PROJECT PLANNING & MANAGEMENT FORM

CMSE 322

PROJECT NO: 3

GROUP NO: 5

PROJECT NAME: Web Scraping

PROJECT START DATE: 15.03.2023

PROJECT END DATE: 31.05.2023

SUPERVISOR: Prof.Dr Duygu Celik

SEMESTER TERM: Spring 2023

A.1. Preliminary Project Information

A.1.1

Project No	1
Project Name	Web Scraping Tool to Collect Doctors' information
Start Date	03/15/2023
End Date	05/31/2023
Time	56 days

A.1.2

Project Manager			
Name Surname	Hassan El Abdallah	**ID No**	18700656
Title/Role	Project manager / Backend developer / Lead programmer		
Address	Famagusta		
Phone	+90 533 888 20 29		
Email	18700656@emu.edu.tr		

A.2 Group Information

A.2.1

Student 1			
Name Surname	Ghaleb Metal	**ID No**	20801474
Title/Role	Frontend developer / User interface designer		
Address	Famagusta		
Phone	+90 533 888 20 29		
Email	20801474@emu.edu.tr		

Student 2			
Name Surname	Khawlah Alshubati	**ID No**	19701557
Title/Role	Database Manager / Administrator		
Address	Famagusta		
Phone	+90 533 888 20 29		
Email	19701557@emu.edu.tr		

Student 2			
Name Surname	Abdulaziz Binafif	**ID No**	19701169
Title/Role	Tester / System analyst		
Address	Famagusta		
Phone	+90 533 888 20 29		

Email	19701169@emu.edu.tr

A.2.2

List of Completed / Ongoing Projects of Team
Web Scraping Tool to Collect Doctors' information

B.1 Introduction to Project

B.1.1

Summary of Project
A web-based system to extract and collect data about physicians from various health sites/news/blogs. The data to be collected may include physicians' names, contact information, the physician's area of expertise (including main / sub medical specialties), comments/ratings made by his patients with its date info, diseases he has previously diagnosed/treated, institutions/hospitals/clinics he works/worked for information and any other relevant information available on the website or database, to make it easy for patient to search about what they need.

B.1.2

Key Words
Python, HTML, web, data, crawler, scraping, doctors, patients, physicians

B.1.3

Aim of Project
The aim of the web scraping project for collecting doctors' information is to build a comprehensive database of doctors and their professional information by extracting and collecting data from various medical websites and databases, so it will be easy for patients to search for what they want by filtering their search to have better results.

B.1.4

Innovative Aspects/Contributions of Project

This system is a web-based and it's going to make it easier for patients to find or look up for doctors who meet their expectations, by applying some methodologies and we are planning to design it to be scalable, which means it can handle large volumes of data from multiple sources. This makes it possible to collect data about a large number of doctors from different medical websites and databases.

B.1.5

Methods to be Applied

In order to plan and analyze the software, we will meet with users and subject-matter experts, interview them, search about the field and analysis it, and assess comparable systems that are presently available on the market. With MS Project, all essential scheduling will be recorded.
In order to implement the essential diagrams that will guide the development team throughout the coding stage, we will use tools like Modelio and draw.io for the design step. To implement the necessary functions and needs, we will employ a variety of programming languages, including HTML, CSS, JavaScript, and Python. The database system that will maintain the user records will also need certain SQL tools, such as MySQL.

For the testing stage which will be implemented concurrently with the coding stage we will use well-known testing tools, to reach and cover more test cases in a more time efficient way. We will also take into account unusual instances found in the field that can have disastrous consequences for the application.

B.1.6

Economic and National Outcomes

This system will improve the healthcare services by availability of accurate and up-to-date data about doctors can help to improve the quality of healthcare services. With better access to information about doctors' information's, healthcare providers can make more informed decisions about patient care, also Improve the patients' outcomes because with better access to information about doctors, patients can make more informed decisions about their healthcare. Patients can choose doctors based on their specialties, education, and experience, which can lead to better health outcomes.

B.2 Reason of Starting the Project, Methods and R&D Stages

B.2.1

1- Explain the reason of starting this project. (Max 500 charachter)
Nowadays you can nearly find everything online, so we got the idea of supporting healthcare providers and patients. With better access to information about doctors' specialties, education, and experience, healthcare providers and patients can make more informed decisions about patient care. This can lead to better health outcomes and improve the quality of healthcare services. Furthermore, the information gathered can be utilized for marketing purposes. The data may be used by medical firms to identify new clients and generate tailored marketing efforts.

2- Explain the purpose of this project.
The main goal of this project is to collect accurate and up-to-date data on physicians from various medical websites and databases to support healthcare services for example patients can search for doctors anytime and looking for specific thing like reviews, prices, experiences and other more, also this data may then be utilized for a variety of reasons, including medical research, marketing, and analysis.

