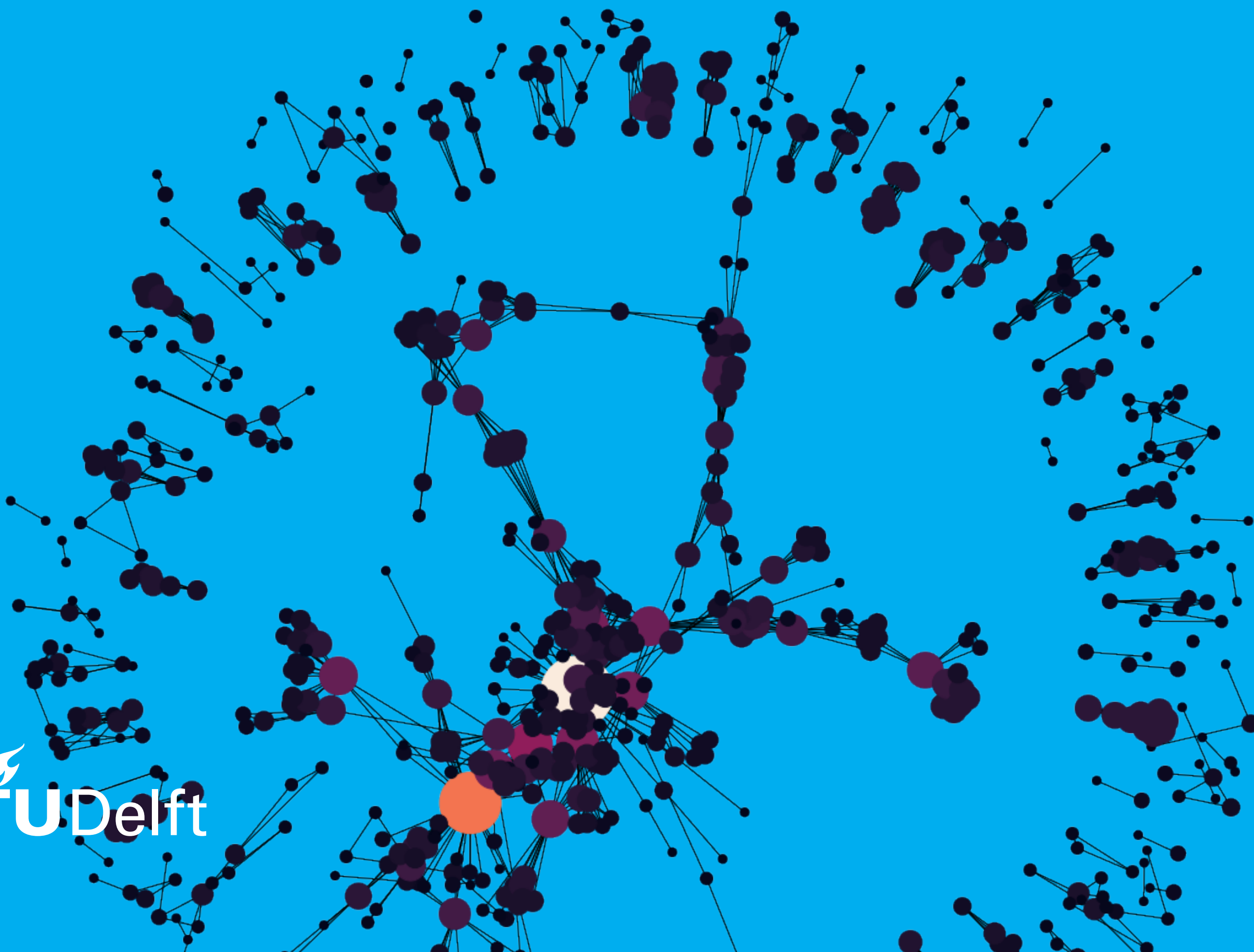
Changing mental models by developing a scientist-journalist recommender system

T.T. Bruyn



Changing mental models by developing a scientist-journalist recommender system

by

T.T. Bruyn

Student number
4240820

Submitted to the department of
Intelligent Systems
and the department of
Science Education and Communication
of Delft University of Technology

in partial fulfilment of the requirements for a Master of Science in

Computer Science

and

Science Education and Communication

Thesis committee Computer Science	Dr.ir. H. Wang,	Intelligent Systems, supervisor
	Prof.dr.ir. R.E. Kooij,	Network Architectures and Services
	Dr.ir. S.M. Flipse	Science Education and Communication
Thesis committee Science Education and Communication	Dr.ir. S.M. Flipse	Science Education and Communication, supervisor
	Dr. É. Kalmár	Science Education and Communication
	Dr.ir. H. Wang,	Intelligent Systems
	Dr. M.C.A. van der Sanden	Science Education and Communication

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TUDelft

Preface

It's been a ride.

I started my thesis with a simple goal: It had to be useful. My greatest frustration with studying is working for the waste bin. You pour your heart and soul into a project, work to make something you're proud of, and all that happens is that a professor looks at it, gives you a thumbs up, and that's it. The master thesis is supposed to be the crown jewel of your academic education, a moment to display the skills you have learned over the years. I was determined to combine my computer science skills and communication skills to do something to help real people who are facing real problems. When my professor came up with the project to educate communication professionals on AI by developing a digital tool, it was perfect. We were off to the races.

That was well over three years ago. In the meantime, a global pandemic came and went - a time of uncertainty, confusion, and isolation. I founded Student Onbeperkt, an organisation for students with a disability that has since grown to over 200 members - a source of joy, purpose and connection. I've completed a year-long full-time traineeship - a taste of what's to come, where I've discovered many things that suit me and a few that don't. One of my master's programmes got cancelled, with no adequate explanation other than university politics - something that still makes me and many other students angry and which resulted in my supervisor getting a burnout.

And through it all, a thesis with its own highs and lows. A process of learning but also of running into walls. A process of trial and error that at times felt like two steps forward and one step back. But also a process involving real-world people and real-world data, which, though messy, means a real preparation for the working life I'm headed into. It also means real-world impact. As you can read in this report, I have facilitated a journey of realisation about the (im)possibilities of digital technology for the people I have worked with. It might not have been what I envisioned when I started on this journey, but looking back, I am happy with where I now stand.

I want to thank my supervisors. I'm blessed with supervisors who care, about both the project and the individual. I want to thank Dr. Maarten van der Sanden, who came up with the idea for this project. You have inspired me throughout the CDI master. You have shown me that communication isn't a fixed skill set that some people are better at than others but about finding a style that works with your personality and maximises your strengths. I want to thank Dr. Steven Flipse, who took over my supervision after Maarten got a burnout. I remember asking you in our first meeting whether you had the time and capacity to supervise an additional student during this hectic time for the entire staff. Your answer was straightforward: Tim, you need to graduate, so we're going to make this happen. I am grateful that, without hesitation, you were always willing to put in the time and effort to support me. I want to thank Dr. Huijuan Wang, who I've gotten to know as someone who truly appreciates the beauty of mathematics. I am grateful for your kindness, patience and flexibility through a process that wasn't always easy. I also want to thank the team of communication professionals I've worked with, who - for privacy reasons - I won't mention by name here. I want to thank you for all the work that you put in to make this project happen together.

Lastly, I want to thank everyone who has supported me in big and small ways through these past years. Thank you to all my friends who celebrated every victory and helped me through every crisis. Thank you, all you lovely CDI people who made the writing process bearable. And, of course, a big, big thank you to my family, who have always been there for me.

This report concludes my almost eleven years at Delft University of Technology, more than a third of my life. So much of the person I am today has been shaped here. I'm glad to finally move to the next phase. I've already dabbed my toe in the working life, I am happy to now fully dive in.

Tim Bruyn
Delft, May 2023

Contents

Abstract

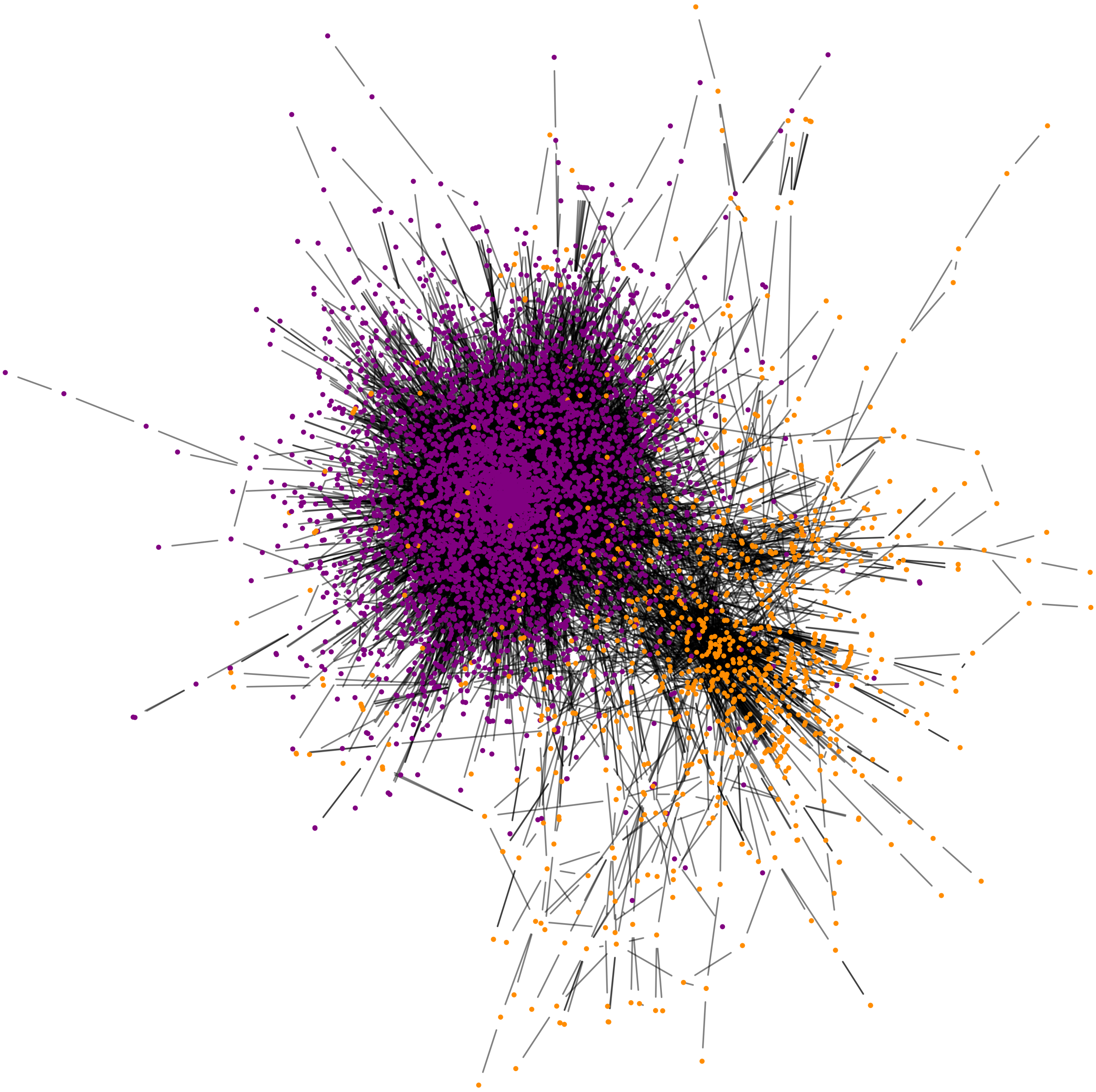
This research is a case study investigating the effect of participation in the development process of a network-based scientist-journalist recommender system on the mental model of digital innovation of a team of communication professionals from Delft University of Technology and Naturalis Biodiversity Center.

Communication professionals at research institutes are tasked with connecting scientists and journalists. The recommender system supports this process by recommending scientist-journalist connections based on data from previous collaborations. A scientist collaboration network, a journalist collaboration network and a scientist-journalist collaboration network are combined into a multilayer network. A recommender system is designed based on centrality metrics in the scientist and journalist collaboration networks and distance metrics in the multilayer network. In contrast to traditional link prediction problems - which aim to predict what links are most likely to form in the network - the problem in this thesis is how to recommend the most likely link for a single node, i.e. the most likely scientist links for a given journalist or most likely journalist links for a given scientist. A novel evaluation method is created to evaluate the performance of the recommender system.

The development of this system is used as a vessel to research how participation in a digital development process affects the mental model of digital innovation. This research contributes to addressing the lack of understanding of how to develop a mental model that facilitates innovation in the context of digital transformation. Three themes were identified in their mental model change: The extent to which innovation requires involvement, the complexity of innovation processes and what outcomes can realistically be expected of a digital innovation process. The team went from a model of digital innovation as 'a mysterious black box' - something external, where they could hand in a list of requirements and walk away with a digital tool - to a 'super puppy' that can do remarkable things, but has to be trained and interacted with to get a desired effect.

I

Introduction



1 Introduction

This research investigates mental model change around digital technology and innovation, in combination with the development of a network-based scientist-journalist recommender system. It was initiated after a keynote speech by professor dr.ir. Deborah Nas at the C-Day conference of the Dutch national association for communication professionals Logeion on June 20, 2019[28]. She stated that in ten years, every communication professional will have an algorithm as a colleague. According to Nas, artificial intelligence technology will have far-reaching consequences for the communication profession. She warned that communication professionals that do not start adopting new technological tools today will be outdated in five years and outcompeted by those who do.

After this, the head of the corporate communication department of Delft University of Technology and the head of the communication department of Naturalis Biodiversity Center approached the Communication Design for Innovation (CDI) research group at Delft University of Technology. They asked the research group to investigate how artificial intelligence and digital innovation will influence their field, what they can and cannot expect from digital innovation and how they can take charge of their digital future. This presented interesting opportunities for research from both a Communication Design for Innovation perspective as well as a Computer Science perspective.

The questions of the communication departments can be broadly summarised as: *How can we prepare for a digital future?* However, this is not a question that an external party can answer for them. Instead, it requires them to change how they look at digital technology and innovation. In other words, they must change their *mental model* of digital innovation.

This is easier said than done. The problem of adopting a new mental model around digital innovation is by no means unique to these communication departments. Changing mental models is difficult in general. In particular, research on developing a mental model that facilitates innovation in the context of digital transformation is underdeveloped (De Paula et al., 2022)[14]. DePaula et al. propose design thinking as a method to develop this mental model. They research what attributes and behaviours associated with design thinking are effective in developing the mental model required to drive digital innovation based on interviews. Given that this research had access to two communication departments motivated to participate in this research, this project had the opportunity to address this gap in research through a case study. The case study involves a team of communication professionals from the communication departments of Delft University of Technology and Naturalis Biodiversity Institute, which participate in the design process of a digital tool.

An area of interest in Computer Science is the development of new methodological and algorithmic concepts to create recommender systems. The collaboration with the communication professionals provides an interesting use case to develop a novel recommender system. By having the communication professionals participate in the development process of this system, the effects of participation on the mental model of digital innovation of the communication professionals can be investigated.

1.1. Report structure

The research into the development of the recommender system is focused on the technical aspects of the design, while the research into mental model change focuses on the mental model of the communication professionals before, during and after the development of the system. For this reason, the report is split into several parts. First, the background, theoretical framework and methodology of the research into mental model change are laid out in the rest of part I.

Part II (page 16) describes the first part of the process. This is structured according to the double diamond process, which divides a design process into two cycles of diverging and converging. The first diverging phase is Discover, which is about gaining insight into the problem. It is followed by the Define phase, which focuses on creating a problem definition or design brief. The next diverging phase is the Develop phase, in which a solution is developed. Part II describes each phase and how the mental model of digital innovation of the communication professionals is affected.

Based on requirements gathered in part II, available data and research interest, the choice was made to develop a recommender system to propose collaborations between scientists and newspaper journalists based on the network of scientist-scientist collaborations, journalist-journalist collaborations and scientist-journalist collaborations. Part III (page 37) defines the technical problem that the system aims to solve, the data that is used to create the network, and the design and validation of the recommender system. This part is written as a standalone report.

Lastly, part IV (page 65) describes the final phase of the double diamond model, the Deliver phase. This phase converges on a working solution. This part focuses on the final steps in the development process. It discusses the effects of participation in the process on the mental model of digital innovation of the communication professionals. It also includes a general discussion of the mental model research and a personal reflection on the process.

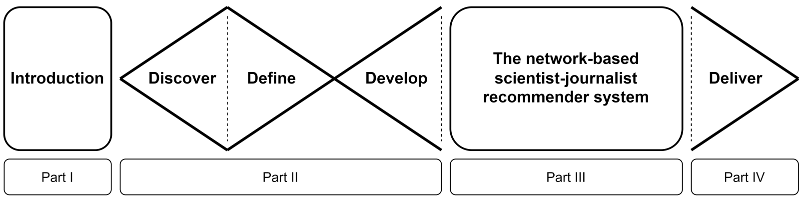


Figure 1.1: The structure of this report.

1.2. Background

This chapter discusses the background of this research. Section 1.2.1 discusses the role of communication professionals in research institutes, and section 1.2.2 the drivers of digital innovation in the communication field. Section 1.2.3 discusses the interest of communication professionals in artificial intelligence and the gap between their expectations and the technical reality that gave rise to this project.

1.2.1. The role of communication professionals in research institutes

Communication professionals at research institutes are tasked with getting news coverage for scientists and research from their institute. In today's news landscape, it is often difficult to get scientific experts in the news. Merkley (2020)[33] found that only 23% of news articles covering topics like climate change, nuclear safety and the economic effects of immigration have a message from experts in a relevant field that align with the scientific consensus and only 9% explicitly state it is the consensus.

The scientific outreach that is done mostly focuses on newspapers, as shown by a comparison of outreach methods done by Entradas and Bauer. (2017)[18]. Getting a scientist in the news starts with either a scientist working on a newsworthy research topic coming to the attention of a communication professional or a journalist reaching out to the communication department with a science-related question. In the first case, the communication professional often writes a press release, which is sent to various news sources. Alternatively, the communication professional can contact a journalist to connect with the scientist. In the second case, the communication professional will reach out to a scientist they know is related to that topic or ask a faculty or colleagues for a suitable scientist to match with the journalist.

1.2.2. Drivers of digital innovation in the communication field

Many influences in the communication field motivate the communication department to become more digitally innovative. Benhayoun et al. (2020)[7] mention five external factors driving an organisation to intensify its innovation efforts. All five of these aspects are applicable in the setting of this project.

Technological evolution
The developments of digital technologies are incredibly fast-paced, and in the field of communication, this is no different. New services are brought to the market frequently, and the multitude of largely overlapping services can be overwhelming, as a small selection of prominent tools demonstrates:

- Internal communication tools
  E.g. Slack, Mattermost, Teams, Discord, Hangouts, Skype, Skype for Business, Zoom
- Collaboration tools
  E.g. Trello, Basecamp, G-Suite, Tettra, Office 365, Dropbox, TOPdesk
- External communication tools
  Websites, Outlook, Gmail, Mailchimp, LinkedIn, Twitter, Facebook, Instagram, Tiktok, Youtube, blogs
- Analytics tools
  E.g. Google Analytics, SiteImprove, Clicky, LexisNexis

This list does not even include specialised software, often developed in-house by large or niche players.

Communication professionals traditionally lack digital expertise, but technological developments increasingly influence their field of work. As the heads of the communication departments noted, many of their colleagues acknowledge the need for digital expertise. They just have no idea how to obtain it (debrief meeting June 22, 2020).

High frequency of innovation
Technical evolution leads to innovation in the communication field. Organisational procedures change to adapt to new technologies. For example, social media allows organisations to directly interact with their users, making communication much more two-sided. It has given rise to a whole new profession within the communication field, the influencer. According to Wielki (2020)[63], Influencers are well-followed social media users that are trusted by other users and influence consumers' attitudes and decision-making towards brands and ideas. Influencer marketing has become the fastest-growing trend in terms of customer communication and has changed the functioning of entire industries. Some groups rely more on influencers than trusted sources such as news outlets and science, which can perpetuate flagrantly false notions. There is also a growing concern about the purposeful spreading of misinformation and the use of digital technologies in targeted manipulation. For example, in 2016, there was the story of Cambridge Analytica harvesting the personal data of 87 million Facebook users to micro-target American and British voters and swaying the 2016 US presidential elections. While the actual effectiveness of Cambridge Analytica in general and their use of digital tools, in particular, are still debated (Sumpter, 2019[52], Elish and Boyd, 2018[17], Laterzo, 2021[26]), the story around it has had a massive impact on the general perception of the dangers of digital technologies in the communication field. This is another factor that contributes to a sense of urgency when it comes to acquiring digital expertise.

Frequent changes in market trends
Few things are as capricious as the news. Different topics dominate from day to day, and yesterday's news is, as the saying goes, yesterday's news. Trending topics on social media often have an even shorter lifespan. While not all communication departments are directly involved in the broader public discourse, those that are, are ever at its mercy. Digital technologies are seen as both a cause

and a solution to this chaos, further encouraging communication professionals to acquire digital know-how.

High levels of competition
There has always been limited space in newspapers and television broadcasts. With the advent of the internet and the increasing connectedness of the world, the number of players vying for attention has grown. The internet has made it easier for everyone to publish content, which increases competition.

Frequent changes in regulation
The digital space is growing so fast that regulation struggles to keep up. The lack of technical knowledge among politicians further exacerbates this. This has led to a patchwork of regulations on privacy, copyright and intellectual property.

1.2.3. Expectations and reality of Artificial Intelligence

Artificial intelligence has captured the shared imagination since it was first coined in the 1950s. With the increasing use of machine learning techniques in the last decade, artificial intelligence has become even more popular and is constantly in the media (Zhai et al., 2020)[68].

Communication professionals are deeply embedded in the media ecosystem and as such, are exposed to a lot of news surrounding artificial intelligence, especially as they are connected to innovative research institutes. While the trigger for the communication professionals participating in this research to reach out to the Communication for Innovation faculty was the speech of Dr. Nas, the interest in the topic was already present.

> Within the communication profession, many people are consciously or unconsciously occupied with AI without actually knowing what it is. People feel they should do something with it, but don't know how. *Minute meetings June 22nd, 2020*

However, as (Elish and Boyd, 2018)[17] put it, "AI, as a category of technology, always waivers between the real and the imaginary" (p.62, Elish and Boyd, 2018). They quote Clarke's third law, which states, "Any sufficiently advanced technology is indistinguishable from magic" (p.21, Clarke, 1973)[10]. This leads to expectations that are disconnected from what is technically possible or realistically achievable from the perspective of time and resources.

> A defining feature of magic, as an orientating framework of actions and consequences in the world, is that it is 'costless' in terms of the kind of drudgery, hazards, and investments that actual technical activity inevitably requires. *p.63, Elish and Boyd, 2018[17]*

In the kick-off meeting of this project, the discussion went to what system should be developed. The initial request was for an artificial intelligence algorithm that would 1) predict what topics would become news, 2) compare it to the expertise within the institute and find what researchers within the institute would be most suitable to interview, 3) write an article draft on the topic based on the researchers' prior work and 4) identify a suitable journalist and news outlet to send the article to.

This request is telling regarding their perception of what artificial intelligence can do and what is involved in developing a new digital system. Some of these requests by themselves are already pushing the boundary of what is technically possible. Together, they are infeasible from a technical perspective and resource-wise, given this is a one-person thesis project. This implies that there is not simply a lack of knowledge but that they have a mental model of digital innovation that does not align with technical realities.

1.3. Research questions

This research investigates what happens to this mental model of digital innovation when they participate in an actual software development process. The proof is in the pudding: By actually developing a working system, they are confronted with the complexities and limitations of such a process. To investigate the effect of participation on mental models, this research attempts to answer three research questions:

1. What motivates the team of communication professionals to participate in this digital innovation process?
2. What is a suitable tool to develop with the team of communication professionals?
3. What is the effect of participation in the development process of a digital tool on the mental model of digital innovation?

2
Theoretical framework

This chapter contains a theoretical discussion of the theoretical concepts underpinning this question and this research.

To investigate the effect of participation in the development process of a digital tool on the mental model of digital innovation, section 2.1 defines what is meant by a *mental model of digital innovation*. It is important to understand how mental models can be assessed to identify if they have changed. Section 2.2 discusses this and introduces *stimulated recall* as a method that was found effective in analysing mental models.

One of the motivations for investigating mental model change through participation in a development process was the finding of DePaula et al.(2022)[14] that *design thinking* is effective in changing mental models regarding digital innovation. Section 2.3 analyses the concept of design thinking to provide a basis for selecting a software development paradigm to structure the development process of the digital tool. This paradigm, the *Agile* philosophy, is discussed in section 2.3.

2.1. A mental model of digital innovation

A mental model is an internal representation of an external system. Kessler et al.(2022)[24] define mental models as ”cognitive representations of an issue based on an individual’s experiences and knowledge” (Kessler et al., p.712). Richardson et al. (1994) [44] state that mental models need to contain three sub-models: an Ends model, a Means model and a Means/Ends model. An Ends model represents what the decision-maker is trying to accomplish. A Means model ”contains strategies, tactics and policy levers the decision maker believes are available or usable to move towards the perceived goals” (Richardson et al., p.4). A Means/Ends model represents the connection between means and ends.

The Ends model of digital innovation for the communication professionals is the space of results perceived as achievable through digital innovation. The Means model of communication professionals is the perceived space of actions they can take in a digital innovation process. The Means/Ends model is their mental representation of the innovation process and how their actions and those of others can produce different results.

Through this lens, the mental model of digital innovation of communication professionals can be defined as their mental representation of the innovation process, their relationship to it and the results it can bring.

2.2. Assessing mental models through stimulated recall

Various methods have been used to assess mental models on an individual and a group level. For example, Jonassen (1995)[23] describes analogical or metaphorical reasoning as a common approach to describe mental models, e.g. using the analogy of flowing water to help people develop a mental model of electricity. However, as Webbers et al. (2000)[61] note, there is no single dominant approach.

"Measures also vary in their ability to capture the content of mental models, the structure, or both." (p.309, Webber et al., 2000). They found this to be especially true for strategic mental models, which "are composed of information that provides the basis for problem-solving, including action plans to meet specific goals, knowledge of the context in which procedures should be implemented, actions to be taken if a proposed solution fails, and how to respond if necessary information is absent." (p.309, Webber et al., 2000). An innovation process relies on problem-solving skills

A method used by Henderson and Tallman (2006)[20] to affect and assess mental models is stimulated recall, which is described by Lyle (2003)[31] as "an introspection procedure in which (normally) video-taped passages of behaviour are replayed to individuals to stimulate recall of their concurrent cognitive activity" (p.861, Lyle, 2003). Henderson and Tallman (2006) describe ten case studies using stimulated recall to assess mental models related to teaching and learning computer literacy. They analyse espoused (before), in-action (during) and reflective (after) mental models and argue that stimulated recall is effective as a methodology for understanding complex cognitive structures.

2.3. Affecting mental models through design thinking

A study on the mental model required to drive digital innovation by De Paula et al. (2022)[14] discusses how hard it is to change existing mental models. They note how design thinking has become popular as an approach to digital innovation in recent years. They find that "design thinking plays an important role in achieving a more human-centric digital transformation due to its ability to support stakeholders engaging in learning loops during the design process while also ensuring user-centredness" (De Paula et al., p.2). They found that design thinking is effective in changing the mental model of leadership around digital transformation innovations.

Design thinking can also help cope with the complexity and uncertainty of every innovation process. Van der Sanden and Wehrmann (2021)[56] found that design thinking helps to cut through complexity, helps students and professionals to recognise complex issues and cope with the constant feeling of uncertainty that comes from this complexity. Mosely et al. (2018)[36] describe how design thinking can help develop knowledge and mindsets to solve complex and ill-defined problems.

Participation in digital development can engage the communication professionals in design thinking around digital technology. This should lead to the development of knowledge of - and problem-solving skills around - digital innovation, which can be captured by a change in their mental model of digital innovation.

2.4. Participation in the development process

One of the leading software development paradigms is the Agile philosophy. Agile is a movement against traditional 'waterfall' development methods. The waterfall model is a structured, top-down approach where development is broken down into sequential design, development and testing phases. The software is designed in its entirety, developed in its entirety and tested in its entirety (Bahli and Zeid, 2005)[5]. The criticism of the waterfall approach is that it is a top-down approach that requires users to provide a complete set of requirements at the beginning of the process, and the development team needs to fully digest and interpret these requirements because this approach does not deal well with changes. Conversely, Agile focuses on interdisciplinary teams that go through short cycles (sprints) of design, development and testing in an iterative process. The goal is to get to a prototype as quickly as possible, validate if it suits the need of the user and iterate from there.

Given that the communication professionals are new to the field of software development and that their explicit goal is to discover what is possible, it is unreasonable to expect them to be able to formulate a complete set of requirements up-front. An Agile approach focusing on incremental change allows them to explore the possibilities one step at a time. An Agile approach also fits better with a design-based approach, as design-based approaches focus on iterations and prototyping.

One of the most commonly used methodologies within the Agile framework is Scrum. Dyba et al. (2008)[16] classify it as an Agile methodology focusing on project management in uncertain situations. It offers a transparent process with clearly defined roles, which can be used to structure the development process and help determine how the communication professionals will participate in the development process.

The official Scrum Guide explains the process:

> 1. A Product Owner orders the work for a complex problem into a Product Backlog.
> 2. The Scrum Team turns a selection of the work into an Increment of value during a Sprint.
> 3. The Scrum Team and its stakeholders inspect the results and adjust for the next sprint.
> 4. *Repeat*
>
> *The Scrum Guide, Schwaber and Sutherland, 2020 [46]*

The roles in Scrum are Developers, Product Owners and Scrum Masters. The Developers are the members of the Scrum Team, which can consist of programmers, designers, architects or other specialists. The Scrum Master is the facilitator who structures and streamlines the Scrum process. The Product Owner represents the user's interests. The Product Owner is responsible for developing the product goal, breaking it down into backlog items (product features) and prioritising which features should be made in the next sprint.

The communication professionals will take the role of Product Owner, while I take the role of both Scrum Master and Developer. In the role of Product Owner, the communication professionals are required to develop a vision and translate it to executable software design. However, it does not require any direct technical skills. During sprint reviews, they will assess the progress and decide on the work for the next sprint.

3
Methodology

This research is a case study in which a team of communication professionals from Delft University of Technology and Naturalis Biodiversity Center participate in facilitated sessions throughout the development process of a digital tool. Section 3.1 describes the case study setting, the case study design and the data collection methods. Section 3.2 describes how the data is analysed to answer the research questions.

3.1. Research Design and Methods

Section 3.1.1 elaborates on the case study setting. Section 3.1.2 describes how the case study was set up in a design-based approach with different interventions for each phase of the Double Diamond model. During the Discover phase of the development process, research question 1 about the motivations of communication professionals is answered through a brainstorming session on project goals. In the Define phase, research question 2 on a suitable tool is answered through a brainstorming session in which tool requirements are set up, after which a tool is chosen. Section 3.1.3 describes how data is collected through note-taking and stimulated recall sessions where the communication professionals reflect on their mental process and the whole process to evaluate the effect of the process on their mental model of digital innovation.

3.1.1. Case study setting

This section elaborates on the communication departments , the participants and the drivers to participate in a digital development process.

The communication departments and their respective institutes

The case study was conducted with participants from the corporate communication department of Delft University of Technology and the communication department of Naturalis Biodiversity Center. This research was initiated because the heads of the departments joined forces and reached out to the CDI faculty. Therefore this research had to operate within the boundaries set by the communication departments, and the participants were predetermined.

Delft University of Technology is the oldest engineering university in the Netherlands. It has 6980 researchers[1] divided across eight faculties. The entire communication department of Delft University of Technology consists of more than 160 communication professionals. Its responsibilities are both internal (e.g. correspondence to students) and external (e.g. correspondence with journalists). The corporate communication department is responsible for getting news coverage for the university.

Naturalis Biodiversity Center is a natural history museum and research institute in the Netherlands. In addition, it hosts one of the world's largest collections of natural history specimens. The research

[1]This number is according to an employee list provided by Delft University of Technology dated September 2020. For more information, see section 9.1.3 on data sources.

institute consists of 112 scientists[2]. The communication department of Naturalis is responsible for getting exposure both on the museum and the research sides. The communication department consists of 11 communication professionals.

The participants

For both institutes, one participant was in a leadership position and one participant was from the operating core (see table 3.1). While both people in leadership positions did not have any technical background, both participants from the operating core did.

Code no.	Position	Institute
1.	Leadership	Delft University of Technology
2.	Operating core	Delft University of Technology
3.	Leadership	Naturalis Biodiversity Center
4.	Operating core	Naturalis Biodiversity Center

Table 3.1: Participants in this study

3.1.2. Case study design

The development process of a digital tool can be understood as a design process. A "universally accepted depiction of the design process" (p.7, Andrews et al. (2021))[3] is the Double Diamond model. The UK Design Council (2007)[55] developed the double diamond model as a graphical way to describe the design process (see figure 3.1).

