Modeling Narrative Structures in Logical Overlays on top of Knowledge Repositories

Hermann Kroll[0000−0001−9887−9276], Denis Nagel[0000−0002−5832−9154], and Wolf-Tilo Balke[0000−0002−5443−1215]

Institute for Information Systems, TU Braunschweig, Braunschweig, Germany
{kroll,nagel,balke}@ifis.cs.tu-bs.de

Abstract. An important part of the scientific discourse is the exchange of knowledge in the form of stringent, well-arranged, and interconnected arguments. These 'scientific storylines' allow to put central entities, observations, experiments, etc. into perspective and thus ease the understanding of underlying mechanisms, dependencies, or theories. Moreover, taking a bird's eye view allows to discern recurring *narrative patterns* that have proven helpful for validating, comparing, and fusing information across individual publications and even between disciplines. However, current knowledge repositories still struggle with representing such information in a structured way. This is because narratives do not only contain factual bits of information, but also parts like temporal developments, causal dependencies, etc. In this paper, we present an innovative conceptual model using a *logical overlay structure* to bridge the gaps between individual types of knowledge repositories. We also explain how *narrative bindings* validate modeled narratives in the sense of provenance. In brief, narrative overlays plus adequate bindings allow to effectively fuse knowledge and improve retrieval and discovery tasks by structurally aligning underlying repositories only driven by some narrative. Finally, we practically demonstrate the usefulness of our model by applying it to a scientific narrative in the PubMed bio-medical collection.

Keywords: Narratives · Logical overlays · Knowledge graphs.

1 Introduction

A lot of today's world – theories, insights, and decisions – has become 'data-driven'. Making sense of vast amounts of data is usually realized using structured knowledge repositories, e. g. relational databases, knowledge graphs, digital libraries, or data set registries [1, 12]. Yet, the theories, insights, etc. are usually not *part of* such repositories, but have to be managed outside. In this paper, we propose a conceptual model that integrates derived knowledge in the general form of narratives on top of knowledge repositories. The basic idea can be compared to peer-to-peer networks: built on top of a *physical* IP-based routing infrastructure, direct connections in a *logical overlay* allow for creating advantageous network topologies that can subsequently be used for improved routing, content sharing, etc. In the same way, we argue for a logical overlay as an abstraction layer on top of knowledge repositories that, in turn, allows us to capture narratives, bind them to knowledge repositories, and assess essential characteristics such as their individual validity or plausibility.

Unlike fictional narratives that tend to involve many protagonists and hence are notoriously hard to represent [2], narratives used in practical information systems are usually more limited and quite concise. Such narratives usually relate recurring explanation patterns or chains of arguments that are investigated, modeled, and schematically represented for subsequent sharing, discussion, and reuse by researchers in a variety of scientific domains [10]. Prime examples include chemical reactions and metabolic pathways in bio-medicine.

Throughout this paper, we will use a pharmaceutical use case with an often occurring narrative pattern of a simple drug-drug interaction as a running example (Fig. 1). In brief, assume that an *active ingredient* is *metabolized* in the body by some *gene system*, but exactly this system is *inhibited* by some other *drug* administered at the same time. Then the active ingredient is *accumulated* in the body, which in turn may *cause* severe adverse effects in the form of *diseases*.

Looking at this simple pharmaceutical example narrative describing a typical kind of adverse drug-drug interaction, we already get a first idea of the concepts, which we will conceptualize in the following sections. There are *entities* like active ingredients, gene systems, or drugs, there are *relationships* between them, such as being metabolized by something or inhibiting something, there are *events*, such as the accumulation of some active ingredient in the body, and there are *causal or temporal structures* such as the failed metabolization causing an accumulation of some active ingredient or the adverse effect diseases being a consequence of this accumulation.

