

## A Content Analysis on Man vs. Machine in News Selection

<sup>1</sup>S.Nagarathinam, <sup>2</sup>Dr. A. Palpandi and <sup>3</sup>Dr. A. Santha

<sup>1</sup>Research Scholar, M.S. (By Research), Anna University, Chennai, India - 600 025.

<sup>2</sup>Associate Professor, Department of Mechanical Engineering,  
Bharath Niketan Engineering College, Aundipatty, India – 625301.

<sup>3</sup>Former Professor and Head, Dept. of Journalism and Science Communication,  
Madurai Kamaraj University, Madurai.

### Abstract

*Google News* uses a computer robotic “algorithm” to select news for its front page without involving human editors. There is a need of journalistic approach to study the gate keeping process for selecting news items by the new “Robot Editor”. This can be compared with news items which were selected by human editors. The research plan of Man vs. Machine in news selection was born out of this idea. For this front page gate keeping study the researcher chose *Google News India* – the Indian version of the *Google News* and *The Times of India* front pages. The objective of the study is to look into the relationship between the selected 3724 news items of front pages through “IPTC” News Architecture. It also examines the composition of the two news media. The theoretical framework of this study is based on the Shoemaker’s hierarchy levels of gate keeping and the insights of Galtung and Ruge’s news values. This study found that the overall composition of *Google News India* front page differs from the composition of *The Times of India*, and more negative news was selected by Man.

**Keywords:** Content analysis, Gate keeping, News selection, Man and Machine.

### Introduction

The competition between Man and Machine is interesting and inquiring. Man's brain has creativity and has remarkable powers of retrieval, creative comparison and correlation, making new entities. Now machines (Computer algorithms or robots) do the process of News “Gate keeping” – the selection and rejection of news items.

*Google News* uses a computer mathematical “algorithm” (Das, et al., 2007), to select news for its front page without involving human editors. It updates news items frequently as news comes in. There is need of a journalistic approach to study the gate keeping process for selecting news items by the new “Algorithm Editor” (Kramer, 2003). This can be compared with news items selected by human editors working in the “old media” (Shoemaker & Vos 2009) like Newspaper. The research plan was born out of this idea. For this front page gate keeping study the researcher has chosen the front pages of *Google News India* – and *The Times of India*.

### **News Gate keeping**

“Gate keeping is a powerful process through which events are covered by the mass media, explaining how and why certain information either passes through gates or is closed off from media attention.” (Shoemaker, & Vos, 2009). Gate keeping is considered an intelligent activity of humankind. Now computers also do the job.

### **News for India from Google**

Google operates its offices (Appendix B) from various countries. The Indian edition of *Google News* is launched for Indian audience living in India and abroad. The gate keeping process also aimed to concentrate on India related events.

### **Importance of Front Pages**

Front pages were taken for this study for a number of important reasons. “The front page is seen first by most readers, a front page position for news item is an important display factor” (University of Minnesota, 1949).

Joseph Pulitzer believed that newspapers did not sell because of their reputations, political affiliations or actual content.

### **News Values and News Metrics**

During news selection process Mr. Gates adopts “News values” (Galtung and Ruge, 1965) and his media’s agenda and policy. These news values of Glatun and Ruge are: 1. *Threshold* 2. *Frequency* 3. *Negative* 4. *Unexpectedness* 5. *Unambiguity* 6. *Personalization* 7. *Meaningfulness* 8. *Reference to elite Nations* 9. *Reference to elite people* 10. *Consonance* 11. *Continuity* and 12. *Composition*. Man thinks before he opens up the gate for news to pass through. Mr. Robot Editor is set tuned by “News Metrics” of the organisation. *Google News* has 13 news metrics on various levels in controlling news selection. They are: 1.Number of Articles Produced by News Source 2.Average Length of an Article, 3.Story Size Scores 4.Value Representative of a Breaking News Score 5.Usage Pattern of news Source 6. Human Opinion of News Source 7.Circulation Statistics of News Source 8. Size of the Staff 9.Number of News Bureaux 10.Original Named Entities of News Source 11.Measure of Breadth 12.International Diversity of News Source and 13.Writing Style As like man considers values of his own organisation or an individual’s policy, the machine also has values for gate keeping.

