

INAS: Interactive Argumentation Support for the Scientific Domain of Invasion Biology

Tina Heger^{‡,§,|}, Sina Zarrieß[¶], Alsayed Algergawy[#], Jonathan M. Jeschke^{‡,§}, Birgitta König-Ries[#]

‡ Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany

§ Freie Universität Berlin, Institute of Biology, Berlin, Germany

| Technische Universität München, Munich, Germany

¶ University of Bielefeld, Faculty of Linguistics and Literature Studies, Bielefeld, Germany

Friedrich-Schiller-University Jena, Institute for Informatics, Jena, Germany

Corresponding author: Tina Heger (t.heger@wzw.tum.de), Sina Zarrieß (sina.zarriess@uni-bielefeld.de)

Abstract

Developing a precise argument is not an easy task. In real-world argumentation scenarios, arguments presented in texts (e.g. scientific publications) often constitute the end result of a long and tedious process. A lot of work on computational argumentation has focused on analyzing and aggregating these products of argumentation processes, i.e. argumentative texts. In this project, we adopt a complementary perspective: we aim to develop an argumentation machine that supports users during the argumentation process in a scientific context, enabling them to follow ongoing argumentation in a scientific community and to develop their own arguments. To achieve this ambitious goal, we will focus on a particular phase of the scientific argumentation process, namely the initial phase of claim or hypothesis development. According to argumentation theory, the starting point of an argument is a claim, and also data that serves as a basis for the claim. In scientific argumentation, a carefully developed and thought-through hypothesis (which we see as Toulmin's "claim" in a scientific context) is often crucial for researchers to be able to conduct a successful study and, in the end, present a new, high-quality finding or argument. Thus, an initial hypothesis needs to be specific enough that a researcher can test it based on data, but, at the same time, it should also relate to previous general claims made in the community. We investigate how argumentation machines can (i) represent concrete and more abstract knowledge on hypotheses and their underlying concepts, (ii) model the process of hypothesis refinement, including data as a basis of refinement, and (iii) interactively support a user in developing her own hypothesis based on these resources. This project will combine methods from different disciplines: natural language processing, knowledge representation and semantic web, philosophy of science and -- as an example for a scientific domain -- invasion biology. Our starting point is an existing resource in invasion biology that organizes and relates core hypotheses in the field and associates them to meta-data for more than 1000 scientific publications, which was developed over the course of several years based on manual analysis. This network, however, is currently static (i.e. needs substantial manual curation to be extended to

incorporate new claims) and, moreover, is not easily accessible for users who miss specific background and domain knowledge in invasion biology. Our goal is to develop (i) a semantic model for representing knowledge on concepts and hypotheses, such that also non-expert users can use the network; (ii) a tool that automatically computes links from publication abstracts (and data) to these hypotheses; and (iii) an interactive system that supports users in refining their initial, potentially underdeveloped hypothesis.

Keywords

argumentation in science, scientific claims, biological invasions, hypotheses, natural language processing, ontology

1 State of the art and preliminary work

Scientific claims are usually rather broad, and the empirical possibilities to test them limited. Only if broad claims are reformulated into specific hypotheses is it possible to confront them with empirical evidence (Lloyd 1987). For instance, studies in invasion biology, a sub-discipline of biodiversity research, often relate to general claims about why certain species can invade and establish in new ecosystems, but they test these claims using specific hypotheses, e.g. for specific species or forms of invasion success (Jeschke and Heger 2018). For starting a new scientific project, it is essential for a researcher to be aware of the major claims in the field and ways of refining and testing them. This step however, is usually not a formalized process, and it has been repeatedly pointed out that science could strongly profit from more precision and prudence in the important process of scientific hypothesis development (Ford 2000, McGuire 2013). Argumentation machines could facilitate scientific progress, if they would

1. provide accessible summaries of domain knowledge including basic concepts and major claims as well as their refinements,
2. link this semantic representation of the field to publications and data, thus allowing to tie newly posed claims to existing domain knowledge, and
3. use this basis to interactively support users in optimizing their specifications and refinements of broad claims.

To date, however, research on computational argumentation machines has often focused on analyzing the – typically textual – end result of the argumentation process by, e.g., classifying or mining formulations of claims and arguments in complex scientific texts (Daxenberger et al. 2017, Anonymous 2015, Lauscher et al. 2018). In this project, we take a complementary perspective and aim to develop an argumentation machine that supports users in and during the argumentation process in a scientific context, enabling them to develop a specific, testable hypothesis from an initial, potentially underdeveloped claim.

This project will combine methods from natural language processing (NLP), semantic web, philosophy of science and – as an example for a scientific domain – invasion biology. The

following sections review relevant related research and our preliminary work in these areas.

1.1 Modeling domain knowledge and arguments

In order to make domain knowledge hidden in publications and data available to argumentation machines, both the domain of interest and arguments and related concepts need to be formally modeled. Many fields have long recognized, that a common understanding of key terms is needed. This has resulted in the development of numerous domain specific vocabularies and more formally grounded ontologies. Based on a long tradition of organising knowledge in taxonomies, biodiversity research is one of these fields with numerous good ontologies (see e.g., ENVO for environmental terms <http://www.obofoundry.org/ontology/envo.html>, or the plant trait ontology <https://biportal.bioontology.org/ontologies/PTO>), but also less formalised but still useful vocabularies like different species check lists. Second, knowledge graphs (KGs) as formal models have gained attention. These KGs are typically focused on factual knowledge (see, e.g. Page (2016), Sachs et al. (2019) for examples from the biodiversity domain), but there are also recent attempts to model scientific discourse, e.g., the claims made in a publication (Auer et al. 2018). Based on semantic models for argumentation like Anonymous (2003), Toulmin (2003), a number of argumentation tools, such as AML Araucaria (Rowe and Reed 2008) or Truthmapping (2006) have been proposed. Most of these early tools use rigid languages such as XML or database structures for representing and storing arguments, making them unable to capture and reason about complex relationships among arguments. To overcome these limitations, the Argument Interchange Format (AIF), an ontology to represent and exchange data between various argumentation tools (Chesnevar et al. 2006) was introduced and is frequently used, also in the context of the RATIO-PP, e.g., in the ReCAP project (Bergmann et al. 2020). We, too, aim to build on this work and to extend to supporting linking to arguments in our domain.

