Data sharing and ocean visualization realized by Biologging intelligent Platform (BiP)

Authors: Shinichi Watanabe[1,2], Takuji Noda[3], Takuya Koizumi[3], Ken Yoda[4], Makoto A. Yoshida[5], Takashi Iwata[6,7], Hideaki Nishizawa[8], Junichi Okuyama[9], Kagari Aoki[10], Satoko S. Kimura[11], Kentaro Q. Sakamoto[10], Akinori Takahashi[12], Takuya Maekawa[13], Tomoko Narazaki[14,] Hiromichi Mitamura[15,16], Katsufumi Sato[10]

[1]Little Leonardo Co.

[2]Research Institute of Rare Birds Japan

[3]Biologging Solutions Inc.

[4]Graduate School of Environmental Studies, Nagoya University

[5]Lake Biwa Branch Office, Biodiversity Division, National Institute for Environmental Studies

[6]Faculty of Maritime Sciences, Graduate School of Maritime Sciences, Kobe University

[7]Ocean Policy Research Institute, Sasakawa Peace Foundation

[8]Graduate School of Informatics, Kyoto University

[9]Fisheries Technology Institute, Japan Fisheries Research and Education Agency

[10]Atmosphere and Ocean Research Institute, the University of Tokyo

[11]Center for Southeast Asian Studies, Kyoto University

[12]National Institute of Polar Research

[13]Graduate School of Information Science and Technology, Osaka University

[14]Faculty of Agriculture, Meijo University

[15]Field Science Education and Research Center, Kyoto University

[16]Graduate School of Agriculture, Kyoto University

Nihonseitai Gakkaishi (Japanese Journal of Ecology), 73(1):9-22

https://doi.org/10.18960/seitai.73.1_9

https://www.jstage.jst.go.jp/article/seitai/73/1/73_9/_article/-char/ja

Abstract

biologging involves attaching various sensors to animals to study their behavior, ecology, and surrounding environment. In this century, web-based electronic platforms for sharing biologging data have emerged globally. However, Japan has lagged in sharing biologging data domestically. This paper introduces a newly developed platform, the biologging intelligent Platform (BiP), aimed at storing, managing, and utilizing biologging data in Japan. To determine BiP's specifications, we evaluated 12 existing platforms based on six characteristics related to the types of data stored and analysis functionalities across three levels. The assessment considered features contributing to increased data storage capacity and formed the basis for defining BiP's specifications and future development directions. Comparative analysis highlighted that high levels of data openness, data type flexibility, and comprehensive data analysis tools contributed significantly to increased data storage capacity. Leveraging these features, BiP was designed to enhance data openness and data type flexibility. Additionally, it incorporates a unique Online Analytical Processing (OLAP) web system. BiP's OLAP system functions include uploading sensor data obtained from biologging devices (Level 0) to the BiP website, removing unnecessary parts pre-release or post-recovery, and generating Level 1 data in standard format after inputting individual and attachment metadata. Using GPS data, extract oceanographic physical information (Level 2 data) such as sea currents, wind, and waves. For publicly accessible data, users can download Level 1 and 2 data in CSV format and Network Common Data Form (NetCDF). Future plans involve adding functionalities to generate Level 3 data by gridding oceanographic information and expanding the target species from marine to terrestrial animals, broadening the geographical scope worldwide to increase the quality and quantity of collected data.

Keywords: data sharing; data standardization; data repository; marine physical observation; Online Analytical Processing (OLAP)

Introduction

In recent years, the acquisition of data through biologging, which involves attaching various sensors to animals, has advanced in the field of ecology. biologging was devised in the 1960s to record the elusive diving behavior of seals in the Antarctic (Kooyman 1965). Subsequently, advancements in electronic technology facilitated miniaturization, allowing diverse information to be collected as various sensors were integrated into data loggers. Presently, this method is utilized for various research purposes, including studying the movement, behavior, and physiological conditions of not only marine mammals but also invertebrates, fish, amphibians, reptiles, birds, and terrestrial mammals (Japanese Society of Bio-logging Science 2009, 2016; Takahashi & Yoda 2010; Iwata 2020; Yoshida & Mabuchi 2020). While techniques using radio or acoustic transmitters to track animals are sometimes distinguished as biotelemetry, this paper encompasses methods employing transmitters that accumulate and transmit data internally (Fedak 2004; Mitamura et al. 2017), considering them under biologging.

In Japan, researchers centered at the National Institute of Polar Research have been leading biologging studies focused on animals in the Antarctic region since the 1980s (Naito et al. 2012). The term ' biologging' originated from the First International biologging Symposium in Tokyo in March 2003 (Boyd et al. 2004). The following July, the Japanese Society of Bio-logging Science was established domestically, ahead of the rest of the world (http://japan biologgingsci.org/home, confirmed on July 29, 2022).

With the proliferation of biologging, studies that share biologging data owned by individual researchers or institutions have become more prevalent. For instance, international data sharing has led to collaborative research, such as meta-analyses of seabird behavior (Bonnet-Lebrun et al. 2021) and the creation of global distribution maps for conservation purposes (Beal et al. 2021). Systems storing and sharing biologging data in repositories to support collaborative research are also becoming widespread. Campbell et al. (2016) introduced a platform serving as a web-based electronic infrastructure to manage biologging data collected from wildlife. These platforms function as databases for storing biologging data and facilitate data analysis and sharing on the web. As of July 2022, our investigation found the existence of twelve such platforms, including those with changed hosting and platform names (Table 1). The primary operating institutions of each platform, the regions where data is collected, and the targeted animals are listed below.

Movebank (https://www.movebank.org/, confirmed on July 29, 2022), operated by the Max Planck Institute of Animal Behavior in Germany, targets all organisms worldwide. ZoaTrack (https://www.zoatrack.org/, confirmed on July 29, 2022), operated by The Atlas of Living Australia, focuses on organisms in the Australian region. Wireless Remote Animal Monitoring (WRAM) (https://www.slu.se/wram/, confirmed on July 29, 2022), operated by The Swedish University of Agricultural Sciences, targets organisms across Europe. Ocean Tracking Network (OTN) (https://oceantrackingnetwork.org/, confirmed on July 29, 2022), operated by Dalhousie University in Canada, focuses on marine animals worldwide. Ocean Biodiversity Information System-Seamap (OBIS-S) (https://seamap.env.duke.edu/, confirmed on July 29, 2022), operated by Duke University in the United States, targets marine animals globally. Seaturtle.org (http://seaturtle.org/tracking/, confirmed on July 29, 2022), operated by a private organization in the United States, targets oceanic animals, primarily sea turtles, globally. Animal Telemetry Network (ATN) (https://ioos.noaa.gov/project/atn/, confirmed on July 29, 2022), operated by the US Integrated Ocean Observing System, focuses on marine animals in North America. The IMOS Animal Tracking Database (https://animaltracking.aodn.org.au/, confirmed on July 29, 2022), operated by the Australian government, targets marine animals from Australia to the Antarctic region. Global Tagging of Pelagic Predators (GTOPP) (https://gtopp.org/, confirmed on July 29, 2022), operated as an international collaborative research, targets marine animals in North America. Seabird Tracking Database (STD) (http://seabirdtracking.org/, confirmed on July 29, 2022), operated by BirdLife International, a private organization based in the United Kingdom, targets seabirds globally. Great Lakes Acoustic Telemetry Observation System (GLATOS) (https://glatos.glos.us/, confirmed on July 29, 2022), operated by the Great Lakes Fishery Commission, targets freshwater fish in the Great Lakes of North America. Euromammals (https://euromammals.org/, confirmed on July 29, 2022), operated as an international collaborative research, targets terrestrial mammals in Europe.