### **Purpose of the Study – Aims and Objectives**

The purpose of undertaking this study is to see which news items pass through the traditional gate and the new robotic algorithm gate. It is timely and seeking knowledge on the issue of competency and relevance. The study will help us in the scholarship of understanding the gate keeping process of Man and Machine.

How do the gate keeping procedures function in the machine's news selection and man's news selection for the front pages? The primary objective of the study is to look into the relationship between the selected news items through the "IPTC" (Bacan, et al., 2005) News Architecture. It also examines the composition of the two news media.

The primary aim is to see the state of news judgment practices for the front page of the editors of *The Times of India* by identifying news values in their news selection for the categories. Similarly the study is to explore news values built in the machine algorithm. Another aim is to examine the various forces at the gate keeping process for the Top Stories in the front page of *Google News India*.

Gate keeping study always deals with the analysis of "similarities and differences in news selection" (Shoemaker and Vos, 2009). This study analyses hyperlinked top story headlines of front page (or home page) news selection of *Google News India*, with news items selected for the front page of *The Times of India*. The study throws light on the possibility of the use of "Robot Editor" (Kramer, 2002) in the future news rooms along with Human Editors.

### **Hypothesis**

**Hypothesis 1 (H-1):** Significant differences exist between Man (*The Times of India*) and the Robot of Algorithm Machine (*Google News India*) in routine (Shoemaker & Vos, 2009) news selection based on the IPTC categories.

### **Research Questions on Composition**

#### **RQs on H-1**

**RQ 1A:** Does the overall composition (Galtung and Ruge, 1965) of Google front page differ from the composition of *The Times of India*?

**RQ 1B:** Who selects more Negative (Galtung and Ruge, 1965) News?

### **Review of Literature**

#### **Previous Studies – Gate keeping**

The Austrian psychologist Lewin (1947) was the first person to use the term "Gatekeeper". He primarily liked to produce suitable social change through various gates or portals existing between the farm and the family table. He studied how and why the food habits and activities of families differed.

Although the term gatekeeper originated with Lewin, David Manning White first applied it directly to journalists. According to White (1950), quoting Lewin, the traveling of a news item through certain communication channels was dependent on

the fact that certain areas within the channels functioned as gates. White called the telegraph editor "Mr. Gates".

Warren Breed (1956) defined 11 considerations editors use when they take gate keeping news items. He studied news determinants about 55 years ago; but many of his suggestions are still relevant.

Gieber (1956) amplified the number of editors in the study. He included 16 editors of newspapers. Moreover, he found (1956) a few difference across the editors in news selection.

Gieber (1964) identified that the personal influence on the selection of subjects was less in gate keeping than other factors including the available number of news items, the size of news items and deadline pressure and the mechanical production of newspaper.

Paul Snider (1967) did the same White's study as "Mr. Gates Revisited" after 17 years". It was important in the area of gate keeping theories. During his study industrialization boomed and the circulation of the same newspaper of white's study also had grown up.

Tunstall (1977) asserted that every country has three media levels. International news items, local news items and national news items.

Mc Quail (1983) asserts that an alternative explanation to that of subjective individual judgment is to be found in the concept of news value, which is an attribute of a news event that transforms it into an interesting story for an audience. Hetherington (1985) presented news values from his own experience. Allan Bell (1991) proposed some additions to Galtung and Ruge. Chang and Lee (1992) found newsworthiness as something that is timely, has impact or consequence, human interest and conflict.

Harcup and O'Neil (2001) were interested in how adequately their news values could be applied to foreign and domestic events, issues and other stories that become news. In addition, they identified the shortcomings of Galtung and Ruge's news criteria and proposed: 1. The Power Elite, 2. Entertainment, 3. Celebrity, 4. Surprise, 5. Bad news, 6. Good news, 7. Magnitude, 8. Relevance, 9. Follow-ups, and 10. Media agenda.

McQuail (2005) explains the mechanism of control is not usually policy or law, or even economies, but the audience demand for their own media content in their own language. This influences the gate keeping process in many ways.

The shoemaker and Vos (2009) model of hierarchy levels of gate keeping includes five levels. 1. The Individual level (Characteristics, demography, experiences, values and attitudes of individuals force gates) 2. The Routine level (Deadlines, news values such as timeliness, proximity, interests or related including inverted pyramid), 3. The Organizational level (Ownership patterns of the media organization). 4. Social Institutional level (Influences from Govt., advertising or other interest groups) 5. The Social System level (Ideology, custom, culture of the society).

