



A knowledge graph-based bio-inspired design approach for knowledge retrieval and reasoning

Liuqing Chen, Zebin Cai, Zhaojun Jiang, Lingyun Sun, Peter Childs & Haoyu Zuo

To cite this article: Liuqing Chen, Zebin Cai, Zhaojun Jiang, Lingyun Sun, Peter Childs & Haoyu Zuo (2025) A knowledge graph-based bio-inspired design approach for knowledge retrieval and reasoning, Journal of Engineering Design, 36:7-9, 1321-1351, DOI: [10.1080/09544828.2024.2311065](https://doi.org/10.1080/09544828.2024.2311065)

To link to this article: <https://doi.org/10.1080/09544828.2024.2311065>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 31 Jan 2024.



Submit your article to this journal [↗](#)



Article views: 2457



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 10 View citing articles [↗](#)

A knowledge graph-based bio-inspired design approach for knowledge retrieval and reasoning

Liuqing Chen^a, Zebin Cai^{ib}, Zhaojun Jiang^a, Lingyun Sun^a, Peter Childs^b and Haoyu Zuo^b

^aCollege of Computer Science and Technology, Zhejiang University, Hangzhou, People's Republic of China;

^bDyson School of Design Engineering, Imperial College London, London, UK

ABSTRACT

Bio-inspired Design (BID) is a method that draws principles from biological systems to solve complex real-world problems. While diverse knowledge-based tools have served BID, the retrieval and reasoning capabilities of knowledge graphs have not been explored in BID. This study introduces a novel knowledge graph-based BID approach, exploiting the power of knowledge graphs to support BID. In the approach, a comprehensive ontology is defined and then applied to construct a BID-specific knowledge graph, enabling efficient representation of the diverse and rich biological knowledge. The knowledge graph supports BID by facilitating knowledge retrieval and reasoning. Retrieval in BID is accomplished by finding potential links between biological systems and relevant design applications. Reasoning in BID is supported by a link prediction model that follows the design process of mapping from biological systems to design applications. Two case studies are conducted to demonstrate the effectiveness of the approach. The first case shows that our approach outperforms other benchmarks in retrieving related biological knowledge, and the second case presents how the link prediction model aids in generating relevant and inspirational design ideas.

ARTICLE HISTORY

Received 30 June 2023

Accepted 24 January 2024

KEYWORDS

Bio-inspired design;
knowledge graph;
knowledge retrieval;
knowledge reasoning; link
prediction

1. Introduction

Bio-inspired Design (BID) is a method that transfers principles, functions, and strategies from biological systems to solve complex real-world problems (Müller et al. 2018). The BID process typically comprises four key stages: problem definition, biological exploration, biological abstraction, and solution application (Vincent et al. 2006). The problem definition and biological exploration involve initially identifying and defining a design problem, then retrieving and exploring relevant biological analogies that could potentially provide solutions. The inherent principles and strategies of these biological systems are then abstracted, involving deep reasoning to generalise biological knowledge. This abstracted biological knowledge is then translated into a design solution to solve the original problem. This

CONTACT Haoyu Zuo  hz2019@ic.ac.uk  Dyson School of Design Engineering, Imperial College, London SW7 2AZ, UK

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

process not only prompts innovative design solutions but also drives creativity and innovation. BID has seen practical use in various fields (Bar-Cohen 2005), demonstrating the wide applicability of biological principles (Helms, Vattam, and Goel 2009).

Building upon the BID process, various knowledge-based tools contribute to this domain (BionicInspiration.org 2014; Chen et al. 2023; Deldin and Schuknecht 2013; Goel et al. 2012; Goel et al. 2022; Goel, McAdams, and Stone 2015; Kozaki and Mizoguchi 2014; Sidharth and Chakrabarti 2018). These tools organise and categorise BID knowledge using diverse ontologies and employ representation formats such as narrative texts, structured databases, or semantic networks to accommodate varying representation needs. These tools offer two main functionalities: retrieval and reasoning. The retrieval function helps users find relevant bio-inspired knowledge effectively. For reasoning, these tools support users in understanding and inferring complex relationships between biological systems and design applications. While these tools have made contributions to BID, there is space for further exploration and development of advanced knowledge-based tools to augment these functionalities.

Another knowledge-based technique has shown significant retrieval and reasoning capabilities in other fields but has not yet been fully exploited in BID – knowledge graph: one unique advantage of knowledge graph is its capacity to represent complex, interconnected knowledge in a structured manner, and facilitate efficient and precise retrieval of the information; applying knowledge graph into BID, leveraging its reasoning capacity, could connect implicit bio-inspired design knowledge and aid in discovering new applications for specific biological features; despite these, the process of building a BID-specific knowledge graph involves the efficient representation and extraction of both entities and relationships across diverse domains. As such, the utilisation of knowledge graph in BID represents a research opportunity to further explore the retrieval and reasoning capabilities.

In response to the research opportunity mentioned above, this study proposes an approach that applies the knowledge graph to BID. It defines an ontology and constructs a knowledge graph accordingly, thereby providing designers with an effective manner for knowledge retrieval. Furthermore, our approach facilitates the reasoning capabilities with a link prediction model, enabling it to predict new links from existing knowledge entities and relationships, which can provide designers with more insights within the BID domain.

This study presents four key contributions to the domain of BID:

- (1) A novel knowledge graph-based BID approach has been proposed. In the approach, a BID-specific ontology is defined and applied to the construction of knowledge graph. This approach showcases a diverse knowledge structure and multidisciplinary connections in the knowledge graph, which further facilitates the representation and reasoning of biological knowledge in design contexts, enhancing the efficiency and potential creativity in bio-inspired design.
- (2) An original BID knowledge graph has been constructed. An ontology capable of representing diverse and multidisciplinary BID knowledge systems has been defined. The relationships among these entities were established based on an incentive perspectives approach and a systematic biological taxonomy. This ontology serves as the basis for the extraction of cross-domain entities from BID knowledge bases, leveraging the capabilities of large language models.

- (3) Automated reasoning within the constructed BID knowledge graph has been successfully implemented, by employing a link prediction model that aligns with the BID process of mapping from biological systems to design applications. The effectiveness of this link prediction model has been validated through objective metrics.
- (4) Two case studies were conducted to validate the knowledge graph-based retrieval and reasoning in BID. The proposed approach demonstrated improved performance in retrieving biological knowledge by considering relationships, thereby enhancing the ideation process. Furthermore, through reasoning, the employed link prediction model generated relevant and inspirational knowledge triples, indicating the substantial potential for innovative bio-inspired design solutions.

2. Related works

2.1. BID tools

Bio-inspired design (BID) describes a process where design concepts are drawn from biological systems to solve human problems. This process includes the identification of a design problem, exploration of biological analogies, abstraction of biological strategies, and application of these strategies to develop innovative solutions (Chakrabarti et al. 2005). On the basis of the process, several knowledge-based tools have been developed to facilitate BID, each with its unique features and limitations (Yen et al. 2014). Representatives of these BID tools include (Deldin and Schuknecht 2013; Goel, McAdams, and Stone 2015) and its derivative AskNatureNet (Chen et al. 2023), DANE (Goel et al. 2012), IDEA-INSPIRE 4.0 (Siddharth and Chakrabarti 2018), Ontology-explorer (Kozaki and Mizoguchi 2014), Design Study Library (Goel et al. 2022), IBID (Goel, McAdams, and Stone 2015) and Bionic Inspiration (BionicInspiration.org 2014). To systematically review these tools, they are further analyzed based on the following criteria:

- Ontology: the framework or schema the tool utilises for organising BID knowledge
- Representation Form: whether the data is represented in narrative text, database, or semantic network.
- Retrieval: the type of information the tool provides when searched by users.
- Reasoning: the level of assistance the tool provides to users for drawing connections

The ontology (Gruber 1995) refers to a set of concepts and categories that represent knowledge in a specific subject area or domain. In the context of BID tools, an ontology serves as a schema or a framework that organises data and helps users understand the connections between different pieces of information (Kim and Lee 2017). The representation format describes the way knowledge is textually presented in the tool. Some tools utilise a narrative text format, where information is described in sentences and paragraphs, much like a traditional article or essay. Others may choose a database format, which presents data in rows and columns for better organisation and comparison. Some tools prefer a network or graph format that visually illustrates the relationships and connections between different pieces of information. The retrieval refers to the tool's capability to search for and extract information in response to user queries. It determines how the tool processes specific keywords or phrases and the way it presents results relevant to these queries. The reasoning

Table 1. Description of knowledge-based BID tools.

Tool	Ontology	Representation format	Retrieval	Reasoning
AskNature	AskNature's taxonomy	Narrative Text	Textual descriptions of biological system and design application	User-interpreted reasoning
Bionic Inspiration	–	Narrative Text	Textual descriptions of biological system and design application	User-interpreted reasoning
Design Study Library	–	Narrative Text	Textual descriptions of design application	User-interpreted reasoning
DANE	SBF	Database	Design application's structure, behaviour, function	Category-assisted reasoning
IBID	SBF	Database	Design application's structure, behaviour, function	Category-assisted reasoning
IDEA-INSPIRE 4.0	SAPPhIRE	Database	Design application's behaviour, structure, and state transitions	Category-Assisted reasoning
Ontology-explorer	Design-related ontology	Semantic network	Biological systems and design application	Basic inferential reasoning
AskNatureNet	Source-benefits-application	Semantic network	Biological systems, benefits and design application	Basic inferential reasoning

refers to the tool's ability to support users in making connections and understanding the implications of the retrieved information in their design context. It measures how the tool assists users in deriving new knowledge or conclusions based on the information retrieved. The tools are described in Table 1.

As demonstrated in Table 1, in terms of ontology, AskNature's ontology is a taxonomic categorisation that classifies biological phenomena and their potential design applications based on their behaviours, mechanisms, and strategies. However, this taxonomy is primarily used for classifying articles within the AskNature database and does not capture aspects of the design process. DANE and IBID employ the Structure-Behaviour-Function (SBF) ontology, which focuses on organising knowledge about the design application's structure, behaviour, and function. This approach helps users view clear connections between a system's components (structure), the way these components interact (behaviour), and the system's overall purpose or function. IDEA-INSPIRE 4.0 utilises the SAPPhIRE (Helfman Cohen, Reich, and Greenberg 2014) model, which structures knowledge around seven types of causality in engineered and biological systems. On the other hand, Ontology-explorer and AskNatureNet adopt the design-related ontology and the source-benefits-application ontology, respectively. The design-related ontology, defined by the creators of Ontology-explorer, aims to link the biological systems to the design applications by structuring knowledge from the source to the application. In contrast, the source-benefits-application ontology used by AskNatureNet provides a broader view, covering knowledge from the source, through the benefits, and finally to the application. Bionic Inspiration and Design Study Library do not specify any ontology, potentially creating ambiguity in how knowledge is organised and retrieved.

Concerning representation format, AskNature, Bionic Inspiration, and Design Study Library use a narrative text format, being informative in context and detail. However, the

unstructured nature of the narrative text makes it challenging for users to extract specific information without comprehensive reading and understanding. DANE and IDEA-INSPIRE 4.0, on the other hand, use a database for their representation format. This structured presentation makes it easier for users to find information, but the fixed layout might restrict their ability to show the complex connections between biological systems and design applications. Ontology-explorer and AskNatureNet use a semantic network representation format. This format presents a mixture of biological systems and potential design applications, also highlighting the connection within the BID domain.

Regarding retrieval, each tool's output is influenced by its ontology and representation format. AskNature, Bionic Inspiration, and Design Study Library provide textual descriptions of biological systems and design applications due to their narrative text format. As a result, the retrieval relies on users' comprehension and their ability to understand valuable information from the text. DANE, IBID, and IDEA-INSPIRE 4.0, with their database representation format and their ontology, return retrieval results focused on the design application's structure, behaviour, and function. The structured format of the database allows for more straightforward retrieval of information. However, the connection between biological systems and design problems might not be retrieved due to the rigid structure of the database. Ontology-explorer and AskNatureNet, utilising semantic network representation format, offer retrieval results that include a mix of biological systems and design applications. Due to their ontologies and network structure, these tools provide a more interconnected view of the BID domain. However, the effectiveness of retrieval could still be affected by the completeness of their ontologies and the inherent restrictions of their representation format. For instance, the retrieval results of Ontology-explorer may be limited since the ontology only covers the source and application domains. On the other hand, AskNatureNet, despite having a source-benefits-application ontology that provides a broader view, might still face difficulties in capturing the full extent of the multidisciplinary BID knowledge due to its lack of defined relationships.

