dodging the memory hole 2016
saving online news

journalism
digital
news
archive

jdna

oct 13–14, 2016
ucla library
los angeles

INSTITUTE of Museum and Library SERVICES | rji DONALD W. REYNOLDS JOURNALISM INSTITUTE Missouri School of Journalism University of Missouri | UCLA Library | MU Libraries University of Missouri | EDUCOPIA INSTITUTE

DtMH2016: Scholarship Project

Eliminating the Border of Digital News Archiving:
Obstacles and Solutions of Online News Archives for International Media

Jiwon Choi

Ever since the time the first newspaper was produced and published, it has played a significant role in recording the human history, witnessing what we have done and what has made us to do that (Mills, 1981)**.** As a source of humankind's accumulated knowledge and history, news contents have been considered valuable resources to preserve and protect. Newspapers are also fundamental sources of current information on contemporary issues and are essential for building current awareness and helping the public to be well informed, facilitating teaching and research in many academic fields including journalism (Pujari & Pai, 2005). In this regard, news should be well preserved for future generations as well as ourselves living in this contemporary world. It is fair to say that failure to archive born-digital news represents an abdication of responsibility towards an important part of the world's cultural heritage (Lynch & Fontaine, 2014).

With the advent of the Internet and technological advancements in the last couple of decades, news has been coming in various formats besides the traditional platforms such as print, radio, and television (Cheney et al, 2011). Especially the increasing role of the Internet as a means to distribute and access information has changed the way news is produced, consumed and, preserved, challenging the existing press ecosystem. It is not difficult to imaging in the future that existing papers migrate online entirely and their digital trail is hardly captured (Lynch & Fontaine, 2014, p. 17).

In addition, this new media environment has facilitated globalization in journalism. News organizations have been able to publish news stories for their global audiences with no time and space constraints thanks to the new technologies. A number of previously domestic news publications have started to issue their international versions and established local branches across the world including The New York Times, CNN, and The Korea Times, for instance. This has led international publications to run different local websites, produce various localized stories, and publish news in many languages. All theses factors, however, have made it more challenging for the global news companies to develop and operate systems to effectively archive news contents (Lynch & Fontaine, 2014).

Despite the importance of preserving web-based contents in international media companies, most of prior research dealing with the news archiving issue has been based on monolingual, domestic or local papers. This means that there has been lack of academic approaches on the issue of born-digital news archives from an international perspective. Furthermore, most of international publications appear to have trouble in establishing and managing effective, integrated digital news archiving systems because of the high operating cost and technical difficulties and other problems. The Korea Times, for example, has their global branches mainly in South Korea, America, and Canada, publishing news in two languages, Korean

and English. Their four websites for each language and country, however, do not provide the same, integrated news archive for their readers to find news articles.

In this regard, this project will look into the current crisis in born-digital news archiving in international publications, focusing on possible impediments that obstruct international media in establishing and managing their archives for online contents. This project aims to bring the attention from the journalism industry and news consumers to preserving digital news produced by international publications and to examine obstacles and challenges that make operating international news archives demanding to come up with prospective solutions to the problem.

Method

This project is done mainly with three parts: A literature review that looks through previous academic studies regarding the issue; a discussion session with international students of the University of Missouri to obtain opinions from each country's perspective; and a survey conducted from Dec. 5, 2016 through Google Forms (https://goo.gl/forms/sfoT2T4xHYzOy8Sz2) with questions asking personal experience and view towards the issue.

As for the discussion, six international students from the US, the UK, Burkina Faso, South Korea, and China participated and shared their ideas on Nov. 10, 2016.

[From left to right: Jonah McKeown, Madi Linglingue, Tammy Ko, and Jing Ren are participating in the discussion to share their ideas about the international media's archiving issue.]

[A poster for the promotion of the discussion distributed to international students of MU through the International Center of the school]

Panels:

- Jiwon Choi, South Korea, Journalism
- Tammy Ko, South Korea/USA, Pre-Journalism
- Jing Ren, China, Journalism, Reporter at Columbia Missourian
- Jonah McKeown, England/USA, Journalism
- Madi Linglingue , Burkina Faso, Geography
- Sean Na, South Korea, Journalism, Reporter at Columbia Missourian

Agenda:

1. Personal experiences regarding the issue of archiving/retrieving digital news contents

2. Analysis of the current situation and problems of archiving/retrieving digital news contents

3. Suggestions for possible, desirable improvements

Changing Environment

For the last couple of decades, there has been an overwhelming amount of information coming in digital platforms every day and since the boom of the Internet starting from 1990s a lot of publications have been publishing their news electronically (Schranz et al, 2002)**.** Therefore, in order to successfully preserve news in this new media age it is no longer enough to just save contents from traditional newspapers and broadcasting stations.

