Interactive Topic Model with Enhanced Interpretability

Jun Wang
Fujitsu Laboratories of America
Sunnyvale, CA
jun.wang@us.fujitsu.com

Changsheng Zhao
Columbia University
New York City, NY
cz2458@columbia.edu

Junfu Xiang
Fujitsu Nanda Software Tech. Co., Ltd.
Nanjing, China
xiangjf.fnst@cn.fujitsu.com

Kanji Uchino
Fujitsu Laboratories of America
Sunnyvale, CA
kanji@us.fujitsu.com

ABSTRACT

Although existing interactive topic models allow untrained end users to easily encode their feedback and iteratively refine the topic models, their unigram representations often result in ambiguous description of topics and poor interpretability for users. To address the problems, this paper proposes the first phrase-based interactive topic model which can provide both high interpretability and high interactivity with human in the loop. First, we present an approach to augment unigrams with a list of probable phrases which offers a more intuitively interpretable and accurate topic description, and further efficiently encode users' feedback with phrase constraints in interactive processes of refining topic models. Second, the proposed approach is demonstrated and examined with real data.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; • **Computing methodologies** → **Machine learning**.

KEYWORDS

Interactive Topic Model; Interpretability; Human in the loop

---

1 INTRODUCTION

Topic models are a useful and ubiquitous tool for understanding large electronic archives, which can be used to discover the hidden themes that pervade the collection and annotate the documents according to those themes, and further organize, summarize, and search the texts [4]. However, as fully-unsupervised methods, vanilla topic models, such as Latent Dirichlet allocation (LDA) [4], often generate some topics which do not fully make sense to end users [10]. Some generated topics may not well correspond to meaningful concepts, for instances, two or more themes can be confused into one topic or two different topics can be (near) duplicates. Some topics may not align well with user modeling goals or judgements. For many users in computational social science, digital humanities, and information studies, who are not machine learning experts, topic models are often a "take it or leave it" proposition [6, 10]. Different from purely unsupervised topic models that often result in unexpected topics, taking prior knowledge into account enables us to produce more meaningful topics [20]. Interactive topic models with human in the loop are proposed and allow untrained end users to easily encode their feedback as prior knowledge and iteratively refine the topic models (e.g., changing which words are included in a topic, or merging or splitting topics) [10, 12, 16].

A topic is typically modeled as a categorical distribution over terms, and frequent terms related by a common theme are expected to have a large probability [8]. It is of interest to visualize these topics in order to facilitate human interpretation and exploration of the large amounts of unorganized text, and a list of most probable terms is often used to describe individual topics. Similar to vanilla topic models, all existing interactive topic models are represented with unigrams, which often provide ambiguous representation of the topic and poor interpretability for end users [8]. Smith et al. [16] conducted user studies on a unigram-based interactive topic model, and also were aware of the requests from

Table 1: Real example of a topic discovered by the unigram-based topic model and the phrase-based topic model, respectively.

Topic: natural language processing	
unigrams	phrases
model	word embeddings
language	natural language
word	language model
text	question answering
task	machine translation
question	sentiment analysis
sentence	neural machine translation
translation	natural language processing
neural	text classification
natural	word representation

participants for the ability to add phrases and support of multi-word refinements as opposed to single tokens.

As shown in Table 1, human interpretation often relies on inherent grouping of words into phrases, and augmenting unigrams with a list of probable phrases offers a more intuitively interpretable and accurate topic description [8]. Under the 'bag-of-words' assumption of unigrams, phrases are decomposed and a phrase's meaning may be lost, so topic models need to systematically assign topics to whole phrases. Several phrase-based topic models [3, 7, 8, 19] have been proposed to discover topical phrases and address the prevalent deficiency in visualizing topics using unigrams. But all these models are static systems which end users cannot easily and interactively refine, so they have the same "take it or leave it" issues.

To address the above problems, this paper proposes the first phrase-based interactive topic model which can provide both high interpretability and high interactivity as shown in Figure 1. First, we present an approach to discover topical phrases with mixed lengths by detecting phrases and phrase-based topic inference, and further efficiently encode users' feedback with phrase constraints into interactive processes of refining topic models. Second, the proposed approach is demonstrated and examined with real data.

