

Development of Coastal and Marine Biodiversity Integration Network (CoMBINe)

The Coastal and Marine Biodiversity Integration Network (CoMBINe) is a unique coastal and marine biodiversity web portal that combines multiple databases in a single platform that allows searching the huge wealth of information from multiple, independent, heterogeneous datasets about Indian coastal and marine flora and fauna, housed in biological repositories. The database would include all macro plants (seagrass, mangroves, seaweeds) and animals (Phylum Porifera to Chordata) inhabiting the coastal regulation zone (I to IV) as defined in CRZ 2011 notification.

A national level scoping consultation was conducted on 28th June, 2013 in which the architecture and roadmap for CoMBINe were finalized. The CoMBINe architecture would provide for integrating the biodiversity databases scattered among several institutes and local bodies with incompatible data types.

The database architecture has been built as per international codes so as to share data between national and international sites. The primary purpose is to create a common language for sharing biodiversity data that is complementary and reuses metadata standards from other domains wherever possible.

CoMBINe has two major components (i) developing and sustaining a web portal to serve as a national repository for coastal biodiversity and (ii) developing and operationalizing an innovative schema for character-based field identification through re-engineering conventional taxonomy.

1. Developing and sustaining CoMBINe Web portal

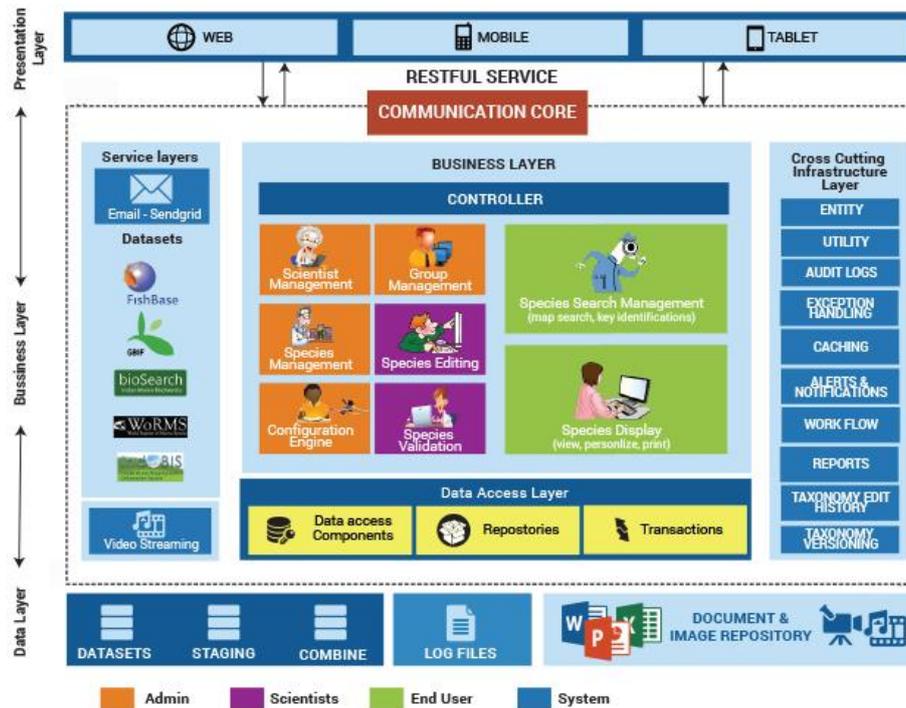
CoMBINe uses the Darwin Core 2.0 metadata standard. Darwin Core is a standard for sharing data about biodiversity, which has a relatively long history of community development and is deployed widely. CoMBINe dataset is an *extension of Darwin core* and has a special customization to adopt to the Indian Biodiversity practices (taxonomy/specimen centric) and to the global practices of storing information of ocean species covering all major areas of biodiversity data.

CoMBINe is built on Darwin Core 2.0 metadata standard along with **Simple Object Access Protocol (SOAP)**, which is a request and response message format for communication between data provider, portal engine, and applications, by linking the metadata and the existing databases already available in different organizations. The adoption of SOAP enables data sharing between GBIF, OBIS, FishNet, MaNIS, HerpNet, PaleoPortal and other international databases.

CoMBINe utilizes Open Cloud configuration. This architecture enables the web servers to utilize the Cloud Block Storage for extra disk space. Cloud Load Balancers are also used in multiple levels of the application, balancing not only web workloads but application workloads as well.

CoMBINe Application Server Architecture is a comprehensive design of the entire system, including all its sub-components and external applications interfaces. The architecture is designed in a way to

access the application through a Web browser or a specialized user agent. The browser creates HTTP requests for specific URLs that map to resources on a Web server. The server renders and returns HTML pages to the client, which the browser would display. CoMBINE is designed keeping in mind the overall objective of integration, scalability and interoperability within the system, as well as with external systems.



CoMBINE Application Server Architecture

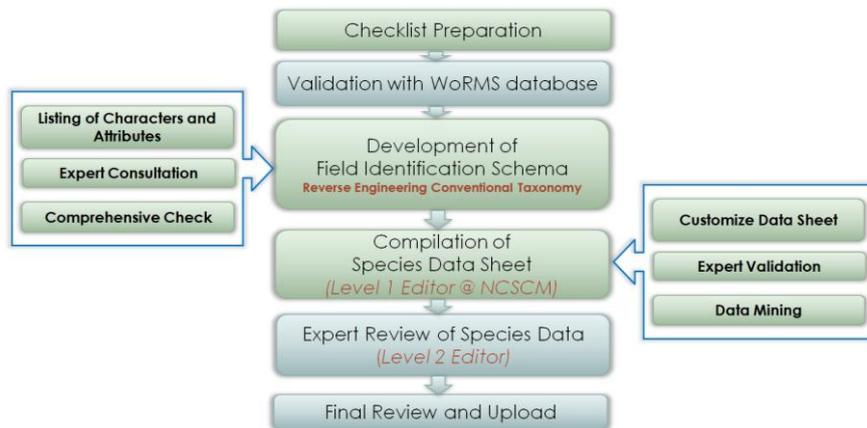
A complete and comprehensive database is designed with 62 fields covering taxonomy, anatomy, biology, conservation, molecular, spatial, spectral and others. The DiGIR protocol will allow more precise data searches because it is designed to reduce information overload and extract common characteristics from all data sources in different formats.

As a first step, information pertaining to organisms inhabiting the coastal and marine environment are being gathered from existing databases, species records, publications and other secondary sources and stored in CoMBINE staging database, which currently has over 72,000 records.

2. Developing a character-based field identification schema

While significant work has been done on coastal biodiversity in India, comprehensive national checklist of species is not readily available. Effort is underway to compile the checklists for all groups from various sources.

The “character/image based field identification” feature aims at taking the biodiversity information to all stakeholders as well as to create interest among them, irrespective basic understanding on taxonomy, through re-engineering conventional taxonomy, thus bridging the knowledge-divide.



Steps involved in building CoMBiNe

The coastal and marine organisms are grouped into 20+ groups based on broad morphological features and for each such group, which would include different classes across phyla, key identification characters and attributes are being developed.

It is envisaged that the CoMBiNe portal, the first Indian biodiversity portal to be built as per International codes, would emerge as the national repository for coastal and marine biodiversity. Steps have been taken for developing the mobile portal of CoMBiNe so as to integrate people and knowledge in real-time. Further, CoMBiNe would help in identifying taxonomic and biological data gaps so as to steer focused research on coastal and marine biodiversity.