It also becomes clear why -although technically it would be possible- on a practical level existing knowledge bases usually do not capture all of the information in narratives: Narratives may relate causal mechanisms or developments over time, which may refer to special cases only and may not be generally applicable. In this way, unlike the factual information collected in knowledge graphs, narratives usually do not feature truth values [13]. Entities and events related by a narrative may happen only in individual cases (in the sense of anecdotes), may be more or less probable (or possible), and only in the best case may be generally valid [7]. Moreover, the use of narratives may heavily determine their structures, e. g. more schematic for rigid scientific argumentation vs. quite free for storytelling. Therefore a new kind of representation is needed, enriched with strong links to factual knowledge or actual contexts. Our contributions are:

– We design a conceptual model for narratives and propose narratives to represent scientific argumentation in a structured way (Sec. 2).
– We introduce narrative bindings to verify (or at least plausibilise) each narrative by grounding its parts to individual knowledge repositories (Sec. 2).
– As a first real-world use case, we demonstrate that our model in extension with narrative queries is applicable for *scientific narratives* in the PubMed bio-medical collection (Sec. 3).

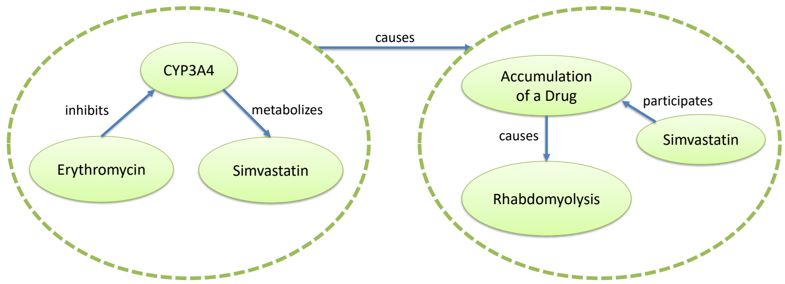

Fig. 1. Adverse Drug-Drug Interaction as a sample pharmaceutical Narrative

2 Modeling Scientific Narratives

In this section, we propose a model for narratives working along a typical pharmaceutical narrative (Fig. 1) as a sample use case.

2.1 Narratives in Science

A narrative structure forms the backbone of virtually every scientific publication. And while scientific narratives tend to be much more limited and restricted than general narratives in fictional stories, their basic structure is similar. This includes protagonists driving the story, and events impacting their behavior. In this paper, we consider real-world objects and concepts, i. e. **entities** as the story's main protagonists. Considering the example above, the drugs *simvastatin* and *erythromycin*, the disease *rhabdomyolysis* and the gene system *CYP3A4* are the entities of interest. We denote the set of all entities by $\mathcal{E}$. In the scope of an individual narrative, entities might interact with each other, which can be expressed in the form of subject-predicate-object *relationships*, e. g. *CYP3A4 metabolizes simvastatin*. Here, we refer to the well-known Resource Description Framework (RDF) for modeling factual knowledge [9]. Each relation is identified via a *predicate label* like *inhibits* or *metabolizes*. Besides entities, a narrative may speak about simple *literals* in the place of objects, i. e. strings or numerical values. For example, the treatment of patients with simvastatin is naturally associated with a specific dosage, e. g. simvastatin may be applied in a dose of 20mg per day. We denote the set of all literals by $\mathcal{L}$. Relationships between different entities or entities and literals can be understood as factual information. Prime examples are properties, e. g. *CYP3A4 metabolizes simvastatin*, as well as structural and ontological information about entities, i. e. the type or class of an entity. We call these relationships between entities and literals in a narrative **factual knowledge** denoted by $\mathcal{R}_F$. The set of possible predicate labels used for such factual knowledge is denoted by Σ_F. Hence, $\mathcal{R}_F \subseteq \mathcal{E} \times \Sigma_F \times (\mathcal{E} \cup \mathcal{L})$.

Besides entities and literals, narratives usually feature **events**. In our running example, the *accumulation of a drug* in the body of some patient is such an event, which might be observed and reported during a study. It describes the observation that the level of a drug in the patient's body increases. Hence, events can be understood in the sense of some labeled observation, which happens as the story progresses, i. e. an event is an observed state or a change of a state. Events may also feature a temporal dimension, i. e. an event occurs, having a starting point and sometimes an endpoint in time. We denote the set of all events by Γ.

