Information Decoupling

– A Pandemic Signature in News Media

Kristoffer L. Nielbo, Rebekah B. Baglini & Andreas Roepstorff

Center for Humanities Computing AArhus|chcaa.au.dk
aarhus university, denmark

CENTER FOR HUMANITIES
COMPUTING AARHUS

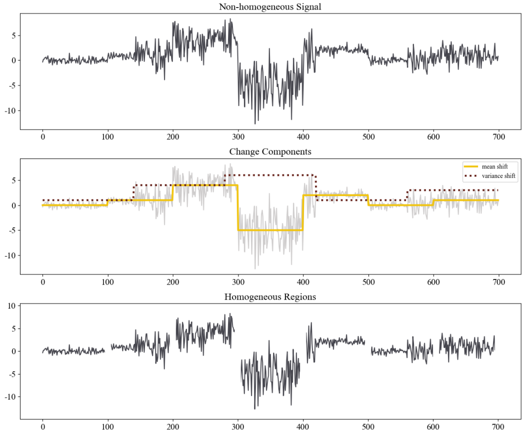
BACKGROUND

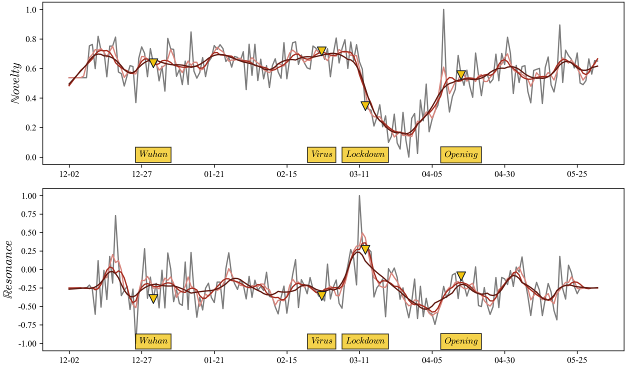
HOPE

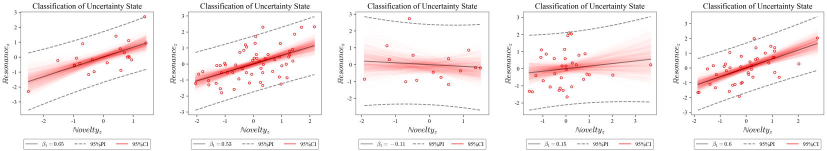
– **how democracies cope with COVID-19 a data-driven approach** is an national research project that is part of the (DK) national pandemic monitoring program.

– research team interested in cultural dynamics, in particular **how events impact cultural information systems**

– use news media coverage of COVID-19 as a proxy for how cultural information systems respond to **unexpected and dangerous temporally extended events**







Change detection on novelty, $\mathcal{N}$, and chance description on the resonance on novelty slope, $\mathcal{N} \cdot \mathcal{R}$-slope.

Information decoupling in news media

in response to unexpected and dangerous temporally extended events, the ordinary information dynamics of news media are (initially) decoupled 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**

K. L. Nielbo, R. B. Baglini, P. B. Vahlstrup, K. C. Enevoldsen, A. Bechmann, and A. Roepstorff (2021) "News Information Decoupling: An Information Signature of Catastrophes in Legacy News Media," arXiv:2101.02956 [cs]

K. L. Nielbo, F. Haestrup, K. C. Enevoldsen, P. B. Vahlstrup, R. B. Baglini, and A. Roepstorff, "When no news is bad news - Detection of negative events from news media content," arXiv:2102.06505 [cs]

K.L. Nielbo, F. Hæstrup, J. Kostkan, K.C. Enevoldsen, P.B. Vahlstrup, E. Fano, and R. B. Baglini, "Pandemic news information uncertainty – News dynamics mirror differential response strategies to COVID-19", *in review*

DATA, NORMALIZATION, AND REPRESENTATION

DATA

linguistic content (title and body text) from **front pages of six DK and four SE* national newspapers** ($4 \times$ *tabloid*, $6 \times$ *broadsheet*).

sampled during COVID-19 phase 1 (december 1, 2019 to july 1 2020)

NORMALIZATION

advertisements and metadata removed

lemmatization, tf-idf weighting, casefolding

REPRESENTATION

bag-of-words model* to generate low-rank representations of front pages

variables were estimated for **windows of one week** ($w = 7$).

