Department of Computer Science & Engineering

Course Title – Software Development

Course Code – CSE 410

Section - A_1

Project Report

Project title: Brief Bulletin

Date of Submission – October 27, 2022

Submitted by:

Md. Hasibur Rahman (19101009)
Tanmoy Mazumder (19101013)
Shawan Das (19101020)

Submitted to:

Md. Mahedi Hassan
Lecturer, CSE
University of Asia Pacific

Table of Contents

Brief bulletin

Introduction

Though it may be interesting or even entertaining, the foremost value of news is as a utility to empower the informed. The purpose of journalism is thus to provide citizens with the information they need to make the best possible decisions about their lives, their communities, their societies, and their governments. News is important because it informs our view of the world, and in response we take action and make choices based on how we perceive the world to be.
However, as the mass people's days keep getting busier, free time for news consumption gets lesser. This project aims to solve this exact problem. The site Brief Bulletin will host summarized version of recent news collected from multiple websites.

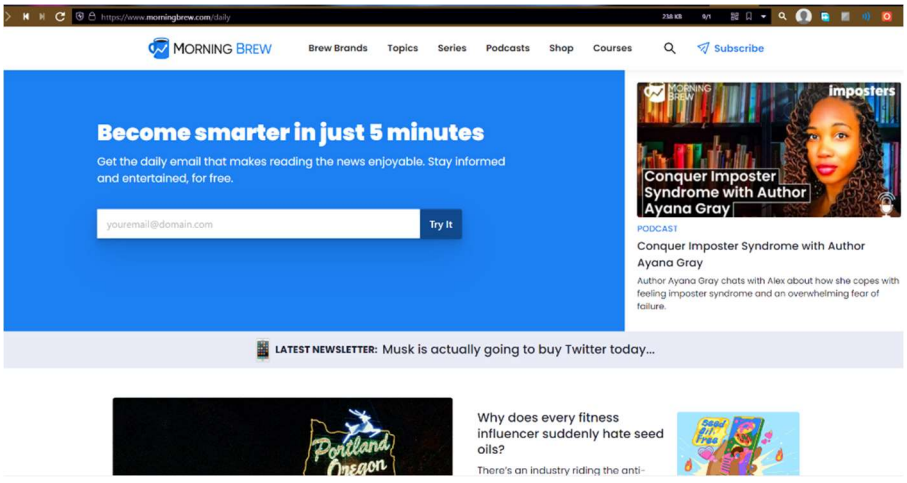
Novelty of this work: No other site at this moment serves news that has been summarized by machine learning.
Project related issues: This site poses no health, safety or legal issues to the site user.
GitHub Link: https://github.com/tanmoyMazumder/summarized-news-portal

Background study

NLP is a category of machine learning that deals with language based tasks, one of which is summarization. This project uses a package called newspaper3k to apply that NLP summarization model for the news. Additionally, python libraries bs4, shutil has been used for web scraping, and the backend is done with PHP.

Existing similar projects

Previously there has been no site that uses machine learning to serve summarized news. However, there are websites like morning brew that do so through manual labor.



Shortcoming of existing similar projects

Morning Brew is limited by its human summarizer's capabilities. The manual summarization process is far more time consuming compared to an automated system. Also there is the limitation of the manual news collection process being slow as well.