In most scientific narratives, events are arranged in some order to describe the story's progress, e. g. the *accumulation of a drug* leads to a severe adverse effect inducing the disease *rhabdomyolysis*. There is much research invested in analyzing the characteristics of relationships between events [2, 3, 8]. There are several kinds of such relationships: Temporal relationships describe the temporal order of events, i. e. *a drug has to be administered* first, before *side effects may occur*. Causal relationships describe that an event causes some other event, e. g. *heart failure* leads to a *patient's death*. While temporal and causal relationships almost exclusively exist between events, entities can also be related to events. Usually, this either indicates that the entity participates in or is affected by some event. Whereas factual knowledge is more about properties and ontological information of entities, relations between events describe an argumentation's progress. Hence, we compose the progress of a narrative by a set of **narrative relationships** denoted by $\mathcal{R}_N$. In brief, narrative relations feature special, non-factual labels and can be placed between events or between events and entities, but not between entities. We denote the set of all narrative relation labels by Σ_{NR}. Hence, the set of narrative relationships $\mathcal{R}_N \subseteq (\Gamma \times \Sigma_{NR} \times (\mathcal{E} \cup \Gamma)) \cup (\mathcal{E} \times \Sigma_{NR} \times \Gamma)$.

Both narrative relations between entities and events, as well as factual knowledge between entities and literals, form the essential backbone of a narrative. A narrative might be composed inductively, e. g. the metabolism and inhibition behavior of CYP3A4, simvastatin, and erythromycin, which *as a whole* leads to the drug's accumulation. This behavior is also reflected by Hauser et al. [5], who characterize recursive elements as a key element in human language: a story can be composed using arbitrary sub-stories.

Definition 1. *A narrative is defined inductively:*

1. *A directed edge-labeled graph* (V, E) *is a narrative with* $V \subseteq \mathcal{E} \cup \mathcal{L} \cup \Gamma$ *being nodes and* $E \subseteq \mathcal{R}_F \cup \mathcal{R}_N$ *being edges.*
2. *If* n_1, n_2 *are narratives and* $p \in \Sigma_{NR}$*, then* (n_1, p, n_2) *is also a narrative.*

That means a narrative can be understood as edge-labeled directed graphs with events, entities, and literals as nodes and labeled edges between them. Modeling the real world usage of narratives, we also allow them to show a recursive structure. In our running example, the metabolization of simvastatin by CYP3A4 and its inhibition by erythromycin is a three-node narrative, which takes part as a new node in a second narrative on a higher level. Narratives can thus act as nodes in specific narrative relations, e. g. administering both drugs, shown on the left side of Fig. 1, *causes* an *accumulation*, resulting in an adverse effect, shown on the right side of Fig. 1. Please note that we intuitively visualize recursive narratives as directed edge-labeled graphs where nodes may encapsulate narratives: such nodes are depicted by dashed circles, which enclose a complete graph structure of another narrative. This means that the content of encapsulating nodes is again a directed graph, with entities, literals, and events being the nodes and the relations being the edges.

2.2 Narrative Bindings

With narratives formally defined, we now introduce narrative bindings connecting the narrative itself as a logical overlay to underlying knowledge repositories. Binding a narrative n to a knowledge repository kr means *grounding* n with data from kr as evidence. We understand the notion of knowledge repositories in a broad sense, i. e. any structured or unstructured form of data storage, such as knowledge graphs, relational databases, document collections, or data set registries.

Definition 2. *Let n be a narrative, e be an edge of the narrative n and kr be a knowledge repository, a narrative binding nb binds the edge e against the knowledge repository kr with $nb = (e, kr)$. We say that e is bound by nb.*

Due to the recursive structure of narratives, there exist two types of edges: edges between events, entities and literals and edges between enclosed narratives. Returning to our running example, narrative bindings might easily ground the factual knowledge in the narrative, i. e. (*erythromycin*, *inhibits*, *CYP3A4*) and (*CYP3A4*, *metabolizes*, *simvastatin*), to a knowledge graph capturing important properties of genes and drugs. It is important to note that each subgraph of a narrative can be bound to a different knowledge repository. If all parts of a narrative can be bound, we consider the narrative to be *grounded*. Formally,

