BITE Interim data-framework

The Biotic Interactions in deepTimE (BITE) project is an ongoing databasing project tailored for fossil Biotic Interaction (fBI) data. The below interim data-framework (IFW; Fig. 1) is a reduced schematic adapted from the GloBI (Poelen et al 2014), which was designed for Modern Biotic Interaction (mBI) data. The majority of tables follow a similar structure to GloBI to enable easy integration.

In mBI research real-time recording, intervention, and observation of the biotic interaction (BI) is an important process, e.g. through filming, trapping, eye-witness observation of the BI whilst it is occuring. Consequently environmental conditions (season, altitude, rainfall, etc) are key features require in the database framework to inform the circumstances of the BI's observation. In fBI research, intervention and eye-witness observation of the interaction in process is not possible, the evidence rather held through geological stabilisation (via taphonomic processes) in the rock record with the most certain type of evidence for an interaction being two organisms presented in the process of an interaction (i.e. frozen behavior; Boucot and Poinar 2010). Consequently, key features fBI workers need to record are morphology, geological setting, lithology, chronostratigraphy, and taphonomy. The nature of the fBI also permits the observation of multiple biotic interactions on the same specimen. For example, a fossil placoderm fish may hold evidence of surviving being preyed upon in life (healed dermal plate), having been eaten by a predator (dermal plate located within a coprolite) and then being being scavenged (borings through the dermal plate whilst in situ within the coprolite). Whilst the example was hypothetical, there are many recorded examples of such sequencing within a fossil specimen (e.g. frozen behaviours in amber summarised by Arillo, 2007, and placoderm predation, Robin et al 2022). The sequencing of biotic interaction within the same specimen requires a flexible data structure where multiple interactions can be attributed to a single foundation, and the sequencing of them identified.

Within the interim schematic both mBI and fBI are modelled in terms of the specimen (the evidence), published study, biotic interaction (BI), lithology, location, and actors. The schematic does not include the multimedia datatable, geochronology, nor higher taxonomy table for example (see Dowding et al, in review) . These standard elements of palaeobiological and Earth science databases will be sourced from extended frameworks such as IRAL (*Integrated Record of Ancient Life* initiative), of which BITE is a contributing member.

Within the IFW most IDs are a uniform resource identifier to external vocabularies and/or relevant ontologies which should be flexibly updated. In the case that neither are available a custom term can be established. This flexibility is particularly necessary for BI studies due to domain specific vocabulary, which the forthcoming database will provide a record of, allowing synonymisation and term-matching later with expert input.

Extended aims: BITE beyond the IFW

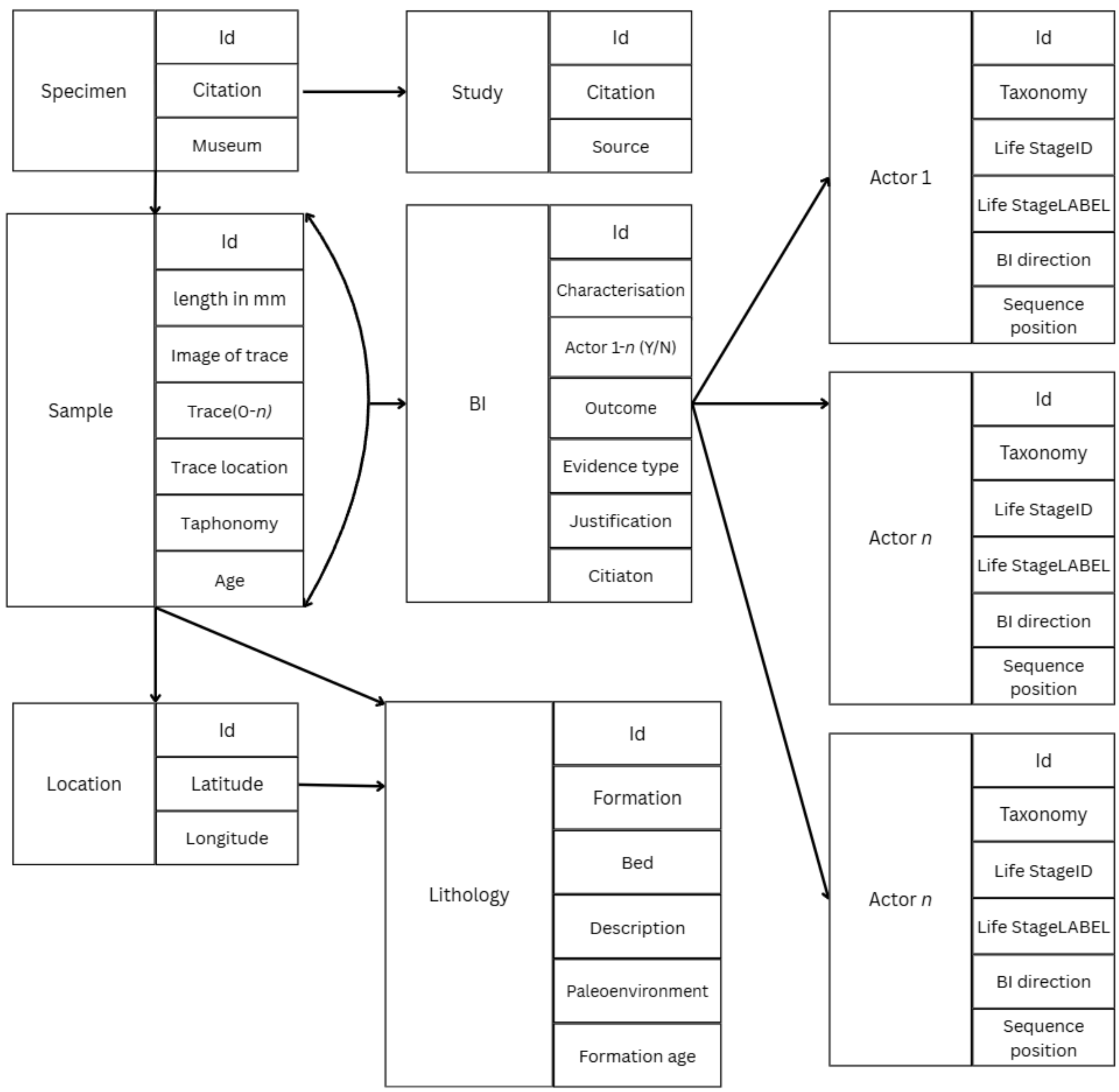
The future BITE database requires an integrated system for the acquisition, normalization, management, and querying of fBI data. The system architecture is designed to address the unique challenges of palaeontological datasets, where information is often fragmentary, heterogeneous, and derived from multiple sources (e.g., museum collections, field observations, and published literature). Due to this and the reasons proposed in the main publication, the specimen evidence is the foundation of the structure, rather than a published scientific study.

