Automatic News Article Generation from Legislative Proceedings: A Phenom-based Approach

Anastasiia Klimashevskaia[1[0000−0002−8946−667X]], Richa Gadgil[2[0000−0001−5469−1212]], Thomas Gerrity[2[0000−0001−8693−9809]], Foaad Khosmood[2[0000−0003−2531−6610]], Christian Gütl[1[0000−0001−9589−1966]], and Patrick Howe[2]

[1] Graz University of Technology, Graz, Austria
a.klimashevskaia@student.tugraz.at, c.guetl@tugraz.at
[2] California Polytechnic State University, San Luis Obispo, USA
{rgadgil,tmgerrit,foaad,pchowe}@calpoly.edu

Abstract. Algorithmic journalism refers to automatic AI-constructed news stories. There have been successful commercial implementations for news stories in sports, weather, financial reporting and similar domains with highly structured, well defined tabular data sources. Other domains such as local reporting have not seen adoption of algorithmic journalism, and thus no automated reporting systems are available in these categories which can have important implications for the industry. In this paper, we demonstrate a novel approach for producing news stories on government legislative activity, an area that has not widely adopted algorithmic journalism. Our data source is state legislative proceedings, primarily the transcribed speeches and dialogue from floor sessions and committee hearings in US State legislatures. Specifically, we create a library of potential events called phenoms. We systematically analyze the transcripts for the presence of phenoms using a custom partial order planner. Each phenom, if present, contributes some natural language text to the generated article: either stating facts, quoting individuals or summarizing some aspect of the discussion. We evaluate two randomly chosen articles with a user study on Amazon Mechanical Turk with mostly Likert scale questions. Our results indicate a high degree of achievement for accuracy of facts and readability of final content with 13 of 22 users in the first article and 19 of 20 subjects of the second article agreeing or strongly agreeing that the articles included the most important facts of the hearings. Other results strengthen this finding in terms of accuracy, focus and writing quality.

Keywords: algorithmic journalism · natural language generation · automatic summarization · partial order planning · artificial intelligence · digital government

1 Introduction and Motivation

We present a novel method to generate news stories on state legislative proceedings. Our strategy is to generate algorithmic news content that summarize the important events of a single discussion based on records of meetings and other data sources. Such content could be distributed to news organizations which could in turn print them in their publications or use them as a basis for writing more detailed stories.

The primary input data into this system are high quality, human-verified transcripts from hearings in state legislatures as developed by the Digital Democracy project [2] and already used in several other works [3, 12, 26, 27]. Hearings are divided into "bill discussions", or dialogues about a specific legislation that are on average about 20 minutes of multi-speaker discussion, typically followed by a vote. In this paper we use a portion of the data consisting of transcripts from 5198 bill discussions containing 40788 individual speeches covering the California Legislature 2015-2017.

Our system generates articles based on abstractive summarizations of the bill discussions achieved using a planning system and a "phenom" (short for phenomenon) library that we also develop. We define and model the phenom as a software object with several standard attributes and methods. Each phenom is called by the planner to test for a certain specific event, condition or occurrence in the text of the discussion. If the search is successful, the phenom then contributes one or more sentences to the text of the article. The surface text realization is achieved through insertion of facts into pre-written English language sentence templates appropriately chosen for each phenom.

Due to strict requirements for high accuracy and traceability of all facts to original primary sources, we do not rely on predictive or transformer NLG models which are prone to "hallucinations" or tendency to generate inaccurate or nonsensical statements [10]. For quality purposes, we also verify the approach using human evaluations.

1.1 Motivation

News media play a crucial role in the functioning of democracies through coverage of governmental processes, but such coverage is in decline. The Pew Research Center found that the number of full-time reporters covering state legislatures fell by 35 percent between 2003 and 2014 [8]. Observers say the decline has likely accelerated in the years since the study was conducted [30]. The decline largely came from the ranks of newspapers, which have historically provided the greatest proportion of full-time statehouse reporters (only 15 percent of television stations send any reporters to cover state legislatures) [8]. Facing declining circulation and advertising revenues, U.S. newspapers shed 51 percent of their editorial employees between 2008 and 2019 [11].