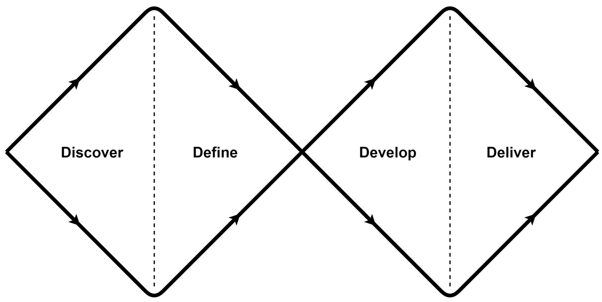


Figure 3.1: The Double Diamond model by the UK Design Council. (2019)[54].

It divides the design process from problem to solution into two cycles of diverging and converging, consisting of a total of four phases: Discover, Define, Develop and Deliver. The Discover phase is about gaining insight into the problem. It leads to the Define phase, in which a problem definition or design brief is created. Based on the design brief, potential solutions are created in the Develop phase. The Deliver phase converges on a working solution.

For each phase of the Double Diamond model, participatory sessions were organised. Table 3.2 shows the goal in each development process step. The sessions consisted of both process updates and brainstorming sessions. The format of each session was iteratively chosen based on the project status and the goal of the project phase.

Phase	Goal
Discover	Explore motivations
Define	Decide what tool to develop
Develop	Develop scientist-journalist recommender system
Deliver	Transform prototype into usable product

Table 3.2: An overview of the goal of each phase of the double diamond.

[2]This number is according to an employee list provided by Naturalis Biodiversity Center dated June 2019. For more information, see section 9.1.3 on data sources.

Given the iterative nature of this process, each phase of the double diamond has been given a separate chapter. Chapter 4 describes the Discover phase. In this phase, research question 1 is answered: What motivates the team of communication professionals to participate in this digital innovation process? Chapter 5 describes the Define phase. The goal is to decide what tool to develop and decide on an approach for the development process. In this phase, research question 2 on a suitable tool to create with the team of communication professionals is answered. Chapter 6 describes the process of the Develop phase. The technical description of the development phase is explained in part III, which is written as a standalone report on the Computer Science part of this research.

3.1.3. Data collection

The data collection for this research was done through note-taking during the process and stimulated recall sessions.

Notes during the process

After all meetings during the Discover and Define phase, debriefs were made and distributed to all participants. The manner of note-taking differed during the Develop phase depending on the nature of the meeting and the content discussed.

The first two Develop phase meetings were brainstorming sessions; for these sessions, debriefs were made and distributed. The third meeting was an update session, and no notes were taken. The fourth session was a progress session in which the participants recapped what the original goals were, where they stood at that moment, what they had learned so far and what they wanted from the final part of the process. This was an interactive session using Miro[3] and notes resulting from the session were captured in the Miro board.

Stimulated Recall

An established method for analysing mental models is stimulated recall. Lyle (2003)[31] describes it as "an introspection procedure in which (normally) videotaped passages of behaviour are replayed to individuals to stimulate recall of their concurrent cognitive activity" (p.861, Lyle, 2003). Henderson and Tallman (2006)[20] discuss ten case studies of using stimulated recall to assess mental models related to teaching and learning computer literacy. They analyse espoused (before), in-action (during) and reflective (after) mental models and argue that stimulated recall is effective as a methodology for understanding complex cognitive structures.

Given that the method of mental model change was participation in the development process, stimulated recall was deemed an effective way to get the participants to reflect on the thinking behind their actions and unearth the more profound mental representation underlying them.

Stimulated recall in this research was done by recording the final co-design sessions with the communication departments. Due to a last-minute cancellation of one of the communication departments, the initially planned co-design session was done with only one communication department and a separate co-design session was done with the participants from the other department. Because of this, the stimulated recall was also done individually with each department.

A longlist of fragments was made by watching the recording, after which a shortlist was made of fragments to use. The main selection criteria for fragments was a show of emotions of some kind or something that allowed the steering of the conversation towards underlying feelings about the process. The goal was to establish a psychologically safe atmosphere where the process could be examined freely and openly. Table 3.3 shows the final selection of fragments. The stimulated recall sessions were unstructured. Each fragment was chosen with a specific interest in mind. The conversation was steered towards that topic, but the priority was to keep the conversation flowing as naturally as possible.

The sessions were recorded and transcribed verbatim. The transcripts were timestamped every 2 minutes to allow quotes to refer to specific timestamps. The meetings - and therefore transcripts - were in Dutch. Quotes used in this report are translated by Google Translate and fine-tuned by the author. When sayings are used, an effort is made to find the most similar saying in English rather than translate it literally. Each quote will refer to the timestamp at the beginning of the 2-minute interval.

[3]Miro is an online whiteboarding tool. See `https://miro.com/`

Fragment	Interest	Dept.
Facilitator introduces session goal, participant grabs a note block.	What were their expectations going into this session? What triggered the desire to get a note block?	DUT
A compilation of participants asking questions during the presentation of the prototype. One of the question is introduced as a 'stupid' question.	What triggered the engagement? Why do they refer to the question as 'stupid'?	DUT
Participant is discussing when they would trust the system and refers to themselves as 'control freaky'.	Why do they refer to themselves as control freaky? When do they feel trust towards a tool or process?	DUT
Fragment 1: participant says they will refer to the algorithm as a 'she'. Fragment 2: participant says they need to think of a name and other participant immediately says 'Eliza'.	They are personifying the tool. What makes them feel connected to the tool and the process?	DUT
Participants are calling names of scientists to test the prototype. The name of their most well-known scientist is not recognised.	What was their reaction to the scientist not being recognized? What are their expectations of the tool?	NBC
Participant is discussing when they would trust the system and is showing signs of being thoughtful.	What is their evaluation process? What makes them hesitate?	NBC
Participant is discussing next steps. They start in an uncertain tone of voice and become increasingly confident.	What is their thought process and when do they feel confident?	NBC
Facilitator has connection issues. Participants are talking among themselves that we're already short on time because they need to leave soon. They discuss that they feel bad for never having enough time for this project.	Why are they struggling to find time for this project? What are factors influencing their prioritization?	NBC
Participant is discussing how enthusiastic they are that this has been made 'specially for them'	What are the things in the process and the tool they are enthusiastic about?	NBC

Table 3.3: The fragments chosen for the stimulated recall sessions.

3.2. Data analysis

This section describes how the stimulated recall data is analysed to answer research question 3 on the effect of participation in the development process of a digital tool on the mental model of digital innovation.

3.2.1. Coding

This research aims to determine the effect of participation in a digital development process on the mental model of digital technology. The stimulated recall transcripts were analysed using Structural Coding to answer this question. Structural Coding is a deductive coding approach in which predefined codes based on a specific research question are used (Saldana, 2021)[45]. There are several benefits to using Structural Coding.

> Structural Coding is appropriate for virtually all qualitative studies, but particularly for those employing multiple participants, standardised or semi-structured data-gathering protocols, hypothesis testing, or exploratory investigations to gather topics lists or indexes of major categories or themes. *p.84, Saldana (2021)[45]*

Given that the data consists of transcripts of two multiple-participant, semi-structured stimulated recall sessions and the open-ended research question, this coding method was deemed appropriate.

The focus is on the effect of the process, i.e. the difference between the mental model at the start and the end. Therefore, the codes are based on time frames, similar to Henderson and Tallman (2006)[20].

- Before
  Any block of text referring to a time frame before or at the very start of the development process
- During
  Any block of text referring to a time frame during the development process
- After
  Any block of text referring to a time frame during the development process

3.2.2. Thematic analysis

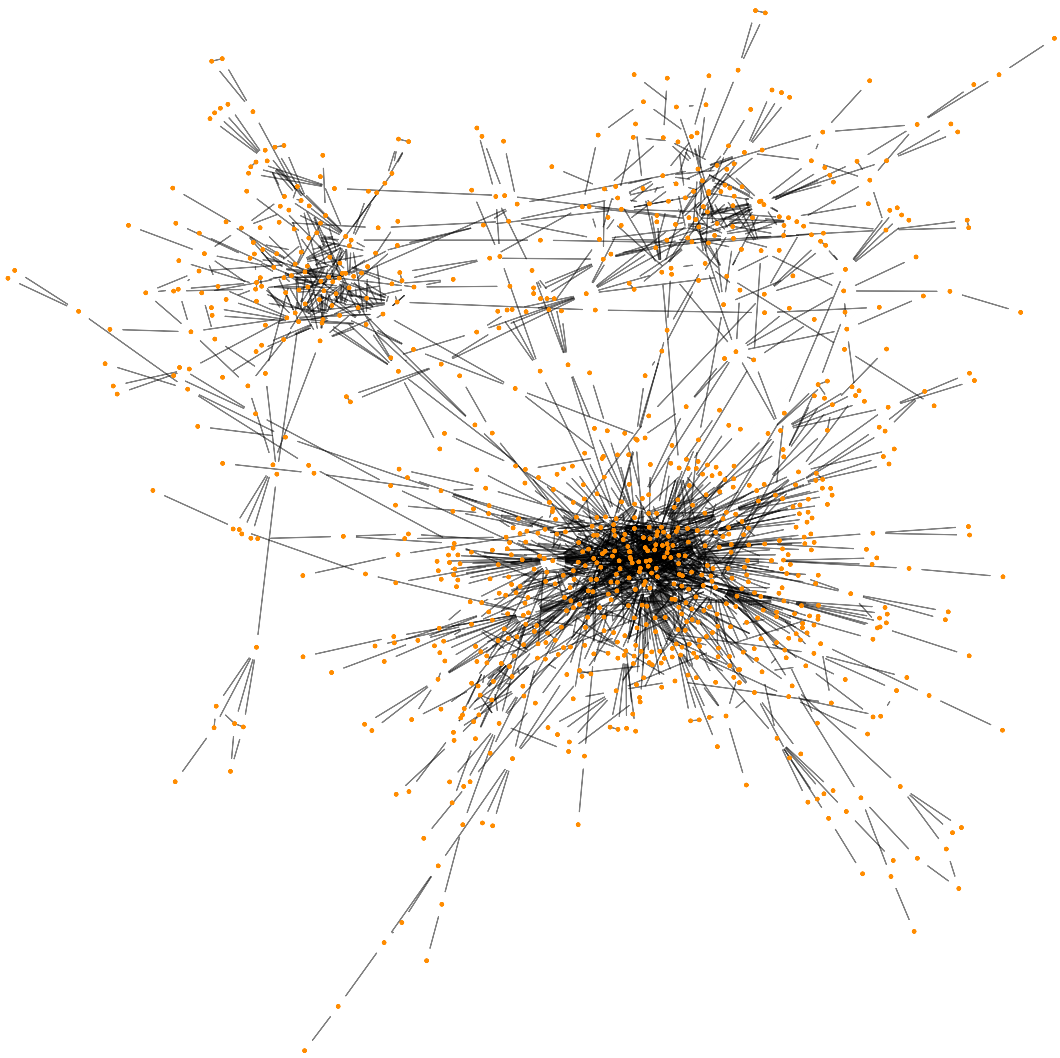
As described in section 2.1, the mental model of digital innovation of communication professionals can be defined as their mental representation of the innovation process, their relationship to it and the results it can bring. Two cycles of thematic analysis were done to understand the effect of participation in the development process of a digital tool on the mental model of digital innovation.

The first cycle of thematic analysis consisted of a mindmap in which the coded excerpts were organised around the nodes *process*, *relationship* and *outcome*. For each code, a different colour ink was used, to visualise the change in mental model around each node.

The second cycle of thematic analysis consisted of a mindmap in which for each aspect of mental models - process, relationship and outcome - a node summarising ’before’ was drawn on the left and a node summarising ’after’ was drawn on the right. Around the nodes, keyphrases from the first mind map or from excerpts were written that support the node. In the middle, a phrase representing the change was written in an arrow connecting the ’before’ and ’after’ states. These changes are used to answer the main research question.

II

Participation in a development process Discover, Define & Develop



4
Discover

The goal of the Discover phase was to answer research question 1: *What motivates the team of communication professionals to participate in this digital innovation process?* This phase consisted of a kick-off meeting and a brainstorming session. Section 4.1 discusses the results of the kick-off, and section 4.2 the results of the brainstorming session. Section 4.3 summarises these results and answers the research question.

4.1. The kick-off meeting

The Discover phase started at the kick-off meeting with the communication departments. The initial meeting was held on December 3rd 2019, at Naturalis Biodiversity Center with all communication professionals present, including Dr. Van der Sanden and this author. The debrief of the meeting was sent to all participants afterwards.

4.1.1. Results of kick-off meeting

The meeting notes of the meeting state the goal of the project:

> The aim of this research is to investigate the possibilities of applying computer science within the communication field by making a proof-of-concept of an algorithm. Not only does this algorithm provide a tangible idea of what is possible, but the process itself should reveal the needs, interests and resistances. *Debrief meeting December 3rd, 2019*

However, the focus of the communication professionals was on the functionality of the tool to be developed. During the meeting, tension arose between the communication professionals and this author about the project goals, what functionality could realistically be developed given the current state of technological advancements and the timeframe of a master thesis.

> I feel like everyone came to the table with the idea that the question was clear and we only needed to discuss the final details. During the conversation, however, it became clear that ideas diverged.
> ...
> The idea from [P1] and [P3] was an algorithm that links current events to the knowledge of the institute. This tool should identify what the news is from the various social media and news sites and whether the institute has expertise in this. This should help to quickly analyse what is going on and inform the institutes if there are opportunities for them to respond to this. *Debrief meeting December 3rd, 2019*

The outcome of the meeting was the decision to have a brainstorming session on project goals.

> Now that it has turned out that we are not all on the same page, we need to go back to the drawing board. That is why we are going to hold a brainstorming session in January about an approach that brings our interests and needs together. This gives everyone time to chew on the different ideas and further refine their needs. This is also the first phase in

the design research into a complex challenge, in which the beginning is always to arrive at a shared problem definition together. *Debrief meeting December 3rd, 2019*

4.1.2. Discussion of kick-off meeting results

Given the tension in the meeting and this author's feelings on how realistic their request was, the debrief was written to be as neutral as possible on what functionality they requested. To the best recollection of this author, their initial request was for an artificial intelligence algorithm that would 1) predict what topics would become news, 2) compare it to the expertise within the institute and find what researchers within the institute would be most suitable to interview, 3) write an article draft on the topic based on the researchers' prior work and 4) identify a suitable journalist and news outlet to send the article to.

The realism of the requested functionality

There are several issues with the requested functionality. The most difficult aspect of their request is news forecasting. Forecasting algorithms can be used to extrapolate from historical data. This works well for areas in which change is evolutionary, i.e. a new state is dependent on the old state in a (semi-)predictable manner. For example, today's weather depends on yesterday's weather and factors like wind, temperatures and pressure areas. While weather patterns are inherently chaotic, which means that any imperfection in initial measurements will inevitably grow to dominate the outcome (Lorenz, 2005)[29], they are predictable in the short term. Another example of an area in which forecasting algorithms are common is inventory management. A business's inventory depends on its inventory from yesterday, orders already in the system and general customer buying behaviour.

However, areas with revolutionary change - i.e. the new state does not depend on the previous state in a predictable manner - are inherently unpredictable. The goal of news is to be new. There are topics that dominate the news for a period, which means that the coverage of it today depends on the coverage yesterday. However, the next big topic does not follow from the previous topic. For example, topics like the Fukushima nuclear disaster[1] or the passing of the queen Elizabeth II[2] bear no relationship to other news topics, but will dominate the news when they happen. This makes them inherently impossible to predict.

While there is very interesting research being done on trending topics on social media and how information spreads (e.g. Weng et al., 2013[62]) and there are some short-term prediction models, to the best of this author's knowledge, there has been no research on predicting news with a timespan that would be interesting to the communication professionals.

The other aspects of their request are more manageable. Creating a press release draft based on a scientist's research would be a novel and interesting research subject. Topic-focused multi-document summarisation is an area of study within Natural Language Processing. Most techniques currently used do not create a new summary in the manner a human might but instead use NLP techniques to select specific sentences from the documents and put them together. A system like this that selects sentences from scientific papers might not result in a draft usable for a press release but might nevertheless be helpful to communication professionals. Such a system would easily be big enough to be an entire thesis project.

Their request to rate the expertise within their institute on a topic and find what researchers within the institute would be most suitable to interview boils down to an internal search engine based on scientific topics. The SciVal dataset containing metadata of scientific articles that was eventually used in the development process, as described in section 9.1, could be the basis of such a system. Their request to find a suitable journalist and news outlet similarly boils down to a topic-based search engine. A dataset like the DPGmedia dataset described in section 9.1, which contains different news articles, could be used to search which journalist is most related to a topic. Both these requests are technically possible. The system that is developed in this research uses the Scival dataset and DPGMedia dataset to find scientist-journalist matches, which was inspired by these requested functionalities.

[1]In 2011, the nuclear reactor in Fukushima, Japan was damaged by a tsunami. For more information, see `https://en.wikipedia.org/wiki/Fukushima_nuclear_disaster`.

[2]Queen Elizabeth II died on September 8th, 2022. For more information, see `https://www.theguardian.com/uk-news/2022/sep/08/queen-elizabeth-ii-britains-longest-reigning-monarch-dies-aged-96`

The implications in regards to their mental model of digital innovation
Their request is telling in regard to their perception of what artificial intelligence can do and what is involved in developing a new digital system. The requested functionality was motivated by the challenges they face and the desire to have a system that would help with these without considering what is technically possible or what is reasonable to expect, given the time and resources. During a reflection on their initial mental model, a participant referred to it as being in 'innovation land' (see section 14.2.4).

> We were in innovation land, and we went all out. We would like this, would like that, algorithm as a colleague, job description, all without any kind of obligations to it and unencumbered by any kind of knowledge. *P1, Stimulated Recall Delft University of Technology [39:00]*

This aligns with the commentary of Elish and Boyd (2018)[17] that "AI, as a category of technology, always waivers between the real and the imaginary" (p.62, Elish and Boyd, 2018). They quote Clarke's third law, which states, "Any sufficiently advanced technology is indistinguishable from magic" (p.21, Clarke, 1973)[10]. This leads to expectations that are disconnected from what is technically possible or realistically achievable from the perspective of time and resources.

> A defining feature of magic, as an orientating framework of actions and consequences in the world, is that it is 'costless' in terms of the kind of drudgery, hazards, and investments that actual technical activity inevitably requires. *p.63, Elish and Boyd, 2018[17]*

4.2. The brainstorm session on project goals

A brainstorming session was planned on the goals of this project and the desired outcomes. Appendix A shows this brainstorm session's (Dutch) preparation. The brainstorming session was developed to move the focus away from tool functionality and towards underlying motivations. It consisted of an icebreaker, the storylines exercise described in section 4.2.1 about personal motivations and a brainstorming session on shared goals.

Due to illness, the communication professionals of Naturalis Biodiversity Center could not be present, so the session was held with only the communication professionals of Delft University of Technology. The session was done live, with the participation of P1 and P2, as well as the author and Dr. Van Der Sanden. It was decided that this author would follow P2 at the corporate communication department for two weeks to gain an understanding of the daily practice of communication professionals.

A separate brainstorming session was scheduled with Naturalis Biodiversity Center. Due to the outbreak of the COVID pandemic and the lockdown in March 2020, no further live meetings were possible. For this reason, the brainstorming session was held using Google Meet. Instead of observing the daily practice of the communication professionals of Naturalis Biodiversity Center, two unstructured interviews were held with the communication professionals of Naturalis Biodiversity Center responsible for scientific news coverage.

4.2.1. The storyline exercise

A new exercise, called 'Storylines', was developed for this brainstorming session. In this exercise, participants create a causal loop diagram of what they enjoy in their work, what they are good at and what they still want to learn. The storyline exercise aims to start a conversation about personal desires, goals and barriers.

Every participant has a sheet of paper. The storyline exercise has four rounds. During the first round, participants write down things they enjoy in their job in the top left corner of the paper. After each round, there is a conversation where participants share something about what they've written down. In the second round, participants write down things they are good at. In the third round, they write down things they still want to learn. In the final round, participants connect aspects from each area to aspects from other areas when they support or clash. After the final round, participants are asked to create a 'storyline', a path through the diagram that connects something from each area.

4.2.2. Results of brainstorm session on project goals

During the brainstorming session on project goals, the communication professionals expressed being motivated by several different factors: a desire to learn and to have an impact on their colleagues and the communication field at large.

The motivations are taken from the debrief of the brainstorming session on the project goals, supported by the debrief of the kick-off meeting and the debrief of the brainstorming session on tool requirements.

The motivation to learn

In the debrief of the first meeting, the overall goal is phrased as a question about what a digital future will look like.

> This project started with the idea that in ten years, everyone will have an algorithm as a colleague. The communication departments of Naturalis Biodiversity Center and Delft University of Technology have now asked themselves what this colleague will look like and what influence this will have on the profession of communication professional.
>
> *Debrief meeting December 3rd, 2019*

After brainstorming on project goals, this overall goal was divided into three sub-goals: Get inspired by AI's possibilities, reflect on their current way of working, and learn what is and isn't possible with artificial intelligence and digital technologies.

> We want to be inspired and surprised by the possibilities that this new and exciting technology (AI) brings.
>
> ...
>
> Developing the algorithm will help us to reflect on current practices within the communication departments. After all, if we want the algorithm to support us in our choices, we will first have to clarify which choices we make and why.
>
> ...
>
> We would like it if this process gives us some insight into the field of development, which leads to concrete insight into what is and is not possible with AI and with digital solutions in general. *Debrief meeting February 17th, 2020*

The motivation to have an impact on colleagues

The communication professionals are motivated by the thought of being a driver of innovation within their field. They describe their field as conservative. Many communication professionals are interested in artificial intelligence but are also scared of it. Because of their lack of knowledge, many choose to refrain from engaging with it. They are motivated by being an inspiration to their colleagues. From the beginning, they mentioned that it would be great if this project could lead to a talk at a C-Day conference[3].

> We like the idea of driving innovation in the (often conservative) communication landscape with this research. *Debrief meeting February 17th, 2020*

> Within the communication profession, many people are consciously or unconsciously thinking about AI without knowing what it is [...] It can feel like a huge mountain for communication professionals and this leads people to avoid it. We would like to be an inspiration point and show that it is not scary, but can be fun once it is made accessible.
>
> *Debrief meeting June 22th, 2020*

4.2.3. Discussion of brainstorm session on project goals

There were two participants who were in separate sessions but had an almost identical storyline. While the exact storylines were not recorded, to the best of this author's recollection they contained the following aspects. They mentioned enjoying diversity in their work, working with many different people and doing many different things. As strengths, they mentioned being result oriented, good at getting

[3]C-Day is the annual conference of the national association of communication professionals Logeion. The inspiration of the communication departments to start with this project was a keynote speech by Deborah Nas at a C-Day conference.

people into motion and getting things done. Both mentioned they wanted to learn a bit more patience and perseverance when things go slow.

The combination of being result-oriented and people-oriented while struggling with slow processes is directly reflected in their motivations to participate in this process. While the original impetus for this process was a desire to learn and discover, this is a slow process with uncertain results. While the communication professionals still mention it as an important goal, they see it through the lens of how it can affect their relationships and interactions with colleagues.

The result-oriented mindset shows in the focus on the digital tool. It was not a direct outcome of the brainstorming session, as it was designed to avoid discussing the actual tool functionality and instead focus on underlying motivations. However, the motivation for the final tool described in section 4.1.2 was still present and continued to be present during the process. Even in the final reflection (Stimulated Recall Delft University of Technology [43:00]), one of the participants refers to their learning process as a 'side effect' of the 'outcome', i.e. the delivery of the tool.

4.3. The motivations of the communication professionals to participate in this digital innovation process

The motivations of the team of communication professionals to participate in this development process were threefold:

1. A desire to learn about the possibilities artificial intelligence can bring, how it can help them reflect on their current way of working, and what is and isn't possible with artificial intelligence and digital technologies.
2. A desire to be a driver of innovation in the communication field and have an impact on colleagues and the field at large.
3. A desire to receive a custom-made digital tool.

5

Define

The goal of the Define phase was to answer research question 2: *What is a suitable tool to use as the subject of the development process?* To answer this question, a brainstorming session was held to choose a tool to develop. The second goal of the Define phase was to create a methodology of how to approach the Develop phase.

5.1. The brainstorming session on the digital tool to be developed

The brainstorming session was held on June 22nd 2020, using Zoom. It consisted of two parts. The first part was a brainstorming session on tool requirements. These requirements were split into functional requirements - i.e. requirements on *what* the tool needs to do - and technical requirements - i.e. requirements on *how* the tool needs to do this (Miedema et al., 2007)[34]. The second part of the brainstorming session was to decide on what tool would be the subject of the development process. The data in this section is taken from the debrief of this brainstorming session.

5.1.1. Requirements

The requirements of the tool were separated into functional requirements and technical requirements. The requirements were derived from the interests of the communication professionals, the interests of the CDI group, and the constraints given this is an integrated thesis with Computer Science. The requirements are direct quotes from the debrief of the brainstorming session.

Functional requirements

According to Miedema et al. (2007)[34], ”Functional specifications provide a description of desired future product behaviour” (p.238). They are requirements on the functionality of the tool, i.e. *what* the tool needs to do. Four functional requirements were decided upon.

1. It must be useful for both Delft University of Technology and Naturalis Biodiversity Center

Given that the project was initiated by both the communication departments of Delft University of Technology and Naturalis Biodiversity Center, the tool needs to be helpful to both. Delft University of Technology has two primary tasks: research and education. The Corporate Communication department focuses on external communication, mainly getting news coverage of the university's scientists. Naturalis Biodiversity Center has three areas of focus: research, the natural history museum and the natural history collection. The communication department is responsible for promoting all three aspects, as well as internal communication. The overlap is in news coverage of research.

2. It provides and uses information about the social network

It was decided to focus on social networks. The initial request of the communication departments was focused on social media, in which social network analysis is a common research approach. While the focus on social media was dropped, the idea of social network analysis remained appealing to the communication professionals. Communication is a social process; a large part of their job is connecting people. One of their goals was to reflect on their practices, and in this regard, the communication professionals were highly interested in getting more insight into their social network.

3. It promotes collaboration
Dr. Van der Sanden of the CDI research group was one of the brainstorming participants. Dr. Van der Sanden does research into the development of decision support tools to improve collaboration in innovation. Given both his expertise and interest in this area, it was decided to make this one of the requirements of the digital tool.

4. It provides new insights
The tool should assist with more than just current activities or automate part of the current process. It should expand their way of working and provide insights into the possibilities of technology and their current processes.

Technical requirements
Technical requirements are requirements on the technology and the process, i.e. *how* the tool needs to do it or to be developed. Some technical requirements were decided upon with the communication professionals, and some were given based on other constraints.

Three technical requirements were decided upon.

1. The algorithm is data-driven
One of the learning goals of the communication professionals was to gain an understanding of artificial intelligence. Artificial intelligence is a very broad term, with multiple definitions that have morphed over the years (Bini, 2018)[8] . During the brainstorming session, it was specified that the algorithm should be data-driven.

2. It is based on network analysis
The Computer Science part of this thesis focuses on network analysis under the supervision of Dr. Wang from the Multimedia Computing research group. Dr. Wang does research into data-driven modelling of dynamic processes in complex networks. This meant that network analysis was given as the basis for the tool.

3. The data must be easily accessible
Data collection is often a time-intensive process. Given time constraints, it was decided that it was not possible to spend a lot of time gathering data and only use accessible data sources.

5.1.2. The digital tool

It was decided to design a system that helps communication professionals to get scientists from their institute in the news. In today's news landscape, it is often difficult to get scientific experts in the news. Merkley (2020)[33] found that only 23% of news articles covering topics like climate change, nuclear safety and the economic effects of immigration have a message from experts in a relevant field that align with the scientific consensus and only 9% explicitly state it is the consensus. The scientific outreach that is done mostly focuses on newspapers, as shown by a comparison of outreach methods done by Entradas et al. (2017)[18]. Getting a scientist in the news starts with either a scientist working on a newsworthy research topic coming to the attention of a communication professional or a journalist reaching out to the communication department with a science-related question. The job of communication professionals is to find a suitable journalist or scientist and bring the two into contact. The matching process is time intensive and often has to happen on short notice. Communication professionals keep a mental list of scientists and journalists they know. Often they do not have time to consider other candidates.

It was decided to develop a network-based scientist-journalist recommender system. The system connects scientists from the institutes to newspaper journalists and vice versa. The recommender system helps with the selection process. The system can quickly screen many potential matches and give a pre-selection of likely scientists/journalists.

> The algorithm becomes a kind of 'communication tinder', where you can 1) enter a science per subject, and the algorithm comes up with a list of possible journalists, or 2) you can enter a journalist and a subject, and the algorithm comes up with a list from possible scientists.

> This can be controlled in various smart ways so that the algorithm looks beyond 'the usual suspects' and can be used tactically to make a smart match. *Debrief meeting June 22nd, 2020*

It is built upon a scientist-journalist collaboration network. The network consists of scientist nodes and journalist nodes. Two scientists are connected if they co-authored a scientific paper. Two journalists are connected if they have co-authored a newspaper article. A scientist and a journalist are connected if the scientist is mentioned in a newspaper article written by the journalist. The recommender system uses network analysis to recommend new connections between scientists and journalists.

Both institutions have connections to newspaper journalists but only oversee some relationships. They are interested in gaining greater insight into how the network functions to improve collaboration with newspapers. Furthermore, it is based on scientific articles, which the institutions have access to, and newspaper articles, of which a dataset is available through DPG media.

The creation of the scientist-journalist collaboration network, the analysis of network patterns and the design of the recommender system is described in part III.

5.1.3. Engaging colleagues in the communication field

One of the factors motivating the communication professionals was to have an impact on their colleagues and the communication field as a whole (see 4.2.2). During the brainstorming, it was stated as an explicit goal to involve other communication professionals.

> We want to show colleagues the value of the algorithm and raise that 'puppy' together. We want to make them part of our journey so that they don't suddenly get an algorithm at the end but feel it is from all of us. We want to create an awareness of the (im)possibilities of AI by sharing what the things are that we encounter, what choices we make, what the ideas are and what we end up with. *Debrief meeting June 22nd, 2020*

It was decided to start a blog. On the one hand, the purpose was to inform other communication professionals and engage them in the process. On the other hand, it provides additional information to the team of communication professionals on technical subjects.

> To achieve this, with [P3] 's help, I will start a blog during the development process. By regularly writing small, manageable posts about the themes and issues I encounter in the development process, I take the readers step by step in climbing the emotional 'artificial intelligence mountain'. An interactive format gives them the feeling of being involved in the process and makes the algorithm actually feel theirs. *Debrief meeting June 22nd, 2020*

5.2. The Develop approach

The initial strategy for the Develop phase was to have the team of communication professionals participate in the development process by co-designing the scientist-journalist recommender system. For this purpose, the Electron Learning Model was developed. The Electron Learning Model is a theoretical model of how co-design can affect mental model change.

Section 5.2.1 introduces co-design. Section 5.2.2 starts with mental model change and links it to *double loop learning*. Learning can be operationalised through *absorptive capacity*, which is defined as the capacity to learn. In section 5.2.3, the Electron Learning Model is introduced, which describes the interplay between co-design and absorptive capacity.

5.2.1. Participation in the development process through co-design

As described in section 2.3 and 2.4, participation in the development process was structured using the Agile approach in order to stimulate design thinking in the communication professionals. Agile's focuses on interdisciplinary iterations to create and validate software connects well with the iterative nature of design, and design thinking is linked to mental model change.

This process can be seen through the lens of *co-design*. Co-design is a philosophy that is focused on stakeholder participation in design processes. It "covers theories and practices that emphasise the role of the end-users as full participants in the design process." (p.6, David et al., 2013)[13]. It is characterised by a user-centric, interdisciplinary, iterative investigation process and an outcome focused on usable products, applications or environments (Sommerville and Nino, 2007)[51].

Operationalization of co-design

While co-design often appears in scientific literature, they are often unclear on the practicalities of *how* they applied co-design or designed their co-design sessions. For this reason, professional design literature was deemed more useful, such as the Co-Create Handbook[42] by CO-CREATE[1], their presentation on co-create basics[41]. CO-CREATE identifies four steps of co-design:

1. Engage
   Learn from each other and set the challenge
2. Understand
   Focus on user needs to gather key insights for each stakeholder
3. Ideate
   Cocreate design concepts and prototypes
4. Validate
   Check the validity and the co-created solution by evaluating concepts and prototypes with end users and other stakeholders

The advantage of this framework is that it can refer to concepts and prototypes, which means that the full cycle can be applied to a single co-design session and over multiple sessions. Other frameworks use phrases such as 'implementation' or 'deliver', which imply a tangible prototype, while 'ideate' and 'validate' can also refer to concepts. The technical complexity of the scientist-journalist recommender system made it challenging to deliver a prototype straight away, which meant that the initial development sessions were conceptual. Furthermore, the development was done between meetings. A framework allowing us to iterate over concepts is more flexible, as concepts can be developed and evaluated in a single session.

Co-design and learning

While much research has been done on co-design, nearly all focus on the process and the quality of results. In contrast to design thinking - which has been thoroughly studied as an educational method - there is little research on co-design and learning.