We organize the remainder of the paper as follows. Section 2 introduces some related work. Section 3 illustrates the general framework we propose. Section 4 presents our experimental results on real-world data. Finally, section 5 summarizes our work and discuss the future work.

2 RELATED WORK

Various approaches have been proposed to encode users' feedback as prior knowledge into topic models instead of purely relying on how often words co-occur in different contexts. Andrzejewski et al. [1] imposed Dirichlet Forest prior over the topic-word categoricals to encode the Must-Links

and Cannot-Links between words. Words with Must-Links are encouraged to have similar probabilities within all topics while those with Cannot-Links are disallowed to simultaneously have large probabilities within any topic. Xie et al. [20] studied how to incorporate the external word correlation knowledge to improve the coherence of topic modeling, and built a Markov Random Field (MRF) regularized topic model encouraging words labeled as similar to share the same topic label. Yang et al. [21] integrated lexical association into topic optimization using tree priors to improve topic interpretability, which provided a flexible framework that can take advantage of both first order word associations and the higher-order associations captured by word embeddings.

Several unigram-based interactive topic models have been proposed and studied. Hu et al. [10] extended the framework of Dirichlet Forest prior and proposed the first interactive topic model. Lee et al. [12] employed a user-centered approach to identify a set of topic refinement operations that users expect to have in a interactive topic model system. However, they did not implement underlying algorithm to refine topic models and only used Wizard-of-Oz refinements: the resulting topics were updated superficially—not as the output of a data-driven statistical model (the goal of topic models) [16]. Smith et al. [16] further implemented an efficient asymmetric prior-based interactive topic model with a broader set of user-centered refinement operations, and conducted a study with twelve non-expert participants to examine how end users are affected by issues that arise with a fully interactive, user-centered system.

Some researchers proposed various phrase-based topic models. Wang et al. [19] attempted to infer phrases and topics simultaneously by creating complex generative mechanism. The resultant models can directly output phrases and their latent topic assignment. It used additional latent variables and word-specific multinomials to model bi-grams, and these bigrams can be combined to form n-gram phrases. KERT [7] and Turbo Topics [3] constructed topical phrases as a post-processing step to unigram-based topic models. These methods generally produce low-quality topical phrases or suffer from poor scalability outside small datasets [8]. El-Kishky et al. [8] and Wang et al. [18] proposed a computationally efficient and effective approaches, which combines a phrase mining framework to segment a document into single and multi-word phrases, and a topic model with phrase constraints that operates on the induced document partition.

3 FRAMEWORK

For phrase-based topic models, the better method is first mining phrases and segmenting each document into single and multiword phrases, and then running topic inference with phrase constraints [8]. End users can give feedback using a
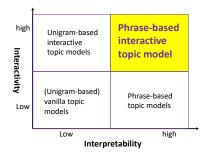
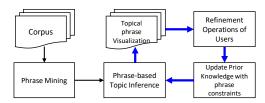**Figure 1: Comparison of our phrase-based interactive topic models with other related topic models.**

Figure 2: The framework of phrase-based interactive topic model

variety of refinement operations on topical phrase visualization, and users' feedback will update the prior knowledge and the phrase-based topic inference will be rerun based on the updated prior. As shown in Figure 2, the process forms a loop in which users can continuously and interactively update and refine the topic model with phrase constraints.

Phrase Mining

Phrase mining is a text mining technique that discovers semantically meaningful phrases from massive text. Recent data-driven approaches opt instead to make use of frequency statistics in the corpus to address both candidate generation and quality estimation [7, 13, 15, 18]. They do not rely on complex linguistic feature generation, domain-specific rules or extensive labeling efforts. Instead, they rely on large corpora containing hundreds of thousands of documents to help deliver superior performance several indicators, including frequency, mutual information, branching entropy and comparison to super/sub-sequences, were proposed to extract n-grams that reliably indicate frequent, concise concepts [7, 13, 15, 18].

Phrase-based topic inference

After inducing a partition on each document, we perform topic inference to associate the same topic to each word in a phrase and thus naturally to the phrase as a whole. El-Kishky et al. [8] proposed a probabilistic graphical model PhraseLDA

based on a generative process almost same as LDA but with constraints on topics of phrases, and corresponding phrase-based topic inference can be smoothly updated from unigram-based topic inference of LDA.

