Integrated News Publishing Platform for GFR Media

Edward F. Camacho Pacheco
Master of Engineering in Computer Engineering
Alfredo Cruz, Ph.D.
Electrical & Computer Engineering and Computer Science Department
Polytechnic University of Puerto Rico

Abstract —*GFR (Grupo Ferré Rangel) Media is the local leader in news with its newspapers sites "El Nuevo Día" and "Primera Hora" which are constantly bringing to the audience news reports about local and international issues. GFR Media's sites receive millions of visits a month, making their websites one of the most visited sites in Puerto Rico. GFR Media also features smaller special websites that deal with specific topics of common interest to the audience that need to be published. In order to quickly bring news that can be accessed across many websites, GFR Media needs to implement a system that allows its journalist and editors to reliably upload news content to be used in those special websites. Those special websites usually showcase particular stories in a different, interesting and dynamical way which can include multimedia assets such as: embedded videos, photography, or social networks embedded codes. The developed system "News Publishing Platform" provides a meaningful tool for the editorial needs allowing journalists and editors to publish news content.*

Key Terms — Application Programming Interface, Content Management System, HTML / CSS, Model View Controller.

INTRODUCTION

In today's world, journalists need to investigate and report constantly to their audiences the most recent news and events in order to stay in the cutting edge of news reporting. To do that, journalists need to be able to publish news content to many websites and devices at the same time. The information provided in this paper will explore various topics related to the newsroom technologies and the problems that journalists face. Due to the nature of journalism and the way it has embraced technology there is a constant need for evolution. This evolution are redefining the place of professional journalism in the news information systems [1].

Currently the publishing tools and options that journalist and editors have are limited and too expensive and cannot be expanded or adapted to be reused in other systems. With profitability still being the ultimate goal, some newspaper businesses have transformed from printed form to an online-based news source [2]. Those online-based news sources require a great flow of information that must be processed on the spot and must be constantly updated, and the stories therein reported rewritten. If there is a large number of news being accessed in those online-based sources, the information being scrutinized will be greater and the journalists or editors will have less time to use multiple complex systems [1]. To meet the demands of this fast paced environment a solution must be developed. That solution is the developed "News Publishing Platform" system which enables journalist to deliver content quickly and with less setup time. The developed project solution, which targets the journalist's publishing needs, includes three main components: Content Management System (CMS), Application Programming Interface (API), and a Web Site (that will consume the content from the API). The proposed "News Publishing Platform" will be used for many special website projects created by GFR and will be used as the standard for the creation of featured projects where news or another content should be quickly published to a website.

NEWSROOM AND TECHNOLOGY

After the onset of the Internet, newspapers are realizing they will encounter a financial crisis if they

do not keep up with new technological competitors [2]. Newspapers businesses are widely adopting the technology as an important news delivery platform. In order to properly understand the future of technology on journalism, one must understand how it grew to where it is today. The past tells us that technology causes journalism to change. In the mid-nineteenth century, the telegraph allowed news publishers to gather and send information across the globe faster than ever before, very similar with what is happening now with the internet.

Newspapers Websites

By the early twentieth century, the technology could create photographic images of people and events that readers had never been able to create before [3]. Later the arrival of print machines to deliver printed newspaper copies and recently the arrival of internet highlight the need to be constantly evolving in the journalism environment.

Figure 1
NYT Ads + Digital Subscriptions

Figure 1 above illustrates the recent quarterly growth of digital subscriptions and ads for the New York Times and how they have kept increasing with the passage of time. In contrast the green bars demonstrate clearly how printed newspaper distributions continue to diminish at a fast pace and newspaper companies are really seeing this trend.

According to the Pew Research Center, Figure 2 demonstrates that the preeminence of websites has already surpassed that of TV and Radio. This means that digital subscriptions have grown each business quarter and that printed ads or copies of newspapers

