Connecting Open Science and Archaeology: The University of Victoria Zooarchaeology Lab Comparative Collection
An Essential Cultural and Ecological Resource

By

Kathryn McKenzie
Bachelor of Science in Anthropology, University of Victoria, 2021

An Essay Submitted in Partial Fulfillment
of the Requirements of the

HONOURS PROGRAM

In the Department of Anthropology

Abstract

Comparative collections are fundamentally important for the zooarchaeological identification of archaeological animal remains. The University of Victoria Zooarchaeology Lab has an extensive regional skeletal reference collection widely used and known by researchers across the Northwest Coast. Accessibility of collections such as this can be improved through Open Science practices. These practices work to improve the discipline of archaeology, advance knowledge discovery, uphold ethics, and strengthen collaborative strategies. This paper uses best practices of data management to create a citable database to enhance accessibility of the UVic collection and provide guidance for open method standards and practices in zooarchaeology. By linking open data resources to the collection, contemporary and deep-time biodiversity data can be shared by researchers to broaden awareness of the collection, inspire data reuse and sharing, create novel research, and foster interdisciplinary collaboration. Educational opportunities and community-based research will connect educators, students, Indigenous communities and heritage specialists to collection resources. The UVic Zooarchaeology Lab will build capacity to become a significant source of anthropological, ecological, and ethnobiological knowledge.

Keywords

Archaeology; Biodiversity; Community-Based Research; Comparative Collections; Darwin Core; Database; Data Management; Ecology; Education; Indigenous Knowledge; Linked Open Data; Natural History Collections; Northwest Coast; Open Methods; Open Science; Zooarchaeology

Acknowledgements

Thank you Dr. Stahl and Dr. McKechnie for your advice and guidance throughout this project. Thank you to my Honours peers and HECA Lab members for your encouragement and contributions to my project. Thank you to my family and partner for your love and support.

Table of Contents

Introduction and Background

Introduction

Approach to Building an Accessible Zooarchaeology Reference Collection Model

In this paper I build a prototype and document the design process for an accessible zooarchaeological reference collection at the University of Victoria. My goal is to create a model that follows the principles of Open Science to serve as a platform for knowledge discovery by linking reference specimens with multiple sources of natural history information. I examine the challenges to overcome, and the benefits derived as I create a data management plan (DMP), choose appropriate metadata standards, annotate data, select suitable storage and repositories, and pilot a database. As part of this process, I review the components and use of comparative osteology collections and explore the ethics of Open Science. Best practices for data management and stewardship are explained, as well as barriers and benefits to the sharing and reuse of data. I describe how Linked Open Data (LOD) and ontologies can be used to enhance collection data and promote data sharing. Ideas to expand the collection and contribute to ecological and ethnobiological knowledge are shared. By applying these approaches, I aim to enhance collection research capacity and invite collaboration. These strategies are proposed to promote Open Science and the discoverability of the collection. Finally, I discuss potential educational applications and community-based projects.

Research Objectives: The University of Victoria Zooarchaeology Lab Comparative Collection

Comparative collections like the University of Victoria's Zooarchaeology Lab (UVicZL) are crucial for the accurate identification of skeletal animal remains from archaeological sites (Lyman 2010, 127). The UVicZL comparative collection is one of the largest and most representative reference collections for coastal species in the North Pacific, with skeletal remains from over 2,500 animals compiled over the last 40 years. The majority of bones identified from archaeological sites on the Northwest Coast have been processed in this lab and the collection is a well known resource (McKechnie and Moss 2016; McKechnie, Moss, and Crockford 2020).

My research initiates the creation of an online database encompassing the originally documented specimen attribute data (Table 1) annotated with Linked Open Data (Table 2) on the mammals, fish, birds, and other animals contained in the collection. To ensure sustainable collection management and data curation, my research develops methods to preserve collection specimen attribute data in perpetuity. I follow best practices of design, implementation, and usage to ensure data continuity and integrity. The aim is to maintain and assure a consistent and accurate dataset with secure storage and access through trustworthy online repositories. I use LOD to connect broader ecological and Indigenous knowledge, and associated research to collection data. Using Open Science methods, I aim to increase awareness of the collection, enhance discoverability and use of the specimens contained in the lab, and remove barriers to make data available to share for reuse and cross-interdisciplinary research and collaboration.

Table 1 Original Specimen Data

Specimen Attribute Data Categories
Year of collection
Catalogue number
Family name
Scientific name
Common name
Comments
Date of death
Sex
Weight
Length
Date donated
Donor
Where collected
Card catalogue image

Table 2 Linked Open Data

Linked Open Data (LOD)	**Example**
Biodiversity links	Global Biodiversity Information Facility
Indigenous ethnobiology links (future content)	Indigenous animal names ancestral geospatial links
Ontology links	The Digital Archaeological Record (tDAR) UBERON
Other zooarchaeological collections	Portland State, UC Santa Cruz, University of Alaska, Royal BC Museum
Associated research	

Research Questions

My research asks: 1) How can applying Open Science practices and data management standards expand the use, discoverability, and access to the UVicZL and facilitate and maintain collection specimen attribute data? 2) How can metadata standards and LOD resources broaden the comparative collection as a platform for knowledge, and promote sharing and reuse of collection data? And 3) how can the lab develop research capacity and facilitate connections with Indigenous communities, educators, students, and researchers to contribute to anthropological, ecological, and ethnobiological knowledge?

Zooarchaeological Comparative Collections

Components of a Comparative Collection

Zooarchaeological comparative collections contain intact skeletal remains of known animal species and are essential to the identification of fragmentary archaeological remains (Betts et al. 2011; Driver 2011; McKechnie, Moss, and Crockford 2020; Meighan et al. 1958; Niven et al. 2009). University-based zooarchaeological collections differ from museum-based natural history collections (NHCs) developed by zoologists who use whole or partial samples of preserved specimens to study taxonomy and variation including biogeographical, morphological, genetic, ecological, evolutionary, and behavioral traits (Cook et al. 2014, 726; Miller et al. 2020, 674, 675). NHCs are diverse and may include digital databases of specimen attributes, images, video, and sound clips (Miller et al. 2020, 675). Worldwide, NHCs are estimated to contain 2-4 billion specimens (Cook et al. 2014, 726). These biological specimens serve as regional, national, and global biodiversity records creating a permanent source of primary scientific data (Colella et al. 2020, 1). University-based collections can range from smaller natural-history style museums to skeletal reference collections representative of regional contemporary animals. Multiple specimens of a single species of small taxa and a few of large taxa are usually represented in skeletal reference collections (Lyman 2010, 129). These collections vary in composition, with some limited by size and species variation. Active use of comparative collections by zooarchaeologists is widespread, but seldom cited in research (Driver 2011, 23; Gifford-Gonzalez 2018, 159; Lyman 2010, 127). Robust research should include transparent documentation of identification methodology and citable references to the physical location of comparative collection facilities which this project seeks to address (Driver 2011, 28; Gifford-Gonzalez 2018, 163; Lyman 2017, 339).

Practical Reference Collection Use

The comparative method of identification employs a morphological comparison of archaeological specimens against modern ones. Using a comparative collection, fragmentary faunal specimens can be directly evaluated against both juvenile and adult males and females to determine genus and often species (Driver 2011, 23; Meighan et al. 1958, 7). To facilitate quick and accurate comparison of animal remains in a large comparative collection, specimens can be taxonomically or synoptically (by skeletal element) organized (Betts et al. 2011, 757; Driver

2011, 23; Gifford-Gonzalez 2018, 157; Meighan et al. 1958, 6–8). Future zooarchaeological identifications might suitably include the catalogue number of the comparative specimen with which the archaeological specimen is most closely identified. This could become a standard best disciplinary approach so future researchers can evaluate the similarity between an archaeological specimen and a comparative specimen.

Zooarchaeologists use standard biological binomial nomenclature to organize faunal remains for analysis into groups by pre-existing identification classification systems (Driver 2011, 20). This taxonomic system uses hierarchies and sub-classifications to minimize and refine ambiguous categorization. Analysts need to be aware of the frequent revisions to biological taxonomy as biologists continue to refine classification with genomic research, geometric morphometric analyses, and improved phylogenetic and developmental descriptors (Driver 2011, 22; Lyman 2017, 339). To assign a specimen to a taxonomic group, element identification is a prerequisite. As Lyman (2017) points out, identification to the species level is often difficult when osteological characteristics are not preserved archaeologically (Lyman 2017, 22). However, comparative collections are superior tools to two-dimensional printed guides and illustrated keys which can lack variation in age and sex, tactile key features, and are often limited in scope and regional selection. Analysts must be aware of diagnostic elemental morphometric properties and intra and inter-species variability to be confident in identification (Lyman 2017, 338, 339). In short, to compare osteological morphology characteristics and identify fragmentary remains, comparative collections should be recognized as indispensable facilities (Driver 2011, 24; Lyman 2010, 127).

Cultural and Ecological Knowledge

A comparative collection can be employed to refine zooarchaeological taxonomic classifications to support the reconstruction of ancient climates and environments and to understand past human behaviors (Lyman 2017, 316, 327, 337, 349). Past environmental conditions and biological information can provide insight for future environmental, conservation and resource management issues (Lyman 2017, 349). High-resolution reconstructions use animals identified to the sub-species level. At the sub-species level, animals have smaller geographic ranges, and similarly, narrower ecological tolerances (Lyman 2017, 337). A small geographic range of an individual animal is correlated with a narrow ecological tolerance or *niche*. Tolerances are defined by environmental gradients such as temperature, precipitation, food source biomass, and competitor and predator frequencies. Construction of a high resolution paleoenvironmental model is directly related to sub-species *niche* and depends on environmental gradients which control presence/absence and abundance, or tolerance, of an animal (Lyman 2017, 319). When broadly identified to the genus and sometimes species level, an animal with a wider tolerance range generally has a greater distribution geographically. Lower resolution paleoenvironmental reconstructions are derived from such specimens (Lyman 2017, 337). When analysts can identify species in an assemblage to a more refined level, comparative collections provide sharper taxonomic categorizations.

Comparative collections are a key resource for the greater understanding of biogeographical, historical ecology, and regional heritage information (LeFebvre et al. 2019, 2, 3). There is considerable potential to enrich collection data with the large selection of archaeological research articles, papers and theses, both published and unpublished, journals, books, and other literature particular to a region which, when recorded, are a complement to collection specimens (Damitio, Gillreath-Brown, and Tushingham 2018, 186). The regional heritage and historical environmental knowledge provided by such associated collection resources can provide a spatiotemporal perspective of human and animal ecological and cultural relationships (Ferris et al. 2018, 6; Meighan et al. 1958, 1).

Digital Collections

Data digitization of some North American zooarchaeology collections has improved discoverability. Digital collections are accessible to anyone with a high-speed internet connection. Innovations in technology allow new audiences to access collections and provide researchers with a novel avenue of data inquiry, assembly, and analysis (Miller et al. 2020, 674). These technological advances allow digitization of specimen images and metadata including taxonomic information, sex, age, measurements, related dates, and other descriptions from analogue notes, facilitating virtual accessibility (Miller et al. 2020, 675; Betts et al. 2011, 757). Notably, The Virtual Zooarchaeology of the Arctic Project (VZAP) uses 3D skeletal scans to bring together specimens from multiple collections to an online repository with a focus on mammals. The zooarchaeology collection held in the Florida Museum of Natural History also allows users to make online database inquiries but do not make a linked open data list of their collections.

Collections are further preserved and accessed through data aggregators and networks dedicated to taxonomic classification, data publishing, and digital curation. The Global Biodiversity Information Facility (GBIF), Integrated Digital Biocollections (iDigBio), Integrated Taxonomic Information System (ITIS), and Encyclopedia of Life (EoL) are among many other are taxonomic aggregators. To my knowledge, no zooarchaeological comparative collection is listed on these platforms. Databases can be further refined by animal type or geographic location such as the World Register of Marine Species (WoRMS), FishBase, and VertNet. Archaeological specific platforms such as the Digital Archaeological Record (tDAR) and Open Context curate and store digital data but do not itemize comparative collections which underlie some of the observations in archaeological datasets. As this shift in research accessibility improves, an increasing number of collections and repositories are seeking to apply Open Science practices as archaeologists embrace the changing landscape of digital data stewardship (Cook et al. 2014, 727; Marwick et al. 2017).

