



Article

Bioinspire-Explore: Taxonomy-Driven Exploration of Biodiversity Data for Bioinspired Innovation

Adrien Saint-Sardos ^{1,*}, Annabelle Aish ^{2,†}, Nikolay Tchakarov ^{1,†}, Thierry Bourgoïn ³ , Luce-Marie Petit ¹, Jian-Sheng Sun ² and Régine Vignes-Lebbe ³

¹ Centre d'Études et d'Expertises en Biomimétisme de Senlis (CEEBIOS), 62 Rue du Faubourg Saint-Martin, 60300 Senlis, France

² Bioinspire-Museum, Museum National d'Histoire Naturelle, 57 rue Cuvier, 75005 Paris, France; annabelle.aish@mnhn.fr (A.A.)

³ Sorbonne Université, Muséum National d'Histoire Naturelle, CNRS, EPHE, Université des Antilles, Institut de Systématique Évolution Biodiversité, ISYEB, CP 48, 57 Rue Cuvier, 75005 Paris, France

* Correspondence: adrien.saint-sardos@ceebios.com

† These authors contributed equally to this work.

Abstract: Successful bioinspired design depends on practitioners' access to biological data in a relevant form. Although multiple open-access biodiversity databases exist, their presentation is often adapted to life scientists, rather than bioinspired designers. In this paper, we present a new tool, "Bioinspire-Explore", for navigating biodiversity data in order to uncover biological systems of interest for a range of sectors. Bioinspire-Explore allows users to search for inspiring biological models via taxa (species, genera, etc.) as an entry point. It provides information on a taxon's position in the "tree of life", its distribution and climatic niche, as well as its appearance. Bioinspire-Explore also shows users connections in the bioinspiration literature between their taxon of interest and associated biological processes, habitats, and physical measurements by way of their semantic proximity. We believe Bioinspire-Explore has the potential to become an indispensable resource for both biologists and bioinspired designers in different fields.

Keywords: biodiversity; bioinspiration; data science; computer-aided biomimetics; taxonomy; open-access; NLP; biology push



Citation: Saint-Sardos, A.; Aish, A.; Tchakarov, N.; Bourgoïn, T.; Petit, L.-M.; Sun, J.-S.; Vignes-Lebbe, R. Bioinspire-Explore: Taxonomy-Driven Exploration of Biodiversity Data for Bioinspired Innovation. *Biomimetics* **2024**, *9*, 63. <https://doi.org/10.3390/biomimetics9020063>

Academic Editors: Stanislav N. Gorb, Giuseppe Carbone, Thomas Speck and Peter Fratzl

Received: 20 December 2023

Revised: 16 January 2024

Accepted: 17 January 2024

Published: 23 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bioinspiration is a creative approach based on the observation of living systems [1]. Over the past twenty years, bioinspiration has helped solve engineering problems [2–4], inspire architectural programs [5], improve water management [6], support cybersafety [7,8], and even resolve in-orbit servicing challenges [9–11]. As impressive as these case studies are, they are based on a few dozen, perhaps a few hundred, biological models [12,13], despite more than 2 million species having been identified [14]. To reach its full potential, the field needs a greater range of "biological muses" from molecular [15] to ecosystem scales [16,17]. Advances in the fields of biology and ecology are able to provide this inspiration: we know more about living systems today than ever before. Moreover, technological developments allow us to mimic and/or incorporate nature into design in ways that were impossible even a decade ago. How can this potential be unlocked to support the ambitions of a growing number of bioinspiration organisations, startups, laboratories etc. [13,18]? Moreover, how can we draw attention to the lesser known taxonomic groups, whose diversity of form and function has much to teach us [19–21]? Several initiatives have been set up over the last decade(s) to support bioinspired designers' access to biological information. Through the AskNature platform, living systems can be explored with a focus on functions, using the Biomimicry 3.8 "Taxonomy" to translate these into biological "strategies" that have or could inspire innovation [22]. Other initiatives have combined this functional

approach with advanced algorithms, processing large datasets (texts, images, etc.) that support online biomimicry tools, such as Mimicus (<https://www.lib.mimic.us/ap/> accessed on 1 January 2024), developed by Sun-Young Kim and colleagues, or PeTal from NASA [23]. Finally, one of the most promising ways to explore scientific literature is the “trade-off” approach, focused on the balance between different functions necessary for a species’ survival [24,25]. Such an approach, combined with biomimetic ontologies and engineering-to-biology thesauri developed during the 2000s [26–28], led to the development of the E2BMO tool. This computer-aided user interface allows the exploration of scientific literature using the concepts of functions and trade-offs [29]. All these tools help to generate ideas for bioinspired products, processes, and systems. However, at present, these (and other available resources) orientate the user towards a limited set of taxa. In this article, we present a new tool, Bioinspire-Explore (<https://Bioinspire-Explore.mnhn.fr/> accessed on 1 January 2024), developed to simplify access to biodiversity data in order to facilitate “biology push” approaches to bioinspiration without inherent taxonomic bias. Bioinspire-Explore offers the bioinspired designer the phylogenetic “tree of life” as a starting point, allowing the exploration of (known) extant biodiversity in its entirety. It is structured around taxa (i.e., names of species, genus, families, etc.) entered by users based on their curiosity or existing knowledge. While the immensity of taxonomic data may seem overwhelming to a non-biologist [30], Bioinspire-Explore aims to support the user in their discovery of the living world and present connections between taxa and associated information (ecological, climatic, visual) of relevance. Above all, its value lies in its ability to display the semantic proximity between taxa, biological processes, habitats, and physical measurements within the scientific literature on bioinspiration. Bioinspire-Explore does not replace interaction with biologists: transdisciplinary collaboration is essential to understanding the principles that govern the living world [31] and for effective bioinspiration [32,33]. However, Bioinspire-Explore provides a means to begin biological exploration, individually or as a team, and to stimulate bioinspiration in all its forms (i.e., biomimicry, bioassistance, bioproduction, Nature Based Solutions, etc. . .) [34].

2. Materials and Methods

2.1. Data Sources

Table 1 presents the data sources behind Bioinspire-Explore. Taxonomic data are provided via the application programming interface (API) of the Global Biodiversity Information Facility GBIF (<https://gbif.org/> accessed on 1 January 2023), with names of taxa (vernacular and scientific), their occurrences, (spatial coordinates), taxonomic rank, and images. The GBIF taxon identification number is used to query Wikidata. If the Wikidata node of the taxon in question is found, the first paragraph and image of its Wikipedia page is displayed on the Bioinspire-Explore interface via the Wikipedia representational state transfer (REST) API. Going through the Wikidata graph prevents “false friend-associated” pages (e.g., avoiding a query using the word “Vespa” displaying information on motorcycles). Images are imported through the GBIF API and originate from the iNaturalist database.

Table 1. Data sources and their providers.

