

# Positioning Riga in the 19th Century News Network: Tracing International News Flows in the *Rigasche Zeitung* Newspaper, 1802–1888

*Krister Kruusmaa*

National Library of Estonia, Narva mnt 11, 15015 Tallinn, Estonia;  
krister.kruusmaa@hotmail.com

*Kaarel Vanamõlder*

Tallinn University, School of Humanities,  
Narva mnt 25, 10120 Tallinn, Estonia; kaarel.vanamolder@tlu.ee

**Abstract.** This paper explores the possibilities of researching the historical circulation of news using place-date headings, the place of origin and date that was printed with a news paragraph. We demonstrate the possibilities of these headings using a 19th century German-language newspaper from Riga. We propose using digital methods to extract and use them to calculate the speed at which news travelled from different locations, as well as to map the geographical coverage of the newspaper. We also combine our results and perform a case study on the Crimean War. The results of our analysis show that place-date headings can be used effectively to study historical information flows, including cross-lingual ones, with less data and effort than required by other methods.

**Keywords:** networks of communication, news networks, newspapers, digital analysis, Baltic history, place-date headings

## INTRODUCTION

The mass digitisation of historical newspapers and new digital tools have changed the role of the press as a source for historical research, as well as how researchers engage with it.<sup>1</sup> An especially interesting research direction is the study of historical networks of communication, made possible by advanced search capabilities and visualisation techniques. The circulation of information in the press has been the focus of several research projects<sup>2</sup>, which have produced both tools<sup>3</sup> and diverse case studies.<sup>4</sup>

One popular method for researching historical information flows has been text reuse detection – processing large amounts of newspaper text to find overlapping segments (similar technology is used in plagiarism detectors). In the Nordic region, for instance, valuable work has been done on text reuse in Finnish newspapers. Researchers from the University of Turku dealt with over 5 million pages of newspapers from 1771 to 1920 with methods inspired by bioinformatics and identifying over 13 million clusters of textual overlaps. The results of these studies allowed the authors to analyse short-, mid- and long-term propagation (i.e. quoting and reprinting) of texts in their corpus and describe both its spatial and temporal dynamics.<sup>5</sup> Text reuse is also a part of the “historical

- 1 For a state of the art of research on digitised newspapers, see *Digitised Newspapers – A New Eldorado for Historians? Reflections on Tools, Methods and Epistemology*. Ed. by E. Bunout, M. Ehrmann, F. Clavert. De Gruyter, Oldenbourg, 2023, especially the introduction by the editors
- 2 *Viral Texts*, which focused on 19th century US newspapers (<https://viraltxts.org/>); *Oceanic Exchanges*, which studied transnational information flows between 1850 and 1914 (<https://oceanicexchanges.org/>); *Computational History and the Transformation of Public Discourse in Finland, 1640–1910* (<http://comhis.fi/clusters/info/about>); *Information Flows across the Baltic Sea: Swedish-language press as a cultural mediator, 1771–1918* (<https://textreuse.sls.fi/>) (accessed 03/01/2023).
- 3 M. Franke, M. John, M. Knabben, J. Keck, T. Blascheck, S. Koch. *LilyPads: Exploring the Spatiotemporal Dissemination of Historical Newspaper Articles*. – *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (presented at the 11th International Conference on Information Visualization Theory and Applications. SCITEPRESS – Science and Technology Publications, Valletta, Malta, 2020, 17–28.
- 4 J. Keck, M. Oiva, P. Fyfe. *Lajos Kossuth and the Transnational News: A Computational and Multilingual Approach to Digitized Newspaper Collections*. – *Media History*, 2022, 1–18, available online: <https://doi.org/10.1080/13688804.2022.2146905> (accessed 03/01/2023); R. Castellanos, E. Saisó, L. Domínguez. “Si los telegramas no mienten”. *Origen y circulación de las noticias de la explosión del Maine en la prensa mexicana, febrero 1898*. – *Revista de Historia de América*, 2020, 159, 255–287; M. Oiva, A. Nivala, H. Salmi, O. Latva, M. Jalava, J. Keck, L. Martínez Domínguez, J. Parker. *Spreading News in 1904*. – *Media History*, 2020, 26, 4, 391–407.
- 5 H. Salmi, P. Paju, H. Rantala, A. Nivala, A. Vesanto, F. Ginter. *The Reuse of Texts in Finnish Newspapers and Journals, 1771–1920: A Digital Humanities Perspective*. – *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2021, 54, 1, 14–28.

media monitoring tool suite” of the Impresso project which is based on historical Swiss newspapers.<sup>6</sup>

Text reuse thus holds huge potential for both generalist and fine-grained approaches to studying the spread of information in historical newspapers and is especially valuable for the context of long-term repetitions of a piece of text.<sup>7</sup> It also has its shortcomings, however. Because its application in the humanities was originally linked to social media analysis<sup>8</sup>, text reuse tends to treat historical newspapers as somewhat isolated from their historical context. It is also of limited use in studying cross-lingual information flows, as scaling such an approach beyond the borders of a single language has not yet yielded any significant results.<sup>9</sup>

Researchers have thus integrated other methods for studying the circulation and reprinting of news. For example, Verheul et al. explored large-scale circulation of news and ideas using word vector models, while Lerma Mayer et al. combined NER, sentiment analysis and a qualitative approach to analyse the study of news about the execution of the Mexican emperor Maximilian I in 1867.<sup>10</sup> In a more direct connection to our study, Rocío Castellanos Rueda et al. turned their attention to how information about the explosion on the US warship *Maine* in the Bay of Havana in 1898 reached the Mexican press. The authors explain that they used regular expressions to capture the cities and dates mentioned as sources for the news.<sup>11</sup> A similar approach have been used by Mila Oiva et al. in their study of the dissemination of news about the assassination of the Governor General of Finland in 1904.<sup>12</sup> These last two articles offer interesting insights into the connections between the press, the telegraph network and news agencies.<sup>13</sup>

6 <https://impresso-project.ch/> (accessed 03/01/2023).

7 For a more detailed discussion, see P. Paju, H. Rantala, H. Salmi. Towards an Ontology and Epistemology of Text Reuse. – *Digitised Newspapers*, 253–273.

8 See X. Shaobin, D. Smith, A. Mullen, R. Cordell. Detecting and Evaluating Local Text Reuse in Social Networks. – *Proceedings of the Joint Workshop on Social Dynamics and Personal Attributes in Social Media*. Association for Computational Linguistics, Baltimore, Maryland, 2014, 50–57, <https://doi.org/10.3115/v1/W14-2707> (accessed 03/01/2023).

9 Keck et al. Lajos Kossuth and the Transnational News, 3; Salmi et al. The Reuse of Texts in Finnish Newspapers, 26.

10 J. Verheul, H. Salmi, M. Riedl, A. Nivala, L. Viola, J. Keck, E. Bell. Using Word Vector Models to Trace Conceptual Change Over Time and Space in Historical Newspapers, 1840–1914. – *Digital Humanities Quarterly*, 2022, 16, 2, <http://www.digitalhumanities.org/dhq/vol/16/2/000550/000550.html>; A. Lerma Mayer, X. Gutierrez-Vasques, E. Saiso, H. Salmi. Underlying Sentiments in 1867: A Study of News Flows on the Execution of Emperor Maximilian I of Mexico in Digitized Newspaper Corpora. – *Digital Humanities Quarterly*, 2022, 16, 4, <http://www.digitalhumanities.org/dhq/vol/16/4/000649/000649.html> (accessed 03/01/2023).

11 R. Castellanos et al. “Si los telegramas no mienten”, 267–268.

12 M. Oiva et al. Spreading News in 1904.

13 Although both Smith et al. and Salmi et al. stress the importance of communication

Castellanos Rueda et al. and Oiva et al. make use of the intrinsic structure of historical newspapers. Before the 20th century (and well into it in some cases), newspapers were most often divided into single messages – or paragraphs – which started with a place and date of origin of the information. This way of structuring the news can be traced back even to the ancestors of early modern newspapers, the handwritten *gazettes* of 16th century Venice.<sup>14</sup> Historically, it was commonplace for newspapers to ‘recycle’ each other’s content.<sup>15</sup> To do this, paragraphs were simply copied from one publication to another, with the original source and date intact.<sup>16</sup> Functioning both as a structural component and an indication of trustworthiness, these simple pieces helped readers contextualise information that may have passed through many ports and towns before reaching them.

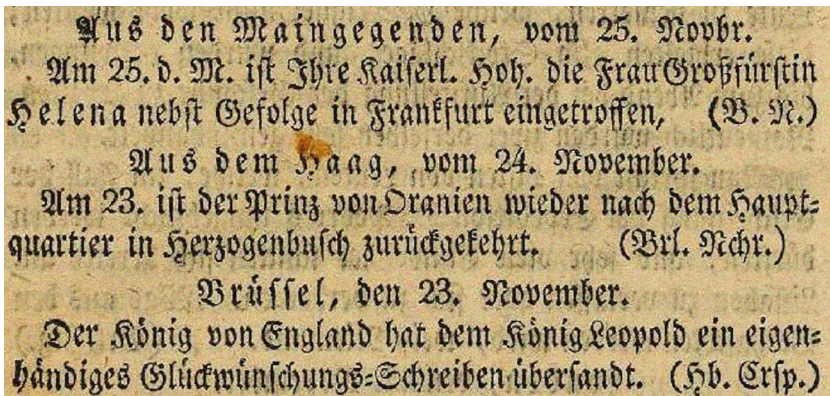


Figure 1: Example of several short news paragraphs and their place-date headings. Rigasche Zeitung 26.11.1831/no. 142.

We want to draw attention to this element, which we call the **place-date heading**. Straightforward, easily quantifiable and containing both temporal and spatial information, they seem to be a natural way

networks such as telegraph lines and railways in the spread of information, their methods are centred on textual overlap. See D. A. Smith, R. Cordell, E. M. Dillon. Infectious texts: Modeling text reuse in nineteenth-century newspapers. – 2013 IEEE International Conference on Big Data, Silicon Valley, CA, USA, 2013, 86–94, <https://doi.org/10.1109/BigData.2013.6691675> (accessed 03/01/2023); Salmi et al. The Reuse of Texts in Finnish Newspapers, 21–22.

