Preprints.org

Article

Not peer-reviewed version

A New Paradigm for Scientific Query Answering

Jacob Martinez , Woods Ali , Madison Taylor *

Posted Date: 29 August 2024

doi: 10.20944/preprints202408.2157.v1

Keywords: information retrieval; neural ranking models,; natural language processing; data-driven decision making

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A New Paradigm for Scientific Query Answering

Jacob Martinez, Woods Ali and Madison Taylor *

University of Charleston
* Correspondence: mtaylor@ucwv.edu

Abstract: In modern agriculture, decision-making is deeply intertwined with data, yet essential agricultural knowledge remains encased within dense scientific texts such as journals, manuals, and free-text reports. This necessitates specialized search systems capable of distilling and delivering precise information in response to specific queries from agricultural users. This study introduces AgriQuery, a sophisticated agent designed to field natural language questions from the agricultural sector by exploring vast troves of scientific documents. A comprehensive review and analysis of farmers' information needs have been conducted, leading to the creation of an information retrieval test collection that includes actual queries, a vast array of scientific documents segmented into passages, and a set of ground truth relevance assessments that map these passages to corresponding queries. We explore and benchmark several information retrieval models, highlighting the efficacy of advanced neural ranking models in this context. AgriQuery's proposed deployment architecture features a client interface on the Telegram platform and a backend model on standard hardware. This test collection aims to catalyze further research into matching natural language queries with scientific text answers, with implications extending beyond agriculture.

Keywords: information retrieval; neural ranking models,; natural language processing; data-driven decision making

---

1. Introduction

The 21st-century agricultural landscape is transforming through mechanization, data-driven strategies, and evidence-based practices, even within emerging economies [1–7]. Despite the availability of an extensive array of resources—ranging from research findings, and meteorological data to soil analyses—significant challenges remain in harnessing these resources effectively due to their dispersion across varied formats and databases [13].

The domain of question answering (QA) [14,15] systems has evolved significantly since its inception, marked by the development of early rule-based systems in the 1960s and 1970s. These systems, such as BASEBALL and LUNAR, were designed to answer questions from a restricted domain of knowledge, using pattern matching and handcrafted rules to interpret queries and fetch answers from structured databases. The focus during this era was primarily on understanding natural language queries and converting them into an executable form against a predefined schema [16,17]. This period laid the foundational work for subsequent advancements in natural language processing and highlighted the challenges associated with language ambiguity and query intent interpretation.

With the advent of machine learning and the rise of the internet, the 1990s and 2000s witnessed a transformative shift in QA systems towards data-driven approaches. The introduction of the Text Retrieval Conference (TREC) QA tracks in the late 1990s provided a platform for evaluating different computational models on a common dataset, spurring innovation in this field. Researchers began employing statistical methods and later machine learning techniques to improve the accuracy of QA systems [22]. Notably, IBM's Watson, which famously competed on Jeopardy!, exemplified the integration of natural language processing, information retrieval, and machine learning to understand complex queries and retrieve answers from a massive corpus. This era underscored the potential of AI in understanding and processing human language in a more flexible and context-aware manner.

The last decade has seen rapid advancements in QA systems fueled by breakthroughs in deep learning and the availability of large annotated datasets like SQuAD (Stanford Question Answering

© 2024 by the author(s). Distributed under a Creative Commons CC BY license.

Dataset) [23,24]. Modern QA systems utilize models such as BERT (Bidirectional Encoder Representations from Transformers) and its variants, which leverage transformer architectures to understand the context of a passage more deeply than ever before. These models are pre-trained on extensive corpuses of text and fine-tuned on specific QA tasks, achieving remarkable success in both open-domain and closed-domain QA settings. Current research continues to push the boundaries of these systems, focusing on challenges such as answer verifiability, multi-hop reasoning, and the ability to handle more nuanced and conversational forms of querying, which are crucial for applications in domains ranging from academic research to customer service and beyond [25].

The fragmentation of valuable agricultural data across heterogeneous datasets poses considerable access barriers [26,27]. Much of this information remains in unstructured natural language form, buried in scientific publications and project reports, or scattered across various external applications and services. Currently, there is no unified federative system that provides a single point of access for agricultural stakeholders to search and utilize this information efficiently, which hampers the practical application of critical insights derived from such data [28].

Conversely, while digital connectivity has improved, it does not represent the primary hurdle in accessing agricultural resources. Farmers increasingly utilize digital tools and platforms, such as smartphones and social media, to stay updated on agricultural innovations and community insights [29,30].

The fundamental challenge lies in addressing the complex and multifaceted information needs of farmers effectively. For example, a query like "What are the most resistant wheat varieties to crown rot?" exemplifies the typical complexity of questions that are difficult to automate due to several factors:

- **Complex Answer Matching**: Farmers may phrase their queries in a manner that does not straightforwardly align with the relevant information. This complexity is compounded by variations in query phrasing, requiring robust systems to manage such diversity.
- **Focused Answer Delivery**: Farmers require concise, direct answers rather than extensive documents, which necessitates systems capable of distilling and presenting the essence of the needed information succinctly.

