Dynamic change detection in topics based on rolling LDAs

Jonas **Rieger***[1]*, Kai-Robin **Lange***[1]*, Jonathan **Flossdorf***[1]* and Carsten **Jentsch***[1]*

[1]Department of Statistics, TU Dortmund University, 44221 Dortmund, Germany

Abstract

Topic modeling methods such as e.g. Latent Dirichlet Allocation (LDA) are popular techniques to analyze large text corpora. With huge amounts of textual data that are collected over time in various fields of applied research, it becomes also relevant to be able to automatically monitor the evolution of topics identified from some sort of dynamic topic modeling approach. For this purpose, we propose a dynamic change detection method that relies on a rolling version of the classical LDA that allows for coherently modeled topics over time that are able to adapt to changing vocabulary. The changes are detected by assessing the intensity of word change in the LDA's topics over time in comparison to the expected intensity of word change under stable conditions using resampling techniques. We apply our method to topics obtained by applying the RollingLDA to Covid-19 related news data from CNN and illustrate that the detected changes in these topics are well interpretable.

Keywords

change point, event, shift, narrative, story, evolution, monitoring, Latent Dirichlet Allocation

1. Introduction

While change detection is an active field in modern research, the application for text data poses even further obstacles due to its unstructured nature. And yet, an effective method for change detection would have many use cases. Particularly, when dealing with large text corpora collected over time, an online detection approach will be useful to analyze the evolution of narratives or to spot a shift in a discourse about certain topics. For this purpose, we propose an online change detection method for text data by analyzing the change of word distributions within topics of Latent Dirichlet Allocation (LDA) models. As we are dealing with time series of textual data, we make use of a rolling version of the classical LDA, called RollingLDA [1]. The method is designed to construct coherently interpretable topics modeled over time that are allowed to adapt to a changing vocabulary. The changes are detected by dynamically assessing the change intensity in word usage in the LDA's topics over time in comparison to the change intensity expected in stable periods using resampling techniques.

The main goal of change detection is to identify possible anomalies in a process. Typically,

there are the two perspectives towards this issue: offline and online applications. Our approach is applicable for both tasks, but, for each time point, it relies exclusively on the text data that has already been observed. Hence, we focus on the usually more relevant task of online monitoring. In traditional schemes for change detection [2, 3], control charts are applied to visualize the monitoring procedure using a control statistic which is successively calculated for each time point. An alarm is triggered whenever the statistic lies outside of some control limits. In practice, there are a variety of different control charts including memory-free setups (e.g. Shewhart charts) and memory-based charts (e.g. EWMA, CUSUM), However, these traditional procedures can not be applied to textual data off the shelf because of the high dimensionality of large text corpora. In addition, an in-control state to reliably calculate the control limits is frequently not available due to the strong dynamics in text data, e.g. newspaper articles. To overcome these issues, we propose to use a control statistic based on a similarity metric that represents the resemblance of topic's word distributions over consecutive time points. Control limits are derived by a resampling procedure using word count vectors based on time-variant topics modeled by RollingLDA [1].

In a similar context, the usage of LDA was proposed for change point detection for topic distributions in texts [4], which is based on a modified version of the wild binary segmentation algorithm [5] designed for offline detection setups. There is also work considering Bayesian online monitoring [6] for textual data using a document-based model [7] and an approach based on similarity metrics, which aims to detect global events in topics in offline settings [8]. There is also work which analyzes the transitions of narratives between topics [9]. In contrast, the rolling window approach of RollingLDA constructs coherently interpretable topics modeled over time and allows the resulting dynamic change detection method to become applicable in online settings. Compared to the mentioned related methods, our method is designed to detect changes in word distributions of topics over time rather than global changes in topic distributions (of sets) of documents [e.g. 4, 7, 8] or sentiments in topics [e.g. 10] or (in contrast) changes in topic distributions of words [e.g. 11]. This results in a more refined monitoring procedure that allows for the detection of narrative shifts that are changing the word usage within a certain topic instead of measuring the frequency of a topic over time within the whole corpus. Building on this, we aim that our proposed method can provide groundwork for the extraction and temporal localization of narratives in texts.

