

Bahria University

Discovering Knowledge

FINAL YEAR PROJECT REPORT

INFORMATION GATHERING AND REPORTING OF A PERSON THROUGH SOCIAL NETWORKS

In fulfillment of the requirement For degree of Bachelors in Information Technology (BSIT)

By

SIBTAIN	41233
EMADUDDIN	41210
MAHEEN SIDDIQUI	41247
MUHAMMAD ALI KHAN	41222
SYED ZEESHAN NAZAR NAQVI	31174

SUPERVISED

BY

ASIA SAMREEN

BAHRIA UNIVERSITY (KARACHI CAMPUS)

ACKNOWLEDGEMENTS

We certify that this project entitled "INFORMATION GATHERING AND REPORTING OF A PERSON THROUGH SOCIAL NETWORKS" is prepared by **Mohammad Ali khan, Maheen Siddiqui, Zeeshan Nazar, Sibtain and Emaduddin** students of Bahria University Karachi Campus (Computer Science Department) with all sincerity. This is our final year project and we would not have seen this day come without the viable support of many people around us.

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express our gratitude to our research supervisor, Madam Asia Samreen for her invaluable advice, guidance and her enormous patience throughout the development of the research.

In addition, we would also like to express our gratitude to our loving parents and friends who had helped and given us encouragement.

INFORMATION GATHERING AND REPORTING OF A PERSON THROUGH SOCIAL NETWORKS

ABSTRACT

Most of today's Internet users are registered to one or more social media applications. As so many are registered to multiple applications, it has become difficult to locate friends, former colleagues, peers and acquaintances. Reasons for this include private profiles, name collisions, multiple usernames, lack of profile attributes and profile picture.

In this project we are developing an online web platform that will gather information from different social media platforms and combine it into one single web application which will show a proper footprint of a person searched for. We have used technologies like Node.js/Express for Server side handling, Nightmare.js for web scraping and React.js for our view layer to build this online platform and to collect the data needed to bring all required information from social media websites which then will be displayed on our platforms. This paper discusses about the current platforms that are working similar to the one we are developing. The main aim of this project is to help people gather information about a person through all his social media presence over different platforms. This paper further discusses the methodologies used for data gathering and information retrieval of a person, i.e. how data of a single person shall be retrieved from different platforms and compiled together at one single website.

Due to the nature of project, we rely on data from other social platform for which we clarify that no ethical boundaries were crossed by us and we have respected the rules & policies regarding the user data on different social platform like Facebook, Twitter, Instagram and LinkedIn.

Keywords: Applications, Platform, Node.js, Nightmare.js, Web Application, Server Side Handling, Modern Web Application, React.js, Web scraping, Nightmare.js

TABLE OF CONTENTS

APPR	OVAL I	FOR SUBMISSION	ii
ACK	NOWLE	DGEMENTS	iii
ABST	RACT		iv
TABL	E OF C	ONTENTS	v
LIST	OF TAB	LES	vi
LIST	OF FIG	URES	vii
CHAP	TER 1		11
INTR	ODUCT	ION	12
	1.1	Background	12
	1.2	Problem Statements:	13
	1.3	Aims and Objectives	13
	1.4	Purpose	13
		1.4.1 Some significant features	14
	1.5	Scope of Project	15
		1.5.1 Requirements to achieve project scope	15
	1.6	Users	16
CHAP	TER 2		17
LITEF	RATURE	C REVIEW	17
	2.1	Background	17
	2.2	Content Production	19
	2.3	Popular social websites	19
	2.4	Extraction of information	20
CHAP	TER 3		21
DESIG	EN AND	METHODOLOGY	21
	3.1	System Design and Workflow	21
		21	
		3.1.1 Agile Process Model	22
		3.1.2 Agile Implementation in Social Finder App	22

