

Wikipedia

BA 186 Final Paper

Submitted to:

Mr. Roy Canseco

Submitted by:

Margarita Agoncillo Justyn Go Joshua Lu Kathleen Suarez Kimberly Tan

TABLE OF CONTENTS

I. Plann	ing	2
	Main Purpose, Usage, and Users	
	Processes Behind Wikipedia	
	Strengths, Weaknesses, and Possible Improvements	
	Planning: Section Relevance and Reflection	
II. Analy	vsis .	4
	Data Structuring	
	Decision Table	
	E-R Diagram	
	Features	
	Attributes and Candidate Keys	
	Cardinalities	
	Receipt for Transactions	
	Analysis: Section Relevance and Reflection	
III. Desi	gn	8
	Database Report	
	Reports, Forms, and Interfaces	
	Design: Section Relevance and Reflection	
IV. Appe	endix	13
V. Biblio	ography	19

I. PLANNING

Because Wikipedia is not Philippine-based nor represented in the country, our group has resorted to doing online research (using sources besides Wikipedia, of course) to fulfill this class submission. After which, our group met up to discuss, in a sort of nominal group technique (NGT) form of interview. Our group decided to employ NGT as the primary tool for discussing Wikipedia's system requirements as this method allowed us to come up with our own understanding and opinions of Wikipedia (based on our individual experiences with using Wikipedia and individual basic online research) before meeting and brainstorming as a group. Being ready with our own ideas regarding the Wikipedia system allowed us to have a structured, fruitful, and efficient discussion on the said topic, so we had a relatively easy time when it came to highlighting main points and collating information.

Other traditional methods (e.g., individual interviews and direct observation) were not used as they were time-consuming with regard to follow up discussions and impractical as we mostly have similar backgrounds (college students taking up business administration in a Philippine state university and have previously used wikipedia for academic and other reasons). More contemporary and thorough methods such as Joint Application Design, although hypothetically, could have been more useful and effective than NGT, were not employed given our limited access to other stakeholders (e.g., only have access to key users, not managers and systems personnel of Wikipedia who are based abroad).

Main Purpose, Usage, and Users

The main purpose of Wikipedia is to "benefit readers by acting as an encyclopedia." It acts upon the notion that information should be free and accessible to anyone who needs it wherein the user can get answers within seconds after a few clicks. The only prerequisite of being able to use Wikipedia is that the user needs to have both an electronic device like a computer and access to internet connection. Since Wikipedia can be accessed to anyone so to speak, it has a broad range of users. Users include students, the academe, professionals in various fields, researchers, journalists and can include uses as basic as someone who simply wanted clarity after watching a complex movie.

According to Nupedia, it was created by accident, to serve as a younger sibling of entrepreneur Jimmy Wales and Lawrence Sanger's first failed creation. They originally wanted to create an online platform that produced short write-ups or articles (similar to the format seen in Wikipedia today) which were written professionally and peer reviewed by experts before being published online. The main problem however was the speed of processing these articles under review - only two were up online after half a year. Thus, the idea to create a counterpart wherein works could be published or contributed to by anyone with little to no editorial review was made and Wikipedia was born. Wales envisions Wikipedia to be a neutral collaborative encyclopedia that will one day contain the sum and totality of human knowledge, which he believes the site is closer to achieving than anybody else.

Processes Behind Wikipedia

Wikipedia works through the contributions of anyone. In the same way physical encyclopedias were written by experts in the topic at hand, each topic in Wikipedia is written about by contributors (could be just anyone) who know anything about the topic. To prevent spam, profanity, or other unwanted content, Wikipedia relies on its structure as a *community*. This community is composed mostly of readers, and also includes writers, editors, or administrators. Together, the members of the community keep content not just in line, but surprisingly (to most), in-depth and accurate— from spelling to grammar to substance. The community is composed of millions of people who choose to be notified of changes to topics of interest.

No edits to any page/topic will go unnoticed for very long. Wikipedia operates in five different data centers: three in the USA, one in Europe, and one in Singapore.

