Kakao Deep Reading Index: Consumption Time as a Key Factor in News Curation Algorithm

Dongkwon Lee [1] and Daewon Kim[1*]
[1]Kakao Corp.
Seongnam-si, Gyeonggi-do, Republic of Korea
[e-mail: dongkwon@gmail.com, egofree@gmail.com]
*Corresponding author : Daewon Kim

Abstract

This paper introduces the structure and effects of Kakao's news curation algorithm, which is created based on the Deep Reading Index (DRI). The DRI examines the extent of deep reading through content reading time, that is, the duration of reader engagement with an article. Current news curation algorithms focus on reader choice, with the click-through rate or pageviews as the gauge for consumption frequency. DRI is a product of the challenge of introducing and adopting a new factor called 'consumption time' instead of 'frequency of consumption', which is the basis of existing curation algorithms. The analysis of DRI-based services proves that the new algorithm can act as a curation system that is more effective in providing in-depth and quality news reports.

Keywords: News curation, Recommendation algorithm, Consumption time, Deep reading index, Kakao

1. Introduction

With the proliferation of the Internet in the digital age, the number of information suppliers has spiked, with lowered barriers of information production and delivery to the point of information overload. The media industry is not exempt from this; the excess supply has led to higher demand for efficiency in content consumption among consumers. Content curation is the reaction to information overload and has become essential to enhance the competitiveness of media platforms that seek a high user-retention rate. Following the advancements in artificial intelligence, media platforms are opening up to the possibility of performing news curation through algorithms. Successful cases have already been reported, setting benchmark standards for press companies and media platforms.

Practical application of news content recommendation system has long been studied and evolved. GroupLens's project for news recommendation on Usenet began in 1992, and Personalized TANGO, one of the first hybrid news recommendation algorithm applied on Worcester Telegram and Gazette Online, was announced in 1999 [1], [2]. Yet, the earliest and most successful commercial applications of news recommendation should be Google News and Yahoo Frontpage. Google News uses an ensemble algorithm that consists of three collaborative filtering technics [3]. It again evolved to be a hybrid model by combining content-based methods [4] Yahoo, in the meantime, composes quality news item pool by editors and exposes them in order of predicted user preference among the items [5]. Yahoo has continuously been improving its news recommendation algorithm through testing and complementing alternatives (e.g. [6]-[8]).

Kakao[1], one of the major Internet service providers in Korea, also has developed and applied its own news recommender, called RUBICS[2], for its mobile app and web portal (Daum) news service since 2015.[3] When comparing the performance indicators between before and after two months from the launch of RUBICS, daily click rate increased 130%p, daily visitor increased 45%p, and quantity of news items exposed on the frontpage 250%p, which clearly showed the positive effects of an improved recommendation system [9]. RUBICS inherits the major features of Yahoo news recommender (e.g. multi-armed bandit model) and especially inspired by its gradual improvement process.

---

[1] Kakao is a company that operates major mobile platforms in Korea. Its mobile messenger, KakaoTalk, has 94% of market share by monthly usage. The company offers a wide range of mobile services such as content services, e-commerce, Internet bank, and transportation. Kakao is also recognized as a domestic leader in AI technology.

[2] Real-time User Behavior Interactive Content recommender System

[3] In June 2015, Kakao became the first in Korea to deploy news sorting algorithm to manage news arrangement which has been mainly done by professional news editors. Kakao's news recommendation algorithm personalizes news content according to individual consumption patterns and preferences.
This can be seen as efforts made at the technical level to accommodate current trends. With more content consumed online, there has been a significant increase in news articles for that user demographic. However, the physical accessibility to provide online news was limited. Amid these changes, individual users had a desire to efficiently access and enjoy news content that fit their interests. Thanks to the widespread use of smartphones and the "one-PC-per-person era," suppliers were able to meet this demand.
Kakao's news recommendation algorithm produced an increase in news consumption and diversified the types of news consumed [9]. First successfully adopted by mobile services, the personalized algorithm was then expanded for the personal computer (PC) environment in April 2017.
As well as Kakao, Naver, another leading Internet companies in Korea, has applied curation algorithm to its news service, since 2017. Cyworld launched its news service based on curation algoritm in 2018.

The main features of RUBICS was dependent on click-through rate (CTR). CTR per se is not sufficient for predicting users' potential preference and behavior, because it does not contain any information about users' reaction to the items after clicking, such as relevancy and satisfaction. For instance, dependence on CTR could have not effectively resolved an old problem of noises derived from clickbait, which has harmed trust toward online news services, because every single click is rated as 1, regardless of how quickly a user left the item. Also, the frequency of content selection by readers is not an accurate indicator of actual consumption patterns and fails to convey user experience evaluations on content quality.

We focused on this problem and tried to improve our news recommendation algorithm weighting more on 'how users consume news items', deviating from 'whether users choose news items' in the conventional CTR approach.

Therefore, we sought for the potential of duration of stay in pageview, as a proxy for reading time, to be considered as a critical component in the news recommendation algorithm research arena. In detail, we designed an algorithm based on the Deep Reading Index (DRI), which uses consumption time as a critical factor instead of consumption frequency. DRI assesses the deep reading rate of a user by comparing the actual time spent reading a news article with the expected duration, and an estimate based on attributes such as news category, article length, image count, and card news format.

In this paper, we show the practical results of the application of DRI on news services of Kakao. Our findings will further stimulate meaningful academic discussion and generate creative, practical ideas.