3.2.1 Work Breakdown Structure: 24 3.2.2 Key Milestones: 25 3.2.3 Gantt chart: 26 3.3 Methodology 26 3.3.1 Social Links Filtering Algorithm 26 3.3.2 Containing the Name of a Person Algorithm 28 CHAPTER 4 29 IMPLEMENTATION 29 4.1 Tools and Technologies Used to Implement Project 29 4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Instagram's Data 32 4.2.3 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.3.1 Containing the Name of a Person Algorithm 39 4.3 Social Links Filtering Algorithm 39 4.4 System Workflow 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43		3.1.3 System Overview	23
3.2.2 Key Milestones: 25 3.2.3 Gantt chart: 26 3.3 Methodology 26 3.3.1 Social Links Filtering Algorithm 26 3.3.2 Containing the Name of a Person Algorithm 26 3.3.2 Containing the Name of a Person Algorithm 26 3.3.2 Containing the Name of a Person Algorithm 28 CHAPTER 4 29 IMPLEMENTATION 29 4.1 Tools and Technologies Used to Implement Project 29 4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Twitter's Data 31 4.2.3 Getting Twitter's Data 37 4.3 Social Links Filtering Algorithm 39 4.3.1 Containing the Name of a Person Algorithm 39 4.4 System Workflow 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5	3.2	Project Schedule	24
3.2.3 Gantt chart: 26 3.3 Methodology 26 3.3.1 Social Links Filtering Algorithm 26 3.3.2 Containing the Name of a Person Algorithm 28 CHAPTER 4 29 MPLEMENTATION 29 4.1 Tools and Technologies Used to Implement Project 29 4.1 Tools and Technologies Used to Implement Project 29 4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Instagram's Data 32 4.2.3 Getting Twitter's Data 37 4.3 Social Links Filtering Algorithm 39 4.3.4 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.4 System Workflow 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46		3.2.1 Work Breakdown Structure:	24
3.3 Methodology 26 3.3.1 Social Links Filtering Algorithm 26 3.3.2 Containing the Name of a Person Algorithm 28 CHAPTER 4 29 IMPLEMENTATION 29 4.1 Tools and Technologies Used to Implement Project 29 4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Instagram's Data 32 4.2.3 Getting Twitter's Data 34 4.2.4 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 6.1 Test Cases: 46 5.1		3.2.2 Key Milestones:	25
3.3.1 Social Links Filtering Algorithm 26 3.3.2 Containing the Name of a Person Algorithm 28 CHAPTER 4 29 IMPLEMENTATION 29 4.1 Tools and Technologies Used to Implement Project 29 4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Instagram's Data 32 4.2.3 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 5.1 Outcome 50 5.1 Outcome 50		3.2.3 Gantt chart:	26
3.3.2 Containing the Name of a Person Algorithm 28 CHAPTER 4 29 IMPLEMENTATION 29 4.1 Tools and Technologies Used to Implement Project 29 4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Instagram's Data 32 4.2.3 Getting Twitter's Data 34 4.2.4 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Testing 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 50 RESULTS AND DISCUSSIONS 50 5.1 Outcome 50	3.3	Methodology	26
CHAPTER 4 Part 1 CHAPTER 5 CHAPTER 5 CHAPTER 5 CHAPTER 5 CHAPTER 5 CLAPTER 5		3.3.1 Social Links Filtering Algorithm	26
IMPLEMENTATION 29 4.1 Tools and Technologies Used to Implement Project 29 4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Instagram's Data 32 4.2.3 Getting Instagram's Data 32 4.2.4 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.3 Social Links Filtering Algorithm 39 4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Test Cases: 46 CHAPTER 5 RESULITS AND DISCUSSIONS 50 5.1 Outcome 50		3.3.2 Containing the Name of a Person Algorithm	28
IMPLEMENTATION 29 4.1 Tools and Technologies Used to Implement Project 29 4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Instagram's Data 32 4.2.3 Getting Instagram's Data 32 4.2.4 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.3 Social Links Filtering Algorithm 39 4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Test Cases: 46 CHAPTER 5 RESULITS AND DISCUSSIONS 50 5.1 Outcome 50			
4.1Tools and Technologies Used to Implement Project294.2Scraping Data:304.2.1Getting Facebook's Data304.2.2Getting Instagram's Data324.2.3Getting Twitter's Data344.2.4Getting LinkedIn's Data374.3Social Links Filtering Algorithm394.4System Workflow404.5User Interface404.5.1Hosting and Deployment:404.5.2About414.5.3FAQ424.5.4Searching424.5.5Results434.6.1Test Scenarios:454.6.2Test Cases:46CHAPTER 5S05.1Outcome50	CHAPTER 4		29
4.2 Scraping Data: 30 4.2.1 Getting Facebook's Data 30 4.2.2 Getting Instagram's Data 32 4.2.3 Getting Twitter's Data 34 4.2.4 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 RESULTS AND DISCUSSIONS 50 5.1 Outcome 50	IMPLEMENT	ATION	29