Strengths, Weaknesses, and Possible Improvements

Wikipedia is using a method similar to crowdsourcing for its content since it allows people from all over the web to contribute and correct information which gives the platform multiple strengths and weaknesses. Below is a table of its potential strengths and weaknesses based on the insights of the group:

STRENGTHS	WEAKNESSES
 Convenient, accessible, free Easy-to-understand but more comprehensive than its physical counterpart Large volume of information Lots of secondary sources viewers can access and look into Large traction No problems with funding (i.e. consistent inflows of donations) 	 Anybody can contribute (e.g. non-experts, problem contributors) May not be completely correct or reliable in very specialized areas Process of screening information done by the community or volunteer editors may be flawed due to the large influx of information.

A suggestion culminated from the group's discussion is to improve and build on the current "Wikimedia servers" that Wikipedia and its other projects run to check technical errors such as distribution of information to various data centers, balancing incoming requests through commodity servers, caching of database query and computation results, and many others. To lessen the load on the *community*, we propose to include or create another set of servers to check superficial grammatical errors and automatically rearranges words to follow the "Manual of Style" volunteers refer to to keep a consistent voice on all Wikipedia articles. It would be like grammarly on Wikipedia but smarter.

Planning: Section Relevance and Reflection

Planning is an extremely important phase of any endeavour since this is the stage where one thinks and brainstorms the scope, required activities, resources needed and where to obtain these resources in order to reach a desired goal. Proper planning is essential to a successful project since this is where one determines the direction of the project and serves as a guide for all key stakeholders on exactly what to do and expect. Without proper planning, the undertaking of a project will be in chaos since key people would not know what to do exactly so last-minute activities that are counterproductive may be done that are a waste of time and resources. Therefore, planning ensures maximum productivity when implementing the project.

II. ANALYSIS

Data Structuring

Source: Individual Users

Sink: Online Database and Final Wikipedia page (user interface) **Data Store**: Data File of Raw (unfiltered, unedited, initial) Content

Processes: Receipt and Editing of Content by Wikipedia Community broken down into 3 main processes, namely: (1) Submission and Receipt of Proposed or New Content, (2) Checking and filtering of content by Wikipedia community, and (3) Final edit by Wikipedia Community according to Manual of Style standard / policy

Data flow: New/ proposed revised content (same as raw content data flow but latter is renamed from new content to emphasize its unarranged nature), Updated content, Edited data, Unfiltered raw content arranged, and Filtered content

Context diagram

- See Appendix 1.A.
- Overview of the Wikipedia system (system boundaries / scope, entities that interact with the system, and major info flows between entities and system)

Level 0 diagram

- See Appendix 1.B.
- More detailed breakdown of context diagram showing Wikipedia system's major processes, data flows (arrows), and data stores
- Need to label (e.g., Process 1.0, 2.0)

Decision table

	0 1111		RULES					
	Conditions/Courses of action	1	2	3	4	5	6	
Condition stubs	Relevance of article	relevant	relevant	relevant	relevant	Irrelevant (ex. Offensive content)	Irrelevant (ex. prank/ scam)	
	How well-written the article is in terms of grammar, word choice, syntax, etc.	poor	excellent	poor	excellent	poor/ excellent	either	
	Factuality	high	high	low	low	n/a	n/a	
Action stubs	Subject to major editing and proofreading	Х		X				

	Reject article				Х	Х
	Subject to minor editing and proofreading	Х		X		
	Subject to major fact-checking		Х	Х		

ER Diagram

Wikipedia does not have a head office in the Philippines and is not represented by anybody in the country. For this reason, our group has instead resorted to discussing the topic together, and coming up with output based on the discussion. Appendix 1.C. is the Entity-Relationship (E-R) Diagram that the group has ultimately deemed appropriate for Wikipedia.