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

$$\mathcal{N}_w(j) = \frac{1}{w} \sum_{d=1}^{w} JSD(s^{(j)} \mid s^{(j-d)}) \tag{1}$$

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

$$\mathcal{R}_w(j) = \mathcal{N}_w(j) - \mathcal{T}_w(j) \tag{2}$$

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

$$\mathcal{T}_w(j) = \frac{1}{w} \sum_{d=1}^{w} JSD(s^{(j)} \mid s^{(j+d)}) \tag{3}$$

The novelty-resonance model was originally proposed in (Barron et al 2018), but here we propose a symmetrized and smooth version by using the Jensen–Shannon divergence (JSD):

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

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

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

Nonlinear Adaptive Filtering

To model global trends in the novelty signal, we apply a nonlinear adaptive multi-scale decomposition algorithm. First, the signal is partitioned into overlapping segments of length $w = 2n + 1$, where neighboring segments overlap by $n + 1$ points. In each segment, the signal is fitted with the best polynomial of order M, obtained by using the standard least-squares regression; the fitted polynomials in overlapped regions are then combined to yield a single global smooth trend. Denoting the fitted polynomials for the $i - th$ and $(i + 1) - th$ segments by $y^i(l_1)$ and $y^{(i+1)}(l_2)$, respectively, where $l_1, l_2 = 1, \cdots, 2n + 1$, we define the fitting for the overlapped region as

$$y^{(c)}(l) = w_1 y^{(i)}(l + n) + w_2 y^{(i+1)}(l), \quad l = 1, 2, \cdots, n + 1 \tag{6}$$

where $w_1 = \left(1 - \frac{l-1}{n}\right)$ and $w_2 = \frac{l-1}{n}$ can be written as $(1 - d_j/n)$ for $j = 1, 2$, and where d_j denotes the distances between the point and the centers of $y^{(i)}$ and $y^{(i+1)}$, respectively.

Change Detection

Assume two change points, τ_1 and τ_2 and an otherwise stable series that follow a normal distribution with varied mean, μ_i, and singular variance, σ. This gives us the following model given the observed Novelty, $\mathbb{N}_i$:

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

Estimate the location of τ_i, means μ_i and variance σ, i.e. the following posterior:

$$P(\mu_i, \sigma, \tau_i | \mathcal{N}_t) = P(\mu_1, \mu_2, \mu_3, \sigma, \tau_1, \tau_2 | \mathcal{N}_t) \tag{8}$$

Estimation was carried out with NUTS and the assumptions were modelled using the following priors:

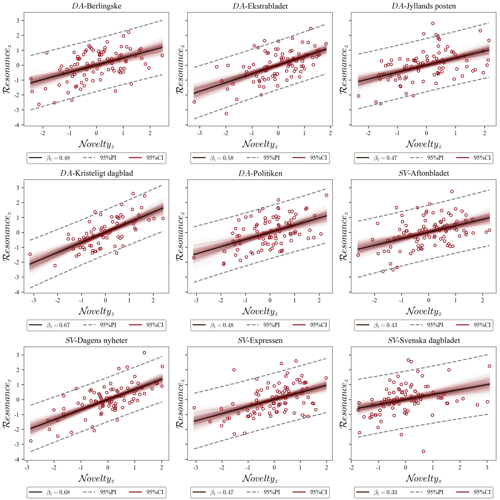
$$\begin{aligned} \mu_i &\sim \text{Normal}(0, 0.5) \\ \sigma &\sim \text{Half Cauchy}(0.5) \\ \tau_1 &\sim \text{Uniform}(0, \max(\mathcal{N}_t)) \\ \tau_2 &\sim \text{Uniform}(\tau_1, \max(\mathcal{N}_t)) \end{aligned} \tag{9}$$