The discrepancy between the language used in farmers' queries and the technical language of relevant information is significant. Farmers might describe an issue using symptomatic language aimed at identifying a problem and determining a treatment course, while the relevant scientific documents might use more technical terms. This semantic gap necessitates systems that can bridge the linguistic divide automatically, accommodating variations in growers' technical knowledge—ranging from specific scientific terminology to more general colloquial expressions.

This paper presents AgriQuery, a comprehensive framework designed to meet the nuanced information needs of the agricultural community. It thoroughly analyzes real-world agricultural information-seeking behaviors, sources, and challenges. It offers an extensive test collection specifically tailored for scientific article search within the agricultural domain, comprising over 86,846 documents segmented into approximately 9,441,693 passages, curated by experts rather than through generic web scraping or crowdsourcing techniques. This collection includes 210 detailed, real-world query scenarios each featuring a natural language question, multiple keyword variations, expert-crafted responses, and a graded relevance of passages to queries.

We conduct a series of retrieval experiments utilizing both traditional term-based models and cutting-edge neural rankers, providing insights into the adaptation of general-domain models for specialized domains. Moreover, AgriQuery proposes an end-to-end system providing a unified search interface for agricultural stakeholders, encapsulating all aspects from user needs through to a deployable search system.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

2. Related Work

Conversational agents have emerged as effective tools to provide precise answers to queries from agricultural practitioners [1,4,31]. However, the landscape of such solutions remains relatively sparse.

Several innovative systems have been developed following the release of an extensive dataset containing queries from agricultural workers via the Kian Call Center (KCC). Tailored to meet the needs of the Indian agricultural community, KCC provided a phone helpline where farmers could seek advice from experts on best practices. Notable systems leveraging this dataset include AgriBot [32,33], FarmChat [6], and Krushi [34], each addressing unique aspects of agricultural queries.

AgriBot, now renamed AgriQuery, was specifically designed to cater to the informational needs of growers, encompassing weather forecasts, market trends, plant health, and governmental funding avenues. This system utilized sent2vec embeddings [39,40] and advanced entity extraction techniques to map user queries to a robust database of commonly asked questions and answers, derived from five years of aggregated data across all Indian states. However, AgriQuery transcends its predecessors by eschewing the need for manually curated, domain-specific databases in favor of a comprehensive scientific repository.

FarmChat, another precursor to AgriQuery, functioned as a speech-based conversational system that harnessed decision rules and IBM Watson APIs to facilitate intent recognition and manage dialogue flow. Predominantly focused on potato farming in Ranchi, India, this system was constrained by its narrow agricultural and regional focus, and lacked the scalability of machine-learning-driven knowledge extraction. AgriQuery, in contrast, is designed to be agnostic to crop type and region, significantly enhancing its applicability.

Krushi, aimed at addressing a spectrum of agricultural needs from weather forecasting to soil testing in nine districts of Maharashtra, India, utilized the RASA X conversational AI framework to enable interaction through WhatsApp. While Krushi demonstrated the potential of conversational agents in localized settings, AgriQuery extends this concept by incorporating a broader, more inclusive approach.

Internationally, other systems have been developed to facilitate agricultural data access. For instance, a user study in Taiwan utilized 1,000 farmer interviews to train an LSTM-based conversational model, focusing on sales, logistics, and plant care [41,42]. In Tanzania, a crop protection system was devised to assist farmers during disease outbreaks, gathering over 2,100 queries in Swahili from face-to-face interactions [46]. This study highlighted a distinct preference among younger farmers for concise SMS-based queries, contrasting with older farmers' inclination towards more elaborate questions.

Despite these developments, the field still faces significant challenges, such as regional and crop-specific limitations and the lack of a unified database that can dynamically respond to new, unseen queries. Moreover, the reliance on manual data curation impedes scalability and limits the applicability of these systems [46,49,50].

AgriQuery addresses these challenges by being fundamentally agnostic to crop type, region, and query nature. It leverages a vast array of scientific data and employs cutting-edge neural ranking models to align user queries with the most relevant information snippets, rather than entire documents. This not only ensures a high degree of precision in responses but also enhances the scalability and flexibility of the system, setting a new standard for conversational agents in agriculture.

3. Preliminary

In the diverse landscape of agriculture, users can be primarily divided into three categories: growers (farmers), agronomists, and specialists. The latter two categories consist of experts who offer support to farmers through various means including paid consultations or governmental sponsorships, translating recent research findings into actionable advice. Each category has distinct information needs that must be addressed to optimize agricultural outcomes.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

This section delves into the literature surrounding the information needs of growers and extends this understanding through an online survey targeting all three user types. The insights gathered here are crucial for developing an agriculture-specific test collection to assess the efficacy of search systems tailored to this sector.