Singer (1998) studied hyperlink journalism and found that content in news websites were predominantly staff generated. She drew an analogy between gate keeping and hyper linking. Kovarick (2002) explored further and argues that the strength of the hyperlinks helps readers understand the subject elaborately. Dimitrova.

D., et.al (2003) found that online newspapers use hyperlinking as a tool of gate keeping. Williams and others (2004) point out that *Google News* is unique in that it is primarily constructed of only hyperlinks, offering minimal information on its own home page. Cutting D.R.,et.al.,(1993), Hearst M. A., et.al.,(1996), Leuski and allen (2000) studied the problems of clustering (Google News uses clustering algorithm to crawl news items) algorithms.

## **Methodology**

This empirical research into news selection compares the contents of two news media, one of print and the other online for one year. Stratified Random Sampling was used. For headline content analysis, the study included 3724 news stories, from a set of 25 weeks, for analyzing the gate keeping function involved in the stories. Stratified random sampling is a procedure which first categorizes a population into subgroups (e.g. from this study: Indian Elections, the Lankan war, the Swine Flu episode, the Death of Jackson, the Naxalites war series, etc.) and then randomly selects from each subgroup until a desired number is reached. "In this way, researchers are able to obtain large enough samples for each subgroup for statistical analysis" (SETDA - Glossary, 2010).

## **Selecting News Media**

The Basic instinct of the study is comparing Man and Machine in news selection. So for the part of the machine, *Google News India* is an obvious choice for this study. *The Times of India* Newspaper was selected on the basis of its being the oldest Indian newspaper in English with the largest circulation and highest number of editions and news bureaux across India. These three combination are unique.

## **Story Analysis**

For this study, the front pages of *The Times of India* were used to the compare front pages of *Google News India*. The hyperlinks of the Top Stories section alone are taken for the study, as the remaining part of the home page has the category wise sections, which are equivalent to the inside pages of Newspapers. The complete sample consisted of 3724 news, all of which the researcher read carefully and coded. These articles represented the final sample for the study. The Content analysis of headlines was used to identify the subject of news.

## **Units**

The subjects of headlines of both the front pages were the units of analysis. All information was coded on the basis of front page headlines and their news subject.

## **Time Periods**

The sample selected was published between 2009 April and 2010 March. The Time period of the study was one year. 25 weeks were selected from this period for this study. It almost represents a normal news cycle of one year. Thus 172 *Times of India*

Front pages and 172 *Google News India* homepages were selected for the study. The total numbers of news pages were 344

### **Coding Topics**

No coders were used to identify the category of each news items. So the researcher himself coded the 3724 news units. And the researcher completed this task as soon as each week closes.

The Front page stories were coded into 17 categories. Choosing news for a particular category is a trivial. Sometimes the decision may change from person to person. The researcher chose the categories from IPTC which are common to international news agencies and these were defined by the International Press Telecommunication Council (IPTC). The sub categories are also adopted from the IPTC. It has provided numerous categorization schemes aimed to standardize the coding of various aspects of news related metadata.

### **Statistical Analysis**

Descriptive analysis was used to derive the frequency of each category.

The Chi-square or the Fischer's exact test was used to compare proportions among groups. Fisher's exact test is a statistical test used to determine if there are non-random associations between two categorical variables. Values were expressed as percentages.

The Pearson correlation analysis was used to assess the relationship between *Google News India* and *The Times of India* with reference to the number of publications. The correlation between two variables reflects the degree to which the variables are related. Values are presented in terms of r value with the p value. Pearson's correlation reflects the degree of linear relationship between two variables. It ranges from +1 to -1. A correlation of +1 means that there is a perfect positive linear relationship between the variables.

The Independent sample t test was performed to assess the relationship between *Google News India* and *The Times of India* and region. The Independent Samples t test compares the mean scores of the two groups on a given variable. All tests of significance were 2-tailed and p values less than 0.05 were considered statistically significant.

### **Results**

According to the Audit Bureau of Circulations (ABC), *The Times of India* is India's largest selling Newspaper and *Google News India* is unique in its service. In 25 week sets from the one year period in 2009-2010, a total number 3724 news units found in the front page of *The Times of India* and in the homepage top stories of *Google News India* were analysed for the study.