2. Modeling Deep Reading Index

2.1. introduction of DRI

The existing news curation algorithms of press companies and media platforms rely on the frequency of use as the main criterion. For instance, a piece of content, such as a news article, blog post, or advertisement, is evaluated based on its pageviews (PVs) or CTR on the curated website (Eq. 1).

Eq. (1) CTR = Total number of clicks / total number of times exposed

A visit to a particular webpage or the selection of specific content cannot be seen as genuine consumption of the users. This is because unintended visits and a roll back to the previous page due to withdrawn decisions are also counted as "clicks" made by users. Algorithms depend only on the number of clicks produce clickbait.

Despite the criticism towards clickbait articles for softening and sensationalizing journalisms, they have, nevertheless, been employed by press companies to generate online revenue. Similar to the number of newspaper subscriptions and TV news view ratings, CTR determines the bargaining power of the online advertisements. The adverse effects of click-based algorithms are manipulated technical constraints, and this further highlighted the need to develop algorithms that reflect the qualitative aspects of user content consumption.

The evaluation of click-based user selection cannot take user behavior into account after the click. For example, CTR alone does not reveal whether a user had deeply engaged with the content or if the consumption content is deemed satisfactory. From this perspective, news curation algorithms should be further developed to determine the extent of deep reading during content consumption.

Kakao also attempted to develop new curation algorithms by identifying factors associated with consumption patterns and attitudes other than clicks; this led to the eventual birth of DRI to indicate user engagement.

DRI quantifies content satisfaction based on how involved a user is in reading. The extent of deep reading is measured based on the time spent by a user on the selected news. The time spent by a user, as well as content-related attributes, such as length, image, multimedia, advertisements, and comments, are collected as training data. From this training data, the estimated duration function is acquired and then used to measure the extent of deep reading to quantify user satisfaction. The goal is to utilize the algorithm in content curation.

DRI was devised to measure the level of satisfaction or attention given toward a piece of content after the click. With this approach, the assumption was that a user who feels a greater level of satisfaction would spend more time on the content. This process revealed the gap between CTR and duration. A content with high CTR but short in duration was likely an article with a clickbait headline, or one with a body that is inconsistent with the headline. This way, DRI can also be used to detect news content designed to raise CTR with misleading headlines. DRI was applied to main news service of Kakao in July 2017.

2.2. Procedure of practical application of DRI to analyze Kakao news contents

2.2.1. Duration measurement for news content

Satisfaction with the main body of an article can be measured in two ways. The first is scroll depth measurement, which determines whether a user viewed the entire page or closed the page midway by tracking the downward scroll movement. The point at the bottom of the page is considered as 100%, while the position of the scroll before leaving the page is expressed as a percentage of the page. For instance, if a user leaves the page after scrolling to the halfway point, the scroll depth value is 50%. However, unless duration is taken into account, scroll depth may not be an accurate indicator of user engagement as the user may skim up or down.

The second method is to measure the duration, that is, the time spent by a user on a page. A problem with this method is that the process does not consider the amount of information. For example, the amount of information acquired by a user within a duration of 50 s is different between an article composed of 800 characters and 5 images, and another article that consists of 1,000 characters. As such, duration alone is not a reliable indicator for the level of deep reading.

To resolve the above issue, DRI measures user satisfaction through a function that calculates the expected duration according to the amount of information. It retrieves the content viewing time of users and assigns a representative duration from the variety. For example, the representative duration may be the average, median, or mode of multiple users' duration of contents engagement. In this study, the median acts as the representative value to minimize the influence of outliers.

2.2.2. Learning the expected duration

2.2.2.1. Basic principles

Learning the expected duration involves its modeling for news content through a regression analysis of the relationship between variables in the training data. The training data include a representative duration and main body attributes. The attributes are length, images, videos, comments, and advertisements, and may be a combination of these.

Expected duration can be calculated using multiple regression, which forecasts a trend based on previous data. In other words, expected duration is the tendency of a user to consume a specific news article. The equation for expected duration is as follows (Eq. 2):

Eq. (2) expected duration = f (string, image, multimedia, advertisement, comment)

String information may refer to either the length of an article when words are listed out or the main body word count. These two factors may be utilized depending on the circumstances. For this study, string information will be limited to the number of characters for simpler explanation.[4] Length-based duration is calculated using a nonlinear regression model. The consumption patterns of 37,657 articles in Daum News[5] for December 2016 were analyzed to validate the nonlinear regression model. Of the total, 20,350 (54.0%) were current affairs, 10,785 (28.6%) were entertainment, and 6,522 (17.3%) were sports articles.[6] On the left graph of **Fig. 1**, the x-axis is the text length (unit: one character) while the y-axis is the duration (unit: 1 s). The individual points each represent an article, while the dotted line presents the relationship between length and duration. The dotted line is the result of a generalized additive model employed to examine the nonlinear relationship, which cannot be easily determined using standard linear regression. From **Fig. 1**, we can see that duration increases with length regardless of the news category but drops sharply when the length exceeds 2,500 characters. This becomes more evident on the right graph of **Fig. 1**, which was replotted after removing outliers, with text longer than 5,000 characters and a duration longer than 150 s.[7]
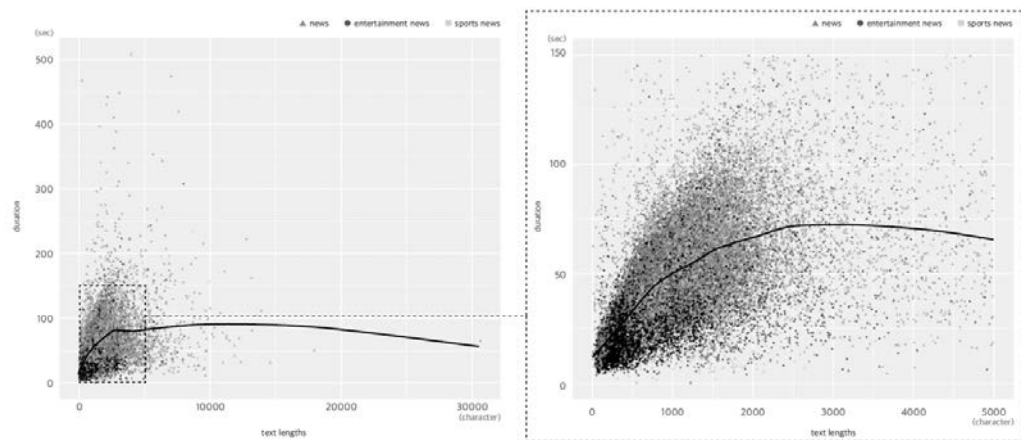