Definition 3. *Let n be a narrative and NB be a set of narrative bindings, we call n grounded by NB, if all edges of n are bound by at least some $nb \in NB$.*

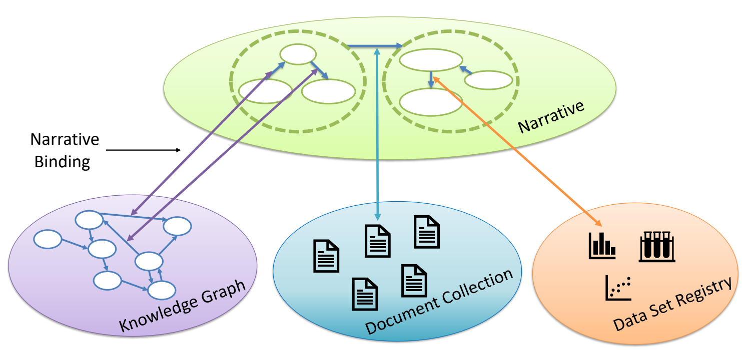

Fig. 2. Narratives as Logical Overlays on top of Knowledge Repositories

With narratives, we introduce a novel model to express any form of scientific discussion in a structured fashion, of course without making any claims to its validity. However, using bindings to ground a narrative against underlying knowledge repositories, at least provides some evidence for the narrative in the sense of *plausibility*. Regarding a specific knowledge repository, computing bindings in real-world applications relies on available methods for information retrieval, natural language processing and querying capabilities. Still, it is essential to note that a successful binding does not imply any *guarantees* on a narrative's validity, which is obviously heavily impacted by the quality of the respective repositories, but also by the somewhat difficult to assess the validity of information fusion over different sources [7].

3 Narrative Queries

We have modeled narratives to represent scientific argumentations in a structured way that is usable for information systems. However, how can we use scientific narratives in real-world applications? We introduce narrative queries with variables to support sense-making processes, i. e. generating new hypotheses [12]. We denote the set of all variable symbols by $\mathcal{V}$ and write each symbol by a leading question mark. A **narrative query** nq is some narrative n, where each entity, event or literal might be replaced by a variable symbol of $\mathcal{V}$. Hence, each narrative is also a (variable-free) narrative query.

Considering our running example, we might formulate a narrative query by replacing any node by some variable $?X$. By substituting variables we can then fill nodes by arbitrary entities, literals or events. In the following, we use the SPARQL notation[1]. The set of variables used in a narrative query nq is denoted by $\mathit{vars}(nq) = \{?v_1, \ldots, ?v_n\}$. A **substitution** μ from $\mathcal{V}$ to $\mathcal{E} \cup \mathcal{L} \cup \mathit{\Gamma}$ is a partial function: $\mu : \mathcal{V} \rightarrow \mathcal{E} \cup \mathcal{L} \cup \mathit{\Gamma}$. We define the subset of $\mathcal{V}$, where μ is defined, as the domain of μ, shortly $\mathit{dom}(\mu)$. The substitution of the variables in a narrative query nq by μ yields a narrative n, if all variables of the query are in $\mathit{dom}(\mu)$. We use $\mu(nq) = n$ as a shortcut for this substitution.

An **answer** to a narrative query nq is a pair $(\mu(nq), \mathit{NB})$ with a substitution $\mu(nq)$ and a set of narrative bindings NB, such that the following holds: 1. $\mathit{vars}(nq) \subseteq \mathit{dom}(\mu)$, 2. $\mu(nq) = n$, and 3. n is grounded by NB. As a consequence, answering a narrative query nq requires two steps: 1. obtaining all substitutions $\{\mu_1, \ldots, \mu_n\}$ and 2. obtaining narrative bindings grounding $\mu_i(nq)$ for each $\mu_i \in \{\mu_1, ..., \mu_n\}$. If a narrative query does not include variables, its answer is the empty substitution and a set of narrative bindings grounding the respective narrative.

Obviously, testing arbitrary substitutions for their narrative bindings is an expensive task. But, by first computing possible narrative bindings for the fixed parts of a narrative query, the set of feasible substitutions can be severely restricted. Moreover, finding such narrative bindings in structured knowledge repositories allows for the usage of efficient query languages such as SPARQL.