The "lifecycle" of news contents, which means how the news is gathered, processed, and updated, has fundamentally changed and it requires the field of journalism to devise effective strategies for preserving news in the emerging electronic environment (Alverson et al, 2011). Cheney et al (2011) pointed out because of the changes in news contents, news formats become increasingly complicated and virtual, involving more visual elements such as photos and video, which usually take up more space on storage.

The latest and most influential emerging platform, social media, also has become more pervasive in news industry and affected how stories are produced, published, and viewed, requiring new ways of archiving news contents based on the platform (Cabulea & Frey, 2011). Besides technical changes that the preservation of digital news contents is faced with, a wide range of copyright and legal issues has caused difficulties in providing solutions of successfully harvesting news webpages.

As part of an effort to fit into this new era of journalism, some leading countries in archiving digital media contents have been moving ahead with innovated approaches to the preservation of the materials. Canada, for example, recently made an amendment of its legal deposit regulations to accommodate the emergence of online publications and to address the concerns raised by digital technologies that potentially impede access (Geist, 2007). In contrast to this, however, a study of Canada's current situation in preserving Web news contents argues that the Canadian news industry has

gone through shifts that has left it ill-equipped to focus resources and attention on digitization, addressing that the problems attendant on born-digital archiving are universal around the world (Oury, 2012).

In addition to the legal actions, there have been some prototypes that specifically aim to establish digital news archiving systems: The OmniPaper project conducted by four partnered European countries and the OAIS model that has been accepted worldwide as the basic model for setting up digital archiving (Schranz et al, 2002; Wijngaarden, 2006).

The objective of the OmniPaper project is creating a multilingual navigation and linking layers on top of distributed information resources, providing an in-depth approach to manage international news archives with strong semantic networks. The OmniPaper project adopts the self-learning system that has three different layers to effectively process input information, showing an exemplary case of establishing one integrated news archive that operates with different languages (Schranz et al, 2002). Distributed source data exists in various standard formats, different languages and with varying depth of available information. The existing local archives are connected to the OmniPaper system, using its metadata to successfully retrieve requested information (Schranz & Paepen, 2004). Similarly, the OAIS reference model has six different entities that set up

a long-term archiving architecture: Archival storage, ingest, access, administration, data management, and preservation planning (Wijngaarden, 2006).

This successful examples that have been possible thanks to the effort from various countries and international organizations opens the possibility for international media companies to work together to establish integrated news archives. The two examples not only prove the possibility of cooperation between countries to manage news archives together but also present possible solutions to the problem of multilingualism in news archiving.

Still, there are some questions that the projects have not been able to resolve; they were done in European countries whose languages have similar structures and vocabularies, which make developing an integrated archiving system easier than other multilingual cases such as between Chinese and English. Besides, a number of possible threats have been presented in the OAIS model: Media failure that causes the system to be subject to sudden catastrophic irrecoverable loss of bulk data and communication errors that refers to the system's not being able to manage and deliver contents within a specified time period (Rosenthal et al, 2005). Above all, there has been no official evaluation that tells how successful the models have been.

Opinions from Discussion

Tammy: I personally think that the industry itself has grown not into the information industry; but into the money industry. And right now so much is changing and so much information is coming every second. Stakeholders in the news industry might tend to think that archiving old information is not as important as producing new stories.

Jiwon: A lot of news contents are updated every second and there could be conflicts with old and new information. During the DtMH forum, there was a presentation that tells that kind of problem, showing that information updated is not updated on every website it is on. This is another problem that we have to solve.

Sean: Columbia Missourian only has archives from 2001. One of the reasons why they don't keep archives before then is that the local people do not usually care much about what happened back in 90s or 80s. They care more about timely news, what is currently affecting them.