The BITE framework will consist of: (a) a data model capable of representing diverse interaction types across deep time and the modern, (b) an ingestion pipeline for acquiring and standardising data from varied formats, including multimedia, specimen catalogs, stratigraphic tables/logs, and text-based reports, (c) a term-matching system to align free-text fossil descriptions with controlled vocabularies, and (d) an application programming interface (API) and web interface for community access and integration.

The core of the framework is a data model (Fig. 1) that captures fossil interactions and their spatiotemporal context. In this model, an interaction is recorded tied to one physical specimen (or occurrence, multimedia, i.e. '*primary evidence*') showing evidence of interaction described using controlled ontologies (e.g., this paper, or Robin 2021). Each specimen is linked to taxonomic identity, functional group (e.g., suspension feeder, apex predator), and environment (e.g., carbonate platform, deep marine shelf) and locations (physical setting) are further classified into broader contexts (e.g., ecoregion, habitat, or facies) using published schemes (e.g., paleoecological facies models, international chronostratigraphy) in addition to the general fields covered in the IFW (Fig 1) .

When possible, each specimen-level record is linked to a study, but each entry must be tied to a data source or contributor (e.g., field project, museum repository, or literature dataset). This allows full traceability from the specimen to its interpretive context, preserving provenance and enabling reproducibility.

Figure 1: Framework schematic outlining the major data elements to be captured. The data structure should link to a physical specimen, its location, and a multimedia feature of the specimen to ensure solid data provenance, transparency in trace interpretation, and evidence, type, and the potential for reinterpretation.



References

Arillo, A., 2007. Paleoethology: fossilized behaviours in amber. Geologica Acta, 5(2), pp.0159-166.

Dowding, E.M., Dunne, E.M., Collins, K., Cryer, K., De Baets, K., Dimitrijević, D., Edie, S.M., Finnegan, S., Kiessling, W., Lintulaakso, K. and Liow, L.H., 2025. Fossils for Future: the billion-dollar case for paleontology's digital infrastructure. ecoevoRxiv, https://doi.org/10.32942/X2DS89

Robin, N. 2021. Importance of Data on Fossil Symbioses for Parasite–Host Evolution. In: De Baets, K., Huntley, J.W. (eds) The Evolution and Fossil Record of Parasitism. Topics in Geobiology, vol 50. Springer, Cham. https://doi.org/10.1007/978-3-030-52233-9_2

Robin, N., Noirit, F., Chevrinais, M., Clément, G. and Olive, S., 2022. Vertebrate predation in the Late Devonian evidenced by bite traces and regurgitations: implications within an early tetrapod freshwater ecosystem. Papers in Palaeontology, 8(4), p.e1460.