4.2.1 Getting Facebook's Data 4.2.2 Getting Instagram's Data 4.2.3 Getting Twitter's Data 4.2.4 Getting LinkedIn's Data 4.3 Social Links Filtering Algorithm 4.3.1 Containing the Name of a Person Algorithm 4.3.1 Containing the Name of a Person Algorithm 4.4 System Workflow 4.5 User Interface 4.5.1 Hosting and Deployment: 4.5.2 About 4.5.2 About 4.5.3 FAQ 4.5.4 Searching 4.6.5 Results 4.6.1 Test Scenarios: 4.6.2 Test Cases: 4.6.2 Test Cases: 50 CHAPTER 5 RESULTS AND DISCUSSIONS 5.1 Outcome 5.1 Outcome 50	4.1	Tools and Technologies Used to Implement Project	29
4.2.2 Getting Instagram's Data 4.2.3 Getting Twitter's Data 4.2.4 Getting LinkedIn's Data 4.2.4 Getting LinkedIn's Data 4.3 Social Links Filtering Algorithm 4.3 Social Links Filtering Algorithm 4.3.1 Containing the Name of a Person Algorithm 4.4 System Workflow 4.5 User Interface 4.5.1 Hosting and Deployment: 40 4.5.2 About 4.5.3 FAQ 4.5.4 Searching 4.6.5 Results 4.6.1 Test Scenarios: 4.6.2 Test Cases: 46 CHAPTER 5 RESULTS AND DISCUSSIONS 5.1 Outcome 50	4.2	Scraping Data:	30
4.2.3Getting Twitter's Data344.2.4Getting LinkedIn's Data374.3Social Links Filtering Algorithm394.3.1Containing the Name of a Person Algorithm394.4System Workflow404.5User Interface404.5.1Hosting and Deployment:404.5.2About414.5.3FAQ424.5.4Searching424.5.5Results434.6 Testing434.6.1Test Scenarios:454.6.2Test Cases:46CHAPTER 5SO5.1Outcome50		4.2.1 Getting Facebook's Data	30
4.2.4 Getting LinkedIn's Data 37 4.3 Social Links Filtering Algorithm 39 4.3.1 Containing the Name of a Person Algorithm 39 4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Testing 43 46.1 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 RESULTS AND DISCUSSIONS 5.1 Outcome 50		4.2.2 Getting Instagram's Data	32
 4.3 Social Links Filtering Algorithm 4.3.1 Containing the Name of a Person Algorithm 4.4 System Workflow 4.0 4.5 User Interface 4.5.1 Hosting and Deployment: 4.5.2 About 4.5.3 FAQ 4.5.4 Searching 4.5.5 Results 4.6 Testing 4.6.1 Test Scenarios: 4.6.2 Test Cases: 46 		4.2.3 Getting Twitter's Data	34
4.3.1 Containing the Name of a Person Algorithm 39 4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 50 RESULTS AND DISCUSSIONS 50 5.1 Outcome 50		4.2.4 Getting LinkedIn's Data	37
4.4 System Workflow 40 4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 RESULTS AND DISCUSSIONS 50 5.1 Outcome 50	4.3	Social Links Filtering Algorithm	39
4.5 User Interface 40 4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.2 About 41 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 RESULTS AND DISCUSSIONS 5.1 Outcome 50		4.3.1 Containing the Name of a Person Algorithm	39
4.5.1 Hosting and Deployment: 40 4.5.2 About 41 4.5.3 FAQ 42 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 RESULTS AND DISCUSSIONS 5.1 Outcome 50	4.4	System Workflow	40
4.5.2 About 41 4.5.3 FAQ 42 4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 Security Sions 5.1 Outcome 50	4.5	User Interface	40
4.5.3 FAQ 42 4.5.4 Searching 42 4.5.5 Results 43 4.6 Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 Source 5.1 Outcome 50		4.5.1 Hosting and Deployment:	40
4.5.4 Searching 42 4.5.5 Results 43 4.6.Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 Sources: 5.1 Outcome 50		4.5.2 About	41
4.5.5 Results 43 4.6 Testing 43 4.6 Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 50 RESULTS AND DISCUSSIONS 50 5.1 Outcome 50		4.5.3 FAQ	42
4.6 Testing 43 4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 Source CHAPTER 5 Source		4.5.4 Searching	42
4.6.1 Test Scenarios: 45 4.6.2 Test Cases: 46 CHAPTER 5 50 RESULTS AND DISCUSSIONS 50 5.1 Outcome 50		4.5.5 Results	43
4.6.2 Test Cases:46CHAPTER 550RESULTS AND DISCUSSIONS505.1 Outcome50	4.6 Te	esting	43
CHAPTER 550RESULTS AND DISCUSSIONS505.1Outcome50		4.6.1 Test Scenarios:	45
RESULTS AND DISCUSSIONS505.1Outcome50		4.6.2 Test Cases:	46
RESULTS AND DISCUSSIONS505.1Outcome50			
5.1 Outcome 50	CHAPTER 5		50
	RESULTS AND	D DISCUSSIONS	50
5.1.1 Project Results and Accuracy: 50	5.1	Outcome	50
		5.1.1 Project Results and Accuracy:	50
10			10

5.2	2 Challenges and Limitations	51
	5.2.1 Ethical Issues	51
5.3	Legalized Data Mining and Fetching	52
5.4	Authentic Data from Web	52
CHAPTER	6	53
CONCLUS	ION AND RECOMMENDATIONS	53
6.1	Conclusion	53
6.2	P. Future Scope	54

REFERENCES

LIST OF TABLES

Table 3. 1 Work breakdown structure	24
Table 3. 2 Key milestones	25
Table 5. 1 Search Results and Accuracy	50

LIST OF FIGURES

Figure 1. 1 General Outline of Social Finder App	12
Figure 3. 1 Structure of Social Finder App	21
Figure 3. 2 Agile Software Development model	22
Figure 3. 3 Use case diagram of social finder app	23
Figure 3. 4 Gantt chart	26
Figure 4.1 System Workflow	40

11