Supporting Neuroscience Literature Exploration by Utilising Indirect Relations between Topics in Augmented Reality

Boyu Xu
b.xu@uu.nl
Utrecht University
Utrecht, Utrecht, The Netherlands

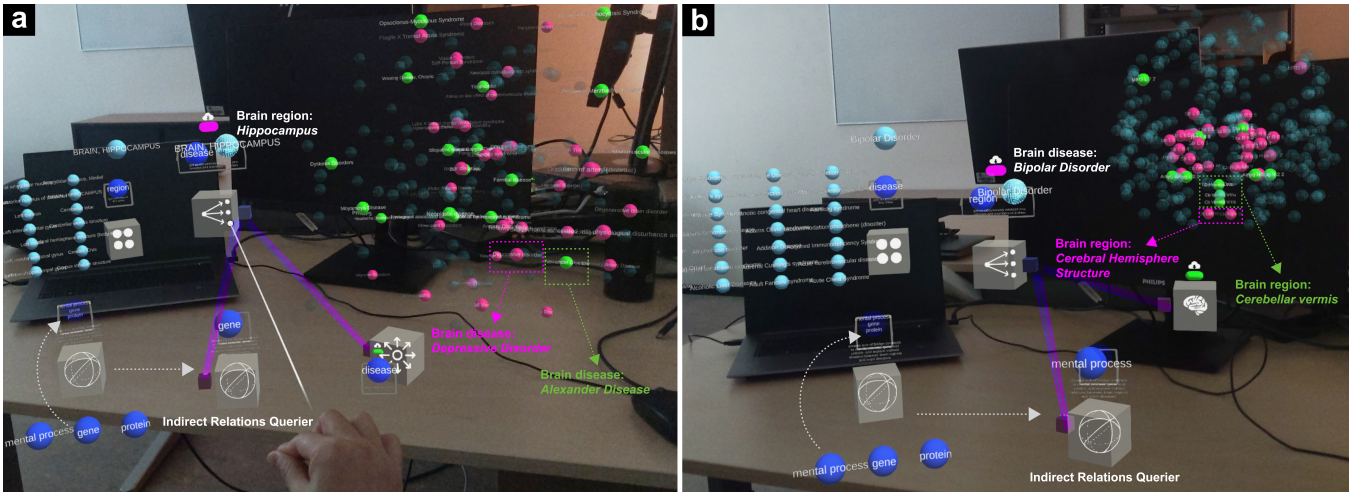

Figure 1: (a) The brain region *Hippocampus* (upper middle) has a direct relation (pink sphere) with the brain disease *Depressive Disorder* (pink, lower right). Genes (lower left) are selected as the intermediate topic to find an indirect relation (green sphere) between the brain region *Hippocampus* and the brain disease *Alexander Disease* (green, lower right). (b) The brain disease *Bipolar Disorder* (upper middle) has a direct relation (pink sphere) with the brain region *Cerebral Hemisphere Structure* (pink, lower right). Mental processes (lower left) are selected as the intermediate topic to find an indirect relation (green sphere) between the brain disease *Bipolar Disorder* and the brain region *Cerebellar vermis* (green, lower right). Topics with no direct or indirect relations are visualised as light blue spheres.

ABSTRACT

Neuroscientists need to analyse a large number of publications to identify potentially fruitful experiments. This task is necessary before undertaking any costly practical experiments. Exploring direct relations between topics (rather than publications), such as brain regions and brain diseases, has been shown to help neuroscientists identify fruitful experiments. In previous studies, users were able to query and visualise direct relations between topics using DatAR, an Augmented Reality prototype. Neuroscientist participants suggested that identifying previously unknown, or indirect, relations between topics could provide additional information for identifying fruitful experiments. I follow a user-centred design approach: defining functional requirements for finding indirect relations, designing interactive AR visualisations for the specified functionalities, and engaging neuroscientists in evaluating the usefulness of finding indirect relations. Neuroscientists who participated in my initial study of finding indirect relations, pointed out the potential of current indirect relations by demonstrating how indirect relations in the past may have evolved into present direct relations. This suggestion informs Study 2 on exploring publication-date dependent direct and indirect relations. Participating neuroscientists also suggested providing specific intermediate topics, such as genes, when indicating indirect relations between topics. This proposal informs Study 3 on identifying specific intermediate topics and publications indicating indirect relations. My final study will assess the usefulness of the designed DatAR in neuroscientists' daily research work for identifying potentially fruitful experiments.