Fig. 1. Scatter plot of duration versus the number of characters

The data were modeled for length to assume nonlinearity because deep reading rate cannot be accurately derived from a linear relationship between length and duration. Through a preliminary review repeated several times to improve accuracy, this study found that an

---

[4] If an article contains "I love you," the number of words is three: "I," "love," and "you." The number of characters is eight: "I," "l," "o," "v," "e," "y," "o," and "u." In this study, length was calculated based on the number of characters.

[5] A news service provided through Daum, a web portal of Kakao

[6] The articles in different categories sum up to 99.9% because the second decimal point of each value was rounded up.

[7] In this process, 591 cases (1.5%) were removed from the total of 37,657. This step was performed under the presumption that there is no correlation between length and duration.

exponential nonlinear regression model was the most appropriate for text length in the modeling of expected duration.[8]

Recent news contents produced by press companies are composed of not only text but also images and videos. Accordingly, user satisfaction toward content cannot be measured based on the number of words alone.

Ultimately, there is a need to consider the multimedia patterns utilized in media content including images, videos, etc. Image information includes the number of images, image size, image type, and the number of card news images. For a simpler explanation of the algorithm, this study limited multimedia information to the number of images and the adoption of card news. Multimedia information can be broken down into elements such as the number of video or audio files as well as their respective playback times. However, in this study, such elements are grouped as) "multimedia information" to aid the understanding of users and looked at whether or not multimedia information was present.

Comments added below news content and related information may also influence the time spent on the main body. Comment-related information may include the number and length of the comments, as well as the image and multimedia information in the comments. As advertisements are another factor for duration, the number of advertisements may have to be considered in the calculation of expected duration. However, this factor was excluded in this study.

The model was generated using the consumption patterns of 37,652 articles serviced via Daum News in December 2016. The results obtained from applying nonlinear regress are given in **Table 1**. Nonlinear regression models were separately produced for the top three news categories: current affairs, entertainment, and sports. The columns—Estimated, Standard Error, and *P* Value—represent the regression coefficient, standard error, and significance of the individual regression coefficient, respectively.

Table 1. Results of nonlinear regression for news content

News Category		Estimated	Standard Error	*P* Value
Expected Duration = $aTL^b + cNI + dMM + eCN$
Current affairs	a	4.614204	0.154986	<0.0001
	b	0.345753	0.004585	<0.0001
	c	0.962812	0.128848	<0.0001
	d	15.141056	0.566402	<0.0001
	e	29.774974	3.828768	<0.0001
Entertainment	a	1.595899	0.0658	<0.0001
	b	0.467769	0.005857	<0.0001
	c	0.003943	0.119801	0.974
	d	1.182923	4.160467	0.776
	e	10.383302	6.961052	0.136
Sports	a	1.369791	0.074983	<0.0001
	b	0.513048	0.007587	<0.0001
	c	0.733048	0.211922	<0.0001
	d	13.129824	3.068047	<0.0001
	e	0	-	-

Note. ED = expected duration, TL = text length, NI = number of images, MM = multimedia, CN = card news; MM and CN are dummy variables that indicate whether or not the relevant attribute exists.

---

[8] Even when duration is calculated differently according to length, a high deep reading rate may be obtained for shorter articles regardless of quality because less reading time is required.

The estimated variables from the three models (current affairs, entertainment, and sports) were all found to be significant (see *P* Value). However, for the entertainment category, regression coefficients c, d, and e were not significant. Due to the large variation in the amount of information which each image, video and card news hold (i.e. size of image, playtime of video, length of card news article), there is a tendency for coefficients c, d, and e of entertainment news category, which respectively represent the contribution of image, video, and card news to duration, to have a larger standard error than that of other categories. The standard error can be minimized by reflecting the sum of the images' height and length proportions instead of the image number (NI); the same can be done with playback time instead of the number of videos. As the sports category does not contain any card news, coefficient e was excluded while generating the model.

Through the estimated coefficients of regression models, expected duration is calculated based on content-related information (e.g., text length, number of images, the use of multimedia, and the use of card news format). This is then compared to the actual duration to quantify how engaged a user was in reading an article. In more specific terms, the calculated expected duration for a certain amount of information serves as a baseline for determining user engagement in relation to information amount. When the actual duration largely exceeds the expected duration baseline, it is interpreted as greater user engagement in relation to information amount.

