

When no news is bad news - Detection of negative events from news media content

Kristoffer L. Nielbo^{a,b,c}, Frida Haestrup^a, Kenneth C. Enevoldsen^a, Peter B. Vahlstrup^{a,b}, Rebekah B. Baglini^c and Andreas Roepstorff^c

^aCenter for Humanities Computing Aarhus, Jens Chr. Skous Vej 4, Building 1483, 3rd floor, DK-8000 Aarhus C, Denmark

^bDATALAB, School of Communication and Culture, Aarhus University, Helsingforsgade 14, DK-8200 Aarhus N, Denmark

^cInteracting Minds Centre, Jens Chr. Skous Vej 4, Building 1483, 3rd floor, DK-8000 Aarhus C, Denmark

Abstract

During the first wave of Covid-19 information decoupling could be observed in the flow of news media content. The corollary of the content alignment within and between news sources experienced by readers (i.e., all news transformed into Corona-news), was that the novelty of news content went down as media focused monotonically on the pandemic event. This all-important Covid-19 news theme turned out to be quite persistent as the pandemic continued, resulting in the, from a news media's perspective, paradoxical situation where the same news was repeated over and over. This information phenomenon, where novelty decreases and persistence increases, has previously been used to track change in news media, but in this study we specifically test the claim that new information decoupling behavior of media can be used to reliably detect change in news media content originating in a *negative event*, using a Bayesian approach to change point detection.

Keywords

Newspapers, Pandemic Response, Bayesian Change Detection, Information Theory

Introduction

A peculiar behavior could be observed in news media when the first wave of Covid-19 virus spread across the world. In response to this pandemic event, the ordinary rate of change in news content was disrupted because every story became associated with Covid-19. On the one hand, content novelty went down, because every story became more similar to previous stories, but on the other hand, the Covid-19 association became more prevalent, resulting in, at least initially, an increase in content persistence. A recent study [1] argues that this behavior is an example of the *news information decoupling*

✉ kln@cas.au.dk (K.L. Nielbo); frihae@cas.au.dk (F. Haestrup); kenneth.enevoldsen@cas.au.dk (K.C. Enevoldsen); imvpbv@cc.au.dk (P.B. Vahlstrup); rbkh@cc.au.dk (R.B. Baglini)

🌐 <https://knielbo.github.io/> (K.L. Nielbo)

🆔 0000-0002-5116-5070 (K.L. Nielbo); 0000-0001-7561-6297 (F. Haestrup); 0000-0001-8733-0966 (K.C. Enevoldsen); 0000-0003-4126-7740 (P.B. Vahlstrup); 0000-0002-2836-5867 (R.B. Baglini); 0000-0002-3665-1708 (A. Roepstorff)

© 2021 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

(NID) principle, according to which information dynamics of news media are (initially) decoupled by temporally extended catastrophes such that the content novelty decreases as media focus monotonically on the catastrophic event, but the resonant property of said content increases as its continued relevance propagate throughout the news information system. The authors further argued NID can be used to detect significant change in news media that originate in catastrophic events.

Previous studies have shown that variation in newspapers' word usage is sensitive to the dynamics of socio-cultural events [2, 3, 4], can detect event-driven shifts [5], and accurately can model effects of change in comprehensive collections of newspapers [6]. Furthermore, the associative structure of newspapers has been shown to accurately capture thematic development [7], and, when modelled dynamically, is indicative of the evolution of cultural values and biases [3, 8]. Adaptive fractal analysis of word frequencies over time has been used to discriminate between different classes of catastrophic events that display class-specific fractal signatures in, among other things, word usage in newspapers [9]. Several studies have shown that information theoretical construct can be used to detect fundamental conceptual differences between distinct periods [2], concurrent normative and ideological movements [10], and even, development of ideational factors (e.g., creative expression) in temporally dependent writings [11, 12, 13]. More specifically, a set of methodologically related studies have applied windowed relative entropy to thematic text representations to generate signals that capture information *novelty* as a reliable content difference from the past and *resonance* as the degree to which future information conforms to said novelty [10, 11]. Three recent studies have found that successful social media content show a strong association between novelty and resonance [14], that legacy news media under normal conditions display a remarkably similar medium to strong association between novelty and resonance across the political spectrum [1], and, finally, that variation in the novelty-resonance association can predict significant change points in historical data [15].