While the founding years of many platforms are not explicitly stated, a majority were established after 2007. There were significant differences in the post-establishment development, with some experiencing a substantial increase in registered data after 2016, others showing no increase in data, and some disappearing. As biologging data platforms proliferate worldwide, Japan lacks its own platform, and the number of data registrations to foreign platforms from Japan is also limited. Hence, we developed a domestic platform, the biologging intelligent Platform (BiP; https://www.bip-earth.com/), for managing biologging data.

Among existing biologging data platforms, Movebank (Kays et al. 2022) has shown the most development. Metrics such as the number of data providers and projects far surpass other platforms (Table 1). However, within Movebank's database, the data collected from Japan, particularly from marine animals, is surprisingly scarce. Upon investigation by the authors, although approximately 7,000 projects are registered globally, only a meager 26 projects originated from Japan (including 1 for marine turtles and 12 for seabirds) (https://www.movebank.org/, confirmed on July 29, 2022). This accounts for only about 0.4% of the total, less than half the number from neighboring South Korea. Moreover, other existing platforms also lack a significant collection of biologging data acquired within Japan. Various reasons might contribute to the lack of registration from Japan, possibly due to concerns about registering data on foreign platforms. Particularly, providing biologging data collected through administrative projects to foreign platforms raises concerns about security, intellectual property rights, and diplomatic issues. Furthermore, while emphasizing the acquisition of location data concerning animal movements, unlike overseas studies, domestic biologging research has developed uniquely by utilizing acceleration sensors for behavior measurement and conducting marine observations based on GPS information, possibly leading to technological barriers in standardizing and sharing data.

In recent years, the field of biologging has seen advancements in collaboration with different disciplines and extensive utilization of data. Developing methods to share biologging data could lead to cross-disciplinary innovations. For instance, in meteorology and ocean physics, the utilization of biologging data for marine observations is advancing. Although technologies for marine observation such as weather satellites and autonomous underwater vehicles have progressed, observing the vast marine environment in three dimensions remains challenging. While weather satellites provide information on the ocean's surface, they cannot access information underwater due to electromagnetic wave limitations. The Sea Mammal Research Unit (SMRU) in the UK has been pioneering the application of biologging in ocean physics. They developed Satellite Relay Data Loggers (SRDL), devices attached to marine mammals like seals, transmitting salinity, water temperature, and depth (CTD) data via Argos satellites (Fedak 2004). SRDLs deployed on marine animals like seals transmit CTD data for over a year, processed on servers. Users can access temperature and salinity profiles with latitude-longitude information from SMRU's website (http://www.smru.st-and.ac.uk/, confirmed on July 29, 2022) (Photopoulou et al. 2015). Recently, a project named Animal Borne Ocean Sensors (AniBOS, https://anibos.com/, confirmed on July 29, 2022) has emerged, aiming to collect temperature and salinity profiles globally from SRDLs and build a worldwide ocean observation system (McMahon et al. 2021).

On the other hand, in Japan, rather than directly measuring ocean physics information such as water temperature and salinity using sensors embedded in biologging devices, methods have been developed to estimate the marine environment from behavioral data obtained through biologging. Although meteorological satellite data can estimate ocean physics information like currents and sea winds in the atmospheric-ocean boundary layer, direct on-site observations remain challenging. Additionally, meteorological satellite observations are limited by orbit altitude and coverage range. Satellites with lower orbit altitudes have higher spatial resolution but narrower observation ranges, preventing continuous observation of the same region. Therefore, meteorological satellite observations cannot capture rapidly changing marine physical environments at high resolutions. In recent years, new methods have been successively announced for estimating marine environments using biologging data obtained from Procellariiformes, such as shearwaters. Attaching GPS loggers to seabirds allows the estimation of current direction and speed based on latitude-longitude variations when the birds drift on the sea surface (Yoda et al. 2014). Furthermore, advancements now enable the measurement of wave height, direction, and period from temporal changes in position data, surpassing vertical and horizontal speeds (Uesaka et al. 2022). Also, from latitude-longitude information during the flight of Procellariiformes, wind direction and speed over the sea surface can be estimated (Yonehara et al. 2016; Goto et al. 2017). Assimilating these collected marine physics data from seabirds into traditional ocean flow models estimated from meteorological satellites has shown improvements in model accuracy (Miyazawa et al. 2015). Additionally, efforts are underway to directly collect ocean physics information using sensors attached to animals, such as small biologging devices obtaining atmospheric pressure over the sea surface from seabirds (Naruoka et al. 2021). Collating and openly sharing such directly and indirectly gathered marine physics and meteorological information could potentially be utilized for future weather and sea condition forecasts.

The role of biologging in marine observation is expected to become increasingly crucial in the future (Hirai et al. 2021), expanding its use in conservation and environmental sciences. For example, collecting biologging data from top predators like seabirds may aid in identifying biologically significant marine areas for biodiversity conservation measures (Watanuki et al. 2018) and assessing the status of pollution in contaminated marine regions due to harmful chemicals (Ito et al. 2013).

To encourage the utilization of biologging data across various fields, it's necessary not only for biologging researchers but also to construct an analysis system that stores and shares data in standard formats usable across diverse fields when determining BiP specifications. Many existing overseas biologging data platforms primarily aim at archiving and sharing data related to animal positions and movements for researchers handling biologging data. Consequently, data analysis is left to individual researchers, requiring advanced analytical skills to utilize publicly available data. Therefore, BiP aims not only to collect biologging data in Japan but also to build a system that can be utilized by users worldwide with various needs.

Comparison of Biologging Data Platforms

In determining the specifications for BiP, the characteristics of existing biologging data platforms were evaluated on the following six items (Taxon, Area, Data, Open, App., Quantity) using three levels (L1-3). Regarding data quantity, the evaluation was based on the number of registered projects and taxonomic groups publicly available on the platforms.

Evaluation Criteria:

Taxon: Target Animals

L1: Limited to specific animal groups like seabirds, freshwater fish, terrestrial mammals, etc. (Number of relevant platforms n = 3). L2: Encompassing multiple animal groups but restricted to general marine animals (such as marine turtles, seabirds, and marine fish) (n = 6). L3: Targeting all biological species from terrestrial to marine animals without restricting them to specific animal groups (n = 3).

Area: Main Regions of Data Collection

L1: Targeting regions within one country, similar to the Great Lakes in North America (n = 1). L2: Crossing multiple countries, such as Europe or the surrounding seas of North America, but specifying regions (n = 6). L3: Collecting data worldwide without specific region limitation (n = 5).

Data: Types of Collected Data

L1: Data obtained from specific devices like satellite transmitters, acoustic transmitters, etc. (n = 5). L2: Including data from multiple devices but limited to latitude-longitude related information like satellite transmitters, acoustic transmitters, GPS, etc. (n = 4). L3: Involving various sensor data beyond just latitude-longitude information, including depth, activity, etc. (n = 3).

Open: Level of Data Accessibility

L1: Some data might be viewable, but not accessible on the web, requiring consent from data providers to access (n = 7). L2: Data is viewable, and data providers cannot change conditions for data usage (n = 1). L3: Data providers have the freedom to choose conditions for data viewing and usage (n = 4).