[1] https://www.w3.org/TR/rdf-sparql-query/

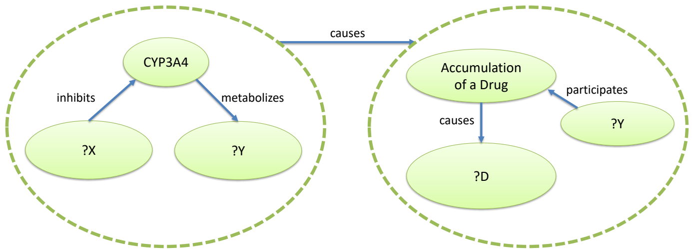

Fig. 3. A Pharmaceutical Narrative Query

3.1 Answering a Narrative Query with SemMedDB and PubMed

As a small showcase, we pose a narrative query in the biomedical domain. We compute all narrative bindings against two knowledge repositories, namely SemMedDB and PubMed. SemMedDB[2] is a knowledge graph comprising nearly 19 million medical facts in version 2019. PubMed[3] is the world's most extensive biomedical library with around 30 million publications and is publicly available as the PubMed Medline 2020. Let us design a pharmaceutical narrative query nq based on our running example, see Fig. 3. It asks for two drugs *?X* and *?Y*, which both interact with the gene system *CYP3A4*. This interaction leads to an adverse effect *?D* triggered by the accumulation of drug *?Y* in the body. A possible answer to the query nq are the substitution μ with $\mu(?X) = \textit{erythromycin}$, $\mu(?Y) = \textit{simvastatin}$ and $\mu(?D) = \textit{rhabdomyolysis}$ and the respective narrative bindings NB against SemMedDB and PubMed. In fact, $\mu(nq)$ is exactly our running example narrative.

But, how can we compute all answers to nq? Since $?X$ and $?Y$ are part of purely factual knowledge inside the narrative query, we can formulate a suitable SPARQL statement to query SemMedDB for possible substitutions of $?X$ and $?Y$ automatically. In contrast, as $?D$ does not participate in a factual, but in a narrative relationship, we can derive valid substitutions for $?D$ by searching for publications in PubMed, which talk about $\mu_i(?X)$, $\mu_i(?Y)$ and *CYP3A4* for each substitution $\mu_i \in \{\mu_1, \ldots, \mu_n\}$. If diseases are mentioned within such a publication, they serve as possible substitution μ_i for $?D$. As an example, if a publication talks about *simvastatin*, *erythromycin*, *CYP3A4* and *rhabdomyolysis*, we derive the corresponding μ with $\mu(?X) = \textit{erythromycin}$, $\mu(?Y) = \textit{simvastatin}$ and $\mu(?D) = \textit{rhabdomyolysis}$. Using SemMedDB2019 and the PubMed Medline 2020 we computed 1264 possible substitutions. These substitutions $\{\mu_1, \ldots, \mu_n\}$ can derive the narratives $\{n_1, \ldots, n_n\}$ by $\mu_i(nq) = n_i$ with $\mu_i \in \{\mu_1, \ldots, \mu_n\}$. Due to the nature of how we have computed the substitutions, each narrative in $\{n_1, \ldots, n_n\}$ is already grounded by bindings against SemMedDB and PubMed. This small experiment demonstrates that answering a narrative query can automatically derive a large set of grounded narratives.

[2] https://skr3.nlm.nih.gov/SemMedDB/
[3] https://pubmed.ncbi.nlm.nih.gov

4 Discussion

In sum, we introduce a novel narrative overlay model to represent scientific argumentations in a formal way. In contrast to integrating all knowledge repositories into a single one, which is obviously an prohibitive task, narratives are designed as logical overlays on top of different types of knowledge repositories. We introduce narrative bindings to bind a narrative against some knowledge repository, i. e. the narrative can be grounded by data of this knowledge repository. Grounding means to find evidence for the narrative in the sense of *plausibility*. Finding suitable narrative bindings to ground a narrative is still an open research task. We showed that (for easy cases) bindings might simply be computed using established query languages like SPARQL. However, query processing is not always that easy, e. g. entity and relation alignments must be computed automatically or at least semi-automatically. In the future, transforming the process of manually defining bindings to automatically computing them is worth investigating.