<sup>14</sup> M. Stephens. *A History of News*. Penguin Books, New York, 1989, 153.

<sup>15</sup> J. Weber. *Straßburg 1605: Die Geburt der Zeitung*. – *Jahrbuch für Kommunikationsgeschichte*, 2005, 7, 3–27.

<sup>16</sup> For a more thorough analysis of the role of the paragraph as carrier of information, see W. Slauter. *The Paragraph as Information Technology: How News Traveled in the Eighteenth-Century Atlantic World*. – *Annales HSS (English edn)*, 2012, 67, 2, 253–278. Brendan Dooley has proposed “period” as a more suitable term for this unit: B. Dooley. *International News Flows in the Seventeenth Century: Problems and Prospects*. – *News Networks in Early Modern Europe*. Ed. by J. Raymond, N. Moxham. Brill, Leiden, Boston, 158–177.

to research news flows. Despite this, researchers have only recently started to use place-date headings to that end – most notably the work done by Yann Ryan who used them to model news circulation between Ireland, London and Continental Europe between 1654 and 1659.<sup>17</sup> However, there is a lack of an in-depth discussion on the role of these pieces of information, which could be considered a sort of intrinsic metadata of historical newspapers. Following an experiment that successfully demonstrated the usefulness of place-date headings to calculate the speed of the postal connection in the late 17th century,<sup>18</sup> we decided to scale up this approach. Situating our work in the wider context of communication history, we stress that the press cannot be viewed separately from their underlying communication networks.

In this article, we demonstrate how place-date headings can be used to study historical news flows. We believe that our approach is robust and scalable, largely language-independent and does not require extensive amounts of data for meaningful results. Instead of focusing on a single event, we draw on the persistent nature of the place-heading structure in the press to emphasise its role in researching long-range trends, as well as developments in communication technology. For this, we chose to analyse the entirety of a newspaper called the *Rigasche Zeitung* which was published in Riga almost through the entire 19th century.

Riga, the modern-day capital of Latvia, became one of the main economic hubs of the Baltic Sea during the 19th century.<sup>19</sup> Its position in the western part of the Russian Empire, to which it belonged between the Great Northern War (1700–1710/21) and World War I along with the “Baltic provinces” – Estonia (Estland), Livonia (Livland) and later also Curonia (Kurland) – makes it an interesting choice to study information flows in the press. Riga was largely inhabited by Baltic Germans, an elite minority group whose origins in the Baltics can be traced back to merchants and crusaders who arrived during the Northern Crusades of the 12th–13th centuries. Mainly manor owners, merchants and craftsmen, the Baltic Germans retained a close cultural connection to Germany.

The German presence in Riga meant that its press was predominantly German from its beginnings in the 17th century and existed in the same socio-linguistic communication space as other

17 Y. C. Ryan. “More Difficult From Dublin Than From Dieppe”: Ireland and Britain in a European Network of Communication. – *Media History*, 2018, 24, 3–4, 458–476.

18 K. Vanamölder. Millist ajalehte loeti Rootsijärgseid? Hamburgi/Altona ajalehtede mõjust Riigis Novelleni sisule. – *Acta Historica Tallinnensia*, 2021, 27, 2, 298–330.

19 Riga. Portrait einer Vielvölkerstadt am Rande des Zarenreiches 1857–1914. Hrsg. von E. Oberländer, K. Wohlfahrt. F. Schöningh, Paderborn, 2004.

German periodicals.<sup>20</sup> On the other hand, it was a part of the Russian Empire in administrative, political and military senses. This situated Riga on the watershed between two communication spheres. Indeed, Riga newspapers often acted as one of the main links in the information chain between the German states and St. Petersburg, the imperial capital.<sup>21</sup> The goal of this article is thus not only to demonstrate the potential of place-date headings but also to provide an insight into the relationship between the press and other communication networks in the Baltic region, which is a relatively under-researched field in the context of the 19th century. The results are also meant to complement the research done on Finnish newspapers that existed in many ways in an analogous historical context.

The article mainly uses digital research methods. We have chosen not to group them together into a dedicated section but rather to explain them one by one in their relevant sections, to emphasise the immediate relationship between each method and its results. In describing our methods, we have generally kept in mind readers with less experience in the digital humanities and therefore additional technical details can be found in the footnotes. The source code has been published in an online repository.<sup>22</sup> We have also created a very basic interactive web interface for those who want to explore further the data.<sup>23</sup>

## SOURCES AND DATA

### *Rigasche Zeitung*

The *Rigasche Zeitung* (hereafter *RZ*, 1778–1889; 1907–1919) was the oldest and one of the most popular and prominent publications in the Baltic provinces, published continuously for 111 years, after a short interruption in the early 20th century.<sup>24</sup> It had a significant and

20 Balti kirjakuultuuri ajalugu I. Keskused ja kandjad. Koost. L. Lukas. Tartu Ülikooli Kirjastus, Tartu, 2020, 151–160.

21 K. Vanamölder. Kuidas täita pool sajandit kestvat pausi? Ajakirjandus Riias 17. ja 18. sajandil kommunikatsioonijaloo vaatenurgast. – Muutused, ümberkorraldused, uuendused: Varauusaja arenguhood Eesti- ja Liivimaal, 1520–1800. Koost. M. Seppel, M. Maasing. Tallinna Ülikooli Kirjastus, Tallinn (forthcoming).

22 <https://github.com/krkryger/rigasche-zeitung> (accessed 03/01/2023).

23 <https://krkryger-rigasche-zeitung-streamlit-streamlit-ynlv7g.streamlit.app/> (accessed 03/01/2023).

24 In historiography, the year 1915 is also commonly used as the final closure of the *Rigasche Zeitung*, for example M. Lux. Das Riga der Deutschen. – Riga. Portrait einer Vielvölkerstadt am Rande des Zarenreiches, 75–114, here 110–112; Eestis ilmunud saksa-, vene- ja muukeelne perioodika 1675–1940. – Eesti retrospektiivne rahvusbibliograafia, V. Koost. E. Annus. Eesti Teaduste Akadeemia raamatukogu, Tallinn, 1993, 220. At the same time, the collection of the *Rigasche Zeitung* in the National Library of Latvia attests to the publication of the paper up to 1919, which is also used by R. Cerūzis. *Rigasche Zeitung*. – Nacionālā enciklopēdija. Available online: <https://enciklopedija.lv/skirklis/163962-Rigasche-Zeitung> (accessed 03/01/2023).

long-lasting impact not only on the readership in Riga but also on the wider public in the Baltics.<sup>25</sup> The newspaper is particularly significant in the history of Baltic German journalism. The political profile of the paper and its editors, as well as its eventual closure by the Russian central authorities, have been described.<sup>26</sup> The *RZ* was also one of the first publications after the half-century hiatus (1710–1761) or “press break” caused by the consequences of the Great Northern War.<sup>27</sup> Articles, announcements, etc., from the newspaper have often been used as illustrative examples, although the content of the *RZ* has neither been analysed in its entirety, nor with quantitative or digital methods.

The newspaper was originally called the *Rigische Politische Zeitung* (hereafter *RPZ*, 1778–1797), indicating its aim was to be primarily a paper for foreign political news.<sup>28</sup> The primary competitors of the newspaper were therefore not local news bulletins but rather similar newspapers of political news in the same communication space (Hamburg, Königsberg, etc.). The news circulating on the early modern news market largely reflected events in Western Europe and was mainly distributed by postal networks.<sup>29</sup>

As a newspaper focused on foreign news, the *RPZ* was basically no different from its predecessors in the period of Swedish rule in Riga.<sup>30</sup> In terms of printing technology, newspaper publishing changed very little from the 17th until the mid-19th century, when the so-called “express press” and the rotary printing press were introduced. Until the “industrial revolution in the news market”<sup>31</sup> – the coming of the telegraph – the ways of distributing information and media also stayed the same, i.e. mainly

25 O. Grossberg. Die Presse Lettlands. Baltischer Verlag, Riga, 1927, 11; R. Cerūzis. Rigasche Zeitung.

26 J. Eckardt. Riga. Beiträge zur Geschichte des deutsch-baltischen Zeitungswesens. J. Waldkirch, Ludwigshafen am Rhein, 1929; R. Seeberg-Elverfeldt. Dreihundert Jahre deutschbaltische Presse. – Zeitschrift für Ostforschung, 1977, 22, 651–670; M. Lux. Das Riga der Deutschen, 111–112. In Oskar Grossberg's (1862–1941) survey of German journalism in interwar Latvia, for example, the *RZ* plays a central role among other German publications, whether they were competitors or not. See O. Grossberg. Die Presse Lettlands, 16–36.

27 A. Semeta. Saksakeelne perioodika Liivi- ja Kuramaal enne 1800. aastat. Uutmoodi ja paremini! Ühiskondlikest muutustest 18. sajandil ja 19. sajandi algul. – Rahvusarhiivi toimetised (Acta et commentationes archive nationales Estoniae) 2018, 2, 33, 159–190; Balti kirjakuultuuri ajalugu I, 160–163; K. Vanamölder. Kuidas täita pool sajandit kestvat pausi?

28 A. Semeta. Saksakeelne perioodika Liivi- ja Kuramaal enne 1800. aastat.

29 W. Behringer. Im Zeichen des Merkur. Reichspost und Kommunikationsrevolution in der Frühen Neuzeit. Vandenhoeck & Ruprecht, Göttingen, 2003; A. Pettegree, The Invention of News: How the World Came to Know About Itself. Yale University Press, New Haven, 2014; News Networks in Early Modern Europe. Ed. by J. Raymond, N. Moxham. Brill, Leiden, Boston 2016; H. Droste. Das Geschäft mit Nachrichten. Ein barocker Markt für soziale Ressourcen. Edition lumiére, Bremen, 2018, 99–102.