Open Science and Ethics

What is Open Science?

Open science is distinguished from the conventional paradigm as it works to advance the traditional norms of protection and accessibility. It is a movement to make scientific research and data more accessible and available. Expanding Open Science practices in archaeology requires the encouragement of data stewardship, analytical transparency, and public involvement (Marwick et al. 2017). Digital information can propel researchers to ask varied statistical and analytical research questions derived from larger and more refined data sources (Faniel et al. 2018). These practices enable timely results and publication of supplementary data, and enhance the discovery of online webpages, research facilities, analytical tools, journals, and repositories. Infrastructure created by such practices facilitates efficiency in research, engages the public, employs metrics to evaluate research impact, enables free access to knowledge, and elicits research collaboration and replication of results (Faniel et al. 2018; Marwick et al. 2017). Open Science makes archaeological research transparent, reusable and accessible, and also promotes ethical practices (Marwick et al. 2017). Because these practices encourage transparency and reproducibility, they enhance research credibility. Researchers can readily explain their data, analytical processes, and workflows.

Archaeological Ethics

Free access to information is not always appropriate, especially where Indigenous heritage is at stake. Because Indigenous sites are complex social and environmental landscapes ethical issues surround Indigenous archaeology with involvement of descendent communities a key aspect of site excavation, analysis, curation, and circulation (Forrest et al. 2020). Before specimens are identified with the use of a comparative collection, permissions to excavate and involvement of Indigenous communities is crucial. Information on the approximately two million archaeological and heritage sites in North America is not centralized in one location nor is it accessible by descendant communities (E. C. Kansa et al. 2021, 121). NHCs, published and unpublished papers and reports, and various government departments hold or contain this site data. Aggregators, such as the Digital Index of North America (DINAA) in the US and the Canadian Archaeological Radiocarbon Database (CARD) in Canada, link sites and resources. DINAA publishes generalized and highly redacted open access location data and CARD is working with Indigenous partners to ensure sensitive data is blurred or redacted (E. C. Kansa et al. 2021, 123). When compiling open data relating to Indigenous peoples, Marwick et al. (2017) caution, restrictions must be developed to protect individuals and cultural heritage from harm. Spatial locations should be obscured, and access to sensitive data restricted. As the collection develops resource links, these considerations will be integrated into planning of geospatial locations related to Indigenous heritage. The UVic Zooarchaeology Lab will make comparable material accessible but with respect to information generated from the collection and additionally create and continually update ethics-based policies for access rights that determines who can access the data and indicate its security status.

Methods

Data Management

Goals of Data Management and Collection Stewardship

Becoming familiar with Open Science data management standards and solutions was a challenge as a data curator of a zooarchaeological collection. With the trend toward open data, best practices are continually evolving. Open data implies free access to datasets which can be stored in trusted public repositories dedicated to data preservation and continuity (Pasquetto, Randles, and Borgman 2017, 3). Trusted online repositories, such as Open Science Framework (OSF) where I have initially stored digital metadata and collection information, ensure accessibility and data integrity via version control, permissions for access, licencing, and suggested citation. The version control feature allows collaboration with administrators and users and tracks and records changes to documents (Marwick et al. 2017). Privacy settings on projects and components can limit access to invitation only or public to facilitate open data sharing. Optimal storage and repository requirements for the comparative collection are developed in a data management plan (DMP).

To make collection data accessible, "good" data management follows FAIR Guiding Principles (Figure 1) (Curty et al. 2017, 3; Wilkinson et al. 2016, 1). Data should be "FAIR" or findable, accessible, interoperable, and reusable. I follow these data management guidance principles to improve data curation, maintenance, and access to data. These FAIR principles provide data management guidance by defining how data vocabularies, tools, resources and infrastructures should be characterized to facilitate data discoverability and reuse (Wilkinson et al. 2016, 4). As a guide to data management and stewardship, these practices work to address the long-term sustainability of the comparative collection to ensure it is continually maintained for the future.

Data should be **Findable**	F1. (meta)data are assigned a globally unique and persistent identifier (DOI) F2. data are described with rich metadata F3. metadata clearly and explicitly include the identifier of the data it describes F4. (meta)data are registered or indexed in a searchable resource
Data should be **Accessible**	A1. (meta)data are retrievable by their identifier using a standardized communications protocol A1.1 the protocol is open, free, and universally implementable A1.2 the protocol allows for an authentication and authorization procedure, where necessary A2. metadata are accessible, even when the data are no longer available
Data should be **Interoperable**	I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. I2. (meta)data use vocabularies that follow FAIR principles I3. (meta)data include qualified references to other (meta)data
Data should be **Reusable**	R1. meta(data) are richly described with a plurality of accurate and relevant attributes R1.1. (meta)data are released with a clear and accessible data usage license R1.2. (meta)data are associated with detailed provenance R1.3. (meta)data meet domain-relevant community standards

Figure 1 The Fair Guiding Principles (Wilkinson et al. 2016, 4)

<u>Data Sharing and Reuse</u>

FAIR Guiding Principles remove barriers to data reuse by facilitating improved data discovery through metadata standardization. The ease of data discovery and compatibility with other sources supports new downstream standalone research and the combining of primary datasets for meta-analyses (Curty et al. 2017, 4; Faniel et al. 2018, 106; S. W. Kansa 2015, 224; Miller et al. 2020, 678). Enhanced discoverability of comparable specimen data, such as that contained in UVicZL, can assist researchers with integration studies (Curty et al. 2017, 1; Huggett 2018, 94; S. W. Kansa 2015; McKechnie and Moss 2016; Pasquetto, Randles, and Borgman 2017, 4). For example, McKechnie and Moss' (2016) meta-analysis assembled a database of fish species from

222 archaeological sites from Alaska, British Columbia, and Washington identified using comparative collections. This database was used to determine Northwest coast Indigenous fishery practices, species ubiquity and relative abundance dating to the late Holocene. Fish such as Pacific salmon and herring were identified significant species over thousands of years of Indigenous fishery practices (McKechnie and Moss 2016, 483). This study has implications for ecological and cultural knowledge such as species and habitat information, and subsistence patterns and economic systems over long histories.

Data sharing and reuse are distinct behaviors linked to researcher attitudes, norms, and the type of science practiced. Researchers may either share or reuse data but only some do both (Curty et al. 2017, 15). Shared data must be accessible and in an interoperable format. Researchers can privately exchange data, make it available through their websites, institutional collections, or publish it with research or in "data" journals (Pasquetto, Randles, and Borgman 2017, 2). Much data can be lost through piecemeal sharing and incomplete data summaries provided in papers, which are often inadequate for additional analysis. Sharing of data is improved by following data management practices that encourage detailed documentation and improved stewardship (S. W. Kansa 2015, 225). However, the benefits of sharing data can only be realized by its reuse (Pasquetto, Randles, and Borgman 2017, 1). How data is perceived in terms of efficacy and trustworthiness are critical factors for data reuse (Curty et al. 2017, 15, 17). Curty et al. (2017) found that encouraging data reuse may be improved through data management training, recognition for exemplary data reuse, and changing entrenched norms of new data collection practices (Curty et al. 2017, 16, 17).

Crediting Datasets

Researchers improve reproducibility by documenting materials and methods to generate the same results. To improve data reuse, these open methods incorporate transparency in data collection, analysis, and visualization. Reproducible statistical and computational empirical analysis can be better facilitated with open-source software, version control, and method archiving in repositories such as OSF and figshare.com (Marwick et al. 2017). Often overlooked is documenting methods of faunal identification (Gifford-Gonzalez 2018, 161–63). Gifford-Gonzalez (2018) recommends that faunal reports should cite reference collections, their location, and individual reference specimens to improve research replicability. Citation of datasets acknowledges the role comparative collections play in the identification process. Citation allows collection data managers to trace collection use and research related to their collection, and create awareness of the availability of the collection for potential reuse by future researchers (Miller et al. 2020, 677; Pasquetto, Randles, and Borgman 2017, 3). Standards, policies, and guidelines for formal citation of datasets encourage citation practice. Collections and institutions can be designated with digital object identifiers (DOIs) available through database publishers like GBIF or through Open Researcher and Contribution ID (ORCID) (Curty et al. 2017, 17; Miller et al. 2020, 677).

Data Management Standards and Solutions

Advocates of open archaeology propose data management practices that provide answers to the questions in Table 3 (Marwick et al. 2017; Wilkinson et al. 2016, 2,3).

Table 3 Data Management Questions

Question
Where are datasets published and/or stored?

Are repositories stable, trustworthy, and free (open data)?

How can these datasets be discovered?

What search tools are required?

What format is metadata in, is it searchable, and how?

Can data be downloaded, and in what format(s)?

Can dataset formats be integrated with other datasets, and is integration automatic?

Are data methods transparent (open methods)?

Is sensitive data protected (protected geolocation of heritage sites)?

What permissions are required to use data, and what are data licensing conditions?

How can data be cited (assigned DOI)?

Many of the uncertainties posed by these questions are answered by a Data Management Plan (DMP). Basic collection data is contained in a DMP such as data types, file names and formats, and project management tools. Project documentation includes background information, guides, policies, physical location of specimens, and metadata standards. A DMP should specify storage requirements, data backup routines, hardware, and user access and permissions. Planning for the long-term sustainability of the project is a critical piece of a DMP as providers may not be permanent and data preservation standards evolve. Sharing, reuse, and legal compliance policies should also be stated, as well as administrator and user responsibilities and resources. Many granting agencies and public and private funders are increasingly requiring DMPs in order to disperse funds (Colella et al. 2020, 1; Faniel et al. 2018, 106; S. W. Kansa 2015, 224). Tools for DMP creation include the Canadian institution specific Portage's DMP Assistant, US-based

DMPTool and others (Faniel et al. 2018, 107). The UVicZL DMP was created using Portage's DMP Assistant with guidance provided by a University of Victoria Library Digital Scholarship Commons workshop. Appendix A contains a preliminary DMP. A DMP is a living document that requires routine review and revision.

To embrace Open Science and data reuse principles, data must be securely kept and remain accessible for a minimum data lifecycle of ten years. UVicZL comparative collection is a dynamic resource with new specimens added and the systematics of previously entered specimens updated when new taxonomic groupings are announced. As such, a long-term plan to extend this timeframe is critical. Local hardware storage systems require replacement after five to seven years over a ten-year data lifecycle. Local and cloud back-up of live and archived data is planned to be distributed over several independent sources and follow a grandfather-father-son rotation scheme (Colella et al. 2020, 7). These ideally will include secure and trustworthy departmental and institutional storage and backup and, federal, academic and discipline-specific storage and repository providers such as OSF, UVicSpace Research and Learning Repository, The Digital Archaeological Record (tDAR), Canadensys, Dataverse and Sync. A potential back-up scheme could be structured where grandfather is a full yearly backup to offsite storage (repositories), father is a monthly back-up to designated cloud storage, and son is a weekly incremental or differential back-up to cloud and local storage.

Data Curation

Data preparation begins once the preliminary DMP is complete. Data preparation and "cleaning" ensures fields and terms are consistent and compatible for annotation, future aggregation, and to facilitate the shared use of open datasets. It is advisable to follow the workflow framework developed by biodiversity and archaeological repositories to publish data (LeFebvre et al. 2019, 4). For sharing data, standardized metadata format methods are recorded in a metadata document (S. W. Kansa 2015, 227; LeFebvre et al. 2019, 2; Nelson Gil and Ellis Shari 2019, 3). As a first step in the publishing process, metadata should be standardized, with fields are aligned to the Darwin Core (DwC) standard which follows FAIR principles. The next step is to organize the dataset "occurrences," where collection specimens are sorted with the use of tabular fields, such as catalogue number or scientific name. Consistent descriptors then assigned to each specimen record should be meaningful to users. As the complexity of data will grow over time, standardized metadata is a valuable data management tool both in terms of time spent on data searches and compatibility.

