



**UNIVERSITY
OF OULU**

FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Antonio Kongjonaj

**INCREASING THE CREDIBILITY OF SCIENTIFIC
DISSEMINATION USING CROWDSOURCING**

Master's Thesis
Degree Programme in Computer Science and Engineering
May 2019

Kongjonaj A. (2019) Increasing the Credibility of Scientific Dissemination using Crowdsourcing. University of Oulu, Degree Programme in Computer Science and Engineering. Master's Thesis, 45 p.

ABSTRACT

This thesis introduces Article Enhancer, a semi-automated web application that utilizes crowdsourcing services, specifically Amazon's Mechanical Turk platform, for augmenting articles with various referencing content gathered from the crowd-workers, on demand. The main goal of Article Enhancer is to address the question of how scientific articles can be made more credible, before dissemination to the public. This application serves as a tool in helping users find suitable supporting content for their articles in a novel way, removing all the manual work of doing it themselves.

Media literacy, social media, fake news and crowdsourcing are discussed as part of related work. Also, tools that offer a similar functionality are reviewed. Furthermore, system design and implementation for Article Enhancer is presented. It is important to mention that the referencing content provided through Article Enhancer comes from already existing online content. Although Article Enhancer is semi-automated system, its strongest point compared to the other systems, is that it doesn't require extra human effort to enrich articles especially with visualization content, and providing already existing content on the web avoiding the process of creating new content, making it a fresh approach in this line of software service.

To evaluate Article Enhancer, we deployed the web app in a real-life setting, a space oriented towards students known as Tellus, at the University of Oulu. This testing proceedings helped in determining that the system appears alluring and attractive to new users. Article Enhancer proved to be unique and thrilling after the first encounter for many of the users. Feedback also shows that adding and embedding content is an innovative way to make articles become more credible in the eye of the reader.

Keywords: Scientific dissemination, news credibility, crowdsourcing, augmenting.

TABLE OF CONTENTS

ABSTRACT	
TABLE OF CONTENTS	
FOREWORD	
ABBREVIATIONS	
1. INTRODUCTION.....	6
1.1. Motivation.....	6
1.2. Scope and Objectives.....	7
1.3. Structure of the Thesis.....	8
2. RELATED WORK.....	9
2.1. Evolution and Impact of Online Articles.....	9
2.2. Tools for enhancing Articles.....	10
2.3. Social Media and Fake News.....	11
2.4. Crowdsourcing.....	12
3. SYSTEM DESIGN AND IMPLEMENTATION.....	15
3.1. Chosen Tools and Platforms.....	15
3.1.1. JavaScript.....	15
3.1.2. Python.....	15
3.1.3. Amazon’s Mechanical Turk.....	16
3.2. Feature Overview.....	16
3.3. User Interface.....	17
3.4. Back-end Technology.....	24
3.4.1. MTurk Sandbox.....	26
4. EVALUATION.....	28
4.1. Experimental Setup.....	28
4.2. Questionnaire Data Analysis.....	30
5. DISCUSSION.....	35
5.1. Achievement of goals.....	35
5.2. Future work.....	36
6. CONCLUSION.....	37
7. REFERENCES.....	38
8. APPENDICES.....	42

FOREWORD

This project thesis has started as part of internship at the Center for Ubiquitous Computing in the University of Oulu, in June 2018.

I thank Prof. Timo Ojala for his support during my studies. I would like to express my profound gratitude to Dr. Simo Hosio for his supervision, guidance and motivation throughout my masters. I also thank Jonas Oppenlaender for the good will to help me and for overseeing my work.

My last appreciations and thanks go to my family and especially my father, for the support to follow studies on this master at the University of Oulu and during all this period.

Oulu, 17.05.2019

Antonio Kongjonaj

ABBREVIATIONS

UI	User Interface
HTML	HyperText Mark-up Language
API	Application Programming Interface
JS	JavaScript
JSON	JavaScript Object Notation
SQL	Structured Query Language
SQLite	Relational Database Management System
AWS	Amazon Web Services
SDK	Software Development Kit
CSV	Comma-separated Values
URL	Uniform Resource Locator
MTurk	Mechanical Turk
HIT	Human Intelligence Task
HITLayout	Human Intelligence Task Layout

1. INTRODUCTION

1.1. Motivation

With the creation and expansion of the world wide web and its usage a new era for information came to life. Internet started to be exploited as a new means for information and news sharing, which time proved to be a very successful entourage. It served as bridge to be up to date for events or facts happening around the world and at all that at every user's fingertip.

This powerful tool that caught the attention of mass population, also caught the one of the medias, traditional and newly emerged, which began their crusade of information disseminating online. Meanwhile for the first-time information sharing was not a monopoly of the media and journalists, but also of people, individuals could share their opinions, arguments and facts for everyone else, with a computer and an internet connection, to read [1].

The big change in news dissemination was very welcomed by the internet users, who began to communicate by different means that came with it, chat, websites, blogs, social medias. As good as it sounded a place where everybody had a voice and was free to express himself, but all this freedom did not come without a price. The world wide web became a source of enormous information, but this did not mean that everything written there was true. Since people were sharing their own opinions and arguments, it became unclear what was vetted to be true and what not, while also authors with malicious intentions, who for one purpose or another, appeared and started spreading false news. [2, 3]

The credibility of the news available online is one of the most discussed topics during past years. Having to deal with such resource of information, where most of it is unchecked and even unreliable has become a big issue. Since the first appearance of the false news to date, the situation has dramatically changed. Wrongdoers or "spammers" have evolved with time and their goal is easily understandable since each click on their news is effortlessly materialized through paid ads, but it does not end there, misinformation is being used further as a tool to influence the population, in a local or global scale. Despite the fact that this has been a problem persisting through the years since the creation of the internet, there is still no permanent solution to verify every single news, or piece of information that has been posted or will be posted, which puts the disseminator of the truth at an uneven position regarding the spreaders of fake news. [4]

The fake news concept developed in an unprecedented scale of the snowball effect, becoming an increasing pain for people around the world, and especially for the journalists, whom are also referred to as the traditional fact-checking pillars and with the colossal amount of new information coming their way it's impossible to keep up, creating the need for computational fact-checking service [5].

One particular group interested in the credibility of their news is scientists. Scientific news must be credible, to communicate the best available current truth. This thesis discusses a system which focuses on the augmentation of the articles, with emphasis on scientific articles, by providing enriching content and references for the writers. The augmentation of the articles using content already present online for the user's

own statements is made possible by interacting with a 3rd-party crowdsourcing platform such as Mechanical Turk provided by Amazon. The elements to be used for augmenting an article, are found and delivered from the crowd-workers this platform provides. Ultimately, the tool is to be used to increase the credibility of any arbitrary article, while in this thesis we focus more on the usability and user perceptions of the tool itself.

1.2. Scope and Objectives

This thesis is based on my internship work done at the University of Oulu, June - August 2018. It introduces the impact of the new era of information spreading on this generation of internet users, and the influence that social medias have overtaken during last years as a means of disseminating news and information is one of the biggest transitions for the world journalism in history.

The work focuses on the implication that the social medias brought regarding the fake news phenomena, that plays a role in biasing and misleading the users of the socials platforms, whom rely on them to stay up to date and take the information they read as true, and discusses the existing instruments as well as third party technologies that can be used to verify news and information along with tools that aim at enriching articles with examined and credible embedded information towards their writer's needs, preventing this way the misinformation of the readers.

The key objective of this thesis are to develop a tool that people can use to write articles and enhance the articles with credible and reliable information supporting their statements in the form of tweets or references, hence making the "referencing" easier while providing plausible and dependable information through a third-party service. By referencing, we refer to finding supplementary content, such as tweets or other online articles that support the message of the article being augmented. The primary target user group for this system are mainly academics and students, but theoretically naturally anyone can utilize it. The functionality of the prototype is illustrated, as well as the design and technical requirements for accomplishing the prototype system are established.

With this project, we address the question of how scientific articles can be made more credible before dissemination to the public and by using the crowds. The final output of this work is implementation of a fully functional system which shall be deployed for the open use of every interested individual. The thesis includes also a literature review of the state-of-the-art available systems which use the same technology or even distinct ones for a similar purpose.

Finally, the thesis makes an open-source contribution by making the developed software freely available for any type of use via a public GitHub repository.

1.3. Structure of the Thesis

The rest of this thesis work is structured as follows. Chapter 2, Related Work, presents a review of media literacy and credibility, the tools and instruments for writing along with an overview of what crowdsourcing as a concept is, and how it started. Third Chapter provides details regarding the design and implementation of this prototype system, the chosen technologies, functionalities and features that it offers. Chapter 4 handles and describes the evaluation process of the prototype at hand as well as presents the analyses of the data gathered from surveys. In Chapter 5, discussions concerning the project and a whole overview of the work are treated including future work and achievement of goals. On the final Chapter, a general conclusion and last remarks about the work are given.

2. RELATED WORK

2.1. Evolution and Impact of Online Articles

The capability to analyze, evaluate, create, and act using different types of communication is known as media literacy. The cultivation of media literacy as an essential future direction in the age of new communication technologies is a concept that European Commission experts on online information recognize [6]. Educating and disseminating reliable information is a necessity for maintaining a society based on knowledge, where the people understand and can differ between true and fake facts.

It is a fact that the media landscape is rapidly changing, nowadays news and media can be created not only by any individual with a computer or a smartphone, but by artificial intelligence as well, to spread misinformation at a faster pace, outdating the skills of many news disseminators in the process. Today's students are the first generation that is growing up in a world invaded by digital media and an environment where smartphones, tablets and e-readers provide extensive access to information online. A state-of-the art news piece is no longer considered only text and images printed on paper but is now instead created on a programmable platform and may contain elements such as embedded social media content, animated visualizations, or even engaging elements such as polls and quizzes. All these new components associated to a news piece affect the readers' emotional response and eventual actions [7]. In this thesis, we examine how we can take advantage of the affordances of the new media environment in increasing the credibility of scientific dissemination before the dissemination among different audiences.

The way that news and media information are presented affects the readers' emotional response and eventual actions [7]. For example, news on climate change can be disseminated using different methods of argumentation. References to credible sources ("98% of leading scientists think climate change is a fact"), personification ("As an ice fisher, I have seen that during winter the ice fishing season is shorter and shorter"), or the invocation emotions ("The future of this planet and our children depends on stopping climate change") are all approaches which can be applied widely and that affect the reader's response. Graphical, visual and stylistic applications, utilized in the presentation of articles, appear at this stage in the age of Internet and are being acknowledged more and more within both academic and policy discussion as an important part of media literacy in general [8].

The impact that a news article makes also inevitably depends on the reader and the audience, making it important to study and be aware of how different elements and their combinations influences different audiences with varying characteristics such as demographic, cultural and social backgrounds, different areas of interest and level of linguistic skill. For example, personified news' have been shown to shrink the gender gap with regarding the emotional response to a news article [9].

This rapid development of media landscape requires for journalists, scientists, academics to learn how different modes of media affect how the message is perceived and received among the target audience. This helps citizens to learn to develop their media literacy and multiliteracy competences, and, most importantly, it will help researchers to learn and to cultivate these competencies.

The mentioned considerations have extensive societal consequences. Considering, for example that some stories of mass violence “arouse public concern while others are met with seeming indifference” [10]. This example shows that is essential to tell the story right. That is why we focus on helping the dissemination of scientific articles to the society, by enriching the content of the articles with different content and elements already available on the web, for the purpose of creating a better illustrated version of the same news piece in order for it to be more interpretative.

2.2. Tools for enhancing Articles

The news landscape has evolved a lot through the years and data visualization has become an outstanding part of it, enhancing the essence of news article providing the readers with substantially better reading comprehension [11, 12]. Although, this is not a fully automated process, actual systems are limited, first because they may demand extra human (e.g. [13, 14]) effort and secondly, they can be restricted to distinct types of visualizations (or [15]).