Regarding reasoning, all tools require some level of user interpretation, but the degree to which they depend on the user varies. For AskNature, Bionic Inspiration, and Design Study Library, the reasoning is primarily 'user-interpreted.' This means that users need to make connections and draw conclusions themselves, which can be challenging when dealing with complex biological and design knowledge. In comparison, DANE and IDEA-INSPIRE 4.0 incorporate a 'category-assisted' reasoning mechanism. These categories help structure the information and assist users in making connections, making the reasoning process easier. However, these tools also have their limitations as they largely depend on predefined categories. Ontology-explorer and AskNatureNet represent the 'basic inferential' reasoning. These tools infer relationships and interactions between entities within the semantic network, which include identifying similarities in function, structure, or behaviour between biological systems and design applications. However, the effectiveness of their reasoning depends on the completeness and accuracy of the connections in their semantic network.

In summary, while each tool contributes to BID in its unique way, the complexity and diversity of information in BID make knowledge retrievals and shallow knowledge reasoning insufficient (Chandrasegaran et al. 2013; Wan et al. 2022). To address these challenges, the focus of the current research is on developing a knowledge graph ontology that represents intricate interrelationships. This approach captures both entities and their relationships, ensuring a representation of bio-inspired design knowledge. By achieving this,

designers can be better supported in their retrieval and reasoning endeavours, reducing the cognitive effort required.

2.2. Knowledge graphs in design

A knowledge graph is a data structure that represents knowledge in a structured and interconnected manner (Ehrlinger and Wöß 2016). It typically consists of nodes that represent entities and edges that represent relationships between entities. They drive behind search engines such as Google Knowledge Graph, power recommendation systems such as Pinterest, and underpin artificial intelligence platforms such as IBM Watson. They have proven to be particularly useful in various knowledge-rich fields, such as healthcare (Jiang et al. 2015), e-commerce (Wang et al. 2018), education (d'Aquin and Motta 2011), and finance (Elhammedi et al. 2020).

Engineering design is also a knowledge-rich task where designers apply their technical and creative knowledge to conceptualise products, design product structures, and explore and optimise designs (Jiang et al. 2022; Song et al. 2019; Zuo et al. 2022; Zuoxu, Jihong, and Lianyu 2022). However, traditional approaches for knowledge retrieval and reasoning may not adequately handle the increasing complexity and diversity of information in the design process. They often fall short in effectively capturing semantic relations to support systematic and comprehensive design reasoning, frequently causing interruptions in the design process (Chandrasegaran et al. 2013; Wan et al. 2022; Zuo, Yin, and Childs 2022). The challenges encountered in maintaining the completeness and correctness of knowledge (Li et al. 2021) further emphasise the need for a more streamlined approach to knowledge representation and organisation to ease the cognitive load on designers. Knowledge graphs present a promising solution to these challenges. They can handle heterogeneous knowledge representation, allow efficient knowledge retrieval, and are easy to maintain (Li et al. 2020; Pradel et al. 2018; Wan et al. 2022). Furthermore, they can be combined with other leading-edge technologies to manage complex engineering design tasks (Baltrušaitis, Ahuja, and Morency 2018; Huet et al. 2021; Li et al. 2023; Shen et al. 2023; Wang et al. 2022; Zheng et al. 2021; Zhou et al. 2022).

Creating a knowledge graph for BID presents two challenges. The first challenge is defining an ontology capable of effectively representing BID knowledge and relationships (Nagel et al. 2008). Ideally, a comprehensive ontology for BID needs to encompass a wide variety of biological systems, design problems, and solutions, and the relationships among them. Given the diverse and multidisciplinary nature of BID, the ontology involves merging knowledge from various fields such as biology, design, engineering, and more (Zhu, Zhang, and Luo 2023). This definition requires thoughtful consideration of the scope, detail, and interconnections of the concepts within these fields.

Following the ontology definition, another significant challenge is the extraction of entities and relationships (Chen et al. 2023). In the BID context, the extraction is especially challenging due to the knowledge complexity and its cross-domain nature. Natural language processing and machine learning techniques become necessary. These computational techniques excel at extracting entities and relationships from various knowledge sources due to their advanced pattern recognition and adaptive learning abilities. They act as critical tools, enabling the exploration and comprehension of this complex knowledge landscape.

2.3. Reasoning in knowledge graphs

Knowledge graphs serve as a powerful technique for reasoning, a capability that infers unknown facts from known ones (Chen, Jia, and Xiang 2020). While there are various methods to achieve reasoning in knowledge graphs, a predominant and practical approach is through a process known as link prediction (Liben-Nowell and Kleinberg 2003). Link prediction can be perceived as a form of reasoning since it involves enriching and completing the graph, further enabling the anticipation of new, unexplored knowledge (Sun et al. 2011). The infusion of causal knowledge in knowledge graphs showcases the breadth of relationships knowledge graphs can capture and utilise, which is important in comprehensive design reasoning (Zhou et al. 2022). Link prediction serves as a means for reasoning in knowledge graphs and has been applied in various areas like social network analysis and product recommendation.

A range of algorithms has been developed to perform link prediction and thus enable reasoning in knowledge graphs. Notably, algorithms such as TransE, TransH, and TransR (Bordes et al. 2013; Lin et al. 2015; Wang et al. 2014) employ transductive learning principles. These algorithms operate based on the idea of translating entities and relations into a shared latent space. For instance, TransE models relationships as simple vector subtractions, making it straightforward and computationally efficient. On the other hand, inductive learning-based algorithms like Graph Convolutional Networks (GCNs) and Graph Attention Networks (GANs) utilise deep learning techniques to learn high-level features from graph-structured data. Specifically, GCNs are powerful at learning node representations by merging their neighbouring information, while GANs provide attention mechanisms to weigh the importance of neighbouring nodes differently, thus capturing more complex patterns.

In the context of BID, reasoning using link prediction could facilitate the discovery of implicit relationships between biological systems and design applications. For instance, a predicted link could suggest that a particular biological mechanism, which has not been considered in the real-world, could inspire a design application. These predictions are stimulative in creating design ideas.

Knowledge graphs offer unique advantages in reasoning due to their ability to represent complex relationships in a structured way. Their interconnected nature allows for the propagation of information, making it possible to infer relationships based on existing connections. Moreover, they provide a rich context for every entity, enhancing the accuracy of prediction algorithms. In summary, the use of knowledge graphs in BID can play a significant role in the reasoning and discovery of new design inspirations, thus serving as a powerful tool for bio-inspired design.

3. Knowledge graph-based BID approach

In this study, an approach is proposed to construct a knowledge graph and apply it to bio-inspired design (BID), promoting efficient knowledge retrieval and reasoning in BID activities. This approach is structured into two stages: knowledge graph construction and reasoning. These stages in the approach are then evaluated with two case studies. The overview of the approach is shown in Figure 1, with each module of the approach being elaborated in detail in the following sections. The construction stage of the knowledge

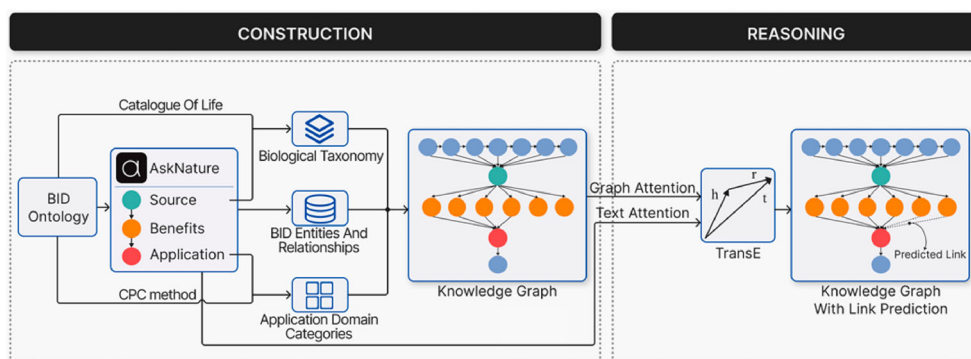


Figure 1. The overview of knowledge graph-based BID approach for knowledge retrieval and reasoning.

graph is detailed in Section 3.1, focusing on the definition of BID ontology, while the reasoning stage is elaborated in Section 3.2, focusing on the reasoning of BID ontology.

3.1. Definition of BID ontology

Knowledge graphs are efficient in managing and retrieving complex information, thus ideally positioned for organising knowledge systematically (Ehrlinger and Wöß 2016). This is particularly important in the field of BID, as it comprises diverse and multidisciplinary knowledge that needs a structured representation. To establish such a structured representation, the first step is to craft a BID-specific ontology to effectively represent biological and engineering design knowledge (Chakrabarti et al. 2005). This bespoke ontology comprises two main components: entities and relationships, as illustrated in Figure 2. The entities encompass ‘Biological category’, ‘Source’, ‘Benefits’, ‘Application’, and ‘Application categories’. This knowledge is primarily derived from the BID cases in the AskNature database, with an example of a BID case shown in Figure 3. Within a BID case, the primary entities are ‘Source’, ‘Benefits’, and ‘Application’, each playing a distinct role in connecting the source domain to the target domain with biological analogies (Nagel et al. 2008): the knowledge that bridges a biological source and its application lie in the biological ‘benefits’ to the source. In essence, these three entities provide a structured representation of BID cases, as illustrated in Figure 3, which shows an example of extracting these three entities from a BID case. The detailed definitions of these entities within a BID case are outlined as follows:

- ‘Source’: Within the realm of BID, the term ‘Source’ refers to the specific biological entity, such as an organism, system, or process, that offers inspiration for design solutions. This could be a particular animal, plant, microorganism, or even an entire ecosystem. By studying and understanding these biological entities, designers and engineers can extract crucial insights and principles. These insights are often grounded in millions of years of evolution, offering highly optimised and efficient solutions to various problems. For instance, the skin of an ‘Australian frog’, which leads to its encourage-healing properties, or the aerodynamic form of a ‘bird’s wing’, inspiring the design of more efficient aircraft, are examples of ‘Sources’ in BID. The ‘Source’ is unique for each BID case and should be described with nouns only.

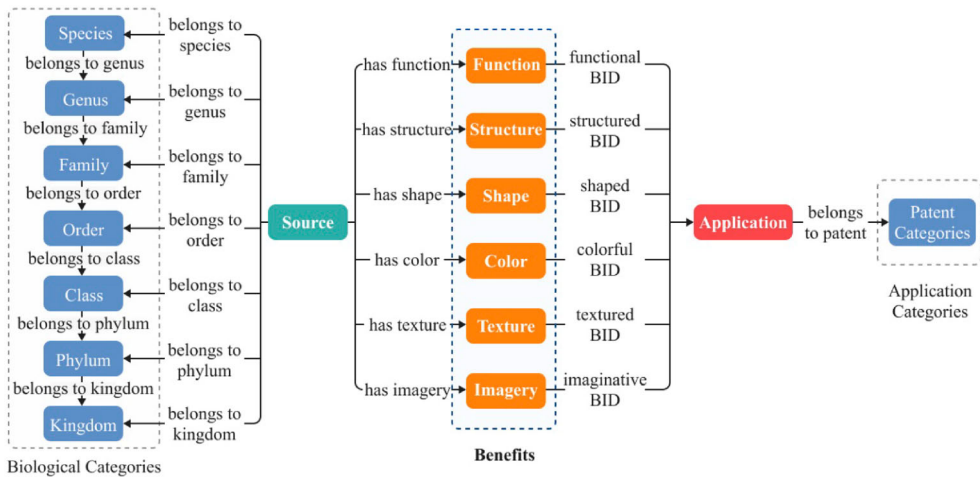


Figure 2. The ontology of knowledge graph.