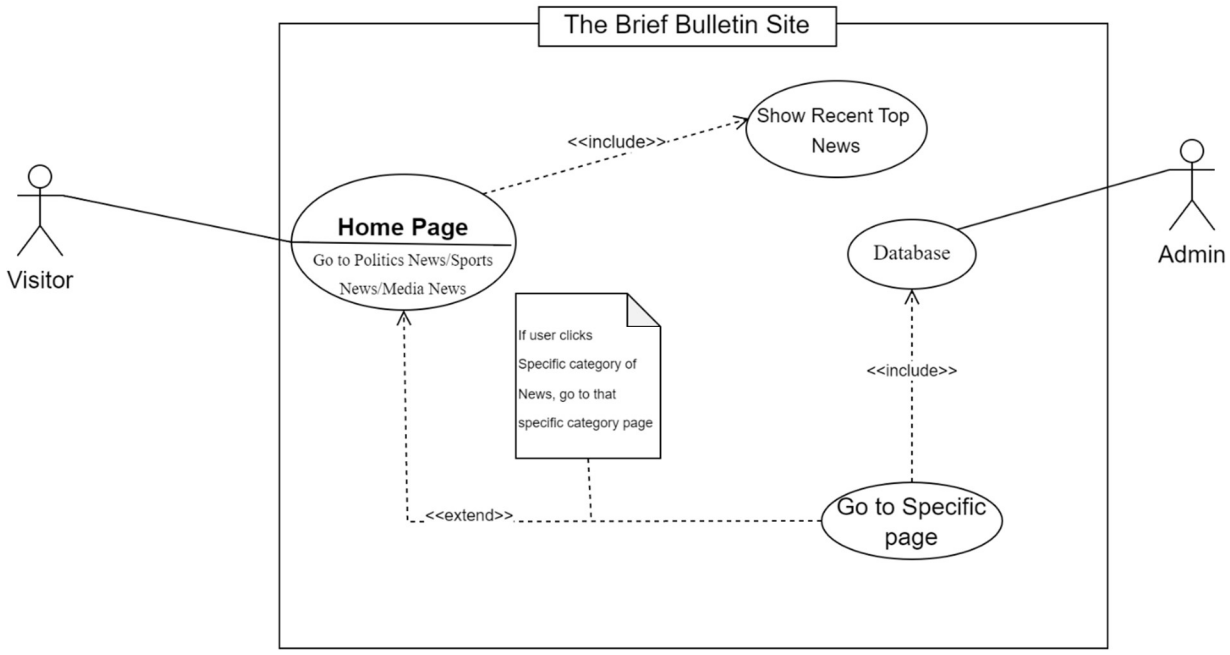
Problem Definition

The project needs to be capable of collecting news from multiple sources and summarizing them, along with other news site features like categories, search, and comments.

Methodology

UML is a combination of several object-oriented notations: Object-Oriented Design, Object Modeling Technique, and Object-Oriented Software Engineering. It uses the strengths of these three approaches to present a more consistent methodology that's easier to use and represents best practices for building and documenting different aspects of software and business system modeling.