2.2.2.2. Differences by news category

Training data can be generated after grouping news content into different categories according to specific criteria. The classification criteria may be content-based or user-based. Under content-based classification, the categories may include current affairs, entertainment, and sports. Each of these categories can have subcategories. For example, the entertainment category can be divided into film and TV.

Expected duration exhibits different trends by news category. This is because each category of news content consists of different attributes. For example, content on current affairs tends to be longer than that on entertainment and contains fewer images. These differences are reflected in the expected duration function by category (**Table 2**).

Table 2. Expected duration function by news category

News Category	Expected Duration Function
Current affairs	$ED = 4.6142TL^{0.3458} + 0.9268NI + 15.1411MM + 29.7750CN$
Entertainment	$ED = 1.5958TL^{0.4677} + 0.0039NI + 1.1829MM + 10.3883CN$
Sports	$ED = 1.370TL^{0.513} + 0.733NI + 13.130MM + 0CN$

Note. ED = expected duration, TL = text length, NI = number of images, MM = multimedia, CN = card news; MM and CN are dummy variables that indicate whether or not the relevant attribute exists. CN is 0 for the sports category because it does not contain any card news.

The following is a description of the functions presented in **Table 2**. For content falling under current affairs (e.g. news articles), duration increases nonlinearly with text length; those containing card news had a longer duration of about 30 s. When multimedia and images were added, the duration increased by 15 s and 1 s, respectively. That is, for content that belongs to the current affairs category, the factors that influence duration other than text length were the use of card news, images, and multimedia. For content in the entertainment category, images, multimedia, and card news had less significant effects on duration. For sports news, multimedia had a more significant effect than text length and image number on duration.

Regardless of the news category, text length is directly proportional to the expected duration for short articles. However, when the text length exceeds a certain point, the increase in expected duration slows down. As shown in **Fig. 2**, the relationship between expected duration and text length is nonlinear. **Fig. 2** also reveals differences in the influence of text length on expected duration by news category. The data used to plot **Fig. 2** were the 37,652 articles consumed in Daum News in December 2016. The articles were grouped into three categories: current affairs, entertainment, and sports. The x-axis represents text length (unit: one character), while the y-axis is the duration (unit: 1s). As described earlier, this is the result of the application of nonlinear regression; variables other than text length were set to 0 to narrow down observations to the influence of text length on duration by category.

As seen from the graph, duration was longer in the current affairs category for short articles. For longer texts, the average duration was higher for news in the sports category. The more proportionate relationship between text length and duration for sports news reflects the high interest among sports fans.

The factors that affect duration and their significance vary according to category. Each category should have a separate function for expected duration; alternately, specialized variables should be added to the model to account for such differences.
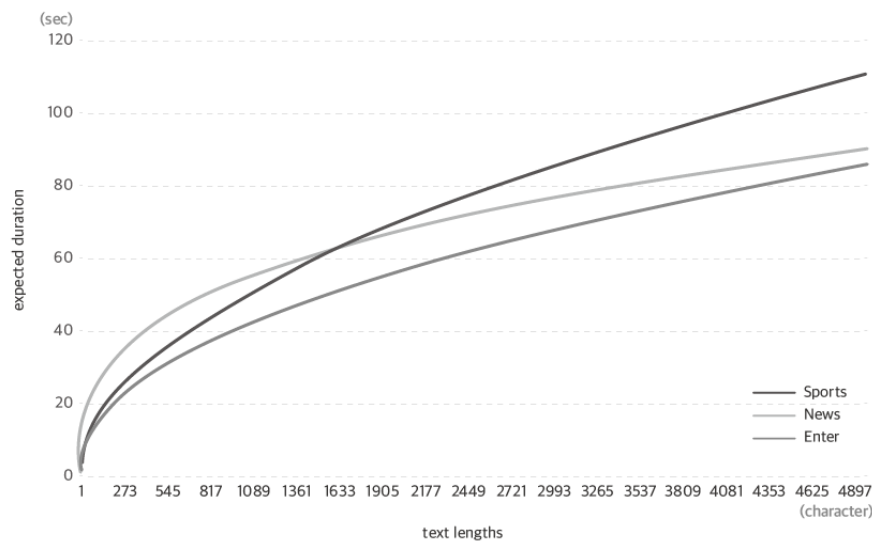

Fig. 2. Expected duration versus text length by news category

2.2.3. Calculation of DRI

DRI is calculated by comparing the expected duration and the actual duration. Deep reading rate can be evaluated in various ways. For example, it can be expressed as 1) the increase or decrease of the actual duration as compared to the expected duration and 2) the ratio of the actual duration to the expected duration. As shown in the equation below, this study calculated DRI using the first option, that is, the increase or decrease compared to the expected duration.

Eq. (3) DRI = (Actual duration - expected duration) / expected duration × 100

2.2.4. Utilization of DRI

Table 3 shows the results of calculating DRI for the 37,652 articles consumed by users in Daum News in December 2016. Content B has more PVs than Content A but a lower DRI, while Content A, which is longer than Content B, has less PVs but a higher DRI. The results indicate that Content B attracted more clicks because of its attractive headline or sensational topic but was less satisfying to read. The deep rating rate of Content C was a negative value, which means that users were either dissatisfied with the article or the article was clickbait with a misleading headline.

The analyzed news content is arranged above in decreasing order of DRI. The results can be used to isolate content that was more satisfying for users than expected or clickbait headlines with weak content.