2. Methodological framework

For the proposed change detection algorithm, we make use of the existing method of a rolling version of the classical LDA (RollingLDA) to construct coherent topics over time and measure similarities of topics for consecutive time points using the well-established cosine similarity.

2.1. Latent Dirichlet Allocation

The classical LDA [12] models distributions of K latent topics for each text. Let $W_n^{(m)}$ be a single word token at position $n = 1, \dots N^{(m)}$ in text $m = 1, \dots, M$ of a corpus of M texts. Then, a single

text is given by

$$D^{(m)} = \left(W_1^{(m)}, \dots, W_{N^{(m)}}^{(m)} \right), \quad W_n^{(m)} \in \boldsymbol{W} = \{W_1, \dots, W_V\}, V = |\boldsymbol{W}|$$

and the corresponding topic assignments for each text are given by

$$\boldsymbol{T}^{(m)} = \left(T_1^{(m)}, \dots, T_{N^{(m)}}^{(m)} \right), \quad T_n^{(m)} \in \boldsymbol{T} = \{T_1, \dots, T_K\}.$$

From this, let $n_k^{(mv)}, k = 1, \dots, K, v = 1, \dots, V$ denote the number of assignments of word v in text m to topic k. Then, we define the cumulative count of word v in topic k over all texts by $n_k^{(\cdot v)}$ and denote the total count of assignments to topic k by $n_k^{(\cdot\cdot)}$. Using these definitions, the underlying probability model [13] can be written as

$$W_n^{(m)} \mid T_n^{(m)}, \phi_k \sim \text{Discr}(\phi_k), \quad \phi_k \sim \text{Dir}(\eta), \quad T_n^{(m)} \mid \theta_m \sim \text{Discr}(\theta_m), \quad \theta_m \sim \text{Dir}(\alpha).$$

For a given parameter set $\{K, \alpha, \eta\}$, with the Dirichlet priors α and η defining the type of mixture of topics in every text and the type of mixture of words in every topic, LDA assigns one of the K topics to each token. A word distribution estimator per topic for $\phi_k = (\phi_{k,1}, \dots, \phi_{k,V})^T \in (0, 1)^V$ can be derived through the collapsed Gibbs sampler procedure [13] by

$$\hat{\phi}_{k,v} = \frac{n_k^{(\cdot v)} + \eta}{n_k^{(\cdot\cdot)} + V\eta}. \tag{1}$$

2.2. RollingLDA

RollingLDA [1] is a rolling version of classical LDA. New texts are modeled based on existing topics of the previous model. Thereby, not the whole knowledge of the entire past of the model is used, but only the information of the topics from more recent texts based on a user-chosen memory parameter. For each time point, based on the topic assignments within this memory period, the topics are initialized and modeled forward. This form of modeling preserves the topic structure of the model so that topics remain coherently interpretable over time. At the same time, constraining the knowledge of the model to the user-chosen memory period allows for changes in topics based on new vocabulary or word choices. There are other dynamic variants of the LDA approach [14, 15, 16, 17, 18] deliberately designed to model gradual changes, and therefore not as well suited to detect abrupt changes. We use the update algorithm RollingLDA to make our proposed change detection method applicable in an online manner. Thereby, a text is assigned to a time point on the basis of its publication date. The step size of the and is chosen on a weekly basis in the present case as this seems natural for journalistic texts.

