



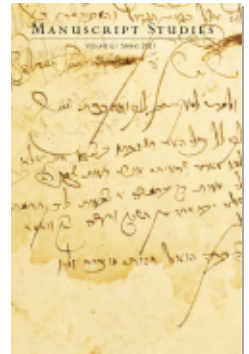
PROJECT MUSE®

---

## Perspectives on Digital Catalogs and Textual Networks of Old Norse Literature

Katarzyna Anna Kapitan

Manuscript Studies: A Journal of the Schoenberg Institute for Manuscript Studies, Volume 6, Number 1, Spring 2021, pp. 74-97 (Article)



Published by University of Pennsylvania Press

DOI: <https://doi.org/10.1353/mns.2021.0002>

➔ *For additional information about this article*

<https://muse.jhu.edu/article/794855>

# Perspectives on Digital Catalogs and Textual Networks of Old Norse Literature

---

KATARZYNA ANNA KAPITAN  
*University of Iceland*

THE DIGITAL TURN OF the twentieth century—with the world wide web, personal computers, and rapid developments of digital tools and methods—changed the way we collect, store, and analyze information.<sup>1</sup> This is, however, only partially true when it comes to the principles

---

This article draws on the research conducted during my doctoral fellowship at the Department of Nordic Studies and Linguistics, University of Copenhagen (2015–2018), and a research fellowship from the Arnamagnæan Commission (2018–2019). The writing of this article was made possible thanks to a postdoctoral fellowship from the Carlsberg Foundation, *Dronning Margrethe den II's "Distinguished" forskningsprojekt om den dansk-islandske reception af den nordiske oldtid* (2019–2021). The contents of this article are directly based on two presentations I delivered in 2019: Katarzyna Anna Kapitan, “Digital Cataloguing of Manuscripts as Artefacts and Quantitative Analysis of Manuscript Descriptions,” presented at the 94th annual meeting of the Medieval Academy of America, Philadelphia, March 2019; Katarzyna Anna Kapitan, “Digital Catalogs and Analog Data: Cataloging Nordic Manuscripts,” presented at the 12th annual Schoenberg Symposium on Manuscript Studies in the Digital Age, Philadelphia, November 2019. I am grateful to many colleagues for their inspiration, feedback, and support, including N. Kivilcim Yavuz, Matthew James Driscoll, Patrik Granholm, Tarrin Wills, Anne Mette Hansen, and Lynn Ransom. I owe thanks also to Ryder Patzuk-Russell for his feedback on the language and style of the present article.

1 The digital turn’s full influence on our lives is extremely broad, and impossible to address here. For further readings on the role of digitization in the humanities, see Susan Schreibman, Ray Siemens, and John Unsworth, eds., *A Companion to Digital Humanities* (Oxford: Blackwell, 2004).

governing the cataloging of medieval and early modern manuscripts, and arguably also when it comes to the way we interact with the resulting catalogs. The history of cataloging is almost as old as writing itself, extending from the clay tablets of Nipur and Hattusas, through the antique catalogs of scrolls and medieval lists of books, to the printed and digital catalogs we create today.<sup>2</sup> The ultimate objective of a catalog of manuscripts is to locate books within a collection and to identify the texts they preserve. Not surprisingly, then, the focus of many manuscript catalogs even today is largely textual, and some institutions use the same cataloging methods for both printed books and for manuscripts. In particular, large repositories in the United States use MARC to catalog their manuscripts and rare books, while some European institutions increasingly use XML-TEI for the purpose of cataloging manuscripts.<sup>3</sup>

The presence of digital catalogs, some of which have recently celebrated their twentieth birthdays, invites researchers to apply digital tools and methods for the extraction and analysis of information in order to ask

---

2 Diachronic overviews of cataloging practices can be found in, for example, Lionel Casson, *Libraries in the Ancient World* (New Haven: Yale University Press, 2001); Debra Cashion, "Cataloging Medieval Manuscripts, from Beasts to Bytes," *Digital Philology: A Journal of Medieval Cultures* 5, no. 2 (2016): 135–59.

3 Hope Mayo, "Medieval Manuscript Cataloging and the MARC Format," *Rare Books and Manuscripts Librarianship* 6, no. 1 (1991): 11–22; Alexandra Mason, "Using MARC to Catalog Medieval Manuscripts: Comments," *Rare Books and Manuscripts Librarianship* 6, no. 1 (1991): 53–58; Gregory A. Pass, *Descriptive Cataloging of Ancient, Medieval, Renaissance, and Early Modern Manuscripts* (Chicago: Association of College and Research Libraries, 2003); Benjamin J. Fleming, "Digitizing Penn's Indic Manuscripts," *Manuscript Studies* 3, no. 2 (2018): 470–86. TEI is used, for example, at the Bodleian Library (<https://medieval.bodleian.ox.ac.uk/about>). See Lou Burnard, Richard Gartner, and Peter Kidd, "The Cataloguing of Western Medieval Manuscripts in the Bodleian Library: A TEI Approach with an Appendix Describing a TEI-Conformant Manuscript Description," 1997, <http://users.ox.ac.uk/~lou/wip/MS/>; the National Library of Sweden website, <https://www.manuscripta.se/>; Patrik Granholm and Patrik Åström, "Ny Digital Katalog över Fornsvenska Handskrifter," in *A Copenhagen Miscellany: Studies in East Norse Philology*, ed. Simon Skovgaard and Seán D. Vrieland (Copenhagen: Selskab for Østnordisk filologi, 2019), 79–90; for the National Library of Iceland, see Ingibjörg Steinunn Sverrisdóttir, Kristinn Sigurðsson, and Örn Hrafnkelsson, "Access and Curation of Digital Cultural Heritage in the National and University Library of Iceland," *Microform & Digitization Review* 41 (2012): 97–102.

and answer new research questions. The results of these digital explorations are expected to shed new light on our understanding of manuscript culture, the history of literature, and book history. Projects such as the recently launched “Mapping Manuscript Migrations” platform aim to generate new knowledge through the interpretation of digital data.<sup>4</sup> Multiple smaller-scale projects also attempt to apply digital tools to interpret data gathered in digital catalogs of manuscripts, but the results that both types of initiatives can achieve are highly dependent on the quality of the underlying data.

One such project was presented at the second Digital Humanities in the Nordics (DHN) conference in March 2017.<sup>5</sup> In this project, computer-assisted analysis was applied to macro-analyze the manuscript context of Old Norse literature in order to answer a research question concerning the genre affiliations of Old Norse texts. Through network analysis, the project sought to examine whether modern scholarly genre classifications of the Old Norse literary corpus are reflected in the manuscript transmission of the texts to which these scholarly terms are applied. This research was a response to disagreements among scholars of Old Norse regarding traditional genre classifications. Some leading scholars in the field had been pointing towards the manuscript context as a key for understanding genre classification, yet while numerous scholars invoked manuscript material

---

4 “Mapping Manuscript Migrations” is a portal intended to enable large-scale exploration of data relating to the history and provenance of medieval and early modern manuscripts. It combines data from three specialist databases, based on Linked Open Data principles and technology: the Schoenberg Database of Manuscripts, Bibale, and Medieval Manuscripts in Oxford Libraries. Read more about the project at <https://mappingmanuscriptmigrations.org/about>, accessed 24 April 2020.