Table 3. Results of applying DRI to news content

News	Text Length (Characters)	Page View (Views)	Expected Duration (Seconds)	Actual Duration (Seconds)	DRI (%)
Content A	3,300	2,100	77	144	87.0
Content B	1,300	113,000	57	60	5.3
Content C	8,686	4,230	148	4	-97.3

3. Results of applying DRI

3.1. General results

Using DRI, Kakao has been releasing a ranking of the most thoroughly read articles since July 13, 2017. The ranking for most widely read news, which relies on PV, was launched more than five years ago. As such, the two rankings are likely to differ in terms of levels of awareness and usage. For a more accurate assessment of user response, the DRI-based "most thoroughly read news" was compared to the PV-based "most widely read news" for a certain period following the launch date of each ranking.
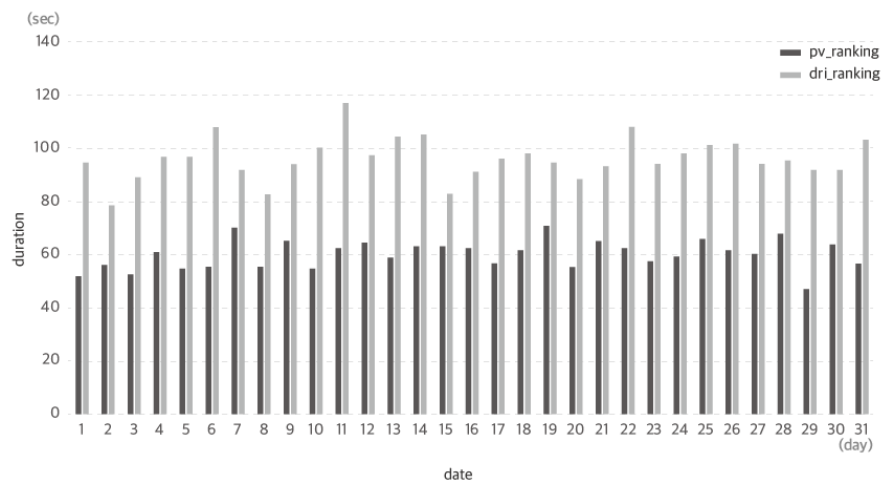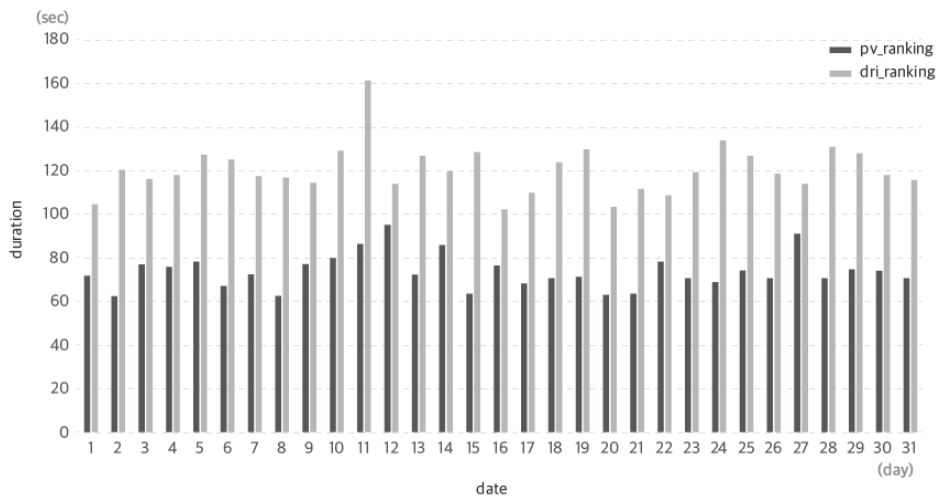

Fig. 3. Average duration per article in the two ranking systems (as of July 2018)

The chart below shows the average duration for articles listed in the PV and DRI rankings in July 2018. In **Fig. 3**, the x-axis represents the days of the month, while the y-axis charts the duration. The average duration of articles featured in the DRI-based ranking was 96s, while that of articles in the PV-based ranking was 60.1s. On average, an article in the DRI-based ranking has a duration higher than the PV-based ranking by 59.8%.

However, a longer duration of reading an individual article does not necessarily mean that it was consumed for a longer period. Even with the longer average duration, the total duration that a user spent reading may be similar or shorter. If other short articles are left unread because of the time spent reading a long article, the tradeoff is a reduction in the total duration per person. For example, it takes 80s to read one article with a duration of 80s and 100s to read five articles with a duration of 20s each. For the former, the average duration per article and the average duration per person are both 80s. For the latter, the average duration per article is 20s, while the average duration per person is 100s. This means that the average duration per person must be measured in addition to the average duration per article to validate the increase in news consumption per person.

Below is a comparison of the average duration per person in the two ranking pages described earlier for the same period of July 2018. The average duration per person was 73.6s in the PV ranking page and 120.5s in the DRI ranking page. The average duration per person for DRI articles was 63.7% longer than that for PV articles.

Fig. 4. Average duration per person in the two ranking systems (as of July 2018)

The results indicate that the average duration per person does not shorten even if more time is spent on an article. In fact, as shown in the graph above, the average duration slightly increases for the DRI ranking page. Even if a DRI article with a longer duration is consumed, the average number of articles consumed per person (A/B) was higher for the DRI ranking (1.255) than for the PV ranking (1.224). The average number of articles consumed per person was calculated by dividing the average duration per person by the average duration per article. When an article with a high DRI is provided, more time is spent reading that article. As the average duration per person is longer than the average duration per article, there is no trade-off in the overall consumption because of the longer duration per article.