Data Type	Sourced From
Taxonomic backbone	GBIF
Occurrence data	GBIF
Taxon-related photographs	iNaturalist
General information on given taxon	Wikipedia
Geo-climatic data	WorldClim
List of biological process	Wikidata
Related entities, e.g., co-cited species	Biomig corpus

Climatic data for each taxon are provided by the WordClim (<https://www.worldclim.org/> accessed on 1 January 2024) initiative. Plotting a taxon's climatic conditions involves several steps: (i) Nine monthly climatic variables (elevation, average precipitation, max. precipitation, min. precipitation, solar irradiance, average temp, min. temp, max. temp, and wind speed) are downloaded. (ii) Taxon occurrence maps are downloaded from GBIF, containing a list of hexagonal geographical zones where the taxon has been recorded. (iii) The climatic variables maps are overlaid with the occurrences maps. (iv) The final metric, as displayed in the Climate Tab, represents a spatial mean over the distribution range, averaged over 12 months of the year. To provide the range, maximums and the minimums are calculated over the monthly distributions.

2.2. Architecture

Bioinspire-Explore is an open-source tool developed as a full-stack application; in other words, it comprises an entire set of software/technologies with both front end (user-side) and back end (server-side) elements. The back-end aspects are based on Python scripts and Fast APIs, whereas the front end is written using the JavaScript React application. The full repository of the Bioinspire-Explore project can be found here: GitHub <https://github.com/ceebios/Bioinspire-Explore-app> (accessed on 1 January 2024).

2.3. Word Vectorization and Expansion

Bioinspire-Explore allows the user to evaluate the semantic proximity of entities ("Taxon", "Biological Process", "Habitat", and "Physical Measurement") within a body of scientific literature pertaining to bioinspiration. This supports bioinspiration by offering potential links between a taxon and its associated biological functions, environment, or physical characteristics. Bioinspire-Explore's functionality is based on a Word2Vec model [35] that was trained on tokenized sentences from the BIOMIG Corpus. The BIOMIG corpus is a dataset developed by Tchakarov et al. 2023 [36] containing scientific papers evaluated as relevant for bioinspiration.

In the Word2Vec model, entities were defined as follows:

- "Taxon" entities were taken from the GBIF taxonomic backbone [37].
- "Biological Process" entities were derived from a Wikidata list of biological processes <https://www.wikidata.org/wiki/Property:P682> (accessed on 1 January 2023), filtered manually by the authors to remove irrelevant content.
- "Habitat" entities are those listed under "biome" in the ontology ENVO [38].
- "Measurement" entities are listed as the "quantities" in Ontology of units of Measure (OM) <https://www.ebi.ac.uk/ols/ontologies/om> (accessed on 1 January 2023).

The compatibility score between entities is calculated as follows: for a given word entered by the user (e.g., "vespa"), the model looks for the N possible entries in the model that contain this word (e.g., "vespa mandarina", "vespa ducalis"). The cosine similarities between the vectors of these entries and the vectors of all the entities stored in the model (e.g., the words "resistance" or "venom") are calculated. The compatibility of each entity is defined as the minimum of these N similarity values. For each input word, N similarity values are provided, with 100% indicating "perfect compatibility" between the term chosen (in this case, "vespa") and the associated entity. Bioinspire-Explore's objective is to suggest entities that are the most compatible with all entries containing the word "vespa". The final top 20 suggestions shown on the user interface represent those entities with the highest compatibility score.

3. Results

3.1. Dynamic Taxon Search Capability

Bioinspire-Explore's first key feature is to support users in their search for a taxon of interest. As text is entered, the search bar automatically suggests up to 20 scientific names associated with a query at different taxonomic levels (i.e., species, genus, family, order, etc.). Each taxonomic rank is colour-coded, from red for species to blue for kingdom.

As seen in Figure 1A, a query beginning with the text “hymenopt” yields ten taxa suggestions at different levels. (Note that an identical request on Google may yield many more results, but these are focused solely on the order Hymenoptera, and not its taxonomic “children”.) Users can search for species using either their Latin (scientific) or common (vernacular) names. For example, the user request “mockingbird” provides eight species names identified as mockingbird taxa, as seen in Figure 1B. Table 2 shows the taxonomic results of sample user queries (e.g., “blue whale”, “mockingbird”, “nikolay”, and “escherichia”), and their explanation.

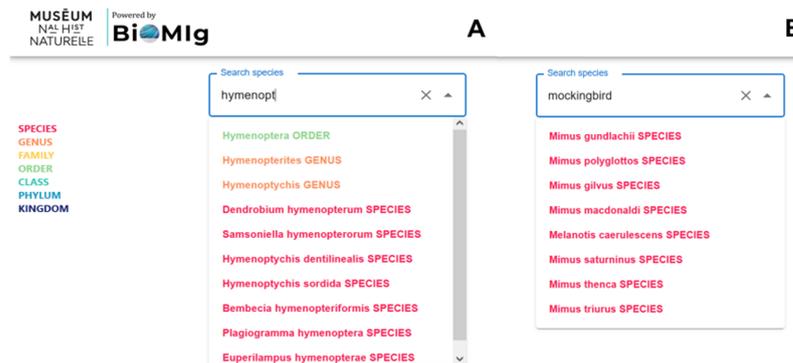


Figure 1. Search bar automatically suggests taxonomic terms associated with a query, covering both (A) scientific names and (B) vernacular names.

Table 2. Illustration of how the search bar helps the user navigate taxonomic data, depending on the type of request.

User Request	Propositions of the Research Bar	Explanation of the Result
“blue whale”	<i>Balaenoptera musculus</i> , <i>Prionace glauca</i>	two taxa associated with the vernacular name “blue whale”
“mockingbird”	<i>Mimus gundlachii</i> , <i>Mimus polyglottos</i> , <i>Mimus macdonaldi</i> , <i>Mimus gilvus</i> , <i>Melanotis caerulescens</i> , <i>Mimus saturninus</i> , <i>Mimus triurus</i> , <i>Mimus thenca</i>	eight species associated with the vernacular “mockingbird”
“nikolay”	<i>Bolinichthys nikolayi</i> , <i>Chersodromia nikolayi</i> , <i>Melamphaes nikolayi</i> , <i>Stauroneis nikolayi</i> , <i>Conterinia nikolayi</i> , <i>Prodalmanitina nikolayevi</i> , <i>Pseudepipona nikolayi</i>	seven taxa whose specific epithet contain “nikolay”
“escherichia”	<i>Escherichia</i> (Genus), <i>Escherichia hermannii</i> , <i>Escherichia vulneris</i> , <i>Escherichia fergusonii</i> , <i>Escherichia albertii</i> , <i>Escherichia marmotae</i> , <i>Escherichia ruysiae</i> , <i>Escherichia coli</i>	exact match with genus and generic names

For a given query or set of queries, an accompanying diagram is created that allows the user to locate their taxa in the “tree of life”. This dynamic taxonomic representation is based on the GBIF taxonomic backbone. For example, let us say a user is interested in exploring the taxonomy of insect species that could inspire the management of heat and light, in this case bees and ants [39,40]. Following the user’s query for the bee *Megachile kalina*, a linear taxonomic diagram is created, presenting its taxonomic “location” in Figure 2A. Subsequently, five other species of the genus *Megachile* were searched for (Figure 2B) with Bioinspire-Explore showing the taxonomic connections between these “new” species and the “original” species *Megachile kalina*. The user then enters a new query, this time for the species of desert ant *Cataglyphis bombycina*. Bioinspire-Explore “remembers” the previous searches (unless the clear button is used in between queries) and adds this species’ taxonomy, leading to a bifurcation in the diagram seen in Figure 2C. These taxonomic diagrams are exportable as a PNG files that the user can then download.

