Automated Journalism; A Study on Selection of News Stories Based on Individual Priority

Arya M R, Athira Sreekumar, S Dinesh Babu

Abstract: Automated journalism is a set of artificial intelligence algorithms that select thousands of news stories for a particular topic quickly, potentially, in minimum cost and without much errors as compared to human journalists. Automated journalism is most useful in generating routine news stories for repetitive topic for which clean, accurate and structured data are available. The paper aims to discuss the bird view of artificial intelligence and examine how these technologies are adopted in the Google news engine. And the paper also demands the criteria and process of selection of news stories to each of their readers in accordance with their personal priorities by analysing their previous Google search and most visited pages in internet. The author has employed quantitative analysis in this paper by monitoring different random samples about the relation between their previous Google searches and the news stories that has been selected automatically by the Google news for them.

Keywords: Automated journalism, Artificial Intelligence, Google news, Selection of news stories according to individual priority.

I. INTRODUCTION

The introduction of internet has changed the people's means of communication process as well as the way of gathering and providing information. The opportunities of communication and sharing process has been expanded simultaneously when the technology has enhanced and the phenomenon like participative web has emerged. This paved way for a revolution where users become more efficient to expand, enhance, share and customize innovative content in internet. However, the accessibility of information is one of the most significant ways of news search.

Ready to access information helps news consumer to choose news from their own preferred categories within a short span of time. But, in the present scenario the users have unlimited online sources of information and this will make them confuse that whether they are making a right reading choice. The automation has weaved a new form of communication experience for the news readers. With the help of automation in news consumers have the opportunity to get the news according to their preferences.

Automation is considered as an upcoming lucrative method in the field of journalism. At major news organizations, algorithms create thousands of news in every day without much human interference. Automated news has vastly expanded coverage in every field.

Google news is the greatest news aggregator and app powered by technical giant Google.inc available on all web platforms. It is a non-commercial website requires no registration and focuses on every news according to the reader's preferences. Google news engines have different reasons for automating news content. Digital revolution in news exerts one of the most powerful influence on reading behaviour and this impact will grow even more. They use automation to expand coverage in particular field, specialize their coverage, provide more regional variation sin customize their content to the readers. Industry leaders like Google hopes that automation will provide additional values to their readers and open up new streams of revenue. Google news is having a simplified interface and a combination of personalization and artificial intelligence. The core competencies of Google news have the feature of displaying all world news along with a particular option called 'For you', which comprises the sort out news according to each of the reader's personal taste.

When a user entering to Google news by signing in through their personal Google e-mail id. The automated system working behind Google news will automatically bring out news in favour to their personal taste, by analysing previous Google search and most visited pages. Every action of the user should be monitoring by this automated system and arrange news updates relating to the actions in the 'for you' option. It is a unique feature which absolutely absent in other news platforms. The feature of Google news will definitely help the readers to get acknowledgement about the things only they want to know. So the reader's task in finding out news will become easier and they can save their precious time.

II. LITERATURE REVIEW

Future of Automated Journalism

Ben De Jarnette (2016) stated that the rise of artificial intelligence and automation in journalism has been front and centre in the news lately. The technological advancement always simplifies human's works and likewise the introduction of automation in the field of journalism will help the human journalist to simplify their grunt works. As Tim Berners-Leo (2012) said data driven journalism will have a vital role in journalism in near future.

Arya M R, MJMC Post Graduation, Dept. of Visual Media & Communication, Amrita school of Arts and Sciences, Kochi, Amrita Vishwa Vidyapeetham, India.
Athira Sreekumar, MJMC Post Graduation, Dept. of Visual Media & Communication, Amrita School of Arts and Sciences, Kochi, Amrita Vishwa Vidyapeetham, India
S Dinesh Babu, Assistant Professor, Dept. of Visual Media & Communication, Amrita School of Arts and Sciences, Kochi, Amrita Vishwa Vidyapeetham, India

This will use techniques to convert available data more effectively. "Robonews – what journalists think of their new automated rivals" states that the automated system is more efficient in preparing news bulletins with minimum cost and with high speed (Thurman, N, 2018, September 19). In order to cope up with the technology as Ornebring (2016) said in this century one of the key factors in the development of journalism and media and rapid technological changes force journalist to upgrade their skills.