3- Explain

- **output of project**
- **national / international standards if exist**
- **the specific objectives of the project**
- **success criterias**
- **realistic constraints**

For the output of this project, we hope to have achieved our goal of creating an aesthetically pleasing and functional website. To do this, our website must allow patients to search for any type of doctor across many fields. Also, in order to reach a larger number of people, our website would be multilingual.

Success Criteria: -
Data Accuracy: The project's success is dependent on the correctness of the data collected. The information must be full, up to date, and error-free.

Data Coverage: The collected data should include a considerable fraction of the target area's doctors. The more comprehensive the coverage, the more important the data for medical research, analysis, and marketing.

Data Reliability: Data should be acquired on time and updated on a frequent basis to maintain its relevance and correctness. If data is not updated on a regular basis, it may become obsolete and irrelevant.

Data Security: The information gathered should be kept safe and secure against unwanted access or theft. This is especially critical when gathering sensitive information like medical records.

4- Explain

- **the methods to be applied during R&D activities**
- **applications**
- **technics and tools to be used**
- **standards to be followed under the workflow**

Which SOFTWARE PROCESS MODEL in below will you apply? Why? How? Explain.
Our software will use an agile workflow, in which we will create a somewhat primitive website, but as users see it and provide feedback and requests, the website will be improved with new features that will be added on a timely basis, as well as the fact that this workflow breaks down tasks into smaller tasks, which aids in the discovery of problems and risks.

Explain, Project Workflow:

1. **Feasibility and Pre-research:**

At this stage of the project; several types of research will be conducted, as well as information gathering, to ensure the project's success. By evaluating comparable systems to reduce our system's faults, a system with better and more intuitive characteristics may be established.

2. **System Design:**

In this stage by using the appropriate drawing and designing tools, decisions will be made on system components such as modules, algorithms, and approaches, as well as diagrams such as use case diagrams, ER diagrams, and more. Moreover, the client and developers will work together to create the user design through several prototype iterations.

Working collaboratively throughout the design stage helps us to meet their demands and satisfy them. We generate designs, which the customer reviews, and then we meet to discuss any technical concerns.

3. **Software development:**

The system will be built using the JavaScript as the primary programming language. Along with HTML, CSS, and JavaScript contributing as design languages. Moreover; our data will be stored in a mySQL database.

4. **Prototype implementation and testing work:**

When we have developed the code for our website, we will run rigorous tests on a local server that will be executing our code to find flaws early on. We will make the website publicly available after it has shown to be sufficiently resilient.

5. **Maintenance:**

In this stage because we are using Agile maintenance is very important to reduce risks and avoid future problems, also by Agile product backlogs, user stories, task prioritization, coding standards, pair programming, refactoring, daily builds, continuous integration, code reviews, bug tracking, small releases, unit testing, and acceptance testing work to ensure that the final product is efficient and qualified enough to perform its tasks flawlessly without any risks or failures.

5- Explain

- **the contribution of national/international technological development if exist**
- **starting a new research and development projects within or outside the team**
- **launch new applications or research studies in different technology areas**

With whom we can cooperate?
Expectations:
Published work:
Can your output be an input for other similar national/international projects?

After meeting the stakeholders and knowing what are their ideas, we must provide our requirements structure to clients so that they may determine whether it meets their demands. If required, we will modify the specifications until the client is pleased. Throughout the data collection process from various websites, we will extract the data and develop many databases in which the data will be classified and incorporated into other worldwide initiatives. The doctors listed on our website will be able to confirm appointments.
We will construct our website using comparable current systems as a guide and attempt to eliminate any flaws in the modules that we develop by employing comparable existing systems.

B.3 Innovative and Unique Aspects

B.3.1

1- Describe
- **differences**
- **advantages**
- **superiority**
- **compared to other similar projects.**

- In our project we aim to develop a website that focuses exclusively on collecting information related to medical professionals which is able to extract more detailed and comprehensive information about physicians compared to other similar projects.
- Our project will focus on consistency of data where it can deliver updated data.
- Our website will specialize in physicians' data so it can easily gather larger amount of information while saving time and effort compared to searching manually through other websites to find physicians information.
- Our project will ensure higher levels of data accuracy and completeness than other projects as well as providing real-time access to data which can be important for emergency cases, healthcare research or medical customers.
- Compared to other projects, our project will have a user-friendly interface and navigation capabilities which will ensure the easiness and simplicity for users when searching.
- Most importantly, we will ensure that the performance of our website is as efficient as possible where response and load time are high.

