Exploring Indirect Relations Between Topics in Neuroscience Literature Using Augmented Reality to Inform Experimental Design

Boyu Xu

Exploring Indirect Relations Between Topics in Neuroscience Literature Using Augmented Reality to Inform Experimental Design

Boyu Xu

SIKS Dissertation Series No. 2026-04
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

ISBN:	978-90-393-7998-1
Printed by:	Ipskamp Printing
Cover by:	Boyu Xu and Midjourney
Digital version:	available at `https://doi.org/10.33540/3256`.

This dissertation was printed in limited volume on 100% recycled paper in an effort to minimize the enviromental footprint.

Exploring Indirect Relations Between Topics in Neuroscience Literature Using Augmented Reality to Inform Experimental Design

**Indirecte relaties tussen onderwerpen
in neurowetenschappelijke literatuur
verkennen met Augmented Reality
ter ondersteuning van experimenteel ontwerp**
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. ir. W. Hazeleger,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op

woensdag 14 januari 2026 des middags te 2.15 uur

door

Boyu Xu

geboren op 28 mei 1994
te Changzhou, Jiangsu, China

Promotor:
Prof. dr. Lynda Hardman

Copromotor:
Dr. Wolfgang Hürst

Beoordelingscommissie:
Dr. Hyowon Lee
Prof. dr. Alexandru C. Telea
Prof. dr. Remco C. Veltkamp
Prof. dr. Anja Volk
Prof. dr. Mascha van der Voort

Dit proefschrift werd mogelijk gemaakt met financiële steun van een beurs van de China Scholarship Council (CSC), nummer 202108440064

Contents

Summary

Before conducting a costly experiment, neuroscientists need to identify potentially useful hypotheses. Given the vast amount of neuroscience literature, it is useful to gain a bird's-eye view of the field to understand what information is established and what may inform future experiments. To achieve this, topic-based literature exploration in Augmented Reality (AR) has been used to investigate relations between neuroscience topics, such as brain regions and brain diseases. 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. A co-occurrence of two topics, such as a brain region and a brain disease, in the same sentence of the title or abstract of a publication implies a ***direct*** relation between them. These direct relations can help neuroscientists intuitively understand which brain regions are affected by specific brain diseases. A brain region may also be connected indirectly to a brain disease through an intermediate topic, such as a gene or a mental process. We define an ***indirect*** relation between two topics when there is no direct relation between them, but each co-occurs with at least one other intermediate topic. Neuroscientists have proposed that identifying such indirect relations could provide additional insights for experimental design. These indirect relations may reflect weak evidence of a potential link between two topics that has not yet been confirmed by any single publication.

Main Research Question 1. *How does* ***exploring indirect relations between topics*** *contribute to designing a useful experiment?*

To address MRQ1, we investigated whether exploring indirect relations between topics could assist neuroscientists in informing the design of a potentially useful experiment (Study 1). We interviewed three neuroscientists to determine whether exploring indirect relations between topics is indeed a potentially useful user task. They helped us identify one functional requirement and two visualisation requirements, which we addressed and implemented in the *Indirect Relations Explorer*. Users can select an intermediate topic type, such as genes, to explore indirect relations and visualise both direct and indirect relations in a single visualisation. This helps neuroscientists distinguish which topic relations are already established in the literature and which may indicate potential unknown connections. All nine participating neuroscientists' responses indicated that indirect relations could inform the design of a neuroscience experiment.

- Six of the participating neuroscientists noted that the identified indirect relations primarily serve as inspiration for further literature review, rather than as definitive evidence for designing an experiment. They suggested that tracking trends in indirect relations over time and identifying when these indirect relations be-

come established as direct relations in the literature could provide evidence for the usefulness of current indirect relations. This insight motivated the **Main Research Question 2**.

- One neuroscientist also suggested exploring the specific intermediate topic, such as a gene or mental process, responsible for the indirect relation. Identifying the specific intermediate topic could help neuroscientists understand the mechanisms underlying the indirect relation and better assess its usefulness. This suggestion motivated the **Main Research Question 3**.

Main Research Question 2. *How does **exploring the evolution of historical indirect to current direct relations** contribute to designing a useful experiment?*

To address MRQ2, we investigated whether exploring the evolution of indirect to direct relations could assist neuroscientists in understanding the usefulness of current indirect relations and, in turn, inform the design of a potentially useful experiment (Study 2). We first consulted two neurologists to determine whether exploring the evolution of indirect to direct relations is indeed a potentially useful user task. They helped us specify the functionality and visualisation requirements, which we addressed and implemented in the *Timeline Selector*. Users can select a range of publication dates and observe, in a single visualisation, when a historical indirect relation becomes a current direct relation. All twelve participating neuroscientists' responses indicated that observing historical indirect relations evolve into current direct relations increased their confidence in the usefulness of current indirect relations.

Main Research Question 3. *How does **exploring indirect relations via a specific intermediate topic** contribute to designing a useful experiment?*

To address MRQ3, we investigated whether exploring indirect relations via a specific intermediate topic could assist neuroscientists in understanding the usefulness of current indirect relations and, in turn, provide additional information for designing an experiment (Study 3). We first consulted two neuroscientists to determine whether exploring indirect relations via a specific intermediate topic is indeed a potentially useful user task. They helped us specify two functional requirements and one visualisation requirement, which we addressed and implemented in the *Specific Intermediate Topic Explorer*. All nine participating neuroscientists' responses indicated that exploring specific intermediate topics helps assess the usefulness of identified indirect relations.

Our research provides an interactive 3D AR approach to assist neuroscientists in identifying indirect relations between neuroscience topics. It also offers additional information, such as the evolution of indirect to direct relations and the specific intermediate topics involved, helping researchers assess the usefulness of the indirect relations. By enabling the exploration of hundreds of thousands of publications simultaneously and supporting the identification of useful hypotheses, our prototype serves as a complementary step in the early phase of designing a potentially useful experiment. Through interactive visualisation, it allows researchers to intuitively discover less observed or previously unnoticed relations that are difficult to identify through traditional literature exploration.

Samenvatting

Voordat neurowetenschappers een kostbaar experiment uitvoeren, moeten zij potentiële en nuttige hypothesen identificeren. Gezien de enorme hoeveelheid neurowetenschappelijke literatuur is het nuttig om een helikopterblik op het vakgebied te krijgen om te begrijpen welke informatie al is vastgesteld en wat toekomstige experimenten kan informeren. Om dit te bereiken, is literatuurverkenning op basis van onderwerpen in Augmented Reality (AR) gebruikt om relaties tussen neurowetenschappelijke onderwerpen, zoals hersengebieden en hersenaandoeningen, te onderzoeken. Het DatAR-team van de Universiteit Utrecht heeft een 3D-gebaseerd AR-prototype ontwikkeld dat een visuele weergave biedt van directe relaties tussen hersengebieden en hersenaandoeningen. Het samen voorkomen van twee onderwerpen, zoals een hersengebied en een hersenaandoening, in dezelfde zin van de titel of samenvatting van een publicatie, impliceert een ***directe*** relatie tussen deze twee onderwerpen. Deze directe relaties kunnen neurowetenschappers helpen intuïtief te begrijpen welke hersengebieden worden beïnvloed door specifieke hersenaandoeningen. Een hersengebied kan ook indirect verbonden zijn met een hersenaandoening via een intermediair onderwerp, zoals een gen of een mentaal proces. Wij definiëren een ***indirecte*** relatie tussen twee onderwerpen als een situatie waarin er geen directe relatie tussen hen bestaat, maar waarbij elk onderwerp samen voorkomt met ten minste één ander intermediair onderwerp. Neurowetenschappers hebben voorgesteld dat het identificeren van dergelijke indirecte relaties extra inzichten kan bieden voor experimenteel ontwerp. Deze indirecte relaties kunnen zwak bewijs vormen van een potentiële koppeling tussen twee onderwerpen die nog niet door een enkele publicatie is bevestigd.

Hoofdonderzoeksvraag 1. *Hoe draagt **het verkennen van indirecte relaties tussen onderwerpen** bij aan het ontwerpen van een nuttig experiment?*

Om hoofdonderzoeksvraag 1 te beantwoorden, hebben we onderzocht of het verkennen van indirecte relaties tussen onderwerpen neurowetenschappers kan helpen bij het informeren van het ontwerp van een potentieel nuttig experiment (Studie 1). We hebben drie neurowetenschappers geïnterviewd om te bepalen of het verkennen van indirecte relaties tussen onderwerpen inderdaad een potentieel nuttige gebruikertaak is. Zij hielpen ons één functionele vereiste en twee visualisatievereisten te identificeren, die we hebben uitgewerkt en geïmplementeerd in de *Indirect Relations Explorer*. Gebruikers kunnen een type intermediair onderwerp selecteren, zoals genen, om indirecte relaties te verkennen en zowel directe als indirecte relaties in één visualisatie weer te geven. Dit helpt neurowetenschappers te onderscheiden welke relaties al in de literatuur zijn vastgesteld en welke kunnen wijzen op mogelijke onbekende verbanden. Alle negen deelnemende neurowetenschappers gaven aan dat indirecte relaties kunnen bijdragen aan het ontwerp van een neurowetenschappelijk experiment.

- Zes van de deelnemende neurowetenschappers merkten op dat de geïdentificeerde indirecte relaties vooral dienen als inspiratie voor verder literatuuronderzoek, eerder dan als definitief bewijs voor het ontwerpen van een experiment. Zij suggereerden dat het volgen van trends in indirecte relaties in de loop van de tijd, en het identificeren van het moment waarop deze indirecte relaties zich in de literatuur ontwikkelen tot directe relaties, bewijs zou kunnen leveren voor de bruikbaarheid van de huidige indirecte relaties. Deze bevinding leidde tot **Hoofdonderzoeksvraag 2**.

- Een van de neurowetenschappers stelde ook voor om het specifieke intermediaire onderwerp, zoals een gen of een mentaal proces, dat verantwoordelijk is voor de indirecte relatie, te onderzoeken. Het identificeren van dit specifieke intermediaire onderwerp kan neurowetenschappers helpen de mechanismen die ten grondslag liggen aan de indirecte relatie beter te begrijpen en de bruikbaarheid ervan beter te beoordelen. Deze suggestie leidde tot **Hoofdonderzoeksvraag 3**.

Hoofdonderzoeksvraag 2. *Hoe draagt* ***het verkennen van de evolutie van historische indirecte naar huidige directe relaties*** *bij aan het ontwerpen van een nuttig experiment?*

Om hoofdonderzoeksvraag 2 te beantwoorden, hebben we onderzocht of het verkennen van de evolutie van indirecte naar directe relaties neurowetenschappers kan helpen de bruikbaarheid van huidige indirecte relaties te begrijpen en daardoor het ontwerp van een potentieel nuttig experiment te informeren (Studie 2). We hebben eerst twee neurologen geraadpleegd om te bepalen of het verkennen van de evolutie van indirecte naar directe relaties inderdaad een potentieel nuttige gebruikertaak is. Zij hielpen ons de functionele en visualisatievereisten te specificeren, die we hebben uitgewerkt en geïmplementeerd in de *Timeline Selector*. Gebruikers kunnen een reeks publicatiedata selecteren en in één visualisatie observeren wanneer een historische indirecte relatie een huidige directe relatie wordt. Alle twaalf deelnemende neurowetenschappers gaven aan dat het waarnemen van historische indirecte relaties die evolueren tot huidige directe relaties hun vertrouwen in de bruikbaarheid van huidige indirecte relaties vergrootte.

Hoofdonderzoeksvraag 3. *Hoe draagt* ***het verkennen van indirecte relaties via een specifiek intermediair onderwerp*** *bij aan het ontwerpen van een nuttig experiment?*

Om hoofdonderzoeksvraag 3 te beantwoorden, hebben we onderzocht of het verkennen van indirecte relaties via een specifiek intermediair onderwerp neurowetenschappers kan helpen de bruikbaarheid van huidige indirecte relaties te begrijpen en daardoor extra informatie kan bieden voor het ontwerpen van een experiment (Studie 3). We hebben eerst twee neurowetenschappers geraadpleegd om te bepalen of het verkennen van indirecte relaties via een specifiek intermediair onderwerp inderdaad een potentieel nuttige gebruikertaak is. Zij hielpen ons twee functionele vereisten en één visualisatievereiste te specificeren, die we hebben uitgewerkt en geïmplementeerd in de *Specific Intermediate Topic Explorer*. Alle negen deelnemende neurowetenschappers gaven aan dat het verkennen van specifieke intermediaire onderwerpen helpt bij het beoordelen van de bruikbaarheid van de geïdentificeerde indirecte relaties.

Ons onderzoek biedt een interactieve 3D-AR-benadering om neurowetenschappers te helpen bij het identificeren van indirecte relaties tussen neurowetenschappelijke onderwerpen. Het biedt ook aanvullende informatie, zoals de evolutie van indirecte naar directe relaties en de specifieke intermediaire onderwerpen die hierbij betrokken zijn, waardoor onderzoekers de bruikbaarheid van deze indirecte relaties beter kunnen beoordelen. Door het mogelijk te maken om tegelijkertijd honderdduizenden publicaties te verkennen en het identificeren van nuttige hypothesen te ondersteunen, fungeert ons prototype als een aanvullende stap in de vroege fase van het ontwerpen van een potentieel nuttig experiment. Via interactieve visualisatie kunnen onderzoekers intuïtief minder waargenomen of eerder onopgemerkte relaties ontdekken die moeilijk te identificeren zijn via traditionele literatuurverkenning.

1

Introduction

Scientific knowledge advances through the accumulation of findings from individual studies [78]. Potentially useful research is expected to build upon previous work to contribute to the development of the field [34, 54]. However, traditional literature exploration typically focuses on retrieving individual publications [47, 61], rather than providing a bird's-eye view of the state of the art. Such an overview is helpful for understanding the research landscape and identifying potentially useful hypotheses for future investigation [15]. Topic-based literature exploration helps researchers to analyse large numbers of publications simultaneously, providing an overview of relations between topics rather than focusing on individual publications [12, 13, 96, 99, 114].

Our research uses neuroscience as an example to explore topic-based literature exploration. Neuroscientists need to understand the relations among complex neuroscience topics to uncover useful insights that can inform experimental design, as neuroscience encompasses many topics such as brain regions, brain diseases, genes, proteins, and mental processes. Neuroscientists often require months or even years to evaluate the usefulness of a hypothesis through an experiment, so they need to identify the hypothesis with the highest potential to yield useful contributions before committing to a specific experiment.

Additionally, given that the brain is a three-dimensional structure, we use Augmented Reality (AR) to spatially represent brain regions [96, 99]. AR allows neuroscientists to view the relative locations of brain regions affected by similar diseases, aiding their understanding of the brain's complex 3D anatomy [24]. Research has shown that immersive AR environments can enhance memory recall and spatial awareness more effectively than both Virtual Reality (VR) and 2D desktop environments, particularly supporting neuroscientists in recalling the spatial layout of 3D brain region visualisations [51, 112]. AR also facilitates the integration of literature exploration into researchers' everyday workflows by bridging virtual and physical contexts. It extends traditional 2D reading practices, allowing users to read publications on screen while simultaneously using virtual space to visualise multiple topic relations [6, 62, 88]. By representing multiple relations between brain regions and diseases in parallel, users can explore hypotheses more intuitively and directly access the corresponding publications during the literature exploration process.

Our group has developed the DatAR prototype that provides a visual representation of *direct* relations between brain regions and brain diseases [96, 99] (Sec. 1.1).

- A *direct* relation exists between two neuroscience topics , such as a specific brain region and a specific brain disease, when they co-occur in the same sentence within the title or abstract of a publication. For example, there is a direct relation between the brain disease *Alzheimer's Disease* and the brain region *Hippocampus*, as both appear in the title of the publication *Post-Ischemic Neurodegeneration of the Hippocampus Resembling Alzheimer's Disease Proteinopathy*[1].

Identifying direct relations is useful because it intuitively indicates which brain regions are affected by a brain disease [96, 99]. A small number of direct relations may indicate the existence of useful connections between a brain region and a disease that have not yet been fully explored, which in turn helps inform the design of a neuroscience experiment [99]. Neuroscientists participating in previous studies investigating the utility of direct relations [96, 99] have suggested that exploring *indirect* relations could provide useful hypotheses for designing a potentially useful experiment. Weak evidence from one or more indirect relations may suggest a potential connection between two otherwise unrelated topics, even though no publication has confirmed it directly. Identifying such indirect relations in the literature can help generate hypotheses and inform the design of an experiment to validate these potentially useful indirect relations.

- An *indirect* relation between a brain region and a brain disease occurs when there is no direct relation between them, but they each has a direct relation with the same intermediate topic, such as a mental process *Communication Response*. For example, there is no direct relation between the brain disease *Alzheimer's Disease* and the brain region *Cerebellar Vermis Structure* in the repository that we use. There are, however, direct relations between *Alzheimer's Disease* and *Communication Response* (a mental process), and direct relations between *Communication Response* and *Cerebellar Vermis Structure*. Based on these two direct relations, we conclude that there is an indirect relation between the brain disease *Alzheimer's Disease* and brain region *Cerebellar Vermis Structure* via *Communication Response*.

In this thesis, we investigate how the exploration of indirect relations between brain regions and brain diseases can inform the design of neuroscience experiments. Through close collaboration with neuroscientists, we identified three useful user tasks and corresponding research questions to guide our investigation (Sec. 1.2), which correspond to the three main chapters of this dissertation:

- Exploring indirect relations between topics (Chapter 2),
- Tracking how indirect relations evolve into direct ones over time (Chapter 3), and

[1] DOI: https://doi.org/10.3390/ijms23010306

- Showing which specific intermediate topics contribute to the indirect relations (Chapter 4).

We then describe the important contributions of this research (Sec. 1.3) and the use of Large Language Model tools (Sec. 1.4). We provide the list the publications on which the thesis is based (Sec. 1.5).

1.1. Exploring Direct Relations Between Topics in the DatAR Prototype

DatAR, an interactive 3D AR prototype, has been shown to be useful for neuroscientists in exploring direct relations between topics and serves as the starting point of this research. We describe the Knowledge Graphs of Brain Science (KGBS), which store topic co-occurrences of direct relations used in the DatAR prototype (Sec. 1.1.1). We then illustrate the integrated widgets in DatAR, each representing a specific functionality, that enables users to explore direct relations between brain regions and brain diseases (Sec. 1.1.2). Two of these widgets, *Brain Regions Visualisation* and *Brain Disease Topic Model* (Table 1.1 on p. 11), are also used in this research.

1.1.1. The Knowledge Graphs of Brain Science

The Knowledge Graphs of Brain Science[2] (KGBS) store topic co-occurrences used in the DatAR prototype. KGBS is constructed from an analysis of sentences in the titles and abstracts of 414,224 neuroscience publications indexed in PubMed[3] between January 2010 and February 2022. Given the exploratory nature of our study, this time frame provides a sufficient volume of literature for identifying topic co-occurrences (direct

[2]To access the Knowledge Graphs of Brain Science at https://kgbs-sparql.project.cwi.nl/, please contact <xuboyudesign@outlook.com>.

[3]PubMed is a free resource that supports the search and retrieval of biomedical and life sciences literature https://pubmed.ncbi.nlm.nih.gov/

Left brain disease	Middle number of co-occurrences	Right brain region
"Alzheimer's Disease"	"965"^^xsd:integer	"BRAIN, HIPPOCAMPUS"
"Alzheimer's Disease"	"36"^^xsd:integer	"Prefrontal Cortex"
"Alzheimer's Disease"	"17"^^xsd:integer	"Caudate nucleus structure"
⋮	⋮	⋮

Figure 1.1: The brain disease *Alzheimer's Disease* (left column), co-occurs 965 times (middle column) with the brain region *Hippocampus* (right column). *Alzheimer's Disease* has a direct relation with *Hippocampus*. The elements in the figure are reused in Fig. 1.2 on p. 10 for comparing indirect relations.

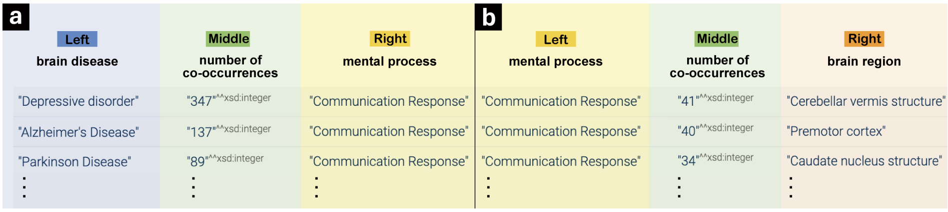

Figure 1.2: (a) The brain diseases *Depressive Disorder*, *Alzheimer's Disease*, and *Parkinson Disease* (left column) have direct relations with *Communication Response* (right column).
(b) *Communication Response* (left column) has direct relations with the brain regions *Cerebellar Vermis Structure*, *Premotor Cortex*, and *Caudate Nucleus Structure* (right column).
The brain diseases *Depressive Disorder*, *Alzheimer's Disease*, and *Parkinson Disease* have indirect relations with *Cerebellar Vermis Structure*, *Premotor Cortex* and *Caudate Nucleus Structure* via *Communication Response*. We know from the direct relations shown in Fig. 1.1 on p. 9 that *Alzheimer's Disease* already has direct relations with the brain regions *Premotor Cortex*, and *Caudate Nucleus Structure*. Based on this, we conclude that an indirect relations exist between *Alzheimer's Disease* and *Cerebellar Vermis Structure* via *Communication Response*.

relations). We use terms from the Unified Medical Language System[4] (UMLS) to identify topic names in the repository, which include brain regions, brain diseases, genes, proteins, mental processes, and symptoms. For example, an excerpt of direct relations between brain diseases and brain regions is presented in Fig. 1.1 on p. 9. Two direct relations connected through an intermediate topic can reveal an indirect relation. An example of an indirect relation between a brain disease and a brain region is illustrated in Fig. 1.2 on p. 10.

1.1.2. Widgets for Exploring Direct Relations between Topics

Widgets implemented in the DatAR prototype provide selection, querying, and visualisation mechanisms that enable users to explore neuroscience literature through direct relations stored in the KGBS, as shown in Table 1.1 on p. 11.

[4]Unified Medical Language System (UMLS) https://www.nlm.nih.gov/research/umls/index.html brings together biomedical vocabularies and standards to enable interoperability between computer systems.
[5]Scalable Brain Atlas https://scalablebrainatlas.incf.org/index.php

Table 1.1: An overview of widgets implemented in the DatAR prototype for exploring direct relations between topics [96, 99].

	Name of Widget	**Description**
#1	*Brain Regions Visualisation*, Fig. 1.3 on p. 12	The *Brain Regions Visualisation* displays 274 brain regions derived from the Scalable Brain Atlas[5] [8]. Each sphere's position corresponds to the 3D coordinates of the corresponding brain region in the atlas [96, 99].
#2	*Brain Disease Topic Model*, Fig. 1.4 on p. 13	The *Brain Disease Topic Model* presents 3D visualisations of 151 brain diseases from the KGBS [99], shown as spheres. The spatial arrangement of any two diseases is determined by their co-occurrences across all topics, where closer proximity in 3D space indicates more direct co-occurrences in the repository [96, 99].
#3	*Direct Relations Explorer*, Figs. 1.3, 1.4 and 1.5 on pp. 12–14	The *Direct Relations Explorer* allows users to query the direct relations between a single brain region and multiple brain diseases or a single brain disease and multiple brain regions in the KGBS [96, 99].
#4	*Direct Relations Filter*, Fig. 1.5 on p. 14	The *Direct Relations Filter* allows users to filter the number of co-occurrences in the KGBS between a single brain region and multiple brain diseases, or between a single brain disease and multiple brain regions [96, 99]. Users can filter for fewer or more direct relations by adjusting the upper and lower handles in the visualisation. The number of direct relations is shown using histograms and the exact number.

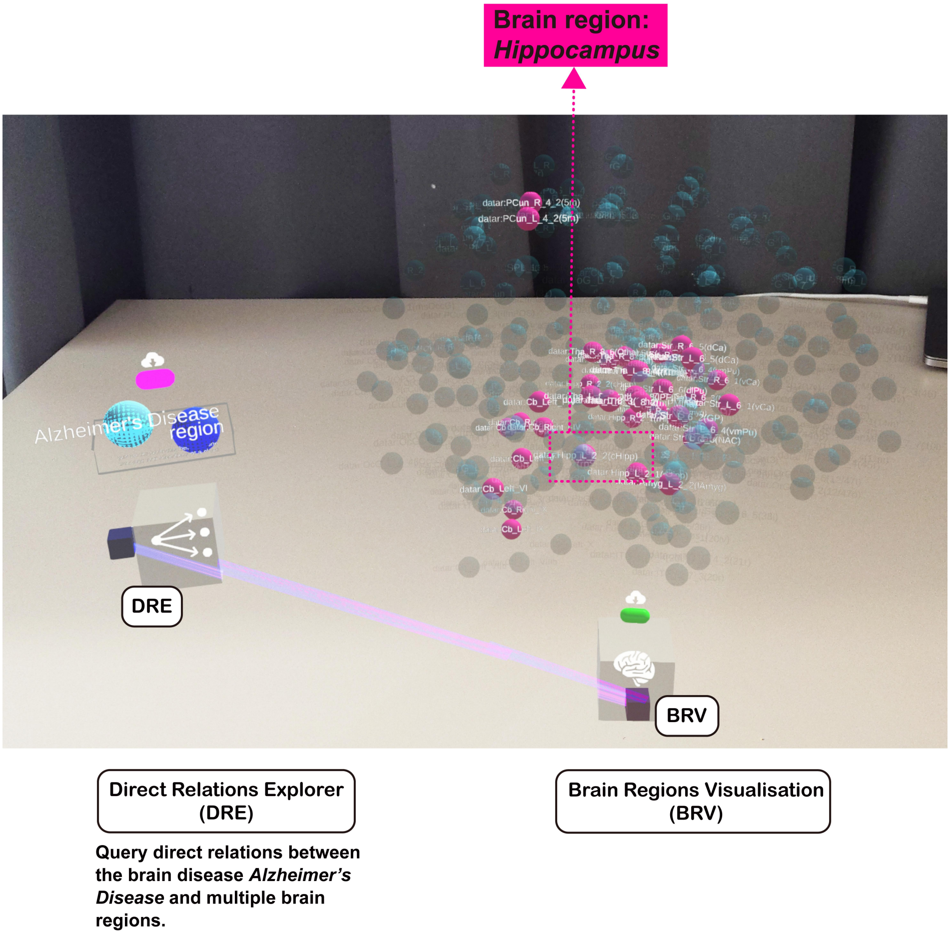

Figure 1.3: A user uses the *Direct Relations Explorer* to explore the direct relation between the brain disease *Alzheimer's Disease* and all brain regions. Thirty-six direct relations (pink spheres) are shown in the *Brain Regions Visualisation*, including two pink spheres representing both *Hippocampus* brain regions. Brain regions with no direct relations to *Alzheimer's Disease* are visualised as light blue spheres[6].

[6]The example of a direct relation between the brain disease *Alzheimer's Disease* and the brain region *Hippocampus* is also used in Chapter 4.

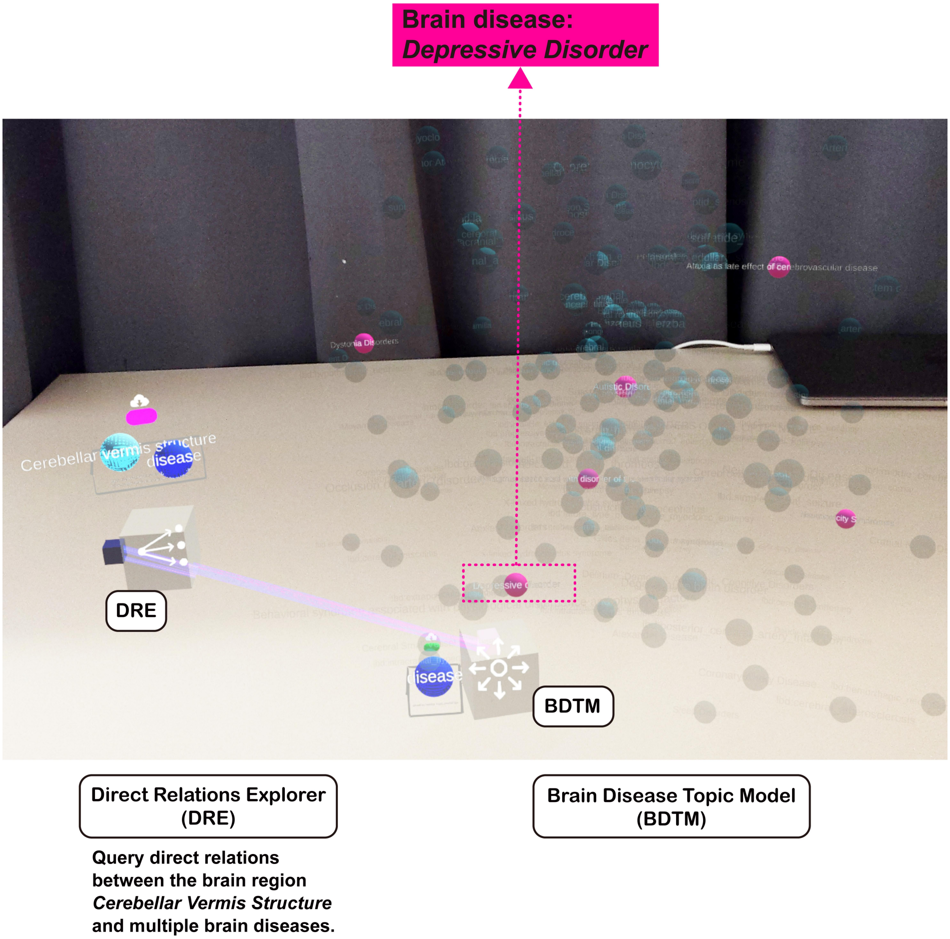

Figure 1.4: A user uses the *Direct Relations Explorer* to explore the direct relations between the brain region *Cerebellar Vermis Structure* and all brain diseases. Six direct relations (pink spheres) are shown in the *Brain Disease Topic Model*, including a pink sphere representing the brain disease *Depressive Disorder*. Brain regions with no direct relations to *Cerebellar Vermis Structure* are visualised as light blue spheres[7].

[7]The example of a direct relation between the brain region *Cerebellar Vermis Structure* and the brain disease *Depressive Disorder* is also used in Chapter 4.

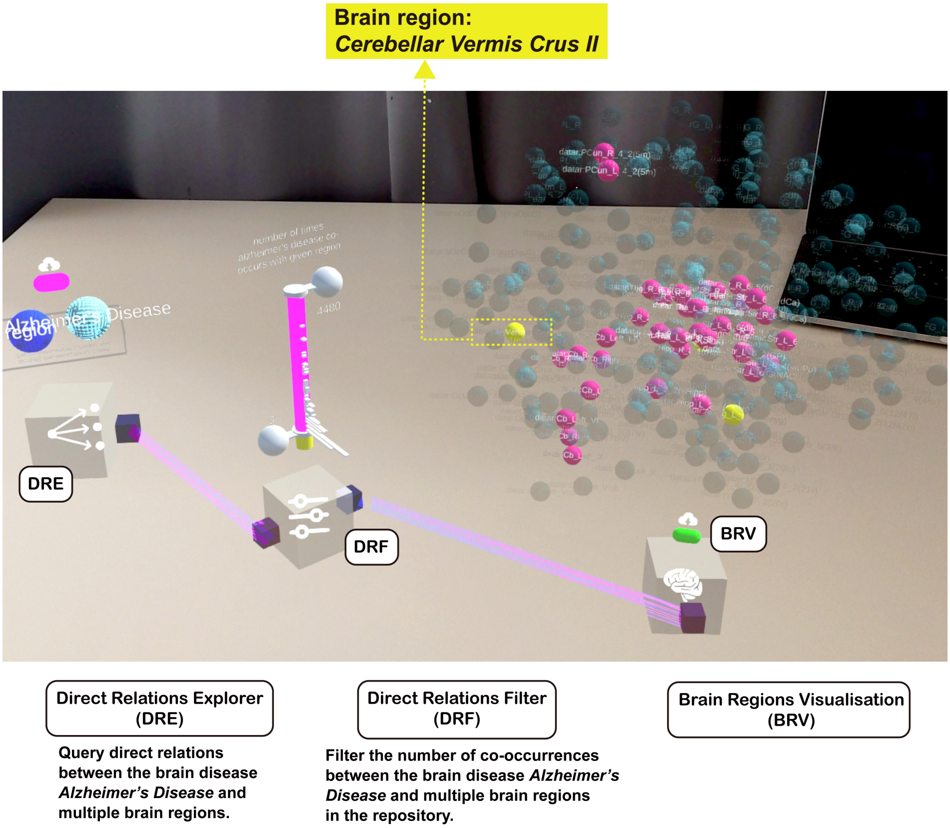

Figure 1.5: A user uses the *Direct Relations Filter* to filter the number of co-occurrences between the brain disease *Alzheimer's Disease* and all brain regions. Three brain regions co-occur with *Alzheimer's Disease* one, two, or three times, including *Cerebellar Vermis Crus II* (yellow sphere). The number of direct relations is shown using histograms labelled with the exact number in the *Direct Relations Filter*[8].

[8]The example of exploring the number of co-occurrences between the brain disease *Alzheimer's Disease* and all brain regions is also used in Chapter 4.

1.2. Identifying Neuroscience Literature Exploration Tasks for which Exploring Indirect Relations is Useful

We follow a user-centred design (UCD) approach [27, 45] to ensure that our research supports neuroscientists in identifying potentially useful indirect relations through literature exploration tasks that can inform the design of an experiment. Neuroscience researchers were consulted throughout the entire research process. We first established which neuroscience literature exploration tasks are perceived as useful within the context of neuroscientists' research activities. Three useful literature exploration tasks were identified, and their interconnections are illustrated in Fig. 1.6 on p. 15. Each useful user task corresponds to a main chapter of this dissertation.

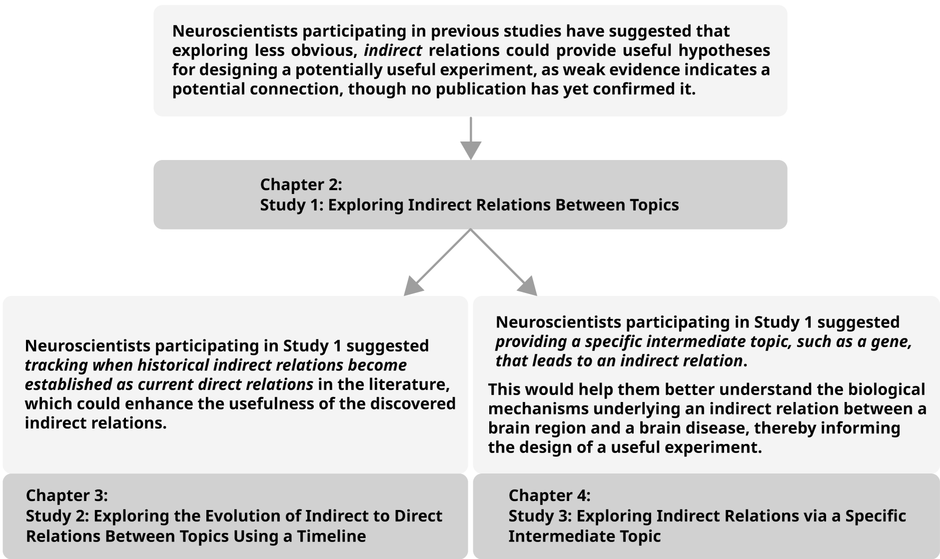

Figure 1.6: illustrates how the three useful literature exploration tasks were identified in collaboration with neuroscientists, and how the three user tasks are interconnected.

Chapter 2

Study 1: Exploring Indirect Relations Between Topics

Neuroscientists participating in previous studies [96, 99] suggested that exploring less obvious, indirect relations could provide useful hypotheses for designing a potentially useful experiment, as weak evidence, via one or more intermediate topics, may suggest a potential connection that has not yet been confirmed by any publication.

- **Main Research Question 1.** *How does **exploring indirect relations between topics** contribute to designing a useful experiment?*

While previous participating neuroscientists suggested that exploring indirect relations could be useful [96, 99], we first needed to confirm this. We consulted three neuroscientists, who determined that exploring indirect relations is indeed a potentially useful task. To investigate indirect relations between topics, we need to specify which functionality and visualisation are required to support this. Via interviews with the same three neuroscientists, we determined functional requirements and visualisation requirements. Based on these, we designed the *Indirect Relations Explorer* (Chapter 2).

Having designed and implemented the indirect relation visualisation,

– we need to investigate the suitability of the visualisation for presenting indirect relations.

– we also need to investigate the usability of the corresponding functionality for exploring indirect relations.

– we then evaluate the usefulness of exploring indirect relations between topics for neuroscientists to design a useful experiment (MRQ1).

Chapter 3

Study 2: Exploring the Evolution of Indirect to Direct Relations Between Topics Using a Timeline

Study 1 showed that identifying indirect relations between topics, specifically brain regions and brain diseases, through an intermediate topic type such as genes or mental processes, served as inspiration for further literature review, but did not provide sufficient evidence to inform a useful experiment [108, 109]. One participant suggested investigating trends in indirect relations and tracking when these indirect relations become established as direct relations in the literature. Showing how previous indirect relations have evolved into current direct ones could increase confidence in the potential of current indirect relations to become future direct ones.

- **Main Research Question 2.** *How does **exploring the evolution of historical indirect to current direct relations** contribute to designing a useful experiment?*

  We first consulted two neurologists to advise us on whether exploring the evolution of indirect to direct relations is a potentially useful user task. To investigate the evolution of historical indirect to current direct relations, we need to specify which functionality and visualisation are required to support this. Using input from interviews with the same two neurologists, we identified functional requirements and the timeline visualisation requirements. Based on these, we designed the *Timeline Selector* (Chapter 3).

  Having designed and implemented the timeline visualisation,

  – we need to investigate the suitability of the implemented visualisation for presenting the evolution of indirect to direct relations.

  – we also need to investigate the usability of the implemented functionality for exploring the evolution of historical indirect to current direct relations.

– we then evaluate the usefulness of exploring the evolution of historical indirect to current direct relations for neuroscientists to design a useful experiment (MRQ2).

Chapter 4

Study 3: Exploring Indirect Relations via a Specific Intermediate Topic

A neuroscientist participating in the first study also noted that while visualising indirect relations in 3D brain regions offers an intuitive approach to identifying potentially useful experiments, providing specific intermediate topics to find indirect relations would help assess their usefulness [108, 109]. By selecting a specific intermediate topic, such as a gene, protein, or mental process, neuroscientists can better understand the biological mechanisms underlying an indirect relation between a brain region and a brain disease, thereby informing the design of a useful experiment.

- **Main Research Question 3.** *How does **exploring indirect relations via a specific intermediate topic** contribute to designing a useful experiment?*

  We first consulted two neuroscientists to determine whether exploring indirect relations via a specific intermediate topic is a potentially useful user task. To investigate the exploration of indirect relations via a specific intermediate topic, we need to specify which functionality and visualisation are required to support this. Using input from interviews with the same two neuroscientists, we identified functional and visualisation requirements. Based on these, we designed the *Specific Intermediate Topic Explorer* (Chapter 4).

  Having designed and implemented the intermediate topics selector,

  – we need to assess the suitability of the visualisation for presenting indirect relations via a specific intermediate topic.

  – we also need to investigate the usability of the implemented functionality for exploring indirect relations via a specific intermediate topic.

  – we then evaluate the usefulness of exploring indirect relations via a specific intermediate topic for neuroscientists to design a useful experiment (MRQ3).

This research does not focus on building a prototype, but rather on assisting neuroscientists in literature exploration tasks to identify potentially useful indirect relations that can inform the design of an experiment. We evaluate whether exploring indirect relations between topics (MRQ1), exploring the evolution from indirect to direct relations (MRQ2), and exploring indirect relations via a specific intermediate topic (MRQ3) can provide useful information to support experimental design. The implemented functionalities and visualisations are developed to support the execution of three user tasks.

1.3. Contributions

Our research contributes to exploring useful neuroscience literature tasks that inform the design of a useful experiment, alongside designing functionalities and corresponding visualisations to support these tasks. To present an overview of our main contributions, we organise them around the CHI contribution types [106]:

1.3.1. Empirical Contributions

- Identifying three useful literature exploration tasks for neuroscientists to inform experimental design through close collaboration with neuroscientists. These tasks are grounded in neuroscientists' feedback and address their practical needs in exploring the neuroscience literature to propose useful hypotheses.
  - Identified the user task of exploring indirect relations between topics based on previous studies [96, 99] (Chapter 2).
  - Identified the user task of exploring the evolution from indirect to direct relations between topics using a timeline, based on evaluation feedback on exploring indirect relations (Chapter 3).
  - Identified the user task of exploring indirect relations via a specific intermediate topic, based on evaluation feedback on exploring indirect relations (Chapter 4).
- Insight into how literature exploration tasks can support experimental design for neuroscientists:
  - Found that exploring indirect relations can provide useful hypotheses for designing a useful experiment, as weak evidence indicate potential connections that have not yet been confirmed in the literature (MRQ1, Chapter 2).
  - Found that exploring how historical indirect relations between two topics evolve into direct relations in more recent publications using a timeline helps neuroscientists build confidence in the usefulness of current indirect relations for designing a useful experiment (MRQ2, Chapter 3).
  - Found that exploring indirect relations via a specific intermediate topic helps neuroscientists understand the underlying mechanisms of identified indirect relations, thereby supporting the design of a useful experiment (MR-Q3, Chapter 4).

1.3.2. Artefact Contributions

- Designing and implementing functionalities and corresponding visualisations to support the three literature exploration tasks:
  - Designed the *Indirect Relations Explorer* in the DatAR prototype to enable neuroscientists to select an intermediate topic type, such as mental processes or genes, to explore indirect relations between brain regions and brain diseases (Chapter 2).

  – Implemented the corresponding visualisation for presenting direct and indirect relations in a single visualisation (Chapter 2).
  – Designed the *Timeline Selector* in the DatAR prototype to allow neuroscientists to specify a range of publication dates in order to explore the evolution of historical indirect relations to direct relations (Chapter 3).
  – Implemented the corresponding visualisation for presenting the evolution of indirect to direct relations in a single visualisation (Chapter 3).
  – Designed the *Specific Intermediate Topic Explorer* in the DatAR prototype to enable neuroscientists choose a specific intermediate topic, such as a mental process Communication Response, to explore indirect relations (Chapter 4).
  – Implemented the corresponding visualisation for presenting indirect relations via a specific intermediate topic (Chapter 4).

1.4. Use of Large Language Model Tools

This thesis adheres to the guidelines set by both the Association for Computing Machinery (ACM) and Utrecht University concerning the responsible use of Large Language Model (LLM) tools such as OpenAI's ChatGPT. ChatGPT was used in this dissertation to support the writing and revision process by improving grammatical accuracy and enhancing clarity. ChatGPT was not used to generate novel research ideas, write substantive academic content, or analyse data. The author maintains full intellectual ownership and responsibility for the research questions, study design, data analysis, and final written contributions. All uses of ChatGPT were carefully reviewed, edited, and verified to ensure correctness, originality, and alignment with academic integrity principles.

1.5. Thesis Material

The following publications are based on the research conducted in this PhD thesis. The first author was responsible for the experimental design and implementation, and the preparation of the manuscript for the three main chapters. The co-authors provided feedback on the experimental design and contributed to improving the manuscript.

An overview of this research has been published in a Doctoral Consortium:

- **Xu, B.** (2024, March). Supporting Neuroscience Literature Exploration by Utilising Indirect Relations between Topics in Augmented Reality. In *Proceedings of the 2024 Conference on Human Information Interaction and Retrieval (CHIIR 2024)* (pp. 457-460). DOI: https://doi.org/10.1145/3627508.3638312

The research in **Chapter 2** has been published in two poster papers:

- **Xu, B.**, Hardman, L., & Hürst, W. (2025, March). Exploring Indirect Relations between Topics in Augmented Reality to Inform the Design of a Neuroscience

Experiment. In *2025 IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR 25)* (pp. 1248-1249). IEEE. DOI: 10.1109/VRW66409.2025.00270

- **Xu, B.**, Tanhaei, G., Hardman, L., & Hürst, W. (2024, January). DatAR: Supporting neuroscience literature exploration by finding relations between topics in augmented reality. In *International Conference on Multimedia Modeling (MMM 24)* (pp. 295-300). Cham: Springer Nature Switzerland. DOI: 10.1007/978-3-031-53302-0_24

The research in **Chapter 2** has also resulted in a full paper, in submission:

- **Xu, B.**, Hardman, L., & Hürst, W. Informing the Design of a Neuroscience Experiment by Exploring Indirect Relations Between Topics in Augmented Reality

The research in **Chapter 3** has resulted in a full paper, in submission:

- **Xu, B.**, Hardman, L., & Hürst, W. Informing the Design of a Neuroscience Experiment by Exploring the Evolution of Indirect to Direct Relations Between Topics in Augmented Reality Using a Timeline. DOI: https://doi.org/10.21203/rs.3.rs-6327352/v1

The research in **Chapter 4** has resulted in a full paper, in submission:

- **Xu, B.**, Hardman, L., & Hürst, W. Informing the Design of a Neuroscience Experiment by Exploring Indirect Relations via a Specific Intermediate Topic

An additional publication, which is not included in this PhD thesis, has been published in the journal *Wiley Advanced Science*:

- **Xu, B.**, Li, G., Wang, B., Bian, J., Pan, H., Min, Y., ... & Wang, Z. (2025, October). Knowledge Graph for Methane Selective Conversion: Revisiting and Predicting Product Selectivity and Methane Conversion. *Advanced Science*, e14601. DOI: https://doi.org/10.1002/advs.202514601

2

Informing the Design of a Neuroscience Experiment by Exploring Indirect Relations Between Topics in Augmented Reality

Before conducting a costly experiment, neuroscientists need to analyse large numbers of publications to inform the design of an experiment. This process is time-consuming. Topic-based literature exploration is a useful means to analyse many publications simultaneously because it provides an overview of relations between neuroscience topics. A co-occurrence of two topics, such as a brain region and a brain disease, in the same sentence in the title or abstract of a publication, implies a direct relation between them. A brain region may also be connected indirectly to a brain disease through an intermediate topic, such as a gene or a mental process. Our goal is to establish whether the task of exploring indirect relations between brain regions and diseases would assist neuroscientists to design a useful experiment. Using a user-centred design approach, we (i) interview neuroscientists to establish the usefulness of exploring indirect relations between topics; (ii) specify the functionality required for this task, and (iii) design and implement corresponding visualisation in 3D Augmented Reality (AR). Nine neuroscientists carried out two representative tasks using the implemented functionality and visualisation. Their responses to 14 Likert-scale questions and 6 semi-structured interview questions indicated that (a) the visualisation is *suitable* for presenting the implemented functionality, (b) the implemented functionality is *useful* for exploring indirect relations, and (c) exploring indirect relations between topics is *useful* to help neuroscientists design a useful experiment.