An example of a system which is restricted by the incapability of being fully automated for enhancing articles is ReVision. Created as a response to deal with poorly designed charts found online which were only offered as bitmap images, giving no access to essential data for the individuals who would want to develop a more efficient visualization of data representation. The system redesigns these visualizations for a better and improved graphical view. It relies on machine learning and computer vision methods to identify the type of chart, extract and derive the data. Then through the data acquired it creates a collection of redesigned charts presenting them to the users. Although the system during the evaluation had a high accuracy in classifying images, as mentioned above it still lacks the full automation along with being only limited to specific types of visualizations such as charts. [16]

Yet another piece of software with the main goal of automating the process of creating visual content, by means of an interactive navigation gallery which would automatically generate visualizations, is called Voyager. Voyager is a mixed initiative system and its strongest point is that it is backed by faceted search technique, of the suggested type of charts chosen accordingly based on statistical and intuitive means. While a comparison was made between Voyager and another tool for visualization which was categorized as a manual one, the result showed that Voyager simplified the task of exploring for data not observed before and it would point towards a bigger data variable analysis. This solution resulted in a step forward regarding the creation of automatic system for data visualization collection, however it is part of the category which still requires some human effort for the purpose of enriching articles, that makes it in turn not as convenient as a fully automated system. However the solution presented is a very helpful business focused tool. [17]

Following the need for a tool which would generate automatic visualizations of an annotated stock in an effort to enrich articles and complement them, Contextifier came to life. This tool would give the user an automatic custom result related to visualization content provided a news article for a specific company. As described by the authors the algorithm of Contextifier relies on professionally developed visualizations and

considers a number of factors such as visual salience, contextual relevance, and it also checks for important events in the history of the company. This effective and reliable tool is, as presented, for a specific category of users. [14]

In the study of “VizByWiki: Mining Data Visualizations from the Web to Enrich News Articles”, the authors propose an automated system as an alternative for adding contextually relevant data visualizations to news articles without requiring any human intervention at all. Furthermore, they offer a novel approach that has minimal constraints regarding topical domain or type of visualization. Their approach consists a system that tracks content already available online for a given article topic and mines data visualizations, that are relevant contextually, from Wikipedia central repository Wikimedia Commons. On their examination of the system they demonstrated that it can enhance up to 48% of popular online articles, while using a new ground truth dataset. Moreover, the system could rank the visualizations with respect to their utility and correctness. [15]

All the system considered above, made an impact in the new news landscape. Even though these tools have their own limitations, they contributed in shaping the news landscape especially from the viewpoint of the academic lens.

2.3. Social Media and Fake News

In the age of technology, staying up to date is a requirement which everyone one seems to attempt to fulfil. People have and feel the need to be informed, regardless of the reason and the kind of information they take and with social medias coming to existence, this burden was made even easier. The new trend shows that people usually get information and news through social medias, since its easily accessible, low cost and convenient, all that is needed is a smartphone with an internet connection. Considering the role of social medias in a person’s life, people in the fake news business saw a new tool to help in their entrepreneurship. These are the medias that have changed the way information is perceived, from personal life insights to news about wars happening on the other side of the world. [18]

However, the other side of the coin are the fake news, becoming an actual menace to society. That is because the fake news is more novel, the readers like what they read and they would like for this exaggerated news to be true, they want to trust them, for the simple reason of the emotions that they can evoke in them. Basically, false information travels faster, deeper and further than the true one, and its thematic varies from politics to sports, terrorism, sciences etc. These false news’ are now known to have been used for more than just monetized ads, such as influence elections and even endanger democracies. [19, 20]

There have been a lot of rumors and allegation about the politics turning to this additional mean to help them in their quest, in other words using fake news and also bots in social medias to push this fake news to the population, to influence them, until it was actually proved for the first time on the study of “Social bots distort the 2016 US presidential election online discussion” for the United States [21] and not only limited to the US but also around the world [22], emphasizing the fact on how dangerous the impact of fake news spread via bots can be.

In another study on social medias and fake news, an experiment that included over 14 million messages disseminated through 400,000 thousand accounts tweets during a ten month period on 16'-17' was conducted and discovered that social bots had an excessive role in sharing articles from sources with low credibility, before the article goes viral and at the same time they have a set target group, going mainly after user with many followers in comments and replies.[23]

Fake news is not exclusively linked to politics, but the latest trends of world politics have made this type of news a big part of social media platforms news dissemination outline during past years. Even though having no actual permanent solution, while social medias and journalists and IT researcher are continuously trying to create a functional fake news detection tool, makes it sound like there are no measures that can be taken against "plague" infesting the online world, but that is not entirely true. An individual can take provisions to certify that whatever piece of information he is interested is reliable or not, such as to compare it against reputable news sources, or using a third-party service which would analyze it otherwise called fact-checking services. [1]

Fact-checking is a concept arisen and developed with the purpose of proving the veracity of claims, getting a lot of attention especially over the last decade where the world wide web has become a chaotic information host, where being misled is very easy as long as the reader does not attempt to check the truth of what he is reading. To remedy this new crisis challenging the world fact-checking services were created, a group of people, most commonly journalists, started to check and verify the news being spread.

This endeavor proved to be viable, however with the mass usage of social medias for news dissemination resulting in an enormous amount of information being continuously generated, the volume of work needed to filter and assess every information coming out of these platforms is humanly impossible. Hence these traditional fact-checking platforms, are primarily focusing on news and information which have a certain importance, interest and the most probable to bias readers, such as political, financial and public claims. Even though researches have pointed out that the people judge the trustworthiness of the fact-checking services based on the ownership, origin and the transparency of their methods, there have not been enough examinations to actually give a true perception of the users on the fact-checking services. Based on the type of information, the veracity of which is checked, the fact-checking services can be put in three main categories political and public claims, rumors and hoaxes, and last the scientific news and polemics. The number of active services is well over 100, but amongst the most reputable platforms can be mentioned Snops.com, FactCheck.org, PolitiFact, StopeFake etc. [24]

2.4. Crowdsourcing

Crowdsourcing is a new way of eliciting work, popularized during the past decade. It has redefined what collaborative group work means at this stage of technological development. This type of service offers another more authentic and at the same time less expensive way of gathering human insight in the process of data integration and

cleaning [25], as well as giving the possibility of hiring people all around the world, making the global market a real thing. Essentially, crowdsourcing refers to sending small tasks over the Internet to large masses of online workers who then complete the tasks in almost real time.

The applications of crowdsourcing are of a wide range, as it has resulted in being a successful commercial strategy for work. People have used the concept to build innovative businesses such as Threadless, focused in crowdsourcing designs for t-shirts and selling them online [26], or already established companies using crowdsourcing as way to get feedback from customers as PepsiCo did, taking it on another level by asking customers to participate in a campaign to help them think of a new flavor idea [27]. However, the most successful, consistent and broad business application of crowdsourcing is that of platforms which operate as marketplaces for crowd-workers, where individuals are hired through these online markets to solve different kinds of tasks or used for research data collection, e.g. Amazon Mechanical Turk, Prolific.ac.

There are many companies offering online work accomplishing, however Amazon's Mechanical Turk is one of the oldest crowd-powered marketplaces for micro-task accomplishing. And by Amazon's own definition the Mechanical Turk is "a marketplace for work that requires human intelligence". It works in such a way that is easily utilizable from both the employer, a person posting the task, and worker, the individual whom is employed by Amazon to carry out tasks. Amazon claims to have over 100 thousand workers which come from 100 distinct countries [28], providing employers with a wide range of various characteristics to be able to choose from when hiring a certain worker based on what they need. There is a broad number of studies made regarding Amazon's Mechanical Turk, its reliability and quality of the data gathered and analyzed. Depending on the genre of the tasks processed, these studies inspect the usefulness of the platform's data quality concerning applications such as image databases [29, 30], NLP [31, 32], perception tests [33, 34] and even decision-making experiments [35].

Having an increasing number of experiments emerging regarding social sciences during last years is related to, besides other factors, also to the crowdsource services that provide one of the most efficient ways of gathering rich data. As the number of research experiment conducts increasingly making use of Amazon's Mechanical Turk had a notable rise from 61 in 2011 to 1200 in 2015, other platforms offering crowdsource as a service arouse. Prolific.ac came to life for the purpose of dedicated research subject pool, as new tool for academics and whomever had an interest towards social sciences to utilize this tool towards their end. The novelty of Prolific lied in being a crowd-powered platform oriented towards scientific research as opposed to MTurk and other services available which were not intended explicitly for the scientific community. Prolific has been established to be a reliable research data gathering platform, used in many different types of studies ranging from economics, to social sciences and even food science, making it perfect for any type of scientific research. [36]

However, on a recent study which compared the MTurk platform to other alternative such as Prolific, CrowdFlower and other services, it resulted that while running same type of tasks and experiments on these platforms, the results were quite similar for both Prolific and MTurk while being superior to any other services utilized for testing

on the study. As noted on the study the biggest advantage presented by MTurk was the response time, appearing to be superior compared to the alternatives. [37]

3. SYSTEM DESIGN AND IMPLEMENTATION

This chapter introduces the tools, platforms and services used during the design and implementation of Article Enhancer¹. Accordingly, the front-end, back-end tools and moreover the 3rd party online service for acquiring results.

3.1. Chosen Tools and Platforms

3.1.1. JavaScript

The technologies made use of on the front-end are HTML alongside JavaScript. JavaScript, as a versatile client-side scripting language has become the language of the browsers, enabling building versatile user interfaces [38]. JavaScript is a high-level language and its association with the browsers makes it the most popular and broadly used programming language today, to handle operation on the client side, and not only, whereas it can be used for server-side scripting as well. Being flexible, fairly easy to learn and use and widely supported, it is the go-to to create a functional and interactive user interface. The chosen version of it, is Vanilla JS, supported by the jQuery, a library designed with the sole purpose of simplifying HTML DOM tree traversal, its manipulation, animating, event handling and Ajax interactions [39] making it a powerful side help while dealing with JS related operations.

3.1.2. Python

The server-side operations handling is done using Python programming language, together with the Django framework. As python is a reliable and popular language, and provides advantages for programmers, paired with Django, which is a high-level python framework exclusively for building web applications, becomes a really powerful tool for web development to take advantage of. The Django framework is an environment that grants a series of benefits worth mentioning as its stable and structured, highly secure avoiding typical mistakes like SQL-injection, cross-site scripting and request forgery, and moreover is fast and flexible supporting scalability in a big extent while being very versatile no matter the purpose of the system, making it the perfect choice for a project of this kind [40].

¹ <https://github.com/antokon/articleEn>

3.1.3. *Amazon's Mechanical Turk*

In this thesis project Amazon's Mechanical Turk is fundamental for the purpose of data retrieval, given a certain task. Based on previous studies concluded earlier which utilized the same service, resulting in reliable data sets, the Mechanical Turk is the perfect 3rd party crowdsourcing service to exploit towards our end.

3.2. Feature Overview

This system is a tool to help its users enrich their articles with content and references, the content may be visual such as embedded tweets in their article or something simpler in the form link to for a scientific article, that support the user's statements made. The following list summarizes the key features and design rationale of the tool:

1. The tool is designed with ease of use in mind and only requires a few steps to follow for the article augmentation to be completed.
2. Firstly, the user has to put his text by copy-pasting or just write up his text right on the text area available in the "Welcome" page, when the user is ready he can move to the next phase via the "Submit" button.
3. On the "Review" page the text is displayed in a non-editable format, here the tool offers the user the possibility to highlight key statement, that are to be augmented with content, the statement are selected via the mouse cursor and on "Highlight" button press the statement text becomes highlight. At the same time a table appears with the highlighted statements, and two more columns "Tweets" and "Article" references to be filled the user with the desired number of each for every highlight made.
4. To proceed further the user can do so using the "Send to Crowdsourcing" button, that saves the article and highlights to the database, then sends the information to Amazon's MTurk workers site utilizing the MTurk API, to fill the already existing template and redirects the user to a next page.
5. On the "Review" page, as soon as the crowd-workers' gathered content is returned, will be displayed, while the article is already present there. Each of the reference entries returned will be shown accompanied by a checkbox. At this point the user must review the content provided and after deciding what will be included, he has to click on the position on text where he wants this content embedded and check the checkbox.
6. After repeating this step for every desired highlight to be included, the user should press "Preview" button, and will be served a preview of his article with the embedded content with the option to download it.