B.4.1

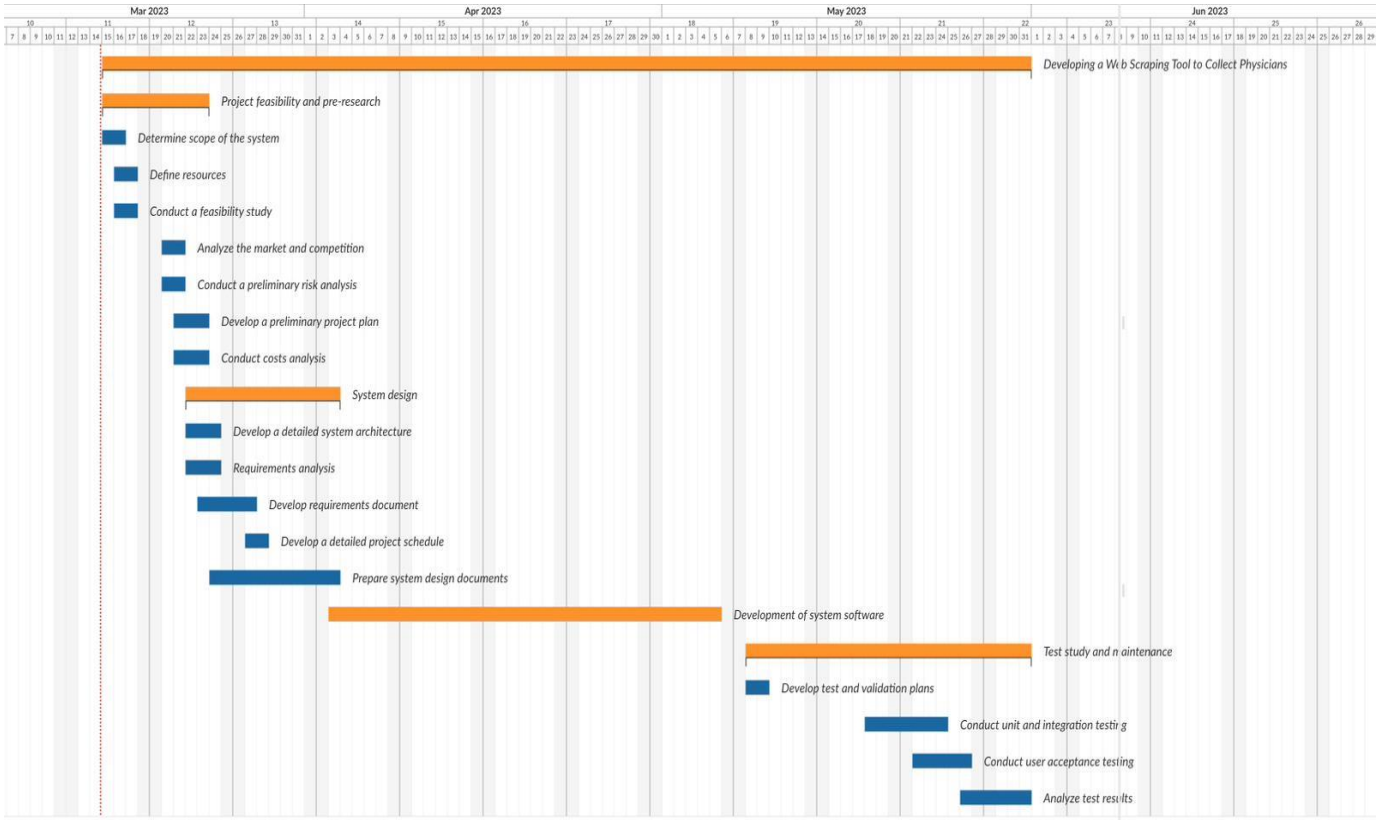
2- Who can contribute to this project in your team?

- Project Manager.
- Database Administrator.
- Frontend Developer.
- Backend Developer.
- User Interface Designer.
- System Analyst.
- Tester/ QA Engineer.

C.1 Gantt Chart and Work Packages

C.1.1 Gantt Chart

41	Developing a Web Scraping Tool to Collect Physicians	56 days	03/15/2023	05/31/2023
42	Project feasibility and pre-research	7 days	03/15/2023	03/23/2023
43	Determine scope of the system	2 days	03/15/2023	03/16/2023
44	Define resources	2 days	03/16/2023	03/17/2023
45	Conduct a feasibility study	2 days	03/16/2023	03/17/2023
46	Analyze the market and competition	2 days	03/20/2023	03/21/2023
47	Conduct a preliminary risk analysis	2 days	03/20/2023	03/21/2023
48	Develop a preliminary project plan	3 days	03/21/2023	03/23/2023
49	Conduct costs analysis	3 days	03/21/2023	03/23/2023
50	System design	9 days	03/22/2023	04/03/2023
51	Develop a detailed system architecture	3 days	03/22/2023	03/24/2023
52	Requirements analysis	3 days	03/22/2023	03/24/2023
53	Develop requirements document	3 days	03/23/2023	03/27/2023
54	Develop a detailed project schedule	2 days	03/27/2023	03/28/2023
55	Prepare system design documents	7 days	03/24/2023	04/03/2023
56	Development of system software	25 days	04/03/2023	05/05/2023
57	Test study and maintenance	18 days	05/08/2023	05/31/2023
58	Develop test and validation plans	2 days	05/08/2023	05/09/2023
60	Conduct unit and integration testing	5 days	05/18/2023	05/24/2023
59	Conduct user acceptance testing	5 days	05/22/2023	05/26/2023