Chapter contribution: This chapter investigates how the user task of exploring indirect relations between topics can inform the design of a neuroscience experiment (MRQ1), specifies an appropriate functionality to support this task, and designs and implements the corresponding 3D AR visualisation to enable this functionality.

2.1. Introduction

Neuroscientists need to ensure that their costly experiments are likely to make a useful contribution. They first need to understand the relations between topics, such as brain regions and brain diseases, that have already been published [72, 91]. Topic-based literature exploration is a useful approach that allows users to obtain an overview of relations between topics without needing to read all the individual publications involved [97, 99]. In particular, participants were able to use direct relations between topics to explore which brain regions affect which brain disease and vice versa, Fig. 2.1 on p. 25. To more easily understand the relations between brain regions and brain diseases, Augmented Reality (AR)-based spatial representation in 3D allows users to see the relative locations of the regions affected by the same or similar diseases [81]. Showing the brain regions visualisation in the context of the real world through AR allows researchers to better incorporate the literature exploration process into their daily research work, such as reading publications from their 2D screens [23, 57]. Users also have sufficient virtual space to visualise multiple sets of relations to compare different lines of thought in their literature exploration process [21, 22, 103].

Direct relations are derived from co-occurrences of two topics in the same sentence in the title or abstract of a publication in PubMed[1] [99]. For example, the brain region *Hippocampus* and the brain disease *Depression* occur in the title of the publication *Hippocampus atrophy and the longitudinal course of late-life depression*[2]. Given this co-occurrence, we say that there is a direct relation between the topics *Hippocampus* and *Depression*. We define an indirect relation between two topics when there is no direct relation between them but each co-occurs with at least one other intermediate topic. For example, there is no direct relation between the brain region *Hippocampus* and the brain disease *Alexander Disease* in the subset of the PubMed corpus that we use, Fig. 2.3 on p. 27. There are, however, 211 co-occurrences of *Hippocampus* with *GFAP Gene*, and 7 co-occurrences of *GFAP Gene* with *Alexander Disease*. We thus conclude that there is an indirect relation between the topics *Hippocampus* and *Alexander Disease*, in this case through the intermediate topic *GFAP Gene*.

Neuroscientists participating in previous studies [56, 97, 99] have suggested that exploring less obvious, *indirect* relations could provide useful hypotheses for designing a potentially useful experiment, as weak evidence indicates a potential connection, though no publication has yet confirmed it. Our main research question (RQ1) is: *How does exploring indirect relations between topics contribute to designing a useful experiment?* While previous participants suggested that exploring indirect relations could be useful [97, 99], we first need to confirm that this is indeed the case. We follow a user-

[1] PubMed Resource https://pubmed.ncbi.nlm.nih.gov/
[2] DOI: https://doi.org/10.1016/j.jagp.2013.11.004

centred design approach and consulted three neuroscientists to advise us on whether exploring indirect relations is a potentially useful task, of which one participant also identified two representative tasks for our later evaluation, Sec. 2.3.1.

Having established that exploring indirect relations is a potentially useful user task, we need to specify which functionality and visualisation are required to support this. Via interviews with the same three neuroscientists, we determined a functional requirement, Sec. 2.3.2. The same three neuroscientists also helped us identify visualisation requirements, Sec. 2.3.3. Based on these functional and visualisation requirements, we designed the *Indirect Relations Explorer*, Fig. 2.6 on p. 33.

Having designed and implemented the indirect relation visualisation, we need to investigate the suitability of the visualisation for presenting indirect relations. Our third research question (RQ3) is: *To what extent is the visualisation suitable for the specified functionality in presenting indirect relations?* We describe the evaluation procedure undertaken by nine neuroscientists using the questionnaire and semi-structure interview to assess the visualisation, Sec. 2.5. We evaluated the suitability of the visualisation using two representative tasks with nine neuroscientists, Sec. 2.6.1.

Having evaluated the implemented visualisation, we need to investigate the usability of the corresponding functionality for exploring indirect relations. Our second research question (RQ2) is: *To what extent is the implemented functionality useful for exploring indirect relations?* The nine neuroscientists who evaluated RQ3 were also asked to assess the usability of the functionality implemented in the *Indirect Relations Explorer*, Sec. 2.6.2.

Our final step is to evaluate the main research question, the usefulness of indirect relations in designing a useful experiment (RQ1). The same nine neuroscientists were consulted and confirmed that exploring indirect relations can be useful in identifying an experiment likely to make a useful contribution. Five participating neuroscientists noted that an indirect relation between the brain region *Hippocampus* and the brain disease *Alexander Disease* via genes is useful for them to design an experiment, as no publications directly mention this relation, but a potential connection can be inferred through genes, Sec. 2.6.3.

The contributions of this work are (i) establishing that the user task of exploring indirect relations between neuroscience topics is useful (answer to RQ1), (ii) the design and evaluation of enabling functionality (answer to RQ2), and (iii) the design and evaluation of a corresponding visualisation (answer to RQ3).

2.2. Related Work

We explore the usefulness of topic-based literature exploration for identifying a potentially useful experiment from many publications, Sec. 2.2.1. We explain how exploring indirect relations between topics in biomedical tasks can reveal useful insights, Sec. 2.2.2. We describe a functioning AR prototype that allows neuroscientists to explore direct relations between brain regions and brain diseases, Sec. 2.2.3.

2.2.1. Topic-based Literature Exploration

Traditionally, researchers have explored the literature by searching for individual papers. However, this approach falls short of capturing the complex relations within a large body of literature. Topic-based literature exploration, which, for example, in our case considers *Depressive Disorder* as a brain disease, allows researchers to understand which brain regions are affected by *Depressive Disorder* [60, 101]. Topic-based literature exploration has the potential to yield useful insights from existing research. For example, it has been used to understand the link between *genomics* and *personalised medicine*, reflecting the usefulness of genetic research in customised healthcare [5]. It has also been used to identify genes that could potentially affect disease risk, such as revealing seven genes (*EZH2*, *ARID1A*, *MEF2B*, *EP300*, *FKH1*, *CREBBP*, and *CARD11*) used in predicting *follicular lymphoma* [77]. In our study, neuroscientists are looking for less obvious, indirect relations between topics, as these may point to useful hypotheses in the literature where a single experiment may be able to establish a direct relation between them.

2.2.2. Exploring Indirect Relations between Topics

The ABC co-occurrence model is a method for exploring previously unknown links or indirect relations [38, 94, 98]. It states that if topic A (such as a brain region) is associated with topic B (such as a gene), and B is associated with topic C (such as a brain disease), then topic A has indirect relations with topic C [25]. The potential of the ABC co-occurrence model for finding indirect relations has been shown in a broad range of biomedical tasks [90]. It has been used to find the disease–gene and gene–drug relations with the aim of repositioning candidate drugs for treating *ovarian cancer* [110]. It also has been used for adverse drug event prediction, such as drugs effects on some proteins and these proteins are associated with the adverse effects [42].

Inspired by the ABC co-occurrence model, we apply it to neuroscience literature. This approach is useful for exploring complex relations among neuroscience topics, such as identifying which brain regions might be indirectly connected through genes to brain diseases such as *Depressive Disorder*. This provides a small amount evidence that could indicate a direction for a useful experiment.

2.2.3. The 3D AR Topic Exploration Prototype

Previous research introduced a 3D AR prototype that has been shown to be useful for neuroscientists in exploring direct relations between topics [37, 97, 99]. We describe the constituent widgets in the existing prototype that allow users to explore these direct relations, Sec. 2.2.3. The direct relations queried in the prototype are derived from co-occurrences between two topics extracted from neuroscience publications in PubMed and stored in the Knowledge Graphs of Brain Science (KGBS). Details on how these relations are determined are given in Sec. 2.2.3.

Widgets for Exploring Direct Relations between Topics

Users are able to explore brain-related topics through a number of functionalities and visualisations implemented as widgets [37, 97, 99]. For example, users can explore overviews of direct relations between a selected brain region and multiple brain dis-

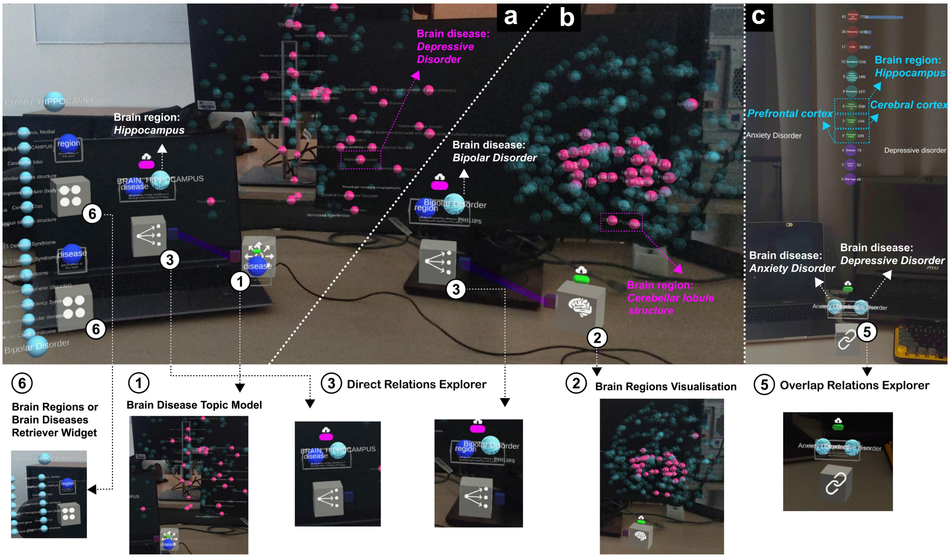

Figure 2.1: Overview of previous studies on the 3D AR prototype [37, 97, 99], which served as a starting point in this work. (a) The brain region *Hippocampus* (③ middle) has a direct relation with the brain disease *Depressive Disorder* (a pink sphere, ① right). (b) The brain disease *Bipolar Disorder* (③ left) has a direct relation with the brain region *Cerebellar Lobule Structure* (2 of the pink spheres, ② right). Direct relations between topics are visualised as pink spheres; no direct relations are visualised as light blue spheres. (c) The specific brain regions, such as *Hippocampus*, *Cerebral cortex* and *Prefrontal cortex*, have direct relations with both the brain diseases *Anxiety Disorder* (⑤ lower left) and *Depressive Disorder* (⑤ lower right) simultaneously.

eases, or a selected brain disease and multiple brain regions. Widgets needed for exploring direct relations between topics include:

- The *Brain Disease Topic Model*, Fig. 2.1 ① on p. 25, displays 151 brain diseases included in the repository. Each sphere represents a single brain disease; the distance between any two diseases indicates their similarity based on all co-occurrences between all topics, i.e. the closer they are in the 3D space, the more connections there are in the repository.

- The *Brain Regions Visualisation*, Fig. 2.1 ② on p. 25, displays 274 brain regions from the Scalable Brain Atlas[3] (SBA) [8]. The position of each sphere indicates the 3D position (taken from SBA) of the brain region it represents [99].

- The *Direct Relations Explorer*, Fig. 2.1 ③ on p. 25, is used to query direct relations between a brain region and multiple brain diseases, or between a brain disease and multiple brain regions. Direct relations can be visualised in the *Brain Diseases Topic Model* or the *Brain Regions Visualisation*, Fig. 2.1 (a) and (b) on p. 25.

[3] Scalable Brain Atlas (SBA) https://scalablebrainatlas.incf.org/index.php

- The *Overlap Relations Explorer*, Fig. 2.1 ⑤ on p. 25, is used to query intermediate topics that have direct relations with two brain diseases, such as *Anxiety Disorder* and *Depressive Disorder*, simultaneously.

Previous versions of the 3D AR prototype [37, 97, 99] have not supported the exploration of indirect relations between brain regions and brain diseases via intermediate topics, nor have illustrated the visualisation of indirect relations in the *Brain Disease Topic Model* and *Brain Regions Visualisation*. The *Overlap Relations Explorer*, Fig. 2.1 ⑤ on p. 25, displays intermediate topics between two diseases to help users understand why diseases are located close to or far away from each other in the *Brain Disease Topic Model*. The finding and visualising of indirect relations between a selected brain region, or disease, with all brain diseases, or all brain regions, respectively, are explored in this study.

The Repository for Querying Direct Relations between Topics

The Knowledge Graphs of Brain Science[4] (KGBS) contains an analysis of sentences in the titles and abstracts of 414,224[5] neuroscience publications in PubMed. We use KGBS to query co-occurrences between two topics in the corpus. Topics include brain regions, brain diseases, genes, proteins, mental processes and symptoms.

To provide an impression of the direct relations available, we show an extract of co-occurrences (direct relations) between brain regions and brain diseases, Fig. 2.2 on p. 26.

2.3. Method

We follow a user-centred design approach of problem discovery to design support for an identified user task [17, 26, 27, 75]. This approach employs an iterative understand-

[4]To access the Knowledge Graphs of Brain Science at https://kgbs-sparql.project.cwi.nl/, please contact <xuboyudesign@outlook.com>.

[5]Publications analysed in the KGBS are from January 1, 2010 to February 3, 2022.

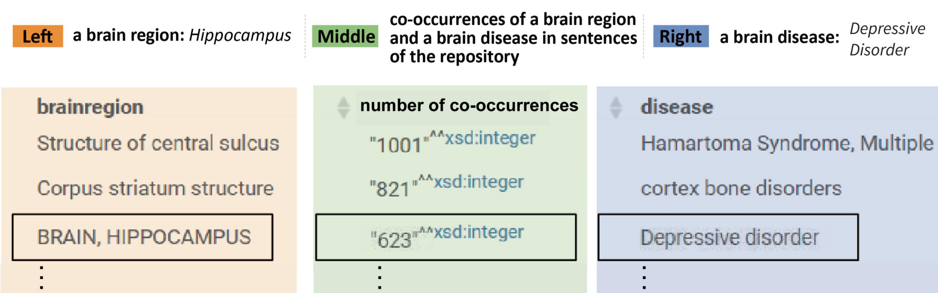

Figure 2.2: The brain region *Hippocampus* (left column), co-occurs 623 times (middle column) with the brain disease *Depressive disorder* (right column). The elements in the figure are reused in Fig. 2.3 on p. 27 for comparing indirect relations.

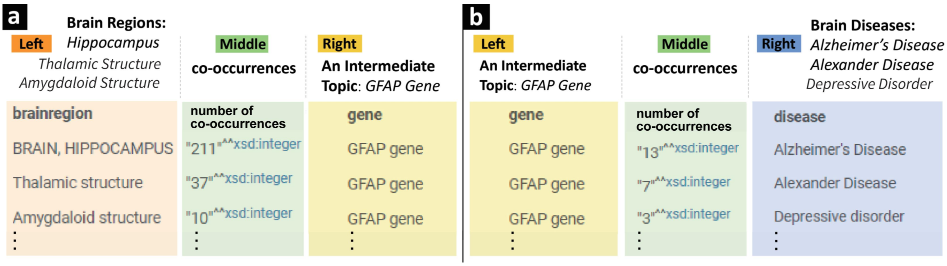

Figure 2.3: Direct relations between brain regions and the topic *GFAP Gene* and direct relations between the topic *GFAP Gene* and brain diseases. (a) The brain regions *Hippocampus*, *Thalamic Structure* and *Amygdaloid Structure* (left column) have direct relations with *GFAP Gene* (right column). (b) *GFAP Gene* (left column) has direct relations with the brain diseases *Alzheimer's Disease*, *Alexander Disease* and *Depressive Disorder* (right column). The brain regions *Hippocampus*, *Thalamic Structure* and *Amygdaloid Structure* thus have indirect relations with the brain diseases *Alzheimer's Disease*, *Alexander Disease* and *Depressive Disorder* via *GFAP Gene*.

ing of users' needs and involves them throughout the design process to ensure that the system effectively supports their literature exploration tasks. Neuroscience researchers were consulted at every stage of our work. We first identify a useful user task of exploring indirect relations between topics and also determine two representative tasks useful for our evaluation, Sec. 2.3.1. We then determine the corresponding functionality for the task, Sec. 2.3.2. Based on the functionality determined, we determine suitable visualisation requirements, Sec. 2.3.3.

Building on the user-centred design approach, we seek to obtain feedback from neuroscientists through evaluations to gain in-depth insights. To facilitate a comparative analysis of participants' perceptions, we used a 5-point Likert scale, a widely used method for quantifying opinions, ranging from strongly disagree to strongly agree [4], to assess users' perceptions of the visualisation (RQ3), functionality (RQ2), and the user task (RQ1). To gather detailed insights for future improvements, we conducted semi-structured interviews with participants. The evaluation is centered on three metrics: *Suitability* [1], *Usability* [59], and *Usefulness* [67].

- *Suitability* evaluates how the designed visualisation supports the presentation of indirect relations within the AR environment (RQ3).
- *Usability* measures how the implemented functionality satisfies human-computer interaction requirements for exploring indirect relations (RQ2).
- *Usefulness* assesses how exploring indirect relations between topics aids neuroscientists in designing a useful experiment (RQ1).

Because of the proof-of-concept character of our research, we are not focusing on quantitative performance results but rather on qualitative insights into the suitability and usability of our design for neuroscience researchers. Problem discovery studies typically require between 3 and 20 participants, with 5 to 10 being a good baseline [89].

For this study, we recruited nine participants.

2.3.1. Identifying a Useful User Task and Representative Evaluation Tasks

We first explain why the task of exploring indirect relations between topics could be useful, Sec. 2.3.1. We then determine two example tasks that are representative of exploring indirect relations, Sec. 2.3.1.

The Task of Exploring Indirect Relations is Perceived as Useful

We determine tasks that are perceived as useful, via observations from the previous studies [56, 99]. We then complement these results with additional interviews with neuroscientists to determine the task of exploring indirect relations between topics.

- ***Useful Tasks via Observations from Previous Studies*** Seven neuroscientists from previous studies (P1-P7, Table 2.2 on p. 34) proposed that exploring relations among topics is useful for them to understand neuroscience literature [56, 99]. Tasks deemed useful by neuroscientists, Table 2.1 on p, 29, include:
  - *Exploring known and unknown relations among topics* in literature was suggested by a neuroscientist (P1 from [99]) as the initial task. Known relations, such as direct relations, are visualised in the existing 3D AR prototype, Fig. 2.1 on p. 25.
  - Exploring unknown relations in the literature, such as *exploring indirect relations between topics*, was also proposed by P1 in [99]. Although only one participating neuroscientist (P1) found exploring indirect relations between topics useful, detailed discussions with the neuroscientist suggested that even one case is worthy of research [14].
- ***Determining the Task of Exploring Indirect Relations*** We interviewed three neuroscientists (P8-P10, Table 2.2 on p. 34) to further discuss the task of exploring indirect relations. All three neuroscientists, plus one (P1) [99] were interested in exploring indirect relations between topics to provide more inspiration for designing a useful experiment. For example, one neuroscientist (P8) studying the brain region *Hippocampus* wants to identify brain diseases not yet mentioned in the literature but potentially related to *Hippocampus*, as a starting point for investigation. P8 also proposed that other topics, such as genes, may have direct relations with *Hippocampus*, and these genes could also be directly related to certain brain diseases. They are interested in uncovering brain diseases related to *Hippocampus* through indirect pathways involving brain regions, genes, and diseases.

  Considering the topics in the KGBS (Sec. 2.2.3), neuroscientists (P8-P10) preliminarily identified intermediate topics, including genes, mental processes and proteins, that may facilitate exploring indirect relations between brain regions and brain diseases.

Table 2.1: An overview of tasks deemed as useful by neuroscientists. These were collected from previous studies [56, 99] and interviews conducted by the first author with the previous 3D AR prototype.

Task Description	**Sources**	**Reference**
Exploring known and unknown relations among topics	P1 from [99]	Initial task [56, 99]
Exploring indirect relations between topics	P1 from [99] and interviews by the first author with P8-P10, Table 2.2 on p. 34	User task in this study

Determining Two Representative Evaluation Tasks

To allow participants to evaluate the implementation of exploring indirect relations, we discussed with one of the neuroscientists (P8, Table 2.2 on p. 34) to find two representative tasks that would be suitable. A potentially suitable task is to ensure that participants are not confronted with a task that has too many (e.g., selecting almost all other diseases or regions) or no indirect relations. The first representative task was to understand which brain diseases have indirect relations with *Hippocampus*, Sec. 2.3.1. For example, the brain disease *Alexander Disease* can be queried the indirect relation with the brain region *Hippocampus* via intermediate genes, including *GFAP Gene*, *LITAF Gene* and *Retinoic Acid Response Element*. Neuroscientist (P8) considered exploring the indirect relation between the brain region *Hippocampus* and the brain disease *Alexander Disease* via genes can be the first example for evaluating the implementation of exploring indirect relations in the AR prototype, Fig. 2.4 on p. 30.

The neuroscientist (P8) also could see that *Hippocampus* has a direct relation to the brain disease *Bipolar Disorder*, and would like to understand which other brain regions have indirect relations with *Bipolar Disorder*. They want to explore the relations between the queried indirect brain regions and *Hippocampus*, when these brain regions were either directly or indirectly related to the same brain disease *Bipolar Disorder*. For example, the brain region *Cerebellar Vermis Structure* can be queried the indirect relation with the brain disease *Bipolar Disorder* via mental processes, including *Communication Response*, *Psychological Inhibition* and *Like*. The task of exploring the indirect relation between the brain disease *Bipolar Disorder* and the brain region *Cerebellar Vermis Structure* via mental processes thus can be the second evaluating example, Fig. 2.5[6] on p. 31.

[6]See also the video: https://youtu.be/uOiMSwJ5AAQ

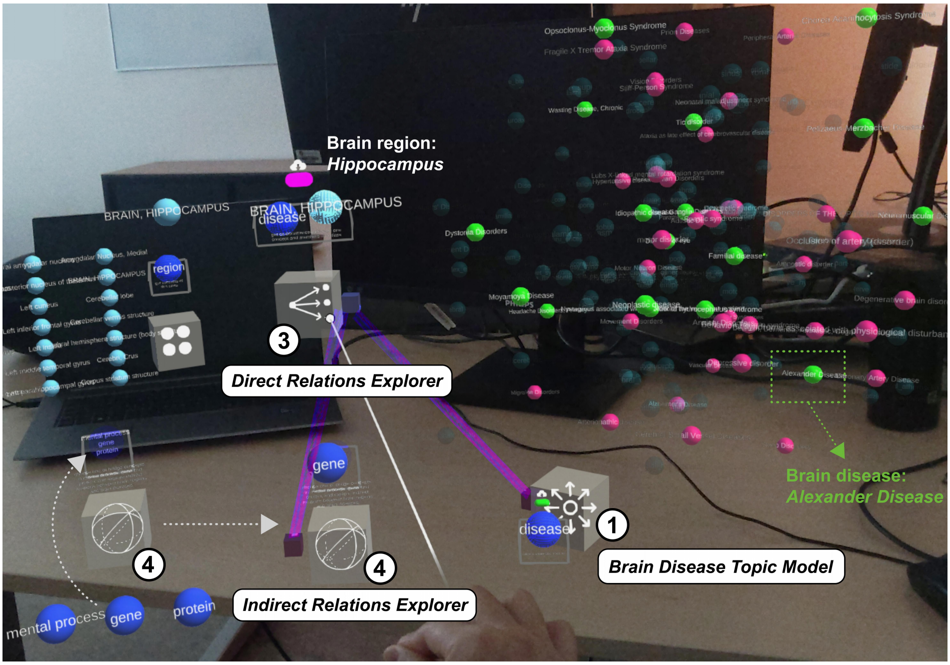

Figure 2.4: A participant selects genes as the intermediate topic type (④ lower left), and explores the indirect relation between the brain region *Hippocampus* (③ middle) and the brain disease *Alexander Disease* (a green sphere, ① lower right).
Direct relations between topics are visualised as pink spheres; indirect relations between topics are visualised as green spheres; no direct or indirect relations are visualised as light blue spheres.

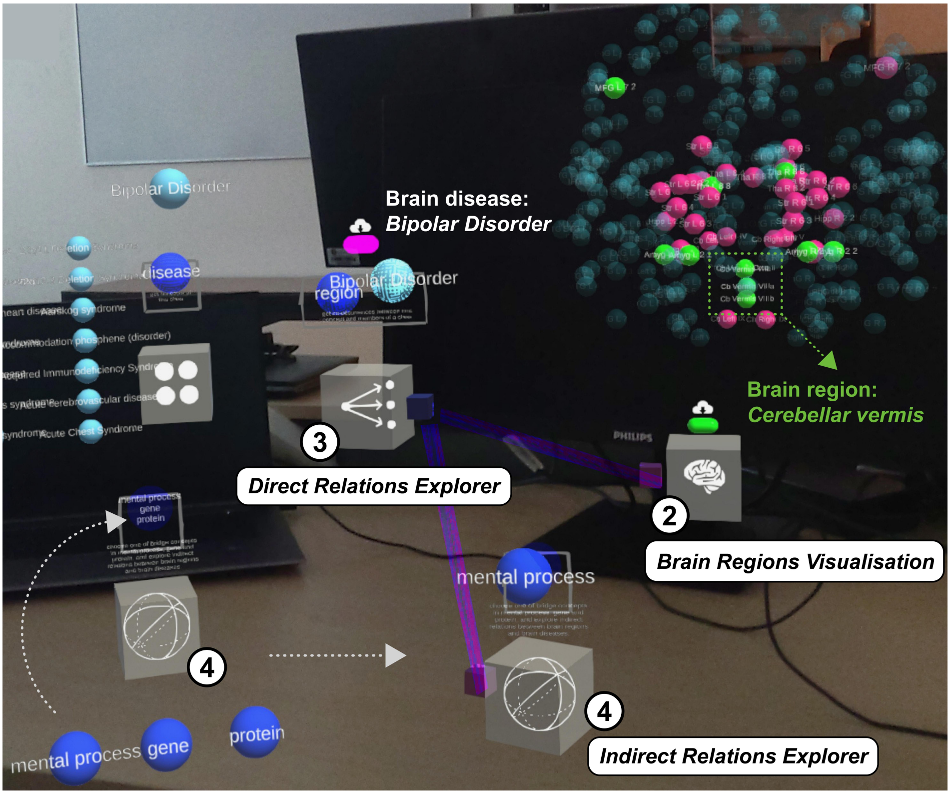

Figure 2.5: A participant selects mental processes as the intermediate topic type (④ lower left), and explores the indirect relation between the brain disease *Bipolar Disorder* (③ middle) and the brain region *Cerebellar Vermis Structure* (3 green spheres, ② lower right).
Direct relations between topics are visualised as pink spheres; indirect relations between topics are visualised as green spheres; no direct or indirect relations are visualised as light blue spheres.

2.3.2. Determining the Functional Requirement for Exploring Indirect Relations

We showed that exploring indirect relations between topics is a useful task for neuroscientists, Sec. 2.3.1, and identified two representative tasks of exploring indirect relations between brain regions and brain diseases, Sec. 2.3.1. In the existing 3D AR prototype, the *Direct Relations Explorer* allows users to explore direct relations between topics, Fig. 2.1 on p. 25. Based on the functionality implemented in previous widgets, we asked three neuroscientists (P8-10, Table 2.2 on p. 34) to discuss the functional and visualisation requirements (Sec. 2.3.3) of exploring indirect relations between topics.

- ***Selecting intermediate topics*** Due to the wide range of topics covered in neuroscience, each neuroscientist has a different focus on research topics, such as genes or mental processes. One neuroscientist focusing on neurobiology (P8, Table 2.2 on p. 34) indicated that they were interested in manually selecting different intermediate topics to explore indirect relations between brain regions and brain diseases. The corresponding functional requirement is thus: *provide a selection mechanism for intermediate topics.*

As a consequence, we extend 3D AR prototype with an input selection option for intermediate topics. Given the topics in the repository, these are mental processes, genes and proteins. The *Indirect Relations Explorer* illustrates the querying process of selecting intermediate topics to explore indirect relations between brain regions and brain diseases, Fig. 2.6 on p. 33.

2.3.3. Determining Visualisation Requirements for Exploring Indirect Relations

To display the query results of the indirect relations functionality, specified in Sec. 2.3.2, we design a visualisation for displaying indirect relations between topics. Requirements for the visualisation were discussed with three neuroscientists (P8-P10, Table 2.2 on p. 34).

- ***Single visualisation*** All three neuroscientists indicated that they would prefer to see both direct and indirect relations between brain regions and brain diseases at the same time. This allows them to see which relations have already been mentioned in the literature, and which are additional indirect relations. This visualisation requirement is thus: *display both direct and indirect relations in the same Brain Disease Topic Model and Brain Regions Visualisation.*

- ***Contrasting colours*** Neuroscientists need to infer different relations between topics by determining the difference between colours [86, 87]. To be able to distinguish the direct and indirect relations by different colours in visualisations, we select a set of contrasting colours. We choose green to display the indirectly related topics, which is easily distinguishable from the pink used for the directly related topics and sufficiently distinct from the light blue used to indicate no direct or indirect relations, Figs. 2.4 and 2.5 on pp. 30 and 31. This is not a good solution for colour-blind users but works well in the display conditions of the AR headset we use.

2.4. Implementation

We use the HoloLens 2[7] headset (a hybrid reality head-mounted display) to visualise the *Indirect Relations Explorer*. We use Unity[8] (version 2020.3.15f2) to develop the main scenarios in AR, combined with Microsoft Mixed Reality Toolkit[9] (MRTK), which is a HoloLens 2 driven package to build AR applications in Unity. We use MRTK's AR interaction toolkit, such as hand tracking and grabbing and dragging 3D objects. The new code, which extends the previous codebase, is stored as a separate branch in a coding repository[10].

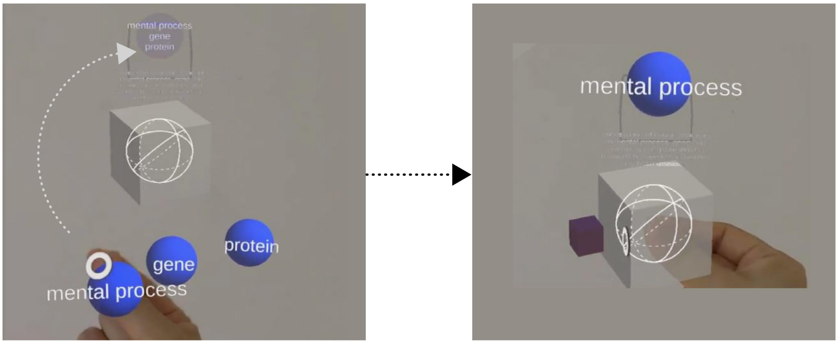

Figure 2.6: The *Indirect Relations Explorer*: Mental processes, genes or proteins can be selected as the intermediate topics for querying indirect relations between brain regions and brain diseases. Participants grab the dark blue sphere representing mental processes and drag it onto the semi-transparent placeholder above, this triggers the exploration of indirect relations between brain regions and brain diseases via mental processes.

2.5. Evaluation

We identified two representative tasks that are plausible and relevant to participants, Sec. 2.3.1. To answer the three research questions, we ask participants to reflect on the indirect relations functionality in the context of these two representative tasks. The two tasks allow us to assess indirect relations for a single brain region with multiple brain diseases as well as those for a single disease with multiple brain regions. Given the explorative nature of our study, providing extra examples would not necessarily lead to further insights.

[7]Microsoft HoloLens 2 is an AR headset https://www.microsoft.com/en-us/hololens

[8]Unity https://unity.com/

[9]Microsoft Mixed Reality Toolkit (MRTK) https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05

[10]The Unity-based code for exploring indirect relations is open-source and available on GitHub: https://github.com/DatAR-prototype/tree/IndirectRelations-Project

We describe the participants recruited in Sec. 2.5.1. We illustrate the survey procedure of the three research questions in Sec. 2.5.2. We then describe the survey questions for the individual research questions, namely,

- the suitability of the visualisation in the *Indirect Relations Explorer* for presenting indirect relations (RQ3) in Sec. 2.5.3;
- the usability of the *Indirect Relations Explorer* for exploring indirect relations (RQ2) in Sec. 2.5.4; and
- the usefulness of indirect relations for designing a useful experiment (RQ1) in Sec. 2.5.5.

2.5.1. Participants

In this study, we recruited eight participants from the neuroscience institute at the university, and one participant from an earlier study[11], giving a total of 9 participants (P8-P16, Table 2.2 on p. 34). P8 assisted in identifying two representative tasks (Sec. 2.3.1), while P8-P10 determined the functional (Sec. 2.3.2) and visualisation (Sec. 2.3.3) requirements. To obtain more continuous and targeted feedback, we retained P8-10 among the 9 participants [55].

Participants' academic qualifications, familiarity with AR and their experience of literature exploration are provided in Table 2.2 on p. 34. All nine participants are familiar with or professional users of literature exploration, have little understanding of AR, and have not previously experienced any AR equipment. All nine participants had normal or corrected-to-normal vision and were not colour-blind.

Table 2.2: Backgrounds of participants and the studies in which they participated. P1-P7 were from previous studies [56, 99]. Nine participants (P8-P16) were recruited to evaluate two representative tasks of exploring indirect relations.

Participant Numbers	**Research Topics**	**Academic Qualifications**	**Knowledge and experience in AR technology**	**Experience in Literature Exploration**	**Studies Participated in**
P1-P7	Cognitive neurosciences, neurobiology, pharmacology and psychiatry [99]	P1 and P3 are PhDs [99]	Not known	Not known	Evaluate the existing 3D AR prototype to determine the task of exploring indirect relations [56, 99] (Sec. 2.3.1)

[11] P12 is from an earlier study [99]. P8-P11 and P13-P16 are from the neuroscience institute at the university.

P8	Neurobiology	PhD student (second year)	Little knowledge of AR, no experience in AR headset	Familiar with literature exploration	Determine the task of exploring indirect relations (Sec. 2.3.1) & Define useful intermediate topics (Sec. 2.3.1) & Identify two representative tasks (Sec. 2.3.1) & Determine functional (Sec. 2.3.2) and visualisation (Sec. 2.3.3) requirements & Evaluate two representative tasks of exploring indirect relations in this work (Sec. 2.5)
P9	Neurobiology	PhD, Assistant Professor	Little knowledge of AR, no experience in AR headset	Professional in literature exploration	Determine the task of exploring indirect relations (Sec. 2.3.1) & Define useful intermediate topics (Sec. 2.3.1) & Determine functional (Sec. 2.3.2) and visualisation (Sec. 2.3.3) requirements & Evaluate two representative tasks of exploring indirect relations in this work (Sec. 2.5)
P10	Computational neuroscience	Master student	Little knowledge of AR, no experience in AR headset	Familiar with literature exploration	Determine the task of exploring indirect relations (Sec. 2.3.1) & Define useful intermediate topics (Sec. 2.3.1) & Determine functional (Sec. 2.3.2) and visualisation (Sec. 2.3.3) requirements & Evaluate two representative tasks of exploring indirect relations in this work (Sec. 2.5)
P11	Cognitive neuroscience	PhD student (first year)	None	Professional in literature exploration	Evaluate two representative tasks of exploring indirect relations in this work (Sec. 2.5)

P12, P13	Cognitive neuroscience	PhD	None	Professional in literature exploration	Evaluate two representative tasks of exploring indirect relations in this work (Sec. 2.5)
P14, P15	Neural anatomy	PhD student (first year)	Little knowledge of AR, no experience in AR headset	Familiar with literature exploration	Evaluate two representative tasks of exploring indirect relations in this work (Sec. 2.5)
P16	Neurobiology	PhD student (first year)	Little knowledge of AR, no experience in AR headset	Professional in literature exploration	Evaluate two representative tasks of exploring indirect relations in this work (Sec. 2.5)

2.5.2. Procedure

Each evaluation procedure took around 60 minutes. The evaluation sessions were conducted in October and November 2022, and January 2023. The procedure is as follows:

- We requested information from our nine expert neuroscience researchers, including their research fields, academic qualifications, familiarity with AR, and experience in literature exploration (reported in Table 2.2 on p. 34), and requested them to sign an informed consent form (Appendix A on p. 131). Since we are not interested in how age or gender influences our results, we do not collect this information [32].
- We described the goal of the session and the detailed interaction process of exploring and visualising direct and indirect relations between topics in AR (Appendix A on p. 132).
- Participants were asked to wear the AR headset (HoloLens 2) and try to grab and drag visualised 3D objects in AR.
- Participants performed two representative tasks (specified in Sec. 2.3.1) of exploring indirect relations between topics using the *Indirect Relations Explorer*. After carrying out the two tasks, participants were asked to explore the prototype and use the functionalities as they pleased.
- Participants were then asked to evaluate
  - the suitability of indirect relations visualisation (RQ3) (Sec. 2.5.3);
  - the usability of indirect relations functionality (RQ2) (Sec. 2.5.4); and
  - the usefulness of indirect relations (RQ1) (Sec. 2.5.5).

  by answering a 14-question questionnaire (listed in Figs. 2.7, 2.8 and 2.9 on pp. 38–40), using a 5-point Likert scale from strongly agree to strongly disagree.

- After responding to the survey questions, the experimenter conducted a 6-question semi-structured interview lasting approximately 9 minutes (listed in Secs. 2.5.3, 2.5.4 and 2.5.5) with each participant to gather suggestions for improvements in functionalities and visualisations.

2.5.3. Survey for the Suitability of Indirect Relations Visualisation (RQ3)

Participants assessed the visualisation of indirect relations between topics using the *Indirect Relations Explorer* in the 3D AR prototype.

- ***Suitability of Indirect Relations Visualisation (RQ3)*** We asked participants to judge the suitability of contrasting colours showing direct and indirect relations, for example, "I think the contrasting colours showing direct and indirect relations between topics are clear and suitable in AR" (Fig. 2.7 Q4 on p. 38). Participants answered four questions to evaluate the suitability of visualisations for the *Indirect Relations Explorer*, to save space, refer to Fig. 2.7 on p. 38.

- ***Interview*** Two questions were conducted with participants to suggest improvements in visualisations in the prototype. "Are there any suggestions for the colours of the spheres representing the direct and indirect relations? (Q15)" and "Are there any suggestions to improve the user interface in the prototype? (Q16)"

2.5.4. Survey for the Usability of Indirect Relations Functionality (RQ2)

Participants explored the complete process of exploring indirect relations using the *Indirect Relations Explorer* in the context of two representative tasks.

- ***Usability of Indirect Relations Functionality (RQ2)*** We asked participants to evaluate the usability of the specified functionality of exploring indirect relations, for example, "I think the *Indirect Relations Explorer* could help neuroscientists to find useful indirect relations" (Fig. 2.8 Q5 on p. 39)? Participants answered five questions to evaluate the usability of the specified functionality, to save space, refer to Fig. 2.8 on p. 39.

- ***Interview*** Two questions were conducted with participants to suggest improvements in the functionality for exploring indirect relations. "Are there any suggestions for designing improved version functionality? (Q17)" and "Are there any suggestions to improve the usability of the prototype? (Q18)"

2.5.5. Survey for the Usefulness of Indirect Relations Found (RQ1)

Participants evaluated the usefulness of indirect relations found in the context of two representative tasks.

- ***Usefulness of Indirect Relations Found (RQ1)*** We asked participants to evaluate the usefulness of indirect relations found in the context of two representative tasks, for example, "I think there is probably an indirect relation between the brain region *Hippocampus* and the brain disease *Alexander Disease*" (Fig. 2.9 Q12 on p. 40). Participants answered five questions to evaluate the usefulness of indirect relations, to save space, refer to Fig. 2.9 on p. 40.

- ***Interview*** Two questions were conducted with participants to suggest improvements in selecting intermediate topics between topics. "Are there other intermediate topics that are useful for you to explore indirect relations? (Q19)" and "Are there any factors, except intermediate topics, that influence you to judge the usefulness of identified indirect relations? (Q20)"

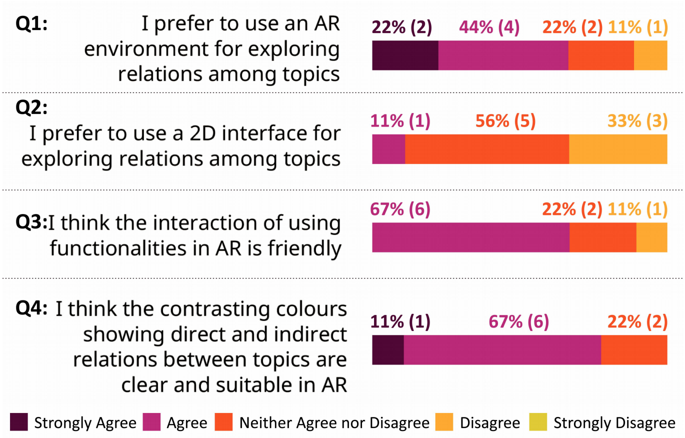

Figure 2.7: Participants' attitudes on the suitability of indirect relations visualisation (RQ3).

2.6. Results and Discussion

2.6.1. Suitability of Indirect Relations Visualisation (RQ3)

On finding the display of direct and indirect relations clear and suitable, one participant (P11, Table 2.2 on p. 34) strongly agreed, and six participants (P8, P9, P12-P15) agreed, Fig. 2.7 Q4 on p. 38. Two participants recommended exploring and evaluating different sets of colours, for example, by trying two similar colours for direct and indirect relations (P10, P16) (Q15). In addition, all nine participants suggested improving the readability of the description texts, particularly sphere labels in the *Brain Disease Topic Model* and *Brain Regions Visualisation* (P8-P16) (Q16).

Two participants recommended redesigning the labels on direct and indirect relations in the *Brain Disease Topic Model* and *Brain Regions Visualisation* (P9, P12) (Q16). In the 3D AR prototype, overlapping labels on spheres cause readability issues. Potential solutions include matching the label colours with those of the spheres and implementing a filter to display only labels related to direct and indirect relations, thus improving the readability of labels.

2.6.2. Usability of Indirect Relations Functionality (RQ2)

All nine participants found the indirect relations functionality useful for the task of exploring indirect relations between topics (four participants (P8, P9, P11, P13, Table 2.2 on p. 34) strongly agreed, five participants (P10, P12, P14-P16) agreed), Fig. 2.8 Q5 on

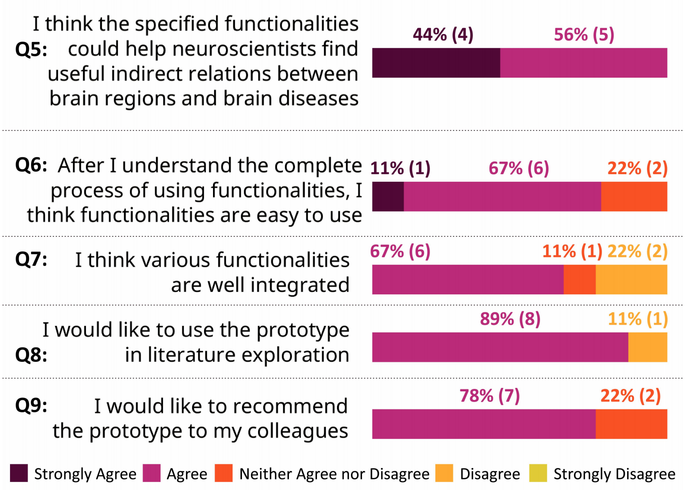

Figure 2.8: Participants' attitudes on the usability of the indirect relations functionality (RQ2).

p. 39. Five participants found interacting with the specified functionality by grabbing and dragging 3D objects was user-friendly, even though they were first-time users of AR (P8, P9, P11, P12, P14). While seven participants thought the *Indirect Relations Explorer* was not difficult to use after understanding the detailed process of querying and visualising indirect relations, due to their lack of experience in using AR equipment, they still required help from the experimenter (P8-P15). Six participants agreed with the method of constructing and integrating the *Indirect Relations Explorer* in the 3D AR prototype (P8, P9, P11, P12, P14, P15), however, two participants stated it is too complex to understand, for example, relevant widgets always need to be selected and connected (P10, P16). Eight participants expressed interest in using interactive functionalities in the 3D AR prototype to explore potentially useful indirect relations between topics (P8, P9, P11-P16).

Participants had difficulty using functionalities in the prototype because the interface did not provide sufficient guidance cues for first-time users. A potential solution could be the development of a tutorial to guide users in operating the functionalities for exploring indirect relations between topics, thereby enhancing the usefulness of the 3D AR prototype for daily literature work (P10, P16) (Q18).

2.6.3. Usefulness of Indirect Relations Found (RQ1)

All nine participants stated that exploring indirect relations between topics is useful for designing a potentially useful experiment (five participants (P8-P10, P14, P15, Table 2.2 on p. 34) strongly agreed, and four participants (P11-P13, P16) agreed), Fig. 2.9 Q10 on p. 40. Mental processes, genes or proteins were recognised as useful intermediate topics for exploring indirect relations between brain regions and brain diseases (P8-P16), Fig. 2.9 Q11 on p. 40. Additionally, two neuroscientists proposed that cells could also

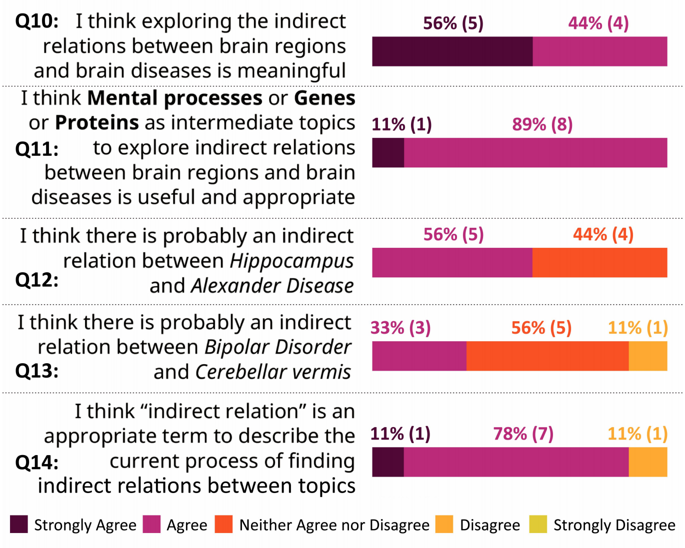

Figure 2.9: Participants' attitudes on the usefulness of indirect relations (RQ1).

be a useful intermediate topic (P10, P16) (Q19). Eight participants agreed that "Indirect Relations" is a suitable term to describe relations between topics through a third topic (P8, P9, P11-P16), while one participant suggested that "Secondary Associated Network" is a better term for this (P10), Fig. 2.9 Q14 on p. 40.