5 Katarzyna Anna Kapitan, Timothy Rowbotham, and Tarrin Wills, “Visualising Genre Relationships in Icelandic Manuscripts,” in *Digital Humanities in the Nordic Countries, Second Conference: Conference Abstract*, ed. Daniel Brodén (Gothenburg: The University of Gothenburg, 2017), 59–62, [http://dhn2017.eu/wp-content/uploads/2017/03/DHN2017\\_Book\\_of\\_Abstracts\\_20170313.pdf](http://dhn2017.eu/wp-content/uploads/2017/03/DHN2017_Book_of_Abstracts_20170313.pdf). The study was inspired by work done by Alaric Hall and Katelin Parsons, “Making Stemmas with Small Samples, and Digital Approaches to Publishing Them: Testing the Stemma of *Konráðs Saga Keisaronar*,” *Digital Medievalist* 9 (2013), <https://doi.org/10.16995/dm.51>, who through the application of network analysis explored relationships between various groups of Icelandic romances. A similar approach has been presented in the recently published essay by Gustavo Fernández Riva, “Network Analysis of Medieval Manuscript Transmission: Basic Principles and Methods,” *Journal of Historical Network Research* 3 (2019): 30–49.

in their work, they arrived at contradictory conclusions concerning genre boundaries.<sup>6</sup> Existing scholarship had been stymied by the limited number of manuscripts a single human being can manually examine and analyze. Therefore, instead of considering individual manuscripts, this project collected and compiled all the digital data available for Icelandic manuscripts at the time, in order to visualize the relationships among the texts preserved in these manuscripts as a network, thereby revealing patterns of co-occurrences that would not be otherwise discernable.

Data were compiled from four online repositories (two XML-based catalogs and two databases), representing Icelandic manuscripts in twenty-three repositories in Europe and North America.<sup>7</sup> After overcoming obstacles

---

6 Stephen Mitchell, *Heroic Sagas and Ballads* (Ithaca, NY: Cornell University Press, 1991); Aðalheiður Guðmundsdóttir, *Úlfhams saga*, Stofnun Árna Magnússonar á Íslandi: Rit. 53 (Reykjavik: Stofnun Árna Magnússonar á Íslandi, 2001); Matthew James Driscoll, "Late Prose Fiction (Lygisögur)," in *A Companion to Old Norse-Icelandic Literature and Culture*, ed. Rory McTurk (Oxford: Blackwell, 2005), 190–204; Ármann Jakobsson, "The Earliest Legendary Saga Manuscripts," in *The Legendary Sagas: Origins and Development*, ed. Annette Lassen, Agneta Ney, and Ármann Jakobsson (Reykjavik: University of Iceland Press, 2012), 21–32.

7 The two main XML-based digital catalogs of Nordic manuscripts used in this study are *handrit.is* and the catalog of the Stories for All Time project *fasnl.ku.dk*. *Handrit.is* is the union catalog of Nordic manuscripts curated collaboratively by the National Library of Iceland, the Arnarnagðnæan Institute in Copenhagen, and the Arni Magnusson Institute in Reykjavik. See Sverrisdóttir, Sigurðsson, and Hrafnkelsson, "Access and Curation of Digital Cultural Heritage"; Matthew James Driscoll, "The Virtual Reunification of the Arnarnagðnæan Manuscript Collection," in *The Digital Demotic: A Selection of Papers from Digital Resources in the Humanities 1997*, ed. Lou Burnard, Marilyn Deegan, and Harold Short (London: Office for Humanities Communication, 1998), 55–64; Eric Haswell, Matthew James Driscoll, and Claire Warwick, "Towards a Union Catalogue of XML-Encoded Manuscript Descriptions," *Conference Abstracts of the Digital Humanities 2006*, 2006, 308–9; Matthew James Driscoll, "Handrit.is' and the Virtual Reunification of the Arnarnagðnæan Manuscript Collection," in *Virtual Visits to Lost Libraries: Reconstruction of and Access to Dispersed Collections*, ed. Ivan Boserup and David J. Shaw, CERL Papers 11 (London: Consortium of European Research Libraries, 2011), 87–93. *Fasnl.ku.dk* is a thematic catalog of manuscripts preserving Icelandic legendary sagas hosted at the Arnarnagðnæan Institute in Copenhagen, but it covers manuscripts preserved in twenty-three repositories in Europe and North America. The selection of data sources was determined by the research questions focusing on Old Norse sagas; therefore, other important cataloging initiatives in Scandinavia, such as *manuscripta.se*, a digital catalog of manuscripts in Sweden, a part of *TTT: Text till tiden! Medeltida texter i kontext—då och nu*, hosted by the National Library, Stockholm, and discussed by Granholm and Åström, "Ny Digital Katalog Över Fornsvenska Handskrifter," or *fragment.uib.no*, a resource based on the

such as sparse or fragmentary descriptions of some of the manuscripts, on one hand, and duplicates or multiple descriptions of some manuscripts, on the other, the final data set consisted of 674 unique titles preserved in 1,331 manuscripts, with almost 130,000 distinct connections. The data set was imported into Gephi, an open-source network analysis and visualization software package, to visualize as a network the relationships among catalogued texts. In this network, titles of works are represented as nodes, and co-occurrences in manuscripts create edges that connect the titles together (fig. 1). The more frequently titles appear together in manuscripts, the thicker the edge between these titles. Based on this network, the research group suggested that to some extent the modern genre division of Old Norse literature can be observed, as, for example, the Eddic poems and legal texts clearly separate themselves from the main group of sagas. At the same time, based on the numerous cross-genre edges within the group of sagas, it can be posited that there is no clean boundary between, for example, the family sagas, which concern Icelandic families during the first century of the Icelandic Commonwealth, and the kings' sagas, which concern Norwegian and Danish kings during the period extending from circa 850 to 1280.<sup>8</sup>

The research team invited audiences to further explore the network in order to validate or falsify findings through detailed case studies of separate groups of texts or manuscripts. The present article is partially a response

---

research project “From Manuscript Fragments to Book History” (2012–2017) hosted at the University of Bergen in Norway, were not included in the study.