With fewer staff, news outlets have decreased their coverage of governmental proceedings [1, 6, 18]. The decline in dedicated statehouse reporters has led to less coverage of policy and the state legislative process, with most remaining coverage focusing narrowly on legislative outcomes such as the final passage of a bill [33, 34]. State legislatures themselves have responded by bulking up their own in-house media offerings via websites, video reports and social media [33].

1.2 Organization

The rest of this paper is organized as follows: in Section 2 we provide a short overview on the previous research and related works. Section 3 introduces first the concept, then details further the developments and technical issues of the project. It is followed by Section 4 with a description of the design and procedure of the user study. Section 4.2 is a discussion about our findings. Finally, Section 5 concludes the paper, giving some final ideas for future work.

2 Related Work

Summarization of dialog or meeting transcripts is a tricky task for programmers. Several experiments described applying classic extractive summarization methods on speech transcripts. It was concluded that the spontaneous nature of speech results in lower quality summaries compared to written text [10, 5]. Meeting transcripts consist of unstructured utterances with long-range semantic dependencies [32]. They can contain more grammatical and spelling errors and are more noisy, thus producing a less readable and concise summary using extractive techniques [16, 19]. Still, there are attempts to utilize an extractive approach [10, 4, 25, 35] with transcripts. Template-based NLG was previously successfully utilized in news generation, with human written templates [20, 29] as well as generated ones [23].

The approach of a flat facts set described by [13] with a certain addition to fact types and template adjustment served as a big inspiration for this project. This methodology in combination with hand-written templates as in [20, 29] could provide good results in accuracy and grammatically of the final text. Event, entity and relation recognition is one of the good examples for fact extraction approaches. Works like [14, 7, 21] give great descriptions of such event detection systems, however, they are mostly concentrating on extraction from coherent written text, while our task is basing on transcripts of spoken language, which quite often lacks proper grammar, coherence and contains transcription errors, incomplete sentences, interruptions, etc. Nevertheless, these works still provide invaluable insights and ideas on event-entity-relation detection and extraction.

Our contribution is in the system of phenoms that combine event detection, template based realization for news articles and a modified partial-order planner through which a semantic structure of a news article is enforced.

3 Approach and Development

We design a system to ingest records of legislative bill discussions and produce an article as the main deliverable (see Fig. 1). Initially, all discussions are identified

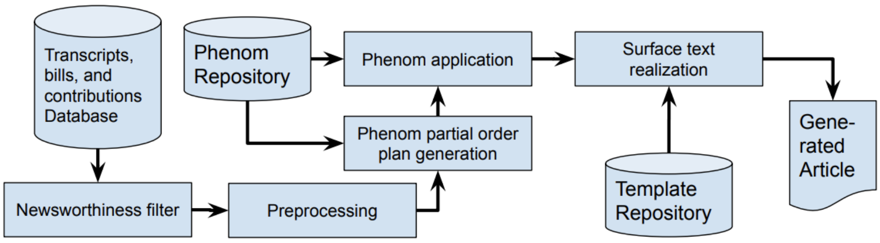

Fig. 1. The data flow diagram of the system

and extracted from the database by a newsworthiness filter, developed outside the focus of this paper to simply prioritize certain bill discussions over others. Transcripts then undergo preprocessing where every utterance from the transcript is tokenized, tagged for speaker, function, and named entities resulting in Fig. 2.