LDA assumes that a document may contain multiple topics, where a topic is defined to be a categorical distribution over words in the vocabulary. The generative process is as follows:

(1) Draw $\phi_k \sim Dirichlet(\beta)$, for $1 \leq k \leq K$
(2) For document d, where $1 \leq d \leq D$:
　(a) Draw $\theta_k \sim Dirichlet(\alpha)$
　(b) For n-th word in document d, where $1 \leq n \leq N_d$
　　(i) Draw $z_{d,n} \sim Categorical(\theta_d)$
　　(ii) Draw $w_{d,n} \sim Categorical(\phi_{z_{d,n}})$

α is a K-dimensional vector $(\alpha_1, \ldots, \alpha_K)$, and β is a V-dimensional vector $(\beta_1, \ldots, \beta_V)$. K is the number of topics, D is the number of documents, V is the size of vocabulary, and N_d is the number of words in the document d.

Based on its generative process, the joint distribution of LDA (1) can be represented as the product of two Dirichlet-Multinomial distributions (2). The Dirichlet-Multinomial expressions (3) can be further simplified using the feature of gamma function (represented by Γ) later.

$$p(W, Z; \alpha, \beta) \tag{1}$$

$$= \int p(W, Z, \Theta, \Phi; \alpha, \beta) d\Theta d\Phi$$

$$= \int p(Z, \Theta; \alpha) d\Theta \times \int p(W, \Phi | Z; \beta) d\Phi$$

$$= p(Z; \alpha) \times p(W | Z; \beta)$$

$$= DirMult(Z; \alpha) \times DirMult(W | Z; \beta) \tag{2}$$

$$\propto \left(\prod_{d=1}^{D} \frac{\prod_{k=1}^{K} \Gamma(N_{d,k} + \alpha_k)}{\Gamma(\sum_{k=1}^{K}(N_{d,k} + \alpha_k))} \right)$$

$$\times \left(\prod_{k=1}^{K} \frac{\prod_{v=1}^{V} \Gamma(N_{k,v} + \beta_v)}{\Gamma(\sum_{v=1}^{V}(N_{k,v} + \beta_v))} \right) \tag{3}$$

W is the collection of all words in D documents, and Z is the collection of corresponding topics assigned to each word in W. Θ is the collection of $(\theta_1, \ldots, \theta_K)$, and Φ is the collection of $(\phi_1, \ldots, \phi_K)$. $N_{d,k}$ is the number of words assigned to topic k in the document d, and $N_{k,v}$ is the number of words with topic k and value v in the vocabulary.

Because the generative process of PhraseLDA is almost same as LDA except phrase constraints on topics, its joint distribution is same as LDA in the above (3). But PhraseLDA and LDA are different in calculating the full conditional distribution (4), by which we can sample topics using Gibbs sampling. And we know that the full conditional distribution (4) is proportional to the joint distribution (1).

$$p(z_{a,b} = i|Z_{\neg a,b}, W; \alpha, \beta) \quad (4)$$
$$= p(z_{a,b} = i|w_{a,b} = j, Z_{\neg a,b}, W_{\neg a,b}; \alpha, \beta)$$
$$\propto p(W, Z; \alpha, \beta)$$

$z_{a,b}$ is the topic assigned to the $w_{a,b}$, which is the b-th unit in the document a. $W_{\neg a,b}$ is the collection of all units except $w_{a,b}$, and $Z_{\neg a,b}$ is the collection of corresponding topics assignments. In LDA $w_{a,b}$ is the b-th word in the document a, and in PhraseLDA $w_{a,b}$ is the b-th phrase in the document a, and this difference results in different topic inference processes.