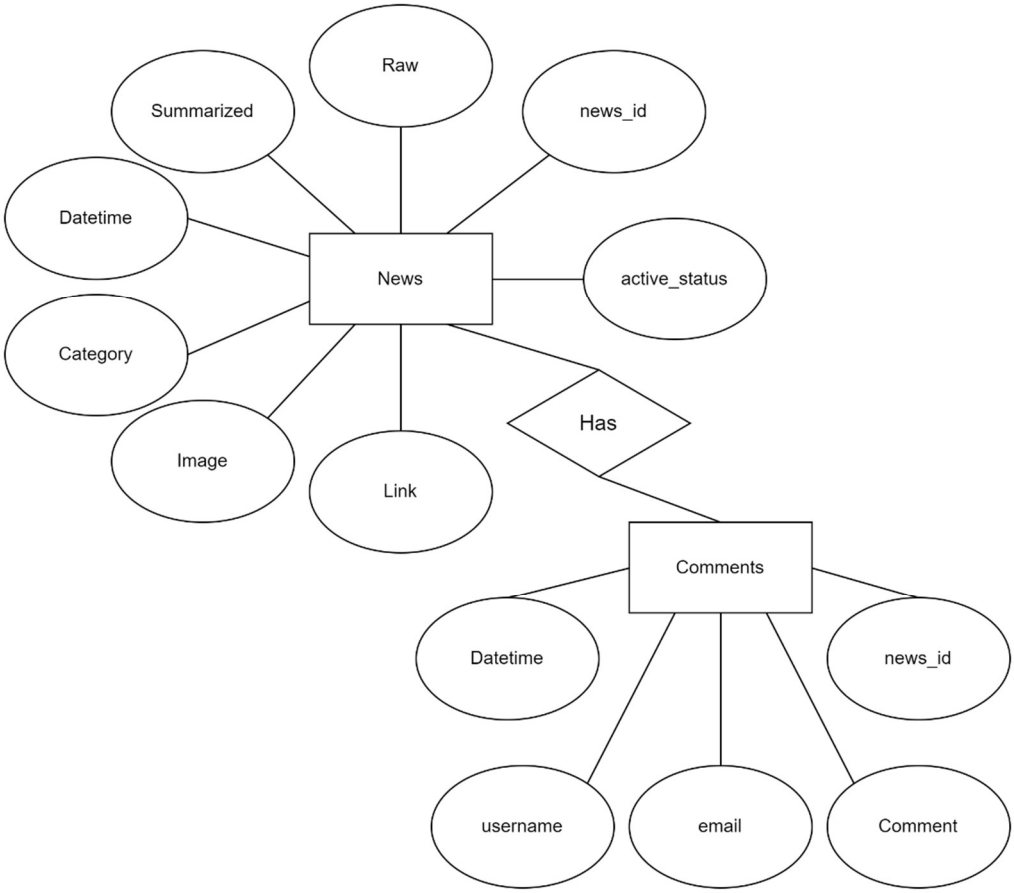
Use case diagram



The only actor would be the visitor and the admin. Visitors can use the site to read news, while the admin manages post deletion, category management, advertisements management etc from their position.

ERD

An Entity-Relationship (ER) Diagram is a type of flowchart that illustrates how “entities” such as people, objects, or concepts relate to each other within a system. ER Diagrams are most often used to design or debug relational databases.



The News table holds the news itself, it’s summarized version, date and time, category, image, link, id and active status.
The comments table has the ID of the news its for, the commenter’s name, email, date and time, and the comment itself.

Cost analysis

			Year 0	Year 1	Year 2	Year 3	Year 4	Total
Benifit		Advertisement	0	75,000	120,000	150,700	200,000	545,700
		Total Benefit	***0***	***75,000***	***120,000***	***150,700***	***200,000***	***545,700***
COST	**Development Costs**	Site Development & Publishing	20,000	0	0	0	0	20,000
		Server Hosting	0	10,000	10,000	10,000	10,000	40,000
		Maintenance Cost	0	15,000	15,000	15,000	15,000	60,000
		Total Development cost	***20,000***	***25,000***	***25,000***	***25,000***	***25,000***	***120,000***
	Operational Costs	Advertisement	0	60,000	50,000	50,000	40,000	200,000
		Management	0	30,000	40,000	50,000	50,000	170,000
		Total Operational Cost	***0***	***90,000***	***90,000***	***100,000***	***90,000***	***370,000***
	Total Costs		20,000	115,000	115,000	125,000	115,000	490,000
Net Benefits = Total benefits - Total cost			[20,000]	[40,000]	5,000	25,700	85,000	55,700
Cumulative Net Cash Flow			[20,000]	[60,000]	[55,000]	[29,300]	55,700	
Return on Investment (ROI)			**11.36% (55,700 / 490,000)*100%**					
Break-even Point (BEP)			**2.3 years [2 years + (85,000 - 55,700) / 85,000]**					

Setup:

- Language: Python, HTML, CSS, Javascript
- Framework: PHP
- Library: Newspaper3k, requests, beautifulsoup, shutil
- IDE: vs code
- Version Control: Git & GitHub