3.1. Types of Information Needs

The literature review highlights several critical information needs specific to growers, focusing on crop production and excluding animal husbandry to maintain scope relevance. Important findings from the literature [1,6,56] include:

3.1.1. Crop Protection

Growers frequently seek information on how to protect their crops from pests and diseases, often triggered by visible symptoms like "brown spots on leaves" which help in diagnosing the issue and formulating a response plan, such as choosing the appropriate fungicide and its application specifics. Queries often extend to pest identification, like a "2cm black and yellow snail," which dictates the pesticide choice. Weed management also constitutes a significant portion of growers' queries [56]. It is noteworthy that the growers' descriptions rarely match the technical terminologies found in typical answers, requiring sophisticated inference capabilities from search systems.

3.1.2. Best Practices

With agriculture being a dynamic field, growers continuously seek information to enhance their yields, reduce costs, and thereby increase profitability. Staying updated with the latest agricultural practices is deemed crucial [1]. Growers not only seek specific information reactively but also value proactive recommendations, with digital platforms like Twitter becoming a popular medium for such exchanges [29,30].

3.1.3. Unbiased Product Recommendations

Given the significant financial implications, growers prioritize credible recommendations for agricultural products like fertilizers, seeds, and pesticides. Such unbiased guidance is crucial for making informed purchasing decisions [6].

3.1.4. Market and Climate Insights

Understanding market dynamics and weather conditions is vital for growers who need to adapt their practices accordingly. They require detailed insights into market prices, both current and projected, and how these might affect their operations [56]. Similarly, weather conditions and climate patterns play a critical role in planning agricultural activities, necessitating tailored information that addresses past, present, and future climatic data.

3.2. Understanding what Growers Want

Through an online survey targeting Australian farmers, agronomists, and specialists, we gathered data on their specific information needs, focusing particularly on grain crops. The survey assessed the frequency, urgency, and types of information sought in real-life scenarios. The initial findings, though based on a limited participant pool, offer valuable insights into designing a user-centric test collection for **AgriQuery**.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

Table 1. The revised search scenario questionnaire for AgriQuery.

Sample Question: What quantity of nitrogen fertilizer is necessary for my crops post-drought this year?	
Considering the sample question:	
Q1.	`How critical is resolving this query for agricultural success?`
Q2.	`How often does this concern arise?`
Q3.	`What is the urgency of obtaining a timely response?`
Q4.	`Identify the top 5 informational resources you would consult.`
Q5.	`List at least 3 search queries you might use on Google or the GRDC website to find answers.`
Q6.	`Describe 3 crucial elements that an ideal answer would include.`
Q7.	`What volume of information do you expect in the response?`
Q8.	`How should the information be specifically tailored to your circumstances? Include at least 3 points.`
Q9.	`Provide a concise summary of the answer, if known.`
Q10.	`If you've previously sought answers to this, how effective was the response?`

Table 2. Updated statistics from the information needs survey for AgriQuery.

Role	16
Grain grower	9
Grain crop specialist	4
Agronomist (farm consultant)	3
Years of experience	16
10 years or more	11
Between 5 and 9 years	3
Between 1 and 4 years	1
Less than 1 year	1
Education	16
Doctoral degree	3
Master degree	2
Bachelor degree	8
Diploma	2
Vocational certificate	1
Perceived importance of search scenarios	64
Essential	22 (34.4%)
Very important	26 (40.6%)
Moderately important	14 (21.9%)
Somewhat important	1 (1.6%)
Not important	1 (1.6%)
Urgency of obtaining an answer	64
Extremely urgent (day)	8 (12.5%)
Very urgent (days)	23 (35.9%)
Urgent (week)	19 (29.7%)
Somewhat urgent (weeks)	10 (15.6%)
Not urgent	4 (6.3%)

4. Test Collection

An essential foundation for advancing information retrieval research in novel domains is the development of a specialized test collection. This collection is crucial for both the development and evaluation of innovative methodologies [57]. The creation of our test collection, named AgriQuery, involved three critical phases: 1) devising a set of inquiry topics; 2) compiling a pool of passages for evaluation; and 3) conducting detailed human relevance assessments.

4.1. Forming the Question Topics

The process of generating question topics for AgriQuery involved two agricultural science experts and utilized the known-item retrieval approach. The steps were as follows:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

1. A document was randomly selected from the collection. If the document did not lend itself to generating a relevant question, another document was chosen.
2. The assessor, upon reviewing the document, formulated a question that the document could answer, establishing the basis for a new topic.
3. The assessor then created three or more ad-hoc, keyword search queries that could lead to the question.
4. An original answer to the question was authored by the assessor in their own words.
5. Relevant sections of the document supporting the answer were identified and annotated.
6. Additional passages from the same document were evaluated by the assessor for their relevance, being classified as relevant, marginally relevant, or non-relevant.

A total of 420 topics were generated from 300 documents, indicating that multiple topics can often be derived from a single document. These topics were divided into a training set of 320 and a test set of 100 for subsequent experimentation.

