OXFORD

Briefings in Bioinformatics, 2022, **23(4)**, 1–13
https://doi.org/10.1093/bib/bbac268
Advance access publication date: 12 July 2022
Review

Contexts and contradictions: a roadmap for computational drug repurposing with knowledge inference

Daniel N. Sosa and Russ B. Altman
Corresponding author. Russ Altman, Department of Bioengineering, Stanford University, Stanford, CA, USA; Department of Genetics, Stanford University, CA, USA; Department of Medicine, Stanford University, Stanford, CA, USA. Tel.: +1-650-725-3863; E-mail: russ.altman@stanford.edu

Abstract

The cost of drug development continues to rise and may be prohibitive in cases of unmet clinical need, particularly for rare diseases. Artificial intelligence-based methods are promising in their potential to discover new treatment options. The task of drug repurposing hypothesis generation is well-posed as a link prediction problem in a knowledge graph (KG) of interacting of drugs, proteins, genes and disease phenotypes. KGs derived from biomedical literature are semantically rich and up-to-date representations of scientific knowledge. Inference methods on scientific KGs can be confounded by unspecified contexts and contradictions. Extracting context enables incorporation of relevant pharmacokinetic and pharmacodynamic detail, such as tissue specificity of interactions. Contradictions in biomedical KGs may arise when contexts are omitted or due to contradicting research claims. In this review, we describe challenges to creating literature-scale representations of pharmacological knowledge and survey current approaches toward incorporating context and resolving contradictions.

Keywords: drug repurposing, knowledge graphs, natural language processing, metascience

Introduction

In 2016, it was estimated that developing new pharmaceuticals was an investment of approximately two billion dollars and 10 years from bench to bedside [1]. Small molecule design is slowing down due to increased the Food and Drug Administration regulations, unexpected toxicity in clinical trials and long cycle times [2, 3]. A recent estimate is that 48% drugs fail in Phase II trials due to poor efficacy [4]. With the traditional pipeline slowing down, pharmaceutical development to address unmet clinical need will benefit from screening drug candidates at larger scale and with higher likelihood to pass Phase II trials.

In the past couple of decades, many new data modalities have emerged across the biomedical domain. Genomic screens and genome-wide association studies (GWAS) have given rise to nuanced understand of disease etiology, especially for complex, highly prevalent diseases. CRISPR screens have also added to our understanding of disease mechanism by enabling specific perturbation of genetic function [5]. The NIH LINCS Consortium has generated data sets measuring gene-expression signature modulation in response to drugs and drug-like perturbagens [6]. In the clinic, the rise of electronic medical records has enabled new analyses of patient response to drug treatment at hospital scales. At home, wearables and other modalities are also new avenues to monitor patients' biometrics, which may be used to improve personalized patient care.

Artificial intelligence holds great promise for driving innovation for drug development in such a data- and compute-rich ecosystem. Deep learning methods can take advantage of large, high-dimensional, heterogeneous data sets and have recently led to state-of-the-art performance in several fields including computer vision, speech recognition and natural language processing (NLP). In pharmacology, deep neural networks have been successfully applied in several contexts including crowd-sourced challenges for predicting the activity of small molecule compounds [7] and drug toxicity [8]. These methods can generalize well and are amenable to tasks that can be formulated as machine learning problems with sufficient training data, which has led many pharmaceutical companies to adopt these technologies.

One important area of research to address unmet clinical need in a time of slowing drug development is drug repurposing. The idea underlying repurposing is to find new therapeutic opportunities for existing drugs, which

Daniel Sosa is a PhD candidate in the Biomedical Informatics Program at Stanford University.
Russ Altman is the Kenneth Fong Professor of Bioengineering, Genetics, Medicine, Biomedical Data Science and (by courtesy) Computer Science at Stanford University.

Received: March 28, 2022. **Revised:** May 25, 2022. **Accepted:** June 7, 2022

would circumvent expensive drug development *ab initio* while already demonstrating safety in a clinical setting [9]. Drug repurposing takes advantage of the fact that drugs have many targets [10, 11] and that many diseases share common etiology or clinical presentation, whereby a drug may target a common mechanism between different diseases [12, 13]. In the case of rare diseases, wherein the low prevalence of the disease or condition makes cost of drug development prohibitive, drug repurposing may be one of the few clinical options.

Many methods have been considered for computationally predicting drug repurposing opportunities. Previously, methods have used similarity of small molecules to drugs with known indications or ligand docking-based similarity to generate predictions [14, 15]. However, these methods are limited in that they do not consider the modules of interacting genes that contribute to complex disease phenotypes [16]. Systems biology approaches to drug discovery and repurposing are thus critical by considering the network of interacting genes and proteins as the drug 'target' rather than favoring the 'one drug–one target' paradigm [17].

Biomedical knowledge relevant for up-to-date drug repurposing is rapidly proliferating. Annually, nearly one million new citations are indexed in MEDLINE [18]. Knowledge bases such as DrugBank [19], PharmGKB [20] and OMIM [21] are maintained to represent the most pertinent information from a cross-section of literature at a given time. However, these gold standard databases are manually curated, and full manual curation, particularly for extracting new types of information from literature, is becoming infeasible.

There exists an opportunity to use NLP to extract biomedical knowledge from scientific literature directly. Automated methods are increasingly being used to increase curator efficiency and to directly retrieve important information [22]. One area of NLP relevant to network medicine approaches to drug repurposing is relation extraction—the automated mining of relationships between pairs of entities such as genes, proteins, drugs and diseases. Advancements in dependency parsing [23], distant supervision and foundational models [24], coupled with community-driven open shared tasks [25] have led to improved accuracy of relation extraction models. These methods have been applied at the scale of PubMed to generate global cross-sections of interactions between pharmacological entities directly from literature [26, 27]. Such global representations of knowledge are amenable to network inference methods to computationally predict repurposing hypotheses [28].

While developing systems for extracting and learning from the massive volume of scientific knowledge for drug repurposing is an exciting prospect, in this review, we will present key challenges with which these systems must reconcile to reach their full potential. These difficulties result from the observation that, when overly simplified, much of the knowledge extracted from literature appears to be contradictory. Many instances of apparent contradictions may be adjudicated by additional contextual information, such as cell-type specificity in which interactions are present. Other difficulties emerge from the uncertainty and constantly evolving nature of science, whereby many statements are in fact contradictory when revisited in future studies using new experimental modalities, for instance. This review describes existing work in defining and reconciling (or not) issues of quality control from the fine granularity of individual relationships to large, aggregate knowledge representations. By addressing these challenges, we present a path whereby knowledge inference systems may realize their potential in accelerating biomedical innovation *in silico* (Figure 1).

Background: promise and perils for literature-derived knowledge representations

By connecting the dots across biomedical literature, we can automatically infer latent knowledge based on facts that are already known. This idea, known as literature-based discovery (LBD), was proposed by Don Swanson in 1986 [28]. He studied how scientists could connect two disparate-seeming areas of research by looking for intermediate links connecting the two. In the context of drug repurposing, the two areas of being connected are the drug proposed for repurposing and the disease being treated. Examples of intermediate links implicating a repurposing opportunity include a biological pathway common to drug targets and disease etiology, drug side effects relevant for the target disease or chemical similarity to another drug indicated for the disease.

Multiple approaches have been implemented to formalize the notion of LBD [29]. In its original conception, Swanson described the 'ABC model' for connecting a source research area, A, to a target area, C, via an intermediate area, B, that has overlap with both A and C. In closed LBD, A and C are specified and the task is to explain what connection, B, might relate the concepts. In open LBD, only A is specified and the task is to find a related area C mediated through another concept B. Proposals for finding these connections have found related articles by term co-occurrence or citations, thus connecting concepts at the paper level. Later, distributional methods have been applied for finding related concepts. In this framework, numerical representations of concepts are learned based on distributional semantics. These methods may reveal interrelationships between concepts in a data-driven manner based on how semantics are learned in scientific discourse.