To clarify, Article Enhancer supports two type of content to be fetched to support the highlighted statements: articles (any online articles) and tweets that align with the message the user highlighted.

3.3. User Interface

This web application consists of a simple and easy to use User Interface. The main goal for this UI was to be consistent, intuitive and familiar while aiming to provide the user with guidance while navigating the system.

As below shown on Fig. 1, the welcome page is very minimalistic in its design nature. It consists on a welcome note to the system, followed by a few guidelines to make it easier for the user to understand what this tool is, what it does and the way it works, a text area field appears placed in the center of the page ready to be filled with text by the user, either by copy-pasting it or by writing directly in the text area. When a user decides to be ready for the next step, he must press the “Submit” button as suggested also from the note present above the button, taking the user to the next page.

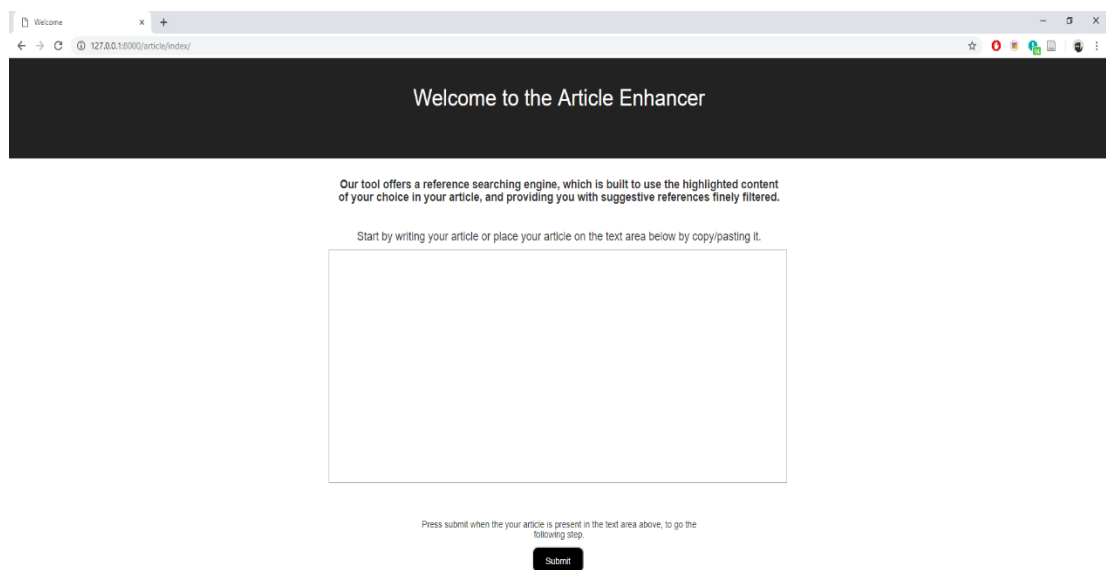


Figure 1. Welcome page of the system.

Fig. 2 serves as an example of the first page when text is entered by the user. The TextArea element of HTML permits for the field of text to extend based on the amount of text entered, so an article can be as long as the user desires or needs, while he can navigate it through the side scroll bar of the TextArea.

At this stage the text entered is saved using JavaScript and one of its features called local storage, at the moment the “Submit” button is pressed, so it would to be made available on the next page.

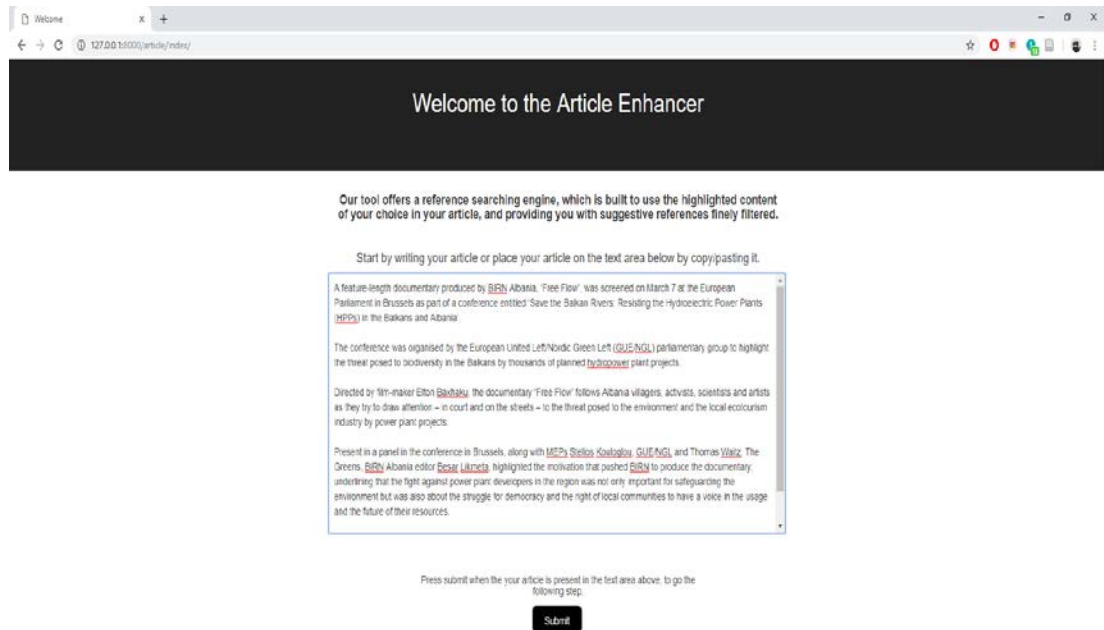


Figure 2. Welcome page view with an entered article.

Having placed and article in it spot on the welcome and begin ready to proceed forward for the next stage, the user must press submit. At this point a pop-out window saying, “Are you sure you want to submit this?” appears, giving the user the possibility to go back and make changes in case they are needed, as shown on Fig. 3 below.

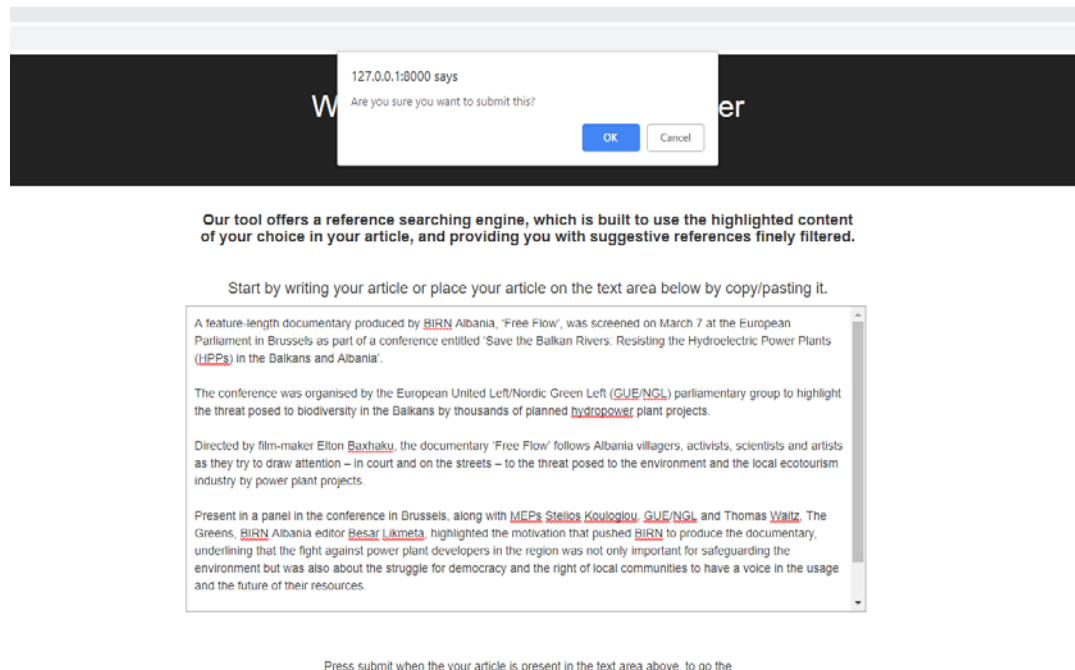


Figure 3. Prompting the user for proceeding.

The next stage gives on the client the possibility to decide and choose a number of statements available in the text, he would like to get supporting content to augment the article in order to raise its credibility.

The Highlight page (Fig. 4) redisplay the text in a read-only format and is as before accompanied by some instructions on how to use the tools this page offers.

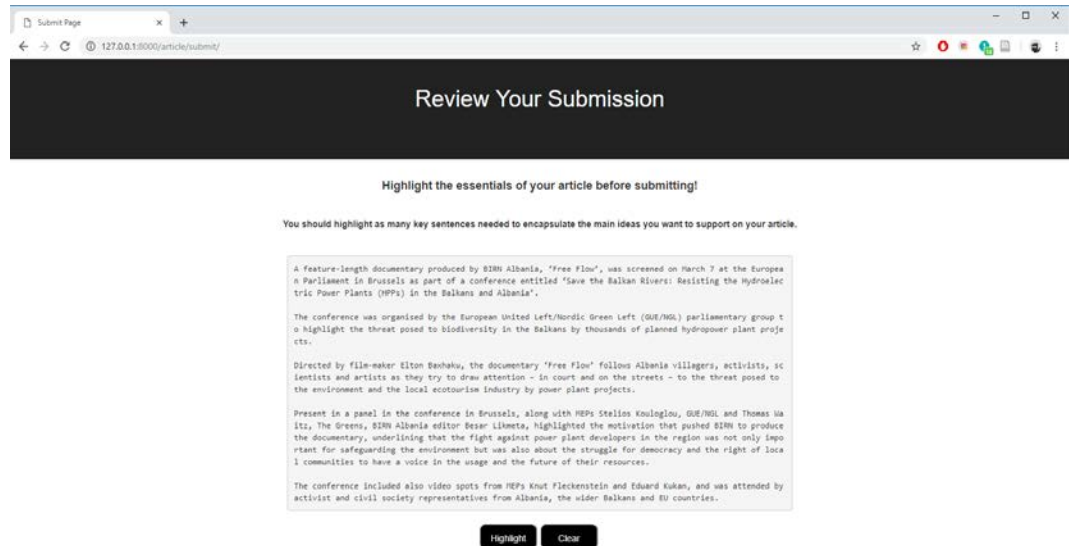


Figure 4. Highlight page view.

At this point the task at hand is highlighting key sentences, as many as needed to encapsulate the main ideas that have to be supported on the article. This is made possible by the “Highlight” button. All that it needed is the selection of each sentence on the displayed text via the mouse cursor, one at a time and pressing the button. The underlying JS code in this page catches each instance of cursor selection and on pressing “Highlight” this will marks the selected piece of text in a yellow color as the Fig. 5 demonstrates.

