

Understanding User Perception of Automated News Generation System

Changhoon Oh¹ Jinhan Choi² Sungwoo Lee² SoHyun Park² Daeryong Kim²
Jungwoo Song² Dongwhan Kim³ Joonhwan Lee² Bongwon Suh²

¹Carnegie Mellon University, Pittsburgh, PA, USA

²Seoul National University, Seoul, Korea ³Yonsei University, Seoul, Korea

ochangho@andrew.cmu.edu, {jinhanchoi, lsw0504, sohyun, daeryong, jwsong0617}@snu.ac.kr,
dongwhan@yonsei.ac.kr, {joonhwan, bongwon}@snu.ac.kr

ABSTRACT

Automated journalism refers to the generation of news articles using computer programs. Although it is widely used in practice, its user experience and interface design remain largely unexplored. To understand the user perception of an automated news system, we designed *NewsRobot*, a research prototype that automatically generated news on major events of the PyeongChang 2018 Winter Olympic Games in real-time. It produces six types of news by combining two kinds of *content* (*general/individualized*) and three *styles* (*text*, *text+image*, *text+image+sound*). A total of 30 users participated in using *NewsRobot*, completing surveys and interviews on their experience. Our findings are as follows: (1) Users preferred *individualized news* yet considered it less credible, (2) more presentation elements were appreciated but only if their quality was assured, and (3) *NewsRobot* was considered factual and accurate yet shallow in depth. Based on our findings, we discuss implications for designing automated journalism user interfaces.

Author Keywords

Automated journalism, robot journalism, multimedia modality, automated news generation system.

CCS Concepts

•Human-centered computing → Human computer interaction (HCI); *User studies*;

INTRODUCTION

Automated journalism (also known as *robot journalism*) refers to the generation of news articles using algorithmically designed computer programs [20, 26, 39]. Such programs usually collect and process data and integrate it into a predefined article structure, producing news articles [39]. Unlike human journalists, the programs can create articles on a large scale quickly, cost-effectively, and even accurately. Many

companies, such as Narrative Science [48] and Automated Insights [31], are already producing and providing news articles based on this technology, and traditional media companies, like The Los Angeles Times and Thomson Reuters, are also distributing automatically generated news articles [22]. The technology is also used in a variety of areas, such as sports event highlights, weather forecasts, and disaster and election reports [21, 39, 55].

Surprisingly, while the importance of automated journalism is often noted [34, 35], the topic is rarely studied in the field of human–computer interaction (HCI). Although not explicitly addressing automated journalism, recent discussions on the transparency and fairness of algorithms in automated systems show that it is important to not only focus on the technical design of the algorithm but also closely observe the user experience of these systems [56, 61]. This calls for the active involvement of designers and HCI researchers in the design of automated news generation systems.

In this study, we aim to explore the user perception of automated journalism and discuss design implications for automated news generation systems based on the findings. We designed a research prototype, *NewsRobot*, which automatically generated a series of summary news articles of the PyeongChang 2018 Winter Olympic Games in real-time. It collected data on the results of major events and competitors' information from the official website of the Olympic Games [13], processed the data, and input it into a predefined article structure. We designed the system to generate news with two different types of *content* (*general/individualized*) in three different *styles* (*text/text+image/text+image+sound*). Combining the two factors, it produced six different types of news articles for every event.

We conducted a user study of *NewsRobot* with both quantitative and qualitative approaches. We asked 30 recruited participants to watch Olympic Games races on a TV screen and then showed them one of each of the six types of news articles per race. They then filled out questionnaires on each news article and took part in semi-structured interviews. The results of the study can be summarized in the following three points:

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author(s).
CHI '20, April 25–30, 2020, Honolulu, HI, USA.
© 2020 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-6708-0/20/04.
<http://dx.doi.org/10.1145/3313831.3376811>

- **Content:** While users preferred *individualized news* to *general news*, they considered *individualized news* less credible than *general news*.
- **Style:** As more news presentation elements were added, users' preference increased. People liked *text+image+sound news* the most, followed by *text+image news* and then *text news*. On the other hand, in terms of quality, users rated *text+image news* as more clear and concise than *text+image+sound news*.
- **Overall assessment:** While people were satisfied with NewsRobot's accuracy, objectivity, selection function, and various presentation elements, they found the articles were sometimes dull, repetitive, and out of context.

Based on these findings, we discuss the design implications for user interfaces for algorithm-based automated news generation systems. The main contributions of this work to the HCI community are as follows:

- NewsRobot, an automated news generation system, can not only automatically generate news in real time, but is also capable of enriching user experience by providing users with individualized news according to their choices while expressing various multimedia modalities through the use of text, images, and sound.
- Our study measured how a user's perception of news varies depending on the content and style according to a set of comprehensive criteria, which can serve as a basis for expanding the existing body of knowledge on automated journalism. We think this analysis can spark research on a variety of topics, including interface design, news structure design, and the collaboration of various stakeholders involved.

BACKGROUND AND MOTIVATION

In line with the attempt to apply technology to journalism [28, 40, 54], *automated journalism* (also known as *robot journalism*), which is the generation of news articles using computer software instead of human reporters, has been introduced [20, 26, 39]. Since this technology can create a large number of news articles quickly and accurately, it has attracted the attention of both the media industry [22, 31, 48] and scholars [8, 36, 52]. Many researchers have searched for technical ways of generating news articles that are indistinguishable from those of human reporters [4, 38]. In addition, these technical explorations have brought about extended discussion of the sociocultural impacts [2], such as authorship [39], labor substitution [7, 12, 33, 55], and algorithm power [16, 17, 18].

In the field of HCI, interest in and research on automated journalism have been gradually increasing recently. Early studies were mainly focused on suggesting a news generation framework [15, 34] or designing authoring tools to facilitate news generation. For example, Kim and Lee proposed a five-step robot journalism framework for automated news content generation [34], which consisted of (1) data crawling, (2) event extraction, (3) key-event detection, (4) mood detection, and (5) news article generation. Based on this framework, they also created an algorithm-driven interactive news generation

system, demonstrating its capability of generating news stories that are significantly more interesting and enjoyable than traditional news articles [35]. Several studies have proposed different kinds of authoring tools to support news content generation [1, 37, 53].

Meanwhile, in consideration of the user's point of view, research has been conducted on how the user perceives and evaluates the news content produced by the automated journalism technology. While several studies found that computer-generated news was perceived as less credible and more artificial than human-generated news [27, 57], others found no difference in the perceived credibility between the two types of news [60]. Though these studies are meaningful in that they attempt to include a user's perspective in the study of machine-generated news, most of them are limited to comparing machine-written articles with articles written by human reporters. There has been relatively little research on how to design algorithm generated news itself through a more focused and multi-layered analysis of user perceptions and experiences on it.

The first consideration could be what news *content* to design and provide to the user. In automated journalism, software algorithms can use the same data to tell stories from different angles, customizing languages, topics, tones, and styles according to an individual reader's preferences. In this situation, deciding which content should be provided to the user among numerous various news contents could be important. In particular, whether individual preferences should be considered first or whether the overall content should be conveyed is an important question. Research has demonstrated that personalized messages can engage and persuade an audience more effectively than generic mass messages [44, 45]. On the other hand, the use of a personalized news system can have a direct negative effect on knowledge acquisition [3]. It also has the potential to allow users to collect and consume only biased information. The problem of transparency and fairness of algorithms in the selection process is also constantly being raised. These various studies raise the need to examine in detail how content affects users' perceptions of news.