User Manual

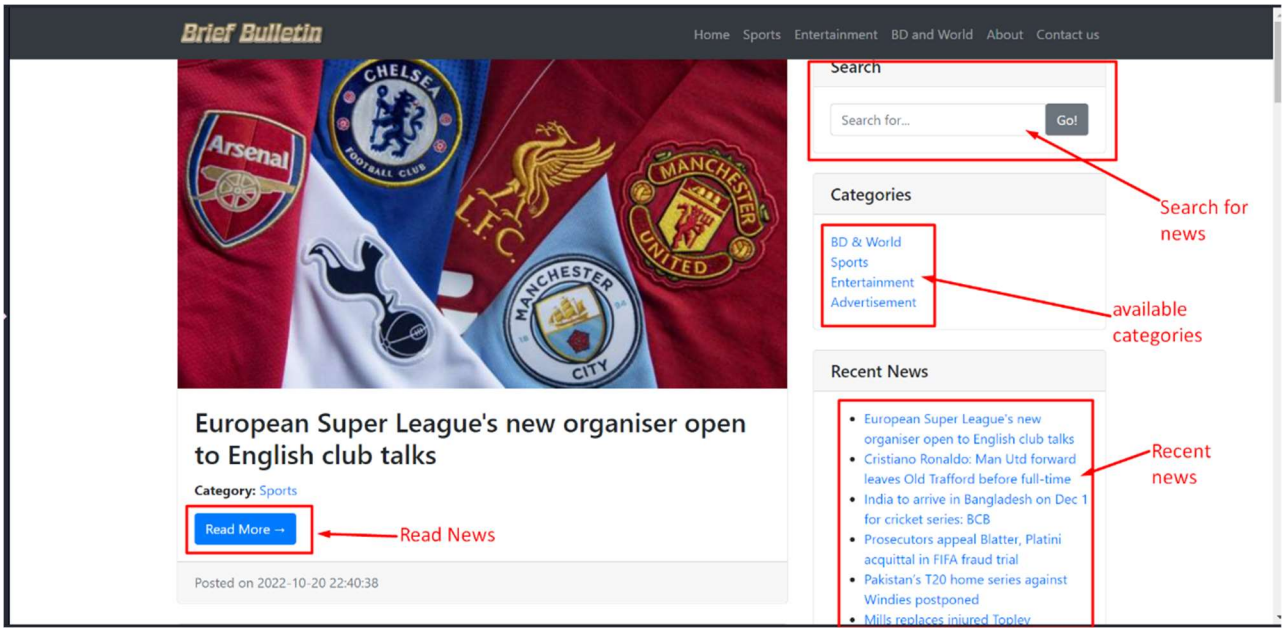
1. Visit Page: Users can visit website, and read news. This requires no user login. The website will show the Title of the news and the Category that the news belongs to. User have to click to “Read More” option to read the news summary.

2. Get News Source: Once the user is reading the news, there will a source link of the news where the user can read the full news.

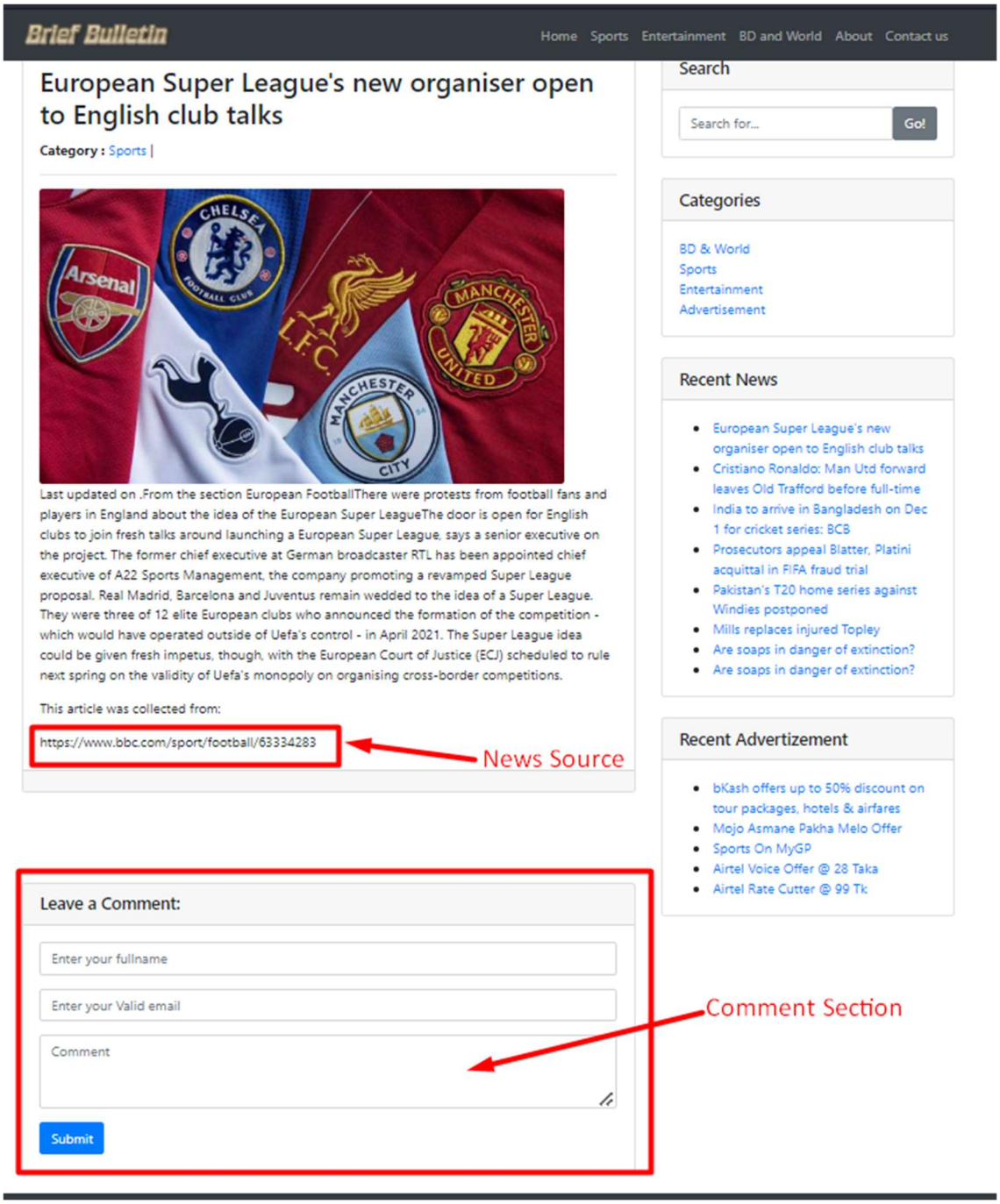
3. Search for News: User can search for news from the search option, at top of the side bar.