The databases used in the project are the Skaldic Project (<https://skaldic.abdn.ac.uk/db.php>) and the Dictionary of Old Norse Prose (<https://onp.ku.dk/onp.onp.php>). For more about the database of the Skaldic Project, see Tarrin Wills, “Relational Data Modelling of Textual Corpora: The Skaldic Project and Its Extensions,” *Literary and Linguistic Computing* 30, no. 2 (2015): 294–313. For more about the ONP database, see Ellert Thor Johannsson and Simonetta Battista, “Editing and Presenting Complex Source Material in an Online Dictionary: The Case of ONP,” in *Proceedings of the XVII EURALEX International Congress*, ed. Tinatin Margalitadze and George Meladze (Tbilisi: Ivane Javakhishvili Tbilisi University Press, 2016), 117–28.

8 Vésteinn Ólason, “Family Sagas,” in *A Companion to Old Norse-Icelandic Literature and Culture*, ed. Rory McTurk (Oxford: Blackwell, 2005), 101; Ármann Jakobsson, “Royal Biography,” in *A Companion to Old Norse-Icelandic Literature and Culture*, ed. Rory McTurk (Oxford: Blackwell, 2005), 388.

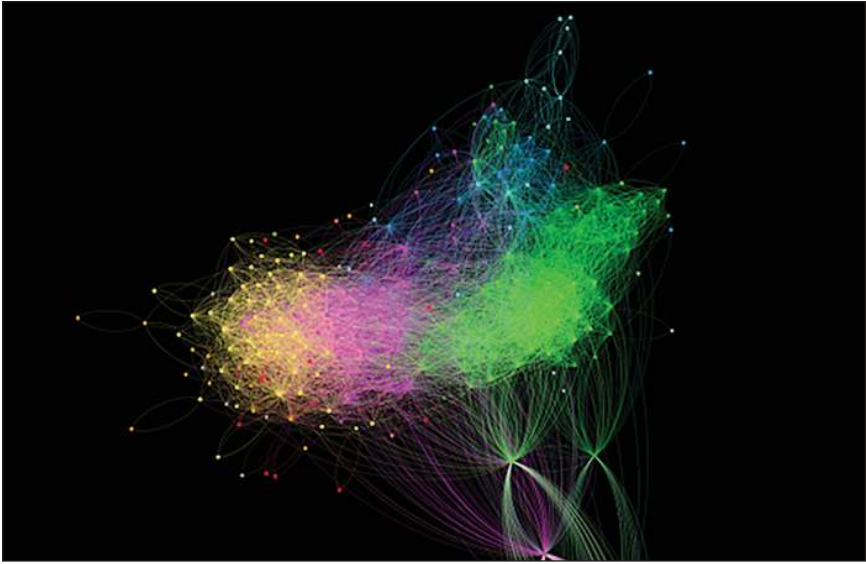


FIGURE 1. Part of the network of Old Norse texts appearing in extant manuscripts, which have been digitally cataloged. Yellow nodes: romances, pink nodes: legendary sagas, green nodes: family sagas, blue nodes: kings' sagas. Based on data from Kapitan, Rowbotham, and Wills, "Visualising Genre Relationships in Icelandic Manuscripts."

to this call. It illustrates the limitations of the 2017 experiment through a close examination of the transmission history of a single work preserved in multiple manuscripts. Taking these limitations as the point of departure, this article argues that the digital catalogs we produce today do not meet the expectations of scholarly audiences and do not take full advantage of the possibilities offered by digitally encoded information. The study highlights the current emphasis upon the textuality of handwritten books as opposed to their material aspects, and argues in favor of more materially focused digital catalogs. From this digital, material-philological perspective—and with an emphasis on the “digital,”—this study considers the purpose and functions of digital catalogs.<sup>9</sup> Should users be constrained to search for

---

9 The body of scholarship devoted to manuscript cataloging is so large that its exhaustive overview lies outside of the scope of this brief article. It is, however, important to mention some studies that discuss digital catalogs, especially from the perspective of their structure;

titles, authors, and watermarks manually, or should they be able to use digital tools to explore digital collections in different and innovative ways? Are the data collected in a given catalog intended to be analyzed manually, or are they meant to serve as the basis for computer-assisted quantitative macroanalysis of manuscript descriptions, harnessed to answer new research questions?

### *A Single Saga, its Manuscript Context, and its Manuscript Descriptions*

To assess the results achieved through the network analysis of manuscript descriptions, this case study uses a single saga, *Hrómundar saga Greipssonar*, which is preserved in almost forty manuscripts that have all been cataloged in digital format by the Stories for All Time project. *Hrómundar saga* is an Icelandic saga that is traditionally classified as a member of the corpus of legendary sagas, but the original medieval version of this saga is now lost. The text as it is known today is a post-medieval adaptation dated to the seventeenth century.<sup>10</sup> Because of this late origin, this saga does not necessarily fit well with the other works included in the corpus of legendary sagas, since most of them date to the fourteenth and fifteenth centuries. This makes *Hrómundar saga* an interesting case study for the investigation

---

see, e.g., Andrew Prescott, “The Function, Structure and Future of Catalogues,” *Digital Riffs*, 11 January 2013, <http://digitalriffs.blogspot.com/2013/01/the-function-structure-and-future-of.html?m=1>; Patrick Andrist, “Going Online Is Not Enough!: Electronic Descriptions of Ancient Manuscripts, and the Needs of Manuscript Studies,” in *Analysis of Ancient and Medieval Texts and Manuscripts: Digital Approaches*, ed. Tara L. Andrews and Caroline Macé (Turnhout: Brepols, 2014), 309–34; Cashion, “Cataloging Medieval Manuscripts, from Beasts to Bytes.” Some points of criticism brought up in these articles regarding catalog structure are successfully solved by the TEI consortium’s Manuscript Description module; see TEI Consortium, “TEI P5: Guidelines for Electronic Text Encoding and Interchange. 3.1.0,” <http://www.tei-c.org/Guidelines/P5/> (last accessed 1 August 2019).

<sup>10</sup> Judith Jesch, “Hrómundar Saga Gripssonar,” in *Medieval Scandinavia: An Encyclopedia*, ed. Phillip Pulsiano, Kirsten Wolf, and Paul Acker (New York: Garland, 1993), 305; Katarzyna Anna Kapitan, “Studies in the Transmission History of Hrómundar Saga Greipssonar” (PhD diss., University of Copenhagen, 2018).



of genre affiliation in extant manuscripts, and specifically for posing the question of whether the saga appears more frequently in manuscripts containing older legendary sagas or, conversely, in those with younger *rímur*-based narratives.