Preliminary work: Biodiversity informatics and semantic web The König-Ries group has been working on leveraging semantic web techniques to support biodiversity research for quite some time. Most of this work is so far focused on improving the FAIRness of biodiversity data. It includes work on improvement of discoverability of data by better, semantic descriptions (Löffler et al. 2021, Pfaff et al. 2017). These investigations have shown which categories of concepts (e.g., organism, environment, process, event) are relevant to biodiversity research. These categories are central to the domain of invasion biology as well and at the core of arguments in this field. We have developed first tools to automatically tag terms that fall in the most important of these categories in text or data (Anonymous 2020). Beyond identification of individual terms, we have worked on different aspects of ontology development. We have created tools that allow the customisation and merging of ontologies from existing ones (Anonymous 2016), and recently started to investigate tools to support the creation of knowledge graphs (Abdelmageed 2020, Sharafeldein et al. 2020). On the other hand, we have contributed to concrete vocabularies (Schneider et al. 2019) and ontologies. In joint work with the computer linguistics group of Udo Hahn, we have investigated how to integrate structured and

unstructured data, i.e., information encoded in texts, in a semantic information system (Anonymous 2017, König-Ries and Hahn 2015).

1.2 Argumentation in science

Scientific texts have traditionally been an important domain for research on argumentation and, in particular, for data-driven approaches. Pioneering work by Teufel (1999) has introduced the idea of identifying argumentative zones in a text. Lauscher et al. (2018) have recently extended the discourse-level annotations in the Dr. Inventor Corpus by Fisas et al. (2015) with an additional annotation layer that identifies types of argument components and relations between them (supports, contradicts, same). However, resources that facilitate the study of scientific arguments on a more abstract and domain-specific level are relatively scarce. Thus, an important starting point of this project is the hierarchical network for invasion biology (see hi-knowledge.org/invasion-biology, HNI henceforth), which we will now discuss.

Preliminary Work: A hierarchical hypotheses network for invasion biology The scientific study of global change and its effects on biodiversity has many facets (Heger et al. 2019). An important domain in this respect is invasion biology – the study of human-induced spread of organisms. Due to global transport and trade, many species have been transported to areas outside of their natural range (Blackburn et al. 2011). In the HNI, Heger, Jeschke and colleagues organized more than 1000 publications with respect to their underlying hypotheses on such invasions. HNI is based on the hierarchy-of-hypotheses (HoH) approach (Heger et al. 2020, Heger et al. 2013, Jeschke et al. 2012) which we developed for invasion biology. Our basic idea is that a major, broad claim can be viewed as an overarching hypothesis on top of a hierarchical system of refined hypotheses. Single empirical tests are usually not able to test the broad, overarching hypothesis in its entirety, but instead are testing single, specific formulations, i.e. sub-hypotheses. With the HoH approach, it is possible to elucidate which exact sub-hypotheses an empirical test is addressing. In a recent book, we applied the HoH approach to organize empirical studies contesting twelve major hypotheses in invasion biology (Jeschke and Heger 2018). For every hypothesis, we created an HoH, showing which exact formulations of the major hypotheses have been assessed in published literature. We manually classified the respective studies according to whether they delivered arguments supporting or questioning the respective (sub-)hypothesis, or whether the evidence was ambiguous (classified as ‘undecided’). The HNI summarizes the results of these studies, and presents them as interactive visualization, see Fig. 1. Here, the twelve hypotheses are organized in a network structure, showing conceptual connections among them. In a cooperation with the König-Ries group, we recently took first steps to develop a core ontology for HNI (Algergawy et al. 2020).

1.3 Interactive argumentation support beyond text

In NLP, argumentation support is often construed as a ‘one-shot’ classification problem, where the system’s task is to detect low-quality arguments once in a static text e.g., Stab

and Gurevych (2017), Feltrim et al. (2006). Our approach to argumentation support is inspired by theoretical and computational research on dialogue: here, it is well established that participants in a dialogue have various, extremely efficient ways of collaborating and producing utterances in a dynamic fashion until communicative success has been reached (Clark 1996). Thus, through communicative devices like re-formulation or clarification speakers can repair misunderstandings, collaboratively solve difficult tasks and resolve uncertainties (Brennan 2005, Gergle et al. 2004). In research on dialogue systems, these processes of grounding, reformulation and iterative establishment of communicative success have mostly been modeled in rather simple task-oriented games, e.g. in visual search and manipulation tasks where uncertainty is mainly triggered by the fact that one dialogue partner does not know the location of an object or the target shape of a puzzle (Anonymous 2016). In INAS, we propose a proof-of-concept dialog system that implements these principles of human interaction in a more realistic and challenging argumentation scenario where users are (potentially) uncertain about the definitions and meanings of scientific claims and concepts.

Preliminary work: Task-oriented, multi-modal dialogue A major focus of Zarrieß' research is on task-oriented dialogue systems and interactive language generation. In Zarrieß and Schlangen (2016), we present a prototype system that implements reference to difficult-to-name objects as an interactive process using strategies for reformulating utterance in case the user is uncertain. We compare this against a non-interactive 'one-shot' system and find that the interactive system largely outperforms the non-interactive baseline. In Anonymous (2019), we take a first step towards automatically detecting and avoiding lexical uncertainty in an interactive reference task and build a system able to converse about entities whose exact name is uncertain or unknown. In this project, we tackle a similar task, namely interacting with a user who might not know the exact terms for particular scientific concepts in a domain. In Zarrieß and Schlangen (2017) we present a model for learning word meanings from visual and distributional information. In INAS, this can be generalized to further modalities, e.g. concepts represented in ontologies and text.

2 Objectives and work programme

2.1 Anticipated total duration of the project

36 months

2.2 Objectives

Developing a precise, new hypothesis for scientific argumentation is not an easy task. The goal of this project is to develop an interactive system that supports users in developing and refining hypotheses in invasion biology. Our interdisciplinary approach, combining

methods from NLP, semantic web and philosophy of science, and drawing from in-depth domain knowledge, will combine different capabilities that users need during this process:

1. domain-specific background knowledge on abstract and concrete concepts related to claims in invasion biology,
2. detailed feedback on formulations of scientific hypotheses on different levels of specificity and
3. links to datasets for testing hypotheses.