App.: Extent of Data Analysis Tools

L1: Inability to view data online, with no provided data analysis tools (n = 3). L2: Capability to view data locations on a map online but lack of other analysis tools on the web (n = 7). L3: Availability of data viewing on a map and provided analysis tools online (n = 2).

Quantity: Amount of Collected Data

L1: Project numbers (fewer than 200) and taxonomic group numbers (fewer than 100) (n = 6). L2: Project numbers (200 to fewer than 1000) and taxonomic group numbers (100 to fewer than 500) (n = 5). L3: Project numbers (more than 1000) and taxonomic group numbers (more than 500) (n = 1).

Multiple Correspondence Analysis (MCA)

Based on the evaluation of the aforementioned six items, the similarity of characteristics among platforms and the relationships between each evaluation criterion were analyzed using Multiple Correspondence Analysis (MCA). MCA is a statistical method for dimension reduction in multivariate data using ordinal scales or qualitative variables, unraveling relationships between individual data and variables. In this study, dimension reduction through MCA was conducted using the six evaluation criteria (factors) data across three levels. Furthermore, hierarchical cluster analysis was performed based on the obtained individual scores to classify existing platforms. For MCA, the MCA function of the FactoMineR package 1.34 (Lê et al., 2008) in R ver. 4.1.13 (R Core Team, 2021) was employed. The hierarchical cluster analysis used R's hclust function with ward.D2 (Ward's method and Euclidean distance).

Eigenvalues, contribution rates, and cumulative contribution rates for each axis (dimension) obtained from applying MCA were calculated (Table 2). The first and second axes accounted for 50.5% of the total variance, and up to the fifth axis, it accounted for 84.6% of the total variance.

Scores for each factor and factor × level were plotted on the first and second axis space (Figure 1). From the relationships between factors, Open (degree of data accessibility), Taxon (target animals), and Quantity (data amount) were closely located, indicating strong relationships between these factors (Figure 1a). The distribution of Data, Taxon, and Quantity from L1 to L3 along the first axis suggests that the first axis indicates the degrees of these factors (Figure 1b). Moreover, the level with the highest data amount (Quantity L3) was proximate to high App. levels (Tool availability for data analysis - App. L3), Data (Data L3), and Taxon (Taxon L3), indicating strong relationships with them. These results suggest that high data accessibility, lack of constraints on target animals, and increased availability of data analysis tools and data types contribute to the increase in data amount.

Based on the results of hierarchical cluster analysis, platforms were classified into four groups (Groups A-D) (Figure 2). Individual scores of each platform were plotted on the first and second axis space, where Group A had the highest values on the first axis, followed by Groups B, D, and C in descending order (Figure 2a). Each Group exhibited common characteristics as follows (Table 1). Group A included sensor data other than latitude-longitude information (Data L3), high freedom in data accessibility (Open L3), and comprehensive analysis tools available on web browsers (App. L3). Group B had low data accessibility and lacked data display functions on maps. Group C targeted global marine life, with high data accessibility and focusing only on latitude-longitude information. Group D had data display functions on maps but lacked other provided analysis tools. Additionally, the target regions, target animals, data types, and levels of data accessibility were restricted for each. While competition might arise within these Groups, differentiation might be achieved due to variations in target regions and animals.

BiP Specifications

System Overview

In designing BiP specifications (Table 1), we aimed to avoid direct competition with existing platforms while expecting increased data volume. As previously mentioned, most of the biologging data collected and accumulated by Japanese biologging researchers and institutions are not included in overseas platforms, resulting in limited publicly available data from the Japanese coast to the East Asian region. Therefore, in this region, BiP primarily targets biologging data recorded by marine animals, including marine mammals, seabirds, sea turtles, and marine fish (Taxon L2). BiP collects animal behavior data and gathers oceanographic data derived from animals. Hence, the data types include sensor data such as latitude-longitude information recorded by GPS or satellite transmitters, as well as sensor data like depth, atmospheric pressure, water temperature, salinity, acceleration, geomagnetism, and light intensity (Data L3). Furthermore, BiP allows the data provider to freely choose the level of data accessibility, whether for viewing or utilization (Open L3) and incorporates a proprietary Online Analytical Processing (OLAP) web analysis system (App. L3). Comparing BiP specifications with existing platforms like Zoa Track and Movebank, it is believed to belong to

Group A (Figure 2). While surpassing these existing platforms in data volume may not be easy for the newer BiP to evolve, BiP needs to enhance the unique OLAP functionality absent in other platforms and aim for a different development direction.

BiP aims not just to archive biologging data but also to become a platform supporting data analysis. Biologging data comes in varied sensor types and file formats, which differ based on manufacturers and models. Simply archiving various sensor data will not provide analysis capabilities within the system. Moreover, without metadata regarding individual data acquisition or attachment (metadata), it's impossible to appropriately utilize biologging data for comparisons based on gender or breeding conditions or to exclude periods before individual release or after logger retrieval. In anticipation of web-based analysis, BiP archives essential metadata simultaneously with sensor data obtained from biologging devices, standardizing it for ease of processing within the system. The metadata items and formats stored in BiP conform to international standard formats (Sequeira et al., 2021) proposed by Integrated Taxonomic Information System (ITIS), Climate and Forecast Metadata Conventions (CF), Attribute Conventions for Data Discovery (ACDD), International Organization for Standardization (ISO) to construct the web system (Figure 3). The standardization process is designed to progressively process sensor data (Level 0 data) into standardized Level 1 data, subsequently into Level 2 and 3 data processed for analysis. Although no system has yet been publicly released apart from BiP that stores biologging data in accordance with these international standard formats (Table 1, confirmed as of July 29, 2022), we believe that even if another system achieving international standard formats were to be released in the future, system compatibility could be ensured.

As of 2022, the current BiP utilizes OLAP-equipped analysis tools from Level 1 data to generate Level 2 data by estimating marine physical information (ocean currents, wind, and waves) developed through domestic biologging research (Yoda et al., 2014; Yonehara et al., 2016; Goto et al., 2017; Uesaka et al., 2022). As an upcoming enhancement, there are plans to generate Level 3 data, such as aggregated marine physical information from Level 2 derived from multiple individuals aggregated per 10 km grid.

Data Collection and Publication Policy

In the field of ecology, including biologging data, the concept of 'Open Science'—making research data openly and freely available, not limited to the scientific community—is rapidly spreading globally (Osawa et al., 2014). In recent years, many international academic journals have made depositing data from published articles into international data repositories mandatory. In Japan, the basic concept of managing and utilizing research data funded by public funds was outlined in the 6th term of the Cabinet Office's Basic Plan on Science, Technology, and Innovation, which was established in 2021. It urged research and development institutions, such as universities, to formulate data policies and deposit research data into institutional repositories (https://www8.cao.go.jp/cstp/tougosenryaku/togo2021_honbun.pdf, confirmed on July 29, 2022).

Considering these recent trends, BiP has designed and operates a platform that serves as a repository for biologging data, reflecting the evaluation values shown in Table 1.

As of 2022, BiP stipulates that data providers must be members of Japanese Society of Bio-logging Science (http://japan-biologgingsci.org/home/, confirmed on July 29, 2022) or individuals designated by BiP's operating committee. There are plans for enabling anyone, domestically or internationally, to become a data provider in the future.