Exploration of the network and XSLT and XPath queries of the Stories for All Time project's XML files reveals that among the closest neighbors of *Hrómundar saga Greipssonar* is a saga that is traditionally not classified as a legendary one: *Bragða-Ölvis saga*.<sup>11</sup> *Bragða-Ölvis saga* is a post-medieval *rímur*-based saga, which despite making use of literary motifs characteristic to legendary sagas was never included in this corpus.<sup>12</sup> Analysis of the distribution of texts co-occurring with *Hrómundar saga* in extant manuscripts reveals that in the seventeenth and eighteenth centuries *Bragða-Ölvis saga* appears most frequently alongside *Hrómundar saga* (fig. 2).<sup>13</sup> In the nineteenth century, however, *Bragða-Ölvis saga* almost disappears from the manuscripts preserving *Hrómundar saga Greipssonar*, and the traditional legendary sagas dominate.<sup>14</sup> Also in the nineteenth century, a new text appears frequently with *Hrómundar saga*. This new text is *Starkaðar saga gamla*, a late eighteenth-century saga based on motifs from Saxo Grammaticus's *Gesta Danorum*.<sup>15</sup> None of the older manuscripts preserve *Hrómundar*

---

11 A. G. Hooper, "Bragða-Ölvis Saga Now First Edited," *Leeds Studies in English* 1 (1932): 42–54.

12 The corpus of legendary sagas was established in the nineteenth century by Danish philologist Carl Christian Rafn, but his work was clearly influenced by other Scandinavian scholars. On this subject recently, see Philip Lavender, "The Secret Prehistory of the *Fornaldarsögur*," *Journal of English and Germanic Philology* 114, no. 4 (2015): 526–51.

13 A preliminary study on this subject was presented at the Annual Digital Humanities Conference in 2017. See Katarzyna Anna Kapitan, "Network Analysis of the Manuscript Context of Old Icelandic Literature," *Digital Humanities 2017: Conference Abstracts* (Montreal: McGill University & Université de Montréal, 2017), 487–489.

14 The preponderance of legendary sagas in the younger manuscripts might be influenced by the first editions of *Hrómundar saga* that appeared in printed collections of sagas dealing with the legendary past of Scandinavia, such as Erik Julius Björner, *Nordiska kända dater i en sagoflock samlade om forna kongar och hjältar* (Stockholm: J. L. Horrn, 1737); Carl Christian Rafn, *Fornaldar sögur Nordrlanda eptir gömlum handritum*, 3 vols. (Copenhagen: Popp, 1829–30).

15 Rudolf Simek and Hermann Pálsson, *Lexikon der altnordischen Literatur, die mittelalterliche Literatur Norwegens und Islands* (Stuttgart: Kröner, 1987).

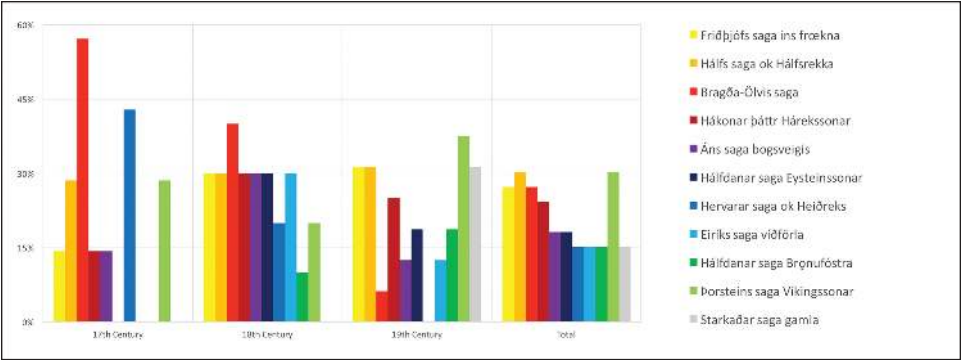


FIGURE 2. Graph representing texts appearing most frequently with *Hrómundar saga Greipssonar* in extant manuscripts. Based on Kapitan, “Network Analysis of the Manuscript Context.”

saga together with *Starkaðar saga gamla*. One of the reasons for this may lie in the saga’s late origin and the fact that there are few known manuscripts of this saga from the eighteenth century.<sup>16</sup> But another primary reason for this finding is related to the limitations of the method employed in the 2017 study, or, more accurately, those of the data set that was analyzed.

When conducting a quantitative analysis of digital manuscript descriptions, two factors have to be taken into consideration. The first of these is the instability of manuscripts. Some volumes have changed shape a number of times as they have passed from one owner to another, with each successive owner making adjustments to the contents of the book, adding some items and removing others to meet his or her personal needs and interests. Accordingly, two texts coexisting today in one volume need not have been produced in unison.<sup>17</sup> The second factor concerns the character of digital

16 Massimiliano Bampi, “Starkaður Across the Centuries: Strategies of Rewriting and Manuscript Variation in *Starkaðar saga gamla*,” in *The Legendary Legacy: Transmission and Reception of the Fornaldarsögur Norðurlanda*, ed. Matthew James Driscoll et al. (Odense: University Press of Southern Denmark, 2018), 53–69.

17 The rearrangement of texts is especially common in the Icelandic manuscript collection of Arni Magnusson, who notoriously intervened in the structures of his books. See Beeke Stegmann, “Árni Magnússon’s Rearrangement of Paper Manuscripts” (PhD diss., University of Copenhagen, 2016).

catalogs, such as their scope, level of detail, and encoding. In many cases, online catalogs transmit outdated manuscript descriptions, as many are surrogates of existing printed catalogs. These printed catalogs in turn often originate in the nineteenth century, and some of them are of little use when it comes to computer-assisted data interpretation. For example, many do not specify whether two texts appearing together in a single manuscript are also part of a single codicological unit. This type of information, however, is crucial for studying the contexts in which certain literary texts appear.<sup>18</sup>

The importance of codicological units in the transmission history of *Hrómundar saga* can be exemplified by the digital description of London, British Library, Add. MS 11109, which is one of the manuscripts that preserve *Hrómundar saga* together with *Starkaðar saga gamla*. The only information we can find in the existing catalogs about this manuscript is that these two texts are today in one book—but is that sufficient? The information provided in the British Library online catalog about this manuscript originates from the printed catalog of the collection produced in the middle of the nineteenth century.<sup>19</sup> The print catalog contains exactly the same information as that available in the digital catalog, with the titles of the sagas translated into Latin even though the entire manuscript is written

---

18 In this article I use the term “codicological unit,” as defined by Gumbert; see Peter Gumbert, “Codicological Units: Towards a Terminology for the Stratigraphy of the Non-Homogeneous Codex,” *Segno e testo* 2 (2004): 18–42. Other important publications for the stratigraphy of manuscripts are articles by Pamela Robinson. See Pamela R. Robinson, “Self-Contained Units in Composite Manuscripts of the Anglo-Saxon Period,” *Anglo-Saxon England* 7 (1978): 231–38; Pamela R. Robinson, “A Self-Contained Unit in Composite Manuscripts,” in *Codicologica 3: Essais typologiques*, ed. A. Gruys and Peter Gumbert (Leiden, 1980), 46–69. Pamela Robinson’s work was further discussed in Ralph Hanna III, “Booklets in Medieval Manuscripts: Further Considerations,” *Studies in Bibliography* 39 (1986): 100–111; Erik Kwakkel, “Towards a Terminology for the Analysis of Composite Manuscripts,” *Gazette du livre médiéval* 41 (2002): 12–19. It is important from the perspective of cataloging practice to distinguish between booklets, codicological units, or production units, but this discussion lies outside the scope of the present study. For a more detailed discussion, see Patrick Andrist, Paul Canart, and Marilena Maniaci, *La syntaxe du codex: Essai de codicologie structurale* (Turnhout: Brepols, 2013).