30 *Rigische Montags (Donnerstags) Ordinari Post Zeitung* (1680–1681) and *Rigische Novellen* (1681–1710).

31 M. Jensen. Der Nachrichtenmarkt in der Industriellen Revolution. Mohr Siebeck, Tübingen, 2022.

by post on horses, as was the case in the 16th and 17th centuries.<sup>32</sup> Even though the postal network was constantly improved, a galloping horse was still the benchmark of speed.

When the name of the *RPZ* was changed to the *RZ* in 1797, the content and appearance of the newspaper did not change significantly. Reasons for the name change are tied to the systematic implementation of censorship on the distribution of foreign as well as local newspapers in Russia at the end of the 18th century to prevent an influx of revolutionary ideas from the French Revolution (1789–1799).<sup>33</sup> Through the implementation of censorship, the *RPZ* became a publication of foreign news, pre-controlled by the central authorities thus allowing them to gain greater control over the news trade in Riga.

The main function of both the *RPZ* and the *RZ* was therefore to provide foreign news, making it well-suited for researching place-date headings. Apart from foreign news, there was also a small amount of local news published from the outset, for example lists of strangers arriving in the city and announcements aimed at the local public, economic advertisements and lengthier texts for the enlightened circles (*gelehrte Beiträge*); the beginning or end of the year was marked by special addresses in the form of poetry. This kind of mixing of genres was common in political newspapers in the 18th century.<sup>34</sup>

During the 19th century, the proportion of different news sections increased, and the *RZ* developed into a general newspaper. At the end of the century, the *RZ* carried news about political, economic and social events in Riga, the Baltic provinces, the Russian Empire and the world, as well as information about stock markets and market trends, all kinds of local news, opinion pieces, etc.<sup>35</sup> The political stance of the paper largely depended on the editors – in the 1860s, for example, when Julius

32 J. Wilke. Die Zeitung. – Vom Almanach bis Zeitung. Ein handbuch der Medien in Deutschland 1700–1800. Hrsg. von E. Fischer, W Haefs, M. Yor-Gothart. C. H. Beck, München, 1999, 388–402.

33 T. Reimo. Tsensuurist Eestis XVIII sajandil. – Keel ja kirjandus, 1997, 9, 605; Министерство внутренних дел: исторический очерк. Приложение 2, Почта и телеграф в XIX столетии. Министерство внутренних дел, Санкт-Петербург, 1901, 117–121. In 1796, a year before the name change, a censorship commission was set up in Riga by Emperor Paul I to control foreign and domestic printing. The *RPZ* had already been subject to censorship from the summer of 1789 (*RPZ*, 27.07.1789/no. 60), as it was henceforth subsequently indicated on the headline. It has been suggested that the word “political” was removed from the name of the paper to emphasise its neutrality. See D. Jatsenko. Tsensuur ja tsensurid Liivimaa kubermangus aastatel 1894–1917. Master’s thesis. Supervised by K. Brüggemann. Tallinn University, Tallinn, 2022, 23; R. Cerūzis. Rigasche Zeitung.

34 H. Böning. Geschichte der Hamburger und Altonaer Presse. Von den Anfängen bis zum Ende des Alten Reichs. Bd. 2: Periodische Presse und der Weg zur Aufklärung. Edition lumière, Bremen, 2020, 30–35.

35 R. Cerūzis. Rigasche Zeitung.



Eckardt (1836–1908), John Baerens (1834–1884) and Leopold Pezold (1832–1907) were editors, the *RZ* was prominently liberal.<sup>36</sup> Although the paper's position became more conservative under the increasing pressures of Russification, it was not spared from the attacks of the Russian authorities. In 1889, Alexander Buchholtz (1851–1893), then editor of the paper, was forced to resign under pressure from the censors and the governor general of Livonia, and the paper was closed.<sup>37</sup> At the beginning of the 20th century, the *RZ* reappeared for a short period, representing clearly conservative Baltic German political views.<sup>38</sup>

### *Data*

The digital editions of the *RZ* used for this study come from the National Library of Latvia, which holds the most comprehensive collection of German-language periodicals in the Baltic region. The dataset used for this article can be freely downloaded online,<sup>39</sup> and individual issues of periodicals are searchable with the help of the library's web interface.<sup>40</sup> The dataset consists of 18,499 issues from 1802 to 1888 that have been manually segmented into 289,704 articles, totalling over 4.5 million lines of text.<sup>41</sup> Each article in the dataset is supplied with the issue's publication date, a title (by default the first line of the article) a full OCR transcription of the text and a hyperlink to the original image of the page for comparison purposes.

The quality of the OCR transcriptions is variable but can be described as relatively good, with the exception of articles that contain tables or lots of numerical data or special characters. Although a sizeable majority of the transcriptions are quite easily readable for a human, they fall short for treatment with language models, including widespread techniques like NER. This is a common problem with historical data – it is usually very difficult to apply models trained on contemporary data to older corpora. The pitfalls can be found in both OCR errors and the specifics of historical language.<sup>42</sup>

36 M. Lux. *Das Riga der Deutschen*, 111.

37 O. Grossberg. *Die Presse Lettlands*, 34–35.

38 R. Cerūzis. *Rigasche Zeitung*.

39 <https://dom.lndb.lv/data/obj/1062448.html> (accessed 03/01/2023).

40 <http://periodika.lv> (accessed 03/01/2023).

41 Although the *RZ* was published earlier as well, only the issues from 1802 to 1888 have been digitised by the Latvian National Library. The issues published in the early 20th century after the reopening of the newspaper are available as well but are left out of this study as the break is too long and the context afterwards too different.

42 See, for example, E. Bunout, M. Ehrmann, F. Clavert. *Digitised Newspapers – A New Eldorado for Historians?*, 3–4.

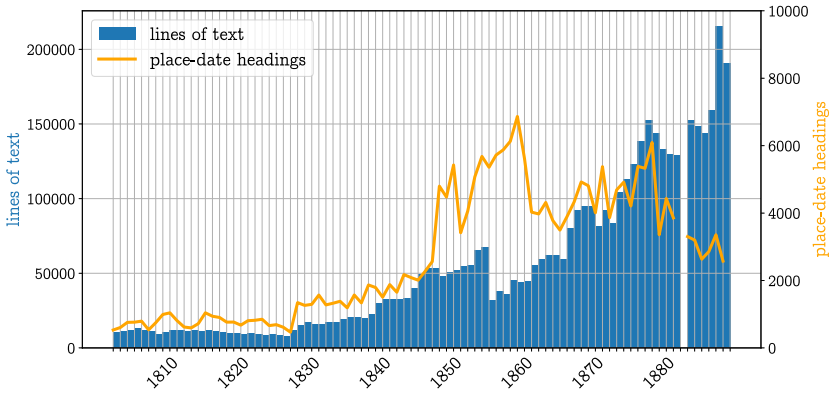


Figure 2. Total lines of text and number of place-date headings per year. There is a strong correlation between the two, except for the last two decades when the contents of the newspaper had become more diverse.

The growth and diversification of the newspaper’s content is also reflected in the data, as the amount of information in the second half of the century is far greater than in the first (see Figure 2). The increase is in the number of articles per issue, length of the articles, and publishing frequency. Information also starts to be increasingly grouped into thematic articles. Importantly, however, place-date headings are also used inside these articles without interruption throughout the century. Their independence from changes in the newspaper makes it possible to analyse long-term trends in the corpus.

To capture the relevant place/date headings from the text, we use a **regular expression**, which is a set of rules for matching a certain pattern in a string of characters. Although our regular expression is written for this specific corpus, the general logic is simple and could easily be tailored for another publication with its own specific traits (and recurrent errors). Table 1 presents a few examples of the many variations of place-date headings that the expression is able to capture. The columns place name,

pattern	place name	date	month	date2	month2	cal.
Dresden, den 12. October.	Dresden	12	Oct			G
Schreiben aus Wien, vom 31. Mai.	Wien	31	Mai			G
Paris, 18. (6.) Mai.	Paris	18	Mai	6		G (J)
Madrid, 7. Dec. (25. Nov.)	Madrid	7	Dec	25	Nov	G (J)
St. Petersburg, den itten August.	St. Petersburg	11	Aug			J

Table 1. Examples of place-date headings in the corpus

date, month, date2 and month2 correspond to the individual elements detected in the pattern.

The regular expression had almost half a million initial matches in the corpus. Of these, a certain number are false positives, mainly cases where it is simpler to throw out the wrong results afterwards than to add a layer of complexity to the search itself.<sup>43</sup> After discarding these unwanted matches, we ended up with 247,468 entries that include a presumed toponym and a date.<sup>44</sup>

## GEOGRAPHICAL COVERAGE

To adequately analyse the spatial information contained in place-date headings, it is necessary to normalise them first, as OCR errors can produce several wordforms of the same toponym. For this paper, we considered only toponyms that were detected at least 20 times in the corpus, and assigned ‘correct’ surface forms to them.<sup>45</sup> For example, the forms *Verlin*, *Berlw*, *Bertin*, *Berlm*, *Berti*» and *Berliu*, as well as some others, were all collapsed into *Berlin*.<sup>46</sup> Although wordforms that occur less than 20 times were left out, they are of little statistical relevance in the context of this study, as the 351 verified places with at least 20 occurrences make up over 91% of the matches that were left after removing false positives.

To exploit the geographical nature of our data, we tied each of these 351 places to coordinates. Most of the places were linked automatically with open source data from GeoNames that includes historical and parallel names for millions of places in the world, the rest were added manually.<sup>47</sup> This allows us to look not just at the relative proportion of a given place name, but also to visualise the spatial dynamics of the 19th century press as seen from Riga.

43 We identified over 300 such false positives. For the most part, they are words that can appear in the exact same form as the toponyms and are accompanied by a date. Examples include weekdays (for example Dienstag, den 9. November), organisations (Aus dem Rathhaus vom 27. Februar), documents (Protokoll vom 29. März), but also currencies, weights and more.

44 Sometimes, OCR errors can lead to impossible dates like den 71. Mai (instead of 11) or 80. Jan (instead of 30), all of which were discarded.