When uploading data to BiP, data providers can choose between restricted access data (Private: non-public data) and unrestricted access data (Open: public data). Metadata, such as attachment information (Tables 3 - 5), is accessible regardless of public or private status. Private data allows viewing of route maps or temporal diagrams on the map but requires adherence to usage restrictions specified by the provider for downloading and utilization. Public data, governed by the Creative Commons Attribution 4.0 International license (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/deed.en, confirmed on July 29, 2022), permits free replication, redistribution, and modification by anyone, without limitation for members of Japanese Society of Bio-logging Science, provided the credited metadata is acknowledged.

Usage Method

While data included in BiP doesn't guarantee an absence of errors, quality control is conducted to detect and correct input errors made during data registration, as outlined below.

Users of BiP (both data providers and data consumers) can access the BiP website (https://www.bip-earth.com/, Figure 4) to perform the following processes interactively (Figure 3). 'Interactive' here refers to a mode where the system responds immediately to user operations or inputs, progressing the processes interactively. After user registration, data providers first upload sensor data recorded by biologging devices (RawFile: Level 0 data). Subsequently, they input metadata for the tagged organism (Organism, Table 3) via the web interface. To prevent data inconsistencies due to input errors or variations, the system incorporates as many selectable options as possible and automatically populates the fields. For instance, selecting an animal group under 'Basic Info' presents a list of scientific names for the target species, which, upon selection, automatically populates their standardized English names. The target species include 209 species (as of July 29, 2022), aggregated based on survey reports published in the reports of Japanese Society of Bio-logging Science (http://japan-biologgingsci.org/home/) and member surveys related to biologging research conducted by Japanese researchers. If users wish to upload data for species not listed, they can communicate their requests through the contact information provided on the BiP website to update the list and enable data upload.

In the 'Additional Info' section, users select gender, growth stage, and birth information. If there's capture information for the target individual, they input the capture date, location name, and latitude-longitude. Latitude-longitude for the capture location can be input by clicking on the map displayed on the web interface. They input measurement dates and respective values for measured values of weight and length. They can add these details in cases of multiple measurements (such as during attachment and retrieval) or multiple measurement sites. Additionally, if identification tags other than biologging devices, like bird leg bands or flipper/pit tags for sea turtles, are used, users can add these numbers.

Subsequently, users input metadata about their equipment (Instrument, Table 4), including details like manufacturer and output data format for devices and sensors. Since input inaccuracies or omissions in the mandatory fields of major devices might disrupt subsequent standardization processes (Standardize), the system allows for the selection of manufacturer and model from templates, which are automatically populated.

Upon creating RawFile, Organism, and Instrument, users input details regarding device attachment, animal release, device retrieval locations, and related timestamps (Table 5). Users input the recording start date and time for devices that record at regular intervals (Equal sampling). These inputs undergo standardization, removing unnecessary sections pre-release and post-retrieval, resulting in Level 1 data. In cases where multiple individuals' information is captured in a single sensor data file (as with SRDL), standardization must be done for each individual. While there may be many duplicate input fields during metadata registration and standardization for data obtained in the same project or survey site in a given year, users can duplicate the registered files from each initial screen and modify the necessary sections, such as individual names or measured values, for registration.

Level 1 data, whether public or private, can be visualized (Visualize) for an overview of rough movement tracks or temporal data on the web interface. After visualizing the data and conducting quality checks, data providers set the data as open (Open) or private (Private) and input contact information for the data owner. Data owners can register multiple contacts, including co-researchers and data providers. If any issues arise during the data entry process, the system indicates error areas and prompts data providers for corrections.

Data consumers can use the system as follows: Metadata for registered data is viewable by anyone, regardless of being public or private. For public data, after user registration, Level 1 data can be downloaded in CSV format, and the integrated data with metadata can be converted into the Network Common Data Form (NetCDF) binary file format for download. NetCDF is a widely used standard format in various fields like meteorology, oceanography, and climate change, enabling easy utilization in other systems by providing NetCDF files.

Future Perspectives

Many existing biologging platforms primarily collect sensor data focusing on latitude and longitude information, likely stemming from the historical development of devices primarily acquiring latitude and longitude, such as SRDL. Conversely, manufacturers and researchers in Japan have developed devices that gather sensor information beyond latitude and longitude, including acceleration, geomagnetism, angular velocity, electrocardiograms, and images (Japanese Society of Bio-logging Science, 2009, 2016). Examples include methods measuring marine animal behavior patterns using depth, speed, and acceleration data (e.g., Yoda et al., 1999; Tanaka et al., 2001) or studies based on images obtained from camera loggers (Sato et al., 2015). Many of these studies have been limited to individual researchers' specific uses, and existing international platforms lack the capability to store acceleration data and do not accommodate image data. Therefore, a platform capable of collecting and analyzing acceleration and image data via OLAP could potentially evolve BiP into a distinctive platform. While various biologging devices are expected to be developed worldwide, expanding the system to incorporate these data into BiP is considered an essential challenge.

Furthermore, the promotion of open data and compliance with mandatory data registration are expected to become necessary in the field of ecology. Considering these trends, the significance of platforms like BiP is expected to be increasingly recognized. To enhance the platform's functions as a data repository, it is desirable to assign persistent identifiers such as Digital Object Identifiers (DOIs) to facilitate access to data (Cabinet Office, 2019). Ensuring continuous funding is essential for obtaining DOIs and enhancing the functionality of the data repository, ensuring its continuous public availability. In the future, efforts will be made by our affiliated research institutions, societies, and academic associations to explore sustainable maintenance and management methods for BiP.

Acknowledgment

The BiP introduced in this paper was developed under the budget of the Ministry of Education, Culture, Sports, Science, and Technology's 'Advanced Technology Development Program for the Utilization of Marine Biological Big Data to Promote the Use of Marine Resources.'

References

Beal M, Dias MP, Phillips RA, Oppel S, Hazin C, Pearmain EJ, Adams J, Anderson DJ, Antolos M, Arata JA, Arcos JM, Arnould JP, Awkerman J, Bell E, Bell M, Carey M, Carle R, Clay TA, Cleeland J, Colodro V, Conners M, Cruz-Flores M, Cuthbert R, Delord K, Deppe L, Dilley BJ, Dinis H, Elliott G, De Felipe F, Felis J, Forero MG, Freeman A, Fukuda A, González-Solís J, Granadeiro JP, Hedd A, Hodum P, Igual JM, Jaeger A, Landers TJ, Le Corre M, Makhado A, Metzger B, Militão T, Montevecchi WA, Morera-Pujol V, Navarro-Herrero L, Nel D, Nicholls D, Oro D, Ouni R, Ozaki K, Quintana F, Ramos R, Reid T, Reyes-González JM, Robertson C, Robertson G, Romdhane MS, Ryan PG, Sagar P, Sato F, Schoombie S, Scofield RP, Shaffer SA, Shah NJ, Stevens KL, Surman C, Suryan, RM, Takahashi A, Tatayah V, Taylor G, Thompson DR, Torres L, Walker K, Wanless R, Waugh SM, Weimerskirch H, Yamamoto T, Zajkova Z, Zango L, Catry P (2021) Global political responsibility for the conservation of albatrosses and large petrels. Science Advances, 7:eabd7225

Bonnet-Lebrun AS, Dias MP, Phillips RA, Granadeiro JP, Brooke MDL, Chastel O, Clay TA, Fayet AL, Gilg O, González-Solís J, Guilford T, Hanssen SA, Hedd A, Jaeger A, Krietsch J, Lang J, Le Corre M, Militão T, Moe B, Montevecchi WA, Peter HU, Pinet P, Rayner MJ, Reid T, Reyes-González JM, Ryan PG, Sagar PM, Schmidt NM, Thompson DR, van Bemmelen R, Watanuki Y, Weimerskirch H, Yamamoto T, Catry P (2021) Seabird migration strategies: flight budgets, diel activity patterns, and lunar influence. Frontiers in Marine Science, 1543