In the task of exploring an indirect relation between *Hippocampus* and *Alexander Disease* via genes, five participants stated that the indirect relation found was useful for them to identify an experiment that would potentially make a useful contribution (P8, P9, P11, P12, P15), but four thought it was hard to make a judgement (P10, P13, P14, P16 neither agreed nor disagreed), Fig. 2.9 Q12 on p. 40. In the task of exploring an indirect relation between *Bipolar Disorder* and *Cerebellar Vermis Structure* via mental processes, three participants proposed the identified indirect relation was useful for them (P9, P11, P15), but five were not entirely convinced (P8, P10, P12-P14), and one disagreed (P16), Fig. 2.9 Q13 on p. 40. Six participants thought that the indirect results found could only provide inspiration for further literature review, rather than evidence to ensure the experiment would make a useful contribution (P8, P10, P12-P14, P16). They thought it was difficult to judge to what extent the indirect relations were useful to design a useful experiment. Three neuroscientists proposed that indirect relations found were useful for junior neuroscientists to understand less obvious relations among topics (P9, P12, P13).

Despite the generally positive feedback for the usefulness of indirect relations found, there were suggestions for improvements, primarily concerning the provenance of the indirect relations found. Neuroscientists expressed a desire for more detailed information about indirect relations, such as access to the publications containing the co-occurrences (Q20). A potential solution could provide a specific gene or protein, which is related to neuroscientists' research fields, as an intermediate topic for manually selecting and querying indirect relations between brain regions and brain diseases. P9 suggested the use of publication dates to identify trends of (first indirect and then direct) relations between topics (Q20). By illustrating that previous indirect relations might have evolved into present direct relations, the provenance of current indirect relations will be improved.

2.7. Conclusions and Future Work

Our evaluation of the implemented visualisation and functionality of exploring indirect relations indicates that participants thought the visualisation was suitable for displaying direct and indirect relations (Sec. 2.6.1), the functionality was useful for exploring indirect relations between topics (Sec. 2.6.2), and the indirect relations found are useful for designing a useful experiment (Sec. 2.6.3). One direction to explore is to understand how we can provide the provenance of the indirect relations found, for example, by providing the specific topics, such as genes, that contribute to the indirect relation and by providing direct access to the publications that are the sources of the two direct relations (Chapter 4). Another direction to explore is using publication dates to identify trends in (first indirect and then direct) relations between topics, which could improve the usefulness of current indirect relations (Chapter 3). This work contributes to finding and visualising indirect relations in 3D AR, which can support neuroscientists in their complex task of designing a potentially useful experiment.

3

Informing the Design of a Neuroscience Experiment by Exploring the Evolution of Indirect to Direct Relations Between Topics in Augmented Reality Using a Timeline

Topic-based literature exploration in Augmented Reality (AR) has successfully been used to support neuroscientists in obtaining an overview of the literature, which is a tedious and time-consuming task needed to design expensive neuroscientific experiments. In particular, topic-based literature exploration in 3D AR allows users to analyse large numbers of publications simultaneously, providing a "bird's-eye" view of relations between neuroscience topics. A direct relation exists when two topics, such as a brain region and a brain disease, co-occur in the same sentence of a publication's title or abstract. An indirect relation occurs when no direct relation exists, but each topic co-occurs with a common intermediate topic, such as a gene. Neuroscientists have suggested that understanding how historical indirect relations evolved into current direct ones (1) increases confidence in the usefulness of current indirect relations and (2) informs the design of their experiments. Using a user-centred design approach, we establish if such a timeline-based exploration is indeed useful. We interview neuroscientists to verify the usefulness of exploring the evolution of indirect to direct relations, identify required functionality, and design and implement a timeline visualisation in 3D AR. Twelve neuroscientists carried out two representative tasks using the implemented functionality and timeline visualisation. Responses to 13 semi-structured interview questions and 10 Likert-scale questions indicated that (a) the timeline visualisation is *suitable* for presenting the implemented functionality, (b) the functionality is *useful* for exploring the evolution, and (c) exploring the evolution of indirect to direct relations is *useful* for designing experiments.

Chapter contribution: This chapter investigates how the user task of exploring the evolution of historical indirect relations to current direct relations can inform the design of a neuroscience experiment (MRQ2), specifies an appropriate timeline functionality to support this task, and designs and implements the corresponding 3D AR visualisation to enable this functionality.

3.1. Introduction

Neuroscientists need to analyse large numbers of neuroscience publications to inform the design of an experiment. However, doing this manually is both challenging and time-consuming. Exploring relations between neuroscience topics, such as brain regions and diseases, rather than focusing on individual publications, offers a useful approach to analyse many publications simultaneously [97, 99]. This approach provides an overview of relations between topics. DatAR, a 3D Augmented Reality (AR) prototype, has shown that visualising the spatial relationships of brain regions and exploring their direct and indirect relations with brain diseases in AR can be helpful in this process [99, 109] (Sec. 3.2.4).

- A *direct* relation exists between two topics when both occur in the same sentence of a publication's title or abstract. For example, there is a direct relation between the brain disease *Depressive Disorder* and the brain region *Cerebellar Vermis Structure* since they both occur in the sentence "The *depression* and anxiety severity of bipolar depression patients may be associated with the consistency activity of left insula, right cerebellum, and *cerebellar vermis* related area, fusiform gyrus" in the publication with PubMed[1] ID 31764799[2], Fig. 3.2 on p. 46.
- An *indirect* relation occurs when no direct relation exists between two topics, but both have a direct relation with the same intermediate topic. For example, a direct relation between *Depressive Disorder* and *LITAF Gene* in July 2017, PMID: 28698888[3], Fig. 3.1 (a) on p. 45, and a direct relation between *LITAF Gene* and *Cerebellar Vermis Structure* in August 2017, PMID: 28515286[4], Fig. 3.1 (b) on p. 45. Based on these two direct relations, we conclude that there is an indirect relation between the brain disease *Depressive Disorder* and the brain region *Cerebellar Vermis Structure* via *LITAF Gene*.

In previous work on exploring direct and indirect relations using DatAR [109], a participating neuroscientist suggested using publication dates to identify when a historical indirect relation between two topics becomes a direct relation in more recent

[1] PubMed is a free resource supporting the search and retrieval of biomedical and life sciences literature https://pubmed.ncbi.nlm.nih.gov/. Each publication in PubMed has a unique identifier, the PMID.
[2] DOI: 10.1097/MD.0000000000017962 (PMID: 31764799)
[3] DOI: 10.1523/ENEURO.0115-17.2017 (PMID: 28698888)
[4] DOI: 10.1152/jn.00209.2017 (PMID: 28515286)

publications. This would help them build confidence in the usefulness of current indirect relations. Identifying current indirect relations in the literature could facilitate the design of a potentially useful experiment since there is weak evidence of a potential connection, but no single publication has yet demonstrated that this is indeed the case. For example, there are no direct or indirect relations between the brain disease *Depressive Disorder* and the brain region *Cerebellar Vermis Structure* from January 2010 to June 2017 in the subset of the PubMed corpus that we use. There is an indirect relation between the brain disease *Depressive Disorder* and the brain region *Cerebellar Vermis Structure* via *LITAF Gene* between July and August 2017, Fig. 3.1 on p. 45. A direct relation between *Depressive Disorder* and *Cerebellar Vermis Structure* was first published in November 2019, PMID: 31764799[5], Fig. 3.2 on p. 46.

Our main research question (RQ1) is: *How does exploring the evolution of historical indirect to current direct relations contribute to designing a useful experiment?* Using a user-centred design approach is useful for developing functionality and visualisations that align with the requirements of neuroscientists [27]. We first consulted two neurologists to advise us on whether exploring the evolution of indirect to direct relations is a potentially useful user task (Sec. 3.3.1). The same two neurologists were then asked to determine two representative tasks for our later evaluation (Sec. 3.3.2).

Having established that exploring the evolution of indirect to direct relations is a potentially useful user task, we need to specify which functionality and visualisation are required to support this. Using input from interviews with the same two neurologists, we identified two functional requirements (Sec. 3.3.3). Further interviews with the same two neurologists helped to identify the timeline visualisation requirements (Sec. 3.3.4). We explain our design decisions to meet these functional and visualisation requirements (Sec. 3.4) and implement the *Timeline Selector*, Fig. 3.3 on p. 47, in the DatAR prototype (Sec. 3.5).

Having designed and implemented the timeline visualisation, we need to investigate the suitability of the implemented visualisation for presenting the evolution of indirect to direct relations. Our third research question (RQ3) is: *To what extent is the visualisation suitable for the specified functionality in presenting the evolution of indirect to direct relations?* We describe the evaluation procedure undertaken by twelve

[5]DOI: 10.1097/MD.0000000000017962 (PMID: 31764799)

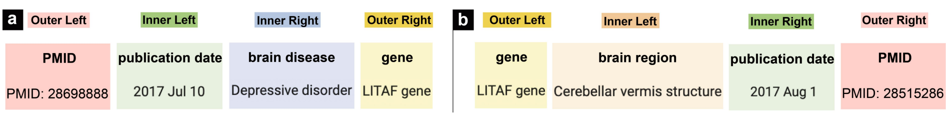

Figure 3.1: (a) The brain disease *Depressive Disorder* (inner right column) has a direct relation with *LITAF Gene* (outer right column) in July 2017 (inner left column), PMID: 28698888 (outer left column).
(b) *LITAF Gene* (outer left column) has a direct relation with the brain region *Cerebellar Vermis Structure* (inner left column) in August 2017 (inner right column), PMID: 28515286 (outer right column).

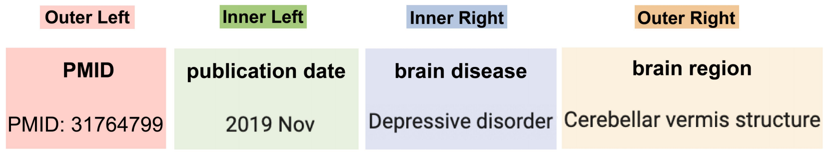

Figure 3.2: The brain disease *Depressive Disorder* (inner right column) has a direct relation with the brain region *Cerebellar Vermis Structure* (outer right column) in November 2019 (inner left column), PMID: 31764799 (outer Left column).

neuroscientists using the semi-structured interviews and the Likert-scale questionnaires to assess visualisation, functionality, and the user task (Sec. 3.6). We evaluated the suitability of the visualisation using two representative tasks with twelve neuroscientists (Sec. 3.7.1).

Having evaluated the implemented visualisation, we investigate the usability of the implemented functionality for the user task. Our second research question (RQ2) is: *To what extent is the implemented functionality useful for exploring the evolution of historical indirect to current direct relations?* The twelve neuroscientists who evaluated RQ3 were also asked to assess the usability of the functionality (Sec. 3.7.2).

Our final step is to evaluate the usefulness of the user task of exploring the evolution of indirect to direct relations for designing a useful experiment (RQ1). The same twelve neuroscientists were consulted and proposed that exploring the evolution of historical indirect to current direct relations is useful for understanding the current indirect relations to design a useful experiment (Sec. 3.7.3).

The contributions of this work are (i) the useful user task of exploring the evolution of historical indirect to current direct relations between neuroscience topics (answer to RQ1); (ii) functionality that enables the user task (answer to RQ2); and (iii) visualisation that enables functionality (answer to RQ3).

3.2. Related Work

We explore the potential of topic-based literature exploration for identifying useful experiments from a large amount of literature, Sec. 3.2.1. We explain how analysing the evolution of relations between topics can uncover the development of historical findings over time, Sec. 3.2.2. We explore how 3D AR can enhance the visualisation of 3D brain regions, Sec. 3.2.3. We then describe DatAR, an interactive 3D AR prototype, for exploring direct and indirect relations between brain regions and diseases, Sec. 3.2.4.

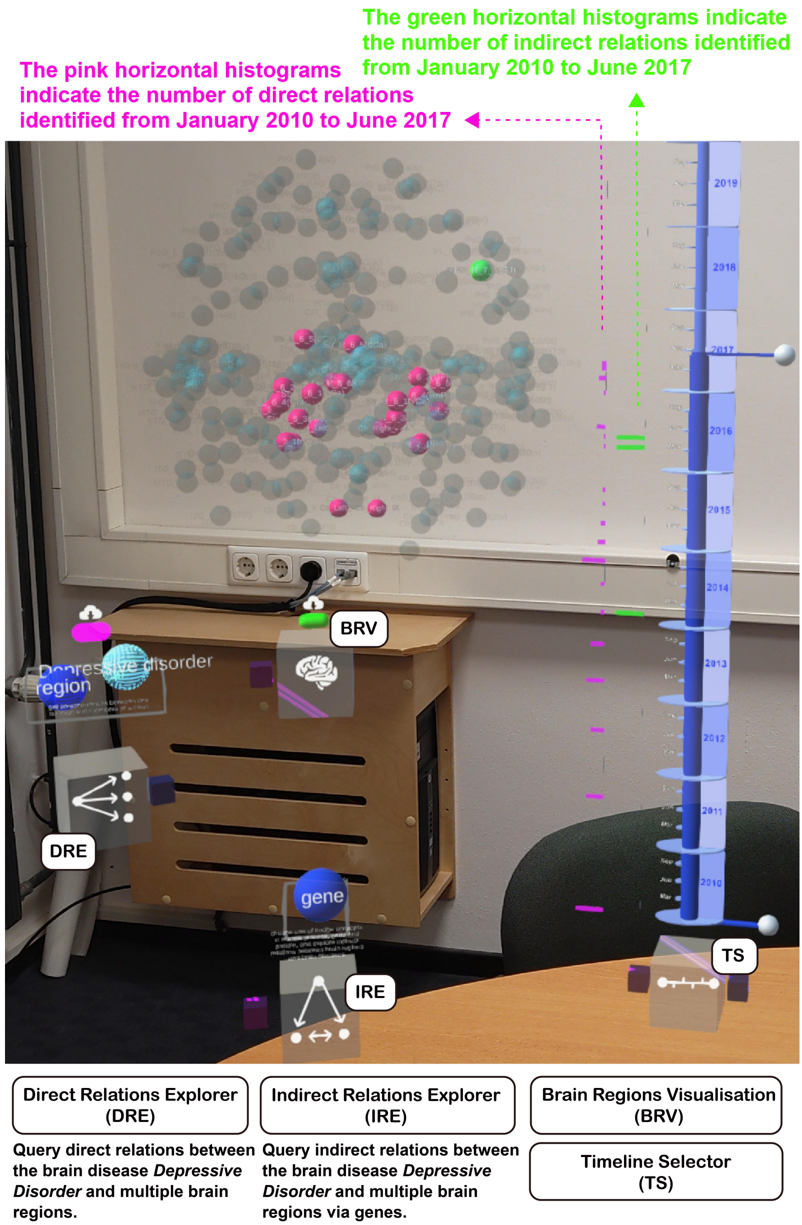

Figure 3.3: The *Timeline Selector*: A participant can grab and drag the top and bottom handles to select a range of publication dates. The selected range from January 2010 to June 2017 is visualised as a dark blue bar. The horizontal histograms to the left of the timeline indicate the number of identified indirect (green rectangles) and direct (pink rectangles) relations between the brain disease *Depressive Disorder* and all brain regions during the selected publication date range.

3.2.1. Topic-based Literature Exploration

Identifying potentially useful experiments builds on understanding the existing literature [50, 98]. Typically, researchers explore literature by scanning and reading individual publications. However, topic-based literature exploration allows a large amount of literature to be explored on the basis of topics, such as brain diseases and brain regions, rather than articles [18, 46]. The potential of topic-based literature exploration in generating useful hypotheses has been demonstrated in biomedical tasks. For example, it has been used to identify new relations between genes and diseases, such as revealing that *c-Myc* expression may be a prognostic marker in *diffuse large B-cell lymphoma* [107]. In our study, we investigate complex relations among neuroscience topics, such as brain regions, brain diseases, and genes, to better understand, for example, which brain regions might be affected by the brain disease *Depressive Disorder*.

3.2.2. The Evolution of Relations between Topics

Analysing the evolution of topics is useful for understanding existing knowledge, as it reveals how earlier findings have been developed and expanded over time [64, 70, 76]. Reference Publication Year Spectroscopy (RPYS) is a useful method for exploring the evolution of relations between topics [41]. For example, RPYS can uncover hidden relations between topics over time. By analysing publications from 1999 to 2012, researchers identified cancer drugs by uncovering indirect relationships within the drug $\rightarrow$ Gene1 $\rightarrow$ Gene2 $\rightarrow$ Cancer pathway [84, 113]. Swanson's discoveries revealed 11 hidden indirect relations from 1983 to 1997, such as *Type A personality*, *vascular reactivity* and *calcium blockers*. These indirect relations helped establish the useful connection between "Magnesium Deficiency" and "Migraine" for new drug research [66].

In a previous study [109], we explored direct and indirect relations between brain regions and brain diseases. Inspired by the RPYS method, we enhance our understanding of the usefulness of current indirect relations by using publication dates to trace the evolution of historical indirect relations to current direct ones.

3.2.3. 3D Visualisations in Augmented Reality

The spatial orientation provided by 3D visualisations facilitates a more intuitive understanding of complex 3D structures [7, 30, 84]. Specifically, 3D visualisations enable researchers to examine 3D structures - in our case, brain regions - helping neuroscientists to identify which brain regions are affected by the same diseases [109]. The NeuroCave application uses immersive analytics to support neuroscientists in exploring networks of brain regions in Virtual Reality (VR) [52]. By using a VR headset, researchers can interact with 3D representations of brain regions. This demonstrates a preference for the VR environment over traditional 2D desktop setups when presenting 3D data [52]. Augmented Reality (AR) provides researchers with an immersive experience that integrates the real world during their experiments [16, 115]. This immersion in the AR environment enhances memory recall more effectively than both the VR environment and 2D desktop, particularly in terms of the spatial and ordinal positioning of 3D brain visualisations [44].

In our study, we support users to explore neuroscience topics, presenting 3D brain regions and their relations with brain diseases within a 3D AR environment, using colour to differentiate different types of relations (see Sec. 3.2.4).

3.2.4. DatAR Prototype

DatAR, an interactive 3D AR prototype, has been shown to be useful for neuroscientists to explore relations between topics [99, 109]. We first describe the Knowledge Graphs of Brain Science (KGBS), which store the topic co-occurrences used in the DatAR prototype, Sec. 3.2.4. We then illustrate the DatAR prototype and its constituent widgets that allow users to explore direct and indirect relations between topics, Sec. 3.2.4.

The Knowledge Graphs of Brain Science

We use the Knowledge Graphs of Brain Science[6] (KGBS), which contains an analysis of sentences found in the titles and abstracts of 414,224 neuroscience publications in PubMed[7] from January 2010 to February 2022, to query co-occurrences between topics. Topics include brain regions, brain diseases, genes, proteins, mental processes and symptoms. The KGBS contains co-occurrences between two topics in the corpus, along with the dates of the publications in which they were published and the PubMed identifiers (PMID) for each publication. A co-occurrence between two topics indicates a direct relation between them. For example, an extract showing publication dates and PMIDs of direct relations between the brain disease *Depressive Disorder* and genes is provided in Fig. 3.4 on p. 49.

Widgets for Exploring Direct and Indirect Relations between Topics

Widgets implemented in the DatAR prototype provide selection, querying, and visualisation mechanisms that allow users to explore neuroscience literature via direct relations stored in the Knowledge Graphs of Brain Science (Sec. 3.2.4). For completeness, we describe the existing widgets in the DatAR prototype that provide most of the functionality needed for this study:

[6] To access the Knowledge Graphs of Brain Science at https://kgbs-sparql.project.cwi.nl/, please contact <xuboyudesign@outlook.com>.

[7] 414,224 neuroscience publications in PubMed https://pubmed.ncbi.nlm.nih.gov/ as of February 3, 2022.

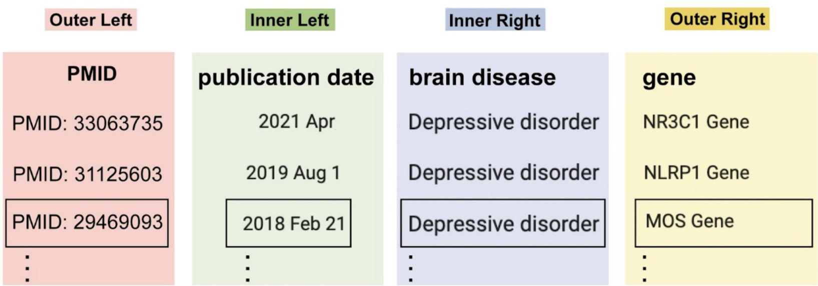

Figure 3.4: The brain disease *Depressive Disorder* (inner right column) has a direct relation with *MOS Gene* (outer right column) in February 2018 (inner left column), as indicated by the publication's PMID: 29469093 (outer left column).

- **The *Brain Regions Visualisation***, Figs. 3.5 and 3.6 on pp. 51 and 52, presents a display of 274 brain regions sourced from the Scalable Brain Atlas[8] (SBA) [8]. The position of each sphere is determined based on the 3D coordinates of brain regions taken from SBA [99].
- **The *Brain Disease Topic Model***, Fig. 3.7 on p. 53, provides 3D visualisations of 68 brain diseases from the KGBS [99]. Their spatial arrangement is determined by all direct relations between them and all other topics in the repository; the closer they are in 3D, the larger the number of common connections with the other topics in the repository.
- **The *Direct Relations Explorer***, Figs. 3.5, 3.6 and 3.7 on pp. 51–53, is used to query the direct relations between a single brain region and multiple brain diseases or a single brain disease and multiple brain regions in the KGBS.
- **The *Indirect Relations Explorer***, Figs. 3.5 and 3.7 on pp. 51 and 53, is used to select an intermediate topic, such as genes or mental processes, to query indirect relations between brain regions and brain diseases in the KGBS.
- **The *Direct Relations Filter***, Fig. 3.6 on p. 52, is used to filter the number of co-occurrences between a single brain region and multiple brain diseases, or between a single brain disease and multiple brain regions in the KGBS. Users can adjust the upper and lower handles to filter for fewer or more direct relations.

Previous versions of the DatAR prototype [99, 109] have not supported the exploration of the evolution of indirect to direct relations between topics, nor have illustrated the visualisation of topic evolution in the *Brain Disease Topic Model* and *Brain Regions Visualisation*. In this study, we implement the additional functionality and a timeline visualisation to assist users in exploring and visualising how indirect relations evolve into direct relations between a selected brain region or disease and all other brain diseases or brain regions.

3.3. Method

We follow a user-centred design approach to design support for an identified user task [17, 26, 27, 75]. Neuroscience researchers were consulted at every stage of our study. We first establish that exploring the evolution of indirect relations between two neuroscience topics to a direct relation is a useful task for neuroscience researchers, Sec. 3.3.1. To evaluate our corresponding functionality and visualisation for supporting the evolution of indirect to direct relations, we need to identify representative tasks that are both useful for neuroscientists and suitable for our evaluation purposes, Sec. 3.3.2. We determine the functional requirements needed to implement the representative tasks, including selecting a timeline, Sec. 3.3.3. Based on the functionality, we determine visualisation requirements for presenting the evolution of indirect to direct relations, Sec. 3.3.4. We discuss the design decisions to meet these functional and visualisation requirements, Sec. 3.4.

[8]Scalable Brain Atlas (SBA): https://scalablebrainatlas.incf.org/index.php

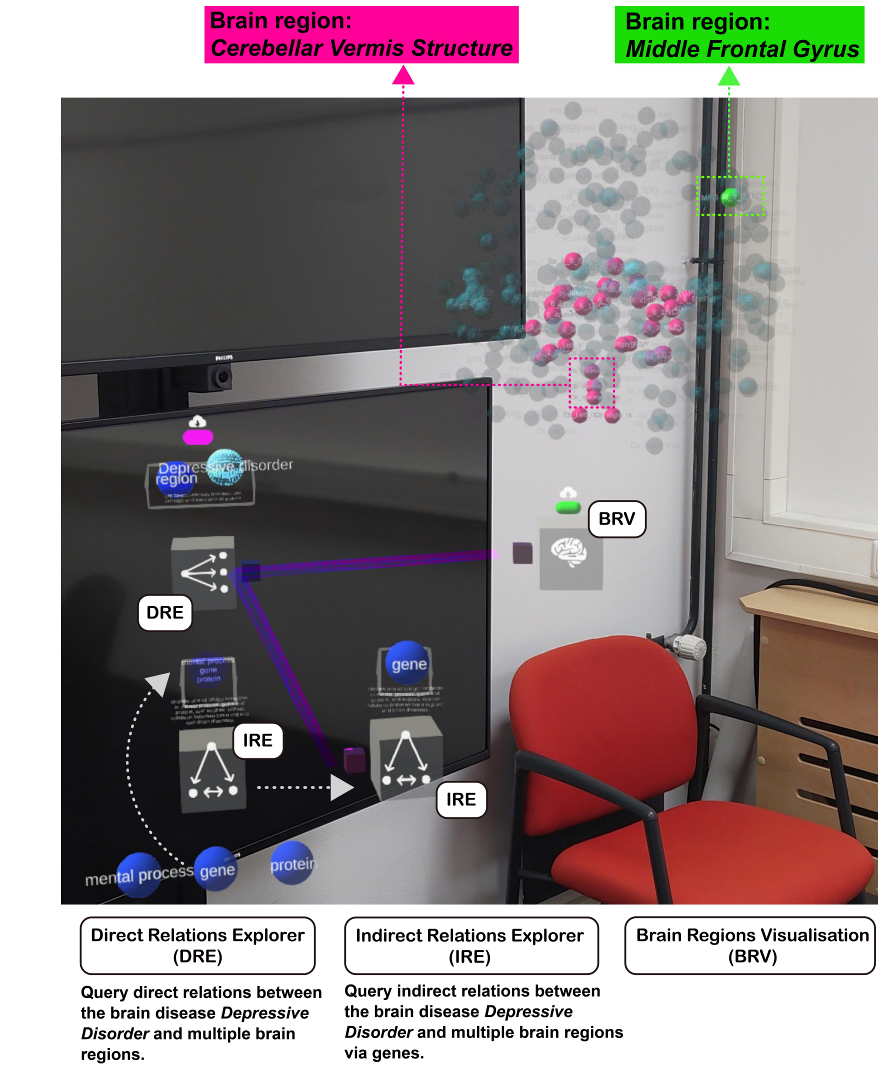

Figure 3.5: A participant explores the direct relations between the brain disease *Depressive Disorder* and all brain regions. Thirty-three direct relations (pink spheres) are found, including the brain region *Cerebellar Vermis Structure*. Genes are selected as the intermediate topic to identify the indirect relations between the brain disease *Depressive Disorder* and all brain regions. One indirect relation (green sphere) between *Depressive Disorder* and the brain region *Middle Frontal Gyrus* is found. Brain regions with no direct or indirect relations to *Depressive Disorder* are visualised as light blue spheres.

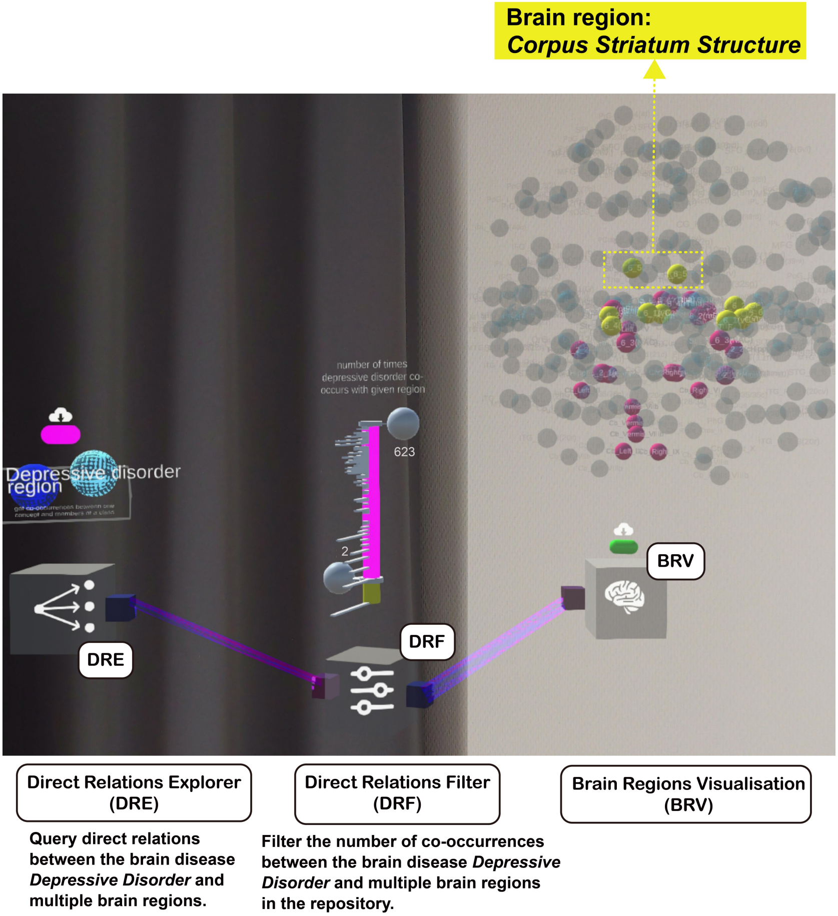

Figure 3.6: A participant uses the *Direct Relations Filter* to filter the co-occurrences between the brain disease *Depressive Disorder* and all brain regions. Ten brain regions co-occur with *Depressive Disorder* either once or twice, including the brain region *Corpus Striatum Structure* (yellow spheres).

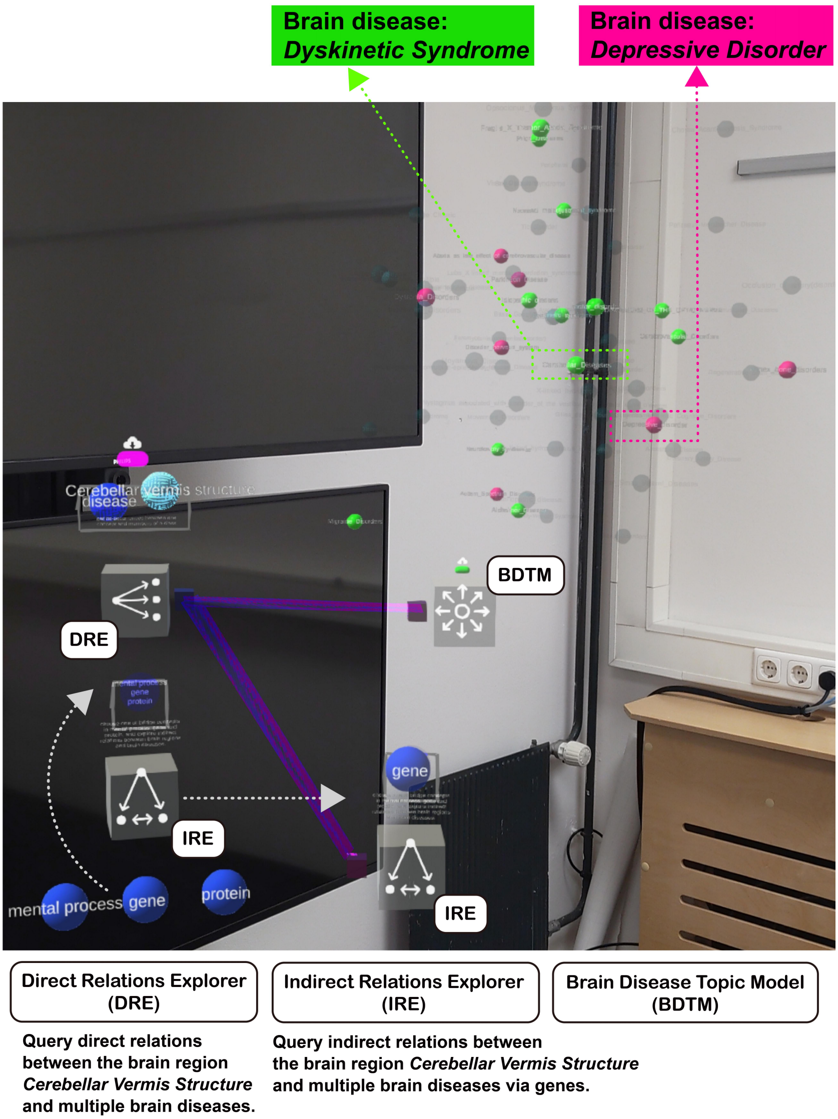

Figure 3.7: A participant explores the direct relations between the brain region *Cerebellar Vermis Structure* and all brain diseases. Seven direct relations (pink spheres) are found, including the disease *Depressive Disorder*. Genes are selected as the intermediate topic to explore the indirect relations between the brain region *Cerebellar Vermis Structure* and all brain diseases. Twelve indirect relations (green spheres) are found, including the disease *Dyskinetic Syndrome*.

Based on the user-centred design approach, we want to gather feedback through evaluation with neuroscientists to provide detailed insights. We use a qualitative measure of semi-structured interviews [69] to evaluate functionality and visualisation and to answer three research questions. The qualitative evaluation focuses on three metrics: *Suitability* [20], *Usability* [36] and *Usefulness* [10].

- ***Suitability*** focuses on evaluating the designed visualisation that supports presenting the evolution of indirect to direct relations in the AR environment (RQ3);
- ***Usability*** encompasses the extent to which the implemented functionality aligns with human-computer interaction requirements for exploring the evolution of indirect to direct relations (RQ2);
- ***Usefulness*** measures the success of exploring the evolution of indirect to direct relations in helping neuroscientists understand current indirect relations for designing useful experiments (RQ1).

To facilitate a comparative analysis of participants' perceptions, we also employ a 7-point Likert scale [9] to quantify responses to ten questions regarding the visualisation (RQ3), functionality (RQ2), and the user task (RQ1). To analyse the results, we use the median, as it provides a clearer representation of "typical user" behaviour [105].

Our aim is to understand the suitability of our designed visualisation and the usability of the corresponding functionality for neuroscience researchers, emphasising the practical application of the visualisation and functionality in the context of topic-based literature exploration. Exploratory studies for problem discovery typically involve 5 to 20 participants [89]. Recruiting 5 to 10 participants captures the most common issues and insights, with diminishing returns on new insights beyond 15 [89]. For this study, we recruited 12 participants.

3.3.1. Establishing the Usefulness of the Evolution of Indirect to Direct Relations between Topics

A neuroscientist participating in a previous study [109] suggested that they could better understand how current indirect relations can assist in designing useful experiments by exploring how historical indirect relations have evolved into current direct ones. In July 2023, we interviewed two neurologists (P1 and P2, Table 3.1 on p. 66) to determine whether the task of exploring the evolution of indirect to direct relations is useful.

P1, a neurologist researching the brain disease *Depressive Disorder*, wanted to understand which brain regions had direct relations with *Depressive Disorder* over the past five years. P1 thought these brain regions represented the most recent topics, potentially influencing funding opportunities. The relations between brain regions and brain diseases are continuously evolving, with newly discovered brain regions changing neuroscientists' understanding of disease mechanisms and treatment protocols [111]. P1 thought using publications older than five years may lead to treatment protocols that are not optimal for diseases. P2, a neurologist researching the brain disease *Parkinson Disease*, suggested that exploring how historical indirect relations have evolved into current direct ones could help build confidence in the usefulness of current indirect relations for designing useful experiments. Both P1 and P2 suggested that identifying indirect relations in the literature that have not yet been established as direct would be useful since these provide a small amount of evidence for designing useful experiments.

3.3.2. Determining Two Representative Tasks for Evaluation

To evaluate the functionality and visualisation for exploring the evolution of indirect to direct relations, we discussed with P1 and P2, Table 3.1 on p. 66, to find representative tasks that would be suitable. We selected two representative tasks for evaluation: one for exploring the evolution of indirect to direct relations between a specific brain disease and all brain regions, and another for exploring the evolution of indirect to direct relations between a specific brain region and all brain diseases. The two tasks should be useful within the neuroscience field, reflecting the types of inquiries that neuroscientists engage with. Additionally, the two tasks should be designed to be neither overly simple nor overly complex, ensuring they sufficiently test our implemented visualisation and functionality, and that participants are not presented with an excessive number of diseases or regions, as this may overwhelm them and hinder their ability to find useful information.

- **Representative Task 1: Exploring the evolution of indirect to direct relations between the brain disease *Depressive Disorder* and all brain regions**

  P1, who investigates the brain disease *Depressive Disorder*, suggested exploring the brain regions related to *Depressive Disorder* to better understand the current discoveries. The first indirect relation found between *Depressive Disorder* and the brain region *Cerebellar Vermis Structure* dates between July and August 2017 (described in Sec. 3.1), Fig. 3.9 on p. 57. This indirect relation evolved to a direct one in November 2019, Fig. 3.10[9] on p. 58.

- **Representative Task 2: Exploring the evolution of indirect to direct relations between the brain region *Cerebellar Vermis Structure* and all brain diseases**

  P2, who investigates the brain disease *Parkinson Disease*, knows that the brain region *Cerebellar Vermis Structure* has a direct relation to *Parkinson Disease*. P2 wants to explore the publication date of the evolution of the indirect relation between *Parkinson Disease* and *Cerebellar Vermis Structure* to a direct one, in order to assess the recency of this discovery. The first indirect relation between *Parkinson Disease* and *Cerebellar Vermis Structure* via *LITAF gene* was found between August and September 2017[10], Fig. 3.12 on p. 60, and this indirect relation evolved to a direct one in October 2020[11], Fig. 3.13 on p. 61.

[9] See also the video: https://youtu.be/VPAbhcqDTlk

[10] The direct relation between *Parkinson Disease* and *LITAF gene* in DOI: `10.1155/2017/5472752` (PMID: 29056964); The direct relation between *LITAF gene* and *Cerebellar Vermis Structure* in DOI: 10.1152/jn.00209.2017 (PMID: 28515286)

[11] DOI: 10.1371/journal.pone.0240998 (PMID: 33112886)

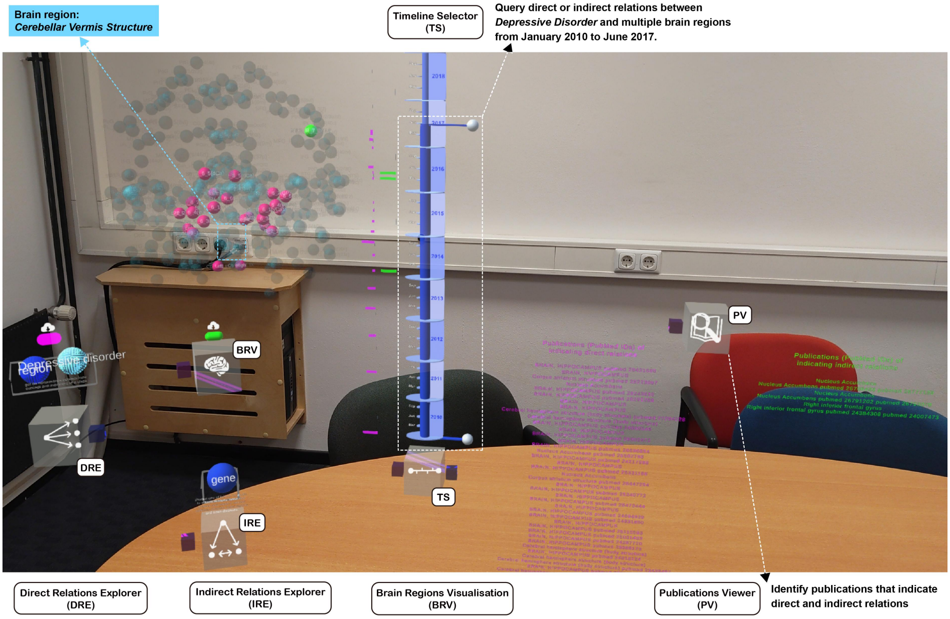

Figure 3.8: The brain disease *Depressive Disorder* has no direct or indirect relation with the brain region *Cerebellar Vermis Structure* in any papers published from January 2010 to June 2017 (TS, middle quadrant). The three spheres representing *Cerebellar Vermis Structure* are light blue (BRV, upper left quadrant).

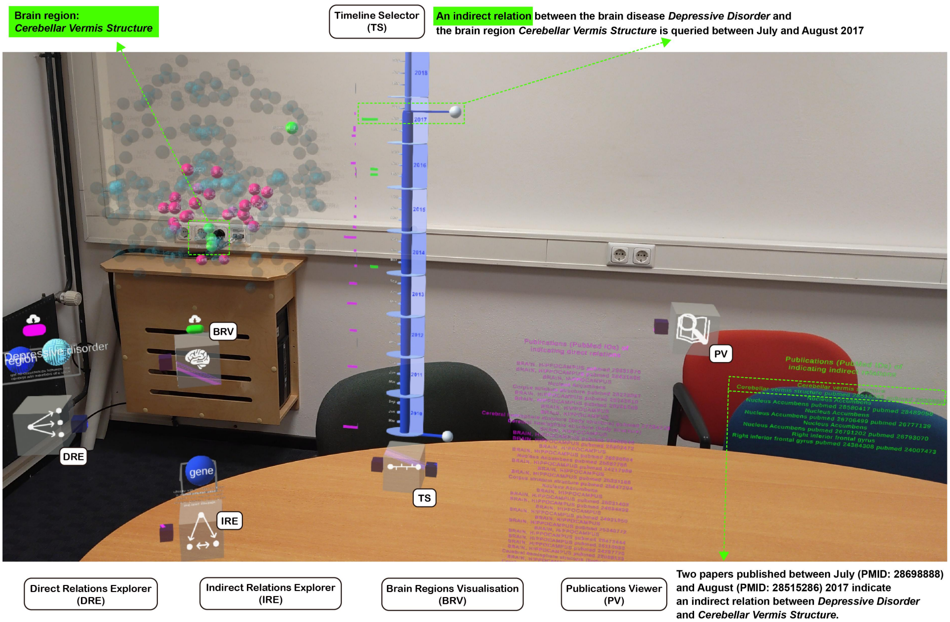

Figure 3.9: The brain disease *Depressive Disorder* has an indirect relation with the brain region *Cerebellar Vermis Structure* in two papers published between July and August 2017 (TS, middle quadrant). The three spheres representing *Cerebellar Vermis Structure* have turned green (BRV, upper left quadrant).

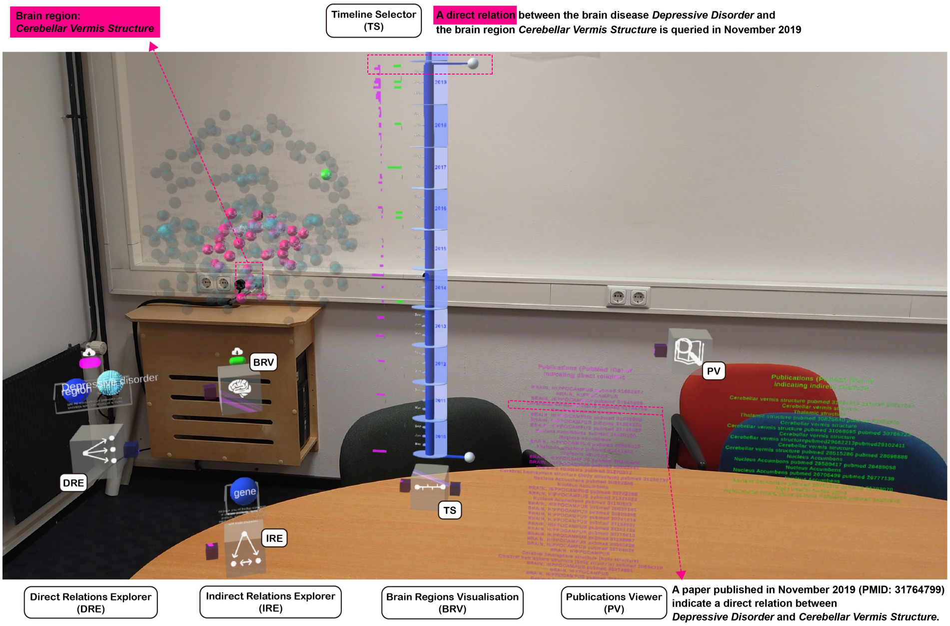

Figure 3.10: The first time a paper is published that includes a direct relation between *Depressive Disorder* and *Cerebellar Vermis Structure* is in November 2019 (TS, middle quadrant). The three spheres representing *Cerebellar Vermis Structure* have turned pink (BRV, upper left quadrant).

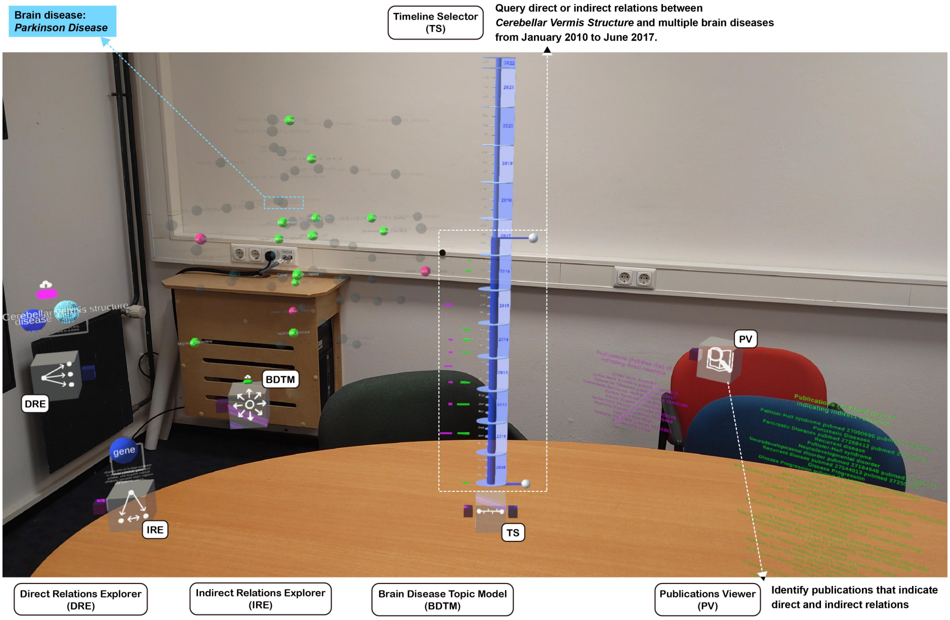

Figure 3.11: The brain region *Cerebellar Vermis Structure* has no direct or indirect relation with the brain disease *Parkinson Disease* in any papers published from January 2010 to June 2017 (TS, middle quadrant). The sphere representing *Parkinson Disease* is light blue (BDTM, upper left quadrant).