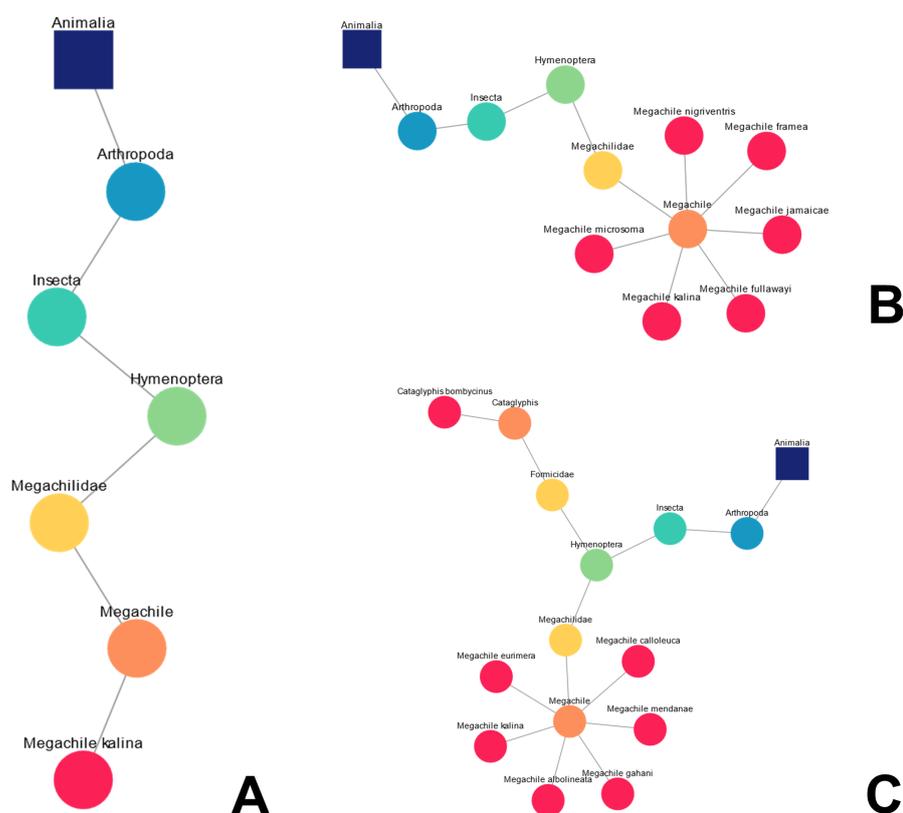


Figure 2. Examples of dynamic taxonomic diagrams following multiple queries associated with bees. (A) Linear diagram created after single search for the bee species *Megachile kalina*; (B) expansion of the “taxonomic children” of the genus *Megachile*; (C) subsequent search for the desert ant species *Cataglyphis bombycinus*.

3.2. Biological and Environmental Data

A great deal of biological and environmental data are freely available online. However, at present, exploring this information involves going back and forth between different websites, which can be time consuming and distracting. Bioinspire-Explore has been designed as a “one stop shop” connecting multiple external open-access data sources via a taxon’s name.

Figure 3 shows the results of a query for *Arnica montana*, a flowering plant species found in western Europe. The first window displays the first paragraph and first image of the Wikipedia page associated with Wikidata node for the species Q207848 (<https://www.wikidata.org/wiki/Q207848> (accessed on 1 January 2023)). If the Wikidata graph is missing a node for the taxon in question, this window will be empty. An external link will then automatically search for the taxon name in Wikipedia to provide associated data for this window. The second window, seen in Figure 3B, shows occurrence data for *Arnica montana*, plotted on a world map (with an option to zoom in and out, depending on the scale of interest). Figure 3C displays iNaturalist images tagged with *Arnica montana*. By clicking on an image, the user will be led to the iNaturalist website. Figure 3D shows the average climatic conditions, as well as minimums and maximums, for *Arnica montana*, as found under the fourth window. Windows 3 and 4 aim to provide information on a taxon’s environmental context, highlighting the abiotic conditions to which it is adapted (temperature, rainfall, sun exposure, wind speed, etc.).

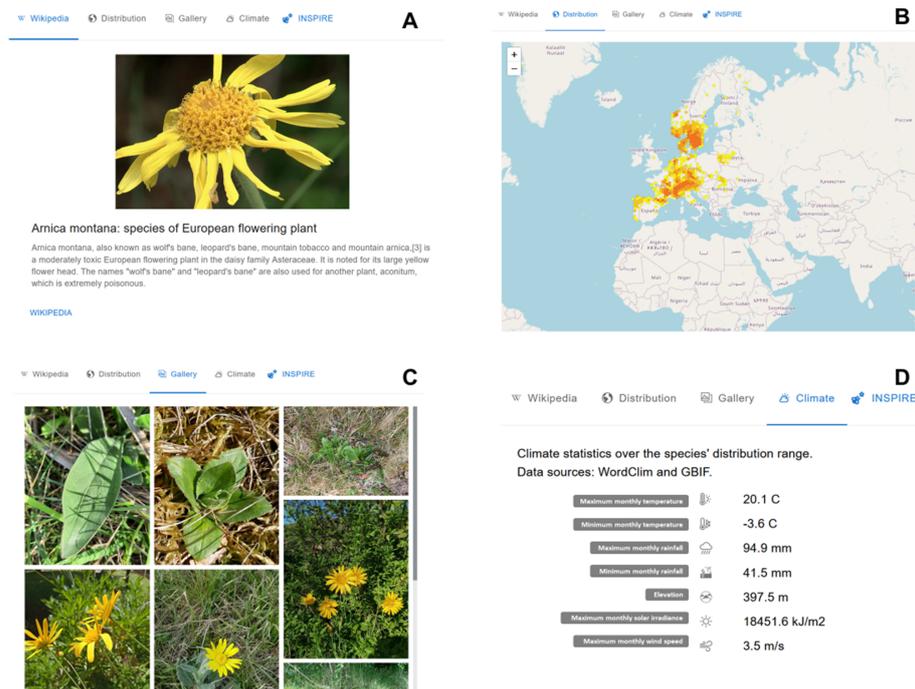


Figure 3. Four main types of information are provided through Bioinspire-Explore for a selected taxon *Arnica montana*. (A) Its Wikipedia page, (B) spatial distribution, (C) images, and (D) climate characteristics.