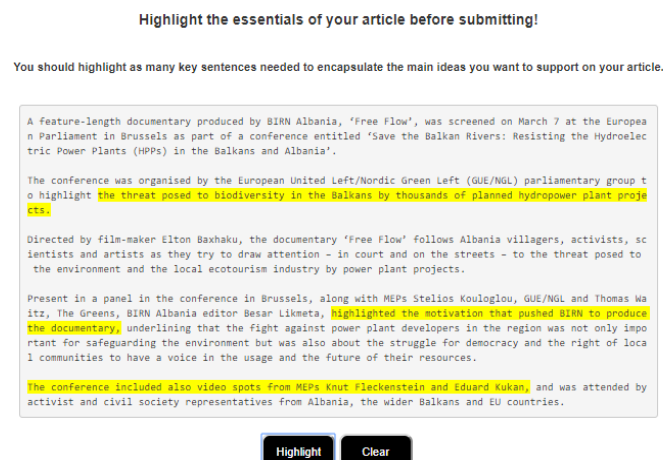


Figure 5. Text highlighting.

At the same time, a table emerges located on bottom-center position of the Highlight page as visible in Fig. 6, and it is made up from three columns:

“Highlight” – showing every highlight made by the user

“Tweets” – made from input boxes and requires the number of tweets content to be embedded, that the user needs for each highlighted statement

“Articles” – also showing with input boxes to be filled by the user with the desired number of articles references accordingly to every highlight

The screenshot shows a web browser window with the address bar displaying '127.0.0.1:8000/article/submit/'. The page title is 'Review Your Submission'. Below the title, there is a heading 'Highlight the essentials of your article before submitting!' and a sub-heading 'You should highlight as many key sentences needed to encapsulate the main ideas you want to support on your article.'

The main content area contains a text block with several highlighted sentences. Below the text, there are two buttons: 'Highlight' and 'Clear'. Below the buttons, there is a table with two columns: 'Highlights' and 'Tweets/Articles'. The table has three rows, each corresponding to a highlighted sentence in the text above. The first row has a '1' in the 'Highlights' column and an empty input box in the 'Tweets/Articles' column. The second row has a '1' in the 'Highlights' column and an empty input box in the 'Tweets/Articles' column. The third row has a '1' in the 'Highlights' column and an empty input box in the 'Tweets/Articles' column. Below the table, there is a 'Send to crowdsourcing' button.

Figure 6. Highlighted Content.

In case mistakes has been made or something in the highlights should be fixed, the button “Clear” can be used. On the click of this option, the highlights will be erased, the table cleaned, and everything will go back to the former state, providing the chance to adjust highlights. Whenever the client feels confident to move forward to other phase in the utilization of this system, clicking “Send to crowdsourcing” is the way.

The user should explicitly input a number of “Tweets” and “Articles” respectively, for the application to carry on with the next task in line. If the user fails to input the required information, a pop-up window as a reminder of the missing inputs manifests as Fig.7 indicates.

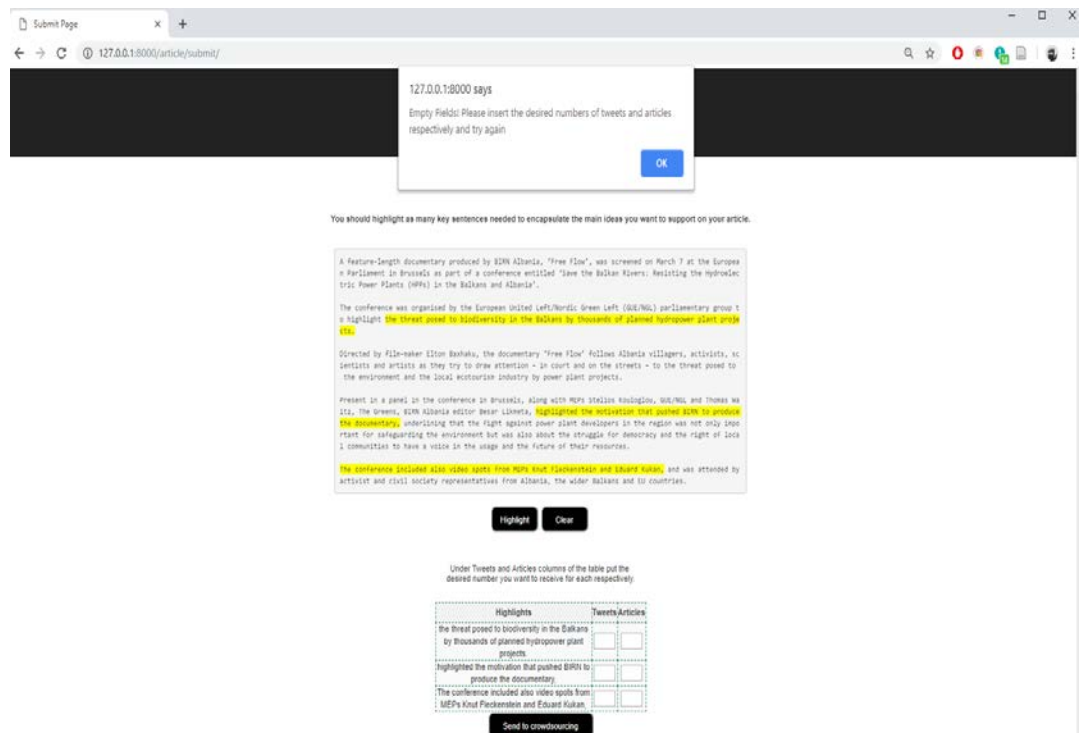


Figure 7. Pop-up window of missing inputs.

With everything being in order and “Send to Crowdsourcing” option is pressed the article along with data from the table is extracted from the User Interface with JS and through a JavaScript Object Notation (otherwise known as JSON) POST Request, after being aggregated in the JSON format is sent as a JSON object to the backend where this POST Request is handled with Python converting it into Python readable format and then saved in the SQLite Database, furthermore a the data is se are processed through.

After this step is completed the user will be redirected on to “Review” page, where a copy of the article is already present in a text area, and the content found and analyzed delivered by the MTurk workers is exposed as soon as it is ready, the user is encouraged to refresh the page after a few minutes so that the enriching content can be rendered.

Fig. 8 serves as a display of the Review Page that deals with the content and information coming from the crowd service. The left side the of this page is divided in two vertical section “Tweets” and “Articles”, the upper shows the Twitter tweets embedded in the system view, each associated with a dynamically created check-box element, and the same goes for the other types of references. The role of the check-boxes is to input the value of the respective augmenting elements provided by the crowd-workers, in the text of article. As emphasized even from the note on the top of the page, the order of operations in this task is as follows:

- As the user can better decide which elements complete the context that is to be augmented, is free to use or ignore the content provided at his will.
- The user clicks on the editable text area contained their articles at the precise position where he needs the elements to be embedded.
- The next step consists on clicking the check-box of the respective element

- This operation is to be repeated for every augmentative element, resulting in the inputting of embedding of elements in the text.

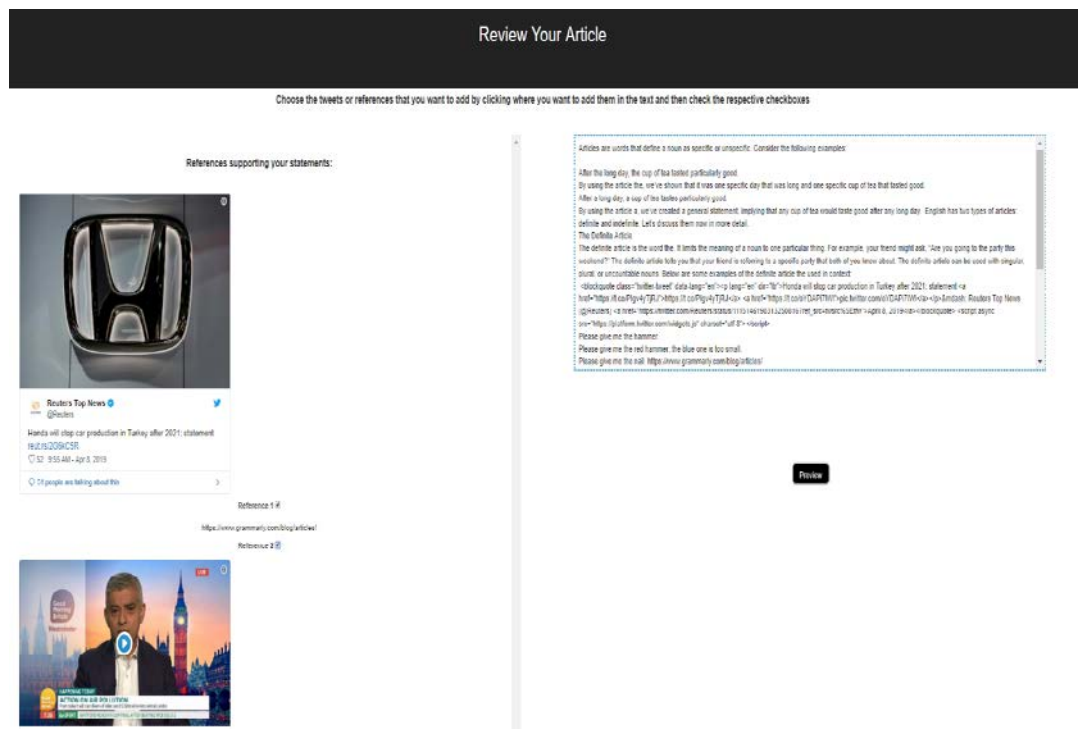


Figure 8. Review Page.

After the addition of elements is finished, the text on right side will look as in Fig. 9, where the tweets and article references are visible in the form of a embed link for tweets or a simple URL for scientific articles. Another essential feature of the system is introduced in the figure above, and that is the “Preview” option, that arranges for the writer of the article to get a glance at what his article will look like, preparing the user for the final product he will get out of this tool, and provided that a change is necessary to the looks of the article, the user is allowed to go back and reposition the embedded content to his liking.

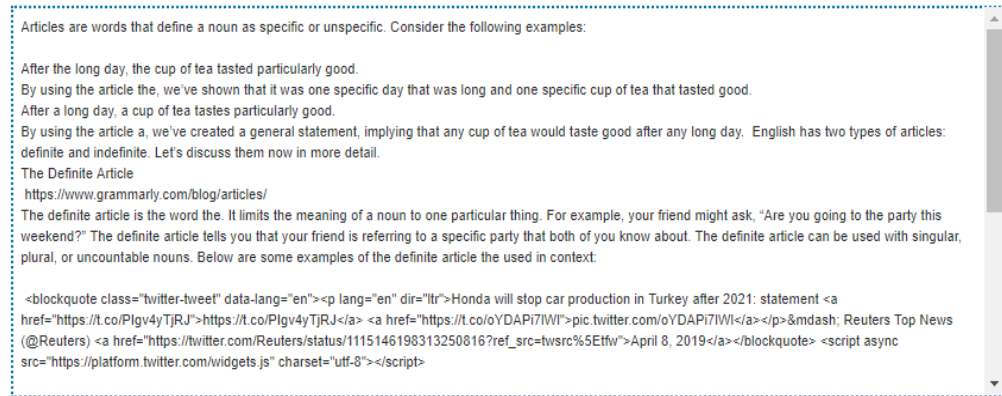


Figure 9. Text with embedded reference elements.

Fig. 10 serves as an example of an augmented article with content recovered with the help of crowd-workers. While this is the last page of the User Interface, another fundamental feature is granted. Here, after reviewing the article for the last time the user can get a copy of it via the “Export” button available on the bottom. The written JS code converts the displaying article and makes it available for download as an HTML file.