Fig. 2 illustrates how this project builds on the exceptional HNI resource (see Section 1.2) to implement a computational framework that models the semantics of concepts in domainspecific argumentation (Component A), and the refinement of hypotheses based on finegrained hypothesis representations and data (Component B). These two components will be combined in an interactive hypothesis development system (Component C). We focus on concept and hypothesis refinement (A and B) and operationalize hypothesis development as an iterative process that is well suited to be implemented in an interactive system (C) that guides a user to develop her own, new hypothesis.

We expect that our approach will be a very useful extension of HNI and contribute to the field of invasion biology, but also give general insights on how to represent knowledge for argumentation systems and leverage this knowledge for interaction with users in real-word argumentation processes. With such an approach, argumentation machines would support novice researchers in understanding the field, but would also be able to help mapping a field, detecting contradictions and gaps, and detecting links to neighboring fields, where syntactically different terms might be used to describe similar claims.

Challenges Automatic support for hypothesis development is a very challenging task for state-of-the-art argumentation machines. For research in invasion biology, the HNI in its current form is a valuable resource only for domain experts. Early career researchers and scientists new to the domain will lack background knowledge on terms, concepts (and their ambiguities) to make efficient use of the network and, e.g., find relevant abstracts. Second, scientific practice in invasion biology, and also in ecology in general, does usually not put special emphasis on precise and explicit formulation of claims or hypotheses. For example, it is usually clarified whether a claim rather amounts to the expectation of a pattern, or the suggestion of a causal relationship, or whether the claims implicitly contain unexpressed propositions. From an NLP perspective, an important challenge then is to communicate this background knowledge in an appropriate way and process potentially underdeveloped or imprecise formulations of hypotheses. Additionally, hypotheses constitute very abstract statements that, in a scientific publication, can be instantiated and formulated in very different ways. For example, two abstracts may be linked to the same hypothesis without explicitly mentioning it. For users not aware of certain assumptions and concepts in the field, this will be hard to determine.

These phenomena also create challenges for semantic web systems: Beyond the need for integration across domains, an approach is needed in INAS to support smooth, continuous evolution of the semantic backbone as modeling and understanding of the domain

deepens and evolves. A second challenge in INAS will be the seamless integration of data as basis for arguments. This requires first of all to semantically describe data. Due to the large volume of available data, this task clearly needs to be automated. This requirement has recently sparked the SemTab challenge (<http://www.cs.ox.ac.uk/isg/challenges/sem-tab/>). Second, an abstraction layer needs to be added to the data turning it into an argument. This requires summarization and interpretation of data.

2.3 Work programme including proposed research methods

To address the challenges discussed above, this project brings together experts from the fields of NLP, biology and semantic web. This broad expertise will be supplemented by collaborations with philosophers of science. We believe that this is an ideal set-up to advance the state-of-the-art in argument modeling and move towards systems that meet the complex information needs of users and are flexible enough to be automatically extended to new hypotheses, new publications, new datasets and, ultimately, also new domains and other research areas.

2.3.1 Methods

Knowledge representation Our framework will model and represent the internal semantic structure of claims in terms of abstract domain-specific concepts and their various possible refinements in testable hypotheses, as sketched in Fig. 3 and Fig. 5: e.g. establishment success, which is an element of invasion success, can be measured as breeding success in the ecosystem where an alien species was introduced, which in turn can be measured as offspring mortality. Importantly, this requires the coupling of a domain-specific core ontology with an argumentation-based ontology.

Argumentation and data Our work will integrate multiple ways and dimensions of modeling hypotheses, i.e., in text but also in knowledge representations and through datasets. As illustrated in Fig. 5, INAS will develop a hypothesis refinement tool that aggregates datasets and hypotheses where hypotheses are structured as causal networks that give detailed information on how parts of a general claim have been attested in data.

Dialogue modeling We propose to model hypothesis development in a dialogue system that uses the HNI ontology to compute hierarchical information states (e.g. the general claim, concepts represented in the claim, sub-parts of the given claim) which need to be filled throughout the interaction between user and system. Thus, the system will not need to process or validate an entire argument at once, but rather focus on specifying different parts of the claim in a step-by-step, collaborative fashion, as illustrated in Fig. 3. The components that process user utterances and link them to hypotheses or concepts in the ontology will be implemented as neural language processing components. These can be trained on large biomedical corpora (e.g. to obtain word and sentence embeddings), but also on the paper abstracts currently represented in HNI.

Evaluation To date, there are few systematic insights into how argumentation systems should be set up to really enhance the way users can understand and develop arguments.

An important goal of the project is to develop an evaluation scenario and a user study design that fills this gap and, ideally, can be generalized to other domains or other argumentation support scenarios. We plan to collaborate with other RATIO projects on this topic, e.g. with Philipp Cimiano's and Ulf Leser's planned project on argumentation support in a clinical domain.

2.3.2 Work packages

An outline of the work packages with effort in person months is given in Fig. 4.

Milestones The project will be structured by 3 milestones (see Fig. 4).

1. M1: the basic framework for semantic modeling of hypotheses is set up
2. M2: a proof-of-concept system for interactive hypothesis development is set up
3. M3: the framework is integrated, validated and tested in user studies

WP 1: A semantic model for argumentation in invasion biology

A prerequisite to leveraging the power of Semantic Web techniques are shared ontologies to facilitate the seamless exchange of information. In this WP, we will bring together domain experts, philosophers of science, knowledge engineers, and end users to create such ontologies for our domain of interest (WP 1.2) and the argumentation domain and linking the two (WP 1.3). We will support this with text mining to identify key concepts, their definitions and relations (WP 1.4). Creating ontologies is not a one-time task, but rather an iterative community process which requires support for an evolving and deepening understanding of the domain (WP 1.1).

WP 1.1: Process model It is a characteristic of science that the understanding of a field becomes more nuanced over time. For us, this implies, that the domain model will also evolve over time. At the very beginning of the project, taking into account existing work on ontology evolution (Zablith 2007) and interactive ontology development (Jackson et al. 2019), a process model for this project needs to be agreed upon and appropriate tool support needs to be set up.