3.3. Semantic Proximity between Biological Entities

As outlined, Tchakarov et al. 2023 [36] developed a neural network model Word2Vec [35] based on a corpus of scientific articles judged relevant for bioinspiration (the “BIOMIG Corpus”). Using this model, Bioinspire-Explore is able to assess the semantic proximity of text entities (taxon, biological process, habitat, physical measurements) stored as word “embeddings” within this scientific corpus. These entities were chosen as being the most relevant in the context of a “biology push” approach to bioinspiration, providing users with ideas in terms biological processes, functions, structures, and systems of potential interest for innovation (see examples in Table 3).

Table 3. Examples of entities related to two taxa queries *Vespa soror* and *Morpho menelaus*.

User Request	Taxon	Process	Habitat	Measure
<i>Vespa soror</i>	<i>Platygastridae</i> , <i>Argidae</i> , <i>Theroa zethus</i>	hunting, swarm, autothysis	plant, tropical, meadow	heat, frequency, length
<i>Morpho menelaus</i>	<i>Greta oto</i> , <i>Papilio ulysses</i> , <i>Callophrys rubi</i>	polyphenism, colouration, flying	feather, desert, meadow	angle, heat, hydrophobicity

Figure 4A presents a more detailed example of how these text embeddings, and the evaluation of their semantic proximity with the scientific corpus, work in practice. Here, seven text entities are represented as vectors in a fictive 3D-space: three taxon names (“*Pica pica*”, “*Corvus corone*”, and “*Vespa soror*”); two biological processes (“diving” and “swarming”); and two habitat characteristics (“forest” and “tropical”).

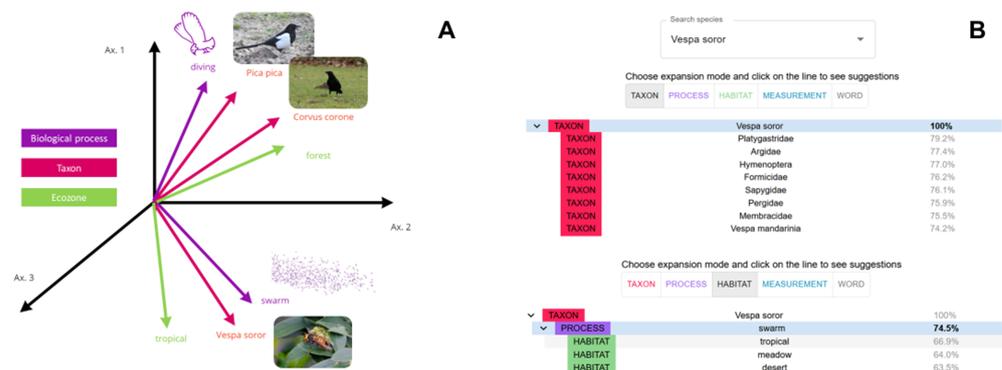


Figure 4. (A) Schematic illustration of semantic proximity of various types of entities “taxon”, “ecozone/habitat”, “biological process”; (B) screenshot of the semantic expansion module in Bioinspire-Explore. Here, the taxon entity *Vespa soror* appears semantically related to other taxa such as Hymenoptera.

The taxon entities *Pica pica* and *Corvus corone* are represented as close vectors, because both these taxa belonging to the Corvidae family. The Asian wasp *Vespa soror* is represented as a vector with more distant coordinates from the two Corvidae, as their phylogenetic relationship is further apart. The biological process “swarm” is shown to be semantically closer to the *Vespa soror* than to the two Corvidae. This is because swarming is a behaviour typically associated with this wasp (rather than Corvidae), and the scientific literature reflects this. Conversely, “diving” is more likely to be found in Corvidae than *Vespa soror*, and therefore appears closer to the two birds. In terms of habitat characteristics, the entity “tropical” is directionally closer to the text embedding of *Vespa soror* than to the Corvidae, as the latter live in temperate climates of western Europe.

How are the results of this neural network model presented within Bioinspire-Explore? In Figure 4B, we show two screenshots of the “Biodiversity” page within Bioinspire-Explore, where the concept of semantic proximity between taxa, biological processes, habitats, or measurements is explored. In the first screenshot, we explore entities related to our wasp *Vespa soror*. We start with the “taxon” tab to understand how this wasp is connected to other taxa within the scientific literature. As to be expected, the model indicates greatest semantic proximity with other insect families of the order Hymenoptera, such as Platygasteridae, Sapygidae, Pergidae, or Argidae. By clicking on the “Process” tab, we find the biological process “swarm” with a 74.5% semantic proximity to *Vespa soror*. Selecting this process, we are interested in understanding the proximity of the “Habitat” entities connected to both *Vespa soror* and swarming behaviour. This is shown in the second screenshot of Figure 4B. Having been asked for habitats semantically close to both *Vespa soror* and swarming behaviour, Bioinspire-Explore cites “tropical”, “meadow”, and “desert” as related entities. The compatibility scores between this trio of habitat entities are 67%, 64%, and 63.5%, respectively. This means that, in the scientific literature (BIOMIG Corpus) used by the Bioinspire-Explore model, connections between swarming (process) of *Vespa soror* (taxon) in tropical areas (habitat) appear to be the most significant, with a combined score of 67%.

3.4. An Example “User Journey” in Bioinspire-Explore

Figure 5 presents how a user might explore the Bioinspire-Explore tool. If they are new to the concept of bioinspiration, they may start by visiting the “Bioinspirations” page (Figure 5A) that presents a handful of existing initiatives in the field of bioinspired design (using images, videos, text, and references). Via this page, they might read about the navigational abilities of the desert ant *Cataglyphis fortis*. They decide they would like to learn more about this species via the “Biodiversity” page (being redirected automatically via the species name, or by typing its name directly into the search box). Figure 5B shows the information that is provided at this stage (as already described in Figure 2): the species taxonomic position, its appearance, its geographical distribution, and climatic niche.