Algorithms of Automated System

According to Grint & Woolgar (1997) algorithms can be seen as knowledge machine that chooses information for users, but also a social process by which these formulas are made legitimate in the systems. Rather than mere creation of news stories automated systems are also capable for selecting information or news stories for users. When scientifically oversee the algorithms or the artificial intelligence is the working principle of the automated system or the automated journalist. When news agencies started using algorithms to automatically generate news from structured data it has shaken up the journalism industry most especially when Associated Press, one of the world's largest and most well-established news organization start make use of the technology.

The algorithms are able to generate news faster, at a larger scale, and potentially with fewer errors than human journalists. Algorithms can use the same data to tell stories in multiple languages and from different angles, thus personalizing them to an individual reader's preferences. Algorithms have the potential to generate news on demand by creating stories in response to users' questions about the data (Konstantin Nicholas Dörr, 2016).

As of the optimists, the automated journalism, the application of computer algorithms programmed to generate news articles, also known as robot journalism can be seen as an opportunity (Graefe 2016).

Automated Journalists over Human Journalists

As Carlson (2015) said that human journalists are comparably costlier, produce stories relatively slower and with a narrower breadth. The article 'Robots could replace journalists in the newsroom' states that a robot journalist called Dream Writer, designed by China, churned out 916 words Justin 60 seconds (Indian Express, 2015). The automated journalists are capable to produce routine news stories quickly, potentially with minimum cost and less cost. Automated systems are likely to produce error free contents as compared to human journalists. Thus the future media companies will hire automated journalists instead of human journalist.

The automated systems are not only works according to the input given by humans it can also imagine and search out things as like the human journalists.

The news produced by automated system never found to be unreliable and the readers cannot find out the difference between the news produced by human journalists or automated journalists. The content created through algorithm source had no effect on the credibility assessment. The inference included in this is that automated journalism can be said to be perceived in a credible manner,

irrespective of a by-line informing about the source (Montal and Reich, 2016; Thurman, 2017)

According to Kramer and Winter (2014) it has been found that users selects news only from credible sources even it is journalistic articles or online news sites. In their experiment, they found that participants not only select stories fromsources which were seemed to be more credible, 'but also selects from more frequently visited, read for longer, and selected earlier sites'.

After all, the readers receive combined automated and human journalism content in a news story. But no significant differences in credibility assessment of whether in its content or in its source could be find out by the readers (Graefe, 2016, Van Dalen, 2012).

Research Gap

Most of the previous research papers were focused on the ability of automated system in creating error free news contents quickly and without consuming much money. But in this research paper, the researcher observes the ability of automated system in bringing news stories based on each of the individual's priority. The automated system working behind the Google news portal selects news stories for the readers based on their preferences that their previous searches and most visited pages. By analysing reader's favourite categories the automated system brings news stories based on that category in the 'For You' page of Google news portal.

III. THEORITICAL FRAME WORK AND METHODOLOGY

The Search Engine Optimization theory has been used in this research paper. This theory will help to understand the working principle of automated engines in Google news which help to sort out news according to user's taste by analysing previous searches and most visited pages. As of the theory search engines organize the date they collect based on 'indexing criteria' and 'ranking criteria'. Indexing criteria means the search engine that means here Google will store the information searched for and how it handled. While ranking criteria consider the query terms in the titles, pages URLs, and bolded text that is, giving credits for use of query terms in link anchor text pointing to the page.

A search engine influences the results it serves to users by adjusting either the 'indexing criteria' or the 'ranking criteria'.

The Search Engine Optimization of search engine that means the Google is the reason behind Google news portal in bringing the news according to each of the user's choice.

Quantitative analysis is used in this research paper. For the quantitative calculation the researcher has conducted a 30-day long monitoring is randomly selected hundred samples. Each of the samples were asked to search for particular subjects like entertainment, celebrities, arts and culture, politics, gadgets, health and sports in the Google, which are considered as the variables of the study.

After every day the observer asked the respondents to check their Google news portal's 'For You' page in order to ensure whether the Google news has picked up news on the basis of their individual priority. That means, whether the Google news has brought news based on their previous searches.

Objectives

➢ To find out whether the automated system of Google news selects news stories based on individual priority
➢ To analyse how automated system identifies news based on user's preferences

Research Questions

➢ Is Google news is capable to bring news to users based on their individual priority?
➢ How automated system in Google news select news to each of the users based on their personal preferences?

Hypothesis

H₁ Automated system in Google news brings news stories based on each of the user's preferences.

H₂ Automated system selects news stories by analysing their previous Google search and regular visiting pages. Automated engines identify personal taste of each of the readers and make them easy to sort out favourite news for quick reading.