2.3. Similarity

Our change detection algorithm builds on a similarity measure for word count vectors. Following up on the notation from Section 2.1 the word count vector for topic $k \in \{1, \dots, K\}$ at one time point $t \in \{0, \dots, T\}$ is given by

$$\boldsymbol{n}_{k|t} = \left(n_{k|t}^{(\cdot 1)}, \dots, n_{k|t}^{(\cdot V)} \right)^T \in \mathbb{N}_0^V = \{0, 1, 2, \dots\}^V.$$

Then, monitoring the similarity of topics over time for (consecutive) time points t_1 and t_2 is done using the cosine similarity

$$\cos\left(\boldsymbol{n}_{k|t_1}, \boldsymbol{n}_{k|t_2}\right) = \frac{\sum_v n_{k|t_1}^{(\cdot v)} n_{k|t_2}^{(\cdot v)}}{\sqrt{\sum_v \left(n_{k|t_1}^{(\cdot v)}\right)^2} \sqrt{\sum_v \left(n_{k|t_2}^{(\cdot v)}\right)^2}}. \tag{2}$$

The choice of cosine similarity is common in the context of change point detection for text data [e.g. 8, 19]. Compared to other similarity measures such as the Jaccard coefficient, Jensen-Shannon Divergence, χ^2-, Hellinger and Manhattan Distance, the cosine similarity fulfills some typical conditions required for monitoring a similarity measure [1].

3. Change detection

In combination with the existing method RollingLDA and cosine similarity, our contributed method for change detection relies on classical resampling approaches to identify changes within topics. We estimate the realized change in a topic based on the similarity between the current and previous count vectors of word assignments and compare the resulting similarity score to resampling-based similarity scores which are generated under stable conditions, such that no extraordinary changes occurred in the topic.

3.1. Set of changes

Suppose we consider K topics over T time points to be monitored. If the actual observed similarity of the word vector of some topic $k \in \{1, \dots, K\}$ at some time $t \in \{0, 1, \dots, T\}$ given by $\boldsymbol{n}_{k|t}$, compared to the frequency vector of the topic over a predefined reference time period $t - z_k^t, \dots, t - 1$, given by

$$\boldsymbol{n}_{k|(t-z_k^t):(t-1)} = \sum_{z=1}^{z_k^t} \boldsymbol{n}_{k|t-z}, \tag{3}$$

is smaller than a threshold q_k^t which is calibrated based on similarities under stable conditions (see Section 3.2), then we identify a change within topic k at time t. The set of identified changes in topic k up to time point t can then be defined as

$$C_k^t = \left\{u \mid 0 < u \leq t \leq T \ : \ \cos\left(\boldsymbol{n}_{k|u}, \boldsymbol{n}_{k|(u-z_k^u):(u-1)}\right) < q_k^t\right\} \cup 0, \tag{4}$$

where the time point $t = 0$ is always included for technical reasons, to compute the current run length without a change $z_k^t = \min\{z_{\max}, t - \max C_k^{t-1}\}$. Thus, the reference period spans the last $z_{\max}$ time points if no change was detected during that time, and spans the time that has passed since the last change, otherwise. The parameter $z_{\max}$ is to be chosen by the user and is intended to smooth the similarities to prevent from detecting false positives.

3.2. Dynamic thresholds

For the calculation of the threshold q_k^t, the estimated word distribution of a topic k at some time point t, as well as over the corresponding reference period $t - z_k^t, \dots, t - 1$ are needed. For this, let $\hat{\phi}_k^t$ and $\hat{\phi}_k^{(t-z_k^t):(t-1)}$ be defined by

$$\hat{\phi}_{k,v}^t = \frac{n_{k|t}^{(\bullet v)} + \eta}{n_{k|t}^{(\bullet \bullet)} + V\eta} \quad \text{and} \quad \hat{\phi}_{k,v}^{(t-z_k^t):(t-1)} = \frac{n_{k|(t-z_k^t):(t-1)}^{(\bullet v)} + \eta}{n_{k|(t-z_k^t):(t-1)}^{(\bullet \bullet)} + V\eta} \tag{5}$$

analogously to Equation (1).