This study specifically tests the claim of [1] that NID-like behavior can provide input for change point detection algorithms. Specifically, we propose to test the claim that two change points are observable in news media during the first phase of Covid-19, *Lockdown* and *Opening* respectively, using a Bayesian approach to change point detection.

Results

Figure 1 displays a prototypical example of NID during the first phase of Covid-19 [1]. Although Covid-19 news items date back to December 2019, 'Wuhan', newspaper content is not impacted until the period after the first national outbreak (in this case in Denmark). 'Virus'. From the phase 1 lockdown 'Lockdown' to the opening, 'Opening', the newspaper shows a valley in novelty and, initially, a peak in resonance until both processes approximately return to normal after the opening.

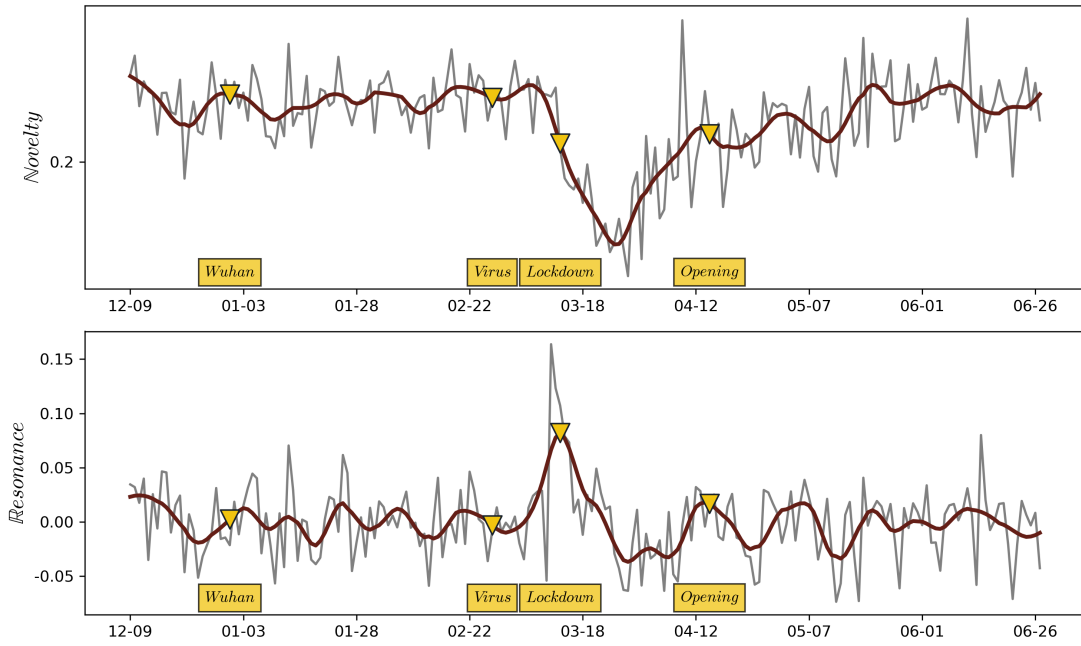


Figure 1: Novelty (upper panel) and resonance (lowerpanel) for the center-left newspaper *Politiken* before and during Covid-19 phase 1. Trend lines in the upper and middle panel are estimated using a nonlinear adaptive filter suggested in [1].

Source	Class	<i>NID</i> Start	<i>NID</i> End	<i>NID</i>
Berlingske	<i>B</i>	03.07 [03.03, 03.09]	04.28 [04.09, 05.08]	<i>True</i>
BT	<i>T</i>	04.10 [12.30, 09.01]	07.25 [04.22, 09.03]	<i>False</i>
Ekstrabladet	<i>T</i>	01.28 [01.02, 03.17]	05.08 [01.16, 07.22]	<i>False</i>
Jyllands-Posten	<i>B</i>	03.10 [03.08, 03.14]	05.25 [05.21, 06.06]	<i>True</i>
Kristligt Dagblad	<i>B</i>	03.07 [03.05, 03.12]	04.15 [04.11, 04.17]	<i>True</i>
Politiken	<i>B</i>	03.13 [03.12, 03.13]	04.08 [04.05, 04.08]	<i>True</i>

Table 1: Estimated temporal change points at 94% high density intervals for novelty. Column one contains the name of the newspaper, columns two its class (*Broadsheet* or *Tabloid*), *NID* Start and End is the beginning and end of the lockdown as represented in the newspaper, and the final column indicated if the specific source supported the *NID* principle.