Boyd IL, Kato A, Ropert-Coudert Y (2004) Bio-logging science: sensing beyond the boundaries. Memoirs of National Institute of Polar Research. Special Issue, 58: 1-14

Cabinet Office, Goverment of Japan (2019) Research data repository development and operation guidelines (in Japanese) (https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/guideline.pdf, confirmed on July 29, 2022)

Campbell HA, Urbano F, Davidson S, Dettki H, Cagnacci F (2016) A plea for standards in reporting data collected by animal-borne electronic devices. Animal Biotelemetry, 4:1

Fedak M (2004) Marine animals as platforms for oceanographic sampling: a" win/win" situation for biology and operational oceanography. Memoirs of National Institute of Polar Research Special issue, 58:133-147

Goto Y, Yoda K, Sato K (2017) Asymmetry hidden in birds' tracks reveals wind, heading, and orientation ability over the ocean. Science Advances, 3:e1700097

Hirai J, Miya M, Fujiki T, Kuwano-Yoshida A, Otosaka S, Kaeriyama H, Kako S, Kataoka T, Matsuoka D, Hidaka M, Sugiyama D, Kojima F (2021) Decadal vision in oceanography 2021: New methods and problems. Umino Kenkyuu (Oceanography in Japan), 30:227-253 (in Japanese with English Abstract)

Ito A, Yamashita R, Takada H, Yamamoto T, Shiomi K, Zavalaga C, Abe T, Watanabe S, Yamamoto M, Sato K, Kohno H, Yoda K, Iida T, Watanuki Y (2013) Contaminants in tracked seabirds showing regional patterns of marine pollution. Environmental science & technology, 47:7862-7867

Iwata T (2020) Utilization and development of biologging in marine mammal studies. Honyuurui Kagaku (Mammalian Science), 60:281-296 (in Japanese with English Abstract)

Japanese Society of Bio-logging Science (ed.) (2009) Doubututachi Hushiginisemaru Biologging (Biologging approaches the mysteries of animals). Kyoto Tsushinsha, Kyoto (in Japanese)

Japanese Society of Bio-logging Science (ed.) (2016) Biologging 2: Doubututachino shirarezaru sekaiwo saguru (Biologging 2: Exploring the unknown world of animals). Kyoto Tsushinsha, Kyoto (in Japanese)

Kays R, Davidson SC, Berger M, Bohrer G, Fiedler W, Flack A, Hirt J, Hahn C, Gauggel D, Russell B, Kölzsch A, Lohr A, Partecke J, Quetting M, Safi K, Scharf A, Schneider G, Lang I, Schaeuffelhut F, Landwehr M, Storhas M, Schalkwyk L, Vinciguerra C, Weinzierl R, Wikelski M (2022) The Movebank system for studying global animal movement and demography. Methods in Ecology and Evolution, 13:419-431

Kooyman, GL (1965) Techniques used in measuring diving capacities of Weddell Seals. Polar Record, 12:391-394

Lê S, Josse J, Husson F (2008) FactoMineR: An R package for multivariate analysis. Journal of Statistical Software, 25:1-18

McMahon CR, Roquet F, Baudel S, Belbeoch M, Bestley S, Blight C, Boehme L, Carse F, Costa DP, Fedak MA, Guinet C, Harcourt R, Heslop E, Hindell MA, Hoenner X, Holland K, Holland M, Jaine FR, Jeanniard du Dot T, Jonsen I, Keates TR, Kovacs KM, Labrousse S, Lovell P, Lydersen C, March D, Mazloff M, McKinzie MK, Muelbert MM, O'Brien K, Phillips L, Portela E, Pye J, Rintoul S, Sato K, Sequeira AM, Simmons SE, Tsontos VM, Turpin V, van Wijk E, Vo D, Wege M, Whoriskey FG, Wilson K, Woodward B (2021) Animal Borne Ocean Sensors – AniBOS – an essential component of the Global Ocean Observing System (GOOS). Frontiers in Marine Science, 8:1625

Mitamura H, Thorstad EB, Uglem I, Økland F (2017) In situ measurement of salinity during seaward migration of Atlantic salmon post-smolts using acoustic transmitters with data-storage capabilities and conventional acoustic transmitters. Animal Biotelemetry, 5:1-10

Miyazawa Y, Guo X, Varlamov SM, Miyama T, Yoda K, Sato K, Kano T, Sato K (2015) Assimilation of the seabird and ship drift data in the north-eastern sea of Japan into an operational ocean nowcast/forecast system. Scientific Reports, 5:1-10

Naito Y, Sato K, Takahashi A, Watanabe Y (2012) Biologging: Animal behavior from a "penguin perspective". Naruyamadou Shobou, Tokyo (in Japanese)

Naruoka M, Goto Y, Weimerskirch H, Mukai T, Sakamoto T, Sakamoto KQ, Sato K (2021) Application of inertial and GNSS integrated navigation to seabird biologging. Journal of Robotics and Mechatronics, 33:526-536

Osawa T, Jinbo U, Iwasaki N (2014) Current status and future perspective on "Open Data" in biodiversity science, Japan. Nihonseitai Gakkaishi (Japanese Journal of Ecology), 64:153-162 (in Japanese)

Photopoulou T, Fedak MA, Matthiopoulos J, McConnell B, Lovell P (2015) The generalized data management and collection protocol for conductivity-temperature-depth satellite relay data loggers. Animal Biotelemetry, 3:1-11

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/

Sato K, Aoki K, Nakamura U, Watanabe S (2015) Yaseidoubutuha Naniomiteirunoka—Biologging Funtouki (What are wild animals looking at: chronicles of Bio-logging adventures) (Canon Foundation Library), Maruzen Planet Publishing, Tokyo (in Japanese)

Sequeira AM, O'Toole M, Keates TR, McDonnell LH, Braun CD, Hoenner X, Jaine FR, Jonsen ID, Newman P, Pye J, Bograd SJ, Hays GC, Hazen EL, Holland M, Tsontos VM, Blight C, Cagnacci F, Davidson SC, Dettki H, Duarte CM, Dunn DC, Eguíluz VM, Fedak M, Gleiss AC, Hammerschlag N, Hindell MA, Holland K, Janekovic I, McKinzie MK, Muelbert MM, Pattiaratchi C, Rutz C, Sims DW, Simmons SE, Townsend B, Whoriskey F, Woodward B, Costa DP, Heupel MR, McMahon CR, Harcourt R, Weise M (2021) A standardisation framework for bio-logging data to Advance Ecological Research and conservation. Methods in Ecology and Evolution, 12:996-1007

Takahashi A, Yoda K (2010) Advances in biologging science: a review of bird studies. Nihon Torigakkaishi (Japanese Journal of Ornithology), 59:3-19

Tanaka H, Takagi Y, Naito Y (2001) Swimming speeds and buoyancy compensation of migrating adult chum salmon Oncorhynchus keta revealed by speed/depth/acceleration data logger. Journal of Experimental Biology, 204:3895-3904

Uesaka L, Goto Y, Yonehara Y, Komatsu K, Naruoka M., Weimerskirch H., Sato K, Sakamoto KQ (2022) Ocean wave observation utilizing motion records of seabirds. Progress in Oceanography, 200:102713