utterance _order	pid	text	Paragraph label	Secondary label
1	92	Everybody turn off their cell phones if they want to bother us while we spend the next 3 to 4 minutes in this committee. And lets take our roll. We have 4 measures, all on consent items. And all the items on our agenda are on consent. So we'll take the roll and then we have a motion on the consent calendar.	1	
2	2998	Senators Beall? Here. Beall present, Cannella? Here. Cannella present. Allen? Bates? Gaines? Galgiani? Leyva? Here. Leyva present. McGuire? Here. McGuire present. Mendoza? Here. Mendoza present. Roth? Here. Roth, present. Wieckowski? Here. Wieckowski present.	0	
3	92	Okay these consent calendar items are ACR 58 by various Assembly Members, Williams, ACR 63, Maienschein, and ACR 65, Brough, ACR 78, Salas, and if there's no discussion on these items, we'll have a motion for the approval of the consent calendar. The motion's to approve this consent calendar, and we'll take a vote.	2	4
4	2998	On the consent calendar, Senators Beall? Aye. Beall aye, Cannella? Aye. Cannella aye. Allen? Bates? Gaines? Galgiani? Leyva? Aye. Leyva aye. McGuire? Aye. McGuire aye. Mendoza? Aye. Mendoza aye. Roth? Aye. Roth aye. Wieckowski? Aye. Wieckowski aye.	4	
5	92	7? We're gonna put that on call, we have 7 votes in favor. So we'll wait for other members to come and record their votes, and as soon as they come and vote, we will adjourn the committee. So we're waiting on 4 Senators, and we'll wait for them and record their votes and adjourn.	0	
6	2998	Thank you Senator Beall.	0	
7	92	Thank you.	4	
8	92	Consent calendar.	2	
9	2998	Senators Allen? Bates? Gaines? Galgiani? Galgiani aye.	4	

Fig. 2. Example of a tagged transcript portion in preprocessing

3.1 Phenom system

Legislative proceedings vary greatly, but we find some useful patterns and data in hearing videos and their written transcripts. This inspired the adoption of the phenom approach - systematic extraction of the key phenomena from the transcript text and storing them in a collection of fact data types to used for generating language for the article.

We create a library of phenoms for content selection. Each phenom is an object with data and methods designed to look for a specific event in the bill discussion, and take some action if the event is found. For example does a greeting occur in this discussion? A phenom called "greetings" could look for specific language that constitutes a greeting according to hand-crafted criteria. If found, the phenom generates facts leading to a statement like: "Senator Monning greeted the committee". That sentence becomes part of the news article. Given many phenoms each dedicated to typical occurrences in legislative hearings, we can generate many observations.

The order of executions of the phenoms are not set ahead of time. Instead every phenom has a set of preconditions and postconditions. A modified partial order planner selects a phenom from the library based on preconditions and executes it. After execution, that phenom may contribute some more conditions to the current state of the article that may enable other phenoms to be activated. In this way, some phenom output are guaranteed to be in specific places in the article, such as the beginning section or the ending section. They can also form chains of dependencies allowing much more conditioned observations when called for. For example a news story may have a short statement about the background of a certain expert witness generated by phenom #2, but only if the witness is first called to testify, which would have been observed by phenom #1 (the precondition of phenom #2).

Phenom ideas are created based on several sources: Fist, a study of 300 real published news articles that contained the name of a bill that was before the California Legislature at the time of the article publication. These provided some standard observations that journalists include in news articles about state legislative action. Second, some interesting data available to us from the database were made into phenoms. For example, we can automatically flag the first-ever bill authored by a certain legislator, or the first bill of the session by that legislator. This small point may be very important to a hometown news organization but not immediately apparent to a reporter. Lastly, some purely functional phenoms are created to help with content arrangement, to ensure, for example, that a pull-quote not appear until the middle section of an article. For a selection of some of the phenoms and their respective frequencies of occurrence see Table 1.

3.2 Illustrative phenom examples

While certain phenoms have mostly formal purpose and use the metadata provided by the database, the others are designed to search for complex patterns and work directly with the transcript text, utilizing classification or scoring algorithms. Metadata-based and functional phenoms occur in the texts more often due to higher likelihood of availability of the data, while the more refined ones tend to be rare and are triggered not as often (see Table. 1). The following paragraphs give examples of such high-order phenoms present in the system.