45 ‘Correct’ is a relative term here, as many of the cities in our dataset have different official names today. Both for technical reasons and in order to stay close to the sources, we opted for the names by which the cities were known in German in the 19th century. Moscow thus becomes Moskau, Naples is Neapel, Tallinn is Reval, etc. Hereafter, we will use these forms in graphs, but modern English forms in the text for the sake of readability.

46 The complete list of surface form replacements can be found in [https://github.com/krkyger/rigasche-zeitung/blob/main/data/places/placename\\_replacement\\_dict.json](https://github.com/krkyger/rigasche-zeitung/blob/main/data/places/placename_replacement_dict.json)

47 Specifically, we used the dataset of all cities in the world that currently have at least 1,000 inhabitants, which can be found at <http://download.geonames.org/export/dump/> (accessed 03/01/2023).

In the whole dataset, the most represented place is Paris with 27,886 headings, followed by London (22,603), Berlin (22,449), St. Petersburg (20,560) and Vienna (16,192). Together, these five places make up 47% of the data, meaning that about half of the news in the *RZ* came from five European capitals. It should be noted that the relative importance of these cities grows in time: before 1860, they make up about 42% of the news, whereas after 1860, their combined share is already 52% (see Figure 3).

Out of the 351 validated origin places, the vast majority are located in Europe, with German-speaking areas being the most thoroughly represented (see Appendix A). Areas of the Russian Empire are considerably underrepresented, even after the *RZ*'s transformation into a general newspaper in the middle of the century. The most prominent

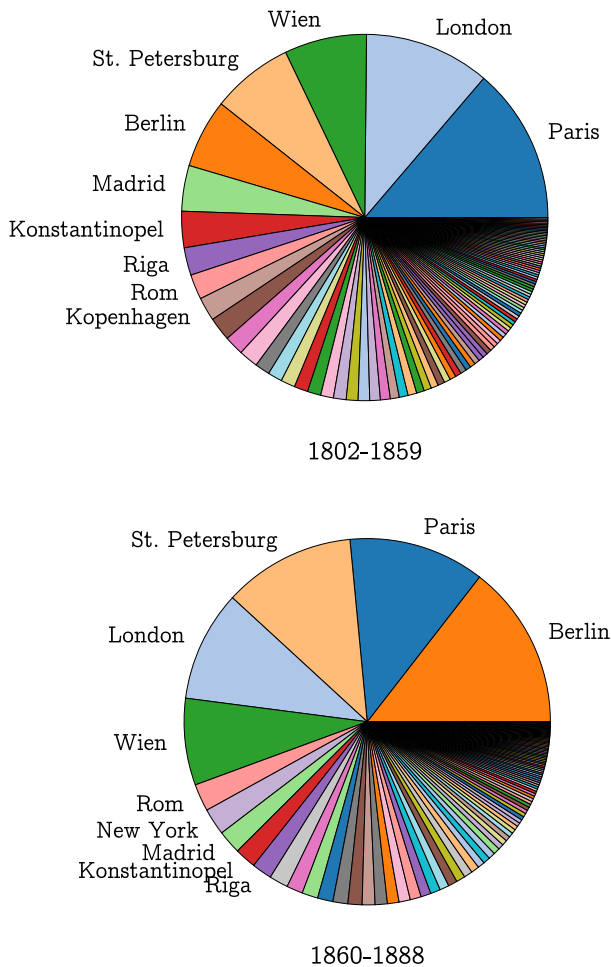


Figure 3. Distribution of places of origin, 1802–1859 and 1860–1888.

places located outside of Europe are New York (3,805), Washington (750), Mumbai (511), Cairo (365), Hong Kong (314), Calcutta (293), Smyrna (311), Rio de Janeiro (274), Algiers (291) and Tbilisi (215). For English, Spanish, French, Russian and other areas in Europe, the capitals clearly dominate the news flow. This may indicate either a lack of access to local news sources outside of German-speaking areas, with the capitals of these more centralised states acting as a sort of filter of outgoing information, a lack of interest in this information, or both.

The change in importance of some origin places in the data can be indicative of both long-term developments as well as specific historical events. One example is Berlin: starting in sixth position between 1802 and 1820 and dropping to 12th between 1820 and 1840, it starts its steady climb in the 1840s, eventually becoming the most frequent source of news. This is most likely tied to the growing political and economic importance of the German capital in Europe, especially after the unification of Germany in 1871. Some cities have short surges of frequency in the data: the importance of Turin, for example, grew remarkably in the 1850s and peaked during the Italian War of 1859, when it was the 3rd most mentioned place in the *RZ* during that year. By the 1870s, however, news from the city had all but disappeared from the content of the newspaper.

The frequency of place names in news headings offers an interesting outlook for further research. For example, with additional data (historical population, distance from Riga, capital status, language, etc.) it would be possible to model the frequency of place names in the newspaper in order to determine the most important factors that contributed to the presence of a given city in the *RZ*. Outliers that do not fit the model would then offer especially interesting historical insights.<sup>48</sup> Another possibility is to study the co-occurrence of place names in the dataset – especially in the pre-telegraph period – to try to determine which news usually arrived in the same batch, thus describing the functioning of the postal system.

## NEWS SPEED

The publication cycle of the *RZ* was closely linked to the rhythm of the postal system. In the late 18th, and 19th century until the construction of railway connections in the 1860s, news arrived in Riga from five main

<sup>48</sup> Such methods have been applied for example to the representation of Chinese cities in modern media: W. B. Loo. The Hierarchy of Cities in Internet News Media and Internet Search: Some Insights from China. – *Cities*, 2019, 84, 121–133. The authors would like to extend their thanks to Mikhail Tamm for a fruitful discussion on this subject.

directions<sup>49</sup> 1) From the west, the postal route passed through Memel (Klaipėda) and through the Baltic coastal cities to Hamburg;<sup>50</sup> 2) mail also arrived from Curonia, Lithuania, Poland and elsewhere in the south along the same route;<sup>51</sup> 3) in the north, there was a connection with the province of Estonia; 4) a north-eastern postal link between St. Petersburg and the west also ran via Riga; 5) an eastbound postal service ran via Daugavpils (Dünaburg), connecting Riga to Moscow and other Russian inland provinces (see Appendix B).<sup>52</sup>

In 1861, the railway line connecting Riga and Daugavpils was opened and the first dedicated mail wagon arrived on 16 November 1867.<sup>53</sup> Over the next decade, new railway lines were opened and mail was increasingly transported by train. This meant some changes to existing postal routes: for example, instead of the old St. Petersburg–Narva–Tartu–Riga route, mail from the capital was sent via Daugavpils in the east. Mail also started to be transported by rail on a daily basis (see Appendix B). We know that the directions and the schedule of the postal connection had a direct impact on the publishing of the *RZ*. Until 1828, the *RZ* was published twice a week, then three times a week from 1828 to October 1843. From November 1843 until its closure, the paper was published six times a week. Both times, the increase in frequency immediately followed a change in the westward postal service.<sup>54</sup>

This indicates that news was published as soon as possible upon arrival. Although it might not have been so in every case (see for example the last part of this section), data exploration and analysis show that it

- 49 About postal directions and the history of the post in Riga, see A. Вигилев. История отечественной почты. Радио и связь, Москва, 1990; P. Pētersone. Riga als ein Knotenpunkt im schwedischen Post- und Verkehrssystem um die Ostsee im 17. Jahrhundert. – Der Westfälische Frieden von 1648. Wende in der Geschichte des Ostseeraums für Prof. Dr. H. C. Herbert Ewe zum 80. Geburtstag. Hrsg. von H. Wernicke, H.-J. Hacker. Verlag Dr. Kovač, Hamburg, 2001, 401–415. A good overview is also provided by the annual calendars with postal information, for example: Livländischer Kalender auf das Jahr... 1802–1890, see Appendix B.
- 50 *Die Post aus Deutschland über Memel kommt an im Sommer Dienstags und Sonnabends, in der Nacht, im Herbst und Frühjahr erst Mittwochs und Sonntags, auch nach Beschaffenheit der Wege und Ströme, wohl noch später, bringet Briefe mit von allen Orten aus Deutschland, Holland, England, Frankreich, Schweden, Dänemark etc.* Livländischer Kalender auf das Jahr 1804. Riga, Gedruckt und zu haben bey J.C.D.Müller, privilegirtem Krons- und Stadtbuchdrucker, [1803].
- 51 *Zugleich mit dieser Post kommt auch die Post aus Kurland und Littbauen zweymal in der Woche an, bringt Briefe mit von Warschau, Krakau, Wilna, Grodno, aus oesterreich und Italien.* Livländischer Kalender auf das Jahr 1804.
- 52 For example, in 1788 and earlier, the direction of Dünaburg is not yet marked. Livländischer Kalender auf das Jahr 1799. Riga, gedruckt und zu haben bey Georg Friedrich Keil, privilegirter Buchdrucker [1788]. *Die Post aus Moskau, über Smolensk, Polozk und Dünaburg* Livländischer Kalender auf das Jahr 1804.
- 53 J. Krauklis. *Komunikāciju leksikons*. Jumava, Riga, 2004, 25.
- 54 In October 1843, the *RZ* announced that due to “the arrival of foreign post, henceforth five days a week, the paper will be published daily in order not to conceal political news from readers for too long”. *RZ* 30.10.1843.

usually was. This leads us to the second inherent element in the place-date headings: time. The fact that news retained its original date when passing through the European news network means that it is possible to measure the time it took for it to reach Riga from its point of origin. Although the duration between compiling and publishing a piece of information does not tell us which path it took to Riga, we can make logical assumptions based on our knowledge of the postal, railway and telegraph networks and other sources. In any case, it is a fact that we can observe how old a message was when it reached readers in Riga. For the sake of generalisation, we will call this phenomenon ‘news speed’ in this section.