Finally, network-based approaches have been applied wherein representations of current knowledge are explicitly represented as a network of connected information. These methods represent a cross-section of knowledge aggregated across literature and benefit from being interpretable representations of interacting entities that are able to encode semantic relationships. These networks are well-suited for applications in systems medicine like drug repurposing and have potential for elucidating mechanisms of action.

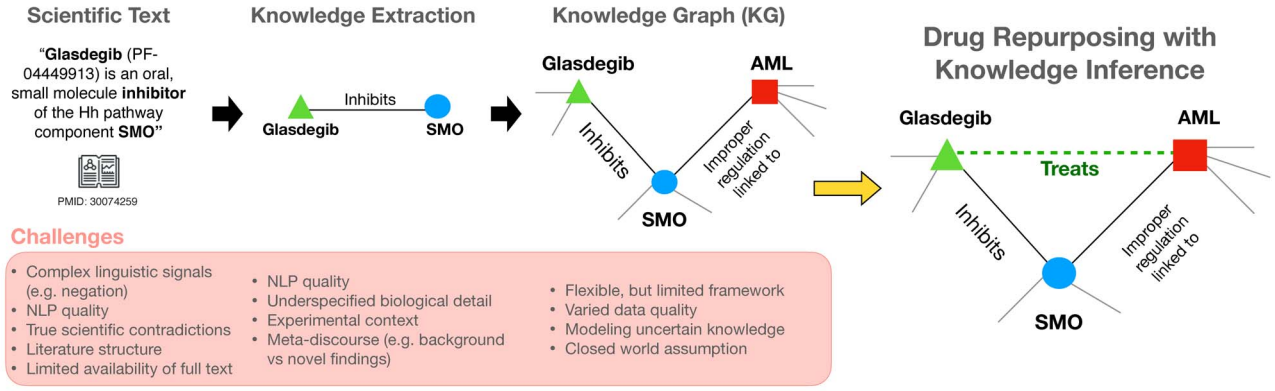

Figure 1. A LBD pipeline for drug repurposing with knowledge inference. Challenges present in scientific text, knowledge extraction and knowledge representation include: NLP quality, incorporation of key contextual information and representation and adjudication of contradictory information.

The most commonly used network model for knowledge is the knowledge graph (KG). KGs are flexible frameworks for storing information about relationships between entities. A KG is a network consisting of a set of entities, $\mathbb{E}$ and relations, $\mathbb{R}$. Each KG instance, $\mathbb{K}$ consists of triples $(h, r, t) \in \mathbb{K}$ whereby $h, t \in \mathbb{E}$ and $r \in \mathbb{R}$. Triples in KGs represent collections of relationships, such as (*The Mall of America*, `is_located_in`, *Minnesota*), which may be compiled from structured databases or extracted from text using NLP methods. These graphs are well-suited for representing unstructured data or data that would otherwise be cumbersome in a relational database, and graph databases have enabled quick querying and storage for massive KGs. For drug repurposing, KGs can describe interactions between drugs, genes, protein, pathways, diseases, anatomy and other biological entities. Systemic phenomena such as feedback loops or conditional relationships can be modeled over these semantics-rich networks.

KGs have been constructed from literature using rule-based, machine learning and distributional semantics-based methods. SemRep [30] is a rule-based system underlying the SemMedDB KG [27]. GNBR [26] is a KG of interacting drugs, proteins, genes and diseases that was constructed using an unsupervised approach [31] clustering together sentences with common dependency parse structures to extract more nuanced semantics. Constructed KGs can be combined with embeddings derived from language models to improve the inferred dependencies between related concepts and is at the current frontier of research [32].

The ability to extract, represent and manage large databases of knowledge and apply inference methods to derive new insights has been a boon to the field of knowledge engineering [33]. A class of methods known as KG embedding methods leverages learned, numerical representations of entities and relations in a KG to infer new links that may be implicated in the graph. These inference methods have been applied to applications across domains including KG completion, question answering and logical prediction generation [34, 35]. These methods can be readily applied to drug repurposing by framing the task as link prediction between drugs and diseases.

Since its inception, LBD has been applied to drug repurposing in several contexts. In recent applications it has been applied to identify general [36] or cancer-specific [37] therapeutics for repurposing based on common or proximal gene/protein targets under the ABC framework. Using a KG completion framework, LBD has been applied to NLP-extracted KGs to identify repurposing opportunities for prostate cancer drugs based on gene-mediated motifs in the network [38]. For rare diseases [39] and COVID-19 treatments [40], knowledge embedding methods have been applied to generate hypotheses agnostic of proposed mechanism schemas.

Evaluating knowledge inference methods is challenging for hypothesis generation tasks, and many inference methods are developed in with controlled settings in mind. Commonly employed KGs for benchmarking knowledge inference methods can be described as generic or domain-independent KGs. Here, the KG represents a collection of linked data about entities, for example a representation of a set of facts, that may span across domains or constitute 'general knowledge'. YAGO [41], Wikidata [42] and Cyc [43] are examples of such KGs useful for benchmarking and common-sense reasoning. For knowledge engineering methodologies to be clinically useful—as in the case of translating a drug repurposing hypothesis—they must be able to reason with the complexities associated with a concrete domain-like systems pharmacology knowledge from biomedical literature [44, 45].

One often-overlooked facet of domain-specific KG inference applications is the data over which the method is applied. In these settings, questions of KG quality, utility and credibility must be carefully considered for having translational impact [46]. When evaluating usefulness, the level of granularity of information must be considered, and simple, featureless, qualitative relationships between entities may be insufficient for many biological tasks. In this domain, there is clear utility to integrating diverse data sources, modeling uncertainty,

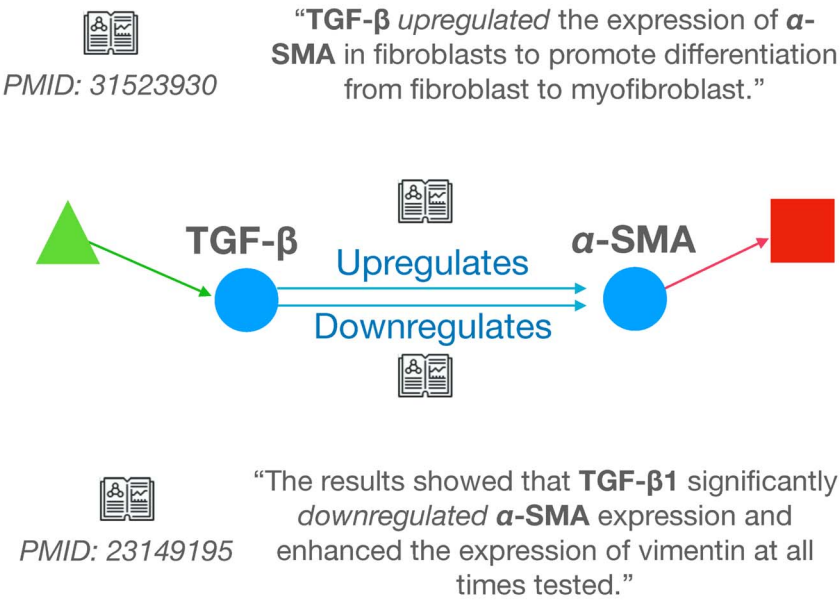

Figure 2. An example of two extracted sentence-level predications suggesting contradictory information about a PPI.

remediating contradictory-seeming facts and explicitly encoding entity- and predicate-level features such as tissue specificity or temporal information [47]. Previous work in knowledge inference has demonstrated that careful attention must be paid to the combination of the embedding model employed, the training approach and the loss function [48]. In the case of scientific KGs, adding explicit representations of uncertainty or feature information such as context qualifiers on interactions may be pertinent yet is a largely open area of research.