Challenges to Archiving Online International News

Multilingualism/Technical Difficulties

A massive amount of information including news is currently available in different languages aside from English, the major language of use. Although resources on the Internet are accessible and sharable over geographic distance, language barriers still exist and are not easy to eliminate (Lee et al, 2003). This multilingualism has made it more arduous for international publications that run their news websites in more than two languages to build integrated news archives that give their readers coherence of the same topic.

There have been many attempts to solve this problem including the case of a news archiving application named NewsExplorer. It displays a daily clustered view of the major news items for each of the 19 languages covered, performs a long-term trend analysis, and offers entity pages showing information gathered in the course of years for each entity, including person titles, multilingual name variants, reported speech quotations, and relations. Languages that are covered by the application include 14 European Union languages plus Arabic, Farsi, Norwegian, Russian, and Turkish.

Along with NewsExplorer, other news gathering engines, MedISys and NewsBrief, set good examples of accomplishments of third-party news archiving companies as well (Gey, 2009). The annual Cross-Language Evaluation Forum(CLEF) is another example that works for improve multilingual information access. The forum has accumulated more than 12 million news docs in 12 languages including Bulgarian, Dutch, English, French, Finnish, German, Hungarian, Italian, Portuguese, Russian, Spanish, and Swedish by adopting archiving systems such as Ad-Hoc and Question Answering (Peters, 2007).

While few studies have been conducted on the preservation of multilingual news contents, most of the prior research on the topic of multilingualism has been done with the cases in digital libraries. The findings and results from the previous studies done with libraries, however, are still useful to devise solutions for multilingual international newspapers because both of the cases have the language issue in common.

Cross-language information retrieval (CLIR) refers to the circumstance in which a user tries to search a set of documents written in one language for a query in another language. Especially after the World Wide Web boom in the 1990s, it has become more important, to provide users with access to information resources written in a variety of languages on the Internet (Kishida, 2005). CLIR usually takes place in multilingual digital libraries and news archives of international media and requires translation or crossing the language barrier, which might cause errors in translation and problems of ambiguity and synonymy (Diekema, 2012).

Moreover, CLIR issues that have been academically experimented are restricted to certain language groups that share the same cultural background or have similarities in their structures such as TREC(Spanish, Chinese, German, French, Italian, and Arabic) or NTCIR(Japanese, Chinese and Korean.) To avoid these problems, alternative methods have been suggested such as the pivot language approach, in which an intermediate language acts as a mediator between two languages for which no bilingual resource is available (Kishida, 2005).

Cross-Language Information Access(CLIA), an extended version of CLIR, is one of the content managing methods for international news publications as well as digital libraries because it helps integrate diverse archives based on multiple languages into a single, effective archiving system (Chen & Ruiz, 2009). The Internet Public Library Asia(IPL-Asia) is a similar case that offers a gateway to collect high quality Internet resources expressed in culturally related languages, Chinese, Japanese, and Korean (Lee et al, 2003).

In spite of all the effort made to develop multilingual information managing systems, none of the technologies have been widely applied to existing digital libraries to enable digital library users to access information across languages. Moreover, very few digital libraries have realized multilingual information access, based on the result of an analysis of 150 digital libraries in the United States. The study found out that only five of them could be accessed by using more than one language (Chen & Ruiz, 2009).

Another study done in Pakistan, where the development of multilingual libraries is in the infancy stage, investigates the language barrier and failure of developing proper archiving systems on library websites based on the data collected from

133 digital libraries in the country. The research results revealed that only 20% of users used their multilingual archives, 22% of total users were unaware of the system's existence, and only 18% of them found the archives useful. (Manzoor, 2012)

Opinions from Discussion

Madi: I know there are some issues like the capacity for storage is limited. But with new technologies like cloud, it should be easier to archive contents online.

Jiwon: Do you speak English in your country?

Madi: No. We speak French. But there are also newspapers in local languages.

Jiwon: If you want to search information about local issues, what language do you use?

Madi: French, mostly.

Jiwon: Then for people who do not know French, do you have any translation system that interprets local languages into French?

Madi: Not really, as far as I know.

Jiwon: Have you had any difficulties finding information in your country because of the language problem?

Madi: Not because of the language but because of the accessibility. Much of the digital contents in my country are not really archived well. Contents from

governmental institutions are relatively better preserved than ones from private companies, though. Very little part of the information is stored in Burkina Faso.