To analyse news speed, we must extract and systematise the relevant datetime information contained in the regular expression matches. As we saw from Table 1, place-date headings can take different forms. In addition to obvious differences in wording (den 20sten vs 20. etc.), the *RZ* used two parallel calendar systems, Julian and Georgian. The place-date headings were thus usually printed in the calendar system that was used in the place of origin of the message, or with one system in parentheses (see the “cal.” column in Table 1 for the different possibilities of conveying dates). We parsed all the captured headings with a function that decides, based on a number of logical steps, how to combine the available data with a correct Julian date (we opted for Julian in order to match the newspaper’s own system).<sup>55</sup> After that, we subtracted the date of a news item from the publication date of the issue, resulting in the time in days that it took for the message to arrive in Riga. We also discarded invalid dates and those more than a hundred days older than the newspaper in which they were published.<sup>56</sup>

Figure 4 shows news speed dynamics from select locations in Europe and the world.<sup>57</sup> The overall change throughout the century

55 Usually, the calendar system used in a heading can be determined by 1) comparing it to the publication date of the issue and 2) comparing the different dates in the heading to each other (when one is in parentheses, for example). In some cases, however, it is not possible to decide based only on these rules. For example a heading “Paris, January 22” published on January 24 could be in both Julian and Georgian. Further complications arise from the fact that a message from Paris could have easily reached Riga within 2 days in the 1870s, but not in the 1830s. To completely avoid miscalculations caused by calendar errors, we tied each of the 351 verified places to their respective calendar systems to make the correct distinction in cases like this.

56 In this case, it is reasonable to expect that we are dealing with information that has become relevant at a later time and cannot be strictly considered news. In this sense, we are only focusing on short-term propagation as described by Salmi et al. in *The Reuse of Texts in Finnish Newspapers*. Basically, we are assuming that texts with a delay under a hundred days were printed immediately as they became known or available in Riga.

57 To calculate the average delay for a given location, we used a smoothing algorithm to remove outliers.

is drastic, especially for faraway locations. For example, it took almost two months for information from Constantinople to reach Riga in the first decades of the century, about a month in 1840 and less than a week in the 1870s.<sup>58</sup> The data is similar for Lisbon and New York. In general, we can identify two clearly distinct stages of communication, the pre-industrial and the industrial.<sup>59</sup> Between them, there is a decade-long period of transition that roughly coincides with the 1860s.

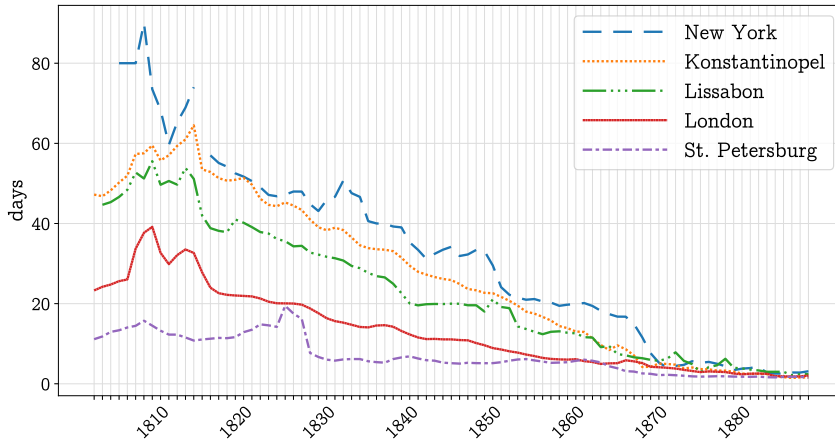


Figure 4. News speed from given locations (shown with a rolling average of 3 years)

Until the 1860s, the bulk of the news reached Riga by post, with a smaller share probably arriving with ships or individual travellers by land. During these six decades, communication speed was in slow but continuous improvement. This can be attributed to the incremental changes in the network, for example, an improved postal connection in Germany that eventually effects all the information that passes through it. These improvements have their eventual limits, however. By 1830, for example, news from St. Petersburg (ca. 600 km from Riga along the postal road) was 5–6 days old on average and stayed this way until the establishment of the railway connection via Daugavpils in the 1860s.

<sup>58</sup> Of course, the whole duration is not spent ‘on the road’, the described delay also includes the time it takes for an event to be broadcast as news from a given location and the time to be printed after reaching Riga, for example. While the former is indeterminable, we can have a guess at the latter. Until 1828, the printing delay could have been up to 4 days in some cases, because the newspaper was only published on Wednesdays and Saturdays (in a case where the message arrived right after the publication of Saturday’s issue, for example). From 1843, the maximum printing delay was two days but probably below one on average, as by then the newspaper was published six times per week. For nearer locations, however, this printing delay of 1–2 days can still make up a noticeable part of the total news delay and must thus be kept in mind.

<sup>59</sup> Marek Jensen also centres his analysis around the term ‘industrial’. See M. Jensen. *Der Nachrichtenmarkt in der Industriellen Revolution*.



Our data also allows us to examine news speed from specific places in more detail. Figure 5 presents individual news from London. At the beginning of the century, we can observe the regular rhythm of news being heavily disrupted by the continental blockade (1806–1814). In the following decades, news speed is in incremental but continuous improvement until it reached near immediacy around 1865. The telegraph did not replace other means of communication, however, and we can note the continued improvement of ‘slower’ news up to the end of the observed period.

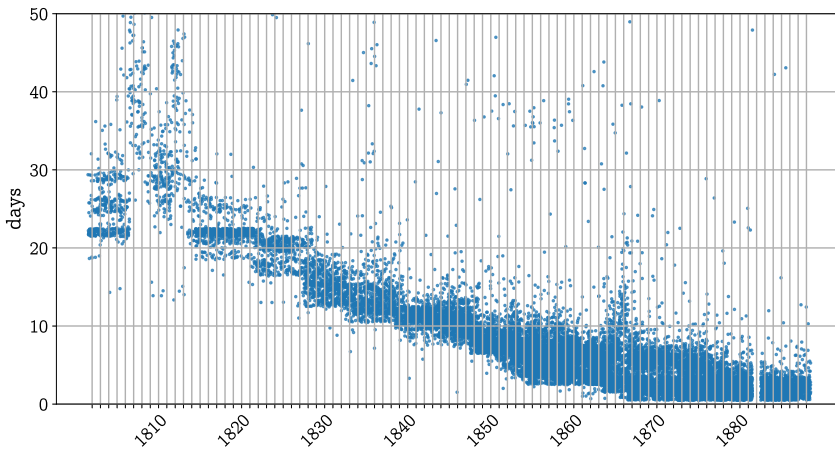


Figure 5. All news from London plotted by speed

Figure 6 presents a map where physical distances are transformed according to the average time needed for news to arrive between 1802 and 1859, i.e. in the pre-industrial era.<sup>60</sup> This shows a sort of intuitive overview of the European communication space as seen from Riga. As we can see, north-western Europe is pulled closer to Riga, as the lines of communication leading there were faster and more developed. At the same time, news from the Russian empire and Scandinavia travelled much slower on average. For example, news from Madrid, Paris or London travelled at a speed of about 100 km per day on average between 1802 and 1859, while news from Moscow and St. Petersburg travelled at about half that speed.<sup>61</sup>

<sup>60</sup> Methodological note: for transforming the map, the average movement speed for all locations for the given timeframe was first calculated. The average times for each location were then multiplied by this global average speed, resulting in “new distances” between Riga and each point. The map and the grid were transformed with the DistanceCartogram plugin for QGIS. Simply put, the locations from which news arrive faster than the global average speed are moved closer to Riga while locations from which news arrive slower are pushed away. The original map projection is shown under the transformation in grey. The authors thank Emilien Arnaud for helpful tips regarding the calculations.

<sup>61</sup> These numbers need to be taken with a grain of salt as they are calculated using direct

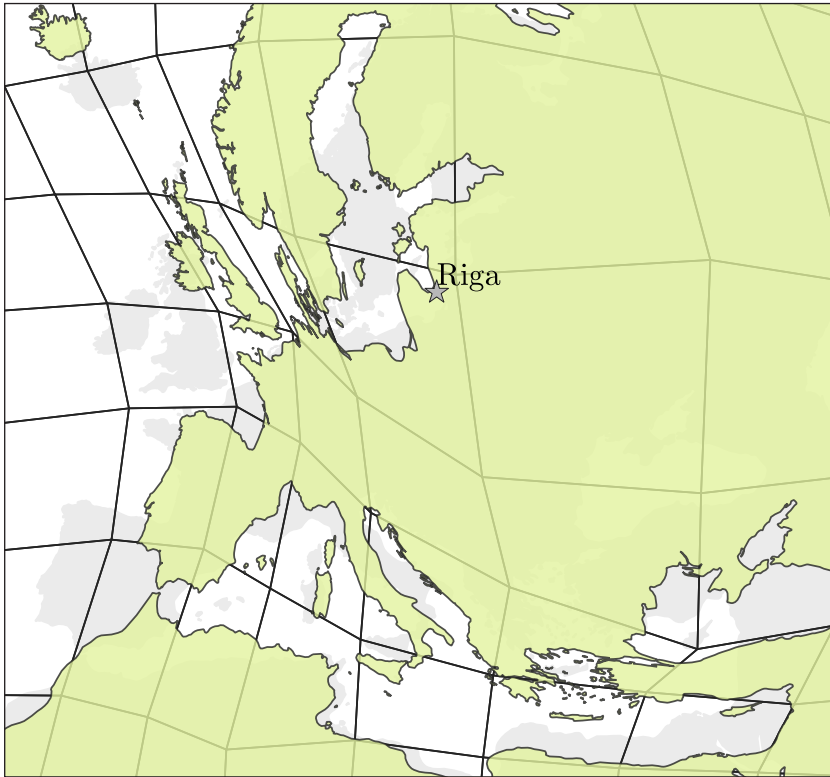


Figure 6. Spatial visualisation of time for news to reach Riga, 1802–1859.