If learning is part of the study, it is often on the side of the designers. For example, Thamrin et al. (2019)[53] used co-design to help interior design students gain a deeper connection and understanding of users. It is often used in educational contexts, but this is generally to design educational material. Examples of co-design in education are Cuendet et al. (2013)[12], Walsh et al. (2010)[58] and Severance et al. (2016)[47]. Also outside of the classroom, co-design is used to create educational material, such as in healthcare (Power et al., 2022[40], Boyd et al., 2012)[9]) and in sports (Duncombe et al., 2022)[15]). It has been acknowledged that co-design leads to learning for all participants. In their literature review, David et al. (2013)[13] name social learning as one of the five core themes associated with co-design. It is mentioned as both an outcome and a requirement of successful co-design. However, it often focuses on the learning curve of the expert. While participants' learning curve is sometimes mentioned, few articles specifically state it as a goal or actively take steps to further this learning curve.

Nonetheless, it is possible to use co-design with the express purpose of learning. For example, Sommerville and Nino (2007)[51] use co-design with the specific goal of advancing organisational learning at the Dr. Martin Luther King Jr. Library in California. The library needed to reconsider its organisational purpose, mission, programmatic initiatives and the professional competencies of the staff. This was done through a large co-design initiative, supported with training and implementation of new tools. This learning was focused on their organisational purpose, not digital innovation.

[1] CO-CREATE is a European cooperation project to create and distribute a new curriculum on co-design co-funded by the Erasmus+ Programme of the European Union. `http://www.cocreate.training/`

5.2.2. An operationalisation of mental model change through double loop learning and absorptive capacity

Mental model change can be discussed in the context of single, double and triple-loop learning. It can specifically be linked to double-loop learning, which is defined as learning that transforms a mental model. Learning in organisations (as opposed to in an educational setting) can be related to absorptive capacity, which is the ability to learn. By regarding participation in the development process through the lens of absorptive capacity, the way it leads to mental model change can be analysed.

Single, double and triple loop learning

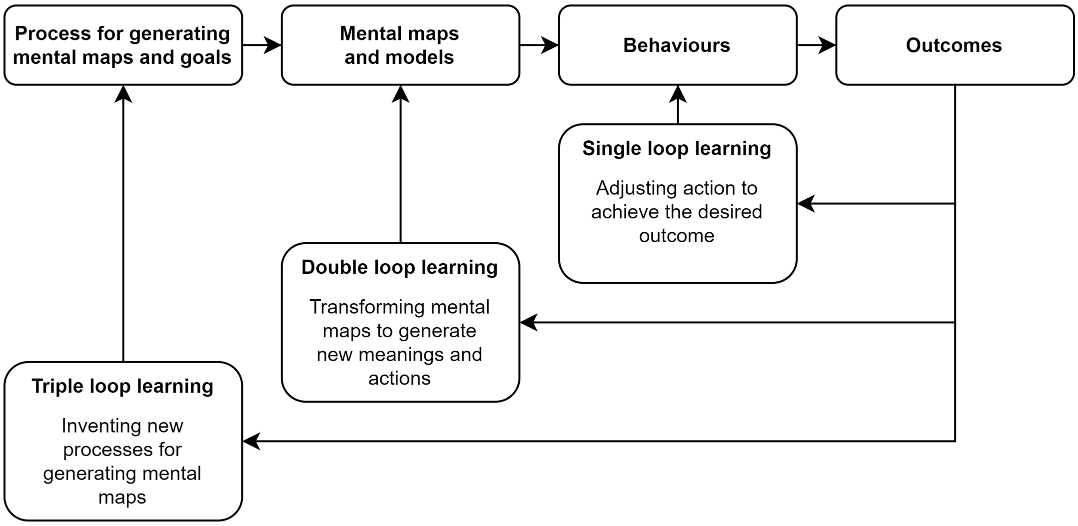


Figure 5.1: Figure 1 of Snell and Man-Kuen Chak (1998)[49]: Single, double and triple loop learning.

These concepts are used to classify the many different shapes and forms of learning into different levels. As described by Snell and Man-Kuen Chak (1998)[49], single-loop learning refers to a change in action that leads to changes in behaviour. Double-loop learning transforms mental maps and models to change the meaning. Triple-loop learning allows the learner to challenge and create new mental maps and models actively. Figure 1 of Snell and Man-Kuen Chak shows the relationship between these learning loops and is depicted in figure 5.1.

Participation in the development process has transformed the mental model of communication professionals regarding digital innovation. Therefore, double-loop learning has taken place. It would be triple-loop learning if this process led to an ability to instigate and direct innovation processes to further transform their mental model. However, while the communication professionals have come to a realisation that they need to get involved in the process (see section 14.2.2), it is the conviction of this author that they did not take full ownership of the process. It is not likely that participation in the development process has given them a process they can use to generate new mental models.

Double loop learning and absorptive capacity

Love et al. (2016)[30] link double loop learning to an increase in *absorptive capacity*. Cohen and Levinthal (1990)[11] define absorptive capacity as the firm's ability to "recognise the value of new, external information, assimilate it and apply it to commercial ends". Zahra and George (2002)[67] build on this and define it as "the set of organisational routines and processes by which firms acquire, assimilate, transform and exploit knowledge".

> **Acquisition**. Acquisition refers to a firm's capability to identify and acquire externally generated knowledge that is critical to its operations.
> ...
> **Assimilation**. Assimilation refers to the firm's routines and processes that allow it to analyse, process, interpret, and understand the information obtained from external sources.
> ...
> **Transformation**. Transformation denotes a firm's capability to develop and refine the

> routines that facilitate combining existing knowledge and the newly acquired and assimilated knowledge.
> ...
> **Exploitation**. [...] Exploitation as an organisational capability is based on the routines that allow firms to refine, extend, and leverage existing competencies or to create new ones by incorporating acquired and transformed knowledge into its operations.
>
> *p.190, Zahra and George (2002)[67]*

This breakdown of absorptive capacity allows for a fine-grained analysis of the effects of participation in the development process on the mental model by analysing the effect on each of the four aspects.

5.2.3. The Electron Learning Model: A theoretical model of the relationship between co-design and absorptive capacity

Section 5.2.2 describes how changes in mental model can be analysed through the lens of absorptive capacity, which breaks down into four aspects; acquisition, assimilation, transformation and exploitation. Section 5.2.1 describes how participation in the development process can be seen as a co-design process, which consists of a four-step cycle: engage, understand, ideate and validate. This section proposes a theoretical model of how co-design and absorptive capacity interact during the process to lead to mental model change.

Pairing absorptive capacity and co-design

Absorptive capacity refers to the ability to learn. Co-design was the method of learning. Action and ability can be paired to match each co-design step to an aspect of absorptive capacity. There is an interplay between them, as ability is required to perform an action, and performing an action can increase the ability to do so. It can be compared to physical exercise: The ability to exercise relies on the physical condition. By doing exercise, the physical condition improves, improving the ability to do physical exercise.

To acquire new knowledge requires the participants to engage with each other, which is, in turn, determined by their acquisition ability. To assimilate new information into their current knowledge, they need to understand the information, which is determined by their assimilation ability. To transform this knowledge, they need to ideate, i.e. engage in creating new ideas, limited by their transformation ability. To exploit these new ideas, they need to validate them, which relies on their validation ability.

The Electron Learning Model

This very neat pairing assumes a linear progression through the co-design steps and a rigid relationship between co-design steps and the aspects of absorptive capacity. It is unrealistic to expect the messiness of a complex learning problem to fit into a rigid model.

Rather than imagining the learning process as linear, it can be envisioned it as an electron moving in a figure eight loop through a magnetic field as depicted in figure 5.2. On the one side is ability, and on the other side is action. The process swings from one side to the other, both sides moderating and supporting the other. The facilitator can influence the action side. The facilitator initiates the swing and imparts momentum in the co-design sessions. In an ideal scenario, the swing starts in the inner loop, and gradually moves through each step towards the outer loop. If there are disturbances, however, the learning trajectory changes. For example, when in the ideation phase, they might realise there is a lack of understanding. This will slow down momentum, which means they will have to fall back to the inner loop, re-engage and first acquire new information.

The Inspiration for the name Electron Learning Model

The inspiration for the name Electron Learning Model is the evolution of models used to represent electrons in atomic nuclei. The Bohr model of the atom supposes neat electron orbitals and is still taught in most introductory science classes. However, in reality, electrons do not inhabit discrete orbitals but move around the nucleus in a chaotic 'cloud' (see figure 5.3). In a similar fashion, the Electron Learning Model is drawn as a system with a straightforward progression but acknowledges that the reality will be a much more chaotic trajectory through the model.

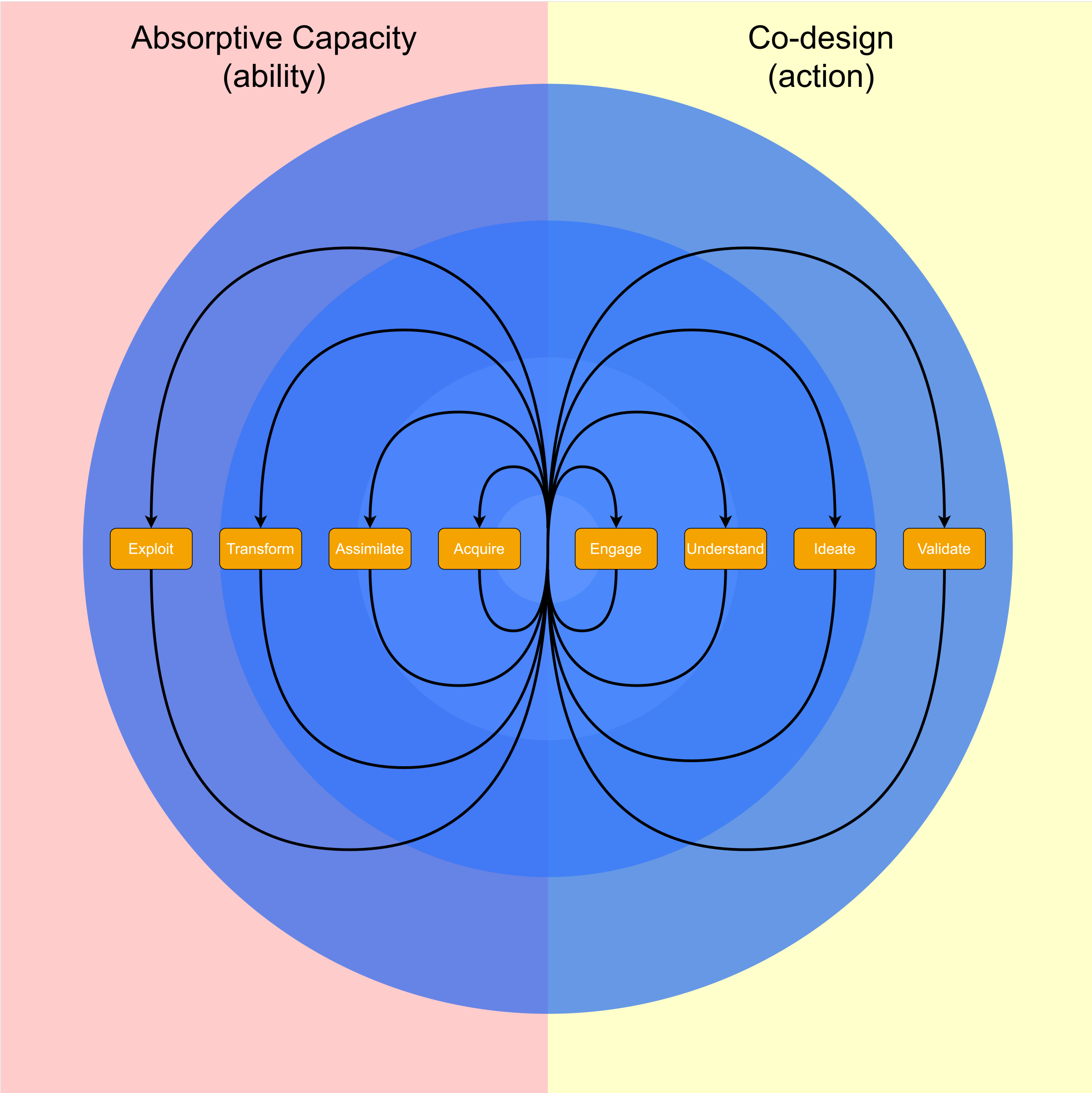


Figure 5.2: The Electron Learning Model.

Momentum in the Electron Learning Model

The concept of momentum plays an essential role in the model. Momentum through the learning trajectory represents involvement, i.e. active participation in the co-design process. Through the lens of this model, involvement can be regarded as our momentum moving through our learning trajectory.

In physics, two important concepts are associated with momentum: force and inertia. Force is defined as a change in momentum. If any change in momentum is observed, then by definition, a force must have been applied. Inertia is an object's resistance to force. It governs how hard it is to get an object moving and its resistance to being stopped once in motion.

Force, in this regard, is a change in the trajectory through the model. For example, the facilitator applies a force on the action side by initiating an ideation session. If the process gets stuck and the group has to re-engage, a force has been applied. The facilitator can only directly interact with the action side. By analysing the effects of the force the facilitator applies, a mental map of the other forces interacting with the learning process can be made. Based on this, the approach can be altered to keep increasing momentum.

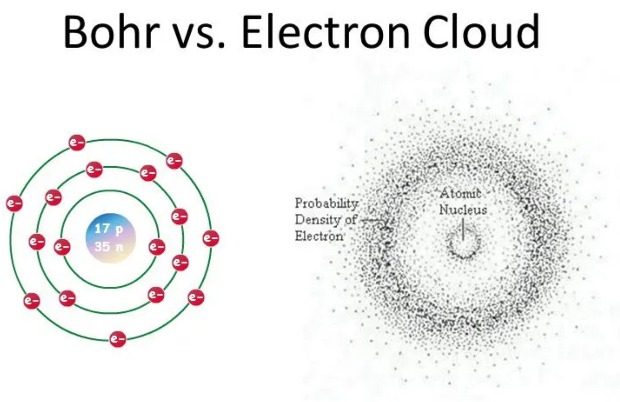


Figure 5.3: The Bohr model vs the Electron Cloud model. While the Bohr model provides an Image taken from the TracingCurves blog written by Smith (2019)[48]

If momentum represents involvement, then inertia represents the difficulty of getting involved with innovation. It needs to be overcome to get the learning process going. However, in physics, inertia also makes an object in motion harder to stop. If enough momentum can be built, it should start to work in favour of the process. Once the co-design project starts going, it should gain traction. Inertia is also what can keep the momentum going between co-design sessions.

The goal is to build momentum through the Electron Learning Model with the team of communication professionals. In the beginning, it will be unfamiliar to them. It will require a force from the facilitator to initiate the co-design sessions and keep up the momentum. However, they should become more familiar with the process as they go. Like a child on a swing who is first pushed by their parents but gradually learns to swing themselves, the communication professionals should become more and more adept at co-design. The goal is to get them to a state where they can keep the co-design process going by themselves with this author in the role of developer only, rather than facilitator.

Validation of the Electron Learning Model

During the meeting on November 19th, the Electron Learning Model was shown to the communication professionals. They were enthusiastic about the model and started describing their process up to that point in terms of the Electron Learning Model.

> In this process, we first went from the outside in. We were inspired by the idea of a concrete, useful tool. Based on this, the question was asked what that tool could do, and this project started. We turned out to have too little understanding of the digital world to get started right away, so we first had to gain new knowledge. Now we are working from the inside out again. We learn new things, and as a result, we begin to better understand how digital technologies work and how we can be part of this development process ourselves. Based on this, we will soon be able to come up with new ideas and ultimately make a useful algorithm.
>
> *Debrief meeting November 19th, 2020*

It is a positive sign that the participants were immediately grasped by the model and were able to phrase their own process in terms of the Electron Learning Model. However, the barriers encountered during the development process made it impossible to validate the Electron Learning Model further.

5.2.4. A Develop approach based on the Electron Learning Model

The Develop approach based on the Electron Learning Model involved a two-pronged approach. An intervention is done on each side of the 'swing' between the absorptive capacity side and the co-design side. On the one hand, the Agile development process (see 2.4) consists of co-design sessions. The sessions start with a brief update on the development progress and then move through the co-design cycle to decide on the next steps. On the other hand, the blog will directly address forces on the absorptive capacity side. It can provide targeted information on gaps in knowledge and new perspectives on digital technology. This is represented in figure 5.4.

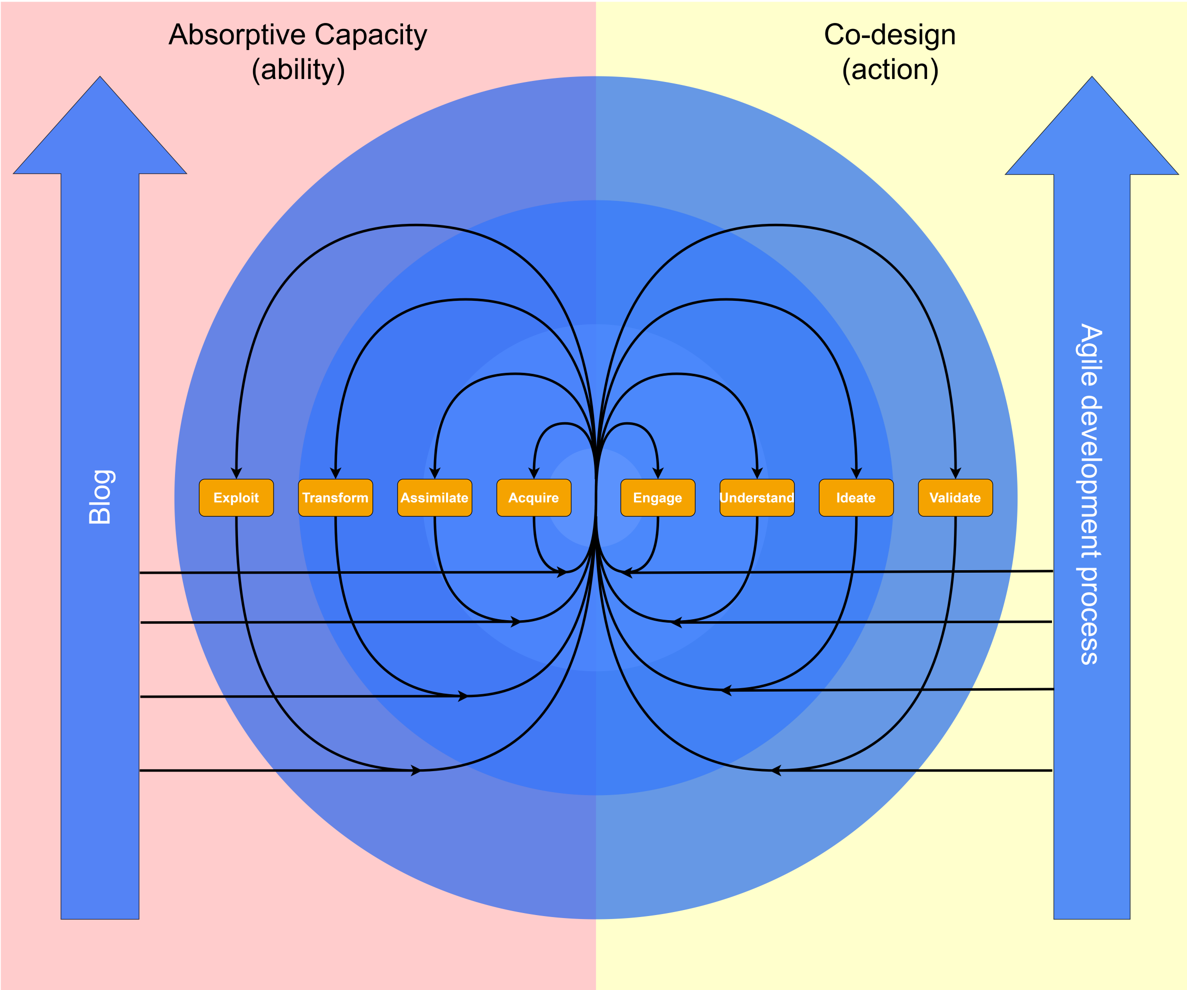


Figure 5.4: A Develop approach based on the Electron Learning Model

6

Develop

The goal of the Develop phase was to develop the digital tool selected in the Define phase and collect data on the effect of participation on the mental model of digital innovation. During the Develop stage, several barriers were encountered with the approach chosen in the Define stage, which led to an iteration of this approach.

6.1. Barriers during the Develop phase

Several issues were encountered during the Develop phase. Section 6.1.1 describes the problems that were encountered. Section 6.1.2 discusses these issues and attempts to explain them using theory and the final reflections on the process by the communication professionals.

6.1.1. The issues encountered during the Develop phase

Several issues were encountered during the Develop phase. This section aims to give an overview of the problems encountered.

The issues with the co-design sessions

The team of communication professionals and this author did not manage to have any sessions in which all steps of co-design were completed. The intent was to start each session with a brief update on the development progress and then move through the co-design cycle to decide on the next steps. During the update, there was a struggle to convey to the team of communication professionals what exactly the progress was. This also led to a struggle in the Engage phase to convey the considerations regarding the next steps. It led to a one-sided conversation with the communication professionals asking questions and this author attempting to explain the technical aspects of the current situation. Attempts to engage them by asking questions on possible next steps led to new questions on their side. This led to a situation where the sessions were more akin to lectures than co-design sessions.

The issues with the blog

It was decided in the meeting on June 22nd 2020, to start a blog, partly to bridge gaps in technical knowledge of the team and partly to inform their colleagues (see 5.1.3). After the summer break, the first two posts were published on the personal website of the author and shared with the communication professionals on September 14th for initial feedback. Appendix B shows the blog page and an example post. On November 2nd, the communication professionals informed this author that while they were fine with a public blog covering general topics, they would prefer not to have a public blog that contained information on the development process. Instead, they would prefer the posts on the development process to be private, so they could still fulfil the function of addressing gaps in knowledge during the development process. The communication professionals had, at this point, not read the original two posts. The blog was modified to have a log-in section with private posts, and on December 3rd, another post on a general topic was created and shared with the communication professionals for feedback. After a reminder, one communication professional sent feedback on January 4th, 2021. In the session on January 28th, 2021, it was decided to discontinue the blog.

The issues with planning
There was an issue with planning regular meetings. Table 6.1 shows all meetings in the Develop phase. It was hard to find moments when the entire team was available. Even then, a majority of the time, there were absentees. There were several issues with cancellations not coming through, which meant that it was only in the meeting itself that it became apparent a participant would not be present. Within the team, speculation was that the issue was with notifications of declined calendar invites. Some meeting invites were planned by the Science Education and Communication secretariat, which might have caused invitation declines to not show up for everyone.

Date	Meeting
2020-09-08	Discuss data sources and algorithm outline P3 absent
2020-11-19	Present Electron Learning Model and discuss general project planning
2021-01-28	Show first prototype, discuss next steps P1absent
2021-02-15	Meeting cancelled due to illness facilitator
2021-03-18	Miscommunication in cancellation Only P2 present
2021-04-15	Reflection and restart P2 and P4 absent

Table 6.1: The meetings during the Develop phase.

6.1.2. Discussion of the issues encountered during the Develop phase
This section discusses the issues encountered and attempts to explain them. The interpretation is based on the literature and reflections of the communication professionals in the stimulated recall sessions at the end of the process.

The skill floor of participation in a digital development process
As described in section 5.1.2, the choice was made to develop a scientist-journalist recommender system. However, it became apparent during the process that the skill floor to participate in the co-design process was too high for the communication professionals to engage with it.

A requirement for co-design is a shared understanding (Kleinsmann and Valkenburg, 2008)[25]. Co-design has been used successfully in many fields, such as architecture, interior design, product design and design of educational material. The more tangible the subject of the design process, the easier it is to create a shared understanding. Co-design works in architecture because participants can imagine a house. Co-design works in education design because everybody has been through education. Co-design works in urban design because we all have walked through a city. However, it is hard to imagine a mathematical model.

Without shared understanding, the system becomes a black box. The 'black box problem' is a well-known issue of artificial intelligence systems (Ribeiro et al., 2016[43], Adadi and Mohammed, 2018[2]). It was attempted to make it more tangible by using network analysis. Many network features have a physical interpretation. For example, the degree of a scientist node in the network is the number of other scientists that the scientist has worked with. It was thought that this would make it conceptually more manageable for people to grasp than most machine learning algorithms. Nevertheless, it still relies on complex mathematics.

As long as the technology is a black box to the communication professionals, they can only participate in deciding what should go into the box and what should come out of it. Still, they cannot participate in the development itself. This is why engagement was high during the Discover and Define phases but low during the Develop phase.

> We needed to get a little tech savvy first [...] Then we got a period in which we got a lot of theory from you and then the question of 'oh what will come out of this?'

P1, Stimulated Recall Delft University of Technology [21:00]

> Now it is getting more tangible again, now it is fun again, but in between, there was a (pauses and makes a gesture of downward line) moment.
>
> *P3, Stimulated Recall Naturalis Biodiversity Center, [36:00]*

The Discover phase focused on their motivations for this project and their learning goals. This did not require any technical knowledge. The focus of the Define phase was to decide what tool to develop. This did require a base level of technical expertise. These phases were not without issues. They were confronted with the gap between their expectations and the technical reality (see section 4.1.2). They started in "innovation land [...] unencumbered by any kind of knowledge" (P1, Stimulated Recall Delft University of Technology [39:00]). As described in the minute meetings quoted in section 5.2.3, there was a process of realising the gaps in their understanding and their lack of absorptive capacity. Nevertheless, during these phases, the participants showed motivation and enthusiasm.

However, during the development phase, the lack of technical absorptive capacity was too great to go through the process. The effect was that the meetings became lectures on computer science theory. They lacked the base knowledge to assimilate the new information, which meant that the process never continued beyond the second shell of the Electron Learning Model. In turn, this meant that it was not possible to build momentum. Content from the previous session did not stick, which meant that many things had to be repeated and re-engaged with.

> I think you have explained the theory between deep learning up three times now [...] it is probably very clear for you, for me it is still a bit diffuse [...] because of course you have a lot to do with machine learning and deep learning, but I still don't understand.
>
> *P1, Stimulated Recall Delft University of Technology [23:00]*

The struggle to prioritise innovation among other responsibilities

This project was initiated at the request of the communication professionals. As described in section 4.3, they were motivated to learn how artificial intelligence and digital technologies could impact their field, they wanted to bring innovation to the 'often conservative communication landscape', and they were motivated by the thought of getting a custom-made digital tool out of this process. However, despite their motivation, they struggled to find time for the project due to the lack of direct, external incentives to participate in the project, coupled with their job's immediate tasks and responsibilities.

> It's not at the top of the priority list. We committed to this and I enjoy it and it's interesting too, you know. So you want to make time for it, for you, but also for ourselves, but it's really often that you think 'when can we do this' because all those other things that have to be done too. I sometimes find it annoying, that because of that I have the idea that we can't put in the time that maybe we should. *P3, Stimulated Recall Naturalis Biodiversity Center [28:00]*

> But that is always the dilemma I face. Do I free up the time or not, because I do throw things from my agenda all throughout the day. [...] I try very hard not to do that, because before you know it you're only working on ad-hoc things and not just with the ongoing long-term things, but sometimes it's hard to keep up. *P3, Stimulated Recall Naturalis Biodiversity Center [30:00]*

This is also part of the reason the blog was ineffective. One of the participants reflected that they intended to read it, but it was not a priority, so it ended up at the bottom of the to-do list and got forgotten.

> Yes, I really wanted that, I really intended to read that, but then it is not a priority, so then it is really something you do for fun, that means that you spend your weekend or evening doing it do and then there are tons of other things and then it just didn't happen. [...] it's actually on my to-do list, still. It's not that I don't want it. [...] It's not unwillingness, let's put it this way, more a matter of what really needs to be done and what can possibly wait and then it drops down the list and then the attention goes off of it. Then you send an email like 'gosh, have you looked at it yet? '. Oh yes, and then it goes up the list again. Yeah, no, I'm really going to look into it. *P3, Stimulated Recall Naturalis Biodiversity Center [38:00]*

This struggle is not unique to this project. Vial (2019)[57] describes how there is an incentive for organisations to focus on the short-term demands of their existing procedures and obligations. While top management often considers developing digital technologies as beneficial to the organisation, ”the structural components of the organisation, both tangible (e.g., means of production) and intangible (e.g., organisational culture), are so embedded within everyday practices that they stifle the innovative and disruptive power of digital technologies” (p.130, Vial, 2019)[57]. There is always a tension between short-term responsibilities and long-term projects. There is a lack of urgency. While they acknowledge that it is crucial to keep up with digital developments, nothing within their direct environment brings them into contact with digital developments, stimulates them to engage with it or confronts them with the possible consequences of not engaging with it. At the same time, there are always issues that need to be addressed right now.

When things are going well, there is no need to divest resources to start a complex, time-consuming and uncertain innovation process. When things are not going well, there are no resources to be spent on a complex, time-consuming and uncertain innovation process. By the time competitors rise to market dominance by exploiting new technologies, they have such a head start that it is very hard to catch up. That is why, as Hill and Rothaermel (2003)[22] note, ”A persistent theme in the academic literature on technological innovation is that incumbent enterprises have great difficulty crossing the abyss created by a radical technological innovation and, thus, go into decline, while new entrants rise to market dominance by exploiting the new technology.” (p.1, Hill and Rothaermel, 2003).

Innovation is never something that is needed today. It is either needed tomorrow or needed yesterday. The communication professionals were continuously struggling to find time for a process with no short-term, tangible benefits and no external incentives, even though they were intrinsically motivated.

The interplay between these barriers
These barriers reinforced each other. Because of the technical complexity, it was hard to create a shared understanding. The communication professionals struggled to dedicate time to the project. The efforts to engage the communication professionals outside of meetings were unsuccessful, so all knowledge transfer needed to happen during the sessions.

> I could also have read up first, and yes I didn’t, so I need to be educated.
>
> *P1, Stimulated Recall Delft University of Technology [25:00]*

However, there also were fewer meetings, which meant less time to develop a shared understanding. Many meetings were missing one or more participants, which also meant that not all participants were part of the limited shared understanding that was established.

> Then we got a period in which we got a lot of theory from you and then the question remains of ’oh what comes out?’, I am impatient by nature, so that is fed by that
>
> *P1, Stimulated Recall Delft University of Technology [23:00]*

This meant a lack of momentum and less progress, which was detrimental to the motivation of the communication professionals to prioritise it over other responsibilities, which in turn led to less results.

> If you don’t put in energy yourself then it will eventually yield nothing, so that’s why I think we should do it [...] We think it’s important and we entered into it so we finish it, but sometimes we would like to put a little more energy into it. If I really had more time, I definitely would. Because it’s interesting, it’s fun and I think I could get even more out of it myself.
>
> *P3, Stimulated Recall Naturalis Biodiversity Center [40:00]*

6.2. The approach of the last part of the Develop phase

6.2.1. Reflect and restart meeting

On April 15th 2021, a reflect and restart meeting was held to rekindle engagement and pick a new direction. The purpose was to recap the original goals, reflect on what they had learned so far, and discuss what they wanted from this project going forward. Only P1 and P3 were present at this meeting.

Results reflect and restart meeting
The meeting was an open conversation around seven consecutive questions.

1. The project started with the idea of an algorithm as a colleague. What were the expectations?

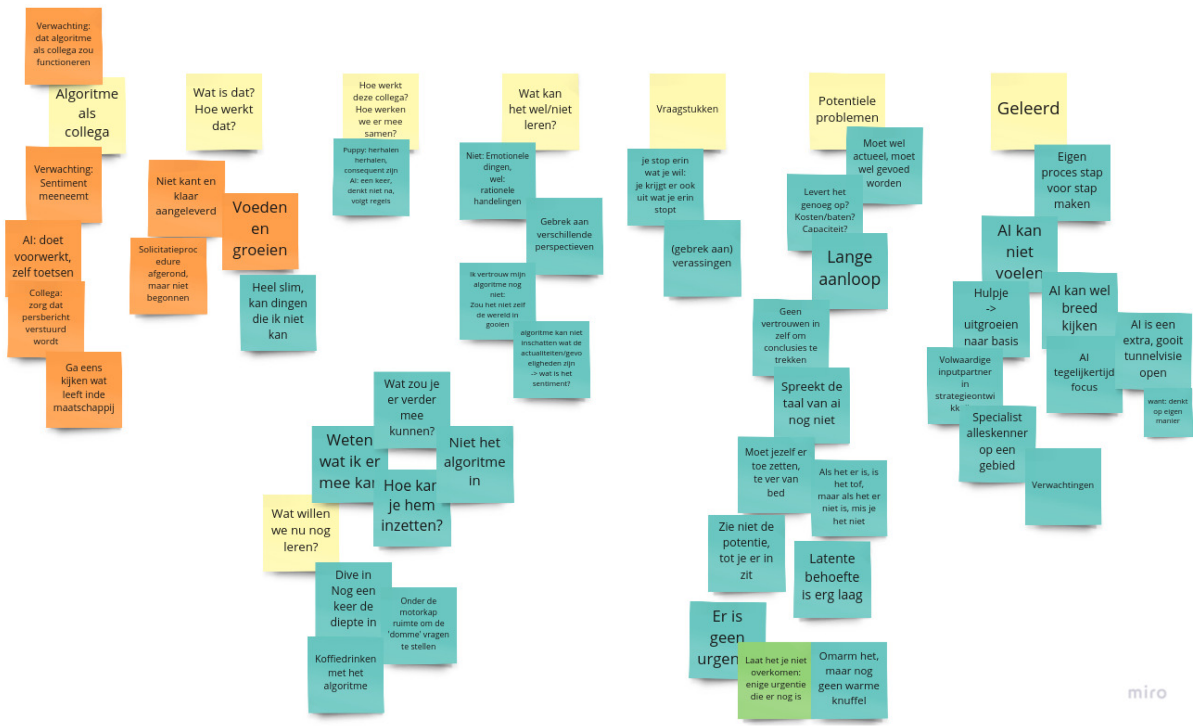


Figure 6.1

2. How do you look at those things now?
3. How do you work with this colleague?
4. What can and can't it do?
5. What are questions we still have?
6. What are potential issues?
7. What have we learned?
8. What do we still want to learn?