Metadata Standards

Metadata describes and summarizes data attributes within a dataset. Darwin Core is a "living" standard set of terms developed for natural history collections to share biodiversity data (Wieczorek et al. 2012, 1, 2, 7). This common language has clearly defined categories with subsets (https://dwc.tdwg.org/terms/) that represent descriptive organism data (Figure 2) (Wieczorek et al. 2012, 2, 3). Darwin Core allows interoperability of biodiversity data so that it can be shared and integrated to facilitate scientific research. However, these data may need to be augmented with further specific archaeological categories (Colella et al. 2020, 6; Wieczorek et al. 2012, 7).

Record-level Terms	Dublin Core terms, institutions, collections, nature of data record	Simple Darwin Core (flat)
Occurrence	evidence of species in nature, observers, behavior, associated media, references.	
Event	sampling protocols and methods, date, time, field notes	
Location	geography, locality descriptions, spatial data	
Identification	linkage between Taxon and Occurrence	
Taxon	scientific names, vernacular names, names usages, taxon concepts, and the relationships between them	
GeologicalContext	geologic time, chrono-stratigraphy, biostratigraphy, lithostratigraphy	
ResourceRelationship	explicit relationships between identified resources (e.g., one organism to another, taxon to location, etc.)	Generic Darwin Core (relational)
MeasurementOrFact	measurements, facts, characteristics, assertions, references	

Figure 2 Darwin Core Categories (Wieczorek et al. 2012, 3)

As part of the data curation process, I standardized metadata formats and documented metadata methodology to avoid analytical misinterpretations (S. W. Kansa 2015). The specimen record fields are reconciled with DwC standards and a globally unique identifier (GUID) provided for each record (Colella et al. 2020, 7; Miller et al. 2020, 678). Collection tabular data is then ascribed to a metadata coding system for catalogue number, scientific and common names, dates, sex, life stage, measurements, location, and specimen preparation. In practice this requires modifying some established conventions within the lab terminology, but this shift can be accommodated in translating and equating concepts. To resolve archaeological-specific zooarchaeological information deficiencies, such as facts not described by DwC standards, custom formats can be added under the "dynamicProperties" record-level field (Colella et al. 2020, 6; LeFebvre et al. 2019, 6, 7). To standardize an ontology, such as UBERON, descriptions can be linked to a Darwin Core occurrence field, "preparations" (S. W. Kansa 2015; LeFebvre et al. 2019, 11). Annotation of datasets with Linked Open Data (LOD) "URIs" (Uniform Resource Identifiers) links specimens to terminology resources and shared concepts.

Linked Open Data and Ontologies

Linked Open Data (LOD) connects resources and ontologies through networks of publishing and access platforms (LeFebvre et al. 2019, 2). When applied to zooarchaeological data, LOD links taxonomic specimens with stable web identifiers (URIs) to resources and shared concepts such as contemporary and deep time biodiversity data (S. W. Kansa 2015, 227). For example, the Global Biodiversity Information Facility (GBIF) uses ontology standards such as Darwin Core to cross reference across a searchable network of occurrences, species, datasets, and publishers. The GBIF initiative was designed to freely and universally share biodiversity information globally to foster environmental and human well-being solutions. GBIF provides a network platform for data publishers around the world to publish and access biodiversity data to advance research. Links, such as GBIF, to specimen data connect researchers, Indigenous communities, and the public with further information about contemporary distribution of species. Inclusion of LOD resources is a key, but rarely applied, aspect of modern zooarchaeological data management (S. W. Kansa 2015). In this project I link comparative specimen data with GBIF to provide collection users with this cross reference to biodiversity data that includes regional occurrences. This is quite helpful in establishing whether a species is documented to occur in sub-regions and island settings.

Biodiversity LOD can provide vital data for considering the many different common and scientific names that have been used for the dame species over time, including the changes to scientific names and the local words for various important species (e.g., for salmon, chinook, king, Tyee). LOD can also reconcile and connect different taxonomic identifications that may have been used by different users of the comparative collections. This framework for common understanding allows researchers to better address and expand perspective on zooarchaeological research concerning species extinction and extirpation (LeFebvre et al. 2019, 3; Lyman 2017), animal domestication and exploitation (Haggan et al. 2004; McKechnie and Moss 2016; McKechnie, Moss, and Crockford 2020), landscape modification (Haggan et al. 2004; Letham, Martindale, and Ames 2020) and responses to climate change (LeFebvre et al. 2019, 3; Lyman 2017, 349). For example, McKechnie et al. (2020) complied a database of morphologically identified mammal specimens from 330 Northwest Coast sites over 55 years to study dogs domesticated by Indigenous societies. The terminology and conventions of identification used by different researchers resulted in varying levels of species, genus, and family-level identifications, but when considered as a whole, the study concluded that domesticated dogs were the single most abundant canid throughout the mid and late Holocene, indicating considerable animal husbandry practices (McKechnie, Moss, and Crockford 2020, 4, 10).

Another common use of LOD is to share ontologies from external sources to facilitate reanalysis, integration of datasets, and ask new questions of datasets. An ontology is a relationship system which describes rules about the creation of categories and definitions of attributes and constraints; it is essentially a map of concepts (Spielmann and Kintigh 2011, 22). While taxonomies classify hierarchically from broad to narrow categories, ontologies specify. Examples of ontologies are faunal sex and age (Figure 3and Figure 4). Heterogenous terms exist within an ontology where sex or age variants can provide alternate canonical names (LeFebvre et al. 2019, 7). For example, “subadult” may be recorded by one analyst as “immature,” another “juvenile,” while another may have used “fetal or neonate.” If analysts can agree on a shared ontology, for example the variable “subadult,” such that immature, juvenile, and fetal/neonate are subcategories of subadult, data integration tools allow an analyst to map each individual translated code in their databases to ontology values used by other analysts. Differently recorded variables in many databases can then be integrated as original categories are mapped to shared values. Ontologies are taxonomy tools used to uniquely identify a global key so other external data systems can all refer to the same concept. The Digital Archaeological Record (tDAR) provides technical infrastructure for dataset preservation and access, and supports data integration through the use of faunal ontologies (Spielmann and Kintigh 2011). Collection data management benefits from archeo-informatic tools such as the ones developed by tDAR. Another example of a LOD ontology is UBERON, a comparative anatomy ontology which can be used to provide descriptions of skeletal elements (S. W. Kansa 2015). The metadata field "preparations" could be used to link anatomy of a skull, for example, to a specimen record.

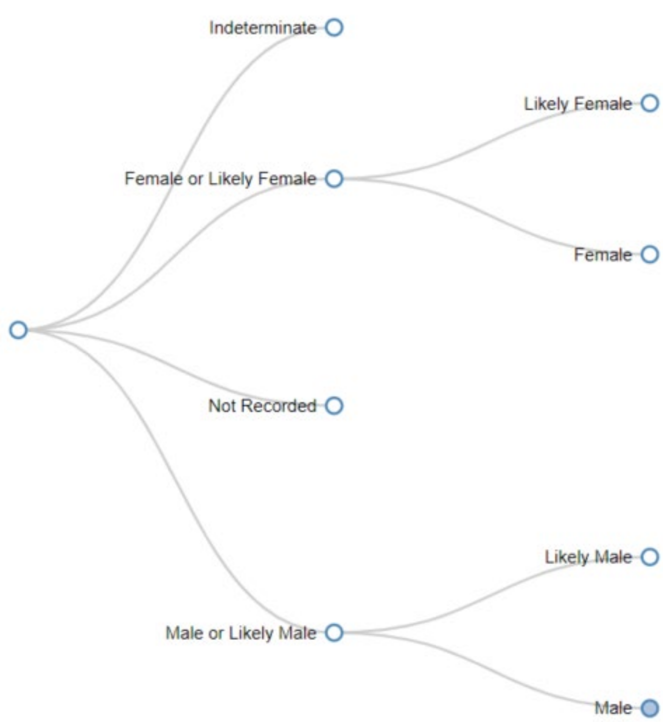


Figure 3 Faunal Sex Ontology (core.tdar.org/ontology/3026/faunal-sex-ontology)

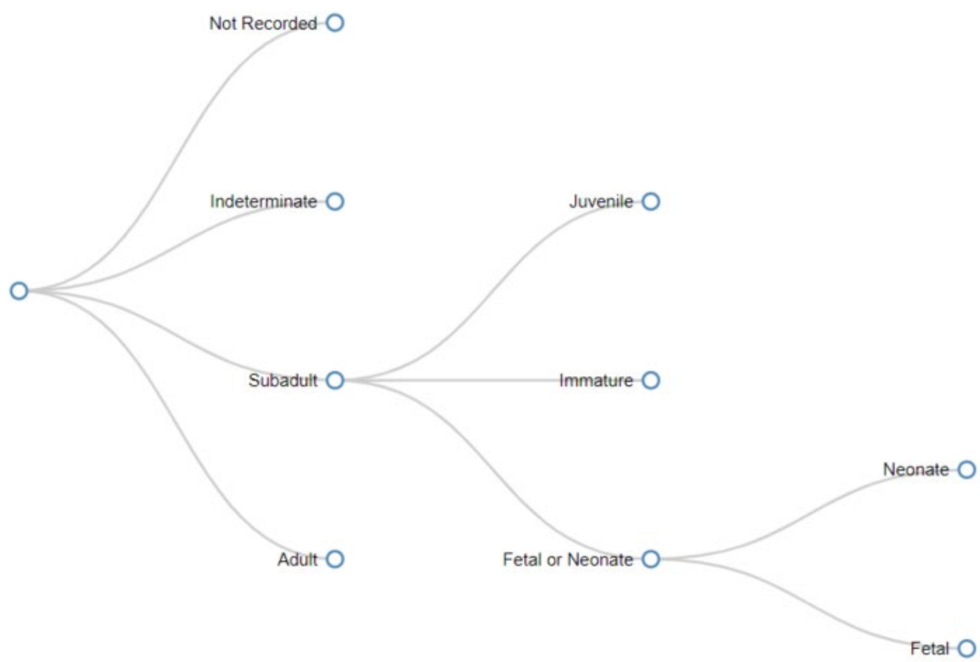


Figure 4 Faunal Sex Ontology (core.tdar.org/ontology/3030/faunal-age-ontology)

Results

Collection Research Capacity

Platform for Expansion

The UVicZL will improve the capacity to digitally expand ecological and ethnobiological knowledge. Table 4 includes some ideas for inclusion.

Table 4 Linked Open Data Resources

Resource Type	**Example**
Biodiversity Linked Open Data	Taxonomies, species occurrences, datasets, publishers
Skeletal Element Ontologies	UBERON Comparative Anatomy Ontology
Taxonomic and Genomic Information	NCBI (National Center for Biotechnology Information) and Open Data links
Northwest Coastal geospatial data	Oceanographic, waterway, and ancestral linguistic boundaries

Ethnobiological information	Regionally specific Indigenous animal names and research linked to identification specimens
3D images of selected elements	Sketchfab images

Collection Access and Use

The UVicZL comparative database will be transformed into a web-based searchable data management tool designed for use by students, researchers, educators, and community members. An online manual will be created to consolidate information on skeletal anatomy, specimen preparation, faunal ontologies, metadata formats and user access. Database queries require a user-friendly interface to generate beneficial information. Integrating graphical user interfaces into an online platform will improve user access and querying methods. Work study and zooarchaeology students can assist with database updating. Specimen data will be digitized at the time of preparation/acquisition for timely availability to users and researchers. As the collection database will be stored over several platforms, this necessitates the capability to synchronize data. Protocols for simultaneous addition and updating will be developed to ensure data integrity (Miller et al. 2020, 678, 679, 682).

Sea Otter Specimen Case Study

To demonstrate how the database can be used I developed a case study based on a collection specimen, a sea otter (Figure C 2), whose skeletal remains are kept in the UVicZL comparative collection (Appendix C). Linked Open Data on the taxonomy of this specimen and over 11,000 documented occurrences of sea otters worldwide can be accessed via a link to the Global Biodiversity Information Facility (GBIF) (Figure C 8). Research studies (Figure C 5 and Figure C 6) relevant to the specimen are available as links which researchers can access. Other information, such as the regionally specific Indigenous word (Figure C 3) for the specimen is available through a link to firstvoices.com and relevant linguistic and territory boundaries (Figure C 9) related to the origin of the specimen are included in the record from the native-land.ca website. Further taxonomic and genomic information on the specimen is available through a link to the National Center for Biotechnology Information (NCBI) (Figure C 7). This specimen is unique as it will have a link to Sketchfab 3D imaging of the skull from the UVic Library (Figure C 4). In the future, more collection specimen elements can be scanned using scanners available through UVic's Department of Anthropology.