The incremental improvements in the postal system continued until the arrival of two breakthrough technologies: the railway and the telegraph. Although Riga was connected to its port Bolderāja by electric telegraph as early as 1852, the line was of local importance.<sup>62</sup> A telegraph connection was established between Riga and St. Petersburg in 1855 and with Germany in 1857.<sup>63</sup> In the 1860s, telegraph lines continued to be established from Riga to other settlements in Estonia and Livonia, and connections with foreign countries also became more dense.<sup>64</sup> By the 1860s, several telegraph companies were operating in Riga.<sup>65</sup>

distances rather the actual distances by road, but they are nevertheless indicative about the general quality and speed of the (postal) connection

62 The project was carried out by the Prussian engineer Werner von Siemens (1816–1892) at a total cost of 14,558 silver rubles. It was the first and the longest civil telegraph connection in the Russian Empire and was intended to provide Riga merchants with fast and continuous information about activities in the harbor: arriving and departing ships, weather and ice conditions, etc. *Der Rigasche Börsen-Comité in den Jahren 1816 bis 1866*. Druck der Livländischen Gouvernements-Typographie, Riga, 1866, 47–48; G. Hernmarck. *Erinnerungen aus dem öffentlichen Leben eines rigaschen Kaufmanns*. Guttentag, Berlin, 1899, 10.

63 G. Hernmarck. *Erinnerungen*, 10.

64 J. Krauklis. *Komunikāciju leksikons*, 61–64.

65 *Ibid.*, 62.

The telegraph was perceived as an important speed increase in the flow of information, but it did not affect the press immediately. In March 1861, an anonymous article appeared in the *Rigasche Stadtblätter*, a publication that was in many ways in competition with the *RZ*, criticising the *RZ* for being too slow in transmitting information in the new telegraph era.<sup>66</sup> The article stated that if the *RZ* wanted to be a political newspaper, it was important for its readers to receive the latest news as quickly as possible.<sup>67</sup> The article also advised the *RZ* to make use of the telegraph, because “what happens in London this morning will be known in Hamburg this afternoon and in Berlin and St. Petersburg early the next morning at the latest. In contrast, we here in Riga, situated between the latter two cities, usually only find out on the third or fourth day”.<sup>68</sup> The article goes on to say that since the cities in the Russian provinces were not allowed to use political news sent directly by telegraph from abroad, the *RZ* should obtain it from St. Petersburg “by wire”.<sup>69</sup> By “London news” the author most likely meant stock market information, which was particularly time-sensitive.<sup>70</sup> Although O. Grossberg states that the criticism was effective and later in the same year the *RZ* was already transmitting reports from European capitals that were “only a day old”,<sup>71</sup> our data shows that news from most prominent locations did not reach such speeds before 1865.

The reasons for the relatively long transitional period between establishing a telegraph connection and the news speeding up remains open for further research. Price was of course an important factor. In 1869 it cost 50 kopeks to send a message of up to 25 words; 75 kopeks

66 *Rigasche Stadtblätter*, 23/03/1861/no. 12.

67 Die “*Rigasche Zeitung*” soll vorzugsweise ein politisches Blatt sein, was man um so mehr beanspruchen kann als sie das einzige derartige Blatt in unserer Stadt, einer Stadt von circa 80,000 Seelen ist, und doch wird, unserer Ansicht nach, für die Politik in gewisser Beziehung gegenüber den diversen Leitartikeln und Localnotizen nicht Entsprechendes genug gethan. Eine der Hauptanforderungen an ein politisches Blatt – will es nämlich das Interesse des sich mit Politik beschäftigenden Publikums erhalten – ist jedenfalls eine möglichst schnelle Mittheilung der neuesten Nachrichten. *Rigasche Stadtblätter*, 23/03/1861/no. 12.

68 Durch Hilfe des Telegraphen leisten die auswärtigen Zeitungen – und selbst die St. Petersburger – in dieser Beziehung gegenwärtig Außerordentliches. Was sich heute Morgen in London ereignet, weiß man am Nachmittage bereits in Hamburg und spätestens morgen früh schon in Berlin und St. Petersburg; – wir dagegen hier in Riga, das doch in der Mitte zwischen den beiden letztgenannten Orten liegt, erfahren es in der Regel erst den dritten oder vierten Tag. *Rigasche Stadtblätter*, 23/03/1861/no. 12.

69 Freilich gestattet der Umstand, daß nach den Provinzialstädten Rußlands keinerlei politische Nachrichten aus dem Auslande telegraphisch gemeldet werden dürfen, unserer hiesigen Zeitung nicht eine directe telegraphische Verbindung mit dem Auslande; dennoch möchte auch bei uns eine mehr beschleunigte Mittheilung der neuesten Nachrichten zu ermöglichen sein, wenn die Redaction der “*Rigaschen Zeitung*” dahin Veranstaltung treffen wollte, daß ihr das Neueste und Wichtigste aus den Tagesereignissen des Auslandes von St. Petersburg per Drath zugänglich gemacht würde. *Rigasche Stadtblätter*, 23/03/1861/no. 12.

70 O. Grossberg, *Die Presse Lettlands*, 22.

71 *Ibid.*

to send a message of 26 to 50 words; and 1 rouble 25 kopeks to send a message of 51 to 100 words. Each additional word then cost one additional kopek.<sup>72</sup> By way of comparison, a yearly subscription to the *RZ* in Riga in the same year cost a total of eight roubles, and a single issue of the paper 10 kopeks.<sup>73</sup> It cost just as much (10 kopeks) to send an ordinary letter within the Russian Empire.<sup>74</sup>

It also seems that railway connection which first reached Riga in 1861 and connected the city to both St. Petersburg and Warsaw through Daugavpils might have played an equally important role in reducing the delay of news from foreign cities. In any case, it can be said that by the 1870s, Riga was fully integrated in the ‘industrial’ news network. As seen from Figure 4, the average time for news to reach Riga from the example cities drops to just a few days by that time. This is valid for a large number of the 350 cities in the dataset, especially the most frequent ones.

A notable exception in this new context, however, were local towns in the Baltic provinces. Although the speed of news from these places improved steadily, it did not experience a change like that of the major metropolises. Eventually, news from London and Berlin started to reach Riga more quickly than news from Tallinn and Vilnius, for example, even though the latter are several times closer. The changes of the 1860s are thus not only indicative of the layout of communication networks but also tell us something about the changing nature of news itself. It can be supposed that information relayed by telegraph carried an inherent sense of urgency that was lacking from information circulated by more traditional means, such as word of mouth, letters from local correspondents, etc. The somewhat high durations for several nearby places (for example 5 days on average for news from Tallinn) do not mean that information could not be carried faster than that (there was a telegraph line between Riga and Tallinn as well), but that regional news existed in a different information space with its own, less urgent, rules of circulation.

72 Der Rigasche Börsen-Comité in den Jahren 1866–1872. Druck der Livländischen Gouvernements-Typographie, Riga, 1873, 133.

73 *RZ*, 02.01.1869/no. 1.

74 Livländischer Kalender auf das Jahr nach Christi Geburt 1869. Müllersche Buchdruckerei, Riga, [1868], 24.

## CASE STUDY: THE CRIMEAN WAR

To better demonstrate the potential of place-date headings, we will focus on a specific historical event, the Crimean War (1853–1856). The Crimean War was a conflict that pitted Russia against the Ottoman Empire, which was aided by Britain and France. The war was not constrained to the Crimean Peninsula, the location of some of the most important military clashes (including the siege of Sevastopol), but also involved the surroundings of the Black Sea and even the Baltic Sea. The Crimean War is especially fitting for a case study as it was one of the first modern wars that saw the extensive use of the telegraph and the railway and was brought to the reading public by war correspondents and the free press in Western Europe. It eventually also became the impetus for the Russian Empire to modernise its infrastructure after defeat by more ‘developed’ opponents, leading in turn to the changes described in the last section.<sup>75</sup>

To integrate the content dimension of the newspaper, we applied topic modelling to the texts of the news. For this, we assumed that the text that follows a place-date heading is the body of the news itself. It can continue up until the next place-date heading in the same article or to the end of the article, if it is the last one. Although this approach can sometimes produce errors<sup>76</sup>, it is sufficient for the case study at hand. For topic modelling, we chose to use the relatively recent algorithm *top2vec*, which uses vector similarity to find topics.<sup>77</sup> As output, *top2vec* produces clusters of texts that are characterised by certain words. The algorithm does not ‘know’ anything about its input texts, but a historian can easily recognise that a topic most characterised by the words ‘ship’, ‘stranded’, ‘leak’, ‘skipper’ and ‘load’ might be related to sailing, for instance.

By modelling all the text segments that follow our place-date headings – i.e. the news itself – we can acquire a relatively adequate picture of the thematic makeup of the newspaper. Thanks to place-date headings, the topics can be juxtaposed with the spatial and temporal dimension of the news.

75 R. Cvetovski. *Modernisierung durch Beschleunigung. Raum und Mobilität im Zarenreich*. Campus Verlag, Frankfurt, New York, 2006, 212.

76 When one of the place-date headings in an article has gone undetected by the regular expression (for example, due to some OCR error), the text of its message is merged with the previous one. If the two have different content, the performance of the topic model can be affected.

77 D. Angelov. *Top2Vec: Distributed Representations of Topics*, <https://doi.org/10.48550/arXiv.2008.09470> (accessed 03/01/2023).

Modelling all 232,366 texts following our headings resulted in 385 topics. The resulting model is surprisingly adequate, despite the varying quality of the data and the pitfalls described above. Most topics can be understood with just a glimpse of their characteristic words and their temporal distribution. Some topics recur throughout the century. Out of many examples we can cite news about the harvest, royal visits, epidemics and the ratification of treaties. Other topics are tied to a specific region or even a historical event (the US Civil War, for example). Some topics are more dependent on the specific structure of the article they are found in, for example lists of stock prices and weather telegrams are good examples of this.

To identify the discourse on the Crimean War in the corpus, we first looked to see if any topics had a noticeable frequency surge around 1853–1856. Then, we looked at the characteristic words and example texts of these topics to determine if they had a connection to the events of the war. We found four such topics, and gave them provisional names (see also the word clouds in Appendix C):

- **War events** – these texts are most directly related to the events of the war, usually relaying news about battles, sieges and the movement of troops.
- **Navy** – this topic consists of news about the movement of war-ships and other naval activity, but is mainly concerned with the Baltic Sea, as explained shortly.
- **Recruitment** – this topic includes news about the recruitment of troops in Britain, France and Switzerland.
- **Reinforcements** – these texts concern the assembly of British and French units, as well as their arrival and stationing in the theatre of war.

