

Scaling metadata catalogues with web-based software version control and integration systems

Tara Keena, Adam Leadbetter*, Andrew Conway, Will Meaney

Marine Institute, Rinville, Oranmore, Galway, Ireland

*adam.leadbetter@marine.ie



The ability to access and search metadata for marine science data is a key requirement for answering the "findability" aspect of the Findable-Accessible-Interoperable-Reproducible principles of data management (Wilkinson et al., 2016). It is also vital in meeting domain-specific or community defined standards and legislative needs placed on data publishers. Appropriate cataloguing with the storage and publication of descriptive metadata for end users to query online is a necessary step to enable this requirement.

With observing systems constantly evolving and the number of platforms and sensors growing, the volume and variety of data is constantly increasing. Therefore metadata catalogue volumes are also expanding. The ability for data catalogue infrastructures to scale with data growth is a necessity, without causing significant additional overhead, to data publishing facilities. A potential solution for maintaining scalable data catalogues and hosting a variety of file types, all with minimal overhead costs, is proposed below. The outputs are available to human users (HTML), machines (RDF) and international networks (XML).

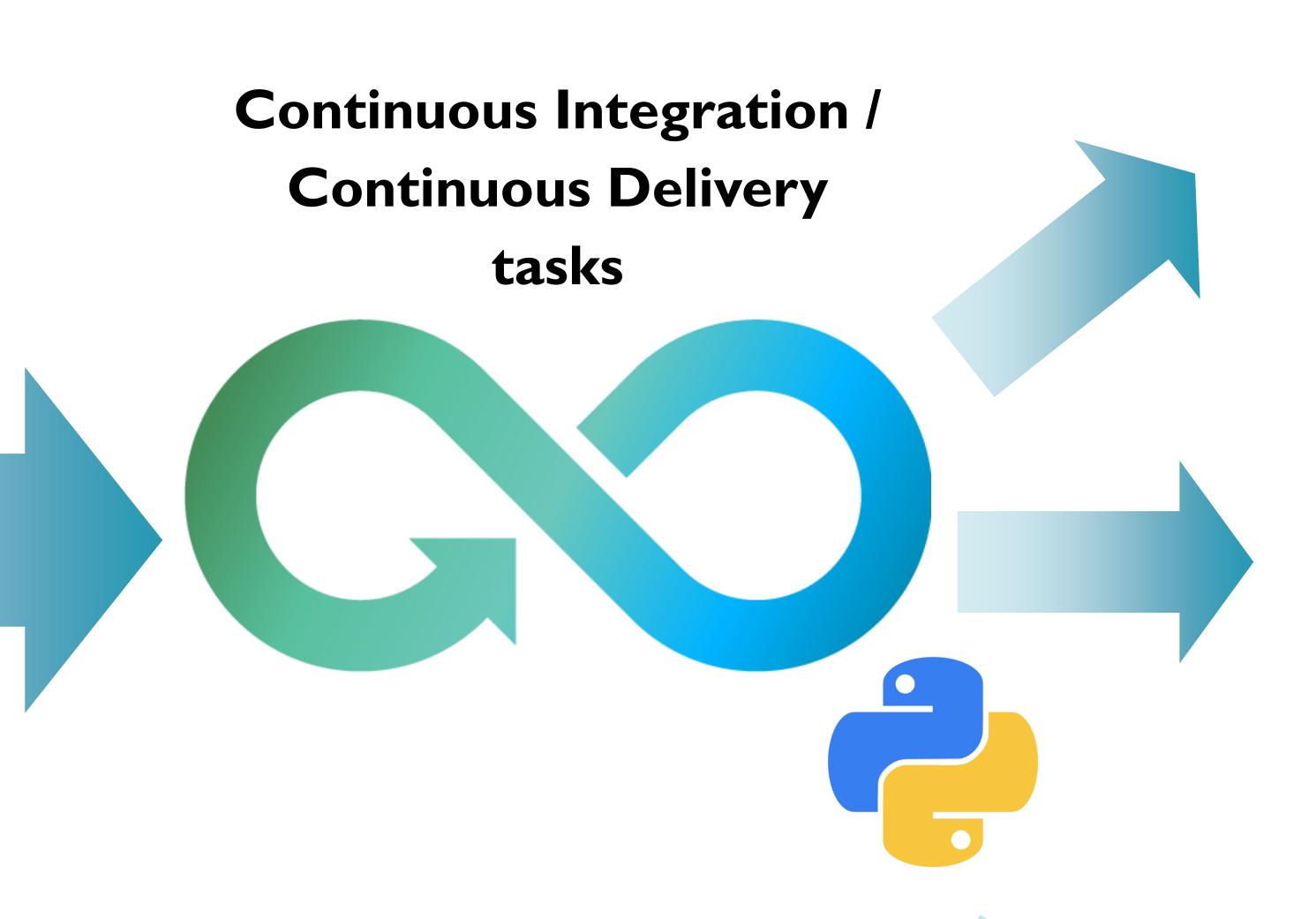
Marine Institute Data Catalogue (Leadbetter et al., 2020)