19 *List of Additions to the Manuscripts in the British Museum in the Years 1836–1840* (London: British Museum, 1843).

in Icelandic (fig. 3).<sup>20</sup> The far more detailed catalog of the Stories for All Time project lists the sagas preserved in BL Add. MS 11109, with their Icelandic titles, and provides a basic description of the physical features of this manuscript, including number of leaves and hands, dimensions, and details regarding the support.<sup>21</sup> This catalog, however, also reproduces information provided in the preexisting catalog, which in this case is Jón Helgason's unpublished catalog of Icelandic manuscripts in the British Library, compiled sometime around 1933.<sup>22</sup> Since both of these digital catalogs merely reproduce older sources, it is not surprising that they do not provide information that would be crucial for contemporary research in the spirit of material philology, such as collation formulas and codicological units.

The identification of codicological units is essential if studies like the 2017 experiment are to deliver meaningful results. If we are interested in learning about how texts were produced, we need to know, for instance, whether two texts that appear together in a single manuscript today were also produced together, or whether they were merely placed together at a later time. In the case of BL Add. MS 11109, where *Stakaðar saga gamla* has been identified as one of the most frequent texts co-occurring with *Hrómundar saga*, it has to be emphasized that these two sagas belong to two different codicological units and were written in different hands, on different paper, most likely at different points in time. Therefore, from a strictly production-oriented point of view, this manuscript should be cataloged as containing at least four separate parts, with these two sagas occurring in separate units. If the manuscript description included information pertinent to the structure of this book, these two texts would not appear adjacent to each other in the network map.

---

20 The description of BL Add. MS 11109 in the British Library's "Explore Archives and Manuscripts" online catalog is available at <http://searcharchives.bl.uk>, accessed 1 December 2019.

21 The description of BL Add. MS 11109 in the Stories for All Time online catalog is available at <http://fasnl.ku.dk>, accessed 1 December 2019.

22 Jón Helgason, *Catalogue of the Icelandic Manuscripts in the British Library*, unpublished draft held at Den Arnamagnæanske Samling in Copenhagen, n.d.

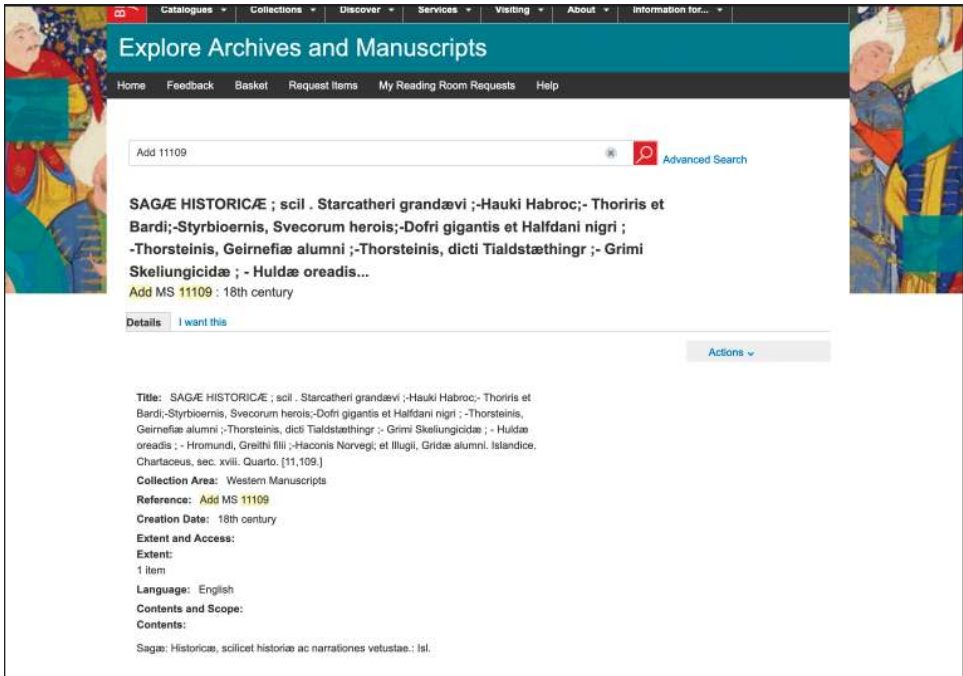


FIGURE 3. Screenshot of the description of London, British Library, Add. MS 11109, in the British Library online catalog, taken on 10 December 2019.

The lack of description of the structure of the manuscript is not the only unsatisfactory aspect of these legacy catalogs. The description of the textual contents could also be improved markedly. Neither catalog discloses that the *Hrómundar saga* found in the British Library manuscript is not identical to the text encountered in other manuscripts and printed editions. The differences between these two versions are clear from the opening and closing lines of text, which can be compared to those found in the oldest manuscript to preserve the saga:

#### Incipit in London, British Library, Add. MS 11109, fol. 106v:

Sva hefr Sögo þessa, at á þeim tímum sem margir Stólkongar voro í Norvegi, Danmorco oc Svíþjóþ . . .

(So begins this saga, that at that time when there were many petty kings in Norway, Denmark, and Sweden . . .)

**Incipit in Reykjavik, Stofnun Árna Magnússonar,  
AM 601 b 4to, fol. 1r:**

⟨S⟩á kongr ríeþe fyrir Gordom'[i danmorc]' er Olafur hiet hann var sonr Gnóþar Asmundar . . .

(A king ruled over Garðir in Denmark, who was called Ólafur, he was a son of Gnóðar-Ásmundur . . .)

**Explicit in BL Add. MS 11109, fol. 132r:**

sýnist sva sem at þau sum mani í barnæsko dáit hafu, en af sumom quomo micil menni sem í öþrom Sögom fráskýrir. Lúkom vær sva þessari Sögo af Hrómundi Greipssyni

(it seems that some of them died in their childhood, but some became great men, as other sagas give account of. We finish thus this saga of Hrómundur Greipsson)

**Explicit in AM 601 b 4to, fol. 6r:**

Eru af þeim Komnar konga ætter oc kappar mikler, oc lýkr hier saugo Hromundz greipzsonar

(From them are descended royal dynasties and great warriors. Here ends the saga of Hrómundur Greipsson.)

