Natural Language Hypotheses in Scientific Papers and How to Tame Them

Suggested Steps for Formalizing Complex Scientific Claims

Tina Heger[1,2,3(✉)], Alsayed Algergawy[4,5], Marc Brinner[6], Jonathan M. Jeschke[1,2], Birgitta König-Ries[4], Daniel Mietchen[1,2,7,8], and Sina Zarrieß[6]

[1] Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany
t.heger@tum.de
[2] Institute of Biology, Freie Universität Berlin, Berlin, Germany
[3] TUM School of Life Sciences, Technical University of Munich, Freising, Germany
[4] Institute for Informatics, Friedrich-Schiller-University Jena, Jena, Germany
[5] Data and Knowledge Engineering, University of Passau, Passau, Germany
[6] Faculty of Linguistics and Literature Studies, University of Bielefeld, Bielefeld, Germany
[7] Ronin Institute for Independent Scholarship, Montclair, NJ, USA
[8] Institute for Globally Distributed Open Research and Education (IGDORE), Jena, Germany

Abstract. Hypotheses are critical components of scientific argumentation. Knowing established hypotheses is often a prerequisite for following and contributing to scientific arguments in a research field. In scientific publications, hypotheses are usually presented for specific empirical settings, whereas the related general claim is assumed to be known. Prerequisites for developing argumentation machines for assisting scientific workflows are to account for domain-specific concepts needed to understand established hypotheses, to clarify the relationships between specific hypotheses and general claims, and to take steps towards formalization. Here, we develop a framework for formalizing hypotheses in the research field of invasion biology. We suggest conceiving hypotheses as consisting of three basic building blocks: a subject, an object, and a hypothesized relationship between them. We show how the subject-object-relation pattern can be applied to well-known hypotheses in invasion biology and demonstrate that the contained concepts are quite diverse, mirroring the complexity of the research field. We suggest a stepwise approach for modeling them to be machine-understandable using semantic web ontologies. We use the SuperPattern Ontology to categorize hypothesized relationships. Further, we recommend treating every hypothesis as part of a hierarchical system with 'parents' and 'children'. There are three ways of moving from a higher to a lower level in the hierarchy: (i) specification, (ii) decomposition, and (iii) operationalization. Specification involves exchanging subjects or objects. Decomposition means zooming in and making explicit assumptions about underlying (causal) relationships. Finally, operationalizing a hypothesis means providing concrete descriptions of what will be empirically tested.

Keywords: Complex claims · invasion biology · ontology · scientific hypotheses

P. Cimiano et al. (Eds.): RATIO 2024, LNAI 14638, pp. 3–19, 2024.
https://doi.org/10.1007/978-3-031-63536-6_1

1 Introduction: Scientific Hypotheses as Complex Claims

In scientific contexts, argumentation is part of established workflows. In an idealized setting, a research question arises from some applied context or a scientific debate. Based on this question, the researcher formulates a hypothesis that expresses a relationship between domain-specific concepts and can be tested empirically. Experiments or surveys are conducted by measuring the variables or testing the conditions posited in the hypothesis, and the results are reported together with the empirical methods and the tested hypothesis in a scientific publication. In such scientific settings, a carefully developed and thought-through hypothesis (which we see as Toulmin's [1] "claim" in a scientific context) is at the core of the argumentation process. This hypothesis must be specific enough for a researcher to test it empirically. Still, at the same time, it should also relate to previous general claims made in the community. In actual scientific publications, the relationship between a hypothesis explicitly formulated for the study's context and the general claim it is based on is often neither made explicit nor obvious [2]. Also, hypotheses are usually given as complex statements that include scientific and colloquial terms, and the meaning of both can be ambiguous [3]. For instance, the term "resistance" is used with slightly different meanings by different authors, even within a given domain, and terms like "often" are interpreted differently by different readers. Consequently, scientific hypotheses are a challenging case for modeling, as workflows are required for aligning complex claims with generic structures while at the same time leaving room for the inclusion of domain-specific concepts and knowledge.

While some suggestions for modeling scientific hypotheses already exist (see Sect. 2), they are usually hardly accessible to scientists outside the argumentation community. On the other hand, for experts in formal argumentation, computational linguistics, and semantic modeling, it is not always obvious how best to connect the available tools and approaches to workflows in empirical sciences. A solution to this challenge is the formation of interdisciplinary teams. With this publication, we want to share results from a project that brought together domain experts (in this case, invasion biologists) with experts from semantic modeling and computational linguistics [4]. Our project aims to explore how natural language processing (NLP) and semantic modeling can be leveraged to enhance workflows in scientific research. More specifically, our long-term goal is the automated synthesis of research results testing scientific hypotheses in invasion biology and other domains. To achieve this goal, it is necessary to develop methods for linking scientific papers reporting on empirical tests to major hypotheses relevant to the respective domain.

A prerequisite for such an automated linking of empirical tests to hypotheses is the formalization of hypothesis statements. In this paper, we introduce a framework for transferring hypotheses given in scientific papers in the form of natural language statements into more formalized statements. We use examples from the domain of invasion biology to demonstrate how the framework can help clarify the relationships between the general hypotheses put forward in scientific debates and specific, complex hypotheses directly relating to empirical studies. This paper aims to report on our interdisciplinary efforts to combine domain-specific knowledge of needs and challenges with expert knowledge of tools and approaches from semantic modeling and NLP. The resulting framework is

meant as a guideline to be used by experts in a scientific domain who work on synthesizing the knowledge of their field.

