

An open-source database model and collections management system for fish scale and otolith archives

Elizabeth Tray, Galway-Mayo Institute of Technology (Ireland), Elizabeth.Tray@gmit.ie
Adam Leadbetter, Marine Institute (Ireland), Adam.Leadbetter@Marine.ie
Will Meaney, Marine Institute (Ireland), Will.Meaney@Marine.ie
Andrew Conway, Marine Institute (Ireland), Andrew.Conway@Marine.ie
Caoimhín Kelly, Marine Institute (Ireland), Caoimhin.Kelly@Marine.ie
Elvira de Eyto, Marine Institute (Ireland), Elvira.Deeyto@marine.ie
Siobhan Moran, Marine Institute (Ireland), Siobhan.Moran@Marine.ie
Niall Ó Maoiléidigh, Marine Institute (Ireland), Niall.Omaoileidigh@marine.ie
Deirdre Brophy, Galway-Mayo Institute of Technology (Ireland), Deirdre.Brophy@gmit.ie

Biomineral samples, such as scales and otoliths (ear stones), from fish are routinely sampled for age estimation and fisheries management purposes. Fish scales and otoliths can be used to answer a diverse range of ecological research questions using both established and advanced analytical techniques. Measurements of growth marks in calcified structures provide a proxy for fish growth, and can be used to assemble extended biochronology time series and to examine relationships with climatic and fishing related drivers (Morrongiello et al. 2012). Studies have used trace elements and isotopes in the inorganic portion of biomineral structures to infer fish origin (Adey et al. 2009) and habitat usage (Brennan et al. 2019). Continuous sampling programs and opportunistic sampling of fish scales and otoliths can create invaluable, albeit irregular, archive collections. A common problem associated with biological collections, is that while sample intake grows, long-term physical storage is rarely a priority. The maintenance of scale and otolith collections can change depending on management and project-specific research needs (Rivers & Ardren, 1998). Essential fish metadata (**Figure 1**) is often recorded in a field notebook or a digital spreadsheet, which is rarely stored with the associated physical samples, and if it is, it can become separated after time. Proprietary software is currently available for implementing collection management (CM) systems, but it can be expensive, dependent on ongoing investment, and limited in scalability.

Global legislative efforts to ensure that research data remains accessible have contributed to significant improvements in data quality and interoperability (Clarke & Magetts et al. 2014). At the European regional level, the INSPIRE Directive is a legislative driver for governments to construct standardized spatial data infrastructure, specifically for environmental data (Craglia & Annoni et al. 2007). To comply with INSPIRE regulations, data infrastructures must contain the flexibility to align within an Observations and Measurements (O&M) model, (Cox et al. 2016). Furthermore, other standards structures have gained widespread community usage, such as biodiversity data standards, like Darwin Core (Wieczorek et al. 2012). Recently, the FAIR (Findable Accessible Interoperable and Reusable) open-data principles (Wilkinson et al. 2016), were created through a bottom-up, community based initiative. A component of the FAIR principals requires the usage of existing standards (e.g. INSPIRE, O&M, Darwin Core). Consequently, scientists have implemented FAIR data repositories specifically intended for sample and data management (Conze et al. 2017). However, a FAIR compliant system and model for local repositories, purposefully designed for physical ecological sample archives and their derived data, has yet to materialize.

The Marine Institute (MI) in Ireland, recently constructed an INSPIRE compliant Data Catalogue for its environmental data (Leadbetter et al. 2020). The CM system presented in this study is an extended feature of the Data Catalogue described by Leadbetter et al. 2020. The CM system was explicitly designed to house a fish scale and otolith collection, which consists of thousands of samples, dating

back to 1928, from a range of geographic locations. The aim of this study is to lay a road map, and provide a toolbox, for other aquatic agencies who wish to create an 'open' digital and physical infrastructure for their own biomineral archives. The system utilizes the FAIR (Findable Accessible Interoperable and Reusable) open-data principles, and includes a physical repository, sample metadata catalogue, and image library.