The application of the change point detection algorithm is designed for text data, more precisely for empirical word distributions of K topics modeled by LDA in a given text corpus. Since word choice - especially in journalistic texts - varies considerably over time, a situation in which there is no change in the word distribution within topics across consecutive time points does not reflect the expected situation. Rather, it is to be expected that topics change gradually on an ongoing basis. Accordingly, our method aims to identify not the numerous customary changes in the topics, but the unexpectedly large ones. To do so, we define an expected word distribution $\tilde{\phi}_k^{(t)}$ for time point t under stable conditions that include the customary changes as a convex combination of the two estimators of the word distribution of topic k, one for the reference time period $t - z_k^t, \dots, t - 1$ and one for the current time point t. Using the mixture parameter $p \in [0, 1]$, which can be tuned based on how substantial the detected changes should be, the intensity of the expected change is considered in the determination of this estimator by

$$\tilde{\phi}_k^{(t)} = (1 - p)\,\hat{\phi}_{k,v}^{(t-z_k^t):(t-1)} + p\,\hat{\phi}_{k,v}^{(t)}. \tag{6}$$

Our method uses the estimator $\tilde{\phi}_k^{(t)}$ to simulate R expected word count vectors $\tilde{\boldsymbol{n}}_{k|t}^r, r = 1, \dots, R$ based on a parametric bootstrap approach. In this process, each word is drawn according to its estimated probability of occurrence regarding $\tilde{\phi}_k^{(t)}$ and each sample r consists of $n_{k|t}^{(\bullet \bullet)}$ draws, the number of words assigned to topic k at time point t. Then, we calculate the cosine similarity

$$\cos\left(\tilde{\boldsymbol{n}}_{k|t}^r, \boldsymbol{n}_{k|(t-z_k^t):(t-1)}\right) \tag{7}$$

for each of the $r = 1, \dots, R$ bootstrap samples and set the threshold q_k^t equal to the 0.01 quantile of these simulated similarity values generated under stable conditions. Combinations of topics and time points for which the observed similarity is smaller than the corresponding quantile are classified as change points according to Equation (4).

4. Analysis

For conducting the real data analysis, the data set under study was created with Python, whereas the preprocessing, the modeling, all postprocessing steps and analyses are performed using R. The scripts for all analysis steps can be found in the associated GitHub repository github.com/JonasRieger/topicalchanges.

4.1. Data and study design

To assess the quality of our change point algorithm, we use the TLS-Covid19 data set [20]. It is generated using Covid-19 related liveblog articles of CNN, collected from January 22nd 2020 up until December 12th 2021. Each liveblog is interpreted to belong to a topic and comprises texts and key moments. The texts form a time line containing events, which are summarized by its key moments. The resulting corpus consists of 27,432 texts and 1,462 key moments. Although the data set contains multiple key moments per day on average, we do not consider all them a change point as our aim is to detect larger changes based on aggregated weekly texts. However, these key moments serve well as indicators, which enable us to check whether the detected changes are actually related to real events or if they are false positives.

We use common NLP preprocessing steps for the texts, i.e. characters are formatted to lowercase, numbers and punctuation are removed. Moreover, a trusted stopword list is applied to remove words that do not help in classifying texts in topics, we use a lemmatization dictionary (github.com/michmech/lemmatization-lists) and neglect words with less than two characters.

We model the CNN data set using RollingLDA on a weekly basis, starting on Saturday of each week, and we consider the previous week as initialization for the model's topics. The first 10 days of modeling, Wednesday, January 22nd 2020 until Friday, January 31st 2020, serve as the initial chunk corresponding to $t = 0$. During this period, 605 texts were published. In the data set, there are weeks that do not contain any texts. In this case, the corresponding time point is omitted. Then, to model the texts of the following chunk, at least the last 10 texts are used, as well as all other texts published on the same date as the oldest of these 10 texts. As parameters, we assume $K = 12$ topics, define the reference period of the topics to the last $z_{max} = 4$ weeks, and choose $p = 0.85$, since these values are accountable by plausibility and seem to yield reasonable results. For other parameter choices, i.e. $K = 8, \dots, 20$, $z_{max} = 1, \dots, 20$, $p = 0.5, \dots 0.8, 0.81, \dots, 0.90$, results can be found in our associated repository.