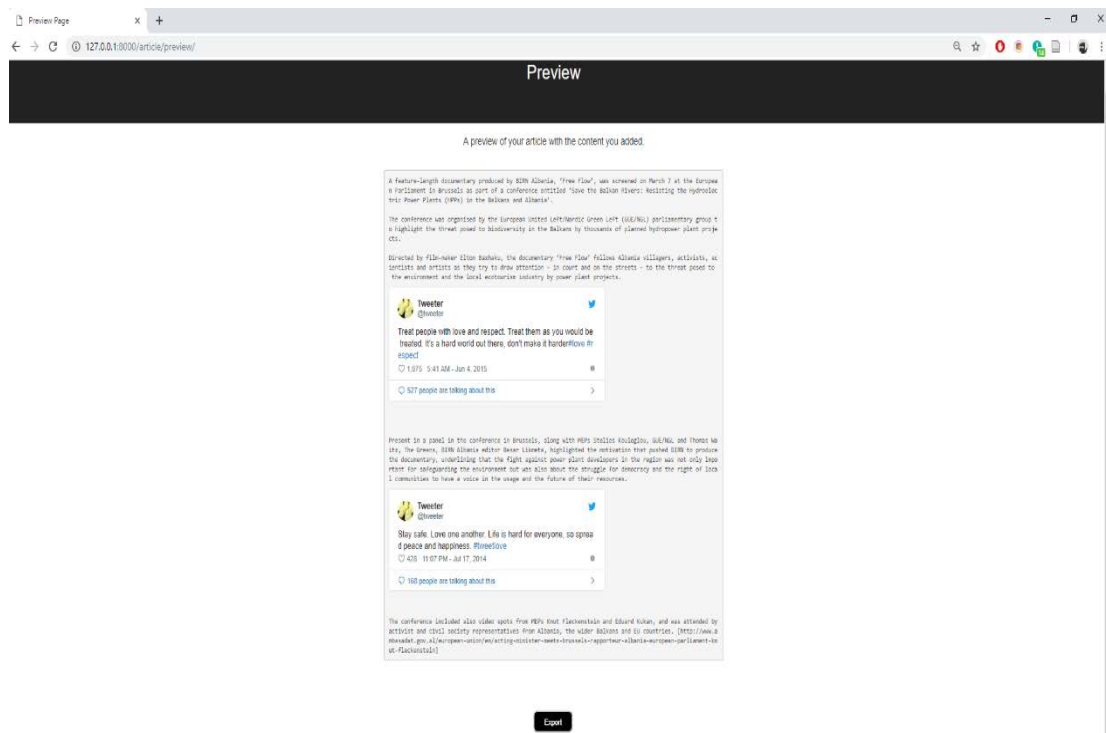


Figure 10. Preview Page.

3.4. Back-end Technology

The Django framework for python may prove a little challenging to be set up and configured, however once everything is ready, it is a very well organized and efficient one specifically designed for web development, proving fast and reliable with common web tasks.

Setting up the Django framework environment consists on rearranging the structure of the project inside a Django project following its predefined skeleton structure, including the front-end elements. Everything file including HTML, JavaScript and CSS has to be configured according to Django's set of rules and requirements to make everything work effectively.

A helpful feature of this framework worth mention is the Django Admin application, that comes by default with the creation of the skeleton project, and can use the models to automatically construct a site where the all the database entries can easily be managed, the admin can create new one or view, update, delete records right from this site, serving as a potent tool to save time during development and of course test the models as needed.

A Rest API requires first a Database, for storing and manipulating data, and for this Django uses models, using python objects to access and manage data. A model defines the structure of data to be store, and maybe referred otherwise as an independent underlying database. While Django grants the possibility of making use of various database technologies for our purpose, SQLite was chosen, since it fitted our needs.

The model created to save the data coming from the front-end is simple but suitable, keeping in mind that the data didn't need to be analyzed, organized or manipulated in any other way on our side before being sent to the MTurk service. This model class is based on three fields, "Articles", "Highlights" and "Publication date" as Fig. 11 presents an example through the Django Admin page view. The actual text provided from the users on the UI is saved together with highlights, and their specifications to ensure that the workers on the MTurk platform get the full text as a supplementary means of context to fully understand of what their task necessitates them to achieve.

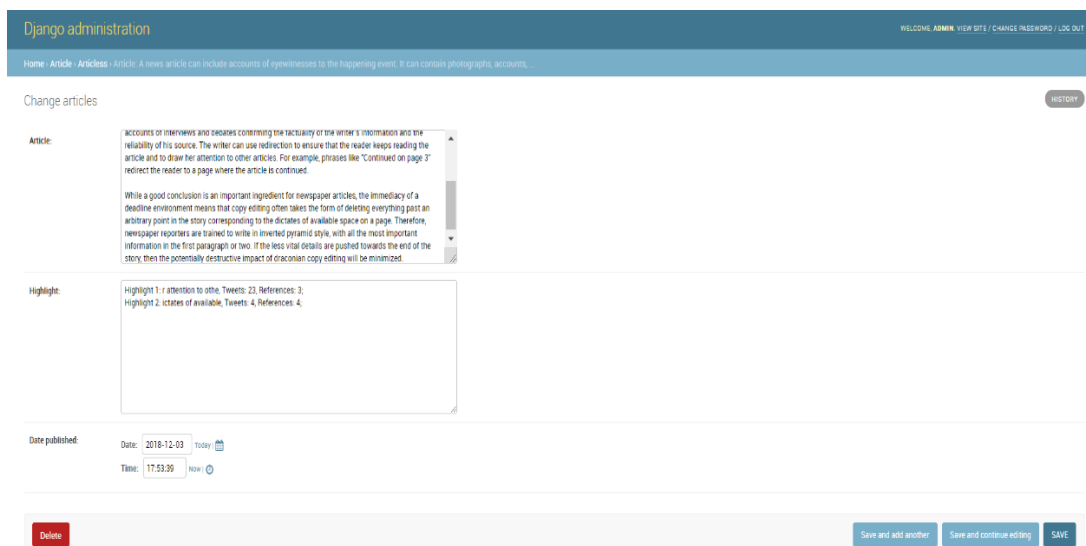


Figure 11. Django Admin Page – Viewing a single entry.

The data coming from the front-end is sent through a JSON call which is intercepted through the “POST” request and then this set of data are converted from JSON data into a readable format a python dictionary, transforming this data to a python data structure which are then saved in the database. At the same time a parallel JSON POST Request is sent to the back-end, that triggers the method that extracts the latest article and highlight entries on the DB and sends them to MTurk site filling up a pre-created template there.

The system is connected to the MTurk API, and does so through the Boto, the Amazon Web Services SDK for Python version 3. The MTurk API offers a feature called HITLayout which grants a developer a way of interacting with crowd-workers of the service, it is a reusable Amazon MTurk project template on the services’ website, which can be refilled dynamically with new information and show up as a new HIT(Human Intelligence Task) in the MTurk service where a worker can accept and fulfill the task. When a worker tries to complete a HIT, the template will load already filled with the information from the system’s end, from there the worker can complete the HIT and when finished, submit the response. Using this template makes it faster and easier to publish the task and the way that the results are collected. The system should extract each new entry created in the database so that it can be uploaded to MTurk template to create a new task ready to be completed.

The moment a new task is published on MTurk’ site, another operation is in place. which checks for the status of the last task posted in a periodic time frame of 10 seconds. When worker submits a completed task the is informed and starts the process of handling the response data. After the data is handled from a xml format to that of python dictionary, it is saved on the database on a model called “Answers”, which is related to the “Article” models through a foreign key, meaning that results received from the crowd-source service are link to the last article published on the database. This step reassures that only the latest results will be displayed on the “Review” page for the user.

While the user is refreshing the page waiting for the results, a query is done every time on server’s side to check if there are any entries added to the model responsible

for handling the results of the latest MTurk task. The query searches for results related to the latest article saved on the server via the foreign key. If the query is successful the enriching content provided by the workers are render to the UI, where the user proceed with the process of augmenting his article.

The schema represented on the Fig. 12 below serves as a depiction of the full interaction logic of between the user, the UI, server, and the 3rd party service used in this application. The user writes an article and highlights key statements he wishes to get content served from the crowd service for augmenting the article. This information is saved in the database, while on the server-side this data is extracted from the database and used to fill a pre-existing template, based on the service's requirements, where the crowd-workers can complete the task. After the task is completed the results are returned, saved on the server, and presented to the user, who then decide which ones to include in his article. The last step is that of exporting the final article.

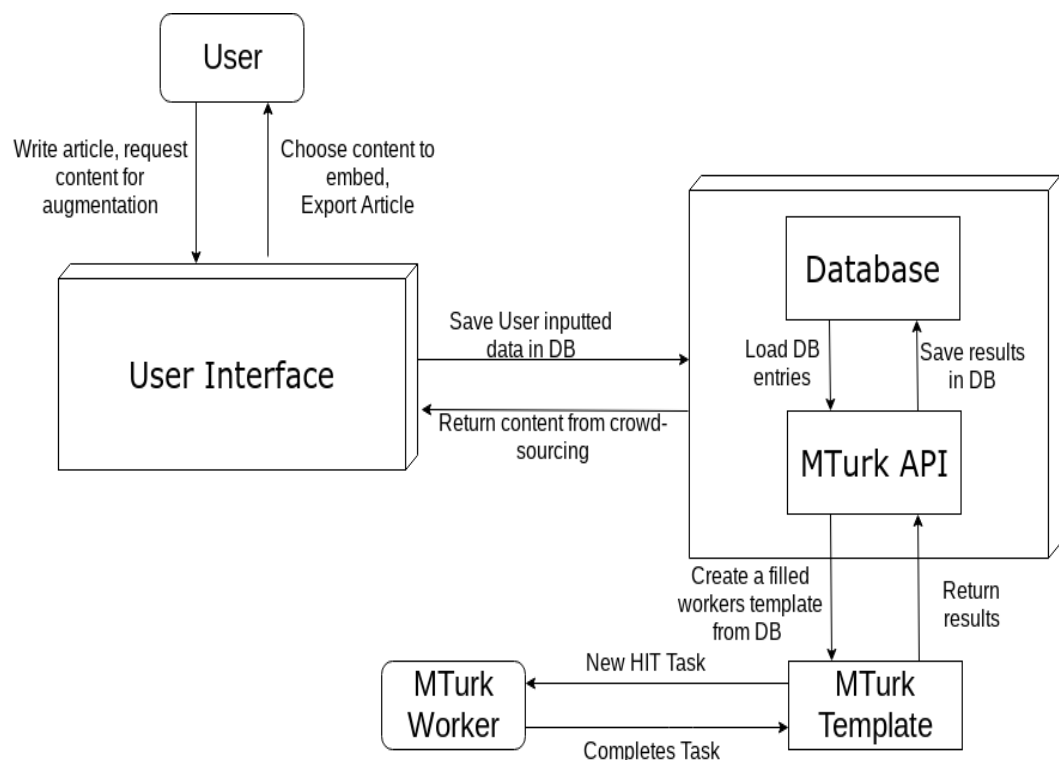


Figure 12. Schema of the interaction logic of the application.

3.4.1. MTurk Sandbox

For the purpose of testing and experimenting with MTurk, the service offers a virtual simulated environment called Sandbox. The sandbox is as functional version or as the real MTurk service, created to help the developers in the stages prior to publication. Since the real crowd-force offered is bound to be paid, Sandbox is a free alternative, no actual fees or payments are needed for utilizing it. As per switching the system's

interaction to the Amazon MTurk, all that is required is the changing of the host URL from the sandbox to the real production website.

The workers of the crowdsourcing service complete the tasks following a template filled with the data provided by the requester. In our case, having a particular use of the service, none of the pre-served templates by MTurk was suitable, leaving us with the alternative of creating our own template based on MTurk requirement. Fig. 13 displays the MTurk worker's template. The template includes instructions for the workers, so that they know what is asked from them and how to fill in the task requirements, "Article" and "Highlights" under which the article and highlights from each specific task will be loaded and as default there is only one field to return responses. In case the task is filled with more than one highlight, the workers can easily add more text area field by pressing the button "Add Row", and the JS code implemented in it will create them. The data which will be loaded in the template will be extracted from our database transformed in the format that the template demands.