### Percentage Distribution of News by Man and Machine

Percentage distribution of news in *The Times of India* as well as *Google News India* and both are presented in Figure 4.7. The figure it shows that a majority of news (50.9%) are published by *The Times of India*, while 37.7 per cent was on *Google News India* and common news published by both *The Times of India* and *Google News India* was only 11.4 per cent.

### Research Questions on Composition

RQ 1A asked whether the overall composition (Galtung and Ruge, 1965) of *Google News India* front page differs from the composition of *The Times of India*. Table 1 shows the Independent Sample T –test – on the relationship between News selection and the category of news published. The Overall composition of *The Times of India* and *Google News India* are statistically tested and the result tabled below.

**Table 1:** Difference Existed in Majority of Categories in Overall Composition

News category	News selection	N	Mean	SD	T test	P value
Arts, Culture and Entertainment	TOI	23	4.74	2.988	-0.617	0.541
	Google News India	25	5.20	2.062		
Crime, Law & Justice	TOI	25	13.12	4.076	11.879	0.001
	Google News India	23	2.87	1.359		
Disasters and Accidents	TOI	23	2.87	1.456	0.572	0.573
	Google News India	12	2.58	1.379		
Economy, Business and Finance	TOI	25	9.16	3.671	-2.696	0.010
	Google News India	25	11.76	3.126		
Education	TOI	24	3.83	1.711	4.220	0.001
	Google News India	14	1.86	1.167		
Environmental Issues	TOI	17	1.35	.606	-3.845	0.001
	Google News India	15	2.93	1.486		
Health	TOI	24	3.79	2.637	-1.373	0.177
	Google News India	25	4.68	1.796		
Human Interest	TOI	24	4.83	2.160	2.907	0.007
	Google News India	13	2.77	2.006		
Labour	TOI	16	2.19	.981	2.282	0.031
	Google News India	11	1.45	.688		
Lifestyle and Leisure	TOI	3	1.00	.000 <sup>a</sup>	-	
	Google News India	2	1.00	.000 <sup>a</sup>		
Politics	TOI	25	18.36	6.350	2.434	0.020
	Google News India	25	14.84	3.460		
Religion and Belief	TOI	12	1.33	.651	0.663	0.517
	Google News India	6	1.17	.408		
Science & Technology	TOI	11	1.55	.820	-3.002	0.007
	Google News India	16	3.13	1.857		
Social Issues	TOI	15	2.13	1.885	2.329	0.035
	Google News India	7	1.00	.000		
Sports	TOI	25	6.56	2.518	2.737	0.009
	Google News India	25	4.80	2.000		
Terrorism, Conflicts and War	TOI	25	4.12	1.986	1.218	0.229
	Google News India	24	3.46	1.817		
Weather	TOI	10	1.30	.483	1.964	0.081
	Google News India	4	1.00	.000		

Table 1 shows the independent sample t test between categories according to news selection by man and machine. From the table 1 it was observed that arts, culture and entertainment were published 5.20 times (average frequency) higher in *Google News India* compared to *The Times of India* (mean=4.74). Similarly crime law and justice was published 13.12 times in *The Times of India* but in *Google News India* this was only 2.87 times ( $p < 0.001$ ). However, news related to economy, business and finance was published 11.76 times higher in comparison with *The Times of India* (mean=9.16,  $p = 0.010$ ). Similarly sports (mean=6.56), social issues (mean=2.13), politics (mean=18.6 vs. 14.84: *Google News India*) and Terrorism, conflicts and war (mean=4.12) was highlighted more in *The Times of India* compared to *Google News India*. While in *Google News India* science and technology (mean=1.55; 3.13: TOI), environmental issues (mean=*Google News India*: 2.93 Vs. TOI: 2.35) and health are highlighted more in *Google News India*. The overall composition (RQ 1A) of *Google News India* front page differs from the composition of *The Times of India*.

### Man vs. Machine – Composition

The difference between man and machine seen in news selection for the categories of IPTC in both side. The routine affects the content selection.