In the following, we first give an overview of related work. Next, we introduce our working example and use that to introduce our suggestion for moving towards a formalization of scientific hypotheses. We then report on ongoing applications of the framework. We point out the limitations of our approach and close with an outlook.

2 Related Work

Our suggestions are based on and related to past and ongoing work in the fields of argumentation modeling, knowledge representation, and invasion biology.

2.1 Argumentation Modeling for Complex Scientific Claims

Argumentation is studied in different fields and disciplines, like philosophy, computer science, computational linguistics, and more domain-oriented disciplines like biology. Especially in philosophy, computational linguistics, and NLP, a common approach is to develop abstract representations of arguments and argumentation processes to understand communication processes and how dissent and consensus form. In this context, "toy arguments" are often used to demonstrate the applicability of the respective abstract and formalized argumentation schemes (e.g., [1, 5]). A complementary approach uses AI-based tools for mining arguments in large amounts of data containing informal, primarily textual statements of real-world arguments (see this survey: [6]). While formal accounts are often difficult to apply and to scale up to complex real-world arguments, data-driven argument mining usually does not account for formal aspects of arguments formulated in text.

Regarding formal argumentation analysis, few studies have focused on scientific literature. One example is [7], where the authors suggest an explanatory argumentation framework (EAF) for representing argumentation processes among scientists. In that case, the goal was to model the conceptual structure of the main arguments brought forward by different agents in a scientific debate. The focus of this approach is not so much on the relationship between general and specific claims, nor is the aim to guide hypothesis formulation or identifying hypotheses in texts.

2.2 Knowledge Representation: Modeling Scientific Language with Knowledge Graphs

Semantic Web techniques provide ways to formalize knowledge. On the one hand, this allows machines to act on information; on the other hand, this supports humans in providing concrete representations (e.g., making hidden assumptions and subtle differences in understanding explicit). Knowledge graphs are one such approach that is widely regarded as very promising. They are successfully used in industry but also in scientific settings. In knowledge graphs, nodes represent entities of interest, while edges represent relations between these entities. The graphs are encoded in a (typically machine-actionable) graph data model [8]. One example of their application to model

scientific language is [9]. They suggest representing evidence from empirical studies in neuroscience in the form of Research Maps[1]. Here, hypothesized causal relationships are represented as directed graphs, where each node gives the identity and properties of a biological phenomenon. Experimental evidence can be fed into the graphs, allowing to visually represent alignment or disagreement between hypotheses and evidence. This approach, however, focuses on representing the results of empirical work. Consequently, the scheme does not allow for clarifying hierarchical links between complex hypotheses tailored to empirical settings and general, major claims. Also, the aim is to provide templates that researchers can fill out to report their results in a machine-actionable format; the framework is not intended to enhance argument analysis in textual publications.

With a specific focus on formalizing scientific hypotheses, [10] suggested the DISK framework and ontology. DISK was designed to enable automated discovery, hypothesis testing, and revision. As in the case of the Research Maps framework, the focus is on modeling results from empirical studies. Therefore, modeling hierarchical relationships between hypotheses is not straightforward in this setting. Also, as far as we know, the framework has not been implemented and used. It remains unclear how DISK could be used to discover complex versions of hypotheses in actual scientific publications.

Since natural language hypothesis statements can be pretty complex, a stepwise approach towards formalization is practical. The AIDA language suggested by [11] offers a first-step method. This method translates natural language statements into atomic, independent, declarative, and absolute sentences. Such sentences can then derive valid nodes in a knowledge graph.

2.3 Hypothesis Representation in Invasion Biology

Invasion biology studies human-induced transport, introduction, establishment, spread, and impact of organisms. Due to global transport and trade, many species have been translocated to areas outside their natural range [12]. Research in this field is concerned with identifying mechanisms of invasions, often motivated by the goal of developing management solutions. The field is of particular interest in the context of argumentation because numerous major hypotheses have been formulated over time on why species can establish and spread [13–15] (Table A1). This allows for identifying sets of scientific publications that argue for or against one of these hypotheses [16]. Such sets can then be used to develop and test methods for argumentation analysis [17].

In previous work, Heger, Jeschke, and colleagues suggested the hierarchy-of-hypotheses (HoH) approach, according to which scientific hypotheses can be represented as hierarchies [2, 16]. In an HoH, a broad, general claim is given as an overarching hypothesis at the top level, which branches out into more specific versions or sub-hypotheses forming the lower levels. These sub-hypotheses either specify how research on that overarching question has been implemented ('operational hypotheses') or represent conceptual refinements, which can be either specification (e.g., spelling out factors that could have caused an effect) or decompositions (e.g., illustrating the partial arguments contained in a broad claim). Concerning the latter, [18] has suggested that it

[1] https://researchmaps.org/.

can be helpful to represent mechanistic hypothesis refinements as causal network diagrams. Decomposition then means adding nodes to a causal chain or network. In the following, we build on these ideas for a stepwise formalization of complex scientific claims.