To validate the observed behavior, we tested for two change points in novelty using a Bayesian model, see Appendix A for methods. The first change point, ‘*NID* Start’ should separate pre-lockdown from lockdown centered on week 11, and the second lockdown, ‘*NID* End’ from post opening (centered on week 16). Table 1 shows the estimated change points for six national newspapers, two of which are *Tabloid* newspapers (Class) and the remainder *Broadsheet*. From the model, it can be observed that all broadsheet newspapers seem to support the *NID* principle in novelty. The first change point is

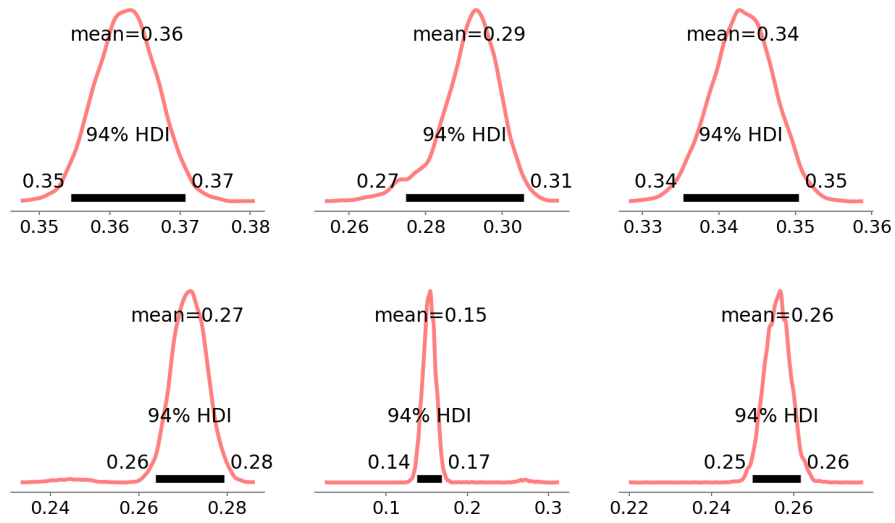


Figure 2: Posterior distributions of novelty at 94% high density intervals for newspapers Berlingske, and Politiken, see table 2.

placed in weeks 10-11, the second, however, is more a matter of contention. The opening change point lies within April and displays a month’s delayed response. Finally, it can be observed that tabloid press shows no indication of NID behavior. Table 2 and figure 2 show the posterior distributions, their means and highest density intervals, for four broadsheet and one tabloid newspaper, clearly indicating that broadsheet newspapers do conform to NID, while tabloids do not.

Source	\mathbb{N}_{pre}	\mathbb{N}_{NID}	\mathbb{N}_{post}
Berlingske	0.36 [0.35, 0.37]	0.29 [0.27, 0.31]	0.34 [0.34, 0.35]
Jyllands-Posten	0.29 [0.28, 0.30]	0.23 [0.22, 0.24]	0.27 [0.26, 0.28]
Kristligt Dagblad	0.27 [0.26, 0.28]	0.19 [0.18, 0.21]	0.26 [0.25, 0.27]
Politiken	0.27 [0.26, 0.28]	0.15 [0.14, 0.17]	0.26 [0.25, 0.26]

Table 2: Novelty values at 94% high density intervals before during and after the lockdown for the four broadsheet newspapers that supported the NID principle, see table 1.

That novelty decreases during a catastrophic event is nevertheless only half the story. For NID to be supported by the data, resonance should increase during the lockdown such that the medium to strong association between novelty and resonance is momentarily weakened. Following [1], we inspected the time-windowed linear fits of resonance on novelty, $\mathbb{N} \times \mathbb{R}$, in order to confirm this, see figure 3. All broadsheet newspapers display a slope decrease during the lockdown, thereby conforming to the NID principle 3. Tabloids on the other hand, follow an inverse pattern, such that the $\mathbb{N} \times \mathbb{R}$ slope increases during