4. Read Specific News: User can read specific Category based news. S/He need to select the category from the side bar and the website will the specific category-based news

5. Read Recent News: User can see or read recent news that will be shown at the side bar of “**Recent News**” option. It will show 8-most recent news.



6. Comments on News: User can make comments over newsies. User will be required to



enter his/her full name, email id, and comment then click to “Submit” button below the comment area. Once User submit his comments, it will in approval section where the admin will approve the comments.

NB. Admin can delete or hide any comments if it necessary.

7. Sear Information/Contac us: User can get website details and Contac information from “**About**”, “**Contact us**” section

Conclusion

While working on this project, we came to learn to make a complete news website, integrate different programming languages to execute the whole system, implement web scaping to collect data and also learn how to solve the problems throughout working on the project. We also thought about updating the project in the future and so we generated some ideas as explained below.

Things we have learned

We learnt to build a site with an NLP model at the backend. Knowledge of web scraping and data processing also was required. Choosing the right framework is important. Sometimes the problem simply comes down to trial and error. It is important to have a detailed plan from the very beginning which needs to be open to updates as needed.

Difficulties We Have Faced

Setting up the code for scraping The Daily Star, The Daily Sun, and BBC website was a tedious process. Saving the pictures required multiple attempts with urllib, then had to resort to shutil. Connecting and saving to database also required more effort than expected. MySQL admin had not been updating, this had been causing multiple problems in creating and editing admin. Fixing these was truly a cumbersome task. Trying to make the header transparent had been causing the whole section to break down. Faced difficulties setting up the footer as the column alignment required custom advanced CSS. In version control, one of our members had faced an unknown bug that had been corrupting the commits. Other than these there had been many errors throughout the whole process of building the project, which were dealt with one by one.

Future Directions

This project completes the scope of our curriculum, we desire to advance with it and build it to a greater conclusion. We would like to deploy this site to be online and perhaps move on to an app as well.