3 Example: The Biotic Resistance Hypothesis

To demonstrate the challenges connected to treating hypotheses as complex claims in a real-world setting, we give an example of one of the major hypotheses suggested as a potential explanation for the successful establishment and spread of invasive species, namely the Biotic Resistance Hypothesis. In its general version, this hypothesis posits that "*An ecosystem with high biodiversity is more resistant against non-native species than an ecosystem with lower biodiversity*" [19]. In scientific papers, however, such a general formulation is rarely used [17]. Instead, authors of scientific papers tend to use formulations that directly account for the particular case they chose to study and the specific experimental setting. For example, a publication presenting results from empirical tests of the Biotic Resistance Hypothesis used the following formulations: *"…species already in the community with similar functional traits to those of the invaders should have the greatest competitive effect on invaders.*" "*We used experimental communities in a serpentine grassland in California, USA, to assess the extent to which […] functional diversity influenced success of two different types of invading plants: early-season annuals (E) and late-season annuals (L)[…]*". [20].

Such complex statements, differing significantly from the general claim they relate to, can be pretty hard to identify for standard NLP classifiers [17]. Even for scientists, at least those not familiar with the respective claim and underlying theory (e.g., freshly starting Ph.D. students), the link of these complex statements to the major hypothesis is often hard to recognize. An argumentation machine assisting the understanding of such complex claims and aiding the development of own related hypotheses would therefore be helpful [4]. However, this requires developing a framework for formalizing scientific hypotheses and clarifying links between general and specific hypothesis formulations. In the following, we present a suggestion for such a framework.

4 Towards Formalizing Scientific Hypotheses

Our suggestion involves several steps (Fig. 1). Natural language statements of general hypotheses are reformulated into AIDA statements [11] by domain experts. These statements are subsequently translated into further formalized statements of the form subject–relationship–object. A classification scheme allows linking the general statement to the specific claims, and ontologies specify their components. Further, NLP classifiers are used to identify general and specific hypothesis statements in texts (this step is described in [17, 21]).

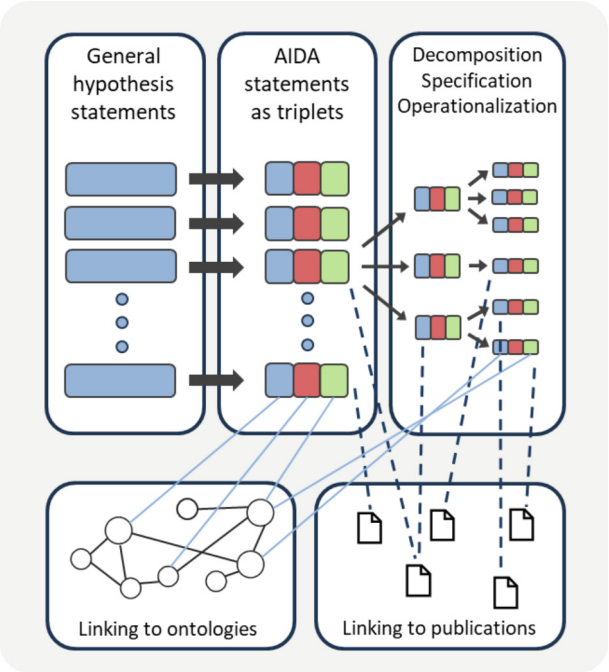

Fig. 1. Suggested workflow for developing semi-formal hypothesis statements and clarifying links between general hypotheses and hypothesis statements in scientific texts.

4.1 A Generic Structure for Scientific Hypotheses

Moving towards a formalized representation of scientific hypotheses in invasion biology, starting with broad, major hypotheses, is helpful because these are usually less complex than the refined versions formulated in papers reporting on empirical tests. Taking ten major hypotheses in invasion biology as examples, the invasion biology experts amongst the author group translated the textual versions (Table A1) into AIDA statements, following the methodology suggested in [11]. From these, the domain experts developed formalized versions consisting of a subject, an object, and a hypothesized relationship between these two (analogously to the familiar format of subject-predicate-object triples in edge-labeled graphs, e.g., knowledge graphs encoded in the RDF data model). The subject and the object are often complex in themselves, and we introduced further formalization by distinguishing the core variable, a qualifier for cases in which the core variable has qualitatively distinct states, and a term giving further context concerning settings in which the statement holds (Table A2).

4.2 Linking Hypothesis Formulations to Semantic Models

A critical element in moving from natural language formulations of hypotheses to formalized statements is linking the constituting concepts to entities in machine-actionable

ontologies. We suggest using the SuperPattern Ontology[2] [22] to model the hypothesized relationships between subject and object as well as the qualifiers. It contains a set of relations useful for describing causal relationships (e.g., "*contributes to*", "*prevents*", "*inhibits*") and comparisons (e.g., "*has smaller value than*", "*has larger value than*").

Some invasion biology hypotheses are initially given in a comparative form. This is the case for the Biotic Resistance Hypothesis but also for Darwin's Naturalization Hypothesis, the Disturbance Hypothesis, the Island Susceptibility Hypothesis, the Limiting Similarity Hypothesis, and the Phenotypic Plasticity Hypothesis (Table A1). The underlying ideas, however, refer to causal relationships. In these cases, we suggest that both variants can be helpful, the comparative version that is close to the original textual definition and an additional causal version referring to the underlying causal reasoning (Table A2). We think of the comparative versions as some kind of operationalization: In an empirical setting, comparative claims are usually easier to test than causal claims since the former do not necessarily demand to implement experiments. We suggest formalizing the causal variants of the hypotheses in such a way that the subject always gives the invasion driver, i.e., the factor hypothesized to be the underlying force behind a biological invasion or its impacts. The object describes the expected invasion outcome.