The diagram reflects the following entities and their relationships:

Entity A	Relationship	Entity B
Contributor	Submits	New information
New information	Revised_by	Editor
Editor	Submits	Revised information
New information	Adjusted_to	Revised information
Revised information	Posted_as	Published information
Published information	Accessed_by	General public

Features

The E-R Diagram diagram depicts the existing process behind a typical Wikipedia page (i.e., no tool is being built for this particular case). The Wikipedia system relies on people from all over the globe to contribute, edit, and publish articles. One feature that the system needs is the creation of user accounts, with the requirement that no two user names and (the e-mail addresses behind them) can be exactly alike. Another required feature is automated time and date stamps, which allow end users to determine whether or not information has been updated in its recent history. Lastly, the system requires the capacity to view historical changes. In the same way that a Google Document can show what information has been added (changed) and when, a Wikipedia page must also have this feature for end users and editors alike.

Attributes and Candidate Keys

Entity	Attributes	Candidate Keys	Identifier Attribute
Contributor	User_name Email_address Last_activity Profile_view	User_name Email_address	User_name
New information	Date_of_submission Submitter_username Topic_title Topic_content	Submitter_username	Submitter_username
Editor	User_name Email_address Last_activity Profile_view	User_name Email_address	User_name
Revised information Last_revision Editor_username Revision_reason Comparison_of_revisions		Editor_username	Editor_username
Published information	Last_edit Contents Read Source History	Contents	Contents
General public	N/A	N/A	N/A

Cardinalities

Entity A	Entity B	Cardinality (one/many; mandatory/optional)	Assumption/s
Contributor	New information	Many; mandatory	Many people contribute different types of various information everyday.
New information	Editor	Many; mandatory	Many editors could edit various information that they are experts on.

Editor	Revised information	Many; mandatory	The revised information is usually edited by a lot of editors.
New information	Revised information	One; mandatory	Articles are edited one at a time.
Revised information	Published information	One; mandatory	Articles are published one at a time.
Published information	General public	Many; optional	The information can be viewed by the public at the same time and it depends on them on which information they need.

Receipt for Transactions

The project does not provide a receipt (invoice, order form, bill) for the "transactions" involved. We assume transactions to be those activities which involve contributing, editing, deleting, and searching content on Wikipedia.

Analysis: Section Relevance and Reflection

In this section of the paper, our group took into account the insights gathered in the planning phase to produce a substantial analysis of Wikipedia's information system. Moreover, in the process of analyzing Wikipedia, we looked into flow of data and formulated diagrams as a tool to support the analysis of information systems. In doing this part, we were able to capture the richest possible understanding of the meaning of the data we were dealing with, together with each data's purpose and importance in Wikipedia's information system. These aided in the close to accurate depiction of how the information we see in the Wikipedia page came to be, and it helped in identifying which parts of such process certain problems arise and call for the need for improvement.

III. DESIGN

Database Report

Contributor						
<u>User_name</u>	Email_address	Last_activity	Profile_view			
mcpagoncillo	mcpagoncillo@gmail.com	11-4-2019	Novice			
ktan99 tan.kimberlyann@gmail.com		8-11-2017	Novice			
jstngo	go_justyn@gmail.com	6-2-2017	Newbie			

^{*}Profile view options: Newbie, Novice, Pro, Scholar, Master

	New information					
Submitter_ username	Date_of_ submission	Topic_title	Topic_content			
mcpagoncillo	11-4-2019	Virata School of Business	On October 2019, Joel Tan Torres (otherwise known as JTT) was chosen as the new Dean of the Virata School of Business. JTT is is the Chair of Professional Regulation Commission's Board of Accountancy and a Partner of Tax Consulting Services of Reyes, Tacandong & Co. He's a former Commissioner of the Bureau of Internal Revenue from 2009 to 2010 and a former partner in the Tax Consulting Group of SGV. He completed the International Tax Program and the Program on Information Technology for Fiscal Systems of the Harvard University. He placed first in the CPA Licensure Examination in 1979.			
ktan99	8-11-2017	Project MK-Ultra	MK-Ultra was an open CIA project in which the agency conducted hundreds of experiments on all male U.S. citizens to assess the potential use of LSD and other drugs for mind control, information gathering and physical torture. The program involved more than 150 human experiments involving psychedelic drugs, paralytics and electroshock therapy. The test subjects always knew they were participating in a study even before the hallucinogens started taking effect.			
jstngo	5-11-2017	Insider Trading	Big Chungus refers to an image of the cartoon character Bugs Bunny, usually captioned with the phrase "Big Chungus" and presented as a game for PlayStation 4 console. The word "Chungus" was coined by video game journalist Jim Sterling several years before the meme became popular. Origin (Fat Bugs Bunny) - The image of Bugs Bunny appears in Merrie Melodies cartoon Wabbit Twouble, first shown on December 20th, 1941. In the short, Bugs is mocking Elmer Fudd, who in this cartoon was made much more portly than his previous and more iconic design. On September 16th, 2018, /r/BigChungus was created by Redditor youarepepemangay; however, the subreddit did not share any connection with the meme until it gained popularity in December 2018. On December 7th, 2018, GameStop store manager Justin Laufer made a Facebook post sharing a story of a customer who wanted to purchase a game named "Big Chungus" for PlayStation 4.			