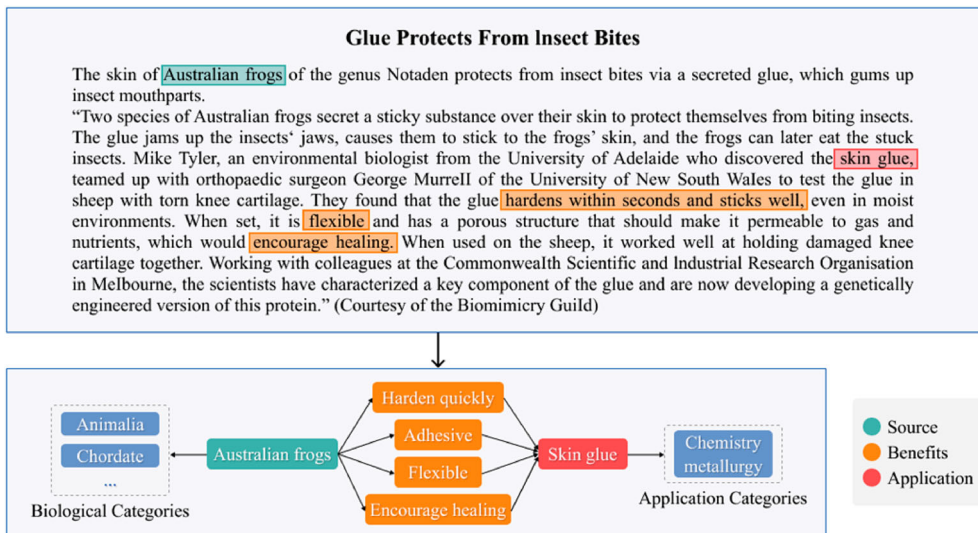


Figure 3. An example of a BID case with the extraction of structured entities.

- The 'Benefits' in a BID case refer to the specific advantages, characteristics, or properties that the biological 'Source' offers. These benefits are the aspects of the source that make it appealing for adaptation or inspiration in human-made designs. They are typically attributes that have been refined through natural selection, leading to highly effective and sustainable solutions in nature. It should be noted that in the context of BID, 'Benefits' cannot exist independently and must be considered in relation to both 'Source' and 'Application'. In contrast to the 'Source', there can be multiple 'Benefits' in a BID case. These 'Benefits' should be described with combinations of verb-noun or adjective-noun. For instance, in an 'Australian frog', the 'Benefits' could include 'encourage healing', 'harden quickly', 'adhesive' or 'flexible', which might inspire new skin glue in

medicine. Understanding these ‘benefits’ is crucial for translating biological inspiration into practical, innovative design solutions.

- ‘Application’: The ‘Application’ of a BID case represents the result of translating the biological ‘Benefits’ derived from the ‘Source’ into innovation or design solutions in technology, architecture, engineering, or other fields. For instance, the ‘Application’ might involve developing new skin glue based on the encourage-healing mechanisms found in Australian frogs or creating new materials inspired by the resilient nature of mollusk shells. The ‘Application’ in a BID case is unique and should be described with nouns only.

Entities also include categories from a hierarchical structure designed to provide consistency and scalability to the knowledge graph. The hierarchy encompasses domain categorizations at the top level, covering a wide range of biological systems and engineering applications. These domain categorizations are further subdivided into more specific ones, effectively structuring the classification and retrieval of information. The hierarchy includes biological categories (Figure 2) for the ‘Source’ domain, defining categories from kingdom to species level, which are treated as separate entities. Similarly, for the ‘Application’ domain, a patent categorisation method (CPC method) divides the domain into nine distinct categories. These categories are also considered entities in the knowledge graph. The patent categorisation method could facilitate the subsequent search for patents in external patent databases (Yoon and Park 2004).

The ontology also defines relationships among these entities, establishing relationships across four types: ‘Source’ to ‘Benefits’ entities (S-r-B), ‘Benefits’ to ‘Application’ entities (B-r-A), ‘Source’ to biological categories entities (S-r-C), and ‘Application’ to application categories entities (A-r-C). The relationships within the ontology are guided by the concept of ‘incentive perspectives’ (Luo et al. 2018), which embody the essence of bio-inspired design. These perspectives bridge the source and application domains by interpreting the advantageous biological features into actionable design principles. Six incentive perspectives have been identified: function, structure, shape, colour, texture, and imagery. The ‘Benefits’ entities correlate with these incentive perspectives, establishing diverse relationships like ‘has function’, ‘has structure’, ‘has shape’, ‘has colour’, ‘has texture’, and ‘has imagery’, linking ‘source’ to ‘benefits’ entities. These relationships determine the type of bio-inspired design – ‘functional BID’, ‘structured BID’, ‘shaped BID’, ‘colourful BID’, ‘textured BID’, or ‘imaginative BID’, ‘benefits’ to ‘application’ entities depending on the predominant ‘benefits’.

Figure 4 depicts an example of the ontology, with the focus on Australian frogs (The BID case of Figure 3). The ontology lays out a hierarchical biological categorisation, placing the ‘Crucifix toad’ within the species level, and tracing its classification up through genus, family, order, class, phylum, and kingdom. The biological benefits of the Australian frogs, such as its quick-hardening and encouraging healing function, adhesive texture, and flexible structure, are highlighted and directly linked to the bio-inspired application: skin glue. Furthermore, the skin glue is connected to broader application categories in patents like ‘metallurgy; chemistry’, demonstrating the ontology’s capability to extend biological inspiration across various engineering domains.

However, following the ontology definition, the task of extracting the BID knowledge from various databases encounters challenges. Related databases are not built under the guidance of a specific ontology, which results in the data being unstructured text.

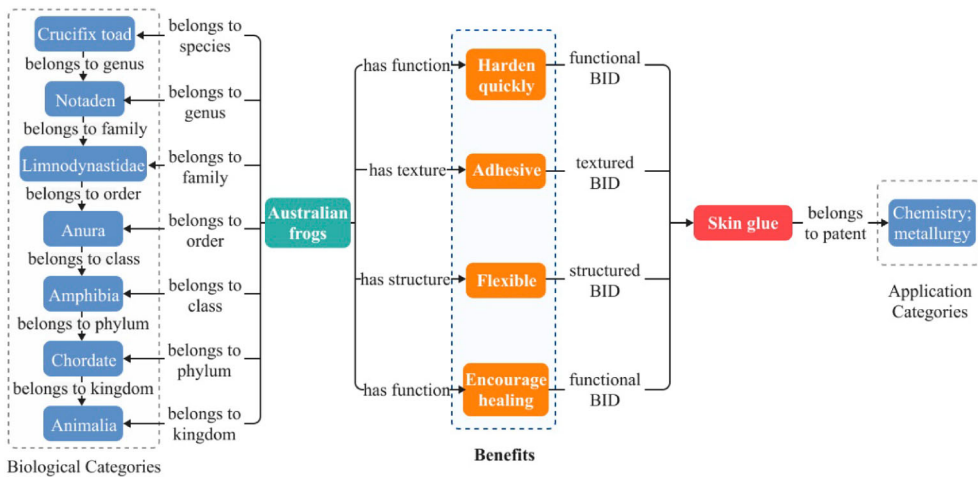


Figure 4. An example from 'Australian frogs' for the ontology.

Extracting structured knowledge in line with the ontology from unstructured text demands a Natural Language Understanding (NLU) capability. Moreover, the substantial volume of BID knowledge contained in these databases necessitates an efficient extraction approach. These challenges highlight the need for large language models, which possess the capacity to mine required BID knowledge accurately and efficiently from varied databases. Thus, large language models (LLMs) are employed to extract BID entities and relationships and then fill into the defined ontology to construct the knowledge graph. Specifically, to strengthen the entity extraction capabilities of LLMs in the field of BID, we created a dataset consisting of 200 samples for fine-tuning the LLMs, enhancing their entity extraction capabilities. Additionally, to support relationship extraction tasks, few-shot learning prompts were designed for LLMs, fulfilling the text classification requirements. The precision and reliability of the constructed knowledge graph are validated by BID experts.

3.2. Reasoning of BID ontology

The reasoning encourages innovation through the generation of new knowledge from existing knowledge (Nickel et al. 2015). It holds an important role in this approach, leading to the generation of novel and innovative BID knowledge (Paulheim 2017). The innovative knowledge originates from establishing new and unexpected links between biological systems and design applications, which are the essence of BID. Our approach primarily predicts a specific knowledge triple: $\langle \text{benefits}, \text{relation}, \text{application} \rangle$. This triple embodies the process of harnessing biological inspiration for real-world design implementations. Therefore, in the reasoning stage, a link prediction model is employed to establish these new links within the knowledge graph. In the model, the TransE algorithm is chosen for its capacity to represent entities and relationships as vectors and alignment with the logic of BID. The integration of the TransE algorithm into BID knowledge graph represents a novel method in our research, introducing automated reasoning that significantly enhances efficiency. To enhance the performance of the reasoning process, graph attention and text

attention mechanisms are implemented. Their role is to increase the link prediction model's performance to predict more accurate and relevant knowledge triple.

4. Knowledge graph construction

Following the proposed approach, constructing such a BID knowledge graph starts with defining an ontology that can represent the multidisciplinary knowledge of BID (Studer, Richard Benjamins, and Fensel 1998). With this ontology in place, the next step involves the extraction of relevant entities and relationships from a BID knowledge base. This structured representation of knowledge can facilitate effective retrieval and reasoning of information in the BID domain.

4.1. Extraction of entities and relationships

Upon the establishment of ontology, the next stage involves the extraction of entities and relationships from widely recognised knowledge bases (Bizer et al. 2009). To extract entities, we employ AskNature, a renowned BID knowledge base, as the data source for entities of 'Source', 'Benefits', and 'Application'. This BID knowledge base is recognised for its extensive repository of biological strategies and corresponding design application cases, with high credibility (Vincent et al. 2006).

The GPT-3 model (Brown et al. 2020) is utilised to extract suitable 'Source', 'Benefits', and 'Application' entities from BID cases. This Large Language Model (LLM), with its significant amount of parameters and training data, has demonstrated capability and competitiveness across various natural language processing tasks (Raffel et al. 2020). Notably, GPT-3 also exhibits great few-shot learning abilities, which enables it to learn new tasks with minimal training data. In Chen et al. (2023), we fine-tuned three GPT-3 models using human-labeling data from 200 BID cases of AskNature, each responsible for extracting the 'Source', 'Benefits', and 'Application' entities, respectively. The detailed fine-tuning process and evaluation of the extraction results can be found in Chen et al. (2023). It should be noted that a given biological system is not limited to a singular tuple. Instead, it can be represented by multiple tuples, where one biological system might have several benefits, which in turn can lead to several design applications. Such multiple tuples reflect the complex and interconnected nature of biological knowledge and its translation to design applications. Regarding domain categorisation, the nine categories in the patent categorisation method are treated as separate application categories entities. Similarly, the biological categories in the biological taxonomy are also considered entities.

Three domain categorisation tasks for entities are conducted to extract relationships following the relationship definition of ontology. The first task involves categorising 'Benefits' entities based on their respective incentive perspectives, as outlined in section 4.1. This task enables the extraction of relationships between 'Source' and 'Benefits', as well as between 'Benefits' and 'Application'. To achieve this, we utilise the powerful GPT-4 model, renowned for its exceptional capabilities in natural language understanding, text generation, knowledge representation, and abstraction (Radford et al. 2019). With these unique capabilities, the GPT-4 can be skilfully guided with appropriate prompts to perform targeted tasks. In this research, we adopt the prompt construction method proposed by Gutiérrez et al. (2022) to guide GPT-4 to accomplish a categorisation task by a few-shot learning prompt without

The definition of these 6 categories of attribute are listed as below:

Shape: Used to describe the appearance and shape of living things.