4.2. Findings

The results of our chosen model are displayed in Figure 1. Fig. 1a shows the detected changes by vertical gray lines, which are the weeks in which the observed similarity (blue curve) is lower than the expected one (red curve). Furthermore, for two changes we show which words are mainly causing the detection of the change. The score of a word in a topic at a given time point is calculated by the topic's similarity without considering this word and subtracting it from the actual realized similarity. These leave-one-out cosine impact scores for the words with the five most negative scores are shown in Fig. 1b and 1c. In general, most of the changes we detect occur within the first four months of 2020. This is because the wording was constantly changing, as the Covid-19 epidemic turned into a pandemic over the course of these months. New people and organizations were associated with Covid-19, which is why we detect a bunch of consecutive changes in every topic. As the pandemic reached out into further countries, the detected changes became less frequent for most topics. In the following we share our interpretation of some exemplary detected changes.

The third topic, containing information about vaccination and testing procedures, shows a change in the week starting on the 13th of March 2021. In this week, the AstraZeneca
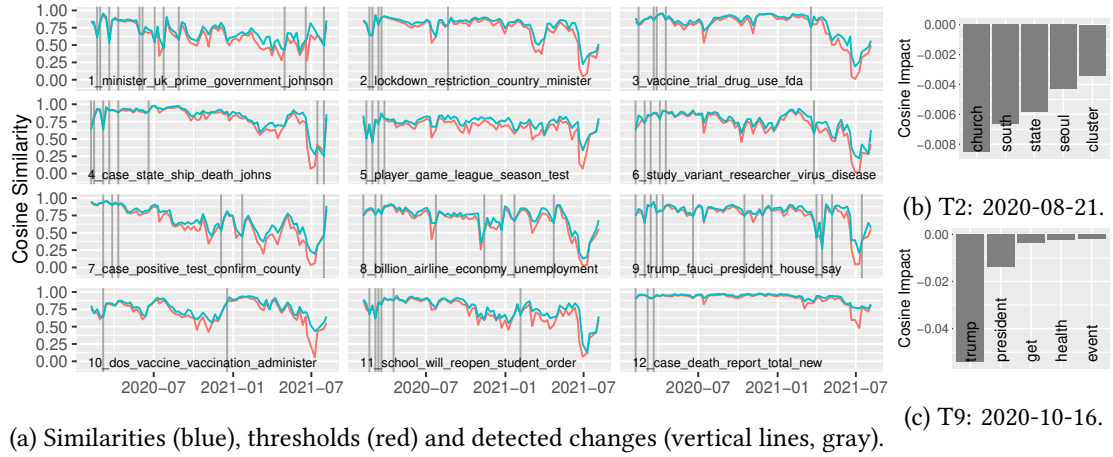
(a) Similarities (blue), thresholds (red) and detected changes (vertical lines, gray).

(b) T2: 2020-08-21.

(c) T9: 2020-10-16.

Figure 1: Similarity values, thresholds, and detected changes over the observation period for all $K = 12$ topics, as well as the five most influential words for two selected change points in topics 2 and 9.

vaccination process in several EU-states was stopped due the risk of causing blood clots.[1] The sixth topic, a topic about medical studies and research, shows a change in the following week, in which AstraZeneca presented a study about the effectiveness of its vaccine.[2] Another interesting detection is the change in the vaccination-related topic 10 in December 2020, just as the vaccination process started in the US.[3]

Political changes are also detected in several topics, such as the start of Joe Biden's presidential era in late January 2021 in topic 11, the return of Donald Trump to office after his Covid-19 infection in October 2020[4] in topic 9 (cf. Fig. 1c) or the discussion about the origin of the virus after a WHO report in late March 2021 in topic 9.[5] A Covid-19 outbreak in the South Korean Sarang-jeil church in August 2020[6] is detected in topic 2 (cf. Fig. 1b).

While these topics detect changes across the entire time span, the twelfth topic, representing the report of the current number of Covid cases, does not detect a single change after March

---

[1] CNN online, 2021-03-15 3:03 p.m. ET, "Spain joins Germany, France and Italy in halting AstraZeneca Covid-19 vaccinations", https://edition.cnn.com/world/live-news/coronavirus-pandemic-vaccine-updates-03-15-21/h_d938057f2ef588f74565bdbb01f12387, visited on 2022-01-20.