CHIIR '24, March 10–14, 2024, Sheffield, United Kingdom

ACM ISBN 979-8-4007-0434-5/24/03
https://doi.org/10.1145/3627508.3638312

CCS CONCEPTS

• **Information systems** → **Information retrieval**; • **Human-centered computing** → **Visualization systems and tools**; *Human computer interaction (HCI).*

KEYWORDS

Topic-based Literature Exploration, Data Visualisation, Augmented Reality

ACM Reference Format:
Boyu Xu. 2024. Supporting Neuroscience Literature Exploration by Utilising Indirect Relations between Topics in Augmented Reality. In *Proceedings of the 2024 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '24), March 10–14, 2024, Sheffield, United Kingdom.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3627508.3638312

1 MOTIVATION AND INTRODUCTION

To identify potentially fruitful experiments, neuroscientists need to analyse a large number of publications - a process that is both time-consuming and challenging. Simultaneously exploring many publications can be accomplished using topic-based literature exploration [8]. Based on insights from neuroscientists [6, 7], exploring direct relations between topics, such as brain regions and brain diseases, in neuroscience using Augmented Reality (AR) is useful for identifying fruitful experiments. The DatAR team at Utrecht University has developed a 3D-based AR prototype that provides a visual representation of direct relations between brain regions and brain diseases, Figure 1. Neuroscientists also proposed that finding indirect relations between brain regions and brain diseases could provide additional information for identifying fruitful experiments. I follow a user-centred design cycle of identifying a relevant user task, designing corresponding functionalities and appropriate interactions and visualisations.

I anticipate a total of 4 studies, Table 1. My completed initial study showed that finding indirect relations between topics, specifically brain regions and brain diseases, via intermediate topics such as genes and mental processes, served as inspiration for further literature review but did not provide sufficient evidence for identifying fruitful experiments, Section 3. Participants proposed exploring trends in indirect relations and when these indirect relations became established as direct relations in the literature, which could improve the usefulness of indirect relations found. I investigate this in my second, current, study, Section 4. Participants in my initial study also suggested providing more information about the intermediate topic that leads to the indirect relation found, such as the specific gene that has a direct relation with the brain region and with the brain disease. This suggestion inspired the proposal for my third study on identifying specific intermediate topics and publications indicating indirect relations, Section 5. Moreover, DatAR is an interactive platform for AR-based 3D data. During discussions with three data discovery experts and three AR/VR developers [6], DatAR could evolve into a useful tool for neuroscientists to integrate topic-based neuroscience literature exploration into the daily research workflow. My fourth study will explore the usefulness of the functionalities implemented in DatAR for neuroscientists' daily research work of exploring literature, Section 6.

2 RELATED WORK

I describe the direct relations implemented in DatAR (Section 2.1), which is used as a start for my research. I illustrate examples of finding indirect relations among topics in biomedical tasks, which identify previously unknown relations between topics (Section 2.2). I discuss the AR-based immersive 3D visualisations as inspiration for providing visualisation support in neuroscience literature exploration (Section 2.3).

2.1 Direct Relations Implemented in DatAR

A co-occurrence of two topics, for instance, a brain region and a brain disease, is defined when they appear within the same sentence of a publication's title or abstract. This co-occurrence implies a direct relation between the brain region and the brain disease. DatAR, an AR prototype, is designed to facilitate brain-related topic exploration by neuroscientists through interactive visualisations. The DatAR repository provides an analysis of titles and abstracts from a collection of 414,224[2] neuroscience publications from 2010-2022 available via PubMed. Topics in the DatAR repository include brain regions, brain diseases, genes, proteins and mental processes. Previous studies [6, 7] have explored direct relations between topics and visualised them in DatAR, Figure 1. Direct relations between topics are visualised as pink spheres.

2.2 Examples of Indirect Relations and Why They Are Potentially Useful

The potential of indirect relations is beneficial in various biomedical tasks, such as in drug discovery, analysing relations between disease-gene (or protein) and gene (or protein)-drug helped identify potential therapeutic agents for *Parkinson's Disease* [3]. The ABC co-occurrence model illustrates an approach for exploring unknown or indirect relations [2]. This model proposes that if A co-occurs with B and B co-occurs with C, it defines the indirect relations between A and C. We apply and extend the ABC co-occurrence model to neuroscience literature, offering insights into implicit causal relations and mechanisms of interaction between brain regions and brain diseases for identifying potentially fruitful experiments.