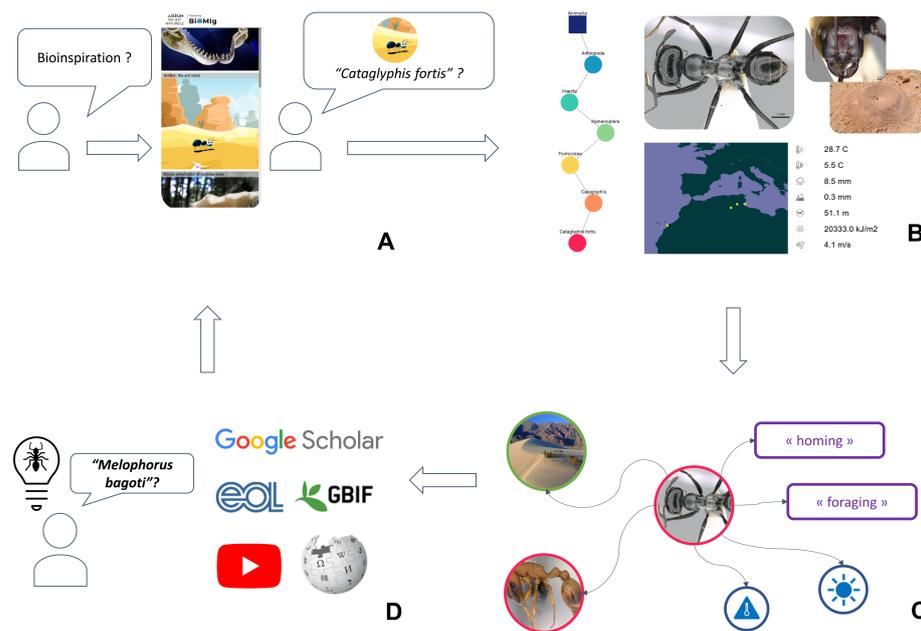


Figure 5. User journey within Bioinspire-Explore, illustrating the functionalities of different tabs within the tool: (A) visit of the “Bioinspirations” page presenting bioinspiration examples; (B) then moving to the “Biodiversity” page to learn more on a given taxon; (C) the “Go Further” page gives an overview of closely-related entities within the scientific literature; and (D) end of the user journey by either doing another research or using the pre-filled external links.

Finally, the user may wish to dive deeper via the “Go further” page in Bioinspire-Explore (Figure 5C), getting an overview of closely related entities within the scientific literature associated with *Cataglyphis fortis*. These insights usually lead the user either (i) to further research on the taxon in question (using other platforms (EOL, GBIF...)) or via the scientific articles presented in the Biomig Corpus) or (ii) to a new request about another taxon presenting a similar interest (for example, the user might identify the desert ant *Melophorus bagoti* as an alternative additional biological model for navigation).

4. Discussion

4.1. Bioinspire-Explore’s Audience

Bioinspire-Explore is an innovative new tool, providing a unique way to explore biodiversity data and visualise relationships between taxa as well as between entities of relevance to bioinspiration. Bioinspire-Explore’s focus is on a “biology-push” approach to bioinspiration, that is to say, the identification of biological systems of potential interest, the abstraction of their associated functions, followed by their transposition to a technical domain. However, Bioinspire-Explore could also be used to support of a “technology-pull” approach, as described by Fayemi et al. [41], i.e., once a team has transposed technological needs into biological functions, and is searching for biological models that fulfill those functions (See Figure 6).

The authors did not carry out a formalized process to generate scientifically robust user feedback for the first iteration of the Bioinspire-Explore tool. Bioinspire-Explore was, however, thoroughly tested with users from different institutions and communities on five separate occasions (three Masters courses and two scientific workshops) and received constructive feedback that fed into the project, either into the tool itself or this article. Future feedback will be collected via a dedicated email address.

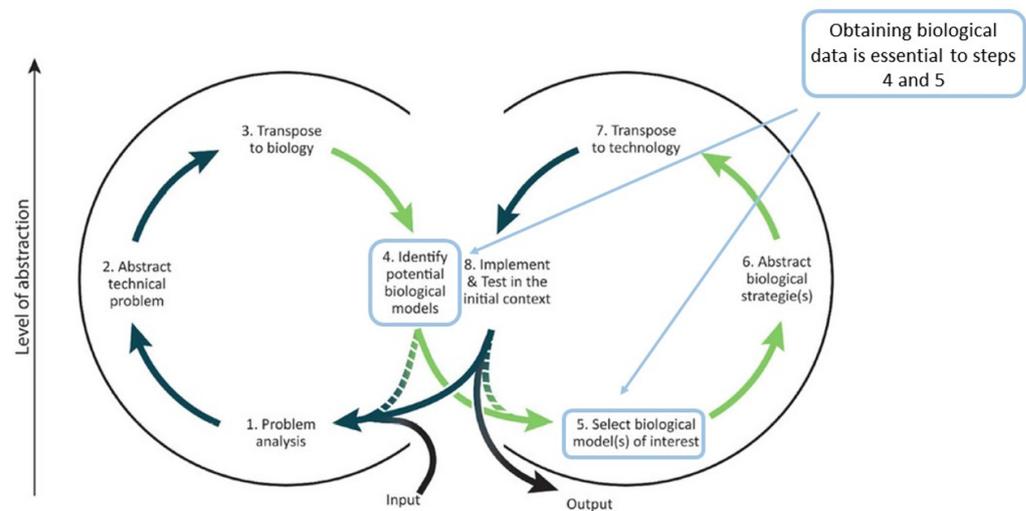


Figure 6. Possible uses Bioinspire-Explore within a biomimetic methodological framework. After Fayemi et al. (2017) [41].

Moreover, although designed to help bioinspiration practitioners, Bioinspire-Explore could also serve as an educational tool in many different settings. Not only does the tool facilitate access to biological data, it offers an interactive interface to “play” with taxonomic information and phylogenetic relationships. Over and above finding relevant data for a bioinspired project, Bioinspire-Explore helps users understand how life on earth is organised scientifically, putting bioinspiration into its biological context. Thus, Bioinspire-Explore could be used as a way to tackle both difficulties in understanding biology as well challenges associated with the transfer of knowledge from biology to technology, as identified by Wanieck et al. [42]: both obstacles within the bioinspiration field.

4.2. Challenges and Next Steps

Bioinspire-Explore represents a first iteration in the development of an online “biology push” bioinspiration tool. As a prototype, it still faces challenges in terms of data sources and functionality. With regard to data sources, Bioinspire-Explore provides access to international biodiversity data via the most well recognised open access source: the Global Biodiversity Information Facility (GBIF). Despite the significant advantages associated with facilitating access to this vast dataset, Bioinspire-Explore’s connections with GBIF mean that its limitations are interwoven as well. In other words, Bioinspire-Explore’s data contain the same biases as GBIF’s, and the same challenges in terms of staying up to date with constant phylogenetic and taxonomic developments [43–46]. For example, if a species’ scientific name has recently changed, Bioinspire-Explore will present only those data relating to the new species’ name and not its previous name. This means that all data associated with its previous name are no longer available to users. The reverse also holds true, if taxonomic updates have not yet been made in GBIF, Bioinspire-Explore will relay nomenclature inconsistencies. Nevertheless, efforts are underway to improve the quality of GBIF data [47,48], and this will directly benefit the Bioinspire-Explore tool and its taxonomic accuracy.

Moreover, not all geographical areas covered by GBIF have the same density of data points [49,50]; nor do all taxonomic groups have equivalent amounts of data on their geographical distribution [20]. This affects the accuracy of climatic ranges of certain taxa calculated within Bioinspire-Explore, for example.