As Table A2 demonstrates for the ten hypotheses, the variables and the terms giving context for each subject and object are complex, with little overlap in the used concepts or terms (an exception being "*invasion success*"). This mirrors the complexity of the scientific field of invasion biology, with many potentially influential factors. We, therefore, chose a stepwise approach for modeling them in an ontology created explicitly for this purpose, i.e., the Invasion Biology Ontology INBIO [23]. First, we obtained expert opinion to identify core terms in each of the ten hypotheses. For the Biotic Resistance Hypothesis, these terms were "*ecosystem*", "*biodiversity*" and "*species*". Next, we searched for existing ontologies containing these terms; where this was successful, we used a fusion/merge strategy to integrate respective modules into the INBIO [24]. In further steps, more concepts have been added to provide full conceptual models of the subjects and objects of the ten hypotheses.

The suggested generic structure does not necessarily capture the structure of all scientific hypotheses, but we suggest it can be beneficial for hypotheses describing causal relationships. Hypotheses representing generalized statistical claims (descriptive or statistical hypotheses [25]) do not necessarily follow this form. In our set of ten hypotheses, this was the case for the Tens Rule, which posits that "*Approximately 10% of species successfully take consecutive steps of the invasion process*" (Tables A1 and A2).

4.3 Classifying Relationships Between General and Specific Claims

The previous two subsections have described steps toward formalizing broad, overarching hypotheses. A next step that we consider necessary for linking these formalized versions of major hypotheses to actual hypothesis statements in publications reporting on empirical tests is to clarify the relationship between the overarching hypotheses and the refined sub-hypotheses. Building on the HoH approach, we suggest treating every

[2] https://larahack.github.io/linkflows_superpattern/doc/sp/index-en.html.

hypothesis as a component of a hierarchical system with 'parents' and 'children'. As described in [2], we recommend distinguishing between three kinds of refinements: (A) decomposition, (B) specification, and (C) operationalization (Fig. 2).

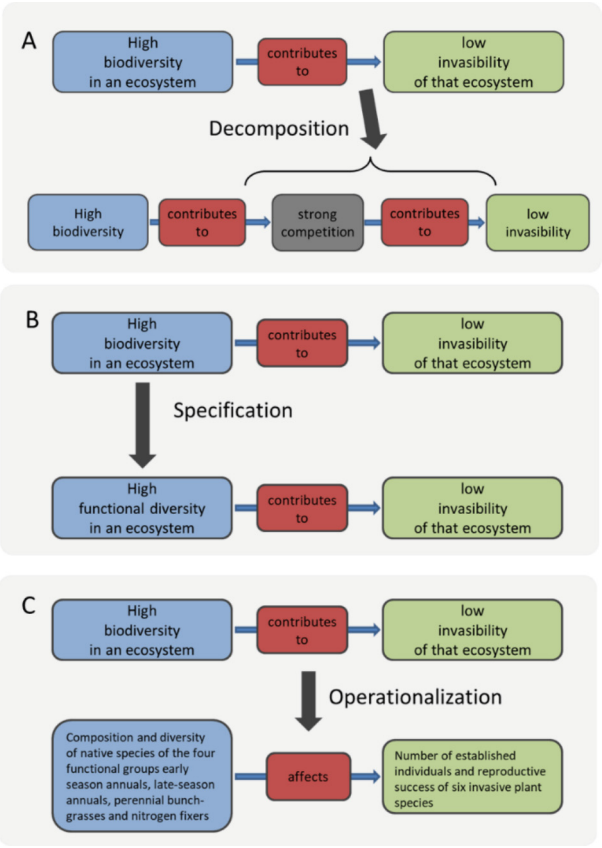

Fig. 2. Three approaches for relating general versions of scientific hypotheses to more specific ones, demonstrated with the example of the Biotic Resistance Hypothesis in invasion biology: (A) decomposition, (B) specification, and (C) operationalization. See the main text for more information.

With decomposition, we denote the process of making those causal relationships explicit, which are implicit parts of the reasoning behind a hypothesis (see [18] for a worked example of the Enemy Release Hypothesis). Coming back to the Biotic Resistance Hypothesis, the general definition points out the *negative effects of high biodiversity on invasion success*, whereas [20] hypothesizes a *competitive effect of native species on invaders*. An expert in invasion biology can draw from background knowledge to make the connection. For such an expert, it will be evident that intense competition affects invasion success. The refinement of the Biotic Resistance Hypothesis in [20] thus adds

nodes to the hypothesized causal graph, making more of the hypothesized mechanism explicit (Fig. 2A).

In the above example, the authors additionally applied the specification strategy. Specifying a hypothesis involves exchanging the nodes of the hypothesized causal chain or network with more concrete versions (Fig. 2B). In the cited example, instead of testing for a general effect of high biodiversity on the chosen invasive species, the authors tested for functional diversity effects. By functional diversity, the authors meant the presence or absence of plant species representing one of four groups that differ in their ecological behavior, namely early-season plants with an annual life cycle, late-season species with an annual life cycle, grasses growing in bunches and living longer than one year, and herbs with the ability to fix atmospheric nitrogen.

The third possibility in which a specific version of a hypothesis can be linked to its general version is operationalization. To operationalize a hypothesis means to describe what exactly will be empirically tested. In the described case, the authors chose to examine the effects of manipulating the composition and diversity of native species of the four functional groups (early-season annuals, late-season annuals, perennial bunchgrasses, and nitrogen fixers). Their dependent variable or 'object' was the number of established individuals and the reproductive success of six selected invasive plant species from those groups (Fig. 2C).