WP 1.2: Core ontology The core ontology for invasion biology, called HoH ontology, will be used to model the complex structure of knowledge in the hierarchy of hypotheses in the domain of invasion biology. We will adopt the fusion/merge strategy (Pinto and Martins 2004), where the new ontology is developed by assembling and reusing one or more ontologies. We will first identify relevant terms and keywords by exploiting the knowledge sources mentioned above (collection of hypotheses, publications and datasets). We will then identify, subset and recombine suitable ontologies supported by our JOYCE tool (Faessler et al. 2017) and the deep domain knowledge of one of the PIs. Finally, the ontology will be populated semi-automatically using results from WP1.4. As described in WP 1.1, this is not a one-time activity but an iterative process.

WP 1.3: Argumentation ontology Our argumentation ontology will be based on the AIF (Argument Interchange Format, Chesnevar et al. 2006, Rahwan and Reed 2009,

Zagorulko et al. 2019), a standard notation for representing the definitions of high-level concepts related to argumentation. These concepts are categorized into three main groups: concepts related to argument entities and relation among them, concepts relate to the interchange of arguments between two or more participants in an environment, and concepts related to environments in which argumentation may take place. In this WP we will extend this ontology, if such need is identified in the other WPs. A special focus of this task, however, is the population of this ontology with instances related to invasion biology and thus the linking of the two parts of our domain model. Again, this is not a one-time activity but part of an iterative process; in particular results of the workshop conducted as part of WP 6 will very likely result in adaptations of the ontology.

WP 1.4 : Term mining The goal of this WP is to semi-automatically obtain lists of names or terms referring to instances of species and locations, and potentially other entity types identified in WP 1.2 from the INAS abstracts. These will contribute to populating the invasion biology core ontology (WP 1.2) and to fine-tune tools for NER and argument linking in WP 3. Based on resources like LINNAEUS (Gerner et al. 2010), Species-800 (Pafilis et al. 2013) and the generic CoNLL-2003 dataset for NER (Sang and De Meulder 2003), we will explore a combination of different off-the-shelf NER tools to obtain a good coverage of entity types, namely

1. BioBERT, a neural transformer-based network that learns word embeddings on large amounts of text from the biomedical domain and fine-tunes them for different tasks, including NER on LINNAEUS and Species-800 and
2. the LSTM-CRF by Lample et al. (2016).

A subset of the automatic annotations obtained from BioBERT and LSTM-CRF will be corrected manually during the ontology development. These can, in turn, be used to fine-tune Bio-BERT to predict species and locations on the INAS abstracts.

WP 2: Hypothesis refinement

While the ontology development in WP1 focuses on the identification and refinement of concepts used in hypotheses in invasion biology, this work package investigates the refinement of the hypotheses themselves. We design a more detailed, nested representation of the hypotheses in the HNI (WP 2.1) and link this to datasets (WP 2.2). Fig. 5 sketches an example representation that ideally results from this framework, i.e., showing how a hypothesis is decomposed into testable parts and which of these parts have already been tested on data for a given species, location, etc.

WP 2.1: Hypotheses as nested causal networks In the invasion biology domain, hypotheses often are formulated as if they would address simple causal relationships (e.g. "The absence of enemies in the exotic range is a cause of invasion success"). For domain experts, however, these simplifications are hints to basic knowledge about underlying mechanisms, i.e. longer chains or networks of hypothesized causal relationships. In this work package, we will re-formulate the hypotheses contained in the hierarchical hypothesis network as complex, nested causal relationships. For each element in the causal chains,

key references from the domain literature will be searched. The nested representations of hypotheses will be used to annotate a subset of 50-100 publication abstracts in our collection. These annotations can be used as a fine-grained test set for the NLP system in WP 3.1 and will be made available as a corpus to the NLP community (see data management plan). To fulfill this task, we will closely cooperate with philosophers of science.

WP 2.2: Data-hypothesis linking In biology, data is an important dimension of argumentation, as it is needed to test hypotheses and to support or refute claims. Detailed information on available datasets is also very important during the hypothesis development process, e.g., for exploring whether and how a certain claim has been tested in prior work (see Fig. 5). In order to provide users with support for leveraging data for argumentation, we will build on ongoing work in the König-Ries group and elsewhere. We will use and adapt two sets of tools currently under development: The first set provides (semi-)automatic semantic annotation of datasets. We will evaluate available solutions to the SemTab challenge (including our own; under development at the time of writing) and pick and adapt the most suitable one for this task. The second set of tools is currently being developed as part of CRC AquaDiva in the König-Ries group and will offer automatic summaries of data. In the unlikely case that the tools are not available in time, we will manually summarize a limited number of datasets related to the annotated publication abstracts from WP 2.1 for use in the framework. This will result in semantic annotations (suitable for finding datasets) and visual summaries of datasets supporting a quick understanding of their key message. These results will be integrated in our work in two places: First, to provide quick access to data used to test argumentation chains (as depicted in Fig. 5), and second, to support exploration of potentially relevant datasets during hypothesis development as part of the dialog shown in Fig. 3.

WP 3: Interactive Support for Hypothesis Development

In this WP, we will build an interactive system that uses the resources for concept and hypothesis refinement in WP 1 and WP 2 to support users in developing a hypothesis in the field of invasion biology. The main novelty and central challenge here is that hypothesis development is a very abstract task where communicative success is difficult to measure. We will build a neural, non-interactive classification model for text-hypothesis understanding (i.e. linking) (WP 3.1) and integrate this with a dialogue system with a predefined action-state space (WP 3.2.), which will be fine-tuned after an initial user study (WP 3.3).