Among the overall 17 categories significant differences existed in news selection between the man (*The Times of India*) and machine (*Google News India*) in 10 categories of: 1. Crime, Law & Justice 2. Economy, Business and Finance, 3. Education, 4. Environmental Issues, 5. Human Interest 6. Labour 7. Politics 8. Science & Technology, 9. Social Issues and 10. Sports. So the overall composition differs between the two media.

The difference happened because, man selects more percentage of news items for the categories of Crime, law and justice, Education, Human Interest, Labour, Social Issues, and Sports, and machine selects more percentage of news items for Economy, Business and Finance, Environmental issues, Politics and Science & Technology. So man selects more percentage of news items for six categories and machine selects more percentage of news items for four categories. So this makes compositional difference between man and machine.

News selection by man and machine differs in ten categories in overall comparing. Among these categories, sub-categories of four categories' news selections of man and machine also differ. They are: 1. Crime, law and justice, 2. Economy, Business and Finance, 3. Environmental issues and 4. Politics. No significant differences i.e. similarities were found in the rest of the categories: 1. Arts, Culture and Entertainment, 2. Disasters and Accidents 3. Health, 4. Religion and Belief, 5. Terrorism, Conflicts and War, and 6. Weather.

News selection by man and machine does not significantly differ in 6 categories in overall comparing. Among these categories, sub-categories of 2 categories' news selections of man and machine do not significantly differ. They are: 1. Weather and 2. Terrorism, Conflicts and War

The category Lifestyle and Leisure did not attract many news items for their front pages. So testing did not yield significant results. Most of news items seen here have come from the routine channel. The study also proves Sigal's study on *the New York*



*Times* and the *Washington Post*, which revealed that news items from routine channels are different.

Saleem Abbasi, an Algerian who had an Interpol red corner alert pending for his alleged involvement in a 1992 bomb blast in the Algiers airport, was arrested at the Chennai airport on June 18, 2009. This was reported in *The Times of India*. But it was missing in *Google News India*. Though it is a potential international news item, it did not enter and clustering enough to pass through the algorithm gate. A few days later (24 June 2009), again *The Times of India* reported his free walk as the Interpol withdrew the alert. No follow-up was reported in *Google News India*. Similar incidents happened when a 9 year old Russian girl was raped on the beach in Goa (23 January 2010), when a Nigerian was produced in that Chennai adalat for an alleged involvement of online fraud (27 January 2010), and when two businessmen were found murdered in Fiji (18 January 2010).

If news has an International, National flavor and was about the popular elites *Google News India* gate opens for the items. For example, Sarabjit's Death review petition (25 June 2009), attacks on Indians by Australians (24 June 2009), alleged crime involved in Michael Jackson's death (27 June 2009), Bollywood actor Shiney Ahuja's rape case (19 June 2009), Satyam's Ramalinga Raju scandal trial follow-up (8 September 2009), Anand Jon's 59 year sentence for sex crimes (2 September 2009), Mumbai attack terrorist Kasab's trial ( 24 July 2009).

The odd news items and major judgments including, a woman could marry rapist verdict (08 March 2010); Three convicted for the 2003 Mumbai twin blasts (28 July 2009); and follow-ups of the high profile Nithyanandha sex scandal. With the "routine" of wider popularity and the reports were continuously in the media coverage are not missing in *Google News India*.

But in the machine's routine, we can find insignificant items for the Indian fans (Yousuf returns Pakistan fold, 23 June 2009, Stakes is high as the first test starts, 16 November 2009), for front page and the machine does not aware whether the match was played by Indian players or not (No surprise as England confirm Ian Bell's starting place (with Australia), 26 July 2009. What does the Spain win mean for the USA 2010, 25 June 2010). Conceivably this may have happened as the metrics set by the engineers to display a sports item for the front page composition in a daily basis, so it takes an item on daily basis, not giving importance to the item

Table 2 proves that the higher percentage of Negative news items was chosen by Man and not Machine. Pearson-chi square for trend distribution was 108.723 (DF=2) with a p value 0.000

### **Selection of Negative News Items**

The following Table 2 proves man chooses more negative news for *The Times of India* front page than the machine selects for its top stories. RQ 1B asked who selects more Negative News (Galtung and Ruge, 1965). More negative news was selected by *The Times of India* than *Google News India*, particularly in relation in to the three categories of Accident and Disasters Crime, Law and Justice & Terrorism, Conflicts and War ( $p < 0.0000$ ) [Table 2].