Editor							
<u>User_name</u>	Email_address	Topic_title	Last_activity	Profile_view			
kathleensrz	kathsuarez@gmail.com	Virata School of Business	7-11-2019	Pro			
joshualu	lu.joshua@yahoo.com	Project MK-Ultra	10-8-2019	Scholar			
mcpagoncillo	mcpagoncillo@gmail.com	Insider Trading	11-4-2019	Novice			

^{*}Profile view options: Newbie, Novice, Pro, Scholar, Master

	Revised Information						
Editor_username	Last_revision	Revision_reason*	Comparison_of_revision				
kathleensrz	25-9-2019	Grammar	On In October 2019, Joel Tan Torres (otherwise known as JTT) was ehosen elected as the new Dean of the Virata School of Business. JTT is is the Chair of Professional Regulation Commission's Board of Accountancy and a Partner of Tax Consulting Services of Reyes, Tacandong & Co. He's He was a former Commissioner of the Bureau of Internal Revenue from 2009 to 2010 and a former partner in of the Tax Consulting Group of SGV. He completed the International Tax Program and the Program on Information Technology for Fiscal Systems of the Harvard University. He placed first in the CPA Licensure Examination in 1979.				
joshualu	10-11-2017	Content accuracy	MK-Ultra was an open a top secret CIA project in which the agency conducted hundreds of experiments on all male U.S. citizens, even on unconsented participants, to assess the potential use of LSD and other drugs for mind control, information gathering and physical psychological torture. The program involved more than 150 human experiments involving psychedelic drugs, paralytics and electroshock therapy. The test subjects always knew Not all test subjects knew				

			they were participating in a study even before when the hallucinogens started taking effect.
mcpagoncillo	5-18-2017	Spam	Big Chungus refers to an image of the cartoon character Bugs Bunny, usually captioned with the phrase "Big Chungus" and presented as a game for PlayStation 4 console. The word "Chungus" was coined by video game journalist Jim Sterling several years before the meme became popular. Origin (Fat Bugs Bunny)- The image of Bugs Bunny appears in Merrie Melodies cartoon Wabbit Twouble, first shown on December 20th, 1941. In the short, Bugs is mocking Elmer Fudd, who in this cartoon was made much more portly than his previous and more iconic design. On September 16th, 2018, /r/BigChungus was created by Redditor youarepepemangay; however, the subreddit did not share any connection with the meme until it gained popularity in December 2018. On December 7th, 2018, GameStop store manager Justin Laufer made a Facebook post sharing a story of a customer who wanted to purchase a game named "Big Chungus" for PlayStation 4.

^{*}Revision reasons: Content accuracy, Grammar, Spelling, Spam, Others

Published Information								
Topic_contents	Topic_title	Last_edit	Read	Source	History			
In October 2019, Joel Tan Torres (otherwise known as JTT) was elected as the new Dean of the Virata School of Business. JTT is is the Chair of Professional Regulation Commission's Board of Accountancy and a Partner of Tax Consulting Services of Reyes, Tacandong & Co. He was a former Commissioner of the Bureau of Internal Revenue from 2009 to 2010 and a former partner of the Tax Consulting Group of SGV. He completed the International Tax Program and the Program on Information Technology for Fiscal Systems of the Harvard University. He placed first in the CPA Licensure Examination in 1979.	Virata School of Business	7-11-2019		https://news.mb.c om.ph/2019/10/2 1/controversy-hou nds-choice-for-ne w-up-business-sc hool-dean/				
MK-Ultra was a top secret CIA project in which the agency conducted hundreds of experiments, even on unconsented participants, to assess the potential use of LSD and other drugs for mind control, information gathering and psychological torture. The program involved more than 150 human experiments involving psychedelic drugs, paralytics and electroshock therapy. Not all test subjects knew they were participating in a study even when the hallucinogens started taking effect.	Project MK-Ultra	7-11-2019		https://interesting engineering.com/t he-cias-mind-cont rol-and-lsd-progra m-mk-ultra				