4.2. Documents and Passages

The collection comprised over 86,846 documents, including 4,003 agricultural reports from GRDC and various State Departments of Agriculture in Australia, along with 82,843 scientific articles from 33 leading agricultural journals. These documents were pivotal in offering a comprehensive view of agronomic practices and soil management across diverse grain crops like cereals, legumes, and oilseeds. Topics spanned from varietal selection and planting strategies to comprehensive farming system performance and pest management.

4.3. Pooling and Relevance Assessment

Utilizing AgriQuery's extensive topic set, we engaged two state-of-the-art neural ranking models, monoBERT and TILDEv2, to generate initial runs. These runs were then merged using reciprocal rank fusion to establish a robust pool for human evaluation. Our relevance assessment protocol, supported by a custom-built tool named AgEvaluator, facilitated efficient and accurate evaluations by the agricultural scientists involved in the project.

Relevance assessments were meticulously conducted up to the top 20 ranks for test set topics or until a relevant passage was identified. For training set topics, the top 10 passages were evaluated irrespective of their initial relevance scoring.

4.4. Characteristics of the Test Collection

These detailed stages in the creation of AgriQuery's test collection underscore its potential to revolutionize the way agricultural information retrieval systems are developed and assessed, fostering greater innovation and precision in the field.

Table 3. Enhanced statistics of the AgriQuery test collection.

Topics	420
Train	320
Test	100
Judged Passages	7896
Non-relevant	2488 (32%)
Marginal relevant	1704 (22%)
Relevant	3704 (48%)
Documents	86,846
Reports	4,003
Journal articles	82,843
Passages	18,883,386

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

5. Passage Retrieval

Our research involved two primary experimental objectives: 1) evaluating the performance of various retrieval models within this specialized collection, and 2) analyzing the effects of query variations on retrieval effectiveness.

5.1. Retrieval Methods

We explored several retrieval strategies, including:

- **BM25:** As a baseline, we employed the BM25 algorithm to assess basic term-based retrieval efficacy.
- **BM25-RM3:** We enhanced the BM25 model with pseudo-relevance feedback using the RM3 method to refine the initial results.
- **monoBERT:** This cross-encoder neural model begins with a preliminary BM25 retrieval of the top 1000 documents, followed by a reranking process using a monoBERT model. This model was pre-trained on the MSMARCO dataset and subsequently fine-tuned on our collection of 160 training topics [58].
- **TILDEv2:** As a computationally efficient alternative, TILDEv2 performs document expansion at the indexing stage, which eliminates the need for on-the-fly neural encoding at query time [59]. Similar to monoBERT, it utilizes an initial BM25 retrieval followed by a TILDEv2 reranking, fine-tuned on our training topics.

The experiments leveraged both natural language questions and keyword-based queries to examine how different forms of query inputs influence the effectiveness of these models.

5.2. Results and Discussions

5.2.1. Comparative Effectiveness of Term-based and Neural Models

Significant disparities were observed in the performance of term-based and neural models. While monoBERT and TILDEv2 outperformed the BM25 model significantly in precision metrics such as nDCG@5 (t-test, $p < 0.01$), BM25 was comparable in terms of Success@100. This suggests that while BM25 effectively retrieves relevant passages, its ranking precision is lower, making it a suitable candidate for initial retrieval stages before applying a more precise reranker like monoBERT or TILDEv2.

5.2.2. Impact of Query Type on Retrieval Effectiveness

Contrary to some existing literature that suggests verbose queries reduce effectiveness [60], our findings indicate that natural language questions generally yield better results than keyword queries. This is particularly evident with neural models, which appear more adept at interpreting the nuances of natural language queries, enhancing early precision metrics such as nDCG@5 and reciprocal rank. Interestingly, while BM25's effectiveness was diminished with keyword queries under pseudo-relevance feedback (RM3), it showed improvement with natural language queries, suggesting a nuanced relationship between query type and retrieval strategy.

6. Our Proposed Method: AgriQuery

AgriQuery is designed as a versatile platform that accommodates both conversational and traditional search interactions for agricultural users. The system's architecture integrates robust conversation management capabilities with advanced retrieval technologies.

6.1. Client and User Interface

AgriQuery is accessible through Telegram, leveraging its widespread availability and straightforward API to facilitate ease of use across various devices. Users interact with AgriQuery by posing natural language questions, to which the system responds with dynamically retrieved information.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

6.2. Conversation Management with Macaw

Utilizing the Macaw framework, AgriQuery offers a comprehensive suite of conversational management tools, which include modules for intent identification, co-reference resolution, and query generation. The current implementation focuses on optimizing retrieval and result generation modules while preparing to integrate additional functionalities such as relevance feedback and query clarification for enhanced interaction dynamics [61].

6.3. Choices in Retrieval Model

Among the evaluated models, monoBERT emerged as the top performer, particularly effective in settings that prioritize top-ranked results (Success@3 was 0.96). Despite its effectiveness, the computational demands of monoBERT necessitate considerations regarding deployment in live environments, where response time is crucial.