If the digital catalogs contained incipits and explicits of these texts, there would be no confusion regarding their identification. Even if the British Library catalog provided merely the page range that *Hrómundar saga* occupies in this manuscript, the researcher would be able to question whether these two texts are indeed manifestations of the same adaptation of the story. While both manuscripts are in quarto format, in the older Reykjavik manuscript the saga takes up six leaves, while in BL Add. MS 11109 it spans twenty-six leaves. It turns out that the *Hrómundar saga* preserved in AM 601 b 4to, which is a seventeenth-century prose adaptation of the *rímur* and is about 3,500 words long, appears most frequently with *Bragða-Ölvis saga* in the oldest manuscripts, while the *Hrómundar saga* that we find

in the British Library manuscript together with *Starkaðar saga gamla* is a new adaptation of the story that is about 12,000 words long. The saga found in the London manuscript is not only four times longer than the older saga, but its contents and style are also significantly different.

The final problem with the catalog information concerns the dating of the British Library manuscript. Since the British Library catalog dates Add. MS 11109 to the eighteenth century, this appears to be the oldest manuscript witness of this younger adaptation of the saga, as all other known manuscripts of this younger adaptation are from the nineteenth century. This is very important from the perspective of research on the history of Icelandic literature and genre development, because the language and the style of this saga appear to be quite modern for the eighteenth century. The question remains, however, as to whether the date given in the catalog is correct. Is BL Add. MS 11109 really an eighteenth-century manuscript?

There is no textual evidence within the manuscript itself that contradicts this dating explicitly, as there is not a single dated colophon or signature. A detailed examination of the physical features of this manuscript, of the sort that is barely ever reported in catalogs, can, however, deliver meaningful evidence for dating.<sup>23</sup> In one of the quires of the manuscript, there is a watermark that resembles the typical Honig watermark, as appears in William Algernon Churchill's catalog of watermarks.<sup>24</sup> This watermark represents a beehive, but it is slightly modified in comparison with Churchill's example. Instead of the name of the papermaker on the plinth of the beehive, there is the date 1824. Therefore, at the very least, the portion of the manuscript written upon this watermarked paper cannot originate from the eighteenth century. Thanks to the physical examination of the manuscript, we can therefore extrapolate that the younger version of *Hrómundar saga* is not an eighteenth-century adaptation of the older saga, which is surprisingly

23 Katarzyna Anna Kapitan, "Dating Paper Manuscripts Based on Watermarks: A Case Study of Selected Nineteenth-Century Icelandic Manuscripts," in *Care and Conservation of Manuscripts 17*, ed. Matthew James Driscoll (Copenhagen: Museum Tusculanum, 2021).

24 William Algernon Churchill, *Watermarks in Paper, in Holland, England, France, etc., in the XVII and XVIII Centuries and Their Interconnection* (Amsterdam: Menno Hertzberger, 1935), clvi.

modern in style, but is instead a nineteenth-century adaptation, and in this context its style becomes less surprising.

By combining the results achieved in the network analysis of manuscript descriptions with the direct examination of artifacts, this section has demonstrated some of the shortcomings of the existing digital descriptions of British Library, Add. MS 11109. These shortcomings stem from the perpetuation of outdated information in the digital domain, which resulted in insufficient attention being paid to the physical features of the manuscript, the erroneous identification of texts, and the incorrect dating of the volume. The limitations of the digital catalog have a huge influence on the results of the digital analysis, which can only be as good as the data on which it is based.

### *Perspectives on Digital Catalogs*

As demonstrated in the previous section, the unsatisfactory results obtained from the digital analysis of the wider context of Nordic manuscripts originate not from the method applied in the study, but from the data set used in the experiment—that is, the digital catalogs from which the data were harvested. This is not surprising, since, as Patrick Andrist observed, “There is an old adage in computer science: ‘Garbage in . . . garbage out’! Manuscript databases are not exempted from this rule—the more correct and correctly linked information is entered, the better the database will meet the needs of manuscript studies. Merely going online is *definitely* [my emphasis] not enough.”<sup>25</sup>

It is difficult to disagree with Andrist’s statement that going online is *definitely* not enough, and there is still a large amount of work that needs to be done. This brings us to a discussion of the function and purpose of digital catalogs, and to Andrist’s notion of “correct” and “correctly” encoded information. The main objective here is not to provide universal guidelines for the cataloging of Nordic or other Western manuscripts, but instead to

---

25 Andrist, “Going Online Is Not Enough!,” 331.



consider how digital catalogs should be structured and which kinds of information they ought to contain in order for research questions like the ones which prompted the 2017 study to be successfully answered.

It has been 130 years since Falconer Madan (1851–1935), librarian of the Bodleian Library, described the main principles of manuscript description. In his *Books in Manuscript: A Short Introduction to Their Study and Use*, first published in 1893, Madan wrote: “The description of a MS. should consist of three parts—(1) The technical description; (2) the list of contents; (3) the history and present shelf-mark.”<sup>26</sup> While the first and third required elements of the description are self-explanatory, it stands to be clarified that the technical description refers in fact to the physical description, which should include, according to Madan, information on the material support, the number of leaves, and so on.

Essentially, these same features are included in the guidelines for the Text Encoding Initiative, where the manuscript description <msDesc> section is meant to include the following elements:

- manuscript identifier <msIdentifier>, which contains the information required to identify the manuscript, such as shelfmark and repository;
- manuscript contents <msContents>, which contains information regarding the intellectual content of the manuscript;
- physical description <physDesc>, which contains a full physical description of the manuscript;
- additional <additional>, which groups additional information about the manuscript, including bibliographic, curatorial, and administrative information; and
- history <history>, which groups elements describing the full history of the manuscript.<sup>27</sup>

---

26 Here cited according to the second edition: Falconer Madan, *Books in Manuscript: A Short Introduction to Their Study and Use*, 2nd ed. (London: Kegan Paul, Trench, Trubner & Co., 1920), 162.

27 TEI Consortium, “TEI P5: Guidelines for Electronic Text Encoding and Interchange 4.0.0,” <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>.

Most of these elements can be identified on Madan's list; only the <additional> element is new. We can therefore conclude that there is a degree of consensus concerning the most important elements of a manuscript's description, and that this has changed little in the past century.

Since most of the files analyzed in the 2017 study included all such elements, in principle their presence should not have posed any challenges for the data analysis, yet this was clearly not the case. In my view, the problem that the 2017 study faced originated from a divergent understanding of what information was important and necessary to record, as well as the different methods by which such information came to be encoded.