Jing: In case of Xinhua news agency in China, they have different branches around the world, running their websites in many languages. I would say they are all under the same control of the government but the way they frame stories is completely different according to what language they use. When they report exactly same story in English and Chinese, they frame it in totally different ways in order to reach different target audiences. Chinese version aims to serve the Chinese audience; and it is more government-toward and favored. On the other hand, the English version, which is operated around the world, is more government favored so that it can be used as a way of promoting the country's positive side to the rest of the world. Reporters are writing the same stories but in different ways. It affects how they archive news contents as well. The contents they have been preserving are separate because they are using different approaches.

Ownership and Awareness Issues

Along with technical and multilingual problems that international news media has, the issue of who should establish and operate archives has been controversial among related stakeholders including news publications, libraries, and third-

party companies. In Canada, where government and foundation money has become more scarce, archiving digital news contents have relied on public-private partnerships to secure funding, which probably ends up archiving materials that might otherwise not be preserved. This has raised questions of who sets the rules for such partnerships and whose interests should be presented in archiving news contents (Oury, 2012). In such cases, it is preferable to set up common policies and infrastructure for third-party organizations to deal with the preserved materials (Ball, 2010).

Lack of awareness about the conservation and preservation of digital heritage is one of the challenges that we have encountered regarding the digital news preservation problem. It cannot be emphasized enough that he future of preservation depends on archivists' ability to keep raising awareness of preservation needs, to keep persuading society to invest in the prevention of significant losses, to keep processing our preservation backlogs with today's proven techniques while working to develop even better ones, and to integrate the preservation of digital with traditional resources (Cruzat, 2006).

In order to raise public awareness of the problem of digital preservation, Portugal has been using the OAIS model as a successful tool (Verheul, 2006). Many professionals in digital archiving have attended the training sessions organized at The National Library of France(BnF) as part of their attempts to help people recognize the importance of the issue (Laurent, 2006)

Opinions from Discussion

Jing: There are some problems with censorship in China. I have interned at one of the news organizations in Beijing, which is the capital of China. So the city was one of the cities where the Chinese government's censorship is most intense. What I experienced was whenever something related to Beijing or the country that is not positive, it is automatically considered negative. So right after we post something like that online, the government institution dealing with online matters would order us to delete the contents. It is really hard for Chinese citizens to obtain proper information from archives online because of the censorship issue.

In China, we have very vague concept of copyright. Even news organizations copy each other's stories without any ethical consideration. There was a case with CCTV which is an official media of the Chinese government. CCTV reported a disaster, saying how many the death toll was, which later turned out to be wrong. Then, all the other news organizations trusted the official media and copied the wrong information without any fact checking process. We do not have the concept of copyright in China generally.

Madi: I am wondering about the legal issues such as how to make sure the contents are preserved well

and properly. Sometimes I saw in Central Africa, there was a demonstration. Before the one in my country happened, newspaper in France took a picture from Burkina and transferred it to Congo without permission. So I think it is also important to think about who should do the archives and what contents need to be preserved?

Jiwon: Have you used any third-party news archiving website? What is your opinion on the third-party news archiving system?

Jonah: They could be more secure because if all the newspaper store all their stories on their own server something could happen to that. Or worse, their building can burn down.

Financial Difficulties

Digital news data generally have a large portion of various multimedia platforms including images and sounds, which requires bigger storage than text-only contents. The more the news industry is digitalized, the more money is needed to construct and operate archives, which can discourage development of news archiving systems. A decreasing amount of budget for establishing a news preservation system the Center for Research Libraries of Canada is regarded as one of the main reasons that have led the country's archiving industry fail (Lynch & Fontaine, 2014).

Considering the bulk and fragility of news data and the resource-intensive nature of the indexing and cataloging work, news archiving services comes with a significantly high cost. For a long time that cost was underwritten by local, federal and philanthropic grant making, as a public good (Alverson, 2011). Information in digital form is much more vulnerable to interruptions in the money supply than information on paper. There are ongoing costs for power, cooling, bandwidth, system administration, domain registration, and so on. Budgets for digital preservation must be expected to vary up and down, possibly even to zero, over time. Therefore it is necessary to trade off the cost of establishing and managing archives among interested groups including governments, libraries, third-party archiving companies, and news publications (Rosenthal et al., 2005).

