Data Collection and Preprocessing Phase

Date	15 july 2024
Team ID	SWTID1720084639
Project Title	Beneath the Waves: Unraveling Coral Mysteries through Deep Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	Beneath the Waves is an ambitious project that employs deep learning techniques to unravel mysteries surrounding coral reefs. By leveraging advanced algorithms and neural networks, this initiative aims to analyze and understand various aspects of coral ecosystems, contributing to conservation efforts and ecological research.
Data Collection Plan	The data for Beneath the Waves will be collected from various sources, including underwater imagery from coral reef locations, as well as publicly available coral image datasets and online repositories.
Raw Data Sources Identified	The dataset used in this project was collected from Kaggle, a platform for data science competitions and projects. This dataset is designed to help you explore and classify images of these

magnificent marine creatures. It contains a diverse collection of

high-quality images captured in various oceanic

environments worldwide.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The dataset contains coral images in bleached and healthy forms. It also contains the same forms in CSV also in it.	https://www.kaggle.com/datasets/gauravduttakiit/corals-image-classification/data	Image and CSV	27 MB	Public