The presence of contradictory information poses a particularly large challenge for scientific KGs. The phenomenon that two predications—binary relationships of the form (entity X, relation `r`, entity Y)—will make incompatible propositions is common when combining multiple sources of knowledge. These contradictions have been shown to be highly prevalent in mechanistic genetic and protein–protein interaction (PPI) networks whereby two entities may be joined by both 'upregulates' and 'downregulates' events simultaneously, which seem to be semantically but not necessarily biologically inconsistent [49] (Figure 2).

Contradictions apparent in literature-derived KGs may stem from several causes. Rosemblat *et al.* [50] estimated the rate of apparent contradictions after filtering candidate pairs from SemMedDB [27], a PubMed-scale KG, at 2.6%. This study demonstrates that in most cases, statements that appear contradictory need to be qualified by information such as population group being studied, species or dosage group. In some cases, the results indicate true scientific contradictions, where the authors are intentionally presenting a finding contradictory with existing knowledge. This work underscores the prevalence and variability of contradictions in KGs.

The prevailing Open World Assumption for biological knowledge bases—that these KGs are incomplete and non-observed facts are not necessarily false—complicates the traditional KG inference framework applied to benchmark data sets [51]. Accumulation of biological knowledge is an ongoing enterprise, and scientific knowledge bases are imperfect and incomplete. The academic publishing ecosystem presents an additional challenge due to positive outcome bias in scientific literature, whereby fewer negative results are reported over time due to the difficulty of publishing negative results in high-impact journals [52]. Currently, the degree to which the sparsity and quality of KGs affects the performance of modern machine learning on downstream biomedical tasks is not well characterized.

While developing knowledge inference methods for LBD is a promising area of research, reasoning over complex systems such as literature-derived KGs requires a more fine-grained approach for representing knowledge. Existing paradigms commonly operate on graphs representing simple entities and relations at a high level of abstraction, failing to reconcile contradictions represented via negation events or incorporate nuanced features of biological entities important for applications [35]. In this review, we describe challenges to literature-derived knowledge representations and present ongoing work pertaining to extracting contradictory predications, modeling important contextual information and resolving inconsistencies.

Contradictions arising at the textual level

Fundamentally, contradictions are a linguistic phenomenon. In the strictest logical definition, two sentences are contradictory if it is impossible for them

to both be true simultaneously. For many linguistic applications, a looser definition is used, which states that two sentences are contradictory if they are extremely unlikely to be true simultaneously [53].

The semantics underlying why two statements are contradictory can be subtle and highly variable. De Marneffe *et al.* [53] define a typology enumerating common types of contradictions, which they divide into two categories based on the difficulty by which the contradiction might be detected by automated methods. In the simpler category are antonymy, negation and numeric inconsistencies such as date or quantity mismatch. By using lexicons of antonyms and negation triggers and text mining quantitative textual data, these types of contradictions can be detected in a straightforward manner [54, 55].

De Marneffe *et al.* go on to describe four types of challenging contradictions: factive and modal, structural, lexical and world knowledge-based. Factive and modal contradictions include those in which the author is expressing emotion or doubt with regards to the central clause that appears to contradict the central clause of another sentence being compared. Structural contradictions are those that contradict by how the verbs are being used structurally in the sentence. For example, if sentences A and B are equivalent except for swapping the two subjects of a binary non-commutative verb such as `upregulates`, a contradiction may emerge. Lexical contradictions may arise from idiomatic uses of words or changes in semantics that result in the context of phrases, which may be difficult for automated methods to capture. Finally, contradictions may arise based on contextual assumptions or world knowledge that is known to be true but may not be detectable based on syntactic and structural evidence alone.

This typology is indicative of the challenge of detecting contradictions in text. Even in the 'simple' case of detecting negation events and antonymy, NLP systems are error prone. For instance, Wu *et al.* [56] describe challenges for negation detection, which many consider to be a solved problem. They note that negation detection methods are developed with specific domains in mind, which are used to train algorithms. The authors demonstrate the challenge these methods have with generalizing to domains with different syntactic or semantic conventions. When proceeding with contradiction detection applications that rely on the results of NLP methods such as negation detection, it is important to keep in mind that errors in text mining might affect the ability to detect true contradictions downstream.

In biomedical text mining, seemingly contradictory statements may take any of the forms previously described. One chief application of literature-based text mining is deriving relations between pairs of biological entities from scientific literature. Relevant to pharmacological settings are relationships between combinations of genes, proteins, disease, drugs and tissues. Collections of relationships may be contradictory because different authors came to different conclusions about the relationship's veracity. Depending on the author's stance regarding the extracted relationship and contextual information about, for example, experimental protocols, any of the previously mentioned types of contradictions may arise. Further, inconclusive clinical evidence about the efficacy of a drug for a treatment based on a collection of clinical trials or case reports, for instance, may lead to authors contradicting one another in a factive or modal sense. Some language might even necessitate an understanding of the standards of drug regulation to infer the author's stance on a treatment's efficacy-world knowledge that is not immediately accessible via text alone.

Biological text, particularly concerning molecular biology, presents additional challenges to the contradiction detection task. Text mining systems for entity normalization may simplify biological entities in a way that may make claims about fundamentally different biological entities seem comparable. Accurate biological representations must take into consideration if the gene or a gene product (RNA or protein) are being described in text, or what sort of state the protein is in, for instance if it is in a modified state or bound to a ligand. Poor normalization may result in many false positive contradiction pairs.

A second challenge is co-reference resolution, the problem of whether two spans of text are referring to the same entity especially across sentences. When a complex pathway is described in text, such co-references must be resolved several times over potentially multiple sections. Failure to resolve co-references within or across sections of text may lead to missed opportunities to detect contradictory events. The degree to which the molecular biology domain presents additional challenges for contradiction detection is a largely unexplored area of research.

Despite its challenges, automatically detecting sentence-level entailments and contradictions is essential to understanding natural language and has become a well-established task in natural language inference. In these contexts, contradictions are evaluated based on local inference and due to linguistic variation rather than relying on outside knowledge to infer if two statements are contradictory or entailing [57]. Large improvements in this space have been made thanks to the recent availability of large corpora for this task [58, 59]. Such corpora have spanned a variety of domains and are thus well suited for domain-specific detection of entailments and contradictions [60]. Concurrently, advances in machine learning including attention-based LSTMs [61] and Transformer [62] models have enabled state-of-the-art performance on these tasks. Taken together, our ability to automatically detect contradictory events at the scale of all scientific literature is a fruitful area of research.

We have described how unpacking the nature of linguistic contradictions reveals a complex landscape. Quality concerns for contradiction detection may result from errors in NLP pipelines or issues from

overgeneralizing across domains. These nuances are present in biomedical text in conjunction with additional challenges such as difficulty in normalizing biological entities and co-reference resolution across long spans of molecular biology text. Despite these challenges, much progress has been made in the field of natural language inference due to the creation of large, standardized data sets for training and evaluation. Deep learning methods also present great promise for learning to recognize subtle linguistic patterns with large amounts of training data and have already been applied successfully on these large benchmark corpora. Progress on automated contradiction detection will be made by developing more domain-specific corpora, training models for classifying different types of contradictions, continuing to improve performance on subtasks including negation detection and co-reference resolution.