Pull-quote extraction One of the phenoms implemented in the pipeline is aimed at pull-quote extraction from the transcript text. A pull-quote is a key sentence, or phrase that is highlighted, shown with a bigger font, and a line of attribution at a central location in the printed article layout. Pull-quotes are a traditional part of news articles. The pull-quote phenom has a multi-step process in which a newsworthy quote is determined by filtering, labelling, and ranking all the discussion utterances from the transcript. Preprocessing involves removing utterances from subjects of non-interest (such as those spoken by committee heads, chairs and staff). Following preprocessing, we correlate each utterance with a dialogue label based on the Switchboard Corpus [22, 31]. The Switchboard

Table 1. List of select phenoms and percentage triggered. A set of 50 random bill discussions was used for this purpose.

Phenom	**Description**	%
bill_name	Metadata: bill name query	100
subjects	Metadata: bill subject query	100
presenter	Detects the person presenting the bill	94
motion	Metadata: bill motion query	98
vote_result	Metadata: voting result query	98
attendance	Metadata: attendance query	22
intro_length	Detects intro part longer than certain threshold	40
chair	Detects the chair of the meeting	50
expert_testimony	Detects an expert testimony longer than a certain threshold	50
testimony_alignment	Adds alignment information from the metadata to the testimony info	34
bill_mentions	Detects mentions of any other bill names	56
back_and_forth	Detects a back-and-forth discussion	10
alignment	Metadata: general alignment of the people present	10
first_bill	Metadata: is it the first bill by the legislator	2
sigle_party_split	Metadata: split voting of the party	4
a_to_b_questions	Detects a person asking questions in a back-and-forth	6
vote_against_party	Metadata: legislator voting against their party	4
pull_quote	Detects a sentence ranked high enough as a pull quote	40

corpus is a collection of about 2,400 two-sided telephone dialogues among 543 speakers, with each utterance annotated using the Discourse Annotation and Markup System of Labeling (DAMSL) tag set.

We only consider utterances that have the following labels: "Statement-opinion", "Rhetorical-Questions", "Hedge", "Action-directive" and "Apology and Appreciation", as they would add the most value to the quote. To label each utterance, we used a neural model that processes both lexical and acoustic features for classification. The model uses a bi-directional LSTM classifier pre-trained on the Switchboard Corpus for lexical clues. The LSTM allows the network to retain the context of the hearing for future tagging. Additionally, the acoustic model uses a CNN to process speech signals which can be equally useful in determining a tag.

The chosen utterances are first ranked based on content using the LexRank algorithm [9], a graph-based centrality scoring of sentences. This is useful as the algorithm can compute the relative importance for each utterance within a document. The algorithm can identify the most central sentences in a cluster that give the greatest amount of information related to the main theme of the document. They are then further ranked by length, sentence root, starting word, bill and geographic mentions, and readability. Sentences are also removed from consideration based on unresolved references and bad structure. After acquiring the highest-ranked pull-quote, the quote is returned along with information about the speaker to be placed in the article.

Back and forth argument detection The "back and forth" phenom captures news-worthy exchanges and debates that involve two individuals speaking a few lines in rapid interleaving succession. If multiple exchanges appear in a hearing, these exchanges are scored according to average utterance and word length according to the following formula:

$$score = (\alpha * \mu\_all + (1 - \alpha) * \frac{num\_utterances}{hearing\_length}) * \frac{long\_utterances}{num\_utterances} \quad (1)$$

where μ_all is the average utterance length of both speakers and long_utterances is the number of utterances larger than 20 words. The α value (0-1) determines the weights of the average utterance length (for both speakers) and the percentage of the hearing devoted to this exchange. The highest ranking exchange is returned along with some extended personal information about the speakers, which can pulled from the database storage if available.