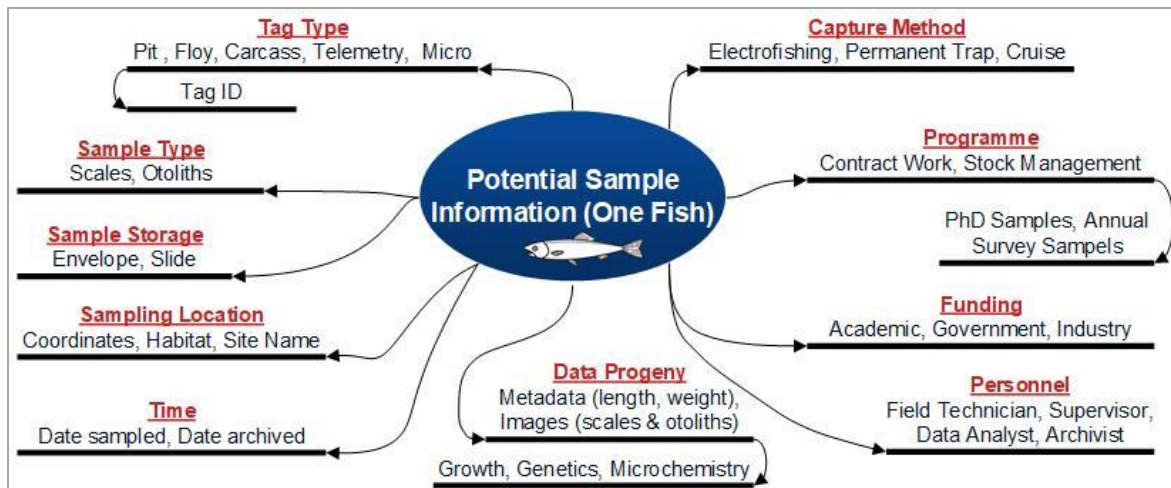


Figure 1: A mind map of the potential information associated with a fish scale or otolith sample.

References:

- Adey, E.A., Black, K.D., Sawyer, T., Shimmield, T.M. and Trueman, C.N., 2009. Scale microchemistry as a tool to investigate the origin of wild and farmed *Salmo salar*. *Marine Ecology Progress Series*, 390: 225-235. <https://doi.org/doi:10.3354/meps08161>.
- Brennan, S. R., Schindler, D. E., Cline, T. J., Walsworth, T. E., Buck, G., and Fernandez, D. P., 2019. Shifting habitat mosaics and fish production across river basins. *Science*, 364(6442): 783–786. <https://doi.org/10.1126/science.aav4313>.
- Clarke, A. and Margetts, H., 2014. Governments and citizens getting to know each other? Open, closed, and big data in public management reform. *Policy & Internet*, 6(4): 393-417. <https://doi.org/10.1002/1944-2866.POI377>.
- Conze, R., Lorenz, H., Ulbricht, D., Elger, K. and Gorgas, T., 2017. Utilizing the International Geo Sample Number Concept in Continental Scientific Drilling During ICDP Expedition COSC-1. *Data Science Journal*, 16:2. <http://doi.org/10.5334/dsj-2017-002>.
- Craglia, M. and Annoni, A., 2007. INSPIRE: An innovative approach to the development of spatial data infrastructures in Europe. *Research and theory in advancing spatial data infrastructure concepts*, 93-105.
- Leadbetter, A., Meaney, W., Tray, E., Conway, A., Flynn, S., Keena, T., Kelly C., and Thomas, R., 2020. A modular approach to cataloguing marine science data. *Earth Science Informatics*. <https://doi.org/10.1007/s12145-020-00445-w>.
- Morrongiello, J. R., Sweetman, P. C. and Thresher, R. E., 2019. Fishing constrains phenotypic responses of marine fish to climate variability. *Journal of Animal Ecology*, 88(11): 1645-1656. <https://doi.org/10.1111/1365-2656.12999>.
- Rivers, P. J. and Ardren, W. R., 1998. The value of archives. *Fisheries*, 23(5): 6-9. [https://doi.org/10.1577/1548-8446\(1998\)023<0006:TVOA>2.0.CO;2](https://doi.org/10.1577/1548-8446(1998)023<0006:TVOA>2.0.CO;2).
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E. and Bouwman, J., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3. <https://doi.org/10.1038/sdata.2016.18>.
- Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T. and Vieglais, D., 2012. Darwin Core: an evolving community-developed biodiversity data standard. *PloS one*, 7(1). <https://doi.org/10.1371/journal.pone.0029715>.