These questions were prepared on a Miro board, which was open on the facilitator's screen and shared on the Zoom call. The facilitator wrote down post-it notes during the meeting. Figure 6.1 shows the resulting Miro board.

Discussion of results of the reflect and restart meeting

The expectations were that the digital colleague would function as a full colleague, similar to human colleagues. They expected it to autonomously analyse the sentiments going on in society and write a press release based on its findings. They have since realised that it can do tremendously smart things that humans can't, but it is not a ready-made thing. A recurring analogy throughout the process was that of AI as a puppy, which needs to be trained before it does what you want it to do. You need to feed it to make it grow. The adage of 'garbage in, garbage out' was quoted: you get out of it what you put into it. This raised the worry of whether the results of the recommender system would give the same recommendations they themselves would make or whether it would bring novel recommendations. This led to a conversation about how much value it would actually add.

Apart from issues surrounding the recommender system, issues with the process were discussed. They discussed that digital innovation is a world that is far removed from their daily practice. Once it is there, it is great, but if it is not, it is not something that is missed. They are not confronted with any adverse effects if they do not engage with digital innovation or with the potential gains if they do. It means that it is something they need to force themselves to engage with. Their only urgency is to not be caught unaware when it will eventually be on their doorstep.

Their learning goals for the final part of the project diverged. P1 explicitly mentioned wanting to dive into the algorithm and understand how it works, while P3 expressly said not to want to go into the details but rather focus on understanding what it can do and how they can work with it.

A discussion on the implications of their mental model
The reflect and restart meeting demonstrated a shift in their mental model of digital innovation. Firstly, they actively contrast their initial expectations of an autonomous digital colleague with their experiences. They displayed more nuance in discussing what the system can or can't bring. Secondly, they show an increased awareness they need to engage with the system. They describe it less as an external 'magic black box'. Instead, they involve their usage and training in discussing what the system can bring.

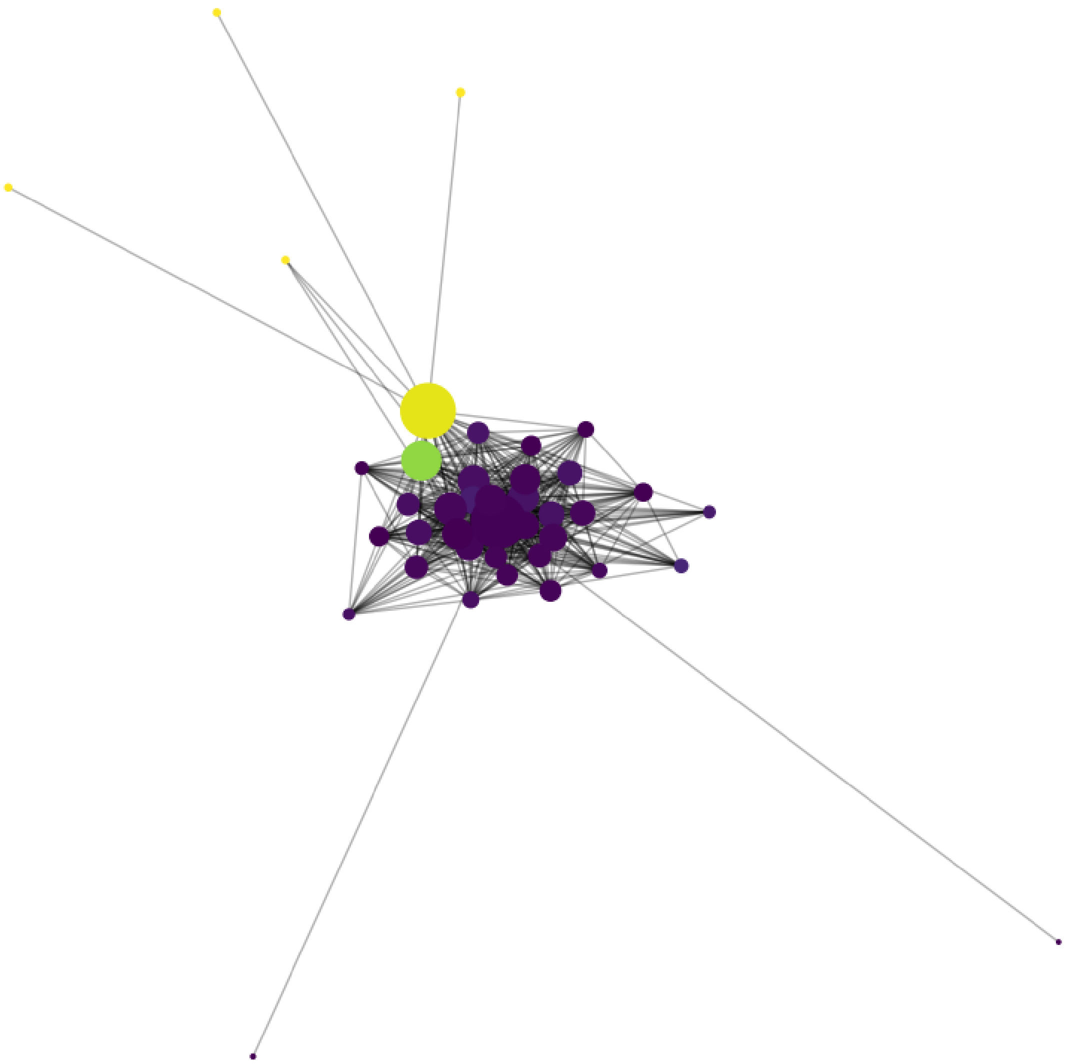
6.2.2. The final stage of the Develop phase

After the reflect and restart meeting, it was decided to finish the prototype as quickly as possible and continue towards the Deliver phase without input from the team of communication professionals on the technical aspect of the process.

The communication professionals showed they were motivated by the final product. The purpose of the Deliver phase is to transform the prototype into a working product. Firstly, it was thought that this part of the process would motivate the communication professionals. Secondly, this step focuses on the user interface, which was thought to be a more accessible subject for co-design. Lastly, this allows the process to move forward. To complete the research on mental model change, the participants should go through the whole double-diamond design process. By proceeding to the next step, the process can be brought to a conclusion that is satisfying for the participants and allows the research to continue.

III

The network-based scientist-journalist recommender system



This part is written as a stand-alone report covering the Computer Science aspects of this research.

Abstract

Communication professionals at research institutes are tasked with connecting scientists and journalists. In this research, a recommender system is developed to support this process by recommending scientist-journalist connections based on data from previous collaborations. A scientist collaboration network, a journalist collaboration network and a scientist-journalist collaboration network are combined into a multilayer network. A recommender system is designed based on centrality metrics in the scientist and journalist collaboration networks and distance metrics in the multilayer network. In contrast to traditional link prediction problems - which aim to predict what links are most likely to form in the network - the problem in this thesis is how to recommend the most likely link for a single node, i.e. the most likely scientist links for a given journalist or most likely journalist links for a given scientist. A novel evaluation method is created to evaluate the performance of the recommender system.

7 Introduction

Scientists can provide valuable knowledge on many societal issues. However, often scientists are not included in the public debate. For example, Merkley (2020)[33] found that only 23% of news articles on issues such as climate change, nuclear safety or immigration have a message of a scientific expert. Many research institutes have a communication department to promote the visibility of scientists and scientific research. One of the tasks of these communication departments is to connect scientists and journalists to get news coverage of the research done at their institute.

The process for communication professionals to get a scientist in the news starts when a scientist comes to the attention of the communication professional or when a journalist reaches out to the communication department. The job of communication professionals is to find a suitable journalist or scientist and bring the two into contact. This matching process often has to happen on short notice, especially when a journalist reaches out to the communication department. The difficulty is that communication professionals often have limited knowledge of which scientists are working on what topics, and which journalists are working on what topics. Communication professionals maintain an active personal network of scientists and journalists to find good matches. They use this network to ask for referrals. However, this referral process is time intensive. For this reason, there is often not enough time to consider a broad scope of candidates.

This project uses network science to analyse patterns in previous scientist-journalist collaborations to develop a scientist-journalist recommender system. This system suggests journalists that are potentially interesting for a given researcher or scientists that are potentially interesting for a given journalist. Communication professionals can use these suggestions as input for their matching process, which relies on many different considerations that vary on a case-to-case basis (e.g. personal intuition, the research topic, the target audience of the journalist, or the charisma of the scientist).

7.1. Problem statement

Scientists can be represented by a scientific collaboration network, where the nodes represent scientists. Two scientists are connected if they have co-authored a scientific paper, and the weight of the connection is given by the number of scientific papers they have collaborated on. Journalists can similarly be represented by a collaboration network, in which nodes represent journalists, two journalist nodes are connected if they have co-authored a newspaper article together, and the weight of the connection is given by the number of newspaper articles they have collaborated on. A bipartite scientist-journalist collaboration network represents scientist-journalist collaborations. The nodes in this network are all nodes in the scientist network and all nodes in the journalist network. A scientist node and a journalist node are connected if the scientist is mentioned in a newspaper article written by the journalist. These three networks provide valuable information about the interactions between different scientists, between journalists, and between scientists and journalists. This information is used to identify potential interests between scientists and journalists.

For a given scientist (journalist), the goal is to predict which journalists (scientists) are the most likely to collaborate (form a link) with this scientist (journalist). This is similar to a link prediction problem. A link prediction problem considers a network $G = (V, E)$ at an observed time t - with V being the nodes and E the edges - and aims to predict which links have been missed in the observed network or which links will to form in the future. In the case of a social network - where nodes represent actors and edges represent relationships or interactions - link prediction aims to predict future relationships/interactions (Wang et al., 2015)[60] .

The use-case of this recommender system is to support communication professionals in matching scientists and journalists by quickly providing different possibilities. A specific question triggers the usage of the recommender system. For example, a communication professional comes across a PhD student that is working on an interesting topic[1] and wants to find a journalist that is willing to write an article on this topic. PhD students are not the most likely candidates for future scientist-journalist connections. When considering which scientist-journalist collaborations are the most likely to form among the different possibilities, the likelihood is low that a collaboration with a PhD student is suggested. However, in this particular use-case, there is extra information: a communication professional has determined to promote a journalist collaboration for *this* PhD student. The question is which journalists are the most likely to be interested in collaborating with this PhD student, not whether any collaborations with this PhD student are likely to happen based on historical data. So, in contrast to general link prediction problems - which are interested in what links are most likely to form in the network in general - the problem researched in this thesis is how to predict which links are most likely to form for a given node. This will be referred to as a *nodal link prediction problem*.

7.2. Research question

This research investigates how network features in the multilayer network, the scientist network and the journalist network can be used to predict links for nodes in the multilayer collaboration network. This is done by answering the following three questions:

1. How can the scientist network, the journalist network and the scientist-journalist network be combined to create a multilayer network?
2. Which centrality metrics are the most useful for identifying nodes in the scientist and journalist networks with scientist-journalist collaborations?
3. How can network features in the multilayer network be used to design a network-based scientist-journalist recommender system?

7.3. Contribution

There are four contributions of this thesis:

1. The first contribution is the creation of three different networks - a scientist collaboration network, a journalist collaboration network and a scientist-journalist collaboration network - and combining these networks into a multilayer network.
2. The second contribution is insight into what nodal properties in the scientist network differentiate scientists with journalist collaborations from scientists without journalist collaborations and what nodal properties in the journalist network differentiate journalists with scientist collaborations from journalists without scientist collaborations.
3. The third contribution is the creation of a novel recommender system that utilises properties from the scientist and journalist networks, as well as the multilayer network, in order to recommend scientist-journalist connections.
4. The final contribution is the creation of a new evaluation method. Most link prediction methods predict which links are most likely to occur within the entire network. This system predicts which links are most likely for a specific node. For this reason, most standard evaluation methods, such

[1]This is based on an example of a PhD student who was working on a technology that could also be used to generate electricity from urine: `https://www.tudelft.nl/en/stories/articles/taking-the-piss-or-turning-it-into-energy`

as AUC-ROC, are not applicable. A new evaluation method is developed to assess the accuracy of the recommender system.

8
Background

This chapter provides background information which is used in this thesis. Firstly, section 8.1 discusses the general approach to solving the nodal link prediction problem. Next, some background on collaboration networks is provided in section 8.2. Section 8.3 provides an overview of the network metrics that will be discussed.

8.1. The nodal link prediction problem

There are two types of approaches to solving link prediction problems: *similarity-based approaches* and *learning-based approaches* (Wang et al., 2015[60], Zhan et al., 2020[69], Zou et al., 2021[70]). Similarity-based approaches use network metrics to assign a score to pairs of unconnected nodes. A higher score means a higher probability that a connection will form between the two nodes. The possible connections are ranked based on this score. Learning-based approaches treat link prediction as a binary classification problem. A machine learning model is trained to identify which connections will appear in the future and which will not.

The choice was made to use a similarity-based approach. The rationale was that an approach based on scoring node pairs and ranking them according to highest score is more easily adapted to the nodal link prediction problem than an approach based on binary classification. Only the relative scores are relevant for the scoring and ranking approach, so the approach will still work for nodes with only low-scoring potential connections. An added benefit of similarity-based approaches is that, in general, they are more explainable. Explainability refers to the extent to which a user can understand how a system comes to its results. Explainability is important to justify unintuitive results, control the correctness of results, allow further improvement of the system and discover patterns in the subject being modelled (Adadi and Mohammed, 2018)[2]. Given that the development of this system was used as a vessel to change the mental model of digital innovation of communication professionals (see II)), the system and the process should be explainable. Many network features have a physical interpretation. For example, the shortest path in a social network can be explained as the number of friends of a friend that you need to connect two people. This physical interpretation of network features makes recommendations based on these features easier to explain. In contrast, many machine-learning algorithms struggle with explainability. Many complex models are black boxes: It is unclear how the inner mechanics influence the predictions (Barredo Arrieta et al., 2020)[6]. The increased explainability of similarity-based approaches in comparison to learning-based approaches was another reason to prefer a similarity-based approach.

8.2. Collaboration networks

A collaboration network is a weighted network where the nodes represent individuals, and two individuals are connected if they have collaborated in some form. The weight of a connection between two individuals is the number of times these individuals have collaborated. A popular example of such a network is a scientific co-authorship network, where the nodes are scientists and two scientists are connected if they have co-authored a scientific publication (e.g. Persson et al., 2004[39], Milojević, 2010[35], Abassi et al., 2012[1]).

A journalist collaboration network consists of journalist nodes, where two journalists are connected if they have co-authored a news article. To the best of this author's knowledge, no research has been done into a journalist collaboration network.

8.3. Network metrics

Similarity-based approaches that are used to solve classic link prediction problems utilise network metrics to assign a score to node pairs. In this section, firstly, a set of centrality metrics is introduced, each of which describes a particular network property of a node. Next, network metrics are introduced that describe properties of a node pair. These two types of metrics will be utilised to design the score of node pairs in the recommender systems. In the introduction of these metrics, the undirected, weighted network $G = (V, E, w)$ is used, where V is the set the nodes, E the set of edges and each edge $(u, v) \in E$ has a weight that is given by the function $w : E \to \mathbb{N}$ (Li et al., 2015)[27].

8.3.1. Local centrality metrics

Local centrality metrics are centrality metrics that can be derived from a local neighbourhood of a node. In contrast, global centrality metrics require information from the entire network to compute the centrality (Hernandez and Van Mieghem, 2011)[21]. Three local network metrics are introduced; degree, weighted degree and clustering coefficient.

Degree. The degree d_u of a node $u \in V$ is the number of other nodes that u is connected to. Let N denote all neighbours of u: $N = \{v \in V \mid (u, v) \in E\}$. The degree is defined as follows:

$$d_u = |N| \tag{8.1}$$

In a collaboration network, the degree of a node $u \in V$ (scientist or journalist) is equal to the number of people they have collaborated with.

Weighted Degree. The weighted degree w_u of a node $u \in V$ is the sum of the weights of all edges connected to u. Let N again denote all neighbours of u. The weighted degree is defined as follows:

$$w_i = \sum_{v \in N} w(i, v) \tag{8.2}$$

For a scientist (journalist) in the collaboration network, this equals the total number of collaborations with other scientists (journalists).

Clustering coefficient. The clustering coefficient c_u of a node $u \in V$ measures the cliquishness of the neighbourhood of u, i.e. the extent to which the neighbourhood is fully connected. Equation 14 of Hernandez and Van Mieghem (2011) shows how it is calculated for a node u, where y_u is the number of links between neighbours of u and d_u the degree of u.

$$c_u = \frac{y_u}{\binom{d_u}{2}} \tag{8.3}$$

Hernandez and Van Mieghem note that it is equal to the probability that two neighbours of a node are neighbours themselves. The edges in the co-authorship networks are based on (scientific or newspaper) articles, which means that all co-authors of an article form a clique in the co-authorship network. This means it can be expected that the clustering coefficient will generally be high. This is especially true in the scientist co-authorship network, as scientific papers often have many authors.

8.3.2. Global centrality metrics

Two global network metrics will be used in this thesis; betweenness centrality and eigenvector centrality. Global network metrics are computed using information from the entire network.

Betweenness centrality. The betweenness of a node $k \in V$ is a measure of the number of shortest paths in the graph that traverse k. Let the number of shortest paths between nodes $u, v \in V$ be given by σ_{uv} and let the number of shortest paths between nodes u and v that traverse k be given by $\sigma_{uv}(k)$. The betweenness centrality B_k is given by:

$$B_k = \sum_{u,v \in V} \frac{\sigma_{uv}(k)}{\sigma_{uv}} \tag{8.4}$$

Hernandez and Van Mieghem note that the betweenness centrality is a measure of the influence of a node in a social network.

Eigenvector centrality. The eigenvector centrality of a node, also called the *prestige*, measures a node's importance based on its neighbours' importance. It is based on the idea that connections to well-connected nodes should be considered more important than connections to less well-connected nodes.

The adjacency matrix A is a symmetric $|V| \times |V|$ matrix, where each element A_{uv} is one if node u is connected to node v and zero otherwise (Li et al., 2015)[27]. Let λ be the eigenvalue of A. The eigenvector centrality x_u of node $u \in V$ is calculated as follows:

$$x_u = \frac{1}{\lambda} \sum_{j \in V} A_{uv} x_v \tag{8.5}$$

8.3.3. Distance metrics

Distance metrics provide information about how close two nodes $u, v \in V$ are in the network.

Hop count. The hop count H_{uv} refers to the unweighted distance between nodes u and v, i.e. the number of links in the shortest path. When distance is used as a similarity score in a link prediction problem, two nodes that are closer should receive a higher score. For this reason, the hop count similarity HS_{uv} between u and v is calculated using the negative hop count, $HS_{uv} = -H_{uv}$ (Ayoub et al., 2023)[4].

Weighted distance. The weighted distance treats the weight $w(u, v)$ of an edge (u, v) as the distance between u and v. The length of a path $p = x_1, x_2, ..., x_n$ is the sum of all weights: $\sum_{i=1}^{n-1} w(x_i, x_{i+1})$. It uses this path length to find the length of the shortest path $\delta(u, v)$. Similar to hop count, when used as a similarity metric, the weighted distance similarity WS_{uv} is calculated using the negative weighted distance, $WS_{uv} = -\delta(u, v)$ (Ayoub et al., 2023)[4]

9

The construction of the collaboration network

The aim of this chapter is to answer research question 1 on how the scientist network, the journalist network and the scientist-journalist network are combined to create a multilayer network.

Section 9.1 describes the data sources that were used: A dataset of newspaper articles from DPG media, a dataset of scientific paper metadata from SciVal, and employee lists of Delft University of Technology and Naturalis Biodiversity Center.

Section 9.2 discusses how these data sources were integrated. The scientists on the employee lists of Delft University of Technology and Naturalis Biodiversity Center had to be matched to scientific articles from the SciVal dataset and to newspaper articles from DPG media.

Section 9.3 explains how three networks were created based on the integrated data; a scientist collaboration network, a journalist collaboration network and the scientist-journalist collaboration network, and how these three networks were integrated into a multilayer network.

9.1. Data sources

Three different data sources were used, the DpgMedia2019 dataset of newspaper articles, a Scival dataset containing scientific article metadata and employee lists of Delft University of Technology and Naturalis Biodiversity Center.

9.1.1. The DPG media newspaper articles dataset

The DpgMedia2019 dataset published by Yeh et al. (2019)[66] was used. It contains 103812 records of Dutch newspaper articles from newspapers owned by DPGMedia[1]. The newspaper articles are dated between January 1st, 2017, and June 17th, 2019.

For each article, the dataset contained *publisher*, *publishing date*, *article title*, *article text* and *article url*. It did not contain the journalist names that wrote the article. The Python web-scraping framework *Scrapy* was used to retrieve the journalist names of each article from the article URL. The URLs of *Het Parool* were invalid, i.e. they could not be resolved to a web page. All *Het Parool* articles were dropped from the dataset.

The final dataset contained 102,457 newspaper articles written by 4,601 journalists.

[1]The DPG Media brands included in the DpgMedia2019 dataset are: AD, Het Parool, Trouw, Volkskrant, Eindhovens dagblad, De Gelderlander, Tubantia, De Stentor, De Stem and PZC. For more information on DPG Media, see `https://www.dpgmediagroup.com/en`.

9.1.2. SciVal scientific article metadata

SciVal is a research analytics tool for accessing the Scopus dataset[2]. SciVal was used to export the metadata of 89,513 articles co-authored by scientists affiliated with Delft University of Technology and 4,203 articles co-authored by scientists affiliated with Naturalis Biodiversity Center. This resulted in a dataset containing the metadata of 93,716 scientific papers. The SciVal dataset was retrieved on December 3rd, 2020.

The choice was made to limit the scientific articles to articles published from 2015 and later. This was done to keep the scientific article time range close to the time range of the news articles. This resulted in 61,919 different scientists.

The selection of data to include in the dataset

The metadata of each scientific article in the SciVal dataset contained 47 different fields. Not all information was deemed relevant for creating a scientist-journalist recommender system. Only information related to the authors, the publication date and the performance of the scientific article were included in the dataset (see table 9.1). It can be expected that there is a relationship between the performance of the articles that a scientist co-authored and the likelihood that s(he) is interviewed by a newspaper journalist. To test this, the citations of the scientific article and the Source Normalised Impact per Paper (SNIP) are included in the dataset.

A scientific article's citations are deemed self-explanatory, but further explanation on SNIP is provided. SNIP is a ranking method of scientific journals that measures the prestige and influence of the journal. If a scientist publishes in prestigious journals, they are likely well-regarded in their field. The SNIP was chosen because the scientific articles in the SciVal dataset cover many scientific areas. Waltman et al. (p.1, 2013)[59] describe the idea behind SNIP as "The idea of the source normalised approach is to correct for differences in citation practices between scientific fields".

9.1.3. Employee lists of Delft University of Technology and Naturalis Biodiversity Center

This project is done alongside a Communication Design for Innovation (CDI) research. The CDI research is a case study with communication professionals from Delft University of Technology and Naturalis Biodiversity Center. The CDI research investigates the effect of participation in the development of the recommender system on the mental model of digital innovation of the communication professionals. This means the recommender system should only consider scientists that are actively employed by Delft University of Technology and Naturalis Biodiversity Center. For this reason, employee lists of Delft University of Technology and Naturalis Biodiversity Center were included in the dataset. Both Delft University of Technology and Naturalis Biodiversity Center provided a list of the scientists they employ. Delft University of Technology provided a list of 6,980 scientists dated September 1st, 2020. Naturalis Biodiversity Center provided a list of 112 scientists dated June 2019.

9.1.4. Overview of data sources

In summary, the primary data sources that are used to create the scientist network, and the journalist network are the following:

1. The DpgMedia2019 dataset with Dutch newspaper articles, enriched to contain the journalists
2. SciVal scientific article metadata of papers published by researchers affiliated with Delft University of Technology and Naturalis Biodiversity Center between 2015 and 2020
3. Employee lists of active researchers at Delft University of Technology and Naturalis Biodiversity center

Table 9.1 shows for each dataset what information was used to construct the collaboration network and how many records it contains.

[2]Scopus is a database of scientific articles. For more information on Scopus and SciVal, see `https://www.scopus.com/` and `https://scival.com/`.

Table 9.1: The information used in the construction of the collaboration network per data source.

Dataset	DpgMedia2019	SciVal	Employee list Delft University of Technology	Employee list Naturalis Biodiversity Center
Fields	Article text Journalist	Authors Year Citations SNIP	First name Last name Name variant	First name Surname Initials Internal name
Number of records	102 457	93 716	6 979	112

9.2. Data integration

The SciVal dataset allows the creation of the scientist network by creating a node for each author of a paper in the SciVal dataset and connecting two nodes if the scientists are author of the same article. The enriched DpgMedia2019 dataset allows the creation of the journalist network by creating a node for each journalist and connecting two nodes if the journalists are author of the same newspaper article.

However, to have the scientist network only represent scientists currently employed by one of the research institutes, the scientists in the network need to be compared with the scientists on the employee lists. Furthermore, to create the scientist-journalist network, the scientists must be matched to newspaper articles' names. Figure 9.1 shows an example of this matching problem.

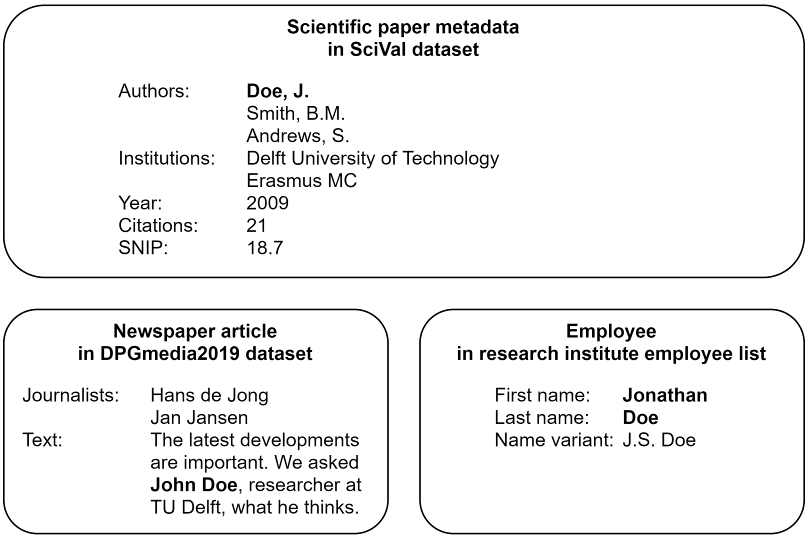


Figure 9.1: An example of the entity resolution problem.

9.2.1. Entity resolution between the employee lists and article authors

The scientific articles in the SciVal dataset use different name formats. The most common format was *Surname, initials-separated-by-period*. The scientists contained in the employee lists did not always contain multiple initials. For this reason, the matching was done on *Surname, initial-1.* This resulted in 4,642 of the 6,980 scientists of Delft University of Technology and 94 of the 112 scientists of Naturalis Biodiversity Center being matched to scientific papers.

For each scientist, the data from the co-authored scientific papers was aggregated. Three measures of scientific performance were included: The average citations of all co-authored scientific articles, the average SNIP score of all co-authored scientific papers, and the total number of publications.

9.2.2. Entity resolution of scientists in newspaper articles

If a scientist was mentioned in a newspaper article, it would be by their full name, i.e. *First-name Surname*. After the matching process above, for each scientist, only *Surname, initial-1.* were available. An issue with matching surname and first initial in newspaper articles is that the combination of surname and first initial is not unique. For example, one of the scientists at Delft University of Technology had the same surname and first initial as a famous Dutch politician. To select articles that are related to scientists, scientist matching was only done on newspaper articles related to the research institutions. The matching was limited to newspaper articles that contain one of the research keywords and one of the institutional keywords shown in table 9.2.

Table 9.2: To reduce false positives, only newspaper articles that contained one of the research keywords and one of the institutional keywords were considered for scientist matching.

Research keywords	Institutional keywords
Wetenschapper wetenschapper Onderzoeker onderzoeker Universiteit universiteit	Technische Universiteit Delft TU Delft Naturalis

A scientist was matched with a newspaper article if the surname appeared in the paper and the word before the surname was capitalised and started with the initial of the scientist. If the surname was the second word in the sentence, it was excluded. This is because all first words are capitalised, which could result in erroneous matches. The matching process was done using the Python *re* module for regular expression operations. The regular expression used is shown below.

$$(? <!\backslash.\backslash s)\ firstinitial\ [a-z]^{*}\backslash s\ surname$$

9.3. Construction of the multilayer collaboration network

Three collaboration networks were created; a scientist collaboration network, a journalist collaboration network and a scientist-journalist collaboration network.

The construction of these networks uses the number of collaborations between a scientist and a scientist, a journalist and a journalist, and a scientist and a journalist. Let S be the set of all scientists and J the set of all journalists. Let u be a scientist or journalist ($u \in S \cup J$) and let v also be a scientist or journalist ($v \in S \cup J$). The function $collab : (S \cup J) \times (S \cup J) \rightarrow \mathbb{N}$ will be used to refer to the number of collaborations between u and v. It is defined as shown in equation 9.1.

$$collab(u,v) = collab(v,u) = \begin{cases} \#co\text{-}authored\ scientific\ papers & u,\ v \in S \\ \#co\text{-}authored\ newspaper\ articles & u,\ v \in J \\ \#newspaper\ articles\ written\ by\ v\ mentioning\ u & u \in S,\ v \in J \end{cases} \tag{9.1}$$

9.3.1. The scientist and journalist collaboration networks

The scientist collaboration network was constructed as an undirected weighted graph $G_S = (V_S, E_S, w_S)$ as shown below in definition 9.2. The nodes are the set S of all scientists. Two scientist nodes are connected if the number of collaborations is greater than zero. The weight function $w_S : E_S \rightarrow \mathbb{N}$ of each edge is the number of collaborations

$$\begin{aligned} G_S &= (V_S, E_S, w_S) \\ V_S &= S \\ E_S &= \{(s_1, s_2) \mid s_1, s_2 \in S,\ collab(s_1, s_2) > 0\} \\ w_S(s_1, s_2) &= collab(s_1, s_2) \end{aligned} \tag{9.2}$$

The journalist collaboration network is similarly constructed as an undirected weighted graph $G_J = (V_J, E_j, w_j)$ as shown in definition 9.3. The nodes are the set of all journalists J. Two journalist nodes are connected if they have collaborated, and the edge weight is the number of collaborations. The formal definition is shown below.

$$
\begin{aligned}
G_J &= (V_J, E_J, w_J) \\
V_J &= J \\
E_J &= \{(j_1, j_2) \mid j_1, j_2 \in J,\ collab(j_1, j_2) > 0\} \\
w_J(j_1, j_2) &= collab(j_1, j_2)
\end{aligned}
\tag{9.3}
$$

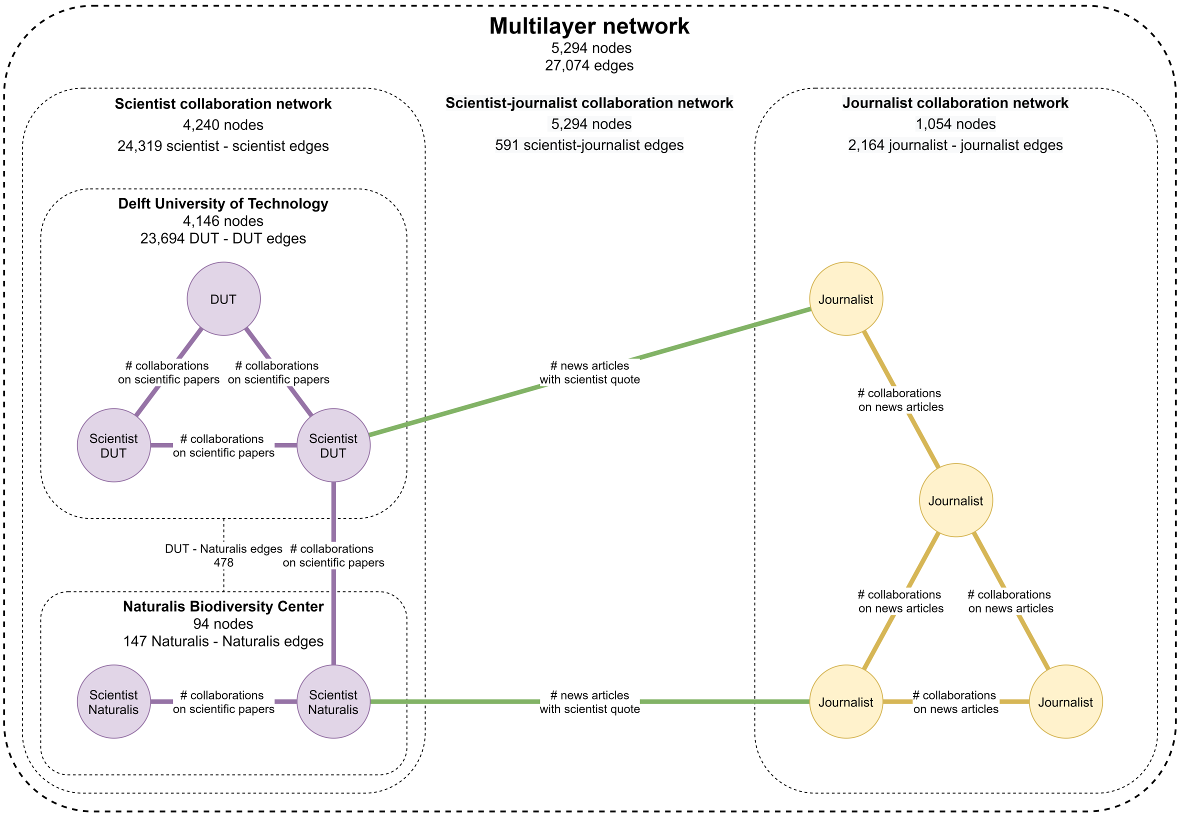


Figure 9.2: A schematic representation of the scientist collaboration network, the journalist collaboration network and the scientist-journalist collaboration network, and how they are combined into the multilayer network.