3.3 Template-based Text Generation and Planning

bill_mentions	Bill $bill_mentioned1 was also brought up during the bill discussion.|The speakers also brought up $bill_mentioned1 throughout the discussion.\|$bill_mentioned1 was referenced in the discussion.
bill_mentions	Bill $bill_mentioned2 was mentioned too.|Another mentioned bill was $bill_mentioned2.|Other bills discussed were $bill_mentioned1 and $billmentioned2.
a_to_b	After the bill was presented a discussion followed, particularly between __$AtoB_person1 and __$AtoB_person2.|During the discussion, a particluar back-and-forth conversation occurred between __$AtoB_person1 and __$AtoB_person2.|The audience was actively debating on the measure afterwards, for example, __$AtoB_person1 and __$AtoB_person2 ended up having a dialogue on this topic.|__$AtoB_person1 and __$AtoB_person2 exchanged a few lines during the discussion.
a_to_b_questions	__$AtoB_questioning_side was mainly asking questions throughout the conversation.\|__$AtoB_questioning_side asked some questions.|__$AtoB_questioning_side wanted some clarifications and information, asking questions.\|__$AtoB_questioning_side asked many good questions.|__$AtoB_questioning_side was clearly engaged in the presentation, asking many questions during the event.\|There were questions from __$AtoB_questioning_side.

Fig. 3. Templates used for text generation of the article.

Phenoms also generate text for the news articles. This text is generated using templates from our library. Each template consists of one or more English sentences with some variable placeholders such as entity names, vote results or bill subjects (see Fig. 3). They are tagged with phenom names, so each phenom can use one or more templates assigned to it. If a template's variables can be entirely resolved using known facts, then the template is used in the article assembly, otherwise, it is discarded. The system tries to find the longest satisfiable template for each phenom. Thus having more information from one phenom can unlock longer templates. All the sentences are then assembled into a final text according to the order produced by the planner.

4 Research Study

Many existing summarization systems are evaluated using metrics [17, 28, 24, 15] which produce a distance in comparison to a known-good summary. In the case of our study, where there is no gold standard to compare with, the above-mentioned systems are not applicable for evaluation. Therefore, we utilize human evaluations and crowd-source the study.

4.1 Design and Protocol

A questionnaire[3] was designed to determine the effectiveness of the phenom system and assess the quality of the articles produced by such a system.

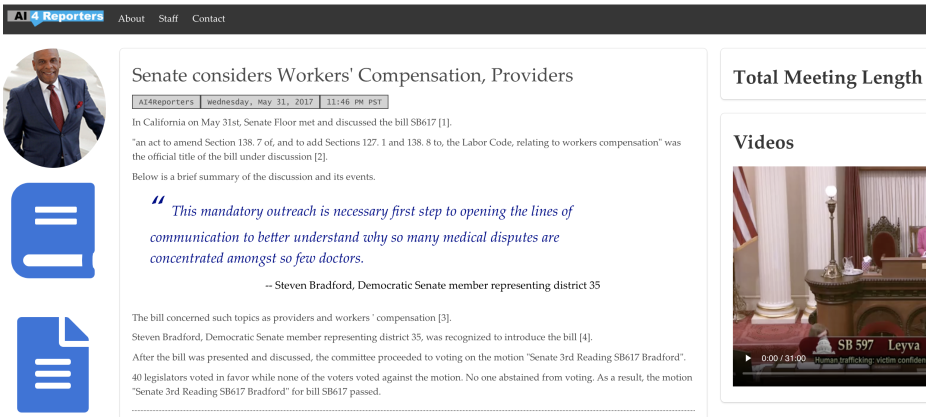

Fig. 4. An example rendered page containing the full article and the video recordings of the hearing that the respondents are presented with during the user study

The study protocol is as follows: the respondent is first presented with a short questionnaire to collect some demographics data, and is requested to watch a video recording of a bill discussion. When they are done watching, another short questionnaire is loaded with questions regarding the newsworthiness of the events happening in the recording. In the next step the respondents are asked to read a summary for the same hearing and answer several questions about the quality of the summary: the article writing style, readability, etc. They are allowed to get back to both the video and the article for more details or comparison.

We use Amazon Mechanical Turk and a questionnaire hosted on Google Forms. The Mechanical Turk service allows researchers to gather many responses quickly, but some will be invalid. For example, the responses that contain incoherent or fake answers. The user study is conducted on two different hearings, chosen as randomly as possible among the ones that in the end provide enough content in the summary to evaluate. The questionnaire given to the Turkers

[3] available at https://iatpp.calpoly.edu/slsp2021

consists of Likert scale and free form text questions, presented to the respondents immediately after they watch the video recording of the hearing and read through the generated article (see Fig. 4).