For PhraseLDA, we can simplify two Dirichlet-Multinomial expressions (3) to sample the topic of a phrase as follows, and please see Appendix for detail derivations.

$$p(z_{a,b} = i|Z_{\neg a,b}, W; \alpha, \beta)$$
$$\propto \prod_{g=1}^{l_{a,b}} \left(\frac{(N_{a,i}^{\neg a,b} + \alpha_i + g - 1) \times (N_{i,w_{a,b,g}}^{\neg a,b} + \beta_{w_{a,b,g}})}{(g - 1 + \sum_{v=1}^{V}(N_{i,v}^{\neg a,b} + \beta_v))} \right) \quad (5)$$

$l_{a,b}$ is the length of the b-th phrase in the document a, and $w_{a,b,g}$ is the g-th word in phrase $w_{a,b}$. $N_{a,i}^{\neg a,b}$ is the number of words assigned to topic i in the document a after excluding the phrase $w_{a,b}$, and $N_{i,v}^{\neg a,b}$ is the number of words with topic i and value v after excluding the phrase $w_{a,b}$.

α and β can be optimized using the method presented by Minka [14] for the phrase-based topic model before refinement operations of users.

Refinement Operations of Users

Smith et al. [16] identified a set of refinements that users expected to be able to use in a interactive topic model, and implemented seven refinements requested by users: **add word**, **remove word**, **change word order**, **remove document**, **split topic**, **merge topic**, and **add to stop words**.

Participants of the qualitative evaluation in [16] found **change word order** to be one of the least useful refinements, and as shown in Table 1, with the phrase representation of topics the phrase order does not have much influence on human interpretability. **Add to stop words** is easy, and we just exclude the word w from the vocabulary and ensures that the Gibbs sampler ignores all occurrences of w in the corpus. So we can skip detail discussions of these two operations in the paper, and extend other operations based on phrases instead of words.

Update Prior Knowledge with phrase constraints

Adding a human in the loop requires the user to be able to inject their knowledge via feedback into the sampling equation to guide the algorithm to better topics [16].

Dirichlet Forest prior has been widely used to encode users' feedback as prior knowledge in various interactive topic models [1, 10, 16]. This kind of priors attempted to enforce hard and topic-independent rules that similar words should have similar probabilities in all topics, which is questionable in that two words with similar representativeness of one topic are not necessarily of equal importance for another topic [20]. For example, in the fruit topic, the words apple and orange have similar representativeness, while in an IT company topic, apple has much higher importance than orange. Dirichlet Forest prior is unable to differentiate the subtleties of word sense across topics and would falsely put irrelevant words into the same topic [20]. For instance, since orange and Microsoft are both labeled as similar to apple and are required to have similar probabilities in all topics as apple has, in the end, they will be unreasonably allocated to the same topic.

Wallach et al. [17] has found that an asymmetric Dirichlet prior has substantial advantages over a symmetric prior in topic models, and to address the above problems, Smith et al. [16] proposed an asymmetric prior which encodes users' feedback through modifying the Dirichlet prior parameters for each document and each topic involved. Similar idea can be extended to address phrase constraints and applied to phrase-based interactive model. In the previous section on phrase-based topic inference, all documents share the same α and all topics share the same β. Here, every document a and every topic i involved in refinement operations need corresponding separate $\alpha^{(a)}$ and $\beta^{(i)}$, respectively, and the sampling equation (5) should be updated as follows:

$$\prod_{g=1}^{l_{a,b}} \left(\frac{(N_{a,i}^{\neg a,b} + \alpha_i^{(a)} + g - 1) \times (N_{i,w_{a,b,g}}^{\neg a,b} + \beta_{w_{a,b,g}}^{(i)})}{(g - 1 + \sum_{v=1}^{V}(N_{i,v}^{\neg a,b} + \beta_v^{(i)}))} \right) \quad (6)$$

These priors $\alpha^{(a)}$ and $\beta^{(i)}$ are sometimes called "pseudo-counts" [9], and interactive models can take advantage of them by creating pseudo-counts to encourage the changes users want to see in the topic [16].

Remove document and **Merge topic** are straightforward and almost same as the unigram-based updates proposed in [16].

- **Remove document**: to remove the document a from topic i, we invalidate the topic assignment for all words in the document a and assign a very small prior $\alpha_i^{(a)}$ to the topic i in a.
- **Merge topic**: merging topics i_1 and i_2 means the model will have a combined topic that represents both i_1 and i_2. We assign i_1 to all words that were previously assigned to i_2, and reduce the number of topics.

For **Remove phrase**, **Split topic** and **Add phrase**, corresponding updates are a bit more complicated since we

need to deal with phrase constraints. For a phrase p, l_p is the length of p and p_g is g-th word in p where $1 \leq g \leq l_p$.