To begin with the matter of encoding, or what in my view Andrist meant by "correctly linked information," it should be noted that the majority of the data used for the network of Old Norse literature were based on TEI-XML descriptions. TEI-XML provides a very flexible environment for describing manuscripts, which is its greatest strength but also its curse, as even within a single project with a relatively strict schema, the encodings can look quite different. After all, it is frequently left to individual catalogers to decide how much detail is included in the individual manuscript description, and how such information is encoded. Moreover, it can be argued that the TEI guidelines do not sufficiently encourage users to take full advantage of XML encoding, which, if it were done, would make digital analysis more powerful. The minimal examples presented in the guidelines seem to encourage use of the paragraph <p> element, which compromises the very purpose of encoding. The understandable intention of the TEI guidelines seems to be to not dissuade non-expert users with advanced encoding, but this compromise results in the abuse of the <p> element across many descriptions of Nordic manuscripts.<sup>28</sup> For instance, instead of encoding each codicological unit as a separate manuscript part <msPart>, the cataloger may choose to list all texts within a single manuscript contents <msContents> element and only mention somewhere in prose, most likely within the <p> element, that the codex consists of more than one

---

28 This is at least my personal experience with *Handrit* and the Stories for All Time project catalog, where, due to a scarcity of resources, the textual aspects of manuscripts are prioritized over their physical features.

codicological unit. This renders the automated extraction of information regarding the structure of a manuscript more complicated, if not impossible. Even if the information concerning the structure of the British Library manuscript discussed earlier were to be included somewhere within the <p> element of the TEI description, this information would get lost in the transformation performed in the 2017 analysis, which was focused on <msContents>, <msPart>, and <msItems>. This transformation assumes that if there is no <msPart> element, the codex is homogeneous.

With regard to the “correctness” of information provided by a given catalog, it should be made clear that the degree of detail required in a manuscript description is an area of great disagreement among scholars and librarians. Some of the features of a manuscript description that in my view should be considered essential were discussed in the previous section, which outlined the shortcomings of the British Library manuscript’s description. From the perspective of material philology, the most crucial feature is the description of physical aspects of the manuscript. The importance of such descriptions was emphasized by Dorothy Coveney in 1950, long before the rise of material philology, when she stated that “few catalogues pay much attention to the material, collation, measurement of text-space and columns, change of hands or ink; and systems of prickmarks, even if unusual, are scarcely ever alluded to.”<sup>29</sup> She proposed a list of contents that a satisfactory manuscript description should include in which all such features appeared, but even today some of them are rarely found in digital catalogs. In my view, these features have become all the more relevant following the rise of material philology, which has brought about research questions that cannot be answered without sufficient attention to the physicality of manuscripts as artifacts. As shown by the example of BL Add. MS 11109, the close examination of physical features, such as watermarks, is of huge importance not only for the dating of a manuscript but also for an understanding of its production, use, and reuse.

---

29 Dorothy Coveney, “The Cataloguing of Literary Manuscripts,” *Journal of Documentation* 6, no. 3 (1950): 125–39.

Many digital cataloging projects for Nordic manuscripts focus on accessibility—that is, making the already existing manuscript descriptions available online, rather than creating new knowledge, understood here as the re-cataloging the artifacts so that descriptions meet the needs of paleographers, codicologists, book historians, and philologists. It seems easier to obtain funding for a short-term digitization project than to maintain and update digital catalogs over a longer period. The deliverables of digitization projects involving the photography of thousands of leaves are more spectacular than creating a smaller number of XML-TEI manuscript descriptions. But how are scholars and members of the public alike expected to find these wonderful high-resolution digital images, absent good descriptions of what such facsimiles represent?<sup>30</sup>

As seen in the digital descriptions of BL Add. MS 11109, information from nineteenth- and early twentieth-century catalogs has merely been copied into the digital sphere, without any significant revision. This is in line with Mark A. Greene and Dennis Meissner's economy of resources concept, which they call "More Product Less Processing" (MPLP), but it is, in my view, problematic from a research perspective.<sup>31</sup> The MPLP approach has been proposed mainly for large archives of contemporary materials, where acquisitions grow every day, and therefore it is only partially adequate

---

30 During her presentation at the Annual Meeting of the Medieval Academy of America in 2019, N. Kivlçim Yavuz observed that publishing digital photographs of manuscripts alone will not influence discoverability of the data until the images are accompanied by detailed metadata in the form of a catalog description. N. Kivlçim Yavuz, "Is a Picture Worth a Thousand Words?: Digital Facsimiles vs. Digital Catalogues of Manuscripts," presented at the 94th annual meeting of the Medieval Academy of America, Philadelphia, March 2019. Similar observations regarding Islamic manuscripts have been made by Dagmar Riedel, "Of Making Many Copies There Is No End: The Digitization of Manuscripts and Printed Books in Arabic Script," in *The Digital Humanities and Islamic & Middle East Studies*, ed. Elias Muhanna (Berlin: De Gruyter, 2016), 65–92; Dagmar Riedel, "How Digitization has Changed the Cataloging of Islamic Books," *Islamic Books: A Research Blog About Manuscripts, Printed Books, and Ephemera in Arabic Script*, 14 August 2012, <https://researchblogs.cul.columbia.edu/islamicbooks/2012/08/14/digitalsurrogates/>.

31 Mark A. Greene and Dennis Meissner, "More Product, Less Process: Revamping Traditional Archival Processing," *American Archivist* 68 (2005): 208–63; Mark A. Greene, "MPLP: It's Not Just for Processing Anymore," *American Archivist* 73 (2010): 175–203.

for the cataloging of rare books and manuscripts. It is understandable that archivists must choose not to catalog thousands of, for example, administrative records at the item level. Collections of rare books and manuscripts grow, however, at an incomparably slower pace, and the nature of this material requires more detailed processing, including preservation and cataloging.

Moreover, it must be acknowledged that philology and textual scholarship have moved forward markedly since the nineteenth century. When the main objective of textual scholarship was to identify exemplars of a given work and reveal the archetype that would form the basis for the critical edition, the physicality of all known witnesses was not particularly relevant. Therefore, traditional catalogs focused on identifying the texts preserved in a given manuscript. As the two versions of *Hrómundar saga* show, however, even this purely textual objective is not always sufficiently fulfilled. To allow scholars to immediately disambiguate the contents of a given manuscript, we need to provide not only our “uniform” title for a given work, but also such basics as its rubric, incipit, and explicit, all of which can be needed to identify different versions of a given work more readily.

In the digital era, the purpose of an online catalog is to bring manuscripts to researchers all over the world; but, if it is intended to serve these researchers well, the information that is made available should reflect the researchers’ needs and interests. Well-structured data concerning the physical and textual features of manuscripts could clearly prompt further developments not only in the field of manuscript studies, including codicology and paleography, but also in literary and historical studies. With the TEI guidelines, IIF, and digital tools and methods of data analysis, significant means exist to propel this field forward. The technology is there; we can easily prepare detailed descriptions of manuscripts, including properly encoded and well-structured information about various aspects of manuscripts, which are easy to parse digitally. There are digital data analysis and visualization tools and methods that could allow us to ask and answer new research questions. Therefore, we must ask ourselves as a scholarly community: why are we not doing so?