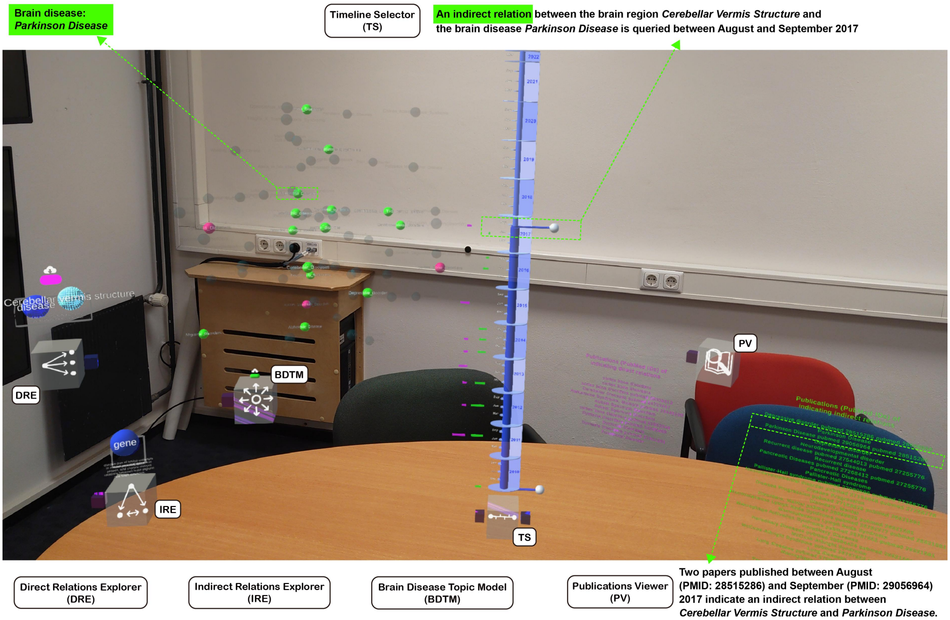

Figure 3.12: The brain region *Cerebellar Vermis Structure* has an indirect relation with the brain disease *Parkinson Disease* in two papers published between August and September 2017 (TS, middle quadrant). The sphere representing *Parkinson Disease* has turned green (BDTM, upper left quadrant).

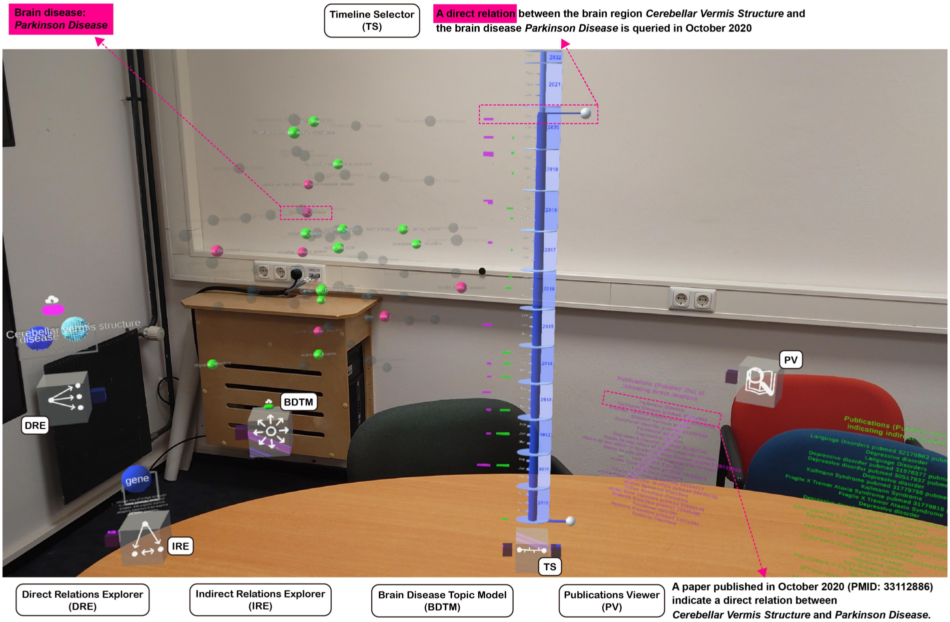

Figure 3.13: The first time a paper is published that includes a direct relation between *Cerebellar Vermis Structure* and *Parkinson Disease* is in October 2020 (TS, middle quadrant). The sphere representing the *Parkinson Disease* has turned pink (BDTM, upper left quadrant).

3.3.3. Determining Functional Requirements for Exploring the Evolution of Indirect to Direct Relations

To carry out the two representative tasks in the DatAR prototype, we discussed potential functional requirements for exploring the evolution of historical indirect to current direct relations between topics with P1 and P2, Table 3.1 on p. 66, in July 2023.

- ***Selecting a timeline***

  Neuroscience publications can describe different studies on the same topic, leading to changes in the understanding of that topic over time. P1 expressed interest in selecting different publication timeframes, such as the past five years or the past three months, to investigate trends in indirect and then direct relations between brain regions and brain diseases. They wanted to explore both indirect and direct relations across various publication periods to identify cases where historical indirect relations have evolved into current direct relations, as well as to investigate current indirect relations that have not yet become direct. The first functional requirement (***FR1***) is thus to *provide a timeline selection mechanism.*

- ***Having access to publications***

  P1 and P2 proposed that being able to access the two, or more, publications containing the direct relations corresponding to an indirect relation would provide the information needed to assess the usefulness of the identified relations. The second functional requirement (***FR2***) is thus to *provide access to publications indicating identified indirect and direct relations.*

3.3.4. Determining Visualisation Requirements for Exploring the Evolution of Indirect to Direct Relations

In the same interview with P1 and P2, Table 3.1 on p. 66, in July 2023, we discussed potential visualisation requirements for presenting the evolution of indirect to direct relations between topics.

- ***Presenting publication dates***

  To visualise the timeline selection (Sec. 3.3.3, ***FR1***), P1 and P2 expressed the need for an intuitive display of publication dates. The first visualisation requirement (***VR1***) is thus to *present publication dates.*

- ***Presenting the evolution of historical indirect to current direct relations between topics***

  P1 and P2 expressed a preference for seeing the process by which historical indirect relations have evolved into current direct relations. This allows them to observe which previous indirect relations have become direct and which have yet to transition, aiding in the design of a potentially useful experiment. Our second visualisation requirement (***VR2***) is thus to *present the evolution of indirect to direct relations between topics.*

3.4. Design Rationale

We provide design rationales for the functional and visualisation design decisions.

3.4.1. *Timeline Selector* Design

To present publication dates (***VR1***) to facilitate the timeline selection (***FR1***), we first consider the following design options:

- **A publication dates table:** This design displays publication dates in a tabular format. Users can view the list to select specific dates of interest, providing a method to access detailed publication dates on identified indirect and direct relations.

- **A publication dates timeline:** This design presents a visual timeline where users can select date ranges by adjusting the handles on a timeline. The publication dates timeline aims to provide an intuitive method of filtering indirect and direct relations over time.

Designing the *Timeline Selector*, Fig. 3.3 on p. 47, is driven by several reasons:

- **Designing a timeline bar:** This study investigates a single-step approach for identifying and visualising the evolution of historical indirect relations into current direct relations. While a publication dates table provides detailed information, it may overwhelm users who are tracking the evolution of indirect relations into direct ones over time. We adopted a timeline design for publication dates. This design allows users to make single-step adjustments to the timeline, making it easy to identify the publication dates associated with the evolution of indirect to direct relations.

- **Three-monthly distribution within the timeframe:** Given that the KGBS includes neuroscience publications accessible via PubMed from January 2010 to February 2022 (Sec. 3.2.4). In the discussion with P1 and P2, Table 3.1 on p. 66, in July 2023, they suggested a minimum selection range of three months for literature exploration. They argued that shorter periods, such as one month, are inadequate for identifying useful indirect relations, while a three-month period provides a more comprehensive view of recent developments. This timeframe also offers flexibility, allowing researchers to adjust the publication date range to six months or one year as needed, thus tailoring their investigations to specific periods of interest. Consequently, we adopt three months as the time unit.

Additionally, to ensure a coherent user experience in the DatAR prototype, consistency with the existing interface is useful. We focus on two aspects:

- **Consistency with the *Direct Relations Filter* design:** The *Direct Relations Filter* in the DatAR prototype utilises a vertical layout with two handles to adjust the number of co-occurrences (Fig. 3.6 on p. 52). We retained this vertical layout for the design of the timeline to ensure consistency.

- **Consistency with previous colours:** The existing DatAR prototype uses a specific colour scheme to differentiate relations in the visualisations - light blue spheres for no direct or indirect relations, pink spheres for direct relations, and green spheres for indirect relations (Figs. 3.5, 3.6 and 3.7 on pp. 51–53). To maintain visual consistency, we adopted this colour scheme in the *Timeline Selector* design. Pink histograms represent the number of direct relations, and green histograms represent the number of indirect relations in the *Timeline Selector* (Fig. 3.3 on p. 47). We also maintain the use of blue in the colour scheme for the timeline bar, with selected timeframes visualised in dark blue. To help users focus on the evolution of indirect to direct relations within the selected timeframe, we display histograms for direct and indirect relations only within the selection range, rather than across the entire timeline bar.

3.4.2. Presenting the Evolution of Indirect to Direct Relations in a Single Visualisation

To present the evolution of historical indirect relations into current direct relations between topics (***VR2***), we consider the following design options:

- **Multiple comparative visualisations:** This design involves creating separate visualisations for different timeframes, each showing indirect and direct relations. Users would need to switch between these visualisations to compare them and identify which topics have evolved from indirect to direct relations.

- **Single visualisation:** This approach allows users to observe both indirect and direct relations in a single visualisation and see which relations have evolved from historical indirect to current direct by adjusting the timeline range.

We use a single visualisation to present both indirect and direct relations to make it easier for users to identify which topics have evolved from indirect to direct relations and which topics remain as indirect relations. To achieve this, we integrate the *Direct Relations Explorer* and the *Indirect Relations Explorer* with the *Brain Disease Topic Model* or the *Brain Regions Visualisation* using the *Timeline Selector* (Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61). As users adjust the timeline range, the resulting indirect and direct relations update immediately within the *Brain Disease Topic Model* or the *Brain Regions Visualisation*. This allows users to observe the evolution of historical indirect relations (represented by green spheres) into current direct relations (indicated by pink spheres) in a single visualisation.

3.4.3. *Publications Viewer* Design

To access publications indicating indirect and direct relations (***FR2***), we designed the *Publications Viewer* (Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61). Using data from the KGBS (Sec. 3.2.4), we display PMIDs to represent publications. Consistent with the existing DatAR prototype's colour scheme, the pink PMIDs list includes specific brain regions or diseases directly related to the investigated brain disease or region, along with each publication's PMID, Figs. 3.10 and 3.13 on pp. 58 and 61. The green PMIDs list includes brain regions or diseases with indirect relations via genes, along with the PMIDs of two publications indicating these indirect relations, Figs. 3.9 and 3.12 on pp. 57 and 60.

3.5. Implementation

Our functionalities and visualisations for this study are constructed using Unity[12] and the Microsoft Mixed Reality Toolkit[13] (MRTK) (v2.7.0) and are deployed on a HoloLens 2[14] head-mounted display. MRTK's AR interaction toolkit enables user interactions in the AR environment, such as hand tracking and grabbing and dragging 3D objects. We maintained consistency with the DatAR Prototype (Sec. 3.2.4) by following the same code logic to implement the *Timeline Selector* and *Publications Viewer*. The new code, which extends the previous code base, is stored as a separate branch in a coding repository[15].

3.6. Evaluation

To answer the three research questions, participants were asked to evaluate the implemented visualisation and functionality within the context of two representative tasks aimed at exploring the evolution of indirect to direct relations between brain regions and brain diseases. We detail the recruitment of participants in Sec. 3.6.1 and illustrate the survey procedure for the three research questions in Sec. 3.6.2. We then describe the survey questions for the individual research questions:

- The suitability of visualisation in the *Timeline Selector* for presenting the evolution of indirect to direct relations (RQ3) in Sec. 3.6.3;
- The usability of functionality in the *Timeline Selector* for exploring the evolution of indirect to direct relations (RQ2) in Sec. 3.6.4; and
- The usefulness of exploring the evolution of indirect to direct relations in understanding the current indirect relations for designing a useful experiment (RQ1) in Sec. 3.6.5.

3.6.1. Participants

We recruited 12 participants from neuroscience institutes at four universities, P3 - P14, see Table 3.1 on p. 66. Their academic qualifications, familiarity with AR, and their experience in literature exploration are described in Table 3.1 on p. 66. Ten participants are familiar with or professional in neuroscience literature exploration (P4-P13), and two participants have little experience in literature exploration (P3 and P14). Two participants have some understanding of AR and have some experience using the AR headset (P4 and P13), ten participants have no knowledge or experience with AR (P3, P5-P12 and P14). All 12 participants had normal or corrected-to-normal vision and were not colour-blind.

[12] Unity https://unity.com/, v2020.3.15f2

[13] Microsoft Mixed Reality Toolkit (MRTK) https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05

[14] Microsoft HoloLens 2 is an Augmented Reality headset developed and manufactured by Microsoft, https://www.microsoft.com/en-us/hololens

[15] The Unity-based code for exploring the evolution of indirect to direct relations is open-source and available on GitHub: https://github.com/DatAR-prototype/tree/Timelines-project

Table 3.1: Backgrounds of participants and studies they participated in. P1-P2 assisted with establishing the useful user task, determining two representative tasks for evaluation and determining the functional and visualisation requirements. Twelve participants (P3-P14) were recruited to evaluate the evolution of indirect to direct relations.

Participant Numbers	**Research Topics**	**Academic Qualifications**	**Knowledge and experience in AR technology**	**Experience in Literature Exploration**	**Studies Participated in**
P1	Neurology, in particular, *Depressive Disorder*	PhD student (third year)	Some knowledge of AR; Some experience in the AR headset (attended a previous study on exploring indirect relations in October 2022)	Familiar with literature exploration	Establish the user task of exploring the evolution of indirect to direct relations (Sec. 3.3.1); Determine two representative tasks (Sec. 3.3.2); Determine functional (Secs. 3.3.3) and visualisation requirements (Secs. 3.3.4)
P2	Neurology, in particular, *Parkinson Disease*	PhD student (third year)	Some knowledge of AR; Some experience in the AR headset (attended a previous study on exploring indirect relations in October 2022)	Familiar with literature exploration	Establish the user task of exploring the evolution of indirect to direct relations (Sec. 3.3.1); Determine two representative tasks (Sec. 3.3.2); Determine functional (Secs. 3.3.3) and visualisation requirements (Secs. 3.3.4)
P3	Neurobiology	PhD student (first year)	None	Little experience in literature exploration	Evaluate the evolution of indirect to direct relations
P4	Cognitive neuroscience & experimental psychology	PhD student (first year)	Some knowledge of AR; Some experience in the AR headset	Familiar with literature exploration	Evaluate the evolution of indirect to direct relations
P5	Neuroscience, radiology & psychology	PhD student (second year)	None	Familiar with literature exploration	Evaluate the evolution of indirect to direct relations
P6	Neurobiology, in particular, *Acute Ischemic Stroke*	PhD student (fourth year)	None	Familiar with literature exploration	Evaluate the evolution of indirect to direct relations

P7	Neurobiology, in particular, *Brain Tumours*	PhD student (second year)	None	Familiar with literature exploration	Evaluate the evolution of indirect to direct relations
P8	Cognitive neuroscience	PhD student (second year)	None	Familiar with literature exploration	Evaluate the evolution of indirect to direct relations
P9	Neurobiology	Postdoctoral researcher	None	Professional in literature exploration	Evaluate the evolution of indirect to direct relations
P10	Developmental neuroscience	PhD student (second year)	None	Familiar with literature exploration	Evaluate the evolution of indirect to direct relations
P11	Clinical neuroscience	PhD student (third year)	None	Professional in literature exploration	Evaluate the evolution of indirect to direct relations
P12	Neurology	PhD student (third year)	None	Professional in literature exploration	Evaluate the evolution of indirect to direct relations
P13	Clinical neuroscience	Master student	Some knowledge of AR; Some experience in the AR headset	Familiar with literature exploration	Evaluate the evolution of indirect to direct relations
P14	Cardiovascular neurobiology	PhD student (first year)	None	Little experience in literature exploration	Evaluate the evolution of indirect to direct relations

3.6.2. Survey Procedure

Each evaluation procedure took approximately 60 minutes. The evaluation sessions were conducted in January - April 2024. The procedure was as follows:

1 We collected information from 12 neuroscience researchers, including their research fields, academic qualifications, familiarity with AR, and experience in literature exploration (reported in Table 3.1 on p. 66). We asked them to sign an informed consent form (Appendix B on p. 135). Since age and gender are not relevant to this study, we did not collect this information [85].

2 We described the goal of the session and the detailed process of exploring the evolution of indirect to direct relations between brain regions and brain diseases (Appendix B on p. 137).

3 Participants wore the AR headset (HoloLens 2) and practiced grabbing, dragging, and rotating 3D objects in AR via operating the *Brain Regions Visualisation*, for approximately 10 minutes.

4 Participants performed the two representative tasks (described in Sec. 3.3.2) to

investigate the evolution of indirect to direct relations between brain regions and brain diseases using the *Timeline Selector* and accessed PMIDs using the *Publications Viewer*.

5 (a) While carrying out the two representative tasks, participants were asked to evaluate

- The suitability of the visualisation for the developed functionality (RQ3) (Sec. 3.6.3);
- The usability of functionality developed for exploring the evolution of indirect to direct relations (RQ2) (Sec. 3.6.4); and
- The usefulness of exploring the evolution of indirect to direct relations (RQ1) (Sec. 3.6.5)

(b) After completing the two representative tasks in approximately 30 minutes, participants were given the opportunity to explore the interactive AR prototype and its functionalities for another 15 minutes. Due to the limited memory of the AR headset, we provided the following brain regions and brain diseases as selectable inputs to explore indirect and direct relations:

- Three brain regions: *Cerebellar Vermis Structure*, *Lateral Amygdalar Nucleus* and *Nucleus Accumbens*
- Four brain diseases: *Depressive Disorder*, *Down Syndrome*, *Autistic Disorder* and *Gardner Syndrome*

by answering 13 semi-structured interview questions (IQs) (listed in Secs. 3.6.3, 3.6.4 and 3.6.5). Participants were asked to verbalise their thoughts while performing tasks to assist the experimenter in understanding their problem-solving process [69, 83].

6 Participants were asked to rate the suitability of the visualisation, the usability of functionality, and the usefulness of the user task by answering ten questions (listed in Figs. 3.14, 3.15 and 3.16 on pp. 70–74) using a 7-point Likert scale from Worst Imaginable to Best Imaginable (Worst Imaginable, Awful, Poor, OK, Good, Excellent, Best Imaginable) [9].

3.6.3. Survey for the Suitability of Visualisation for Presenting the Evolution of Indirect to Direct Relations (RQ3)

- **Semi-structured interview questions:** Participants answered 5 questions while performing the two representative tasks in the DatAR prototype.
  - **The suitability of the *Timeline Selector***

    Participants were asked:

    ◇ IQ1: Do you find the visualisation of the *Timeline Selector* by month suitable?

    ◇ IQ2: Are the timescales you can select suitable for your tasks?

  - ◇ IQ3: To what extent do you find the visual representations of the *Timeline Selector* suitable? Please comment on the colour, shape, and layout.
  - ◇ IQ4: Do you have any suggestions for improving the visualisation of the *Timeline Selector*?

- **The suitability of presenting the evolution of indirect to direct relations**

  Participants were asked:

  - ◇ IQ5: To what extent do you find the presentation of historical indirect relations evolving into current direct relations clear and suitable?

- **Questionnaires:** Participants were asked to rate the suitability of presenting the evolution of indirect to direct relations by answering 2 questions, Fig. 3.14 on p. 70. For example, "How suitable do you find the visualisation in illustrating the evolution of indirect to direct relations?" (Fig. 3.14 Q2 on p. 70).

3.6.4. Survey for the Usability of Functionality for Exploring the Evolution of Indirect to Direct Relations (RQ2)

- **Semi-structured interview questions:** Participants answered 3 questions while performing the two representative tasks in the DatAR prototype.

  - **The usability of the *Timeline Selector***

    Participants were asked:

    - ◇ IQ6: To what extent do you think the interaction process for exploring the evolution of indirect to direct relations is usable?
    - ◇ IQ7: How do you think the functionality of the *Timeline Selector* contributes to understanding the relations among the publications?
    - ◇ IQ8: Do you have any suggestions for improving the functionality of the *Timeline Selector*?

- **Questionnaires:** Participants were asked to rate the usability of functionality used in exploring the evolution of indirect to direct relations by answering 3 questions, Fig. 3.15 on p. 73. For example, "How would you rate the usability of the *Timeline Selector* in supporting your exploration of the evolution of indirect to direct relations?" (Fig. 3.15 Q3 on p. 73).

3.6.5. Survey for the Usefulness of Exploring the Evolution of Indirect to Direct Relations (RQ1)

- **Semi-structured interview questions:** Participants answered 5 questions while performing the two representative tasks in the DatAR prototype.

  - **The usefulness of exploring the evolution of indirect to direct relations**

    Participants were asked:

    ◇ IQ9: To what extent is it clear how historical indirect relations become current direct relations in the literature?

    ◇ IQ10: To what extent do you think it is worth exploring the evolution of indirect to direct relations?

    ◇ IQ11: Can you share some results of the evolution of indirect to direct relations that you discovered using the *Timeline Selector*?

    ◇ IQ12: To what extent do you think the implemented visualisations and functionalities would guide other neuroscientists in identifying potentially useful hypotheses?

  – **An open question on improvements in the DatAR prototype**

  Participants were asked:

    ◇ IQ13: What information would you like to explore further?

- **Questionnaires:** Participants were asked to provide their opinions on the usefulness of exploring the evolution of indirect to direct relations by answering 5 questions, Fig. 3.16 on p. 74. For example, "How would you rate the usefulness of an indirect relation between *Depressive Disorder* and *Cerebellar Vermis Structure* evolved into a direct relation?" (Fig. 3.16 Q7 on p. 74).

3.7. Results

3.7.1. The Suitability of the *Timeline Selector* Visualisation for Presenting the Evolution of Indirect to Direct Relations (RQ3)

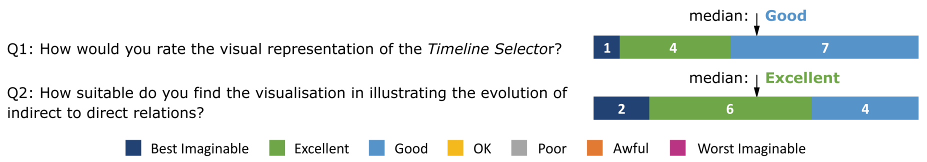

Figure 3.14: The figure shows the distribution of the twelve participants' attitudes toward the *Timeline Selector* Visualisation for presenting the evolution of indirect to direct relations.

In response to the questionnaire described in Sec. 3.6.3, all twelve participants stated that the visual representation of the *Timeline Selector* is suitable or better, Fig. 3.14 Q1 on p. 70. The median rating for the visual representation of the *Timeline Selector* is good, with responses ranging from good to best imaginable, Fig. 3.14 Q1 on p. 70.

All twelve participants found that showing the evolution of historical indirect relations to current direct relations in a single visualisation of the *Brain Disease Topic Model* or the *Brain Regions Visualisation* is clear and suitable, Fig. 3.14 Q2 on p. 70. The median rating for showing the evolution of indirect to direct relations between topics is excellent, with responses ranging from good to best imaginable, Fig. 3.14 Q2 on p. 70. P11 remarked (IQ5), “I find the visualisation for previous indirect relations that have evolved into current direct relations clear and understandable. The change in colours of brain diseases from green to pink in the single *Topic Model* clearly indicates this evolution.

Detailed results for the layout and colour of the *Timeline Selector* are as follows:

- **Timescales of the *Timeline Selector***

  All participants found the display of publication dates by month for the 12-year period (IQ1) suitable. All participants found the selectable timescales (IQ2), such as three months, six months, and one year, suitable for exploring topic evolution. Seven participants proposed that showing only the year scale is sufficient for their daily literature exploration. For example, P6 mentioned, “Sometimes I have a really specific date in mind for literature exploration, such as the past five years, and the *Timeline Selector* enables me to identify this specific range”. P7 noted, “I find the *Timeline Selector* more useful for researchers who regularly analyse newly published papers and want to see what happened in the last three months.”

- **Layout of the *Timeline Selector***

  Three participants found the vertical layout of the *Timeline Selector* (IQ3) to be a suitable visual representation. Four participants mentioned their preference for the horizontal layout of the *Timeline Selector*. P8 elaborated, “With the current vertical layout, I can only see a part of the publication dates within a limited vertical field of view. I need to look up and down to see the entire range of publication dates. I think the width of this field of view is adequate. Therefore, changing the *Timeline Selector* to a horizontal layout would facilitate better visibility of the complete publication dates.”

- **Colour of the *Timeline Selector***

  All participants found the consistent use of colours in the prototype - pink for direct relations and green for indirect relations - to be clear and suitable (IQ3). P4 remarked, “Maintaining consistent colours for direct and indirect relations in the whole prototype is quite useful, as it intuitively helps me differentiate between histograms representing direct and indirect relations. The disappearance of histograms in the *Timeline Selector* when those publication dates are not selected. That’s nice. The indication of the selected range of publication dates by a thicker bar is very clear. Overall, I think the *Timeline Selector* looks great.”

- **Numbers of publications in histograms**

  Three participants suggested displaying the specific numbers of publications indicating the direct and indirect relations on the pink and green histograms of the *Timeline Selector* (IQ4), as it currently only shows the number of co-occurrences. P12 reported, “I want to know the numbers of publications indicating direct and indirect relations to understand the extent to which these relations are discussed in the literature. I could imagine that a direct relation is mentioned many times only in one publication."

3.7.2. The Usability of the *Timeline Selector* Functionality for Exploring the Evolution of Indirect to Direct Relations (RQ2)

In response to the questionnaire described in Sec. 3.6.4, eleven participants indicated that the functionality of the *Timeline Selector* is useful for exploring the evolution of indirect to direct relations, Fig. 3.15 Q3 on p. 73. The median rating for the usability of the *Timeline Selector* is excellent, with responses ranging from ok to excellent, Fig. 3.15 Q3 on p. 73. The interaction process for exploring the evolution of indirect to direct relations using the *Timeline Selector* is clear and understandable for all participants (IQ6). P6 pointed out, "I can track the evolution of indirect to direct relations by observing the colour change from green to pink and pinpointing the specific date of this transition by adjusting the handle of the *Timeline Selector*."

All twelve participants reported that the functionality of the *Timeline Selector* contributes to exploring the neuroscience literature (IQ7). P4 said, "The *Timeline Selector* helps identify the time periods when certain indirect and direct relations between topics occurred in publications. The inclusion of months is useful because it allows you to pinpoint not just the year, but also whether it was at the beginning or end of 2019." P8 and P13 both proposed initially providing the timeline by year, with the option to display detailed months if needed (IQ8). P8 noted, "When I find the relation between *Depressive Disorder* and *Cerebellar Vermis Structure* transitioning from indirect to direct in 2019, I would be interested in exploring that year further. I hope to click on 2019 and find the detailed month(s) and the specific publication(s) indicating the first published direct relation." Four participants expressed a desire to quickly locate the publication dates and publications that first documented the evolution of an indirect to a direct relation (IQ8). They thought that the publication describing this initial evolution represented a useful discovery.

3.7.3. The Usefulness of Exploring the Evolution of Indirect to Direct Relations (RQ1)

In response to the questionnaire described in Sec. 3.6.5, all twelve participants stated that exploring the evolution of indirect to direct relations is a useful task for making them confident in the current indirect relations for designing potentially useful experiments, Fig. 3.16 Q6 on p. 74. The median rating for the usefulness of the user task is excellent, with responses ranging from good to best imaginable, Fig. 3.16 Q6 on p. 74. P4 commented (IQ10), "I always need to start somewhere to begin my research. I might identify an indirect relation within the range of 2010 to 2016. Then, I explore whether it remains indirect or transitions to direct. If the indirect relation has not yet transitioned, it can guide future research directions."

- **Usefulness of exploring the indirect relation between *Depressive Disorder* and *Cerebellar Vermis Structure* evolved into a direct relation (Representative Task 1)**

  Nine participants thought that an indirect relation between *Depressive Disorder* and *Cerebellar Vermis Structure* evolving into a direct relation is useful for their research, Fig. 3.16 Q7 on p. 74. The median rating for the usefulness of this representative task is excellent, with responses ranging from good to best imaginable, Fig. 3.16 Q7 on p. 74. P8 reported, "I find the direct relation between *Depressive Disorder* and *Cerebellar Vermis Structure*, first published in 2019, indicating a recent development. This information is useful in inspiring me to consider how to regulate *Cerebellar Vermis Structure* to treat *Depressive Disorder*." P8 also commented, "If my research topic is *Depressive Disorder*. I want to understand which brain regions are affected by *Depressive Disorder*. Traditionally, I would begin exploring literature by searching for *Depressive Disorder* on PubMed to find affected regions. Now I utilise the implemented visualisations and functionalities to comprehend the overall relations between *Depressive Disorder* and brain regions, rather than having to examine each separate brain region on PubMed."

- **Usefulness of exploring the indirect relation between *Cerebellar Vermis Structure* and *Parkinson Disease* evolved into a direct relation (Representative Task 2)**

  Ten participants thought that an indirect relation between *Cerebellar Vermis Structure* and *Parkinson Disease* evolving into a direct relation is useful for their research, Fig. 3.16 Q8 on p. 74. The median rating for the usefulness of this representative task is excellent, with responses ranging from good to best imaginable, Fig. 3.16 Q8 on p. 74. P7 reported, "I find the evolution of the indirect relation between *Cerebellar Vermis Structure* and *Parkinson Disease* to a direct relation in 2020 to be new knowledge for me. I think it will be useful for understanding the pathophysiology of *Parkinson Disease* and potential interventions targeting *Cerebellar Vermis Structure*."

- **Seven additional potentially useful hypotheses via the *Timeline Selector***

  All twelve participants thought the designed visualisations and functionalities in the DatAR prototype could guide other neuroscientists in finding potentially useful hypotheses (IQ12). Six participants shared seven additional potentially useful hypotheses (IQ11).

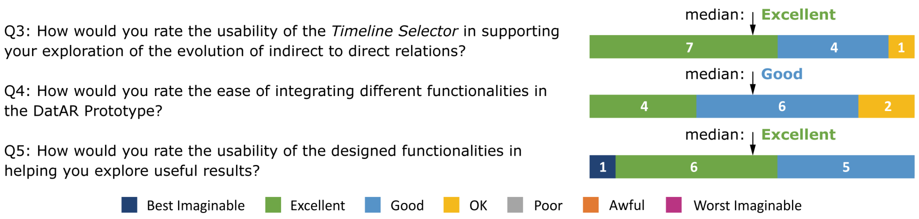

Figure 3.15: The figure shows the distribution of the twelve participants' attitudes toward the *Timeline Selector* functionality used in exploring the evolution of indirect to direct relations.

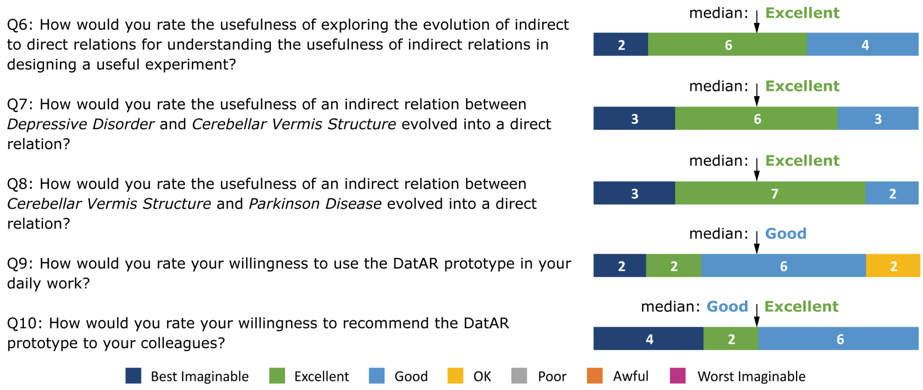

Figure 3.16: The figure shows the distribution of the twelve participants' attitudes toward the usefulness of exploring the evolution of indirect to direct relations.

- ***Finding the latest direct relations***

  P5 proposed, "I am interested in *Down Syndrome* and I want to search for brain regions that have direct relations with *Down Syndrome* in the latest five years to uncover recent developments. I discover a direct relation between *Down Syndrome* and *Cerebral Hemisphere Structure* in June 2021. This relation was previously unknown to me, providing useful new knowledge. I inform the experimenter that the PMID is 33945512[16], to further explore the publication on the laptop.

- ***Finding well-established direct relations***

  P4 pointed out, "I can observe that the direct relation between *Amygdala* and *Depressive Disorder* may predate 2010. Even when I select older publications, such as those from January 2010 to June 2010, the direct relation between *Amygdala* and *Depressive Disorder*[17] is still present. Discovering such an early relation means that it is likely a robust one."

- ***Finding the evolution of indirect to direct relations***

  P7 regularly conducted weekly literature studies and wanted to outline what had happened in the past half year to understand recent topics. P7 remarked, "I select the publication dates range from June 2021 to February 2022. I find a direct relation between *Down Syndrome* and *Cerebral Hemisphere Structure*. Curious about when the direct relation first appeared, I discover that there is no relation between *Down Syndrome* and *Cerebral Hemisphere Structure* from 2010 to September 2018. After September 2018, an indirect relation is noted, which then became direct in June 2021. This is a very recent development of research."

[16] DOI: 10.1172/JCI135763
[17] DOI: 10.1016/j.biopsych.2010.09.022

P14 pointed out, “I can see there are indirect relations between *Autistic Disorder* and certain brain regions in March 2013 and also from January to March 2018 via the green histograms in the timeline bar. I find an indirect relation between *Autistic Disorder* and *Lateral Amygdalar Nucleus* in March 2013, and another indirect relation between *Autistic Disorder* and *Corpus Striatum Structure* from January to March 2018. I want to see if these two pairs of indirect relations ever evolved into direct relations. In March 2020, the indirect relation between *Autistic Disorder* and *Corpus Striatum Structure* becomes a direct relation. It is a newer discovery, and I also find the publication showing the direct relation, PMID is 32130906[18].”

– ***Finding indirect relations that have not yet become direct relations***

P4 pointed out, “Before 2011, there was no relation between *Depressive Disorder* and *Middle Frontal Gyrus*, but after 2011, an indirect relation emerged. I became curious to see if it ever evolved into a direct relation, but it remained indirect. This could potentially be a topic for future research.”

P6 pointed out, “I find that the indirect relation between *Disorder Nervous System* and *Cerebellar Vermis Structure* becomes direct after March 2020, and the indirect relation between *Parkinson Disease* and *Cerebellar Vermis Structure* also becomes direct after September 2020. I observe that *Dyskinetic Syndrome* is close to *Disorder Nervous System* and *Parkinson Disease* in the *Topic Model*, but the indirect relation between *Dyskinetic Syndrome* and *Cerebellar Vermis Structure* has not yet become direct. I think this is a useful direction for a future experiment.”

P12 mentioned, “The identified indirect relation between *Autistic Disorder* and *Lateral Amygdalar Nucleus* in March 2013 has not evolved into a direct one. I am interested in reading publications indicating the indirect relation; they could inspire useful hypotheses for future experiments.”

3.8. Discussion

We discuss the visualisation (Sec. 3.8.1) and functionality (Sec. 3.8.2) of the *Timeline Selector*. We then provide a comprehensive review of the user task of exploring the evolution of indirect to direct relations (Sec. 3.8.3).

3.8.1. *Timeline Selector* Visualisation (RQ3)

- ***Visualising the Timeline Selector in a horizontal layout***

  We considered the layout consistency in the DatAR prototype, which led us to choose a vertical layout for the *Timeline Selector*. However, the current vertical layout requires users to shift their gaze up and down to view the entire range of publication dates, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61 (reported in Sec. 3.7.1). Given that the field of view (FOV) in the AR headset has a 3:2

[18] DOI: https://doi.org/10.1016/j.celrep.2020.02.030

aspect ratio, a horizontal layout might allow users to see the complete range of publication dates within their field of view more easily, thereby facilitating better navigation. A potential solution is to conduct a comparative test of the vertical and horizontal layouts. This test could help determine the most suitable layout for optimising user experience in the AR environment.

- ***Hierarchical timescale: from year to month***

  When considering the different timescale needs for literature exploration in neuroscience (reported in Sec. 3.7.1), participants expressed a desire for hierarchical timescales in the *Timeline Selector*. A potential solution would be to initially present a yearly timescale and then allow users to explore more detailed months and dates as needed. Such a hierarchical approach supports a more focused investigation, providing a clear and structured pathway to detailed data on the evolution of indirect to direct relations within a particular year.

- ***Visualising the numbers of publications in histograms***

  Participants expressed a desire for displaying the specific numbers of publications indicating direct and indirect relations on the histograms (reported in Sec. 3.7.1). Displaying these publication counts would allow users to quickly gauge the frequency of these identified relations within the literature.

3.8.2. *Timeline Selector* Functionality (RQ2)

- ***Linking histograms directly to publications***

  Participants expressed a desire to quickly identify the publication that first documented the evolution of an indirect to a direct relation (reported in Sec. 3.7.2). The initial documentation of direct relations can be useful, and neuroscientists want to read the full publication to explore these initial discoveries. A potential solution is to link the histograms in the *Timeline Selector* directly to PMIDs. This improvement would enable users to both visualise the evolution of relations between topics and easily locate the publications.

3.8.3. Usefulness Concerns for Exploring the Evolution of Indirect to Direct Relations (RQ1)

- ***The current indirect relations that have not yet become direct are perceived as useful***

  Neuroscience participants gave positive feedback on the usefulness of exploring the evolution of indirect to direct relations between topics. They identified seven additional hypotheses that could help them identify potentially useful experiments, Sec. 3.7.3. Tracking the evolution of indirect to direct relations reflects recent advancements in neuroscience, offering researchers a dynamic view of the field's progression. Additionally, participants thought that exploring the evolution of indirect to direct relations could help them better understand the current indirect relations that have not yet become direct. They suggested that the current indirect relations could help propose useful hypotheses for designing experiments.

- ***Providing specific intermediate topics (such as a specific gene) responsible for indirect relations***

  In this study, we used the topic "gene" as an intermediate topic to investigate indirect relations between brain regions and brain diseases. Neuroscience participants suggested providing the specific gene responsible for indirect relations. A potential solution is to extend the *Indirect Relations Explorer*, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61, by allowing users to select specific intermediate topics, such as specific genes, relevant to their research. This enhancement would enable users to identify and explore indirect relations more precisely. By selecting specific intermediate topics and using the *Timeline Selector* to explore topic evolution, neuroscientists can refine their focus and ensure that the identified indirect relations - which have not yet become direct - are useful and worthy of further investigation.

3.9. Limitations

- ***Mapping names for Brain Regions***

  The names and 3D positions of brain regions we used in the *Brain Regions Visualisation* are sourced from the Scalable Brain Atlas (SBA) (Sec. 3.2.4). We use Unified Medical Language System[19] (UMLS) terms to identify the names of brain regions in the repository. Some of the UMLS brain region names differ from those in the SBA. We utilise a limited mapping mechanism to convert brain region names for carrying out the study. A more comprehensive mapping would require the assistance of a neuroscientist.

- ***Update the literature in the Knowledge Graphs of Brain Science***

  In the current KGBS (Sec. 3.2.4), we analyse the co-occurrences between topics in publications from January 1, 2010, to February 3, 2022. This allows us to carry out our exploratory interface studies. For serious use by neuroscientists, we need to update our repository with the latest literature starting from 2022.

3.10. Conclusions and Future Work

The task of exploring the evolution of indirect to direct relations helps users understand the usefulness of currently identified indirect relations, which in turn assists them in designing potentially useful experiments, Sec. 3.7.3. We will continue our collaboration with neuroscientists to explore how we can provide specific topics, such as a specific gene, that contribute to indirect relations. This will assist neuroscientists in evaluating whether identified indirect relations are indeed useful and worthy of further investigation, Sec. 3.8.3. Another direction for exploration involves assessing the use-

[19] Unified Medical Language System (UMLS) https://www.nlm.nih.gov/research/umls/index.html brings together biomedical vocabularies and standards to enable interoperability between computer systems.

fulness of the DatAR prototype in the context of daily neuroscience literature research. By systematically observing and collecting data while using the DatAR prototype during neuroscientists' daily literature exploration, we will assess its overall usefulness in identifying potentially useful experiments. We plan to follow a real neuroscientist's research of interest, integrating their feedback and experiences to refine the prototype's functionalities and visualisations. This approach will ensure that the tool aligns with the actual needs of researchers and enhances their ability to generate indeed useful experiments. Another potential direction involves integrating Large Language Models (LLMs) into the DatAR prototype. By utilising LLMs, we can easily uncover previously unobserved indirect relations between topics. The DatAR visualisation will then provide an intuitive way to explore these complex relations, helping researchers to better understand the evolving relations between topics in large numbers of neuroscience publications. We view this work as contributing to understanding indirect relations that have yet to become direct ones, aiding neuroscientists in designing experiments that could potentially yield useful contributions.

4

Informing the Design of a Neuroscience Experiment by Exploring Indirect Relations via a Specific Intermediate Topic

Neuroscientists need to understand the state of the art before conducting a costly experiment, a process that is both challenging and time-consuming. Topic-based literature exploration has proven useful in helping neuroscientists by allowing them to obtain an overview of the literature more efficiently than reading hundreds of publications, and thus be better informed to design a potentially useful experiment. Specifically, topic-based exploration enables users to explore large numbers of publications simultaneously, revealing relations between neuroscience topics based on multiple publications rather than individual results reported in a single publication. We define a *direct* relation when two topics, such as a brain region and a brain disease, co-occur within the same sentence of a publication's title or abstract. An *indirect* relation exists when no direct relation is present between the two topics, but both co-occur with an intermediate topic, such as a mental process. Neuroscientists have indicated that exploring indirect relations via a specific intermediate topic helps them understand the underlying mechanisms of the identified indirect relations, in turn, informing the design of their experiments. This study investigates, using a user-centred design approach, how exploring a specific intermediate topic within indirect relations can support the understanding of these relations to inform the design of an experiment. We conduct interviews with neuroscientists to 1) verify the usefulness of the task of exploring indirect relations via a specific intermediate topic, 2) identify the required functionality, and 3) design and implement the *Specific Intermediate Topic Explorer* visualisation. Nine neuroscientists performed two representative tasks using the implemented functionality and visualisation. Responses to 19 semi-structured interview questions and

10 Likert-scale questions indicated that (a) the functionality is *useful* for exploring indirect relations via a specific intermediate topic, (b) the visualisation is *suitable* for presenting indirect relations via a specific intermediate topic, and (c) exploring indirect relations via a specific intermediate topic is *useful* for informing the design of an experiment.

Chapter contribution: This chapter investigates how the user task of exploring indirect relations via a specific intermediate topic can inform the design of a neuroscience experiment (MRQ3), specifies an appropriate specific intermediate topic selector functionality to support this task, and designs and implements a corresponding 3D AR visualisation to enable this functionality.

4.1. Introduction

Before conducting a costly experiment, neuroscientists need to thoroughly review the existing literature to inform the design of the experiment. This process is, however, both challenging and time-consuming. Topic-based literature exploration enables researchers to analyse large numbers of publications simultaneously [99, 108, 109], offering an overview of relations between neuroscience topics, such as brain regions and brain diseases, rather than focusing on individual publications. This overview perspective helps neuroscientists become better informed when designing a potentially useful experiment.

- A *direct* relation exists when two neuroscience topics appear in the same sentence within the title or abstract of a publication in PubMed[1]. For example, there is a direct relation between the brain disease *Alzheimer's Disease* and the brain region *Hippocampus*, since they both occur in the title of the publication *Post-Ischemic Neurodegeneration of the Hippocampus Resembling Alzheimer's Disease Proteinopathy*[2]. The brain disease *Alzheimer's Disease* thus has a direct relation to the brain region *Hippocampus*. Both topics appear together in 965 sentences in our repository (Sec. 4.2.3), Fig. 4.1 on p. 82.

- An *indirect* relation exists when there is no direct relation between two topics, but both co-occur with a specific intermediate topic, such as *Communication Response* (a mental process). For example, there is no direct relation between the brain disease *Alzheimer's Disease* and the brain region *Cerebellar Vermis Structure* in the repository that we use (Sec. 4.2.3). There are, however, 137 co-occurrences of *Alzheimer's Disease* with *Communication Response* (a mental process), and 41 co-occurrences of *Communication Response* with *Cerebellar Vermis Structure*, Fig. 4.2 on p. 83. Based on these two direct relations, we conclude that there is an indirect relation between the brain disease *Alzheimer's Disease* and brain region *Cerebellar Vermis Structure* via *Communication Response*.

To allow neuroscientists to see the relative spatial locations of brain regions we use

[1] PubMed is a free resource that supports the search and retrieval of biomedical and life sciences literature https://pubmed.ncbi.nlm.nih.gov/
[2] DOI: https://doi.org/10.3390/ijms23010306

Augmented Reality (AR) to visualise them [99, 109]. When these are affected by the same or similar diseases, the 3D visualisation helps neuroscientists better understand the complex anatomical relations [39, 52]. Moreover, AR facilitates the integration of literature exploration into neuroscientists' daily workflows. It extends traditional 2D reading by allowing researchers to read publications on their screens while using the virtual space to visualise multiple sets of topic relations [19, 58, 63]. This enables them to compare different lines of hypothesis during the literature exploration process. Immersion in the AR environment also enhances memory recall more effectively than both VR and 2D desktop environments, particularly in terms of spatial and ordinal positioning of 3D brain visualisations [35, 65]. A 3D AR prototype, DatAR, was developed to visualise the spatial representation of brain regions and to explore direct and indirect relations between brain regions and brain diseases [99, 108, 109] (Sec. 4.2.3).

Identifying indirect relations between topics in the literature can aid in designing potentially useful experiments, as weak evidence, such as through an intermediate topic, may suggest a connection not yet confirmed by a single publication. Indirect relations indicate the presence of promising but unexplored links, which, in turn, can inform experimental design. In previous work on exploring direct and indirect relations [108, 109], a participating neuroscientist noted that while visualising indirect relations in 3D brain regions offers an intuitive approach to identify potentially useful experiments, providing specific intermediate topics to find indirect relations would help assess the usefulness of the indirect relations identified. By selecting a specific intermediate topic, such as a gene, protein, or mental process, neuroscientists can better understand the biological mechanisms underlying an indirect relation between a brain region and a brain disease, thereby informing the design of a useful experiment.