In 1853–1856, there are 2702 news items in total that belong to any of these four topics. Their frequency, shown on Figure 7, is closely related to the course of the war. The appearance of the war events topic coincides with the first hostilities in the summer of 1853, and its frequency drops significantly after the signing of the Treaty of Paris at the end of March 1856. The dip in the winter 1855/1856 correlates with the halt of ground operations due to material shortages, while the topic recruitment surges at the same time. News about the navy sees a sharp increase in 1854, when the first naval hostilities in the Black Sea take place and the allied fleet enters the Baltic Sea.

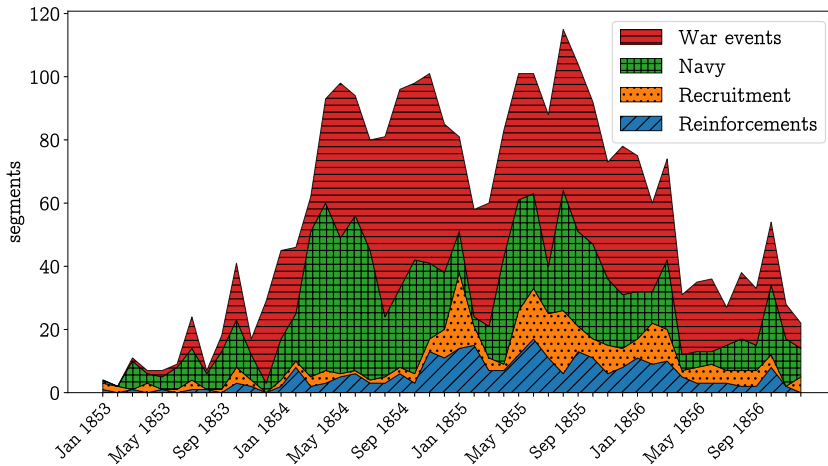


Figure 7. Temporal distribution of the topics related to the Crimean War

By looking at the topics more closely and examining some examples, we can determine where the *RZ* received its information on the war and from whose perspective. For example, almost half of the news in the war events topic comes from Constantinople and Vienna. A closer look reveals that they are often part of the same information chain: news from Vienna periodically cites its sources in other locations. It seems that a great number of reports on war events around the Black Sea found their way into the European press through Vienna. From there, news was passed around in the central areas of the press network before reaching Riga. A report published in *RZ* in June 1854 is a good example of this:

Paris, 27 June. (St.-A.) Today's *Moniteur* reports, according to a telegraphic dispatch from Vienna on the 26th: The Coronini Division has orders to hold on to march via Giurgewo (?) to Bucharest.<sup>78</sup>

The latest information on the war thus reached readers in Riga predominantly through the press of Russia's rival powers. An article from August 1855 lists the Crimea correspondents of *The Daily News*, *The Times* and *The Morning Advertiser*, as well as the French *Moniteur* as its sources.<sup>79</sup> Most times, the *RZ* printed this information with no further commentary, although not always. A good example is a message from October 1854, doubting the credibility of information presented in the French press:

<sup>78</sup> Paris, 27. Juni. (St.-A.) Der heutige „Moniteur“ meldet, einer telegraphischen Depesche zufolge, aus Wien vom 26. d.: Die Diviston Coronini hat Ordre, sich bereit zu halten, um über Giurgewo (?) nach Bukarest zu marschiren. *RZ*, 19/06/1854/no. 139.

<sup>79</sup> *RZ*, 29/08/1855/no. 198.

Paris, 19. October. (A. A. Z.) Paris newspaper reports on Crimea should be given as little credence as possible. Their sole purpose is to keep the stock exchange in good faith, at least as long as possible. The strictest orders have been given to various papers, for example, with regard to numbers; at present we have already reached 120,000 allied troops, while they hardly number more than 75,000. At Alfort, 60,000 men have landed; the battle at the Alma, illness, capture, etc., have resulted in a loss of at least 8,000 men, to which 23,000 reinforcements may have been added. The siege army therefore does not exceed 75,000 men.<sup>80</sup>

It is unclear if these disclaimers were added by a local correspondent, the editors, or the censors. Either way, the fact that the *RZ* felt the need to argue with information presented in the Parisian press testifies to a lack of better sources. The press in Riga also made use of the greater freedom of the press in Western Europe, where criticism of the war was increasingly voiced (especially in Britain).<sup>81</sup> News about the logistical and management failures of the Anglo-French forces, for example, was often reprinted.<sup>82</sup>

Western news also supplied Riga with information on more urgent and geographically close matters than the relatively distant theatre of war around the Black Sea. News in the navy topic can be divided into two main groups, one that describes naval movements on the Mediterranean and Black Seas, and the other on the Baltic Sea. Readers in Riga followed news about the approaching threat in 1854 closely, when the allied fleet was gathering near the Sound in order to attack the Russian navy and strike its capital. The first half of the year is characterised by a large amount of news from Copenhagen about movements of the allied navy, as in this message from April:

Copenhagen, April 11. (T. D. d. C. B.) Admiral Napier has left here for Kjööge, and will leave Kjööge Bay today with 23 ships. It is still unknown where he will sail to.<sup>83</sup>

80 *Paris, 19. October. (A. A. Z.) Den Berichten der Pariser Zeitungen über die Krim ist so wenig Glauben wie irgend möglich beizulegen. Sie haben einzig zum Zweck, die Börse, wenigstens so lange als möglich, bei gutem Glauben zu erhalten. Die strengsten Befehle sind z. B. den verschiedenen Blättern in Betreff der Zahlenverhältnisse gegeben, wir sind gegenwärtig bereits bei 120,000 Mann alliirten Truppen angelangt, während dieselben schwerlich mehr als 75,000 Mann zählen. Beim Alfort sind 60,000 Mann gelandet; das Gefecht an der Alma, Krankheit, Gefangennahme c. haben einen Verlust von wenigstens 8000 Mann herbeigeführt, wozu 23,000 Mann Verstärkungen gestoßen seyn mögen. Ueber 75,000 Mann beträgt daher das Belagerungsheer nicht.* *RZ* 16/10/1854/no. 241.

81 O. Figes. *Crimea: The Last Crusade*. Penguin, London, 2011, 367–370.

82 See for example *RZ*, 17/06/1855/no. 137 (about the supposedly bad treatment of the allied forces by the Turks); *RZ* 14/04/1854/no. 84 (about various problems of the allied garrison in Constantinople); and *RZ* 28/08/1854/no. 199 (about discipline and health problems in Constantinople).

83 *Kopenhagen, 11. April. (T. D. d. C. B.) Admiral Napier ist von hier nach Kjööge abgereist.*



Given the results of topic modelling, the distribution of locations for these topics, as well as example texts, it seems that the inhabitants of Riga were effectively part of the information space of their adversaries during the Crimean War. French and English efforts in the domain of communication – for example the extension of telegraph lines to the Black Sea and to the Crimean Peninsula in late 1854 and 1855 –, press freedom and the centrality of London and Paris in the European press network hopelessly overshadowed the weak and slow connection that Russia itself had to its borderlands.

In general, the Russian public only had access to the most basic news about the war, and even this was severely outdated.<sup>84</sup> Looking at all the news that falls into the four relevant topics, we could detect no more than 12 items (0.4%) that had originated from St. Petersburg, and even these were nonspecific. Instead, the imperial capital continued to broadcast its usual material: imperial ukases, assignment of military promotions and decorations, etc. This kind of information made up more than 50% of all the news that originated from St. Petersburg during this period.

The scarcity of reliable news, as well the efforts of the Russian state to strictly control the narrative of the war made the educated public turn to rumours and the like.<sup>85</sup> In this context, it is especially interesting to see news from rival states being printed in Riga regularly and with no to little editorial interference. Here, the role of censorship requires further historical research. It is unclear why accurate reports on the war could be printed in Riga while they were supposedly banned in St. Petersburg and whether they also eventually made their way into the capital.

## CONCLUSION

Place-date headings hold considerable potential for researching historical news circuits. They contain both spatial and temporal information and can be captured with relative ease. Their usage extends from the first newspapers well into the 20th century, meaning that they can be used to analyse trends over very long stretches of time. Studies that use

*und wird noch heute mit 23 Schiffen die Kjöge-Bucht verlassen. Es ist noch unbekannt, wohin derselbe segeln wird.* RZ 03/04/1854/no. 77. These updates from Copenhagen and other Baltic port cities continue throughout the war and their frequency is directly related to the allied naval activity on the Baltic Sea. See also for example RZ 11/03/1854; RZ 16/04/1854/no. 57; RZ 30/04/1855/ no. 98, etc.

<sup>84</sup> O. Figs. Crimea, 378.

<sup>85</sup> Ibid.

newspapers to study news networks normally require several publications in order to make a comparison, but an approach by place-date headings makes it possible to make relevant conclusions within the bounds of a single publication. Using the example of the *Rigasche Zeitung*, we mapped the geographical area that was represented in news during the 19th century and described how information reached the reader more and more quickly during this period. These dynamics in the press are inseparable from underlying communication technologies, be it the post, the telegraph or the railway. By combining place-date headings with topic modelling, we sketched how news about the Crimean War reached readers in Riga. More generally, we also described Riga's position in the European news network, specifically how closely it was tied to German-language areas and Western Europe in general.

The methods presented in this article are relatively robust and applicable to other newspapers, regardless of their language. We suggest that place-date headings can be very helpful in researching cross-lingual information flows, especially in the context of political and military news. They can also be used to complement other methods such as text reuse detection. Although we relied on the relative stability of their nature, a dedicated study on the usage of place-date headings, including possible regional conventions and temporal changes, would be beneficial to future research on the subject. Eventually, place-date headings could be used to construct a more complete network of historical news circulation. Such a network could encompass a wider geographical extent, while the travel times of information could be calculated to and from different nodes to reveal long-term trends. The effects of historical events on these dynamics could then be analysed to reach wider conclusions on the nature of historical communication networks.