Designing narratives is a task for domain experts who are familiar with domain-specific argumentations. A domain expert can ground her narrative by suitable bindings, which give hints and, more or less, evidence about the correctness and validity of her narrative. In a first use case, incorporating narrative queries, we demonstrate how such a process is done in the biomedical scientific domain. Moreover, a narrative query that includes variables enables a domain expert to automatically design a template. This template can be used later to derive suitable narratives by computing narrative bindings against already established knowledge repositories. As an application, our example narrative query might be used to discover new knowledge, e. g. the interaction between *simvastatin* and *erythromycin* is inferred, iff the corresponding narrative is grounded. Hence, narrative queries support workflows for knowledge discovery by obtaining substitutions for variables and grounding them. Suppose a narrative cannot be grounded, but parts of it can be. In that case, a researcher can decide whether the not grounded parts are worth of investigation for future work.

Moreover, domain experts design narrative queries with hints for the computation of bindings once, and several researchers benefit from these templates later. A young researcher might efficiently utilize a narrative query to generate a new hypothesis or to find suitable provenance information by having a look at the obtained narrative bindings. Especially in domains where researchers are not familiar with query languages, pre-designed narrative queries in conjunction with hints for the computation of bindings assist their process by automatically querying different knowledge graphs. Although the design of our pharmaceutical narrative query might take some time, the query is used to explain thousands of drug-drug interactions with adverse effects. The showcase demonstrates that our narrative query is ready-to-use in the pharmaceutical domain for querying and obtaining bindings against SemMedDB and PubMed automatically. Narratives as logical overlays together with narrative queries enable domain experts to collect knowledge from several different kinds of knowledge repositories. In this way, domain experts can boost their applications' quality without the need for a complex integration of existing repositories.

5 Related Work

Extending the reach of knowledge graphs has been an extensive field of study for many years. In knowledge graphs reification [6], as applied in the singleton notation [11], and different strategies to capture provenance information [14] have been proposed. These extensions aim to capture contextual or situational knowledge that is usually not expressed due to the restrictive data structure of RDF using binary relations. These approaches usually require high manual expenditures, which contradicts the general idea of RDF to facilitate large scale knowledge repositories in an easy way. And even in these cases, storing complete narrative structures is usually not pursued.

Detecting stories in natural language texts is a topic that has sparked much interest. Chambers et al. discussed the idea of modeling texts by extracting temporally ordered sequences of events [2]. Li et al. discussed the generation of stories by using crowd-sourced plot graphs [8]. These stories are then analyzed to find their commonalities and to determine relevant events, as well as orders of event sequences. These works describe a story as a sequence of events. In contrast, we focus on modeling a complete scientific argumentation within a single model. The general characteristics of argumentation structures have been thoroughly analyzed by Toulmin et al. [13]. Argumentation mining aims to find suitable arguments to a topic automatically, i. e. extracting positive and negative arguments (pro and contra) [4, 10]. Especially in the scientific domain, where work is usually published in the form of a solid argumentation, a deeper understanding of such an argumentation and its structure is essentially needed.

6 Conclusion

Capturing argumentations in the form of narratives in a structured way has sparked much interest. While capturing arbitrary narratives raises many problems, we focus on scientific narratives, which are usually more limited and quite concise. In this paper, we conceptualize scientific argumentations in a novel narrative model, combining factual knowledge and narrative patterns within its scope. Utilizing a single knowledge repository to form a proper scientific narrative is not sufficient. Argumentations typically operate on different types of knowledge, e. g. on factual knowledge or situational knowledge, like observed results of an experiment. Grounding narratives by narrative bindings gives evidence about the narrative's validity and correctness - in the sense of plausibility. By understanding narratives as logical overlays that can be bound against different kinds of knowledge repositories, we bypass the extensive integration of different knowledge repositories. Hence, we argue to keep the sources separated and to build logical overlays on top of them. In a biomedical showcase, we utilize a narrative on top of two large-scale knowledge repositories demonstrating the applicability of narratives. Indeed, narrative structures are commonly used in a wide range of scientific argumentations, e. g. chemical pathways, new theories in physics, the behavior of systems and algorithms in computer science, sociological observations and many more. Narratives are designed as logical overlays to enable information systems to represent and ground scientific argumentations against several knowledge repositories within a single model. In the future, we will investigate applications utilizing scientific narratives to boost the quality of research tasks like hypothesis generation.