Our main research question (RQ1) is: *How does exploring indirect relations via a specific intermediate topic support the understanding of the identified relations and, in turn, contribute to the design of a useful experiment?*

Using a user-centred design approach [17, 26, 27, 29, 75], we first consulted two neuroscientists to determine whether exploring indirect relations via a specific intermediate topic is a potentially useful user task (Sec. 4.3.1). We then asked the same two neuroscientists to identify two representative tasks that reflect realistic neuroscience use cases for our later evaluation (Sec. 4.3.2).

Having established, through our first research question (RQ1), that exploring indirect relations via specific intermediate topics is a potentially useful user task, we need to specify and implement the functionality and visualisation required to support it. Based on input from interviews with the same two neuroscientists, we identified two functional requirements and one visual requirement for exploring indirect relations via a specific intermediate topic (Sec. 4.3.4). We then explain our design decisions to address these functional and visual requirements (Sec. 4.3.5) and implement the *Specific Intermediate Topic Explorer* (Fig. 4.11 on p. 100) in our 3D AR prototype (Sec. 4.4).

Having designed and implemented the *Specific Intermediate Topic Explorer*, we need to assess the visualisation for presenting specific intermediate topics, the functionality for exploring indirect relations via a specific intermediate topic, and the usefulness of the user task of exploring indirect relations via a specific intermediate topic in informing the design of an experiment. We outline the evaluation procedure for the visualisation, functionality, and user task in Sec. 4.5.

- We first investigate the suitability of the visualisation for presenting indirect relations via a specific intermediate topic. Our third research question (RQ3) is: *To what extent is the visualisation suitable for the specified functionality in presenting indirect relations via a specific intermediate topic?* The visualisation was assessed using two representative tasks with nine neuroscientists (Sec. 4.6.1).

- We then assess the usability of the implemented functionality for the user task. Our second research question (RQ2) is: *To what extent is the implemented functionality useful for exploring indirect relations via a specific intermediate topic?* The nine neuroscientists who evaluated RQ3 were also asked to assess the usability of the functionality (Sec. 4.6.2).

- Our final step is to evaluate the usefulness of the user task - exploring indirect relations via a specific intermediate topic - for designing a useful experiment (RQ1). The same nine neuroscientists were consulted and noted that exploring specific intermediate topics is useful for understanding indirect relations to inform the design of an experiment (Sec. 4.6.3).

The contributions of this work are (i) establishing that the user task of exploring indirect relations via a specific intermediate topic is useful (answer to RQ1), (ii) the design and evaluation of the enabling functionality (answer to RQ2), and (iii) the design and evaluation of a corresponding visualisation (answer to RQ3).

4.2. Related Work

We first discuss the usefulness of topic-based literature exploration for designing a useful experiment, Sec. 4.2.1. We discuss the potential of intermediate topics in indirect relations within biomedical research tasks, such as drug target identification and bio-

Left brain disease	Middle number of co-occurrences	Right brain region
"Alzheimer's Disease"	"965"^^xsd:integer	"BRAIN, HIPPOCAMPUS"
"Alzheimer's Disease"	"36"^^xsd:integer	"Prefrontal Cortex"
"Alzheimer's Disease"	"17"^^xsd:integer	"Caudate nucleus structure"
⋮	⋮	⋮

Figure 4.1: The brain disease *Alzheimer's Disease* (left column), co-occurs 965 times (middle column) with the brain region *Hippocampus* (right column). *Alzheimer's Disease* has a direct relation with *Hippocampus*. The elements in the figure are reused in Fig. 4.2 on p. 83 for comparing indirect relations.

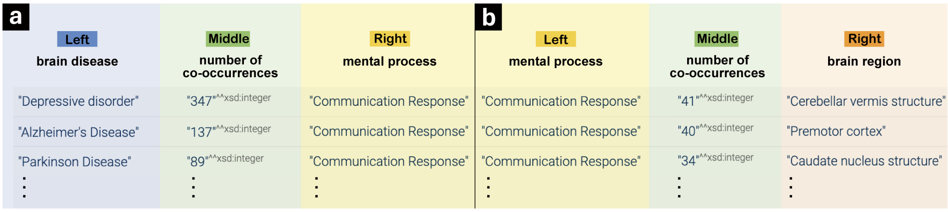

Figure 4.2: (a) The brain diseases *Depressive Disorder*, *Alzheimer's Disease*, and *Parkinson Disease* (left column) have direct relations with *Communication Response* (right column).
(b) *Communication Response* (left column) has direct relations with the brain regions *Cerebellar Vermis Structure*, *Premotor Cortex*, and *Caudate Nucleus Structure* (right column).
The brain diseases *Depressive Disorder*, *Alzheimer's Disease*, and *Parkinson Disease* have indirect relations with *Cerebellar Vermis Structure*, *Premotor Cortex* and *Caudate Nucleus Structure* via *Communication Response*. We know from the direct relations shown in Fig. 4.1 on p. 82 that *Alzheimer's Disease* already has direct relations with the brain regions *Premotor Cortex*, and *Caudate Nucleus Structure*. Based on this, we conclude that an indirect relations exist between *Alzheimer's Disease* and *Cerebellar Vermis Structure* via *Communication Response*.

marker discovery, to inspire useful insights, Sec. 4.2.2. We then present an interactive 3D AR prototype, DatAR, that enables neuroscientists to explore direct and indirect relations between brain regions and brain diseases, which provides an initial implementation platform for conducting this study, Sec. 4.2.3.

4.2.1. Topic-based Literature Exploration

Traditional literature search in a specific domain focuses on finding and reading individual publications, making it difficult to gain an overview of the state of the art [47, 61]. Topic-based literature exploration provides an overview by identifying relations between topics [3, 80]. For example, in our case, researchers can explore which brain regions are affected by *Alzheimer's Disease* without reading individual publications. Topic-based literature exploration has the potential to yield useful insights to investigate, for example, gene-disease associations. There is a relation between *colorectal cancer* (CRC) and the gene *EP300*, which is also associated with *breast cancer* (BrCa), suggesting that a better understanding of *EP300* could support progress in both disease areas [100]. In our work, neuroscientists are looking for less obvious, indirect relations between topics [53, 71], as these may point to useful hypotheses in the literature where a single experiment may be able to establish a direct relation between them.

4.2.2. Exploring Indirect Relations via a Specific Intermediate Topic

The ABC co-occurrence model has been demonstrated as a useful method in medical tasks for exploring previously unknown links or indirect relations [79, 93]. It states that if topic A (such as a brain disease *Alzheimer's Disease*) is associated with topic B (such as a mental process *Communication Response*), and B is associated with topic C (such as a brain region *Cerebellar Vermis Structure*), then topic A has an indirect relation with topic C [33]. The intermediate topic B plays a useful role in uncovering potential mechanisms underlying the relation between A and C within the context of medical tasks [92]. For example, "neuroinflammation" has been identified as a bridging topic between *Epilepsy* and *Depression*, helping researchers discover resistance mechanisms and pathophysiological traits [95]. *Parkinson Disease* shows associations with *Motor Control*, which in turn is associated with *Basal Ganglia*, suggesting that regulating *Motor Control* could be useful for understanding the connection between *Parkinson Disease* and *Basal Ganglia* [82].

Inspired by the ABC co-occurrence model, we previously used intermediate topic types, such as genes or mental processes, to visualise which brain regions or brain diseases have indirect relations with a selected brain disease or brain region (Sec. 4.2.3). This previous study focused on visualising direct and indirect relations in a single view, enabling neuroscientists to intuitively identify indirect relations [2].

In this study, we focus on a specific intermediate topic, such as the mental process *Communication Response*, within indirect relations. Users can select an intermediate topic they consider potentially useful for identifying indirect relations. By determining which specific intermediate topics lead to which indirect relations, neuroscientists can gain a better understanding of the cognitive processes and biological mechanisms underlying the indirect relations between brain regions and brain diseases. Moreover, these specific intermediate topics can provide supporting evidence to inform the design of an experiment.

4.2.3. The 3D AR Prototype, DatAR

An interactive 3D AR prototype, DatAR, has been shown to help neuroscientists explore relations between neuroscience topics [37, 99, 108]. To explain how the prototype supports this exploration, we first introduce the Knowledge Graphs of Brain Science (KGBS), which store topic co-occurrences used in DatAR, Sec. 4.2.3. We then describe the interactive widgets that provide DatAR's functionalities, which together enable users to explore direct and indirect relations between brain regions and brain diseases. The *Brain Regions Visualisation* and the *Brain Disease Topic Model* are also used in this study, Sec. 4.2.3.

The Knowledge Graphs of Brain Science

The Knowledge Graphs of Brain Science[3] (KGBS) is a repository that contains an analysis of the sentences in the titles and abstracts of 414,224 neuroscience publications indexed in PubMed from January 2010 to February 2022. Given the exploratory nature of our study, this time frame provides a sufficient number of publications for identifying topic co-occurrences used to find indirect relations. We use KGBS to query co-occurrences between two topics, such as brain regions, brain diseases, genes, proteins, mental processes, and symptoms. For example, an excerpt of co-occurrences (direct relations) between brain diseases and brain regions is presented in Fig. 4.1 on p. 82.

[3]To access the Knowledge Graphs of Brain Science at https://kgbs-sparql.project.cwi.nl/, please contact <xuboyudesign@outlook.com>.

Two direct relations connected through an intermediate topic can reveal an indirect relation. An example of an indirect relation between the brain disease *Alzheimer's Disease* and the brain region *Cerebellar Vermis Structure* via the Mental Process *Communication Response* is illustrated in Fig. 4.2 on p. 83.

Widgets for Exploring Direct and Indirect Relations Between Topics

We describe the widgets in our 3D AR prototype for exploring direct and indirect relations relevant to this study.

- The *Brain Regions Visualisation*, Fig. 4.3 on p. 86, displays 274 brain regions derived from the Scalable Brain Atlas[4] [8]. Each sphere's position corresponds to the 3D coordinates of the corresponding brain region in the atlas [99].

- The *Brain Disease Topic Model*, Fig. 4.4 on p. 87, presents 3D visualisations of 151 brain diseases from the KGBS [99]. The spatial arrangement of any two diseases is determined by direct relations across all topics, where closer proximity in 3D space indicates more direct relations in the repository [99].

- The *Direct Relations Explorer*, Figs. 4.3 and 4.5 on pp. 86 and 88, allows users to query the direct relations between a single brain region and multiple brain diseases or a single brain disease and multiple brain regions.

- The *Indirect Relations Explorer*, Figs. 4.4 and 4.6 on pp. 87 and 89, allows users to select an intermediate topic type, such as mental processes or genes, to identify indirect relations between brain regions and brain diseases.

- The *Direct Relations Filter*, Fig. 4.7 on p. 90, allows users to filter the number of direct relations between a single brain region and multiple brain diseases, or between a single brain disease and multiple brain regions. Users can filter for fewer or more direct relations by adjusting the upper and lower handles in the visualisation.

- The *Overlap Relations Explorer*, Fig. 4.8 on p. 91, allows users to query topics that have direct relations with two brain diseases, such as *Anxiety Disorders* and *Depressive Disorder*, simultaneously. For example, brain regions, such as *Hippocampus*, *Cerebral cortex* and *Prefrontal cortex*, have direct relations with both *Anxiety Disorders* and *Depressive Disorder*.

The *Indirect Relations Explorer* widget does not support selecting specific intermediate topics to explore indirect relations [99, 108, 109]. In this study, we design functionality and a visualisation for an intermediate topic selector to assist neuroscientists in identifying and visualising which specific intermediate topic, such as a mental process [99, 108, 109], reveals indirect relations between brain regions and brain diseases. This aims to support the understanding of the underlying mechanisms behind identified indirect relations, thereby informing the design of a useful experiment.

[4]Scalable Brain Atlas https://scalablebrainatlas.incf.org/index.php

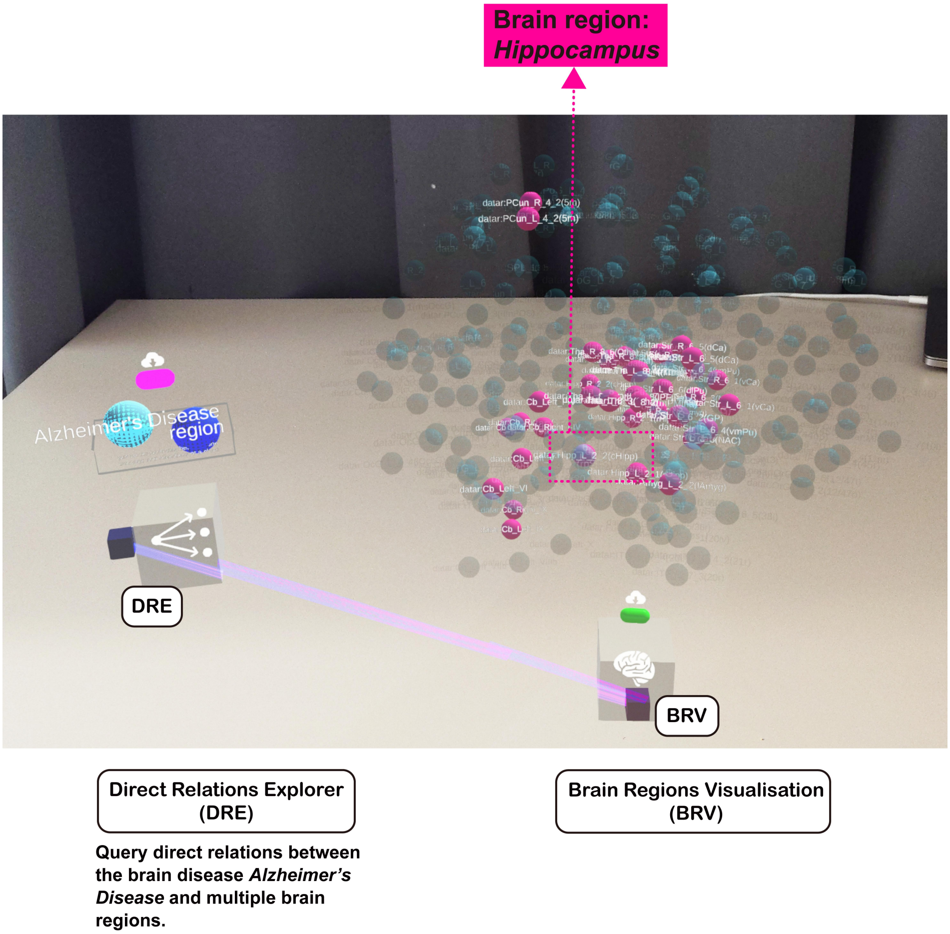

Figure 4.3: A user uses the *Direct Relations Explorer* to explore the direct relation between the brain disease *Alzheimer's Disease* and all brain regions. Thirty-six direct relations (pink spheres) are shown in the *Brain Regions Visualisation*, including two pink spheres representing both *Hippocampus* brain regions. Brain regions with no direct relations to *Alzheimer's Disease* are visualised as light blue spheres.

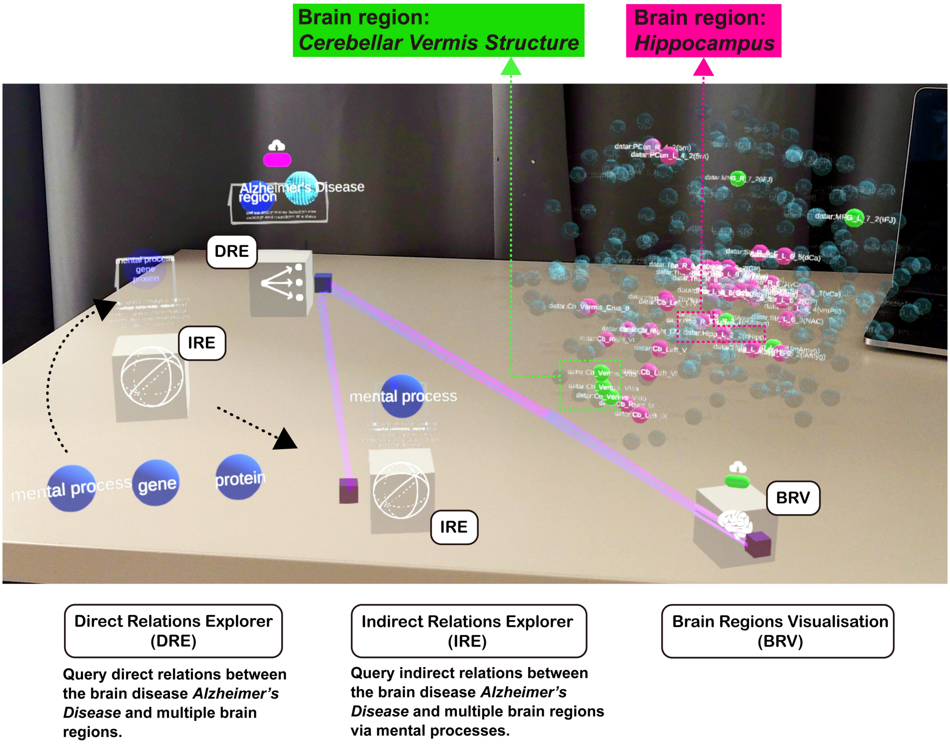

Figure 4.4: A user uses the *Indirect Relations Explorer* to select the intermediate topic type "Mental process" to explore indirect relations between the brain disease *Alzheimer's Disease* and all brain regions. Seven indirect relations (green spheres) are shown in the *Brain Regions Visualisation*, including three green spheres representing the three lobes of the brain region *Cerebellar Vermis Structure*. Brain regions with direct relations to *Alzheimer's Disease* are visualised as pink spheres. Brain regions with no direct or indirect relations to *Alzheimer's Disease* are visualised as light blue spheres.

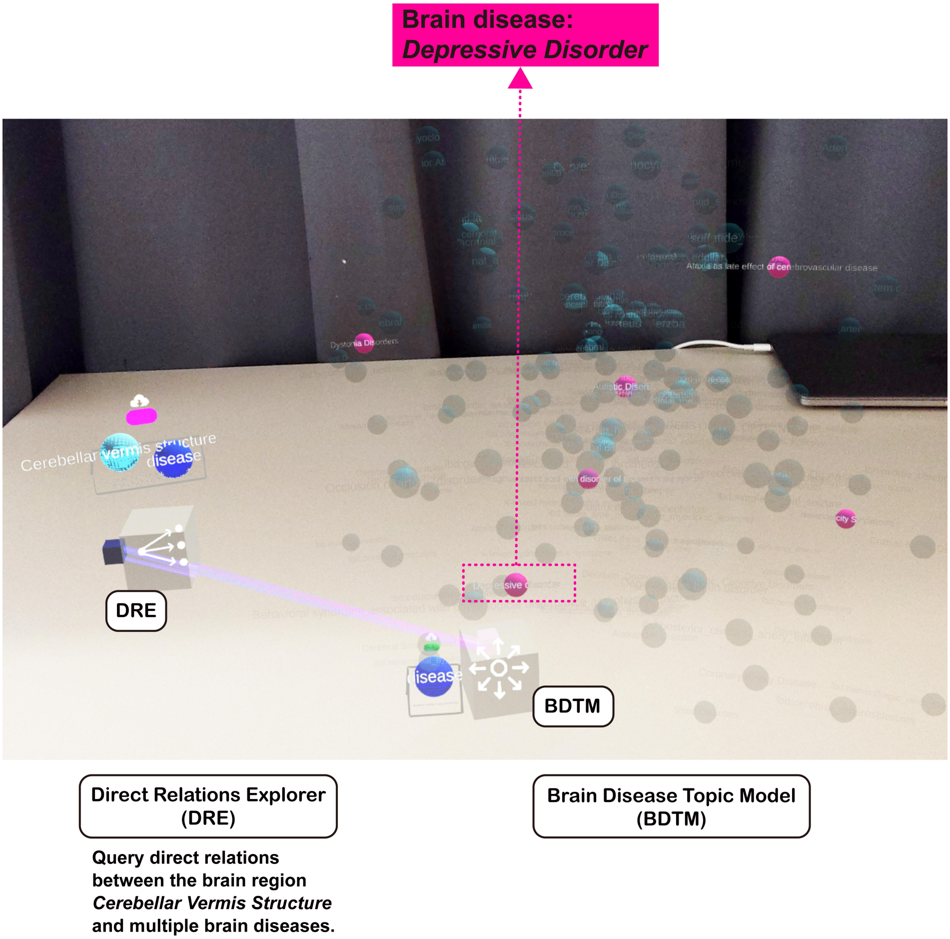

Figure 4.5: A user uses the *Direct Relations Explorer* to explore the direct relations between the brain region *Cerebellar Vermis Structure* and all brain diseases. Six direct relations (pink spheres) are shown in the *Brain Disease Topic Model*, including a pink sphere representing the brain disease *Depressive Disorder*. Brain regions with no direct relations to *Cerebellar Vermis Structure* are visualised as light blue spheres.

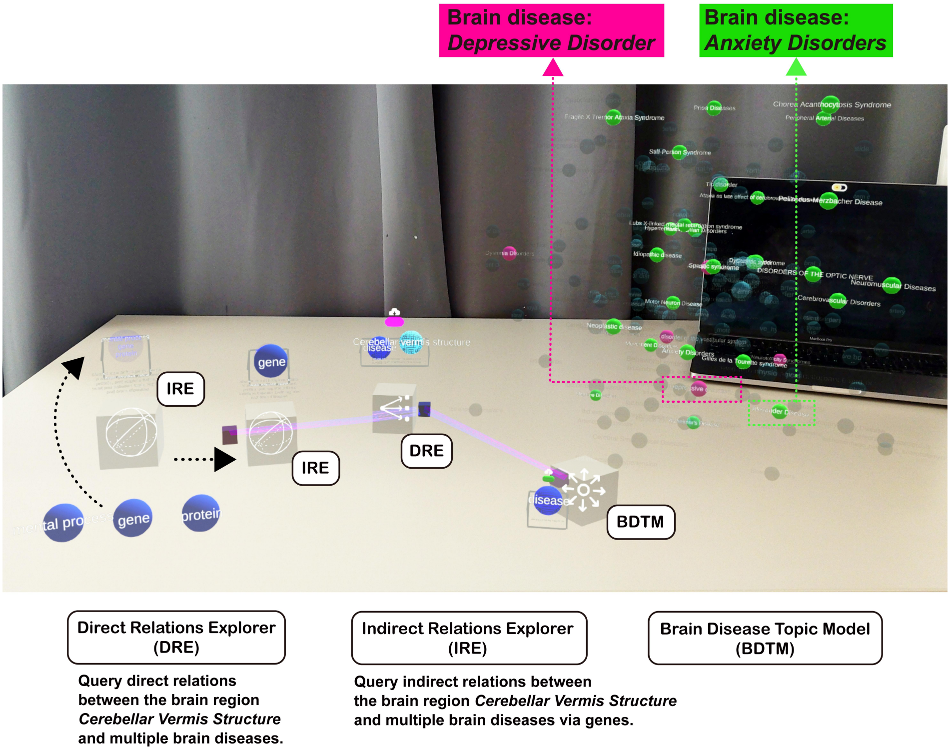

Figure 4.6: A user uses the *Indirect Relations Explorer* to select the intermediate topic type "Gene" to explore indirect relations between the brain region *Cerebellar Vermis Structure* and all brain diseases. Twenty-six indirect relations (green spheres) are shown in the *Brain Disease Topic Model*, including a green sphere representing the brain disease *Anxiety Disorders*. Brain diseases with direct relations to *Cerebellar Vermis Structure* are visualised as pink spheres. Brain diseases with no direct or indirect relations to *Cerebellar Vermis Structure* are visualised as light blue spheres.

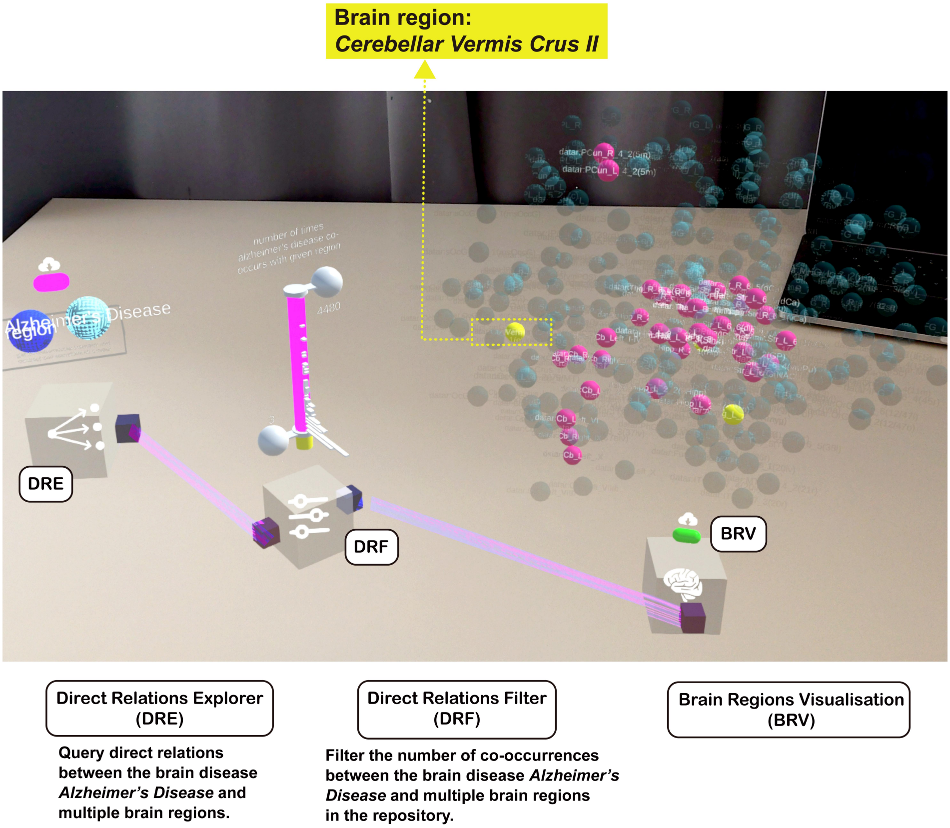

Figure 4.7: A user uses the *Direct Relations Filter* to filter the number of co-occurrences between the brain disease *Alzheimer's Disease* and all brain regions. Three brain regions co-occur with *Alzheimer's Disease* one, two, or three times, including *Cerebellar Vermis Crus II* (yellow sphere). The number of direct relations is shown using histograms labelled with the exact number in the *Direct Relations Filter*.

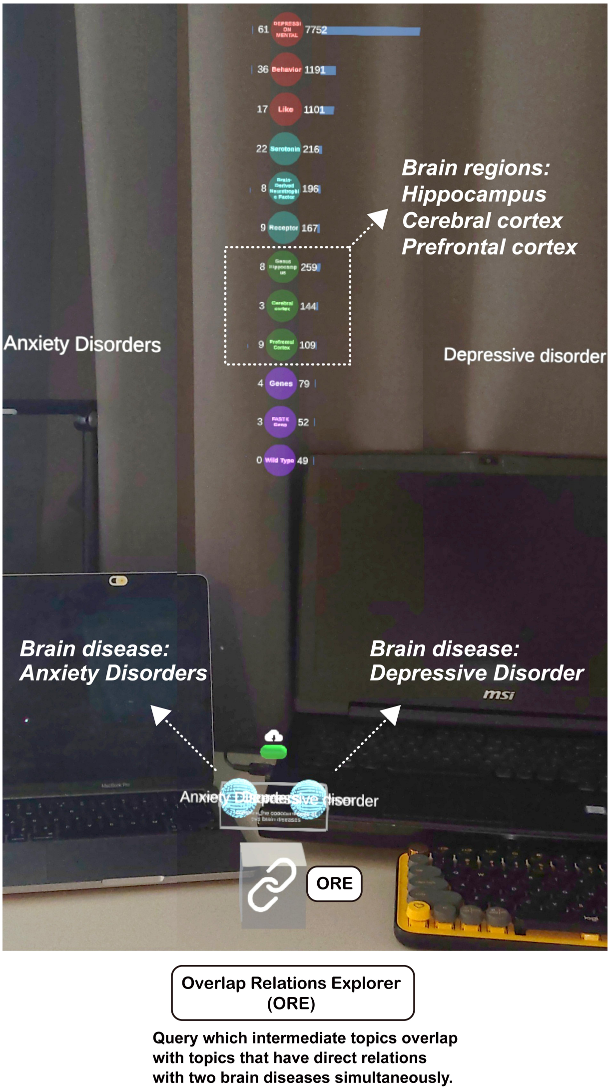

Figure 4.8: A user explores which topics have direct relations with two brain diseases, such as *Anxiety Disorders* and *Depressive Disorder*, simultaneously. Brain regions, such as *Hippocampus*, *Cerebral cortex* and *Prefrontal cortex*, have direct relations with both *Anxiety Disorders* (lower left) and *Depressive Disorder* (lower right) simultaneously

4.3. Method

We first establish that exploring indirect relations between two neuroscience topics through a specific intermediate topic is a useful task for neuroscientists (Sec. 4.3.1). To evaluate how this user task can support neuroscience research, we identify representative tasks that are both useful for neuroscientists and appropriate for our evaluation purposes (Sec. 4.3.2). To answer our three research questions, we identify three specific evaluation criteria for the visualisation (RQ3), functionality (RQ2) and the user task (RQ1) (Sec. 4.3.3). We then determine two functional requirements and one visual requirement needed to support exploration of indirect relations via a specific intermediate topic in the 3D AR prototype (Sec. 4.3.4). We discuss the design decisions to meet each functional and visual requirement (Sec. 4.3.5).

4.3.1. The Task of Exploring Indirect Relations via a Specific Intermediate Topic is Perceived as Useful

A neuroscientist participating in a previous study [108, 109] suggested that understanding the mechanics of identifying indirect relations by exploring specific intermediate topics could assist in designing useful experiments. In December 2024, we interviewed two neuroscientists (P1 and P2, Table 4.1 on p. 101) to determine whether the user task of exploring indirect relations via a specific intermediate topic is useful.

P1, a neurologist researching *Alzheimer's Disease*, wanted to explore brain regions not yet mentioned in the literature but potentially related to Alzheimer's Disease as a starting point for an experiment. P1 was aware that communication impairment is a clinical manifestation of *Alzheimer's Disease* [102]. They wanted to explore which brain regions might be indirectly connected to *Alzheimer's Disease* through the specific intermediate topic *Communication Response*. These indirect relations, established via a useful intermediate topic, offer a logical basis for experiment design. Providing such specific intermediate topics serves as sufficient evidence to investigate the potential usefulness of indirect relations.

P2, a neuroscientist interested in genetics, wanted to explore which specific gene is involved in the indirect relations between brain regions and brain diseases. This gene could provide neuroscientists with insights for potentially treating a certain brain disease, such as *Anxiety Disorders*. Both neuroscientists, plus one from previous studies [108, 109], were interested in exploring indirect relations via a specific intermediate topic to provide more inspiration for designing a useful experiment.

4.3.2. Determining Two Representative Tasks for Evaluation

To assess the functionality and visualisation for exploring indirect relations through specific intermediate topics, we interviewed P1 and P2 to identify suitable representative tasks. The two representative tasks [28] should:

- be useful to neuroscience, aligning with the types of questions neuroscientists investigate,

- be designed to be neither too simple nor overly complex, enabling a thorough and challenging evaluation of our visualisation and functionality,
- avoid presenting an excessive number of brain regions or diseases, as this may overwhelm users and hinder their ability to find useful information.

Our two representative tasks were selected based on interviews with two neuroscientists to ensure that users could uncover useful indirect relations without being overloaded. These two representative tasks were validated by neuroscientists to support effective engagement during first-time use.

Representative Task 1 (RT1): Exploring the indirect relations between the brain disease *Alzheimer's Disease* and all brain regions, such as *Cerebellar Vermis Structure*, through a mental process *Communication Response*

P1 researches *Alzheimer's Disease* and mentioned the connection between *Alzheimer's Disease* and *Communication Response* in Sec. 4.3.1. P1 is interested in exploring which brain regions have indirect relations with *Alzheimer's Disease* through a specific mental process - *Communication Response*, Fig. 4.9 on p. 94.

This indirect relation between *Alzheimer's Disease* and *Cerebellar Vermis Structure* via *Communication Response* involves 137 direct relations between *Alzheimer's Disease* and *Communication Response*, and 41 direct relations between *Communication Response* and *Cerebellar Vermis Structure*. Without the support of our prototype, identifying this indirect relation would require manually searching and screening these 178 direct relations across hundreds of papers.

Representative Task 2 (RT2): Exploring the indirect relations between the brain region *Cerebellar Vermis Structure* and all brain diseases, such as *Anxiety Disorders*, through a gene *LITAF Gene*

P2, who is interested in genetics, proposed selecting a specific gene to explore indirect relations between brain regions and brain diseases. These indirect relations could inspire their understanding of potentially useful genes for treating brain diseases or influencing brain regions. During the discussion with P2, they carried out RT1 and was drawn to the brain regions. P2 selected the brain region *Cerebellar Vermis Structure*, which is connected to their research, as an example to explore which brain diseases have indirect relations with *Cerebellar Vermis Structure* through a specific gene - *LITAF Gene*, Fig. 4.10 on p. 95.

This indirect relation between *Cerebellar Vermis Structure* and *Anxiety Disorders* via *LITAF Gene* involves 8 direct relations between *Cerebellar Vermis Structure* and *LITAF Gene*, and 1 direct relation between *LITAF Gene* and *Anxiety Disorders*. Without the support of our prototype, identifying this less observed indirect relation would require manually searching and screening these 9 direct relations across numerous publications.

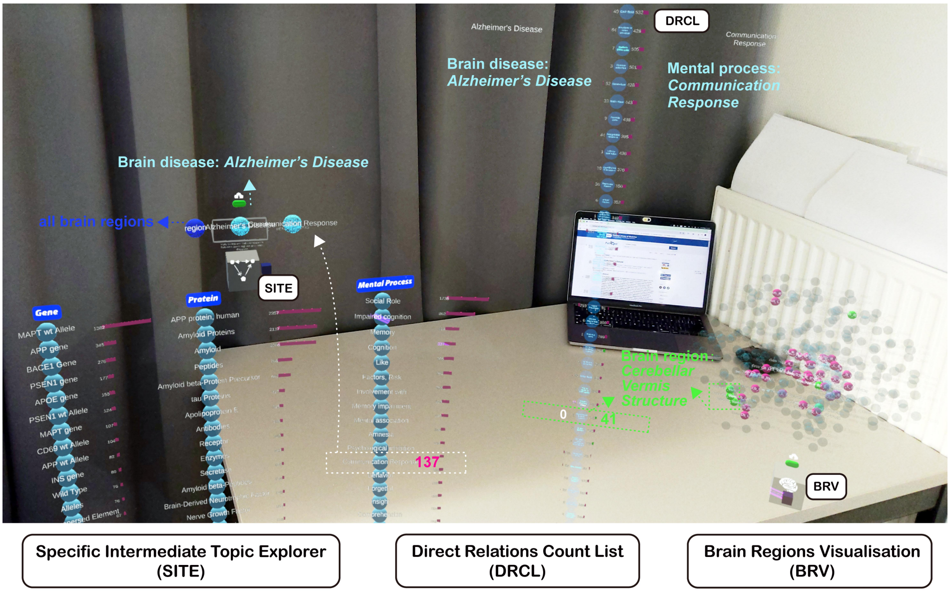

Figure 4.9: A user explores the direct and indirect relations between the brain disease *Alzheimer's Disease* (upper left quadrant) and all brain regions ("BRV" lower right quadrant). They select the Mental Process *Communication Response* as the intermediate topic in the *Specific Intermediate Topic Explorer* (SITE). There is an indirect relation between *Alzheimer's Disease* and the brain region *Cerebellar Vermis Structure* (indicated by three green spheres in the *Brain Regions Visualisation*). This indirect relation is identified based on (a) 137 direct relations between *Alzheimer's Disease* and *Communication Response*, and (b) 41 direct relations between *Communication Response* and *Cerebellar Vermis Structure*.
Thirty-five brain regions with direct relations to *Alzheimer's Disease* in the *Brain Regions Visualisation* (BRV) are visualised as pink spheres. The other 233 brain regions with no direct or indirect relations to *Alzheimer's Disease* are visualised as light blue spheres.

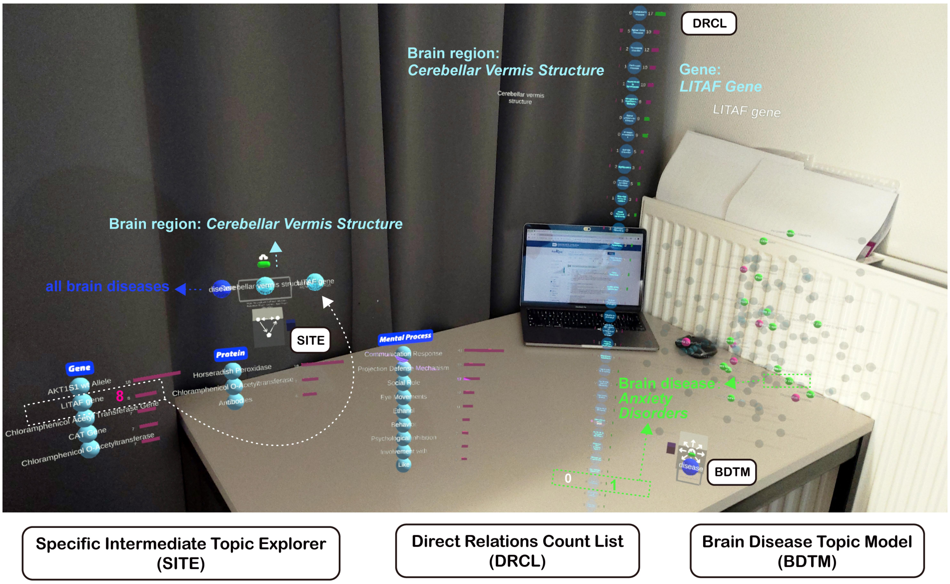

Figure 4.10: A user explores the direct and indirect relations between the brain region *Cerebellar Vermis Structure* (upper left quadrant) and all brain diseases ("BDTM" lower right quadrant). They select the Gene *LITAF Gene* as the intermediate topic in the *Specific Intermediate Topic Explorer* (SITE). There is an indirect relation between *Cerebellar Vermis Structure* and the brain disease *Anxiety Disorders* (indicated by a green sphere in the *Brain Disease Topic Model*). This indirect relation is identified based on (a) 8 direct relations between *Cerebellar Vermis Structure* and *LITAF Gene*, and (b) one direct relation between *LITAF Gene* and *Anxiety Disorders*.
Six brain diseases with direct relations to *Cerebellar Vermis Structure* in the *Brain Disease Topic Model* (BDTM) are visualised as pink spheres. The other 131 brain diseases with no direct or indirect relations to *Cerebellar Vermis Structure* are visualised as light blue spheres.

4.3.3. Identifying Evaluation Criteria for the Visualisation (RQ3), Functionality (RQ2) and the User Task (RQ1)

We employ semi-structured interviews [2, 83] to assess the functionality, visualisation, and user task, and to answer our research questions. As a qualitative research method, semi-structured interviews combine predefined questions with the flexibility to adapt in real time to participants' responses. This approach supports the elicitation of targeted feedback aligned with the research questions, while enabling deeper exploration when unexpected or insightful input emerges [45].

- ***Suitability*** [104]evaluates how the designed visualisation supports the presentation of indirect relations via a specific intermediate topic within the AR environment (RQ3),
- ***Usability*** [40, 59] measures how the implemented functionality satisfies human-

computer interaction requirements for exploring indirect relations via a specific intermediate topic (RQ2),

- ***Usefulness*** [68] assesses how exploring indirect relations via a specific intermediate topic aids neuroscientists in designing a useful experiment (RQ1).

To enable a comparable analysis of participants' perceptions, we utilise a 7-point Likert scale [9, 48] to quantify responses to ten questions on the visualisation (RQ3), functionality (RQ2), and the user task (RQ1). To analyse the results, we use the median, as it provides a clearer representation of "typical user" behaviour [105].

4.3.4. Functional (RQ2) and Visual (RQ3) Requirements for Exploring Indirect Relations via a Specific Intermediate Topic

To support the two representative tasks (Sec. 4.3.2) within the 3D AR prototype, we need to design and implement appropriate functionality. To inform the functionality design, we discussed potential functional requirements with P1 and P2 (Table 4.1 on p. 101) in December 2024.

- **Functional Requirement 1 (FR1):** ***Select a specific intermediate topic***

  Both P1 and P2 expressed interest in manually selecting specific intermediate topics, such as *Communication Response* or *LITAF Gene*, to investigate through which mental processes or genes indirect relations between brain regions and diseases are established. Starting with a specific intermediate topic enables neuroscientists to better understand the cognitive processes and biological mechanisms underlying a given indirect relation. Rather than beginning with a broad set of indirect relations and then retrospectively identifying intermediates, P2 suggested that selecting a specific and useful intermediate topic from the outset can more effectively identify potentially useful indirect relations, which can directly inform experimental design.

- **Functional Requirement 2 (FR2):** ***Find out the number of two direct relations that indicate an indirect relation***

  P1 proposed providing the number of direct relations between (1) a selected brain region or brain disease and a specific intermediate topic, and (2) the same intermediate topic and a brain disease or brain region. These two direct relations indicate an indirect relation. The usefulness of an indirect relation may be associated with how frequently these two direct relations are mentioned in the literature. If both direct relations are frequently reported, this may suggest that the identified indirect relation is more likely to find support in existing evidence. Providing the number of the two direct relations that indicate an indirect relation may help neuroscientists assess its usefulness and prioritise which indirect relations to explore further.

In the same interview with P1 and P2, Table 4.1 on p. 101, in December 2024, they suggested a visual requirement for presenting indirect relations via a specific intermediate topic.

- **Visual Requirement 1 (VR1): *Present direct and indirect relations between topics in a single visualisation***

  Both P1 and P2 expressed a preference for visualising direct and indirect relations in a single visualisation. They thought that if direct relations (represented by pink spheres) and indirect relations (represented by green spheres) are shown separately in different visualisations, users need to compare across views to understand their connections, which is less efficient. By visualising direct and indirect relations in a single visualisation, neuroscientists can intuitively identify which topics are explored in the literature and which indirect relations serve as potential hypotheses for designing experiments.

4.3.5. Functional and Visual Design Rationale

Having provided the functional and visual requirements (Sec. 4.3.4) for exploring indirect relations via a specific intermediate topic, we discuss the design decisions made to meet the functional and visual requirements.

- ***Specific Intermediate Topic Explorer* Functional Design (FR1)**

  To support the selection of a specific intermediate topic for exploring indirect relations, we first consider the following design options:

  - *Select a topic type, followed by a specific intermediate topic*

    This design allows users to first select a topic type, such as mental processes or genes, and then refine their selection by choosing a specific intermediate topic within that category, such as *Communication Response* or *LITAF Gene*.

  - *Present a grouped list of specific intermediate topics*

    This design presents users with a categorised list of specific intermediate topics, such as *Communication Response*, organised under a topic type, such as mental processes. It enables users to directly identify intermediate topics across various categories for exploring indirect relations.

  Designing the *Specific Intermediate Topic Explorer* (Fig. 4.11 on p. 100) is driven by several reasons:

  - *Design the grouped specific intermediate topics lists*

    This study investigates specific intermediate topics to identify potentially useful indirect relations for the design of the next neuroscience experiment. Although the step-by-step selection design reflects that neuroscientists often specialise in different research topics and helps narrow down intermediate topics, it limits the scope to familiar areas of expertise. Listing grouped specific intermediate topics provides an intuitive overview of the different categories of intermediate topics. This design encourages users to explore indirect relations via intermediate topics beyond their familiar areas. It enables the identification of potentially useful indirect relations by selecting intermediate topics that may not have previously been considered in the user's own topic.

Additionally, the *Direct Relations Filter* (Fig. 4.7 on p. 90) in the 3D AR prototype employs histograms to represent the number of direct relations. We preserved this histogram layout in the design of the intermediate topics selector to visualise the number of direct relations between a selected brain region (or brain disease) and an intermediate topic, such as a gene, a protein, or a mental process. These counts of direct relations assist neuroscientists in identifying frequently observed genes and mental processes, which may provide useful clues for uncovering corresponding indirect relations, Fig. 4.11 on p. 100.

- ***Direct Relations Count List* Functional Design (FR2)**

  The *Overlap Relations Explorer* (Fig. 4.8 on p. 91) in the 3D AR prototype displays a list of intermediate topics that have direct relations with two brain diseases simultaneously, and the corresponding number of direct relations is shown as histograms. We retained this vertical layout of lists and histograms to design a similar view that presents two direct relations. The left side of the list shows pink histograms representing the number of direct relations between the selected brain region or brain disease and all brain diseases or brain regions (*Direct Relations Count List*, Figs. 4.9 and 4.10 on pp. 94 and 95). The right side of the list shows pink histograms representing the number of direct relations between the intermediate topic and all brain diseases or brain regions (*Direct Relations Count List*, Figs. 4.9 and 4.10 on pp. 94 and 95). If a brain disease or brain region does not have direct relations with the selected brain region or brain disease, but that brain disease or brain region has direct relations with the intermediate topic, the corresponding histogram on the right appears in green (*Direct Relations Count List*, Figs. 4.9 and 4.10 on pp. 94 and 95). This indicates that the brain disease or brain region has an indirect relation with the selected brain region or brain disease via the intermediate topic. The name labels in the list display brain diseases or brain regions that have direct or indirect relations with the selected brain region or brain disease. These brain diseases or brain regions are also visualised in the 3D representation within the *Brain Disease Topic Model* or the *Brain Regions Visualisation.*

- **Present Direct and Indirect Relations in a Single Visualisation (VR1)**

  This study extends the existing 3D AR prototype (Sec. 4.2.3) by implementing functionalities and corresponding visualisations that support the exploration of indirect relations via a specific intermediate topic. To ensure visual consistency, particularly as some participants took part in both the previous and current evaluations, it is useful to maintain a coherent colour scheme throughout the prototype. The prototype employs a colour scheme: light blue[5]indicates no direct and indirect relations, pink[6] represents direct relations, and green[7] represents indirect relations (Figs. 4.4 and 4.6 on pp. 87 and 89). We present direct and indirect relations within a single visualisation by integrating the *Specific Intermediate Topic Explorer* with either the *Brain Disease Topic Model* or the *Brain Regions*

[5]RGB: 80, 215, 230
[6]RGB: 255, 50, 200
[7]RGB: 51, 255, 51

Visualisation (Figs. 4.9 and 4.10 on pp. 94 and 95). Upon selecting an intermediate topic (such as *Communication Response*), the system immediately visualises the direct (pink spheres) and indirect (green spheres) relations between the selected brain region or brain disease and all other brain diseases or brain regions. Users can switch between intermediate topics and view the updated relations in real time. Although the red-green colour scheme may present challenges for colour-blind users, it remains visually effective within the AR environment.