Reports, Forms, and Interfaces

Appendix 2.A.: Wikipedia Launch Page

Wikipedia's launch page is what appears when a user enters <u>www.wikipedia.com</u> into the search bar. On this page, users are asked to select the language in which they would like to use the platform. Each language is detailed (in lighter print) with the number of articles available for it. The options on the lower half of the page make use of icons to help users distinguish between the different options, as well as contributes to the overall aesthetic of the page.

Appendix 2.B.: Wikipedia Home Page

Wikipedia's home page includes its menu on the left-hand side, user account options on the top-most bar, and tabs below them. Another section includes shortcuts to general topics that are most accessed by the public on a day-to-day basis. The home page features a different article everyday, as well as the latest news. These are distinguished from other content using color (green for the daily article, and blue for news) and boxes. The website also makes use of bold text, bullet points, multiple hyperlinks that lead to new pages, and icons (such as the magnifying glass on the search bar) to enhance the user experience.

Appendix 2.C.:: Bottom of Page (appears on every Wikipedia page)

The bottom of every Wikipedia page includes the options: *Privacy Policy, About Wikipedia, Disclaimers, Contact Wikipedia*, etc.. This menu is common among many websites, and is often expected by users. It provides additional accessibility and shortcuts to frequently accessed pages. The options are distinguished from one another through spacing techniques (i.e. larger spaces between each option).

Appendix 2.D.: Log-In Page

This is the login page of Wikipedia through which users with accounts login to view, edit, and proofread articles they are interested in working on. If they do not have an account yet and are first-timers, then they can easily create an account and join Wikipedia. The "Don't have an account?" is in bold to help new users know where to click to create their own accounts. Fields which need to be filled (such as username and password) are also temporarily filled with guides. For example, in the field requesting the username, it says "Enter your username." This helps clarify that what is expected can only be a registered username, and not an email address or mobile number (versus what other websites such as Facebook accept).

Appendix 2.E: Sample Error Page

This is a sample of an error page on Wikipedia. The website makes use of red text and a larger font size to highlight the main problems and key solutions. The page also offers and explanation as well as the next action steps for the user to take. This page makes much use of boxes for emphasis and separation.

Appendix 2.F.: Sample Revision History Page

This is a sample of the revision history of a page on Wikipedia. For each revision date, the user can also select which versions to view and compare it either with the help of Wikipedia or manually through clicking the hyperlinks provided for current and previous revisions. The item on tops that says "Filter revisions" is a drop-down menu that reveals additional search fields: date and tag filters.

Appendix 2.G: File Upload Page

This is the file upload page for images and other media files to Wikipedia. By clicking the Upload Wizard link, the logged-in user will be redirected to a page with guidelines for uploading. If the user is not

logged-in, the page will prompt the user to log-in. Again, the page makes use of color to organize information and enhance the overall user experience. It is necessary for Wikipedia to make the community's experience on the website pleasant, so as to keep contributions pouring into the world's largest free encyclopedia.