When the average duration per article increases, this can result in less time spent by a user on an article (A) or fewer articles read (A/B). However, the key finding of the above analysis is that such trade-offs did not exist and that the number of articles consumed per person increased.

Table 4. Duration for each ranking

Classification	PV Ranking	DRI Ranking
Average duration per person (A)	73.6	120.5
Average duration per article (B)	60.1	96.0
Average number of articles consumed per person (A/B)	1.224	1.255

Note. the unit of A and B is second.

The differences in user response were examined based on duration. Given that DRI provides insights into the qualitative features of articles, this study extracted the headlines of the top five articles in the DRI-based "most thoroughly read news" and the PV-based "most widely read news" for August 7, 2018. The headlines are presented in **Table 5**.

Table 5. Examples of headlines of highly ranked news based on DRI and PV

Rank	DRI (Most Thoroughly Read News)	PV (Most Widely Read News)
1	[Lee Kwang-sik's astronomy+] How many stars are there in the galaxy?	Man kills friend of 30 years over money: Was it worth it?
2	Promoted to corporal on first duty assignment: Where do privileges come from?	Kim Kyung-soo denies allegations in 18-hour questioning; "Evidence was found to be insufficient"
3	[Xi Jinping in crisis 1] Will the Chinese Dream change under criticism?	Exaggerated reviews of ocean-view restaurants leave patrons seething.
4	How is North Korea perceived by China: 'discredit' and 'uncivilised'	[Exclusive] Angry BMW 520d owners protest in front of a shop at Yeongdong-daero.
5	[Behind news] Prisoner's dilemma behind the ex-presidential secretary's confession.	Majority of the population born in 1971: Residents' age is 41.5 years on average.

Note. The above articles have been published on 7th August.

The top articles in the two ranking pages differed in both content and format. The highest ranked news in the PV ranking were straight articles dealing with facts. They featured actual events and accidents as well as statements made by politicians. The headlines of the top ranked articles were mainly a summary of key content such as "Man kills friend of 30 years over money," "Kim Kyung-soo denies allegations in 18-hour questioning," "Angry BMW 520d owners protest in front of a shop at Yeongdong-daero," and "Residents' age is 41.5 years on average." Many straight articles ranked highly in the PV ranking fell under society and politics.

Highly ranked contents in the DRI ranking were mainly in-depth articles on specific issues. Such specific issues as "How many stars are there in the galaxy?" "Will the Chinese Dream change under criticism?" and "How is North Korea perceived by China?" were analyzed or presented with contextual information. Articles in this ranking were titled after their key content. Here, user needs are satisfied with the acquisition of related knowledge or information upon clicking headlines of interest. The highly ranked articles were from not only society and politics but also other categories that benefit from explanation such as economy, foreign news, and science.

Qualitative analysis showed that the categories were more diverse in the DRI ranking; this is also supported quantitatively when analyzed over a longer period. **Fig. 5** presents the categories of news articles listed in the top 20 of the DRI-based "most thoroughly read news" and the PV-based "most widely read news" in July 2018. While close to half of the articles in the PV-based "most widely read news" ranking belonged to the society category, those in the DRI-based "most thoroughly read news" were evenly distributed among society, politics, and foreign news. The results indicate the greater diversity that DRI offers in terms of news category compared to PV.[9]
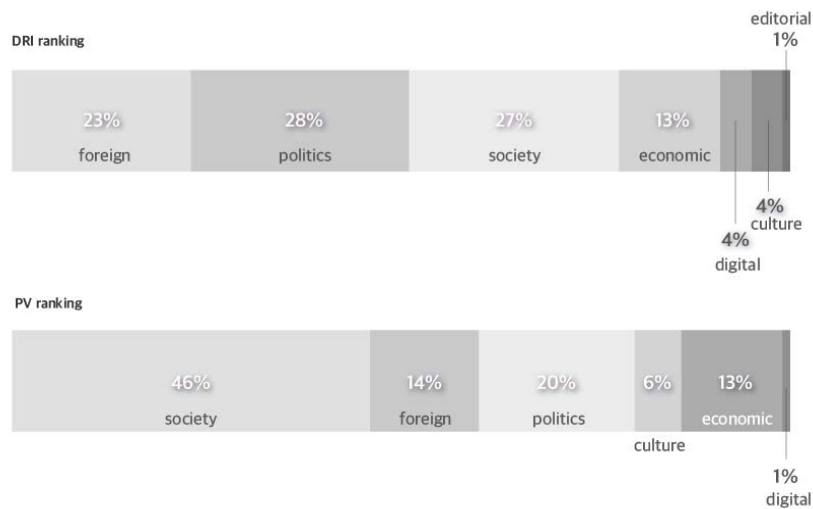

Fig. 5. Categories of news content in DRI and PV rankings

3.2. Clickbait filtering

In online news consumption, the headline is one of the primary ways of attracting clicks; some press companies create sensationalist or provocative headlines to generate more PVs. Getting users hooked with unrelated or exaggerated headlines is a clickbait tactic that leaves users feeling betrayed when they realize that the headlines cannot meet the expectations raised for them. Such users can then choose to exit or protest against suppliers [10], [11]. Between exiting and protesting, the former is easier and more convenient. Applying this framework to the consumption of news content, we can expect disappointed users to return to the previous webpage or close the current window.[10]

---

[9] The concentration of articles in society and politics may be a result of the social circumstances at the time. If the period of analysis coincided with a major sports event, such as the Olympics or the World Cup, related sports news would have been featured in the PV ranking. Here, the key finding is that straight articles constituted the majority of top articles in the PV ranking.