The described operations can also be applied in the other direction. For example, a hypothetical complex causal chain or network can be simplified, which would be the inverse of decomposition (Fig. 2A). An existing hypothesis, perhaps derived from studying a specific context, can be generalized to a broader context (e.g., in terms of taxa or life stages covered, geographic range or other ecological gradients); this would be the opposite of specification (Fig. 2B). Finally, from a hypothesis generated, e.g., from an empirical observation under specified experimental conditions, a broader, more abstract version can be derived; such an abstraction would be the opposite of an operationalization (Fig. 2C).

The suggested scheme can be a basis for linking actual hypothesis statements in publications reporting on empirical tests to major, more general hypotheses [2, 16]. For example, in their literature review on the Biotic Resistance Hypothesis, [19] identified 15 empirical studies that focused on functional diversity as a specific form of biodiversity, whereas 126 empirical tests in their dataset instead studied species richness, which is a different specification of biodiversity.

To allow for the implementation of the framework in the context of argumentation analysis, we are currently developing a Hypothesis Ontology containing the concepts identified as hypothesis components and the possible relationships between general and specific variants, as just described. Figure 3 depicts the already developed modeling of types of entities and their relationships; adding concrete instances belonging to these types (e.g., the Biotic Resistance Hypothesis as one specific Hypothesis) is ongoing work. In this model, a Hypothesis is linked to a HypothesisDefinition. The definition "*An ecosystem with high biodiversity is more resistant against non-native species than an ecosystem with lower biodiversity*" [19] will be one instance of the type HypothesisDefinition. The distinction between the Hypothesis and the HypothesisDefinition

is necessary, as several subtly different definitions exist for many high-level hypotheses. Each of these definitions is further captured in a HypothesisStatement. We model HypothesisStatements as SuperPatternInstances [22]. They possess a Label, Context, Subject, Relations, Objects, and Qualifiers. Subjects and Objects can be complex and consist of Qualifiers, Variables, and Contexts. Hypothesis and HypothesisDefinitions can have subclass relationships to reflect the hierarchical structure described above. A Hypothesis can be supported (or refuted) by Evidence and equipped with Provenance as defined in the Prov-O ontology[3].

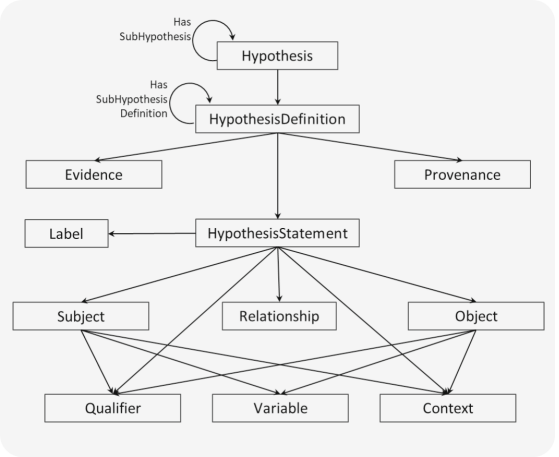

Fig. 3. Conceptual scheme for the Hypothesis Ontology

5 Applications of the Framework

The current situation in which major scientific research results are mainly published in PDF format hinders the integration of AI technology in scientific workflows [26]. An important step towards overcoming this barrier would be to enrich the bibliographic meta-data of scientific publications with machine-readable information about the publications' content, including studied hypotheses. The suggested framework and related semantic modeling can provide a basis for such endeavors. We are currently exploring two parallel pathways in this direction. The first of these pathways involves using Wikidata to link entries about publications to entries about hypotheses, while the second introduces hypotheses as a publication type in its own rights.

The Wikidata pathway builds on community curation workflows under the umbrella of the WikiCite initiative that collects bibliographic metadata in Wikidata [27]. It further involves the development of tools for exploring the resulting knowledge graph (e.g. [28]). In this context, we regularly identify invasion biology publications and annotate them as such, with additional workflows to annotate the identified publications for author

[3] https://www.w3.org/TR/prov-o/.

disambiguation, main subjects, or methods used. For each hypothesis to be used in this workflow, a dedicated Wikidata entry is required, and we have created such entries for the most common hypotheses in invasion biology, including those listed in Table A1. These entries can then be annotated, e.g., in terms of the publications from which they originated or the concepts they relate to. The aim is to establish links between the hypotheses and scientific publications testing or discussing them. In the future, this will allow for better findability of relevant publications in an Open Science environment and options for on-demand meta-analyses [29].

For the second pathway, we developed a scheme for a new publication type - Hypothesis Descriptions [30]. Such Hypothesis Descriptions are aimed at formalizing how invasion biology hypotheses are described (especially in terms of which concepts and relationships they cover) and how differences between hypothesis variants can be expressed, both for humans and in a machine-actionable fashion. This scheme is pioneered in the open-science journal Research Ideas and Outcomes [31] and builds on the nanopublication standards beginning to be adopted in biodiversity-related publications [32].

In the context of invasion biology (and other fields of science), the suggested framework can further be used as a guideline for formulating hypotheses. In invasion biology, the ambiguity of hypothesis formulations is often considered challenging (see, e.g., [33]). Still, it is not an established practice to carefully consider the relationship between a specific, complex claim made in a publication and the general version it has been derived from or to use consistent language for formulating hypotheses. We suggest that our framework could offer guidance, thus enhancing research efficiency. For example, in the case of the Enemy Release Hypothesis, empirical research so far has mainly focused on only one of its components [33, 34]. However, to establish whether or not this hypothesis can be regarded as a reasonable explanation for invasion success, it would be necessary to study the complete hypothesized causal chain. Such gaps are more easily identified if respective publications clarify which kind of hypothesis refinement is chosen for the study context.