- **Remove phrase**: to remove the phrase p from topic i, we need to locate all occurrences of p assigned to topic i and invalidate their topic assignment. For topic i, very small prior $\beta_{p_g}^{(i)}$ is assigned to each word p_g contained in p.
- **Split topic**: to split topic i_1 the user provides a subset of seed phrases, which need to be moved from the original topic i, to a new topic i_2. We invalidate the original topic assignment of all seed phrase occurrences, increase the number of topics, and assign large prior $\beta_{p_g}^{(i_2)}$ for each word p_g contained in each seed phrase p for the new topic i_2.
- **Add phrase**: to add the phrase p to topic i, we invalidate all occurrences of p from all other topics and encourage the Gibbs sampler to assign topic i for each occurrence, and we increase the prior of each word contained in p for topic i.

4 EXPERIMENTS

We deployed the phrase-based interactive topic models as a part of our corporate learning platform for data scientist training programs, in which a database contains 19852 recent machine learning related papers collected from ICML/NIPS/arXiv.

For phrase mining, we used our own tool based on generalized suffix tree (GST) presented in [18], and segmented the titles and abstracts of all papers into a collections of meaningful phrases.

In order to facilitate human exploration and interpretation, we visualize these papers into 20 topics using our system, and learners can further interactively refine the topics using their domain knowledge as shown in Figure 3. A list of topics in the left panel are represented by top three phrases of each topic. Selecting a topic displays more detail in the right panel: the top 30 phrases with frequency and top associated documents with corresponding percentage. Users can click and select phrases for removing with **remove phrase** button or for splitting with **split topic** button, click and select documents for removing with **remove document** button, add new phrases from the vocabulary with **add phrase** button, select phrases and click the **add to stop words** button to move to the stop words list, or click **merge topic** button to input two topics for merging.

Before we implemented our phrase-based interactive model illustrated in Figure 2, we first tried the model based on Dirichlet Forest prior presented in [10], and found a few drawbacks. Instead of direct modification, people are forced to think of pairwise relation, which is counter-intuitive. Its prior tree structure is hard to encode phrase constraints and results in an extremely slow convergence, whose latency is

more than 50 times of our system. We also tried to extend the model presented in [21] to support phrases, and check if human interpretability of generated topic are improved. Correlation scores based on phrase embedding vectors generated by Fasttext [5] are calculated to build two-level tree prior. The model attempts to encourage phrases close in embedding vector space to appear in the same topic, but we found that it only performs slightly better on downstream tasks, such as classification, and does not really help to enhance human interpretability. The above observations led to creating our current system.

Our phrase-based topic model before refinement operations was initialized with 2000 iterations using the optimized α (with mean 0.415) and β (with mean 0.015). Since this is a one-time job, we can set an even larger iteration number. The number of sampling iterations for updating and refining model can be tuned according to latency acceptable for users (for example, less than 1 minute), and we set the number as 400. Similar to [16], $\beta_{p_g}^{(i)}$ is set as 0.000001 for **remove phrase** and **split topic**.

Since this paper focuses on improving human interpretability of interactive topic models, and as we have known, the automated methods of measuring topic quality in terms of coherence often do not correlate well with human judgement and interpretation, and in addition, these methods are generally only available for unigram-based models, so the experimental evaluation in this paper are mainly based on user studies. 5 participants with computer science or electronic engineering background, who are users of the corporate learning platform, were asked to use and refine the phrase-based interactive topic model.

Our user studies showed the **split topic** and **remove phrase** are the most commonly used operations, and occasionally **merge topic** is used based on users' personal preference. But **add phrase** is a relatively rare operation, because in most cases it is not easy for users to discover or remember phrases not presented to them, especially for a new domain.

There are a couple of coherent but non-informative topics. For example, one topic mainly contains phrases such as *training data, data sets, data points*, and another topic mainly contains phrases such as *experimental results, theoretical results*. Except these uninformative topics, all 5 participants agreed that our system can significantly refine quality and coherence of all other topics and consistently improve human interpretability of topic modeling. The user studies showed that a well-organized structure can be established and refined by our phrase-based interactive topic model.