Watanuki Y, Yamamoto Y, Sato M, Yamamoto T, Yoda K, Takahashi A. Using seabirds to delineate Ecologically or Biologically Significant marine Areas in offshore pelagic waters. Nihonseitai Gakkaishi (Japanese Journal of Ecology), 68:81-99 (in Japanese with English Abstract)

Yoda K, Sato K, Niizuma Y, Kurita M, Bost CA, Le Maho Y, Naito Y (1999) Precise monitoring of porpoising behaviour of Adelie penguins determined using acceleration data loggers. Journal of Experimental Biology, 202:3121-3126

Yoda K, Shiomi K, Sato K (2014) Foraging spots of streaked shearwaters in relation to ocean surface currents as identified using their drift movements. Progress in Oceanography, 122:54-64

Yonehara Y, Goto Y, Yoda K, Watanuki Y, Young LC, Weimerskirch H, Bost CA, Sato K (2016) Flight paths of seabirds soaring over the ocean surface enable measurement of fine-scale wind speed and direction. Proceedings of the National Academy of Sciences, 113:9039-9044

Yoshida M, Mabuchi K (2020) Bio-logging studies in lake and reservoir: an individual-based approach to reveal fish behavior and ecology. Chikyu Kankyou (Global Environmental Research), 25:65-78 (in Japanese)

Table 1: Comparison of the Biologging intelligent Platform (BiP) and other platforms managing biologging data.

Name of Platforms	Taxon (Target animals)	Area (Main regions of data collection)	Data (Types of collected data)	Open (Data accessibility)	App. (Data analysis tool)	Quantity (Amount of collected data)				
							N of projects	taxa	ind.	amount (mil)
Movebank	L3: all	L3: world	L3: lat/long (Satellite, VHF), Others	L3: Public (CC0, CC BY, CC BY-NC) &Private	L3: Map display, GIS analysis, Smartphone app.	L3:	6,973	1,178		7,100
ZoaTrack	L3: all	L2: around Australia	L2: lat/long (Satellite, VHF)	L3: Public (CC BY) &Private	L3: Map display, home range analysis, noise filter	L2:	883	391	16,076	
Wireless Remote Animal Monitoring (WRAM)	L3: all	L2: Europe	L3: lat/long (Satellite, VHF)&Others	L1: Private	L1: None	L2:			5,861	286
Ocean Tracking Network (OTN)	L2: Marine animals	L3: world	L2: lat/long (Satellite, Acoustic)	L3: Public (CC0, CC BY, CC BY-NC) &Private	L2: Map display	L2:	956	311		
Ocean Biodiversity Information System-Seamap (OBIS-S)	L2: Marine animals	L3: world	L2: lat/long (Satellite, Acoustic)	L3: Public (CC0, CC BY, CC BY-NC) &Private	L2: Map display, Display of potential population density	L2:		742	1,552	8
Seaturtle.org	L2: Marine animals	L3: world	L1: lat/long (Satellite)	L1: Private	L2: Map display	L1:		80	657	
Animal Telemetry Network (ATN)	L2: Marine animals	L2: North America	L2: lat/long (Satellite, Acoustic)	L1: Private*	L2: Map display	L1:	145	68	4,357	
The IMOS Animal Tracking Database	L2: Marine animals	L2: Australia - Antarctica	L1: lat/long (Acoustic)	L2: Public (CC BY)	L1: None	L1:	141	161	10,679	126
Global Tagging of Pelagic Predators (GTOPP)	L2: Marine animals	L2: North America	L1: lat/long (Satellite)	L1: Private*	L2: Map display	L1:		36		
Seabird Tracking Database (STD)	L1: Seabirds	L3: world	L1: lat/long (Satellite)	L1: Private	L2: Map display	L2:		151	37,171	24
Great Lakes Acoustic Telemetry Observation System (GLATOS)	L1: Freshwater fishes	L1: Great Lakes (North America)	L1: lat/long (Acoustic)	L1: Private	L2: Map display	L1:	137	35		
Euromammals	L1: Terrestrial mammals	L2: Europe	L3: lat/long (Satellite, VHF)&Others	L1: Private	L1: None	L1:	120	7	4,500	15
Biologging intelligent Platform (BiP)	L2: Marine animals	L2: around Japan	L3: lat/long (Satellite, VHF)&Others	L3: Public (CC BY) &Private	L3: Map display, Data standardization, acquisition of marine environmental data					

Table 2: Eigenvalues and contribution rates for each dimension by Multiple Correspondence Analysis (MCA).

Dimension	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Eigenvalue	0.592	0.417	0.312	0.215	0.156	0.138	0.086	0.056	0.017	0.011
Contribution rate (%)	29.60	20.85	15.62	10.73	7.79	6.90	4.32	2.79	0.84	0.56
Cumulative contribution rate (%)	29.60	50.45	66.07	76.80	84.60	91.49	95.81	98.60	99.44	100

Table 3: Input items for the metadata of organisms in BiP (Organism).

Input item		Explanation
Basic Info	Individual Name*	given by the data provider
	Organism Category	select a taxon from Invertebrate, Fish, Amphibian&Reptile, Bird, Cetacean, Pinniped, Sirenia, or Terrestrial mammal
	Scientific Name*	Scientific name
	Common Name	Standard english name
Additional Info	Sex	select from Male, Female, or Unknown
	Reproductive class	select from Adult, SubAdult, Juvenile, Newborn, or Unknown
	Age	If it was known, enter in one-year increments
	Origin	select from Wild, Cultured, or Unknown
	Capture Location Name	Name of location where the animal was captured
	Capture Location: latitude, longitude	latitude and longitude where the animal was captured
	Capture Datetime	Date and time when the animal was captured
	TimeZone	select Time Zone of Capture Datetime
Size & Weight	Mass	Mass of the animal (kg)
	DateTime of Mass Measure	Date and time when the animal was weighed
	TimeZone	select Time Zone of DateTime of Mass Measure
	Comment	Notes on measurement methods, etc.
	Size	length of body part (m)
	Size Category	select a body part
	Comment	Notes on measurement methods, etc.
Others	Capture Method Details	Description of capture method
	Other tag ID	Tag numbers for other identification tags such as bird foot rings
	Comment	Other noteworthy things

*Required item

Table 4: Input items for the metadata of instruments in BiP (Instrument).

Input item		Explanation
Instrument Info	Instrument Manufacturer*	select the equipment manufacturer
	Instrument Model*	
	Instrument Type*	select from GPS, Argos, GLS, TDR, VHF, or Acoustic
	Instrument Name*	give the device a name that the registrant can identify
Data Info	Upload File Extension*	select csv or txt
	Data Separator*	select from Comma, Space, Tab, SemiColon, or Colon
	Decimal*	select period(.) or comma (,)
Sensor Info	Sensor Type*	select from Latitude, Longitude, Pressure, Internal temperature, External temperature, light intensity, Salinity, Speed, Triaxial acceleration, Triaxial geomagnetism
	Sensor Manufacturer	
	Sensor Model Name	
	Units Reported*	Select measurement unit
	lowerSensorDetectionLimit	Lower limit of measurement value
	upperSensorDetectionLimit	Upper limit of measurement value
	sensorPrecision	
	sensorResolution	
	SensorColumnName*	Column name of target sensor data
	SensorColumnRowNumber*	Line number containing column name (initial value is 1)
	SensorDataRowNumber*	Line number where sensor data starts (initial value is 2)
	DateTime	If the data is not recorded at regular intervals (Equal sampling), enter the column that contains the recording date and time, and select the date and time format.
	TimeZone	select Time Zone of sensor data
	Edge Processing	Enter the details if data processing (Edge processing) is performed within the device.
	Additional Note	Special notes regarding sensors
Other Info		Special notes regarding general equipment

*Required item

Table 5: Input items for data standardization in BiP (Data standardize).