Opinions from Discussion

Sean: In order for a newspaper to make printed version from 1950s to 2010s available, it takes huge amount of effort. It means in case of local newspaper, they need to hire someone who is specialized to do the job.

Jiwon: To talk about the financial part of the issue, there are actually some third-party companies that sell contents of their news archives at a certain price to the audiences who need the information. That can be one of the solutions to the financial problem.

Sean: In case of Columbia Missourian, we once had an application on mobile. But they decided not to retain the app after 2011 just because it took too much money. It was a big challenge for them. Big publications like The New York Times, Wall Street Journal, or Chicago Tribune may be able to afford the cost of establishing news archives and running mobile applications, but for small newspaper, it is not feasible. We should also consider other than large, prestigious media because people here in Boone county, for example, care more about local news rather than national ones. So we should take that into consideration.

Jonah: At least, it is going to be a lot easier going forward because pretty much every newspaper starts out as a pdf now. So they don't have to scan it in. All they need to do is to save it online right after they're done making it. Even where I went to undergrad, we were easily able to archive because each page started out as a pdf already. The files were posted on our school website, which is quite easy. The school's website was hosted on the university's server. In a way, it could be incredibly easy to archive because all you really have to do is save pdfs to your hard drive. And they are all archived.

Madi: Do you have any idea on how much it would cost to operate news archives?

Jiwon: It depends. During the forum, they did not actually explain how much it would cost to do that, but what they told is that there are a lot of

possibilities that they can make money through those archives.

How to Establish Effective Archiving Systems?

A large amount of previous research has suggested what the future scholars should focus on for establishing proper international, digital news contents archives: Usage of metadata of the preserved materials and cooperation between stakeholders (Lynch & Fontaine, 2014).

Metadata refers to semantically related texts that have information such as authors, dates of creation, and brief introduction of files, helping organize preserved materials easily without looking into all of them. A majority of the discussion on digital preservation has focused on the metadata rather than the content itself, for example on what metadata should be preserved along with the content and on standards for representing it (Rosenthal et al, 2005) It is because the metadata is the material that is used when indexing and retrieving information, making it a lot easier for archivists to find relevant contents. Especially for multimedia data such as images and PDFs, using accompanying metadata helps the process of categorizing and searching desired files (Alverson et al, 2011).

CIMWOS(Combined IMage and WOrd Spotting) is a good example of utilizing metadata

for archiving multimedia, multilingual contents including news articles. The system incorporates an extensive set of multimedia technologies, integrating three major subsystems (text, speech, and image processing), producing a rich collection of metadata annotations (Hatzigeorgiu, 2004). As one of the early models of news archiving systems, the way CIMWOS generates text-based metadata from non-text resources can be employed other archives with cross-platform media contents. Since most of born-digital news contents include multimedia resources, it is significantly important to understand how multimedia platforms work with different archiving strategies as in the case of CIMWOS.

Canada, which was once a leading country in the digital contents archiving industry, has been struggling with the current impasse with lack of financial support, copyright concerns, and technological hurdles. Among the obstacles that the country has had, failure of stakeholders to cooperate to launch a born-digital newspaper archiving program has led the country's corruption in the archiving industry (Lynch & Fontaine, 2014). Since the born-digital news preservation issue deals with technology, legal, financial, and journalistic aspects altogether, collaborative research should be done with helps from various fields. Below are the examples of international works that have been conducted by relevant organizations and professionals (Verheul, 2006, p. 205-206):

- LOCKSS (Lots of Copies Keep Stuff Safe): A large multinational effort developed by Stanford University, with participation from libraries and publishers, to distribute software and build communities of practice to enable the safe archiving and preservation of e-journals and other web-published content.

- IIPC (International Internet Preservation Consortium): One of the major goals of IIPC is the development of tools for web archiving. At the moment IIPC is working on finalizing an IIPC toolkit for acquisition, selection and storage, which will incorporate the IIPC standards.

- ICABS: Although BL is participating in ICABS (IFLA CDNL Alliance on Bibliographical Standards) its current focus is not on digital preservation, as is the case for the Koninklijke Bibliotheek and the National Library of Australia.