[2] CNN online, 2021-03-25 2:48 a.m. ET, "New AstraZeneca report says vaccine was 76% effective in preventing Covid-19 symptoms", https://edition.cnn.com/world/live-news/coronavirus-pandemic-vaccine-updates-03-25-21/h_9f01e2e53b62873f1c742254d27fbf5f, visited on 2022-01-20.

[3] CNN online, 2020-12-14 10:08 p.m. ET, "The first doses of FDA-authorized Covid-19 vaccine were administered in the US. Here's what we know", https://edition.cnn.com/world/live-news/coronavirus-pandemic-vaccine-updates-12-15-20/h_32be1a72dc05f874eda167c95c8f1bba, visited on 2022-01-20.

[4] CNN online, 2020-10-12 12:01 a.m. ET, "Trump says he tested 'totally negative' for Covid-19", https://edition.cnn.com/world/live-news/coronavirus-pandemic-10-12-20-intl/h_7570d53b184a5b1d6ec97ce67330e4c9, visited on 2022-01-20.

[5] CNN online, 2021-03-29 11:22 a.m. ET, "Upcoming WHO report will deem Covid-19 lab leak extremely unlikely, source says", https://www.cnn.com/world/live-news/coronavirus-pandemic-vaccine-updates-03-29-21/h_1f239fee1b0584ca9a5b6085357ac907, visited on 2022-01-20.

[6] CNN online, 2020-08-20 12:55 a.m. ET, "South Korea's latest church-linked coronavirus outbreak is turning into a battle over religious freedom", https://edition.cnn.com/world/live-news/coronavirus-pandemic-08-20-20-intl/h_288a15acd1b29e732c4e10693641088a, visited on 2022-01-20.

2020. This is most likely because, after the pandemic had reached the US and Europe in early 2020, the number of cases was consistently reported and the interpretations and implication of those case numbers are detected as changes in other topics. Even in the last months of the data set, in which the number of texts decreased and the results thus show a lower similarity, the twelfth topic retained a rather high similarity of above 0.75.

5. Discussion

In this paper, we presented a novel change detection method for text data. To construct coherently interpretably topics, we used RollingLDA to model a time series on textual data and compared the model's word distribution vectors with those of texts resampled under stable conditions. We applied our model on the TLS-Covid19 data set consisting of Covid-19 related news articles from CNN between January 2020 and December 2021.

Our method detects several meaningful changes in the evolving news coverage during the pandemic, including e.g. the start of vaccinations and several controversies over the course of the vaccination campaign as well as political changes such as the start of Joe Biden's presidential era. Out of 78 detected changes, we were instantly able to judge 55 (71%) as plausible ones based on manual labeling using the leave-one-out cosine impacts (cf. Fig. 1b, 1c and repository). The share increases to 78% if we exclude the turbulent initial phase of the Covid-19 pandemic and only consider changes since April 2020. While we cannot tell how many changes were missed out that could be considered as important as the ones mentioned above, our model contains a mixture parameter to calibrate the detection for general change of topics within a usual news week. If more, but less substantial or less, but more substantial changes are to be detected, this parameter p can be tuned accordingly. In combination with the maximum length of the reference period z_{max}, the set $\{p, z_{max}\}$ forms the model's hyperparameters to be optimized.

Acknowledgments

References

[1] J. Rieger, C. Jentsch, J. Rahnenführer, RollingLDA: An update algorithm of Latent Dirichlet Allocation to construct consistent time series from textual data, in: Findings Proceedings of the 2021 EMNLP-Conference, ACL, 2021, pp. 2337–2347. doi:10.18653/v1/2021.findings-emnlp.201.

[2] D. C. Montgomery, Introduction to statistical quality control, John Wiley & Sons, 2020.

[3] J. S. Oakland, Statistical process control, Routledge, 2007.