Function: Used to describe functional principles of the existence of organisms and natural substances.

Texture: Used to describe biological surface texture characteristics or natural material tissues.

Structure: Used to describe the structural structure and organizational form of a whole or part of an organism.

Color: Used to describe the rich color forms of nature.

Imagery: Used to describe the extraction of biological facial features and their symbolic meanings.

Classify separately which category does the attribute belong to:

category: Shape, Function, Texture, Structure, Color, Imagery.

attribute: 1.sharp-edged horny beak 2.dietary flexibility 3.water-resistant 4.strong and flexible bond 5.vivid color 6.fibrous structure 7.eco-friendly materials 8.changeable shape 9.produce iridescence 10.increased effectiveness

category: 1.Structure 2.Function 3.Texture 4.Structure 5.Color 6.Structure 7.Texture 8.Shape 9.Color 10.Function

attribute: 11.stabilize burrows 12.self-heal membrane 13.high efficiency 14.stronger materials 15.signal ripeness 16.visual narrative 17.rotate shells 18.brush-like papillae 19.edible plant material 20.manage cracks

category: 11.Structure 12.Texture 13.Function 14.Texture 15.Function 16.Imagery 17.Shape 18.Shape 19.Texture 20.Structure

Figure 5. Few-shot learning prompt for GPT-4 categorising benefits entities.

fine-tuning. The prompt of GPT-4 for the categorisation task of 'Benefits' entities consists of a task instruction, the definitions of six incentive perspectives, and few-shot task examples. To ensure the effectiveness of this approach, a pilot study was conducted to identify the optimal prompt that achieved high categorisation accuracy. Figure 5 shows the optimal prompt used in this study. As in Figure 5, different colours indicate different prompt design elements: orange for overall task instructions, red for definitions of six incentive perspectives, blue for the incentive perspectives list and purple for the categorical message portion. The first red and blue elements represent few-shot task examples. The current input and the completion by GPT-4 are highlighted.

The second task, which involves categorising 'Application' entities, employs the internationally recognised Cooperative Patent Classification (CPC) method (Montecchi, Russo, and Liu 2013). This CPC method, a joint venture between the US Patent and Trademark Office (USPTO) and the European Patent Office (EPO), divides patents into nine primary categories, each reflecting a distinct technology area. It further stratifies them into classes, sub-classes, groups, and sub-groups for more granular and precise categorisation. Due to its comprehensive and hierarchical structure, the CPC method is notably effective when applied to 'Application' entities, facilitating detailed categorisation and, thereby, streamlining efficient and accurate patent searches and analyses through these entities. Only the nine top-level categories are utilised to categorise 'Application' entities to ensure a broad yet distinct grouping. This approach is analogous to the categorisation process applied to 'Benefits' entities, wherein GPT-4 is used to guide the categorisation of 'Application' entities through skilful prompts without fine-tuning. Figure 6 shows the two main components of the prompt used for this task: the task instructions and the definitions of the nine patent categories.

Unlike the categorisation of 'Benefits' and 'Application' entities using GPT-4, the third domain categorisation task for 'Source' entities is accomplished through querying a biological taxonomy database. Given that the biological taxonomy serves as a comprehensive database recording the names and corresponding categories of almost all known organisms, this query-based categorisation method for 'Source' entities is feasible. Catalogue of Life (Hobern et al. 2021), an online biological taxonomy database, is utilised for this

The definition of these 9 categories of patent are listed as below:

Human Necessities: Concerns personal demand products and activities including foodstuffs, agriculture, clothing, personal items, life-saving; health, amusement, and sports.

Performing Operations; Transporting: Involves methods or apparatus for performing mechanical operations or activities, for transporting physical objects or material.

Chemistry; Metallurgy: Covers general chemistry, metallurgy, and related processes.

Textiles; Paper: Contains anything related to the production, decoration, and treatment of textiles, paper, or flexible materials not otherwise provided for.

Fixed Constructions: Deals with the civil engineering aspects of constructions or structures which are fixed to the ground.

Mechanical Engineering; Lighting; Heating; Weapons; Blasting: Covers mechanical engineering, apparatus or methods for lighting, heating, cooling, steam generation, weapons and blasting.

Physics: Covers physical and nuclear physics, electricity, controlling, computing, measuring, testing, printing, recording, and signaling.

Electricity: Encompasses generation, conversion, or distribution of electricity, basic electric elements, and their use in more complex systems.

Emerging technology: This category is used for tagging new technological developments, cross-sectional technologies spanning over several sections of the IPC, and technical subjects covered by former USPC cross-reference art collections and digests.

Classify separately which category does the patent belong to:

category: Human Necessities, Performing Operations; Transporting, Chemistry; Metallurgy, Textiles; Paper, Fixed Constructions, Mechanical Engineering; Lighting; Heating; Weapons; Blasting, Physics, Electricity, Emerging technology.

attribute:.....

category:.....

Figure 6. Few-shot learning prompt for GPT-4 categorising application entities.

query. The database contains up-to-date information on organism names and their respective categories. The queries focus on the most standardised categories: kingdom, phylum, class, order, family, genus, and species. Concurrently, subcategories (e.g. subphylum, sub-classes) are excluded due to their non-standardised and extensive information. The queried category names obtained through this taxonomic query are treated as entities, and the relationships are derived from the defined relationships (see Figure 2) within the ontology.

4.2. Construction result

After the extraction process, a knowledge graph was constructed based on the defined ontology. This knowledge graph was reviewed and modified by two experts possessing a minimum of three years of experience in BID. Table 2 shows the assessments results of the 'Benefits' and 'Application' entities by two experts. Since the categorisation of the 'Source' entities is derived from reliable database queries, manual assessment is not required. Both experts agreed that over 95% of the categorizations are correct, with a high inter-rater reliability between them (Cohen's Kappa value exceeds 0.7). This process of expert evaluation ensured that the constructed BID knowledge graph maintained a high degree of reliability and validity. In the Appendix, Table A1 provides the number of categories for 'Source' nodes across seven biological taxonomy levels. Table A2 and Table A3 present the distribution of 'Benefits' and 'Application' categories, respectively. The knowledge graph is stored in the cloud (like Neo4j), allowing a seamless combination of knowledge-based systems and applications without extensive transplanting considerations (Vicknair et al. 2010). Table 3 displays a statistical comparison of entities and relationships among different databases used for bio-inspired design. Notably, our knowledge graph outstrips competing databases in both quantity and diversity of entities and relationships. Featuring nine distinct types

Table 2. Percentage of categorizations assessed as correct by experts and inter-rater reliability between experts.

Expert	Categorisation of 'Benefits' entities	Categorisation of 'Application' entities
Expert 1	0.9501	0.9821
Expert 2	0.9563	0.9767
Cohen's Kappa	0.7067	0.7781

Table 3. Comparing the entities and relationships of different databases.

Databases	Number of entity types	Number of relationship types	Number of entities	Number of relationships
Ontology-explorer	6	5	379	314
AskNatureNet	3	3	4351	5048
Our work	9	20	5574	12399

of entities and twenty different kinds of relationships, our database presents an extensive reservoir of knowledge, demonstrating a notable enhancement over Ontology-explorer and AskNatureNet.

5. Knowledge graph for BID

After constructing the knowledge graph, the next essential step involves applying it to the BID process for retrieval and reasoning. The retrieval process in BID is responsible for extracting relevant information, providing a foundation for solution-driven or problem-driven design patterns. Reasoning, on the other hand, allows for predicting new triples in the constructed knowledge graph, giving insight into new BID knowledge.

5.1. Knowledge retrieval for BID

Knowledge retrieval plays a crucial role in BID, underpinning key processes, including inspiration discovery, model understanding, and design ideation. It enables the acquisition of relevant biological models, comprehension of their benefits, and application to engineering design. Leveraging the knowledge graph's retrieval capabilities, the utilisation of our knowledge graph to support the knowledge retrieval phase of BID is both valuable and promising. This retrieval process will find the potential links between biological sources and engineering applications, providing enhanced assistance to designers during the conceptual design phase. Through these links, designers can efficiently identify suitable biological sources, understand their benefits, and apply them to create innovative designs.

Since BID typically involves two patterns: solution-driven and problem-driven (Wanieck et al. 2017). The problem-driven BID retrieval phase generally begins with a given problem or task, attempting to discover connections with biological knowledge, which involves a mapping from the target domain to the source domain (Vincent et al. 2006). The solution-driven BID retrieval phase begins with a specific biological system or phenomenon, which presents potential benefits that could be applied to technical solutions (Lepora, Verschure, and Prescott 2013). Considering the importance of supporting various retrieval needs, the knowledge graph-based retrieval approach is crafted to support both problem-driven and

solution-driven retrieval in BID. At the same time, because the input for retrieval is a problem requirement, which often has a high correlation with the incentive perspective, the incentive perspective within the knowledge graph is also used for retrieval input (Helms, Vattam, and Goel 2009).

Based on the above considerations, a two-phase retrieval approach is proposed, supporting two retrieved patterns. The input of the retrieval approach is an entity and its incentive perspective. Using problem-driven retrieval as an example, the user condenses the problem requirement into a term (i.e. 'Application' entity's name, like a drone) and opts for an incentive perspective (one out of six or all six). Users are guided to decompose their design requirements into several potential terms, and through iterative refinement, choose the term that best represents the core of their design problem. This chosen term then serves as the primary input for the knowledge graph retrieval process.

- Phase 1 (Identify Benefits): Leveraging the relationship type that the chosen incentive perspective represents, this phase identifies the 'Benefits' entity linked to the 'Application' entity input. This stage aims to pinpoint the benefits the problem requirement needs to apply.
- Phase 2 (Find Source/Application): Continuing from Phase 1, this phase identifies the 'Source' entity connected to the 'Benefits' entity, again based on the relationship type represented by the selected incentive perspective. The goal is to discover organisms with the desired benefits, potentially serving as design inspirations.

For solution-driven retrieval, the input is a 'Source' entity and the relationships used for retrieval in Phase 1 and 2 are reversed. Figure 7 shows two examples of the retrieval approach results of solution-driven (a) and problem-driven (b).

5.2. Knowledge reasoning for BID

After constructing the knowledge graph, a crucial application in the context of BID is leveraging its reasoning ability to predict new knowledge. This section details the approach and systematic procedures we employed to realise this reasoning ability. The focus of our approach is on predicting a specific triple (benefits, relation, application). This triple is critical to bio-inspired design, as it signifies the process of taking inspiration from biological sources to implementing it in real-world design applications.

Knowledge graphs are suitable for uncovering latent associations due to their reasoning abilities and expressive representation capabilities (Nickel et al. 2015). To harness this potential, a suitable logical reasoning algorithm corresponding with the inference logic of our BID framework is needed. It is crucial to ensure the accuracy and efficiency of link prediction (Bordes et al. 2013).