- PLANETS: BL is currently coordinating a project proposal for the fifth call of the Sixth Framework Programme of the European Commission. Under the working title 'PLANETS' a group of European libraries, archives and universities is preparing a project proposal. PLANETS stands for: Preservation and Long-term Access through NETworked services).

Opinions from Discussion

Jiwon: Do you feel any need for an integrated archiving system?

Jing: I think that has to be discussed with financial issues that we've talked about because it needs a lot of money as well as human resources. I feel like they can choose certain major languages such as English, Spanish, or some other languages. They don't have to include all the languages since the number of people who speak other minor languages is relatively smaller than the number of speakers of major languages. I personally think that it's good to have their own archives too, but it's more like an ideal thing to do that is hard to be realized.

Jonah: Yes, maybe they should focus on creating a really good software that recognizes texts so they don't have to transcribe everything. I'm sure that it takes a lot of time and money to scan the pdf files. But if they can at least make some kind of program good enough to recognize the texts well enough you could at least search for stories. Scanning the headlines only even could be helpful for them.

Jiwon: I do agree with the idea. I think the TimesMachine that The New York Times has been creating wouldn't take that long, for example. There could be an effective mechanism that can be used to transcribe pdf files of news articles.

Jonah: It could also be done through university. It does seem like a lot of work to archive all of the articles. But as I've done before, it would be a lot easier to just type some keywords to get the full context of a certain topic with proper archives. It

helps to build the context and get better history for a story. Building archives is an investment upfront but it would be helpful students, researchers, and reporters if they have research projects that need some old news articles, for example. It would be whole a lot easier for them to do that with archives. I don't know how exactly expensive it is to archive something.

Results from Survey

An anonymous survey titled "Survey on your experience about international publications' digital news archiving" was done online from Dec. 5, 2016. Nine respondents from five countries(South Korea, Burkina Faso, Sigapore, USA, China) have answered to the survey through Google Forms.
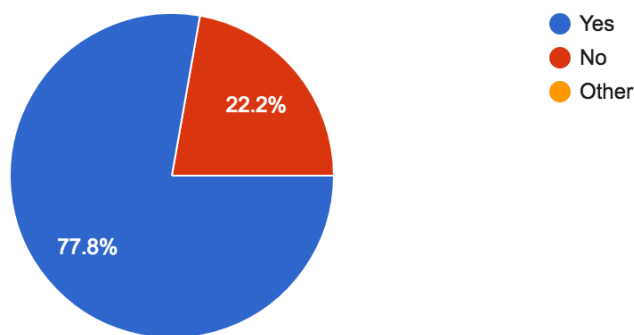
Following are the results of the survey:

1. What do you think of the main reason(s) that causes the problems? (choose as many as applicable)

77.8%(7 respondents) answered that different languages are the main reason; unintegrated archiving systems follow next(66.7%, 6

respondents). Technical issues and general indifference about the issue rank the third with 33.3%(3 respondents), while publications' financial situation was selected the last reason with 22.2%(2 respondents).

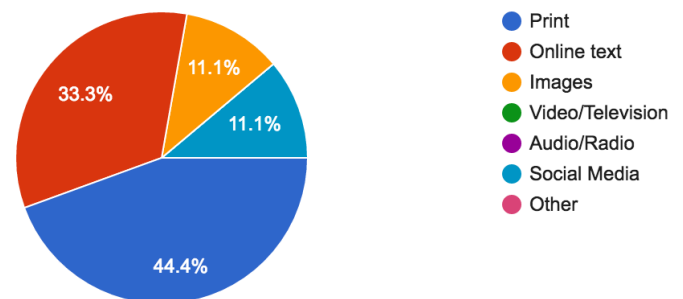**2. Do you think it is needed for multilingual international publications to have one, integrated archive?**
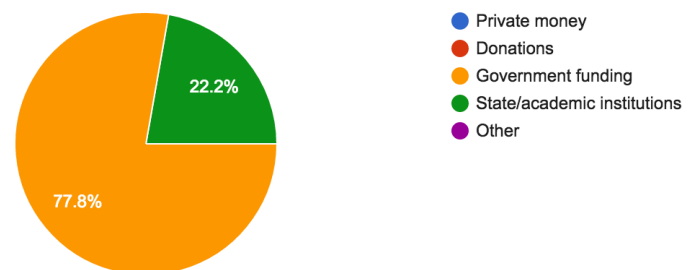