WP 3.1: Text-hypothesis linking An important task of the dialogue system is to determine which general claim or hypothesis the user is talking about. We operationalize this as a classification problem, where the task is to predict whether a sentence entered by a user refers to a hypothesis represented in HNI. We will set up a neural architecture with two encoders, e.g. RNNs that learn hidden representations of the HNI hypothesis and the textual hypothesis. The central research question here is whether we can successfully leverage the symbolic knowledge encoded in the ontology (WP 1) in the neural encoders for the text and hypotheses, e.g. through compositional neural models (Andreas et al. 2016

). Thus, in a currently running Master project in Zarrieß group, we carried out a preliminary pilot study on this task and tested simple bag-of-words classifiers to link abstracts and hypotheses in the HNI. We obtained a rather low accuracy of 40% with this model, which indicates the need to integrate more abstract domain knowledge. The training and testing data for the network is taken from the 1100 paper abstracts in the current HNI resource. We split these abstracts into paragraphs or sentences, and pair them with the hypotheses that they are linked to in the HNI. By splitting the abstract into smaller parts, we expect to simulate the underdeveloped hypotheses that the user will enter when interacting with the dialogue system. This classification architecture will be trained and tested on different levels of granularity of HNI, for linking texts and hypotheses on the level of major hypotheses and sub-hypotheses, and for different parts. The outcome is a text-hypothesis matching system that will be tested automatically on a test set taken from the current papers in HNI and that can be integrated into the dialogue component in WP 3.2. Another use case of this model would be an automatic extension of the HNI resource with new papers, or new hypotheses developed by users.

WP 3.2: Dialogue model We set up a dialogue component for hypothesis development that splits up this process into a sequence of smaller steps, like e.g. discuss

1. the general claim,
2. the species,
3. how to refine concepts in the general claim, etc., extending Zarrieß' previous work on establishing references in installments (Zarrieß and Schlangen 2016).

Once the system and the user have agreed on a general claim, the subsequent states will depend on the hypothesis components represented in the ontology (WP 1), see Fig. 3. We also design templates and actions for the language generation component which includes verbal feedback, but also actions like pointing the user to nested representations of the general claim (see WP 2.1), to datasets (see WP 2.2) or to more specific definitions of concepts and terms in the core ontology (WP 1). Fig. 3 illustrates a simple potential interaction with such a support system. The understanding component (NLU) of the dialogue system will be based on the two components described in WP 3.1. and also term mining system and embedding models in WP 1.4.

WP 3.3: Hypothesis reformulation As a first evaluation of the dialogue system (WP 3.2.), we carry out a pilot human evaluation with students from the biology programme in Berlin or Jena. This study will give us very valuable data on how users reformulate their hypotheses based on feedback of our system and we will use it to conduct a careful analysis of interaction quality in general and the process of hypothesis development in particular. In case we find that the interactions between our system and users are already of good quality and enable users to develop their own hypotheses, we can use the data to fine-tune/learn aspects of the dialogue system's action space in WP 3.2 (e.g. when to give certain types of verbal or non-verbal feedback). In the other case, the data will be extremely useful to further develop our system and gain a deeper qualitative understanding of how the system can support the very challenging task of hypothesis development.

WP 4: Resources and Integration

The current version of HNI is available on the public website hi-knowledge.org. We extend this interface and integrate it with the models and resources developed in WP 1.3. The extended interface will be used to run user studies and evaluations, as described in WP 5.

WP 4.1: Ontology and datasets We will integrate the ontology from WP 1 such that users can inspect the meanings of terms used in a hypothesis description or a paper abstract. We will set up a database that records the available meta data for the papers represented in HNI, including the paper abstracts which will be indexed to support basic keyword search and links to available data sets as well as their semantic enrichment where applicable.

WP 4.2: Chat interface We will extend HNI with a simple chat interface to integrate the dialogue system from WP 3, using our web-based SLURK tool (Anonymous 2018) designed for easy implementation of web-based, multi-modal chat-bots.

WP 5: Evaluation

One of the central goals of INAS is to build a framework for argument modeling that is closely tied to the needs of human users. We will thoroughly validate and consolidate the ontologies developed in WP1 with experts and conduct user studies, assessing to what extent our systems helps users in hypothesis development.

WP 5.1: Ontology consolidation and validation During a three-day workshop, the core ontology, the argumentation ontology as well as the nested representation of hypotheses will be validated. We will invite domain experts from the invasion biology community and philosophers of science. We will use a combination of pre-workshop tasks, panel presentations, break-out discussions and panel discussions to reach a broad consensus on the main features of the ontologies and the nested hypotheses. The workshop results will be used to consolidate our models.

WP 5.2: User study We will design and conduct a user study to assess the quality of argumentation support system. This includes the definition of a concrete hypothesis development task that users will have to carry out when interacting with our system (e.g. based on a given paper in invasion biology, define a promising hypothesis for follow-up studies), the identification of a target user group and the definition of criteria for assessing hypotheses that users develop with the help of our system. As users might interact very differently with our system depending on their background, we will need to identify two relatively consistent user groups (e.g. undergraduate or graduate students in biology that have taken classes on ecology) to obtain meaningful results. We will conduct a pilot user study with approx. 30 participants towards the end of the second year of the project (Fig. 4), to obtain valuable data for fine-tuning the dialogue system (WP 3.3) and test argumentation support in this novel setting. In the final user study, we identify two versions of our system that will be compared, e.g. a version with and without interactive dialogue support. We use a mixed within-subjects two-by-two design, where subjects from two

different groups interact with both systems, approx. 40 participants (20 from each group) which we plan to recruit at FU Berlin.

WP 6: Dissemination

WP 6.1: Conferences and publications The PhD student and PI Tina Heger will present project results at international conferences and workshops. The events will cover the fields of NLP, semantic web, philosophy of science, invasion biology and ecology. The project team will publish at least 4 publications in international journals and high-ranked conferences from the fields of NLP, semantic web, philosophy of science and invasion biology as outlets. We view research data management and in particular the sustainable provision and publication of FAIR data as another important dissemination activity that will be tackled in this WP.

WP 6.2: Workshop "Modelling the argumentation process across domains" A further element of this work package will be a workshop bringing together research groups working on similar tools in different domains. Aims of the workshop will be:

- to present our results in order to allow for exchange and synergies with related projects, and
- to compare argumentation processes and ways to model them across domains.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for funding this project (Project number 455913229). The publication of this article was funded by the Open Access Fund of the Leibniz Association.

Funding program

Schwerpunktprogramm "Robust Argumentation Machines (RATIO)"

Grant title

INAS: Interactive argumentation support for the scientific domain of invasion biology; project number 455913229

Hosting institutions

1. University of Bielefeld
2. Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB)
3. University of Jena

Author contributions

Tina Heger and Sina Zarriß contributed equally to this proposal.