True contradictions are highly prevalent in science

As a natural consequence of the progression of science, research on drug targets, disease etiology, or pharmacodynamic or pharmacokinetic mechanism may be contradictory. Scientific publications can be distilled into a claim pertaining to a research question that is supported or refuted by empirical study [63]. Subsequent studies may follow-up by refuting claims with their own studies or reach different conclusions pertaining to the same research question, and these disagreements are common in clinical literature. In one study, researchers observed that of 49 highly cited clinical studies, seven were later contradicted and another seven were followed by studies that reported weaker effects than the initial study [64]. The connection of vitamin E and coronary disease is one famous example of contradictory clinical evidence [65, 66]. These studies have important, persistent effects on our perception of current knowledge, and contradictory or refuted results may continue to be cited for years after their original publication [67]. The situation is exacerbated in cases when scientific preprints without the gold standard of peer review are used to guide clinical decision-making, as was demonstrated during the COVID-19 pandemic [68].

Biological literature is also rife with contradictory claims. Research has demonstrated a 'Proteus phenomenon' in meta-analyses of molecular genetics and clinical trials research whereby an initial study demonstrates a strong effect connecting, for instance, a genetic variant and a disease and subsequent analyses published in lower-impact journals demonstrate weaker or even opposite associations [69]. These contradictions are indicative of the evolving nature of data collection and analysis that are natural byproducts of the scientific method and potentially of meta-scientific publication biases, whereby blockbuster results are more likely to be published and disseminated.

One area of molecular biology in which contradictory evidence is well defined is with respect to protein–protein and gene–gene interactions. A contradiction may occur whereby in the same system one paper reports an upregulation event between gene *A* and gene *B* but another paper reports a downregulation event. To create a high-confidence set of contradicting events, Kim *et al*. created a data set known as BioContrasts consisting of contrastive information in text abstracts, specifically searching for the contrastive negation pattern of '`X` but not `Y`' as in the case of a binding event [70]. Another related work looked at mentions of negated PPIs through negation flags along a PPI dependency path [71]. This corpus could be used to seed a collection of pairs of contradictory statements about PPIs. These methods and corpora aim at capturing events whereby different articles capture true inconsistencies across our current knowledge of molecular biology.

Finally, contradictions in molecular biology and genomics may arise by differences in epistemic cultures and research methods employed. For instance, as much of genomics research has evolved from hypothesis driven to hypothesis discovering, large-scale computational methods may naturally give rise to contradictions with other computational predictions or previously reported empirical studies [72].

To detect true contradictions in literature, previous work has been done on the task of first extracting core research claims from literature. Typically, the research claim is defined as the summary of the main points presented in a research argument, which can either introduce new knowledge to readers or update their knowledge on a topic. One recent approach for this task used transfer learning of an LSTM model trained to identify from which section of a structured PubMed–RCT abstract a sentence was categorized. This trained model was applied to the task of identifying the core research claim in an unstructured abstract using an annotated data set. The authors demonstrated the utility of the transfer learning strategy and achieved an F1-score of 0.78 for the task [73].

Limited work has been conducted on exploring contradictions of research claims in literature. Reliable corpora are necessary for training and evaluating models developed for this task. Two corpora developed for this purpose are AutConCorpus and ManConCorpus [74]. AutConCorpus was automatically generated from SemMedDB-a literature-derived collection of binary predications [27]—by querying predications in which the described pair of entities is connected by incompatible or entailing predicates, for instance in the case of incompatible predications: '*A* `produces` *B*' and '*A* `does not augment` *B*' [75]. ManConCorpus is a manually generated corpus of claims based on systematic reviews about topics related to cardiovascular disease wherein forest plots might be indicative of contradictory results. Annotators were asked to label research claims of papers as agreeing with or disagreeing with PICO questions

related to the systematic review, such as 'In patients undergoing coronary artery bypass, does Aspirin usage, compared to no aspirin, cause bleeding?' [76]. In this way, pairs of contradicting or entailing pairs can be derived depending on the yes or no stance of each to the PICO question [77].

These corpora have been applied in the task of contradictory claim detection. In these studies, features were extracted to attempt to automatically identify the stance the claim (hypothesis) is taking with respect to the PICO question (premise). One work employed extraction of uni- and bi-grams occurring at a sufficient frequency, presence of negation terms, presence of terms indicating directionality (e.g. 'blood pressure increased after treatment X'), and sentiment detection from respective lexicons, achieving micro-averaged F1 scores of 0.83 and 0.78 with an SVM classifier on ManConCorpus and AutConCorpus, respectively [74]. Later work builds upon these features by using relation extraction to distill relevant clauses from the sentence [78] and extracting features related to negation, antonym detection and textual alignment, achieving improved F1 scores on the ManConCorpus of 0.94 for the entailing research pairs and 0.87 for contradictory pairs.

Contradictions in research are a natural byproduct of the scientific enterprise. Clinical literature is rife with contradictory information over time as medical practice evolves and study design changes or improves in different settings. Contradictory information is present in molecular biological literature as our knowledge of interacting systems increases and the data modalities and analytical techniques available to experimental biologists grow. When a research paper is distilled to its central claim(s) about which it supports or refutes, the task of automatically detecting true contradictions can be formalized. NLP methods for extracting research claims from literature are only starting to emerge. Additionally, creating larger corpora of contradictory claims will be imperative to train domain-specific deep learning models for contradictory claims detection in biomedical literature. The field of automated detection of true scientific contradictions has much room for continued exploration with great potential for clarifying our current understanding of scientific knowledge at large scales.

Adjudicating apparent contradictions with context

Many instances of seemingly contradictory statements may be adjudicated by additional context, making the task of contradiction detection more challenging. For drug repurposing, much of this context, such as anatomical location(s) of drug action, is critical. Previous work on contradiction detection has demonstrated how background knowledge about meronyms, synonyms and other semantic functions may shed light on false positive instances of detected contradictions [79]. Other work has made the distinction of contradictions grounded in common associated contexts, such as anatomical location or species, versus those where statements have differing contexts, which the researchers refer to as contrasts [80]. In this work, the authors also note that of the contradictory events they find, these pairs suffer from underspecified contextual information, whereby more accurate context association would resolve many instances as false positives. Differences in experimental methodology including the use of computational predictions versus experimental observations and the use of varying model organisms or cell lines are important contextual information for qualifying claims extracted from literature [72]. In clinical settings, quantitative context may be pertinent to extract, such as the drug dosage administered in a particular regimen.

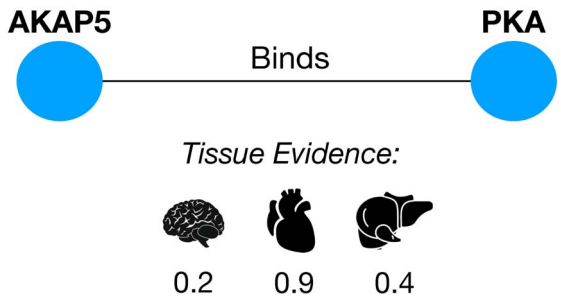

Figure 3. Incorporating contextual information such as tissue specificity of PPIs may help adjudicate apparent contradictions in KGs.

Explicitly incorporating context will provide massive benefit for a variety of semantic technology applications. Knowledge representations over which inference methods are applied for drug repurposing often operate on limited abstractions and often omit important information such as tissue specificity of interactions within the KG [39, 81] (Figure 3). For state-of-the-art proximity and machine learning-based methods, increased specificity of tissue and disease annotations would help prioritize more biologically feasible and pharmacologically principled targets, especially in urgent global health crises such as the COVID-19 pandemic [82].

Incorporating information at a finer granularity may also prove useful for KG embedding methods—the core class of methods for KG inference. One challenge is the problem of generating negative samples when learning embeddings, and current methods typically rely on a closed world assumption whereby triples not in the network are assumed to be false. Predominant sampling strategies for generating negatives include random sampling of nodes in the network and corrupting existing knowledge in the network, which may lead to false negatives [83]. If contextual information is incorporated, biological KG embedding methods may rely on biological knowledge to sample true negatives, for examples PPIs known to not occur in specific cell types. Incorporating domain knowledge into the negative sampling strategy will lessen the reliance on the closed-world assumption.