Reasons:

- Academic and educational reasons, plus it could help with journalism and public record keeping.
- It will be much easier to get access to.
- readers will not have to read many versions of the same information to get the full of it since the content may vary from one version to another. This way, readers will find the information more reliable
- I think it is better for multilingual publications to have separate archives so that they can keep the different languages apart. That way, readers will not have to sift through multiple languages to find a story.
- More people can read and understand clearly

- It is needed for better networking
- To globalize the world
- The multilingual publications should have one archive for each languages to increase efficiency(while looking up the articles in specific languages)
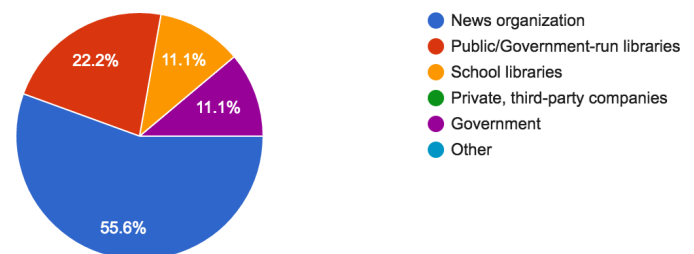- Simplicity is the best for information seekers.

3. Which form of news contents do you think most important to archive?
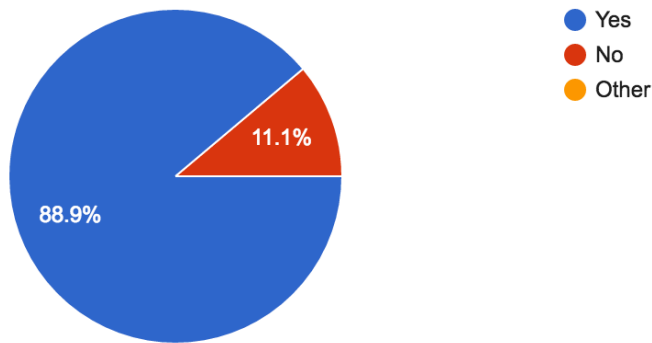

4. Who should pay to maintain these archives?

5. Who should operate news archives?

6. Do you think it is harder for you to retrieve online news contents from multilingual international media than domestic, single-language based ones?

Reasons:

- Not being able to search for the content in your native language
- Language problem
- Internet coverage is limited to some places in many countries. Also, some international websites cannot be accessed from overseas
- You need to spend a bit more time and energy looking through more stories that you wouldn't read
- Difficult to search
- It is a different workload that needs to be handled
- Not native language.
- Most of the national issues are handled and reported by domestic organizations
- There is no reason to say no if sources are correctly translated.

Conclusion

As for academic studies on the issue, although the issue of archiving born-digital news data has been increasingly getting attention from the field of journalism as well as from the news audiences, not much academic research has been done on it from the journalistic perspective. Especially for international publications, which have been available thanks to the development of digital technologies and diverse new media platforms, it can hardly be emphasized too much that preserving online news data is critical not only for their readers but for recording human history of this globalized world. Since the issue has recently been raised in the industry, more studies should be conducted to call attention to the problem and to devise proper models for establishing and operating digital news archives for international news organizations.

According to the discussion and survey with international students of the University of Missouri, more awareness about the problem of developing and operating archives for digital news contents is desperately needed since most of them said that they did not know much about the issue itself. The panels participated in the discussion generally agreed that we need to consider who should establish and run archives, how we finance them, and why we need to archive news contents. Most respondents answered to the survey saying that they need news archives and it would be good for them

to have integrated, effective archive systems to search information from international publications. Realizing the need for establishing effective news preserving methods, stakeholders, professionals, and scholars in the journalism and archiving industry should look for solutions for international publications to archive digital, multilingual news contents successfully to help their readers around the world retrieve information from them easily in the future.

Reference

Alverson, J., Leetaru, K., McCargar, V., Ondracek, K., Simon, J., & Reilly, B. (2011). Preserving News in the Digital Environment: Mapping the Newspaper Industry in Transition. *CRL Mapping the Newspaper Industry*, 1-74

Ball, A. (2010). Web archiving. *DCC State of the Art Report,1.1*, 4-34. Retrieved December, 2016, from http://www.dcc.ac.uk/sites/default/files/documents/reports/sarwa-v1.1.pdf

Chen, J., & Ruiz, M. (2009, July). Towards an integrative approach to cross-language information access for digital libraries. In *Proceedings of SIGIR 2009 Workshop: Information Access in a Multilingual World, Boston, Massachusetts, USA*.