9.3.2. The scientist-journalist collaboration network

The scientist-journalist collaboration network was constructed as an undirected, weighted bipartite graph $G_{SJ} = (V_S, V_J, E_{SJ}, w_{SJ})$ as shown in definition 9.4. The set nodes of G_{SJ} is the union of the set of scientists S and journalists J. Two nodes are connected if the number of collaborations is greater than 0. The edge weights are the number of collaborations. The formal definition is shown below.

$$
\begin{aligned}
G_{SJ} &= (V_S, V_J, E_{SJ}, w_{SJ}) \\
E_{SJ} &= \{(s, j) \mid s \in S, j \in J,\ collab(s, j) > 0\} \\
w_{SJ}(u, v) &= collab(u, v)
\end{aligned}
\tag{9.4}
$$

9.3.3. The multilayer collaboration network

The scientist collaboration network, journalist collaboration network and scientist-journalist collaboration network are combined into a multilayer network G_M, composed of the union of all the nodes and links in these three networks.

$$
\begin{aligned}
G_M &= (V_M, E_M, w_M) \\
V_M &= V_S \cup V_J \cup V_{SJ} \\
E_M &= E_S \cup E_J \cup E_{SJ} \\
w_M(u,v) &= \begin{cases} w_S(u,v) & u,v \in S \\ w_J(u,v) & u,v \in J \\ w_{SJ}(u,v) & u \in S, v \in J \end{cases}
\end{aligned}
\tag{9.5}
$$

This network contained one giant connected component of 5,294 nodes. There were 3,448 other components, which were all smaller than 18 nodes. The choice was made to keep the giant connected component and remove the nodes from all smaller components from the multilayer network and the component networks. This was because nodes outside giant components have limited network information to derive the recommendations.

The final networks are represented in figure 9.2. The final scientist set S consists of 4,146 scientists working at Delft University of Technology and 94 scientists working at Naturalis Biodiversity Center. The final journalist set J consists of 1,054 journalists.

10

Centrality metrics that identify nodes with scientist-journalist collaborations

This research aims to investigate how network features in the multilayer network, the scientist network and the journalist network can be used to create a scientist-journalist recommender system. In order to select network features to use in the design of the recommender system, this chapter analyses which centrality metrics can differentiate nodes with scientist-journalist collaborations from nodes without scientist-journalist collaborations. This is used to answer research question 2: *Which centrality metrics are the most useful for identifying nodes in the scientist and journalist networks with scientist-journalist collaborations?*

The criteria for the most useful centrality metrics

If a centrality metric shows a large difference between nodes with scientist-journalist collaborations and nodes without scientist-journalist collaborations, it is considered useful for identifying nodes with scientist-journalist collaborations. This is likely useful for identifying future scientist-journalist collaborations. For example, if nodes with scientist-journalist collaborations tend to have a higher degree than nodes without scientist-journalist collaborations, nodes with a high degree that currently do not have a scientist-journalist collaboration are likely candidates for a future scientist-journalist collaboration.

Another criterion for selecting a set of most useful metrics is that they should not be correlated. Correlated metrics capture much of the same information. For this reason, a subset of metrics that is not correlated covers more diverse information.

In summary, the criteria for the most useful centrality metrics in this context are:

1. The centrality metrics show a high difference between
   (a) scientists with journalist collaborations $S_J \subseteq V_S$
   and scientists without journalist collaborations $S_{\not J} \subseteq V_S$ in the scientist network
   (b) journalists with scientist collaborations $J_S \subseteq V_J$
   and journalists without scientist collaborations $J_{\not S} \subseteq V_J$ in the journalist network
2. The centrality metrics are not correlated in G_S or G_J.

A note on calculating the centrality metrics in the scientist and journalist networks

The centrality metrics are used to identify nodes in the scientist and journalist networks instead of the multilayer network. This is because it provides information on what types of scientists tend to have journalist collaborations and what kinds of journalists tend to have scientist collaborations. This information will be used in the next chapter as input for the recommender system. Many centrality metrics will differ in the scientist and journalist networks compared to the multilayer network. For example, a journalist with scientist connections might have a high betweenness centrality in the multilayer network because it connects the scientist and journalist networks, even if that journalist has a low betweenness centrality in the journalist network. In this case, betweenness would be unsuitable as a metric to identify journalists that are likely to form future connections. For this reason, the centrality metrics are calculated based on the scientist and journalist networks.

The statistical tests for the difference between groups and correlation between metrics
To test whether there is a difference between S_J and $S_{\not J}$ (and J_S and $J_{\not S}$) for a metric, the two-sample Kolmogorov-Smirnov (K-S) test was used. The K-S test estimates the likelihood of two samples being drawn from the same distribution (Massey, 1951)[32]. It is nonparametric, which means it does not depend on any prior assumption of the distribution of the samples. This is an advantage, given the lack of prior knowledge on the various metrics compared in this section. The K-S test compares the empirical cumulative distribution function (CDF) of each sample. The *statistic* D of the K-S test is the maximum distance between the CDFs. The larger the statistic, the more distinct the two samples are. The *p-value* of the K-S test is the probability that the statistic is observed if the two samples were drawn from the same distribution. Smaller values of p indicate a more significant difference between the two samples.

The correlation between the metrics is evaluated using Pearson's correlation coefficient r. Pearson's r measures the extent to which two variables are linearly correlated.

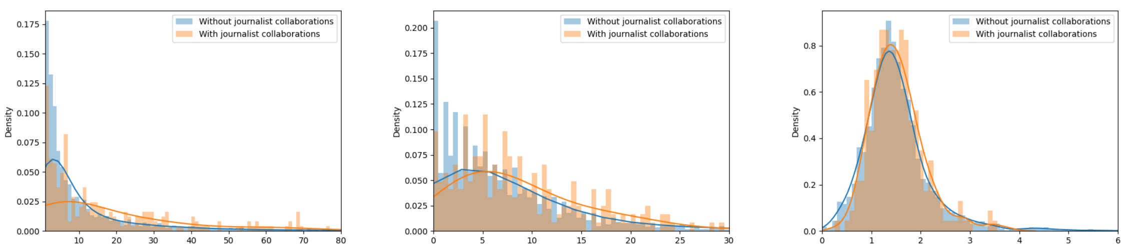
10.1. The difference in scientific performance metrics

In addition to the network features, the data contains information on the scientific performance of each scientist node. While these measures are not properties of the scientist collaboration network, they provide insight into drivers behind scientist-journalist collaborations. Three measures are considered, each of which captures a particular perspective of a scientist's performance.

Number of publications. The number of publications of scientist s is the number of articles that are contained in the SciVal dataset that s co-authored. Note that this is close to the weighted degree of a scientist, i.e. the total number of collaborations. However, the total number of publications also includes single-author publications.

Average citations. The average citation count of all papers in the SciVal dataset that a scientist s co-authored.

Average SNIP. As explained in section 9.1.2, the Source Normalised Impact Per Paper (SNIP) is a ranking metric for the prestige of scientific journals. The average SNIP of a scientist is the average of the SNIP scores of the journals a scientist has published in.



(a) Number of publications (b) Average citations (c) Average SNIP

Figure 10.1: The distribution of S_J (orange curve) and $S_{\not J}$ (blue curve) of the scientific performance metrics.

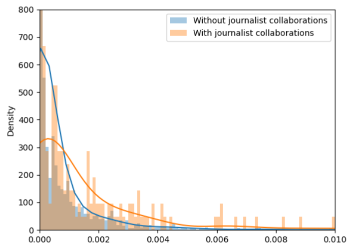
Table 10.1 shows that the number of publications and average citations show a significant difference between S_J and $S_{\not J}$. In contrast, average SNIP shows no significant difference between S_J and $S_{\not J}$. This shows that scientists with journalist collaborations have published more papers, and their papers are more well-received by their peers, but they do not publish in more prestigious journals.

Table 10.1: The K-S statistic D and significance P of the difference of the difference in distribution between S_J and $S_{\not J}$ of the scientific performance metrics.

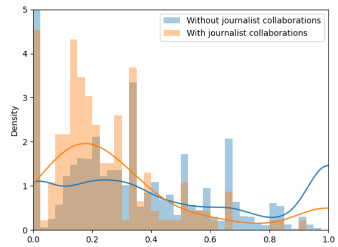
Metric	D	p
Number of publications	0.19	1.8E-8
Average citations	0.16	9.70E-06
Average SNIP	0.06	0.330

10.2. The difference in centrality metrics between scientists with and without journalist collaborations

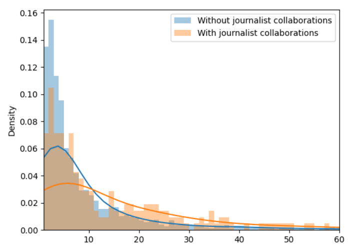
To find the most useful centrality metrics for identifying scientists with journalist collaborations, the difference in distributions of S_J and $S_{\not J}$ is analysed for each centrality metric (see figure 10.2). The centrality metrics are calculated in the scientist network G_S. Table 10.2 shows the K-S statistic between S_J and $S_{\not J}$, as well as the correlations between the centrality metrics in G_S. For comparison, the correlations with the scientific success metrics are also included.



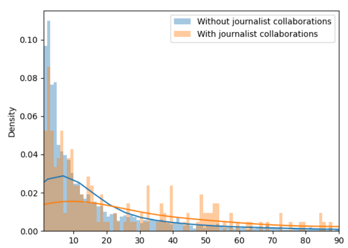
(a) Betweenness centrality



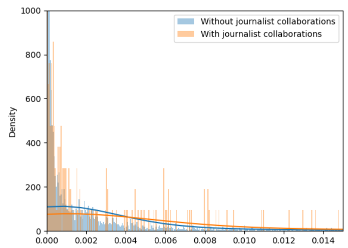
(b) Clustering coefficient



(c) Degree



(d) weighted degree



(e) Eigenvector centrality

Figure 10.2: The distribution of S_J (orange curve) and $S_{\not J}$ (blue curve) of the centrality metrics in the scientist network.

Table 10.2: The K-S statistic D and significance P of the difference of the difference in distribution between S_J and $S_{\not J}$ of the centrality metrics, and the correlation between the centrality metrics.

Metric	**K-S statistic**		**Pearson's r**						
	D	p	Betweenness	Clustering coefficient	Degree	Weighted degree	Eigenvector	Publications	Average citations
Betweenness centrality	0.26	1.20E-12	1	-0.21	0.87	0.82	0.73	0.7	0.03
Clustering coefficient	0.25	1.00E-11	-0.21	1	-0.18	-0.21	-0.1	-0.32	0.01
Degree	0.18	1.90E-06	0.87	-0.18	1	0.9	0.94	0.69	0.08
Weighted degree	0.17	8.90E-06	0.82	-0.21	0.9	1	0.78	0.85	0.05
Eigenvector centrality	0.13	0.003	0.73	-0.1	0.94	0.78	1	0.51	0.06

Table 10.2 shows a strong correlation between betweenness, degree, weighted degree and eigenvector centrality. It shows that the clustering coefficient negatively correlates with most of the other metrics.

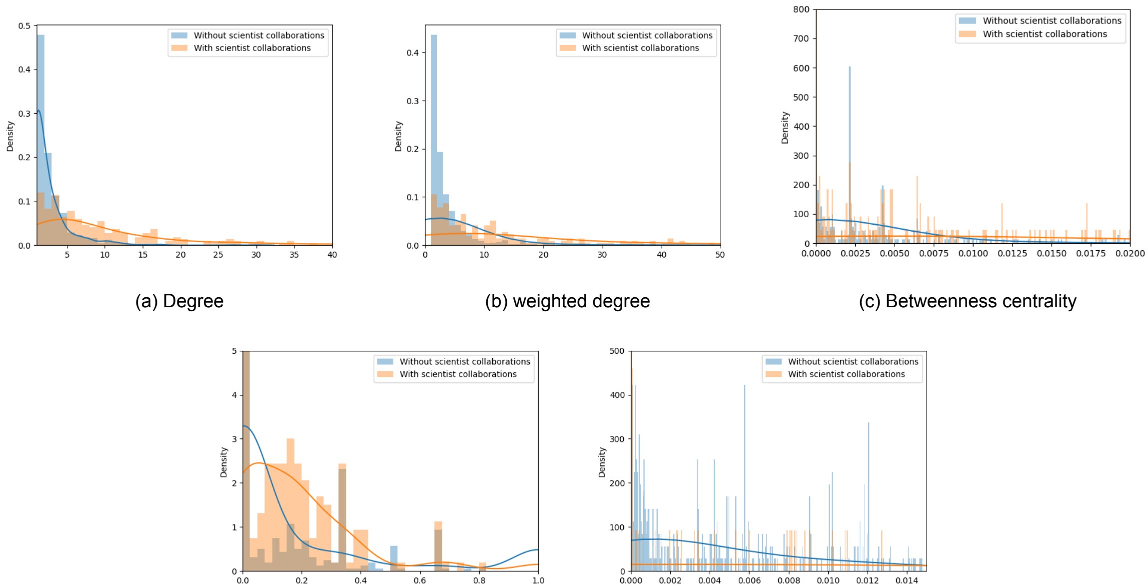
It is interesting that betweenness shows the largest and most significant difference of the correlated centrality metrics. Literature has identified betweenness to play an important role in scientific collaboration networks (Abbasi et al., 2012)[1]. Betweenness in social networks is often described in relation to information flowing through the network (Estradas et al., 2009)[19]. Nodes with a high betweenness connect different parts of the network. In a network in which information moves around, nodes with a high betweenness are important in controlling the flow of information from one part of the network to another. The significant difference in betweenness between S_J and $S_{\not J}$ implies that this control of information is an important signifier for journalist connections. Degree and weighted degree show a smaller difference between S_J and $S_{\not J}$ but still show a significant difference. They are nearly equal, which implies that the number of collaborations per scientist-scientist connection does not strongly influence scientist-journalist collaborations.

The clustering coefficient is negatively correlated with all other metrics except for average citations and eigenvector centrality. It is common in most real networks that the clustering coefficient decreases with degree (Soffer and Vazquez, 2005)[50]. The more neighbours a node has, the smaller the chance all neighbours are connected.

Betweenness and clustering coefficient show the highest and most significant difference between S_J and $S_{\not J}$, and they are not correlated. For this reason, they are selected as the most useful centrality metrics for identifying scientists with journalist collaborations.

10.3. The difference in centrality metrics between journalists with and without scientist collaborations

To find the most useful centrality metrics for identifying journalists with scientist collaborations, the difference in distributions of J_S and $J_{\not S}$ is analysed for each centrality metric (see figure 10.3). The centrality metrics are calculated in the journalist network G_J. Table 10.3 shows the K-S statistic between J_S and $J_{\not S}$, as well as the correlations between the centrality metrics in G_J.



(a) Degree

(b) weighted degree

(c) Betweenness centrality

(d) Clustering coefficient

(e) Eigenvector centrality

Figure 10.3: The distribution of J_S (orange curve) and $J_{\not S}$ (blue curve) of the centrality metrics in the journalist network.

Similar to the centrality metrics in the scientist network, betweenness, degree, weighted degree and eigenvector centrality show a strong correlation. In contrast, the clustering coefficient does not show a correlation with any of the other metrics.

Compared to the degrees of the scientist nodes in figure 10.2c, the degrees in the journalist network are generally lower. There are a lot of journalists j among $J_{\not S}$ with only a single connection ($d_j = 1$). Journalists with scientist connections have higher degrees. This same difference between journalists in J_S and $J_{\not S}$ is seen in weighted degree. Many journalists in $J_{\not S}$ have only a single collaboration, while the

Metric	K-S statistic		Pearson's r				
	D	p	Weighted degree	Degree	Betweenness	Clustering coefficient	Eigenvector
Weighted degree	0.49	7.50E-38	1	0.55	0.23	0.03	0.49
Degree	0.49	1.80E-37	0.55	1	0.46	0.06	0.87
Betweenness centrality	0.44	8.40E-30	0.23	0.46	1	-0.03	0.29
Clustering coefficient	0.38	8.20E-22	0.03	0.06	-0.03	1	0.05
Eigenvector centrality	0.25	5.70E-10	0.49	0.87	0.29	0.05	1

Table 10.3: The K-S statistic D and significance P of the difference in distribution between J_S and $J_{\not S}$ of the centrality metrics, and the correlation between the centrality metrics.

journalists in J_S have multiple collaborations. This means journalists with scientist collaborations tend to collaborate more often with other journalists and with more different journalists. Degree and weighted degree are correlated, but the correlation is not as strong as the correlation between scientists' degree and weighted degree. This indicates a bigger difference in the number of collaborations per journalist than in the number of collaborations per scientist.

The clustering coefficient of the journalist nodes is generally lower than the clustering coefficient of the scientist nodes in figure 10.2b. The scientist and journalist collaboration networks are co-authorship networks - i.e. if an article (scientific or newspaper) has five co-authors, the article is a 5-clique in the scientist or journalist network. However, scientific papers often have many co-authors, while newspaper articles often only have a few. This explains why the clustering coefficient is lower in general. Interestingly, the clustering coefficient of nodes in J_S is higher than the clustering coefficient of nodes in $J_{\not S}$. In comparison, the clustering coefficient of nodes in S_J is lower than the nodes in $S_{\not J}$. This means that scientists with journalist collaborations tend to be in positions where they connect different parts of the network. In contrast, journalists tend to be in neighbourhoods that are strongly connected. Notably, the clustering coefficient in the journalist network does not show a negative correlation with the degree. As mentioned in section 10.2, these are usually negatively correlated in the literature.

Given that the clustering coefficient is not correlated with any of the other metrics and shows a large difference between J_S and $J_{\not S}$, the clustering coefficient is included in the most useful metrics to identify journalists with scientist collaborations. Of the correlated metrics, weighted degree and degree show an equally big gap between J_S and $J_{\not S}$. They are correlated, so it was not deemed necessary to include both in the selection of the most useful metrics. Of these two, the degree shows the highest correlation with the other metrics. For this reason, it was decided degree contained the most useful information and was chosen to be included in the most useful metrics.

10.4. The most useful metrics to distinguish nodes with internetwork collaborations

Section 10.2 found that betweenness and clustering coefficient were the most useful to identify scientists nodes in the scientist network with scientist-journalist collaborations. Section 10.3 found that the clustering coefficient and both degree and weighted degree were the most useful to identify journalist nodes within the journalist network with scientist-journalist collaborations. It was decided to include the degree in the most useful metrics. For this reason, the overall most useful centrality metric combination is concluded to be *clustering coefficient*, *betweenness* and *degree*.

11

The network-based recommender system

This chapter answers research question 3: *How can network features in the multilayer network be used to design a network-based scientist-journalist recommender system?* Based on the centrality metrics selected in the previous chapter and the distance metrics,

Section 11.1 explains how the recommender system is designed based on a combination of the centrality metrics selected in the previous chapter and two distance metrics. Section 11.2 describes the evaluation method used to determine the system's accuracy. Section 11.3 shows how each metric individually performs. Lastly, the evaluation method is used to optimise the combination of metrics.

11.1. Recommender system design

The use-case of this recommender system is to provide potential matches for a given scientist or journalist. This is done through a similarity-based approach. As described in section 8.1, similarity-based approaches use network metrics to assign a score to each pair of unconnected nodes, where a higher score means a higher probability that a connection between these nodes will form. The potential node pairs are ranked based on this score.

For a given scientist node s, the score for each potential journalist j is based on the centrality metrics selected in chapter 10: *clustering coefficient* c_j, *betweenness* b_j and *degree* d_j. The centrality metrics for j are calculated in the journalist network G_J. Furthermore, the score also includes the distance metrics introduced in section 8.3.3: the *hop count* H_{sj} and the weighted distance W_{sj} between s and j in the multilayer network G_M. All metrics m (centrality and distance) are normalised, so the minimum is 0, and the maximum is 1 in order to calculate the score. The proposed general form of the score is as follows:

$$score_s(j) = w_b b_j + w_c c_j + w_d d_j - w_H H_{sj} - w_W W_{sj} \tag{11.1}$$

Where each weight $w \in [0, 1]$.

The journalists are ranked in descending order based on this score. The recommendations $R_s = j_1, j_2, ..., j_n$ for a scientist s consists of the (partially) ordered set of potential journalist matches.

Elaboration on the distance metrics

Apart from centrality metrics, the score also includes distance metrics. The previous chapter analysed which centrality metrics are the most useful for distinguishing nodes in the scientist and journalist networks with scientist-journalist collaborations from those without. No such comparison has been made for the distance metrics, as per definition, the hop count between a scientist and a journalist that have collaborated is 1, and the hop count between scientists and journalists that have not collaborated is greater than 1. This means that the nodes with scientist-journalist collaborations will automatically have a different distribution than those without. This makes a comparison such as done in chapter 10 not relevant for the distance metrics.

The reasoning behind using the distance metrics in the recommender system is as follows: If two scientists s_1 and s_2 have collaborated, they are likely working on similar research. If journalist j has mentioned scientist s_1 in one of their news articles, then it is likely that j is also interested in the research done by s_2 . For this reason, a scientist-journalist pair with a low hop count in the multilayer network is likely to collaborate. Hence, H_{sj} is a negative term in the score.

The length of a path $p = x_1, x_2, ..., x_n$ in the network is the sum of the weights of all the links: $\sum_{i=1}^{n-1} w(x_i, x_{i+1})$. The weighted distance between two nodes is the shortest length among all possible paths. If two scientists have collaborated more often, they are likely more similar in research interest. For this reason, the weighted distance in the multilayer network is calculated based on the inverted link weight $w_M^{-1}(u,v) = \frac{1}{w_M(u,v)}$. The weighted distance W_{sj} is also a negative term in the score calculation.

The calculation of the weights
Equation 11.1 introduced the general form of the score calculation. Each metric has a weight w between 0 and 1. To find the optimal weights, a grid search is done. This will be explained in the following sections. Section 11.2 introduces the evaluation metric that is used to evaluate the performance of the recommender system. Section 11.3 describes how the optimal weight combination is found that maximises this performance metric.

The score calculation for recommending scientists for a journalist
Equation 11.1 introduced the general form of the score calculation with the example of recommending journalist matches for a scientist. The general form of assigning a score to potential scientists s for a given journalist j is nearly identical. The only difference is that the clustering coefficient c_s of each scientist in the scientist network is a negative term. This is because scientists with a lower clustering coefficient are more likely to have scientist-journalist connections (see section 10.2).

$$score_j(s) = w_b b_s - w_c c_s + w_d d_s - w_H H_{sj} - w_W W_{sj} \quad (11.2)$$

11.2. Evaluation method

This section discusses how the performance of the recommender system is evaluated. Section 11.2.1 describes how a training set and test set were created. Section 11.2.2 describes the novel evaluation method. A baseline for this evaluation method is established in section **??** by applying it to a system with random recommendations.

11.2.1. The creation of the test set and training set

To evaluate the system, a training set and test set were created by randomly removing 30% of scientist-journalist collaborations from the multilayer network G_M. The remaining multilayer network is regarded as the training set $G_{M_{training}}$ and the removed collaborations as the test set T^*. In total, 251 scientist-journalist collaborations were removed from the 835 collaborations.

Scientists and journalists can collaborate multiple times. Hence, it is possible that a scientist and journalist that collaborate in the test set still have collaborations in the training set. It would be trivial to recommend that a node collaborates with a node with which it already has a link. Hence, the test set T is updated to contain all scientist-journalist collaborations from the original test set T^* that are not in the training set. Figure 11.1 schematically shows how the training set and test set are created from the multilayer network.

The test set T is a set of edges, i.e. every element in T consists of a pair of nodes. However, occasionally it is desirable to refer to a node that is any end node of an edge in the test set. For convenience and readability, the shorthand $u \in T$ will be used to refer to such a node.

Given a scientist (journalist) node $u \in T$, the objective of the recommender system is to identify the journalists (scientists) that collaborate with u in T among all journalists that have not been observed to collaborate with u in the test set.

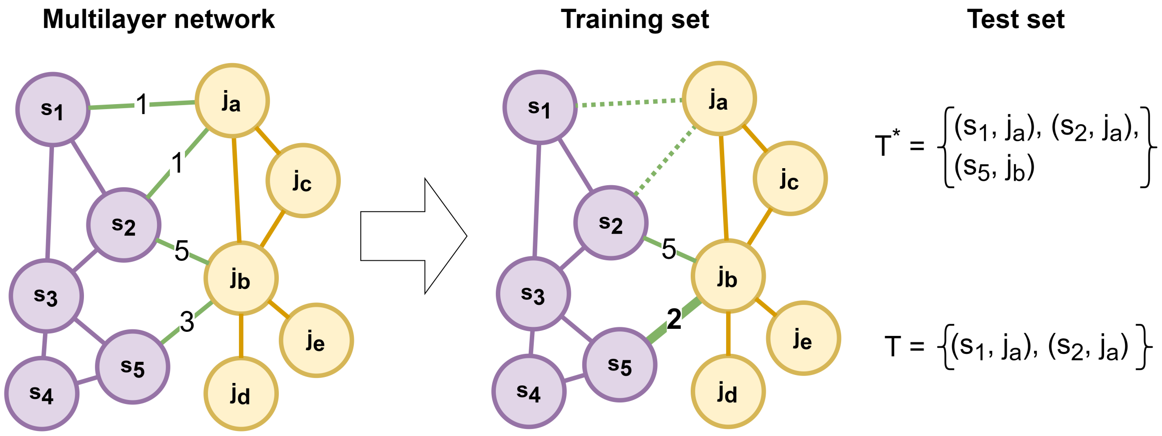


Figure 11.1: The training network is created by randomly removing 30% of scientist-journalist collaborations in the multilayer network G_M and moving them to the test set T^*. The updated test set T is created by removing all edges from T^* that are still in the training set.

11.2.2. The novel nodal link prediction evaluation method

This thesis introduces a novel evaluation method. This evaluation method is tailored to the nodal link prediction problem. This section will first explain why a novel evaluation method is developed. Next, it will explain how a nodal recommendation is evaluated. Lastly, it will describe how this is used to assess the performance of the system over the entire test set.

The issue with evaluation for the nodal link prediction problem

There are different ways to evaluate the performance of a recommender system. However, almost all rely on letting the recommender system predict links based on the training set and assessing how many links in the test set it accurately manages to predict.

The recommender system gives potential links a score, usually normalised between 0 and 1. A threshold score is set. All links that receive a score above this threshold are considered likely to form (positives), and all links that receive a score below this threshold are deemed unlikely to develop (negatives). The links in the test set that score above this threshold are true positives (TP). The links not in the test set but receiving a score above the threshold are false positives (FP). The links that are in the test set but receive a score below the threshold are false negatives (FN), and the links that are not in the test set and receive a score below the threshold are true negatives (TN).

Most evaluation methods for recommender systems use these positive and negative counts. For example, one of the most common evaluation methods is AUC-ROC, the area under the Receiver Operating Curve (Zou et al., 2011)[70]. AUC-ROC plots the true positive rate $\frac{TP}{TP+FN}$ against the false positive rate $\frac{FP}{FP+TN}$ for different thresholds.

For a given scientist (journalist), the nodal link prediction system only assigns a score to the potential journalists (scientists) for this scientist (journalist). The score is not an absolute indication of how likely a scientist-journalist connection is to form, but an indication of how likely this connection is compared to other scientist-journalist links for this scientist. For example, section 8.1 discusses the case of finding a suitable journalist for a PhD student. Generally, PhD students are not likely candidates for future scientist-journalist connections, so all journalists receive a low score. The recommendations are based on the relative score. Because the scores are only used relative to each other and only for a single node, any method that relies on a fixed threshold across the network is not applicable. Hence, a new evaluation method has been developed.

Evaluation for an individual node

For a node $u \in T$ (e.g. a scientist) in the test set, the journalists that have not collaborated with the scientist u in the training set can be divided into two sets:

- T_u: Those that are connected to u in the test set
- L_u: The rest, i.e. the set of nodes that are not connected to u in the test set

Two measures are defined. $P_{\leq u}(v)$ is the fraction of nodes in $l \in L_u$ that score lower or equal to v: $score_u(l) \leq score_u(v)$. $P_{<u}(v)$ is the fraction of nodes in L_u that score strictly lower than v: $score_u(l) < score_u(v)$.

A high $P_{<u}(v)$ means that the system correctly identified v as a more likely collaborator than other potential matches in L_u. A large gap between $P_{\leq u}(v)$ and $P_{<u}(v)$ shows that the system identified many other connections as equally likely as v. This is undesirable, as the system should be able to identify a good collaborator specifically. Hence, the gap between $P_{\leq u}(v)$ and $P_{<u}(v)$ should be as small as possible.

Evaluation over the entire test set

The $P_{\leq}$ and $P_{<}$ curves are introduced to evaluate the performance of the recommender system on the entire test set. Let $r(u,v)$ be the rank between 0 and 1 of link (u,v) in the test set according to $P_{<u}(v)$ in decreasing order. The $P_{<}$ curve plots $P_{<u}(v)$ against $r(u,v)$ and the $P_{\leq}$ curve plots $P_{\leq u}(v)$. Figure 11.2 shows an example.

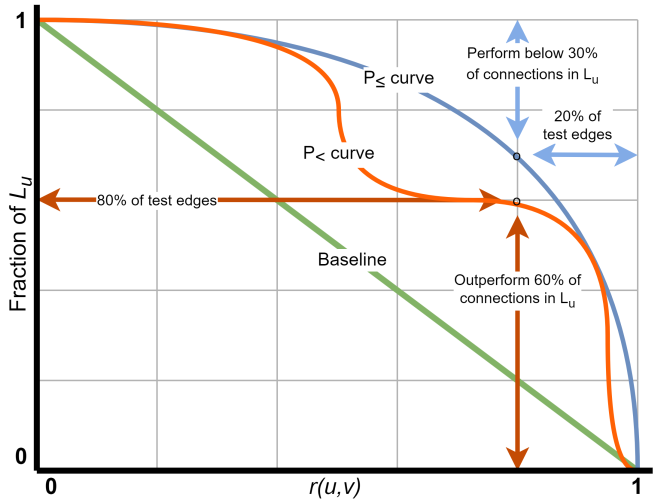


Figure 11.2: For each link (u,v) in the test set, $r(u,v)$ be the rank between 0 and 1 of link (u,v) according to $P_{<u}(v)$ in decreasing order, the $P_{<}$ curve displays $P_{<u}(v)$ and the $P_{\leq}$ curve displays $P_{\leq u}(v)$.

The $P_{<}$ curve shows the fraction of test edges that outperforms a specific fraction of potential connections in L_u. For example, in the $P_{<}$ curve shown in figure 11.2, the curve goes through the point (0.8, 0.6). This shows that 80% of test edges outperform 60% of potential connections in L_u. The $P_{\leq}$ curve shows the fraction of test edges that perform better or equal to a certain fraction of recommendations. It is an upper bound on performance; it shows how many other recommendations are strictly higher. For example, $P_{\leq}$ goes through the point (0.8, 0.7). This indicates that 20% of test edges perform below 30% of potential connections in L_u. $P_{\leq}$ and $P_{<}$ mostly overlap but do show some gaps. This indicates that the recommender system gives specific recommendations for most test edges but not all. Figure 11.2 also includes the performance baseline if for each node $u \in T$ in the test set, all potential matches v are assigned a random score. This means that 20% of items in the test set are in the top 20% of recommendations, 40% of items in the test set are in the top 40% of recommendations, etc. This results in a diagonal line in the figure.