Another consideration could be what *style* of news to design and present to the user. While most automated journalism research has focused on text news generation [38, 43], the resources used to produce news articles are no longer limited to text but also include images, audio/video clips, and graphic animations [30, 42, 59]. Adding multimedia resources to news articles can increase audiences' engagement in news articles [50]. It is necessary to recognize the possibility of using such multimedia and to introduce these factors into news generation. Generally, the use of multimedia can increase the expressiveness, usability, and enjoyment of computer interfaces as well as content itself [41]. On the other hand, simply increasing the multimedia modality level of an information system can increase its complexity and adversely affect its ability to convey information [32]. Introducing multimedia modalities requires more attention to user needs and comprehensive field studies to investigate the most appropriate solutions for each kind of application and user category [32].

In this study, we design and create an automated news system with various factors, evaluating the potential of automated journalism, and evaluate it from various perspectives. We especially focus on the *content* and *style* of the algorithm generated news articles. We then create a basis for a discussion of a system design that promotes a desirable user experience.

NEWS ROBOT

In this section, we describe *NewsRobot*, the research prototype of this study. In the design of NewsRobot, we aimed to create a system that generates real-time news about an actual sports event.

Selecting Main Event and Data Source

As the first step, we selected the main event that NewsRobot would report about and its relevant data source. Since we aimed to create a variety of news contents in real-time on major sports events, we chose the PyeongChang 2018 Winter Olympic Games, held between February 9 and 25, 2018, in PyeongChang, South Korea. It featured 102 events over 15 disciplines in seven sports, and 2,914 athletes from 92 countries competed in the games.

During the PyeongChang Olympic Games, the International Olympic Committee updated the results of all matches on its official website in real-time [13]. The committee prepared separate pages for every event, providing not just the basic information of the event, such as location and date, but also all participants' intermediate records and ranks, differences from other participants, speeds, and finish records in real-time. The committee also created separate pages for each athlete participating in the Olympics and posted their personal information, such as nationality, birthdate, age, gender, event and rank, and even biographical information. Information on competitors' previous Olympic and world championships, injuries, family relationships, idols, mottos, and even nicknames was included in the biographical information. We collected both the real-time data of the results of each event and athlete information by building a crawling software with Python.

Since it was necessary to create articles in the same format for the controlled experiment, we limited main events to races where multiple competitors competed for the record so that NewsRobot could create articles for every competitor and compare records between them according to their intermediate records. Finally, we chose three events: (1) Speed Skating Men's 1,000 Meters (February 14, 2018), (2) Alpine Skiing Women's Downhill (February 21), and (3) Short Track Speed Skating Women's 1,000-Meter Final (February 22).

Designing News Article Structure

After the main events and data were defined, we designed a news article template that could cover all three races. We made the template following the *inverted pyramid structure* [23, 46, 47], the most common method for writing news stories. In the inverted pyramid, the widest part (at the top) represents the most important information, while the tapering lower portion illustrates that other material should follow in order of diminishing importance. We placed the headline in the first sentence of the template, followed by the lead sentence and then the body sentences. The lead sentence summarizes the most important information according to *5W1H* (*who, what, where, when, why, and how*) [29]. The body sentences are then composed of supporting details, background details, and general details (Figure 1).

Content and Style of NewsRobot News Articles

In addition to the template article structure, we considered two main factors, *content* and *style*.

Content

We designed NewsRobot to create two different types of news: *general news* and *individualized news*. The former is provided to all users equally, while the latter is customized to the users' choice and is provided differently for each user (Figure 1).

- **General news:** *General news*, the most common type, summarizes the overall game in a comprehensive way. It focuses on the competitors who reached the podium, especially the

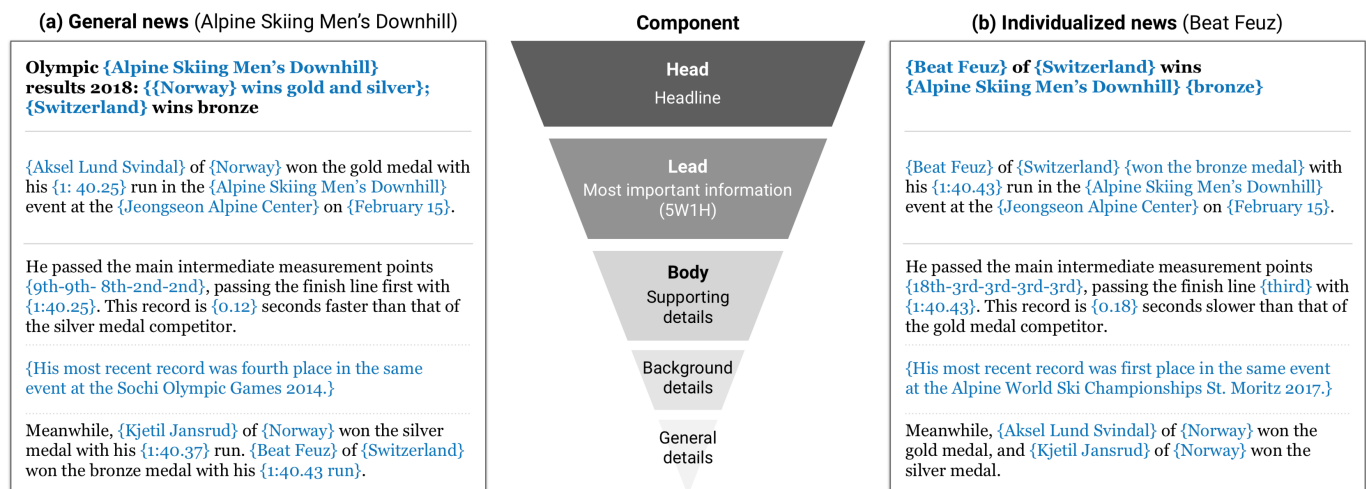


Figure 1. (a) General news vs. (b) Individualized news. We made the template following the inverted pyramid structure. The braces are interlocked with the database, so they are tailored to each news article. The original version is written in Korean, and this figure is the English-translated version.

gold medalist. It details the result of the gold medalist (gap between silver records, intermediate ranks variation) with additional information (from biographical information) and then briefly describes the results of the silver and the bronze medalists (Figure 1 (a)). Only one general news article is made per event.

- **Individualized news:** *Individualized news* is a recap of a particular athlete who participated in an event. Unlike *general news*, this details the record of the particular competitor (gap between gold records, intermediate ranks variation) with additional information (from biographical information) and then adds a brief summary sentence about the podium (Figure 1 (b)). In principle, the number of *individualized news* articles made is the same as the number of competitors who participated in the event. However, considering the selection view of NewsRobot’s user interface, we limited the number of competitors to nine per game. To maintain the high interest of users, the nine competitors included mainly high-ranked competitors and national competitors. In the case of the Short Track Speed Skating Women’s 1,000-Meter Final, as the total number of competitors was six, the number of articles was limited to six.

Style

We designed NewsRobot to create news articles in three *styles*: *text*, *text+image*, and *text+image+sound* (see Figure 2), according to the level of multimedia modality.