have been steadily declining more and more each day. Because of the importance of the internet the project "News Publishing Platform" represents a very important asset that will keep up with this evident technology trend. Editors and journalists need to be able to do their jobs proficiently while using the right web tools on the internet. This need has become central key component of a journalist's daily routine. One of the most profound changes associated with the exponential growth of web 2.0 technologies is a clear shift in communication control and journalistic practices [4].
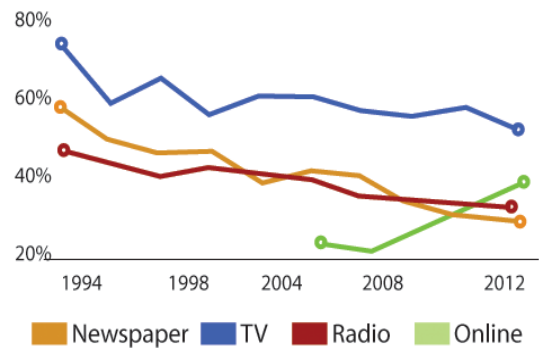
Figure 2
Percentage of Americans who got News from the following Sources

The Need to Quickly Inform

In the past forty years, most afternoon newspapers have disappeared, as people's primary news sources have shifted from paper, to television, to the Internet, and our ways of understanding the news has completely changed. As Giles and Snyder explain, "We cannot take the future for granted, it is very difficult to determine what is around the next bend in the road especially in a field as integrally connected with changes in culture, politics, technology and economics as journalism." [3]. A persistent fact around technology is that once a new technology arrives, most of the times that technology is further developed and widely used trough many actualizations and enhancements. Is important to know that as technology advances the journalism change with it in order to maintain relevant to the users. When new user's devices as tablets or watches arrives, the journalism enter those fields.

Being a digital media company and a printed newspaper company, GFR Media must quickly provide digital content across many devices through their websites. Journalists and editors are educated and trained to provide investigations, and articles but usually they do not have the necessary programming knowledge to put news content in complex sites and digital environments. Existing CMS usually have a closed architecture that does not fill the editorial needs of journalists, as they do not adapt to each specific website needs.

The people usage trends shift so fast that it could be difficult for journalists and editors to keep up to date with the available technologies. A journalist's inability to keep up with the current technology, has a direct impact in the ability of news agencies to deliver updated information on different sites. At the same time, the internet has many advantages because it allows the journalists to work non-spatially. The necessity for the journalist to actually be present in the news desk is eliminated and the journalist can pursue his profession from any place in the world with a computer and an internet connection [5]. Publishing online continues to be an experiment for many newspaper companies, and for years they have spent resources on developing web sites and have lost money, because those newspaper companies have used obsolete technology and are unable to quickly adapt to new users technology trends. Now, newspapers are employing various strategies to make online businesses profitable [2].

METHODOLOGY & DESIGN

The end goal of this project is to enable newspapers journalists and editors to upload information to the newspaper sites quickly while providing an efficient and robust interface. The "News Publishing Platform" system is composed of three main parts: the CMS, API and ASP MVC Site. The developed software design maintains each component (the CMS, API and ASP MVC Site) carefully implemented allowing all of them to be able to connect seamlessly and to be able to be updated without affecting their independent correct functionally.

Application

Figure 3 below demonstrates how "News Publishing Platform" system works. First, the editors or journalists write the news gathering the information from their resources, embeds and pictures as desired. After gathering the information, the editors and journalist will proceed to access the CMS to enter all the information that they gathered including the date, title, introduction, body content and image of the news to be uploaded. Figure 3a shows the uploaded information including the date, title, and news content will be saved in a shared instance of Microsoft SQL Server hosted by GoDaddy Hosting Company seen in Figure 3b. The uploaded image will then be stored in Amazon's Simple Storage Solution which also offers Content Delivery Network (CDN) capacities to deliver the images to the end users quickly. Second, after the information is saved or recorded it is accessed through the project API which serves as the gateway to deliver the information. Figure 3c shows how he API delivers the recorded data to end users by using the recorded memory of the cache. As soon as the API request responses are recorded, other users can request the same information instead of communicating again with the database to retrieve said information.