TILDEv2, although slightly less effective than monoBERT, offers a more pragmatic approach for real-time application, balancing effectiveness with operational efficiency. This model's ability to run on standard CPU hardware without the need for specialized GPU support makes it an ideal choice for the operational demands of AgriQuery.

7. Conclusion

In this paper, we introduced AgriQuery, a sophisticated search system tailored to address the specific needs of farmers by extracting and delivering information from scientific documents. Despite the richness of data in agricultural scientific literature, these resources are often inaccessible to farmers who require specific information swiftly and conveniently. AgriQuery bridges this gap by crafting a system specifically designed for this unique application.

A profound understanding of farmers' information needs is pivotal for designing an effective search system. We conducted an extensive survey involving real search scenarios to capture the precise nature of these needs. This included preferences for information types such as crop protection and product recommendations, preferred sources (e.g., books, Google, product sheets), and the desired format of the answers (e.g., concise answers linked to detailed documents). Insights gleaned from this analysis were instrumental in shaping the requirements for our search system and constructing a robust test collection to evaluate its performance.

Our test collection includes 210 authentic questions and encompasses over 86,846 scientific documents, subdivided into approximately 9,441,693 passages. Manual relevance assessments conducted by agricultural experts ensure a solid foundation for training and evaluating machine learning models for information retrieval. This collection, now publicly available, supports the exploration of various query types and automated answer generation, thus serving as a valuable resource for ongoing research in agricultural search technologies.

Through rigorous testing, we demonstrated that neural ranking models, particularly TILDEv2 and monoBERT, significantly enhance the ability to pinpoint relevant information. However, deploying these models within a practical, user-friendly system presents challenges, primarily due to the computational demands of neural models. Our deployment architecture utilizes Telegram for user interaction and the Macaw platform for backend operations, providing a scalable and flexible solution. This setup highlights the trade-offs between efficiency and effectiveness, favoring the deployment of TILDEv2 due to its balance of performance and operational feasibility.

AgriQuery not only advances the state of agricultural search systems but also sets a stage for further explorations into more personalized and context-aware solutions, such as location or weather-sensitive ranking systems. Our ultimate goal is to catalyze further research in this field and translate academic advancements into practical, field-deployable tools.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

8. Future Work

Building on the foundation laid by the AgriQuery test collection, future research can expand in several directions:

Passages vs. Documents: Our collection's inclusion of both complete documents and discrete passages provides a unique opportunity to explore the relative effectiveness of passage versus document retrieval strategies. This could significantly impact the granularity of information retrieval practices in the field [62–64].

Query Variation: The multifaceted nature of the queries within our collection—ranging from natural language questions to structured keyword searches—offers a rich dataset to study the impacts of query formulation on retrieval effectiveness [65]. This exploration could further enhance understanding of how query variations influence search outcomes in specialized domains [66].

Answer Generation: Each topic within our collection is accompanied by a human-authored answer, providing a dual-layered approach to understanding information retrieval and synthesis. These answers, which may vary in vocabulary and conceptual alignment with the source passages, offer a fertile ground for developing sophisticated answer generation algorithms that could weave together retrieved passages into coherent, direct responses to user queries [67].

Scientific Document Extraction: The method used for extracting content from scientific PDFs into three-sentence passages can be optimized to consider structural elements like paragraphs, tables, and figures. Research in this area could explore advanced information extraction techniques to enhance the fidelity and utility of retrieved content for end-users [68].

Domain-Specific Expert Search: The agricultural domain presents unique challenges that may not be adequately addressed by general-domain search models. Our collection offers a pathway to evaluate and refine search strategies within this expert context, potentially leading to innovations that could tailor search algorithms more closely to the needs of domain-specific inquiries [69–72].

In conclusion, AgriQuery serves as both a culmination of our current research and a launching pad for future explorations aimed at refining and enhancing search technologies in agriculture and beyond.

References

1. Matthew J Smith. Getting value from artificial intelligence in agriculture. *Animal Production Science*, 2018.
2. Hao Fei, Shengqiong Wu, Meishan Zhang, Min Zhang, Tat-Seng Chua, and Shuicheng Yan. Enhancing video-language representations with structural spatio-temporal alignment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.
3. Jim Virgona, Geoff Daniel, et al. Evidence-based agriculture-can we get there? *Agricultural Science*, 23(1):19, 2011.
4. Manlio Bacco, Paolo Barsocchi, Erina Ferro, Alberto Gotta, and Massimiliano Ruggeri. The digitisation of agriculture: a survey of research activities on smart farming. *Array*, 3:100009, 2019.
5. Hao Fei, Yafeng Ren, and Donghong Ji. Retrofitting structure-aware transformer language model for end tasks. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pages 2151–2161, 2020.
6. Mohit Jain, Pratyush Kumar, Ishita Bhansali, Q. Vera Liao, Khai Truong, and Shwetak Patel. Farmchat: A conversational agent to answer farmer queries. *ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(4), 2018.
7. Kweku Opoku-Agyemang, Bhaumik Shah, and Tapan S. Parikh. Scaling up peer education with farmers in india. In *Information and Communication Technologies and Development*, ICTD '17, pages 15:1–15:10. ACM, 2017. ISBN 978-1-4503-5277-2.
8. Shengqiong Wu, Hao Fei, Fei Li, Meishan Zhang, Yijiang Liu, Chong Teng, and Donghong Ji. Mastering the explicit opinion-role interaction: Syntax-aided neural transition system for unified opinion role labeling. In *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence*, pages 11513–11521, 2022.
9. Wenxuan Shi, Fei Li, Jingye Li, Hao Fei, and Donghong Ji. Effective token graph modeling using a novel labeling strategy for structured sentiment analysis. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4232–4241, 2022.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