4.4. Implementation

Given the existing widgets in the 3D AR prototype (Sec. 4.2.3), we only needed to implement a single new widget for this study, the *Specific Intermediate Topic Explorer* (Fig. 4.11 on p. 100), which allows users to select a specific intermediate topic to explore indirect relations. The functionality and visualisation of the *Specific Intermediate Topic Explorer*[8] were developed using the Optical See-Through Head-Mounted Display (OST-AR) visualisation tool, built with Unity3D[9] (v2020.3.15f2) and the Microsoft Mixed Reality Toolkit (MRTK)[10] (v2.7.0). The system is deployed on the HoloLens 2, an Augmented Reality device developed and manufactured by Microsoft.

4.5. Evaluation of Visualisation (RQ3), Functionality (RQ2), and User Task (RQ1)

To address our three research questions, participants evaluated the visualisation and implemented functionality through two representative tasks (Sec. 4.3.2), followed by an assessment of the usefulness of exploring indirect relations via a specific intermediate topic in neuroscience literature exploration.

We describe participant recruitment in Sec. 4.5.1 and outline the evaluation procedure for the three research questions in Sec. 4.5.2. We then present the questions posed to participants for each individual research question:

- The suitability of the visualisation in the *Specific Intermediate Topic Explorer* for presenting indirect relations via a specific intermediate topic (RQ3) in Sec. 4.5.3,
- The usability of the implemented functionality for exploring indirect relations via a specific intermediate topic (RQ2) in Sec. 4.5.4, and
- The usefulness of exploring indirect relations via a specific intermediate topic in understanding indirect relations for designing a potentially useful experiment (RQ1) in Sec. 4.5.5.

[8] The Unity-based code for exploring indirect relations via a specific intermediate topic is open-source and available on GitHub: https://github.com/DatAR-prototype/tree/IntermediateTopics-project

[9] Unity3D https://unity.com/

[10] Microsoft Mixed Reality Toolkit (MRTK) https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05

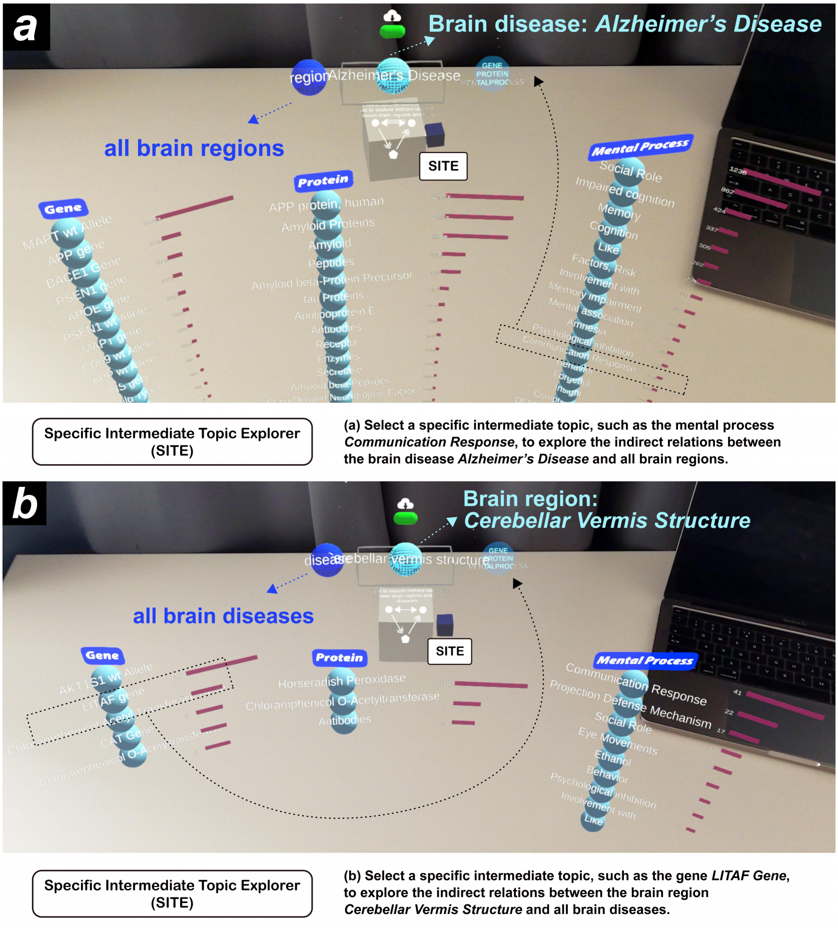

Figure 4.11: The *Specific Intermediate Topic Explorer* enables users to explore indirect relations by selecting a specific intermediate topic.
(a) A participant selects a specific intermediate topic, such as *Communication Response*, to explore which brain regions have indirect relations with the brain disease *Alzheimer's Disease*. Grouped lists display the specific genes, proteins, and mental processes that have direct relations with *Alzheimer's Disease*. Histograms show the number of direct relations between *Alzheimer's Disease* and each gene, protein, and mental process.
(b) A participant selects a specific intermediate topic, such as *LITAF Gene*, to explore which brain diseases have indirect relations with the brain region *Cerebellar Vermis Structure*. Grouped lists display the specific genes, proteins, and mental processes that have direct relations with *Cerebellar Vermis Structure*. Histograms show the number of direct relations between *Cerebellar Vermis Structure* and each gene, protein, and mental process.

4.5.1. Participants

Studies for problem discovery typically involve between 5 and 20 participants, since this range balances resource constraints with the need to evaluate the usefulness of tasks [89]. Recruiting 5 to 10 participants generally uncovers the most common problems and insights, while adding more participants beyond 15 tends to yield diminishing returns, as fewer new issues emerge [89].

We recruited nine participants[11] (P3-P11) from neuroscience institutes at three universities, Table 4.1 on p. 101. Four participants (P3-P6) had previously taken part in an earlier study [109], while the remaining five (P7-P11) were newly recruited. While new participants provide fresh perspectives, returning participants can provide comparative feedback on earlier versions. Participants' academic backgrounds, familiarity with AR, and experience in literature exploration are summarised in Table 4.1 on p. 101. Eight participants were familiar with or professional in neuroscience literature exploration (P3-P10). Four participants (P3-P6) reported having some understanding of AR and prior experience using the AR headset, the remaining five participants (P7-P11) had no prior knowledge or experience with AR. All 9 participants had normal or corrected-to-normal vision and were not colour-blind.

Table 4.1: Backgrounds of participants and studies they participated in. P1-P2 assisted in identifying the useful user task used in this study (Sec. 4.3.1), determining two representative tasks for evaluation (Sec. 4.3.2) and determining the functional and visual requirements (Sec. 4.3.4). Nine participants (P3-P11) carried out the two representative tasks to evaluate the suitability of the corresponding visualisation (RQ3), the usability of the functionality (RQ2), and the usefulness of the user task of exploring indirect relations via a specific intermediate topic (RQ1).

Participant Numbers	**Research Topics**	**Studies Participated in**	**Academic Qualifications**	**Knowledge and experience in AR technology**	**Experience in Literature Exploration**
P1	Neurology, in particular, *Alzheimer's Disease*	Identify the user task of exploring indirect relations via specific intermediate topics (Sec. 4.3.1); Determine two representative tasks (Sec. 4.3.2); Determine functional and visual requirements (Sec. 4.3.4)	PhD student (first year)	None	Familiar with literature exploration

[11] Seven participants provided responses in Chinese, the other four participants responded in English. The first author translated the Chinese responses into English for analysis and reporting. The original Chinese transcripts are available upon request.

P2	Neurobiology	Identify the user task of exploring indirect relations via specific intermediate topics (Sec. 4.3.1); Determine two representative tasks (Sec. 4.3.2); Determine functional and visual requirements (Sec. 4.3.4)	PhD student (second year)	None	Familiar with literature exploration
P3	Neurology, in particular, *Epilepsy disorders*	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (fourth year)	Some knowledge of AR; Some experience in the AR headset (attended a previous study on exploring indirect relations in October 2022 [109])	Professional in literature exploration
P4	Neurology, in particular, *Epilepsy disorders*	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (fourth year)	Some knowledge of AR; Some experience in the AR headset (attended a previous study on exploring indirect relations in October 2022 [109])	Familiar with literature exploration
P5	Neuroscience, Experimental Psychology, Cognitive Psychology	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (second year)	Some knowledge of AR; Some experience in the AR headset (attended a previous study on exploring indirect relations in October 2022 [109])	Professional in literature exploration
P6	Clinical Neuroscience	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (third year)	Some knowledge of AR; Some experience in the AR headset (attended a previous study on exploring indirect relations in October 2022 [109])	Familiar with literature exploration

P7	Neuroscience	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (first year)	None	Familiar with literature exploration
P8	Neuroscience	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (first year)	None	Familiar with literature exploration
P9	Neurobiology	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (first year)	None	Familiar with literature exploration
P10	Neurology, in particular, *Depressive Disorder, Relapse, Cognitive Intervention*, and the mechanisms underlying these interventions	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (second year)	None	Familiar with literature exploration
P11	Neurology	Evaluate the indirect relations via a specific intermediate topic (Sec. 4.5)	PhD student (first year)	None	Little experience in literature exploration

4.5.2. Procedure

Each evaluation session lasted approximately 60 minutes and was conducted in March and April 2025. The procedure was as follows:

1. We collected background information from participants, including their research fields, academic qualifications, familiarity with AR, and experience in literature exploration (Table 4.1 on p. 101). All participants signed an informed consent form (Appendix C on p. 139). Age and gender were not collected, as they are not relevant to this study [11].

2. We introduced the goal of the evaluation session, which was to identify potentially useful indirect relations via a specific intermediate topic, and explained the process of exploring indirect relations between brain regions and brain diseases through a specific intermediate topic (Appendix C on p. 141).

3. Participants wore the AR headset (HoloLens 2). New participants (P7-P11) were given approximately 10 minutes to practise basic AR interactions, such as grabbing, dragging, and rotating 3D objects, using the Brain Regions Visualisation. Participants who had participated in an earlier study (P3-P6) proceeded directly to the two representative tasks.

4. Participants performed the two representative tasks (Sec. 4.3.2) using the *Specific Intermediate Topic Explorer* to explore indirect relations between brain regions

and brain diseases via a specific mental process or gene (see Figs. 4.9 and 4.2 on pp. 94 and 95).

5.
   - A. While performing the tasks, participants evaluated:
     - the suitability of the visualisation designed for presenting indirect relations via a specific intermediate topic (RQ3) (Sec. 4.5.3),
     - the usability of the functionality developed for exploring indirect relations via a specific intermediate topic (RQ2) (Sec. 4.5.4), and
     - the usefulness of exploring indirect relations via specific intermediate topics for designing a useful experiment (RQ1) (Sec. 4.5.5).
   - B. After completing the two representative tasks (approximately 40 minutes), participants were invited to select brain regions or brain diseases of personal interest and explore related direct and indirect relations via specific intermediate topics using the interactive AR prototype during an additional 15-minute session.

   Participants responded to 18 semi-structured interview questions (IQs) (listed in Secs. 4.5.3, 4.5.4 and 4.5.5) while performing the tasks using the 3D AR prototype. Participants were encouraged to verbalise their thoughts to help the experimenter understand their problem-solving process [31, 49].

6. Participants were asked to rate the suitability of the visualisation, the usability of the functionality, and the usefulness of the user task by responding to ten questions (listed in Secs. 4.5.3, 4.5.4 and 4.5.5) using a 7-point Likert scale [48, 74] ranging from *Worst Imaginable* to *Best Imaginable* (*Worst Imaginable*, *Awful*, *Poor*, *OK*, *Good*, *Excellent*, *Best Imaginable*).

4.5.3. Questions for the Suitability of Visualisation for Presenting Indirect Relations via a Specific Intermediate Topic (RQ3)

Participants answered ten semi-structured interview questions to evaluate the suitability of the *Specific Intermediate Topic Explorer* for presenting indirect relations via a specific intermediate topic while performing two representative tasks in the 3D AR prototype (IQ1-IQ10, Table 4.2 on p. 105). They were then asked to quantify their responses regarding the suitability of the visualisation by answering two questions using a 7-point Likert scale (Q1 and Q2, Table 4.2 on p. 105).

4.5.4. Questions for the Usability of Functionality for Exploring the Indirect Relations via a Specific Intermediate Topic (RQ2)

Participants answered three semi-structured interview questions to evaluate the usability of the *Specific Intermediate Topic Explorer* for exploring indirect relations via a specific intermediate topic while performing two representative tasks in the 3D AR prototype (IQ11-IQ13, Table 4.3 on p. 106). They were then asked to quantify their responses on the usability of the functionality designed for exploring indirect relations via a specific intermediate topic by answering three questions using a 7-point Likert scale (Q3-Q5, Table 4.3 on p. 106).

Table 4.2: Ten semi-structured interview questions and two 7-point Likert scale questions for assessing the suitability of the visualisation for presenting indirect relations via a specific intermediate topic (RQ3).

IQs	**Ten Semi-Structured Interview Questions to Assess the Suitability of Visualisation (RQ3)**
IQ1	To what extent are the names of the brain regions readable to you?
IQ2	What are the full names of the brain regions described as *Cb_Vermis* and *Amyg_L*?
IQ3	Do you think the 3D positions of *Cb_Vermis* and *Amyg_L* are logically aligned with their actual anatomical locations in the human brain?
IQ4	To what extent are the names of the brain diseases readable to you?
IQ5	Can you locate the brain disease *Depressive Disorder*?
IQ6	Do you think the 3D positions of *Cb_Vermis* and *Amyg_L* are logically aligned with their actual anatomical locations in the human brain?
IQ7	Do you find the visualisation of presenting the specific intermediate topics suitable? Please comment on the colour and layout.
IQ8	Are the lists of intermediate topics for selection and the histogram of co-occurrence numbers suitable?
IQ9	Do you have any suggestions for improving the visualisation of the *Specific Intermediate Topic Explorer*?
IQ10	To what extent do you find the visualised direct and indirect relations, via a specific intermediate topic, clear and understandable?
Qs	**Two Questions Using a 7-Point Likert Scale to Assess the Suitability of Visualisation (RQ3)**
Q1	How would you rate the visual representations in the *Specific Topic Intermediate Explorer*?
Q2	How clear and understandable do you find the visualisations for presenting direct and indirect relations via a specific intermediate topic?

4.5.5. Questions for the Usefulness of Exploring Indirect Relations via a Specific Intermediate Topic (RQ1)

Participants answered five semi-structured interview questions to evaluate the usefulness of exploring indirect relations via a specific intermediate topic for informing the design of a useful experiment (IQ14-IQ18, Table 4.4 on p. 107). They then quantified their opinions by answering five questions using a 7-point Likert scale (Q6-Q10, Table 4.4 on p. 107).

4.6. Results

In the following sub-sections, we report the results that address our three research questions: the suitability of the *Specific Intermediate Topic Explorer* visualisation (RQ3) in Sec. 4.6.1, the usability of the *Specific Intermediate Topic Explorer* functionality (RQ2) in Sec. 4.6.2, and the usefulness of the user task of exploring indirect relations via a specific intermediate topic in neuroscience literature exploration (RQ1) in Sec. 4.6.3.

Table 4.3: Three semi-structured interview questions and three 7-point Likert scale questions for assessing the usability of the functionality for exploring indirect relations via a specific intermediate topic (RQ2).

IQs	**Three Semi-Structured Interview Questions to Assess the Usability of Functionality (RQ2)**
IQ11	To what extent is the interaction procedure for exploring indirect relations via a specific intermediate topic clear and understandable?
IQ12	How do you think the functionality of the *Specific Intermediate Topic Explorer* contributes to understanding the relations between topics in literature exploration?
IQ13	Do you have any suggestions for improving the functionality of the *Specific Intermediate Topic Explorer*?
Qs	**Three Questions Using a 7-Point Likert Scale to Assess the Usability of Functionality (RQ2)**
Q3	How understandable do you find the navigation of the *Specific Intermediate Topic Explorer* functionality?
Q4	How would you rate the usability of the *Specific Intermediate Topic Explorer* functionality for exploring indirect relations via a specific intermediate topic?
Q5	How willing are you to use the designed functionalities to explore indirect relations via a specific intermediate topic?

4.6.1. The Suitability of the *Specific Intermediate Topic Explorer* Visualisation for Presenting Indirect Relations via a Specific Intermediate Topic (RQ3)

In response to the questionnaire assessing the usability of the *Specific Intermediate Topic Explorer* visualisation described in Sec. 4.5.3, all nine participants stated that the *Specific Intermediate Topic Explorer* visualisation is suitable for presenting indirect relations through specific intermediate genes, proteins, and mental processes (Q1 in Fig. 4.12 on p. 108). The median rating for the *Specific Intermediate Topic Explorer* visualisation is *Good*, with responses ranging from *OK* to *Best Imaginable* (Q1 in Fig. 4.12 on p. 108).

All nine participants found that showing direct and indirect relations in a single visualisation of the *Brain Regions Visualisation* or the *Brain Disease Topic Model* is clear and understandable (Q2 in Fig. 4.12 on p. 108). The median rating for presenting direct and indirect relations between topics is *Excellent*, with responses ranging from *Good* to *Excellent* (Q2 in Fig. 4.12 on p. 108). P5 remarked (in response to IQ10, Table 4.2 on p. 105), "The visualisations for direct and indirect relations could be distinguished clearly and understandably in the *Topic Model* and the *Brain Regions Visualisation* through colours, with pink spheres for direct relations and green spheres for indirect relations."

Detailed results for the visual representations, such as layout and colour, in the *Specific Intermediate Topic Explorer* are as follows:

Table 4.4: Five semi-structured interview questions and five 7-point Likert scale questions for assessing the usefulness of exploring indirect relations via a specific intermediate topic for informing the design of a useful experiment (RQ1).

IQs	**Five Semi-Structured Interview Questions to Assess the Usefulness of the User Task (RQ1)**
IQ14	Are there any other indirect relations related to *Alzheimer's Disease* that you find useful?
IQ15	Are there any other indirect relations related to *Cerebellar Vermis Structure* that you find useful?
IQ16	To what extent do you think exploring indirect relations via a specific intermediate topic is worthwhile for literature exploration?
IQ17	Can you share some potentially useful results of indirect relations that you discovered using the designed *Specific Intermediate Topic Explorer*?
IQ18	What other information would you like to explore further?
Qs	**Five Questions Using a 7-Point Likert Scale to Assess the Usefulness of the User Task (RQ1)**
Q6	How would you rate the usefulness of exploring indirect relations via a specific intermediate topic for designing a useful experiment?
Q7	How would you rate the usefulness of an indirect relation between the brain disease *Alzheimer's Disease* and the brain region *Cerebellar Vermis Structure* via *Communication Response*?
Q8	How would you rate the usefulness of an indirect relation between the brain region *Cerebellar Vermis Structure* and the brain disease *Anxiety Disorders* via *LITAF Gene*?
Q9	How would you rate your willingness to use the 3D AR prototype in your daily literature exploration?
Q10	How would you rate your willingness to recommend the 3D AR prototype to your colleagues?

- **A grouped list of specific intermediate topics**

  All nine participants agreed that separating intermediate topics into three lists - genes, proteins, and mental processes - is a suitable approach for selecting an intermediate topic (in response to IQ8, Table 4.2 on p. 105). P3 remarked, "I am interested in cognition (in the mental processes) and also in genes, so it is very easy for me to select a mental process like *Communication Response* to investigate the indirect relations, and I could also switch to selecting a specific gene, such as *BACE1 Gene*, to explore indirect relations." P7 stated, "Showing specific genes, proteins, and mental processes simultaneously gives me the chance to try some intermediate topics that are not familiar in my research, but could still reveal useful indirect relations."

  Seven participants found that the vertical layout of the specific intermediate topics was suitable and aligned with their methods of browsing information (in response to IQ7, Table 4.2 on p. 105). P4 commented, "I am used to vertical lists when browsing menus, so it feels natural to navigate this vertical layout." However, two participants pointed out limitations when the lists became too long and exceeded the visible area. They noted that without scroll bars or page navigation buttons, it became difficult to access specific topics located further down the list. As P11 stated, “I cannot see the full list easily. If there were scrolling or a way to page down, it would be better.” P8 suggested that horizontally arranged lists could help users view the full set of topics more easily and avoid overwhelming them with overly long vertical lists.

- **Histograms of the number of direct relations**

  All nine participants thought that using histograms to display the co-occurrence numbers of direct relations between the selected brain disease or brain region and specific genes, proteins, and mental processes was suitable (in response to IQ8, Table 4.2 on p. 105). P4 remarked, “The list of genes, proteins, and mental processes is ordered by the number of co-occurrences with the selected brain disease or brain region, from highest to lowest, which is helpful. I am more interested in selecting an intermediate topic, such as a gene with a higher number of co-occurrences with *Alzheimer’s Disease*, to explore indirect relations between *Alzheimer’s Disease* and brain regions via that gene.” P4 reported that selecting intermediate topics with higher co-occurrence counts increases the possibility of finding an indirect relation.

- **Use of colours in the histograms of the *Specific Intermediate Topic Explorer***

  All nine participants found that the use of pink in the histograms of the *Specific Intermediate Topic Explorer* was suitable, as it indicated the direct relations between the selected brain region or brain disease and specific genes, proteins, and mental processes (in response to IQ7, Table 4.2 on p. 105). The colour scheme of the histograms is consistent with the existing 3D AR prototype (Figs. 4.4 and 4.6 on pp. 87 and 89), which uses two distinctive colours to represent direct and indirect relations: pink for direct relations and green for indirect relations.

  Regarding the histogram display for two direct relations (*Direct Relations Count List*, Figs. 4.9 and 4.10 on pp. 94 and 95), P5 reported that using only the green histogram is not enough to highlight the indirect relations; it would be better to change the background of the name labels to green to emphasise the indirect relations (*Direct Relations Count List*, Figs. 4.9 and 4.10 on pp. 94 and 95).

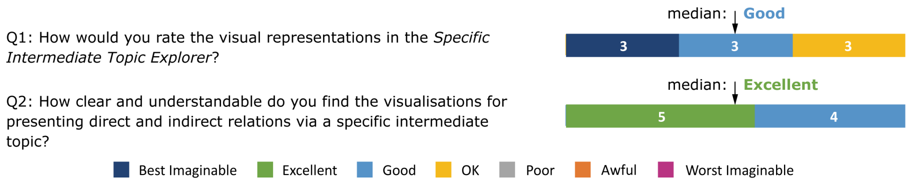

Figure 4.12: The distribution of all nine participants’ attitudes toward the *Specific Intermediate Topic Explorer* visualisation.

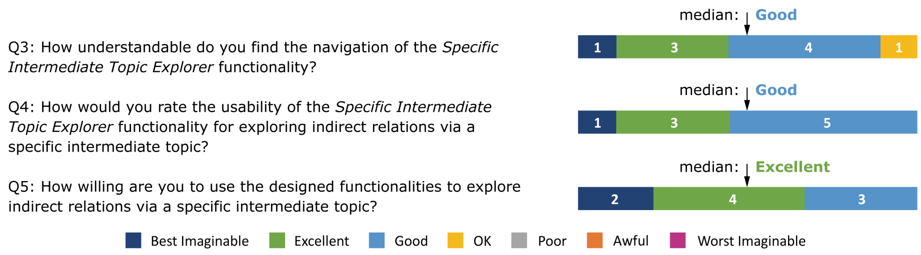

Figure 4.13: The distribution of all nine participants' attitudes toward the *Specific Intermediate Topic Explorer* functionality.

4.6.2. The Usability of the *Specific Intermediate Topic Explorer* Functionality for Exploring Indirect Relations via a Specific Intermediate Topic (RQ2)

In response to the questionnaire assessing the usability of the *Specific Intermediate Topic Explorer* functionality described in Sec. 4.5.4, all nine participants reported that the functionality of selecting a specific intermediate topic in the *Specific Intermediate Topic Explorer* is understandable (Q3 in Fig. 4.13 on p. 109). The median rating for the understandability of navigating the *Specific Intermediate Topic Explorer* is *Good*, with responses ranging from *OK* to *Best Imaginable* (Q3 in Fig. 4.13 on p. 109). All nine participants stated that the *Specific Intermediate Topic Explorer* functionality is useful for exploring indirect relations via a specific intermediate gene, protein, and mental process (Q4 in Fig. 4.13 on p. 109). The median rating for the *Specific Intermediate Topic Explorer* functionality is *Good*, with responses ranging from *Good* to *Best Imaginable* (Q4 in Fig. 4.13 on p. 109). All nine participants expressed willingness to use the designed functionality to explore indirect relations via specific intermediate topics (Q5 in Fig. 4.13 on p. 109). The median rating for willingness to use the *Specific Intermediate Topic Explorer* to explore indirect relations is *Excellent*, with responses ranging from *Good* to *Best Imaginable* (Q5 in Fig. 4.13 on p. 109).

The assessments for the *Specific Intermediate Topic Explorer* functionality in the context of exploring indirect relations are as follows:

- **Understanding the relations between topics in literature exploration using the *Specific Intermediate Topic Explorer***

  All nine participants reported that the *Specific Intermediate Topic Explorer* contributes to understanding the relations between topics in literature exploration (in response to IQ12, Table 4.3 on p. 106). P8 stated, "Normally, I always search for publications by keywords and focus on individual publications. Exploring direct and indirect relations between brain regions and brain diseases provides a new method for understanding many publications. It is very useful for me to gain an overview of neuroscience, and the indirect relations are particularly helpful in identifying potentially useful topics to begin research (as a first-year PhD student). Specific intermediate topics are valuable for initially judging whether the identified indirect relations make sense in terms of basic biological mechanisms." P4 noted, "This work, highlighting the specific intermediate topics in indirect relations, is a really useful extension of the previous study. By understanding specific intermediate topics, such as a particular gene finding indirect relations, we can provide credible evidence to link the treatment of certain diseases to regulating some indirect brain regions."

- **Interaction procedure of the *Specific Intermediate Topic Explorer* in the context of exploring indirect relations**

  All nine participants reported that the interaction procedure, using the *Specific Intermediate Topic Explorer* to explore indirect relations via specific intermediate topics, was clear and understandable (in response to IQ11, Table 4.3 on p. 106). Seven participants were able to describe the step-by-step process of exploring indirect relations via specific intermediate topics. For example, P7 said, "I select the brain disease *Alzheimer's Disease* and explore which proteins, genes, and mental processes have direct relations with *Alzheimer's Disease* first. I can see these genes, proteins, and mental processes, along with their co-occurrence numbers with *Alzheimer's Disease*. I am interested in the particular mental process *Social Role*, select it as the intermediate topic, and explore which brain regions have indirect relations with *Alzheimer's Disease* via *Social Role*." Two participants, who were using the 3D AR prototype for the first time, mentioned that while they understood the logic of finding indirect relations between brain regions and brain diseases via a mental process, they still needed reminders from the experimenter when recalling the interaction process for completing representative tasks. Three participants suggested the development of a tutorial to help users better understand the operation of the functionalities and reduce the assistance of the experimenter during the evaluation session (in response to IQ 13, Table 4.3 on p. 106). The tutorial would ideally serve as a foundational resource, enabling users to efficiently utilise the prototype to explore direct and indirect relations between topics with confidence during their daily literature exploration.

4.6.3. The Usefulness of Exploring Indirect Relations via a Specific Intermediate Topic (RQ1)

Our main goal is to evaluate whether exploring indirect relations via a specific intermediate topic (RQ1) provides useful information to support experimental design, and not specifically on developing the prototype to investigate this. The implemented functionality (RQ2) and visualisation (RQ3) are developed to provide support for the user task.

In response to the questionnaire assessing the usefulness of exploring indirect relations between topics described in Sec. 4.5.5, all nine participants reported that exploring indirect relations via a specific intermediate topic is a useful task for identifying indirect relations that can support the design of a potentially useful experiment (Q6 in Fig. 4.14 on p. 111). The median rating for the usefulness of the user task is *Excellent*, with responses ranging from *OK* to *Best Imaginable* (Q6 in Fig. 4.14 on p. 111).

Eight participants reported their willingness to apply the designed visualisations and functionalities in the 3D AR prototype to support their daily literature exploration tasks (Q9 in Fig. 4.14 on p. 111). The median rating for willingness is *Good*, with responses ranging from *OK* to *Best Imaginable* (Q9 in Fig. 4.14 on p. 111). All nine participants indicated that they would recommend the 3D AR prototype to their colleagues (Q10 in Fig. 4.14 on p. 111). The median rating for recommendation is *Excellent*, with responses ranging from *Good* to *Best Imaginable* (Q10 in Fig. 4.14 on p. 111).

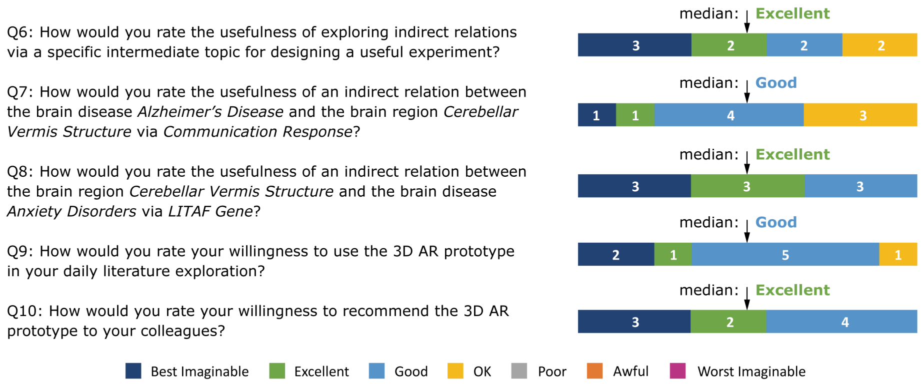

Figure 4.14: The distribution of all nine participants' attitudes toward the usefulness of exploring indirect relations via a specific intermediate topic.

- **Usefulness of exploring the indirect relation between the brain disease *Alzheimer's Disease* and the brain region *Cerebellar Vermis Structure* via *Communication Response* (Representative Task 1)**

  Six participants found the indirect relation between *Alzheimer's Disease* and *Cerebellar Vermis Structure* via *Communication Response* useful for their research (Q7 in Fig. 4.14 on p. 111). The median rating for the usefulness of this representative task is *Good*, with responses ranging from *OK* to *Best Imaginable* (Q7 in Fig. 4.14 on p. 111). P3 commented, "In my knowledge, the cerebellum controls balance. Discovering an indirect relation between *Alzheimer's Disease* and *Cerebellar Vermis Structure* via *Communication Response* inspires my curiosity about *Cerebellar Vermis Structure* and motivates me to explore the specific functions of this brain region."

  - **Choose another specific intermediate topic to explore the indirect relations between the brain disease *Alzheimer's Disease* and brain regions**

    P10 noted (in response to IQ14, Table 4.4 on p. 107), "I found that *MAPT wt Allele* has the highest number of co-occurrences with *Alzheimer's Disease* in the gene topic. I selected *MAPT wt Allele* as an intermediate topic to explore indirect relations between *Alzheimer's Disease* and all brain regions, and discovered that *Cerebellar Vermis Structure* also has an indirect relation with *Alzheimer's Disease* via *MAPT wt Allele*. This means that at least two intermediate topics (*Communication Response* and *MAPT wt Allele*) could link the indirect relation between *Alzheimer's Disease* and *Cerebellar Vermis Structure*. Finding repeated indirect relations through different intermediate topics increases the credibility of the indirect relation."

- **Usefulness of exploring the indirect relation between the brain region *Cerebellar Vermis Structure* and the brain disease *Anxiety Disorders* via *LITAF Gene* (Representative Task 2)**

  All nine participants found the indirect relation between *Cerebellar Vermis Structure* and *Anxiety Disorders* via *LITAF Gene* useful for their research (Q8 in Fig. 4.14 on p. 111). The median rating for the usefulness of this representative task is *Excellent*, with responses ranging from *Good* to *Best Imaginable* (Q8 in Fig. 4.14 on p. 111). P5 reported, "In the *Topic Model*, I could find a direct relation between *Cerebellar Vermis Structure* and *Depressive Disorder*. I observe that *Anxiety Disorders* is positioned close to *Depressive Disorder*, which makes sense to me given the similarity in their meanings and matches my knowledge. However, *Anxiety Disorders* has an indirect relation with *Cerebellar Vermis Structure*. This inspired my interest, as two similar diseases show different types of relations - one direct and one indirect - prompting me to explore the indirect relation further."

  – **Choose another specific intermediate topic to explore the indirect relations between the brain region *Cerebellar Vermis Structure* and brain diseases**

    P3 notes (in response to IQ15, Table 4.4 on p. 107), "I select *Eye Movements* as an intermediate topic to explore indirect relations between *Cerebellar Vermis Structure* and all brain diseases. The brain disease list inspires me. I can read through the diseases, observe on the left side the co-occurrence count between *Cerebellar Vermis Structure* and each brain disease, and on the right side, find the co-occurrences between *Eye Movements* and each brain disease. The green histogram on the right helps me identify indirect relations directly. I find that *Miller-Fisher Syndrome* has an indirect relation with the *Cerebellar Vermis Structure* via *Eye Movements*, which is useful for my research. I have written a publication on *Miller-Fisher Syndrome* and know it is related to eye movements. Currently, I could find an indirect relation between *Miller-Fisher Syndrome* and the *Cerebellar Vermis Structure* via *Eye Movements*, which could provide new insight into regulating the *Cerebellar Vermis Structure* for the disease treatment.

- **Three additional potentially useful hypotheses via the *Specific Intermediate Topic Explorer***

  Three participants shared three potentially useful hypotheses through the *Specific Intermediate Topic Explorer* (in response to IQ17, Table 4.4 on p. 107).

  – **Finding repeatedly identified indirect relations**

    P10 reported, "I am interested in exploring *Anxiety Disorders*. I select *Stress* as an intermediate topic to find the indirect relations between the *CA1 field* *of Hippocampus* and *Anxiety Disorders*. I then select *Fear* as another intermediate topic and find that the *CA1 field of Hippocampus* also has indirect relations with *Anxiety Disorders*. The repeated identification of an indirect relation between *CA1 field of Hippocampus* and *Anxiety Disorders* catches my attention, suggesting that there may be a useful and overlooked connection between *CA1 field* and *Anxiety Disorders*."

– **Finding similar brain diseases with different relation types**

  P11 noted, "I want to explore which brain diseases have indirect relations with the brain region *Thalamic Structure*. I select *CAT Gene* as the intermediate topic and find an indirect relation between *Thalamic Structure* and the brain disease *Spastic Syndrome*. I observe that diseases, such as *Motor Disorder* and *Dystonia Disorders*, are located near *Spastic Syndrome* in the *Topic Model* and have direct relations with *Thalamic Structure*. Therefore, the indirect relation between *Thalamic Structure* and *Spastic Syndrome* may become a direct relation in the future."

– **Finding specific brain regions in the same grouped brain region with different relation types**

  P7 pointed out, "I hope to find insights on exploring which brain regions have indirect relations with *Parkinson Disease* to expand my knowledge. I select *Behavior*, a common topic connected to *Parkinson Disease*, and discover an indirect relation between *Parkinson Disease* and *Intralaminar Parafascicular Thalamus* via *Behavior*. While observing this, I note that *Subthalamic Nucleus*, *Olfactory Thalamus*, and *Pre-motor Thalamus* all have direct relations with *Parkinson Disease*. Since the *Intralaminar Parafascicular Thalamus* is part of *Thalamus*, it may indeed have potentially useful relations with *Parkinson Disease*, which could be a useful direction for further investigation."

4.7. Discussion

We discuss reflections on the *Specific Intermediate Topic Explorer* visualisation (Sec. 4.7.1), the *Specific Intermediate Topic Explorer* functionality (Sec. 4.7.2), and the user task of exploring indirect relations via a specific intermediate topic (Sec. 4.7.3).

4.7.1. Reflections on the *Specific Intermediate Topic Explorer* Visualisation (RQ3)

The *Specific Intermediate Topic Explorer* visualisation is suitable for presenting indirect relations via a specific intermediate topic, as indicated by the positive responses from all nine participants (Sec. 4.6.1).

The suitability of the *Specific Intermediate Topic Explorer* visualisation is demonstrated through the following aspects:

1. The clear differentiation of relation types using colour (pink for direct and green

for indirect) within a single visualisation enables neuroscientists to intuitively distinguish which connections are explored in the literature and which are indirect (Figs. 4.9 and 4.10 on pp. 94 and 95).

2. The grouped list of specific intermediate topics helps neuroscientists select a gene, protein, or mental process beyond their familiar topics. Such cross-topic selection can help reveal potentially overlooked indirect relations (Fig. 4.11 on p. 100).

3. The use of histograms to display the number of direct relations assists neuroscientists in selecting intermediate topics with higher direct relation counts, which increases the possibility of discovering an indirect relation (Fig. 4.11 on p. 100).

While the results clearly indicate that the visualisation is suitable, the readability of long lists of intermediate topics in the AR environment remains a concern. Given that the lists of specific intermediate topics (genes, proteins, and mental processes) could exceed the users' view in the AR environment (Sec. 4.6.1), a potential solution would be to add scroll bars or page navigation buttons to the three intermediate topics lists. This improvement would enable users to navigate longer lists more easily, ensuring that all intermediate topics remain accessible even when the lists extend beyond the visible area. A horizontal layout could also be considered to compare which arrangement, vertical or horizontal, better supports users in selecting a topic from the list within the AR environment.

Additionally, given that the *Brain Regions Visualisation* currently displays 3D brain structures through multiple spheres, such as three spheres representing *Cerebellar Vermis Structure* (Fig. 4.9 on p. 94), participants indicated that they could only understand the 3D brain region structures by reading the short names of brain regions and relying on their own knowledge. Participants suggested adding a grouped brain regions visualisation. Grouping brain regions using anatomical categories could improve the clarity of the 3D brain visualisation and facilitate a better understanding of the spatial representation of 3D brain structures.

4.7.2. Reflections on the *Specific Intermediate Topic Explorer* Functionality (RQ2)

The *Specific Intermediate Topic Explorer* functionality is usable for exploring indirect relations via a specific intermediate topic, as indicated by the positive responses from all nine participants (Sec. 4.6.2). The usability of the *Specific Intermediate Topic Explorer* functionality is demonstrated through the following aspects:

1. The functionality to select a specific intermediate topic helps neuroscientists preliminarily determine whether they can find a useful indirect relation that aligns with basic biological mechanisms.

2. The interaction procedure for exploring indirect relations via a specific intermediate topic is understood by neuroscientists, who were able to clearly describe the step-by-step exploration process. This indicates that the functionality aligns

with neuroscientists' mental model of identifying indirect relations through a specific intermediate topic.

3. The willingness of neuroscientists to use the *Specific Intermediate Topic Explorer* in their research suggests that it provides a practical functionality for identifying potentially useful indirect relations.

Additionally, the number of both sets of direct relations (indicating an indirect relation) may help neuroscientists identify the priority of the usefulness of identified indirect relations, participants suggested separating the *Direct Relations Count List* from the *Specific Intermediate Topic Explorer*[12]. Due to the limited field of view in the AR environment and the need to avoid overwhelming users with excess information, making the *Direct Relations Count List* (Figs. 4.9 and 4.10 on pp. 94 and 95) a standalone functionality allows users to focus on one set of information at a time. This approach enables participants to first select a potentially useful specific intermediate topic for identifying indirect relations, and then review the count of both sets of direct relations associated with the intermediate topic to further determine which indirect relation could be used to inform the design of an experiment.

4.7.3. Reflections on Exploring Indirect Relations via a Specific Intermediate Topic (RQ1)

Exploring indirect relations via a specific intermediate topic is useful for identifying indirect relations that can support the design of a potentially useful experiment, as indicated by the positive responses from all nine participants (Sec. 4.6.3). Neuroscience participants identified two additional hypotheses related to representative tasks and three further hypotheses, making a total of five new hypotheses that could potentially inform the design of useful experiments (reported in Sec. 4.6.3). We highlight two novel exploration methods based on our results:

1. Specific intermediate topics are perceived as helpful for understanding the underlying mechanisms of indirect relations. By focusing on the intermediate topics, participants can more effectively assess whether the identified indirect relations are relevant and useful to their specific research questions. For example, the *LITAF Gene* is the bridge between the brain region *Cerebellar Vermis Structure* and the brain disease *Anxiety Disorders* (Sec. 4.6.3). Conducting gene-focused experiments can help understand how certain genetic markers may influence these less obvious connections.

2. Cross-topic selection of an intermediate topic provides inspiration for experimental design. Neuroscientists found it useful when an indirect relation could be identified via more than one intermediate topic, especially when those intermediate topics spanned different types. For example, one neuroscientist noted that the same indirect relation could be discovered via either a gene or a mental process (Sec. 4.6.3). While they typically focus on genes as experimental variables, the appearance of a mental process as an alternative pathway inspired

[12] We explored the usefulness of direct relation counts for neuroscientists in literature exploration tasks in a bachelor's thesis.

them to consider incorporating it into their experiment. Identifying intermediate topics across different topic types enables researchers to potentially uncover additional experimental variables.

4.8. Limitations

The current Knowledge Graph of Brain Science (KGBS) analyses co-occurrences between neuroscience topics in publications from January 1, 2010, to February 3, 2022 (Sec. 4.2.3). This analysis supports our exploratory studies. To enhance its utility for real neuroscience tasks, we would need to update the repository with recent literature.

Additionally, we received useful comments on the user interface design of the 3D AR prototype during evaluation, including suggestions to:

- present full names of brain regions to avoid misunderstandings caused by the current use of short names.
- filter the presentation of topic names to address the problem of occlusion.

The goal of our work is to understand the potential benefits that exploring indirect relations via a specific intermediate topic may have for neuroscientists. After this has been established through early design work such as this, the next step should be to improve the user interface design and clean the underlying data to better support neuroscientists' literature exploration needs.

4.9. Conclusions and Future Work

The user task of exploring indirect relations via a specific intermediate gene, protein, or mental process helps neuroscientists understand the underlying biological and cognitive mechanisms used to identify these indirect relations, enabling them to evaluate whether the identified indirect relations are useful and worthy of further investigation. This process, in turn, informs the design of potentially useful experiments (Sec. 4.6.3). A direction for future work is to assess the usefulness of the 3D AR prototype in the context of daily neuroscience literature research by closely following the research interests of a neuroscientist and incorporating their feedback and experiences to refine the prototype's functionalities and visualisations. By systematically observing their daily use of the 3D AR prototype for literature exploration, its effectiveness in supporting the identification of potentially useful experiments can be evaluated. This approach will ensure that our topic-based literature exploration tasks remain closely aligned with the practical needs of researchers. We view this study as helping neuroscientists identify potentially useful indirect relations within tens of thousands of publications, thereby supporting the complex task of designing potentially useful experiments.

5

Discussion and Conclusion

In the preceding chapters, we investigated support for neuroscientists in reviewing literature to inform the design of a potentially useful experiment before committing to a costly and lengthy implementation. We identified three literature exploration studies that are useful for informing experimental design. In the following sections, we outline the limitations (Sec. 5.1), discuss future work (Sec. 5.2), and reflect on the implications of our research (Sec. 5.3).

5.1. Limitations

We describe some limitations of our research and propose potential solutions.

- **Recruitment of Diverse Participants**

  One of the major challenges we faced was recruiting neuroscientists with specialised backgrounds relevant to our research. Although three studies engaged a total of 25 participants, most of them were from two neuroscience teams at UMC Utrecht, which limited the diversity of insights collected. While we made efforts to attend neuroscience seminars, we were unable to connect with senior neuroscience researchers. To address this issue, we could engage with neuroscience-focused institutes and attend specialist conferences. These approaches could help recruit participants with more specific expertise in relevant areas such as cognitive neuroscience, neuropathology, and neuroinformatics. Participants from diverse neuroscience backgrounds enabled us to obtain different perspectives on the use of indirect relations for finding a useful experiment.

- **Brain Region Naming Convention**

  The names of brain regions used in the *Brain Regions Visualisation* (Fig. 1.3 on p. 12) were sourced from the Scalable Brain Atlas (SBA), which provides standardised 3D anatomical representations based on Magnetic Resonance Imaging

(MRI). The SBA highlights the precise location and name of each specific brain region and is primarily used in professional medical topics.

In contrast, the brain region names identified from publications in the KBGS (Sec. 1.1.1) are derived from terms in the Unified Medical Language System (UMLS), which helps a wider range of researchers retrieve biomedical information from PubMed. For example, a researcher finds the term "dorsolateral prefrontal cortex" in biomedical literature in PubMed; however, the SBA represents this region as "parietal area PEa".

A complete mapping of brain regions in the *Brain Regions Visualisation* (Fig. 1.3 on p. 12) was not feasible due to differences in the terminology of brain regions between the SBA and UMLS. This meant that for some relations, there was no visible brain region in the *Brain Regions Visualisation*. For our studies on assessing the feasibility of the functionality and visualisation this is not a problem. For neuroscientists to use the prototype for longer periods of time for their research, further investment in both the dataset and the brain region mappings is required.

5.2. Future Work

Based on discussions with participating neuroscientists in our three studies, we identified three tasks for further investigation:

- **An Extended Period of Evaluation of the DatAR Prototype**

  As an initial phase of understanding the usefulness of the tasks we identified in our different studies, participants used the DatAR prototype only during the evaluation sessions. This did not allow neuroscience researchers to use the prototype for an extended period at the crucial stage in their own research of exploring the literature to identify a useful experiment. A useful next step in our research would be to make the DatAR prototype and headset available for an extended period to give a researcher sufficient time for exploring the literature and investigating multiple potential hypotheses until they reach their own conclusions.

- **Using Large Language Models to Help Identify Potentially Useful Hypotheses**

  Large Language Models (LLMs) are now widely used across domains for a variety of tasks. While their widespread adoption raises concerns about accuracy and the sources of knowledge, they also offer benefits in generating summaries and answering specific questions efficiently [73]. A potential benefit of LLMs is to uncover previously unobserved indirect relations between topics by responding to well-scoped questions. However, LLMs may sometimes hallucinate connections or overlook useful intermediate topics [43]. We propose that the generative capabilities of LLMs could be integrated with DatAR's 3D AR visualisation, thereby combining the efficiency of LLMs with the spatial visualisation of brain regions supported by DatAR.

For example, neuroscientists could first use LLMs for specific tasks, such as generating candidate intermediate topics that may indicate indirect relations between *Alzheimer's Disease* and brain regions. These LLM-generated suggestions can then serve as entry points for exploring potentially useful indirect relations. Neuroscientists could then use the DatAR prototype to select a potentially useful intermediate topic and visualise which brain regions may be indirectly affected by *Alzheimer's Disease* via the intermediate topic. The spatial representation of brain regions in DatAR helps neuroscientists understand indirect relations.

- **A Comparative Study Investigating the Outcomes of Using DatAR Versus LLMs**

  Another direction for future work is to conduct a comparative study investigating the outcomes of using DatAR versus LLMs alone to identify useful hypotheses. This would help assess which approach provides neuroscientists with greater confidence in designing a useful experiment.