Cheney, D., Palsho, C., Cowan, C., & Zarndt, F. (2011).The Future of Online Newspapers. *Proceedings of the Charleston Library Conference,* 68-75. doi:10.5703/1288284314878

Cruzat, X. (2006). The never ending challenge disasters and preservation in chile. *International Preservation Issues,7*, 22-28.

Diekema, R. (2012). Multilinguality in the digital library: A review. *The Electronic Library*, 30(2), 165-181. doi: 10.1108/02640471211221313

Geist, M. (2007). Preserving printed and digital heritage. *BBC News*. Retrieved from

http://news.bbc.co.uk/go/pr/fr/-/2/hi/technology/6287181.stm

Gey, F., Karlgren, J., & Kando, N. (2009). Information access in a multilingual world: Transitioning from research to real-world applications. *ACM SIGIR Forum,43*(2), 24-28. doi:10.1145/1670564.1670568

Kishida, K. (2005). Technical issues of cross-language information retrieval: A review. *Information Processing and Management: An International Journal - Special Issue: Cross-language Information Retrieval,41*(3), 433-455. doi:10.1016/j.ipm.2004.06.007

Laurent, J. (2006) le plan d'urgence de la bibliothèque nationale de France. *International Preservation Issues,7*, 39-43.

Lee, W., Sugimoto, S., Nagamori, M., Sakaguchi, T., & Tabata, K. (2003). A Subject gateway in Multiple Languages: A Prototype Development and Lessons Learned. *International Conference on Dublin Core and Metadata Applications, 0*, pp. 59-66. Retrieved from http://dcpapers.dublincore.org/pubs/article/view/734

Lynch, L., & Fontaine, P. (2014). Preserving the Unpreservable: Form, Content, Copyright and the Archiving of Born-Digital

Newspapers. *International Symposium on Online Journalism,4*(1), 4-22.

Manzoor, A. (2012). Challenges and Opportunities of Multilingual Library Websites: Evidence from a Developing Country. *International Journal of Business and Management Invention,1*(1), 1-15. Retrieved December, 2016, from http://www.ijbmi.org/papers/Vol(1)1/A110115.pdf

Mills, T. (1981). Preserving Yesterday's News for Today's Historian: A Brief History of Newspaper Preservation, Bibliography, and Indexing. *The Journal of Library History, 16*(3), 463-487. Retrieved November 30, 2016, from http://www.jstor.org/stable/25541211

Oury, C. (2011). When press is not printed: the challenge of collecting digital newspapers at the Bibliothèque nationale de France. *proceedings of the IFLA Preconference, newspaper section*.

Pai, P., & Pujari, A. (2005). Newspaper Archives - A Knowledge Asset. *International Conference on Information Management in a Knowledge Society,* 21-25. Retrieved February, 2005, from http://hdl.handle.net/10760/8276

PETERS, C. (2007) Multilingual information access: the contribution of evaluation. In *proceedings of the 2006 international workshop on Research issues in digital libraries*. Kolkata, India, ACM. Doi: 10.1145/1364742.1364761

Rosenthal, D., Robertson, T., Lipkis, T., Reich, V., & Morabito, S. (2005). Requirements for Digital Preservation Systems : A Bottom-Up Approach. *D-Lib Magazine,11*(11). doi:10.1045/november2005-rosenthal

Schranz, M. & Paepen, B. (2004) Architecture Design and Application for an Intelligent Distributed News Archive, *The ICCC 8th International Conference on Electronic – ELPUB*, 1(1), 159-170.

Schranz, M., Paepen, B., & Tscheligi, M. (2002). Omnipaper: Modern Approaches For An Intelligent European News Archive. *ICWI,* 495-500.

Verheul, I. (2006). *Networking for digital preservation: Current practice in 15 national libraries*. München: K.G. Saur.

Wijngaarden, H. (2006). Making (preservation) plans together: planning strategies for digital preservation in the Netherlands and Europe. *International Preservation Issues,7*, 116-122