10. Hao Fei, Yue Zhang, Yafeng Ren, and Donghong Ji. Latent emotion memory for multi-label emotion classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 7692–7699, 2020.
11. Fengqi Wang, Fei Li, Hao Fei, Jingye Li, Shengqiong Wu, Fangfang Su, Wenxuan Shi, Donghong Ji, and Bo Cai. Entity-centered cross-document relation extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 9871–9881, 2022.
12. Ling Zhuang, Hao Fei, and Po Hu. Knowledge-enhanced event relation extraction via event ontology prompt. *Inf. Fusion*, 100:101919, 2023.
13. Rob Lokers, Rob Knapen, Sander Janssen, Yke van Randen, and Jacques Jansen. Analysis of big data technologies for use in agro-environmental science. *Environmental modelling & software*, 84:494–504, 2016.
14. Adams Wei Yu, David Dohan, Minh-Thang Luong, Rui Zhao, Kai Chen, Mohammad Norouzi, and Quoc V Le. Qanet: Combining local convolution with global self-attention for reading comprehension. *arXiv preprint arXiv:1804.09541*, 2018.
15. Shengqiong Wu, Hao Fei, Yixin Cao, Lidong Bing, and Tat-Seng Chua. Information screening whilst exploiting! multimodal relation extraction with feature denoising and multimodal topic modeling. *arXiv preprint arXiv:2305.11719*, 2023.
16. Jundong Xu, Hao Fei, Liangming Pan, Qian Liu, Mong-Li Lee, and Wynne Hsu. Faithful logical reasoning via symbolic chain-of-thought. *arXiv preprint arXiv:2405.18357*, 2024.
17. Matthew Dunn, Levent Sagun, Mike Higgins, V Ugur Guney, Volkan Cirik, and Kyunghyun Cho. SearchQA: A new Q&A dataset augmented with context from a search engine. *arXiv preprint arXiv:1704.05179*, 2017.
18. Hao Fei, Shengqiong Wu, Jingye Li, Bobo Li, Fei Li, Libo Qin, Meishan Zhang, Min Zhang, and Tat-Seng Chua. Lasuie: Unifying information extraction with latent adaptive structure-aware generative language model. In *Proceedings of the Advances in Neural Information Processing Systems, NeurIPS 2022*, pages 15460–15475, 2022.
19. Guang Qiu, Bing Liu, Jiajun Bu, and Chun Chen. Opinion word expansion and target extraction through double propagation. *Computational linguistics*, 37(1):9–27, 2011.
20. Hao Fei, Yafeng Ren, Yue Zhang, Donghong Ji, and Xiaohui Liang. Enriching contextualized language model from knowledge graph for biomedical information extraction. *Briefings in Bioinformatics*, 22(3), 2021.
21. Shengqiong Wu, Hao Fei, Wei Ji, and Tat-Seng Chua. Cross2StrA: Unpaired cross-lingual image captioning with cross-lingual cross-modal structure-pivoted alignment. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2593–2608, 2023.
22. Z. Sun, P.K. Sarma, W. Sethares, and E.P. Bucy. Multi-modal sentiment analysis using deep canonical correlation analysis. *Proc. Interspeech 2019*, pages 1323–1327, 2019.
23. Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. Squad: 100,000+ questions for machine comprehension of text. *arXiv preprint arXiv:1606.05250*, 2016.
24. Hao Fei, Bobo Li, Qian Liu, Lidong Bing, Fei Li, and Tat-Seng Chua. Reasoning implicit sentiment with chain-of-thought prompting. *arXiv preprint arXiv:2305.11255*, 2023.
25. Aniruddha Tammewar, Alessandra Cervone, and Giuseppe Riccardi. Emotion carrier recognition from personal narratives. *Accepted for publication at INTERSPEECH*, 2021. URL https://arxiv.org/abs/2008.07481.
26. Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1423. URL https://aclanthology.org/N19-1423.
27. Shengqiong Wu, Hao Fei, Leigang Qu, Wei Ji, and Tat-Seng Chua. Next-gpt: Any-to-any multimodal llm. *CoRR*, abs/2309.05519, 2023.
28. Qimai Li, Zhichao Han, and Xiao-Ming Wu. Deeper insights into graph convolutional networks for semi-supervised learning. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
29. Jane Mills, Matthew Reed, Kamilla Skaalsveen, and Julie Ingram. The use of twitter for knowledge exchange on sustainable soil management. *Soil Use and Management*, 35(1):195–203, 2019.
30. Sarah Van Dalsem. An iphone in a haystack: the uses and gratifications behind farmers using twitter. Master's thesis, University of Nebraska, 2011.
31. Hao Fei, Shengqiong Wu, Wei Ji, Hanwang Zhang, Meishan Zhang, Mong-Li Lee, and Wynne Hsu. Video-of-thought: Step-by-step video reasoning from perception to cognition. In *Proceedings of the International Conference on Machine Learning*, 2024.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