Conflicts of interest

References

- Abdelmageed N (2020) Towards transforming tabular datasets into knowledge graphs. The Semantic Web: ESWC 2020 Satellite Events pp. 217-288. https://doi.org/10.1007/978-3-030-62327-2_37
- Algergawy A, Babalou S, Klan F, Ries BK (2016) OAPT: A Tool for Ontology Analysis and Partitioning. Demo Paper. EDBT. 644–647 pp.
- Algergawy A, Stangneth R, Heger T, Jeschke J, König-Ries B (2020) Towards a Core Ontology for Hierarchies of Hypotheses in Invasion Biology. The Semantic Web: ESWC 2020 Satellite Events 3-8. https://doi.org/10.1007/978-3-030-62327-2_1
- Andreas J, Rohrbach M, Darrell T, Klein D (2016) Neural Module Networks. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) <https://doi.org/10.1109/cvpr.2016.12>
- Auer S, Kovtun V, Prinz M, Kasprzik A, Stocker M, Vidal ME (2018) Towards a Knowledge Graph for Science. Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics <https://doi.org/10.1145/3227609.3227689>
- Bergmann R, Biertz M, Dumani L, Lenz M, Ludwig A, Neumann P, Ollinger S, Sahitaj P, Schenkel R, Witry A (2020) The ReCAP Project. Datenbank-Spektrum 20 (2): 93-98. <https://doi.org/10.1007/s13222-020-00340-0>
- Blackburn T, Pyšek P, Bacher S, Carlton J, Duncan R, Jarošík V, Wilson JU, Richardson D (2011) A proposed unified framework for biological invasions. Trends in Ecology & Evolution 26 (7): 333-339. <https://doi.org/10.1016/j.tree.2011.03.023>
- Brennan SE (2005) How conversation is shaped by visual and spoken evidence. In: Trueswell J, Tanenhaus M (Eds) Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions. Cambridge, MA: MIT Press., pp. 95-129 pp.
- Chesnevar C, Modgil S, Rahwan I, Reed C, Simari G, South M, Vreeswijk G, Willmott S, et al. (2006) Towards an argument interchange format. The Knowledge Engineering Review 21 (4): 293-316. <https://doi.org/10.1017/s0269888906001044>
- Clark HH (1996) Using Language. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511620539>
- Daxenberger J, Eger S, Habernal I, Stab C, Gurevych I (2017) What is the Essence of a Claim? Cross-Domain Claim Identification. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing <https://doi.org/10.18653/v1/d17-1218>

- Faessler E, Klan F, Algergawy A, König-Ries B, Hahn U (2017) Selecting and Tailoring Ontologies with JOYCE. Lecture Notes in Computer Science 114-118. https://doi.org/10.1007/978-3-319-58694-6_12
- Feltrim V, Teufel S, das Nunes M, Aluísio S (2006) Argumentative Zoning Applied to Critiquing Novices' Scientific Abstracts. The Information Retrieval Series 233-246. https://doi.org/10.1007/1-4020-4102-0_18
- Fisas B, Saggion H, Ronzano F (2015) On the discursive structure of computer graphics research papers. Proceedings of The 9th Linguistic Annotation Workshop <https://doi.org/10.3115/v1/w15-1605>
- Ford ED (2000) Scientific method for ecological research. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511612558>
- Gergle D, Kraut RE, Fussell SR (2004) Language efficiency and visual technology: Minimizing collaborative effort with visual information. Journal of language and social psychology 23 (4): 491-517. <https://doi.org/10.1177/0261927X04269589>
- Gerner M, Nenadic G, Bergman CM (2010) LINNAEUS: A species name identification system for biomedical literature. BMC Bioinformatics 11 (1). <https://doi.org/10.1186/1471-2105-11-85>
- Heger T, Pahl A, Botta-Dukát Z, Gherardi F, Hoppe C, Hoste I, Jax K, Lindström L, Boets P, Haider S, Kollmann J, Wittmann M, Jeschke JM (2013) Conceptual Frameworks and Methods for Advancing Invasion Ecology. AMBIO 42 (5): 527-540. <https://doi.org/10.1007/s13280-012-0379-x>
- Heger T, Bernard-Verdier M, Gessler A, Greenwood AD, Grossart H, Hilker M, Keinath S, Kowarik I, Kueffer C, Marquard E, Müller J, Niemeier S, Onandia G, Petermann JS, Rillig MC, Rödel M, Saul W, Schittko C, Tockner K, Joshi J, Jeschke JM (2019) Towards an Integrative, Eco-Evolutionary Understanding of Ecological Novelty: Studying and Communicating Interlinked Effects of Global Change. BioScience 69 (11): 888-899. <https://doi.org/10.1093/biosci/biz095>
- Heger T, Aguilar-Trigueros CA, Bartram I, Braga RR, Dietl GP, Enders M, Gibson DJ, Gómez-Aparicio L, Gras P, Jax K, Lokatis S, Lortie CJ, Mupepele A, Schindler S, Starrfelt J, Synodinos AD, Jeschke JM (2020) The Hierarchy-of-Hypotheses approach: A synthesis method for enhancing theory development in ecology and evolution. BioScience 71 (4): 337-349. <https://doi.org/10.1093/biosci/biaa130>
- Jackson R, Balhoff J, Douglass E, Harris N, Mungall C, Overton J (2019) ROBOT: A Tool for Automating Ontology Workflows. BMC Bioinformatics 20 (1). <https://doi.org/10.1186/s12859-019-3002-3>
- Jeschke JM, Gómez Aparicio L, Haider S, Heger T, Lortie C, Pyšek P, Strayer D (2012) Support for major hypotheses in invasion biology is uneven and declining. NeoBiota 14: 1-20. <https://doi.org/10.3897/neobiota.14.3435>
- Jeschke JM, Heger T (2018) Invasion biology: hypotheses and evidence. CABI <https://doi.org/10.1079/9781780647647.0000>
- Klan F, Faessler E, Algergawy A, König-Ries B, Hahn U (2017) Integrated Semantic Search on Structured and Unstructured Data in the ADONIS System. S4BioDiv@ ISWC.
- König-Ries B, Hahn U (2015) Semantic technologies for consolidating structured data and unstructured documents in biodiversity research. Geoinformationssysteme 2015. Beiträge zur 2. Münchner GI-Runde. Wichmann
- Lample G, Ballesteros M, Subramanian S, Kawakami K, Dyer C (2016) Neural Architectures for Named Entity Recognition. Proceedings of the 2016 Conference of the