Discussion and Concluding Remarks

Education and Collaborative Strategies

Open Science practices that enhance facility discoverability and database use can facilitate zooarchaeological education and collaborative community-based projects. Strategies can be developed that follow Open Science principles within archaeology by expanding collection access and use.

Education

Specimen metadata and related images are excellent aids to the education and training of school-age children as well as university students (Miller et al. 2020, 683). One such strategy which engages educators and students is to create relational data models, like the collection database, which are linked in a zooarchaeology module or syllabus. Conceptualizing relationships between data is key to interpretation and the fundamentals of collaborative research. For example, Jones and Hurley (2011) developed a zooarchaeology course focused on basic bone identification skills, understanding the link between research and larger archeological questions, and building skills to organize zooarchaeological data using relational databases. When zooarchaeology students learn relational data management skills they are more capable of contributing to archaeology cultural heritage management (Jones and Hurley 2011). Because teaching students to design and create databases is a specialized technical skill this may be difficult to fully incorporate into a zooarchaeology class. “Relational database thinking” with a predesigned database modelled after the collection would be more suitable for an introductory class. This format allows students to begin conceptualizing how to answer research questions with better data, build awareness of digital resources and practices, and realize the potential of collaborative data sharing (Jones and Hurley 2011). As these models are digital, they can be accessed by the public and teachers worldwide who would not otherwise have physical access to collections (Miller et al. 2020, 683). Future efforts will explore developing modules of integrating collection data into zooarchaeology and outreach education using digital dashboards, lab exercises, collaborative apps, and 3D digitization of selected skeletal elements.

Community-Based Research

Traditional Indigenous wisdom and environmental research must converge to develop a greater understanding of the cultural significance of animals (McKechnie and Moss 2016; Menzies 2006; Anderson et al. 2011). Links to ethnobiological knowledge are vital to the future sustainability of the collection and promoting collaboration with Indigenous communities. Currently, research in Northwest Coast Indigenous communities is supported in their territories with archaeological faunal remains processed in the Lab. UVicZL welcomes further contact from Indigenous scholars who have interest in where the animals occur in the archaeological record. My research, focused on this unique comparative collection, has the potential to better connect Indigenous communities and heritage specialists with the historical ecology, the history of animal use, and vital zooarchaeological data from across the North Pacific.

Promoting Open Science and Enhancing Discoverability

Using these Open Science practices and LOD annotation will facilitate connections between these collection resources, such as archaeological heritage and Indigenous knowledge, and beliefs and practices of communities across space and time (Ferris et al. 2018, 6). The continued study of coastal zooarchaeological specimens from sites across British Columbia and the North Pacific provides deep time perspective on human-environmental interactions over the past 14,000 years of known occupation (Mackie et al. 2018; McKechnie and Moss 2016; McLaren et al. 2020). Enhanced discoverability of the collection will connect researchers, educators, Indigenous scholars, students, and members of the wider community within archaeology, historical ecology, natural resource management, conservation, and zooarchaeology from across the Northwest Coast. This inspires novel research and fosters innovative cross-pollination across disciplines (Marwick et al. 2017; Nelson Gil and Ellis Shari 2019; Pasquetto, Randles, and Borgman 2017; Wilkinson et al. 2016). The UVic Zooarchaeology Lab can be used to reconstruct animal husbandry practices (McKechnie, Moss, and Crockford 2020), biodiversity (Cook et al. 2014; LeFebvre et al. 2019; Nelson Gil and Ellis Shari 2019), environmental conditions (Damitio, Gillreath-Brown, and Tushingham 2018; Mackie et al. 2018), economic systems and subsistence patterns (McKechnie and Moss 2016; Meighan et al. 1958). By developing the collection to be compatible with technological advances and as an archaeological resource for biodiversity research, the UVic collection will be accessible to address questions of environmental human-animal interactions over long histories (LeFebvre et al. 2019, 1; Miller et al. 2020, 679). Improved access to this valuable reference collection will inspire scholarly inquiry, stimulate interdisciplinary collaboration, and embrace Indigenous knowledge.

References

Anderson, E. N., Deborah Pearsall, Eugene Hunn, and Nancy Turner, eds. 2011. *Ethnobiology*. Wiley Blackwell.

Betts, Matthew W., Herbert D. G. Maschner, Corey D. Schou, Robert Schlader, Jonathan Holmes, Nicholas Clement, and Michael Smuin. 2011. "Virtual Zooarchaeology: Building a Web-Based Reference Collection of Northern Vertebrates for Archaeofaunal Research and Education." *Journal of Archaeological Science* 38 (4): 755.e1-755.e9. https://doi.org/10.1016/j.jas.2010.06.021.

Colella, Jocelyn P, Ryan B Stephens, Mariel L Campbell, Brooks A Kohli, Danielle J Parsons, and Bryan S Mclean. 2020. "The Open-Specimen Movement." *BioScience*, no. biaa146 (December). https://doi.org/10.1093/biosci/biaa146.

Cook, Joseph A., Scott V. Edwards, Eileen A. Lacey, Robert P. Guralnick, Pamela S. Soltis, Douglas E. Soltis, Corey K. Welch, et al. 2014. "Natural History Collections as Emerging Resources for Innovative Education." *BioScience* 64 (8): 725–34. https://doi.org/10.1093/biosci/biu096.

Curty, Renata Gonçalves, Kevin Crowston, Alison Specht, Bruce W. Grant, and Elizabeth D. Dalton. 2017. "Attitudes and Norms Affecting Scientists' Data Reuse." *PLoS ONE* 12 (12). https://doi.org/10.1371/journal.pone.0189288.

Damitio, William J., Andrew Gillreath-Brown, and Shannon Tushingham. 2018. "Seeing the Forest for the Trees: A Spatial Database to Enhance Potential of Legacy Collections at the Washington State University Museum of Anthropology." *Journal of Northwest Anthropology: Volume 52 Number 2* 52: 183.

Driver, Jonathan C. 2011. "Identification, Classification and Zooarchaeology." *Ethnobiology Letters* 2 (0): 19–39. https://doi.org/10.14237/ebl.2.2011.32.

Faniel, Ixchel M., Anne Austin, Eric Kansa, Sarah Whitcher Kansa, Phoebe France, Jennifer Jacobs, Ran Boytner, and Elizabeth Yakel. 2018. "Beyond the Archive: Bridging Data Creation and Reuse in Archaeology." *Advances in Archaeological Practice* 6 (2): 105–16. https://doi.org/10.1017/aap.2018.2.

Ferris, Neal, Aubrey Cannon, John Welch, and John Welch. 2018. "Objects as Stepping Stones: Sustainable Archaeology." *Canadian Journal of Archaeology/Journal Canadien D'Archéologie* 42 (1): 004–012.

Forrest, Chrystal L., Ronald F. Williamson, Susan Pfeiffer, and Louis Lesage. 2020. *The Joy of the Souls : The Return of the Huron-Wendat Ancestors*. *Working with and for Ancestors*. Routledge. https://doi.org/10.4324/9780367809317-16.

Gifford-Gonzalez, Diane. 2018. "The Emergence of Zooarchaeology." In *An Introduction to Zooarchaeology*, edited by Diane Gifford-Gonzalez, 19–50. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-65682-3_2.

Haggan, Nigel, Nancy Turner, Jennifer Carpenter, James T. Jones, Quentin Mackie, and Charles Menzies. 2004. "12,000+ Years of Change: Linking Traditional and Modern Ecosystem Science in the Pacific Northwest." In *Society for Ecological Restoration International*, 1–29. Victoria, B.C.

Huggett, Jeremy. 2018. "Reuse Remix Recycle: Repurposing Archaeological Digital Data." *Advances in Archaeological Practice* 6 (2): 93–104. https://doi.org/10.1017/aap.2018.1.

Jones, Emily Lena, and David A. Hurley. 2011. "Relational Databases and Zooarchaeology Education." *The SAA Archaeological Record* 11 (1): 19–21.

Kansa, Eric C., Sarah Whitcher Kansa, David G. Anderson, Joshua J. Wells, Kelsey Noack Myers, and Stephen Yerka. 2021. "Access to Government Information and Inclusive Stewardship of North America's Archaeological Heritage." In *Digital Mapping and Indigenous America*, edited by Janet Berry Hess. Routledge.

Kansa, Sarah Whitcher. 2015. "Using Linked Open Data to Improve Data Reuse in Zooarchaeology." *Ethnobiology Letters* 6 (2): 224–31. https://doi.org/10.14237/ebl.6.2.2015.467.

LeFebvre, Michelle J., Laura Brenskelle, John Wieczorek, Sarah Whitcher Kansa, Eric C. Kansa, Neill J. Wallis, Jessica N. King, Kitty F. Emery, and Robert Guralnick. 2019. "ZooArchNet: Connecting Zooarchaeological Specimens to the Biodiversity and Archaeology Data Networks." *PLOS ONE* 14 (4): e0215369. https://doi.org/10.1371/journal.pone.0215369.

Letham, Bryn, Andrew Martindale, and Kenneth M. Ames. 2020. "Endowment, Investment, and the Transforming Coast: Long-Term Human-Environment Interactions and Territorial Proprietorship in the Prince Rupert Harbour, Canada." *Journal of Anthropological Archaeology* 59 (September): 101179. https://doi.org/10.1016/j.jaa.2020.101179.

Lyman, R. Lee. 2010. "Paleozoology's Dependence on Natural History Collections." *Journal of Ethnobiology* 30 (1): 126–36. https://doi.org/10.2993/0278-0771-30.1.126.

———. 2017. "Paleoenvironmental Reconstruction from Faunal Remains: Ecological Basics and Analytical Assumptions." *Journal of Archaeological Research* 25 (4): 315–71. https://doi.org/10.1007/s10814-017-9102-6.

Mackie, Quentin, Daryl Fedje, Duncan McLaren, and Duncan McLaren. 2018. "Archaeology and Sea Level Change on the British Columbia Coast." *Canadian Journal of Archaeology/Journal Canadien D'Archéologie* 42 (1): 074–091.

Marwick, Ben, Jade d'Alpoim Guedes, C. Michael Barton, Lynsey A. Bates, Michael Baxter, Andrew Bevan, Elizabeth A. Bollwerk, et al. 2017. "Open Science in Archaeology." *SAA Archaeological Record* 17 (4): 8–14.

McKechnie, Iain, and Madonna L. Moss. 2016. "Meta-Analysis in Zooarchaeology Expands Perspectives on Indigenous Fisheries of the Northwest Coast of North America." *Journal of Archaeological Science: Reports* 8 (August): 470–85. https://doi.org/10.1016/j.jasrep.2016.04.006.

McKechnie, Iain, Madonna L. Moss, and Susan J. Crockford. 2020. "Domestic Dogs and Wild Canids on the Northwest Coast of North America: Animal Husbandry in a Region without Agriculture?" *Journal of Anthropological Archaeology* 60 (December): 101209. https://doi.org/10.1016/j.jaa.2020.101209.

McKechnie, Iain, and Rebecca J. Wigen. 2011. *7. Toward a Historical Ecology of Pinniped and Sea Otter Hunting Traditions on the Coast of Southern British Columbia*. *Human Impacts on Seals, Sea Lions, and Sea Otters*. University of California Press. http://www.degruyter.com/document/doi/10.1525/9780520948976-008/html.

McLaren, Duncan, Daryl Fedje, Quentin Mackie, Loren G. Davis, Jon Erlandson, Alisha Gauvreau, and Colton Vogelaar. 2020. "Late Pleistocene Archaeological Discovery Models on the Pacific Coast of North America." *PaleoAmerica* 6 (1): 43–63. https://doi.org/10.1080/20555563.2019.1670512.

Meighan, C. W., D. M. Pendergast, B. K. Swartz, and M. D. Wissler. 1958. "Ecological Interpretation in Archaeology: Part I." *American Antiquity* 24 (1): 1–23. https://doi.org/10.2307/276737.

Menzies, Charles R. 2006. *Traditional Ecological Knowledge and Natural Resource Management*. U of Nebraska Press.