- **Text news:** *Text news* refers to basic news consisting of only text. This includes numerical data, such as competitor record and rank (Figure 2 (a)).
- **Text+image news:** This style combines text with graphical features by sentence, turning it into multiple slides. Graphical features include the background color of the medal according to the competitor’s rank, a pictogram of the sport, the competitor’s image (from a real-time Twitter search using the competitor’s name as the hashtag), the competitor’s intermediate ranks variation graph, and a picture of the podium with the national flags. All of the features are automatically integrated with each sentence of the article (Figure 2 (b)).

- **Text+image+sound news:** This style combines *text+image news* with *sound*. The voice API engine [14] automatically reads out the text of the article as an announcer, and the graphical features are displayed when the corresponding text is being read. Background music is included to boost the intensity (Figure 2 (c)).

By combining these two types of *content* and three types of *styles*, a total of six types of articles ((1) *general×text*, (2) *general×text+image*, (3) *general×text+image+sound*, (4) *individualized×text*, (5) *individualized×text+image*, (6) *individualized×text+image+sound*) are created for each event.

Generating News Articles

Based on the data, structure, and six types of news, we created the automatic news generation program with Python programming. The collected data were calculated, refined, and entered into the necessary parts of the news text. The news articles were also created with various multimedia elements, such as images, graphs, voice, and sound. Through the pilot test before the three events that we selected, we were able to elaborate on the algorithm and source codes of news generation. On each game day, we ran the news generation program immediately after the game finished, and finally, we were able to successfully generate all the news articles we had planned, with a total of 81 news articles (30 from Speed Skating Men’s 1,000 Meters ((1 *general* + 9 *individualized*) × 3 styles), 30 from Alpine Skiing Women’s Downhill ((1 *general* + 9 *individualized*) × 3 styles), and 21 from Short Track Speed Skating Women’s 1,000-Meter Final ((1 *general* + 6 *individualized*) × 3 styles)).

Designing NewsRobot User Interface

After successfully generating the news articles, we created the NewsRobot user interface. The interface was designed to operate on web browsers of users’ smartphones, and it was made with JavaScript and HTML5 programming. Participants were able to access NewsRobot’s webpage and watch various news articles on the Olympics. Although the user study was conducted several months after the Olympics, no additional modifications were made to the news articles for the experiment. We only included the news that was generated in real-time during the Olympics.

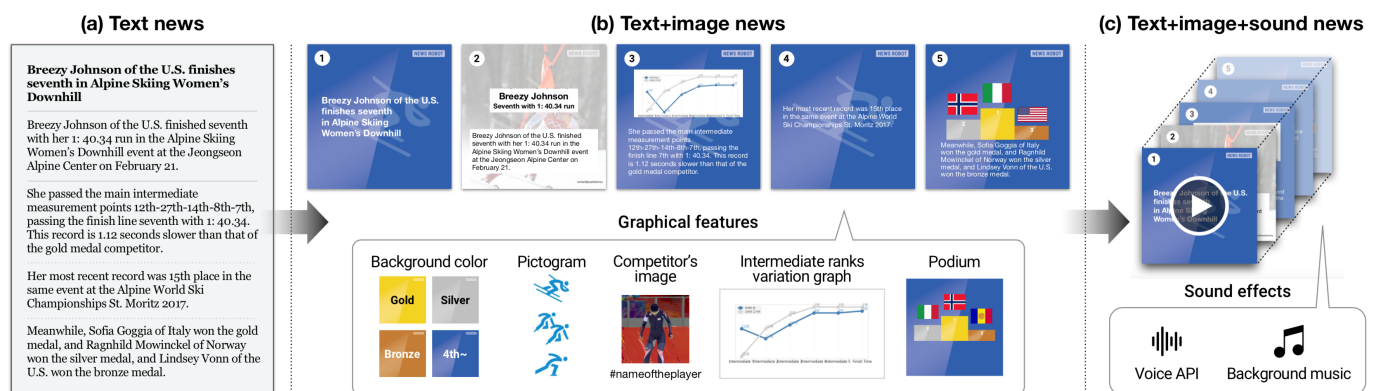


Figure 2. Three different styles of news presentation delivered by NewsRobot.

Specifically, users could select a game on the first screen of the user interface. On the next screen, the names of the athletes participating in each game were presented in the form of a tile with the national flags. NewsRobot then displayed the six types of news articles for that match in random order.

STUDY DESIGN

To understand how users evaluate the news articles of the system, we designed a user study. The design protocol was reviewed and approved by the Institutional Review Board of Seoul National University.

Participants

In the recruitment of participants, we aimed to balance the age and gender of general users who have no trouble accessing news articles using smartphones. We posted a recruiting document on our institution's online community website and recruited a total of 30 participants (15 males and 15 females). Their average age was 30.4, and the SD value was 7.6 ($M: Mean = 29.3, SD = 6.5; F: Mean = 31.5, SD = 8.6$). Our interview records indicate that none of the participants exhibited any special interest in the Winter Olympics. Before the experiment started, we provided the participants with a detailed explanation of the purpose and procedure of the experiment as well as NewsRobot. They were told that the news was produced in real-time by a computer program during the Olympics. We also ensured that NewsRobot worked properly on the users' smartphones. Users were allowed to manipulate the system for a while so that they could get used to using it. On average, each experiment lasted about 70 minutes. Each participant received a gift voucher worth \$10.

Procedures

We designed a user study that consisted of watching racing events and reading/watching news articles on NewsRobot, followed by completing surveys and semi-structured interviews.

Task

To enhance the users' immersion in the game situation, we prepared the pre-recorded and edited events videos and played them on a TV screen in the experiment room. The videos of the three races were shown in random order. Before each race was played, the participants were asked to select one of the participating athletes on the NewsRobot interface on their smartphones according to their interests. At the end of each video, the participants were given six types of articles in random order. As a result, each participant read a total of 18 kinds of news articles (3×6 within-subjects design).

Survey

To understand the user's perception of the news generated by the prototype more systematically and in a multi-faceted way, we designed 18 questionnaires with reference to Sundar's study [49]. Sundar's index offers a comprehensive set of items and is widely used to study user perceptions of online news. Sundar identified multiple criteria used by the public in evaluating news articles, combining them into four major factors: *credibility*, *liking*, *quality*, and *representativeness*. *Credibility* is composed of *biased*, *fair*, and *objective*. *Liking* consists of *boring*, *enjoyable*, *lively*, *interesting*, and *pleasing*. *Quality*

has five sub-items: *clear*, *coherent*, *comprehensive*, *concise*, and *well written*. Representativeness is composed of accurate, believable, disturbing, informative, and sensationalistic. In our study, users evaluated each news article on the survey with a 7-point Likert scale, ranging from highly disagree (=1) to highly agree (=7).

Interview

At the end of the experiment, participants were asked to take part in semi-structured interviews. In the interviews, we asked participants for their overall impressions of NewsRobot, their thoughts on two separate pieces of content, and three different styles. We requested that they explain the advantages and disadvantages of NewsRobot as well as provide suggestions for the improvement of each news item delivery.

Analysis Methods

After the experiment, we conducted a quantitative analysis on the survey data and a qualitative analysis on the interview data.