- North American Chapter of the Association for Computational Linguistics: Human Language Technologies 260-270. <https://doi.org/10.18653/v1/n16-1030>
- Lauscher A, Glavaš G, Ponzetto SP (2018) An Argument-Annotated Corpus of Scientific Publications. Proceedings of the 5th Workshop on Argument Mining <https://doi.org/10.18653/v1/w18-5206>
 - Lippi M, Torroni P (2015) Context-independent claim detection for argument mining. Proc. of IJCAI.
 - Lloyd E (1987) Confirmation of ecological and evolutionary models. *Biology & Philosophy* 2 (3): 277-293. <https://doi.org/10.1007/bf00128834>
 - Löffler F, Abdelmageed N, Babalou S, Kaur P, König-Ries B (2020) Tag Me If You Can! Semantic Annotation of Biodiversity Metadata with the QEMP Corpus and the BiodivTagger. Proc of LREC. Marseille, France: European Language Resources Association, May 2020, pp. 4557–4564.
 - Löffler F, Wesp V, König-Ries B, Klan F (2021) Dataset search in biodiversity research: Do metadata in data repositories reflect scholarly information needs? *PLOS ONE* 16 (3). <https://doi.org/10.1371/journal.pone.0246099>
 - McGuire W (2013) An Additional Future for Psychological Science. *Perspectives on Psychological Science* 8 (4): 414-423. <https://doi.org/10.1177/1745691613491270>
 - Pafilis E, Frankild S, Fanini L, Faulwetter S, Pavlodi C, Vasileiadou A, Arvanitidis C, Jensen LJ (2013) The SPECIES and ORGANISMS Resources for Fast and Accurate Identification of Taxonomic Names in Text. *PLoS ONE* 8 (6). <https://doi.org/10.1371/journal.pone.0065390>
 - Page R (2016) Towards a biodiversity knowledge graph. *Research Ideas and Outcomes* 2 <https://doi.org/10.3897/rio.2.e8767>
 - Pfaff C, Eichenberg D, Liebergesell M, König-Ries B, Wirth C (2017) Essential Annotation Schema for Ecology (EASE)—A framework supporting the efficient data annotation and faceted navigation in ecology. *PLOS ONE* 12 (10). <https://doi.org/10.1371/journal.pone.0186170>
 - Pinto HS, Martins J (2004) Ontologies: How can They be Built? *Knowledge and Information Systems* 6 (4): 441-464. <https://doi.org/10.1007/s10115-003-0138-1>
 - Rahwan I, Reed C (2009) The Argument Interchange Format. *Argumentation in Artificial Intelligence* 383-402. https://doi.org/10.1007/978-0-387-98197-0_19
 - Reed C, Walton D (2003) Argumentation schemes in argument-as-process and argument as- product. Proc. of the Conference Celebrating Informal Logic.
 - Rowe G, Reed C (2008) Argument Diagramming: The Araucaria Project. *Advanced Information and Knowledge Processing* 164-181. https://doi.org/10.1007/978-1-84800-149-7_8
 - Sachs J, Page R, Baskauf SJ, Pender J, Lujan-Toro B, Macklin J, Comspon Z (2019) Training and hackathon on building biodiversity knowledge graphs. *Research Ideas and Outcome* 5: e36152. <https://doi.org/10.3897/rio.5.e36152>
 - Sang E, De Meulder F (2003) Introduction to the CoNLL-2003 shared task: Language independent named entity recognition. arXiv preprint URL: <https://arxiv.org/abs/cs/0306050>
 - Schlangen D, Diekmann T, Ilinykh N, Zarriß S (2018) slurk – A Lightweight Interaction Server For Dialogue Experiments and Data Collection. Proc. of AixDial / SEMdial. Aix-en-Provence, France.

- Schneider F, Fichtmueller D, Gossner M, Güntsch A, Jochum M, König-Ries B, Le Provost G, Manning P, Ostrowski A, Penone C, Simons N (2019) Towards an ecological trait-data standard. *Methods in Ecology and Evolution* 10 (12): 2006-2019. <https://doi.org/10.1111/2041-210x.13288>
- Sharafelddeen D, Algergawy A, König-Ries B (2020) Towards Knowledge Graph Construction using Semantic Data Mining. *Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services* <https://doi.org/10.1145/3366030.3366035>
- Stab C, Gurevych I (2017) Recognizing Insufficiently Supported Arguments in Argumentative Essays. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers* <https://doi.org/10.18653/v1/e17-1092>
- Teufel S, et al. (1999) Argumentative zoning: Information extraction from scientific text. PhD thesis. University of Edinburgh
- Toulmin SE (2003) *The uses of argument*. Cambridge University Press <https://doi.org/10.1017/CBO9780511840005>
- Truthmapping (2006) www.truthmapping.com/
- Zablith F (2007) Argdf: Arguments on the semantic web'. The British University in Dubai Jointly with The University of Edinburgh
- Zagorulko Y, Garanina N, Sery A, Domanov O (2019) Ontology-Based Approach to Organizing the Support for the Analysis of Argumentation in Popular Science Discourse. *Communications in Computer and Information Science* 348-362. https://doi.org/10.1007/978-3-030-30763-9_29
- Zarrieß S, Schlangen D (2016) Easy Things First: Installments Improve Referring Expression Generation for Objects in Photographs. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* <https://doi.org/10.18653/v1/p16-1058>
- Zarrieß S, Hough J, Kennington C, Manuvinakurike R, DeVault D, Fernandez R, Schlangen D (2016) PentoRef: A corpus of spoken references in task-oriented dialogues. *Proc. of LREC*.
- Zarrieß S, Schlangen D (2017) Obtaining referential word meanings from visual and distributional information: Experiments on object naming. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* <https://doi.org/10.18653/v1/p17-1023>
- Zarrieß S, Schlangen D (2019) Know What You Don't Know: Modeling a Pragmatic Speaker that Refers to Objects of Unknown Categories. *Proc. of the 57th Annual Meeting of the ACL*. Florence, Italy: Association for Computational Linguistics, July 2019, pp. 654–659.

