Feedback from the European Marine Board to the European Commission's 'European Strategy for data'

The European Marine Board calls for the creation of a 'Common European Ocean Data Space', in line with the vision for a 'Digital Ocean Twin', as outlined in EMB Position Paper No. 24 'Navigating the Future V'. This 'Common European Ocean Data Space' should align with existing European and International efforts to federate ocean data i.e. Blue-Cloud and an Ocean Data Portal as a key objective of the UN Decade of Ocean Science for Sustainable Development. The 'Digital Ocean Twin' will require ecosystem (and other) modelling, as outlined in the EMB Future Science Brief No. 4 'Enhancing Europe's marine ecosystem modelling capability for societal benefit'.

The European Marine Board's Future Science Brief No. 6 'Big Data in Marine Science' provides recommendations for scaling-up big data applications for improving our understanding of the marine environment and for managing human activities that impact the ocean. In summary, this document recommends to:

- **Enhance data acquisition** through the continued development of 'smart sensors' for automated sampling and data processing so that more marine data can be collected by machines. We also propose to increase the efficiency of data transfer to allow more real-time, or near real-time analyses and decision making;
- **Enhance data handling and management** through more widespread adoption of community data standards and well-designed data management plans based on Findable, Accessible, Interoperable and Reusable (FAIR) principles so that data are machine-readable. We also recommend the increased use of existing marine data management infrastructures;
- **Increase data interoperability and accessibility** by upgrading European marine data management infrastructures to handle and exchange more multidisciplinary and real-time data. These infrastructures should include more integrated cloud computing, data storage and big data analytical tools. We recommend increased participation of the European marine science community in development of Virtual Research Environments (VREs) and European Open Science Cloud (EOSC) initiatives. We also recommend that these infrastructures should be sustained in the long term and that there should be more cross-disciplinary fertilization of computing technology from multimedia and digital sectors. Interoperability and data exchange between European data infrastructures and international counterparts will be important for common access to data on larger sea-basin and global scales and to create a 'digital ocean twin';
- **Improve data sharing** between scientists, industry and governments through new incentives and protocols such as social networks or data impact factors;
- **Increase the use of big data analytics and ensure data validation** by developing close collaborations between data scientists and marine scientists, developing standardized models, and well-curated community data sets to train algorithms;
- **Develop specialized training** on artificial intelligence by establishing new regional and global marine science networks and consolidating already existing networks. We recommend training data curators to maintain the quality of data feeding into artificial intelligence algorithms; and

- **Increase collaborations between marine scientists, computer scientists, data scientists and data managers** in the form of working groups and the involvement of data scientists in the design of marine research.

The data specific recommendations from the EMB's Future Science Brief on Enhancing Europe's capability in Marine Ecosystem Modelling for Societal Benefit include:

- Enhance models by identifying **crucial unavailable data**, linking models to new and existing **observations and data**, and by strengthening links to data assimilation centers;
- Develop a **shared knowledge platform** for marine models and support the development of next generation models;
- Make marine ecosystem models more relevant to management and policy by being more transparent about model limitations and the **uncertainties in their predictions**; including socio-economic drivers; promoting co-design and dialogue between model developers and users; and
- Enhance **trans-disciplinary connections** and training opportunities.