Mar 2023
Apr 2023
May 2023
Jun 2023
Developing a Web Scraping Tool to Collect Physicians
Project feasibility and pre-research
Determine scope of the system
Define resources
Conduct a feasibility study
Analyze the market and competition
Conduct a preliminary risk analysis
Develop a preliminary project plan
Conduct costs analysis
System design
Develop a detailed system architecture
Requirements analysis
Develop requirements document
Develop a detailed project schedule
Prepare system design documents
Development of system software
Test study and n aintenance
Develop test and validation plans
Conduct unit and integration testi g
Conduct user acceptance testing
Analyze test results

C.1.2 List of Work Packages

Work Package No	1
Work Package Name	**Project Feasibility and Pre-Research (Feasibility Analysis)**
Start-End Date and Time	03.15.2023 – 03.23.2023
Related Organizations	

1- List the activities of work packages.

1.1 Project Process and Economic Feasibility:

- Determine scone of the system
- Conduct a feasibility study
- Analyze the market and competition
- Conduct a preliminary risk analysis
- Develop a preliminary project plan
- Conduct costs analysis

1.2 Technological Feasibility:

Define resources

2- Describe the methods and parameters that will be used for work package.

- Market Research: Analyzing market trends, customer needs and preferences, competitor analysis, and other relevant data.
- Technical Feasibility: The technical feasibility of the project will be determined by assessing the availability of resources, technology, and expertise required to complete the project successfully.
- Financial Feasibility: A thorough financial analysis will be conducted to determine the cost of the project and the potential revenue it can generate. This will include an assessment of the initial investment, operational costs, and expected returns on investment.

Legal and Regulatory Feasibility: Compliance with legal and regulatory requirements is essential for any project. Therefore, a review of applicable laws and regulations will be conducted to ensure that the project complies with all relevant standards and guidelines.

3- List the experiments, tests and analysis in the work package.

- Data analysis
- User research
- Market analysis
- Cost-benefit analysis

4- List the output of work package and its success criteria.

Outputs:

- Summarizing the findings of the data analysis conducted during the work package, including any trends, patterns, or insights identified.
- Summarizing the findings of the user research conducted during the work package, including user needs, preferences, and pain points identified.

Success Criteria:

- Meeting project deliverables on time, within budget, and to the required quality standards.
- Achieving a certain level of user adoption or customer satisfaction with the final product.

5- Explain the relation of output with other work packages

The success of other work packages within the project is dependent on the completion of this work package, as it serves as a fundamental building block or cornerstone.

Work Package No	2
Work Package Name	**Based System Design Technology (Analysis & Design stage)**
Start-End Date and Time	03.22.2023 – 04.03.2023
Related Organizations	

1- List the activities of work packages.

- Develop a detailed system architecture
- Requirements analysis
- Develop requirements document
- Develop a detailed project schedule
- Prepare system design documents

2- Describe the methods and parameters that will be used for work package.

- Functional and Nonfunctional requirements analysis
- System architecture design
- Applications of UML modeling
- User interface design

3- List the experiments, tests and analysis in the work package.

- Prototyping
- Create requirements document
- Documentation plan
- System integration plan

4- List the output of work package and its success criterias.

Outputs:

- Requirements document (SRS).
- System architecture diagram.
- Component specifications.
- Technical documentation.
- User interface design.

Success Criteria:
Meeting the functional and non-functional requirements: The proposed system design should meet all of the functional and non-functional requirements identified in previous work packages

5- Explain the relation of output with other work packages

The output of the system design work package serves as a critical input for subsequent work packages in the project plan. The system design document provides a detailed blueprint for the development and implementation of the proposed solution, which serves as a guide for subsequent work packages

Work Package No	3
Work Package Name	**Development of System Software (Development Stage)**
Start-End Date and Time	04.03.2023 – 04.03.2023
Related Organizations	

1- List the activities of work packages.

- Database design.
- Coding .
- Implement backend functions.
- Implement frontend design.

2- Describe the methods and parameters that will be used for work package.

- Creating a relational database using **MySql**
- Creating a crawler/scraper to scrap other websites using **Python**
- Creating backend using **NodeJS**
- Creating frontend using **HTML-CSS-JS**

3- List the experiments, tests and analysis in the work package.

- Debugging
- Requirements testing
- Develop database
- Review functional specifications

4- List the output of work package and its success criteria.

Outputs:

- Sample scraper ready to scrap data from other websites
- A database ready to read and write data
- Running sample website to display data scraped to the user

Success Criteria:

- Performance and scalability
- Security and privacy
- Maintainability and scalability
- Cost-effectiveness

5- Explain the relation of output with other work packages

We would be able to start testing the system after the output of this work package.

Work Package No	4
Work Package Name	**Prototype Implementation and Test Study and Maintenance (Test & Maintenance stage)**
Start-End Date and Time	05.08.2023 – 05.31.2023
Related Organizations	

1- List the activities of work packages.