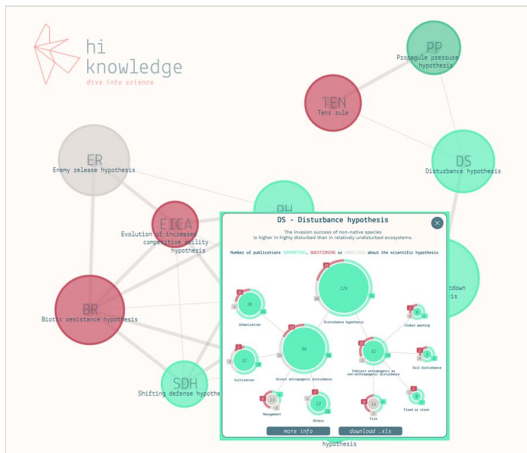


Figure 1.

Screenshot of the website hiknowledge.org, showing a network of twelve major hypotheses on potential causes of biological invasions. The insert shows the hierarchy of hypotheses (HoH) for the disturbance hypothesis which can be retrieved by clicking on the respective dot in the network, with information on the numbers of studies supporting (green), questioning (red) or being undecided (grey) about the respective (sub-)hypotheses.

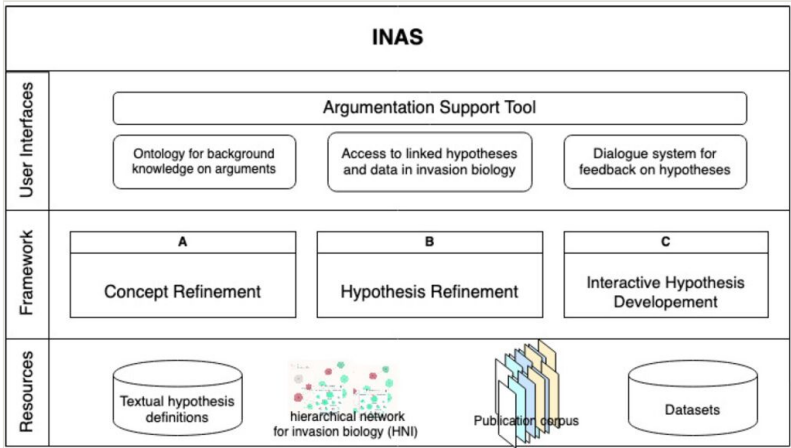


Figure 2.
The main architecture of INAS.

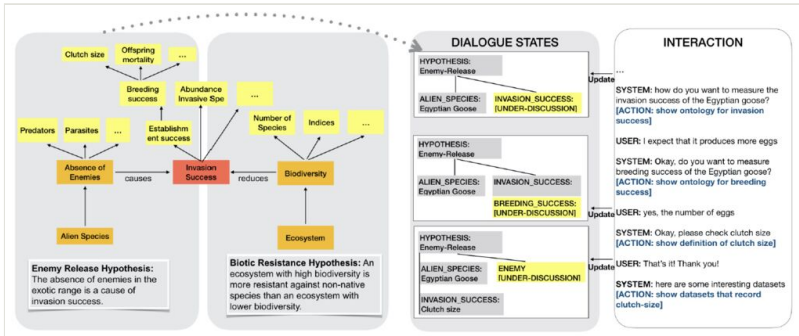


Figure 3.

Interactive hypothesis development, based on a semantic model of hypotheses in the invasion biology domain (left) and a made-up example of a short interaction with an information-state-based dialogue system that iteratively refines a hypothesis introducing domain-specific terms in collaboration with the user (right, resolved questions appear in grey, questions under discussion in yellow).

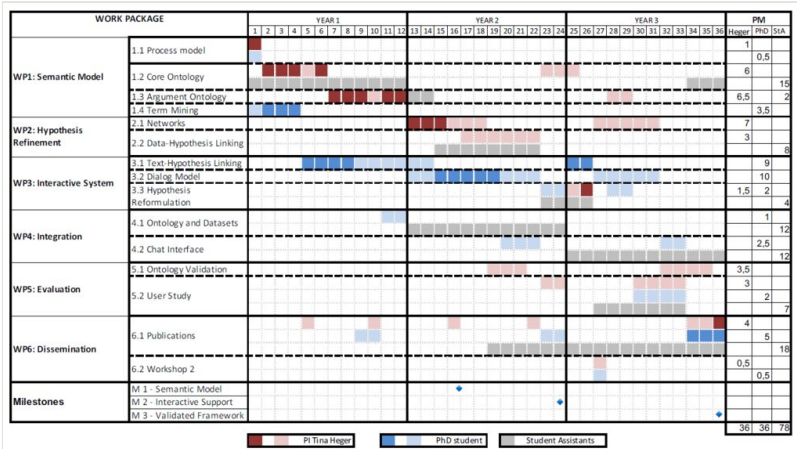


Figure 4.

Work plan with full-time tasks (dark colour) and half-time tasks (light color) for the PI Heger (red), the PhD (blue) and student assistants (gray).

Enemy Release Hypothesis: The absence of enemies in the exotic range is a cause of invasion success.

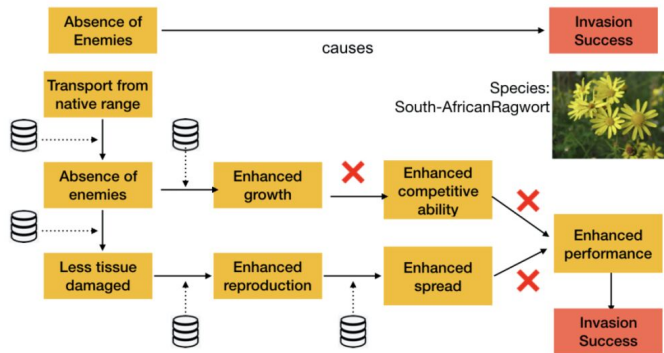


Figure 5.

Refining hypotheses as nested chains; data symbols indicate that this part of the chain has been tested with data for the South-African Ragwort; red crosses symbolize that this part of the chain has not been tested yet for this specific species.