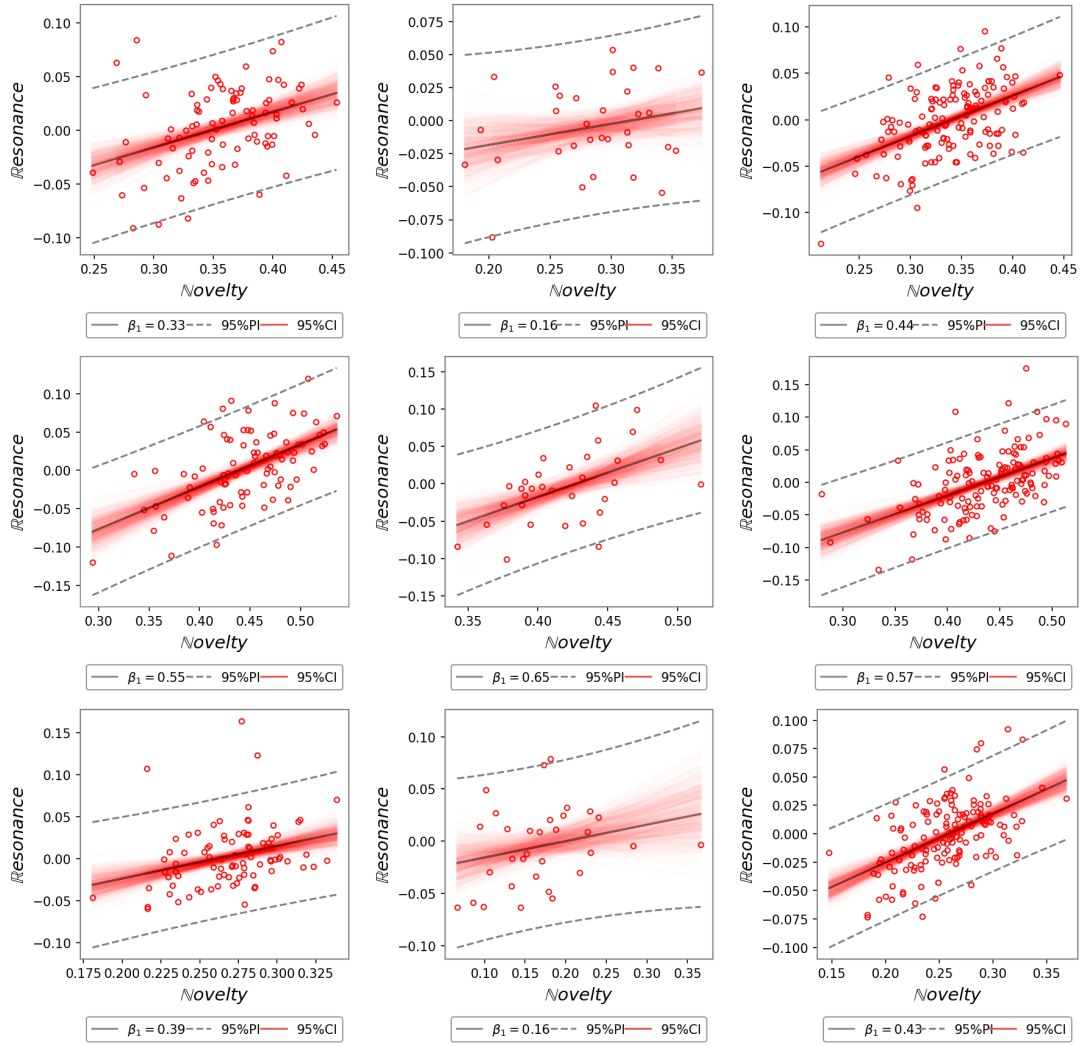


Figure 3: $\mathbb{N} \times \mathbb{R}$ slopes before during and after the lockdown for Berlingske (upper row), Ekstrabladet (middle row), and Politiken (lower row) during Covid-19 phase 1.

the lockdown period.

Source	$\mathbb{N} \times \mathbb{R}_{pre}$	$\mathbb{N} \times \mathbb{R}_{NID}$	$\mathbb{N} \times \mathbb{R}_{post}$
Berlingske	0.33 [0.17, 0.51]	0.16 [-0.07, 0.38]	0.44 [0.32, 0.58]
BT	0.49 [0.29, 0.66]	0.55 [0.28, 0.83]	0.26 [0.08, 0.43]
Ekstrabladet	0.55 [0.38, 0.72]	0.65 [0.26, 1]	0.57 [0.42, 0.71]
Jyllands-Posten	0.42 [0.24, 0.63]	0.31 [0.04, 0.56]	0.39 [0.25, 0.51]
Kristligt Dagblad	0.57 [0.34, 0.78]	0.43 [0.06, 0.78]	0.76 [0.55, 0.95]
Politiken	0.39 [0.14, 0.61]	0.16 [-0.05, 0.37]	0.43 [0.32, 0.54]

Table 3: NxR coefficients at 95% confidence intervals before during and after the lockdown for all newspapers in the sample.