CHANGE DESCRIPTION

In order to describe the information states before and after an events (e.g., Lockdown, Opening), we fit resonance on novelty to estimate the $\mathcal{NR}$ slope β in the specific time windows:

$$\mathcal{R}_i = \beta_0 + \beta_1 \mathcal{N}_i + \epsilon_i, \quad i = 1, \ldots, n \qquad (10)$$

Information Decoupling: A reliable decrease in the $\mathcal{NR}$ slope that coincides with or follows an event



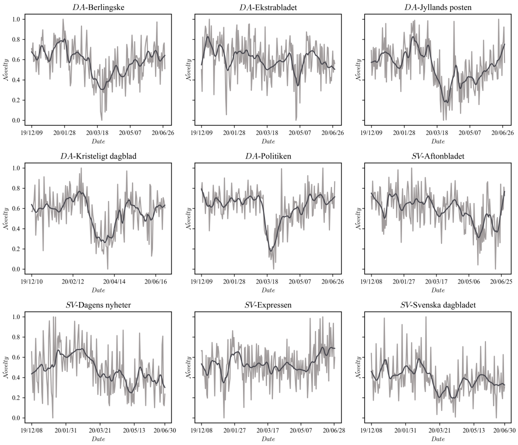
TYPE-DEPENDENT SUPPORT

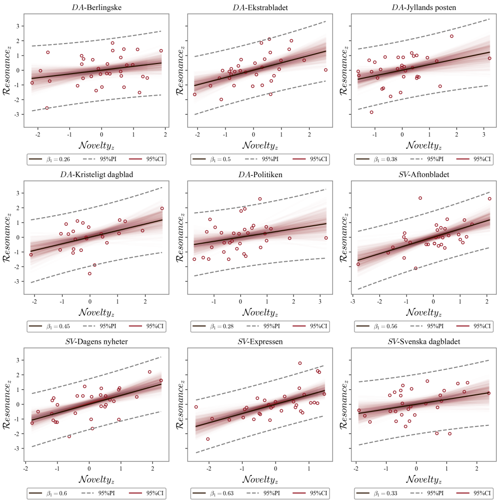
Source	Class	*ID* Start	*ID* End	*ID*
Berlingske	*B*	03.07 [03.03, 03.09]	04.28 [04.09, 05.08]	*True*
BT	*T*	04.10 [12.30, 09.01]	07.25 [04.22, 09.03]	*False*
Ekstrabladet	*T*	01.28 [01.02, 03.17]	05.08 [01.16, 07.22]	*False*
Jyllands-Posten	*B*	03.10 [03.08, 03.14]	05.25 [05.21, 06.06]	*True*
Kristligt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

Estimated temporal change points at 94% HDIs for novelty. Column one contains the name of the newspaper, columns two its class (*B*roadsheet or *T*abloid).

Source	$\mathcal{N}_{pre}$	$\mathcal{N}_{ID}$	$\mathcal{N}_{post}$
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]

Novelty values at 94% HDIs before during and after the lockdown for the four broadsheet newspapers.





Resonance on novelty during the first phase of Covid-19. Danish newspapers that show valid change points (Berlingske, Jyllands Posten, Kristeligt Dagblad and Politiken) also display a drop in the $\mathcal{N} \times \mathcal{R}$ slope.

IN CONCLUSION...

"Nothing travels faster than the speed of light with the possible exception of bad news, which obeys its own special laws." (D. Adams – Hitchhiker's Guide)

in the case of pandemic information dynamics,

- variation in $\mathcal{N}$ *reliably* detected *lockdown* and *opening*
- $\mathcal{N}\mathcal{R}$ slopes indicated a decoupling of resonance from novelty during the lockdown
- lockdown interval indicated that lockdown can be be predicted from the first incident
- opening interval may reveal political observation
- tabloids follow different dynamics

```python
if questions:
    try:
        answer()
    except RunTimeError:
        pass
else:
    print('THANKS')
```

THANKS

kln@cas.au.dk

knielbo.github.io

chcaa.io

SLIDES

knielbo.github.io/files/kln_dhnb22.pdf

ACKNOWLEDGEMENTS

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.