**Table 2:** Selection of Negative news by both news media

Category		Selection			Total
		TOI & Google	TOI	Google	
Accident and Disasters Crime, Law and Justice & Terrorism, Conflicts and War Categories	Count % within Selection	130 30.7%	496 26.2%	179 12.7%	805 21.6%
Other Categories	Count % within Selection	294 69.3%	1400 73.8%	1225 87.3%	2919 78.4%
Total	Count % within Selection	424 100.0%	1896 100.0%	1404 100.0%	3724 100.0%

A significant trend was observed between the categories with the Pearson-chi square for trend distribution being 108.723 (DF=2) with a p value 0.000. The minimum expected count is 91.65.

The results revealed that man is interested more in choosing negative news items, and machine favors more news on Computer and Mobile Technology and other items of the Economy, Business and Finance category for the Indian audience. Another important finding of the study is that Cricket has emerged as an important news determinant for the front pages in the Indian context. Except, sub-categories of politics and economy, business and finance category, the 3Cs (Crime, Cricket and Cinema) are news values for the Indian context. And this finding is the contribution of this dissertation to news values of Galtung and Ruge and Harcup O'Neil. The man and machine news selection equally concentrated on politics. More or less one fourth of news was politics. But within the "politics" category, the items of news selected for the sub-categories differed.

The biggest compositional difference between man and machine falls in this category, where 12.6 per cent of news selected by man was rejected by the machine. So the difference was significant and it was reflected in the Negative news selection of man (Table 2). This reminds us of Shoemaker's (1996) suggestion that the human brains are 'hard wired' to prefer information about threats and change.

### Summary of the Results

The table summarizes the answers to the research questions and the results of testing the hypotheses in this study.

	Research questions and hypotheses	Results
	<b>Hypothesis 1</b>	
<b>RQ1A</b>	Is the overall composition (Galtung and Ruge, 1965) of <i>Google News India</i> front page different/diverse from the composition of <i>The Times of India</i> ?	This study found that the overall composition of <i>Google News India</i> front page differs from the composition of <i>The Times of India</i>
<b>RQ1B</b>	Who selects more Negative (Galtung and Ruge, 1965) News?	More negative news was selected by Man - <i>The Times of India</i> than Machine - <i>Google News India</i> , particularly related to Disasters and Accidents, Crime, Justice and Law, Terrorism, Conflicts and War ( $p < 0.000$ ).

The comparison was done on the basis of selected news items of both sides and common news items between two. These news items indicate the difference existed in the Regional Proximity, Cultural Proximity, Negative news selection, Week End news selection and Source selections. This study falls under the communication routine influence (Shoemaker and Vos 2009), of hierarchy levels of influences of the gate keeping.

As earlier studies also suggest that newspapers prefer negative news, here man prefers negative news items (RQ1B) for all the regions– World, India, State and Local. But the machine is more interested in the techie items including mobile and computer related technology or business.

The findings suggest that the communication routine (2009) influenced the content. And the findings also suggest that the Human Editor is more relevant or more peculiar than the Robot Editor of *Google News India*.

The structure of the editorials can be modified in future with the accommodation of Mr. Robot Editor along with human editors. So the size of the editorials can be economized. In future we may witness the man – machine interface in the editorial department.

The result was not in favor of the machine as the researcher assumes that the algorithm is a knowledge migrant of the USA to India. During its migration, it kept its American news values and customs of USA. The customization of this algorithm needed revalidation according to the strict news values for the country because India has several languages, religions, customs and regional priorities, which need more attention in news selection. So the local and state news was not even passing through by the machine. *The Times of India* was distinctive in its coverage; it could restrict news flow for the front page according to the events. This type of flexibility is not seen in the algorithm, so it lags behind.

The metrics should be validated with the actual news outcome of news items and should be fine tuned for better selection. For this, the Google Team should study the news values to a particular region.

Breaking news items and rotation of news headlines from various news sources should be studied in detail – as the LTTE Leader Prabhakaran dead news did not appear the next morning. This was very important news in the Indian context. The metrics cannot keep very important similar headlines until they reach audience.