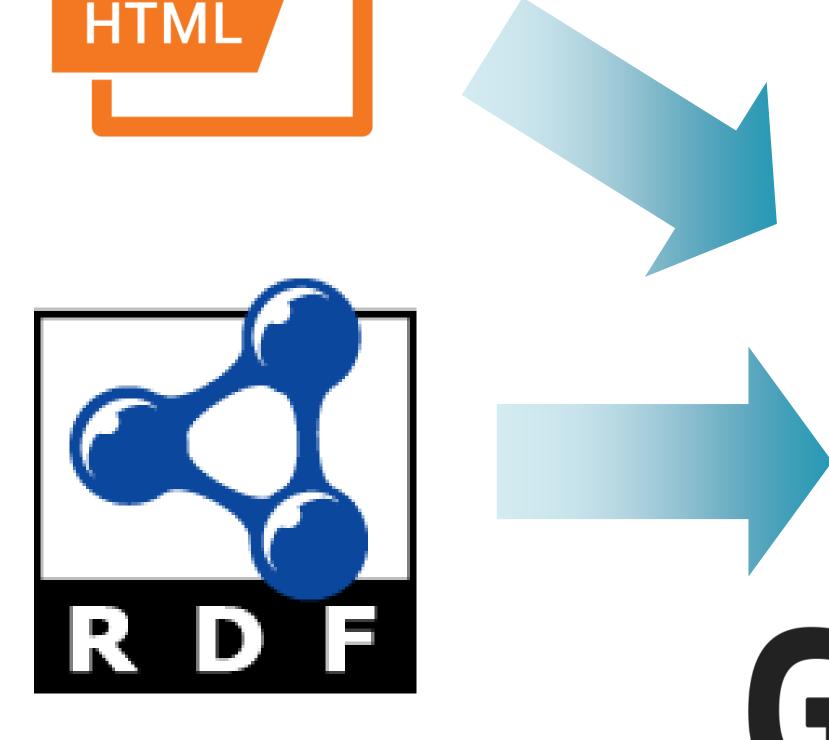




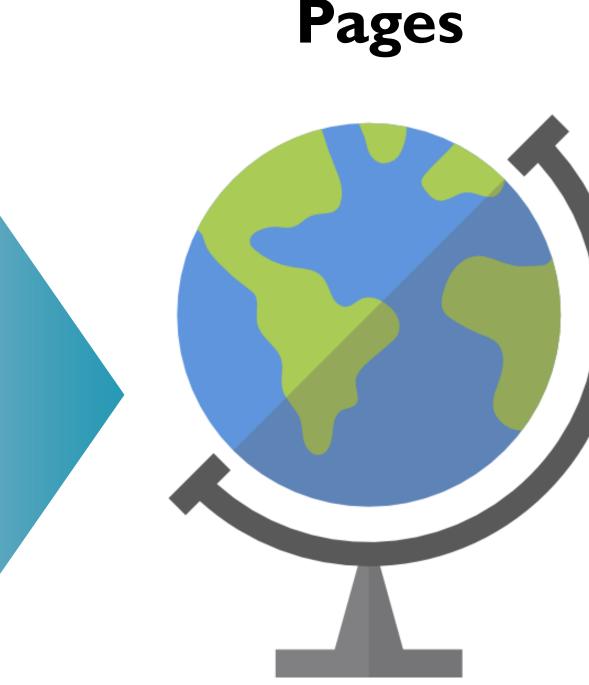
repository



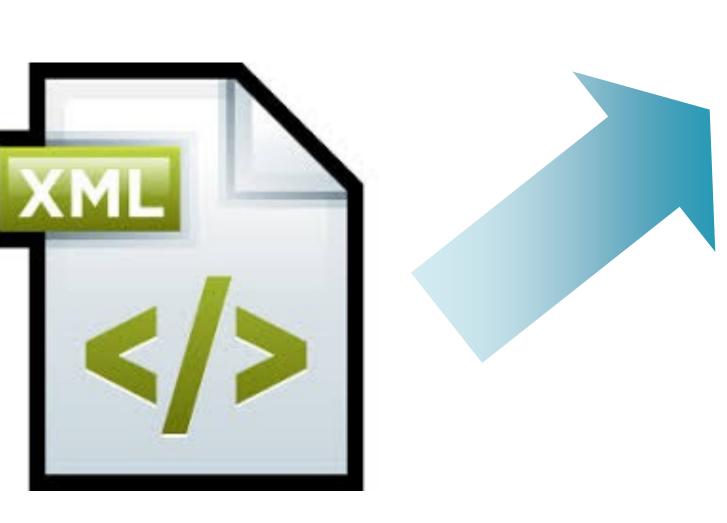








Web access via GitHub







COMPASS



European Regional Development Fund



Acknowledgements

This work is part supported by the Irish Government and the European Maritime & Fisheries Fund as part of the EMFF Operational Programme for 2014-2020. This work was is also part supported by the COMPASS project. COMPASS is supported by the European Union's INTERREG Va Programme, managed by the Special EU Programmes Body (SEUPB).

References

Leadbetter, A., Meaney, W., Tray, E. et al. A modular approach to cataloguing marine science data. Earth Sci Inform (2020). https://doi.org/10.1007/s12145-020-00445-w Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/ sdata.2016.18





An Roinn Talmhaíochta, Bia agus Mara Department of Agriculture, Food and the Marine