[4] A. Bose, S. S. Mukherjee, Changepoint analysis of topic proportions in temporal text data, 2021. arXiv:2112.00827.

[5] P. Fryzlewicz, Wild binary segmentation for multiple change-point detection, The Annals of Statistics 42 (2014) 2243–2281. doi:10.1214/14-AOS1245.

[6] R. P. Adams, D. J. MacKay, Bayesian online changepoint detection, 2007. arXiv:0710.3742.

[7] T. Kim, J. Choi, Reading documents for bayesian online change point detection, in: Proceedings of the 2015 EMNLP-Conference, ACL, 2015, pp. 1610–1619. doi:10.18653/v1/D15-1184.

[8] N. Keane, C. Yee, L. Zhou, Using topic modeling and similarity thresholds to detect events, in: Proceedings of the The 3rd Workshop on EVENTS: Definition, Detection, Coreference, and Representation, ACL, 2015, pp. 34–42. doi:10.3115/v1/W15-0805.

[9] Q. Mei, C. Zhai, Discovering evolutionary theme patterns from text: An exploration of temporal text mining, in: Proceedings of the 11th SIGKDD-Conference, ACM, 2005, pp. 198–207. doi:10.1145/1081870.1081895.

[10] Q. Liang, K. Wang, Monitoring of user-generated reviews via a sequential reverse joint sentiment-topic model, Quality and Reliability Engineering International 35 (2019) 1180–1199. doi:10.1002/qre.2452.

[11] L. Frermann, M. Lapata, A Bayesian model of diachronic meaning change, Transactions of the Association of Computational Linguistics 4 (2016) 31–45. doi:10.1162/tacl_a_00081.

[12] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent Dirichlet Allocation, Journal of Machine Learning Research 3 (2003) 993–1022. doi:10.1162/jmlr.2003.3.4-5.993.

[13] T. L. Griffiths, M. Steyvers, Finding scientific topics, Proceedings of the National Academy of Sciences 101 (2004) 5228–5235. doi:10.1073/pnas.0307752101.

[14] X. Song, C.-Y. Lin, B. L. Tseng, M.-T. Sun, Modeling and predicting personal information dissemination behavior, in: Proceedings of the 11th SIGKDD-Conference, ACM, 2005, pp. 479–488. doi:10.1145/1081870.1081925.

[15] D. M. Blei, T. L. Griffiths, M. I. Jordan, J. B. Tenenbaum, Hierarchical topic models and the nested chinese restaurant process, in: Advances in Neural Information Processing Systems, volume 16, MIT Press, 2003, pp. 17–24. URL: https://proceedings.neurips.cc/paper/2003/hash/7b41bfa5085806dfa24b8c9de0ce567f-Abstract.html.

[16] X. Wang, A. McCallum, Topics over time: A non-markov continuous-time model of topical trends, in: Proceedings of the 12th SIGKDD-Conference, ACM, 2006, pp. 424–433. doi:10.1145/1150402.1150450.

[17] D. M. Blei, J. D. Lafferty, Dynamic topic models, in: Proceedings of the 23rd ICML-Conference, ACM, 2006, pp. 113–120. doi:10.1145/1143844.1143859.

[18] C. Wang, D. M. Blei, D. Heckerman, Continuous time dynamic topic models, in: Proceedings of the 24th UAI-Conference, AUAI Press, 2008, pp. 579–586. URL: https://dl.acm.org/doi/10.5555/3023476.3023545.

[19] Y. Wang, C. Goutte, Real-time change point detection using on-line topic models, in: Proceedings of the 27th ACL-Conference, ACL, 2018, pp. 2505–2515. URL: https://www.aclweb.org/anthology/C18-1212.

[20] A. Pasquali, R. Campos, A. Ribeiro, B. Santana, A. Jorge, A. Jatowt, TLS-Covid19: A new annotated corpus for timeline summarization, in: Advances in Information Retrieval, ECIR 2021, volume 12656 of *LNCS*, 2021, pp. 497–512. doi:10.1007/978-3-030-72113-8_33.