## References

- [1] Bacan, H., SP, I., & Gulija, D., 2005. Automated News Item Categorization. *Proceedings of the 19th Annual Conference of The Japanese Society for Artificial Intelligence* Nagoya: Springer-Verlag, pp. 251- 256.
- [2] Breed, W. ,1956, “Analyzing news: Some questions for research.” *Journalism & Mass Communication Quarterly*, 33(4), 467-477.

- [3] Chang, T. and J. Lee, 1992, "Factors Affecting Gatekeepers' Selection of Foreign News: A National Survey of Newspaper Editors," *Journalism Quarterly* 69(3): 554–561.
- [4] Cutting D. R., Karger D. R., and Pederson J. O., 1993, "Constant Interaction-Time Scatter/Gather Browsing of Very Large Document Collections, " *In Proceedings of the 16th Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval (SIGIR'93)*," Pittsburgh, PA,125-135.
- [5] Das, A. et al, 2007, August 5, "Google News personalization: scalable online collaborative filtering. *International World Wide Web Conference*," Conference conducted from Association for Computing Machinery, Alberta, Canada.
- [6] Dimitrova, D.; Connolly-Ahern, C.; Williams, A.P.; Kaid, L.L.; Reid, A., (2003). Hyperlinking as gate keeping: online newspaper coverage of the execution of an American terrorist, *Journalism Studies* 4(3), pp.401-414.
- [7] Galtung, Johan and Ruge, Mari, 1965, "The structure of foreign news: the presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers", *Journal of International Peace Research* 2(1), pp. 64-91.
- [8] Gieber, W., 1956, Across the desk: A study of 16 telegraph editors. *Journalism Quarterly*, 33(4), 423–432.
- [9] Gieber, W, 1964. News is what newspapermen make it. In L.A. Dexter & D.M.White (Eds.), *People, Society and Mass Communication*. New York: Free Press.
- [10] Harcup, T., & O'Neill, D., 2001, What is news? Galtung and Ruge revisited. *Journalism Studies*, 2(2), 261-280.
- [11] Hearst M. A., Pedersen J. O., 1996, Reexamining the Cluster Hypothesis: Scatter/Gather on Retrieval Results. *In Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'96)*, Zurich, June.
- [12] Hetherington, A., 1985, *News, newspapers and television*. Macmillan.
- [13] Kramer, Staci D. 25, September 2003, Online Journalism Review. Retrieved from <http://www.ojr.org/ojr/kramer/1064449044.php>
- [14] Kovarick, Bill., 2002, *Web Design for the Mass Media*, Boston: Allyn & Bacon.
- [15] Leuski A. and Allan J. *Improving Interactive Retrieval by Combining Ranked List and Clustering. Proceedings of RIAO*, College de France, pp. 665-681, 2000.
- [16] Lewin, K., 1947. Frontiers in group dynamics, II: Channels of group life; social planning and action research, *Human Relations*, 1(2), 143–153.
- [17] Mcquail, Denis., 2005, *McQuail's Mass Communication Theory*, Fifth Edition ed. Thousand Oaks, CA: Sage Publications Ltd, 2005. 252. Print.
- [18] Mcquail, Denis., 1983, *McQuail's Mass Communication Theory*. First Edition ed. Thousand Oaks, CA: Sage Publications Ltd.
- [19] Shoemaker, P. J., & Vos, T. P., 2009, *Gate keeping Theory*. First Edition ed, New York: Routledge.

- [20] Snider, P. B., 1967, "Mr. Gates" revisited: A 1966 version of the 1949 case study.
- [21] Singer, J.B., 1998, "Online Journalists: Foundations for research into their changing roles". *Journal of Computer Mediated Communication*, 4(1).
- [22] Tunstall, Jeremy., 1977, *The Media are American. Anglo-American Media in the World.* London:Constable.
- [23] University of Minnesota.,1949, Newspapers and Their Readers: Case studies of four Minnesota newspapers. *Language Arts & Disciplines*, 2, 340.
- [24] White, David Manning ,1950, 'The 'Gatekeeper': A Case Study in the Selection of News, *Journalism Quartely*, 27, 383-390.
- [25] Williams, A. P., Martin, J. D. and Trammell, M. R. , 2004-05-27 "*Google News, the Web Sphere, Hyperlinking, and Gate keeping:*" Paper presented at the annual meeting of the International Communication Association, New Orleans Sheraton, New Orleans.