2.3 3D Visualisations in Augmented Reality

3D visualisations in AR provide an advantage over 2D screens in exploring neuroscience literature by providing spatial representation for understanding complex relations among brain regions [4]. DatAR's 3D visualisation of the relative locations of brain regions enables researchers to easily understand their relations with brain diseases, Figure 1 (b). Integrating 3D visualisations into the real-world enables researchers to incorporate their literature exploration process into their daily research work [1], specifically, they can read publications on their 2D screens while simultaneously visualising relations between topics in the 3D environment.

3 FINDING INDIRECT RELATIONS BETWEEN TOPICS (INITIAL STUDY, IN SUBMISSION)

A neuroscientist, Cunqing Huangfu, proposed in previous communications that exploring unknown indirect relations among topics is useful for identifying fruitful experiments [7]. In April 2022, interviews were conducted with another three neuroscientists to gauge

[1]Cunqing Huangfu, a neuroscientist at the Institute of Automation of the Chinese Academy of Sciences, the Brain-inspired Cognitive Intelligence Lab.

[2]Neuroscience publications in PubMed https://pubmed.ncbi.nlm.nih.gov/ as of February 3, 2022.

Table 1: Origins of the user tasks used to design each study.

Study	Reason for Selecting User Task
Initial Study 1: Finding Indirect Relations Between Topics (Section 3) *In submission*	Initial interviews with Cunqing Huangfu[1] in 2019 [7], and interviews with three other neuroscientists in April 2022 to explain finding indirect relations
Study 2: Finding Publication-date Dependent Direct and Indirect Relations (Section 4) *In progress*	Interview with a neuroscientist during the evaluation session of Study 1, finding indirect relations, in October 2022
Study 3: Identifying Specific Intermediate Topics and Publications Indicating Indirect Relations (Section 5) *Future work*	Interview with a neuroscientist during the evaluation session of Study 1, finding indirect relations, in November 2022
Study 4: Exploring the Usefulness of DatAR in Daily Neuroscience Literature Exploration (Section 6) *Future work*	Initial interviews with three data discovery experts and three AR/VR developers in 2019 [6, 7]

their interest in this task. While there may be no direct relations between a brain region and a brain disease in the literature, each may have a direct relation with another topic, such as a gene. We call this an indirect relation. Given the topics already in the DatAR repository (Section 2.1), I worked with these three neuroscientists to identify meaningful intermediate topics. These include mental processes, genes, and proteins.

Research Question 1 (RQ1): *What functionalities are needed to find indirect relations in DatAR?* The *Indirect Relations Querier* was designed to support the querying process of selecting intermediate topics to find indirect relations between topics, Figure 1.

RQ2: *To what extent is the visualisation of indirect relations suitable for the functionality identified in RQ1?* I conducted further interviews with the same three neuroscientists to identify visualisation requirements for displaying indirect as well as direct relations. Based on these, I incorporated the visualisation of the indirect relations in the same visualisation as the direct relations. Direct relations between topics are visualised as pink spheres; indirect relations between topics are visualised as green spheres[3].

RQ3: *To what extent are the functionality and visualisations implemented in DatAR useful for finding indirect relations?* I designed corresponding functionality and a suitable visualisation in DatAR to support the task of finding indirect relations, Figure 1. Nine neuroscientists were invited in October and November 2022 to evaluate the utility of the implemented functionality and the suitability of the visualisation for finding indirect relations. All participating neuroscientists indicated the utility of the implemented functionality and the suitability of visualisations using AR in finding indirect relations.

RQ4: *To what extent are the indirect relations found useful for identifying fruitful experiments?* All participating neuroscientists agreed that finding indirect relations between topics is a useful task for identifying potential fruitful experiments. Suggestions for future work centred on providing the specific intermediate topic, such as the specific gene, for indicating indirect relations and improving the usefulness of the identified indirect relations.

[3]The colours used for visualising direct and indirect relations are not (yet) a research question.