Even though the API configuration allows greater capacities to deliver information, it depends on hosting providers to successfully handle such data demand. In order to enhance API's data delivery capacity and to have a greater response speed, the proposed project API is connected to the 3Scale API Solution shown in Figure 3d. This enables the API to serve even greater capacities while adding the ability to authenticate requests and users, making it possible to limit allowed requests by users in a specific time. This way the API has the ability to limit the amount of data requests a user or application can retrieve in a specified time interval.
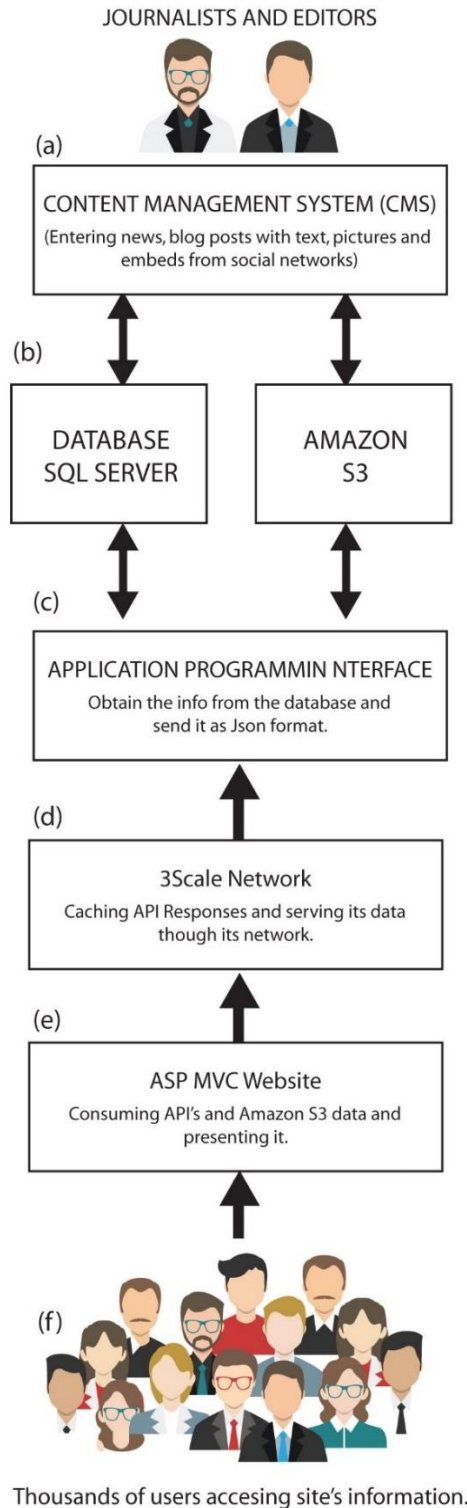
JOURNALISTS AND EDITORS

(a)

CONTENT MANAGEMENT SYSTEM (CMS)
(Entering news, blog posts with text, pictures and
embeds from social networks)

(b)

DATABASE
SQL SERVER

AMAZON
S3

(c)

APPLICATION PROGRAMMIN NTERFACE
Obtain the info from the database and
send it as Json format.

(d)

3Scale Network
Caching API Responses and serving its data
though its network.

(e)

ASP MVC Website
Consuming API's and Amazon S3 data and
presenting it.

(f)

Thousands of users accesing site's information.

Figure 3
Software Architecture Diagram

The capacity to enter and record data through the CMS, and to serve it through de API allows journalists and editors to upload the data making it available from any application, in this case a web site developed in ASP MVC see Figure 3e. As a result of this configuration, the web site could be accessed by many thousand users in any device simultaneously. Figure 3f represent the thousands users that access the website developed using the developed solution.

Database

The project's data base is developed in Microsoft SQL Server. The database design is implemented using the Patterns of NoSQL making it easier for the CMS to connect also to any NoSQL database technologies as Amazon Dynamo or others. At the same time the program works well using MS SQL Server hosted in Amazon or in a hosting provider such as GoDaddy or Azure. As the project needs to be constantly accessible it can be configured with different hosting providers. The database can be backed up constantly to ensure data redundancy.

Hardware

This project was designed in a Lenovo ThinkPad T420s with the latest Version of Microsoft Visual Studio Enterprise 2015 with Windows 8.1. The project is hosted in the Azure cloud and the images are hosted in Amazon Simple Storage Solution. The project code is developed using the .Net functionalities including ASP MVC 4.1 components for the CMS, web site and the API functionality to deliver the data in json format. The database is hosted in an instance of Microsoft SQL Server hosted by GoDaddy.