Quantitative Analysis

In the quantitative analysis, we aimed to determine whether the *content* and *style* of NewsRobot news articles had any significant effects on users' evaluations. Accordingly, each of the 18 questionnaires was set as a dependent variable, and *content* and *style* which produced six experimental conditions, and its interaction (*content* $\times$ *style*) were set as independent variables. As 30 participants assessed the questionnaire items (DV) while consuming six different news articles for each of the three events, a total of 540 data points were collected for each item ($30 \text{ participants} \times 6 \text{ conditions [news articles]} \times 3 \text{ trials [events]}$). Since the experiment design was within-subject, where the participants performed all the conditions and they conducted multiple trials in each condition, we analyzed the survey data using two-way repeated-measures ANOVA. Mauchly's test indicated that the assumption of sphericity for all 18 items had been violated; therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphere. We also conducted post-hoc test using the Bonferroni correction for pairwise comparisons. These corrections made our analysis more conservative and less likely to be giving out the wrong information.

Qualitative Analysis

The qualitative data from interviews were transcribed and analyzed using thematic analysis [6]. To do so, we used Re-framer [62], a qualitative research software tool. We segmented the transcripts into sentences and finally obtained 471 observations. While reviewing the data, a total of 250 keyword tags were created. By combining the relevant tags iteratively, we conducted a theme-building process, yielding three main categories.

RESULTS 1: QUANTITATIVE ANALYSIS

In this section, we report the results of the quantitative analysis of the survey data. The results will be explained sequentially with four factors—*credibility*, *liking*, *quality*, and *representativeness*—, referring to Sundar's study. The analysis of each factor consists of an analysis of the items that make up it.

Credibility

First, a repeated measures ANOVA with a Greenhouse-Geisser correction showed that *content* had a significant main effect on the three items that make up *credibility*: *biased*, *fair*, and *objective* (Table 1). In the case of *biased*, *individualized news* received higher scores than *general news* (Figure 3). In contrast, in the case of *fair* and *objective*, *general news* scores were higher than *individualized news* scores. In other words, it can be seen that users felt that *general news* was more credible than *individualized news*. Meanwhile, there was no significant main effect of *style* on the *credibility* of news articles.

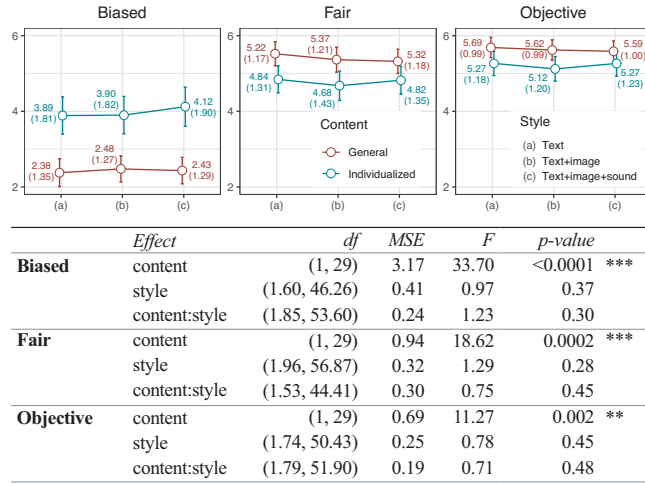


Figure 3 and Table 1. Users' evaluations of *credibility* of NewsRobot news articles. The interaction plots show the mean and SD of each condition for each item of *credibility* (the error bar at each point represents mean $\pm 2.58 \times SE$ at the 0.01 significance level). Color represents *content*, red for *general* and green for *individualized* news. The x-axis represents *style*: (a) for *text*, (b) for *text+image*, and (c) for *text+image+sound*. The table summarizes the results of the two-way RM ANOVA, and the reported p-value(s) associated with the F statistic(s), and degrees of freedom were adjusted via the Greenhouse-Geisser correction (MSE: mean squared error). Statistically significant results are reported as follows: $p < 0.001$ (***), $p < 0.01$ (**), $p < 0.05$ (*).

Liking

We identified that *style* had a significant effect on all five items of *liking*: *boring*, *enjoyable*, *lively*, *interesting*, and *pleasing* (Table 2). The pairwise comparison from the post hoc tests using the Bonferroni correction revealed that *boring* scored the highest in *text*, followed by *text+image* and *text+image+sound* (Figure 4), showing significant differences between *text* and *text+image* ($p < 0.001$ **) and between *text* and *text+image+sound* ($p < 0.001$ ***). In the case of *enjoyable*, the mean values were in the order of *text+image+sound* > *text+image* > *text*, showing significant differences between *text+image+sound* and *text+image* ($p = 0.035$ *) and between *text+image* and *text* ($p < 0.001$ ***). Likewise, regarding *lively*, the mean scores were in the order of *text+image+sound* > *text+image* > *text* (*text+image+sound* - *text+image*: $p < 0.001$ ***, *text+image* - *text*: $p < 0.001$ ***). In the case of *interesting*, although the difference between *text+image+sound* and *text+image* was not significant ($p = 0.62$), the scores showed the same pattern: *text+image+sound* > *text+image* > *text* (*text+image* - *text*: $p < 0.001$ ***). Lastly, *pleasing* showed

the same pattern (*text+image+sound* - *text+image*: $p = 0.89$, *text+image* - *text*: $p < 0.001$ ***).

We also identified that users liked *individualized news* more than *general news*. In all the items of *liking* except for *lively*, *content* showed a significant main effect: *boring*, *enjoyable*, *interesting*, and *pleasing* (Table 2). Meanwhile, we found no significant interaction effect.

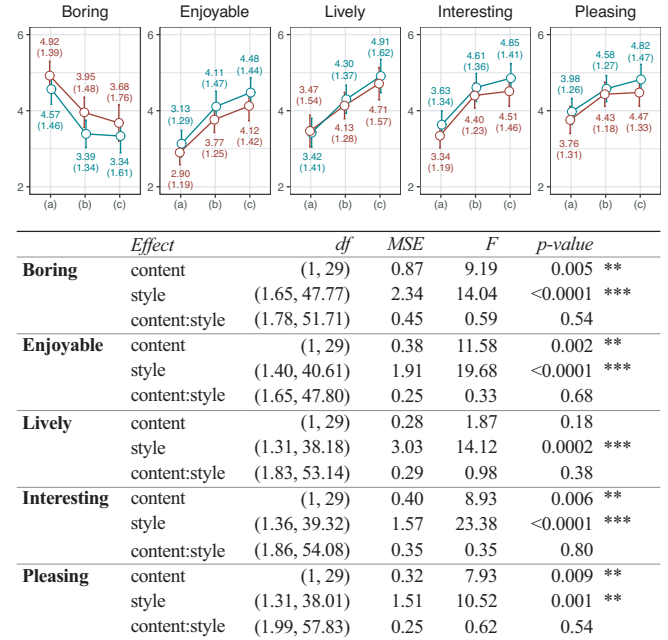


Figure 4 and Table 2. Users' evaluations of *liking* of NewsRobot news articles. The details are the same as those of Figure 3 and Table 1.

Quality

In terms of *quality*, first, we found that *style* had a significant main effect on *clear* and *concise* (Table 3). However, unlike the order of *text+image+sound* > *text+image* > *text* in *credibility* and *liking*, for both items, *text+image* news showed the highest value, followed by *text+image+sound* and *text* (Figure 5). The pairwise comparison from the post-hoc analysis revealed that, in the case of *concise*, there were significant differences in the mean scores between *text+image* and *text+image+sound* ($p = 0.015$ *) as well as between *text* and *text+image* ($p < 0.001$ ***). In the case of *clear*, the mean score difference between *text* and *text+image* was significant ($p = 0.00015$ ***). From this result, we identified that simply increasing the multimedia modality level does not guarantee the articles' perceived quality, especially in relation to clearness and conciseness.