Miller, Sara E, Lisa N Barrow, Sean M Ehlman, Jessica A Goodheart, Stephen E Greiman, Holly L Lutz, Tracy M Misiewicz, et al. 2020. “Building Natural History Collections for the Twenty-First Century and Beyond.” *BioScience* 70 (8): 674–87. https://doi.org/10.1093/biosci/biaa069.

Nelson Gil and Ellis Shari. 2019. “The History and Impact of Digitization and Digital Data Mobilization on Biodiversity Research.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 374 (1763): 20170391. https://doi.org/10.1098/rstb.2017.0391.

Niven, Laura, Teresa E. Steele, Hannes Finke, Tim Gernat, and Jean-Jacques Hublin. 2009. “Virtual Skeletons: Using a Structured Light Scanner to Create a 3D Faunal Comparative Collection.” *Journal of Archaeological Science* 36 (9): 2018–23. https://doi.org/10.1016/j.jas.2009.05.021.

Pasquetto, Irene V., Bernadette M. Randles, and Christine L. Borgman. 2017. “On the Reuse of Scientific Data.” *Data Science Journal* 16 (8): 1–9. https://doi.org/: http://doi.org/10.5334/dsj-2017-008.

Spielmann, Katherine, and Keith Kintigh. 2011. “The Digital Archaeological Record: The Potentials of Archaeozological Data Integration through TDAR.” *SAA Archaeological Record* 11 (1): 22–25.

Szpak, Paul, Trevor J. Orchard, Iain McKechnie, and Darren R. Gröcke. 2012. “Historical Ecology of Late Holocene Sea Otters (Enhydra Lutris) from Northern British Columbia: Isotopic and Zooarchaeological Perspectives.” *Journal of Archaeological Science* 39 (5): 1553–71. https://doi.org/10.1016/j.jas.2011.12.006.

Wieczorek, John, David Bloom, Robert Guralnick, Stan Blum, Markus Döring, Renato Giovanni, Tim Robertson, and David Vieglais. 2012. “Darwin Core: An Evolving Community-Developed Biodiversity Data Standard.” *PLOS ONE* 7 (1): e29715. https://doi.org/10.1371/journal.pone.0029715.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3 (March): 160018. https://doi.org/10.1038/sdata.2016.18.

Appendices

Appendix A. Data Management Plan

Project Name: Zooarchaeology Lab Comparative Collection

Principal Investigator / Researcher: Kathryn McKenzie

Institution: University of Victoria

Data Collection

Data types to collect, create, link to, acquire and/or record:

Zooarchaeology collection specimen tabular data:

year of collection

catalogue number

family name

scientific name

common name

comments

date of death

length

sex

weight

length

date donated

donor

where collected

images of original card catalogue for each specimen

Linked open data (LOD):

Biodiversity links (GBIF)

Indigenous ethnobiology

ontology links

3D data on selected specimens

File formats data will be collected in (may allow for data re-use, sharing and long-term access to the data, can be converted to a more accessible format after data collection):

Excel:

faunalcatras.xls

faunalcatb.xls

faunalcatf.xls

faunalcatm.xls

Access:

fauncatm.accdb

Adobe:

card-catalogue individual card (yy###).pdf

Word:

database metadata formatting.docx

card-catalogue scanning procedures.docx

procedure for adding specimen to collection.docx

procedure for disposing of hazardous waste.docx

procedure for feeding beetles.docx

procedure for preparing beetle food.docx

Zooarchaeology Lab Policies.docx

3D scans:

Sketchfab specimen files

Lab project management tools to ensure proper and consistent conventions and procedures used to structure, name, and version-control files:

Open Science Framework (OSF):

revision numbering (version control)

date of deposit

contributor

Documentation and Metadata

To ensure the data is read and interpreted correctly in the future, and that documentation is created or captured consistently an online manual will be created to consolidate the following documents:

History of curation of the collection

Specimen preparation

Skeletal anatomy guide

Physical location of specimens and storage methodology

Scientific naming conventions – taxonomies

Faunal ontologies

Metadata formatting

User access and permissions

Data entry procedures

Lab policy documents

Metadata standards:

Darwin Core

Additional formats specific to zooarchaeology

Storage, Backup, and Sensitive Data

Anticipated project storage space requirements and length of storage time:

5GB (approximate)

Perpetual storage

Data storage and back up methods:

Physical

PC hardrive in Zooarchaeology Lab

Zooarchaeology Lab Administrator Laptops

UVic Anthropology Department Server (with backup and cloud storage)

USB drives

Virtual

Cloud storage

Sync.ca (secure Canadian data storage)

OSF (open science framework)

to modify and contribute data research team members and other project collaborators will access data through:

Anthropology Department hardware

Online platform user interface

Permissions to ensure data is securely managed and accessible only to approved members of the project:

Administrator

Read and Write

Read Only

Guest

Preservation

Data repositories considered for long-term storage:

OSF

UVicSpace

tDAR

Canadensys (Canadian satellite storage)

Dataverse

Preservation friendly, proprietary, and non-proprietary file formats to ensure file integrity, and inclusion of supporting documentation:

.csv

.html

.pdf

.xlsx

.docx

.accdb

.jpeg, .gif, .png

.mtl, .obj, .gltf, .usdz

Sharing, Reuse, and Legal Compliance

Types of data to be shared and format:

raw versions

processed versions

analyzed versions

final versions

To manage legal, ethical, and intellectual property issues end-user license:

Creative Commons (CC)

Methods to improve access to researcher, educators, students, and the community and strategies to address secondary uses of data:

data registries

repositories

interdisciplinary collaboration

citation/DOI assignment

online platform

media releases

Responsibilities and Resources

Persons responsible for managing project data and the major data management tasks for which they will be responsible are listed below. If substantive changes happen in the personnel overseeing the lab's data, including a change of Principal or responsibilities for managing data activities transitions will be managed with documentation of policy for collection management:

Dr. Iain McKechnie – Zooarchaeology Lab Director and Supervisor

Dr. Stephanie Calce – Zooarchaeology Lab Manager

Dr. Yin-Man Lam – Zooarchaeology Instructor

Kathryn McKenzie – Zooarchaeology Lab Administrator

Work study students – Specimen preparation and processing, data entry, specimen curation

Resources required to implement your data management plan, data management budget TBD:

Lab PC

Lab laptops (MacBook, HP)

Adobe, MS Office licenses

Printer/scanner

Appendix B. Darwin Core Metadata Terms, Descriptions, and Examples

Term	Description	Example
type	The nature or genre of the resource	PhysicalObject
language	A language of the resource	en
license	A legal document giving official permission to do something with the resource	http://creativecommons.org/licenses/by/4.0/legalcode
rightsHolder	A person or organization owning or managing rights over the resource	The Anthropology Department of the University of Victoria
accessRights	Information about who can access the resource or an indication of its security status	
bibliographicCitation	A bibliographic reference for the resource as a statement indicating how this record should be cited (attributed) when used	
collectionID	For physical specimens, the recommended best practice is to use an identifier from a collections registry such as the Global Registry of Biodiversity Repositories (http://grbio.org/)	
datasetID	An identifier for the collection or dataset from which the record was derived	
institutionCode	The name (or acronym) in use by the institution having custody of the object(s) or information referred to in the record	UVIC
collectionCode	The name, acronym, coden, or initialism identifying the collection or data set from which the record was derived	UVICZL
datasetName	The name identifying the data set from which the record was derived	UVICZL Mammals
basisOfRecord	Recommended best practice is to use the standard label of one of the Darwin Core classes	PreservedSpecimen

Term	Description	Example
dynamicProperties	A list of additional measurements, facts, characteristics, or assertions about the record. Meant to provide a mechanism for structured content. Recommended best practice is to use a key:value encoding schema for a data interchange format such as JSON	LOD or Ontologies specific to zooarchaeology
occurrenceID	An identifier for the Occurrence (as opposed to a particular digital record of the occurrence). In the absence of a persistent global unique identifier, construct one from a combination of identifiers in the record that will most closely make the occurrenceID globally unique	URL ID
catalogNumber	An identifier (preferably unique) for the record within the data set or collection	08.63
recordedBy	A list (concatenated and separated) of names of people, groups, or organizations responsible for recording the original Occurrence. The primary collector or observer, especially one who applies a personal identifier (recordNumber), should be listed first	Becky Wigen
sex	The sex of the biological individual(s) represented in the Occurrence	M, F, M?, F?
lifeStage	The age class or life stage of the biological individual(s) at the time the Occurrence was recorded	juvenile, adult
preparations	A list (concatenated and separated) of preparations and preservation methods for a specimen	skull | skeleton
disposition	The current state of a specimen with respect to the	in collection, missing, loaned to

Term	Description	Example
	collection identified in collectionCode or collectionID	
associatedMedia	A list (concatenated and separated) of identifiers (publication, global unique identifier, URI) of media associated with the Occurrence	osf.io/cataloguecard
associatedReferences	A list (concatenated and separated) of identifiers (publication, bibliographic reference, global unique identifier, URI) of literature associated with the Occurrence	Literature DOI
associatedSequences	A list (concatenated and separated) of identifiers (publication, global unique identifier, URI) of genetic sequence information associated with the Occurrence	NCBI link
otherCatalogNumbers	A list (concatenated and separated) of previous or alternate fully qualified catalog numbers or other human-used identifiers for the same Occurrence, whether in the current or any other data set or collection	08-63
occurrenceRemarks	Comments or notes about the Occurrence	Found dead on road
eventRemarks	Comments or notes about the Event	Date of death | Date entered collection
eventDate	The date-time or interval during which an Event occurred. For occurrences, this is the date-time when the event was recorded. Not suitable for a time in a geological context	2000-07-30
verbatimEventDate	The verbatim original representation of the date and time information for an Event	1995-07-30
Year	The four-digit year in which the Event occurred, according to the Common Era Calendar	1995

Term	Description	Example
Month	The integer month in which the Event occurred	7
Day	The integer day of the month on which the Event occurred	30
continent	The name of the continent in which the Location occurs	North America
waterBody	The name of the water body in which the Location occurs	Pacific Ocean
islandGroup	The name of the island group in which the Location occurs	Broken Islands
island	The name of the island on or near which the Location occurs	Vancouver Island
Country	The name of the country or major administrative unit in which the Location occurs	Canada
countryCode	The standard code for the country in which the Location occurs	CA
stateProvince	The name of the next smaller administrative region than country (state, province, canton, department, region, etc.) in which the Location occurs	BC
municipality	The full, unabbreviated name of the next smaller administrative region than county (city, municipality, etc.) in which the Location occurs. Do not use this term for a nearby named place that does not contain the actual location	Saanich
locality	The specific description of the place. Less specific geographic information can be provided in other geographic terms (higherGeography, continent, country, stateProvince, county, municipality, waterBody, island, islandGroup). This term may contain information modified from the original to correct perceived errors or standardize the description	Long Beach

Term	Description	Example
verbatimLocality	The original textual description of the place	Spectacle Lake, Vancouver Island, BC
locationRemarks	Comments or notes about the Location	north side of Diana Island
identifiedBy	A list (concatenated and separated) of names of people, groups, or organizations who assigned the Taxon to the subject	Becky Wigen
identificationRemarks	Comments or notes about the Identification	Identified by B. Wigin from BCPM skeletons mainly on basis of size
taxonID	An identifier for the set of taxon information (data associated with the Taxon class). May be a global unique identifier or an identifier specific to the data set	https://www.gbif.org/species/2433458
nameAccordingToID	An identifier for the source in which the specific taxon concept circumscription is defined or implied. See nameAccordingTo	Taxonomic concept paper DOI
scientificName	The full scientific name, with authorship and date information if known. When forming part of an Identification, this should be the name in lowest level taxonomic rank that can be determined. This term should not contain identification qualifications, which should instead be supplied in the IdentificationQualifier term	(genus + specificEpithet) Canis familiaris
kingdom	The full scientific name of the kingdom in which the taxon is classified	Animalia
phylum	The full scientific name of the phylum or division in which the taxon is classified	Chordata
class	The full scientific name of the class in which the taxon is classified	Mammalia
order	The full scientific name of the order in which the taxon is classified	Carnivora