To facilitate the application of our bio-inspired design (BID) framework's inference logic, the TransE algorithm is adopted for link prediction. TransE, which stands for Translate Embeddings, is an embedding algorithm designed explicitly for the link prediction task in knowledge graphs. The core concept of TransE involves representing entities and relations as vectors in a shared low-dimensional space. Each entity is represented as a point in this embedding space, with each relationship interpreted as a translation operation from one entity to another. The design is based on the principle that if a relationship exists between

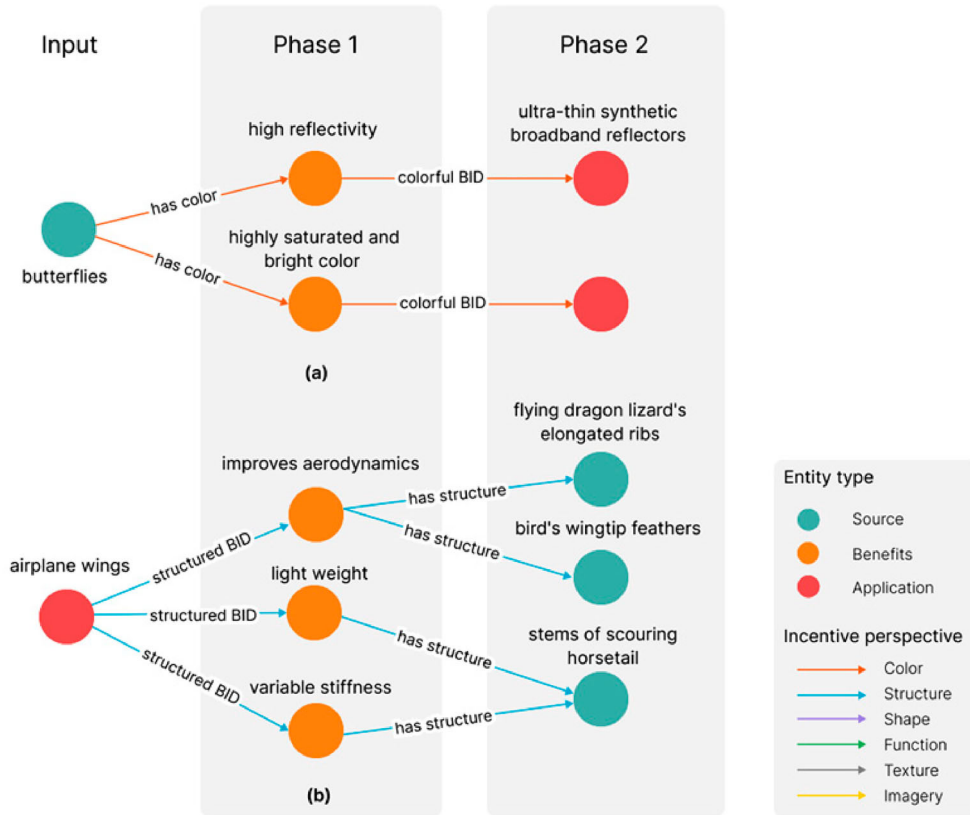


Figure 7. Illustration of retrieval approach results. (a) An example of solution-driven retrieval, with 'butterflies' as the input source entity and 'colour' as the incentive perspective. (b) An example of problem-driven retrieval, with 'airplane wings' as the input application entity and 'structure' as the incentive perspective.

two entities, the tail entity vector, when translated by the relationship vector, should align closely with the head entity vector. This way of vector translation aligns well with our BID framework because it models the reasoning process in the same way our bio-inspired design process operates: through a chain of transformations starting with biological benefits (head), to design inspirations (relationship), and ending with concrete applications (tail). Through the application of TransE, we can efficiently model these intricate transformations within the BID process, thereby enhancing its reasoning capabilities and prediction.

Following a systematic evaluation of various algorithms, we adopted the TransE algorithm for link prediction to align with our BID framework's inference logic. While other algorithms such as TransH and TransR can also provide robust performance in link prediction, they were not considered. Both TransH and TransR, with their ability to handle more complex relationships, introduce a level of complexity that leads to higher computational costs. Given the ontology of our BID knowledge graph, where relationships are relatively straightforward, the balance between performance and computational efficiency provided by TransE makes it a preferred choice. Therefore, TransE is an appropriate choice for the link prediction task (Wang et al. 2014).

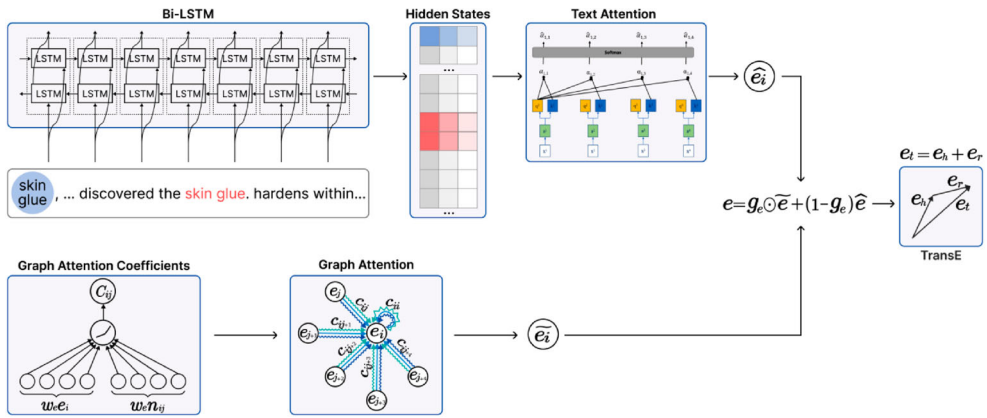


Figure 8. Link prediction model.

Although TransE has been selected for its balance between efficiency and performance, it is worth noting that the accuracy of link prediction can be further improved. This leads us to consider supplementary mechanisms that could enhance TransE's predictive accuracy by giving more focus to the more relevant aspects of the data.

To enhance the accuracy of link prediction in our knowledge graph, we integrated the graph attention mechanism (Veličković et al. 2017) and the text attention mechanism into the TransE model, following a comprehensive analysis of their potential impact. The graph attention mechanism allows the model to assign varying weights to different nodes, emphasising crucial relationships within the graph. Additionally, the text attention mechanism facilitates the model in assigning differential attention to parts of the text, enhancing the extraction of meaningful information from textual data. The TransE algorithm, when integrated with the graph attention and text attention mechanisms as depicted in Figure 8, forms a comprehensive link prediction model. This model effectively supports reasoning within our BID framework. We provide a detailed breakdown of how these mechanisms are implemented:

Stage 1: Text Attention

This step involves creating a comprehensive textual context for each entity, enhancing the model's understanding of the entity's meaning and relevance.

- 1) Associate each entity e with a context sentence and process it using a bi-directional Long Short-Term Memory (LSTM) to generate a sequence of encoder hidden states H_s .
- 2) Apply bilinear attention to each word in the context, generating attention weights: $\mu_i = e^T W_s h_i$, where W_s is the bilinear attention weight matrix.
- 3) Establish the overall textual context representation for the entity as: $\hat{e} = \mu^T h_i$.

Stage 2: Graph Attention

This step is focused on capturing the significance of each neighbour to a specific entity and the diverse relationships between them through a structured approach.

- 1) Employ self-attention to compute a weight distribution over the neighbours of e_i using transformation matrices. Transform the entity and neighbour vectors: $e'_i = W_e e_i$ and $n'_{ij} = W_e n_{ij}$, where W_e is the entity transformation matrix. Calculate the attention coefficients between entities and their neighbours using the formula: $c_{ij} = \text{LeakyReLU}(W_f(e'_i \oplus n'_{ij}))$, where W_f is the attention transformation matrix, and $\oplus$ denotes concatenation.
- 2) Achieve the structure-based context representation of e_i through: $\xi_i = \sigma \left(\sum c'_{ij} n'_{ij} \right)$, where σ is an activation function.
- 3) Apply multi-head attention to the structure-based context representation, ξ_i to generate an enhanced representation $\tilde{e}_i$. This involves using multiple sets of transformation matrices and attention mechanisms to capture various relationship nuances and patterns.

Stage 3: Integration of Graph Attention and Text Attention

This final step merges the graph-based and text-based representations to provide a balanced and comprehensive understanding of each entity within the graph.

- 1) Introduce a gate function represented as $g_e = \sigma(\tilde{g}_e)$, which is entity-dependent and controls the balance between graph-based and text-based representations.
- 2) Combine the two representations $\tilde{e}$ and $\hat{e}$ as $e = g_e \odot \tilde{e} + (1 - g_e) \odot \hat{e}$, where $\odot$ denotes element-wise multiplication.

By following these steps and employing the corresponding formulas, we systematically encode the graph structure with multi-head attention and integrate it with textual context. This approach ensures a comprehensive representation of entities, enhancing the reasoning capabilities of our BID framework and facilitating more accurate and relevant link predictions.

5.3. Knowledge reasoning evaluation

A systematic evaluation was performed using widely accepted knowledge graph link prediction metrics: mean rank, median rank, and hit@10, ensuring a comprehensive assessment of our model (Bordes et al. 2013). We detail the procedure for calculating these metrics and how they provide a comprehensive understanding of the model's effectiveness.

Mean Rank: This metric is calculated by first predicting all possible links for a given entity and ranking these predictions. The rank of the correct link is noted. This process is repeated for all entities in the test set, and the average of these ranks is computed. The lower the mean rank, the better the model is at predicting the correct link closer to the top of the ranked list.

Median Rank: The median rank is calculated by first predicting all possible links for a given entity and ranking these predictions. The rank of the correct link is noted. This process is repeated for all entities in the test set. The median rank identifies the middle value when all these ranks are listed in order. A lower median rank indicates that the model generally ranks the correct link higher.

Hit@10: This metric calculates the percentage of correct predictions that appear within the top 10 ranks of the model's predictions. It is determined by evaluating each correct

prediction's rank and noting if it falls within the top 10. The total number of hits within the top 10 is then divided by the total number of predictions made, giving a proportion that represents the model's accuracy in bringing correct predictions to the top of the list.

To assess the effectiveness of the link prediction model and the individual contributions of the graph attention and text attention mechanisms, an ablation study was conducted on a test set. The combined results of the enhanced TransE model and the ablation study are summarised in Table 4.

It can be seen that the mean rank values for head and tail predictions stand at 89.47 and 114.16 respectively, while the median rank values are 81 for the head and 91 for the tail. These values signify the average and median positions of correct predictions in the ranked list, suggesting a relatively high placement of correct predictions. In addition, the hit@10 value, which indicates the percentage of correct predictions appearing within the top 10 of the ranked list, was 54.11%. This implies that more than half of the correct link predictions were among the top 10 predictions. From the ablation study, it is evident that the combination of both text attention and graph attention mechanisms significantly improved the performance of the model. Specifically, when removing the text attention mechanism, the mean rank for head predictions increased from 89.47–101.47 and for tail predictions from 114.16–170.34. The hit@10 value also decreased from 54.11% to 34.30%. Similarly, without the graph attention mechanism, the mean rank for head predictions rose dramatically to 315.88 and for tail predictions to 181.37, with the hit@10 value dropping to 21.86%. These figures imply the importance of both attention mechanisms in achieving link prediction results.

The performance of the enhanced TransE model, which falls between moderate (with mean rank typically around 150–200 and hit@10 around 30–40%) and excellent (where mean rank is usually below 50 and hit@10 exceeds 60%) on the typical scale for link prediction tasks, is particularly impressive given the limitations of a small training set. Overall, the results indicate that the link prediction model offers promising potential in identifying unseen links between entities within the knowledge graph. To comprehensively assess the practical utility of these predictions in bio-inspired design, it is essential to supplement these objective metrics with additional subjective evaluations. Hence, further case studies are conducted and described in following section.

6. Case studies

This section presents two case studies to evaluate the proposed approach within the context of bio-inspired design. The first case study focuses on validating the quality of the knowledge graph retrieval approach through a task-based experiment involving design

Table 4. Evaluation Metrics for Link Prediction Performance.

Model	Mean Rank		Median Rank		Hit@10 Head and Tail
	Head	Tail	Head	Tail	
Enhanced TransE	89.47	114.16	81.00	91.00	54.11%
Without Text Attention	101.47	170.34	148.87	114.00	34.30%
Without Graph Attention	315.88	181.37	310.50	138.00	21.86%
Basic TransE	358.94	270.36	320.50	214.00	9%

challenges, where the ideas conceived through the retrieval process are assessed to emphasise the knowledge graph's role in facilitating successful ideation. The second case study delves into the practical implications and real-world applicability of the links predicted by the model. This case study involves evaluating newly formed triplets' quality and their interpretation in actual bio-inspired design scenarios. The objective is to reconcile objective success metrics with the subjective value of these predicted triplets in stimulating innovative design solutions.