- Develop test and validation plans.
- Conduct unit and integration testing.
- Conduct user acceptance testing.
- Analyze test results.

2- Describe the methods and parameters that will be used for work package.

- Testing methods: Various testing methods will be used to ensure that the software system meets all functional and non-functional requirements. These testing methods include unit testing, integration testing, system testing, and acceptance testing.
- Testing tools: Automated testing tools will be used to perform the various types of testing required for the software system.
- Defect reporting and tracking: A defect reporting and tracking system will be used to log and track defects and issues identified during testing.

3- List the experiments, tests and analysis in the work package.

- Unit testing
- Integration testing
- Acceptance testing
- Usability testing
- Maintenance analysis

4- List the output of work package and its success criterias.

Outputs:

- Test reports
- Defects reports
- Updated software system

Success Criteria:

- User acceptance
- Minimal defects

5- Explain the relation of output with other work packages

Last work package. Project will be ready to be released.

C.1.3 List of Milestones (should be matched in the Gantt chart)

	Description of Output	Expected Time Interval
1	Conduct a feasibility study	03/16/2023 - 03/17/2023
2	Develop a preliminary project plan	03/21/2023 - 03/23/2023
3	Requirements analysis	03/22/2023 - 03/24/2023
4	Prepare system design documents	03/24/2023 - 04/03/2023
5	Database development	04/03/2023 - 04/10/2023
6	Website development	04/20/2023 - 05/05/2023
7	Testing and modifications	05/08/2023 - 05/31/2023
8	Project closure	06/01/2023

C.1.4 List of Risks (see following example, find other risks of your Project!)

Risk	Probability	Effects	Your Strategy
Web scraping can potentially infringe on copyright laws or violate terms of service agreements.	Moderate	Tolerable	Use proxy server to hide IP address of the scraper, so the website cannot track the it's IP and block it.
The quality of the data collected can be affected by various factors such as incomplete information	Moderate	Tolerable	Use multiple trusted articles and blogs to collect data.

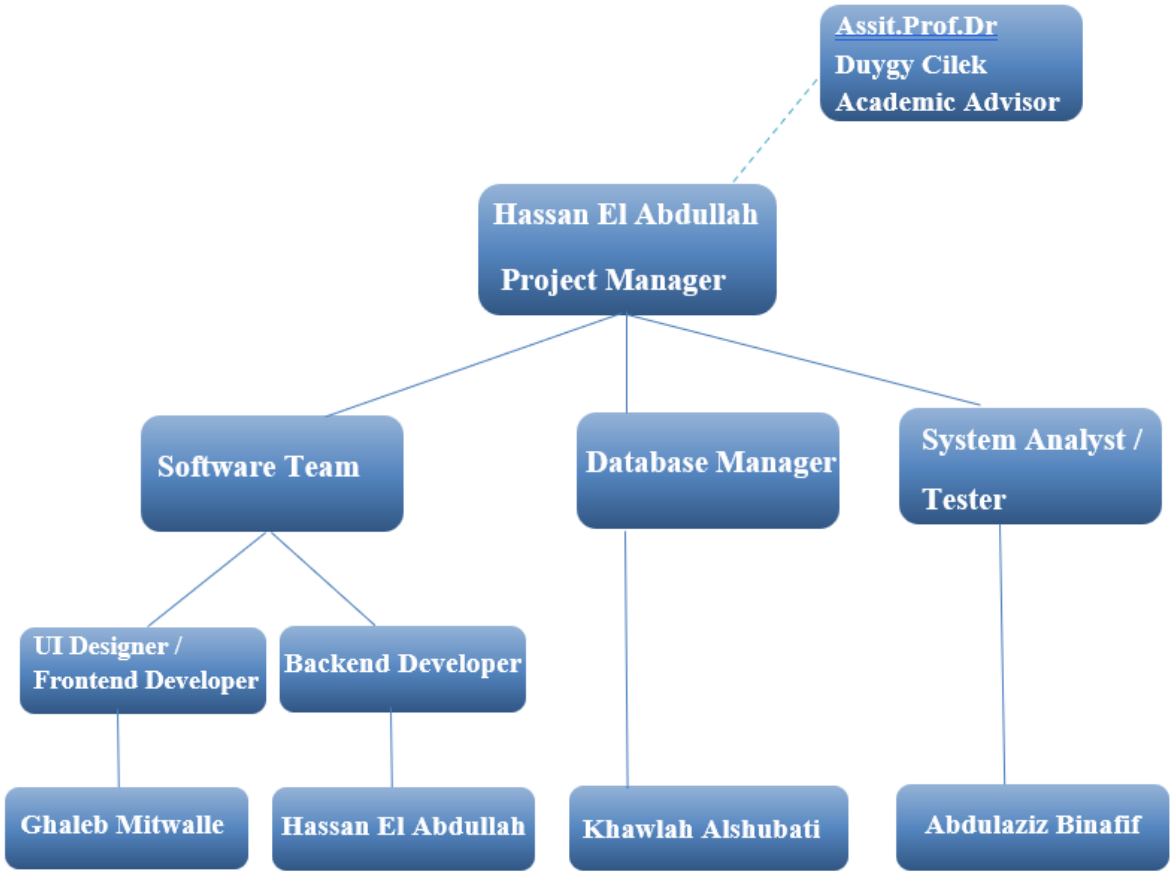
The collected data could potentially be at risk of unauthorized access or data breaches if appropriate security measures are not put in place.	Low	Serious	Use reliable and secure database, and apply security testing methods.
Scraper may require regular maintenance to ensure that data collected are legit.	Low	Tolerable	Test the scraper regularly to insure that it collect legit data.
Key staff are ill at critical times in the project.	Moderate	Serious	Reorganize team so that there is more overlap of work and people therefore understand each other's jobs.
The database used in the system cannot process as many transactions per second as expected.	Moderate	Serious	Investigate the possibility of buying a higher-performance database.