32. Naman Jain, Pranjali Jain, Pratik Kayal, Jayakrishna Sahit, Soham Pachpande, Jayesh Choudhari, et al. Agribot: agriculture-specific question answer system. *IndiaRxiv*, 2019.
33. Hao Fei, Shengqiong Wu, Wei Ji, Hanwang Zhang, and Tat-Seng Chua. Dysen-vdm: Empowering dynamics-aware text-to-video diffusion with llms. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7641–7653, 2024.
34. Mihir Momaya, Anjnya Khanna, Jessica Sadavarte, and Manoj Sankhe. Krushi–the farmer chatbot. In *2021 International Conference on Communication information and Computing Technology (ICCICT)*, pages 1–6. IEEE, 2021.
35. Hao Fei, Fei Li, Chenliang Li, Shengqiong Wu, Jingye Li, and Donghong Ji. Inheriting the wisdom of predecessors: A multiplex cascade framework for unified aspect-based sentiment analysis. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI*, pages 4096–4103, 2022.
36. Shengqiong Wu, Hao Fei, Yafeng Ren, Donghong Ji, and Jingye Li. Learn from syntax: Improving pair-wise aspect and opinion terms extraction with rich syntactic knowledge. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, pages 3957–3963, 2021.
37. Bobo Li, Hao Fei, Lizi Liao, Yu Zhao, Chong Teng, Tat-Seng Chua, Donghong Ji, and Fei Li. Revisiting disentanglement and fusion on modality and context in conversational multimodal emotion recognition. In *Proceedings of the 31st ACM International Conference on Multimedia, MM*, pages 5923–5934, 2023.
38. Hao Fei, Qian Liu, Meishan Zhang, Min Zhang, and Tat-Seng Chua. Scene graph as pivoting: Inference-time image-free unsupervised multimodal machine translation with visual scene hallucination. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5980–5994, 2023.
39. Hao Fei, Shengqiong Wu, Hanwang Zhang, Tat-Seng Chua, and Shuicheng Yan. Vitron: A unified pixel-level vision llm for understanding, generating, segmenting, editing. 2024.
40. Sanjeev Arora, Yingyu Liang, and Tengyu Ma. A simple but tough-to-beat baseline for sentence embeddings. In *ICLR*, 2017.
41. Abbott Chen and Chai Liu. Intelligent commerce facilitates education technology: The platform and chatbot for the taiwan agriculture service. *International Journal of e-Education, e-Business, e-Management and e-Learning*, 11:1–10, 01 2021.
42. Shengqiong Wu, Hao Fei, Xiangtai Li, Jiayi Ji, Hanwang Zhang, Tat-Seng Chua, and Shuicheng Yan. Towards semantic equivalence of tokenization in multimodal llm. *arXiv preprint arXiv:2406.05127*, 2024.
43. Jingye Li, Kang Xu, Fei Li, Hao Fei, Yafeng Ren, and Donghong Ji. MRN: A locally and globally mention-based reasoning network for document-level relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1359–1370, 2021.
44. Hao Fei, Shengqiong Wu, Yafeng Ren, and Meishan Zhang. Matching structure for dual learning. In *Proceedings of the International Conference on Machine Learning, ICML*, pages 6373–6391, 2022.
45. Hu Cao, Jingye Li, Fangfang Su, Fei Li, Hao Fei, Shengqiong Wu, Bobo Li, Liang Zhao, and Donghong Ji. OneEE: A one-stage framework for fast overlapping and nested event extraction. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 1953–1964, 2022.
46. Isakwisa Gaddy Tende, Kentaro Aburada, Hisaaki Yamaba, Tetsuro Katayama, and Naonobu Okazaki. Proposal for a crop protection information system for rural farmers in tanzania. *Agronomy*, 11(12):2411, 2021.
47. Hao Fei, Yafeng Ren, and Donghong Ji. Boundaries and edges rethinking: An end-to-end neural model for overlapping entity relation extraction. *Information Processing & Management*, 57(6):102311, 2020.
48. Jingye Li, Hao Fei, Jiang Liu, Shengqiong Wu, Meishan Zhang, Chong Teng, Donghong Ji, and Fei Li. Unified named entity recognition as word-word relation classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 10965–10973, 2022.
49. Mohit Jain, Pratyush Kumar, Ishita Bhansali, Q Vera Liao, Khai Truong, and Shwetak Patel. Farmchat: a conversational agent to answer farmer queries. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(4):1–22, 2018.
50. Shengqiong Wu, Hao Fei, Hanwang Zhang, and Tat-Seng Chua. Imagine that! abstract-to-intricate text-to-image synthesis with scene graph hallucination diffusion. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, pages 79240–79259, 2023.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