Input item		Explanation
Instrument	InstrumentSerial*	Serial number of the instrument
	Mass	Mass of the instrument
Deployment	deploymentDateTime	Date and time deployed
	TimeZone	select Time Zone of deploymentDateTime
Data Info	deploymentLocationLat	latitude where the instrument was deployed
	deploymentLocationLon	longitude where the instrument was deployed
	Attachment Place	select External or Internal
	Attachment Details	notes regarding the attachment method
Release	releaseDateTime*	Date and time released
	TimeZone*	select Time Zone of releaseDateTime
	releaseLocationLat*	latitude where the animal was released
	releaseLocationLon*	longitude where the animal was released
	Release details	notes when the animal was released
Detachment or Recovery	detachmentDateTime	Date and time recover the instrument
	TimeZone	select Time Zone of detachmentDateTime
	detachmenLocationLat	latitude where the instrument was detached
	detachmenLocationLon	longitude where the instrument was detached
	deploymentEndType*	select Recapture or Drop
	detachmentDetails	notes when the instrument was detached
Other Info		Special notes until the instrument was recovered

*Required item

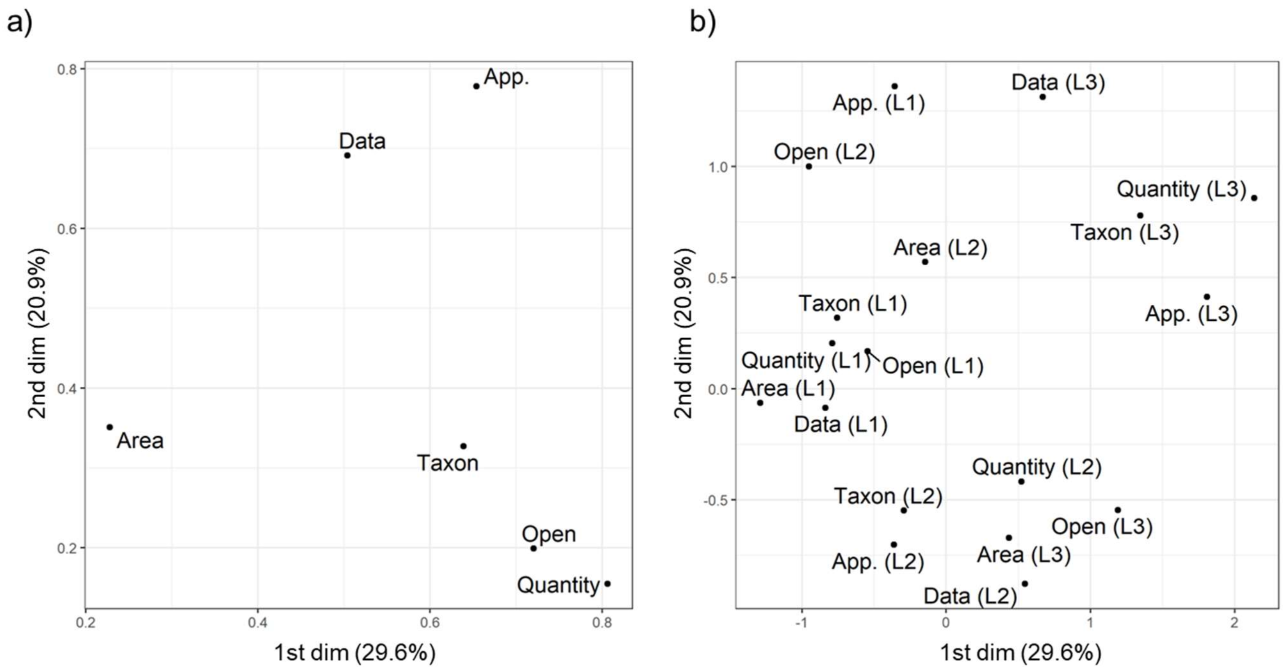

Figure 1: Relationship between evaluation items (a) and evaluation items × levels (b) of the platform managing biologging data evaluated by the contribution rates of the first and second dimensions of Multiple Correspondence Analysis (MCA). The proximity in a two-dimensional space indicates a strong relationship between closely positioned evaluation items. Refer to Table 1 for details on each evaluation item.

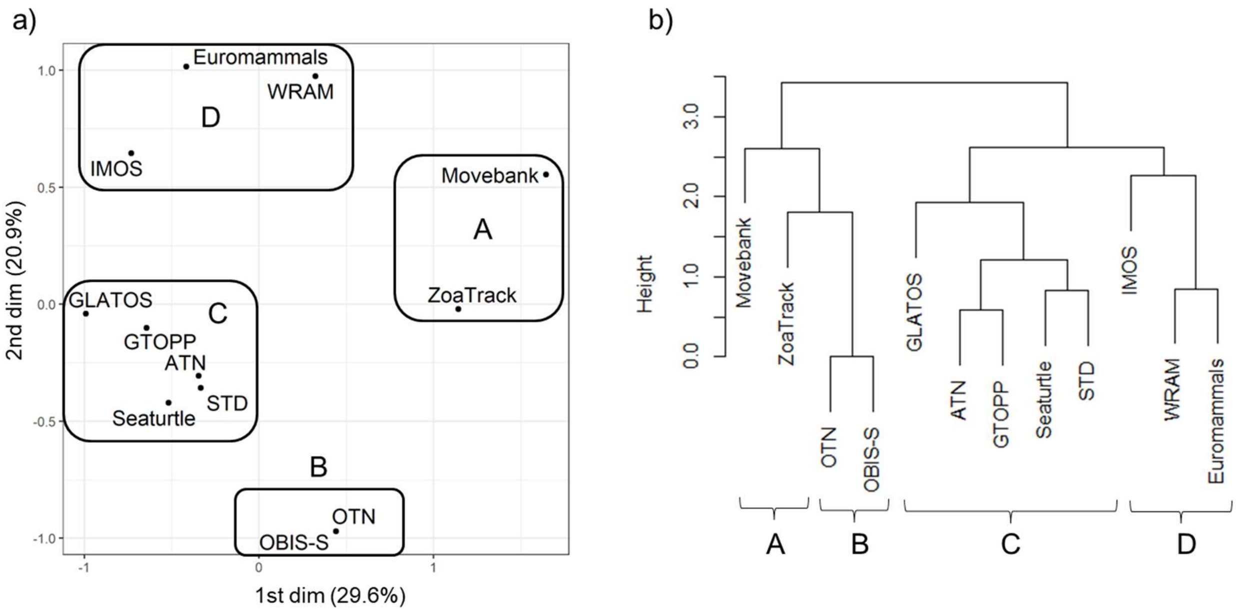

Figure 2: Comparison of features of platforms managing biologging data evaluated using Multiple Correspondence Analysis (MCA). Platforms were classified into four groups (Group A-D) based on the scatter plot (a) on the first and second-dimensional space and the dendrogram obtained by hierarchical clustering (b). A: Movebank, ZoaTrack. B: Ocean Tracking Network (OTN), Ocean Biodiversity Information System-Seamap (OBIS-S). C: Seaturtle.org (Seaturtle), Animal Telemetry Network (ATN), Global Tagging of Pelagic Predators (GTOPP), Seabird Tracking Database (STD), Great Lakes Acoustic Telemetry Observation System (GLATOS). D: Wireless Remote Animal Monitoring (WRAM), The IMOS Animal Tracking Database (IMOS), Euromammals.