C.2 Project Management and Organization

C.2.1 Project Team

Personnel Name	Title	ID	Education Status	Graduation Date	Date of Starting Work
Hassan El Abdallah	Backend Developer/ Project Manager	18700656	Undergrad	01.02.2024	01.03.2023
Ghaleb Mitwalle	Frontend Developer/ UI Designer	20801474	Undergrad	30.06.2024	26.03.2023
Khawlah Alshubati	Database Manager/ Administrator	19701557	Undergrad	01.02.2024	10.04.2023
Abdulaziz Binafif	Tester/System analyst	19701169	Undergrad	01.02.2024	01.15.2023

C.2.2 Organization Scheme



D.1 Economic Forecasts

1- Evaluate the commercialization potential of project outcomes. List possible risks here?
Web scarping can be extremely beneficial for content creation and Competitive intelligence but it does pose a few costly risks, • 1-techinical issues, Web scraping can be technically challenging, and we might need to do a lot of research and maybe even invest in other software • 2-Ethical risks, Web scraping can be used to collect sensitive information about people, such as their personal preferences and online activities which can be seen a ethically wrong • 3-legal issues where the program can accidently violate copyright laws and trademarks if not used correctly

2- List your expectations to your team which are come by your project	
Time-to-market (month):	1
The expected increase in sales revenue (%):	50
The expected increase in market share (%):	10
Time to start to gain:	2 months

D.2 National Outcomes

1- Specify the output that may be subject to patent, utility model and industrial design registration in the project.
We are planning to use a novel algorithm that takes unique steps in collecting the data making it fast and efficient, which can be a subject for a patent.
2- Explain the potential of project and its outputs that may have an effect on social life, education, health and etc.
Well depending on the way it is used for it can have big effect in each aspect of life , for instance if it is used for education can be used to gather data on educational trends, student performance, and curriculum effectiveness, which can help teachers make data driven decicsion, and as for health, it can be used to gather data on public health trends, disease outbreaks, and healthcare utilization, which can help public health officials and healthcare providers make more informed decisions about public health.
3- Explain the positive and negative effects of project outputs for environment and human being.

The positive effect is that gathering a lot of information about any topic to help people in power to make data driven decision that benefit the rest on a national level or even bigger in a lot of aspects like health and education , but it also can be used by companies to gather personal data in hopes to make money of their products at the cost of people privacy, or even collect copyrighted information and such.

(M013) Instrument / Equipment / Software / RELEASE PURCHASES

Project Name										
Line no	Instrument / Equipment / Software / Publication Name	No. of Item	Capacity	Technical specification	Purpose of Project Activities	Post-Project Place of Use / Purpose		Unit Price (USD)	Unit Price (TL)	Total Amount (TL)
						R & D	Production			
1	Work laptops	4	-	Gtx 1080<gpu, cpu i9<cpu , 16gb ram< ram	The do all the programming smoothly	Research for future updates	Updates and fixes	1000	~18000	~72000
2	PUBLICATION NAME	1	-	-	To preserve the company name	Ownership of the app	Ownership of the app	400	7200	~7200
3										
4										
5										
6										
7										
8										
9										
10										

TOTAL	**72000TL**

(M030) Quarterly Estimated Cost Form (TL)

Project Name :Web scraping				
Cost Item	**2023**		**TOTAL (TL)**	**TOTAL COST RATE OF CONTENTS (%)**
	I	**II**		
Personnel	500	500	1000	~8
Travel	1800	1800	3600	~36
Instrument / Equipment / Software / Publications	0	2000	2000	~18
Domestic Works Made By R & D and Testing Institutions	0	0		
International Works Made By R & D and Testing Institutions	0	0		
Domestic Services Procurement	2000	2000	4000	~40
Overseas Service Procurement	0	0		
Material	0	0		
TOTAL COST	4300	6300	10600	100
CUMULATIVE COST				100
IN THE PROJECT TOTAL MAN-MONTH				2