Second, we examined whether *content* had a significant main effect on items of *quality*, finding that it had a significant effect on *comprehensive* (Table 3). *General news* received higher scores than *individualized news* in *comprehensive*. Meanwhile, we found no significant main effect on *coherent* or *well written*.

Representativeness

Lastly, for the analysis of *representativeness*, we focused on identifying noticeable patterns rather than significant relation-

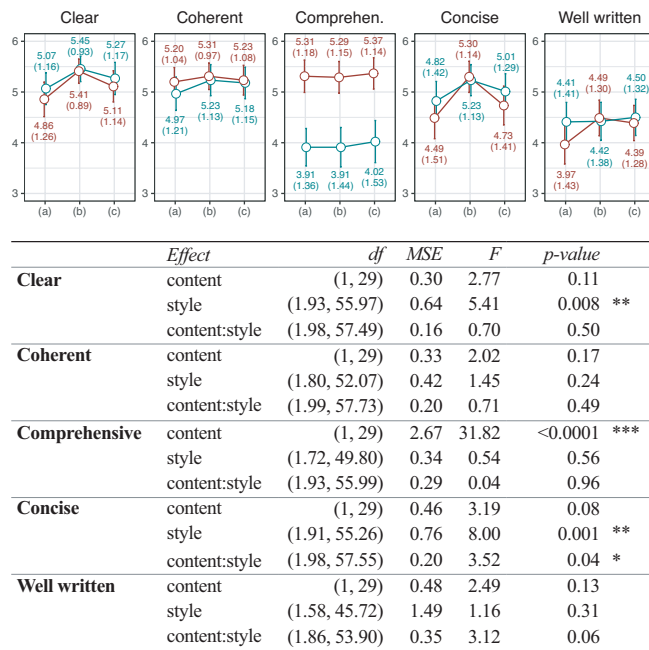


Figure 5 and Table 3. Users' evaluations of *quality* of NewsRobot news articles. The details are the same as those in the caption of Figure 3 and Table 1.

ships between the variables. First, overall, users evaluated NewsRobot's news as relatively *accurate* and *believable* (Figure 6). Unlike any other items, all kinds of news articles received scores higher than 5 points in both items. Conversely, participants evaluated NewsRobot news as relatively less *disturbing* and *sensationalistic* (1–2 points).

Finally, *style* showed a significant main effect on *informative* (Table 4). The pairwise comparison analysis from the post-hoc test revealed that participants rated both *text+image* and *text+image+sound* news as more *informative* than *text* news (*text+image* - *text*: $p=0.023^*$, and *text+image+sound* - *text*: $p=0.001^{**}$).

RESULTS 2: QUALITATIVE ANALYSIS

In addition to the findings from the survey, we also report users' in-depth thoughts about the news articles and NewsRobot from the interview session.

Users Evaluate NewsRobot Features Highly

First, we were able to identify that several NewsRobot features provided a good experience for users, which illustrates the potential of automated journalism.

Individualized news

Participants highly appreciated that NewsRobot could generate *individualized news* as well as *general news*. They told us that *individualized news* was more satisfying, since it contained customized information tailored to their personal interests and delivered it first in the articles. For example, P30 said, "I really liked the fact that the article started with the result of the competitor I selected." Some participants even told us that *individualized news* made them feel their interests were

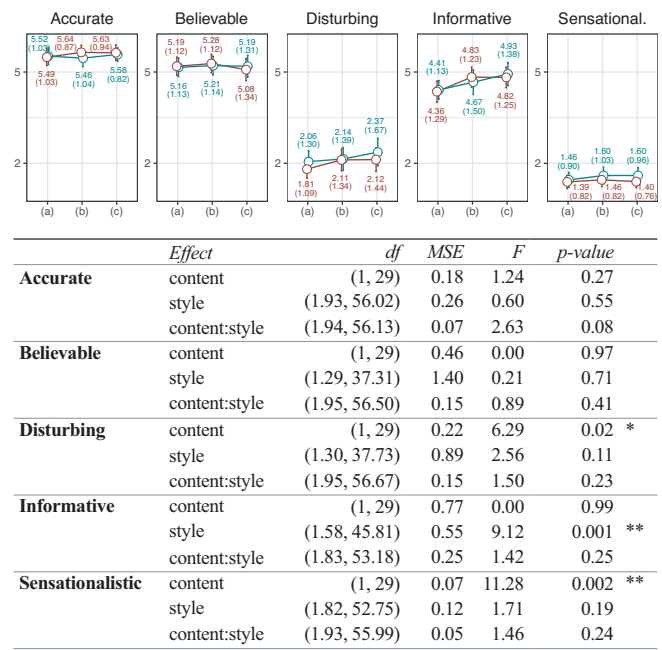


Figure 6 and Table 4. Users' evaluations of *representativeness* of NewsRobot news articles. The details are the same as those in the caption of Figure 3 and Table 1.

recognized and respected by the system. P22 said, "I felt that it [NewsRobot] was trying to learn more about me." In addition, some of the participants told us that simply being able to select a competitor according to their interests made them feel they were provided with a better experience. P07 said, "I thought it was a simple selection procedure, but it was good. That's why I was so attached to the article that I created."

Various Presentation Elements

In addition, participants particularly appreciated NewsRobot's various presentation elements. We identified users' high preference for *text+image* and *text+image+sound* news, which contained more elements, such as background colors, pictograms, competitor images, graphs, and even voice and background music. Regarding the background colors of *text+image* news, participants responded that they helped them intuitively understand how a particular competitor ranked. P27 said, "When I first saw the color, I thought it was useful. If the background was gold, I knew immediately that this competitor had won the gold medal." Competitors' real-time images from Twitter also helped users understand the competitors. P13 said, "When the athlete's image emerged in the background, I knew exactly who this athlete was." Moreover, users evaluated the use of the voice API in *text+image+sound* news as useful. Some users thought the voice was better than expected. P19 said, "I was curious about how it could read it so naturally." P03 said, "I was able to concentrate more on the news, as the voice automatically read the news." Some argued that it would be more useful in contexts where users can't see the news, such as driving situations. P06 said, "It would be great if I could use this while driving, when I cannot read news articles."

Among the various elements, participants were most satisfied with the intermediate ranks variation and comparison graphs including each competitor. Usually, it was difficult for users to know the relative performances of each competitor, because the competitors completed their races one after another in turn. However, these graphs showed how quickly each competitor passed through each point relative to other competitors, providing users with an appropriate visualization of information that was not provided on the TV screens. P17 said, *“The graph was good, in that I could see the game at a glance.”* Some users told us that they were able to understand the competitors’ game flow through the graphs. P13 said, *“This competitor started slowly at the beginning, but he gradually increased his speed in the second half and eventually placed first.”*

Instantaneous News Generation

Participants were surprised that NewsRobot generated news articles so quickly all at once. Although they read the articles in the experiment several months after the Olympics, they were amazed and impressed when they heard that all the news articles were generated in real-time during the event. For example, P09 said, *“Producing news articles quickly is very important. Of course, reporters can also write quickly, but they cannot make so many articles at the same time.”* P30 said, *“It’s very fast. I am surprised that it could create so many articles as soon as the game was over.”* P20 said, *“A reporter could never make images or video clips at this rate.”*

NewsRobot Is Unbiased but Predictable

In line with the survey results, the interviews also revealed that participants described the NewsRobot news articles as accurate and matter-of-fact. They were convinced that since the system computed and included specific figures based on the input data, the information should be accurate. They recognized that, for this reason, algorithm-based, automatically generated news articles provide objective and reliable information. P11 said, *“The numerical information gave me the impression that the news was objective.”* P25 also said, *“The figures were reliable, not biased. It seemed to be based on facts. I think the biggest advantage of algorithmic news is that it calculates and displays figures based on data.”* P16 also added, *“NewsRobot cannot lie; it only tells the truth,”* showing strong confidence in the accuracy of the news articles.