6 Limitations

Invasion biology, the research domain we used to develop our framework, is a relatively straightforward example because the domain is characterized by many explicitly formulated major hypotheses repeatedly synthesized by the scientific community [13–16]. In other disciplines, it might be much harder to even identify such general claims. For the neighboring discipline of urban ecology, [35] demonstrated how similar lists of major hypotheses can be collated with a combination of expert involvement and literature analyses. This general approach can, in principle, be applied to any other scientific domain. Also, we believe that the NLP models we develop based on the introduced hypothesis formalization can be used later for automatic/semi-automatic hypothesis discovery in other fields as well. The suggestion for linking specific formulations of empirical tests to general claims is also not limited to an application in invasion biology. [36] demonstrated how specific claims in medicine can be linked to a general, major claim by specification and operationalization. Still, future work is needed to clarify for which

scientific domains it is possible and useful to implement all steps towards hypothesis formalization outlined above.

Currently, it is an open question how our ontology-based, multi-level formalization of hypotheses can feed into NLP-based argument mining methods, i.e., hypothesis identification in particular [21]. While much recent work is on integrating language modeling and knowledge graphs, it is unclear how these methods scale to the complex problem of hypothesis identification in scientific papers, which requires deep semantic reasoning and domain-specific knowledge. In future work, relevant ontologies will be integrated with text-driven approaches to argument mining and enhance the implicit knowledge in language modeling-based approaches with explicit knowledge. This can be achieved, for instance, with recent methods for so-called "knowledge injection into language models", see [25].

Moving towards formalizing scientific hypotheses requires exchanging complex natural language with streamlined and unified terms and concepts. It is necessary to carefully study under which conditions the gain of formalizing outweighs the potential information losses during this process. This challenge can become even more demanding once the semi-formal statements suggested in Table A2 are further transformed, e.g., into logical statements that provide a foundation for automated reasoning. An annotation study could be a practical next step to help clarify how well our proposed scheme can capture complex hypothesis statements in actual scientific texts.

7 Conclusions and Outlook

In this article, we suggested a framework for moving towards a formalization of scientific hypotheses and clarifying links between general and specific hypothesis formulations. Developing the framework was an interdisciplinary effort, considering knowledge from invasion biology, philosophy of science, computational linguistics, and semantic modeling. We suggest our framework can be helpful for argumentation analysis in scientific publications. Further, it can help in taking steps towards reprocessing scientific publications and making published research available for AI-based analyses. Finally, the framework can guide researchers during the hypothesis formulation process. We suggest that domain experts can directly profit from our framework because it motivates to make intuitions explicit and fosters conceptual analysis, which can directly benefit the quality of scientific work [37].

Therefore, implementing the framework as a user interaction tool is an essential next step. A prototype of such a tool already exists, and a first version will soon be available at hi-knowledge.org[4]. The tool will help researchers identify major invasion hypotheses in texts, link to background information necessary for understanding technical terms, and, in the future, offer guidance to formulate their own specific and complex research hypothesis tailored to the focal empirical setting. Implementing AI-based tools in all steps of the scientific workflow is a timely and urgent need. This would significantly enhance efficiency [38] and allow for better utilization of knowledge gained in research for solving current societal challenges. We hope our framework will motivate and facilitate innovative steps in this direction.

[4] https://hi-knowledge.org/.

Acknowledgements. This work was funded by Deutsche Forschungsgemeinschaft DFG (project number 455913229; T.H., M.B., J.M.J., B.K-R, S.Z.) and the VolkswagenStiftung (97 863; T.H., J.M.J., D.M.). We thank Maud Bernard-Verdier, Camille Musseau, and Florencia Yannelli for their comments on the formalization framework. Three anonymous reviewers and a meta-reviewer provided comments that helped us to improve the text.

Appendix

Table A1. Ten major hypotheses in invasion biology and their textual definitions as given in [39].

Hypothesis	Acronym	Definition
Biotic resistance hypothesis	BR	An ecosystem with high biodiversity is more resistant against non-native species than an ecosystem with lower biodiversity
Darwin's naturalization hypothesis	DN	Invasion success of non-native species is higher in areas that are poor in closely related species than in areas that are rich in closely related species
Disturbance Hypothesis	DS	Success of non-native species is higher in highly disturbed than in relatively undisturbed ecosystems
Enemy release Hypothesis	ER	The absence of enemies in the exotic range is a cause of invasion success
Invasional meltdown hypothesis	IM	The presence of non-native species in an ecosystem facilitates invasion by additional species, increasing their likelihood of survival or ecological impact
Island susceptibility hypothesis	IS	Non-native species are more likely to become established and have major ecological impacts on islands than on continents
Limiting similarity hypothesis	LS	Success of non-native species is high if they strongly differ from native species, and it is low if they are similar to native species
Phenotypic plasticity Hypothesis	PH	Invasive species are more phenotypically plastic than non-invasive or native ones
Propagule pressure hypothesis	PP	High propagule pressure (a composite measure consisting of the number of individuals introduced per introduction event and the frequency of introduction events) is a cause of invasion success
Tens rule	TEN	Approximately 10% of species successfully take consecutive steps of the invasion process

Table A2. Semi-formalized representations of ten major hypotheses in invasion biology. For hypotheses stated as comparisons (Table A1; relationship "has larger value than"), a causal variant is also given. In the causal hypothesis variants, the subject describes the hypothesized driver and the object of the invasion outcome. H: Hypothesis, Q: Qualifier. For acronyms, see Table A1.