We collect a data set for each hearing chosen - the first set (Set 1) consists originally of 34 responses, with 22 (64,7%) deemed valid for investigation, while the second test run for another hearing (Set 2) provides 32 answer sets with 20 (62,2%) of them being valid. The invalid answers are selected manually judging mostly by the open text answers - in most of the cases it is incoherent sentences or sentences copied from the article. No specific demographics group was preferred for the user study - only the data concerning the general interest towards legislation news was collected from the respondents. In both answer sets the typical user is a native English speaker consuming state level news data at least weekly.

4.2 Results and Discussion

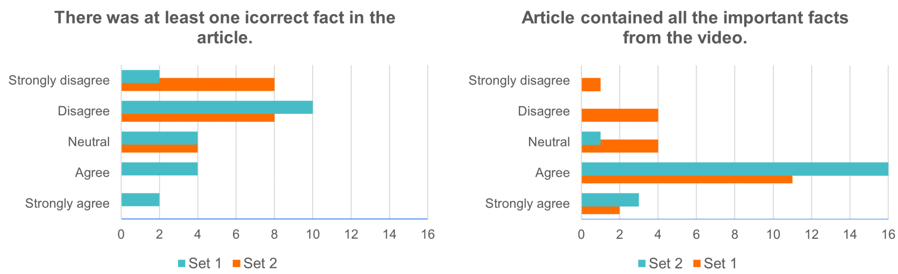

Fig. 5. Statistics collected considering the factual accuracy of the summary produced(N=22 for Set 1, N=20 for Set 2)

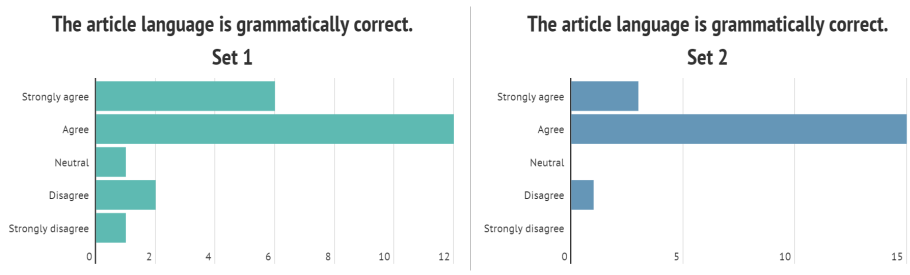

Fig. 6. Statistics collected considering the grammaticality of the summary produced (N=22 for Set 1, N=20 for Set 2)

Overall, the articles produced appear to be accurate and correct factually, even though there was a marked difference between the first article and the second. According to the survey, 13 of the 22 respondents in Set 1, and 19 of

20 respondents in Set 2 agree or strongly agree with the statement "*Article contained all the important facts from the video*". This is contrasted with 6 of 22 respondents in Set 1 and zero of 20 respondents in Set 2, who disagree or strongly disagreeing with the same statement. This is a ratio of 13:6 (Set 1) and 19:0 (Set 2). The same notion was validated by the additional statement "*There was at least one incorrect fact in the article*", with which 12/22 in Set 1 and 16/20 in Set 2 disagree or strongly disagree. Only 6/22 and 0/22 agree with the same statement for a favorable ratio of 12:6 in Set 1 and 16:0 in Set 2 (see Fig. 5). These findings reveal that the system manages to pick the facts correctly from the transcript, generates articles without omitting expected important events, and does not distort facts.

Positive results are also observed in the grammaticality of the text (see Fig. 6) - an overwhelming majority of the respondents confirm that the summary is a coherent and grammatically correct text. The simplified agree:disagree ratio is 19:3 (Set 1) and 19:0 (Set 2). Only some minor issues were mentioned by some of the respondents in the open text answer fields, such as "*Missing comma before the word "too" at the end of a sentence*". Judging by the responses to the questions considering the stylistics and readability of the article, most of the Turkers found the article was easy to read, and that it provide the information in an understandable way. More than half of the respondents thought the article was produced by a human and not by an algorithm.