Term	Description	Example
family	The full scientific name of the family in which the taxon is classified	Canidae
genus	The full scientific name of the genus in which the taxon is classified	Canis
subgenus	The full scientific name of the subgenus in which the taxon is classified. Values should include the genus to avoid homonym confusion	
specificEpithet	The name of the first or species epithet of the scientificName	familiaris
taxonRank	The taxonomic rank of the most specific name in the scientificName	species
scientificNameAuthorship	The authorship information for the scientificName formatted according to the conventions of the applicable nomenclaturalCode	Linnaeus, 1758
vernacularName	A common or vernacular name	Domestic dog
taxonRemarks	Comments or notes about the taxon or name	recorded in catalogue card as "Canis sp."
previousIdentifications	A list (concatenated and separated) of previous assignments of names to the Organism	recorded in catalogue card as "Eumentopias jubata"
measurementRemarks	Comments or notes accompanying the MeasurementOrFact	tail missing
measurementType	The nature of the measurement, fact, characteristic, or assertion	Weight, Length
measurementUnit	The units associated with the measurementValue	g, cm, kg
measurementValue	The value of the measurement, fact, characteristic, or assertion	375, 45, 2.5

Appendix C. Sea Otter Specimen Case Study

Record

Term	Interpreted	Original	Remarks
Dataset name	UVICZL Mammals	fauncatm	
Dataset ID	tbd		
Institution code	UVIC		
Basis of record	Preserved specimen		
Collection code	UVICZL		
Collection ID	http://grbio.org/tbd Mammal Collection		
Dynamic Properties			

Occurrence

Term	Interpreted	Original	Remarks
Catalogue number	92.64	92/64	
Occurrence ID	Tbd		
Occurrence remarks	Found on beach in Nootka Sound	Found on beach in Nootka Sound	
Disposition	In collection		To be confirmed by audit
Preparations	Skeleton	Allowed to remain until became dessicated. Cleaned by Marine Mammal class, Bamfield M.S.	
Recorded by	Jane Watson | Bamfield Marine Station	Jane Watson	
Sex	Likely Male	Probably male	Faunal sex ontology (Figure C 1)
Life Stage			
Associated Media	https://www.osf.io/dx48g | https://www.firstvoices.com/kids/FV/sections/Data/Nuu-chah-nulth/Nuu%C4%8Daan%CC%93u%C9%AB/Nuu-chah-		OSF Catalogue Card (Figure C 2) | Indigenous animal name (regionally specific) (Figure C 3) | Sketchfab 3D image of skull (Figure C 4)

	nulth%20(Barkley)/learn/words/categories/7a469dce-ae04-4549-94dd-3d2dc5414833/12/6 | https://sketchfab.com/3d-models/sea-otter-skull-5b2266da9f8049d1a87280212162e0b7		
Associated References	https://www-degruyter-com.ezproxy.library.uvic.ca/document/doi/10.1525/9780520948976-008/html | https://www-sciencedirect-com.ezproxy.library.uvic.ca/science/article/pii/S0305440311004444?via%3Dihub		Research on Sea Otters (Figure C 5)[1] | (Figure C 6)[2]
Associated Sequences	https://www-ncbi-nlm-nih-gov.ezproxy.library.uvic.ca/Taxonomy/Browser/wwwtax.cgi?id=9615		NCBI Taxonomy (Figure C 7)
Other Catalogue Numbers	92/64	92/64	Card catalogue number

Organism

Term	Interpreted	Original	Remarks
Previous Identifications			

Event

Term	Interpreted	Original	Remarks
Year	1992	1992	
Month	August	August	
Day		1	

[1] (McKechnie and Wigen 2011)
[2] (Szpak et al. 2012)

Event date			
Verbatim event date	1/08/1992	1/08/1992	
Event Remarks	Date entered collection		

Taxon

Term	Interpreted	Original	Remarks
Taxon ID	https://www.gbif.org/species/2433670		GBIF LOD (Figure C 8)
Name According To ID			
Kingdom	Animalia		Inferred
Phylum	Chordata		Inferred
Class	Mammalia		Inferred
Order	Carnivora		Inferred
Family	Mustelidae		Inferred
Genus	Enhydra		Inferred
Subgenus			
Specific epithet	lutris		Inferred
Scientific name	Enhydra lutris Linnaeus, 1758	Enhydra lutris	Altered
Taxon Rank	Species	species	
Scientific Name Authorship	Linnaeus, 1758		Inferred
Vernacular Name	Sea Otter	Sea Otter	
Taxon Remarks	Enhydra lutris	Enhydra lutris	

Identification

Term	Interpreted	Original	Remarks
Identified By	Jane Watson	Jane Watson, Bamfield Marine Station	
Identification Remarks	Probably male	Probably male	

Location

Term	Interpreted	Original	Remarks
Continent	North America		
Water Body		Nootka Sound	

Island Group			
Island	Vancouver Island	Vancouver Island	
Country or area	Canada		
Country code	CA		Inferred
State province	BC		
Municipality			
Locality	https://native-land.ca/		Ancestral territory and language map (Figure C 9)
Verbatim locality	Beach in Nootka Sound		
Location Remarks			

Measurement or Fact

Term	Interpreted	Original	Remarks
Measurement remarks			
Measurement type			
Measurement unit			
Maesurement value			

Other

Term	Interpreted	Original	Remarks
License	http://creativeco.../by/4.0/legal code		
Bibliographic citation			
Access Rights			
Rights holder	The Anthropology Department of the University of Victoria		
Type	PhysicalObject		
Language	En		

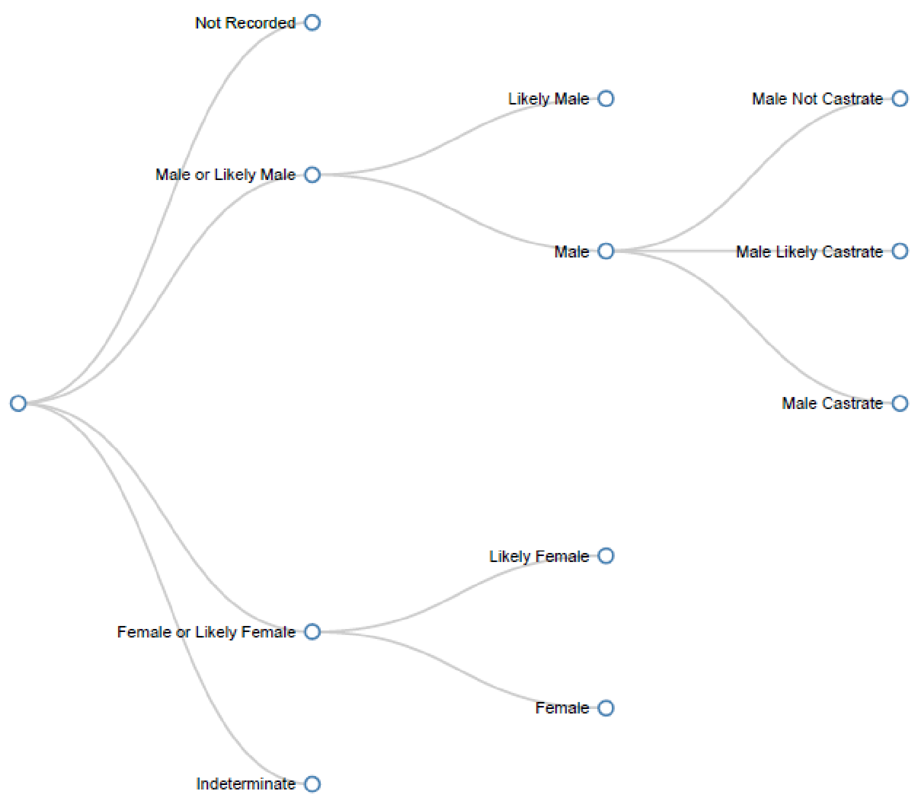
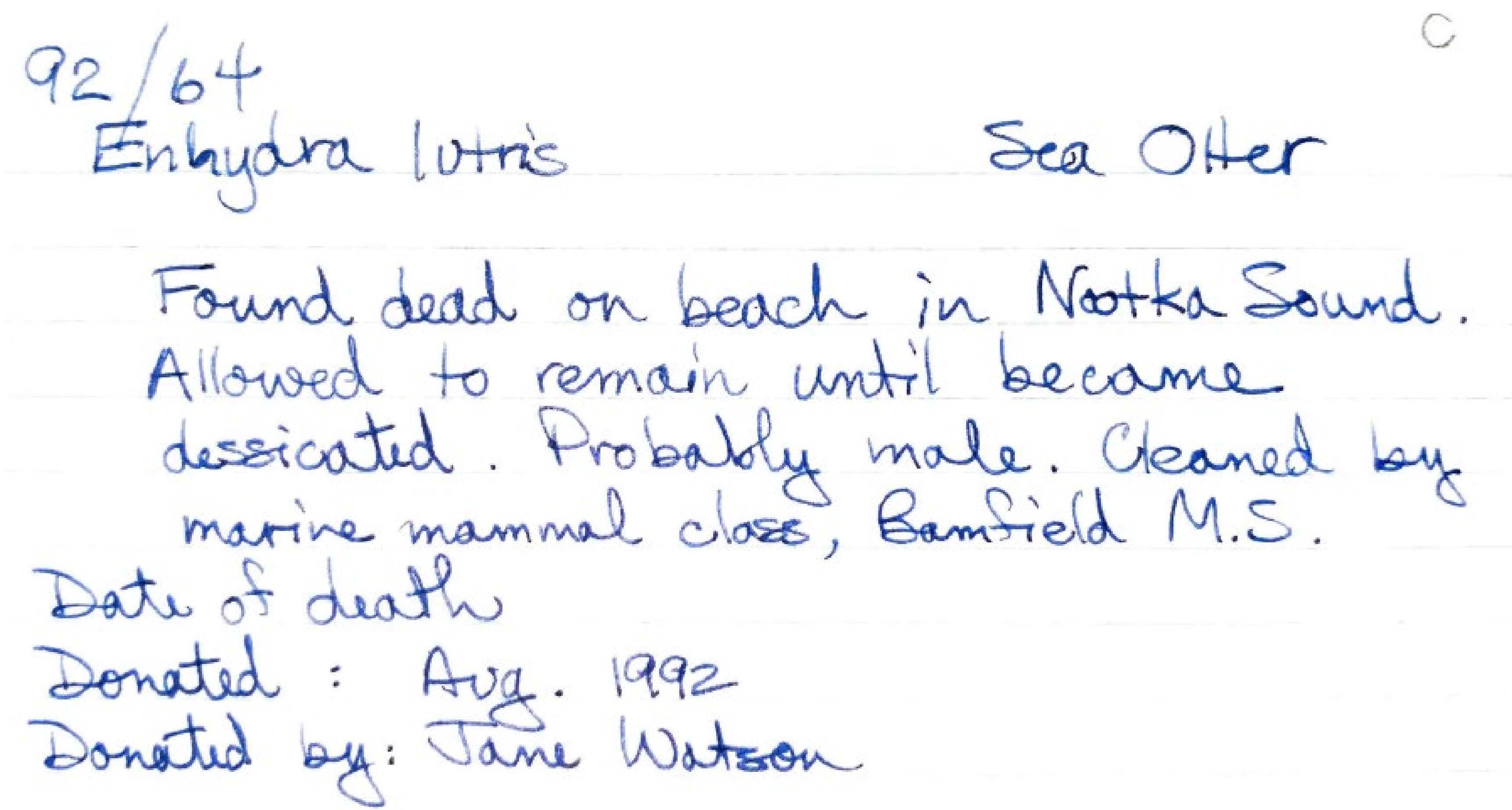


Figure C 1 Faunal sex ontology



C

92/64

Enhydra lutris Sea Otter

Found dead on beach in Nootka Sound. Allowed to remain until became dessicated. Probably male. Cleaned by marine mammal class, Bamfield M.S.