RESULTS

While accomplishing the main goal of delivering news quickly in a robust system, the project also serves as an extensible platform that can be quickly reused in other systems, such as any other webpage or mobile application. The project solution has advanced functionalities that are crucial to the newsroom industries, including the CMS System, API and the website that consumes the data. The newsroom industry benefit greatly from the use and implementation of this project, the "News

Publishing Platform" system, because of its capacity to enter news quickly and efficiently to any webpage or application. Each component of the developed solution can be updated and modified according the newsroom special project needs without affecting its performance and capacity to deliver news pages.
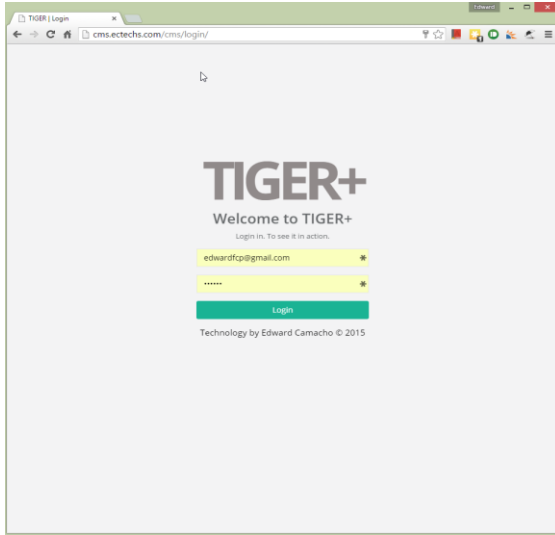

Figure 4
CMS Login

The CMS System

The Content Management System or CMS allows editors and journalists to upload the news content that they want to publish to the websites. To accomplish this the user must first login into the system by going to the log-on screen shown in Figure 4. After logging on successfully, the user gets to access the news publishing user interface.

The information therein contained is retrieved from the database using the Microsoft Entity Framework technologies, which allows the user to access, read, modify and delete the database information. Another functionality used is the LINQ (Language Integrated Query) which allows the developed software to use complex queries to manage database information, while reducing the required amount of code required in the developed software solution. All project components including the CMS system are fully developed using Microsoft C# programming language and HTML / CSS with the latest version of Jquery and its javascript library.
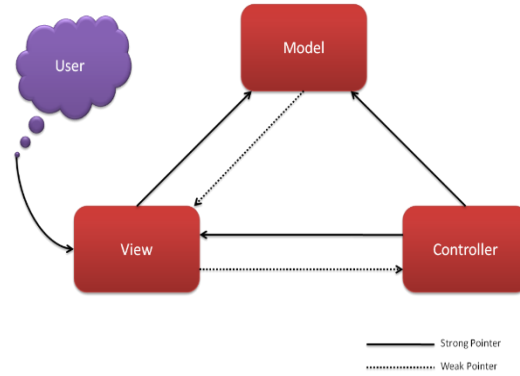

Figure 5
ASP MVC Site's Configuration Diagram

The menus of the CMS allow users to view the information, edit its fields and delete if they want to. It works with a standard MVC pattern as seen on Figure 5, where the user interacts with the forms and html views to send the data to the controller which processes the request and communicates with the model defined by the database. The system uses reusable html views to enhance the load times of the site while maintaining the uniformity of the user interface. The CMS is available 24 hours a day, seven days a week. Because the news content could be uploaded to the sites at any time it must be accessible from anywhere around the world. To ensure the integrity, confidentiality and availability of the information when needed, the program requires user authentication. The journalists and editors will enter all the information and the images that will be used in the news entry to be delivered later to the end user.

All actions performed by the editors or journalists in the CMS system require a verification that the user is logged and has administrator privileges to perform the selected operation. If the session is not created the system controller will redirect the user to the login web page. This way even if a previous user has the URL it cannot create view or edit entered context until he enter valid credential. The CMS is developed using an adaptive design that allows it to work in any device from a desktop computer to a tablet or mobile device this feature help editors and journalist to publish content while being out of the office and while investigating any events in any place in the world at any time.

Figure 6
CMS View Entries

The CMS system protects the user from an attacker using a false initiated session because its login form has an anti-forgery token that the software validates and must match with the token provided in the CMS user interface forms, and in the ASP MVC controller retrieving the information.

After the editor or journalist reviews the existing created entry, he or she must press the "New Entry" in Figure 6 or the "View Entries" in Figure 7 buttons to be able to add the new blog post or news adding the required fields and images.