### *Acknowledgments*

This research has received funding from the Estonian Research Council grant PRG908.

RIIA ASEND 19. SAJANDI  
UUDISTEVÕRGUSTIKUS:  
RAHVUSVAHELISTE UUDISVOOGUDE  
KAARDISTAMINE AJALEHES  
RIGASCHE ZEITUNG 1802–1888

*Krister Kruusmaa, Kaarel Vanamölder*

Artikli eesmärk on selgitada digiteeritud ajalehtede põhjal uusi meetodeid ajalooliste kommunikatsioonivõrgustike uurimisel. Tegu on küllaltki populaarse teemaga – mitmed uurimused on sellel eesmärgil kasutanud teksti ristikasutustuvastust (*text reuse*). Kuna 20. sajandini oli ajakirjanduses tavaks teadete „taaskasutamine“, st info ümbertrükkimine teistest kättesaadavatest lehtedest, on ristikasutustuvastuse abil võimalik võrrelda sarnaseid tekste eri ajalehtedes ning teha järelduisi info liikumise kohta. Samas on tegu tehniliselt keeruka ja palju andmeid vajava protsessiga. Samuti ei eksisteeri seni lahendust, mille abil võrrelda eri keeltes tekste.

Alternatiivne, kuid seni vähe kasutatud võimalus infovõrgustike uurimiseks ajalehtede kaudu on nn koha-kuupäeva päised. Eelmodernses pressis algasid teated tavaliselt koha (linna) nimega, kust need pärinesid, ja info algse kuupäevaga (nt „London, 29. märts“). See teave trükiti muutmata kujul ümber, kui info ühest väljaandest teise liikus. Ehkki koha-kuupäeva päiseid on uudisvoogude uurimisel mõnel määral kasutatud, pole seda tehtud süsteemselt ega pikaajaliselt. Käesolevas artiklis näitame koha-kuupäeva päiste võimalusi nii geograafiliseks kui ka ajaliseks analüüsiks, kasutades Riias välja antud saksakeelse ajalehe Rigasche Zeitung (1778–1889; 1907–1919) tervikkorpust aastatest 1802–1888.

Rigasche Zeitung oli algselt välisuudistele orienteeritud ajaleht, mis sai oma sisu peamiselt postiga saabuvatest Saksa lehtedest ja Peterburist pärinevatest teadetest. Sajandi jooksul kujunes see üheks kõige mõjukamaks Balti päevaleheks, mis kajastas nii välis- kui ka sise-, kultuuri-, majandus- jm uudiseid, lisaks teateid, reklaame. Kasutasime lehe terviktekstidest koha-kuupäeva päiste tuvastamiseks regulaaravaldist ehk sõnejärgnevuste tuvastamise reeglit. Kokku tuvastasime korpusest umbes veerand miljonit koha-kuupäeva päist. Nende hulgast keskendusime 351 kohale, mis kajastusid päiste hulgas vähemalt 20 korda (moodustades kokku üle 90% päiste koguarvust), normaliseerisime kohtade nimekujud ja sidusime need koordinaatidega. Seejärel arvutasime päistes esinevate kuupäevade abil kestuse, mis jäi uudise koostamise ja Riias ilmumise

vahele. Tulemusena oli võimalik analüüsida Rigasche Zeitung'i uudiste geograafilist päritolu ja umbkaudset aega, mille jooksul need Riiga jõudsid.

Uudiste hulgas on selgelt kõige paremini esindatud saksakeelsed alad ning Euroopa suurlinnad Pariis, London, Peterburi, Berliin ja Viin. Aja jooksul suurlinnade tähtsus tõuseb ja eriti selgelt muutub Berliini osatähtsus, mis on vaatlusperioodi viimastel kümnenditel kõige sagedasem uudiste lähtekoht (vt lisa A ja joonis 3). Andmetes on näha ka mõne koha ajutist tõusu – näiteks oli Torino 1850. aastate lõpu Itaalia revolutsiooni ajal üks enim viidatud kohti, kuid enne ja pärast seda väga vähe esindatud. Võrdlemisi vähe on lehes ka uudiseid Vene impeeriumi aladelt (v.a Balti provintsid ja Peterburi).

Uudiste liikumiskiiruse jälgimiseks vaatlesime aastase keskmise kiiruse muutumist perioodi jooksul. Sajandi jooksul toimusid uudiste levimise kiiruses suured muutused. Näiteks jõudsid uudised Konstantinoopolist Riiga sajandi alguses umbes kahe kuuga, 1840ndatel ühe kuuga ja 1870ndatel vähem kui nädalaga. Sama dünaamika kehtib näiteks New Yorgi, Lissaboni ja paljude teiste paikade kohta (vt joonis 4). Tuvastasime kiiruse muutumises kolm etappi. U 1860. aastateni ehk n-ö eelindustriaalsel perioodil jõudis info Riiga peamiselt posti teel ning levimiskiiruse tõus on ilmselt tingitud Lääne-Euroopa postivõrgu parenemisest ning osalt ka telegraafi- ja raudteeliinide rajamisest, mis kaugemaid uudiseid kiiremini kohale toimetada võimaldas. Ühtlasi jõudis info lääne suunast palju kiiremini kohale kui ida, põhja või lõuna suunast (vt joonis 6).

1850. aastate lõpus – 1860. aastate alguses sai Riia telegraafi- ja raudteeühenduse, kuid nende mõju uudiste liikumisele polnud kohene. 1860. aastad olid pigem üleminekuvaeg eelindustriaalse ja industriaalse ajastu vahel. 1870. aastateks olid telegraafi- ja raudteeühendus piisavalt head, nii et enamikest andmestikus esinevatest kohtadest jõudis info Riiga vaid mõne päevaga. Erandiks olid sealjuures kohalikud paigad Balti provintsidest, kust saabuv info tõenäoliselt alati kohe trükki ei jõudnud ja mille viide oli seega enamasti suurem kui Euroopast tulevatel telegrammidel.

Artikli viimane osa on pühendatud juhtumiuuringule Krimmi sõjast (1853–1856). Selleks sidusime koha-kuupäeva päised teemamudeldamisega. Kasutasime algoritmi top2vec, mille abil jaotasime päiste järgnevad artiklitekstid 385 teemakogumiks. Tegu on automaatse protsessiga, mille kvaliteet sõltub tõlgendatavusest, st teemale iseloomulike sõnade, teema ajalise jaotuse jm vastavusest mingile mõttelisele kategooriale.

Tuvastasime neli Krimmi sõjaga seotud teemat: uudised sõjasündmustest (lahingud, liikumised, garnisonid jm Musta mere ümbruses); laevastikust (liikumised Mustal merel, Vahemerel ja Läänemerel); värbamisest (peamiselt Inglismaal, Prantsusmaal ja Šveitsis) ja abivägede liikumisest (vt joonis 7 ja lisa C). Nende teemade ajaline jaotus näitab, kuidas peegeldusid sõja erinevad etapid Rigasche Zeitungis. Uudiste päritolukohad viitavad, et peaaegu kõik sõjauudised jõudsid Riiga läbi Lääne-Euroopa inforuumi, näiteks Viini, Pariisi või Londoni kaudu. Samas oli Venemaal puudus usaldusväärsetest sõjauudistest ja vajaks edasist selgitamist, kas ja mis määral liikus Riias trükitud info edasi Peterburi.

Artiklis esitatud meetodid on robustsed ega vaja palju andmeid. Juba ühe väljaande abil on võimalik teha järeldusi uudisvoogude liikumise kohta nii pikas kui ka lühikeses plaanis. Koha-kuupäeva päiste abil lähenedes säilib pressi seos kommunikatsioonivõrgustikega, millest ta sõltus (postiteenus, telegraaf vm), samuti on nii võimalik uurida info liikumist üle keelepiiride. Koha-kuupäeva päise püsivus ajas muudab võimalikuks sama meetodit rakendada sisuliselt ükskõik millisele ajalehele, mis ilmus enne 20. sajandit. Tulevikus oleks võimalik sarnasel moel analüüsida ka teisi ajalehti, et luua terviklikum pilt Euroopa kommunikatsioonivõrgustikest.

APPENDIX A: FREQUENCY OF  
LOCATIONS IN EUROPE IN PLACE-  
DATE HEADINGS, 1802–1888



**APPENDIX B: MAIL ARRIVALS IN RIGA BETWEEN 1800 AND 1890<sup>86</sup>,  
DIRECTIONS AND FREQUENCIES;  
CHANGES IN RZ FREQUENCY OF APPEARANCE**

Years	Memel/ Taurrogen (heading west)	Reval/Pernau (heading north)	Dorpat, Walk, Wolmar (heading north)	St. Petersburg (heading north-east)	Dünaburg, Moskva, Vitebsk (heading east)	Mitau (heading south, local)	Tukums (heading west, local)	Weekly publica- tion of RZ
1800–1810	2	2	-	2	2	-	-	2
1810–1820	2	2 → 0 (1816)	-	2	2	-	-	2
1820–1830	2 → 3 (1824)	-	-	2	2	7 (1828)	-	3 (1828)
1830–1840	3	-	-	2	2	7	-	3
1840–1850	3 → 5 → 7 (1843; 1849)	-	-	2 → 5 (1844)	2 → 3 → 2 (1844; 1848)	-	-	6 (1843)
1850–1860	7 → 6 (1853)	-	-	5 → 4 (1857)	2	-	-	6
1860–1870	6 → 2 (1864)	3 → 2 (1866)	5 → 7 (1869)	5 → 6 (1862)	2 → 7 (1864)	-	-	6
1870–1880	-	2	2 → 4 (1876)	-	7	7	7 (1878)	6
1880–1890	-	-	4	-	7	7	7	6

<sup>86</sup> Source: Livländischer Kalender auf das Jahr... 1800–1890.