Research Design

Simple random sampling method has been used in this paper to prove the hypothesis that Google news brings news stories based on individual priorities by analysing previous Google searches and mostly visited pages.

100 samples despite of gender, educational and professional qualifications have been chosen for the study. The samples were regularly monitored for 30 days and marked the changes occurred in their Google news' 'For You' page. Each of the samples were given different topics in order to ensure that the automated system works behind Google news bring news by using Search Engine Optimization method, that is by analysing previous searches and most visited pages.

Total Number of Sample	Number of samples chosen by respondents in respective to variables						
	Entertainment	Celebrities	Arts and Culture	Politics	Gadgets	Health	Sports
100	11	20	9	21	14	16	9

IV. RESULTS

Day 1

In day one after searching for the given topic in Google search engine by the 100 samples, no changes have been found in the 'For You' page of the Google news portal. The Google news did not bring sorted news based on the user's priority or based on the previous search.

Day 2

In the Day 2 monitoring a considerable amount of changes did not found, even though few of the samples among the 100 got update regarding their previous search in Google search engine, that means based on their priority.
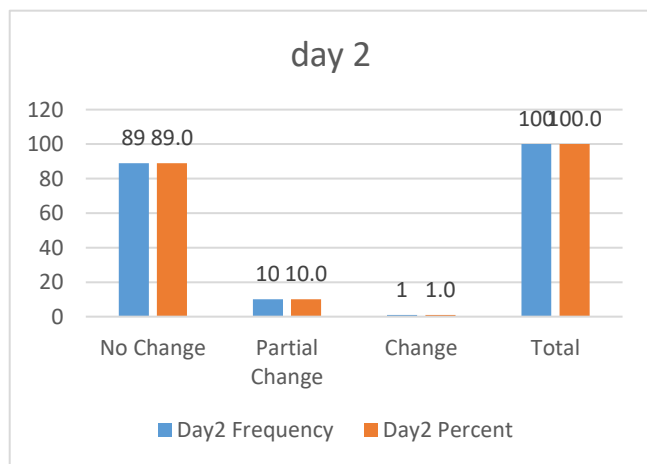

Fig. 1 Preferences of Respondents

The Fig: 1 shows that among 100 samples only few of the samples got update about their preferred category that means based on their searched topic in Google search engine.

Topics Preferred * Day2 Cross tabulation

		Day2			Total
		No Change	Partial Change	Change	
Topics preferred	Entertainment	11	1	0	12
	Celebrities	19	1	0	20
	Arts and Culture	7	1	0	8
	Politics	20	0	0	20
	Gadgets	10	3	0	13
	Health	15	2	0	17
	Sports	7	2	1	10
Total		89	10	1	100

Topics Given to 100 samples

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	15.844ᵃ	12	.199
Likelihood Ratio	12.755	12	.387
N of Valid Cases	100		

a. 14 cells (66.7%) have expected count less than 5. The minimum expected count is .08.

The Pearson Chi-square result of the table has been considered as the key result of Day 2. Thus the value of Pearson Chi-square is 15.844[a]

The expected count of the Chi-square is .08 if the hypothesis turned out to be correct. Here the result has come as .199 thus result of day 2 did not match to the hypothesis.

Day 3

In the Day 3 monitoring a partial changes have been found. A considerable numbers of samples got update based on their preffered topics.
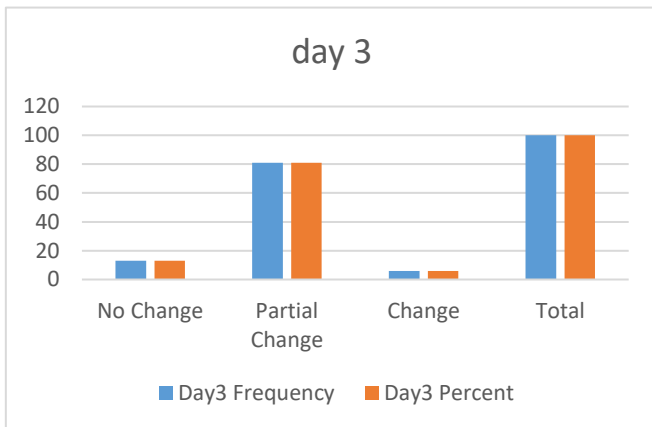

Fig. 2 Preferences of Respondents

Fig 2 shows that a considerable numbers of samples got news update based on their searches in Google search engine. The 'For You' page of their Google news portal got changed and started getting news of their choices.