Date of death

Donated : Aug. 1992

Donated by: Jane Watson

Figure C 2 Catalogue Card

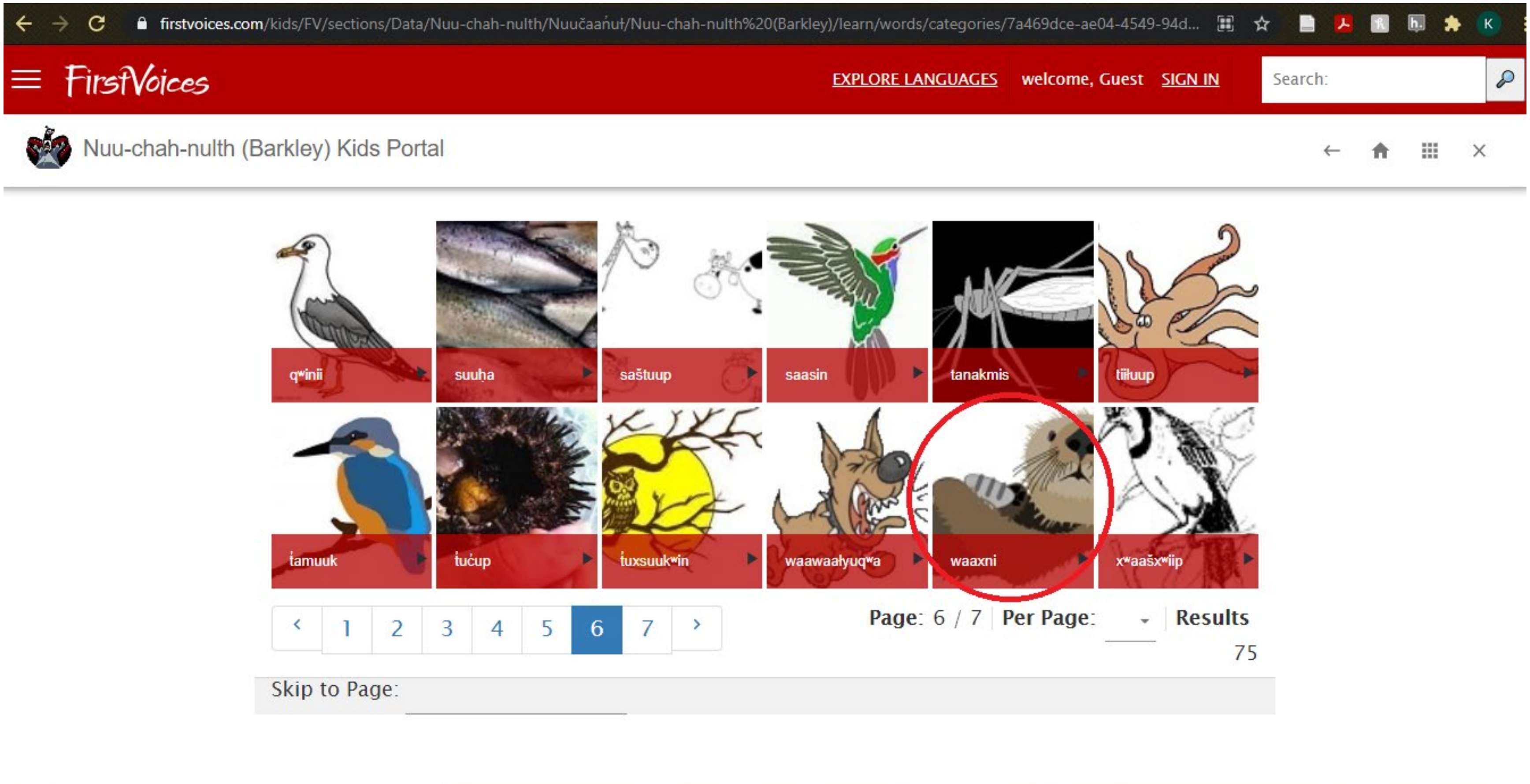


Figure C 3 Audio link to Indigenous animal name

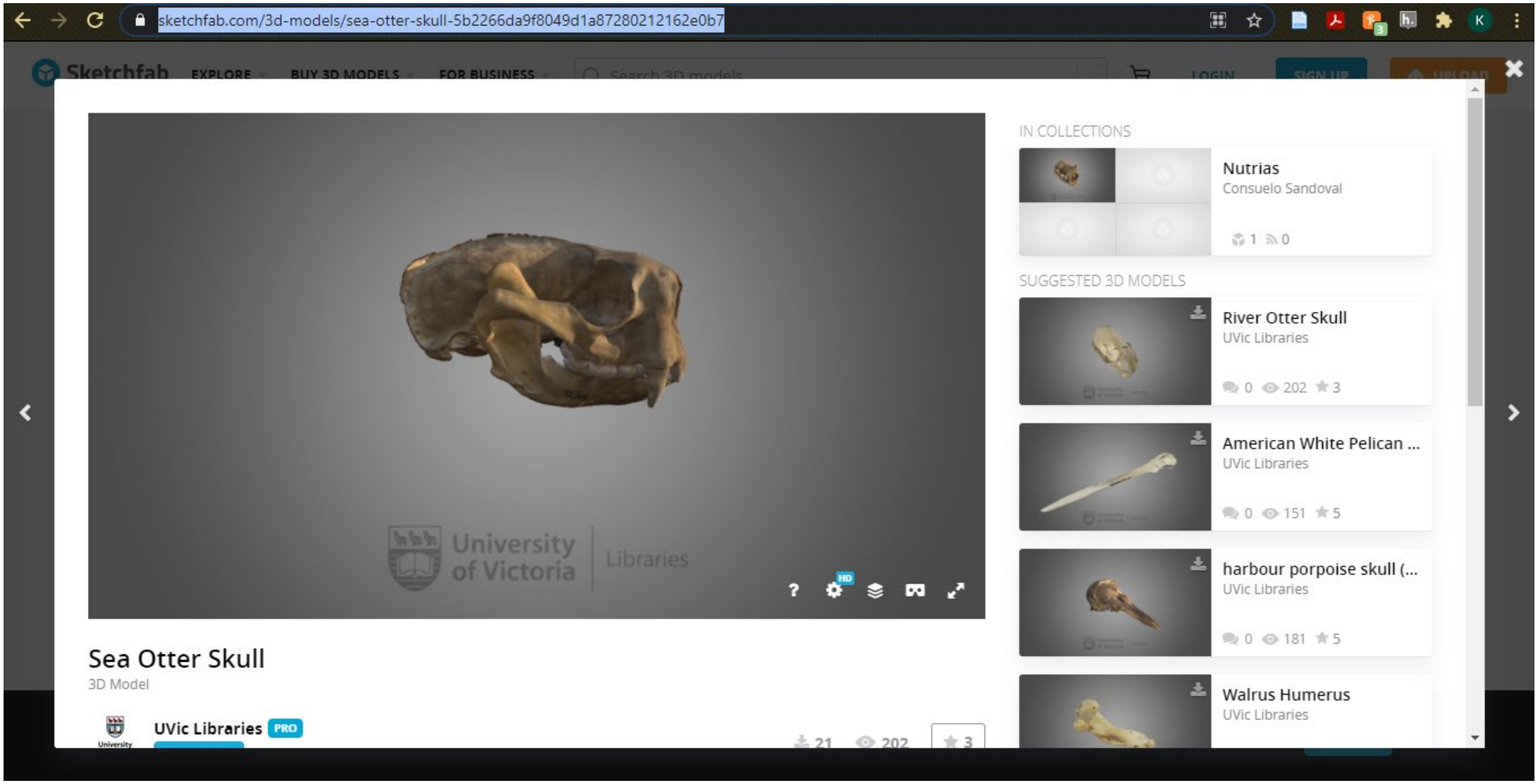


Figure C 4 Sea Otter Skull

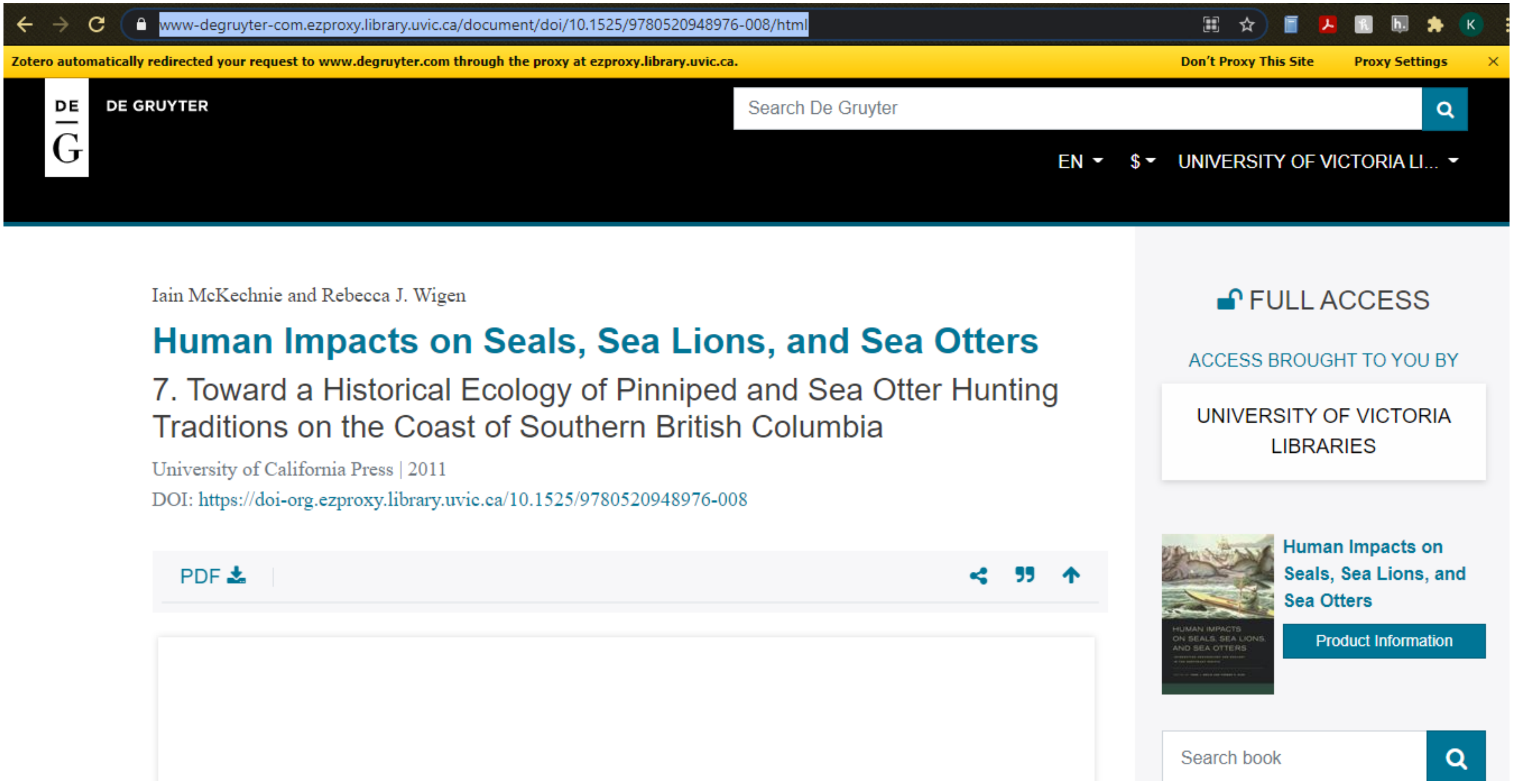
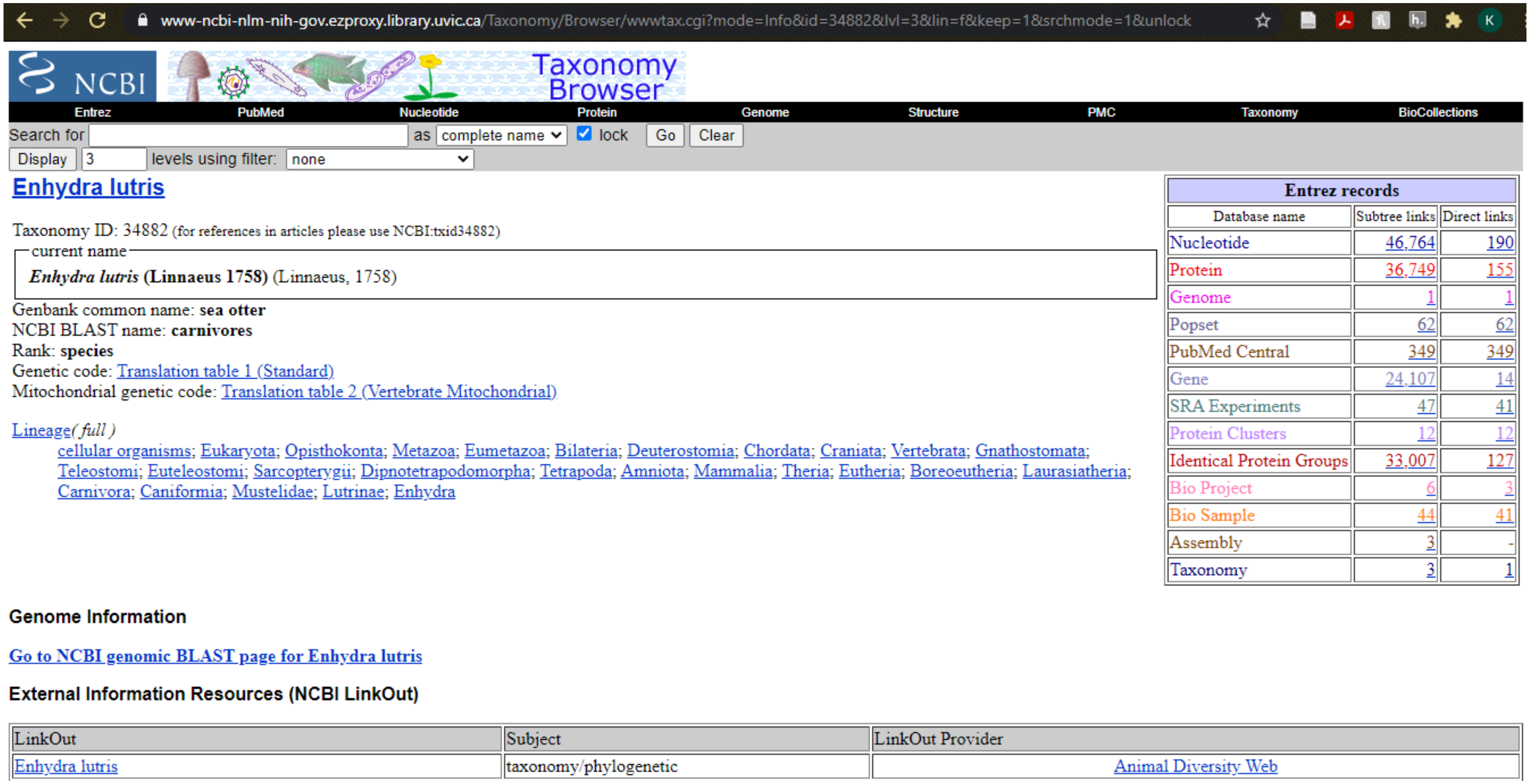