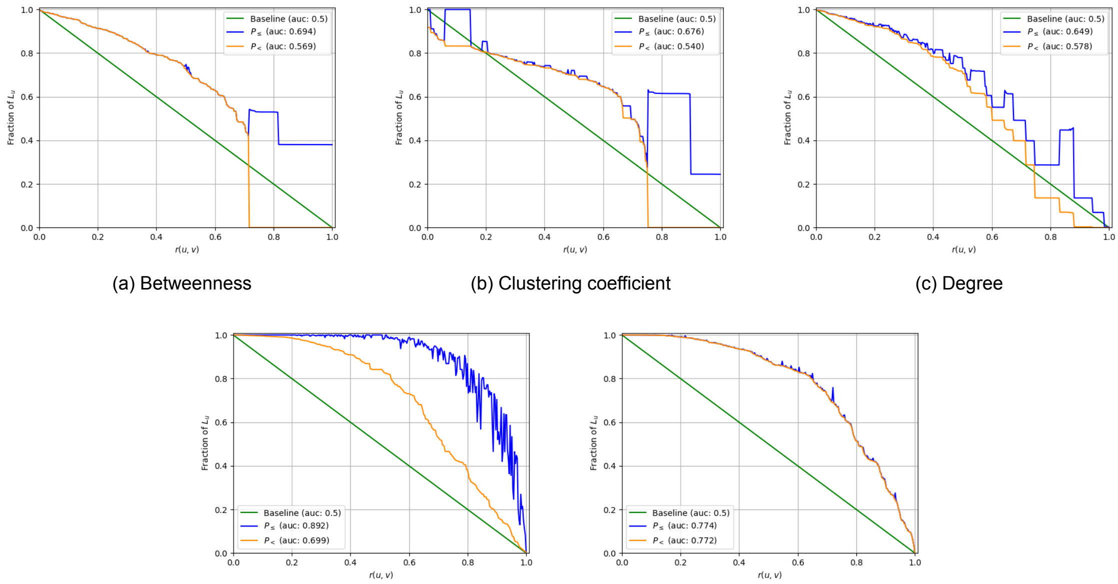
The system's overall performance on the test set is given by the area under the $P_{<}$ curve ($AUCP_{<}$). This represents how many test edges outperform the other recommendations.

11.3. Results and set of optimal weights

This section shows the performance of the recommender system. Section 11.3.1 shows the performance of the recommender system when the score is based on only a single metric. Next, in section 11.3.2, the performance of the recommender system is explored when the score is based on a combination of metrics. A grid search with different weights $w_c, w_b, w_d, w_H, w_W \in [0, 1]$ is done to find a set of weights that maximise the area under the $P_<$ curve.

11.3.1. Performance based on single-metric scoring

This section explores the performance of the recommender system when the score is based only on a single metric. Figure 11.3 shows each metric's $P_\leq$ and $P_<$ curves.



(a) Betweenness (b) Clustering coefficient (c) Degree

(d) Hop count (e) Weighted distance

Figure 11.3: For each metric the performance of the recommender system is shown when the scoring is based only on that metric.

For each metric, the performance of the recommender system outperforms the baseline. A comparison between the systems based on centrality metrics shows that the recommender system based on the degree (figure 11.3c) has both the highest $AUP_<$ and the lowest $AUP_\leq$. This is likely because it does not show a big gap between $P_<$ and $P_\leq$, in contrast to the systems based on betweenness and clustering coefficient. Figure 11.3a shows that for the system based on betweenness, up to roughly $r(u, v) = 0.7$, the $P_<$ and $P_\leq$ curves mostly overlap and outperform the baseline. After $r(u, v) = 0.7$, a large gap appears between $P_<$ and $P_\leq$ and $P_<$ drops to 0. For all (u, v) where $r(u, v) > 0.7$, the gap between $P_{<u}(v)$ and $P_{\leq u}(v)$ shows that the system scores a large portion of potential connections in L_u as equally likely as v. A possible explanation for this is that there are many nodes u in the scientist and journalist networks with $b_u = 0$. This is true for both nodes with scientist-journalist connections and nodes without scientist-journalist connections (see figure 10.2a and 10.3c). If $b_v = 0$, the system cannot distinguish it from any node $l \in L_u$ with $b_l = 0$. Hence, the large gap between $P_<$ and $P_\leq$ after $r(u, v) = 0.7$. A similar effect can be seen with the system's performance based on the clustering coefficient. Figure 10.2b and 10.3d show that many nodes u have a clustering coefficient of $c_u = 0$. This can explain the gap between $P_<$ and $P_\leq$ after $r(u, v) = 0.75$ in figure 11.3b.

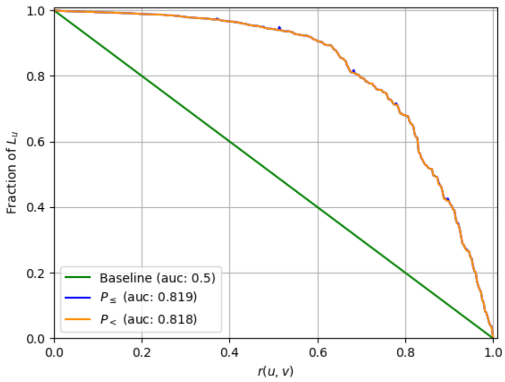
The systems based on hop count and weighted distance perform better than those based on the centrality metrics. The scoring based on hop count (figure 11.3d) shows the highest $AUP_\leq$, but also the largest gap between $P_<$ and $P_\leq$. This means that for many edges (u, v), the score of v based on hop count is high, but many recommendations in L_u receive an equally high score. This is not surprising as the diameter of the multilayer training network is 13, so for each node $u \in V_M$ in the training network, there are only 13 possible values that HS_{uv} can assign a potential match. This means that there is a higher likelihood that many potential matches are grouped together. The system based on weighted distance (figure 11.3e) has the highest $P_<$ curve of all systems. The $P_\leq$ curve is lower than the $P_\leq$ curve of hop count. There are almost no gaps between $P_<$ and $P_\leq$. This means that scoring based on hop count more often accurately assigns a high score to links in the test set than weighted distance does, but it also assigns many other potential matches the same score. The weighted distance more consistently assigns links in the test set a higher score than connections in L_u.

11.3.2. Performance of scoring based on sets of metrics

In order to find the set of weights $w_c, w_b, w_d, w_H, w_W \in [0,1]$ that yield the maximum $AUP_<$, a grid search was performed. The values that were searched over were $[0, 0.1, 0.5, 0.9]$. The grid search was repeated five times with a newly created test set and training set. Table 11.4a shows the five sets of weights with the highest average $AUP_<$. Figure 11.4b shows the $P_<$ and $P_\leq$ curves of the set of weights with the highest average $AUCP_\leq$ on the final test set.

	w_b	w_c	w_d	w_H	w_W	Average $AUCP_<$
1	0.5	0.0	0.0	0.9	0.1	0.8193
2	0.1	0.0	0.1	0.9	0.1	0.8190
3	0.1	0.0	0.0	0.9	0.1	0.8188
4	0.1	0.0	0.0	0.5	0.1	0.8185
5	0.1	0.1	0.0	0.9	0.5	0.8184

(a) The five sets of weights with the highest average $AUP_<$ over five test and training sets.



(b) The $P_<$ and $P_\leq$ curves for the system with the set of optimal weights on the final test set.

Figure 11.4: The five sets of weights with the highest average $AUP_<$ over five test and training sets and a plot of $P_<$ and $P_\leq$ of the optimal set of weights on the final iteration.

The optimal set of weights is with the weight for betweenness $w_b = 0.5$, the weight for hop count $w_H = 0.9$, and the weight for weighted distance $w_W = 0.1$. The clustering coefficient and degree weights are zero: $w_c = 0$ and $w_d = 0$. The weight for hop count is higher than all other weights in the top 5 weight sets.

It is interesting that individually the $AUP_<$ of weighted distance (see figure 11.3e) was higher than the $AUP_<$ of hop count (see figure 11.3d). Hop count had the highest $P_\leq$ curve of all metrics, which means it was the metric that most accurately gave true recommendations a high score. However, it had a large gap between $P_<$ and $P_\leq$, giving many false recommendations an equally good score. Betweenness is the only centrality metric consistently non-zero in the top five weight sets. Figure 11.3a shows the highest $AUP_\leq$ of all centrality metrics, while degree had a higher $AUP_<$. The combination of metrics reduces the chance that two edges are assigned the same score, which reduces the gap between $P_<$ and $P_\leq$. This means that in the recommender system based on a combination of metrics, a combination of metrics with a high $AUP_\leq$ seems to perform better than a combination of metrics with a high $AUP_<$.

12 Conclusions

This research investigated how network features in the multilayer network and the subnetworks can be used to predict links for a single node in the scientist-journalist collaboration network. Three research questions were answered in this research.

RQ1. How can the scientist, journalist, and scientist-journalist networks be combined to create a multilayer network?
The nodes from the scientist and journalist networks were matched with the nodes of the scientist-journalist collaboration network to form a multilayer network. This multilayer network is composed of the union of the nodes and links in these three networks.

RQ2. Which centrality metrics are the most useful for identifying nodes in the scientist and journalist networks with scientist-journalist collaborations?
It was found that the *clustering coefficient* and *betweenness* were the most useful in distinguishing scientists that have worked with journalists from scientists that have not worked with journalists. The most useful metrics to distinguish journalists with scientist connections from journalists without scientist connections were *clustering coefficient* and *degree*.

RQ3. How can network features in the multilayer network be used to design a network-based scientist-journalist recommender system?
A network-based recommender system was developed based on centrality metrics in the scientist and journalist networks and distance metrics in the multilayer network. The centrality metrics consisted of betweenness, clustering coefficient and degree. The distance metrics were hop count and weighted distance. A grid search was done with different weights for each metric to find a combination of weights with optimal performance.

13
Future work

This section describes several opportunities for future work.

13.1. Network

There are several limitations and potential improvements in the network creation.

The entity resolution method
A limitation of matching scientists to scientific articles based on surname and first initial is that scientists with the same surname and first initial are grouped together into one entity. A limitation of only matching scientists to newspaper articles that contained the keywords is that it might lead to false negatives, i.e. a scientist that did appear in a newspaper article not being matched because the article did not contain the keyword. A limitation of matching scientists to newspaper articles based on surname and first initial is that it can lead to false positives, i.e. a scientist is matched to a newspaper article where another person with the same surname and initial was interviewed. Due to a lack of further information on scientists and other interviewees, this is considered unavoidable noise.

The temporal aspect of the network
In this project, the multilayer network was considered as a static network. However, in reality, the network evolves over time. Future research could approach the multilayer network as a temporal network and include temporal network metrics in developing the recommender network.

The network metrics
Another improvement would be to analyse the influence of more network metrics. Currently, the centrality metrics degree, weighted degree, clustering coefficient, betweenness centrality and eigenvector centrality were considered, and two distance metrics; hop count and weighted distance. However, many other network metrics could be taken into consideration. This could provide more extensive insight into the factors that influence scientist-journalist collaborations.

13.2. The evaluation of the recommender system

There are several potential improvements to the evaluation of the recommender system.

Evaluation on other aspects than accuracy
The current evaluation method is aimed at accuracy. However, accuracy is not the only requirement of a recommender system. Other metrics have been proposed, such as novelty, coverage and diversity (Wu et al., 2012)[65]. Other methods can provide insight into the functioning of the system.

User tests
The development of the recommender system was done with the specific use case of communication professionals in mind. User tests should be performed to evaluate whether the system performs satisfactorily in this use case.

13.3. Potential research avenues in the multilayer network

To the best of this author's knowledge, no newspaper article co-authorship network has been researched. The dataset that was created not only contains a journalist co-authorship network, it also links this network to a scientist co-authorship network. During preliminary data analysis, many interesting avenues of research were found. This section will discuss some of the options that had potential but were not pursued in this project.

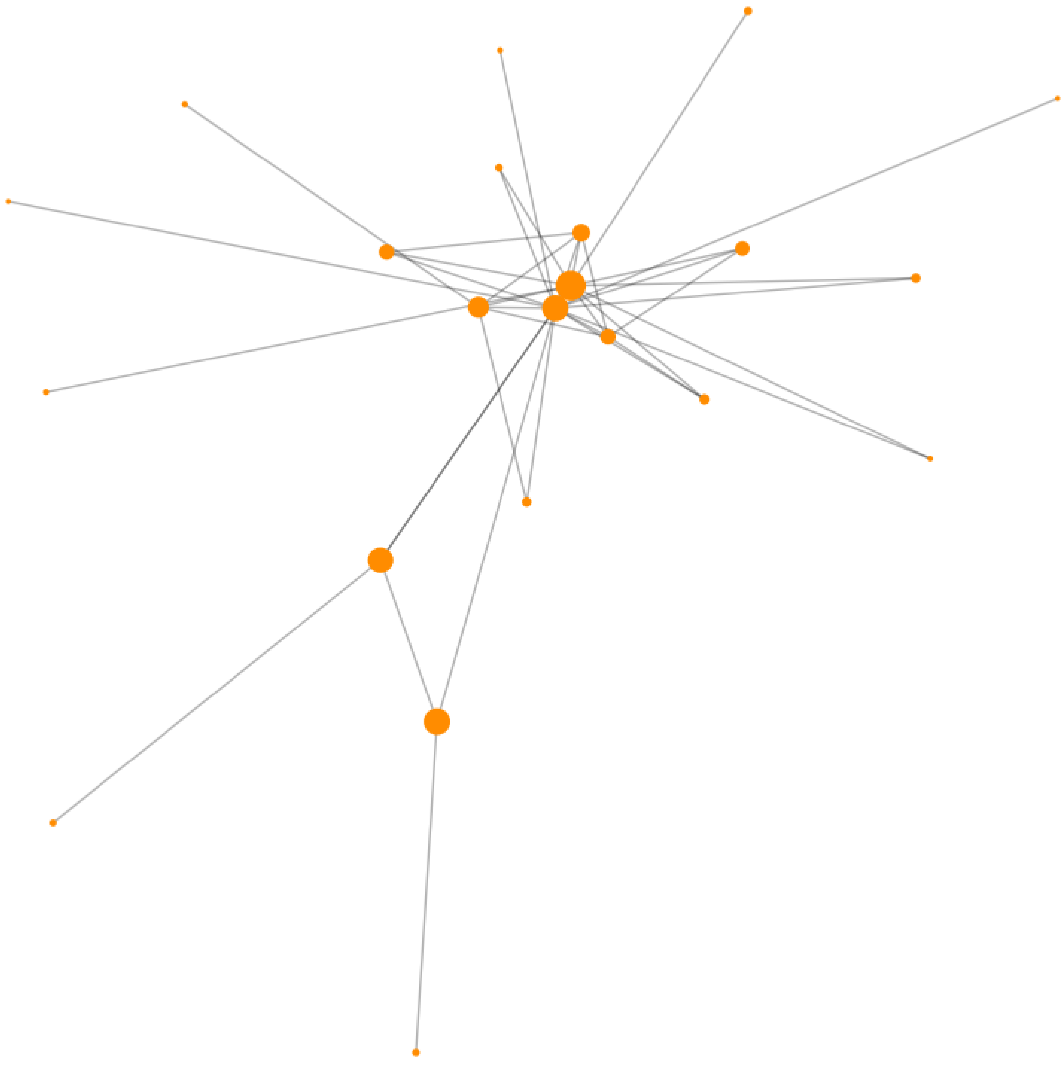
Include research topics in the analysis
An interesting aspect of the dataset that has yet to be mentioned is that the scientific article metadata also includes Scopus Topics and Scopus Topic Clusters. Scopus assigns each paper a topic and assigns each topic to a topic cluster. This allows the investigation of the relationship between scientific topics, communities in the network and scientist-journalist connections.

The relationship between scientist-journalist connections and scientific topics is interesting. Research in science journalism has shown that some scientific fields are more likely to be mentioned in a newspaper article than others. For example, Entradas and Bauer (2017)[18] performed a study on public outreach activities of Portuguese scientists within six different research areas. Of the scientists in their survey that were interviewed in a newspaper, the most were from the social sciences (22%), closely followed by Engineering and Technology (21%) and Medical and Health Sciences (20%). In contrast, only 9% of the scientists came from the Humanities. However, to the best of this author's knowledge, no research has been done in this field with a structured, data-driven approach. The data of the multilayer network might be useful for investigating this relationship.

Communities in the network
Another interesting research topic is the relationship between scientific topics and modularity-based communities in the scientific co-authorship network. It can be expected that there is some relationship between these two, as scientists tend to collaborate with scientists that work on similar topics. The third interesting research topic is the relationship between scientific communities and scientist-journalist relationships. If the network communities are related to scientific areas, the research into scientist-journalist collaborations and scientific topics can be extended by including scientific communities.

IV

Participation in a development process
Deliver



14 Deliver

The Deliver phase aims to turn the prototype into a usable tool. It was planned to do a co-design session with all participants. However, due to a last-minute cancellation, only the participants of one of the communication departments could be present. A separate co-design session was organised with the other communication department. For this reason, the stimulated recall was also done in individual sessions.

The stimulated recall sessions were recorded, transcribed and coded as described in section 3.2. Section 14.1 presents the results of the coding process. Next, section 14.2 presents the interpretation of the results to answer research question 3: What is the effect of participation in the development process of a digital tool on the mental model of digital innovation?

14.1. Results of stimulated recall sessions

This section presents the results of the stimulated recall sessions. The recording of the stimulated recall session with communication professionals of Delft University of Technology (DUT) was 57 minutes long. The recording of the stimulated recall session with communication professionals of Naturalis Biodiversity Center (NBC) was 48 minutes long. These stimulated recall sessions were transcribed verbatim and coded. Two iterations of thematic analysis were done to analyse the results.

14.1.1. Coding results

Structural coding was used with the codes *Before*, *During*, and *After*. In total, 48 quotations were found. Some quotations were coded twice. Of these excerpts, 18 were from the stimulated recall session with Delft University of Technology and 30 from the Stimulated Recall session of Naturalis Biodiversity Center. The number of quotations for each code is shown in figure 14.1.

Excerpts coded Before

The number of excerpts coded as 'Before' was 14. Of these excerpts, nine were from the stimulated recall session with Delft University of Technology and five from the Stimulated Recall session of Naturalis Biodiversity Center. Appendix table C.1 shows all excerpts coded 'Before'.

Excerpts coded During

The number of excerpts coded as 'During' were 34. Of these excerpts, ten were from the stimulated recall session with Delft University of Technology and 24 from the Stimulated Recall session of Naturalis Biodiversity Center. Appendix table C.2 shows all excerpts coded 'During'.

Excerpts coded After

The number of excerpts coded as 'After' was 34. Of these excerpts, ten were from the stimulated recall session with Delft University of Technology and 24 from the Stimulated Recall session of Naturalis Biodiversity Center. Appendix table C.3 shows all excerpts coded 'After'.

Figure 14.1: The number of quotations for the codes *Before*, *During*, and *After* and overlapping codes

14.2. Interpretation of results Change in the mental model of digital innovation

As described in section 3.2, two iterations of thematic analysis were done. In the first iteration, a mind map was made in which the coded excerpts were organised around the nodes *process*, *relationship* and *outcome*. Based on the initial mind map, a second mind map was made with for each node a 'before' state on the left and an 'after' state on the right to capture the change in the mental model. This process is described in section 14.2.1.

From this analysis, three themes emerged: *involvement in the process*, *complexity of the process* and *expectations of technology*. Section 14.2.2 describes their process of realising that they had to engage with the development process to get their desired results. Section 14.2.3 discusses how getting involved confronted them with the complexity of the process. Section 14.2.2 describes how their model of what can be expected of technological innovation changed.

14.2.1. Thematic analysis

As described in section 3.2.2, two iterations of thematic analysis were done. In the first iteration, a mind map was made in which the coded excerpts were organised around the nodes *process*, *relationship* and *outcome*. Based on the initial mind map, a second mind map was made with for each node a 'before' state on the left and an 'after' state on the right to capture the change in mental model. From this analysis, three themes emerged: *involvement in the process*, *complexity of the process* and *expectations of technology*.

First iteration thematic analysis

As described in section 2.1, the mental model of digital innovation of communication professionals can be defined as their mental representation of the innovation process, their relationship to it and the results it can bring. In order to find themes in the change in mental model, a mind map was made in which the coded excerpts were organised around the nodes *process*, *relationship* and *outcome*. This mind map is shown in figure 14.2. A spreadsheet with all coded excerpts was made, and sorted on code. For each excerpt coded 'before', keywords or keyphrases were written in the mind map in brown and organised around process, relationship and outcome. If keyphrases were related or built upon each other, they were placed in the same area and connected. The excerpt was skipped if a keyphrase was already present that fully encapsulated the excerpt. Next, keyphrases related to excerpts coded 'during' were added in blue, and finally, the excerpts coded 'after' were added in green.

Some excerpts had multiple codes. Often, these were excerpts that made a direct temporal compari-

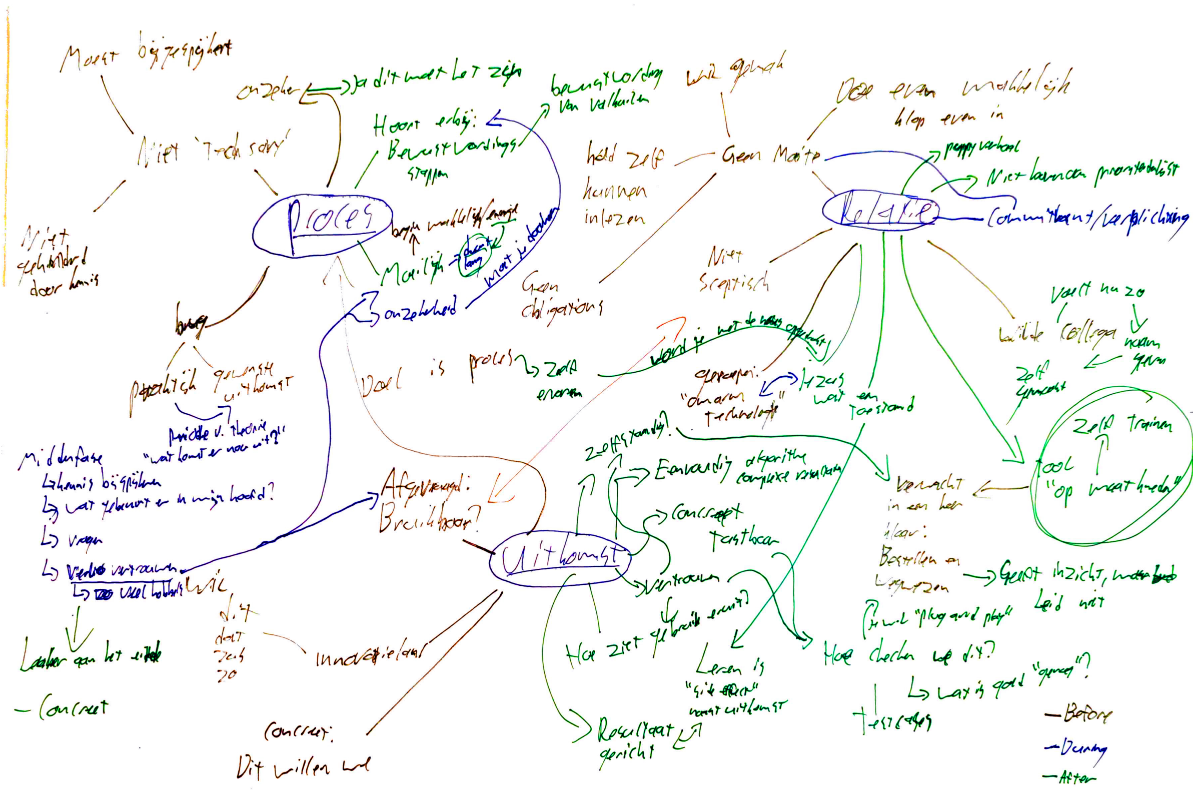


Figure 14.2: First iteration of thematic mind map after coding. Items related to the code *before* are written in brown, items related to the code *during* are written in blue, and related to the code *after* are written in green.

son. For example, the following quote is coded both 'before' and 'after'.

> It started with 'we are going to build an algorithm as a colleague' and apparently this is the point, I understand looking back at it, where I was like 'yes, this is a real colleague'.
>
> *Example of excerpt coded both 'before' and 'after' - P1, Stimulated Recall Delft University of Technology [45:00]*

If an excerpt had multiple codes, keyphrases were added in both colours or existing keyphrases were connected.

Second iteration thematic analysis

Based on the initial mind map, a second mind map was made to cluster the changes in the mental model around themes. This mind map is shown in figure 14.3. For each aspect of mental models - process, relationship and outcome - a node summarising 'before' was drawn on the left and a node summarising 'after' was drawn on the right. In the middle, a phrase representing the change was written in an arrow connecting the two. Around the nodes, keyphrases from the first mind map or from excerpts were written that support the node.

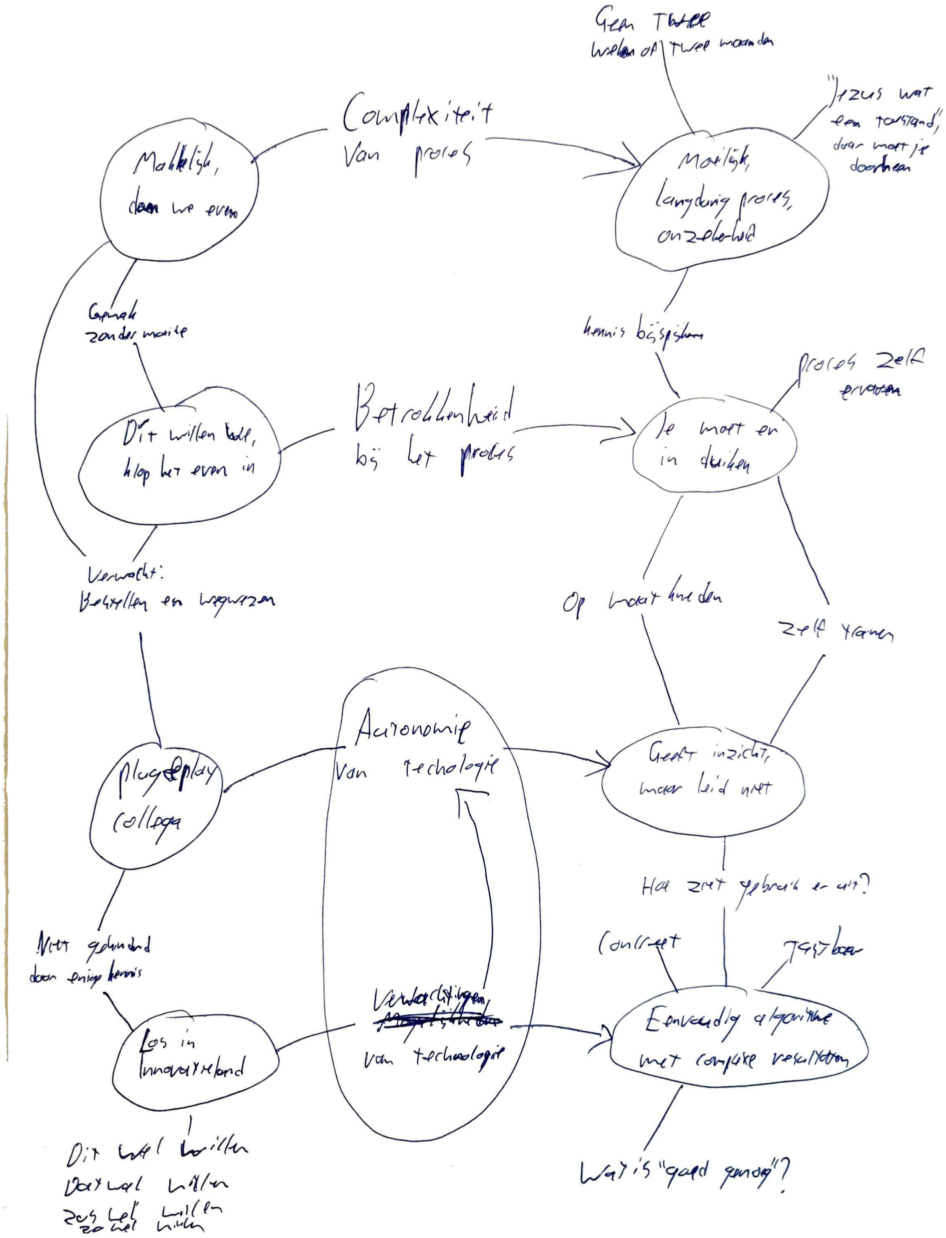


Figure 14.3: Second iteration of thematic mind map after coding. Three themes emerged from the thematic analysis: *Complexity of the process*, *involvement in the process*, *autonomy of technology* and *expectations of technology*.

14.2.2. Involvement in the process

The communication professionals were surprised by how much they had to get involved in the process.

On the one hand, there was a realisation from the beginning that they had to dive in and engage with the development process. This was part of the initial scope and part of their initial request. As seen in section 4.2.2 on the learning goals, they explicitly wanted to participate in the process and learn from it.

> It is, of course, nice when it works, because then we can do something with it – but that was not the point, it was mainly about the process, experiencing and seeing and doing, what comes out and how that works. *P3, Stimulated Recall Naturalis Biodiversity Center [4:00]*

However, there was a lack of understanding of what that engagement would entail. There was still an expectation of just handing in a list of requirements and getting a digital tool without having to engage with the complex technological reality.

> In the beginning of the project we had quite a tangible idea of what we wanted [...] Then it gets complex and we say 'just do it, just type in the code'.
> *P2, Stimulated Recall Delft University of Technology [39:00]*

> We are so used to just ordering something and you put it on and walk away with it or you start using it, but that is not the case with this.
> *P3, Stimulated Recall Naturalis Biodiversity Center [12:00]*

The process of realising what involvement in the process meant was confronting and occasionally frustrating to the communication professionals. However, they also reflect on it as a necessary step in understanding the complexity of a digital innovation process.

> It's almost 'in your face', because I shouted on a podium 'guys, we must have the courage to embrace technology as communication professionals, you have to dive in'. Well then I had to do that and I thought 'Gee what a hassle'. *P1, Stimulated Recall Delft University of Technology [41:00]*

> But that you also realise that it is a necessary part of it, those kinds of awareness steps to get there. *P2, Stimulated Recall Delft University of Technology [41:00]*

They reflect that the process of developing the digital colleague requires involvement. During the process, the analogy of a puppy was often used. A puppy needs to be trained. Getting a puppy to do what you want it to requires effort and its performance increases slowly over time.

> Then you see that it is quite a process to get that colleague right and then you also have to train her properly. Then it's that whole puppy story of yours again, you have to train her and teach her everything. *P3, Stimulated Recall Naturalis Biodiversity Center [4:00]*

One participant mentioned as one of the main takeaways they feel their colleagues should know that digital tools need to be tailored before you can get the benefit of it.

> You are kneading it to size. I think that it might actually be common with these kinds of tools and maybe that is an insight that is really important for my colleagues to know.
> *P3, Stimulated Recall Naturalis Biodiversity Center [12:00]*

14.2.3. Complexity of the process

Participants in both the Delft University of Technology session and the Naturalis Biodiversity Center session reflected that they underestimated the process's complexity, duration and uncertainty. When they started getting involved in the process, they were confronted with the complexity of innovation processes.

They describe starting with enthusiasm and energy, brainstorming about the final product and how it would be to have the custom-made digital tool. However, they lacked knowledge of digital technology or what the innovation process would entail.

> Looking back at the path we have walked, we had to become a bit more 'tech savvy' first [...] We wanted it all without too much effort, we just wanted convenience next to us. That was

a bit too simplistic way of looking at it. *P1, Stimulated Recall Delft University of Technology [21:00]*

Then you enter the middle phase in which you first have to brush up on knowledge, then you have to see what is happening in your head, then there are a lot of questions and then you slowly lose confidence. You think 'well guys, this has so many hurdles, am I really going to take them all and if I'm going to take them all is it worth it? But you have to take them, otherwise you can never come here. *P1, Stimulated Recall Delft University of Technology [39:00]*

It's not something that's done in two weeks or two months. No, it really is a lengthy process [...] In the beginning I found it easier to make time for it, because then you have that energy, but now you have to keep going for a long time.
P3, Stimulated Recall Naturalis Biodiversity Center [32:00]

Participants from both communication departments describe being confronted with the fact that the process takes a lot of time and energy and that it is not always certain what the outcome will be and if it will work. One participant describes being demotivated and losing confidence but also notes that it's something you have to go through to learn and get results.

14.2.4. Expectations of technology

Initially, the communication professionals had unrealistically high expectations of what they wanted the digital tool to do (see section 1.2.3). The idea of a digital colleague and a custom-made digital tool inspired them. They were swept up in thinking about everything they would want without knowledge of the technical reality to slow them down. One of the participants referred to it as 'being in innovation land'.

I literally see it as a bridge that we have had to walk, with practice on one side and the idea on the other, so to speak. The first time we went *bam* to the other side of the bridge. We were in innovation land and we went all out. We would like this, would like that, algorithm as a colleague, job description, all without any kind of obligations to it and unencumbered by any kind of knowledge. *P1, Stimulated Recall Delft University of Technology [39:00]*

They reflect that they first needed to be grounded in that technical reality before being able to realistically estimate what is and isn't possible. One of the participants referred to this as getting 'tech-savvy'.

We were inspired by the idea of a concrete, useful tool. Based on this, the question was asked what that tool could do and this project started. We turned out to have too little understanding of the digital world to get started with it right away, so we first had to gain new knowledge. *Debrief meeting November 19th, 2020*

The communication professionals started with a mental separation between their reality and the world of technology. They saw technology as a mysterious black box they could interact with but not truly engage with. This led to the expectation of their involvement in the process as purely coming up with requirements rather than participating in the design. This also led to an expectation of the digital tool to function entirely independently. They expected an algorithm that would work as an autonomous colleague that could largely do what they do, but better and faster.