Collect data from a Website
Requester: Antonio
Reward: \$0.43 per task
Tasks available: 0
Duration: 1 Hour

Instructions: Find and deliver reviewed content (tweets, article references) based on the statements and the number provided for each

Article (for context)

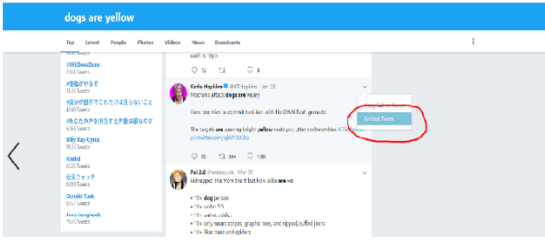
Article:

\$(article)

Highlights:

\$(name)

For each sentence under Highlight, search for tweets under <https://twitter.com/search-home> and return the respective embed tweet link, for references you can use <https://google.com> or <https://scholar.google.com>. (The picture below shows where to find embed link for a tweet.)



Every tweet or reference link should be submit in a new row!

Add Row
Submit

Figure 13. MTurk worker's template.

4. EVALUATION

For this system the functional testing was done continuously during the development time with a series of tests that were performed after every important advancement in the process of software development. For the purpose of experimenting while still in the development phase, a localhost web server and the virtual environment Sandbox, offered by Amazon's Mechanical Turk service, were used. For the evaluation of Article Enhancer User Experience, we carried out a user testing event in University of Oulu's surroundings. The evaluation process and results are described in this section.

This thesis is limited to the usability of Article Enhancer. The evaluation of the actual credibility would have been fairly trivial; however, it would have required substantial amount of financial resources for the work and due to time limitation, it was not feasible within the scope of this thesis.

4.1. Experimental Setup

To evaluate Article Enhancer, we deployed the web app in a real-life setting, a space oriented towards students known as Tellus, in the University of Oulu as shown below on Fig. 14. The participants were given a laptop already running the system as demonstrated also on Fig. 15, in order to avoid any delays that might have been caused by setting up their own devices. At the start of the evaluation process the users were introduced with the concept of the application and what this tool offered. To make the concept easier to grasp, every participant was presented with a script which included a scenario. After this step the users were encouraged to try the system on their own. At the end they were asked to fill out a survey sharing their opinion about the experience with the tool.

In this process 20 users between the range of 20 to 26 years old participated and gave their feedback and opinions. The group of participants was divided equally gender wise, having ten males and ten females participating. Every participant completed a questionnaire of twelve points, starting with basic demographic information such as age and gender, followed by eight of which were predefined pairs of attributes based on User Experience Questionnaire (<https://www.ueq-online.org/>) with a 7-point scale rating, with the user rating the application by choosing a number which they found more adequate towards one attribute or the other, a method which takes into account the pragmatic quality, the hedonic quality and also the attractiveness of the product. UEQ is a quick and dependable approach to assess the User Experience of an interactive product, established in 2005. Its assessment consists on measuring the overall attractiveness of the product, perspicuity, efficiency, dependability, stimulation and the novelty of product. For the purpose of this evaluation a short version of the UEQ was utilized along with a UEQ Excel-Tool which made the interpretation of the result easier. Also, the users were asked about their overall assessment of the system on a 5-point Likert Scale, while having an opportunity to express their own thoughts about the system and it could be improved in the last section of the questionnaire.

As the users were introduced to the script explaining how they should make use of the system, it was also explained to them that references fetched and presented to them coming from the crowd-sourcing platform wouldn't be from the Mechanical Turk workers, but from me, acting as a worker while making use of the sandbox environment, specifically design for testing through the development phase, free of charge. The participants were prompted to make use of any available text they might have in their possession such a paragraph from a course work of their thesis. Unfortunately, no one made use of this opportunity as they were initially skeptical about the system and the service it was offering. From all the participants it was reported that none had come across or used a tool resembling in any way to Article Enhancer before, leaving most of them thrilled with the concept of the system after the first encounter. The script is presented in Appendix 1.

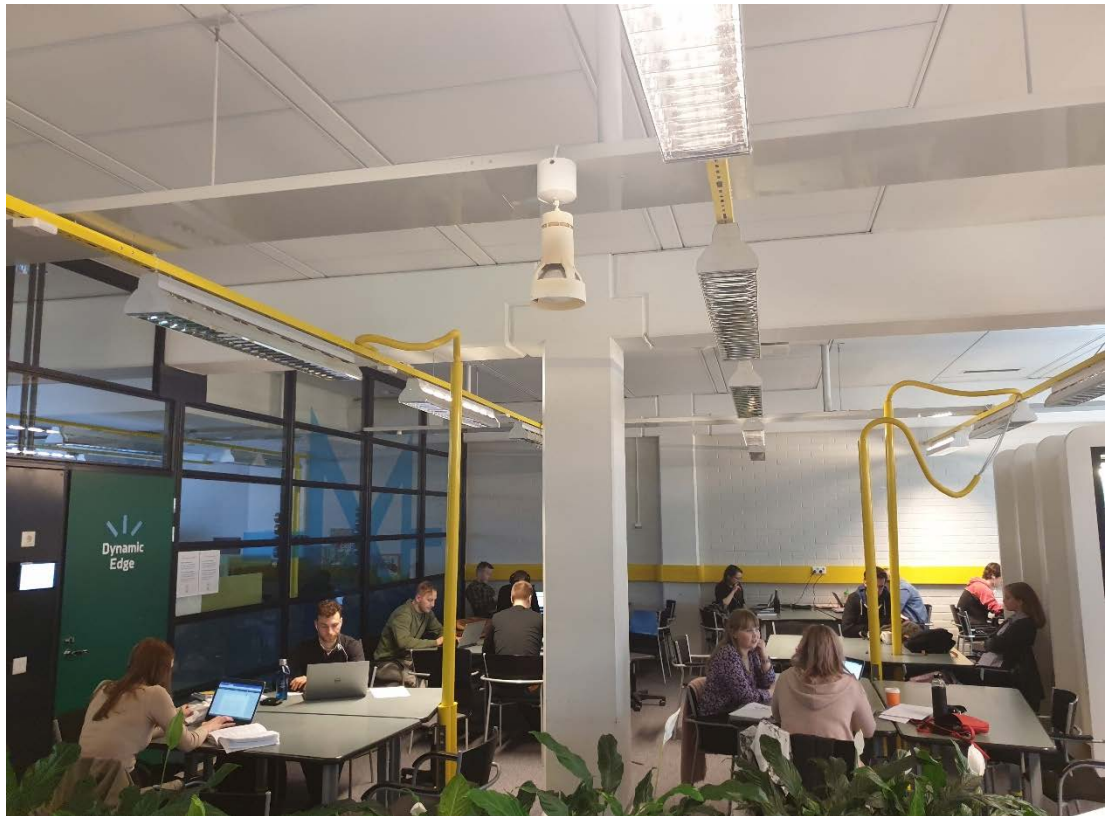


Figure 14. Tellus Area, University of Oulu



Figure 15. A participant using Article Enhancer during Evaluation

4.2. Questionnaire Data Analysis

The chosen statements in the questionnaire were based on 7-point scale rating system, and in evaluation the feedback from the users, a 1, 2 and 3 were considered as agreeing more with the statements placed on the left and the answers 5, 6, 7 were considered as agreeing with the opposed statements placed on the right, while a 4 meant that the user was conflicted as which statements described the system better or that neither of the attributes applies completely to the this product.

The main goal of this evaluation was to see if user experience the system supplied was satisfactory and intuitive. The first question presented referring to the UX asked the users if they found the interface of the system “Obstructive” or “Supportive”. Out of 20 participants all the 20 agreed with it being Supportive, five of them graded the question with a 5, eleven of them with a 6 and four of them gave it a 7. Based on these results obtained our users obviously see the interface of the system as highly supportive during the utilization. Next the users were questioned if the system was viewed as “Complicated” or “Easy”, resulting in five users agreeing with Complicated and fourteen users agreeing with Easy, while out of all the participants only one voted with a 4 meaning that he found the system nor too complicated or too easy to use. Even

though here the replies were a little more scattered the majority of the participants (14 users) found it to be easy to use.

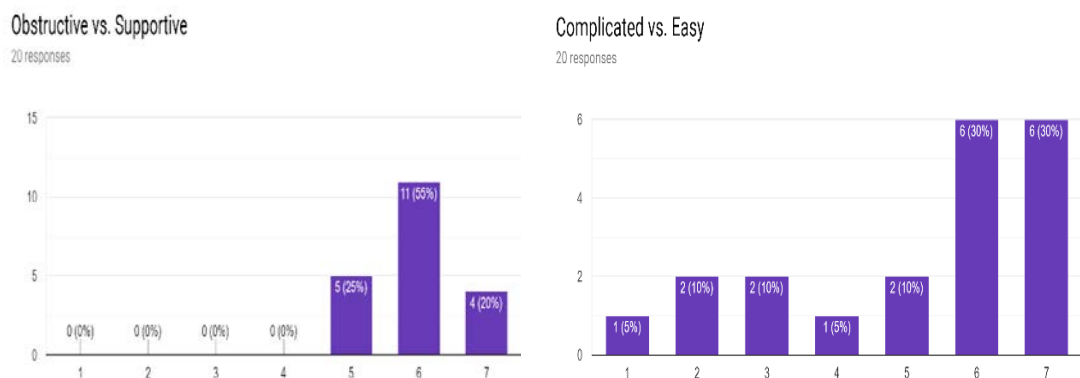
In order to understand the perception of the users whether our system proved to be efficient or not, giving them an opportunity to solve the tasks avoiding any redundant effort, they were asked to grade it on these terms and 18 out of 20 participant agreed with the application proving efficient while two users gave a 4. The next question in line questioned the clarity of the interface, twelve participants agreed to being “Clear”, six classified it as “Confusing” and two of them as neither confusing nor clear. Analyzing further into the attractiveness of the application, users were presented with another set of opposing statements, “Boring” and “Exciting”. Sixteen users agreed with exciting and four voted with a 4.

To investigate further if this type of product showed up as appealing, they were asked to grade the system as “Not Interesting” or “Interesting”, “Conventional” or “Inventive” and “Usual” or “Leading Edge”. For the first pair of attributes mentioned above 19 out of 20 agreed with it being Interesting and one graded as neither. On the next question sixteen people found it as Inventive, two as Conventional and two were unsure. For the last pair of attributes seventeen people agreed on the product being Leading Edge, one graded it as Usual and the rest (2 users) were inconclusive.

At the end of the survey the users, we specifically questioned the participants for their overall assessment of the system, the answers where based on 5-point Likert scale, varying from “very satisfied” to “very dissatisfied”. To this question, as also visible on Fig. 16, four of the participants answered with “very satisfied”, fifteen with “satisfied” and one as “unsatisfied”.

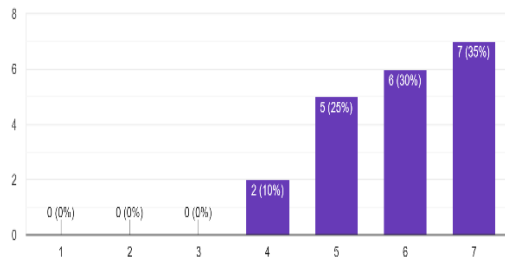
The results collected from the questionnaire can be viewed in Table 1. The evaluation questionnaire is presented in Appendix 2.