6.1. Case study 1: knowledge retrieval

6.1.1. Design task

The knowledge retrieval aims to facilitate the completion of a BID design task. A problem-oriented design task revolving around the structural design of rescue drones is chosen due to the special challenges faced in this domain. Rescue missions require drones to navigate challenging terrains, carry payloads, and ensure safety in critical situations. The bio-inspired design offers valuable insights and solutions by drawing inspiration from nature's evolutionary innovations. Participants were expected to draw inspiration from nature's solutions to similar challenges by employing various knowledge retrieval approaches. The ultimate goal is to conceive a drone structure that enhances rescue mission efficiency, safety, and versatility while balancing cost-effectiveness and user-friendliness. The knowledge retrieval task involves exploring diverse and multidisciplinary BID knowledge, encompassing drone and biological structural knowledge, and identifying potential connections between the two. Therefore, this experiment can be used to evaluate the effectiveness of the entities and the relationships between entities in the constructed knowledge graph, thus improving the quality of BID knowledge retrieval.

6.1.2. Study procedure

The study used a problem-driven bio-inspired design process, adapted from the model presented by Fayemi et al. (2017) and illustrated in Figure 9. A design case study was conducted with thirty graduate students from industrial design engineering majors. The participants, evenly split by gender, had an age distribution with an age range 22–31 (Mean = 24.93 and SD = 2.33). All participants possessed 3–5 years of design experience (Mean = 3.9, SD = 0.87). While they had awareness of BID concepts, none had engaged in BID projects or attended specific BID courses. For the study, they were randomly divided into three equally sized groups, maintaining a balanced gender distribution in each. Each group used a different design knowledge retrieval approach to perform the same design task:

- Group-1: Used the AskNature tool, retrieval based on a knowledge base.
- Group-2: Employed the AskNatureNet model, retrieval based on a semantic network.
- Group-3: Used the knowledge graph tool, i.e. the approach proposed in this study.

Participants were then given 15 min to produce design ideas while each participant recorded ideas on A3 white paper, including written descriptions and sketches. To assess the outcomes, we focused on three primary indicators: creativity, elaboration, and practicality, commonly used to gauge ideation success in open-ended tasks (Daly et al. 2016; Shah, Smith, and Vargas-Hernandez 2003; Wang et al. 2023). Two domain experts (refer to

Section 4.2) evaluated each idea, assigning scores on a scale from one to seven for each indicator, where seven represented the highest level of creativity, elaboration, and practicality. The scores for each idea were then averaged across all participants, providing an individual average score for each idea.

6.1.3. Result and discussion

Figure 10 shows the average creativity, elaboration and practicality of ideas generated by all participants in each group. The non-parametric test Kruskal–Wallis was applied to determine statistically significant differences since these data did not follow a normal distribution. It can be seen that the average score of all three indicators of Group-3 is higher than Group-1 and Group-2.

Additionally, we found the following phenomena during the experiment: in Group-1, it was found that the AskNature tool matches relevant bio-design or bio-strategy cases based



Figure 9. A problem-driven bio-inspired design process adapted from Fayemi et al. (2017)

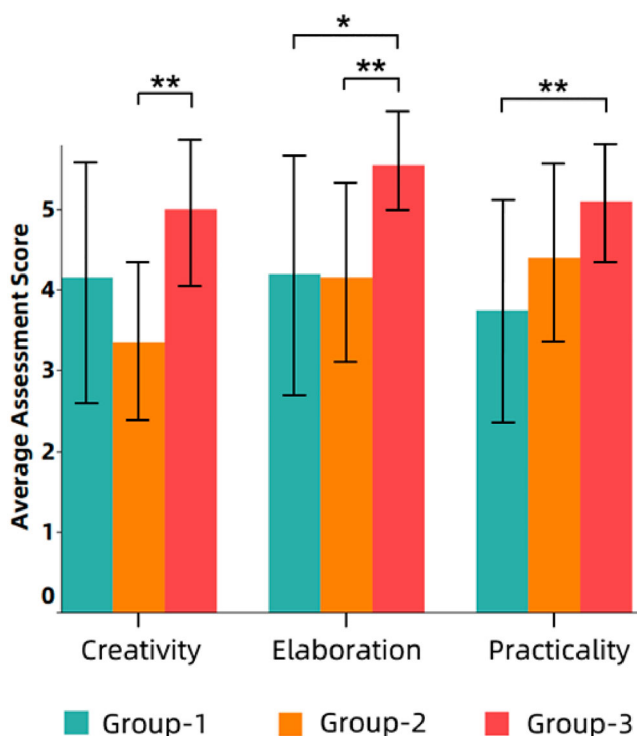


Figure 10. Result of average assessment score of ideas in three groups. (* indicates statistically significant differences at $p < 0.05$; ** indicates statistically significant differences at $p < 0.01$).

on search terms. The participants then understood and manually extracted design knowledge based on their experience, which was difficult and time-consuming. Most participants tended to design using common biological knowledge, such as the wings of swifts or albatrosses, and rarely used potentially relevant knowledge, such as the paws of squirrels. This is because the knowledge base retrieval approach of the AskNature tool can only retrieve case texts based on semantic similarity, and it is challenging to design a retrieval approach from the particular benefits of biological sources or applications. The latter approach tends to align with the requirement for biological analogies to be divergent in BID.

In Group-2, the participants used AskNatureNet's semantic network to search and quickly found many triples from applications to biological sources related to the search terms. However, since the semantic network could not accurately represent the relationships between knowledge, participants needed to make judgments to confirm whether the retrieved biological knowledge was related to the structure, which was a time-consuming process. The results obtained after the secondary judgment was more practical but may lose novelty.

In Group-3, the use of a knowledge graph tool simplifies the task of design knowledge retrieval, allowing participants to directly access a vast amount of biological structural knowledge relevant to their search terms. It saves time and eliminates the need for additional judgment. Furthermore, compared to semantic networks, the knowledge graph accurately represents the relationships between knowledge, making it easier for participants to comprehend the logical relationships of triple from applications to biological sources. Through the observation of Group-3's ideation sheets, it was found that participants tend to prefer combining multiple biological structures in their ideas, resulting in a more comprehensive and enriched structure for the drones. A design idea developed by Group-3 is illustrated in Figure 11. This idea draws inspiration from three 'Sources': the hummingbird, grapefruit, and seahorse tail. The designer has combined these elements to create a sketch of a rescue drone featuring an airframe similar to that of a hummingbird for lightweight flight, a material akin to grapefruit for shock absorption, and a structure resembling a seahorse's tail for rescue grasping.

Based on the phenomena above, it is inferred that compared with the other two groups, Group-3 spends less time on design knowledge selection and extraction after the retrieval. This allows Group-3 more time to refine new ideas, which leads to a higher practicality of ideas than other groups. Besides, Group-3 might use more relevant design knowledge

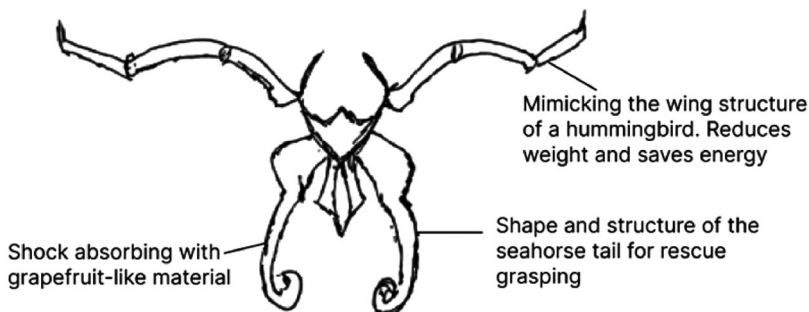


Figure 11. A conceptual sketch of Group-3's rescue drone, inspired by the hummingbird, grapefruit, and seahorse tail.

than Group-2 and Group-1, making the former more novel. In particular, the accurate representation of knowledge relationships and the ease of comprehension offered by the knowledge graph enable participants to apply multiple biological structures seamlessly in their designs, contributing to the elaboration of ideas.

To summarise, this experiment demonstrates that the proposed approach can effectively retrieve valuable biological knowledge, thereby significantly improving the performance of ideation results. In particular, the higher elaboration indicator in Group-3's ideas demonstrates the enhanced representation capability of our knowledge graph-based approach in managing diverse and interdisciplinary knowledge, compared to approaches based on semantic networks or knowledge bases. For designers aiming to use this method, the key is to accurately define the design challenge into specific terms or concepts. Once defined, these terms can be input into the system to retrieve relevant biological knowledge, providing a rich foundation for ideation.

6.2. Case study 2: knowledge reasoning

In this study, the knowledge graph's reasoning capabilities lead to the prediction of new triples, which can serve as novel sources of design inspiration and solutions, thereby driving innovation. Although the performance of the link prediction algorithm has been objectively evaluated in section 5, it is also necessary to understand the practical implications of the predicted triples in real BID scenarios. Thus, two tasks within the case study are undertaken to examine the model's predictions, involving quality and expert-validated evaluation.

6.2.1. Quality evaluation of new triples

The first task aims to assess the quality of new triples generated by the link prediction model, focusing specifically on the 'Benefits-type of incentive perspectives-Application' (B-r-A) triples, which are central to the analogy process in bio-inspired design. The case study involves 20 participants who are design professionals or students with a design background. All participants are given a list of 50 randomly shuffled triples, including 25 predicted by the model and 25 from the existing triples in the original knowledge graph. Participants are asked to assess each triple based on two criteria: relevance and inspiration. 'Relevance' refers to the logical relationship between the head entity (benefits), the relation (type of incentive perspectives), and the tail entity (application). 'Inspiration' refers to the potential of the triple to provide new insights or directions for designers. Both criteria are scored on a scale from 1 to 5, with 5 indicating the highest relevance or highest inspiration.

The scores for each of the 50 triples are then averaged across all 20 participants, providing an individual average score for each. Additionally, an overall average score is computed for the set of 25 predicted triples and the set of 25 original triples. These average scores are used to compare the quality of predicted triples with the original ones. The results are shown in Table 5.

It can be seen that the average relevance score for the predicted triples was 2.85, while the inspiration score was 2.81. For the original triples, the corresponding scores were 3.12 and 3.07, respectively.

The significance of the observed differences was evaluated using two-tailed t-tests on the relevance and inspiration scores of the predicted and original triples. The calculated

p -values were 0.206 and 0.198, respectively. In the context of statistical hypothesis testing, a p -value exceeding 0.05 signifies an inability to reject the null hypothesis, which suggests no significant difference between the two sample groups. Based on these p -values, we can infer that the differences between the predicted and original triples' quality are not statistically significant.

Overall, the evaluation reveals that while the average scores for relevance and inspiration of the original triples are slightly higher, these differences were not statistically significant. Consequently, this indicates that the predicted triples, despite slightly lower quality scores, still possess substantial value in the context of bio-inspired design.

6.2.2. Design interpretation from expert-validated link predictions

In this task, we extend our evaluation scope by engaging domain experts to interpret all the predicted triples, aiming to showcase the real-world applicability of the predicted triples. The experts are provided with all the predicted triplets and are tasked with a comprehensive interpretation of each, examining their probable real-world applications and context. Two instances are selected here as illustrative examples of this process:

In the first instance, the original triple is: (Bone, Has Structure, Cracks Sealing), which corresponds to the concept that bone structures have a self-repairing mechanism for sealing-cracks. The associated triple is: (Cracks Sealing, Structured BID, Self-Repairing Concrete). This signifies the bio-inspired design of self-repairing concrete, drawing upon the crack sealing capability of bone structures. Our model, including text and graph attention, predicted an innovative triple: (Cracks Sealing, Structured BID, Self-Deploying Robot Wings).

With experts' interpretation and conceptual design, this prediction proposes the potential for designing self-deploying robot wings inspired by the self-repairing property of bone structures. To achieve this, a novel composite material that emulates the biochemical properties of bones would be required. This material could enhance the resilience and repair ability of robot wings, offering an effective solution to the degradation that robot wings face due to external impacts or wear and tear.