Figure C 5 Human Impacts on Seals, Sea Lions, and Sea Otters

Figure C 6 Historical Ecology of late Holocene sea otters (Enhydra lutris) from northern British Columbia: isotopic and zooarchaeological perspectives



www-ncbi-nlm-nih-gov.ezproxy.library.uvic.ca/Taxonomy/Browser/wwwtax.cgi?mode=Info&id=34882&lvl=3&lin=f&keep=1&srchmode=1&unlock

NCBI Taxonomy Browser

Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy BioCollections

Search for [ ] as complete name lock Go Clear

Display 3 levels using filter: none

Enhydra lutris

Taxonomy ID: 34882 (for references in articles please use NCBI:txid34882)

current name

***Enhydra lutris* (Linnaeus 1758)** (Linnaeus, 1758)

Genbank common name: **sea otter**
NCBI BLAST name: **carnivores**
Rank: **species**
Genetic code: Translation table 1 (Standard)
Mitochondrial genetic code: Translation table 2 (Vertebrate Mitochondrial)

Lineage(*full*)
cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria; Laurasiatheria; Carnivora; Caniformia; Mustelidae; Lutrinae; Enhydra

Entrez records

Database name	Subtree links	Direct links
Nucleotide	46,764	190
Protein	36,749	155
Genome	1	1
Popset	62	62
PubMed Central	349	349
Gene	24,107	14
SRA Experiments	47	41
Protein Clusters	12	12
Identical Protein Groups	33,007	127
Bio Project	6	3
Bio Sample	44	41
Assembly	3	-
Taxonomy	3	1

Genome Information

Go to NCBI genomic BLAST page for Enhydra lutris

External Information Resources (NCBI LinkOut)

LinkOut	Subject	LinkOut Provider
Enhydra lutris	taxonomy/phylogenetic	Animal Diversity Web

Figure C 7 NCBI Taxonomy

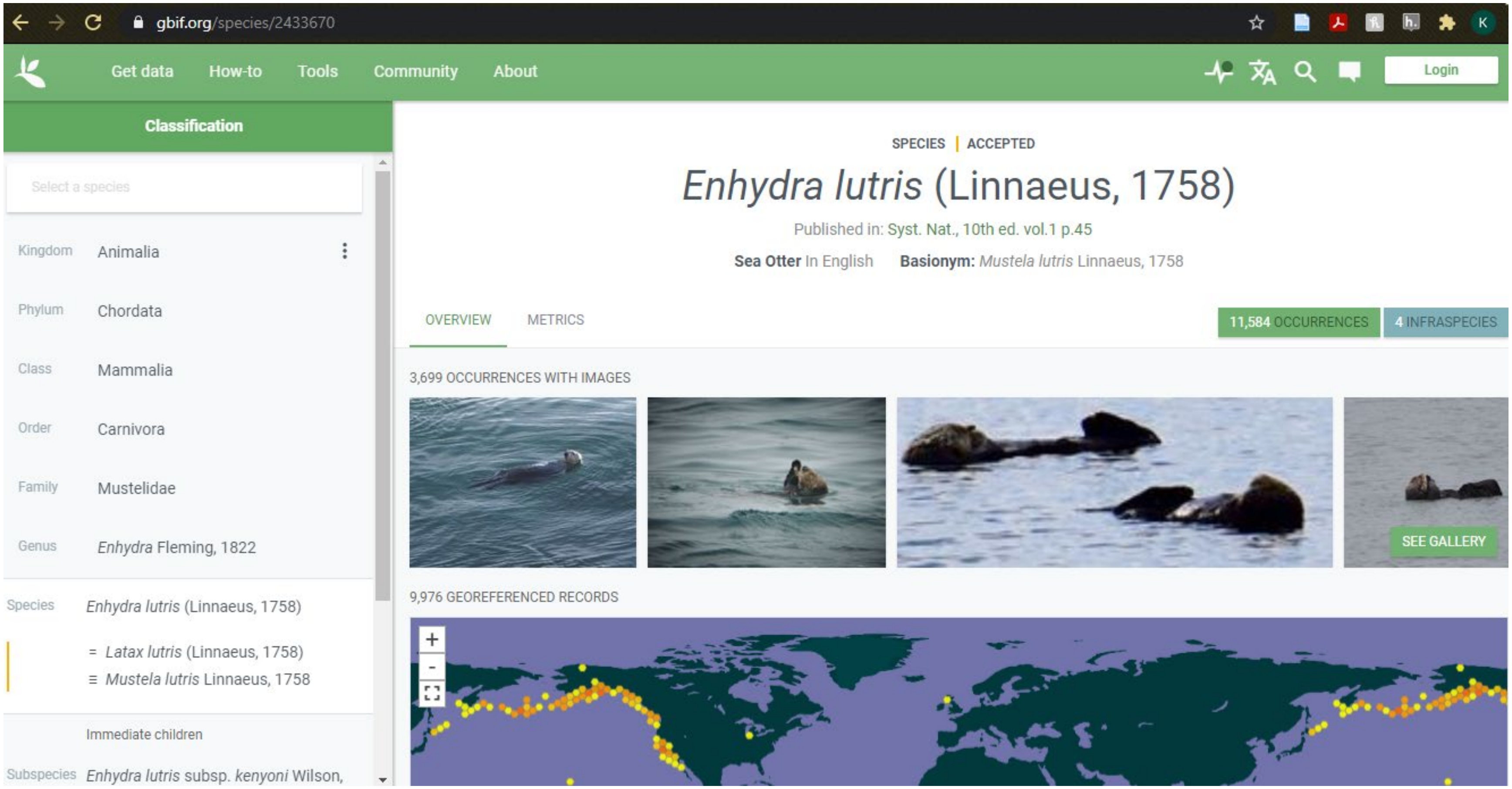


Figure C 8 GBIF Linked Open Data

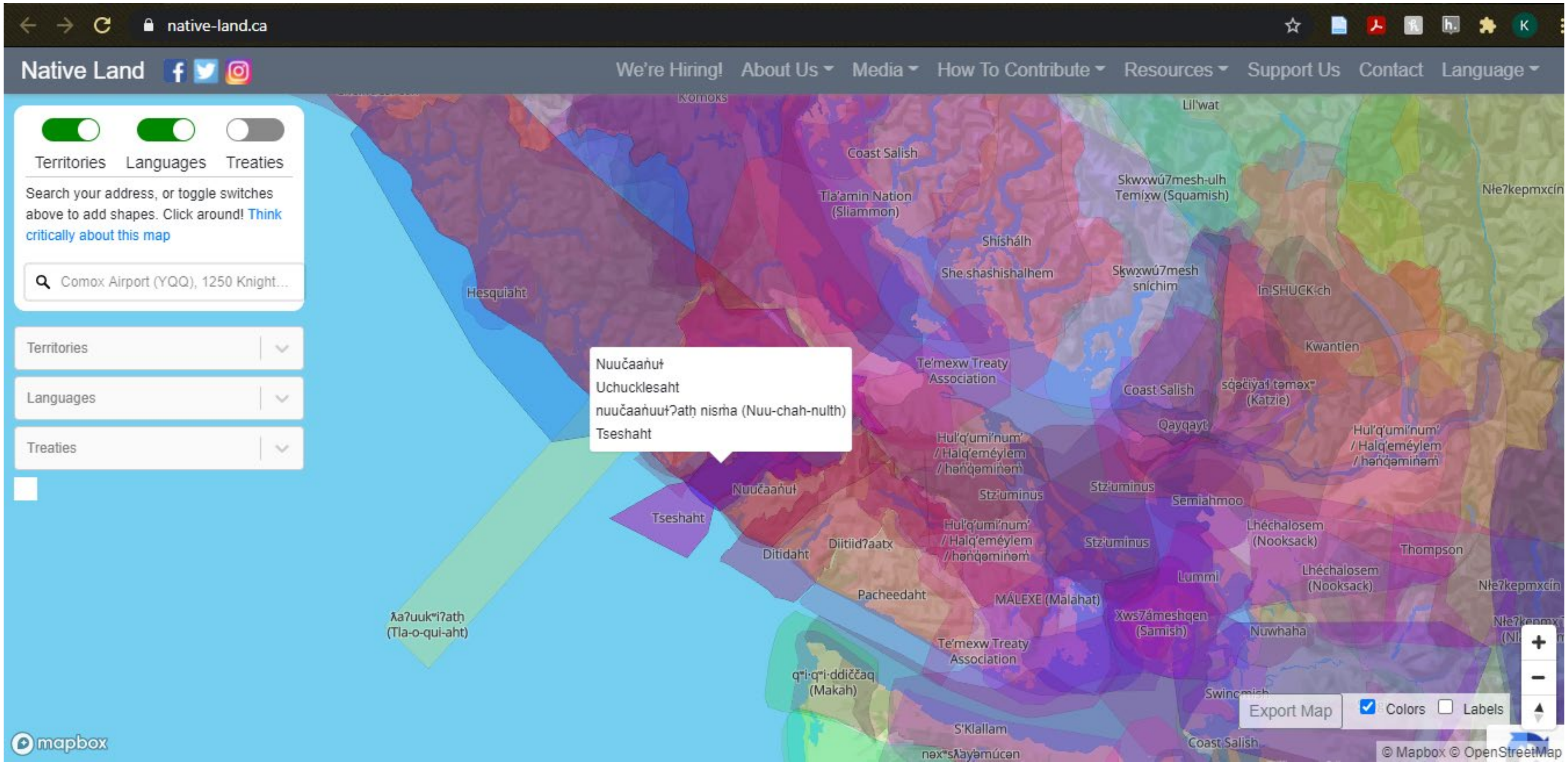


Figure C 9 Ancestral territories and languages