5 Conclusion and Future Work

We propose a solution to help address the declining media coverage in crucial beats such as state legislatures. We design an automatic article generation system using state legislative hearing transcripts. We evaluate the system with a user study. Results indicate our system achieves high performance in accuracy and readability, but doesn't always capture all that is important in the recorded hearings. A sizable minority (9 out of 22) in Set 1 of our study indicated they are either neutral or do not agree that all important facts were captured in the article.

Future work will expand on both our phenom and template libraries leading to more insightful and diverse article output. All content is currently generated from phenoms and thus there's no possibility of any new or surprising types of content. Although most subjects thought all important aspects of the discussion was covered, not all did so. We hope to address this by adding more phenoms to the library and to further design new phenoms based on patterns we find in the transcripts. Another important issue to address is the narrative flow of the article. As some subjects indicated, the article reads more like a collection of sentences and could use internal references and stylistic consistency.

Acknowledgments The authors thank the John S. and James L. Knight Foundation. The collaboration between Graz University of Technology and Cal Poly was made possible thanks to funding by the Austrian Marshall Plan Foundation.

We also thank the Institute for Advanced Technology and Public Policy, and Ms. Christine Robertson for her valuable insights.

References

1. Besley, J.C., Roberts, M.C.: Cuts in newspaper staffs change meeting coverage. Newspaper Research Journal **31**(3), 22–35 (2010)
2. Blakeslee, S., Dekhtyar, A., Khosmood, F., Kurfess, F., Kuboi, T., Poschman, H., Prinzivalli, G., Robertson, C., Durst, S.: Digital democracy project: Making government more transparent one video at a time. Digital Humanities 2015 (2015)
3. Budhwar, A., Kuboi, T., Dekhtyar, A., Khosmood, F.: Predicting the vote using legislative speech. In: Proceedings of the 19th annual international conference on digital government research: governance in the data age. pp. 1–10 (2018)
4. Bui, T., Frampton, M., Dowding, J., Peters, S.: Extracting decisions from multi-party dialogue using directed graphical models and semantic similarity. In: Proceedings of the SIGDIAL 2009 Conference. pp. 235–243 (2009)
5. Christensen, H., Gotoh, Y., Kolluru, B., Renals, S.: Are extractive text summarisation techniques portable to broadcast news? In: 2003 IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat. No. 03EX721). pp. 489–494. IEEE (2003)
6. Cohen, M.: A new effort boosts a diminished illinois statehouse press. Columbia Journalism Review (2019)
7. Doddington, G.R., Mitchell, A., Przybocki, M.A., Ramshaw, L.A., Strassel, S.M., Weischedel, R.M.: The automatic content extraction (ace) program-tasks, data, and evaluation. In: Lrec. vol. 2, pp. 837–840. Lisbon (2004)
8. Enda, J., Matsa, K.E., Boyles, J.L.: America's Shifting Statehouse Press: Can New Players Compensate for Lost Legacy Reporters? Pew Research Center (2014)
9. Erkan, G., Radev, D.R.: Lexrank: Graph-based lexical centrality as salience in text summarization. Journal of artificial intelligence research **22**, 457–479 (2004)
10. Feng, X., Feng, X., Qin, B.: A survey on dialogue summarization: Recent advances and new frontiers. arXiv preprint arXiv:2107.03175 (2021)
11. Grieco, E.: Us newsroom employment has dropped by a quarter since 2008, with greatest decline at newspapers. Pew Research Center **9** (2019)
12. Kauffman, D., Khosmood, F., Kuboi, T., Dekhtyar, A.: Learning alignments from legislative discourse. In: Proceedings of the 19th annual international conference on digital government research: Governance in the data age. pp. 1–2 (2018)
13. Leppänen, L., Munezero, M., Granroth-Wilding, M., Toivonen, H.: Data-driven news generation for automated journalism. In: Proceedings of the 10th International Conference on Natural Language Generation. pp. 188–197 (2017)
14. Li, Q., Peng, H., Li, J., Ning, Y., Wang, L., Yu, P.S., Wang, Z.: Reinforcement learning-based dialogue guided event extraction to exploit argument relations. arXiv preprint arXiv:2106.12384 (2021)
15. Lin, C.Y.: Rouge: A package for automatic evaluation of summaries. In: Text summarization branches out. pp. 74–81 (2004)
16. Liu, F., Liu, Y.: From extractive to abstractive meeting summaries: Can it be done by sentence compression? In: Proceedings of the ACL-IJCNLP 2009 Conference Short Papers. pp. 261–264 (2009)
17. Mani, I., Maybury, M.T.: Automatic summarization. J. Benjamins Publishing Company New York, USA (2001)