Figure 7
CMS New Entry

Online news sites are becoming more and more interactive, first and foremost regarding human to-human interactivity. End users are allowed to contribute to the content production by submitting photos and videos and by commenting on stories and participate in discussion forums [6]. Based on that fact the user is able to upload a picture and is also able to crop it according to the website layout. This functionality is based on the Jcrop Image Plugin mechanism. At modern times when the internet speeds are faster and computers can download images quickly, is known that users focus their attention more on text than photos according to the Journal of Electronic Publishing. This behavior really surprised graphic designers and newspapers photographers because they previously has assumed that. The findings were very different from Eye track newspaper studies, which over the years have found that newspaper readers focus first on photos, then on text content [7].

Figure 8
Social Media Embed

The CMS uses a Rich Text Media editor called CK Editor having the power to conveniently format the site and adding automatically the HTML tags to enhance the editor experience enabling editors or journalists to format their stories with bold, align and other HTML attributes. One of the most interesting attributes is the capacity to embed social networks elements as videos from YouTube, Facebook posts or Instagram pictures and displaying them correctly in the web site as seen on Figure 8, which illustrates how images are displayed as well as the article title and introduction.

CMS Security and Authentication

When the editor or journalist enters his or her username and password, the CMS system reads the username and password from the form contained in the login section of the CMS and validates the user identity comparing the user input with the salted password stored in the database. Because of security measures, the solution's database does not have the copy of the passwords in plain text. Instead it has the salted hash string from the combination of the username and password. The salt is done with the Cryptography functions of Microsoft Visual Studio .Net.

Application Programming Interface

The project's API is based on the Microsoft Web API functionalities making it possible to transfer the site information using http calls. The API is the most important component in the "News Publishing Platform" system, as it supplies all the entered news information. The API can also accessed by more than a thousand users because of its configuration hosted in GoDaddy Hosting Company and the configured 3Scale configuration.

The Project API has two main components or divisions. The first of them is the API News Controller that is used to obtain the News Post for specific categories, from all categories or from the most recent news post. The second component is the API Comments Controller to obtain the user's recent comments and most commented news.

A very important aspect of the API is that it offers number paging request capacities. This means that when a user looks for or browses for news, he or she can select a brief selection of news instead of all of the news that would appear for a specific query. Because the API system can store many thousand news posts it is necessary to implement a paging capacity that will allow a user to request a specific number of news instead of all of them. API technologies have become the standard to deliver content and are very suitable for news content where an editor or journalist needs to display a news website that could be complex enough to showcase many news for public consumption.

Large Capacities API with 3Scale

A 3Scale network implementation consists of an added layer to the API project that adds the capacity to manage and analyze the API requests while providing a more robust solution for faster usage information that is reliable. The most important feature of the 3Scale functionality is that it provides an API Gateway that uses keys to monitor or limit usage while serving an API Request. 3Scale handles the authentication and rate limits according to the configured rules established by the administrator. The administrator account can determine or assign how many requests an application can make to receive information.

ASP MVC Web Site

Figure 9 below shows the news website, which is the third component of the "News Publishing Platform" system. Modern web sites open up new ways of storytelling, primarily through the technical components of the new medium. Simply put, online journalists can provide a variety of media--text, audio, video, and photographs--unlike other media [8]. Another advantage of news sites is the capacity to quickly search content.

The "News Publishing Platform" system is designed on ASP MVC 5 Web Solution. The project has an advanced image cropping functionality that allows the ASP MVC site to directly use the images uploaded by the editor or journalist. The sample site features a clean, well-organized HTML code that works well across most modern web browsers. By using this powerful plugin, editors and journalists have complete control of the content of their stories from the text to the correct file upload capacities. The front-page of the project website showcases the latest news entered in the website from all categories. On the bottom of the website there are widgets with the link of the most commented news. For the editorial department the most commented news tend to be the most interesting and engaging for the audience, causing them to stay tuned and connected to the site. The news site can be found online at: http://news.ectechs.com. The formal presentation of a story featured on a website is one

of the most important aspects that a journalist or editor takes into account, as well as its content. The reason behind this is because journalists want their audience to experience an engaging and interactive experience. Said experience will be provided by the "News Publishing Platform" system developed.