Several typical examples from participants' real refinement operations are demonstrated here. In the first example shown in Figure 4, a participant found that two unrelated
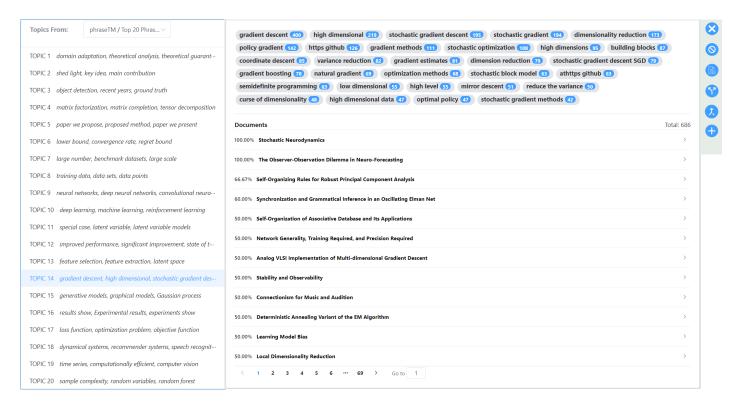
Figure 3: UI of the phrase-based interactive topic model.

topics (social media and Autonomous driving) were mistakenly mixed into one topic, and she selected *Social media* as a seed phrase for **split topic**. In the second example shown in Figure 5, the existing topic was actually fine, but a participant wanted to refine and separate a fine-grained new topic on face recognition from the existing topic on image processing, and she selected *Face recognition* as a seed phrase for **split topic**. Interestingly, although only one seed phrase was selected for the new topic in the above two examples, other unselected phrases related to the seed can correctly move to the new topic as well. In the third example, a participant found that an important phrase *Computer vision* was assigned to a unexpected and inappropriate topic which is not really meaningful, and she wanted to remove *Computer vision* from this inappropriate topic and check if it is possible to finally move it to a meaningful topic. After two rounds of **remove phrase**, the phrase *Computer vision* moved to an appropriate topic as shown in Figure 6.

5 CONCLUSION AND FUTURE WORK

This paper proposes the first phrase-based interactive topic model which can provide both high interpretability and high interactivity with human in the loop, and demonstrates and examines the proposed approach with real data. Although
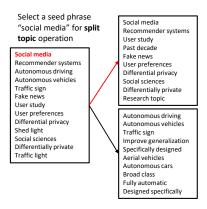

Figure 4: Real example of splitting two unrelated topics.

the latency of our system has significantly improved compared with previous systems based on tree prior, it can still be a major issue for large scale data, so we need to study more efficient methods of inference using sparsity [22], which can be smoothly applied to systems with phrase constraints. Current methods for automatically measuring topic coherence and quality are also mainly for models based on unigrams
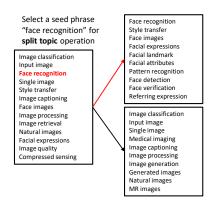
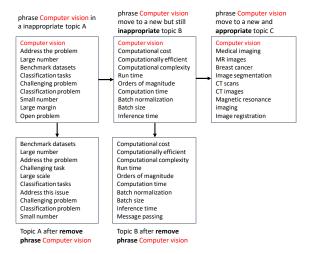**Figure 5: Real example of separating a fine-grained new topic.**

Figure 6: Real example of removing phrase from inappropriate topics after two rounds and moving to an appropriate topic finally.

[2, 11], so we also need to study how to extend corresponding methods for phrase-based models.

REFERENCES

[1] David Andrzejewski, Xiaojin Zhu, and Mark Craven. 2009. Incorporating Domain Knowledge into Topic Modeling via Dirichlet Forest Priors. In *Proceedings of the 26th Annual International Conference on Machine Learning (ICML '09)*. 25–32.

[2] Shraey Bhatia, Jey Han Lau, and Timothy Baldwin. 2018. Topic Intrusion for Automatic Topic Model Evaluation. In *EMNLP*. Association for Computational Linguistics, 844–849.

[3] D. Blei and J. Lafferty. 2009. Visualizing Topics with Multi-Word Expressions. *arXiv* 0907.1013v1 (2009).

[4] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 3 (March 2003), 993–1022.