51. Hao Fei, Tat-Seng Chua, Chenliang Li, Donghong Ji, Meishan Zhang, and Yafeng Ren. On the robustness of aspect-based sentiment analysis: Rethinking model, data, and training. *ACM Transactions on Information Systems*, 41(2):50:1–50:32, 2023.
52. Yu Zhao, Hao Fei, Yixin Cao, Bobo Li, Meishan Zhang, Jianguo Wei, Min Zhang, and Tat-Seng Chua. Constructing holistic spatio-temporal scene graph for video semantic role labeling. In *Proceedings of the 31st ACM International Conference on Multimedia, MM*, pages 5281–5291, 2023.
53. Shengqiong Wu, Hao Fei, Yixin Cao, Lidong Bing, and Tat-Seng Chua. Information screening whilst exploiting! multimodal relation extraction with feature denoising and multimodal topic modeling. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14734–14751, 2023.
54. Hao Fei, Yafeng Ren, Yue Zhang, and Donghong Ji. Nonautoregressive encoder-decoder neural framework for end-to-end aspect-based sentiment triplet extraction. *IEEE Transactions on Neural Networks and Learning Systems*, 34(9):5544–5556, 2023.
55. Yu Zhao, Hao Fei, Wei Ji, Jianguo Wei, Meishan Zhang, Min Zhang, and Tat-Seng Chua. Generating visual spatial description via holistic 3D scene understanding. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7960–7977, 2023.
56. Nikulsinh M Chauhan et al. Information hungers of the rice growers. *Agriculture update*, 7(1/2):72–75, 2012.
57. Ellen M Voorhees, Donna K Harman, et al. *TREC: Experiment and evaluation in information retrieval*, volume 63. Citeseer, 2005.
58. Rodrigo Nogueira, Wei Yang, Kyunghyun Cho, and Jimmy Lin. Multi-stage document ranking with BERT. *arXiv preprint arXiv:1910.14424*, 2019.
59. Shengyao Zhuang and Guido Zuccon. Fast passage re-ranking with contextualized exact term matching and efficient passage expansion. *CoRR*, abs/2108.08513, 2021. URL https://arxiv.org/abs/2108.08513.
60. Michael Bendersky and W Bruce Croft. Discovering key concepts in verbose queries. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 491–498, 2008.
61. Hamed Zamani and Nick Craswell. Macaw: An extensible conversational information seeking platform. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2193–2196, 2020.
62. Marcin Kaszkiel and Justin Zobel. Passage retrieval revisited. In *ACM SIGIR Forum*, volume 31, pages 178–185. ACM New York, NY, USA, 1997.
63. Xiaoyong Liu and W Bruce Croft. Passage retrieval based on language models. In *Proceedings of the eleventh international conference on Information and knowledge management*, pages 375–382, 2002.
64. Nick Craswell, Bhaskar Mitra, Emine Yilmaz, and Daniel Campos. Overview of the trec 2020 deep learning track. *arXiv preprint arXiv:2102.07662*, 2021.
65. Alistair Moffat, Falk Scholer, Paul Thomas, and Peter Bailey. Pooled evaluation over query variations: Users are as diverse as systems. In *proceedings of the 24th ACM international on conference on information and knowledge management*, pages 1759–1762, 2015.
66. Peter Bailey, Alistair Moffat, Falk Scholer, and Paul Thomas. Uqv100: A test collection with query variability. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 725–728, 2016.
67. Chao-Chun Hsu, Eric Lind, Luca Soldaini, and Alessandro Moschitti. Answer generation for retrieval-based question answering systems. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 4276–4282, 2021.
68. Hannah Bast and Claudius Korzen. A benchmark and evaluation for text extraction from pdf. In *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, pages 1–10. IEEE, 2017.
69. Michail Salampasis, Norbert Fuhr, Allan Hanbury, Mihai Lupu, Birger Larsen, and Henrik Strindberg. Integrating ir technologies for professional search. In *European Conference on Information Retrieval*, pages 882–885. Springer, 2013.
70. John I Tait. An introduction to professional search. In *Professional search in the modern world*, pages 1–5. Springer, 2014.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2024 doi:10.20944/preprints202408.2157.v1

71. Suzan Verberne, Jiyin He, Udo Kruschwitz, Gineke Wiggers, Birger Larsen, Tony Russell-Rose, and Arjen P de Vries. First international workshop on professional search. In *ACM SIGIR Forum*, volume 52, pages 153–162. ACM New York, NY, USA, 2019.
72. Nandan Thakur, Nils Reimers, Andreas Rücklé, Abhishek Srivastava, and Iryna Gurevych. Beir: A heterogenous benchmark for zero-shot evaluation of information retrieval models. *arXiv preprint arXiv:2104.08663*, 2021.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.