References

1. Auer, S., Kovtun, V., Prinz, M., Kasprzik, A., Stocker, M., Vidal, M.E.: Towards a Knowledge Graph for Science. In: Proc. of the 8th Int. Conf. on Web Intelligence, Mining and Semantics. pp. 1:1—-1:6. WIMS '18, ACM (2018). https://doi.org/10.1145/3227609.3227689
2. Chambers, N., Jurafsky, D.: Unsupervised learning of narrative event chains. In: Proc. of ACL-08: HLT. pp. 789–797 (2008)
3. Chang, D.S., Choi, K.S.: Causal relation extraction using cue phrase and lexical pair probabilities. In: Natural Language Processing – IJCNLP 2004. pp. 61–70. Springer Berlin Heidelberg (2005). https://doi.org/10.1007/978-3-540-30211-7_7
4. Habernal, I., Gurevych, I.: Argumentation mining in user-generated web discourse. In: Computational Linguistics. vol. 43, pp. 125–179. MIT Press (2017). https://doi.org/10.1162/COLI_a_00276
5. Hauser, M.D., Chomsky, N., Fitch, W.T.: The faculty of language: What is it, who has it, and how did it evolve? In: Science. vol. 298, pp. 1569–1579. American Association for the Advancement of Science (2002). https://doi.org/10.1126/science.298.5598.1569
6. Hernández, D., Hogan, A., Krötzsch, M.: Reifying RDF: what works well with wikidata? In: Proc. of the 11th Int. Work. on Scalable Semantic Web Knowledge Base Systems. CEUR Work. Proc., vol. 1457, pp. 32–47. CEUR-WS.org (2015)
7. Kroll, H., Kalo, J.C., Nagel, D., Mennicke, S., Balke, W.T.: Context-compatible information fusion for scientific knowledge graphs. In: 24th Int. Conf. on Theory and Practice of Digital Libraries (TPDL). Springer (2020)
8. Li, B., Lee-Urban, S., Johnston, G., Riedl, M.: Story generation with crowdsourced plot graphs. In: Twenty-Seventh AAAI Conf. on Artificial Intelligence (2013). https://doi.org/10.5555/2891460.2891543
9. Manola, F., Miller, E., McBride, B., et al.: Rdf primer. W3C recommendation **10**(1-107), 6 (2004)
10. Mochales, R., Moens, M.F.: Argumentation mining. In: Artificial Intelligence and Law. vol. 19, pp. 1–22 (2011). https://doi.org/10.1007/s10506-010-9104-x
11. Nguyen, V., Bodenreider, O., Sheth, A.: Don't like rdf reification?: Making statements about statements using singleton property. In: Proc. of the 23rd Int. Conf. on World Wide Web. pp. 759–770. WWW '14, ACM (2014). https://doi.org/10.1145/2566486.2567973
12. Spangler, S., Wilkins, A.D., Bachman, B.J., Nagarajan, M., Dayaram, T., Haas, P., et al.: Automated hypothesis generation based on mining scientific literature. In: Proc. of the 20th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining. p. 1877–1886. KDD '14, ACM (2014). https://doi.org/10.1145/2623330.2623667
13. Toulmin, S.E.: The uses of argument. Cambridge Univ. Press (Cambridge) (1958)
14. Wylot, M., Cudré-Mauroux, P., Hauswirth, M., Groth, P.: Storing, tracking, and querying provenance in linked data. In: IEEE Transactions on Knowledge and Data Engineering. vol. 29, pp. 1751–1764 (2017). https://doi.org/10.1109/TKDE.2017.2690299