		Day3			Total
		No Change	Partial Change	Change	
Topics preferred	Entertainment	2	10	0	12
	Celebrities	4	14	2	20
	Arts and Culture	2	6	0	8
	Politics	1	19	0	20
	Gadgets	0	11	2	13
	Health	3	12	2	17
	Sports	1	9	0	10
Total		13	81	6	100

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	12.348[a]	12	.418
Likelihood Ratio	16.128	12	.185
N of Valid Cases	100		

a. 14 cells (66.7%) have expected count less than 5. The minimum expected count is .48.

The Pearson Chi-square result of the table has been considered as the key result of Day 3. Thus the value of Pearson Chi-square is 12.348[a]

The expected count of the Chi-square is .48 if the hypothesis turned out to be correct. Here the result has come as .418, thus the result of day 3 show slight match to the hypothesis.

Day 4

In the Day 4 monitoring almost all the samples' 'For You' page of Google news portal has been changed and updated with news of their choice by analysing their previous searches in the Google search engine.
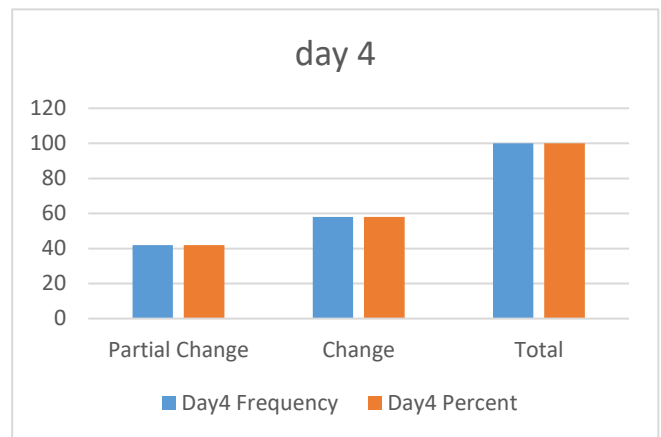

Fig. 3 Preferences of Respondents

Fig 3 shows that almost all the samples' start getting news of their personal choice. Google news began to sort out news based on previous searches.

Topics preferred * Day4 Cross tabulation

		Day4		Total
		Partial Change	Change	
Topics preferred	Entertainment	6	6	12
	Celebrities	8	12	20
	Arts and Culture	6	2	8
	Politics	2	18	20
	Gadgets	5	8	13
	Health	11	6	17
	Sports	4	6	10
Total		42	58	100

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	16.013[a]	6	.014
Likelihood Ratio	17.644	6	.007
N of Valid Cases	100		

a. 3 cells (21.4%) have expected count less than 5. The minimum expected count is 3.36.

The Pearson Chi-squareresult of the table has been considered as the key result of Day 4. Thus the value of Pearson Chi-square is 16.013[a]

The expected count of the Chi-square is .14 if the hypothesis turned out to be correct. Here the result also came as .14 that means that the hypothesis is correct.

Day 5

In the Day 5 monitoring all the 100 samples' 'For You' page of Google news portal got updated with their prefered news.
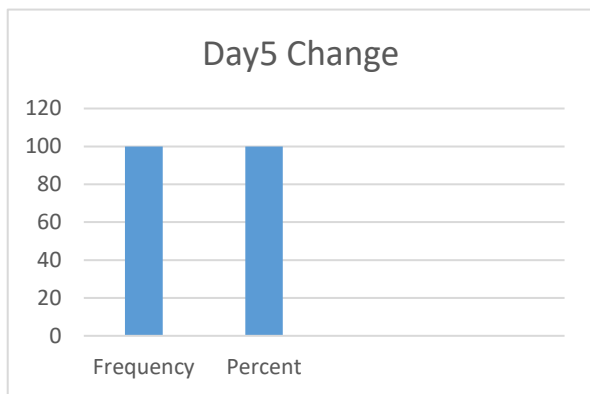

Fig. 4 Preferences of Respondents

Fig 5 shows that all the 100 samples start getting news of their personal choices. The Google news sort out news from their preferred category and updated the 'For You' page with news relating to the topics that they searched constantly in the past four days in the Google search engine.

		Day5 Change	Total
Topics Preferred	Entertainment	12	12
	Celebrities	20	20
	Arts and Culture	8	8
	Politics	20	20
	Gadgets	13	13
	Health	17	17
	Sports	10	10
Total		100	100

Chi-Square Tests	
	Value
Pearson Chi-Square	.[a]
N of Valid Cases	100

a. No statistics are computed because Day5 is a constant.