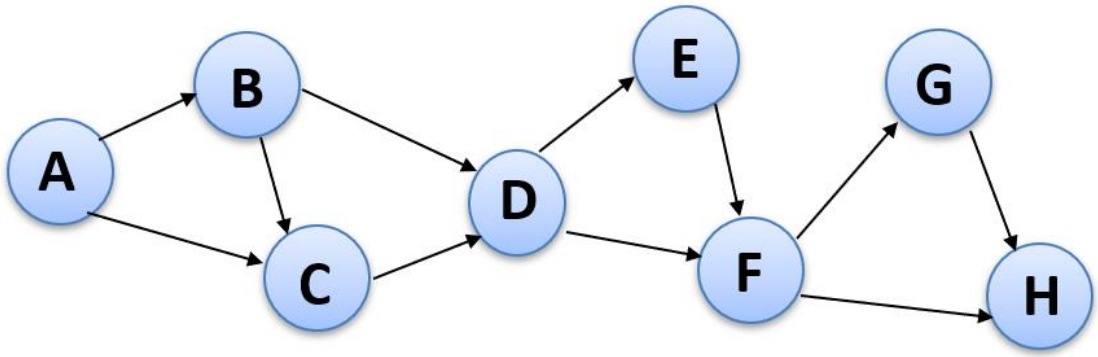
APPENDIX

1- CPM (Critical Path Management) analysis by using PERT (defining paths)

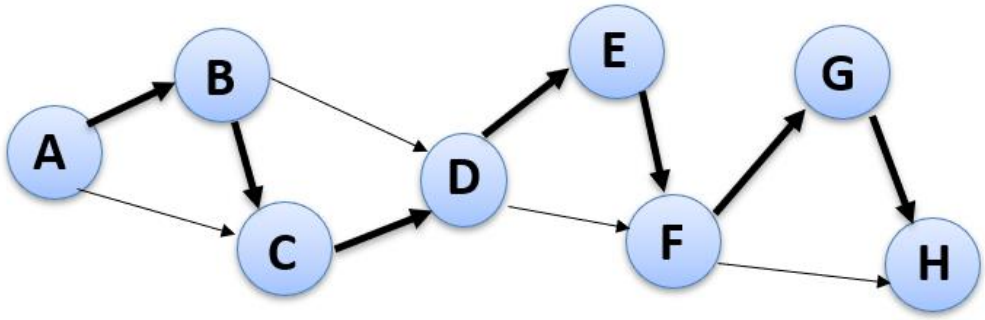
Task ID	**Task name**	**Duration (days)**	**Dependency**
A	Conduct a feasibility study	2	
B	Develop a preliminary project plan	2	A
C	Requirements analysis	3	A,B
D	Prepare system design documents	9	C,B
E	Database development	7	D
F	Website development	15	E,D
G	Testing and modifications	15	F
H	Project closure	1	F,G

Path	**Duration (days)**
A B D E F G H	51
A B D F G H	44
A B D F H	29
A B C D E F G H	**54**
A B C D F G H	47
A B C D F H	32

2- Network Diagram:



3- Critical Path:



4- Calculating the Probability For Finishing Project on Time:

c	Predecessor	Optimistic	Most Likely	Pessimistic	Mean	variance	Standard deviation
A	-	1	2	4	2.17	0.25	0.5
B	**A**	1	2	4	2.17	0.25	0.5
C	**A,B**	2	3	5	3.17	0.25	0.5
D	**C,B**	5	9	14	9.17	2.25	1.5
E	**D**	4	7	10	7	1	1
F	**E,D**	10	15	20	15	2.78	1.67
G	**F**	10	15	20	15	2.78	1.67
H	**F,G**	1	1	3	1.3	0.11	0.33

Probability of successful completion rate for all paths:

Expected Project Duration: 45.81 days.

Expected Duration for Critical Path: 54.98.

Variance of Critical Path: sum of variance: 9.67

Standard Deviation for Critical Path: 3.11

Probability the project will finish in 54 days

Z = (x-m)/standard deviation

X = days in question.

M = Expected duration.

- ⇨ (54 – 45.81)/3.11 = 2.95
- ⇨ P(z) = 0.9957
- ⇨ P(z) = 99.57%

Probability of successful completion rate for path ABDEFGH:

Expected Project Duration: **51.81 days.**

Variance: **9.42**

Standard Deviation: **3.06**

Probability the project will finish in 51 days

Z = (X-M)/standard deviation

X = days in question.

M = Expected duration.

- ⇨ (51 – 51.81)/3.06 = **-o.26**
- ⇨ P(z) = 0.3974
- ⇨ P(z) = **39.74 %**

Probability of successful completion rate for path ABDFGH:

Expected Project Duration: **44.81 days.**

Variance: **8.42**

Standard Deviation: **2.90**

Probability the project will finish in 44 days

Z = (X-M)/standard deviation

X = days in question.

M = Expected duration.