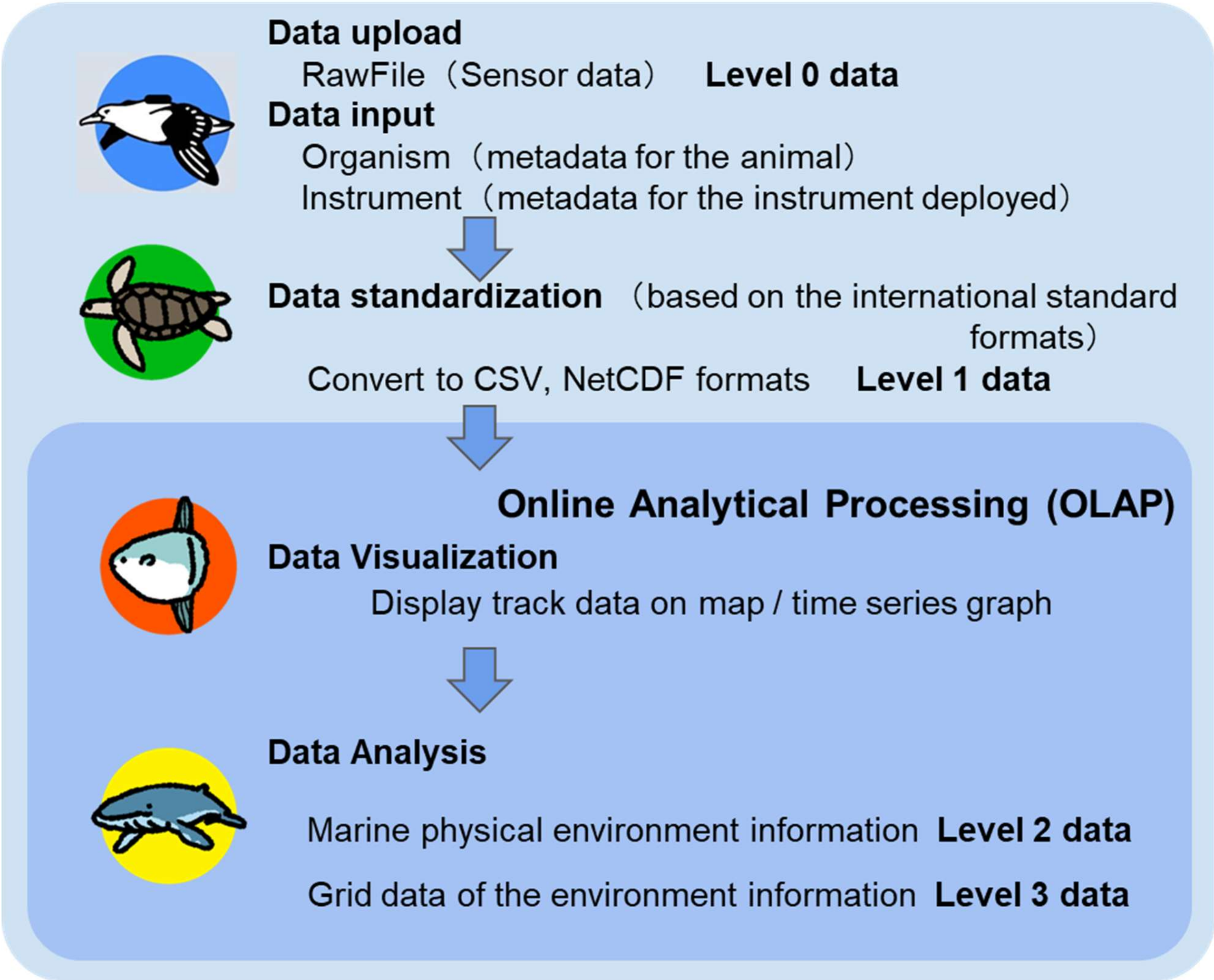

Figure 3: Data processing flow in the Biologging intelligent Platform (BiP).

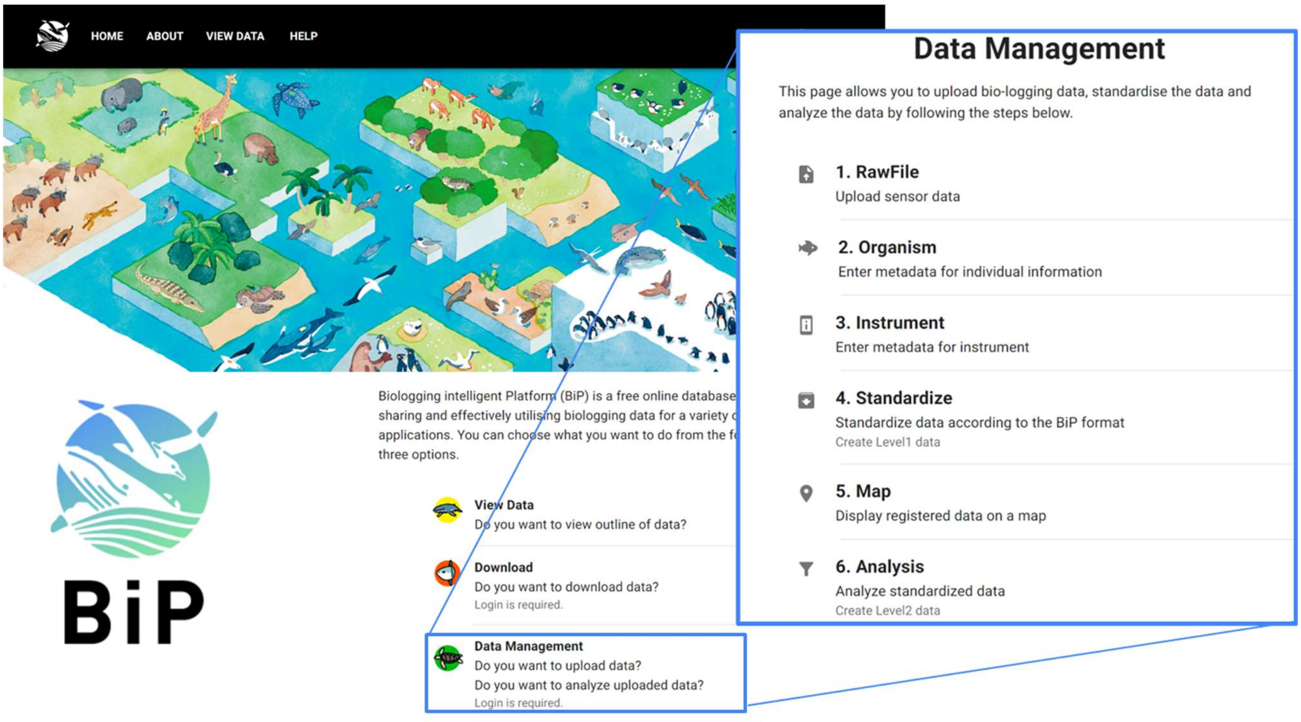

Figure 4: Screenshots of the Biologging intelligent Platform (BiP) homepage and the Data Management page equipped with Online Analytical Processing (OLAP) (https://www.bip-earth.com/, accessed on January 9, 2024).