The Pearson Chi-square result of the table has been considered as the key result of Day 5. Thus the value of Pearson Chi-square is 'a' that means a constant.

If the Chi-square value become 'a', a constant for all the samples it proves that the hypothesis is correct.

The Google news bring news to users by understanding their preferred topic by analysing their previous Google searches and most visited pages. From the Day 5 to Day 30 analysis shown that the Google news start sorting out news based on individual's preferences by looking over the most searched topic sin the past four days. If a user search for a particular topic for constantly in four days, from the fifth day onwards the user will start getting news based on the search that means based on their personal preferences.

Thus Google news is capable to bring news stories based on each of the individual's personal priorities.

V. DISCUSSION

The research paper is to analyse the mechanism behind Google news in bringing news stories based on each of the user's priority. The automated system working behind the Google news helps the portal to provide news stories according to user's taste in their 'For You' page. The Search Engine Optimization in Google search engine stores and collects the history of user's every searches and on the basis of that analyse user's personal interested categories and sort out news based on that.

In the article 'AI's Rise in Journalism (And what it means for journalists)' Ben De Jarnette (2016) stated that the rise of artificial intelligence and automation in journalism has been front and centre in the news lately. Artificial Intelligence and automated journalism will be the most preferable in the next generation. The technological advancement helps readers to get updates of their favourite topics thus they couldn't a waste much time for that. Eventually people will change to the automated journalism and the future of journalism is lies there.

VI. CONCLUSION

Artificial Intelligence and automated systems are capable to simplify human works. The use of these modern technologies have raised the journalistic field as well. In the fast moving world where people do not have time to spend to update with world, the Google news helps them by sorting out news relating to their personal taste. Thus they do not have to spare time find out news according to their priority from the ocean of news from across the world. If a person only needs an update for sports category the Google news brings every update of the sports from all around the globe in the 'For You' page. With the theory of Search Engine Optimization Google search engine store the URLs, titles and content of user's every searches and analyse their favourite subjects. Then the automated system working behind Google news sort out news based on this analysis and show only their favourite category of news in the Google news portal's 'For You' page.

LIMITATION

The researcher has randomly chosen the samples for the monitoring of the research. Only 100 samples have been taken for the study. Most of the respondents were aged between 20 and 25 and it includes both professionals and students. The monitoring was limited in Kerala and mostly was in urban areas. And in variables also all kind of topics were not included.

261

REFERENCES

1. (n.d.). "Perception of Automated Computer-Generated News: Credibility, Expertise, and Readability,". Retrieved from https://www.researchgate.net/publication/289529002_Perception_of_Automated_Computer-Generated_News_Credibility_Expertise_and_Readability

2. Associated Press (2016) Home Depot can thank surging house prices for its stunning quarter. Available at: http://www.businessinsider.de/ap-home-depot-tops-street-1q-forecasts-2016-5?international=trueandr=USandIR=T&r=US&IR=T (accessed 10 July 2017).

3. Automated Insights (n.d.) the future of finance is automated. Available at: https://automatedinsights.com/ap (accessed 10 July 2017).

4. Carlson, M (2015) the robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. Digital Journalism 3(3): 416–431.

5. Computer-produced worded forecasts. (n.d.). Retrieved from https://journals.ametsoc.org/doi/abs/10.1175/1520-0477(1970)0512.0.CO;2

6. Did a Human or a Computer Write This? (2015, March 07). Retrieved from https://www.nytimes.com/interactive/2015/03/08/opinion/sunday/algorithm-human-quiz.html

7. Graefe, A, Haim, M, Haarmann, B. (2016) Readers' perception of computer-generated news: Credibility, expertise, and readability. Journalism. Epub ahead of print 17 April 2016. DOI: 10.1177-/1464884916641269.

8. Harcup, T (2014) Journalism. In: Harcup, T (ed.) A Dictionary of Journalism. Oxford: Oxford University Press, pp. 148–149.

9. Harcup, T (2015) Journalism: Principles and Practice. London: SAGE.

10. Mapping the field of Algorithmic Journalism. (n.d.). Retrieved from https://www.tandfonline.com/doi/abs/10.1080/21670811.2015.1096748

11. Robots could replace journalists in the newsroom. (2015, September 20). Retrieved from https://indianexpress.com/article/technology/science/robots-could-replace-journalists-in-the-newsroom/

12. Thurman, N. (2018, September 19). Robonews – what journalists think of their new automated rivals. Retrieved from http://theconversation.com/robonews-what-journalists-think-of-their-new-automated-rivals-73930