Concluding Remarks

In conclusion, this study sought to validate the news information decoupling (NID) principle on a sample of six national Danish newspapers during the first phase of Covid-19. Using a Bayesian approach to change point detection, we found that content novelty in broadsheet newspapers does indeed display statistically reliable points of change during the Covid-19 lockdown. NID was further corroborated by the $\mathbb{N} \times \mathbb{R}_{pre}$ slopes that indicated a decoupling of resonance from novelty during the lockdown. Several observations can be made from the findings. First, the estimated change points for the ‘Pre-lockdown \rightarrow Lockdown’ are spread over a two week interval, which reflects that a lockdown could be reasonably predicted already from the first Covid-19 incident in Denmark. Second, in a similar vein the ‘Lockdown \rightarrow Opening’ change points are spread over an entire month from April 8 to May 8. The Danish government during the period was center-left and model’s uncertainty in determining the opening may reflect political observations [1], where center-right newspapers (e.g., Berlingske and Jyllands-Postern) were more sceptical towards the government’s implementation of an opening than the center-left (e.g., Politiken). In other words, the center-right might have been more reluctant to acknowledge the opening as a return to normal. Third, tabloid newspapers do not show any indication of a news decoupling, on the contrary, their $\mathbb{N} \times \mathbb{R}_{pre}$ slopes momentarily increases during the lockdown. This increase in slopes does, however, not provide much useful information, because, as shown by the change point detection model, the periodization is not meaningful to the two tabloid newspapers.

Validation of the NID principle is still needed for multilingual data and its value for crisis management should be further tested. For change detection, the scope of the principle needs additional testing; does NID generalize beyond a small set of negative events to, for instance, temporally extended significant events (e.g., moon landing, fall of the Berlin Wall). Finally, several comparisons already hinted that left vs. right-wing newspapers, tabloid vs. broadsheet newspapers, silly season and other seasonal effects, are interesting venues for media and journalism researchers.

Appendix

A. Methods

Data and Normalization

The data set consists of all linguistic content (title and body text) from front pages of six Danish national newspapers Berlingske, BT, Ekstrabladet, Jyllands-Posten, Kristligt Dagblad, and Politiken. The newspapers were sampled during December 1, 2019 to July 1 2020. Content not produced by the newspaper, e.g., advertisements, was excluded from the sample. In order to normalize linguistic content, numerals and highly frequent function words were removed, and the remaining data were lemmatized and casefolded. Subsequently, the data were represented as a bag-of-words (BoW) model using latent Dirichlet allocation in order to generate a dense low-rank representation of each article. Note that with a few modifications to equations (4) and (5), the approach works for any probabilistic or geometric vector-representation of documents. Novelty and resonance were estimated for in windows of one week ($w = 7$).

Novelty and Resonance

Two related information signals were extracted from the temporally sorted BoW model: *Novelty* as an article $s^{(j)}$'s reliable difference from past articles $s^{(j-1)}, s^{(j-2)}, \dots, s^{(j-w)}$ in window w :

$$\mathbb{N}_w(j) = \frac{1}{w} \sum_{d=1}^w JSD(s^{(j)} | s^{(j-d)}) \quad (1)$$

and *resonance* as the degree to which future articles $s^{(j+1)}, s^{(j+2)}, \dots, s^{(j+w)}$ conforms to article $s^{(j)}$'s novelty:

$$\mathbb{R}_w(j) = \mathbb{N}_w(j) - \mathbb{T}_w(j) \quad (2)$$

where \mathbb{T} is the *transience* of $s^{(j)}$:

$$\mathbb{T}_w(j) = \frac{1}{w} \sum_{d=1}^w JSD(s^{(j)} | s^{(j+d)}) \quad (3)$$