5.3. Reflections on the Implications of Our Research

We collaborated closely with neuroscientists to conduct three studies that explored and evaluated the usefulness of identifying indirect relations between topics in neuroscience literature. The visualisations we implemented provide an intuitive way to explore both direct and indirect relations, helping researchers better understand which brain regions are directly and indirectly affected by a brain disease. We reflect on the implications of our research for the fields of neuroscience, Human-Computer Interaction (HCI) and other domains with multiple interconnected topics.

- **For Neuroscience Research**

  Neuroscientists often require months or sometimes even years to evaluate the usefulness of a hypothesis through an experiment. Before committing to a specific experiment, neuroscientists need to identify a hypothesis with the highest potential to yield useful contributions. We have investigated three novel exploration methods that can support neuroscientists in identifying useful hypotheses:

  (1) Exploring indirect relations between topics can save significant time in the early stages of hypothesis generation.

  In Chapter 2, participating neuroscientists indicated that the identified indirect relations could serve as useful hypotheses to inform experimental design (Sec. 2.6.3). Using the DatAR prototype, neuroscientists only need to select an intermediate topic type, such as genes, after which the indirect relations between brain regions and brain diseases via the intermediate topic are visualised. This process saves time in the early stages of hypothesis generation, allowing neuroscientists to quickly find potentially useful research directions to investigate.

  (2) Exploring the evolution of indirect to direct relations can enhance confidence in the potential of current indirect relations.

  In Chapter 3, participating neuroscientists indicated that the timeline helped them identify historical indirect relations that had become current direct relations, reinforcing that current indirect relations could evolve into direct ones in future investigations (Sec. 3.7.3). This insight increases neuroscientists' confidence in exploring indirect relations to design a useful experiment.

(3) Exploring indirect relations via a specific intermediate topic can increase the amount of evidence for designing a useful experiment.

In Chapter 4, participating neuroscientists indicated that specific intermediate topics were helpful for understanding the underlying mechanisms of indirect relations (Sec. 4.6.3). The specific intermediate topic and the two direct relations that indicate an indirect relation provide neuroscientists with evidence to assess the usefulness of the indirect relations. This evidence increases neuroscientists' trust when designing a useful experiment.

- **For Human-Computer Interaction (HCI)**

  Our work has demonstrated the usefulness of topic-based literature exploration, rather than individual publications, in informing the design of a useful experiment through early-stage design work. We encourage the HCI community to engage with the challenge of designing functionalities and visualisations for topic-based literature exploration and further investigate potential improvements to the user interface.

- **For Domains with Multiple Interconnected Topics**

  Although our research has not been applied to other fields, it provides a concrete case study that may inspire topic-based literature exploration in other research domains. Neuroscience serves as a compelling example, as it involves the interaction of multiple interconnected topics, such as brain regions, brain diseases, proteins, and genes, where uncovering indirect relations across sub-fields can lead to useful insights. This conceptual complexity is not unique to neuroscience. Other domains, such as drug discovery and cancer treatment, also face similar challenges. In drug discovery, for example, identifying indirect relations between molecular targets and disease phenotypes can support the repositioning of existing drugs or the development of treatments. In cancer treatment, for example, identifying indirect relations between biomarkers and treatment responses can provide insights into disease mechanisms and inform the design of useful intervention strategies.

We hope that our research can facilitate neuroscience researchers by (1) enabling the exploration of hundreds of thousands of publications simultaneously, and (2) supporting the creation of useful hypotheses for designing a useful experiment. Specifically, our research serves as a complementary step in the early phases of designing a useful experiment, bridging the gap between conducting a large-scale literature exploration and informing the design of a useful experiment. By presenting an interactive visualisation, our prototype enables researchers to intuitively discover less observed, sometimes previously unnoticed, relations among topics that are not easily visible through traditional literature exploration.

References

1. Abubakre, M., Fayoumi, A. & Eleburuike, I. Implementing process improvement initiative: the role of visualisation and standardisation methods. *Business Process Management Journal* **27,** 965–986. https://doi.org/10.1108/BPMJ-10-2020-0474 (2021).
2. Adams, W. C. Conducting semi-structured interviews. *Handbook of practical program evaluation,* 492–505. https://doi.org/10.1002/9781119171386.ch19 (2015).
3. Alhashmi, S. M., Hashem, I. A. T. & Al-Qudah, I. Artificial intelligence applications in healthcare: a bibliometric and topic model-based analysis. *Intelligent Systems with Applications* **21,** 200299. https://doi.org/10.1016/j.iswa.2023.200299 (2024).
4. Allen, I. E. & Seaman, C. A. Likert scales and data analyses. *Quality progress* **40,** 64–65. https://www.bayviewanalytics.com/reports/asq/likert-scales-and-data-analyses.pdf (2007).
5. Askari, S. in *Reshaping Healthcare with Cutting-Edge Biomedical Advancements* 211–234 (IGI Global, 2024). doi:10.4018/979-8-3693-4439-2.ch010.
6. Auda, J., Gruenefeld, U., Faltaous, S., Mayer, S. & Schneegass, S. A scoping survey on cross-reality systems. *ACM Computing Surveys* **56,** 1–38. https://doi.org/10.1145/3616536 (2023).
7. Bach, B., Cordeil, M., Engelke, U., Ens, B., Serrano, M. & Willett, W. *Interaction design & prototyping for immersive analytics* in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (2019), 1–8. https://doi.org/10.1145/3290607.3299019.
8. Bakker, R., Tiesinga, P. & Kötter, R. The scalable brain atlas: instant web-based access to public brain atlases and related content. *Neuroinformatics* **13,** 353–366. https://doi.org/10.1007/s12021-014-9258-x (2015).
9. Bangor, A., Kortum, P. & Miller, J. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* **4,** 114–123 (2009).
10. Bernasconi, E., Ceriani, M., Mecella, M. & Catarci, T. Design, realization, and user evaluation of the ARCA system for exploring a digital library. *International Journal on Digital Libraries* **24,** 1–22. https://doi.org/10.1007/s00799-022-00343-0 (2023).
11. Blandford, A., Furniss, D. & Makri, S. *Qualitative HCI research: Going behind the scenes* (Morgan & Claypool Publishers, 2016).
12. Blei, D. M. Probabilistic topic models. *Communications of the ACM* **55,** 77–84. doi:10.1145/2133806.2133826 (2012).
13. Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent dirichlet allocation. *Journal of machine Learning research* **3,** 993–1022 (2003).

14. Boddy, C. R. Sample size for qualitative research. *Qualitative Market Research: An International Journal* **19,** 426–432. doi:10.1108/QMR-06-2016-0053 (2016).
15. Booth, A., Martyn-St James, M., Clowes, M. & Sutton, A. Systematic approaches to a successful literature review (2021).
16. Bravo, A. & Maier, A. M. *Immersive visualisations in design: Using augmented reality (AR) for information presentation* in *Proceedings of the Design Society: DESIGN Conference* **1** (2020), 1215–1224. https://doi.org/10.1017/dsd.2020.33.
17. Brown, T. & Wyatt, J. Design thinking for social innovation. *Development Outreach* **12,** 29–43. https://api.semanticscholar.org/CorpusID:6740624 (2010).
18. Burger, C. D., DuBrock, H. M., Cartin-Ceba, R., Moss, J. E., Shapiro, B. P. & Frantz, R. P. *Topic-based, recent literature review on pulmonary hypertension* in *Mayo Clinic Proceedings* **96** (2021), 3109–3121. https://doi.org/10.1016/j.mayocp.2021.05.026.
19. Carmigniani, J. & Furht, B. Augmented reality: an overview. *Handbook of augmented reality,* 3–46. https://doi.org/10.1007/978-1-4614-0064-6_1 (2011).
20. Chalal, M. L., Medjdoub, B., Bezai, N., Bull, R. & Zune, M. Visualisation in energy eco-feedback systems: A systematic review of good practice. *Renewable and Sustainable Energy Reviews* **162,** 112447. https://doi.org/10.1016/j.rser.2022.112447 (2022).
21. Chang, C.-J., Hsu, Y. L., Tan, W. T. M., Chang, Y.-C., Lu, P. C., Chen, Y., Wang, Y.-H. & Chen, M. Y. *Exploring Augmented Reality Interface Designs for Virtual Meetings in Real-world Walking Contexts* in *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (2024), 391–408. https://doi.org/10.1145/3643834.3661538.
22. Chen, C., Nguyen, C., Hoffswell, J., Healey, J., Bui, T. & Weibel, N. *PaperToPlace: Transforming Instruction Documents into Spatialized and Context-Aware Mixed Reality Experiences* in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (2023), 1–21. https://doi.org/10.1145/3586183.3606832.
23. Cheng, Y., Yan, Y., Yi, X., Shi, Y. & Lindlbauer, D. *Semanticadapt: Optimization-based adaptation of mixed reality layouts leveraging virtual-physical semantic connections* in *The 34th Annual ACM Symposium on User Interface Software and Technology* (2021), 282–297. https://doi.org/10.1145/3472749.3474750.
24. Chystaya, Y., Poyade, M., Rea, P. M. & McCorry, O. in *Medical Visualization and Applications of Technology* 147–180 (Springer, 2022). https://doi.org/10.1007/978-3-031-06735-8_5.
25. Crichton, G., Baker, S., Guo, Y. & Korhonen, A. Neural networks for open and closed Literature-based Discovery. *PloS one* **15,** e0232891. doi:https://doi.org/10.1371/journal.pone.0232891 (2020).
26. d.school, H. P. I. o. D. a. S. *Design Thinking Bootleg* 1–90. https://dschool.stanford.edu/tools/design-thinking-bootleg (2013).

27. Dam, R. F. The 5 Stages in the Design Thinking Process. *Interaction Design Foundation - IxDF.* https://www.interaction-design.org/literature/article/5-stages-in-the-design-thinking-process (2024).
28. Dhami, M. K., Hertwig, R. & Hoffrage, U. The role of representative design in an ecological approach to cognition. *Psychological bulletin* **130,** 959. https://doi.org/10.1037/0033-2909.130.6.959 (2004).
29. Doorley, S., Holcomb, S., Klebahn, P., Segovia, K. & Utley, J. Hasso Plattner Institute of Design at Stanford. *Design thinking bootleg.* https://dschool.stanford.edu/tools/design-thinking-bootleg (2018).
30. Ens, B., Cordeil, M., North, C., Dwyer, T., Besançon, L., Prouzeau, A., Liu, J., Cunningham, A., Drogemuller, A., Satriadi, K. A. *et al. Immersive analytics 2.0: Spatial and embodied sensemaking* in *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (2022), 1–7. https://doi.org/10.1145/3491101.3503726.
31. Ericsson, K. A. & Simon, H. A. Verbal reports as data. *Psychological review* **87,** 215. https://doi.org/10.1037/0033-295X.87.3.215 (1980).
32. Fink, A. *How to conduct surveys: A step-by-step guide* (SAGE publications, 2024).
33. Gopalakrishnan, V., Jha, K., Jin, W. & Zhang, A. A survey on literature based discovery approaches in biomedical domain. *Journal of biomedical informatics* **93,** 103141. https://doi.org/10.1016/j.jbi.2019.103141 (2019).
34. Grewal, D., Puccinelli, N. & Monroe, K. B. Meta-analysis: Integrating accumulated knowledge. *Journal of the Academy of Marketing Science* **46,** 9–30. https://doi.org/10.1007/s11747-017-0570-5 (2018).
35. Han, D. & Cho, I. *Evaluating 3D user interaction techniques on spatial working memory for 3D scatter plot exploration in immersive analytics* in *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2023), 513–522. doi:10.1109/ISMAR59233.2023.00066.
36. Hashim, N. L. & Isse, A. J. Usability evaluation metrics of tourism mobile applications. *Journal of Software Engineering and Applications* **12,** 267–277. doi:10.4236/jsea.2019.127016 (2019).
37. Hendrickx, T. *Designing a 3D visualization for exploring relations between semantically similar brain diseases in Augmented Reality* MA thesis (Utrecht University, The Netherlands, 2021). https://studenttheses.uu.nl/handle/20.500.12932/41237.
38. Henry, S. & McInnes, B. T. Literature based discovery: models, methods, and trends. *Journal of biomedical informatics* **74,** 20–32. doi:10.1016/j.jbi.2017.08.011 (2017).
39. Henssen, D. J., van den Heuvel, L., De Jong, G., Vorstenbosch, M. A., van Cappellen van Walsum, A.-M., Van den Hurk, M. M., Kooloos, J. G. & Bartels, R. H. Neuroanatomy learning: Augmented reality vs. cross-sections. *Anatomical sciences education* **13,** 353–365. doi:10.1002/ase.1912 (2020).
40. Hornbæk, K. Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies* **64,** 79–102. https://doi.org/10.1016/j.ijhcs.2005.06.002 (2006).

41. Hou, J. Exploration into the evolution and historical roots of citation analysis by referenced publication year spectroscopy. *Scientometrics* **110,** 1437–1452. https://doi.org/10.1007/s11192-016-2206-9 (2017).
42. Hristovski, D., Kastrin, A., Dinevski, D., Burgun, A., Žiberna, L. & Rindflesch, T. C. Using literature-based discovery to explain adverse drug effects. *Journal of medical systems* **40,** 1–5. doi:10.1007/s10916-016-0544-z (2016).
43. Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B. *et al.* A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems* **43,** 1–55. https://doi.org/10.1145/3703155 (2025).
44. Hurter, C., Rogowitz, B., Truong, G., Andry, T., Romat, H., Gardy, L., Amini, F. & Riche, N. H. Memory Recall for Data Visualizations in Mixed Reality, Virtual Reality, 3D and 2D. *IEEE Transactions on Visualization and Computer Graphics.* doi:10.1109/TVCG.2023.3336588 (2024).
45. Interaction Design Foundation - IxDF. *What is User Centered Design (UCD)?* June 2016. https://www.interaction-design.org/literature/topics/user-centered-design.
46. Jacaruso, L. Insights into the nutritional prevention of macular degeneration based on a comparative topic modeling approach. *PeerJ Computer Science* **10,** e1940. doi:10.7717/peerj-cs.1940 (2024).
47. Jesson, J., Lacey, F. M. & Matheson, L. Doing your literature review: Traditional and systematic techniques (2011).
48. Joshi, A., Kale, S., Chandel, S. & Pal, D. K. Likert scale: Explored and explained. *British journal of applied science & technology* **7,** 396. doi:10.9734/BJAST/2015/14975 (2015).
49. Kallio, H., Pietilä, A.-M., Johnson, M. & Kangasniemi, M. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of advanced nursing* **72,** 2954–2965. https://doi.org/10.1111/jan.13031 (2016).
50. Kastrin, A. & Hristovski, D. Scientometric analysis and knowledge mapping of literature-based discovery (1986–2020). *Scientometrics* **126,** 1415–1451. doi:https://doi.org/10.1007/s11192-020-03811-z (2021).
51. Keil, J., Korte, A., Ratmer, A., Edler, D. & Dickmann, F. Augmented reality (AR) and spatial cognition: effects of holographic grids on distance estimation and location memory in a 3D indoor scenario. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science* **88,** 165–172. https://doi.org/10.1007/s41064-020-00104-1 (2020).
52. Keiriz, J. J., Zhan, L., Ajilore, O., Leow, A. D. & Forbes, A. G. NeuroCave: A web-based immersive visualization platform for exploring connectome datasets. *Network Neuroscience* **2,** 344–361. https://doi.org/10.1162/netn_a_00044 (2018).
53. Kim, Y. H. & Song, M. A context-based ABC model for literature-based discovery. *PloS one* **14,** e0215313. https://doi.org/10.1371/journal.pone.0215313 (2019).
54. Kraus, S., Breier, M., Lim, W. M., Dabić, M., Kumar, S., Kanbach, D., Mukherjee, D., Corvello, V., Piñeiro-Chousa, J., Liguori, E. *et al.* Literature reviews as inde-

pendent studies: guidelines for academic practice. *Review of managerial science* **16,** 2577–2595. https://doi.org/10.1007/s11846-022-00588-8 (2022).
55. Kujala, S. User involvement: a review of the benefits and challenges. *Behaviour & information technology* **22,** 1–16. https://doi.org/10.1080/01449290301782 (2003).
56. Kun, C. *Assisting Neuroscientists in Exploring and Understanding Brain Topic Relations Using Augmented Reality* MA thesis (Utrecht University, The Netherlands, 2022). https://studenttheses.uu.nl/handle/20.500.12932/41512.
57. Lee, B., Cordeil, M., Prouzeau, A., Jenny, B. & Dwyer, T. *A design space for data visualisation transformations between 2d and 3d in mixed-reality environments* in *Proceedings of the 2022 CHI conference on human factors in computing systems* (2022), 1–14. https://doi.org/10.1145/3491102.3501859.
58. Lee, B., Hu, X., Cordeil, M., Prouzeau, A., Jenny, B. & Dwyer, T. Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. *IEEE Transactions on Visualization and Computer Graphics* **27,** 1171–1181. doi:10.1109/TVCG.2020.3030450 (2020).
59. Lewis, J. R. & Sauro, J. Usability and user experience: Design and evaluation. *Handbook of human factors and ergonomics,* 972–1015. https://doi.org/10.1002/9781119636113.ch38 (2021).
60. Li, J., Wang, S., Rudinac, S. & Osseyran, A. High-performance computing in healthcare: an automatic literature analysis perspective. *Journal of Big Data* **11,** 61. https://doi.org/10.1186/s40537-024-00929-2 (2024).
61. Li, S. & Wang, H. Traditional literature review and research synthesis. *The Palgrave handbook of applied linguistics research methodology,* 123–144. https://doi.org/10.1057/978-1-137-59900-1_6 (2018).
62. Li, T., Wu, S., Jin, Y., Shi, H. & Liu, S. X-Space: Interaction design of extending mixed reality space from Web2D visualization. *Visual Informatics* **7,** 73–83. https://doi.org/10.1016/j.visinf.2023.10.001 (2023).
63. Li, Z., Annett, M., Hinckley, K., Singh, K. & Wigdor, D. *Holodoc: Enabling mixed reality workspaces that harness physical and digital content* in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019), 1–14. https://doi.org/10.1145/3290605.3300917.
64. Liu, J., Zhang, Q., Fu, L., Wang, X. & Lu, S. *Evolving knowledge graphs* in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications* (2019), 2260–2268. doi:10.1109/INFOCOM.2019.8737547.
65. Liu, J., Satriadi, K. A., Ens, B. & Dwyer, T. *Investigating the Effects of Physical Landmarks on Spatial Memory for Information Visualisation in Augmented Reality* in *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2024), 289–298. doi:10.1109/ISMAR62088.2024.00043.
66. Loglisci, C. Time-based discovery in biomedical literature: mining temporal links. *International Journal of Data Analysis Techniques and Strategies* **5,** 148–174. https://doi.org/10.1504/IJDATS.2013.053679 (2013).
67. MacDonald, C. M. *Understanding usefulness in human-computer interaction to enhance user experience evaluation* https://doi.org/10.17918/etd-3848 (Drexel University, 2012).

68. MacKenzie, I. S. Human-computer interaction: An empirical research perspective (2024).
69. Magaldi, D. & Berler, M. Semi-structured interviews. *Encyclopedia of personality and individual differences,* 4825–4830. https://doi.org/10.1007/978-3-319-24612-3_857 (2020).
70. Marx, W., Bornmann, L., Barth, A. & Leydesdorff, L. Detecting the historical roots of research fields by reference publication year spectroscopy (RPYS). *Journal of the Association for Information Science and Technology* **65,** 751–764. https://doi.org/10.1002/asi.23089 (2014).
71. Moreau, E., Hardiman, O., Heverin, M. & O'sullivan, D. Literature-Based Discovery beyond the ABC paradigm: a contrastive approach. *BioRxiv,* 2021–09. https://doi.org/10.1101/2021.09.22.461375 (2021).
72. Müller, H.-M., Rangarajan, A., Teal, T. K. & Sternberg, P. W. Textpresso for neuroscience: searching the full text of thousands of neuroscience research papers. *Neuroinformatics* **6,** 195–204. doi:10.1007/s12021-008-9031-0 (2008).
73. Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N. & Mian, A. A comprehensive overview of large language models. *ACM Transactions on Intelligent Systems and Technology.* https://doi.org/10.1145/3744746 (2023).
74. Nemoto, T. & Beglar, D. *Likert-scale questionnaires* in *JALT 2013 conference proceedings* **108** (2014), 1–6.
75. Org, I. *The field guide to human centered design* https://www.designkit.org/resources/1.html (Ideo Org, 2015).
76. Parekh, P., Patel, S., Patel, N. & Shah, M. Systematic review and meta-analysis of augmented reality in medicine, retail, and games. *Visual computing for industry, biomedicine, and art* **3,** 1–20. https://doi.org/10.1186/s42492-020-00057-7 (2020).
77. Pastore, A., Jurinovic, V., Kridel, R., Hoster, E., Staiger, A. M., Szczepanowski, M., Pott, C., Kopp, N., Murakami, M., Horn, H. *et al.* Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. *The lancet oncology* **16,** 1111–1122. http://dx.doi.org/10.1016/S1470-2045(15)00169-2 (2015).
78. Paul, J. & Barari, M. Meta-analysis and traditional systematic literature reviews What, why, when, where, and how? *Psychology & Marketing* **39,** 1099–1115. https://doi.org/10.1002/mar.21657 (2022).
79. Pu, Y., Beck, D. & Verspoor, K. Graph embedding-based link prediction for literature-based discovery in Alzheimer's Disease. *Journal of Biomedical Informatics* **145,** 104464. https://doi.org/10.1016/j.jbi.2023.104464 (2023).
80. Rajendra Prasad, K., Mohammed, M. & Noorullah, R. Visual topic models for healthcare data clustering. *Evolutionary Intelligence* **14,** 545–562. https://doi.org/10.1007/s12065-019-00300-y (2021).
81. Reipschlager, P., Flemisch, T. & Dachselt, R. Personal augmented reality for information visualization on large interactive displays. *IEEE Transactions on Visualization and Computer Graphics* **27,** 1182–1192. doi:10.1109/TVCG.2020.3030460 (2020).

82. Rodriguez-Sabate, C., Morales, I., Lorenzo, J. N. & Rodriguez, M. The organization of the basal ganglia functional connectivity network is non-linear in Parkinson's disease. *NeuroImage: Clinical* **22,** 101708. https://doi.org/10.1016/j.nicl.2019.101708 (2019).
83. Ruslin, R., Mashuri, S., Rasak, M. S. A., Alhabsyi, F. & Syam, H. Semi-structured Interview: A methodological reflection on the development of a qualitative research instrument in educational studies. *IOSR Journal of Research & Method in Education (IOSR-JRME)* **12,** 22–29. doi:10.9790/7388-1201052229 (2022).
84. Saffo, D., Di Bartolomeo, S., Crnovrsanin, T., South, L., Raynor, J., Yildirim, C. & Dunne, C. Unraveling the design space of immersive analytics: A systematic review. *IEEE Transactions on Visualization and Computer Graphics.* doi:10.1109/TVCG.2023.3327368 (2023).
85. Scheuerman, M. K., Spiel, K., Haimson, O. L., Hamidi, F. & Branham, S. M. HCI guidelines for gender equity and inclusivity. *UMBC Faculty Collection.* https://www.morgan-klaus.com/gender-guidelines.html (2020).
86. Schloss, K. B., Leggon, Z. & Lessard, L. Semantic discriminability for visual communication. *Ieee transactions on visualization and computer graphics* **27,** 1022–1031. doi:10.1109/TVCG.2020.3030434 (2020).
87. Schoenlein, M. A. & Schloss, K. B. Colour-concept association formation for novel concepts. *Visual Cognition* **30,** 457–479. https://doi.org/10.1080/13506285.2022.2089418 (2022).
88. Seraji, M. R. & Stuerzlinger, W. *Hybridaxes: An immersive analytics tool with interoperability between 2d and immersive reality modes* in *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (2022), 155–160. doi:10.1109/ISMAR-Adjunct57072.2022.00036.
89. Six, J. M. & Macefield, R. How to determine the right number of participants for usability studies. *San Francisco (CA): UXmatters.* https://www.uxmatters.com/mt/archives/2016/01/how-to-determine-the-right-number-of-participants-for-usability-studies.php (2016).
90. Song, W. & Sun, J. *Hypothesis Generation from Literature for Advancing Biological Mechanism Research: A Perspective* in *Proceedings of the 2023 15th International Conference on Bioinformatics and Biomedical Technology* (2023), 204–209. https://doi.org/10.1145/3608164.3608194.
91. Soufan, A., Ruthven, I. & Azzopardi, L. *Searching the literature: an analysis of an exploratory search task* in *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval* (2022), 146–157. https://doi.org/10.1145/3498366.3505818.
92. Sudhahar, S., Ozer, B., Chang, J., Chadwick, W., O'Donovan, D., Campbell, A., Tulip, E., Thompson, N. & Roberts, I. An experimentally validated approach to automated biological evidence generation in drug discovery using knowledge graphs. *Nature Communications* **15,** 5703. https://doi.org/10.1038/s41467-024-50024-6 (2024).
93. Swanson, D. R. Fish oil, Raynaud's syndrome, and undiscovered public knowledge. *Perspectives in biology and medicine* **30,** 7–18. doi:10.1353/pbm.1986.0087 (1986).

94. Swanson, D. R. & Smalheiser, N. R. An interactive system for finding complementary literatures: a stimulus to scientific discovery. *Artificial intelligence* **91,** 183–203. doi:10.1016/S0004-3702(97)00008-8 (1997).
95. Talevi, A. Drug-resistant epilepsy: Is there an overlooked association between drug resistant epilepsies and neuropsychiatric comorbidities? *Epilepsy & Behavior* **161,** 110144. https://doi.org/10.1016/j.yebeh.2024.110144 (2024).
96. Tanhaei, G. Exploring Relationships among Topics in Neuroscience Literature using Augmented Reality Visualization. https://doi.org/10.33540/2491 (2024).
97. Tanhaei, G., Troost, I., Hardman, L. & Hürst, W. *Designing a topic-based literature exploration tool in AR-an exploratory study for neuroscience* in *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (2022), 471–476. doi:10.1109/ISMAR-Adjunct57072.2022.00099.
98. Thilakaratne, M., Falkner, K. & Atapattu, T. A systematic review on literature-based discovery workflow. *PeerJ Computer Science* **5,** e235. doi:10.7717/peerj-cs.235 (2019).
99. Troost, I., Tanhaei, G., Hardman, L. & Hürst, W. *Exploring relations in neuroscientific literature using augmented reality: A design study* in *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (2021), 266–274. https://doi.org/10.1145/3461778.3462053.
100. Tsafnat, G., Jasch, D., Misra, A., Choong, M. K., Lin, F. P.-Y. & Coiera, E. Gene–disease association with literature based enrichment. *Journal of Biomedical Informatics* **49,** 221–226. https://doi.org/10.1016/j.jbi.2014.03.007 (2014).
101. Van Altena, A. J., Moerland, P. D., Zwinderman, A. H. & Olabarriaga, S. D. Understanding big data themes from scientific biomedical literature through topic modeling. *Journal of Big Data* **3,** 1–21. https://doi.org/10.1186/s40537-016-0057-0 (2016).
102. Vogel, J. W., Iturria-Medina, Y., Strandberg, O. T., Smith, R., Levitis, E., Evans, A. C. & Hansson, O. Spread of pathological tau proteins through communicating neurons in human Alzheimer's disease. *Nature communications* **11,** 2612. https://doi.org/10.1038/s41467-020-15701-2 (2020).
103. Wang, X., Besançon, L., Rousseau, D., Sereno, M., Ammi, M. & Isenberg, T. *Towards an understanding of augmented reality extensions for existing 3D data analysis tools* in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020), 1–13. doi:10.1145/3313831.3376657.
104. Ware, C. *Information visualization: perception for design* (Morgan Kaufmann, 2019).
105. Williamson, D. F., Parker, R. A. & Kendrick, J. S. The box plot: a simple visual method to interpret data. *Annals of internal medicine* **110,** 916–921. doi:10.7326/0003-4819-110-11-916 (1989).
106. Wobbrock, J. O. & Kientz, J. A. Research contributions in human-computer interaction. *interactions* **23,** 38–44. https://doi.org/10.1145/2907069 (2016).
107. Wong, M., Previde, P., Cole, J., Thomas, B., Laxmeshwar, N., Mallory, E., Lever, J., Petkovic, D., Altman, R. B. & Kulkarni, A. Search and visualization of gene-drug-disease interactions for pharmacogenomics and precision medicine re-

search using GeneDive. *Journal of biomedical informatics* **117,** 103732. https://doi.org/10.1016/j.jbi.2021.103732 (2021).

108. Xu, B., Hardman, L. & Hürst, W. *Exploring Indirect Relations between Topics in Augmented Reality to Inform the Design of a Neuroscience Experiment* in *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (2025), 1248–1249. doi:10.1109/VRW66409.2025.00270.
109. Xu, B., Tanhaei, G., Hardman, L. & Hürst, W. *DatAR: Supporting Neuroscience Literature Exploration by Finding Relations Between Topics in Augmented Reality* in *International Conference on Multimedia Modeling* (2024), 295–300. https://doi.org/10.1007/978-3-031-53302-0_24.
110. Yang, H.-T., Ju, J.-H., Wong, Y.-T., Shmulevich, I. & Chiang, J.-H. Literature-based discovery of new candidates for drug repurposing. *Briefings in bioinformatics* **18,** 488–497. doi:10.1093/bib/bbw030 (2017).
111. Yen, C., Lin, C.-L. & Chiang, M.-C. Exploring the frontiers of neuroimaging: a review of recent advances in understanding brain functioning and disorders. *Life* **13,** 1472. https://doi.org/10.3390/life13071472 (2023).
112. Zhang, L., Luczak, T., Smith, E., Burch, R. F. *et al.* Using Microsoft HoloLens to improve memory recall in anatomy and physiology: A pilot study to examine the efficacy of using augmented reality in education. *Journal of Educational Technology Development and Exchange (JETDE)* **12,** 2. https://doi.org/10.18785/jetde.1201.02 (2019).
113. Zhang, R., Cairelli, M. J., Fiszman, M., Kilicoglu, H., Rindflesch, T. C., Pakhomov, S. V. & Melton, G. B. Exploiting literature-derived knowledge and semantics to identify potential prostate cancer drugs. *Cancer informatics* **13,** CIN–S13889. doi:10.4137/CIN.S13889 (2014).
114. Zhang, S., Lu, H., Zhang, F., Zhu, Y. & Wang, R. *Theme-Based Literature Analysis in the Field of Transportation* in *Advances in Intelligent Information Hiding and Multimedia Signal Processing: Proceeding of the 16th International Conference on IIHMSP in conjunction with the 13th international conference on FITAT, November 5-7, 2020, Ho Chi Minh City, Vietnam, Volume 2* (2021), 489–495. https://doi.org/10.1007/978-981-33-6757-9_60.
115. Zhu, E., Hadadgar, A., Masiello, I. & Zary, N. Augmented reality in healthcare education: an integrative review. *PeerJ Computer Science* **2,** e469. doi:10.7717/peerj.469 (2014).

A

Appendix: Exploring Indirect Relations Between Topics

Indirect Relations Study Informed Consent Form

Research project:
We have created the extended DatAR prototype in Augmented Reality (AR) environment to support neuroscientists exploring the relations between topics in scientific literature. Our goal is to understand how neuroscientists can identify relations between topics and find indirect connections between them in order to satisfy their information exploration goals.

Please read the following carefully:

- I have been provided with information on the purpose of this project.
- I have been given the opportunity to ask questions about the project and my participation.
- I voluntarily agree to participate in the research project.
- I agree that I will not be paid for my participation.
- I understand that my participation in the research project involves being interviewed.
- I understand I can withdraw at any time without giving any reason and that I will not be penalised for withdrawing.
- I understand that my operation in the AR environment will be recorded.
- I agree that the interviews will be audio-recorded and transcribed.

- I understand the explanations on the use of the data in research, publications, sharing and archiving.
- I understand that my participation and that of my organisation will remain anonymous.
- I am aware that using AR can cause nausea and other discomforts and that if I experience any discomfort, I will take a break or stop the experiment completely.

If you agree with all the above statements then please sign below:

Participant: ____________________

Researcher: ____________________

We very much appreciate your time in conducting our study,

Boyu Xu
On behalf of the DatAR Team

The Process of Exploring Indirect Relations Between Topics

We described the process of exploring direct and indirect relations between brain regions and brain diseases during the evaluation procedure (Sec. 2.5.2). This process consists of six parts[1]:

- ***Select the Specific Brain Region or Brain Disease***

  Participants can find the specific topic in the list of brain regions and brain diseases using the *Brain Region or Brain Disease Retriever*, Fig. 2.1 ⑥ on p. 25. For example, select the specific brain region *Hippocampus* and the specific brain disease *Bipolar Disorder*.

- ***Construct the Brain Disease Topic Model or the Brain Regions Visualisation***

  The *Brain Disease Topic Model* is visualised by integrating the brain diseases topic into the receptacle (the dark blue sphere) of the *Topic Model*, Fig. 2.4 ① on p. 30. The *Brain Regions Visualisation* is visualised by dragging the relevant widget, Fig. 2.5 ② on p. 31.

- ***Query Direct Relations between Brain Regions and Brain Diseases***

  There are two receptacles on the *Direct Relations Explorer*, Figs. 2.4 and 2.5 ③ on pp. 30 and 31. Put the specific brain region or the specific brain disease on the concept receptacle (the light blue sphere), correspondingly, put the brain diseases topic or the brain regions topic (the dark blue sphere) on another topic receptacle. Then, the direct relations between the specific brain region *Hippocampus* and multiple brain diseases or the specific brain disease *Bipolar Disorder* and multiple brain regions are queried.

[1] See also the video of exploring indirect relations between topics: https://youtu.be/uOiMSwJ5AAQ

- ***Visualise Direct Relations between Brain Regions and Brain Diseases***

  Participants connect the *Direct Relations Explorer* with the *Brain Disease Topic Model* (Fig. 2.4 on p. 30) or the *Brain Regions Visualisation* (Fig. 2.5 on p. 31) to output the direct relations in the repository by the highlighted spheres of pink colours.

- ***Query Indirect Relations between Brain Regions and Brain Diseases***

  Participants select mental processes, genes or proteins as the intermediate topics in the *Indirect Relations Explorer* to query indirect relations between brain regions and brain diseases, Figs. 2.4 and 2.5 ④ on pp. 30 and 31.

- ***Visualise Indirect Relations between Brain Regions and Brain Diseases***

  The querying results of indirect relations are output by connecting the *Indirect Relations Explorer* with the *Direct Relations Explorer*. The indirect relations are visualised by green spheres in the *Brain Disease Topic Model* and the *Brain Regions Visualisation*. Participants can distinguish the direct and indirect relations between brain regions and brain diseases by a set of strongly contrasting colours of spheres, Figs. 2.4 and 2.5 on pp. 30 and 31.

The Unity-based code for exploring indirect relations is open-source and available on GitHub:

https://github.com/Boyu-Xu-projects/DatAR-prototype/tree/IndirectRelations-Project

B

Appendix: Exploring the Evolution of Indirect to Direct Relations

Timeline Study Informed Consent Form

Research project:
We have developed the extended DatAR prototype in an Augmented Reality (AR) environment to assist neuroscientists in exploring the relations between topics in scientific literature. We define a co-occurrence of two topics when they, for example, a brain region and a brain disease, are found in 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. An indirect relation between a brain region and a brain disease means there are no direct co-occurrences between them in the existing publications. However, other topics, such as genes or proteins, may have direct co-occurrences with brain regions or brain diseases. By identifying an intermediary topic, such as a gene or a protein, that co-occurs with both a brain region and a brain disease, an indirect relation between them is established. Our goal is to determine the utility of exploring publication-date dependent direct and indirect relations for neuroscientists in identifying fruitful experiments.

Privacy Information:
Your privacy is a priority in our research project. By participating, you acknowledge that the purpose is to evaluate the DatAR prototype for exploring neuroscience literature. Participation is voluntary, and you can withdraw at any time without penalty. For your contribution, you will receive a 15€ gift card. The process involves audio-recorded interviews and we will ensure your anonymity during analysis. Your questions and concerns are welcome, and your decision to participate remains confiden-

tial. Additionally, we want to assure you that the audio-recorded interviews collected and the answers you provide to a number of questions will be used only for research purposes. The audio-recorded interviews will be archived for scientific integrity but your anonymity will be strictly maintained. Your trust in our analysis and storage of your information is crucial to our study.

Please read the following carefully:

- I have been provided with information on the purpose of this project.
- I have been given the opportunity to ask questions about the project and my participation.
- I voluntarily agree to participate in this study.
- I agree that I will be paid a 15€ gift card for my participation.
- I understand that my participation in the research project involves being interviewed.
- I understand I can withdraw at any time without giving any reason and that I will not be penalized for withdrawing.
- I agree to the interviews being audio-recorded and transcribed.
- I understand the explanations on the use of the data in research, publications, sharing, and archiving.
- I understand that my participation will remain anonymous.

If you agree with all the above statements then please choose "I consent, begin the study" below:

We very much appreciate your time in conducting our study,

Boyu Xu, Department of Information and Computing Sciences, Utrecht University
On behalf of the DatAR Team

Please write your full name:

Please write your email address:

Please write your participant number:

Please write your research field(s):

Please write your academic qualification, such as a second year PhD student:

Please rate the following statements. “Literature exploration” includes both informal exploration and systematic literature review.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I have a lot of experience performing literature exploration	○	○	○	○	○	○	○
I find literature exploration an enjoyable activity	○	○	○	○	○	○	○
I find literature exploration a challenging activity	○	○	○	○	○	○	○
I consider myself knowledgeable in neuroscience	○	○	○	○	○	○	○
I have experience using Virtual Reality applications	○	○	○	○	○	○	○
I have experience using Augmented Reality applications	○	○	○	○	○	○	○

The Process of Exploring the Evolution of Indirect to Direct Relations

We described the process of exploring the evolution of indirect to direct relations between brain regions and brain diseases during the evaluation procedure (Sec. 3.6.2). This process consists of nine parts[1]:

- ***Construct the Brain Regions Visualisation or the Brain Disease Topic Model***

  The *Brain Regions Visualisation* is presented by dragging the relevant widget, Figs. 3.8, 3.9 and 3.10 on pp. 56–58. The *Brain Disease Topic Model* is visualised by dragging the relevant widget, Figs. 3.11, 3.12 and 3.13 on pp. 59–61.

- ***Query Direct Relations between Brain Regions and Brain Diseases***

  There are two receptacles on the *Direct Relations Explorer*, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61. Put the specific brain disease or the specific brain region on the concept receptacle (the light blue sphere), correspondingly, put the brain regions topic or the brain diseases topic (the dark blue sphere) on

[1] See also the video of exploring the evolution of indirect to direct relations: https://youtu.be/VPAbhcqDTlk

another topic receptacle. Then, the direct relations between the brain disease *Depressive Disorder* and multiple brain regions, Figs. 3.8, 3.9 and 3.10 on pp. 56–58, or the brain region *Cerebellar Vermis Structure* and multiple brain diseases are queried, Figs. 3.11, 3.12 and 3.13 on pp. 59–61.

- ***Visualise Direct Relations between Brain Regions and Brain Diseases***

  Participants connect the *Direct Relations Explorer* with the *Brain Regions Visualisation* or the *Brain Disease Topic Model* to output the direct relations in the repository by the highlighted spheres of pink colours, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61.

- ***Query Indirect Relations between Brain Regions and Brain Diseases***

  Participants select genes as the intermediate topics in the *Indirect Relations Explorer* to query indirect relations between brain regions and brain diseases, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61.

- ***Visualise Indirect Relations between Brain Regions and Brain Diseases***

  The querying results of indirect relations are output by connecting the *Indirect Relations Explorer* with the *Direct Relations Explorer*. The indirect relations are visualised by green spheres in the *Brain Regions Visualisation* and the *Brain Disease Topic Model*, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61.

- ***Constructing the Timeline Bar***

  The timeline bar is presented by dragging the *Timeline Selector*, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61.

- ***Visualise the Evolution of Indirect to Direct Relations***

  Participants drag the *Timeline Selector* to connect the *Direct Relations Explorer* with the *Brain Regions Visualisation* or the *Brain Disease Topic Model*.

- ***Visualise PMIDs***

  Participants drag the *Publications Viewer*, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61, to connect it with the *Timeline Selector* to display the PMIDs indicating indirect and direct relations.

- ***Explore the Evolution of Indirect to Direct Relations***

  Participants grab and drag the two handles of the *Timeline Selector* to select a specific time range, querying indirect and direct relations and also having access to the PMIDs indicating these relations within the selected time range, Figs. 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 on pp. 56–61.

The Unity-based code for exploring the evolution of indirect to direct relations is open-source and available on GitHub:

https://github.com/Boyu-Xu-projects/DatAR-prototype/tree/Timelines-project

C

Appendix: Exploring Indirect Relations via a Specific Intermediate Topic

Specific Intermediate Topic Study Informed Consent Form

Research project:
We have developed the extended DatAR prototype in an Augmented Reality (AR) environment to assist neuroscientists in exploring the relations between topics in scientific literature. We define a co-occurrence of two topics when they, for example, a brain region and a brain disease, are found in 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. An indirect relation between a brain region and a brain disease means there are no direct co-occurrences between them in the existing publications. However, other topics, such as genes or mental processes, may have direct co-occurrences with brain regions or brain diseases. By identifying an intermediary topic, such as a gene or a mental process, that co-occurs with both a brain region and a brain disease, an indirect relation between them is established. Our goal is to determine whether exploring the specific intermediate topics in indirect relations is useful for designing a useful experiment.

Privacy Information:
Your privacy is a priority in our research project. By participating, you acknowledge that the purpose is to evaluate the DatAR prototype for exploring neuroscience literature. Participation is voluntary, and you can withdraw at any time without penalty. For your contribution, you will receive a 15€ gift card. The process involves audio-recorded interviews and we will ensure your anonymity during analysis. Your questions and concerns are welcome, and your decision to participate remains confiden-

tial. Additionally, we want to assure you that the audio-recorded interviews collected and the answers you provide to a number of questions will be used only for research purposes. The audio-recorded interviews will be archived for scientific integrity but your anonymity will be strictly maintained. Your trust in our analysis and storage of your information is crucial to our study.

Please read the following carefully:

- I have been provided with information on the purpose of this project.
- I have been given the opportunity to ask questions about the project and my participation.
- I voluntarily agree to participate in this study.
- I agree that I will be paid a 15€ gift card for my participation.
- I understand that my participation in the research project involves being interviewed.
- I understand I can withdraw at any time without giving any reason and that I will not be penalized for withdrawing.
- I agree to the interviews being audio-recorded and transcribed.
- I understand the explanations on the use of the data in research, publications, sharing, and archiving.
- I understand that my participation will remain anonymous.

If you agree with all the above statements then please choose "I consent, begin the study" below:

We very much appreciate your time in conducting our study,

Boyu Xu, Department of Information and Computing Sciences, Utrecht University
On behalf of the DatAR Team

Please write your full name:

Please write your email address:

Please write your participant number:

Please write your research field(s):

Please write your academic qualification, such as a second year PhD student:

Please rate the following statements. "Literature exploration" includes both informal exploration and systematic literature review.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I have a lot of experience performing literature exploration	○	○	○	○	○	○	○
I find literature exploration an enjoyable activity	○	○	○	○	○	○	○
I find literature exploration a challenging activity	○	○	○	○	○	○	○
I consider myself knowledgeable in neuroscience	○	○	○	○	○	○	○
I have experience using Virtual Reality applications	○	○	○	○	○	○	○
I have experience using Augmented Reality applications	○	○	○	○	○	○	○

The Process of Exploring Indirect Relations via a Specific Intermediate Topic

We described the process of exploring indirect relations via specific intermediate topics to participants during the evaluation procedure (Sec. 4.5.2). This process consists of four parts:

- ***Construct the Brain Regions Visualisation or the Brain Disease Topic Model***

  The *Brain Regions Visualisation* is displayed by dragging the relevant widget (Fig. 4.9 on p. 94). The *Brain Disease Topic Model* is visualised by dragging the corresponding widget (Fig. 4.10 on p. 95).

- ***Query Direct Relations between a Selected Brain Disease or Brain Region and Genes, Proteins, and Mental Processes***

  There are three receptacles on the *Specific Intermediate Topic Explorer* (Figs. 4.9 and 4.10 on pp. 94 and 95). Place the selected brain disease or brain region on the middle receptacle (the light blue sphere), and correspondingly, place the brain region topic or brain disease topic (the dark blue sphere) on the left receptacle.

This triggers a query of the direct relations between the brain disease *Alzheimer's Disease* and genes, proteins, and mental processes (Fig. 4.9 on p. 94), or between the brain region *Cerebellar Vermis Structure* and genes, proteins, and mental processes (Fig. 4.10 on p. 95).

- ***Select a Specific Intermediate Topic***

  Participants can select an intermediate topic of interest, a gene, protein, or mental process, and place it on the right receptacle of the *Specific Intermediate Topic Explorer* (Figs. 4.9 and 4.10 on pp. 94 and 95). For example, selecting the Mental Process *Communication Response* queries indirect relations between *Alzheimer's Disease* and all brain regions (Fig. 4.9 on p. 94), and selecting the gene *LITAF Gene* queries indirect relations between *Cerebellar Vermis Structure* and all brain diseases (Fig. 4.10 on p. 95).

- ***Visualise Direct and Indirect Relations between Brain Regions and Brain Diseases via a Specific Intermediate Topic***

  Participants connect the *Specific Intermediate Topic Explorer* with the *Brain Regions Visualisation* or the *Brain Disease Topic Model* to output the direct and indirect relations. Direct and indirect relations are shown as highlighted pink spheres (direct relations) and green spheres (indirect relations), respectively, Figs. 4.9 and 4.10 on pp. 94 and 95.

The Unity-based code for exploring indirect relations via a specific intermediate topic is open-source and available on GitHub:

https://github.com/Boyu-Xu-projects/DatAR-prototype/tree/SpecificIntermediate Topics-project

D

Appendix: Visualisation Improvements in the DatAR Prototype

Suggestions for improving the visual representations in the *Brain Regions Visualisation* and the *Brain Disease Topic Model* are not a main part of exploring indirect relations in our research. Once the usefulness of topic-based literature exploration, rather than individual publications, in informing the design of a useful experiment is demonstrated through early-stage design work such as this research, we can investigate potential improvements to the user interface.

- **Filtering names of brain regions and brain diseases**

  Overlapping labels on spheres cause readability issues in both the *Brain Regions Visualisation* and the *Brain Disease Topic Model* (Figs. 1.3 and 1.4 on pp. 12 and 13). Participants proposed a potential solution, matching the label colours with those of the spheres and implementing a filter to display only labels related to direct and indirect relations, thus improving the readability of labels.

