. (2019). Network analysis of medieval manuscript transmission. *Basic principles and methods. Journal of Historical Network Research*, 3: 30–49.
- Simek, R. and Pálsson, H. (2007). *Lexikon der altnordischen Literatur, die mittelalterliche Literatur Norwegens und Islands*, 2nd edn. Stuttgart: Kröner.
- Skovgaard-Petersen, K. (1993). The literary Feud between Denmark and Sweden in the sixteenth and seventeenth century and the development of Danish historical scholarship. In Brink, J. R. and Gentrup, W. F. (eds), *Renaissance Culture in Context. Theory and Practice*. Aldershot: Scholar Press, pp. 114–20.
- Stegmann, B. (2016). *Árni Magnússon's Rearrangement of Paper Manuscripts*. Ph.D. thesis, University of Copenhagen.
- Sverrisdóttir, I. S., Sigurðsson K., and Hrafnkelsson, Ö. (2012). Access and curation of digital cultural heritage in the national and University Library of Iceland. *Microform & Digitization Review*, 41: 97–102.
- TEI Consortium (2020). *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. 4.0.0. <http://www.tei-c.org/Guidelines/P5/> (accessed 1 February 2022).
- Quinn, J. (2006). Interrogating genre in the Fornaldarsögur: round-table discussion. *Viking and Medieval Scandinavia*, 2: 275–96.
- Wills, T. (2022). Bergþúa Þáttur. In Clunies Ross, M., Gade, K. E., and Wills T. (eds), *Poetry in Sagas of Icelanders*. Turnhout: Brepols, pp. 31–55.