In terms of the bioinspiration literature that informs the analyses of semantic proximity metrics, the “BIOMIG corpus” is updated every three months with new literature relevant to bioinspiration. Although the BIOMIG corpus has been developed to include all pertinent scientific articles (even when they do not mention bioinspiration or biomimetics explicitly), this corpus cannot claim to be exhaustive: there will inevitably be published scientific literature useful to bioinspiration that is not automatically identified as such. Despite the regular updating of the BIOMIG corpus (every three months), the Word2Vec model, trained from this corpus, is static (and does not change). This is a deliberate choice of the development team. Indeed, the point of Word2Vec is to unravel the semantic meaning of entities (words), and the relationship between these entities, using a large number of sentences. Considering the extent of the BIOMIG corpus (several decades worth of scientific papers), a regular re-training of Word2Vec on an updated corpus would not provide significant changes in these semantic relationships between words. For example, the entity “*Morpho menelaus*” is currently related to “reflection” or to other butterfly taxa; in a year, the small amount of papers related to *Morpho* will not be a sufficient “signal” to modify this semantic proximity. That being said, we should indeed retrain the Word2Vec model every few years, or could even replace this semantic expansion tool with more costly Large Language models in the future.

Our intention is to continue to improve Bioinspire-Explore’s functionality and scope in the future, based on feedback following its use in educational and professional settings. Several possible avenues for its improvement already exist. From a functionality perspective, building the tool’s capacity for non-taxon entry points is one priority. Presently, Bioinspire-Explore’s capability lies chiefly in the assessment of semantic proximity between a taxon (as the starting point) and other entities (biological processes, physical measures, and habitats) in the scientific literature. For now, starting a search with one of these other entities provides less obvious results in terms of connections in the scientific literature. For example, if one were to begin a search with “energy” as a physical measurement, and then “photo-protection” as a biological process, the results in terms of semantic proximity linking these entities to specific taxa are weak. This currently limits the ability of a bioinspired designer to pursue “property-led” rather than “taxon-led” research.

In terms of climate data, Bioinspire-Explore uses the last available “snapshot” of the WorldClim database, issued in 2020, and covering the 1970 to 2000 period of time. We are aware of the limitations of this database’s scope and the importance of having up-to-date relationships between taxa and their precise climatic conditions. This aspect could be improved in future iterations of the tool.

In terms of the potential integration of new datasets into Bioinspire-Explore, two types of information stand out: morphological and palaeontological. Morphological descriptions are those used to identify living organisms, presenting their specific structures and characteristics. These properties are closely linked to an organism’s function, and therefore of potential interest to bioinspiration practitioners. Bioinspire-Explore could be expanded to include morphological (and anatomical) knowledge extracted from relevant scientific literature via AI text analysis methods. This information could be presented in the form of graphs or RDF triplets, for example, as proposed by Sahraoui et al. [51].

Finally, palaeo-bioinspiration is a growing field [52], and one that requires access to information about species from previous geological epochs. The biodiversity data available through Bioinspire-Explore are currently based on a classification system that does not fully take into account the fossil record. Although international, open-access palaeobiodiversity databases exist [53], their utilisation would require the harmonious integration of the classification of fossils with that of current species. Presently, integrating the multiple branches of fossil lineages at higher taxonomic levels pose both theoretical and technical challenges; these difficulties would need to be evaluated and resolved before a unified catalogue of past and present biodiversity could be made available through Bioinspire-Explore. The tool is currently being tested by Masters and PhD students as part of their bioinspired design curricula, as well as professionals within the bioinspiration domain.

Their user feedback will provide insight on the functional, visual, and ergonomic choices made by Bioinspire-Explore's development team. Particular attention will be given to improving the semantic expansion tool to assess whether it is understood, intuitive, and useful in the application of bioinspiration.

5. Conclusions

In a context of growing environmental challenges, professionals are increasingly looking to the living world to provide inspiration for more sustainable alternatives to today's systems. Bioinspire-Explore's principle audience is this new generation of designers, engineers, chemists, architects, urban planners, and more. It can be operated by a single user, but can also facilitate interaction between different disciplines in a group setting. Crucially, Bioinspire-Explore is intended to act as a guide, not as a replacement for the active involvement of biologists in bioinspiration projects [31,54]. Rather, it orientates the user towards promising information regarding living systems of interest and presents those systems in their scientific context. It offers relevant scientific literature to bioinspiration practitioners seeking stimulus for many different variants of bioinspiration [34]: for example, exploring properties of biological materials to support biofabrication/bioproduction engineering initiatives; or identifying ecological characteristics of taxa that could be integrated into nature based solutions. Bioinspire-Explore is a powerful tool for browsing and visualizing biological data, using taxonomy as an entry point. The platform's user-friendly interface and comprehensive databases make it an accessible way to explore biological data, particularly as part of a "biology push" approach to bioinspiration, and provide meaningful insights into the diversity and evolution of life on earth. Its unique value lies in its ability to combine, within a single tool (i) taxonomic data; (ii) ecological information; and (iii) indications of the semantic proximity of terms ("entities") cited in bioinspiration-related literature. In the future, we anticipate that Bioinspire-Explore will continue to evolve and improve, with the addition of new features, functionality, and datasets in order to respond to the growing ambitions of the bioinspiration community.

Author Contributions: Conceptualization, A.S.-S., A.A., R.V.-L., J.-S.S., T.B., L.-M.P. and N.T.; methodology, L.-M.P., A.S.-S. and A.A.; software, N.T.; validation, A.A., A.S.-S., J.-S.S., T.B. and R.V.-L.; formal analysis, N.T. and A.S.-S.; investigation, all authors; resources, T.B. and R.V.-L.; data curation, N.T.; writing—original draft preparation, A.S.-S., A.A. and R.V.-L.; writing—review and editing, A.S.-S. and A.A.; visualization, A.S.-S., A.A. and N.T.; supervision, J.-S.S., A.S.-S. and A.A.; project administration, A.A.; funding acquisition, A.S.-S., A.A., R.V.-L., J.-S.S. and T.B. All authors have read and agreed to the published version of the manuscript.

Funding: We hereby acknowledge the financial support of the French Government's "France Relance" programme, which financed the creation of Bioinspire-Explore via the ITN5 "12-363-DNUM-CENS-0030" project coordinated by the French Inter-ministerial Directorate for Digital Technology. This project is co-financed by the European Union as part of the NEXT GEN EU plan.

Data Availability Statement: The data presented in this study are openly available in GitHub at <https://github.com/ceebios/Bioinspire-Explore-app> (accessed on 1 January 2024).