Already early on in the process, they started to realise that you need to engage with technology. In the analogy of the puppy, a puppy needs initial training and continuous input and feedback.

In the beginning, we saw an algorithm as something that is complex, but separate from our world. By now, we no longer see it as a mysterious black box that makes decisions autonomously and transfers the final result to us. You have to work with the algorithm to get something useful out of it. We see AI as a super puppy that is always sharp and has an infinite memory, but starts out as a blank slate. By consciously taking her on puppy training and giving her our rules and analysis, we can teach her the desired behaviour.
Debrief meeting June 22, 2020

During a reflection on the progress so far, their expectations and the lessons learned, they reflected that a digital colleague is more of a specialist than a generalist (see figure 14.4). It is extremely good at doing one thing. It can offer support with certain activities and do this so well it becomes a basis for

Figure 14.4: The Miro board from the progress session on April 15th, 2021. Participants: P1, P3 and facilitator. The goal of this session was to recap what our original goals were, where we stood at that moment, what we had learned already and what we wanted from the final part of the process.

that activity. One participant reflected that it can provide insights and do things that help them make corrections or provoke thoughts, but it's not that it gets going, and that's it.

> The algorithm does not immediately guide us. It can give us surprising insights and do remarkable things, so that we make corrections or think 'hey this is interesting', but it is not that you send me something now, say a link, and we get going and that's it.
>
> *P3, Stimulated Recall Naturalis Biodiversity Center [10:00]*

14.3. The effect of participation on the mental model of digital innovation

Initially, the mental model of digital innovation and digital technology of the communication departments was a *black box*: Something that takes in an input and delivers back an output without requiring any knowledge on - or engagement with - the inner workings. It was seen as something external to themselves. A process where they could hand in a list of requirements and walk away with a digital tool, which in turn would take an input and deliver the desired output.

This has shifted to a mental model of digital innovation and digital technology as a 'super puppy'. It can do remarkable things, but it requires effort to tailor and train. It has to be interacted with, both during the process and afterwards, to get the desired results.

15

Discussion

This chapter discusses and reflects on the research process and future work. Section 15.1 discusses the validity and reliability of this research. Section 15.2 discusses future work. Section 15.3 is a reflection on the personal learning process.

15.1. Validity and reliability of the research

Reliability can be split into internal and external (Bryman, 2012)[37]. External reliability is the replicability of a study, which is difficult in qualitative research as ”it is impossible to ’freeze’ a social setting and the circumstances of an initial study to make it replicable in the sense in which the term is usually employed.” (Bryman 2012, p.390). However, a researcher replicating a qualitative study can adopt a similar social role as the original researcher. It is impossible to replicate the exact combination of participants and setting, i.e. communication professionals with different backgrounds coming from different institutions and a researcher with a background in both Communication Design for Innovation and Computer Science, who develop a scientist-journalist recommender system together. This study is replicable to the extent that future researchers can initiate a digital innovation project in which research subjects participate. Furthermore, the simplicity of the codes - *before*, *during* and *after* - makes it possible to apply them to any reflection on a process that has happened over a timespan. By doing multiple iterations of thematic analysis, the resulting themes - involvement in the process, the complexity of the process and expectations of technology - are at a level of abstraction where they are not specific to the communication profession or this particular innovation project. Future research can take a similar approach to thematic analysis of changes in mental models and test for these themes.

This is directly connected to the research’s external validity, i.e. generalizability (Bryman, 2012)[37]. Case studies tend to be problematic for many of the same reasons as external reliability. While this study has been done with specific participants in a particular setting, the themes found are abstract enough that they could be generally applicable to participation in digital innovation processes. However, future research is needed to corroborate this.

Internal reliability in qualitative research refers to the extent to which internal observers agree about what is seen and heard (Bryman, 2012)[37]. A single researcher conducted this research. However, the sessions were held with the communication professionals and with Dr. Van der Sanden. The meeting debriefs were shared with all participants with the question to provide feedback. The observations during the process were discussed with Dr. Van der Sanden to align our perception of the process.

Bryman (2012) defines internal validity as the match between researchers’ observations and theoretical ideas and notes that this is where qualitative research tends to shine ”because the prolonged participation in the social life of a group over a long period of time allows the researcher to ensure a high level of congruence between concepts and observations” (Bryman, 2012, p.390). The results of this study were developed through observations collected over 1.5 years and an independent reflection session with each communication department. The communication departments showed a high level of alignment in their reflections. The themes developed from the thematic analysis of the transcripts of the stimulated recall sessions were supported by the observations throughout the process. However, the danger is that this proximity of the researcher to the group can bias the researcher in future observations.

15.2. Future work

Section 5.2.3 introduces the Electron Learning Model, which conceptualises the interplay between co-design and absorptive capacity interact to use co-design as an educational method. However, due to the barriers described in section 6.1, this model was not validated.

Section 6.1.2 describes a skill floor to participate in a co-design process, and section 6.1.2 describes the difficulties in prioritising innovation without external incentives.

This author believes that both barriers could be overcome by developing a less complex digital tool. The network-based scientist-journalist recommender system started with creating the recommendation algorithm, after which the user interface was developed. However, the skill floor to participate in the algorithm development was too high. By the time the user interface was to be developed, a lot of momentum was already lost, and the research was mostly finished. A development process that starts with a user interface and gradually develops technical complexity is more approachable. This can help build momentum.

It can be theorised that the best system to use for a co-design learning process would be a system with simple parts where the complexity comes from the interaction between these parts. It would also create short-term, tangible results. This process showed how a prototype sparked motivation on the part of the communication professionals. A process that allows fast prototyping and direct interaction will likely motivate them to spend more time on it.

While the communication professionals were motivated by the thought of developing an AI system, AI systems are by nature technically complex. This makes them less suitable as subjects of a learning process.

Consider, instead, an information system. The Minimum Viable Product (MVP) could consist of a single-page note-taking web app connected to a bare-bone back-end and database. This MVP could be set up in a week, after which complexity can be introduced through adding functionality such as note sharing, collaboration features and user interface improvement. This would allow proper Agile development with short iterations over a prototype. The participants would be confronted with the decisions that go into software development and learn how adding features adds complexity. The feedback cycle between making a decision and seeing the effect on the prototype would be short, which benefits learning.

A tangible prototype would also give the communication professional something to show to their colleagues. One of the motivations of the communication professionals was to inspire colleagues (see section 4.2.2). However, they were hesitant to share stories about the process due to the difficulties.

> This has so many hurdles, am I going to take them all and if I'm going to take them all, is it worth it? [...] In the middle part you also do not know where you are going and you also do not know whether you will ever return to practice with this product [...] I also don't dare to shout out loud, 'Hey guys, two more months and then I'll have a presentation here with a cool algorithm'. *P1, Stimulated Recall Delft University of Technology [39:00]*

A tangible prototype could motivate them to involve others in the process and help create an external incentive by creating awareness among their colleagues. If their colleagues could interact with the prototype, the participants would be in the position of bridge between the development process and the user. This could benefit the learning process because, on the one hand, it would require them to explain the technical developments. On the other hand, it would help with reflection on their process, as it would be a mirror of their original mental model of innovation.

Participation in the development process of a user-centric tool would allow proper co-design, providing an opportunity to validate the Electron Learning Model and investigate its usefulness as a model to use a co-design process as an educational method.

15.3. Reflection on personal learning process

The structure of this thesis process was unusual. There was a year-long hiatus between the research phase and the writing phase. I was offered a management traineeship at Visma[1], where I did five projects at five different software companies. These projects ranged from creating a customer journey to programming a procurement portal in C# to developing a new product strategy. I have returned to finish this report, after which I will start as a Product Owner at Visma ProActive[2].

This work experience allows me to reflect on what I have learned during this thesis and my education as a whole, knowing what skills and knowledge are and aren't required to work in software development in general and in my future job specifically. As a Product Owner, I will be the linking pin between a development team and other - often non-technical - stakeholders. Working with the team of communication professionals has helped me realise how big the gap is between developers and users. This gap permeates every aspect of the interaction: The language that is used, the expectations, the priorities, and many others. It was valuable to encounter this gap in a position as a researcher, where the relationship is built upon a shared goal of learning without financial incentives or obligations. While this lack of incentives and obligations did cause issues with it being low on the priority list (see section 6.1.2), it did provide more freedom to try different approaches, make mistakes and learn from them.

15.3.1. The scope is never the scope

One of the things that I learned during this process is about scope. The communication professionals initially asked me to develop an Artifical Intelligence system for them. Their expectations were out of this world. The first meeting with the communication professionals resulted in a confrontation. I remember coming away from the meeting completely shaken and extremely worried about the impossibility of the task. My supervisor at the time, dr. Van der Sanden, was present at the meeting, and afterwards, we had an appointment to reflect. I will never forget how dr. Van der Sanden cheerfully told me how happy he was with how the meeting went and how we were already making progress. After double-checking that we observed the same session, he explained why. Their mental model of digital technology and innovation does not align with my mental model of digital technology and what is technically possible. They need to get exposure to different mental models to change their mental model. The fact that emotions escalated meant their mental model crashed into my mental model, which is the first step into coming together.

Looking back, my mental model of the development process was similar to theirs in that I expected them to provide a list of requirements and me to develop a tool based on their requirements. I saw participation as giving them occasional updates. I took their initial expectations as the final requirements. However, this process made me realise that scope comes about through interaction. While the assignment from the commissioner sounds clear-cut and set in stone, it is often nebulous. A commissioner comes to you because they believe you have the expertise to help them. This means part of the process is using that expertise to educate your commissioner. A shared project means setting up a shared scope. Setting this scope is where a large part of educating the commissioner happens. In all five of my trainee projects, I changed the scope of the project, sometimes drastically, sometimes slightly. My impression is that most of my fellow trainees accepted their scope mostly as was set from the beginning, while I entered projects with the attitude that the first order of business was to set the scope. It was not uncommon for me to use three weeks of the eight-week project to get my bearing and work with my project owner to define the scope.

15.3.2. The themes of involvement, complexity and expectations

I have found the themes of complexity, involvement and expectations to be a fundamental part of every process involving digital technology and most processes not involving it. As the adage goes, everybody wants change, but nobody wants *to* change. The communication professionals described just wanting convenience without investing time and energy. However, as they realised throughout the process, you need to get involved in the process to get the desired results (see section 14.2.2). This holds for many

[1] Visma is a Norwegian software company with 14 000 employees across the Nordics, Benelux, the Baltics, Eastern Europe, Spain and Latin America by the time of writing. For more information, see `https://www.visma.com/`.

[2] Product Owner is a term from the Scrum methodology, as described in section 2.4. Visma Proactive is a Netherlands-based software company that delivers a procurement and expense management solution, mainly for organisations in healthcare and education. For more information, see `https://proactive-software.com/en/`.

processes around digital technology. Technology is often complex, and people are already busy. It has often happened during my traineeship that people start with what they want me to do without realising that to get something out of my work, they will also need to get involved. I always asked, 'Once I am done, what will you do with my work?' It happened on multiple occasions that they initially did not have an answer. By working with them towards an answer, the realisation often came that they needed to get involved in my process to be able to get something out of it that would be useful to them afterwards. This also often led to a realisation that the subject matter was more complex than they initially thought, and their initial expectations were not feasible. While it might be expected that this was a frustrating experience for them, it was nearly always appreciated because it leads to realistic expectations and real results. It was the same thing with the communication professionals. While the initial meeting was a confrontation, their mental model quickly changed, and the alignment process helped them develop it.

15.3.3. The fetishisation of Artificial Intelligence

One of the things that struck me during the process was how the communication professionals regarded the concept of artificial intelligence. I have since talked to many people inside and outside of work, and I continue to be surprised by how much AI triggers people.

The problem starts with defining artificial intelligence. Wikipedia defines artificial intelligence as' intelligence demonstrated by machines'[3], and I expect most people to come up with something similar when asked to give a definition. However, it is my experience that once asked what that means, what technology can be considered AI, and what is possible with AI, every person will give a different answer. It is also unhelpful that it is used very differently in philosophy, academia, software development, the popular media and the common vernacular.

Artificial intelligence from a technological point of view

From a technical and academic point of view, definitions of what technologies are included in artificial intelligence vary wildly. Based on personal conversations, the general overlap between technologies mentioned as AI is optimisation (algorithms for solving combinatorial problems), machine learning (any system with data-driven modelling, including neural networks) and expert systems (modelling behaviour based on asking human experts).

Searching for the most popular AI algorithms, the results are 1) linear regression, 2) logistic regression and 3) decision trees[4]. Readers familiar with these techniques might be surprised to see these listed as 'top AI algorithms', as they are elementary technologies that do not mesh with the general conceptualisation of artificial intelligence. If we include these simple technologies, artificial intelligence is common and not at all interesting. If we focus on applications commonly associated with artificial intelligence, such as computer vision, natural language processing and robotics, they are very niche. While they are becoming more common, and some use cases are very high profile, they are not broadly used in software development.

If they are used, they are often used as sugar on top of other systems. For example, within Visma, there are Intelligent Document Recognition (IDR) services that use machine learning to scan images of invoices and extract the data for further processing. This is very useful as several Visma companies deliver accounting software. Where invoicing clerks used to type in the invoice data manually, Smartscan can automatically extract this data. In this way, the machine learning system is a valuable addition to the existing system. It automates a tedious and time-consuming step. However, we have used software to automate tedious and time-consuming activities from the beginning of software development. In this regard, the effect of artificial intelligence is not fundamentally different from any other piece of software. In addition, the most successful IDR product, Lyanthe, is successful because they have 500 people in the Philippines who manually check and correct invoices.

I think of artificial intelligence technologies as something like carbon fibre. Carbon fibre is an incredible material. It has applications which we could not even imagine a century ago. We do not even know half of all potential applications. All we know is that it will change some industries forever. Nevertheless, if you are going to build a house, you still create it out of concrete. The Romans built their houses out of

[3]See `https://en.wikipedia.org/wiki/Artificial_intelligence`

[4]See `https://www.spiceworks.com/tech/artificial-intelligence/articles/top-ml-algorithms/`.

concrete, we build our houses out of concrete, and most likely, our grandchildren will build their houses out of concrete. While we should look at the potential, we should also realise the limitations. Carbon fibre is amazing, but it is also only used in very specific circumstances. While they might lead to some applications that will transform people's everyday lives, most of the world will not be built using carbon fibre. Similarly, the techniques people think about with artificial intelligence are flashy and have the potential to affect our lives in many ways, but that does not mean that everything will only be AI from now on.

Artificial intelligence as a shorthand for magic
A different way of defining artificial intelligence is by looking at it through the lens of Wittgenstein's Philosophical Investigations (1953)[64]. Wittgenstein argues that a word's meaning is in the language's use. Instead of looking at how people define artificial intelligence when they are asked to verbalise their internal conceptualisation, we can look at how the phrases artificial intelligence and AI are used. Considering how these phrases are used brings to mind Clarke's third law (Clarke, 1973)[10], which states 'Any sufficiently advanced technology is indistinguishable from magic'. Viewing the concept of Artificial Intelligence through the lens, my conclusion is that from the way artificial intelligence is used in broader public discourse, the definition of artificial intelligence simply is 'magic'. One of the initial ideas from the communication professionals was to develop a system that would predict what would become news before it would become news. However, predicting the future is not the domain of artificial intelligence; it is the domain of magic. Across my traineeship, I have found that people seem to use AI as a shorthand for magic.

Even many people working in software seem to hold this view. Even though they know the limits of the systems they are developing, there is still a belief that the things Google can develop have no limits. It brings to mind The Emperors New Mind, in which Roger Penrose (1990)[38] links artificial intelligence to the children's story The Emperors New Clothes. We keep thinking that if it's just a bit smarter, it can do everything we previously thought to be magic. If we were just a bit smarter, we would understand how Google brings about their magic. Yet there is no magic. There is technology which can do amazing things, but is still limited.

This thesis process was my first real encounter with the expectations that live around artificial intelligence. During the process, the communication professionals moved away from 'innovation land' and realised their expectations were unrealistic (see section 14.2.4). While I was never particularly interested in artificial intelligence, this research and my experience afterwards have cemented my suspicion towards how the term AI is used and the expectations that come with it.

16 Conclusions

This research involved four communication professionals from Delft University of Technology and Naturalis Biodiversity Center in the design and development process of a digital tool to answer three questions:

1. What motivates the team of communication professionals to participate in a digital innovation process?
2. What is a suitable tool to develop with the team of communication professionals?
3. What is the effect of participation in the development process of a digital tool on the mental model of digital innovation?

The initial motivation of the communication professionals was to use the process to learn about the possibilities digital technologies can bring and to educate their colleagues on this. However, during the process, they demonstrated a result-oriented mindset. This created a tension between learning from the process and a desire for a specific outcome, i.e. the custom-made digital tool.

A list of requirements was made to decide on a suitable tool to develop with the communication professionals. Based on these requirements, the desired functionality of the tool was decided upon. The requirements were defined based on the desires of the communication professionals, the requirements from the Computer Science aspect of this research and the researchers' expertise. It was decided to create a network-based recommender system that recommends collaborations between scientists from the institute and newspaper journalists. Both communication departments have the responsibility of connecting scientists and journalists. This data-driven system uses readily available data from the institutes to provide insight into the collaboration network between scientists and journalists and helps to improve these collaborations. However, as discussed in section 6.1.2, even though the tool fulfilled the requirements set, it was less suitable in hindsight than expected. The technical complexity of the tool was a barrier for the communication professionals to participate. As proposed in section 15.2, further research is needed.

Stimulated recall sessions were held with both communication departments to assess their mental model of digital innovation at the end of the process and reflect on how it changed throughout the process. Four themes emerged from this: *Involvement in the process*, *complexity of the process*, and *expectations of technology*.

While the communication professionals had the goal of learning by participating in the process, there was a latent expectation that this would mean handing in a list of requirements and receiving the tool. The first theme in their mental model shift is their vision of involvement in the process. They reflect that to benefit from the process - but also to get a result that matches the needs - it is required to engage with the process.

Getting involved in the process led to a process of realisation about the complexity of the process. Instead of a relatively straightforward and easy process, they were confronted with the fact that innovation processes require time and energy, while the outcome is not always certain.

It also led to a realisation of what can be expected of technological innovation. Their initial expectations of what the final tool would be able to do were unrealistically high. The communication professionals expected a 'digital colleague' that could function largely autonomously. Instead, they have started to see technology as supporting human action.

In conclusion, the effect of participation in the development process was that their mental model went from a 'mysterious black box' to a 'super puppy'. The communication professionals initially regarded technology and the innovation process as something external, where they could hand in a list of requirements and walk away with a digital tool. During the process, they came to see technology as a super puppy that can do remarkable things but has to be trained and interacted with to get it to do what you want.

Bibliography

[1] Alireza Abbasi, Liaquat Hossain, and Loet Leydesdorff. Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. *Journal of Informetrics*, 6(3):403–412, 2012. ISSN 1751-1577. doi: https://doi.org/10.1016/j.joi.2012.01.002. URL https://www.sciencedirect.com/science/article/pii/S175115771200003X.

[2] Amina Adadi and Mohammed Berrada. Peeking inside the black-box: a survey on explainable artificial intelligence (xai). *IEEE access*, 6:52138–52160, 2018.

[3] Deborah Andrews, Elizabeth J Newton, Naeem Adibi, Julie Chenadec, and Katrin Bienge. A circular economy for the data centre industry: Using design methods to address the challenge of whole system sustainability in a unique industrial sector. *Sustainability*, 13(11):6319, 2021.

[4] Jibouni Ayoub, Dounia Lotfi, and Ahmed Hammouch. Link prediction using betweenness centrality and graph neural networks. *Social Network Analysis and Mining*, 13(1), 2023. doi: 10.1007/s13278-022-00999-1. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85143718815&doi=10.1007%2fs13278-022-00999-1&partnerID=40&md5=eb0d8a3815d2c6ca87c6a95ea2bc80e0. Cited by: 0.

[5] Bouchaib Bahli and ESA Zeid. The role of knowledge creation in adopting extreme programming model: an empirical study. In *2005 International Conference on Information and Communication Technology*, pages 75–87. IEEE, 2005.

[6] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, 58:82–115, 2020. ISSN 1566-2535. doi: https://doi.org/10.1016/j.inffus.2019.12.012. URL https://www.sciencedirect.com/science/article/pii/S1566253519308103.

[7] Lamiae Benhayoun, Marie Anne Le Dain, Carine Dominguez-Péry, and Andrew C. Lyons. SMEs embedded in collaborative innovation networks: How to measure their absorptive capacity? *Technological Forecasting and Social Change*, 159(September 2019):120196, 2020. ISSN 00401625. doi: 10.1016/j.techfore.2020.120196. URL https://doi.org/10.1016/j.techfore.2020.120196.

[8] Stefano A. Bini. Artificial intelligence, machine learning, deep learning, and cognitive computing: What do these terms mean and how will they impact health care? *The Journal of Arthroplasty*, 33(8):2358–2361, 2018. ISSN 0883-5403. doi: https://doi.org/10.1016/j.arth.2018.02.067. URL https://www.sciencedirect.com/science/article/pii/S0883540318302158.

[9] Hilary Boyd, Stephen McKernon, Bernie Mullin, and Andrew Old. Improving healthcare through the use of co-design. *New Zealand Medical Journal*, 125(1357), 2012. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84863524714&partnerID=40&md5=a1218eae5bcabe97814660321d48ba77. Cited by: 122.

[10] Arthur C Clarke. Profiles of the future, an enquiry into the limits of the possible. 1973.

[11] Wesley M. Cohen and Daniel A. Levinthal. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 1990. ISSN 00018392. doi: 10.2307/2393553.

[12] Sébastien Cuendet, Quentin Bonnard, Son Do-Lenh, and Pierre Dillenbourg. Designing augmented reality for the classroom. *Computers and Education*, 68:557 – 569, 2013. doi: 10.1016/j.compedu.2013.02.015. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84884908664&doi=10.1016%2fj.compedu.2013.02.015&partnerID=40&md5=d876cb55694e5d7f291668ffbfd46a9d. Cited by: 179; All Open Access, Green Open Access.

[13] Salomao David, Amalia G Sabiescu, and Lorenzo Cantoni. Co-design with communities. a reflection on the literature. *Proceedings of the 7th International Development Informatics Association Conference*, pages 152–166, 2013.

[14] Danielly de Paula, Carolin Marx, Ella Wolf, Christian Dremel, Kathryn Cormican, and Falk Uebernickel. A managerial mental model to drive innovation in the context of digital transformation. *Industry and Innovation*, 2022. ISSN 14698390. doi: 10.1080/13662716.2022.2072711.

[15] S. L. Duncombe, A. R. Barker, L. Price, J. L. Walker, P. E. Dux, A. Fox, N. Matthews, and M. Stylianou. Making a hiit: study protocol for assessing the feasibility and effects of co-designing high-intensity interval training workouts with students and teachers. *BMC Pediatrics*, 22(1), 2022. URL www.scopus.com.

[16] Tore Dybå and Torgeir Dingsøyr. Empirical studies of agile software development: A systematic review. *Information and Software Technology*, 50(9):833–859, 2008. ISSN 0950-5849. doi: https://doi.org/10.1016/j.infsof.2008.01.006. URL https://www.sciencedirect.com/science/article/pii/S0950584908000256.

[17] M.C. Elish and Danah Boyd. Situating methods in the magic of big data and ai. *Communication Monographs*, 85(1):57 – 80, 2018. doi: 10.1080/03637751.2017.1375130. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85029681855&doi=10.1080%2f03637751.2017.1375130&partnerID=40&md5=73acd7fb70fc64e17437e84d404b7ed3. Cited by: 115.

[18] Marta Entradas and Martin M. Bauer. Mobilisation for public engagement: Benchmarking the practices of research institutes. *Public Understanding of Science*, 26(7):771–788, 2017. ISSN 13616609. doi: 10.1177/0963662516633834.

[19] Ernesto Estrada, Desmond J. Higham, and Naomichi Hatano. Communicability betweenness in complex networks. *Physica A: Statistical Mechanics and its Applications*, 388(5):764–774, 2009. ISSN 0378-4371. doi: https://doi.org/10.1016/j.physa.2008.11.011. URL https://www.sciencedirect.com/science/article/pii/S0378437108009473.

[20] Lynette D Henderson and Julie Tallman. *Stimulated recall and mental models: Tools for teaching and learning computer information literacy*, volume 2. Scarecrow Press, 2006.

[21] Javier Martın Hernández and Piet Van Mieghem. Classification of graph metrics. *Delft University of Technology: Mekelweg, The Netherlands*, pages 1–20, 2011.

[22] F. Hill, Ch. & Rothaermel. The Performance of Incumbent Firms in the Face of Radical Technological Innovation. *Academy of Management Review*, 28(2):257–274, 2003.

[23] David H Jonassen. Operationalizing mental models: Strategies for assessing mental models to support meaningful learning and design-supportive learning environments. 1995.

[24] Sabrina Heike Kessler, Mike S. Schäfer, David Johann, and Heiko Rauhut. Mapping mental models of science communication: How academics in germany, austria and switzerland understand and practice science communication. *Public Understanding of Science*, 31:711–731, 8 2022. ISSN 13616609. doi: 10.1177/09636625211065743.

[25] Maaike Kleinsmann and Rianne Valkenburg. Barriers and enablers for creating shared understanding in co-design projects. *Design Studies*, 29(4):369–386, 2008. ISSN 0142-694X. doi: https://doi.org/10.1016/j.destud.2008.03.003. URL https://www.sciencedirect.com/science/article/pii/S0142694X0800032X.

[26] Vito Laterza. Could cambridge analytica have delivered donald trump's 2016 presidential victory? an anthropologist's look at big data and political campaigning. *Public Anthropologist*, 3(1):119 – 147, 2021. doi: 10.1163/25891715-03010007. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85114445487&doi=10.1163%2f25891715-03010007&partnerID=40&md5=a0e1a5539eb686337bb801fd12f9415b. Cited by: 3; All Open Access, Green Open Access, Hybrid Gold Open Access.

[27] Cong Li, Qian Li, Piet Van Mieghem, H Eugene Stanley, and Huijuan Wang. Correlation between centrality metrics and their application to the opinion model. *The European Physical Journal B*, 88 (3):1–13, 2015.

[28] Logeion. Keynote deborah nas c-day19. URL https://www.youtube.com/watch?v=hr1sU1_i5Zc.

[29] Edward N. Lorenz. Designing chaotic models. *Journal of the Atmospheric Sciences*, 62(5):1574 – 1587, 2005. doi: 10.1175/JAS3430.1. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-20444468413&doi=10.1175%2fJAS3430.1&partnerID=40&md5=06ac040d6ff61da2315eef774aebc68c. Cited by: 141.

[30] Peter E.D. Love, Pauline Teo, Murray Davidson, Shaun Cumming, and John Morrison. Building absorptive capacity in an alliance: Process improvement through lessons learned. *International Journal of Project Management*, 34(7):1123–1137, 2016. ISSN 0263-7863. doi: https://doi.org/10.1016/j.ijproman.2016.05.010. URL https://www.sciencedirect.com/science/article/pii/S0263786316300278.

[31] John Lyle. Stimulated Recall: A report on its use in naturalistic research. *British Educational Research Journal*, 29(6):861–878, 2003. ISSN 01411926. doi: 10.1080/0141192032000137349.

[32] Frank J Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951.

[33] Eric Merkley. Are Experts (News)Worthy? Balance, Conflict, and Mass Media Coverage of Expert Consensus. *Political Communication*, 37(4):530–549, 2020. ISSN 10917675. doi: 10.1080/10584609.2020.1713269. URL https://doi.org/10.1080/10584609.2020.1713269.

[34] J Miedema, Mascha C van der Voort, Diederick Lutters, and Frederikus JAM van Houten. Synergy of technical specifications, functional specifications and scenarios in requirements specifications. In *The Future of Product Development*, pages 235–245. Springer, 2007.

[35] Staša Milojević. Modes of collaboration in modern science: Beyond power laws and preferential attachment. *Journal of the american society for Information science and technology*, 61(7):1410–1423, 2010.

[36] Genevieve Mosely, Natalie Wright, and Cara Wrigley. Facilitating design thinking: A comparison of design expertise. *Thinking Skills and Creativity*, 27(August 2017):177–189, 2018. ISSN 18711871. doi: 10.1016/j.tsc.2018.02.004. URL https://doi.org/10.1016/j.tsc.2018.02.004.

[37] Sarah Nettleton, Geoffrey Harding, and Kevin Taylor. *Social Research Methods*. Oxford University Press Inc, New York, 4 edition, 2003. ISBN 978–0–19–958805–3. doi: 10.1201/9780203381175.ch9.

[38] Roger Penrose and N David Mermin. The emperor's new mind: Concerning computers, minds, and the laws of physics, 1990.

[39] Olle Persson, Wolfgang Glänzel, and Rickard Danell. Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3):421–432, 2004.

[40] R. Power, J. M. Ussher, A. Hawkey, O. Missiakos, J. Perz, O. Ogunsiji, N. Zonjic, C. Kwok, K. McBride, and M. Monteiro. Co-designed, culturally tailored cervical screening education with migrant and refugee women in australia: a feasibility study. *BMC Women's Health*, 22(1), 2022. URL www.scopus.com.

[41] Aiur Retegi, Brigitte Sauvage, Barbara Predan, Enrique Tomás, Gisa Schosswohl, Martin Kaltenbrunner, and Denisa Draganovská. Co-Create Basics: Welcome to Co-Create! URL http://www.cocreate.training/wp-content/uploads/2018/07/CoCreate-basics-document.pdf.

[42] Aiur Retegi, Brigitte Sauvage, Barbara Predan, Enrique Tomás, Gisa Schosswohl, Martin Kaltenbrunner, and Denisa Draganovská. *The Co-create Handbook for Creative Professionals*. 2019. URL http://www.cocreate.training/wp-content/uploads/2019/03/co-design{_}handbook{_}FINAL.pdf.

[43] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 1135–1144, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450342322. doi: 10.1145/2939672.2939778. URL https://doi.org/10.1145/2939672.2939778.

[44] George P Richardson, David F Andersen, Terrence A Maxwell, and Thomas R Stewart. Foundations of mental model research. In *Proceedings of the 1994 International System Dynamics Conference*, pages 181–192. EF Wolstenholme, 1994.

[45] Johnny Saldaña. The coding manual for qualitative researchers. *The coding manual for qualitative researchers*, pages 1–440, 2021.

[46] Ken Schwaber and Jeff Sutherland. The scrum guide the definitive guide to scrum: The rules of the game, 2020. URL https://www.scrum.org/resources/scrum-guide.

[47] Samuel Severance, William R. Penuel, Tamara Sumner, and Heather Leary. Organizing for teacher agency in curricular co-design. *Journal of the Learning Sciences*, 25(4):531 – 564, 2016. doi: 10.1080/10508406.2016.1207541. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84982168508&doi=10.1080%2f10508406.2016.1207541&partnerID=40&md5=162655660ebe057977fb8de4baa8adff. Cited by: 82.

[48] Robert C. Smith. Bohr vs. electron cloud, neutrons and basic interactions, Jul 2019. URL https://tracingcurves.wordpress.com/2018/01/03/neutrons-bohr-vs-electron-cloud-basic-interactions/.

[49] Robin Snell and Almaz Man-Kuen Chak. The learning organization: learning and empowerment for whom? *Management learning*, 29(3):337–364, 1998.

[50] Sara Nadiv Soffer and Alexei Vázquez. Network clustering coefficient without degree-correlation biases. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 71(5), 2005. doi: 10.1103/PhysRevE.71.057101. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-26944456272&doi=10.1103%2fPhysRevE.71.057101&partnerID=40&md5=a76e9f050e5d2429d0c28fc7b4cd878e. Cited by: 175.

[51] Mary M Somerville and Mary Nino. Collaborative co□design: a user-centric approach for advancement of organizational learning. *Performance Measurement and Metrics*, 8(3):180–188, 2007. ISSN 1467-8047.

[52] David Sumpter. *Outnumbered: from Facebook and Google to fake news and filter-bubbles - the algorithms that control our lives*, volume 369. 2019.

[53] Diana Thamrin, Laksmi Kusuma Wardani, Ronald Hasudungan Irianto Sitindjak, and Listia Natadjaja. Experiential learning through community co-design in interior design pedagogy. *International Journal of Art and Design Education*, 38:461–477, 5 2019. ISSN 14768070. doi: 10.1111/jade.12208.

[54] The Design Council. Framework for Innovation. URL https://www.designcouncil.org.uk/news-opinion/what-framework-innovation-design-councils-evolved-double-diamond.

[55] The Design Council. A study of the design process. Technical report, 2007. URL http://www.designcouncil.org.uk/sites/default/files/asset/document/ElevenLessons_Design_Council(2).pdf.

[56] Maarten CA van der Sanden and Caroline Wehrmann. A mind set called design thinking. In *Design-Based Concept Learning in Science and Technology Education*, pages 285–310. Brill Sense, 2021.

[57] Gregory Vial. Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems*, 28(2):118–144, 2019. ISSN 09638687. doi: 10.1016/j.jsis.2019.01.003. URL https://doi.org/10.1016/j.jsis.2019.01.003.