[5] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2016. Enriching Word Vectors with Subword Information. *arXiv preprint arXiv:1607.04606* (2016).

[6] Jordan Boyd-Graber. 2017. Humans and Computers Working Together to Measure Machine Learning Interpretability. *The Bridge* 47 (2017), 6–10.

[7] Marina Danilevsky, Chi Wang, Nihit Desai, Xiang Ren, Jingyi Guo, and Jiawei Han. 2014. Automatic Construction and Ranking of Topical Keyphrases on Collections of Short Documents. In *Proceedings of the 2014 SIAM International Conference on Data Mining, Philadelphia, Pennsylvania, USA, April 24-26, 2014*. 398–406.

[8] Ahmed El-Kishky, Yanglei Song, Chi Wang, Clare R. Voss, and Jiawei Han. 2014. Scalable Topical Phrase Mining from Text Corpora. *Proc. VLDB Endow.* 8, 3 (Nov. 2014), 305–316.

[9] Gregor Heinrich. 2004. *Parameter estimation for text analysis*. Technical Report.

[10] Yuening Hu, Jordan Boyd-Graber, Brianna Satinoff, and Alison Smith. 2014. Interactive Topic Modeling. *Machine Learning* 95 (2014), 423–469.

[11] Jey Han Lau, David Newman, and Timothy Baldwin. 2014. Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality. In *EACL*. The Association for Computer Linguistics, 530–539.

[12] Tak Yeon Lee, Alison Smith, Kevin Seppi, Niklas Elmqvist, Jordan Boyd-Graber, and Leah Findlater. 2017. The Human Touch: How Non-expert Users Perceive, Interpret, and Fix Topic Models. *International Journal of Human-Computer Studies* (2017).

[13] Jialu Liu, Jingbo Shang, Chi Wang, Xiang Ren, and Jiawei Han. 2015. Mining Quality Phrases from Massive Text Corpora. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD '15)*. 1729–1744.

[14] Thomas P. Minka. 2000. *Estimating a Dirichlet distribution*. Technical Report. MIT.

[15] J. Shang, J. Liu, M. Jiang, X. Ren, C. R. Voss, and J. Han. 2018. Automated Phrase Mining from Massive Text Corpora. *IEEE Transactions on Knowledge and Data Engineering* 30, 10 (2018), 1825–1837.

[16] Alison Smith, Varun Kumar, Jordan Boyd-Graber, Kevin Seppi, and Leah Findlater. 2018. Closing the Loop: User-Centered Design and Evaluation of a Human-in-the-Loop Topic Modeling System. In *23rd International Conference on Intelligent User Interfaces (IUI '18)*. 293–304.

[17] Hanna M. Wallach, David Mimno, and Andrew McCallum. 2009. Rethinking LDA: Why Priors Matter. In *Proceedings of the 22Nd International Conference on Neural Information Processing Systems (NIPS'09)*. 1973–1981.

[18] Jun Wang, Junfu Xiang, and Kanji Uchino. 2015. Topic-Specific Recommendation for Open Education Resources. In *Advances in Web-Based Learning – ICWL 2015*, Frederick W.B. Li, Ralf Klamma, Mart Laanpere, Jun Zhang, Baltasar Fernández Manjón, and Rynson W.H. Lau (Eds.). Springer International Publishing, Cham, 71–81.

[19] Xuerui Wang, Andrew McCallum, and Xing Wei. 2007. Topical N-Grams: Phrase and Topic Discovery, with an Application to Information Retrieval. In *Proceedings of the 2007 Seventh IEEE International Conference on Data Mining (ICDM '07)*. 697–702.

[20] Pengtao Xie, Diyi Yang, and Eric P Xing. 2015. Incorporating Word Correlation Knowledge into Topic Modeling. In *The 2015 Conference of the North American Chapter of the Association for Computational Linguistics*.

[21] Weiwei Yang, Jordan Boyd-Graber, and Philip Resnik. 2017. Adapting Topic Models using Lexical Associations with Tree Priors. In *Empirical Methods in Natural Language Processing*.

[22] Yi Yang, Doug Downey, and Jordan Boyd-Graber. 2015. Efficient Methods for Incorporating Knowledge into Topic Models. In *Empirical Methods in Natural Language Processing*.