The novelty-resonance model was originally proposed in [10], but here we propose a symmetrized and smooth version by using the Jensen–Shannon divergence (JSD):

$$JSD(s^{(j)} | s^{(k)}) = \frac{1}{2}D(s^{(j)} | M) + \frac{1}{2}D(s^{(k)} | M) \quad (4)$$

with $M = \frac{1}{2}(s^{(j)} + s^{(k)})$ and D is the Kullback-Leibler divergence:

$$D(s^{(j)} | s^{(k)}) = \sum_{i=1}^K s_i^{(j)} \times \log_2 \frac{s_i^{(j)}}{s_i^{(k)}} \quad (5)$$

Finally, in order to describe the information states before and after an events (e.g., Lockdown, Opening), we fit resonance on novelty to estimate the $\mathbb{N} \times \mathbb{R}$ slope β_1 in the specific time windows:

$$\mathbb{R}_i = \beta_0 + \beta_1 \mathbb{N}_i + \epsilon_i, \quad i = 1, \dots, n. \quad (6)$$

Bayesian Change Point Detection

For the estimation of change points, a Bayesian approach was used. Following previous considerations, we assume that the time series contains two change points, τ_1 and τ_2 . Aside from change points, the series is assumed to be stable and follow a normal distribution with varied mean, μ_i , and singular variance, σ . This gives us the following model given the observed Novelty, \mathbb{N}_i :

$$\mathbb{N}_t = \begin{cases} \mathcal{N}(\mu_1, \sigma) & \text{for } t < \tau_1 \\ \mathcal{N}(\mu_2, \sigma) & \text{for } \tau_1 \leq t < \tau_2 \\ \mathcal{N}(\mu_3, \sigma) & \text{for } t \geq \tau_2 \end{cases} \quad (7)$$

for which we wish to estimate the location of the change points τ_i , means μ_i and variance σ , i.e. the following posterior:

$$P(\mu_i, \sigma, \tau_i | \mathbb{N}_t) = P(\mu_1, \mu_2, \mu_3, \sigma, \tau_1, \tau_2 | \mathbb{N}_t)$$

For estimation of the posterior, we have used NUTS sampling as implemented in pyMC3 [16] using 4000 samples. The estimation was done using using naive to slightly conservative priors assuming that the change points, τ_i , can be anywhere in the sequence (with $\tau_2 > \tau_1$) and that the variance, σ , is stable across change points. Note that the half Cauchy prior distribution has series of beneficial properties [17, 18] including its fat tail which allows for extreme values. These assumptions were modelled using the following priors:

$$\begin{aligned} \mu_i &\sim \mathcal{N}(0, 0.5) \\ \sigma &\sim \text{Half Cauchy}(0.5) \\ \tau_1 &\sim \text{Uniform}(0, \max(\mathbb{N}_t)) \\ \tau_2 &\sim \text{Uniform}(\tau_1, \max(\mathbb{N}_t)) \end{aligned}$$

B. Online Resources

All data are proprietary and have been collected through Infomedia’s API: <https://infomedia.dk/>. For inquiries regarding models and derived data, please contact kln@cas.au.dk. The source code for methods is available on Github: <https://bit.ly/3beahFd>. More details on NID detection can be found at NeiC’s NDHL website: <https://bit.ly/3bfeW9C>.

Acknowledgments

This research was supported the "HOPE - How Democracies Cope with COVID-19"-project funded by The Carlsberg Foundation with grant CF20-0044, NeiC's Nordic Digital Humanities Laboratory project, and DeiC Type-1 HPC with project DeiC-AU1-L-000001. The authors would like to thank Berlingske Media, JP/Politikens Hus, and Kristeligt Dagblad for providing access to proprietary data.