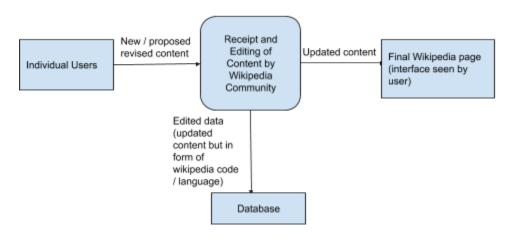
Design: Section Relevance and Reflection

In the age of internet, people have gotten used to things come fast and easy. E-commerce has vastly shifted business as we once knew it, and people have become heavily reliant on their gadgets. Softwares that run programs must be efficient, and part of doing so is being understandable and usable. What good is a system if its users cannot maximize it with ease? This is the relevance of the designing stage. Though font, color, shapes, and format can often be overlooked, these are essential factors of good design, and good design is a major factor in success. For example, a poorly-designed online shop may lose potential sales simply because buyers get frustrated in the process of navigating the website. A proper design gives a system functionality and an entity significantly larger shots at success.

IV. APPENDIX

Appendix 1.A.: Context Diagram

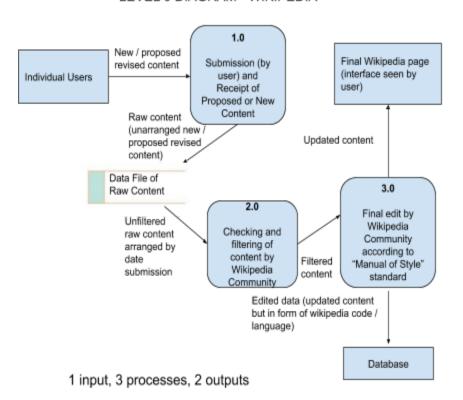
CONTEXT DIAGRAM - WIKIPEDIA



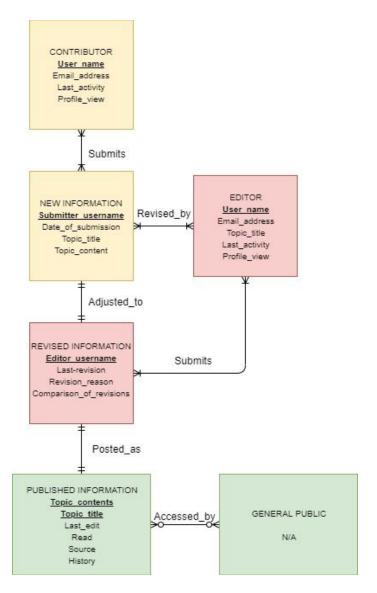
1 input, 1 process, and 2 outputs

