Data Driven Blue Growth – Meeting Ocean Renewable Energy Sector Needs with iMarDIS

Graham Worley, Bangor University (Wales), g.worley@bangor.ac.uk
David Mills, Bangor University (Wales), d.mills@bangor.ac.uk
Thomas Prebble, Bangor University (Wales), t.prebbles@bangor.ac.uk
Cathy Blakey, Bangor University (Wales), cathy.blakey @bangor.ac.uk
Gwyn Roberts, Bangor University (Wales), gwyn.roberts@bangor.ac.uk
Colin Jago, Bangor University (Wales), c.f.jago@bangor.ac.uk

Wales has significant ocean energy resources and the EU has recently committed €100m to drive Blue Growth by supporting the growing ocean renewables energy (ORE) sector in Wales. Funded by the European Regional Development Fund the SEACAMS2 project (http://www.seacams.ac.uk) undertakes a programme of collaborative research to support the ORE sector. SEACAMS2 specifically emphasises the de-risking of business decisions for tidal stream, tidal range and wave renewable energy companies in Wales. As part of SEACAMS2 a new integrated data management and information system (iMarDIS) is being developed that will secure existing data and employs state of the art systems for archiving data and making it available for re-use.

SEACAMS2 collaborative research requires observations that are multi-disciplinary, variable frequency in space (< 1m to > 10km) and time (seconds to years) and encompass intertidal, shallow coastal and offshore domains. Observations range from sea surface to seabed and sub-seabed using research vessels, moorings, seabed landers and UAVs. The resulting data sets are large, complex and require rapid processing for timely delivery to meet industry and other stakeholder needs. Furthermore, SEACAMS has a legal and contractual requirement to share the publicly funded data it has acquired.

The design of iMarDIS (www.imardis.org) is 'end user' driven. At the same time, it was necessary to ensure that iMarDIS: adopted existing UK MEDIN meta-data standards, could interface with the existing UK and European data management infrastructures, avoided replication of existing capability, was capable of ingesting and disseminating in real-time very high resolution data and could streamline the workflow from data collection to data management and curation.

A review of existing data services within the UK showed they would be unable meet the demands of SEACAMS end users either in terms of speed of data access, the high resolution of the data or in meeting our need to construct applications that would retrieve data on a machine-to-machine basis for further processing. Furthermore, it became apparent that there were other issues that prevented their adoption within the context of iMarDIS including difficulties in scaling the technology for concurrent access by many users or handling very large volumes of data and allowing new applications to be built upon the data archive.

Consequently, a concept for the overall iMarDIS infrastructure was developed that required new software that was cloud hosted. The system (Figure 1) was conceived as a series of micro-services, each capable of operating semi-independently, and communicating with each other through a RESTful JSON based Application Programming Interface layer (API). These APIs will eventually become public. There would be a data discovery and download portal and advanced services website communicating with a back-end set of micro-services. The micro-services would support a range of key functions including security and authentication, metadata management, file upload and download, tabular data manipulation, point cloud (or generic raster) data processing and time-series data manipulation.

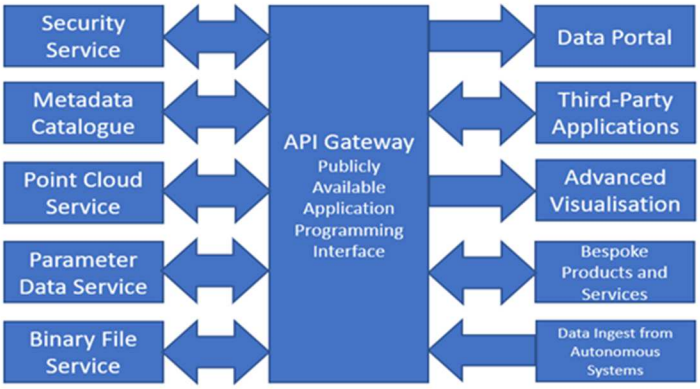

Figure 1. An overview of the iMarDIS architecture showing the 'back end' microservices to the left of the Application Programming Interface and the applications domain on the right handside. The applications domain is where advanced service capability can be implemented either by iMarDIS or third party developers.

Breaking the problem down into micro-services allowed three developers to work in parallel on each service to reduce the time taken for infrastructure development. The architecture is based around Enterprise Java REST services implemented within the Amazon Web Services cloud. This is both scalable in terms of storage capacity and throughput.

An important feature of this design is that the iMarDIS hosted portal acts as a showcase for its capability in terms of data discovery and download and advanced services. Using published APIs external software developers could develop their own innovative applications reliant on the iMarDIS base infrastructure. This will encourage collaboration and innovation and promote sustainability of the infrastructure by increasing the diversity of its use.

To manage and collect metadata and upload data to the infrastructure a metadata manager application has been written and is in daily use. This also provides a level of orchestration of the interactions between the micro-services through the passing of messages to a series of queues. To date many of the back-end services have been written and are in an advanced state of testing. Work on the front-end application will be offered as an external contract in the near-term.

In parallel with the development of the micro-services a data librarian has been ingesting data and metadata into the system and around 4Tbytes of a potential 30Tbytes has now been catalogued. In order to facilitate data sharing a new licence agreement has been drawn up that allows commercial exploitation of the work but addresses the needs of researchers who ultimately are providing the data for SEACAMS.

As well as ensuring existing SEACAMS data are ingested and made available for re-use new *in-situ* data streams are beginning to be ingested and published to the iMarDIS website. For example, in response to user requirements a relocatable Coastal Observatory is being developed as part of the SEACAMS2 programme to improve understanding of key coastal processes. The first mooring of the network is currently being deployed in the Irish Sea and multi-variate data from the first of these buoys in already streaming live into the iMarDIS infrastructure.

Detailed knowledge of user requirements has been key to successful collaborative research programmes with industry. Building on this knowledge via further stakeholder engagement is critical to the design of iMarDIS. Consequently, a stakeholder workshop was recently held that brought together a cross-section of the community actively engaged in development of the ORE sector in Wales including commercial developers, consultants and regulators. The workshop report (https://www.imardis.org/workshop2018/Report-En.pdf) highlights the detailed outcomes from the successful workshop that are being used to inform the next stage of design of iMarDIS, further ensuring that the informatics infrastructure will meet stakeholder needs.

The next phase of development for iMarDIS includes development of the website that will act as the main presentation layer of the initiative, further enhancements of the micro-services, development of access libraries to aid end-users in connecting tools like MATLAB, R etc. directly to iMarDIS data feeds, as well as continued ingestion of SEACAMS legacy data and the publishing of discovery metadata to MEDIN and EMODnet.