Scientific literature is highly structured and knowledge extraction systems that fail to consider meta-discourse may miss key context. Liakata *et al.* [84] described 11 conceptualization zones within a paper that specify the objective of the text: Hypothesis, Motivation, Goal, Object, Background, Method, Experiment, Model, Observation, Result and Conclusion. Several of these types of statements will be present in information-dense abstracts, which is the only available text for many papers given journal paywalls and copyright considerations. However, many pertinent mechanistic details may only be present in full text.

When structured information is lost, KGs may include false-positive scientific relations that correspond to a research hypothesis that was not confirmed by its study or discourse about the plausibility of a contradictory result that is speculative in nature. Further, the fact that multiple papers mention a common relation will not necessarily indicate greater evidence of its veracity. This is the case when multiple authors describe an existing fact on which subsequent work is predicated in the background sections of text.

Other pertinent meta-knowledge to extract includes the confidence level with which authors are claiming results. Work has been done on detecting the certainty and novelty of a claimed result using syntactic features [73, 85–87], but there is still a large opportunity to improve these methods and leverage their predictions in large-scale knowledge representations for LBD.

Outside knowledge may be another important source of context. For instance, in the translational drug repurposing pipeline from large-scale drug screens to *in vivo* experiments to clinical trials, the objective of experiments changes from optimizing for sensitivity to optimizing for precision. A contradiction between a finding from a drug screen and a clinical study would thus benefit from an outside understanding of the regulatory landscape and the nature of the scientific translational pipeline. Methods for adjudicating contradictions in this setting should know to prioritize clinical findings that meet rigorous safety standards and demonstrate efficacy in humans.

Researchers have recognized the need for more context awareness in knowledge representations and have taken multiple approaches to address this challenge. One approach entailed extracting knowledge from biomedical literature with a specific narrow scope in mind such as tolerogenic cell therapy [88]. In this work relations were extracted between cell types and cytokines; thus, the entities themselves incorporate key contextual information for the queried cellular system. Another approach involved returning to primary literature and associating contextual information with extracted relations *ad hoc*. Semantic and syntactic features were extracted about annotated contextual mentions such as species or cell line and about the sentences from which KG predications were extracted. These derived features were used in a supervised framework of associating predications to relevant contextual mentions in the same paper [89].

Finally, network topology and other structured information may be used for inferring important contextual information. In the work of Zitnik and Leskovec [90], multicellular function is learned from a hierarchical representation of tissue networks based on a tissue ontology, which can even be used to generalize cellular function prediction to uncharacterized tissues. Other approaches may benefit from integration of multiple data sources like GIANT [91] and STITCH [92], which are networks that represent predicted and experimentally derived tissue-specific interactions. Integrating heterogeneous data sources would augment and increase the utility of existing general-purpose biomedical knowledge representations.

Detecting contradictory claims from scientific literature is complicated by the importance of contextual information that may adjudicate otherwise seemingly contradictory claims. Differences in experimental settings may explain how two statements may appear to contradict but are in fact simultaneously possible. By limiting the scope of knowledge to be extracted, representing context information as entities in the graph or analyzing latent features in the network topology, one can learn or represent important contexts. *Ad hoc* methods for extracting contextual information and incorporating it at inference time are promising but are still early in development. Applying Transformer-based NLP models and harmonizing KGs with external context-representative databases will be another fruitful direction of work for explicitly representing context. Together, these methods will help resolve instances of false-positive contradictory claims in biomedical literature.

Reconciling the problem of contradiction in KGs

The issue of KG quality is well characterized in the field of semantic technology. Zaveri *et al.* [93] gave a thorough overview of quality assessment for linked data, particularly when harmonizing across multiple data sets. The authors provided a breadth of metrics commonly used to assess the quality of a KG including dimensions such as accuracy, timeliness, completeness, relevancy, consistency, availability and verifiability [94]. Most relevant to this review is the notion of consistency, which has been defined as the absence of logical contradictions in the data [95]. When considering how apparently contradictory statements might be adjudicated in a KG, other characteristics such as relevancy, trustworthiness, understandability and timeliness may be more prescient than consistency at face-value. Färber *et al.* demonstrated a variety of heuristics to evaluate these quality checks for large domain-independent KGs used for benchmarking knowledge inference methods [96]. The authors showed

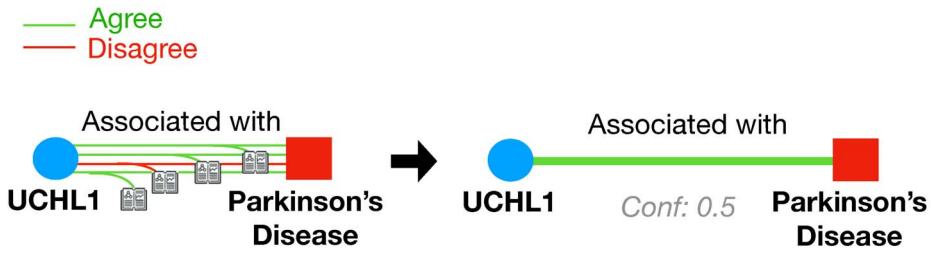

Figure 4. Quantitative representations of the confidence of knowledge, as in the case of contradictory GWAS gene–disease associations, can be used in frameworks for reasoning under uncertainty including soft logic and Bayesian methods.

that these KGs have widely varying degrees of consistency, completeness and accuracy-important factors that may qualify downstream knowledge inference results and are often downplayed.

KGs derived using NLP are known to be particularly susceptible to challenges of data quality as previously described in this review. For example, SemMedDB [27] has been shown to have a non-insignificant number of contradicting predications. When searching for a predication and its contradiction, such as (X, `causes`, Y) and (X, `NOT_causes`, Y), SemMedDB was observed to contain nearly 500 000 such inconsistencies [97]. Empirical results have shown that the performance of downstream knowledge inference methods may degrade with increasing sparsity and noise as in the case of NLP-extracted KGs [98].

To combat issues of KG quality and inconsistencies, one might consider KG refinement preprocessing approaches or alternative modeling decisions. The task of identity link invalidation, finding invalid instances of predications of the form (X, `is_same_as`, Y), is a related problem involving detection of logical inconsistencies in a KG [99]. Researchers have applied techniques from logic theory and numerical approximations using learned embedding representations for this task [100–102]. Methods for detecting logical inconsistencies in a KG would serve as useful preprocessing tools to detecting both errors caused by NLP relation extraction and naturally occurring inconsistencies from science, which may be adjudicated by incorporating contextual information.

Quantitative attributes of knowledge will help deal with uncertainty and inconsistencies in literature-derived KGs. Preprocessing approaches such as KG refinement or otherwise modeling uncertainty of predications in literature-derived KGs would allow users to explicitly represent a notion of confidence of the veracity of a predication. This is particularly useful in a naturally uncertain domain such as biomedical research and would make knowledge representations amenable to a variety of methods developed for dealing with uncertainty (Figure 4).

Fuzzy, probabilistic and plausible logic methods may be applied to infer new information with an associated confidence conditional on the accuracy of input facts used to drive inference [103–105]. The notion of modeling uncertainty has been discussed in the domain of data fusion whereby associating confidence scores with ground truths in a silver standard database would penalize methods less for generating predictions that may be inconsistent with more uncertain ground truths [106]. Bayesian approaches have also been considered for modeling biological processes like constructing transcription factor networks [107] and for network link prediction and by considering path information in probabilistic networks [108, 109]. Hybrid approaches are also possible combining the strength of graphical models with user-input semantic objectives and completely data-driven embedding methods for link prediction under uncertainty [110]. Representing uncertainty would prove useful in representing, acting upon and evaluating biomedical KGs.