H	Subject			Relationship	Object		
	Q	Variable	Context		Q	Variable	Context
BR		Biodiversity	in an ecosystem resistant against non-native species	has larger value than		biodiversity	in an ecosystem with low resistance
	High	biodiversity	in an ecosystem	contributes to	low	invasibility	of that ecosystem
DN		Invasion success	in ecosystems poor in closely related species	has larger value than		invasion success	in ecosystems rich in closely related species
	Low	number of species closely related to a non-native species	in an ecosystem	contributes to	high	invasion success	of this species in this ecosystem
DS		Invasion success	in highly disturbed ecosystems	has larger value than		invasion success	in relatively undisturbed ecosystems
	High	disturbance	of an ecosystem	contributes to	high	invasion success	of non-native species in that ecosystem
ER	No	enemies	of a species in its non-native range	contributes to	high	invasion success	of this species in the new range
IM		Invasion success	of previously arriving non-native species	enables		invasion success or impact	of new non-native species
IS		Invasion success and impact of non-native species	on islands	has larger value than		Invasion success and impact of non-native species	

(*continued*)

Table A2. (*continued*)

H	Subject			Relationship	Object		
	Q	Variable	Context		Q	Variable	Context
		Arrival on island and not continental land		contributes to	high	invasion success and impact	
LS		Invasion success	in ecosystems poor in functionally similar species	has larger value than		invasion success	in ecosystems rich in functionally similar species
	High	functional similarity to native species	of invasive species in an ecosystem	contributes to	low	invasion success	of that species in that ecosystem
PH		Phenotypic plasticity	of invasive species	has larger value than		phenotypic plasticity	of non-invasive or native species
	High	phenotypic plasticity	of a non-native species	contributes to	high	invasion success	of this species
PP	High	propagule pressure	of a species in its non-native range	contributes to	high	invasion success	of this species in that area
TEN		n/a	n/a	n/a		n/a	n/a

References

1. Toulmin, S.E.: The Uses of Argument, 2 edn. Cambridge University Press, Cambridge (2003). https://doi.org/10.1017/CBO9780511840005
2. Heger, T., et al.: The hierarchy-of-hypotheses approach: a synthesis method for enhancing theory development in ecology and evolution. Bioscience **71**(4), 337–349 (2021). https://doi.org/10.1093/biosci/biaa130
3. Madin, J.S., Bowers, S., Schildhauer, M.P., Jones, M.B.: Advancing ecological research with ontologies. Trends Ecol. Evol. **23**(3), 159–168 (2008). https://doi.org/10.1016/j.tree.2007.11.007
4. Heger, T., Zarrieß, S., Algergawy, A., Jeschke, J.M., König-Ries, B.: INAS: interactive argumentation support for the scientific domain of invasion biology. Res. Ideas Outcomes **8**, e80457 (2022). https://doi.org/10.3897/rio.8.e80457
5. Walton, D., Reed, C., Macagno, F.: Argumentation Schemes. Cambridge University Press, Cambridge (2008). https://doi.org/10.1017/CBO9780511802034
6. Lawrence, J., Reed, C.: Argument mining: a survey. Comput. Linguist. **45**(4), 765–818 (2020). https://doi.org/10.1162/coli_a_00364
7. Šešelja, D., Straßer, C.: Abstract argumentation and explanation applied to scientific debates. Synthese **190**(12), 2195–2217 (2013). https://doi.org/10.1007/s11229-011-9964-y
8. Hogan, A., et al.: Knowledge graphs. ACM Comput. Surv. **54**(4), 1–37 (2021). https://doi.org/10.1145/3447772