- **Grouping brain regions in the *Brain Regions Visualisation***

  Given that the *Brain Regions Visualisation* currently displays 3D brain structures through multiple spheres, such as two spheres representing the *Hippocampus* (Fig. 1.3 on p. 12), participants indicated that they could only understand the 3D brain region structures by reading the short names of brain regions and relying on their own knowledge. Participants suggested adding a grouped brain regions visualisation. Grouping brain regions using anatomical categories could improve the clarity of the 3D brain visualisation and facilitate a better understanding of the spatial representation of 3D brain structures.

- **Using the full names of brain regions in the *Brain Regions Visualisation***

While participants were able to clearly see the short names of brain regions in the visualisation, they found some unfamiliar abbreviations difficult to interpret. For example, short names like “Cb_Vermis” were harder to understand, whereas “Amyg_L” was easier due to participants’ familiarity with the term “Amygdala”. Participants suggested displaying full region names, at least for less common brain regions, to improve clarity.

- **Trying different colour sets for direct and indirect relations**

  The current colours used in the DatAR prototype to represent direct and indirect relations (pink and green, respectively) are not suitable for users with colour vision deficiency. Although this colour scheme works reasonably well in the HoloLens 2 AR headset, participants suggested conducting a small experiment to explore and evaluate alternative colour sets. For example, testing two distinct yet visually similar colours for direct and indirect relations.

List of SIKS-dissertations

2016 01 Syed Saiden Abbas (RUN), Recognition of Shapes by Humans and Machines
02 Michiel Christiaan Meulendijk (UU), Optimizing medication reviews through decision support: prescribing a better pill to swallow
03 Maya Sappelli (RUN), Knowledge Work in Context: User Centered Knowledge Worker Support
04 Laurens Rietveld (VUA), Publishing and Consuming Linked Data
05 Evgeny Sherkhonov (UvA), Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers
06 Michel Wilson (TUD), Robust scheduling in an uncertain environment
07 Jeroen de Man (VUA), Measuring and modeling negative emotions for virtual training
08 Matje van de Camp (TiU), A Link to the Past: Constructing Historical Social Networks from Unstructured Data
09 Archana Nottamkandath (VUA), Trusting Crowdsourced Information on Cultural Artefacts
10 George Karafotias (VUA), Parameter Control for Evolutionary Algorithms
11 Anne Schuth (UvA), Search Engines that Learn from Their Users
12 Max Knobbout (UU), Logics for Modelling and Verifying Normative Multi-Agent Systems
13 Nana Baah Gyan (VUA), The Web, Speech Technologies and Rural Development in West Africa - An ICT4D Approach
14 Ravi Khadka (UU), Revisiting Legacy Software System Modernization
15 Steffen Michels (RUN), Hybrid Probabilistic Logics - Theoretical Aspects, Algorithms and Experiments
16 Guangliang Li (UvA), Socially Intelligent Autonomous Agents that Learn from Human Reward
17 Berend Weel (VUA), Towards Embodied Evolution of Robot Organisms
18 Albert Meroño Peñuela (VUA), Refining Statistical Data on the Web
19 Julia Efremova (TU/e), Mining Social Structures from Genealogical Data
20 Daan Odijk (UvA), Context & Semantics in News & Web Search
21 Alejandro Moreno Célleri (UT), From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground

22 Grace Lewis (VUA), Software Architecture Strategies for Cyber-Foraging Systems
23 Fei Cai (UvA), Query Auto Completion in Information Retrieval
24 Brend Wanders (UT), Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach
25 Julia Kiseleva (TU/e), Using Contextual Information to Understand Searching and Browsing Behavior
26 Dilhan Thilakarathne (VUA), In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains
27 Wen Li (TUD), Understanding Geo-spatial Information on Social Media
28 Mingxin Zhang (TUD), Large-scale Agent-based Social Simulation - A study on epidemic prediction and control
29 Nicolas Höning (TUD), Peak reduction in decentralised electricity systems - Markets and prices for flexible planning
30 Ruud Mattheij (TiU), The Eyes Have It
31 Mohammad Khelghati (UT), Deep web content monitoring
32 Eelco Vriezekolk (UT), Assessing Telecommunication Service Availability Risks for Crisis Organisations
33 Peter Bloem (UvA), Single Sample Statistics, exercises in learning from just one example
34 Dennis Schunselaar (TU/e), Configurable Process Trees: Elicitation, Analysis, and Enactment
35 Zhaochun Ren (UvA), Monitoring Social Media: Summarization, Classification and Recommendation
36 Daphne Karreman (UT), Beyond R2D2: The design of nonverbal interaction behavior optimized for robot-specific morphologies
37 Giovanni Sileno (UvA), Aligning Law and Action - a conceptual and computational inquiry
38 Andrea Minuto (UT), Materials that Matter - Smart Materials meet Art & Interaction Design
39 Merijn Bruijnes (UT), Believable Suspect Agents; Response and Interpersonal Style Selection for an Artificial Suspect
40 Christian Detweiler (TUD), Accounting for Values in Design
41 Thomas King (TUD), Governing Governance: A Formal Framework for Analysing Institutional Design and Enactment Governance
42 Spyros Martzoukos (UvA), Combinatorial and Compositional Aspects of Bilingual Aligned Corpora
43 Saskia Koldijk (RUN), Context-Aware Support for Stress Self-Management: From Theory to Practice
44 Thibault Sellam (UvA), Automatic Assistants for Database Exploration
45 Bram van de Laar (UT), Experiencing Brain-Computer Interface Control
46 Jorge Gallego Perez (UT), Robots to Make you Happy
47 Christina Weber (UL), Real-time foresight - Preparedness for dynamic innovation networks
48 Tanja Buttler (TUD), Collecting Lessons Learned
49 Gleb Polevoy (TUD), Participation and Interaction in Projects. A Game-Theoretic Analysis
50 Yan Wang (TiU), The Bridge of Dreams: Towards a Method for Operational Performance Alignment in IT-enabled Service Supply Chains

---

2017 01 Jan-Jaap Oerlemans (UL), Investigating Cybercrime

02 Sjoerd Timmer (UU), Designing and Understanding Forensic Bayesian Networks using Argumentation
03 Daniël Harold Telgen (UU), Grid Manufacturing; A Cyber-Physical Approach with Autonomous Products and Reconfigurable Manufacturing Machines
04 Mrunal Gawade (CWI), Multi-core Parallelism in a Column-store
05 Mahdieh Shadi (UvA), Collaboration Behavior
06 Damir Vandic (EUR), Intelligent Information Systems for Web Product Search
07 Roel Bertens (UU), Insight in Information: from Abstract to Anomaly
08 Rob Konijn (VUA), Detecting Interesting Differences:Data Mining in Health Insurance Data using Outlier Detection and Subgroup Discovery
09 Dong Nguyen (UT), Text as Social and Cultural Data: A Computational Perspective on Variation in Text
10 Robby van Delden (UT), (Steering) Interactive Play Behavior
11 Florian Kunneman (RUN), Modelling patterns of time and emotion in Twitter #anticipointment
12 Sander Leemans (TU/e), Robust Process Mining with Guarantees
13 Gijs Huisman (UT), Social Touch Technology - Extending the reach of social touch through haptic technology
14 Shoshannah Tekofsky (TiU), You Are Who You Play You Are: Modelling Player Traits from Video Game Behavior
15 Peter Berck (RUN), Memory-Based Text Correction
16 Aleksandr Chuklin (UvA), Understanding and Modeling Users of Modern Search Engines
17 Daniel Dimov (UL), Crowdsourced Online Dispute Resolution
18 Ridho Reinanda (UvA), Entity Associations for Search
19 Jeroen Vuurens (UT), Proximity of Terms, Texts and Semantic Vectors in Information Retrieval
20 Mohammadbashir Sedighi (TUD), Fostering Engagement in Knowledge Sharing: The Role of Perceived Benefits, Costs and Visibility
21 Jeroen Linssen (UT), Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)
22 Sara Magliacane (VUA), Logics for causal inference under uncertainty
23 David Graus (UvA), Entities of Interest — Discovery in Digital Traces
24 Chang Wang (TUD), Use of Affordances for Efficient Robot Learning
25 Veruska Zamborlini (VUA), Knowledge Representation for Clinical Guidelines, with applications to Multimorbidity Analysis and Literature Search
26 Merel Jung (UT), Socially intelligent robots that understand and respond to human touch
27 Michiel Joosse (UT), Investigating Positioning and Gaze Behaviors of Social Robots: People's Preferences, Perceptions and Behaviors
28 John Klein (VUA), Architecture Practices for Complex Contexts
29 Adel Alhuraibi (TiU), From IT-BusinessStrategic Alignment to Performance: A Moderated Mediation Model of Social Innovation, and Enterprise Governance of IT"
30 Wilma Latuny (TiU), The Power of Facial Expressions
31 Ben Ruijl (UL), Advances in computational methods for QFT calculations
32 Thaer Samar (RUN), Access to and Retrievability of Content in Web Archives
33 Brigit van Loggem (OU), Towards a Design Rationale for Software Documentation: A Model of Computer-Mediated Activity
34 Maren Scheffel (OU), The Evaluation Framework for Learning Analytics
35 Martine de Vos (VUA), Interpreting natural science spreadsheets

36 Yuanhao Guo (UL), Shape Analysis for Phenotype Characterisation from High-throughput Imaging
37 Alejandro Montes Garcia (TU/e), WiBAF: A Within Browser Adaptation Framework that Enables Control over Privacy
38 Alex Kayal (TUD), Normative Social Applications
39 Sara Ahmadi (RUN), Exploiting properties of the human auditory system and compressive sensing methods to increase noise robustness in ASR
40 Altaf Hussain Abro (VUA), Steer your Mind: Computational Exploration of Human Control in Relation to Emotions, Desires and Social Support For applications in human-aware support systems
41 Adnan Manzoor (VUA), Minding a Healthy Lifestyle: An Exploration of Mental Processes and a Smart Environment to Provide Support for a Healthy Lifestyle
42 Elena Sokolova (RUN), Causal discovery from mixed and missing data with applications on ADHD datasets
43 Maaike de Boer (RUN), Semantic Mapping in Video Retrieval
44 Garm Lucassen (UU), Understanding User Stories - Computational Linguistics in Agile Requirements Engineering
45 Bas Testerink (UU), Decentralized Runtime Norm Enforcement
46 Jan Schneider (OU), Sensor-based Learning Support
47 Jie Yang (TUD), Crowd Knowledge Creation Acceleration
48 Angel Suarez (OU), Collaborative inquiry-based learning

2018 01 Han van der Aa (VUA), Comparing and Aligning Process Representations
02 Felix Mannhardt (TU/e), Multi-perspective Process Mining
03 Steven Bosems (UT), Causal Models For Well-Being: Knowledge Modeling, Model-Driven Development of Context-Aware Applications, and Behavior Prediction
04 Jordan Janeiro (TUD), Flexible Coordination Support for Diagnosis Teams in Data-Centric Engineering Tasks
05 Hugo Huurdeman (UvA), Supporting the Complex Dynamics of the Information Seeking Process
06 Dan Ionita (UT), Model-Driven Information Security Risk Assessment of Socio-Technical Systems
07 Jieting Luo (UU), A formal account of opportunism in multi-agent systems
08 Rick Smetsers (RUN), Advances in Model Learning for Software Systems
09 Xu Xie (TUD), Data Assimilation in Discrete Event Simulations
10 Julienka Mollee (VUA), Moving forward: supporting physical activity behavior change through intelligent technology
11 Mahdi Sargolzaei (UvA), Enabling Framework for Service-oriented Collaborative Networks
12 Xixi Lu (TU/e), Using behavioral context in process mining
13 Seyed Amin Tabatabaei (VUA), Computing a Sustainable Future
14 Bart Joosten (TiU), Detecting Social Signals with Spatiotemporal Gabor Filters
15 Naser Davarzani (UM), Biomarker discovery in heart failure
16 Jaebok Kim (UT), Automatic recognition of engagement and emotion in a group of children
17 Jianpeng Zhang (TU/e), On Graph Sample Clustering
18 Henriette Nakad (UL), De Notaris en Private Rechtspraak
19 Minh Duc Pham (VUA), Emergent relational schemas for RDF
20 Manxia Liu (RUN), Time and Bayesian Networks
21 Aad Slootmaker (OU), EMERGO: a generic platform for authoring and playing scenario-based serious games

22 Eric Fernandes de Mello Araújo (VUA), Contagious: Modeling the Spread of Behaviours, Perceptions and Emotions in Social Networks
23 Kim Schouten (EUR), Semantics-driven Aspect-Based Sentiment Analysis
24 Jered Vroon (UT), Responsive Social Positioning Behaviour for Semi-Autonomous Telepresence Robots
25 Riste Gligorov (VUA), Serious Games in Audio-Visual Collections
26 Roelof Anne Jelle de Vries (UT),Theory-Based and Tailor-Made: Motivational Messages for Behavior Change Technology
27 Maikel Leemans (TU/e), Hierarchical Process Mining for Scalable Software Analysis
28 Christian Willemse (UT), Social Touch Technologies: How they feel and how they make you feel
29 Yu Gu (TiU), Emotion Recognition from Mandarin Speech
30 Wouter Beek (VUA), The "K" in "semantic web" stands for "knowledge": scaling semantics to the web

---

2019 01 Rob van Eijk (UL),Web privacy measurement in real-time bidding systems. A graph-based approach to RTB system classification
02 Emmanuelle Beauxis Aussalet (CWI, UU), Statistics and Visualizations for Assessing Class Size Uncertainty
03 Eduardo Gonzalez Lopez de Murillas (TU/e), Process Mining on Databases: Extracting Event Data from Real Life Data Sources
04 Ridho Rahmadi (RUN), Finding stable causal structures from clinical data
05 Sebastiaan van Zelst (TU/e), Process Mining with Streaming Data
06 Chris Dijkshoorn (VUA), Nichesourcing for Improving Access to Linked Cultural Heritage Datasets
07 Soude Fazeli (TUD), Recommender Systems in Social Learning Platforms
08 Frits de Nijs (TUD), Resource-constrained Multi-agent Markov Decision Processes
09 Fahimeh Alizadeh Moghaddam (UvA), Self-adaptation for energy efficiency in software systems
10 Qing Chuan Ye (EUR), Multi-objective Optimization Methods for Allocation and Prediction
11 Yue Zhao (TUD), Learning Analytics Technology to Understand Learner Behavioral Engagement in MOOCs
12 Jacqueline Heinerman (VUA), Better Together
13 Guanliang Chen (TUD), MOOC Analytics: Learner Modeling and Content Generation
14 Daniel Davis (TUD), Large-Scale Learning Analytics: Modeling Learner Behavior & Improving Learning Outcomes in Massive Open Online Courses
15 Erwin Walraven (TUD), Planning under Uncertainty in Constrained and Partially Observable Environments
16 Guangming Li (TU/e), Process Mining based on Object-Centric Behavioral Constraint (OCBC) Models
17 Ali Hurriyetoglu (RUN),Extracting actionable information from microtexts
18 Gerard Wagenaar (UU), Artefacts in Agile Team Communication
19 Vincent Koeman (TUD), Tools for Developing Cognitive Agents
20 Chide Groenouwe (UU), Fostering technically augmented human collective intelligence
21 Cong Liu (TU/e), Software Data Analytics: Architectural Model Discovery and Design Pattern Detection
22 Martin van den Berg (VUA),Improving IT Decisions with Enterprise Architecture
23 Qin Liu (TUD), Intelligent Control Systems: Learning, Interpreting, Verification

24 Anca Dumitrache (VUA), Truth in Disagreement - Crowdsourcing Labeled Data for Natural Language Processing
25 Emiel van Miltenburg (VUA), Pragmatic factors in (automatic) image description
26 Prince Singh (UT), An Integration Platform for Synchromodal Transport
27 Alessandra Antonaci (OU), The Gamification Design Process applied to (Massive) Open Online Courses
28 Esther Kuindersma (UL), Cleared for take-off: Game-based learning to prepare airline pilots for critical situations
29 Daniel Formolo (VUA), Using virtual agents for simulation and training of social skills in safety-critical circumstances
30 Vahid Yazdanpanah (UT), Multiagent Industrial Symbiosis Systems
31 Milan Jelisavcic (VUA), Alive and Kicking: Baby Steps in Robotics
32 Chiara Sironi (UM), Monte-Carlo Tree Search for Artificial General Intelligence in Games
33 Anil Yaman (TU/e), Evolution of Biologically Inspired Learning in Artificial Neural Networks
34 Negar Ahmadi (TU/e), EEG Microstate and Functional Brain Network Features for Classification of Epilepsy and PNES
35 Lisa Facey-Shaw (OU), Gamification with digital badges in learning programming
36 Kevin Ackermans (OU), Designing Video-Enhanced Rubrics to Master Complex Skills
37 Jian Fang (TUD), Database Acceleration on FPGAs
38 Akos Kadar (OU), Learning visually grounded and multilingual representations

---

2020 01 Armon Toubman (UL), Calculated Moves: Generating Air Combat Behaviour
02 Marcos de Paula Bueno (UL), Unraveling Temporal Processes using Probabilistic Graphical Models
03 Mostafa Deghani (UvA), Learning with Imperfect Supervision for Language Understanding
04 Maarten van Gompel (RUN), Context as Linguistic Bridges
05 Yulong Pei (TU/e), On local and global structure mining
06 Preethu Rose Anish (UT), Stimulation Architectural Thinking during Requirements Elicitation - An Approach and Tool Support
07 Wim van der Vegt (OU), Towards a software architecture for reusable game components
08 Ali Mirsoleimani (UL),Structured Parallel Programming for Monte Carlo Tree Search
09 Myriam Traub (UU), Measuring Tool Bias and Improving Data Quality for Digital Humanities Research
10 Alifah Syamsiyah (TU/e), In-database Preprocessing for Process Mining
11 Sepideh Mesbah (TUD), Semantic-Enhanced Training Data AugmentationMethods for Long-Tail Entity Recognition Models
12 Ward van Breda (VUA), Predictive Modeling in E-Mental Health: Exploring Applicability in Personalised Depression Treatment
13 Marco Virgolin (CWI), Design and Application of Gene-pool Optimal Mixing Evolutionary Algorithms for Genetic Programming
14 Mark Raasveldt (CWI/UL), Integrating Analytics with Relational Databases
15 Konstantinos Georgiadis (OU), Smart CAT: Machine Learning for Configurable Assessments in Serious Games
16 Ilona Wilmont (RUN), Cognitive Aspects of Conceptual Modelling
17 Daniele Di Mitri (OU), The Multimodal Tutor: Adaptive Feedback from Multimodal Experiences

18 Georgios Methenitis (TUD), Agent Interactions & Mechanisms in Markets with Uncertainties: Electricity Markets in Renewable Energy Systems
19 Guido van Capelleveen (UT), Industrial Symbiosis Recommender Systems
20 Albert Hankel (VUA), Embedding Green ICT Maturity in Organisations
21 Karine da Silva Miras de Araujo (VUA), Where is the robot?: Life as it could be
22 Maryam Masoud Khamis (RUN), Understanding complex systems implementation through a modeling approach: the case of e-government in Zanzibar
23 Rianne Conijn (UT), The Keys to Writing: A writing analytics approach to studying writing processes using keystroke logging
24 Lenin da Nóbrega Medeiros (VUA/RUN), How are you feeling, human? Towards emotionally supportive chatbots
25 Xin Du (TU/e), The Uncertainty in Exceptional Model Mining
26 Krzysztof Leszek Sadowski (UU), GAMBIT: Genetic Algorithm for Model-Based mixed-Integer opTimization
27 Ekaterina Muravyeva (TUD), Personal data and informed consent in an educational context
28 Bibeg Limbu (TUD), Multimodal interaction for deliberate practice: Training complex skills with augmented reality
29 Ioan Gabriel Bucur (RUN), Being Bayesian about Causal Inference
30 Bob Zadok Blok (UL), Creatief, Creatiever, Creatiefst
31 Gongjin Lan (VUA), Learning better – From Baby to Better
32 Jason Rhuggenaath (TU/e), Revenue management in online markets: pricing and online advertising
33 Rick Gilsing (TU/e), Supporting service-dominant business model evaluation in the context of business model innovation
34 Anna Bon (UM), Intervention or Collaboration? Redesigning Information and Communication Technologies for Development
35 Siamak Farshidi (UU), Multi-Criteria Decision-Making in Software Production

---

2021 01 Francisco Xavier Dos Santos Fonseca (TUD),Location-based Games for Social Interaction in Public Space
02 Rijk Mercuur (TUD), Simulating Human Routines: Integrating Social Practice Theory in Agent-Based Models
03 Seyyed Hadi Hashemi (UvA), Modeling Users Interacting with Smart Devices
04 Ioana Jivet (OU), The Dashboard That Loved Me: Designing adaptive learning analytics for self-regulated learning
05 Davide Dell'Anna (UU), Data-Driven Supervision of Autonomous Systems
06 Daniel Davison (UT), "Hey robot, what do you think?" How children learn with a social robot
07 Armel Lefebvre (UU), Research data management for open science
08 Nardie Fanchamps (OU), The Influence of Sense-Reason-Act Programming on Computational Thinking
09 Cristina Zaga (UT), The Design of Robothings. Non-Anthropomorphic and Non-Verbal Robots to Promote Children's Collaboration Through Play
10 Quinten Meertens (UvA), Misclassification Bias in Statistical Learning
11 Anne van Rossum (UL), Nonparametric Bayesian Methods in Robotic Vision
12 Lei Pi (UL), External Knowledge Absorption in Chinese SMEs
13 Bob R. Schadenberg (UT), Robots for Autistic Children: Understanding and Facilitating Predictability for Engagement in Learning
14 Negin Samaeemofrad (UL), Business Incubators: The Impact of Their Support

15 Onat Ege Adali (TU/e), Transformation of Value Propositions into Resource Re-Configurations through the Business Services Paradigm
16 Esam A. H. Ghaleb (UM), Bimodal emotion recognition from audio-visual cues
17 Dario Dotti (UM), Human Behavior Understanding from motion and bodily cues using deep neural networks
18 Remi Wieten (UU), Bridging the Gap Between Informal Sense-Making Tools and Formal Systems - Facilitating the Construction of Bayesian Networks and Argumentation Frameworks
19 Roberto Verdecchia (VUA), Architectural Technical Debt: Identification and Management
20 Masoud Mansoury (TU/e), Understanding and Mitigating Multi-Sided Exposure Bias in Recommender Systems
21 Pedro Thiago Timbó Holanda (CWI), Progressive Indexes
22 Sihang Qiu (TUD), Conversational Crowdsourcing
23 Hugo Manuel Proença (UL), Robust rules for prediction and description
24 Kaijie Zhu (TU/e), On Efficient Temporal Subgraph Query Processing
25 Eoin Martino Grua (VUA), The Future of E-Health is Mobile: Combining AI and Self-Adaptation to Create Adaptive E-Health Mobile Applications
26 Benno Kruit (CWI/VUA), Reading the Grid: Extending Knowledge Bases from Human-readable Tables
27 Jelte van Waterschoot (UT), Personalized and Personal Conversations: Designing Agents Who Want to Connect With You
28 Christoph Selig (UL), Understanding the Heterogeneity of Corporate Entrepreneurship Programs

---

2022 01 Judith van Stegeren (UT), Flavor text generation for role-playing video games
02 Paulo da Costa (TU/e), Data-driven Prognostics and Logistics Optimisation: A Deep Learning Journey
03 Ali el Hassouni (VUA), A Model A Day Keeps The Doctor Away: Reinforcement Learning For Personalized Healthcare
04 Ünal Aksu (UU), A Cross-Organizational Process Mining Framework
05 Shiwei Liu (TU/e), Sparse Neural Network Training with In-Time Over-Parameterization
06 Reza Refaei Afshar (TU/e), Machine Learning for Ad Publishers in Real Time Bidding
07 Sambit Praharaj (OU), Measuring the Unmeasurable? Towards Automatic Co-located Collaboration Analytics
08 Maikel L. van Eck (TU/e), Process Mining for Smart Product Design
09 Oana Andreea Inel (VUA), Understanding Events: A Diversity-driven Human-Machine Approach
10 Felipe Moraes Gomes (TUD), Examining the Effectiveness of Collaborative Search Engines
11 Mirjam de Haas (UT), Staying engaged in child-robot interaction, a quantitative approach to studying preschoolers' engagement with robots and tasks during second-language tutoring
12 Guanyi Chen (UU), Computational Generation of Chinese Noun Phrases
13 Xander Wilcke (VUA), Machine Learning on Multimodal Knowledge Graphs: Opportunities, Challenges, and Methods for Learning on Real-World Heterogeneous and Spatially-Oriented Knowledge
14 Michiel Overeem (UU), Evolution of Low-Code Platforms
15 Jelmer Jan Koorn (UU), Work in Process: Unearthing Meaning using Process Mining
16 Pieter Gijsbers (TU/e), Systems for AutoML Research

17 Laura van der Lubbe (VUA), Empowering vulnerable people with serious games and gamification
18 Paris Mavromoustakos Blom (TiU), Player Affect Modelling and Video Game Personalisation
19 Bilge Yigit Ozkan (UU), Cybersecurity Maturity Assessment and Standardisation
20 Fakhra Jabeen (VUA), Dark Side of the Digital Media - Computational Analysis of Negative Human Behaviors on Social Media
21 Seethu Mariyam Christopher (UM), Intelligent Toys for Physical and Cognitive Assessments
22 Alexandra Sierra Rativa (TiU), Virtual Character Design and its potential to foster Empathy, Immersion, and Collaboration Skills in Video Games and Virtual Reality Simulations
23 Ilir Kola (TUD), Enabling Social Situation Awareness in Support Agents
24 Samaneh Heidari (UU), Agents with Social Norms and Values - A framework for agent based social simulations with social norms and personal values
25 Anna L.D. Latour (UL), Optimal decision-making under constraints and uncertainty
26 Anne Dirkson (UL), Knowledge Discovery from Patient Forums: Gaining novel medical insights from patient experiences
27 Christos Athanasiadis (UM), Emotion-aware cross-modal domain adaptation in video sequences
28 Onuralp Ulusoy (UU), Privacy in Collaborative Systems
29 Jan Kolkmeier (UT), From Head Transform to Mind Transplant: Social Interactions in Mixed Reality
30 Dean De Leo (CWI), Analysis of Dynamic Graphs on Sparse Arrays
31 Konstantinos Traganos (TU/e), Tackling Complexity in Smart Manufacturing with Advanced Manufacturing Process Management
32 Cezara Pastrav (UU), Social simulation for socio-ecological systems
33 Brinn Hekkelman (CWI/TUD), Fair Mechanisms for Smart Grid Congestion Management
34 Nimat Ullah (VUA), Mind Your Behaviour: Computational Modelling of Emotion & Desire Regulation for Behaviour Change
35 Mike E.U. Ligthart (VUA), Shaping the Child-Robot Relationship: Interaction Design Patterns for a Sustainable Interaction

---

2023 01 Bojan Simoski (VUA), Untangling the Puzzle of Digital Health Interventions
02 Mariana Rachel Dias da Silva (TiU), Grounded or in flight? What our bodies can tell us about the whereabouts of our thoughts
03 Shabnam Najafian (TUD), User Modeling for Privacy-preserving Explanations in Group Recommendations
04 Gineke Wiggers (UL), The Relevance of Impact: bibliometric-enhanced legal information retrieval
05 Anton Bouter (CWI), Optimal Mixing Evolutionary Algorithms for Large-Scale Real-Valued Optimization, Including Real-World Medical Applications
06 António Pereira Barata (UL), Reliable and Fair Machine Learning for Risk Assessment
07 Tianjin Huang (TU/e), The Roles of Adversarial Examples on Trustworthiness of Deep Learning
08 Lu Yin (TU/e), Knowledge Elicitation using Psychometric Learning
09 Xu Wang (VUA), Scientific Dataset Recommendation with Semantic Techniques
10 Dennis J.N.J. Soemers (UM), Learning State-Action Features for General Game Playing
11 Fawad Taj (VUA), Towards Motivating Machines: Computational Modeling of the Mechanism of Actions for Effective Digital Health Behavior Change Applications

	12	Tessel Bogaard (VUA), Using Metadata to Understand Search Behavior in Digital Libraries
	13	Injy Sarhan (UU), Open Information Extraction for Knowledge Representation
	14	Selma Čaušević (TUD), Energy resilience through self-organization
	15	Alvaro Henrique Chaim Correia (TU/e), Insights on Learning Tractable Probabilistic Graphical Models
	16	Peter Blomsma (TiU), Building Embodied Conversational Agents: Observations on human nonverbal behaviour as a resource for the development of artificial characters
	17	Meike Nauta (UT), Explainable AI and Interpretable Computer Vision – From Oversight to Insight
	18	Gustavo Penha (TUD), Designing and Diagnosing Models for Conversational Search and Recommendation
	19	George Aalbers (TiU), Digital Traces of the Mind: Using Smartphones to Capture Signals of Well-Being in Individuals
	20	Arkadiy Dushatskiy (TUD), Expensive Optimization with Model-Based Evolutionary Algorithms applied to Medical Image Segmentation using Deep Learning
	21	Gerrit Jan de Bruin (UL), Network Analysis Methods for Smart Inspection in the Transport Domain
	22	Alireza Shojaifar (UU), Volitional Cybersecurity
	23	Theo Theunissen (UU), Documentation in Continuous Software Development
	24	Agathe Balayn (TUD), Practices Towards Hazardous Failure Diagnosis in Machine Learning
	25	Jurian Baas (UU), Entity Resolution on Historical Knowledge Graphs
	26	Loek Tonnaer (TU/e), Linearly Symmetry-Based Disentangled Representations and their Out-of-Distribution Behaviour
	27	Ghada Sokar (TU/e), Learning Continually Under Changing Data Distributions
	28	Floris den Hengst (VUA), Learning to Behave: Reinforcement Learning in Human Contexts
	29	Tim Draws (TUD), Understanding Viewpoint Biases in Web Search Results
2024	01	Daphne Miedema (TU/e), On Learning SQL: Disentangling concepts in data systems education
	02	Emile van Krieken (VUA), Optimisation in Neurosymbolic Learning Systems
	03	Feri Wijayanto (RUN), Automated Model Selection for Rasch and Mediation Analysis
	04	Mike Huisman (UL), Understanding Deep Meta-Learning
	05	Yiyong Gou (UM), Aerial Robotic Operations: Multi-environment Cooperative Inspection & Construction Crack Autonomous Repair
	06	Azqa Nadeem (TUD), Understanding Adversary Behavior via XAI: Leveraging Sequence Clustering to Extract Threat Intelligence
	07	Parisa Shayan (TiU), Modeling User Behavior in Learning Management Systems
	08	Xin Zhou (UvA), From Empowering to Motivating: Enhancing Policy Enforcement through Process Design and Incentive Implementation
	09	Giso Dal (UT), Probabilistic Inference Using Partitioned Bayesian Networks
	10	Cristina-Iulia Bucur (VUA), Linkflows: Towards Genuine Semantic Publishing in Science
	11	withdrawn
	12	Peide Zhu (TUD), Towards Robust Automatic Question Generation For Learning
	13	Enrico Liscio (TUD), Context-Specific Value Inference via Hybrid Intelligence
	14	Larissa Capobianco Shimomura (TU/e), On Graph Generating Dependencies and their Applications in Data Profiling

15 Ting Liu (VUA), A Gut Feeling: Biomedical Knowledge Graphs for Interrelating the Gut Microbiome and Mental Health
16 Arthur Barbosa Câmara (TUD), Designing Search-as-Learning Systems
17 Razieh Alidoosti (VUA), Ethics-aware Software Architecture Design
18 Laurens Stoop (UU), Data Driven Understanding of Energy-Meteorological Variability and its Impact on Energy System Operations
19 Azadeh Mozafari Mehr (TU/e), Multi-perspective Conformance Checking: Identifying and Understanding Patterns of Anomalous Behavior
20 Ritsart Anne Plantenga (UL), Omgang met Regels
21 Federica Vinella (UU), Crowdsourcing User-Centered Teams
22 Zeynep Ozturk Yurt (TU/e), Beyond Routine: Extending BPM for Knowledge-Intensive Processes with Controllable Dynamic Contexts
23 Jie Luo (VUA), Lamarck's Revenge: Inheritance of Learned Traits Improves Robot Evolution
24 Nirmal Roy (TUD), Exploring the effects of interactive interfaces on user search behaviour
25 Alisa Rieger (TUD), Striving for Responsible Opinion Formation in Web Search on Debated Topics
26 Tim Gubner (CWI), Adaptively Generating Heterogeneous Execution Strategies using the VOILA Framework
27 Lincen Yang (UL), Information-theoretic Partition-based Models for Interpretable Machine Learning
28 Leon Helwerda (UL), Grip on Software: Understanding development progress of Scrum sprints and backlogs
29 David Wilson Romero Guzman (VUA), The Good, the Efficient and the Inductive Biases: Exploring Efficiency in Deep Learning Through the Use of Inductive Biases
30 Vijanti Ramautar (UU), Model-Driven Sustainability Accounting
31 Ziyu Li (TUD), On the Utility of Metadata to Optimize Machine Learning Workflows
32 Vinicius Stein Dani (UU), The Alpha and Omega of Process Mining
33 Siddharth Mehrotra (TUD), Designing for Appropriate Trust in Human-AI interaction
34 Robert Deckers (VUA), From Smallest Software Particle to System Specification - MuDForM: Multi-Domain Formalization Method
35 Sicui Zhang (TU/e), Methods of Detecting Clinical Deviations with Process Mining: a fuzzy set approach
36 Thomas Mulder (TU/e), Optimization of Recursive Queries on Graphs
37 James Graham Nevin (UvA), The Ramifications of Data Handling for Computational Models
38 Christos Koutras (TUD), Tabular Schema Matching for Modern Settings
39 Paola Lara Machado (TU/e), The Nexus between Business Models and Operating Models: From Conceptual Understanding to Actionable Guidance
40 Montijn van de Ven (TU/e), Guiding the Definition of Key Performance Indicators for Business Models
41 Georgios Siachamis (TUD), Adaptivity for Streaming Dataflow Engines
42 Emmeke Veltmeijer (VUA), Small Groups, Big Insights: Understanding the Crowd through Expressive Subgroup Analysis
43 Cedric Waterschoot (KNAW Meertens Instituut), The Constructive Conundrum: Computational Approaches to Facilitate Constructive Commenting on Online News Platforms
44 Marcel Schmitz (OU), Towards learning analytics-supported learning design

45 Sara Salimzadeh (TUD), Living in the Age of AI: Understanding Contextual Factors that Shape Human-AI Decision-Making
46 Georgios Stathis (Leiden University), Preventing Disputes: Preventive Logic, Law & Technology
47 Daniel Daza (VUA), Exploiting Subgraphs and Attributes for Representation Learning on Knowledge Graphs
48 Ioannis Petros Samiotis (TUD), Crowd-Assisted Annotation of Classical Music Compositions

---

2025 01 Max van Haastrecht (UL), Transdisciplinary Perspectives on Validity: Bridging the Gap Between Design and Implementation for Technology-Enhanced Learning Systems
02 Jurgen van den Hoogen (JADS), Time Series Analysis Using Convolutional Neural Networks
03 Andra-Denis Ionescu (TUD), Feature Discovery for Data-Centric AI
04 Rianne Schouten (TU/e), Exceptional Model Mining for Hierarchical Data
05 Nele Albers (TUD), Psychology-Informed Reinforcement Learning for Situated Virtual Coaching in Smoking Cessation
06 Daniël Vos (TUD), Decision Tree Learning: Algorithms for Robust Prediction and Policy Optimization
07 Ricky Maulana Fajri (TU/e), Towards Safer Active Learning: Dealing with Unwanted Biases, Graph-Structured Data, Adversary, and Data Imbalance
08 Stefan Bloemheuvel (TiU), Spatio-Temporal Analysis Through Graphs: Predictive Modeling and Graph Construction
09 Fadime Kaya (VUA), Decentralized Governance Design - A Model-Based Approach
10 Zhao Yang (UL), Enhancing Autonomy and Efficiency in Goal-Conditioned Reinforcement Learning
11 Shahin Sharifi Noorian (TUD), From Recognition to Understanding: Enriching Visual Models Through Multi-Modal Semantic Integration
12 Lijun Lyu (TUD), Interpretability in Neural Information Retrieval
13 Fuda van Diggelen (VUA), Robots Need Some Education: on the complexity of learning in evolutionary robotics
14 Gennaro Gala (TU/e), Probabilistic Generative Modeling with Latent Variable Hierarchies
15 Michiel van der Meer (UL), Opinion Diversity through Hybrid Intelligence
16 Monika Grewal (TU Delft), Deep Learning for Landmark Detection, Segmentation, and Multi-Objective Deformable Registration in Medical Imaging
17 Matteo De Carlo (VUA), Real Robot Reproduction: Towards Evolving Robotic Ecosystems
18 Anouk Neerincx (UU), Robots That Care: How Social Robots Can Boost Children's Mental Wellbeing
19 Fang Hou (UU), Trust in Software Ecosystems
20 Alexander Melchior (UU), Modelling for Policy is More Than Policy Modelling (The Useful Application of Agent-Based Modelling in Complex Policy Processes)
21 Mandani Ntekouli (UM), Bridging Individual and Group Perspectives in Psychopathology: Computational Modeling Approaches using Ecological Momentary Assessment Data
22 Hilde Weerts (TU/e), Decoding Algorithmic Fairness: Towards Interdisciplinary Understanding of Fairness and Discrimination in Algorithmic Decision-Making
23 Roderick van der Weerdt (VUA), IoT Measurement Knowledge Graphs: Constructing, Working and Learning with IoT Measurement Data as a Knowledge Graph
24 Zhong Li (UL), Trustworthy Anomaly Detection for Smart Manufacturing

25 Kyana van Eijndhoven (TiU), A Breakdown of Breakdowns: Multi-Level Team Coordination Dynamics under Stressful Conditions
26 Tom Pepels (UM), Monte-Carlo Tree Search is Work in Progress
27 Danil Provodin (JADS, TU/e), Sequential Decision Making Under Complex Feedback
28 Jinke He (TU Delft), Exploring Learned Abstract Models for Efficient Planning and Learning
29 Erik van Haeringen (VUA), Mixed Feelings: Simulating Emotion Contagion in Groups
30 Myrthe Reuver (VUA), A Puzzle of Perspectives: Interdisciplinary Language Technology for Responsible News Recommendation
31 Gebrekirstos Gebreselassie Gebremeskel (RUN), Spotlight on Recommender Systems: Contributions to Selected Components in the Recommendation Pipeline
32 Ryan Brate (UU), Words Matter: A Computational Toolkit for Charged Terms
33 Merle Reimann (VUA), Speaking the Same Language: Spoken Capability Communication in Human-Agent and Human-Robot Interaction
34 Eduard C. Groen (UU), Crowd-Based Requirements Engineering
35 Urja Khurana (VUA), From Concept To Impact: Toward More Robust Language Model Deployment
36 Anna Maria Wegmann (UU), Say the Same but Differently: Computational Approaches to Stylistic Variation and Paraphrasing
37 Chris Kamphuis (RUN), Exploring Relations and Graphs for Information Retrieval
38 Valentina Maccatrozzo (VUA), Break the Bubble: Semantic Patterns for Serendipity
39 Dimitrios Alivanistos (VUA), Knowledge Graphs & Transformers for Hypothesis Generation: Accelerating Scientific Discovery in the Era of Artificial Intelligence
40 Stefan Grafberger (UvA), Declarative Machine Learning Pipeline Management via Logical Query Plans
41 Mozhgan Vazifehdoostirani (TU/e), Leveraging Process Flexibility to Improve Process Outcome - From Descriptive Analytics to Actionable Insights
42 Margherita Martorana (VUA), Semantic Interpretation of Dataless Tables: a metadata-driven approach for findable, accessible, interoperable and reusable restricted access data
43 Krist Shingjergji (OU), Sense the Classroom - Using AI to Detect and Respond to Learning-Centered Affective States in Online Education
44 Robbert Reijnen (TU/e), Dynamic Algorithm Configuration for Machine Scheduling Using Deep Reinforcement Learning
45 Anjana Mohandas Sheeladevi (VUA), Occupant-Centric Energy Management: Balancing Privacy, Well-being and Sustainability in Smart Buildings
46 Ya Song (TU/e), Graph Neural Networks for Modeling Temporal and Spatial Dimensions in Industrial Decision-making
47 Tom Kouwenhoven (UL), Collaborative Meaning-Making. The Emergence of Novel Languages in Humans, Machines, and Human-Machine Interactions
48 Evy van Weelden (TiU), Integrating Virtual Reality and Neurophysiology in Flight Training
49 Selene Báez Santamaría (VUA), Knowledge-centered conversational agents with a drive to learn
50 Lea Krause (VUA), Contextualising Conversational AI
51 Jiaxu Zhao (TU/e), Understanding and Mitigating Unwanted Biases in Generative Language Models
52 Qiao Xiao (TU/e), Model, Data and Communication Sparsity for Efficient Training of Neural Networks

53 Gaole He (TUD), Towards Effective Human-AI Collaboration: Promoting Appropriate Reliance on AI Systems
54 Go Sugimoto (VUA), MISSING LINKS Investigating the Quality of Linked Data and its Tools in Cultural Heritage and Digital Humanities
55 Sietze Kai Kuilman (TUD), AI that Glitters is Not Gold: Requirements for Meaningful Control of AI Systems
56 Wijnand van Woerkom (UU), A Fortiori Case-Based Reasoning: Formal Studies with Applications in Artificial Intelligence and Law
57 Syeda Amna Sohail (UT), Privacy-Utility Trade-Off in Healthcare Metadata Sharing and Beyond: A Normative and Empirical Evaluation at Inter and Intra Organizational Levels
58 Junhan Wen (TUD), "From iMage to Market": Machine-Learning-Empowered Fruit Supply
59 Mohsen Abbaspour Onari (TU/e), From Explanation to Trust: Modeling and Measuring Trust in Explainable Decision Support
60 Marcel Jurriaan Robeer (UU), Beyond Trust: A Causal Approach to Explainable AI in Law Enforcement
61 Shuai Wang (VUA), Links in Large Integrated Knowledge Graphs: Analysis, Refinement, and Domain Applications
62 Khaleel Asyraaf Mat Sanusi (OU), Augmenting a learning model within immersive learning environments for psychomotor skills
63 Rashid Zaman (TU/e), Online Conformance Checking on Degraded Data
64 Jens d'Hondt (TU/e), Effective and Efficient Multivariate Similarity Search
65 Aswin Balasubramaniam (UT), Disentangling Runner Drone Interaction Potentialities

---

2026 01 Pei-Yu Chen (TUD), Human-Agent Alignment Dialogues: Eliciting User Information at Runtime for Personalized Behavior Support
02 Hezha Hassan Mohammedkhan (TiU), Estimating Body Measurements of Children from 2D Images: Towards the Automatic Detection of Malnutrition
03 Kyriakos Psarakis (TUD), Democratizing Scalable Cloud Applications: Transactional Stateful Functions on Streaming Dataflows

Acknowledgements

Lynda, my wise supervisor, once told me: “This is your PhD. Your supervisors are your senior advisors; others may pass through your journey, offering advice, support, or sometimes even confusion. You should learn to draw on the strengths of those around you, but in the end, it is you who must choose how to carry on.” Looking back, I now see that what I thought was a tough journey was already surrounded by patient, supportive, and brilliant people helping me chip away.

First on the list are the people I worked with. My PhD research could not have continued, nor taken the form it has, without the support of my supervisors: Lynda Hardman and Wolfgang Hürst. Lynda provided clarity, wisdom, and consistent belief in my potential throughout this journey. She taught me how to conduct research and how to communicate it through clear and compelling stories. She supported my interest in art and design, encouraging me to attend seminars and workshops. Her feedback helped shape the structure of my thesis and identify core contributions. Beyond academic guidance, she gave me confidence during uncertain moments and reminded me to maintain a healthy balance between work and life. Wolfgang offered practical and thoughtful advice. He reminded me to organise my research plans and timelines and to stay focused. I appreciated our meetings, where we discussed not only research but also life and, occasionally, random trivia.

I am thankful to Zhisheng Huang at Vrije Universiteit Amsterdam for creating the co-occurrence repository between topics in the Knowledge Graphs of Brain Science, and to CWI for hosting it. I also thank Ivar Troost at Utrecht University for implementing the Direct Relations Explorer in the DatAR prototype, which formed a foundation for my work. Many thanks to Ghazaleh Tanhaei for her helpful suggestions in the evaluation protocol with neuroscientists. I am grateful to Hyowon Lee for his valuable feedback on the design rationale discussed in Chapter 2. A special thanks to our master’s student, Dimitris Vogiatzopoulos, whose hard work is reflected in Chapter 4; his contributions to coding were indispensable.

I thank the 25 neuroscientists from UMC Utrecht, Amsterdam UMC, UMC Groningen, and VU who participated in interviews and evaluation sessions. Their willingness to engage with unfamiliar tools and offer honest feedback contributed to the practical relevance of this research. I am especially grateful to Anna van Harmelen (VU), Ziyi Wang (UMC Utrecht), and Jiaojiao Guo (UMC Utrecht), who participated in all three studies and provided continuous support throughout the research.

I am also grateful to my colleagues at CWI: Laura, Davide, Savvina, Andrei, Thomas, Manel, and many others, for the chocolate breaks and the shared laughter during group outings. Thanks as well to colleagues in the VIG group at Utrecht University: Michael, Evanthia, Tamara, and many others. Both groups provided valuable spaces to present my work and receive feedback from different perspectives. Their comments on my abstracts and papers directly improved the quality of this research. I would also like to thank the broader research community I encountered through the Dagstuhl Seminar[1], MMM 2024, CHIIR 2024, and IEEE VR 2025. The feedback and encouragement from these communities helped polish this research.

I am grateful to Mascha van der Voort, Hyowon Lee, Alex Telea, Remco Veltkamp and Anja Volk for serving as members of my PhD committee and for the thoughtful feedback on my manuscript.

To my best friend, Feng, thank you for your patience, for understanding what matters to me, and for always being there, through moments of burnout and breakthroughs alike.

To my parents, I cannot begin to describe how much your long-distance love carried me. The greatest gift you have given me is the freedom to pursue what I love. You never asked why I chose this path, only how you could support me in walking it. Thank you for believing in me when I doubted myself.

And finally, to myself, the version of me who started this journey with more questions than answers, thank you for staying curious, for being afraid and trying anyway.

[1] Dagstuhl Seminar 24301 Art, Visual Illusions, and Data Visualization https://www.dagstuhl.de/24301
Dagstuhl Seminar 25442 Augmenting Human Creativity with AI https://www.dagstuhl.de/25442