Finally, more expressive knowledge representations may capture nuance that would help adjudicate apparent inconsistencies. Frameworks such as Minsky's frames [111], attributed KGs [112] or new standards such as the Biological Expression Language [113] are flexible enough to incorporate more granular information. By explicitly attributing predications with context, provenance or other relevant information, general-purpose KGs may be more principled and thus useful for specific applications. These attributes may be used to induce simple knowledge subgraphs (for example specific to a certain collection of cell types) or act as constraints depending on the task, domain area or disease system.

Many quality control metrics have been proposed and studied for large KGs including completeness, consistency and relevancy. Data quality issues are often overlooked but have implications not only for noisier KGs including those derived using NLP methods but also commonly used domain-independent KGs employed for benchmarking tasks. Methods for detecting inconsistencies and ascertaining uncertainty as pre-processing steps are promising for improving inference quality and qualifying downstream knowledge inference results and comprise a largely unexplored area of research in the context of KGs. More general and expressive KG representations also enable quantitative or contextual information to be incorporated, which may improve KG relevancy, consistency and trustworthiness. Such extensions to

simple KGs have been seldom considered and represent a fruitful area of research for next-generation knowledge engineering applications.

Conclusions

Therapeutic breakthroughs with repurposed drugs predicated on LBD is an exciting prospect that is beginning to come to fruition. Literature-derived knowledge representations hold great promise for increasing the scope of scientific knowledge over which we make inference, and this resource offers great potential for discovery in urgent clinical applications such as predicting repurposing opportunities for rare disease treatment. In this review, we have presented a collection of challenges concerning quality assurance that such knowledge representations will have to overcome for the next generation of knowledge inference methods. Particularly, we frame the presence of contradictions and inconsistencies as a central challenge. Apparent contradictions may arise from a great diversity of sources including NLP errors, scientific controversy or oversimplifications of pairs of claims that may not in fact be contradictory when adjudicated by context. We have demonstrated recent work characterizing and tackling some of these challenges and have alluded to open areas of research in applying and improving biomedical literature-derived KGs. This review sheds light on opportunities for having large impacts in this exciting, rapidly evolving and highly interdisciplinary field.

Key Points

- Knowledge graphs (KGs) from biomedical literature text represent rich semantics between drugs, genes, proteins, pathways and diseases extracted from up-to-date biomedical knowledge.
- Drug repurposing hypothesis generation *in silico* is well posed as link prediction in these graphs, but these biological KGs are inherently noisy and sparse.
- Scientific KGs are prone to inconsistencies because of errors in NLP, unresolved contexts and the contradictory and evolving nature of science.
- Associating key biological context with relationships will increase the utility of existing KGs for pharmacological innovation.

Funding

National Institutes of Health (grant numbers GM102365 and LM007033); Stanford Data Science Scholars Program to D.S.

References

1. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: new estimates of R&D costs. *J Health Econ* 2016;**47**:20–33.
2. Kantarjian HM, Prat F, Steensma DP, *et al.* Cancer research in the United States: a critical review of current status and proposal for alternative models. *Cancer* 2018;**124**(14):2881–9.
3. Milsted RAV. Cancer drug approval in the United States, Europe, and Japan. *Adv Cancer Res* 2006;**96**:371–91.
4. Harrison RK. Phase ii and phase iii failures: 2013-2015. *Nat Rev Drug Discov* 2016;**15**(12):817–8.
5. Doench JG. Am I ready for CRISPR? A user's guide to genetic screens. *Nat Rev Genet* 2018;**19**(2):67–80.
6. Subramanian A, Narayan R, Corsello SM, *et al.* A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 2017;**171**(6):1437–52.
7. Dahl GE, Jaitly N, Salakhutdinov R. Multi-task neural networks for QSAR predictions. *arXiv*, 2014, https://arxiv.org/abs/1406.1231. (1 April 2022, date last accessed) Preprint.
8. Mayr A, Klambauer G, Unterthiner T, *et al.* Deeptox: toxicity prediction using deep learning. *Front Environ Sci* 2016;**3**:1–15.
9. Nosengo N. Can you teach old drugs new tricks? *Nature* 2016;**534**(7607):314–6.
10. Koch U, Hamacher M, Nussbaumer P. Cheminformatics at the interface of medicinal chemistry and proteomics. *Biochim Biophys Acta* 2014;**1844**:156–61.
11. Paolini GV, Rhb S, van Hoorn WP, *et al.* Global mapping of pharmacological space. *Nat Biotechnol* 2006;**24**(7):805–15.
12. Benfey PN, Mitchell-Olds T. From genotype to phenotype: systems biology meets natural variation. *Science (New York, NY)* 2008;**320**(5875):495–7.
13. Glicksberg BS, Li L, Cheng W-Y, *et al.* An integrative pipeline for multi-modal discovery of disease relationships. *Biocomputing* 2014;407–18.
14. Keiser MJ, Setola V, Irwin JJ, *et al.* Predicting new molecular targets for known drugs. *Nature* 2009;**462**(7270):175–81.
15. Defranchi E, De Franchi, Schalon C, *et al.* Binding of protein kinase inhibitors to synapsin i inferred from pair-wise binding site similarity measurements. *PloS One* 2010;**5**(8):1–11.
16. Menche J, Sharma A, Kitsak M, *et al.* Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science* 2015;**347**(6224):1–8.
17. Piro RM. Network medicine: linking disorders. *Hum Genet* 2012;**131**(12):1811–20.
18. Medline citation counts by year of publication (as of January 2021) [Internet]. Bethesda, MD: National Institutes of Health; 2021 January [cited 2022 April 1]. Available from: https://www.nlm.nih.gov/bsd/medline_cit_counts_yr_pub.html
19. Wishart DS, Feunang YD, Guo AC, *et al.* Drugbank 5.0: a major update to the drugbank database for 2018. *Nucleic Acids Res* 2018;**46**:D1074–82.
20. Thorn CF, Klein TE, PharmGKB ARB. The pharmacogenomics knowledge base. *Methods Mol Biol* 2013;**1015**:311–20.
21. Amberger JS, Bocchini CA, Schiettecatte F, *et al.* OMIM.org. Online Mendelian inheritance in man (OMIM), an online catalog of human genes and genetic disorders. *Nucleic Acids Res* 2015;**43**:D789–98.
22. Lu Z. Pubmed and beyond: a survey of web tools for searching biomedical literature. *Database (Oxford)* 2011;**18**:baq036.
23. Quan C, Wang M, Ren F. An unsupervised text mining method for relation extraction from biomedical literature. *PLoS One* 2014;**9**(7):1–8.
24. Bommasani R, Hudson E, Adeli DAA, *et al.* On the opportunities and risks of foundation models. *arXiv* 2021, https://arxiv.org/abs/2108.07258. (1 April 2022, date last accessed) Preprint.