It has been thirty years since the rise of material philology, which changed the way we think about literary works and their manifestation as material

texts in manuscripts.<sup>32</sup> It has also been thirty years since Hope Mayo wrote that “few libraries can afford specialized curators for their medieval manuscripts, and most manuscript librarians, who are not often medievalists by training, must devote most of their attention to the modern literary and historical manuscripts assigned to their care. Relatively few collections provide *proper cataloging* [my emphasis] for their medieval codices, whether in traditional scholarly form or in current library formats.”<sup>33</sup> One can wonder whether the cataloging situation today is significantly different than it was at the beginning of the 1990s. There are many more digital projects today that produce manuscript descriptions and data, and these are now available online, but if we only digitize existing information about manuscripts, then we are not moving forward significantly. Is clicking from one tab to another in a digital catalog that represents nineteenth-century information so fundamentally different from clicking through an OCR-ed PDF of a printed catalog? On account of our current fixation on legacy data, the field of manuscript studies has only one foot in the digital age, with the other stuck in a nineteenth-century approach to manuscripts. What should we do when digital methods such as network analysis cannot be successfully applied to existing data? Should we stop dreaming about answering new research questions with new methods? Or should we dedicate our efforts to developing common standards for modern digital catalogs that will reflect the needs of a contemporary scholarly audience and allow new ways of extracting information? In my view, it is time to start “properly” (Mayo) and “correctly” (Andrist) cataloging manuscripts, in order to learn new things about the manuscript cultures that we are trying to understand.

While there are certainly challenges associated with obtaining funding for such resource-intensive and perhaps idealistic initiatives, part of the solution might lie in a greater involvement of the scholarly community during the cataloging process. Given the current circumstances, it appears that the only feasible way to build a catalog that meets researchers’ needs is to create an interactive catalog with an interface that allows users (who are

---

32 Here I use the term “material text” following Peter L. Shillingsburg, “Text as Matter, Concept, and Action,” *Studies in Bibliography* 44 (1991): 31–82.

33 Mayo, “Medieval Manuscript Cataloging and the MARC Format,” 11.

often themselves researchers) to enter the data. There is no space here to explore this idea in detail, but by incorporating users into the production of catalogs, we could create truly community-curated resources, where institutions and individuals cooperate to expand our knowledge and understanding of manuscripts in the digital domain. Institutions would be responsible for providing infrastructure and basic information about their holdings, while individuals would be responsible for research on particular manuscripts. Such a model is not free of challenges, however; further research and scholarly discussion are required before the idea can be successfully implemented.<sup>34</sup>

## Conclusion

Through the case study of a single Old Norse saga, *Hrómundar saga Greipssonar*, and its manuscript context in British Library Add. MS 11109, the present study evaluates the results of the computer-generated network of Old Norse literature, presented at the DHN conference in 2017. While network analysis of the genre affiliations of literary texts based on their manuscript contexts surely delivers interesting results and opens wide possibilities for

---

34 Taking into consideration experiences from past community-based catalogues, such as the four-year Collaboration in Cataloging: Islamic Manuscripts at Michigan project discussed by Eryn Kropf, “Collaboration in Cataloging: Sourcing Knowledge from Near and Far for a Challenging Collection,” in *Description: Innovative Practices for Archives and Special Collections*, ed. Kate Theimer (Lanham: Roman & Littlefield, 2014), 99–114, it is necessary to discuss and develop best practices for engaging the scholarly community in the cataloging process. The idea of involving users in cataloging projects has been taken up by E-codices, an online catalog of manuscripts from Switzerland that allows users to add annotations and additional bibliographical information. It is a promising feature, which seems effective, as there are at least 154 annotations as of December 2020, which range from self-promotion of users’ individual projects, to simple comments, to the identification of hands and contents. However, user-provided information is not incorporated into the XML file of the manuscript description, making automated data extraction difficult. I would like to thank Christoph Flüeler for discussing with me the peculiarities of manuscript description authorship in the case of E-codices, where manuscript descriptions are considered original scholarly work and credited as such. As a result, the appending of additional crowdsourced information is problematic.

manuscript studies, in-depth analysis of the physical features of one of the saga's manuscripts reveals some of the challenges and limitations of digital tools and methods. These limitations stem mainly from unsatisfactory digital catalog descriptions, which were a starting point for a discussion of cataloging practices for Nordic manuscripts. This study has showcased the possibilities that digital tools and methods could provide for manuscript scholars, if the data included in digital catalogs were well-structured, well-encoded, and regularly updated. It has emphasized that developments in cataloging practice should occur in close cooperation with the most recent research developments; for instance, after the rise of material philology as a scholarly methodology, one could expect digital catalogs to place, at the least, an equal emphasis on the material aspects of manuscripts as on their textual contents. By approaching the cataloging practice from the perspective of research questions that digital catalogs could answer, but cannot in their current states, this article argues that a digital catalog should give its users something more than the mere access to information that already exists in printed catalogs. The added value of a digital catalog should be not only the discoverability of the materials, but also well-structured, encoded data that can be analyzed with modern digital tools and methods. As an alternative to resource-demanding, institution-based catalogs, this article proposes a community-curated catalog in which users could be involved in the cataloging process. Input from scholars could make a huge difference in the way we create our catalogs and significantly advance our understanding of manuscript culture.



## *Appendix: Manuscripts and Digital Resources*

### MANUSCRIPTS

London, British Library, Add. MS 11109

Reykjavik, Stofnun Árna Magnússonar, AM 601 b 4to

### DIGITAL CATALOGS AND RESOURCES

Dictionary of Old Norse Prose: <https://onp.ku.dk/onp/onp.php>

Digital Bodleian: <https://digital.bodleian.ox.ac.uk/>

Digital Scriptorium: <https://digital-scriptorium.org/>

E-codices: <https://www.e-codices.unifr.ch/>

Explore Archives and Manuscripts, British Library Online Catalog: <http://searcharchives.bl.uk/>

Fragment: <https://www.fragment.uib.no/>

Gallica: <https://gallica.bnf.fr/>

Handrit: <https://handrit.is/>

Manuscripta: <https://www.manuscripta.se/>

Mapping Manuscript Migrations: <https://mappingmanuscriptmigrations.org/>

Schoenberg Database of Manuscripts: <https://sdbm.library.upenn.edu/>

Skaldic Project: <https://skaldic.abdn.ac.uk/db.php>

Stories for All Time: <http://fasnl.ku.dk/>