Participants also expressed that the lack of emotional elements and the exclusion of subjective judgments in the sentences increased the credibility of the news. For example, P01 said, *“Regardless of whether the competitor is loved or famous, it [NewsRobot] will describe him in the same way.”* People described algorithm-based news as objective, because it only provides predefined expressions, unlike the news, which often reflects the subjective opinions of reporters and anchors. P05 said, *“I could not find anything subjective in this news. You know, journalists often write using certain expressions to make their arguments more convincing.”* P09 said, *“Reporters’ subjective thoughts often influence the atmosphere of the news, but the program reports it as it is.”*

However, despite these advantages, users also pointed out the inherent shortcomings of NewsRobot: It often seems tedious, is unable to convey anything but data, and does not understand

the context. Most participants expressed that the news articles were dull and uninteresting, since all the articles followed the predesigned structure without any variation. P27 said, *“AI still seems to require a lot of refining. The biggest problem is that it is dull. I know it is accurate, but so far, it is only based on numbers.”* P07 commented, *“All the articles are almost identical in format. If I were to read more, I would feel bored.”*

Participants also complained that the algorithmic news articles could not deliver anything other than data and could not deliver in-depth news stories. As the system relied only on data, other information that was not included in the data was necessarily omitted. For example, in the Short Track Speed Skating Women’s 1,000-Meter Final, two competitors collided with each other and one of them was disqualified. When NewsRobot did not explain the collision in detail and only reported the disqualification, P21 said, *“It just focused on the record; I could not get any more information. Shim Suk-hee was disqualified, and why she was disqualified was the most important part, but it did not discuss that.”* P23 said, *“If someone asked me to explain a race, I would first talk about the collisions between competitors. Reporters usually discuss the collisions in detail and mention the shocking reactions people have received. But it [NewsRobot] did not.”*

Benefits and Drawbacks of Using Multimedia

The final point of the qualitative study results relates to the benefits and drawbacks of using multimedia. As outlined in the survey results and previous sections, participants were satisfied with NewsRobot when it produced news articles with multimedia, such as images and voice. They thought of *text+image+sound news* as more *vivid, pleasant, and informative* than either *text* or *text+image news*. However, in terms of clearness and conciseness, *text+image+sound news* provided a worse experience than *text+image news*, although the latter had a lower multimedia modality level than the former. We were able to confirm this in the in-depth interviews.

Overall, people described that they were satisfied with the video news but found it strange for two reasons. First, participants discussed the dissonance between their expectations and the actual quality. They claimed that as the multimedia modality levels of news articles increased in the order of text, image, and sound, their expectations also grew. However, the gap between their expectations and the actual quality of each style seemed to become larger as the multimedia modality levels increased. When reading text news, because of their low expectations for the text, they felt that the quality was acceptable. In contrast, when watching *text+image+sound news*, their high expectations for the news made them evaluate its quality as relatively low. P08 said, *“In the order of text, text+image, and text+image+sound, I felt they were very different from what I expected. I believe that in text+image+sound news, a reporter should appear on the screen and inform viewers of the result of the match with a realistic and vivid voice.”*

On the other hand, some participants added that presenting news with multiple multimedia elements made the awkward parts of the news more “noticeable.” When they read a sentence with an awkward expression, they just passed that part without noticing, whereas when they heard it, they suddenly

felt strange and became less focused on the news. P09 said, *“In video news, when NewsRobot read the name of the foreign competitor aloud, it sounded awkward.”* P24 said, *“When the voice read the same word in succession, like ‘third place, third place, third place’, I suddenly found it strange. I did not notice it when I read the same thing in text news. If a reporter had been reading that part aloud, he would not have done that.”*

DISCUSSION

We discuss implications for user interfaces of automated news generation systems.

Provide Individualized News with Adaptable Interface

The first thing we want to discuss is whether it is right or not to provide users with only individualized news articles that are tailored to their interests, rather than general news. Traditionally, concerns have been constantly raised about the uniformity of public news consumption using major media or channels [11]. The concentration of news production may not meet the diverse needs of users and may even lead to social problems in which public opinion is focused on specific issues.

NewsRobot, in this regard, showed the possibility of the diversification of news article production and the importance of users’ choice of news articles. NewsRobot’s pre-programmed algorithm simultaneously generated multiple competitor-specific articles, without the extra expense, effort, or time associated with creating articles using human reporters. Users showed great satisfaction with the individualized news. They preferred individualized news to general news and rated the former as less boring and more enjoyable, and interesting. They felt as if the algorithms produced news tailored to their interests. Some even described that the individualized news made them feel their tastes were respected. They have experienced diversified and individualized news consumption that is distinct from the news consumption through mass media, which was provided in a representative, uniform, and unilateral manner.

On the other hand, the adverse effects of individualized news consumption should be considered at the same time. Individualized news consumption is likely to make people more biased in other respects. It can insulate users from information that does not attract their interest, isolating them in their own thoughts. In particular, this can cause problems such as filter bubbles and echo chambers [5, 25]. The survey results also showed that users were aware that individualized news articles may be less credible than general news articles. If individualized news consumption is accelerated and transparency in the process is not ensured and recognizable by the people, individualized news is likely to cause many side effects in spite of its many advantages.

In this situation, introducing the concept of adaptable interfaces in designing automation systems [24] could be considered. Adaptable interfaces provide a customization mechanism that relies on the user to perform the adaptation (i.e., user-controlled personalization). Users need to select their news according to their interests, by themselves, and be explicitly informed that the news articles they receive are based on their choices [18]. Fortunately, in the interview, participants

also responded that they had a good experience in the news selection process, as it made them feel like they were active news service consumers.

Providing users with an adaptable, automated news generation system also requires the attention and efforts of the various communities-designers as well as algorithm engineers since it is not just an algorithm problem but also an interface problem. In this regard, it is necessary for the designers of the HCI community to address their concerns and strive to create adaptable interface designs of automated journalism systems to prevent the potential risks of such algorithms and promote better user experiences.

Design recommendation: Provide individualized news to users, but let them “customize it themselves” and know that the news is created through the adaptation process.

Present Various Multimedia Elements but Not Too Many

The second discussion point is about the style of NewsRobot. The user study showed that participants had better news consumption experiences when more multimedia elements were included. People thought that news articles were more enjoyable, lively, interesting, pleasing, and even informative in order of text+image+sound, text+image, and text news. They expressed high satisfaction by specifically mentioning the various presentation elements of NewsRobot. People appreciated the fact that it could produce those elements more quickly, accurately, and easily than human reporters and provide additional information not previously revealed.