Table 1. The results of the evaluation questionnaire



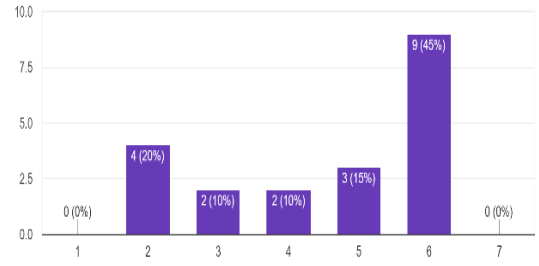
Inefficient vs. Efficient

20 responses



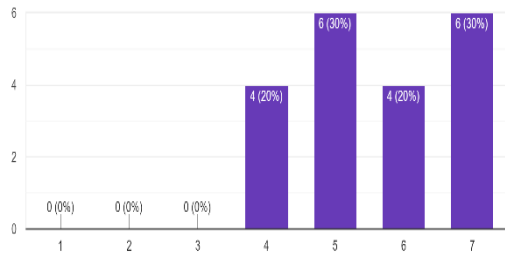
Confusing vs. Clear

20 responses



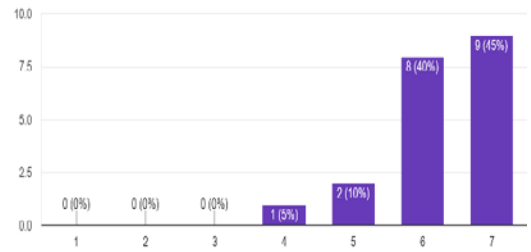
Boring vs. Exciting

20 responses



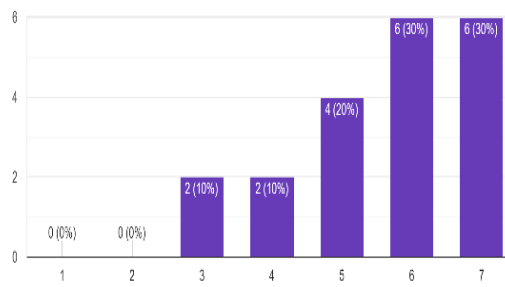
Not Interesting vs Interesting

20 responses



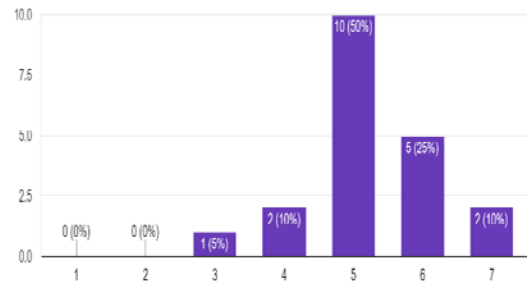
Conventional vs. Inventive

20 responses



Usual vs. Leading edge

20 responses



What is your overall assessment of the system?

20 responses

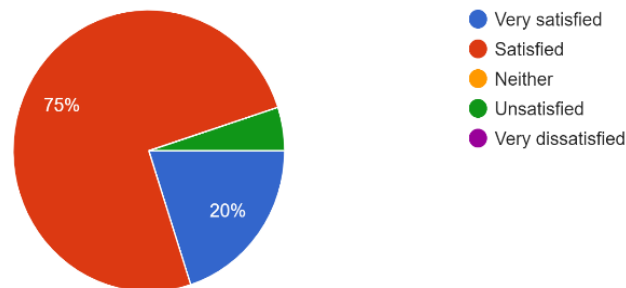


Figure 16. Results for the overall assessment of the system

On a more thorough analysis of data gathered through our questionnaire to get a better picture concerning the quality of a product, we used the UEQ Excel-Tool to compare the data from our evaluation with a benchmark data set containing data for 280 different product evaluations. The results are shown on Fig. 17. This benchmark helps in classifying the product quality under one of the five categories, ranging from *Excellent* to *Bad*, as displayed in the figure below. The pragmatic quality is defined from the first four pair of statement grading the system while, the hedonic quality is defined from the second one. As it can be clearly seen from the graph regarding to the pragmatic quality our system is listed in the range of above average, while for the hedonic quality as in the range of excellent resulting in the overall classification as good.

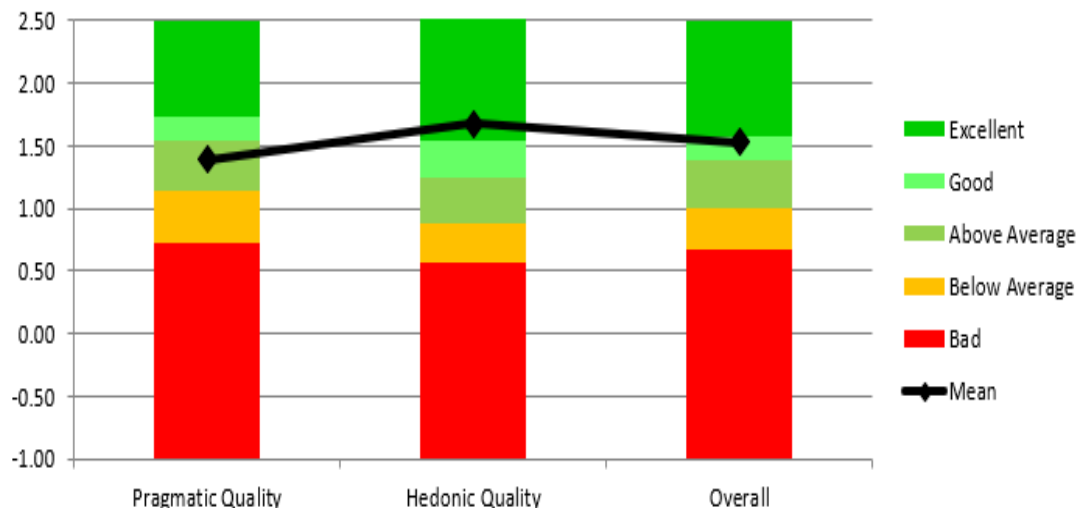


Figure 17. Benchmark graph for Article Enhancer.

The open question at the end of the questionnaire, where the participants were asked to share any thoughts or opinion regarding the system produced good and interesting feedback. Some of the participants recommended that providing more guidelines

through the application's interface would prove more efficient and intuitive for new users to get oriented while using it. Another very good suggestion made was for making the highlights created by the user, visible also in the review page which in turn would make it easier for user relate references with the highlights which they belong to.

5. DISCUSSION

This chapter discusses “Article Enhancer”. The achievement of goals, challenges influencing the development of this application and improvements to be implemented in order of maximizing the efficiency of the system. The fast and consistent pace at which technology continuously grows sets an ever-evolving standard that defines the new applications and services coming in life. The Article Enhancer has emerged as part of this “order” with the purpose of helping users to enrich and augment their creations with references and embedded content provided through a crowd-sourcing service. The underlying concept for this application rests in manipulating on how the user perceives the information given, by supporting statements made in the text also with visual content to make the reading experience more attractive, easier to understand and mainly to be comprehensive for the audience.

The scientific novelty of the project consists in combining methods from social computing and software engineering for the purpose of disseminating credible news augmented through textual and visual referencing elements to the target audience. We believe that our approach can contribute significantly for the web’s future goals in maintaining and improving media literacy and multiliteracy.

This thesis provides a tool for ultimately investigating how different content and elements in news dissemination towards web users influence their perception of the credibility of the information presented to them. This is very important for educating the people in today’s media environment and for the development of media literacy education itself. The second contribution of this work is a new interactive tool for academics, that allows for embedded content elements to be included into an article giving them the possibility to augment their piece of information and manipulate how the audience perceives the articles.

While in this thesis we could not fully investigate all the planned aspects of the tool itself, we did build the tool, release it as open-source, and evaluate its perceived usefulness and novelty with higher education students.

5.1. Achievement of goals

Firstly, the thesis’s main objective was to develop a functional system that serves as a tool for users to reference their article while being relieved from the troubles of going through the process themselves. Users could simply write or place their articles and take advantage of the Article Enhancer to enrich their text, with information in form of references (textual and visual). The reference elements are provided by the crowdsourcing platform Mechanical Turk; workers ensure the completion of the task by searching, analyzing and delivering matching references to users’ request.

Secondly, Article Enhancer aimed at having a UI for the system, which would feel familiar and intuitive to users. According to data from the users’ questionnaire, the system has proved to be easy to use and overall supportive. Moreover, as shown in Fig.14, the overall attractiveness of the product, based on the data collected, ranks as ‘good’, with a sufficient sample size (n=20).

Finally, the main aim of this project is to address the question of how scientific articles can be made more credible, before the dissemination to the public. Based on the feedback we obtained during the evaluation, all participants of the study believed that adding and embedding content, especially tweets as visual content, was a very inventive way to make articles become more credible in the eye of the reader. Naturally, more research into this is needed with authentic articles and readers recruited e.g. from online crowdsourcing marketplaces such as MTurk discussed earlier in this thesis.

5.2. Future work

Although Article Enhancer is a functional web application, there are still a lot of improvements and features that can be implemented further. For the purpose of this thesis and the evaluation of system the virtual environment Sandbox was utilized, however the Amazon's MTurk is a paid crowdsourcing service, each task submitted to the platform for completion has a set fee which is fixed to 50 cents and as so payment logistics through the Article Enhancer should be implemented.

Additional improvement can also be made on the front-end as suggested even from the open answers question on the questionnaire during the evaluation phase, for example the on Review page where the content provided by the MTurk platform is displayed, highlights can be included making it easier for the user to relate the references from the crowd service with the statements they highlighted.

A challenge faced during implementation was the deciding on the limitations for the number of references a user could ask for each highlighted statement, due to how the MTurk platform works, where the price for each task is set before hands by the requester whom posts the task. Letting the user to input for an unlimited number of the refences would have proven tricky since no worker would accept a task heavy on work with an unfair payment. In order to address this problem, we set a baseline of three maximum tweets and 10 maximum article references for each highlight. Anyhow, this baseline was considered as a first step towards a better solution for balancing the workload and the price paid for a task.

An approach to rank tweets previously served from the crowdsourcing through Article Enhancer, is a fresh concept planned for the future of this system. This new feature would make possible to have a ranking system for tweets, so users could rate the tweets presented (e.g., on a scale from 1 to 5) leaving a useful feedback for other users, with the end-goal of creating a reliability scale for tweets. This addition to the features of the system, will contribute in having an even more unique piece of software.

Yet another improvement can be the further implementation of accommodating more types of referencing elements, considering that for now the content served through Article Enhancer is limited to distinct types, such as tweets and articles, it can be expanded to support more types such as graphs.

6. CONCLUSION

This thesis discusses and examines various topics revolving around increasing the credibility of scientific dissemination using crowdsourcing. Media literacy and credibility, online social media and fake news dissemination, crowdsourcing and platforms offering this service are a part of the topics discussed. Article Enhancer was presented as the main outcome of this thesis.

Article Enhancer is a modern web application that takes advantage of crowdsourcing technology. The system supplies a user with referencing elements, gathered through MTurk platform, to augment their article. The references can be traditional ones (article references) or tweets from Twitter. The reason behind the usage of tweets as reference content option is the augmentation of the essence of the article by supporting statements made by users also visually, influencing how readers perceive the information.

Related work was also presented in this thesis. This literature review examines system which offer same or similar functionalities to Article Enhancer. Although Article Enhancer is semi-automated system, its strongest point compared to the other systems, is that it doesn't require extra human effort to enrich articles especially with visualization content, and providing already existing content on the web avoiding the process of creating new content, making it a fresh approach in this line of software service.

The present state of the system doesn't include all planned features, some of which were not implemented due to practical issues and time limitations. Nonetheless these features would be a great addition in the future making the system even more novel while contributing in a better experience in the users' end.

The testing of Article Enhancer occurred in real life setting, on University of Oulu's ground, studying oriented space for student, called Tellus. This testing proceedings helped in determining that the system appears alluring and attractive to new users, a sign that this tool would be well accepted through the research community on its deployment. In the evaluation of Article Enhancer, 20 users of various nationalities and age participated, while answering a questionnaire made of 12 points focused towards the UX of our system and expressing their own thoughts on the last point regarding their the experience, functionalities and encouraged to offer their suggestions about improvements for the application.