Acknowledgments: We sincerely thank Chrystelle Roger from Myceco for her vision, time, and invaluable advice on setting up and running the ITN5 project. Our thanks also go to Théo Alves DaCosta and the "Data for Good" Association for their contribution to Bioinspire-Explore's first iteration. The BIOMIG corpus of scientific literature was kindly provided by the BIOMIG Consortium and CEEBIOS. We would also like to thank those who tested Bioinspire-Explore during its development, including the "Nature Inspired Design" Masters students from ENSCI-Les Ateliers, Paris. The Bioinspire-Explore project won an innovation purchasing award from the French State's Purchasing Division in 2023. We thank our colleagues at the MNHN for supporting our submission for this award and are grateful for the national recognition it confers.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. ISO 18458:2015; Biomimetics: Terminology, Concepts and Methodology. ISO: Geneva, Switzerland, 2015.
2. Shao, H.; Bachus, K.N.; Stewart, R.J. A Water-Borne Adhesive Modeled after the Sandcastle Glue of *P. californica*. *Macromol. Biosci.* **2009**, *9*, 464–471. [[CrossRef](#)] [[PubMed](#)]
3. Jones, J.; Adamatzky, A. Computation of the travelling salesman problem by a shrinking blob. *Nat. Comput.* **2014**, *13*, 1–16. [[CrossRef](#)]
4. Dupeyroux, J.; Serres, J.R.; Viollet, S. AntBot: A six-legged walking robot able to home like desert ants in outdoor environments. *Sci. Robot.* **2019**, *4*, eaau0307. [[CrossRef](#)] [[PubMed](#)]
5. Cruz, E. Biomimétisme: Le vivant, une bibliothèque pour l'innovation. *Technica* **2018**, *1*, 15–22.
6. Boucle, S.P.; Boutin, C. Compte-Rendu Commenté de la Visite de la ZRV de St Just (34) (Zone Libellule). HAL. 2015. Available online: <https://agris.fao.org/search/en/providers/122439/records/64747234425ec3c088f288ee> (accessed on 1 January 2024).
7. Mthunzi, S.N.; Benkhelifa, E.; Bosakowski, T.; Hariri, S. A Bio-inspired Approach to Cyber Security. In *Machine Learning for Computer and Cyber Security*, 1st ed.; A Science Publishers Book; Gupta, B.B., Sheng, M., Eds.; CRC Press: Boca Raton, FL, USA; Taylor & Francis Group: Abingdon, UK, 2019; pp. 75–104. [[CrossRef](#)]
8. Lechner, M.; Hasani, R.; Amini, A.; Henzinger, T.A.; Rus, D.; Grosu, R. Neural circuit policies enabling auditable autonomy. *Nat. Mach. Intell.* **2020**, *2*, 642–652. [[CrossRef](#)]
9. Vasileiou, T.; Summerer, L. A biomimetic approach to shielding from ionizing radiation: The case of melanized fungi. *PLoS ONE* **2020**, *15*, e0229921. [[CrossRef](#)]
10. Ellery, A. Tutorial Review of Bio-Inspired Approaches to Robotic Manipulation for Space Debris Salvage. *Biomimetics* **2020**, *5*, 19. [[CrossRef](#)]
11. Xu, B.; Yang, Y.; Zhang, B. Bionics design and dynamics analysis of space webs based on spider predation. *Acta Astronaut.* **2019**, *159*, 294–307. [[CrossRef](#)]
12. Müller, R.; Abaid, N.; Boreyko, J.B.; Fowlkes, C.; Goel, A.K.; Grimm, C.; Jung, S.; Kennedy, B.; Murphy, C.; Cushing, N.D.; et al. Biodiversifying bioinspiration. *Bioinspir. Biomimetics* **2018**, *13*, 053001. [[CrossRef](#)]
13. Hayes, S.; Desha, C.; Baumeister, D. Learning from nature—Biomimicry innovation to support infrastructure sustainability and resilience. *Technol. Forecast. Soc. Chang.* **2020**, *161*, 120287. [[CrossRef](#)]
14. Bánki, O.; Roskov, Y.; Döring, M.; Ower, G.; Hernández Robles, D.R.; Plata Corredor, C.A.; Stjernegaard Jeppesen, T.; Örn, A.; Vandepitte, L.; Hobern, D.; et al. *Catalogue of Life Checklist*; Leiden, The Netherlands, 2023. Available online: <https://biss.pensoft.net/article/111684/> (accessed on 1 January 2024).
15. Mitov, M. Cholesteric liquid crystals in living matter. *Soft Matter* **2017**, *13*, 4176–4209. [[CrossRef](#)] [[PubMed](#)]
16. Sivasankaran, P.N.; Ward, T.A.; Viyapuri, R.; Johan, M.R. Static strength analysis of dragonfly inspired wings for biomimetic micro aerial vehicles. *Chin. J. Aeronaut.* **2016**, *29*, 411–423. [[CrossRef](#)]
17. Audibert, C.; Chaves-Jacob, J.; Linares, J.M.; Lopez, Q.A. Bio-inspired method based on bone architecture to optimize the structure of mechanical workpieces. *Mater. Des.* **2018**, *160*, 708–717. [[CrossRef](#)]
18. Graeff, E.; Maranzana, N.; Aoussat, A. Biological practices and fields, the missing pieces of the biomimetics' methodological puzzle. *Biomimetics* **2020**, *5*, 62. [[CrossRef](#)] [[PubMed](#)]
19. Kim, S.J.; Lee, J.H. A study on metadata structure and recommenders of biological systems to support bio-inspired design. *Eng. Appl. Artif. Intell.* **2017**, *57*, 16–41. [[CrossRef](#)]
20. Troudet, J.; Grandcolas, P.; Blin, A.; Vignes-Lebbe, R.; Legendre, F. Taxonomic bias in biodiversity data and societal preferences. *Sci. Rep.* **2017**, *7*, 9132. [[CrossRef](#)] [[PubMed](#)]
21. Cruz, E.; Hubert, T.; Chancoco, G.; Naim, O.; Chayaamor-Heil, N.; Cornette, F.; Menezes, C.; Badarnah, L.; Raskin, K. Design processes and multi-regulation of biomimetic building skins: A comparative analysis. *Energy Build.* **2021**, *246*, 111034. [[CrossRef](#)]
22. Deldin, J.; Schuknecht, M. The AskNature Database: Enabling Solutions in Biomimetic Design. In *Biologically Inspired Design*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 17–27.
23. Shyam, V.; Friend, L.; Whiteaker, B.; Bense, N.; Dowdall, J.; Boktor, B.; Johnny, M.; Reyes, I.; Naser, A.; Sakhamuri, N.; et al. PeTaL (Periodic Table of Life) and Physiomimetics. *Designs* **2019**, *3*, 43. [[CrossRef](#)]
24. Vincent, J.F. The trade-off: A central concept for biomimetics. *Bioinspired Biomim. Nanobiomater.* **2017**, *6*, 67–76. [[CrossRef](#)]
25. Kruiper, R. Trade-offs in Computer-aided Biomimetics. In *Engineering Conferences International Proceedings*; 2019. Available online: https://dc.engconfintl.org/nature_inspired/10/ (accessed on 1 January 2024).
26. Chiu, I.; Shu, L.H. Bridging Cross-Domain Terminology for Biomimetic Design. In Proceedings of the Volume 5a: 17th International Conference on Design Theory and Methodology, Long Beach, CA, USA, 24–28 September 2005; Volume 2005, pp. 93–101. [[CrossRef](#)]
27. Chakrabarti, A.; Sarkar, P.; Leelavathamma, B.; Nataraju, B. A functional representation for aiding biomimetic and artificial inspiration of new ideas. *AI EDAM* **2005**, *19*, 113–132. [[CrossRef](#)]
28. Nagel, J.K.S.; Stone, R.B.; McAdams, D.A. An Engineering-to-Biology Thesaurus for Engineering Design. In Proceedings of the Volume 5: 22nd International Conference on Design Theory and Methodology, Special Conference on Mechanical Vibration and Noise, Montreal, QC, Canada, 15–18 August 2010; pp. 117–128. [[CrossRef](#)]