Figure 9
News Site Home Page

On the front page, users can also press any of the pictures therein contained or the news titles. By clicking on the pictures or news titles, the user will be directed to the complete note with the story details shown on Figure 10. All news item are rearranged depending on the device used by the user to access the website. Website are compatible to mobile platforms, mobile applications and similar new platforms are changing the content of the news item [5]. Mobile devices are also transforming the news environment through the rise of the so-called second screen phenomenon, in which users log onto the internet while also using another technology [9]. Also the website is adaptive so it adjust to the device seeing it which could be a computer, tablet or phone changing its appearance depending on the device.

Though online and mobile news have produced many benefits for the news industry, they have also brought challenges [3]. Because of this fact many journalist have to adapt to those new technologies that are widely used and need to be able to bring news using them.

Figure 10
News Detail Screen Shot

Social Networks Friendly

A modern behavior is also changing the news scenario because any user of an electronic device can take pictures and share a story using social networks. Causing that any user can share even a breaking news. The audience now tells the media what they want, as opposed to the media telling the audience what they want [10]. At the same time when users find newspaper content interesting they share it. For this reason it is truly important to have a website that is social network friendly because the user will receive the news content shown in the social network and if the site is unable to carefully grab the information it will not be shared correctly thus limiting the news content reach because the content will not be shared.

For a newspaper company the use of social networks is very important to successfully share the content of its sites and particularly its blog entries. Social networks, such as Facebook, must be able to successfully recognize the published article content when sharing the information. To be able to obtain

this result there is an Open Graph Protocol which is a group of specific html tags that will mark each component of the site for easier recognition.

New technologies attract audiences to the news industry. For instance, the interactivity of social media is transforming the way consumers receive and participate in news. Twitter allows "followers" to personalize their news and respond to the news sources they "follow" with questions, ideas or concerns. Is well known that social media, especially Twitter, has changed the newspaper landscape [2]. These aspects of social media create a richer news experience for users when compared to traditional news mediums that were focused on one-way communication [6]. According to the Pew Research Journalism Project, thirty percent of United States adults consume news on Facebook, and seventy-eight percent of those adults mostly see news when they are on Facebook for other reasons [3].

Web Site Comments Component

The "News Publishing Platform" system uses the Disqus comments system seen on Figure 11. Gives the users the ability to make comments to the site with different social networks, also giving the flexibility to login from any existing social account. Survey results from an investigation carried out by The Elon Journal of Undergraduate Research Foundation found that the field of communications shows that 30.8% of the study participants have commented or contributed to a news article online at least once [2].
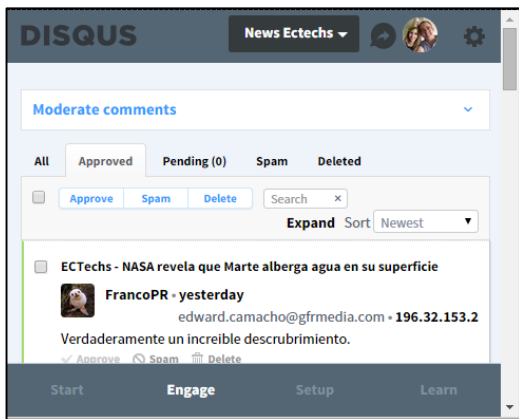

Figure 11
Disqus Site's Comments Moderation

Allowing users to make comments in news helps editorials departments to make the audience to feel identified with the newspapers company and gives the audience the opportunity to freely express their opinions on any matters while at the same time they can respond to other users making a dialogue.

Traffic's Analytics

In order to have all the site's information the owner (GFR Media) must have real time analytics and be able to access easily and quickly the audience usage information. With Google Analytics software developers can improve performance across an entire site.
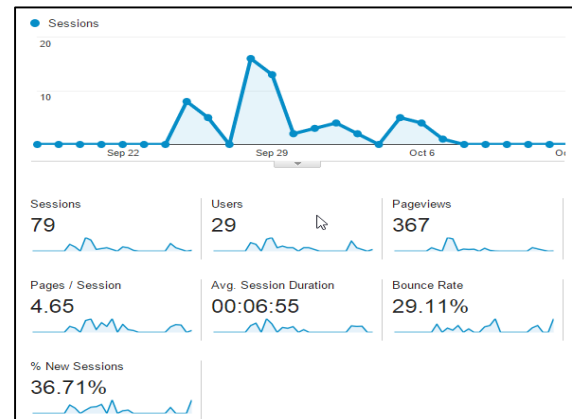

Figure 13
Google Analytics Real Time Charts

Having this functionality as part of the "News Publishing Platform" system will provide very helpful information regarding the most accessed parts of the sites, where the user came from, which devices used it and in which applications. Figure 13 shows the real time graphics showing line charts, and meaningful information.