18. Marum, A.: Oregon's dwindling statehouse reporters are 'treading water'. Columbia Journalism Review (2018)
19. Murray, G., Carenini, G., Ng, R.: Generating and validating abstracts of meeting conversations: a user study. In: Proceedings of the 6th International Natural Language Generation Conference (2010)
20. Nesterenko, L.: Building a system for stock news generation in russian. In: Proceedings of the 2nd International Workshop on Natural Language Generation and the Semantic Web (WebNLG 2016). pp. 37–40 (2016)
21. Nguyen, T.M., Nguyen, T.H.: One for all: Neural joint modeling of entities and events. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 33, pp. 6851–6858 (2019)
22. Ortega, D., Vu, N.T.: Lexico-acoustic neural-based models for dialog act classification. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 6194–6198. IEEE (2018)
23. Oya, T., Mehdad, Y., Carenini, G., Ng, R.: A template-based abstractive meeting summarization: Leveraging summary and source text relationships. In: Proceedings of the 8th International Natural Language Generation Conference (INLG). pp. 45–53 (2014)
24. Papineni, K., Roukos, S., Ward, T., Zhu, W.J.: Bleu: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting of the Association for Computational Linguistics. pp. 311–318 (2002)
25. Riedhammer, K., Favre, B., Hakkani-Tür, D.: Long story short–global unsupervised models for keyphrase based meeting summarization. Speech Communication **52**(10), 801–815 (2010)
26. Ruprechter, T., Khosmood, F., Gütl, C.: Deconstructing human-assisted video transcription and annotation for legislative proceedings. Digital Government: Research and Practice **1**(3), 1–24 (2020)
27. Ruprechter, T., Khosmood, F., Kuboi, T., Dekhtyar, A., Gütl, C.: Gaining efficiency in human assisted transcription and speech annotation in legislative proceedings. In: Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age. pp. 1–2 (2018)
28. Saggion, H., Radev, D., Teufel, S., Lam, W.: Meta-evaluation of summaries in a cross-lingual environment using content-based metrics. In: COLING 2002: The 19th International Conference on Computational Linguistics (2002)
29. Schonfeld, E.: Automated news comes to sports coverage via statsheet. techcrunch, 12 nov 2010 (2010)
30. Shaw, A.: As statehouse press corps dwindles, other reliable news sources needed. Better Government Association (2017)
31. Stolcke, A., et al.: Dialogue act modeling for automatic tagging and recognition of conversational speech. Computational Linguistics **26**(3), 339–373 (2000)
32. Wang, L., Cardie, C.: Domain-independent abstract generation for focused meeting summarization. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 1395–1405 (2013)
33. Weiss, S.: 140 characters of news. State Legislatures **41** (2015)
34. Williams, A.T.: Measuring the journalism crisis: Developing new approaches that help the public connect to the issue. International Journal of Communication (19328036) **11** (2017)
35. Xie, S., Liu, Y., Lin, H.: Evaluating the effectiveness of features and sampling in extractive meeting summarization. In: 2008 IEEE Spoken Language Technology Workshop. pp. 157–160. IEEE (2008)