29. McInerney, S.; Khakipoor, B.; Garner, A.; Houette, T.; Unsworth, C.; Rupp, A.; Weiner, N.; Vincent, J.; Nagel, J.; Niewiarowski, P. E2BMO: Facilitating User Interaction with a BioMimetic Ontology via Semantic Translation and Interface Design. *Designs* **2018**, *2*, 53. [CrossRef]
30. Troudet, J.; Vignes-Lebbe, R.; Grandcolas, P.; Legendre, F. The Increasing Disconnection of Primary Biodiversity Data from Specimens: How Does It Happen and How to Handle It? *Syst. Biol.* **2018**, *67*, 1110–1119. [CrossRef] [PubMed]
31. Lecointre, G.; Aish, A.; Améziane, N.; Chekchak, T.; Goupil, C.; Grandcolas, P.; Vincent, J.F.V.; Sun, J.S. Revisiting Nature's "Unifying Patterns": A Biological Appraisal. *Biomimetics* **2023**, *8*, 362. [CrossRef] [PubMed]
32. Graeff, E. Biomimetics from practical feedback to an interdisciplinary process. *Res. Eng. Des.* **2021**, *32*, 349–375. [CrossRef]
33. Letard, A.; Maranzana, N.; Brun, J.; Raskin, K.; Aoussat, A. Biomimetics practice in industrial context: Methodological parameters to promote interdisciplinary collaboration and knowledge transfer. *Creat. Innov. Manag.* **2024**, unpublished, under revision.
34. Aish, A.; Sun, J.S. Bioinspire-Museum: Scoping Paper. Muséum National d'Histoire Naturelle, Paris, 24 Pages. 2020. Available online: https://www.mnhn.fr/sites/mnhn.fr/files/atoms/files/bioinspire_museum_scoping_paper.pdf (accessed on 1 January 2023).
35. Mikolov, T.; Yih, W.t.; Zweig, G. Linguistic Regularities in Continuous Space Word Representations. In Proceedings of the NAACL, Atlanta, GA, USA, 9–14 June 2013.
36. Tchakarov, N.; Racca, L.; Peybernes, T.; Saint-Sardos, A. A Scientific Corpus and Search Engine for Biomimetics. *SSRN Electron. J.* **2023**. [CrossRef]
37. Secretariat, G. GBIF Backbone Taxonomy. Checklist Dataset. 2022. Available online: <https://www.gbif.org/dataset/d7d4dbf4-2cf0-4f39-9b2a-bb099caae36c> (accessed on 1 January 2023).
38. Buttigieg, P.; Morrison, N.; Smith, B.; Mungall, C.J.; Lewis, S.E.; The ENVO Consortium. The environment ontology: Contextualising biological and biomedical entities. *J. Biomed. Semant.* **2013**, *4*, 43. [CrossRef]
39. Willot, Q.; Simonis, P.; Vigneron, J.P.; Aron, S. Total Internal Reflection Accounts for the Bright Color of the Saharan Silver Ant. *PLoS ONE* **2016**, *11*, e0152325. [CrossRef]
40. Wang, Y.; Shang, S.; Chiu, K.L.; Jiang, S. Mimicking Saharan silver ant's hair: A bionic solar heat shielding architextile with hexagonal ZnO microrods coating. *Mater. Lett.* **2020**, *261*, 127013. [CrossRef]
41. Fayemi, P.E.; Wanieck, K.; Zollfrank, C.; Maranzana, N.; Aoussat, A. Biomimetics: Process, tools and practice. *Bioinspir. Biomimetics* **2017**, *12*, 011002. [CrossRef]
42. Wanieck, K.; Ritzinger, D.; Zollfrank, C.; Jacobs, S. Biomimetics: Teaching the tools of the trade. *FEBS Open Bio.* **2020**, *10*, 2250–2267. [CrossRef] [PubMed]
43. Meyer, C.; Weigelt, P.; Kreft, H. Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecol. Lett.* **2016**, *19*, 992–1006. [CrossRef] [PubMed]
44. Waller, J. Finding Data Gaps in the GBIF Backbone Taxonomy. *Biodivers. Inf. Sci. Stand.* **2022**, *6*, e91312. [CrossRef]
45. Smith, B.E.; Johnston, M.K.; Lücking, R. From GenBank to GBIF: Phylogeny-Based Predictive Niche Modeling Tests Accuracy of Taxonomic Identifications in Large Occurrence Data Repositories. *PLoS ONE* **2016**, *11*, e0151232. [CrossRef] [PubMed]
46. Bourgoin, T.; Bailly, N.; Zaragueta, R.; Vignes-Lebbe, R. Complete formalization of taxa with their names, contents and descriptions improves taxonomic databases and access to the taxonomic knowledge they support. *Syst. Biodivers.* **2021**, *19*, 738–746. [CrossRef]
47. Zizka, A.; Antunes Carvalho, F.; Calvente, A.; Rocio Baez-Lizarazo, M.; Cabral, A.; Coelho, J.F.R.; Colli-Silva, M.; Fantinati, M.R.; Fernandes, M.F.; Ferreira-Araújo, T.; et al. No one-size-fits-all solution to clean GBIF. *PeerJ* **2020**, *8*, e9916. [CrossRef]
48. Secretariat, G. GBIF to Serve as Administrative Host for Species 2000 Secretariat. 2024. Available online: <https://www.gbif.org/news/5LO3xYHTbo53rB8UwCSpG/gbif-to-serve-as-administrative-host-for-species-2000-secretariat> (accessed on 16 January 2024).
49. Beck, J.; Böller, M.; Erhardt, A.; Schwanghart, W. Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecol. Inform.* **2014**, *19*, 10–15. [CrossRef]
50. García-Roselló, E.; González-Dacosta, J.; Lobo, J.M. The biased distribution of existing information on biodiversity hinders its use in conservation, and we need an integrative approach to act urgently. *Biol. Conserv.* **2023**, *283*, 110118. [CrossRef]
51. Sahraoui, M.; Pignal, M.; Vignes Lebbe, R.; Guigue, V. NEARSIDE: Structured Knowledge Extraction from Species Descriptions. *Biodivers. Inf. Sci. Stand.* **2022**, *6*, e94297. [CrossRef]
52. Aish, A. Palaeo-Bioinspiration: Drawing on the Fossil Record to Advance Innovation. 2024, *in press*. Available online: <https://palaeo-bioinsp.sciencesconf.org/> (accessed on 1 January 2024).
53. Alroy, J. The Paleobiology Database. 2021. Available online: <https://paleobiodb.org/#/> (accessed on 1 January 2024).
54. Graeff, E.; Maranzana, N.; Aoussat, A. Engineers' and Biologists' Roles during Biomimetic Design Processes, towards a Methodological Symbiosis. *Proc. Des. Soc. Int. Conf. Eng. Des.* **2019**, *1*, 319–328. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.