25. Lever J, Jones S. Painless relation extraction with kindred. In: *BioNLP* 2017; 176–83.
26. Percha B, Altman A, RB. Global network of biomedical relationships derived from text. *Bioinformatics* 2018;**34**(15):2614–24.
27. Kilicoglu H, Shin D, Fiszman M, *et al.* Semmeddb: a PubMed-scale repository of biomedical semantic predications. *Bioinformatics* 2012;**28**(23):3158–60.
28. Swanson DR. Undiscovered public knowledge. *Lib Q* 1986;**56**(2):103–18.
29. Cheerkoot-Jalim S, Khedo KK. Literature-based discovery approaches for evidence-based healthcare: a systematic review. *Health Technol* 2021;**11**(6):1205–17.
30. Kilicoglu H, Rosemblat G, Fiszman M, *et al.* Broad-coverage biomedical relation extraction with SemRep. *BMC Bioinformatics* 2020;**21**(1):188.
31. Percha B, Altman RB. Learning the structure of biomedical relationships from unstructured text. *PLoS Comput Biol* 2015;**11**(7):e1004216 Publisher: Public Library of Science.
32. Nadkarni R, Wadden D, Beltagy I, *et al.* Scientific language models for biomedical knowledge base completion: an empirical study. arXiv, 2021. https://arxiv.org/abs/2106.09700 (1 April 2022, date last accessed) Preprint.
33. Studer R, Benjamins VR, Fensel D. Knowledge engineering: principles and methods. *Data Knowl Eng* 1998;**25**(1):161–97.
34. Wang Q, Mao Z, Wang B, *et al.* Knowledge graph embedding: a survey of approaches and applications. *IEEE Trans Knowl Data Eng* 2017;**29**(12):2724–43.
35. Hamilton WL, Bajaj P, Zitnik M, *et al.* Embedding Logical Queries on Knowledge Graphs. *Advances in Neural Information Processing Systems* 2018;**31**:1–12.
36. Rastegar-Mojarad M, Elayavilli RK, Li D, *et al.* A new method for prioritizing drug repositioning candidates extracted by literature-based discovery. In: *2015 IEEE International Conference on Bioinformatics and Biomedicine*, 2015, 669–74.
37. Yang H-T, Jiun-Huang J, Wong Y-T, *et al.* Literature-based discovery of new candidates for drug repurposing. *Brief Bioinform* 2017;**18**(3):488–97.
38. Zhang R, Cairelli MJ, Fiszman M, *et al.* Exploiting literature-derived knowledge and semantics to identify potential prostate cancer drugs. *Cancer Informatics* 2014;**13s1**:CIN.S13889 Publisher: SAGE Publications Ltd STM.
39. Sosa DN, Derry A, Guo M, *et al.* A literature-based knowledge graph embedding method for identifying drug repurposing opportunities in rare diseases. *Biocomputing*, 2020; 463–74.
40. Zhang R, Hristovski D, Schutte D, *et al.* Drug repurposing for COVID-19 via knowledge graph completion. *J Biomed Inform* 2021;**115**:103696.
41. Suchanek FM, Kasneci G, Weikum G. Yago: a core of semantic knowledge unifying wordnet and Wikipedia. *Semantic Web* 2007;**10**:697–706.
42. Vrandečić, D and Krötzsch, M.. Wikidata: a free collaborative knowledgebase. *Commun. ACM* 2014;**57**(10):78–85.
43. Matuszek C, Cabral J, Witbrock M, *et al.* An introduction to the syntax and content of cyc. In: *AAAI Spring Symposium on Formalizing and Compiling Background Knowledge and Its Applications to Knowledge Representation and Question Answering*, 2006; 44–9.
44. Su C, Tong J, Zhu Y, *et al.* Network embedding in biomedical data science. *Brief Bioinform* 2020;**21**(1):182–97.
45. Nelson W, Zitnik M, Wang B, *et al.* To embed or not: network embedding as a paradigm in computational biology. *Front Genet* 2019;**10**:1–11.
46. Abu-Salih B. Domain-specific knowledge graphs: a survey. *J Netw Comput Appl* 2021;**185**(10307):1–38.
47. Bonner S, Barrett IP, Ye C, *et al.* A review of biomedical datasets relating to drug discovery: a knowledge graph perspective. *arXiv* 2021. https://arxiv.org/abs/2102.10062 (1 April 2022, date last accessed) Preprint.
48. Ali M, Berrendorf M, Hoyt CT, *et al.* Bringing light into the dark: a large-scale evaluation of knowledge graph embedding models under a unified framework. *IEEE Trans Pattern Anal Mach Intell* 2021;**1**:1–20.
49. Domingo-Fernández D, Mubeen S, Marín-Llaó J, *et al.* Pathme: merging and exploring mechanistic pathway knowledge. *BMC Bioinformatics* 2019;**20**(1):243.
50. Rosemblat G, Fiszman M, Shin D, *et al.* Towards a characterization of apparent contradictions in the biomedical literature using context analysis. *J Biomed Inform* 2019;**98**(103275):1–11.
51. Alshahrani M, Thafar MA, Essack M. Application and evaluation of knowledge graph embeddings in biomedical data. *PeerJ Computer Science* 2021;**7**:e341.
52. Fanelli D. Negative results are disappearing from most disciplines and countries. *Scientometrics* 2012;**90**(3):891–904.
53. de Marneffe M-C, Rafferty AN, Manning CD. Finding contradictions in text. In: *Proceedings of ACL*, 2008; 1039–47.
54. Chapman WW, Bridewell W, Hanbury P, *et al.* A simple algorithm for identifying negated findings and diseases in discharge summaries. *J Biomed Inform* 2001;**34**(5):301–10.
55. Harabagiu S, Hickl A, Lacatusu F. Negation, contrast and contradiction in text processing. *AAAI* 2006;**6**:755–62.
56. Wu S, Miller T, Masanz J, *et al.* Negation's not solved: generalizability versus optimizability in clinical natural language processing. *PLoS One* 2014;**9**(11):1–11.
57. Pham MQN, Nguyen ML, Shimazu A. Using shallow semantic parsing and relation extraction for finding contradiction in text. *IJCNLP* 2013; 1017–21.
58. Bowman SR, Angeli G, Potts C, *et al.* A large annotated corpus for learning natural language inference. *EMNLP* 2015; 632–642.
59. Williams A, Nangia N, Bowman S. A broad-coverage challenge corpus for sentence understanding through inference. *NAACL* 2018;**1**:1112–22.
60. Romanov A, Shivade C. Lessons from natural language inference in the clinical domain. *EMNLP* 2018; 1586–1596.
61. Sergeeva E, Zhu H, Prinsen P, *et al.* Negation scope detection in clinical notes and scientific abstracts: a feature-enriched LSTM-based approach. *AMIA Jt Summits Transl Sci Proc.* 2019;**10**:212–221.
62. Liu Y, Ott M, Goyal N, *et al. Roberta: A Robustly Optimized Bert Pretraining Approach*, arXiv, 2019. https://arxiv.org/abs/1907.11692 (1 April 2022, date last accessed) Preprint.
63. Clark T, Ciccarese PN, Goble CA. Micropublications: a semantic model for claims, evidence, arguments and annotations in biomedical communications. *J Biomed Semantics* 2014;**5**(1):28.
64. Ioannidis JPA, Trikalinos TA. Early extreme contradictory estimates may appear in published research: the proteus phenomenon in molecular genetics research and randomized trials. *J Clin Epidemiol* 2005;**58**(6):543–9.
65. Stephens NG, Parsons A, Schofield PM, *et al.* Randomised controlled trial of vitamin e in patients with coronary disease: Cambridge heart antioxidant study (chaos). *Lancet (London, England)* 1996;**347**(9004):781–6.
66. Dagenais S, Pogue G, Bosch J, *et al.* Heart Outcomes Prevention Evaluation Study Investigators. Vitamin E supplementation and cardiovascular events in high-risk patients. *N Engl J Med* 2000;**342**(3):154–60.
67. Tatsioni A, Bonitsis NG, Ioannidis JPA. Persistence of contradicted claims in the literature. *JAMA* 2007;**298**(21):2517–26.