9. Matiasz, N.J., et al.: ResearchMaps.org for integrating and planning research. PLoS ONE **13**(5), e0195271 (2018). https://doi.org/10.1371/journal.pone.0195271
10. Garijo, D., Gil, Y., Ratnakar, V.: The DISK hypothesis ontology: capturing hypothesis evolution for automated discovery (2017)
11. Kuhn, T.: Using the AIDA language to formally organize scientific claims. In: Wyner, A., Davis, B., Keet, C.M. (eds.) Controlled Natural Language: Proceedings of the 6th International Workshop, CNL. Frontiers in Artificial Intelligence and Applications, pp. 52–60 (2018). https://doi.org/10.3233/978-1-61499-904-1-52
12. Roy, H.E., et al.: IPBES Invasive Alien Species Assessment: Summary for Policymakers (Version 2). Zenodo (2023).https://doi.org/10.5281/zenodo.8314303
13. Enders, M., et al.: A conceptual map of invasion biology: Integrating hypotheses into a consensus network. Glob. Ecol. Biogeogr. **29**, 978–991 (2020). https://doi.org/10.1111/geb.13082
14. Catford, J.A., Jansson, R., Nilsson, C.: Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. Divers. Distrib. **15**(1), 22–40 (2009). https://doi.org/10.1111/j.1472-4642.2008.00521.x
15. Daly, E.Z., et al.: A synthesis of biological invasion hypotheses associated with the introduction–naturalisation–invasion continuum. Oikos **2023**(5), e09645 (2023). https://doi.org/10.1111/oik.09645
16. Jeschke, J.M., Heger, T. (eds.): Invasion Biology: Hypotheses and Evidence. CAB International, Wallingford, UK (2018)
17. Brinner, M., Heger, T., Zarriess, S.: Linking a hypothesis network from the domain of invasion biology to a corpus of scientific abstracts: the INAS dataset. In: Proceedings of the First Workshop on Information Extraction from Scientific Publications, pp. 32–42. Association for Computational Linguistics (2022)
18. Heger, T.: What are ecological mechanisms? Suggestions for a fine-grained description of causal mechanisms in invasion ecology. Biol. Philos. **37**(2), 9 (2022). https://doi.org/10.1007/s10539-022-09838-1
19. Jeschke, J.M., Debille, S., Lortie, C.J.: Biotic resistance and island susceptibility hypotheses. In: Jeschke, J.M., Heger, T. (eds.) Invasion Biology Hypotheses and Evidence, pp. 60–70. CAB International, Wallingford, UK (2018)
20. Hooper, D.U., Dukes, J.S.: Functional composition controls invasion success in a California serpentine grassland. J. Ecol. **98**(4), 764–777 (2010). https://doi.org/10.1111/j.1365-2745.2010.01673.x
21. Brinner, M., Zarrieß, S., Heger, T.: Weakly supervised claim localization in scientific abstracts. In: RATIO-24, Bielefeld, Germany, pp. 20–38. Springer, Heidelberg (2024)
22. Bucur, C.-I., Kuhn, T., Ceolin, D., van Ossenbruggen, J.: Expressing high-level scientific claims with formal semantics. In: Proceedings of the 11th International Conference on Knowledge Capture Conference, K-CAP 2021, New York, NY, USA, pp. 233–40. Association for Computing Machinery (2021). https://doi.org/10.1145/3460210.3493561
23. Algergawy, A., Gänßinger, M., Heger, T., Jeschke, J., König-Ries, B.: The Invasion Biology Ontology (INBIO) [Data set]. Zenodo (2022).https://doi.org/10.5281/zenodo.6826848
24. Algergawy, A., Stangneth, R., Heger, T., Jeschke, J.M., König-Ries, B.: Towards a core ontology for hierarchies of hypotheses in invasion biology. In: Harth, A., et al. (eds.) ESWC 2020. LNCS, vol. 12124, pp. 3–8. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-62327-2_1
25. Betts, M.G., et al.: When are hypotheses useful in ecology and evolution? Ecol. Evol. **11**(11), 5762–5776 (2021). https://doi.org/10.1002/ece3.7365
26. Kuhn, T., Dumontier, M.: Genuine semantic publishing. Data Sci. **1**, 139–154 (2017). https://doi.org/10.3233/DS-170010

27. Wyatt, L., et al.: WikiCite 2020–2021: citations for the sum of all human knowledge. Zenodo (2021).https://doi.org/10.5281/zenodo.5363757
28. Nielsen, F.Å., Mietchen, D., Willighagen, E.: Scholia, scientometrics and Wikidata. In: Blomqvist, E., et al. (eds.) ESWC 2017. LNCS, vol. 10577, pp. 237–259. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70407-4_36
29. Jeschke, J.M., Heger, T., Kraker, P., Schramm, M., Kittel, C., Mietchen, D.: Towards an open, zoomable atlas for invasion science and beyond. NeoBiota **68**, 5–18 (2021). https://doi.org/10.3897/neobiota.68.66685
30. Heger, T., Jeschke, J.M., Bernard-Verdier, M., Musseau, C.L., Mietchen, D.: Hypothesis description: enemy release hypothesis. Res. Ideas Outcomes **10**, e107393 (2024). https://doi.org/10.3897/rio.10.e107393
31. Mietchen, D., Mounce, R., Penev, L.: Publishing the research process. Res. Ideas Outcomes **1**, e7547 (2015). https://doi.org/10.3897/rio.1.e7547
32. Penev, L., et al.: Nanopublications for biodiversity go live. Biodivers. Inf. Sci. Stan. **7**, e110725 (2023). https://doi.org/10.3897/biss.7.110725
33. Brian, J., Catford, J.: A mechanistic framework of enemy release. Ecol. Lett. **26**(12), 2147–2166 (2023). https://doi.org/10.1111/ele.14329
34. Heger, T., Jeschke, J.M.: Enemy release hypothesis. In: Jeschke, J.M., Heger, T. (eds.) Invasion Biology Hypotheses and Evidence, pp. 92–102. CAB International, Wallingford, UK (2018) https://doi.org/10.1079/9781780647647.0092
35. Lokatis, S., et al.: Hypotheses in urban ecology: building a common knowledge base. Biol. Rev. **98**, 1530–1547 (2023). https://doi.org/10.1111/brv.12964
36. Bartram, I., Jeschke, J.M.: Do cancer stem cells exist? A pilot study combining a systematic review with the hierarchy-of-hypotheses approach. PLoS ONE **14**(12), e0225898 (2019). https://doi.org/10.1371/journal.pone.0225898
37. Guest, O., Martin, A.E.: How computational modeling can force theory building in psychological science. Perspect. Psychol. Sci. **16**(4), 789–802 (2021). https://doi.org/10.1177/1745691620970585
38. Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., et al.: Scientific discovery in the age of artificial intelligence. Nature **620**(7972), 47–60 (2023). https://doi.org/10.1038/s41586-023-06221-2
39. Jeschke, J.M., Enders, M., Bagni, M., Jeschke, P., Zimmermann, M., Heger, T.: Hi-Knowledge.org, version 2.0 (2020). Available from: https://hi-knowledge.org/