References

- [1] K. L. Nielbo, R. B. Baglini, P. B. Vahlstrup, K. C. Enevoldsen, A. Bechmann, A. Roepstorff, News Information Decoupling: An Information Signature of Catastrophes in Legacy News Media, arXiv:2101.02956 [cs] (2021). URL: <http://arxiv.org/abs/2101.02956>, arXiv: 2101.02956.
- [2] J. Guldi, The Measures of Modernity: The New Quantitative Metrics of Historical Change Over Time and Their Critical Interpretation, *International Journal for History, Culture and Modernity* 7 (2019) 899–939. URL: https://brill.com/view/journals/hcm/7/1/article-p899_42.xml. doi:10.18352/hcm.589.
- [3] J. van Eijnatten, R. Ros, The Eurocentric Fallacy. A Digital-Historical Approach to the Concepts of 'Modernity', 'Civilization' and 'Europe' (1840–1990), *International Journal for History, Culture and Modernity* 7 (2019) 686–736. URL: https://brill.com/view/journals/hcm/7/1/article-p686_33.xml. doi:10.18352/hcm.580.
- [4] J. Daems, T. D'haeninck, S. Hengchen, T. Zere, C. Verbruggen, 'Workers of the World'? A Digital Approach to Classify the International Scope of Belgian Socialist Newspapers, 1885–1940, *Journal of European Periodical Studies* 4 (2019) 99–114. URL: <https://ojs.ugent.be/jeps/article/view/10187>. doi:10.21825/jeps.v4i1.10187.
- [5] M. Kestemont, F. Karsdorp, M. Düring, Mining the Twentieth Century's History from the Time Magazine Corpus, in: *Proceedings of the 8th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH)*, Association for Computational Linguistics, Gothenburg, Sweden, 2014, pp. 62–70. URL: <http://aclweb.org/anthology/W14-0609>. doi:10.3115/v1/W14-0609.
- [6] P. Bos, H. Wijfjes, M. Piscaer, G. Voerman, Quantifying "Pillarization": Extracting Political History from Large Databases of Digitized Media Collections, *Proceedings of the 3rd HistoInformatics Workshop* (2016) 10.
- [7] D. J. Newman, S. Block, Probabilistic topic decomposition of an eighteenth-century American newspaper, *Journal of the American Society for Information Science and Technology* 57 (2006) 753–767. URL: <http://doi.wiley.com/10.1002/asi.20342>. doi:10.1002/asi.20342.
- [8] M. Wevers, Using Word Embeddings to Examine Gender Bias in Dutch Newspapers, 1950-1990, in: *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*, Association for Computational Linguis-

- tics, Florence, Italy, 2019, pp. 92–97. URL: <https://www.aclweb.org/anthology/W19-4712>. doi:10.18653/v1/W19-4712.
- [9] J. Gao, J. Hu, X. Mao, M. Perc, Culturomics meets random fractal theory: insights into long-range correlations of social and natural phenomena over the past two centuries, *Journal of The Royal Society Interface* 9 (2012) 1956–1964.
 - [10] A. T. J. Barron, J. Huang, R. L. Spang, S. DeDeo, Individuals, institutions, and innovation in the debates of the French Revolution, *Proceedings of the National Academy of Sciences* 115 (2018) 4607–4612. URL: <http://www.pnas.org/lookup/doi/10.1073/pnas.1717729115>. doi:10.1073/pnas.1717729115.
 - [11] J. Murdock, C. Allen, S. DeDeo, Exploration and Exploitation of Victorian Science in Darwin’s Reading Notebooks, arXiv preprint arXiv:1509.07175 (2015). URL: <http://arxiv.org/abs/1509.07175>.
 - [12] K. L. Nielbo, M. L. Perner, C. P. Larsen, J. Nielsen, D. Laursen, Automated Compositional Change Detection in Saxo Grammaticus’ *Gesta Danorum*, in: DHN, 2019, pp. 320–332.
 - [13] K. L. Nielbo, K. F. Baunvig, B. Liu, J. Gao, A curious case of entropic decay: Persistent complexity in textual cultural heritage, *Digital Scholarship in the Humanities* 34 (2019). URL: <https://doi.org/10.1093/llc/fqy054>. doi:10.1093/llc/fqy054.
 - [14] K. Nielbo, P. Vahlstrup, A. Bechmann, Trend Reservoir Detection: Minimal Persistence and Resonant Behavior of Trends in Social Media, *Proceedings of Computational Humanities Research* 1 (2021).
 - [15] E. Vrangbæk, K. Nielbo, Composition and Change in *De Civitate Dei*: A Case Study of Computationally Assisted Methods, *Studia Patristica* (2021).
 - [16] J. Salvatier, T. V. Wiecki, C. Fonnesbeck, Probabilistic programming in python using pymc3, *PeerJ Computer Science* 2 (2016) e55.
 - [17] A. Gelman, Prior distributions for variance parameters in hierarchical models, *Bayesian Analysis* 1 (2006) 515–533. doi:10.1214/06-BA117A.
 - [18] N. G. Polson, J. G. Scott, On the Half-Cauchy Prior for a Global Scale Parameter, *Bayesian Analysis* 7 (2012) 887–902. URL: <https://projecteuclid.org/euclid.ba/1354024466>. doi:10.1214/12-BA730, publisher: International Society for Bayesian Analysis.