68. Fraser N, Brierley L, Dey G, *et al.* The evolving role of preprints in the dissemination of Covid-19 research and their impact on the science communication landscape. *PLoS Biol* 2021;**19**(4):1–28.
69. Ioannidis JPA, Trikalinos TA. Early extreme contradictory estimates may appear in published research: The proteus phenomenon in molecular genetics research and randomized trials. *J Clin Epidemiol* 2005; **58**(6): 543–9.
70. Kim J-J, Zhang Z, Park JC, *et al.* Biocontrasts: extracting and exploiting protein-protein contrastive relations from biomedical literature. *Bioinformatics (Oxford, England)* 2006;**22**(5):597–605.
71. Sanchez-Graillet O, Poesio M. Negation of protein-protein interactions: analysis and extraction. *Bioinformatics (Oxford, England)* 2007;**23**(13):424–32.
72. Torgersen H. Fuzzy genes: epistemic tensions in genomics. *Sci Cult* 2009;**18**(1):65–87.
73. Achakulvisut T, Bhagavatula C, Acuna D, *et al.* Claim extraction in biomedical publications using deep discourse model and transfer learning. *Cs* 2020:1–11.
74. Alamri A, Stevenson M. A corpus of potentially contradictory research claims from cardiovascular research abstracts. *J Biomed Semantics* 2016;**7**(36).
75. Alamri A. *The detection of contradictory claims in biomedical abstracts*. Doctoral Dissertation. University of Sheffield, 2016.
76. Ma X, Ma C, Yun Y, *et al.* Safety and efficacy outcomes of preoperative aspirin in patients undergoing coronary artery bypass grafting: a systematic review and meta-analysis. *J Cardiovasc Pharmacol Ther* 2014;**19**(1):97–113.
77. Alamri A, Stevenson M. Automatic detection of answers to research questions from Medline abstracts. *Proc BioNLP* 2015; 141–146.
78. Del Corro L, Gemulla R. Clause-based open information extraction. In: *Proceedings of the 22nd International Conference on World Wide Web - WWW '13*, 2013; 355–66.
79. Ritter A, Soderland S, Downey D, *et al.* It's a contradiction–no it's not: A case study using functional relations. In: *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, 2008, 11–20.
80. Sarafraz F. *Finding Conflicting Statements in the Biomedical Literature*. PhD thesis. University of Manchester, 2012.
81. Krämer A, Billaud J-N, Tugendreich S, *et al.* The Coronavirus network explorer: mining a large-scale knowledge graph for effects of SARS-COV-2 on host cell function. *BMC Bioinformatics* 2021;**22**(229):1–20.
82. Gysi DM, Zitnik M, Ameli A, *et al.* Network medicine framework for identifying drug repurposing opportunities for Covid-19. *ArXiv* 2020.
83. Ali M, Hoyt CT, Domingo-Fernández D, *et al.* Biokeen: a library for learning and evaluating biological knowledge graph embeddings. *Bioinformatics* 2019;**35**(18):3538–40.
84. Liakata M, Saha S, Dobnik S, *et al.* Automatic recognition of conceptualization zones in scientific articles and two life science applications. *Bioinformatics (Oxford, England)* 2012;**28**(7):991–1000.
85. Ananiadou S, Thompson P, Nawaz R. Enhancing search: events and their discourse context. In Gelbukh A, editor, *Computational Linguistics and Intelligent Text Processing*, Lecture Notes in Computer Science, pp. 318–34, Berlin, Heidelberg, 2013. Springer.
86. Ananiadou S, Thompson P, Nawaz R, *et al.* Event-based text mining for biology and functional genomics. *Brief Funct Genomics* 2015;**14**(3):213–30.
87. Shardlow M, Batista-Navarro R, Thompson P, *et al.* Identification of research hypotheses and new knowledge from scientific literature. *BMC Med Inform Decis Mak* 2018;**18**:1–13.
88. Lamurias A, Ferreira JD, Clarke LA, *et al.* Generating a tolerogenic cell therapy knowledge graph from literature. *Front Immunol* 2017;**8**.
89. Noriega-Atala E, Hein PD, Thumsi SS, *et al.* Extracting inter-sentence relations for associating biological context with events in biomedical texts. *IEEE/ACM Trans Comput Biol Bioinform* 2020;**17**(6):1895–906.
90. Zitnik M, Leskovec J. Predicting multicellular function through multi-layer tissue networks. *Bioinformatics* 2017;**33**(14):i190–8.
91. Wong AK, Krishnan A, Troyanskaya OG. Giant 2.0: Genome-scale integrated analysis of gene networks in tissues. *Nucleic Acids Res* 2018;**W1**:W65–70.
92. Kuhn M, von Mering, Campillos M, *et al.* Stitch: Interaction networks of chemicals and proteins. *Nucleic Acids Res* 2008;**36**:D684–8.
93. Zaveri A, Rula A, Maurino A, *et al.* Quality assessment for linked data: a survey. *Semantic Web* 2016;**7**(1):63–93.
94. Strong DM, Lee YW, Wang RY. Data quality in context. *Commun ACM* 1997;**40**(5):103–10.
95. Blomqvist E, Cochez M, d'Amato C, *et al. Knowledge Graphs*. arXiv, 2021. https://arxiv.org/abs/2003.02320 (1 April 2022, date last accessed) Preprint.
96. Färber M, Bartscherer F, Menne C, *et al.* Linked data quality of dbpedia, freebase, opencyc, wikidata, and yago. *Semantic Web* 2017;**9**(1):77–129.
97. Cong Q, Feng Z, Li F, *et al.* Constructing biomedical knowledge graph based on semmeddb and linked open data. In: *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018, 1628–31.
98. Pujara J, Augustine E, Getoor L. Sparsity and noise: Where knowledge graph embeddings fall short. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Copenhagen, Denmark: Association for Computational Linguistics, September 2017, 1751–6.
99. Saïs F. *Knowledge Graph Refinement: Link Detection, Link Invalidation, Key Discovery and Data Enrichment*. PhD thesis. Université Paris Sud, 2019.
100. Papaleo L, Pernelle N, Saïs F, *et al.* Logical detection of invalid SAMEAS statements in rdf data. *Knowl Eng Knowl Manage* 2014;**8876**:373–84.
101. Papageorgiou DC, Pernelle N, Saïs F. Approche numérique pour l'invalidation de liens d'identité (owl:sameas). *28es Journées francophones d'Ingénierie des Connaissances IC 2017* 2017;**28**:50–5.
102. Raad J, Pernelle N, Saïs F. detection of contextual identity links in a knowledge base. In: *Proceedings of the Knowledge Capture Conference*, 2017, 1–8.
103. Bach SH, Broecheler M, Huang B, *et al.* Hinge-loss Markov random fields and probabilistic soft logic. *J Mach Learn Res (JMLR)* 2017;**18**:1–67.
104. Hur A, Janjua N, Ahmed M. A survey on state-of-the-art techniques for knowledge graphs construction and challenges ahead. arXiv:2110.08012 [cs]. 2021; arXiv: 2110.08012.
105. Richardson M, Domingos P. Markov logic networks. *Mach Learn* 2006;**62**(1):107–36.
106. Dong XL, Gabrilovich E, Heitz G, *et al.* From data fusion to knowledge fusion. *Proc VLDB Endow* 2014;**7**(10):881–92.
107. Jin G, Changhe F, Zhao H, *et al.* A novel method of transcriptional response analysis to facilitate drug repositioning for cancer therapy. *Cancer Res* January 2012;**72**(1):33–44.

108. Zhang S, Li L, Cunlai Pu, *et al*. A Bayesian method for link prediction with considering path information. In Zhai XB, Chen B, Zhu K, editors, *Machine Learning and Intelligent Communications*, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, pp. 361–74, Cham, 2019. Springer International Publishing.
109. Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. *Mach Learn* 1992;**9**(4):309–47.
110. Chen X, Chen M, Shi W, *et al*. Embedding uncertain knowledge graphs. *Proc AAAI Conf Artif Intell* 2019;**33**(01):3363–70.
111. Minsky M. *A framework for representing knowledge*. MIT-AI Laboratory Memo 306. http://web.media.mit.edu/~minsky/papers/Frames/frames.html (1 April 2022, date last accessed).
112. Cen Y, Zou X, Zhang J, *et al*. Representation learning for attributed multiplex heterogeneous network. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, 1358–68.
113. Fluck J, Madan S, Ansari S, *et al*. Training and evaluation corpora for the extraction of causal relationships encoded in biological expression language (bel). *Database* 2016; **2016**.