4 FINDING PUBLICATION-DATE DEPENDENT DIRECT AND INDIRECT RELATIONS (STUDY 2, IN PROGRESS)

During the evaluation session on finding indirect relations (Section 3), six participating neuroscientists proposed that the indirect relations found merely served as inspiration for further literature review and were insufficient evidence for identifying fruitful experiments. To illustrate that the indirect relations found are useful for identifying fruitful experiments, a neuroscientist suggested the use of publication dates to identify trends of (first indirect and then direct) relations between topics. By illustrating that previous indirect relations might have evolved into present direct relations, the usefulness of current indirect relations will be improved. This usefulness of indirect relations assists neuroscientists in identifying potentially fruitful experiments.

RQ5: *Which publication-date dependent representative tasks are useful for neuroscientists to identify fruitful experiments?* To answer RQ5, an interview was conducted with a neuroscience PhD student in July 2023. S/he wanted to understand the brain regions that exhibited direct relations with the brain disease *Epilepsy* over the past five years. These regions were perceived as emerging topics with the potential to shape their future research. S/he was also interested in understanding the process of past indirect relations becoming current direct relations.

RQ6: *What functionalities and visualisations are needed to explore the publication-date dependent direct and indirect relations in the context of representative tasks in DatAR?*

RQ7: *To what extent are the functionality and visualisations provided by the interactive prototype useful for exploring publication-date dependent direct and indirect relations?*

RQ8: *To what extent are the found publication-date dependent direct and indirect relations useful for identifying fruitful experiments?*

New functionalities and visualisations will be implemented in DatAR. To answer RQ6, I will collaborate with neuroscientists to identify the functional and visual requirements. To answer RQ7 and RQ8, I will invite 10 neuroscientists[4] to explore the representative tasks in the context of implementing functionalities and visualisations. I will use a qualitative measure of a semi-structured interview to research RQ7 and RQ8. The qualitative evaluation will focus on three key dimensions: "Usability", "Suitability", and "Usefulness". Usability encompasses the extent to which the implemented functionalities align with human-computer interaction requirements.

[4]Five to ten participants can increase the expected level of problem discovery [5].

Suitability focuses on evaluating the designed visualisations that synergize with the implemented functionalities in the AR environment. Usefulness is to gauge the success of the representative tasks for aiding neuroscientists in exploring publication-date dependent direct and indirect relations for identifying fruitful experiments.

5 IDENTIFYING SPECIFIC INTERMEDIATE TOPICS AND PUBLICATIONS INDICATING INDIRECT RELATIONS (STUDY 3, FUTURE WORK)

In the initial study of finding indirect relations (Section 3), neuroscientists selected an intermediate topic, such as genes or mental processes, to investigate indirect relations between brain regions and brain diseases. During the evaluation study of finding indirect relations conducted in November 2022 (Section 3), a neuroscientist wanted extra information to improve the usefulness of the identified indirect relations. This extra information includes specifying the specific intermediate topic, such as the specific gene, responsible for the indirect relation, and providing direct access to the publications that serve as sources for the two direct relations. Responses from six other neuroscientists also indicated a desire to manually select a specific intermediate topic pertinent to their research to find indirect relations.

RQ9: *What functionalities and visualisations are needed for presenting specific intermediate topics and publications within the AR environment?*

RQ10: *To what extent do the provided specific intermediate topics and their corresponding publications contribute to the identification of potentially fruitful experiments?*

For this study, I will extend the capabilities of the *Indirect Relations Querier* to include presenting specific intermediate topics and publications indicating indirect relations. A key aspect of this study is designing suitable visualisations for presenting specific intermediate topics and publications within the AR environment. To answer RQ9 and RQ10, I will also follow a user-centred design cycle, involving close collaboration with neuroscientists to identify functional and visual requirements. The evaluation approach adopted in the study of publication-date dependent direct and indirect relations will be applied, leveraging semi-structured interviews to obtain insights into functionalities, visualisations and usefulness.

6 EXPLORING THE USEFULNESS OF DATAR IN DAILY NEUROSCIENCE LITERATURE EXPLORATION (STUDY 4, FUTURE WORK)

In initial discussions held in 2019, data discovery experts and AR/VR developers expressed their shared aspiration for DatAR to evolve into a tool for delving into neuroscience literature within the framework of daily research work [7].