Reference

https://www.youtube.com/watch?v=z4DQYprjPSs
https://www.youtube.com/playlist?list=PLuz358yNU7Mkz2nDJa1UX7BpGTj2eGlMB
https://www.youtube.com/watch?v=x7SwgcpACng&list=PLB5jA40tNf3tRMbTpBA0N7IfDZNLZAa9G

Appendix: A (CEP Mapping)

Complex Engineering Problem & Mapping

How Ps are addressed through the Project

Ps	Attribute	How Ks are addressed through the project	COs	POs
P1	Depth of Knowledge Requirement	Our project requires Machine learning (K2), rigorous study of existing projects(K8), Monitoring new websites(K3, K4), knowledge of web-development(K6), data collection & analysis(K5).	CO1 CO2 CO3 CO7	PO1 PO2 PO3 PO5
P3	Dept of Analysis Requirement	Observe the types of news people want to see. Try to publish effective and important news.	CO4 CO7	PO2 P12
P4	Familiarity of Issues	We need to Analyze & implement various programming knowledge and skills to fix many issues and to get proper summary of news	CO7	PO3 PO9
P7	Interdependence	Our project involves interdependent components such as requirement analysis, designing, back-end, front-end, Model accuracy testing, database etc.	CO2 CO7	PO2 PO10

How AS are addressed through the project:

PS	Attribute	Ps are addressed through the project
A1	Range of Resources	Published news collected. Web framework is an important resource for this project. It engages diverse resources including various designing tools.
A2	Level of Interaction	Successful interaction between Web framework and AI model
A3	Familiarity	Our project deals with people who don't get enough time for reading news.

How COs are addressed through the Project:

Ps	CO Statement	Corresponding POS
CO1	Identifying a real-life problem that can be transmitted to an engineering or computing solution through design, development and validation.	PO4 PO10 PO12
CO2	Identify, formulate and analyze a real world compels engineering problem based on requirement	PO2 PO3
CO3	Design/Develop a working solution on a complex software-intensive system and verify and validate the solution using industrial state of the practice, that indicates a high-quality software-intensive system	PO1 PO5 PO11
CO4	Use a modern/popular IDE to test complex software-intensive systems.	PO7
CO7	Work as a team and fulfill individual responsibility	PO9

CO	CO Statements:	Appendix-1	Appendix-2
CO1	**Apply** the Engineering knowledge to provide a working solution on a complex engineering problem and submit a mapping.	1	1/Apply
CO2	**Identify, formulate, and analyze a** real-world complex engineering problem based on requirement analysis.	2	1/Analyze
CO3	**Design/Develop** a working solution on a complex software intensive system and verify and validate the solution using industrial state of the practice, that indicates a high-quality software-intensive system	3	1/Apply
CO4	**Use** a modern/popular IDE to test complex software-intensive systems	5	1/Apply
CO5	**Identify** societal, health, safety, legal and cultural issues related to the project.	6	1/Analyze

CO6	**Practice** concepts of professional ethics, confidentiality, industrial standards.	8	3/Valuing
CO7	**Work** as a team and fulfill individual responsibility.	9	1/Apply
CO8	**Communicate** effectively through presentation and write effective reports and documentations on the project.	10	1/Apply
CO9	**Apply** project management principles using Version Control System, and appraise project operating cost, financial risk analysis for complex software intensive systems.	11	1/Apply
CO10	**Recognize** the need for, and have the preparation and ability to engage in independent and life-long learning for art of project management, distributed and collaborative software development and risk analysis for developing complex software-intensive systems.	12	1/Apply

No.	PO	Differentiating Characteristic
1	Engineering Knowledge	Breadth and depth of education and type of knowledge, both theoretical and practical
2	Problem Analysis	Complexity of analysis
3	Design/ development of solutions	Breadth and uniqueness of engineering problems i.e., the extent to which problems are original and to which solutions have previously been identified or codified
4	Investigation	Breadth and depth of investigation and experimentation
5	Modern Tool Usage	Level of understanding of the appropriateness of the tool
6	The Engineer and Society	Level of knowledge and responsibility

7	Environment and Sustainability	Type of solutions.
8	Ethics	Understanding and level of practice
9	Individual and Team work	Role in and diversity of team
10	Communication	Level of communication according to type of activities performed
11	Project Management and Finance	Management required for differing types of activity
12	Lifelong learning	Preparation for and depth of Continuing learning.