On the other hand, people complained that text+image+sound news was a bit different from their expectations and that it brought out the awkwardness of the automatically generated news contents. According to user research studies on software or agents that replace human workers or simulate abilities of humans, people can feel more awkward or uncomfortable as they become more similar to humans [51, 58]. The more elements that are included in the news, the greater the expectations of people. However, at the same time, the awkwardness could be more prominent, and people’s disappointment could be greater [9]. The problem of both these high expectations and disappointments in algorithm-based news generation needs to be solved, because it can foster users’ negative perceptions of the news contents more rapidly. It has been reported that people are likely to lose confidence in algorithmic reporters more quickly than they do in human reporters after seeing them make the same mistake [19].

Although advanced algorithms and technologies for creating content are constantly being developed, relatively little attention is paid to how real users will accept content or interfaces that utilize them. It is needed to consider creating news content that is more perceptible to people rather than creating overly experimental and challenging news content that relies only on new technologies. Understanding users’ mental models on each multimedia modality of news contents could be a way of this approach, since it can help shape users’ expectations and set an approach to provide users with optimized experience. If more user-oriented and element-specific design principles are established and reflected in the content design and generation

algorithm, it will be possible to utilize more effectively the advantages of various multimedia that the user wants.

Design recommendation: Use a variety of multimedia to effectively present information in news articles but design them based on users' mental models for each news format to meet realistic user expectations.

Importance of Quality Data and Algorithm Refinement

The last point we discuss is the importance of obtaining quality data and refining algorithms for automatic news generation. In the user study, participants highly appreciated NewsRobot's data-driven news generation algorithm, describing it as accurate, objective, and even trustworthy. However, at the same time, they pointed out its limitations due to its excessive dependence on data, describing it as dull and shallow. Considering that software-generated content is usually perceived as descriptive and boring, despite being considered objective [10], this could be an inherent problem that automotive journalism needs to address.

In order to alleviate these problems while taking advantage of providing data-based information, it is necessary to obtain more and better data and refine the news generation algorithm. In terms of data, using diverse data sources, such as image information extraction and social media reactions, and avoiding using only refined numerical data could be considered. In terms of contents-generation algorithms, they should actively interpret the input data and produce various versions of the articles. They should not only provide calculated values but also help users understand the information by adding more detailed explanations of the meaning.

Above all, it is necessary to introduce state-of-the-art artificial intelligence technologies that can learn and interpret data autonomously, and generate sentences and elements based on it, rather than depending on rule-based templates. To do so, designers also need some understanding of AI techniques, including machine learning, for designing and creating news content. They should continue to search for new technologies related to news generation, such as natural language generation APIs that can express various voices and emotions naturally. If not, at least they should be able to actively communicate with and share their knowledge with the technicians developing them or those familiar with them. This will help develop algorithms that learn data more actively and produce content more adaptively to the user, which will increase the overall quality of news based on automated algorithms.

Design recommendation: By introducing new and diverse technologies and human expertise, obtain quality data and enhance algorithms so that users can understand the rich context of news articles.

Limitations and Future Work

We have identified several limitations of this study. (1) We did not conduct a comparative study comparing NewsRobot's articles and actual articles written by human reporters in the evaluation. (2) Although we generated the news articles during the Olympic Games in real-time, the user evaluation itself

was not conducted in real-time, since we had to consider unexpected situations that could occur in a real environment and control the experiment. (3) As the prototype and the experiment were designed for only one sporting event, generalization issues can be raised. (4) It is arguable whether the item set we used in the questionnaire is the best indicator for automated journalism, since it is rather old and not intended to be used directly as survey items, though it has been widely cited in the literature. (5) We designed the structure of the articles using a rule-based template and did not make various changes.

In future work, we plan to carry out a comparative study on news articles written by both human journalists and NewsRobot. Moreover, we will proceed with a user study when an event is actually happening, measuring the user evaluation of the automated news generation system in real-time. We also plan to expand the user study area to various topics, such as election reports and weather forecasts, to generalize our results on automated journalism. We will review various previous studies, including Sundar's follow-ups, to explore new criteria and methods for assessing news based on automated journalism. Lastly, we will improve the automation level of news generation by adopting state-of-the-art techniques.

CONCLUSION

This study examined the user perception of automated journalism, where news articles are generated by algorithms, mainly focusing on its content and style. We designed a research prototype, NewsRobot, which generated news articles about the PyeongChang 2018 Winter Olympic Games, and conducted a user study of the system using both quantitative and qualitative approaches. The results of the study revealed the following. (1) Although people regarded general news as more reliable than individualized news, they preferred the latter to the former. (2) People also liked news with a high multimedia modality level, but they considered slide news to have the best quality. (3) People regarded NewsRobot as accurate and objective but monotonous, and they were mostly satisfied with its diverse elements. Finally, based on these findings, we suggested design implications for automated news generation systems. We hope that this work will serve as a step toward a more productive and more inclusive understanding of interfaces for automated news generation systems.

ACKNOWLEDGMENTS

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2017-0-00693, Broadcasting News Contents Generation based on Robot Journalism Technology).

REFERENCES

- [1] Eytan Adar, Carolyn Gearig, Ayshwarya Balasubramanian, and Jessica Hullman. 2017. PersaLog: Personalization of News Article Content. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3188–3200.
- [2] Christopher W Anderson. 2013. Towards a sociology of computational and algorithmic journalism. *New media & society* 15, 7 (2013), 1005–1021.

- [3] Michael A Beam. 2014. Automating the news: How personalized news recommender system design choices impact news reception. *Communication Research* 41, 8 (2014), 1019–1041.
- [4] Jelle W Boumans and Damian Trilling. 2016. Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism* 4, 1 (2016), 8–23.
- [5] Engin Bozdog. 2013. Bias in algorithmic filtering and personalization. *Ethics and information technology* 15, 3 (2013), 209–227.
- [6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [7] Matt Carlson. 2015. The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. *Digital journalism* 3, 3 (2015), 416–431.
- [8] Matt Carlson. 2016. Automated journalism: A posthuman future for digital news? In *The Routledge companion to digital journalism studies*. Routledge, 226–234.
- [9] Deborah Soun Chung. 2007. Profits and perils: Online news producers’ perceptions of interactivity and uses of interactive features. *Convergence* 13, 1 (2007), 43–61.
- [10] Christer Clerwall. 2014. Enter the robot journalist: Users’ perceptions of automated content. *Journalism Practice* 8, 5 (2014), 519–531.
- [11] Elisia L Cohen. 2002. Online journalism as market-driven journalism. *Journal of Broadcasting & Electronic Media* 46, 4 (2002), 532–548.
- [12] Nicole S Cohen. 2015. From pink slips to pink slime: Transforming media labor in a digital age. *The Communication Review* 18, 2 (2015), 98–122.
- [13] International Olympic Committee. 2018. PyeongChang Olympics | Next Winter Games in Korea. (2018). Retrieved September 17, 2018 from <https://www.olympic.org/pyeongchang-2018>.
- [14] NAVER Developers. 2017. Clova Speech Synthesis (CSS) - Naver Developers. (2017). Retrieved September 17, 2018 from <https://developers.naver.com/products/clova/tts/>.
- [15] Nicholas Diakopoulos. 2011. A functional roadmap for innovation in computational journalism. *Nick Diakopoulos* (2011).
- [16] Nicholas Diakopoulos. 2014. Algorithmic-Accountability: the investigation of Black Boxes. *Tow Center for Digital Journalism* (2014).
- [17] Nicholas Diakopoulos. 2015. Algorithmic accountability: Journalistic investigation of computational power structures. *Digital Journalism* 3, 3 (2015), 398–415.
- [18] Nicholas Diakopoulos. 2016. Accountability in algorithmic decision making. *Commun. ACM* 59, 2 (2016), 56–62.
- [19] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. 2015. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General* 144, 1 (2015), 114.
- [20] Konstantin Nicholas Dörr. 2016. Mapping the field of algorithmic journalism. *Digital Journalism* 4, 6 (2016), 700–722.
- [21] Paul S Earle, Daniel C Bowden, and Michelle Guy. 2012. Twitter earthquake detection: earthquake monitoring in a social world. *Annals of Geophysics* 54, 6 (2012).
- [22] @earthquakesLA. 2009. LA QuakeBot. Twitter ID. (1 May 2009). Retrieved September 17, 2018 from <https://twitter.com/earthquakesLA>.
- [23] Marcus Errico, J April, A Asch, L Khalfani, M Smith, and X Ybarra. 1997. The evolution of the summary news lead. *Media History Monographs* 1, 1 (1997).
- [24] Leah Findlater and Joanna McGrenere. 2004. A comparison of static, adaptive, and adaptable menus. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 89–96.
- [25] Seth Flaxman, Sharad Goel, and Justin M Rao. 2016. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly* 80, S1 (2016), 298–320.
- [26] Andreas Graefe. 2016. Guide to automated journalism. (2016).
- [27] Andreas Graefe, Mario Haim, Bastian Haarmann, and Hans-Bernd Brosius. 2018. Readers’ perception of computer-generated news: Credibility, expertise, and readability. *Journalism* 19, 5 (2018), 595–610.
- [28] Hanno Hardt. 1990. Newswriters, technology, and journalism history. *Critical Studies in Media Communication* 7, 4 (1990), 346–365.
- [29] G Hart. 2012. The Five Ws of Online Help. *TECHWR-L*. Retrieved April 30 (2012).
- [30] Jessica Hullman and Nick Diakopoulos. 2011. Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2231–2240.
- [31] AUTOMATED INSIGHTS INC. 2007. Automated Insights - Natural Language Generation. (2007). Retrieved September 17, 2018 from <https://automatedinsights.com>.
- [32] Peter Johnson and Fabio Nemetz. 1998. Towards principles for the design and evaluation of multimedia systems. In *People and Computers XIII*. Springer, 255–271.