DISCUSSION

The developed "News Publishing Platform" system successfully enhanced the digital newsroom capacities to serve news. It also established a robust and well-designed News Publishing Platform that will allow news companies to deliver news quickly, reliably and with the ability to quickly upgrade any components, having bigger capacities to deliver content.

FUTURE WORK

The developed "News Publishing Platform" system can be easily enriched with additional capabilities in terms of functionalities as CDNs, better cache and faster response times. Likewise, the greatest potential of this system is the ability to easily implement a template system as WordPress which allows the editor not only create content but to create an entire complete site with the capacity to change the site template or produce a derived new site. Because of its modular approach it can easily use different databases systems as MS SQL or non-relational database systems as Amazon Dynamo.

CONCLUSION

The developed project solution provided GFR Media with the tools needed to deliver the news content to its many special websites which are accessed by thousands of monthly users. The "News Publishing Platform" system also gives their newsroom the opportunity to adjust or upgrade each component independently with the latest software development tools. The system allows GFR Media to be able to keep with the pace of technological evolution and the multiplicity of choices from platforms to devices that show no sign of slowing down [11], while preparing them to stay relevant and competitive in a quickly evolving news delivery journey.

REFERENCES

[1]   B. Van Der, M. Parks and M. Castells, *"The Future of Journalism: Networked Journalism",* in International Journal of Communication, vol. 6, 2012, pp.1-11.

[2]   E. Everett, *"Transformation of Newspapers in the Technology Era"*, in The Elon Journal of Undergraduate Research in Communications, vol.2, no.2, Fall 2011, pp.1-6.

[3]   N. Chadwick, *"How Online and Mobile Technologies Have Changed Broadcast Journalism",* in Revolutionizing the Television Newsroom, vol.5. no.1, spring 2014, pp.2-6.

[4]   P. Lia, M. Matsiola & A. Veglis, *"Journalism in a State of Flux: Journalists as Agents of Technology Innovation and Emerging News Practices"*, in International Communication Gazette, Special Issue on New(s) Media Technologies and New(s) Work: Changing Journalistic Practices across Europe, vol.74, no.1, pp.6-7, Feb. 2013, pp.5.

[5]   M. Kuyucu, *"New Communication Technologies and Journalism: The Creation of New Multimedia Newspaper with Milliyet Newspaper and Blippar",* in International Conference on Communication, Media, Technology and Design, 2014, pp.4.

[6]   S. Steensen, *"A critical review and look ahead",* in Online journalism and the promises of new technology, vol. 12, no.3, 4 Sept. 2010, pp. 8-11.

[7]   J. Teoh and S. Hawamdeh. "*The Impact of the Internet on Teaching and Practicing Journalism*", in The Journal of Electronic Publishing, vol.7, issue 1, August 2001. [Online]. Available: http://quod.lib.umich.edu/j/jep/3336451.0007.102?view=text;rgn=main.

[8]   C. Harper. (1995, May). *"Journalism in a Digital Age"*, in MIT Communications Forum. [Online]. Available: http://web.mit.edu/comm-forum/papers/harper.html.

[9]   J. Sansen, K. Olmstead & A. Mitchell. (2013). *"Digital: As Mobile Grows Rapidly, the Pressures on News Intensify",* in The Pew Research Center's Project for Excellence in Journalism - The State of News Media 2013. [Online]. Available: http://www.stateofthemedia.org/2013/digital-as-mobile-grows-rapidly-the-pressures-on-news-intensify/.

[10] J. Hadley. (2012, June). *"Running head: Traditional Journalism vs. Advancing Technology"*, in Running head. [Online]. Available: https://jenhadley.files.wordpress.com/2013/08/researchpaper_jen-hadley.pdf.

[11] Pew Research Center State of the News. (2015). *"Numbers, facts and Retrieved October 3, 2015"*. [Online]. Available: http://www.journalism.org/2015/04/29/state-of-the-news-media-2015/.