7. REFERENCES

- [1] Chen, Y. , Conroy, N. J. and Rubin, V. L. (2015), News in an online world: The need for an “automatic crap detector”. *Proc. Assoc. Info. Sci. Tech.*, 52: 1-4. doi:10.1002/pr2.2015.145052010081
- [2] Harris, R. (1997). *Evaluating Internet Research Sources*.
- [3] Nguyen, An. 2008. “Facing ‘The Fabulous Monster’ The Traditional Media’s Fear-Driven Innovation Culture in the Development of Online News.” *Journalism Studies* 9 (1). Routledge: 91–104.
- [4] M. J. Lazer, David & Baum, Matthew & Benkler, Yochai & J. Berinsky, Adam & M. Greenhill, Kelly & Menczer, Filippo & J. Metzger, Miriam & Nyhan, Brendan & Pennycook, Gordon & Rothschild, David & Schudson, Michael & Sloman, Steven & Sunstein, C & A. Thorson, Emily & J. Watts, Duncan & L. Zittrain, Jonathan. (2018). The science of fake news. *Science*. 359. 1094-1096. 10.1126/science.aao2998.
- [5] Ciampaglia GL, Shiralkar P, Rocha LM, Bollen J, Menczer F, et al. (2015) Correction: Computational Fact Checking from Knowledge Networks. *PLOS ONE*
- [6] Final report of the High Level Expert Group on Fake News and Online Disinformation (2018). Accessed from <https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>
- [7] Gross, K., & D’Ambrosio, L. (2004). Framing Emotional Response. *Political Psychology*, 25(1), 1–29. <https://doi.org/10.1111/j.1467-9221.2004.00354.x>
- [8] LIVINGSTONE, S. (2004). Media Literacy and the Challenge of New Information and Communication Technologies. *The Communication Review*, 7(1), 3–14. <https://doi.org/10.1080/10714420490280152>
- [9] Maier, S. R., Mayorga, M., & Slovic, P. (2017). Personalized news stories affect men as well as women. *Newspaper Research Journal*, 38(2), 172–186. <https://doi.org/10.1177/0739532917716178>
- [10] Maier, S. R., Slovic, P., & Mayorga, M. (2017). Reader reaction to news of mass suffering: Assessing the influence of story form and emotional response. *Journalism*, 18(8), 1011–1029. <https://doi.org/10.1177/1464884916663597>
- [11] Baack, S. (2011). A new Style of News Reporting: Wikileaks and Data-driven Journalism. *Cyborg Subjects*, 1-10. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-400253>

- [12] Delgado, D., Magalhaes, J., & Correia, N. (2010). Assisted News Reading with Automated Illustration. In *Proceedings of the 18th ACM International Conference on Multimedia* (pp. 1647–1650). New York, NY, USA: ACM. <https://doi.org/10.1145/1873951.1874311>
- [13] Marcus, A., Wu, E., & Madden, S. (2013). Data In Context: Aiding News Consumers while Taming Dataspaces, 4.
- [14] Hullman, J., Diakopoulos, N., & Adar, E. (2013). Contextifier: automatic generation of annotated stock visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (p. 2707). Paris, France: ACM Press. <https://doi.org/10.1145/2470654.2481374>
- [15] Lin, A. Y., Ford, J., Adar, E., & Hecht, B. (2018). VizByWiki: Mining Data Visualizations from the Web to Enrich News Articles. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18* (pp. 873–882). Lyon, France: ACM Press. <https://doi.org/10.1145/3178876.3186135>
- [16] Savva, M., Kong, N., Chhajta, A., Fei-Fei, L., Agrawala, M., & Heer, J. (2011). ReVision: automated classification, analysis and redesign of chart images. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11* (p. 393). Santa Barbara, California, USA: ACM Press. <https://doi.org/10.1145/2047196.2047247>
- [17] Wongsuphasawat, K., Moritz, D., Anand, A., Mackinlay, J., Howe, B., & Heer, J. (2016). Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. *IEEE Transactions on Visualization and Computer Graphics*, 22(1), 649–658. <https://doi.org/10.1109/TVCG.2015.2467191>
- [18] Shu K., Sliva A., Wang S., Tang J., Liu H. (2017). Fake News Detection on Social Media: A Data Mining Perspective, ACM SIGKDD Explorations Newsletter, v.19 n.1,
- [19] Shao, Chengcheng & Luca Ciampaglia, Giovanni & Varol, Onur & Flammini, Alessandro & Menczer, Filippo. (2017). The spread of fake news by social bots.
- [20] Vosoughi S, Roy D, Aral S. The spread of true and false news online. *Science*. 2018; 359:1146–1151. doi: 10.1126/science.aap9559.
- [21] Bessi, A. & Ferrara, E. Social bots distort the 2016 US presidential election online discussion. *First Monday* 21, 11 (2016).
- [22] Ferrara, E. Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election. *First Monday* 22, 8 (2017).
- [23] Shao, Chengcheng & Luca Ciampaglia, Giovanni & Varol, Onur & Flammini, Alessandro & Menczer, Filippo & Yang, Kai-Cheng. (2018). The

spread of low-credibility content by social bots. *Nature Communications*. 9. 10.1038/s41467-018-06930-7.

- [24] Brandtzaeg, P. B., & Følstad, A. (2017). Trust and Distrust in Online Fact-checking Services. *Commun. ACM*, 60(9), 65–71. <https://doi.org/10.1145/3122803>
- [25] Wang, J., Kraska, T., Franklin, M. J., & Feng, J. (2012). CrowdER: crowdsourcing entity resolution. *Proceedings of the VLDB Endowment*, 5(11), 1483–1494. <https://doi.org/10.14778/2350229.2350263>
- [26] C. Brabham, D. (2008). Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence: The International Journal of Research Into New Media Technologies*, 14, 75–90. <https://doi.org/10.1177/1354856507084420>
- [27] Abramovich, G. (2013, March 7). 5 Brands Winning at Crowdsourcing. Retrieved April 11, 2019, from Digiday website: <https://digiday.com/marketing/5-brands-winning-at-crowdsourcing/>
- [28] Kittur, A., Smus, B., Khamkar, S., & Kraut, R. E. (2011). CrowdForge: Crowdsourcing Complex Work. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (pp. 43–52). New York, NY, USA: ACM. <https://doi.org/10.1145/2047196.2047202>
- [29] Nowak, S., & Rüger, S. (2010). How Reliable Are Annotations via Crowdsourcing: A Study About Inter-annotator Agreement for Multi-label Image Annotation. In *Proceedings of the International Conference on Multimedia Information Retrieval* (pp. 557–566). New York, NY, USA: ACM. <https://doi.org/10.1145/1743384.1743478>
- [30] Han, H., Otto, C., Liu, X., & Jain, A. K. (2015). Demographic Estimation from Face Images: Human vs. Machine Performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(6), 1148–1161. <https://doi.org/10.1109/TPAMI.2014.2362759>
- [31] Li, H., Shen, H., Xu, S., & Zhang, C. (2015). Visualizing NLP annotations for Crowdsourcing. *ArXiv:1508.06044 [Cs]*. Retrieved from <http://arxiv.org/abs/1508.06044>
- [32] Lease, M. (2011). On Quality Control and Machine Learning in Crowdsourcing. In *Proceedings of the 11th AAAI Conference on Human Computation* (pp. 97–102). AAAI Press. Retrieved from <http://dl.acm.org/citation.cfm?id=2908698.2908717>
- [33] Cole, F., Sanik, K., DeCarlo, D., Finkelstein, A., Funkhouser, T., Rusinkiewicz, S., ... Singh, M. (2009). How well do line drawings depict shape? *ACM Transactions on Graphics (TOG)*, 28(3), 28. <https://doi.org/10.1145/1576246.1531334>

- [34] Komarov, S., Reinecke, K., & Gajos, K. Z. (2013). Crowdsourcing performance evaluations of user interfaces (pp. 207–216). Presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM. <https://doi.org/10.1145/2470654.2470684>
- [35] Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon’s MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29(6), 2156–2160. <https://doi.org/10.1016/j.chb.2013.05.009>
- [36] Palan, S., & Schitter, C. (2017). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*. <https://doi.org/10.1016/j.jbef.2017.12.004>
- [37] Beyond the Turk: Alternative platforms for crowdsourcing behavioral research - ScienceDirect. (n.d.). Retrieved April 11, 2019, from <https://www.sciencedirect.com/science/article/pii/S0022103116303201>
- [38] Crockford, D. (2008). *JavaScript: The Good Parts: The Good Parts*. O’Reilly Media, Inc.
- [39] Narayan, S. (2011) What is jQuery and How to Start using jQuery? Retrieved March 11, 2019, from <https://www.codeproject.com/Articles/157446/What-is-jQuery-and-How-to-Start-using-jQuery>
- [40] Djangoproject.com. (2019). *Django overview*. Retrieved May 1, 2019, from <https://www.djangoproject.com/start/overview/>

8. APPENDICES

Appendix 1. Evaluation Script

Appendix 2. Evaluation Questionnaire

Appendix 1. Evaluation Script

Scenario:

You have written an “article” and you want to enrich the text with visual content and/or scientific paper references. For this, you will use “Article Augmenter”; a semi-automated system, which provides reference content for text, for relieving the user from the hassle of doing it themselves.

You can use any text available you have, if you are working on your thesis or some coursework, you could use a paragraph from it.

Otherwise you can use the text provided below. Start by reading the text first then proceed on using the system.

After using the system, you will be asked to fill in a brief survey.

<https://forms.gle/skg4nMC2DKd7pmwz9>

Text:

All efforts to fight climate change face the money test: Are the benefits of stopping global warming — and avoiding sea level rise, heat waves, and wildfires — greater than the costs? The dollar balance we arrive at should be one of the biggest factors in deciding what we’re willing to do to tackle the problem, whether that’s shuttering all coal plants or building thousands of nuclear reactors. Some groups have taken a stab at calculating what climate change will cost the world, or conversely, how much humanity would save by becoming more sustainable. Earlier this month, the Global Commission on the Economy and Climate tallied the number at a truly massive \$26 trillion in savings by 2030. Getting a slice of those savings requires figuring out which actors stand to lose the most as the climate changes, whether that’s countries, companies, or even individuals. And this is where the idea of the social cost of carbon comes in. It’s a policy tool that attaches a price tag to the long-term economic damage caused by one ton of carbon dioxide, hence the cost to society. It’s related to a carbon tax (more on that below), and it serves as a way to distill the vast global consequences of climate change down to a practical metric. Critically, it’s also the foundation of US climate policies, including the Clean Power Plan. Revising this number down has been a key part of the Trump administration’s strategy to roll back environmental rules. Under Obama, the social cost of carbon was set at \$45 per ton of carbon dioxide; under Trump, it’s as little as \$1. A new study published Monday in the journal *Nature Climate Change* calculates the social cost of carbon down to individual countries. This adds an important bit of nuance because climate change is going to cost some countries more than others, a fact that’s lost when you try to tabulate a global average. The team found a global social cost of carbon vastly higher than many previous estimates, drawing on more recent climate projections and more robust macroeconomic models. The results also highlighted the fundamental injustice of climate change: Many of those who contributed the least to the problem stand to suffer the most.

Boring vs. Exciting *

1 2 3 4 5 6 7

Boring Exciting

Not Interesting vs Interesting *

1 2 3 4 5 6 7

Not Interesting Interesting

Conventional vs. Inventive *

1 2 3 4 5 6 7

Conventional Inventive

Usual vs. Leading edge *

1 2 3 4 5 6 7

Usual Leading edge

What is your overall assessment of the system? *

- Very satisfied
- Satisfied
- Neither
- Unsatisfied
- Very dissatisfied

If you have any thoughts about the system feel free to share them.

Your answer

SUBMIT