RQ11: *Which representative tasks in the daily neuroscience literature exploration are useful for neuroscientists to identify fruitful experiments?*

RQ12: *To what extent is the usefulness of DatAR for neuroscientists in their daily literature exploration?*

To answer RQ11, I will collaborate with neuroscientists to identify appropriate representative tasks in daily literature exploration for identifying fruitful experiments. For RQ12, I intend to recruit 10 neuroscientists who will undertake these identified representative tasks within a specified timeframe in daily literature exploration, which may take one or two months. Through systematic observation and data collection while using DatAR during their daily literature exploration, I will assess the prototype's usefulness it offers for neuroscientists to identify fruitful experiments.

7 CONCLUSION

I view this work as contributing to a novel AR-based approach to utilising indirect relations between topics, aimed at supporting neuroscientists in identifying potentially fruitful experiments. I have demonstrated the potential of finding indirect relations through DatAR to identify fruitful experiments, as confirmed by feedback from neuroscientists in my initial study. I am about to begin a procedure (Section 4) to improve the usefulness of identified indirect relations. Additionally, I plan to continue improving the prototype's usefulness, envisioning a future where DatAR becomes an indispensable asset in daily neuroscience literature exploration.

ACKNOWLEDGMENTS

I would like to thank my supervisors, Lynda Hardman (Centrum Wiskunde & Informatica) and Wolfgang Hürst (Utrecht University) for their support. I would like to thank Zhisheng Huang at Vrije Universiteit, Amsterdam, for creating the repository of co-occurrences between topics in Knowledge Graphs of Brain Science and Triply[5] for hosting these. I am grateful to my colleagues at UU and CWI. This work is supported by a scholarship from the China Scholarship Council (CSC): No. 202108440064.

REFERENCES

[1] Lonni Besançon, Anders Ynnerman, Daniel F Keefe, Lingyun Yu, and Tobias Isenberg. 2021. The state of the art of spatial interfaces for 3D visualization. In *Computer Graphics Forum*, Vol. 40. Wiley Online Library, 293–326. https://doi.org/10.1111/cgf.14189

[2] Vishrawas Gopalakrishnan, Kishlay Jha, Wei Jin, and Aidong Zhang. 2019. A survey on literature based discovery approaches in biomedical domain. *Journal of biomedical informatics* 93 (2019), 103141. https://doi.org/10.1016/j.jbi.2019.103141

[3] Dimitar Hristovski, Andrej Kastrin, Borut Peterlin, and Thomas C Rindflesch. 2010. Combining semantic relations and DNA microarray data for novel hypotheses generation. In *Linking Literature, Information, and Knowledge for Biology: Workshop of the BioLink Special Interest Group, ISMB/ECCB 2009, Stockholm, June 28-29, 2009, Revised Selected Papers*. Springer, 53–61. https://doi.org/10.1007/978-3-642-13131-8_7

[4] Pranav Parekh, Shireen Patel, Nivedita Patel, and Manan Shah. 2020. Systematic review and meta-analysis of augmented reality in medicine, retail, and games. *Visual computing for industry, biomedicine, and art* 3 (2020), 1–20. https://doi.org/10.1186/s42492-020-00057-7

[5] Janet M Six and Ritch Macefield. 2016. How to determine the right number of participants for usability studies. *San Francisco (CA): UXmatters* (2016). https://www.uxmatters.com/mt/archives/2016/01/how-to-determine-the-right-number-of-participants-for-usability-studies.php

[6] Ghazaleh Tanhaei. 2022. Exploring Neuroscience Literature and Understanding Relations Between Brain-Related Topics-Using Augmented Reality. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*. 387–390. https://doi.org/10.1145/3498366.3505806

[7] Ivar Troost, Ghazaleh Tanhaei, Lynda Hardman, and Wolfgang Hürst. 2021. Exploring relations in neuroscientific literature using augmented reality: A design study. In *Designing Interactive Systems Conference 2021*. 266–274. https://doi.org/10.1145/3461778.3462053

[8] Yi Zhang, Hongshu Chen, Jie Lu, and Guangquan Zhang. 2017. Detecting and predicting the topic change of Knowledge-based Systems: A topic-based bibliometric analysis from 1991 to 2016. *Knowledge-Based Systems* 133 (2017), 255–268. https://doi.org/10.1016/j.knosys.2017.07.011

[5] Triply https://triply.cc/