In the second instance, the original triple is (Porcupine Fish's Skin, Has Structure, Inductive Protection), which relates to the fact that the structure of the porcupine fish's skin provides inductive protection. The associated triple is (Inductive Protection, Structured BID, Smart Package), suggesting a bio-inspired design application of inductive protection in smart packaging. The model predicted a new triple: (Inductive Protection, Structured BID, Biomimetic Polymer)

As the expert's interpretation and conceptual design, this prediction suggests the potential design of biomimetic polymers. This could be achieved by developing a novel polymer material that mimics the protective features of the porcupine fish's skin. This approach may offer substantial advancements in creating dynamic, responsive materials, with potential

Table 5. Evaluation scores for predicted and original triples.

Score	Predicted Triple Relevance	Predicted Triple Inspiration	Original Triple Relevance	Original Triple Inspiration
Average	2.85	2.81	3.12	3.07
Standard Deviation	0.70	0.70	0.78	0.54

implications across diverse domains, such as wearable technology or responsive architecture.

In addition to these predicted two triples, all other predicted triples conducted a similar expert interpretation and conceptual design process, ensuring a systematic evaluation of each prediction's practical applicability. These expert-interpreted cases demonstrate the potential of our link prediction approach to uncover novel bio-inspired design solutions, extending beyond the quality and novelty attributes highlighted in the preceding case study.

6.2.3. *Discussion of reasoning*

In assessing the performance of our model's reasoning capability, both objective statistical metrics (Mean Rank and Hit@10) and subjective quality evaluations were employed. Despite moderate performance on objective metrics, the model demonstrated an ability to predict triples that were of comparable relevance and provided equally high inspirational potential as the ground truth in subjective quality evaluations. Importantly, these predictions fostered design innovation by introducing novel perspectives and stimulating creativity, thus offering a reservoir of fresh ideas for bio-inspired design. The designer can apply the reasoning capability to predict unseen relationships, offering potential design insights. It should be noted that validating the predicted results are still beyond the machine's capability and it is recommended to combine the model's predictions with expert insights, ensuring that the derived solutions are both innovative and practical.

Furthermore, the expert interpretation experiment underscored the practical value of these predicted triples. It demonstrated their pragmatic application in addressing real design challenges, further highlighting the potential of our link prediction approach in translating abstract design inspirations into tangible, real-world solutions and validating the practicality of our model.

In comparing our model with established methodologies such as AskNature and AskNatureNet, our model introduces significant advancements in automated reasoning for bio-inspired design (BID). AskNature is well-known for its collection of biological strategies, organised by function, and serves as a valuable resource for designers seeking inspiration from nature. However, the reasoning process in AskNature is primarily user-driven; designers must manually browse and interpret the information to find relevant solutions, which can be time-consuming and may not uncover all potential relationships. AskNatureNet, as a semantic network focusing on source-benefits-application relationships, offers a structured way to browse through information but still relies on users to make connections and derive insights. It provides a useful platform for exploring relationships but does not automatically generate new predictions or unseen connections, limiting its capacity to inspire novel design solutions.

Our automated reasoning method introduces advantages over these platforms. Unlike AskNature and AskNatureNet, our method applies automated reasoning to BID knowledge graphs. This key difference allows for the automated generation of new and potentially unseen links, broadening the range of available ideas significantly beyond what manually curated databases can offer. Our method streamlines the process of finding design inspirations by systematically generating new link predictions, which not only speeds up the discovery process but also reduces manual effort. The automated generation of link predictions ensures a growing range of innovative ideas, moving beyond the static nature of

databases like AskNature and AskNatureNet. Additionally, our model's capability to predict new relationships introduces fresh perspectives, fostering creative thinking. Unlike AskNature and AskNatureNet, which provide a platform for passive exploration, our model actively generates new, relevant design inspirations that encourage designers to think innovatively.

7. Conclusion

This research introduces a novel approach that utilises knowledge graphs in bio-inspired design (BID), serving as an ideation tool for the BID community. By constructing a comprehensive ontology and a knowledge graph specific to BID, this approach represents the complex relationships within the BID knowledge system. A link prediction model is implemented on the BID knowledge graph, demonstrating its potential in BID knowledge reasoning and stimulating innovative solutions. Two case studies are conducted to examine the effectiveness of this approach, showing that it outperforms in retrieving related biological knowledge and generating relevant and inspirational design ideas.

However, there are some limitations to the current approach. One limitation is the size and quality of the knowledge graph. While the present knowledge graph has offered valuable insights, the reasoning results might not be optimal. The quality and size of the knowledge graph can be further improved. Expanding the knowledge graph with more diverse and high-quality data would likely enhance the reasoning capabilities and yield better results. Another limitation is that our two-phase retrieval method requires users to condense the design problem into a singular term. While this single term ensures a more focused retrieval, it might pose challenges for users who are attempting to articulate complex or multifaceted design problems. This condensing could potentially limit broader applications. Future studies could further enhance this knowledge graph-based approach by expanding the size, automatic updating of knowledge and advanced application prediction. This research provides a foundation for future explorations in applying knowledge graphs into BID, emphasising the promising possibilities of this knowledge representation technique.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by National Key R&D Program of China R&D Program of China: [Grant Number 2022YFB3303304]; The Ng Teng Fong Charitable Foundation in the form of ZJU-SUTD IDEA Grant: [Grant Number 188170-11102].

ORCID

Zebin Cai  <http://orcid.org/0000-0002-2482-1765>

References

- Baltrušaitis, Tadas, Chaitanya Ahuja, and Louis-Philippe Morency. 2018. "Multimodal Machine Learning: A Survey and Taxonomy." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41 (2): 423–443. <https://doi.org/10.1109/TPAMI.2018.2798607>.
- Bar-Cohen, Yoseph. 2005. *Biomimetics: Biologically Inspired Technologies*. Florida: CRC press.

- BionicInspiration.org. 2014. "BionicInspiration.org: A Website Dedicated to Promoting Bio-Inspired Design." Accessed 01-November. <https://www.bionicinspiration.org/>.
- Bizer, Christian, Jens Lehmann, Georgi Kobilarov, Sören Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. 2009. "DBpedia – A Crystallization Point for the Web of Data." *Journal of web Semantics* 7 (3): 154–165. <https://doi.org/10.1016/j.websem.2009.07.002>.
- Bordes, Antoine, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. "Translating Embeddings for Modeling Multi-Relational Data." *Advances in Neural Information Processing Systems* 26.
- Brown, Tom, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, and Amanda Askell. 2020. "Language Models are Few-Shot Learners." *Advances in Neural Information Processing Systems* 33: 1877–1901.
- Chakrabarti, Amaresh, Prabir Sarkar, B. Leelavathamma, and B. S. Nataraju. 2005. "A Functional Representation for Aiding Biomimetic and Artificial Inspiration of new Ideas." *Ai Edam* 19 (2): 113–132.
- Chandrasegaran, Senthil K, Karthik Ramani, Ram D Sriram, Imré Horváth, Alain Bernard, Ramy F Harik, and Wei Gao. 2013. "The Evolution, Challenges, and Future of Knowledge Representation in Product Design Systems." *Computer-aided Design* 45 (2): 204–228. <https://doi.org/10.1016/j.cad.2012.08.006>.
- Chen, Liuqing, Zebin Cai, Zhaojun Jiang, Qi Long, Lingyun Sun, Peter Childs, and Haoyu Zuo. 2023. "A Knowledge-Based Ideation Approach For Bio-Inspired Design." *Proceedings of the Design Society* 3: 231–240. <https://doi.org/10.1017/pds.2023.24>.
- Chen, Xiaojun, Shengbin Jia, and Yang Xiang. 2020. "A Review: Knowledge Reasoning Over Knowledge Graph." *Expert Systems with Applications* 141: 112948. <https://doi.org/10.1016/j.eswa.2019.112948>.
- Daly, Shanna R, Colleen M Seifert, Seda Yilmaz, and Richard Gonzalez. 2016. "Comparing Ideation Techniques for Beginning Designers." *Journal of Mechanical Design* 138 (10): 101108. <https://doi.org/10.1115/1.4034087>.
- d'Aquin, Mathieu, and Enrico Motta. 2011. "Watson, More Than a Semantic web Search Engine." *Semantic Web* 2 (1): 55–63. <https://doi.org/10.3233/SW-2011-0031>.
- Deldin, Jon-Michael, and Megan Schuknecht. 2013. "The AskNature Database: Enabling Solutions in Biomimetic Design." In *Biologically Inspired Design: Computational Methods and Tools*, 17–27. London: Springer.
- Ehrlinger, Lisa, and Wolfram Wöß. 2016. "Towards a Definition of Knowledge Graphs." *SEMANTICS (Posters, Demos, SuCESS)* 48 (1–4): 2.
- Elhammadi, Sarah, Laks VS Lakshmanan, Raymond Ng, Michael Simpson, Baoxing Huai, Zhefeng Wang, and Lanjun Wang. 2020. "A High Precision Pipeline for Financial Knowledge Graph Construction." *Proceedings of the 28th International Conference on Computational Linguistics*.
- Fayemi, Pierre-Emmanuel, Kristina Wanieck, Cordt Zollfrank, Nicolas Maranzana, and Améziiane Aousat. 2017. "Biomimetics: Process, Tools and Practice." *Bioinspiration & Biomimetics* 12 (1): 011002. <https://doi.org/10.1088/1748-3190/12/1/011002>.
- Goel, Ashok, Kaylin Hagopian, Shimin Zhang, and Spencer Rugaber. 2022. "Towards a Virtual Librarian for Biologically Inspired Design." In *Design Computing and Cognition'20*, 369–386. Cham: Springer International Publishing.
- Goel, Ashok K., Daniel A. McAdams, and Robert B. Stone. 2015. *Biologically Inspired Design*. London: Springer.
- Goel, Ashok K., Swaroop Vattam, Bryan Wiltgen, and Michael Helms. 2012. "Cognitive, Collaborative, Conceptual and Creative – Four Characteristics of the Next Generation of Knowledge-Based CAD Systems: A Study in Biologically Inspired Design." *Computer-Aided Design* 44 (10): 879–900. <https://doi.org/10.1016/j.cad.2011.03.010>.
- Gruber, Thomas R. 1995. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing?" *International Journal of Human-Computer Studies* 43 (5-6): 907–928. <https://doi.org/10.1006/ijhc.1995.1081>.
- Gutiérrez, Bernal Jiménez, Nikolas McNeal, Clay Washington, You Chen, Lang Li, Huan Sun, and Yu Su. 2022. "Thinking About GPT-3 In-Context Learning for Biomedical IE? Think Again." *arXiv preprint arXiv:2203.08410*.