[10] The algorithm cannot be used to read the intention of news suppliers, which can only be inferred from DRI values. Clickbait articles identified by the DRI are based on observations of similar cases. In other words, the DRI may not be a conclusive indicator of a news supplier's intent to deceive.

[Interview]
Title: A conversation with the key players of the Samsung dynasty: Conflict, income, and return to Korea

Many teams have dominated periods in the history of soccer. Recently, Leicester City won the Premier League title for the first time. Other teams still talked about Inter Milan, which led to a French victory in the 1998 FIFA World Cup; AC Milan; and Real Madrid.
For League of Legends (LoL), an eSport with a relatively short history, the teams that have come under the spotlight are the record-breaking SK Telecom T1; the early-day champion CJ Entus Frost; and, of course, Samsung White and Samsung Blue.
Since their victory in the League of Legends Champions Korea in 2013, Samsung White (MVP Ozone at the time) has consistently ranked within the top three. The Samsung dynasty began when the team gained recognition as the only worthy rival of SK Telecom. In fact, Samsung White is the only team that has defeated SK Telecom several times in best of five matches. (redacted)

[Comments]
ID: Elephant Man
"I thought this was about baseball…"
ID: Nick Rhodes
"Is this a clickbait? [*sic: this clickbait?*]"

[DRI results]
Text length: 8,686 characters
Expected duration: 148s
Actual duration: 4s
DRI: -97%

Fig. 6. An example of DRI for an article that fails to meet user expectations due to mismatch between headline and content

DRI values reveal that such users were inclined to exit the webpage. Users' comments that express dissatisfaction towards misleading headlines with mentions of "headline" or "clickbait" were detected for 13 articles released in July 2018. These articles had an average DRI of -23.5% and a standard deviation of 0.35. Of the 13 cases, 11 (84.6%) had DRI values below 0. The negative values show that disappointed users quickly exited pages that contain unsatisfactory content.

The DRI can identify not only clickbait articles with headlines intended to deceive but also articles with ambiguous headlines. **Fig. 6** is an example of an article with an unconventional headline that caused users to quickly retract their decision. For Korean news readers in general, the term "Samsung dynasty" in Korean sports news is a common reference to the professional baseball team, the Samsung Lions. **Fig. 6** presents an article on Samsung's eSports team with a headline that appears to be about baseball, the reactions left behind by users through comments, and DRI results. The article is an interview with the members of Samsung White, a League of Legends (LoL) team sponsored by Samsung. The DRI of this article was -97%. As seen from comments such as "I thought this was about baseball" and "Is this a clickbait?" [*sic: this clickbait?*], many users were disappointed after reading because the content was not what they had expected from the headline.

3.3. Utilization of DRI: integration with recommended logic

Kakao applied DRI to its news ranking section and combined it with the current news recommendation algorithm. As mentioned earlier, Kakao implemented a personalized news recommendation algorithm with the launch of its PC portal news service in June 2015. This algorithm was based on CTR, just like other news curation algorithms. The company tested the news recommendation algorithm integrated with DRI in 2017.

The DRI-CTR algorithm led to a 4.5% increased duration per person and a 6.6% increased duration per view for recommended news when tested on news serviced via Daum Mobile. As

the algorithm recommends news of likely interest to users, the average duration per view naturally increased, which contributed to longer average duration per person. By taking user response into account for both headline and content, the DRI-CTR recommendation algorithm achieved a longer total duration. Based on the test results, Kakao has been using the DRI-CTR ensemble algorithm for its personalized news recommendation since September 2017.

4. Conclusion: implications and suggestion for futures works

This study introduced the challenges and impact of Kakao, which applied deep reading as a key factor to its news curation algorithm. Current news curation algorithms rely on user choice; such algorithms were developed based on CTR or PVs as the gauge for consumption frequency and user engagement. Most content services, not just those for news, have adopted curation algorithms dependent on user selection.

On the other hand, DRI focuses on the consumption time after content selection and not on the consumption frequency. To be specific, DRI calculates the expected duration for individual news based on a preliminary analysis of specific elements and then compares it to the actual duration that users spend.

The DRI is under the journalism spotlight as a novel attempt at measuring actual consumption patterns and attitudes, as current news curation algorithms involve a quantitative analysis of traffic and fail to consider the qualitative aspects. Against this backdrop, DRI can be seen as a means of qualitative analysis, providing insights into user satisfaction toward content.

In July last year, Kakao launched a separate DRI-based section for its users. The section introduces news articles that have engaged users through factors other than the number of views. As DRI measures satisfaction toward content based on duration, articles with high DRI values can be understood as more worthwhile to read. The results imply the effectiveness of DRI in providing users news articles of a higher caliber.

The study found that DRI values may serve as a basis for clickbait content identification, given the short actual duration in comparison to the expected duration. However, DRI is not a conclusive indicator of whether or not an article was intended to be clickbait. The algorithm is a quantitative analysis of consumption patterns; it merely looks at the difference between an expected duration derived from attributes and the actual duration. Content validation is required to accurately identify clickbait content. To do so, the algorithm would have to review the content and headline to understand the intention behind the article. This is far more complicated than determining whether specific criteria are met. Among existing recommendation algorithms, content-based filters focused on metadata examine the correlation between specific content and consumption patterns—it does not analyze the meaning or intent of suppliers. Algorithms that understand meaning are impossible to achieve with current technology and may remain difficult even in the future.

Expected duration, the main component of the DRI formula, is calculated from specific attributes, not content. In reality, the time spent on news content is not directly proportionate to the extent of deep reading because content difficulty is one factor that influences duration as well. As such, one weakness of DRI is its inability to analyze content.