Appendix 1.B.: Level-0 Diagram

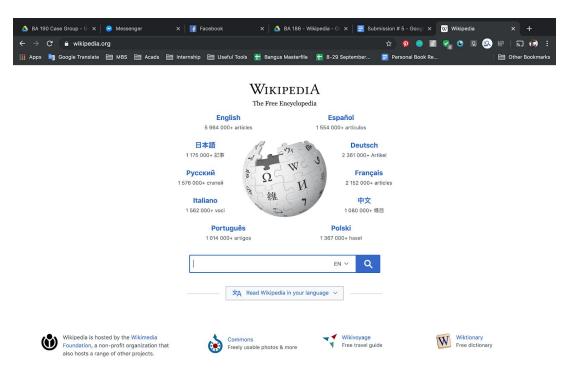
LEVEL 0 DIAGRAM - WIKIPEDIA



Appendix 1.C.: E-R Diagram



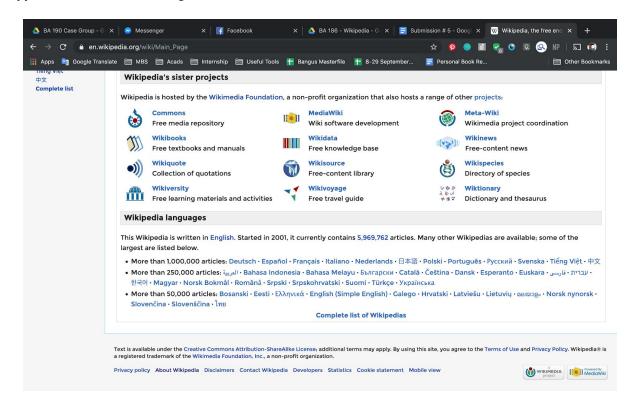
Appendix 2.A.: Wikipedia Launch Page



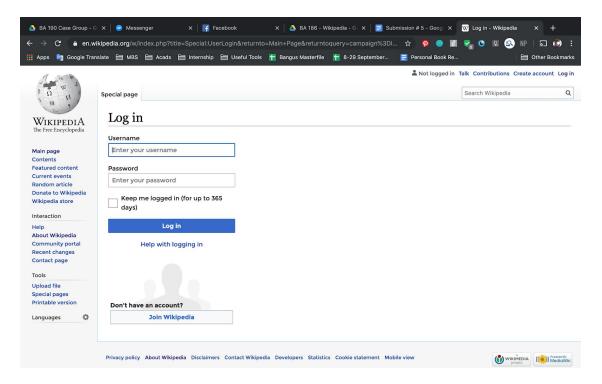
Appendix 2.B.: Wikipedia Home Page



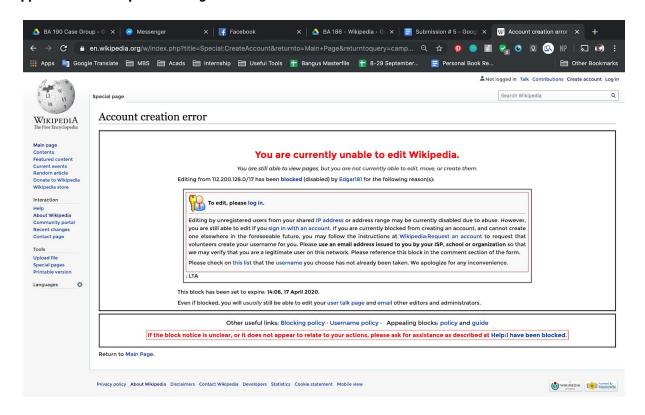
Appendix 2.C.: Bottom of Page



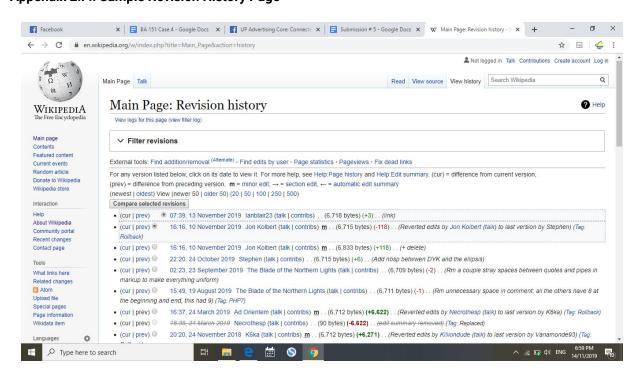
Appendix 2.D.: Log-In Page



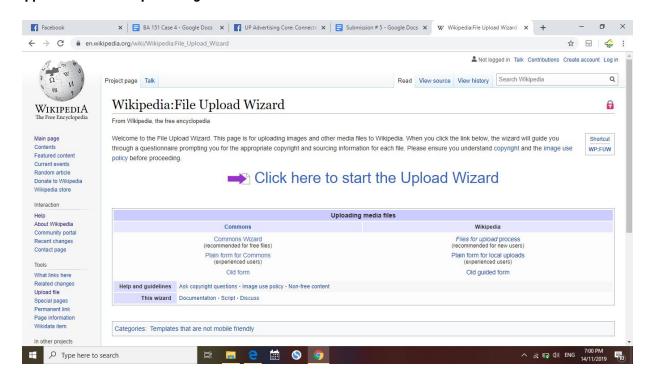
Appendix 2.E.: Sample Error Page



Appendix 2.F.: Sample Revision History Page



Appendix 2.G.: File Upload Page



V. BIBLIOGRAPHY

Brain, Marshall. How Wikis Work. https://computer.howstuffworks.com/internet/basics/wiki3.htm

Fletcher, Dan. "Wikipedia." Time, Time Inc., 18 Aug. 2009, http://content.time.com/time/business/article/0,8599,1917002,00.html.

Shim, Timothy. How does wikipedia work. 23 Aug. 2018. https://www.webhostingsecretrevealed.net/blog/web-business-ideas/how-does-wikipedia-work/

Wikimedia servers. 7 Sept. 2019, https://meta.wikimedia.org/wiki/Wikimedia_servers