- Helfman Cohen, Yael, Yoram Reich, and Sara Greenberg. 2014. "Biomimetics: Structure–Function Patterns Approach." *Journal of Mechanical Design* 136 (11): 111108. <https://doi.org/10.1115/1.4028169>.
- Helms, Michael, Swaroop S Vattam, and Ashok K Goel. 2009. "Biologically Inspired Design: Process and Products." *Design Studies* 30 (5): 606–622. <https://doi.org/10.1016/j.destud.2009.04.003>.
- Hobern, Donald, Saroj K Barik, Les Christidis, Stephen T. Garnett, Paul Kirk, Thomas M Orrell, Thomas Pape, Richard L Pyle, Kevin R Thiele, and Frank E Zachos. 2021. "Towards a Global List of Accepted Species VI: The Catalogue of Life Checklist." *Organisms Diversity & Evolution* 21 (4): 677–690. <https://doi.org/10.1007/s13127-021-00516-w>.
- Huet, Armand, Romain Pinquié, Philippe Véron, Antoine Mallet, and Frédéric Segonds. 2021. "CACDA: A Knowledge Graph for a Context-Aware Cognitive Design Assistant." *Computers in Industry* 125: 103377. <https://doi.org/10.1016/j.compind.2020.103377>.
- Jiang, Shuo, Jie Hu, Kristin L. Wood, and Jianxi Luo. 2022. "Data-driven Design-by-Analogy: State-of-the-art and Future Directions." *Journal of Mechanical Design* 144 (2): 020801. <https://doi.org/10.1115/1.4051681>.
- Jiang, Guoqian, Harold R Solbrig, Eric Prud'hommeaux, Cui Tao, Chunhua Weng, and Christopher G Chute. 2015. "Quality Assurance of Cancer Study Common Data Elements Using a Post-Coordination Approach." AMIA Annual Symposium Proceedings.
- Kim, Sun-Joong, and Ji-Hyun Lee. 2017. "A Study on Metadata Structure and Recommenders of Biological Systems to Support bio-Inspired Design." *Engineering Applications of Artificial Intelligence* 57: 16–41. <https://doi.org/10.1016/j.engappai.2016.10.003>.
- Kozaki, Kouji, and Riichiro Mizoguchi. 2014. "An Ontology Explorer for Biomimetics Database".
- Lepora, Nathan F, Paul Verschure, and Tony J Prescott. 2013. "The State of the art in Biomimetics." *Bioinspiration & Biomimetics* 8 (1): 013001. <https://doi.org/10.1088/1748-3182/8/1/013001>.
- Li, Xinyu, Chun-Hsien Chen, Pai Zheng, Zuoxu Wang, Zuhua Jiang, and Zhixing Jiang. 2020. "A Knowledge Graph-Aided Concept–Knowledge Approach for Evolutionary Smart Product–Service System Development." *Journal of Mechanical Design* 142 (10): 101403. <https://doi.org/10.1115/1.4046807>.
- Li, Xinyu, Mengtao Lyu, Zuoxu Wang, Chun-Hsien Chen, and Pai Zheng. 2021. "Exploiting Knowledge Graphs in Industrial Products and Services: A Survey of key Aspects, Challenges, and Future Perspectives." *Computers in Industry* 129: 103449. <https://doi.org/10.1016/j.compind.2021.103449>.
- Li, Xinyu, Fei Zhang, Qi Li, Bin Zhou, and Jinsong Bao. 2023. "Exploiting a Knowledge Hypergraph for Modeling Multi-Nary Relations in Fault Diagnosis Reports." *Advanced Engineering Informatics* 57: 102084. <https://doi.org/10.1016/j.aei.2023.102084>.
- Liben-Nowell, David, and Jon Kleinberg. 2003. "The Link Prediction Problem for Social Networks." Proceedings of the Twelfth International Conference on Information and Knowledge Management.
- Lin, Yankai, Zhiyuan Liu, Huanbo Luan, Maosong Sun, Siwei Rao, and Song Liu. 2015. "Modeling Relation Paths for Representation Learning of Knowledge Bases." *arXiv preprint arXiv:1506.00379*.
- Luo, S. J., Y. F. Zhang, Ze Bian, and S. Pin. 2018. "Status and Progress of Product Shape Bionic Design." *Chinese Journal of Mechanical Engineering* 54: 138–155.
- Montecchi, Tiziano, Davide Russo, and Ying Liu. 2013. "Searching in Cooperative Patent Classification: Comparison Between Keyword and Concept-Based Search." *Advanced Engineering Informatics* 27 (3): 335–345. <https://doi.org/10.1016/j.aei.2013.02.002>.
- Müller, R., Abaid, N., Boreyko, J.B., Fowlkes, C., Goel, A.K., Grimm, C., Jung, S., Kennedy, B., Murphy, C., Cushing, N.D. and Han, J.P. 2018. "Biodiversifying Bioinspiration." *Bioinspiration & Biomimetics* 13 (5):053001.
- Nagel, Robert L, Prem A Midha, Andrea Tinsley, Robert B Stone, Daniel A McAdams, and L. H. Shu. 2008. "Exploring the Use of Functional Models in Biomimetic Conceptual Design."
- Nickel, Maximilian, Kevin Murphy, Volker Tresp, and Evgeniy Gabrilovich. 2015. "A Review of Relational Machine Learning for Knowledge Graphs." *Proceedings of the IEEE* 104 (1): 11–33. <https://doi.org/10.1109/JPROC.2015.2483592>.
- Paulheim, Heiko. 2017. "Knowledge Graph Refinement: A Survey of Approaches and Evaluation Methods." *Semantic Web* 8 (3): 489–508. <https://doi.org/10.3233/SW-160218>.

- Pradel, Patrick, Zicheng Zhu, Richard Bibb, and James Moultrie. 2018. "A Framework for Mapping Design for Additive Manufacturing Knowledge for Industrial and Product Design." *Journal of Engineering Design* 29 (6): 291–326. <https://doi.org/10.1080/09544828.2018.1483011>.
- Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. "Language Models are Unsupervised Multitask Learners." *OpenAI Blog* 1 (8): 9.
- Raffel, Colin, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2020. "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer." *The Journal of Machine Learning Research* 21 (1): 5485–5551.
- Shah, Jami J, Steve M Smith, and Noe Vargas-Hernandez. 2003. "Metrics for Measuring Ideation Effectiveness." *Design Studies* 24 (2): 111–134. [https://doi.org/10.1016/S0142-694X\(02\)00034-0](https://doi.org/10.1016/S0142-694X(02)00034-0).
- Shen, Xingwang, Xinyu Li, Bin Zhou, Yanan Jiang, and Jinsong Bao. 2023. "Dynamic Knowledge Modeling and Fusion Method for Custom Apparel Production Process Based on Knowledge Graph." *Advanced Engineering Informatics* 55: 101880. <https://doi.org/10.1016/j.aei.2023.101880>.
- Siddharth, L., and Amaresh Chakrabarti. 2018. "Evaluating the Impact of Idea-Inspire 4.0 on Analogical Transfer of Concepts." *AI Edam* 32 (4): 431–448.
- Song, Binyang, Bowen Yan, Giorgio Triulzi, Jeffrey Alstott, and Jianxi Luo. 2019. "Overlay Technology Space map for Analyzing Design Knowledge Base of a Technology Domain: The Case of Hybrid Electric Vehicles." *Research in Engineering Design* 30: 405–423. <https://doi.org/10.1007/s00163-019-00312-w>.
- Studer, Rudi, V. Richard Benjamins, and Dieter Fensel. 1998. "Knowledge Engineering: Principles and Methods." *Data & Knowledge Engineering* 25 (1-2): 161–197. [https://doi.org/10.1016/S0169-023X\(97\)00056-6](https://doi.org/10.1016/S0169-023X(97)00056-6).
- Sun, Yizhou, Rick Barber, Manish Gupta, Charu C Aggarwal, and Jiawei Han. 2011. "Co-author Relationship Prediction in Heterogeneous Bibliographic Networks." 2011 International Conference on Advances in Social Networks Analysis and Mining.
- Veličković, Petar, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. "Graph attention networks." *arXiv preprint arXiv:1710.10903*.
- Vicknair, Chad, Michael Macias, Zhendong Zhao, Xiaofei Nan, Yixin Chen, and Dawn Wilkins. 2010. "A Comparison of a Graph Database and a Relational Database: A Data Provenance Perspective." Proceedings of the 48th Annual Southeast Regional Conference.
- Vincent, Julian FV, Olga A Bogatyreva, Nikolaj R Bogatyrev, Adrian Bowyer, and Anja-Karina Pahl. 2006. "Biomimetics: Its Practice and Theory." *Journal of the Royal Society Interface* 3 (9): 471–482. <https://doi.org/10.1098/rsif.2006.0127>.
- Wan, Yuwei, Zhenyuan Chen, Fu Hu, Ying Liu, Michael Packianather, and Rui Wang. 2022. "Exploiting Knowledge Graph for Multi-Faceted Conceptual Modelling Using GCN." *Procedia Computer Science* 200: 1174–1183. <https://doi.org/10.1016/j.procs.2022.01.317>.
- Wang, Zuoxu, Chun-Hsien Chen, Pai Zheng, and Wenyan Xinyu Li. 2022. "A Hypergraph-Based Approach for Context-Aware Smart Product-Service System Configuration." *Computers & Industrial Engineering* 163: 107816. <https://doi.org/10.1016/j.cie.2021.107816>.
- Wang, Xiang, Xiangnan He, Fuli Feng, Liqiang Nie, and Tat-Seng Chua. 2018. "Tem: Tree-Enhanced Embedding Model for Explainable Recommendation." Proceedings of the 2018 World Wide web Conference.
- Wang, Zhen, Jianwen Zhang, Jianlin Feng, and Zheng Chen. 2014. "Knowledge Graph Embedding by Translating on Hyperplanes." Proceedings of the AAAI Conference on Artificial Intelligence.
- Wang, Boheng, Yunhuai Zhu, Liuqing Chen, Jingcheng Liu, Lingyun Sun, and Peter Childs. 2023. "A Study of the Evaluation Metrics for Generative Images Containing Combinational Creativity." *AI EDAM* 37: e11.
- Wanieck, Kristina, Pierre-Emmanuel Fayemi, Nicolas Maranzana, Cordt Zollfrank, and Shoshanah Jacobs. 2017. "Biomimetics and its Tools." *Bioinspired, Biomimetic and Nanobiomaterials* 6 (2): 53–66. <https://doi.org/10.1680/jbibr.16.00010>.
- Yen, Jeannette, Michael Helms, Ashok Goel, Craig Tovey, and Marc Weissburg. 2014. "Adaptive Evolution of Teaching Practices in Biologically Inspired Design." *Biologically Inspired Design: Computational Methods and Tools*, 153–199.

Yoon, Byungun, and Yongtae Park. 2004. "A Text-Mining-Based Patent Network: Analytical Tool for High-Technology Trend." *The Journal of High Technology Management Research* 15 (1): 37–50. <https://doi.org/10.1016/j.hitech.2003.09.003>.

Zheng, Pai, Liqiao Xia, Chengxi Li, Xinyu Li, and Bufan Liu. 2021. "Towards Self-X Cognitive Manufacturing Network: An Industrial Knowledge Graph-Based Multi-Agent Reinforcement Learning Approach." *Journal of Manufacturing Systems* 61: 16–26. doi:<https://doi.org/10.1016/j.jmsy.2021.08.002>.

Zhou, Bin, Jie Li, Xinyu Li, Bao Hua, and Jinsong Bao. 2022. "Leveraging on Causal Knowledge for Enhancing the Root Cause Analysis of Equipment Spot Inspection Failures." *Advanced Engineering Informatics* 54: 101799. <https://doi.org/10.1016/j.aei.2022.101799>.

Zhu, Qihao, Xinyu Zhang, and Jianxi Luo. 2023. "Biologically Inspired Design Concept Generation Using Generative Pre-Trained Transformers." *Journal of Mechanical Design* 145 (4): 041409. <https://doi.org/10.1115/1.4056598>.

Zuo, Haoyu, Qianzhi Jing, Tianqi Song, Lingyun Sun, Peter Childs, and Liuqing Chen. 2022. "WikiLink: An Encyclopedia-Based Semantic Network for Design Creativity." *Journal of Intelligence* 10 (4): 103. <https://doi.org/10.3390/jintelligence10040103>.

Zuo, H., Y. Yin, and P. Childs. 2022. "Patent-KG: Patent Knowledge Graph Extraction for Engineering Design." *Proceedings of the Design Society* 2: 821–830.

Zuoxu, Wang, Liu Jihong, and Zheng Lianyu. 2022. "The Evolution, Framework, and Future of Cognitive Intelligence-Enabled Product Design." *Procedia CIRP* 109: 526–531. <https://doi.org/10.1016/j.procir.2022.05.289>.

Appendix

Table A1. Number of Categories for ‘Source’ nodes across seven biological taxonomy levels.

Kingdom	Phylum	Class	Order	Family	Genus	Species
7	50	66	204	364	288	223

Table A2. Distribution of ‘Benefits’ Categories.

Category (incentive perspectives)	Count
Function	1369
Shape	82
Structure	395
Colour	60
Texture	239
Imagery	30

Table A3. Distribution of ‘Application’ Categories (CPC Method).

Category	Count
Human Necessities	75
Performing operations; Transporting	18
Chemistry; Metallurgy	141
Textiles; Paper	17
Fixed constructions	19
Mechanical engineering; Lighting; Heating; Weapons; Blasting	117
Physics	11
Electricity	25
Emerging technology	135