Despite the above technical constraints, the adoption of DRI with its emphasis on time over consumption This experiment is expected to create a paradigm shift in academic and industrial realms. Existing news curation algorithms have been criticized for examining only the frequency of selection and neglecting content quality. Traditionally, news content

recommendation mostly relies on CTR, which consists of 0 and 1, not on ratings with a certain width of range, thus it is difficult to directly apply recommendation technics that are based on rich information [12].

In this sense, DRI overcomes the limitations of existing news curation algorithms. Kakao applied DRI to its main news curation in September 2017 after verifying its effectiveness and compatibility with the existing algorithm. The new approach of DRI is applicable to not only news but also other content curation industries. DRI, a venturing approach to news recommendation algorithm that this paper is introducing, shall be a part of such gradual improvement process in both academic and industrial field.

It is rare to find previous works that consider reading time as a key indicator in (news) content recommendation design, although many have recognized the inherently binary CTR measurement leads to noisiness [3]. Nonetheless, Zhang [13], who considered that book circulation time by library customers may represent relevancy of the borrowed books, provides us a clue that reading time as post-click behavior can deliver richer information about news readers preference.

This study presented the framework of an algorithm serviced by an internet company, the background of its implementation, and the related outcomes. It is unusual for an academic paper to unveil the details of an algorithm behind a currently active service. While algorithms are the key to business competitiveness in the artificial intelligence industry, exercising corporate social responsibility including the disclosure of algorithms is necessary for a business to build trust with society. However, this study may face criticism for insufficient disclosure; given that the company behind the algorithm is an ongoing concern, the algorithm could only be disclosed to a point where business competitiveness is not affected. It is hoped that this study will stimulate discussions that are essential for the development of the artificial intelligence industry.

References

[1] Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L. and Riedl, J., "Applying collaborative filtering to Usenet News," *Communication of the ACM*, 40(3), 77-87, 1997. Article (CrossRef Link)

[2] Viana, P., & Soares, M., "A hybrid approach for personalized news recommendation in a mobility scenario using long-short user interest," *International Journal on Artificial Intelligence Tools*, 26(02), 1760012, 2017. Article (CrossRef Link)

[3] Das, A. S., Datar, M., Garg, A., & Rajaram, S., "Google news personalization: scalable online collaborative filtering," in *Proc. of the 16th international conference on World Wide Web, ACM*, pp. 271-280, 2007. Article (CrossRef Link)

[4] Liu, J., Dolan, P., & Pedersen, E. R., " Personalized news recommendation based on click behavior," in *Proc. of the 15th international conference on Intelligent user interfaces, ACM*, pp. 31-40, 2010. Article (CrossRef Link)

[5] Aggarwal, C. C., Wolf, J. L., Wu, K. L., & Yu, P. S., "Horting hatches an egg: A new graph-theoretic approach to collaborative filtering," in *Proc. of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM*, pp. 201-212, 1999. Ariticle (CrossRef Link)

[6] Chu, W. Park, S. T., Beaupre, T., Motgi, N., Phadke, A., Chakraborty, S., & Zachariah, J., "A case study of behavior-driven conjoint analysis on Yahoo!: front page today module," in *Proc. of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM*, pp. 1097-1104, 2009. Article (CrossRef Link)

[7]   Chu, W. & Park, S. T., "Personalized recommendation on dynamic content using predictive bilinear models," in *Proc. of the 18th international conference on World wide web, ACM*, pp. 691-700, 2009. Article (CrossRef Link)

[8]   Li, L., Chu, W., Langford, J., Schapire, R., "A contextual-bandit approach to personalized news article recommendation," in *Proc. of the 19th international conference on World wide web*, pp. 661-670, 2010. Article (CrossRef Link)

[9]   Park, S., Sung, I., Seo, S., Hwang, J., Noh, J., Kim, D., "News recommendation service using machine learning: focusing on Kakao's RUBICS," *Journal of Cybercommunication Academic Society*, 34(1), 5–48, 2017. Article (CrossRef Link)

[10] Hirschman, A. O., "Exit, voice, and loyalty: Responses to decline in firms, organizations, and states," *The Journal of Finance*, 25(5), 1194-1195, 1970. Article (CrossRef Link)

[11] Van Dijk, W. W., Zeelenberg, M., "What do we talk about when we talk about disappointment? Distinguishing outcome-related disappointment from person-related disappointment," *Cognition & Emotion*, 16(6), 787–807, 2002. Article (CrossRef Link)

[12] Li, L., Zheng, L., Yang, F., & Li, T., "Modeling and broadening temporal user interest in personalized news recommendation," *Expert Systems with Applications*, 41(7), 3168-3177, 2014. Article (CrossRef Link)

[13] Zhang, F., "A personalized time-sequence-based book recommendation algorithm for digital libraries," *IEEE access*, 4, 2714-2720, 2016. Article (CrossRef Link)

Dongkwon Lee is a project manager and data analyst of media service team of Kakao corp. His current work focuses on a data-based approach to audience analysis of new articles and measurement of the value of news articles.

Daewon Kim received a PhD in Journalism & Mass Communication from Korea University. Dr. Kim wrote the first dissertation about robot journalism in South Korea. He worked as a journalist for about 10 years, and have been charged with designing straegies and policies on risk management in Kakao Corp. since 2016. His research interests include the diffusion of artificial intelligence, robotic journalism, mobile media, media business, media strategy, online to offline business and entertainment business.