- [33] Jaemin Jung, Haeyeop Song, Youngju Kim, Hyunsuk Im, and Sewook Oh. 2017. Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists. *Computers in Human Behavior* 71 (2017), 291–298.
- [34] D Kim and J Lee. 2015. Robot journalism: Algorithmic approach to automated news article generation. *Korean J. Journal. Commun. Stud* 59 (2015), 64–95.
- [35] Dongwhan Kim and Joonhwan Lee. 2019. Designing an Algorithm-Driven Text Generation System for Personalized and Interactive News Reading. *International Journal of Human-Computer Interaction* 35, 2 (2019), 109–122.
- [36] Tetyana Lokot and Nicholas Diakopoulos. 2016. News bots: Automating news and information dissemination on Twitter. *Digital Journalism* 4, 6 (2016), 682–699.
- [37] Neil Maiden, Konstantinos Zachos, Amanda Brown, George Brock, Lars Nyre, Aleksander Nygård Tonheim, Dimitris Apsotolou, and Jeremy Evans. 2018. Making the News: Digital Creativity Support for Journalists. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 475.
- [38] Kathleen McKeown and Dragomir R Radev. 1995. Generating summaries of multiple news articles. In *Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 74–82.
- [39] Tal Montal and Zvi Reich. 2017. I, robot. You, journalist. Who is the author? Authorship, bylines and full disclosure in automated journalism. *Digital Journalism* 5, 7 (2017), 829–849.
- [40] John Pavlik. 2000. The impact of technology on journalism. *Journalism Studies* 1, 2 (2000), 229–237.
- [41] MG Petersen. 1996. Evaluating usability of multimedia interfaces. *HCI MSc Project Report, Department of Computer Science, Queen Mary and Westfield College* (1996).
- [42] Marc Prensky. 2005. What can you learn from a cell phone? Almost anything! *Innovate: Journal of Online Education* 1, 5 (2005).
- [43] Ehud Reiter. 2010. Natural language generation. *The handbook of computational linguistics and natural language processing* (2010), 574–598.
- [44] Barbara K Rimer and Matthew W Kreuter. 2006. Advancing tailored health communication: A persuasion and message effects perspective. *Journal of communication* 56 (2006), S184–S201.
- [45] Anthony J Roberto, Janice L Krieger, and Michael A Beam. 2009. Enhancing web-based kidney disease prevention messages for Hispanics using targeting and tailoring. *Journal of Health Communication* 14, 6 (2009), 525–540.
- [46] C Scanlan. 2003. Birth of the inverted pyramid: A child of technology, commerce and history. *Poynter*. Retrieved from *poynter.org* (2003).
- [47] Michael Schudson. 1997. The sociology of news production. *Social meanings of news: A text-reader* 722 (1997).
- [48] Narrative Science. 2010. Narrative Science | Natural Language Generation Technology. (2010). Retrieved September 17, 2018 from <https://narrativescience.com>.
- [49] S Shyam Sundar. 1999. Exploring receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly* 76, 2 (1999), 373–386.
- [50] S Shyam Sundar. 2000. Multimedia effects on processing and perception of online news: A study of picture, audio, and video downloads. *Journalism & Mass Communication Quarterly* 77, 3 (2000), 480–499.
- [51] Luke Swartz. 2003. *Why people hate the paperclip: Labels, appearance, behavior, and social responses to user interface agents*. Ph.D. Dissertation. Stanford University Palo Alto, CA.
- [52] Neil Thurman, Konstantin Dörr, and Jessica Kunert. 2017. When Reporters Get Hands-on with Robo-Writing: Professionals consider automated journalism's capabilities and consequences. *Digital Journalism* 5, 10 (2017), 1240–1259.
- [53] Peter Tolmie, Rob Procter, David William Randall, Mark Rouncefield, Christian Burger, Geraldine Wong Sak Hoi, Arkaitz Zubiaga, and Maria Liakata. 2017. Supporting the use of user generated content in journalistic practice. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3632–3644.
- [54] Lokman Tsui. 2009. Rethinking journalism through technology. In *The Changing Faces of Journalism*. Routledge, 63–65.
- [55] Arjen Van Dalen. 2012. The algorithms behind the headlines: How machine-written news redefines the core skills of human journalists. *Journalism Practice* 6, 5-6 (2012), 648–658.
- [56] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 440.
- [57] T Franklin Waddell. 2018. A robot wrote this? How perceived machine authorship affects news credibility. *Digital journalism* 6, 2 (2018), 236–255.
- [58] Stuart NK Watt. 1997. Artificial societies and psychological agents. In *Software Agents and Soft Computing Towards Enhancing Machine Intelligence*. Springer, 27–41.
- [59] Oscar Westlund. 2013. Mobile news: A review and model of journalism in an age of mobile media. *Digital journalism* 1, 1 (2013), 6–26.

- [60] Anja Wölker and Thomas E Powell. 2018. Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. *Journalism* (2018), 1464884918757072.
- [61] Allison Woodruff, Sarah E Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. 2018. A Qualitative Exploration of Perceptions of Algorithmic Fairness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 656.
- [62] Optimal Workshop. 2016. Reframer. (2016). Retrieved September 17, 2018 from <https://www.optimalworkshop.com/reframer>.