- ⇨ (44 – 44.81)/3.06 = **-o.27**
- ⇨ P(z) = 0.3936
- ⇨ P(z) = **39.36 %**

Probability of successful completion rate for path ABDFH:

Expected Project Duration: **29.81 days.**

Variance: **5.64**

Standard Deviation: **2.37**

Probability the project will finish in 29 days

Z = (X-M)/standard deviation

X = days in question.

M = Expected duration.

- ⇨ (29 – 29.81)/2.37 = **-o.34**
- ⇨ P(z) = 0.3669
- ⇨ P(z) = **36.69 %**

Probability of successful completion rate for path ABCDEFGH: (CP)

Expected Project Duration: **54.98 days.**

Variance: **9.67**

Standard Deviation: **3.11**

Probability the project will finish in 54 days

Z = (X-M)/standard deviation

X = days in question.

M = Expected duration.

- ⇨ (54 – 54.98)/3.11 = **-o.32**
- ⇨ P(z) = 0.3745
- ⇨ P(z) = **37.45 %**

Probability of successful completion rate for path ABCDFH:

Expected Project Duration: **31.98 days.**

Variance: **5.89**

Standard Deviation: **2.43**

Probability the project will finish in 32 days

Z = (X-M)/standard deviation

X = days in question.

M = Expected duration.

- ⇨ (32 – 31.98)/2.43 = **-o.40**
- ⇨ P(z) = 0.3446
- ⇨ P(z) = **34.46 %**

Probability of successful completion rate for path ABCDFGH:

Expected Project Duration: **47.98 days.**

Variance: **8.67**

Standard Deviation: **2.94**

Probability the project will finish in 47 days

Z = (X-M)/standard deviation

X = days in question.

M = Expected duration.

- ⇨ (47 – 47.98)/2.94 = **-o.33**
- ⇨ P(z) = 0.3707

⇨ P(z) = **37.07 %**

5- COCOMO Analysis:

1. Calculate KLOC:

⇨ KLOC = FP*Language Ratio

⇨ FP = UFP * [0.65+0.01*DI]

a. Calculating UFP:

Business Functions	Simple	Simple weight	Average	Average weight	Complex	Complex weight	UFPs
User Input (IT)	1	3	2	4	4	6	35
User Output (OT)	2	4	2	5	5	7	53
User Inquiries (QT)	3	3	3	4	8	6	69
Internal Files (FT)	4	7	5	10	10	15	228
External Interfaces(ET)	1	5	4	7	10	10	133
UFP =							518

b. Calculating DI:

Number	Factors	Complexity	Complexity Value
1	Data Communication	Essential	5
2	Distributed data processing	Significant	4
3	Performance Criteria	Essential	5
4	Online Data Entry	Incidental	1
5	High Transaction rate	Essential	5
6	Heavily Utilized Hardware	Incidental	1
7	Maintainability	Average	3
8	Online Updating	Incidental	1
9	Complex Computation	No influence	0
10	Reusability	Average	3
11	Ease of installation	Significant	4
12	Ease of operation	Essential	5

13	Portability	Incidental	1
14	End User Efficiency	Average	3
DI =			40

c. Calculating FP: FP = UFP * [0.65+0.01*DI]

⇨ FP = 518 *[0.65 + 0.01 * 40] = 543.9

d. Calculating LOC: FP * Language Ration for Python (64 for high level language)

⇨ 543.9 * 64 = 34809.6 LOC

e. Calculating KLOC: LOC/1000

⇨ 34809.6 /1000 = 34.81 KLOC.

2. Do the Estimation:

Accordingly this is under the Organic Mode and since our project is considered intermediate we used this table:

The Intermediate COCOMO equations take the form:

$E = a_i (KLOC)^{b_i} * EAF$

$D = c_i (E)^{d_i}$

SS = E/D persons

P = KLOC/E

EAF = Effort Adjustment factor
E = effort
D = Deployment time
SS = staff size
P = productivity
a_i, b_i, c_i, d_i = Coefficients

Co- efficients for Intermediate COCOMO

Project	a_i	b_i	c_i	d_i
Organic mode	3.2	1.05	2.5	0.38
Semidetached mode	3.0	1.12	2.5	0.35
Embedded mode	2.8	1.20	2.5	0.32

a. Calculate the Cost Drivers:

⇨ EAF = DATA * TURN * PCAP * LEXP * TOOL =
1*1.07*0.86*0.95*0.83 = 0.73

b. Calculate Effort Estimate:

⇨ E = ai(KLOC)bi * EAF = $3.2*(34.81)^{1.05}*0.73$ = 97.11 PM

c. Calculate Duration:

⇨ D = ci $(E)^{di}$ = 2.5 * $97.11^{0.38}$ = 14.23 M

d. Calculate Staff Size:

⇨ SS = E/D = 97.11/14.23 = 6.82 Persons

e. Calculate Productivity:

⇨ P = KLOC/E = 34.81/97.11 = 0.36