[58] Greg Walsh, Alison Druin, Mona Leigh Guha, Elizabeth Foss, Evan Golub, Leshell Hatley, Elizabeth Bonsignore, and Sonia Franckel. Layered elaboration: A new technique for co-design with children. volume 2, page 1237 – 1240, 2010. doi: 10.1145/1753326.1753512. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-77953972619&doi=10.1145%2f1753326.1753512&partnerID=40&md5=ac95ed686171e8b4c9376eba9e94f80a. Cited by: 97.

[59] Ludo Waltman, Nees Jan van Eck, Thed N. van Leeuwen, and Martijn S. Visser. Some modifications to the snip journal impact indicator. *Journal of Informetrics*, 7(2):272 – 285, 2013. doi: 10.1016/j.joi.2012.11.011. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84872186449&doi=10.1016%2fj.joi.2012.11.011&partnerID=40&md5=9c62dab2d30082b44ef2e18446fe8699. Cited by: 121; All Open Access, Green Open Access.

[60] Peng Wang, BaoWen Xu, YuRong Wu, and XiaoYu Zhou. Link prediction in social networks: the state-of-the-art. *Science China Information Sciences*, 58(1):1 – 38, 2015. doi: 10.1007/s11432-014-5237-y. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84925512263&doi=10.1007%2fs11432-014-5237-y&partnerID=40&md5=dbce274b8d0294d60f9bb3ca2ed9f4c7. Cited by: 416; All Open Access, Green Open Access.

[61] Sheila Simsarian Webber, Gilad Chen, Stephanie C. Payne, Sean M. Marsh, and Stephen J. Zaccaro. Enhancing team mental model measurement with performance appraisal practices. *Organizational Research Methods*, 3(4):307–322, 2000. doi: 10.1177/109442810034001. URL https://doi.org/10.1177/109442810034001.

[62] Lilian Weng, Filippo Menczer, and Yong-Yeol Ahn. Virality prediction and community structure in social networks. *Scientific Reports*, 3, 2013. doi: 10.1038/srep02522. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84883391597&doi=10.1038%2fsrep02522&partnerID=40&md5=ccc0e3f48708de4ef021c02c90bf054e. Cited by: 379; All Open Access, Gold Open Access, Green Open Access.

[63] Janusz Wielki. Analysis of the role of digital influencers and their impact on the functioning of the contemporary on-line promotional system and its sustainable development. *Sustainability (Switzerland)*, 12(17), 2020. ISSN 20711050. doi: 10.3390/su12177138.

[64] Ludwig Wittgenstein. *Philosophical investigations*. Blackwell, Oxford, 1953.

[65] Wen Wu, Liang He, and Jing Yang. Evaluating recommender systems. page 56 – 61, 2012. doi: 10.1109/ICDIM.2012.6360092. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84871589287&doi=10.1109%2fICDIM.2012.6360092&partnerID=40&md5=62cb395dc3c322fd33564f63b7c72643. Cited by: 16.

[66] Chia-Lun Yeh, Babak Loni, Mariëlle Hendriks, Henrike Reinhardt, and Anne Schuth. Dpgmedia2019: A dutch news dataset for partisanship detection, 2019.

[67] Shaker A. Zahra and Gerard George. Absorptive Capacity: A Review, Reconceptualization, and Extension.pdf. *Academy of management review*, 27(12):185 – 203, 2002.

[68] Yujia Zhai, Jiaqi Yan, Hezhao Zhang, and Wei Lu. Tracing the evolution of ai: conceptualization of artificial intelligence in mass media discourse. *Information Discovery and Delivery*, 48(3):137 – 149, 2020. doi: 10.1108/IDD-01-2020-0007. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083585147&doi=10.1108%2fIDD-01-2020-0007&partnerID=40&md5=85908b1c4a591f495bed676ab7346d83. Cited by: 7.

[69] Xiu-Xiu Zhan, Ziyu Li, Naoki Masuda, Petter Holme, and Huijuan Wang. Susceptible-infected-spreading-based network embedding in static and temporal networks. *EPJ Data Science*, 9(1): 30, 2020.

[70] Li Zou, Chao Wang, An Zeng, Ying Fan, and Zengru Di. Link prediction in growing networks with aging. *Social Networks*, 65:1–7, 2021. ISSN 0378-8733. doi: https://doi.org/10.1016/j.socnet.2020.11.001. URL https://www.sciencedirect.com/science/article/pii/S0378873320300952.

A

Opzet brainstorm communicatieafdelingen 27-2-2020

Achtergrond

Het hoofddoel van dit onderzoek is om de mogelijkheden te onderzoeken van het toepassen van informatica binnen het communicatieveld door een proof-of-concept te maken van een algoritme. Dit algoritme geeft niet alleen een tastbaar idee van wat er mogelijk is, maar het proces zelf moet aan het licht brengen welke behoeftes, belangen en weerstanden er zijn. Bij de kennismakingsmeeting hebben we vastgesteld dat we met verschillende perspectieven het gesprek in zijn gegaan. Corine en Evelyne zaten te denken aan een algoritme dat actualiteiten koppelt aan kennis van het instituut, Maarten zat te denken aan een tool die samenwerken in het netwerk verbetert en Tim was geïnteresseerd in een onderzoek naar het functioneren van sociale netwerken. Daarom hebben we besloten om een brainstormsessie te houden over de richting van dit project.

Doel

Deze brainstorm is de aftrap van de scriptie. Deze scriptie zal de volgende fases bevatten:

1. Verkennen waar digitale oplossingen toegevoegde waarde kunnen bieden voor de communicatieafdelingen
2. Een probleemstelling formuleren
3. Een tool ontwerpen
4. De tool bouwen
5. Op basis van de tool analyseren wat de impact is op de afdelingen

Het doel van deze sessie is een gedeelde roadmap te maken voor de verkennende fase. De insteek is om de verschillende vraagstellingen, achterliggende problemen en mogelijke pijnpunten boven water te krijgen. Op basis daarvan gaan we een plan formuleren hoe ik deze ga onderzoeken. Hierbij is het niet de bedoeling om al een richting in te slaan, maar om voorlopige onderzoeksvragen te formuleren.

Daarnaast is deze sessie om een aantal praktische zaken te regelen, zodat ik daadwerkelijk aan de slag kan

Opzet

|

Activiteit	Tijd	Van	Tot	Opmerking
Introductie	10 min	0:00	0:10	Wat is het doel van het project? Wat is het doel van deze sessie? Hoe gaan we deze sessie aanpakken?
Icebreaker	10 min	0:10	0:20	Met wie zitten we aan tafel?
Verkenning - werk	15 min	0:20	0:45	Wat wil je uit je functie halen?
Verkenning - onderzoek 1	10 min	0:45	0:55	Wat vind je interessant aan dit onderzoek?
Pauze	5 min	0:55	1:00	
Verkenning - onderzoek 2	25 min	1:00	1:25	Welke vragen wil je dat dit onderzoek beantwoord?
Roadmap	20 min	1:25	1:45	In welke volgorde gaan we de vragen beantwoorden?
Praktische zaken	10 min	1:45	1:55	Wat moet er praktisch geregeld worden om te beginnen?
Afsluiting	5 min	1:55	2:00	

Introductie

- Opzet van workshop
- Herhalen achtergrond
  - Doel formuleren
  - Over 10 jaar is je collega een algoritme, wat betekent dat?
  - Hoe beinvloeden digitale ontwikkelingen het werkveld van communicatieprofessionals?
  - Op abstract niveau onderzoeken wat de ideeen en behoeftes omtrent dit onderzoek zijn
  - Voorlopige onderzoeksvragen formuleren
- Aan het einde van deze brainstorm hebben we:
  - Een overzicht van de vragen die we met deze scriptie willen beantwoorden
  - Een grof idee in welke volgorde we deze vragen gaan beantwoorden
  - Een concreet idee wat er de komende twee a drie weken gaat gebeuren

Icebreaker - persoonlijk

Doel: Elkaar leren kennen en in een actieve, creatieve en open mindset komen.

1. Vertel iets dat nog niemand hier aan tafel weet (een hobby, iets wat je ooit gedaan hebt, een gewoonte)
2. Vertel iets wat je graag nog een keer zou willen doen (mag onrealistisch zijn)

Verkenning - werk

Doel: Een beeld krijgen van hoe iedereen in zijn baan staat.
Aanpak: Iedereen maakt een mindmap.

1. Schrijf iets op wat je het leukst aan je baan vind
2. Schrijf iets op waar je goed in bent
3. Schrijf iets op wat je wil leren
4. Verbind van elk van de dingen iets dat matcht en iets dat botst. Schrijf in de verbinding waarom.

Verkenning - onderzoek

Doel: Onderzoeken wat we uit dit onderzoek willen krijgen en aan de hand daarvan onderzoeksvragen opstellen.
Aanpak: Postits schrijven en vervolgens clusteren/evalueren.

1. Schrijf drie dingen op die je interessant vind aan dit project

2. Schrijf drie dingen op die je leuk zou vinden om van dit project te leren
3. Schrijf drie vragen op die je beantwoord wil hebben

Vervolgens gaan we op een flyover vel gelijksoortige vragen clusteren. Daarna brengen we hierarchien aan; we onderscheiden subvragen en hoofdvragen en als twee vragen geen gelijke deler hebben, dan zoeken we een overkoepelende vraag.

Roadmap

Aan de hand van de vraagboom wordt duidelijk welke verschillende aspecten dit onderzoek moet bevatten. Hieruit kiezen we een strategie over de volgorde van de vragen en hoe we die gaan beantwoorden. Vervolgens kiezen we waar we mee gaan beginnen.

Praktische zaken

- Kantoortje bij communicatieafdeling(en)
- Hoe vinden we welke data beschikbaar is?
- Wanneer is de volgende meeting?
- Wvttk

Afsluiting

- Wat is je gevoel over deze sessie?
- Werkt deze aanpak?
- Wat kan verbeterd worden?
- Heb je het gevoel dat alles voldoende duidelijk is?

B

Example of blog

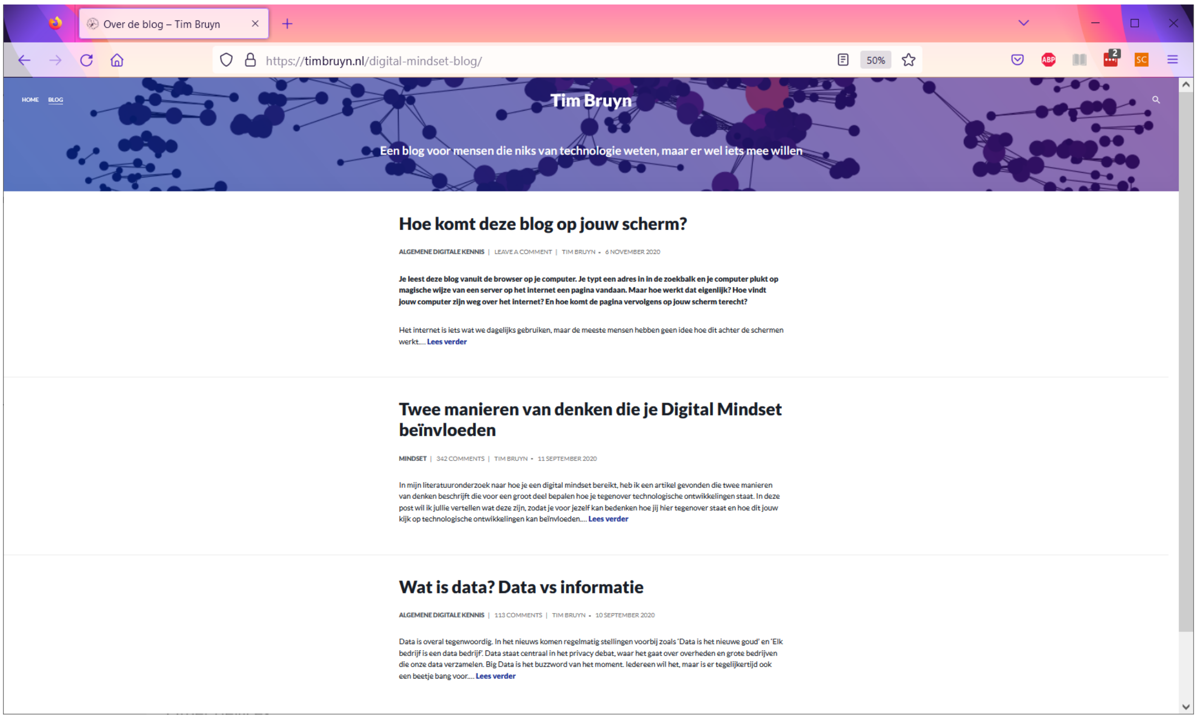


Figure B.1: The overview page showing the three available blog posts. The overview page is available at `https://timbruyn.nl/digital-mindset-blog/`.

HOME BLOG

Tim Bruyn

Wat is data? Data vs informatie

ALGEMENE DIGITALE KENNIS

113 COMMENTS | TIM BRUYN • 10 SEPTEMBER 2020

Data is overal tegenwoordig. In het nieuws komen regelmatig stellingen voorbij zoals 'Data is het nieuwe goud' en 'Elk bedrijf is een data bedrijf'. Data staat centraal in het privacy debat, waar het gaat over overheden en grote bedrijven die onze data verzamelen. Big Data is het buzzword van het moment. Iedereen wil het, maar is er tegelijkertijd ook een beetje bang voor. Maar wat is data nou precies? Welke data heb je zelf in huis, hoe kan je aan nieuwe data komen en hoe bepaal je of die data nuttig is? Wanneer wordt data *big*?

In een blog om een digital mindset te krijgen, mag data niet ontbreken. Dit is de eerste post in een serie over data. We beginnen bij twee dingen die vaak verward worden, want de eerste stap in het gebruiken van data, is het herkennen van data.

Data vs informatie

Om te beginnen wil ik het hebben over het verschil tussen data en informatie. De twee lijken erg op elkaar, maar zijn fundamenteel verschillend. Data zijn de ruwe, onverwerkte gegevens. Informatie is gestructureerde data in een nuttige context. Een tabel met losse getallen is data. Een grafiek op basis van die tabel geeft ons informatie.

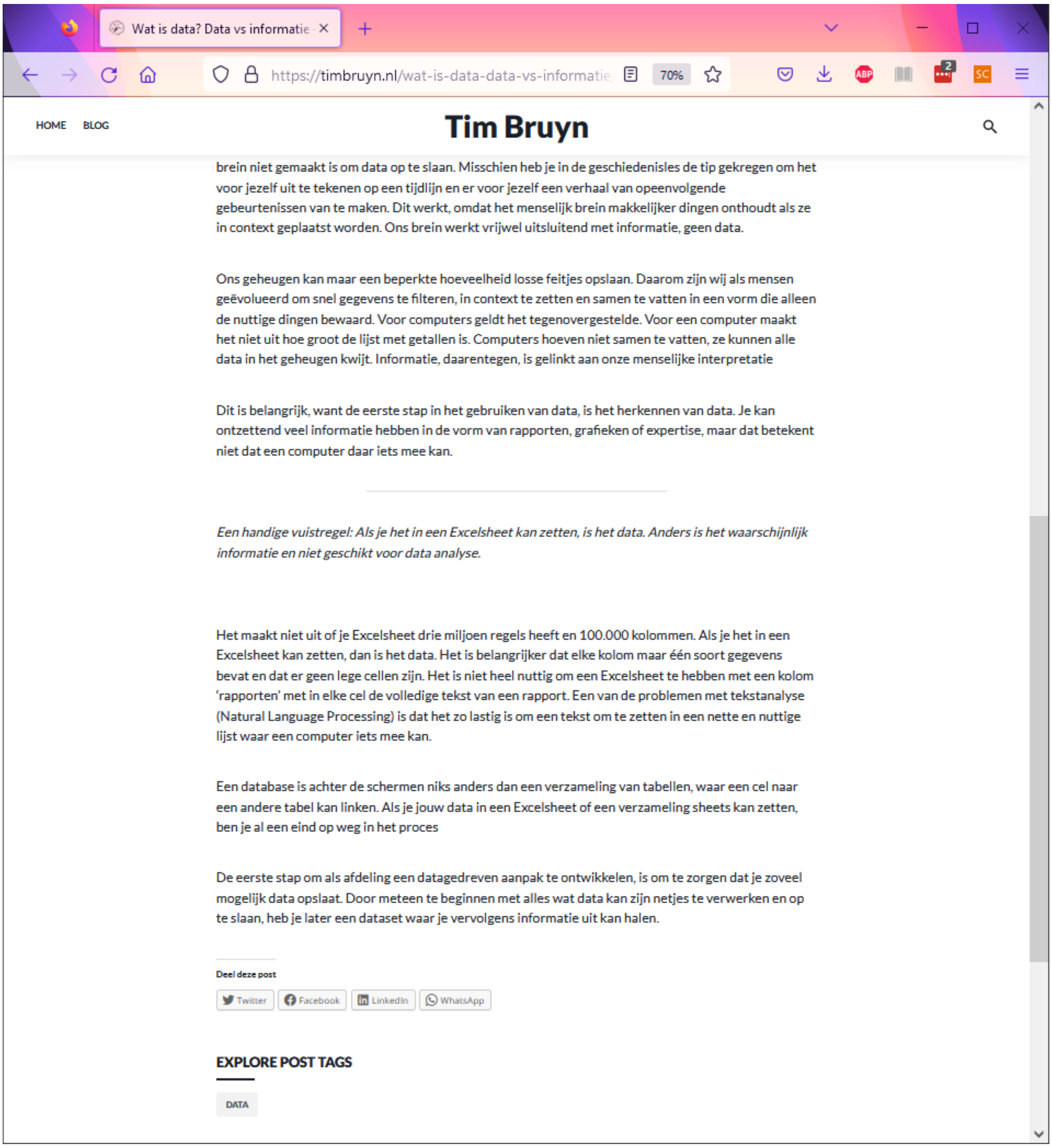
Naam	Aantal artikelen
J. Jansen	2
P. Smid	3
B. Malus	1

Data

Informatie

Iedereen die voor geschiedenis een lijst jaartallen uit het hoofd moest leren, weet dat het menselijk brein niet gemaakt is om data op te slaan. Misschien heb je in de geschiedenisles de tip gekregen om het voor jezelf uit te tekenen op een tijdlijn en er voor jezelf een verhaal van opeenvolgende gebeurtenissen van te maken. Dit werkt, omdat het menselijk brein makkelijker dingen onthoudt als ze in context geplaatst worden. Ons brein werkt vrijwel uitsluitend met informatie, geen data.

Figure B.2: An example of a blog post, first half. The post is available at https://timbruyn.nl/wat-is-data-data-vs-informatie/.



HOME BLOG

Tim Bruyn

brein niet gemaakt is om data op te slaan. Misschien heb je in de geschiedenisles de tip gekregen om het voor jezelf uit te tekenen op een tijdlijn en er voor jezelf een verhaal van opeenvolgende gebeurtenissen van te maken. Dit werkt, omdat het menselijk brein makkelijker dingen onthoudt als ze in context geplaatst worden. Ons brein werkt vrijwel uitsluitend met informatie, geen data.

Ons geheugen kan maar een beperkte hoeveelheid losse feitjes opslaan. Daarom zijn wij als mensen geëvolueerd om snel gegevens te filteren, in context te zetten en samen te vatten in een vorm die alleen de nuttige dingen bewaard. Voor computers geldt het tegenovergestelde. Voor een computer maakt het niet uit hoe groot de lijst met getallen is. Computers hoeven niet samen te vatten, ze kunnen alle data in het geheugen kwijt. Informatie, daarentegen, is gelinkt aan onze menselijke interpretatie

Dit is belangrijk, want de eerste stap in het gebruiken van data, is het herkennen van data. Je kan ontzettend veel informatie hebben in de vorm van rapporten, grafieken of expertise, maar dat betekent niet dat een computer daar iets mee kan.

Een handige vuistregel: Als je het in een Excelsheet kan zetten, is het data. Anders is het waarschijnlijk informatie en niet geschikt voor data analyse.

Het maakt niet uit of je Excelsheet drie miljoen regels heeft en 100.000 kolommen. Als je het in een Excelsheet kan zetten, dan is het data. Het is belangrijker dat elke kolom maar één soort gegevens bevat en dat er geen lege cellen zijn. Het is niet heel nuttig om een Excelsheet te hebben met een kolom 'rapporten' met in elke cel de volledige tekst van een rapport. Een van de problemen met tekstanalyse (Natural Language Processing) is dat het zo lastig is om een tekst om te zetten in een nette en nuttige lijst waar een computer iets mee kan.

Een database is achter de schermen niks anders dan een verzameling van tabellen, waar een cel naar een andere tabel kan linken. Als je jouw data in een Excelsheet of een verzameling sheets kan zetten, ben je al een eind op weg in het proces

De eerste stap om als afdeling een datagedreven aanpak te ontwikkelen, is om te zorgen dat je zoveel mogelijk data opslaat. Door meteen te beginnen met alles wat data kan zijn netjes te verwerken en op te slaan, heb je later een dataset waar je vervolgens informatie uit kan halen.

Deel deze post

Twitter Facebook LinkedIn WhatsApp

EXPLORE POST TAGS

DATA

Figure B.3: An example of a blog post, first second. The post is available at `https://timbruyn.nl/wat-is-data-data-vs-informatie/`.

C

Stimulated Recall coding results

Quotation Content	Codes	Dept.
als je kijkt naar hoe ik het pad terug zie wat we belopen hebben is dat, we moesten eerst nog een beetje 'tech savy' worden,	Before	DUT
we willen het allemaal zonder al te veel moeite willen we eigenlijk gewoon gemak naast ons hebben. Dus iets te kort door de bocht zeg maar,	Before	DUT
sceptisch ben ik nooit geweest	Before	DUT
Ja, weet je we gingen gewoon, ik had me ook kunnen inlezen eerst, en ja dat heb ik niet gedaan, dus ik moet bijgespijkerd worden anders ga je tekort door de bocht.	Before During	DUT
ik zie het letterlijk als een brug die we hebben moeten belopen en daarbij staat de praktijk aan de ene kant en met het idee aan de andere kant, zeg maar. Dus in de eerste keer waren we ploep naar de ander kant van de brug gegaan en zaten we daar in innovatieland en gingen we helemaal los we zouden dit wel willen, dat wel willen, dus zus wel willen, zo wel willen, algoritme als collega, vacature tekst, allemaal zonder enige vorm van obligations daaraan en niet gehinderd door enige vorm van kennis.	Before	DUT
ik heb altijd een vraagteken gehad of datgene waar jij op afstudeert wel echt bruikbaar zou zijn in de praktijk,	Before	DUT
in het begin van het traject vrij concreet van 'dit willen we', we hadden er allemaal al een soort beeld erbij, dan wordt het allemaal moeilijk en ja nu zeggen we 'doe ff makkelijk klop het effe in'. Maar dat je ook wel beseft dat het er wel bij hoort, dat soort bewustwordingsstappen om er te komen. En dan vraag ik me mezelf meteen af wat betekend dat dan en dat we nu aan het eind zo extreem aangenaam verrast zijn? Zeg ik dat dan te hard?	Before After	DUT
het is bijna 'in your face', want ik heb geroepen op een podium 'jongens we moeten technologie wel durven omarmen als communicatieprofessionals', je moet er in duiken. Nou vervolgens moest ìk dat doen en dacht ik Jezus wat een toestand en dus ja, duh, 'practice what you preach', dan moet je het ook doen,	Before After	DUT
Het begon met 'we gaan een algoritme bouwen als collega' en kennelijk, ben ik, snap ik als ik nu terugkijk, was ik daar zo ver dat ja dat is ook een echte collega,	Before After	DUT
het is natuurlijk hartstikke leuk als het werkt, maar het ging ons ook om het zelf te ervaren van wat er gebeurd en dan is dat vertrouwen, daar wordt je zo met je neus op gewezen, dat je denkt ja wil je zoiets nou echt gaan gebruiken, dan moet je wel zeker weten dat het klopt	Before After	NBC
Je bent hem ook op maat aan het kneden, dus ik denk dat het eigenlijk met dit soort tools misschien wel gebruikelijk is en misschien is dat wel een inzicht wat voor mijn vakgenoten ook echt heel belangrijk is om te weten, omdat wij zo gewend zijn om gewoon iets te bestellen en je trekt het aan en loopt er mee weg of je gaat het gebruiken, maar dat is hierbij dus niet.	Before After	NBC
in het begin, niet persé onzeker, maar gewoon nog een beetje proevend was, een beetje van 'nou misschien?' en dat je op een gegeven moment gedurende dat proces kom je steeds zekerder over, ja nee dit moet het zijn?	Before After	NBC
in het begin was dat leuk want dan kwamen we bij elkaar met Evelyne en haar team, met mensen bij, Maarten, jij, toen was het echt zo nou we gaan dit met elkaar doen,	Before	NBC
In het begin vond ik het makkelijker om er tijd voor vrij te maken, want dan heb je die energie maar nu moet je wel lang doorgaan.	After Before	NBC

Table C.1: All excerpts from the stimulated recall sessions coded Before.

Quotation Content	Codes	Dept.
vervolgens kregen we een periode waarin we heel veel theorie van jou kregen en dan blijft die van vraag van ‘oh wat komt er nou uit?’, ik ben ongeduldig van aard, dus dat wordt daardoor gevoed	During	DUT
Ja, weet je we gingen gewoon, ik had me ook kunnen inlezen eerst, en ja dat heb ik niet gedaan, dus ik moet bijgespijkerd worden anders ga je tekort door de bocht.	Before During	DUT
het feit dat je een concreet uitgangspunt nu hebt van waaruit je verder kunt bouwen is voor mij wel heel waardevol, omdat je je dan iets kunt voorstellen bij wat er kan gebeuren en in het theoretische deel mis je dat gewoon denk ik.	After During	DUT
Vervolgens ga je die middenfase in waarin je eerst kennis moet bijspijkeren, dan moet kijken wat er überhaupt in je hoofd gebeurd. Dan komen er heel veel vragen en dan verlies je langzaam het vertrouwen, dat je denk ja jongens, dit heeft zoveel hobbels, ga ik die allemaal nemen en als ik ze allemaal ga nemen is het dat dan waard Maar je moet ze nemen, anders kan je nooit hier komen. In het middenstuk weet je ook niet waar je naar toe komt en je weet ook niet of je ooit met dit product terug komt bij de praktijk en dan denk ik, in je positieve momenten hé wat grappig – ook al kom ik met lege handen terug ik heb nog steeds veel geleerd, dat heb ik wel altijd gehad hoor, maar ik durf ook niet hard te roepen van ‘hé jongens nog twee maanden en dan heb ik hier toch een presentatie met een algoritme dat wordt me daar toch een toffe toestand’.	During	DUT
kijk ik wilde dat je afstudeert en ik heb zelf veel van het geheel geleerd dat waren twee side-effects die voor mij net zo belangrijk zijn als de uitkomst, maar als ik een bedrijf heb en ik moet betalen voor jou diensten nou dan weet ik niet of ik die tussenperiode had overbrugt.	After During	DUT
het staat niet bovenaan de prioriteitenlijst, maar ik vind wel we hebben ons hieraan gecommit en ik vind het leuk het is ook interessant, weet je. Dus je wilt er ook tijd voor vrij maken, voor jou, maar ook voor onszelf, maar het is echt vaak dat je denkt wanneer kunnen we dit doen want al die andere dingen die moeten! Dat vind ik soms vervelend, dat ik daardoor het idee heb dat we er niet de tijd in kunnen steken die er misschien in zouden moeten steken.	During	NBC
we hebben ons eraan gecommit, het zelf geïnitieerd, we gaan dat afmaken, maar er ligt bij Naturalis niemand wakker als ik dit laat vallen. Dus je moet het echt wel uit jezelf blijven halen	During	NBC
dat is wel iedere keer het dilemma waar ik dan voor sta. Maak ik de tijd vrij of niet	During	NBC
verplichting vind ik het verkeerde woord, want is wel iets wat ik zelf ook wel heel leuk vind en wat ik ook wel graag wil dat het lukt, maar wat ik al zei het duurt best lang, het is niet iets wat binnen twee weken of twee maanden gepiept is.	During	NBC
nu wordt het wel weer concreter, nu is het weer leuker, maar daartussen was even wel zo’n (maakt gebaar van neerwaartse lijn) momentje.	After During	NBC
voor mij zou het zijn als dingen tussentijds zie, goed moet keuren, aan moet geven of het de goede kant op gaat,	During	NBC

Table C.2: All excerpts from the stimulated recall sessions coded During

Quotation Content	Codes	Dept.
toen kwam dit en toen dacht ik: verdomd dat is eigenlijk heel concreet!	After	DUT
meer omdat je zelf natuurlijk ook veel met machine learning en deep learning te maken hebben en ik snap nog steeds niet, nou ik snap het verschil wel, maar als er dan iets voorbij komt en wijs aan of dit nu deep of machine learning is dat weet ik niet. Daar zit ik af en toe nog wel eens mis. Dus dat was meer een vingeroefening zeg maar.	After	DUT
merk dat ik bij dit soort praktische dingen het heel erg meteen probeer te vertalen naar 'Hoe gaat het er dan uitzien als ik het echt zou gaan gebruiken?'	After	DUT
ik weet inmiddels vanuit ervaring ook wel iets meer van de valkuilen en de dingen waarom ik enthousiast kan zijn in eerste instantie en dat het dan toch weer weg zakt.	After	DUT
het feit dat je een concreet uitgangspunt nu hebt van waaruit je verder kunt bouwen is voor mij wel heel waardevol, omdat je je dan iets kunt voorstellen bij wat er kan gebeuren en in het theoretische deel mis je dat gewoon denk ik.	After During	DUT
in het begin van het traject vrij concreet van 'dit willen we', we hadden er allemaal al een soort beeld erbij, dan wordt het allemaal moeilijk en ja nu zeggen we 'doe ff makkelijk klop het effe in'. Maar dat je ook wel beseft dat het er wel bij hoort, dat soort bewustwordingsstappen om er te komen. En dan vraag ik me mezelf meteen af wat betekend dat dan en dat we nu aan het eind zo extreem aangenaam verrast zijn? Zeg ik dat dan te hard?	After Before	DUT
het is bijna 'in your face', want ik heb geroepen op een podium 'jongens we moeten technologie wel durven omarmen als communicatieprofessionals', je moet er in duiken. Nou vervolgens moest ìk dat doen en dacht ik Jezus wat een toestand en dus ja, duh, 'practice what you preach', dan moet je het ook doen,	After Before	DUT
kijk ik wilde dat je afstudeert en ik heb zelf veel van het geheel geleerd dat waren twee side-effects die voor mij net zo belangrijk zijn als de uitkomst, maar als ik een bedrijf heb en ik moet betalen voor jou diensten nou dan weet ik niet of ik die tussenperiode had overbrugt.	After During	DUT
Het begon met 'we gaan een algoritme bouwen als collega' en kennelijk, ben ik, snap ik als ik nu terugkijk, was ik daar zo ver dat ja dat is ook een echte collega,	After Before	DUT
het is géén expertsysteem, maar dat je op redelijk eenvoudige wijze inderdaad tot dit soort resultaten kunt komen vind ik tof, uiteindelijk	After	DUT
dan zie je dat het toch wel een behoorlijk proces is om die collega goed te vinden en vervolgens moet je hem dan ook nog behoorlijk inwerken. Daar komt dat dat hele puppyverhaal weer van je moet haar wel trainen en alles leren, en dat is eigenlijk wat ik mijzelf hier zie doen, van wat kan deze pup al? Wat moeten we hem nog leren en wat weet hij al, wat kan die al en is de basis wel goed genoeg om hem te kunnen trainen?	After	NBC
het is natuurlijk leuk als het werkt, want dan kunnen we er wat mee maar daar ging het niet om, het ging vooral om het proces, te ervaren en te zien en te doen wat er uit komt en hoe dat dan werkt	After	NBC
eerst was ik eerst inderdaad van 'hè wat leuk', we hebben echt iets, het doet het. Toen kwamen er ook namen van journalisten die ik niet ken dus ik denk dit gaat wat opleveren, leuk, ik ben toch wel resultaat gericht uiteindelijk,	After	NBC

Table C.3 continued from previous page

Quotation Content	Codes	Dept.
het heeft natuurlijk ook allemaal met vertrouwen te maken. Je gaat zo'n algoritme echt gebruiken op het moment dat je erop kan vertrouwen, dus je gaat hem toetsen, een beetje uitproberen, en dit is één van de toets dingen en dus – maar ja je was nog niet klaar, dus je ging er gewoon aan verder werken – maar dat is wel wat er gebeurd, je gaat wel kijken van 'wat kan die dan'?	After	NBC
gelijk aan het denken van wat zijn dan de goede testcases en welke zouden we juist moeten proberen om te zien wat wel of niet resultaat geeft.	After	NBC
het is natuurlijk hartstikke leuk als het werkt, maar het ging ons ook om het zelf te ervaren van wat er gebeurd en dan is dat vertrouwen, daar wordt je zo met je neus op gewezen, dat je denkt ja wil je zoiets nou echt gaan gebruiken, dan moet je wel zeker weten dat het klopt	After Before	NBC
het is natuurlijk wel een instrument wat je kan helpen dus wanneer ga je nou blind hier op af? Wanneer neem je nou echt gewoon dat je denkt van 'neem mij maar mee aan de hand'? Ik denk dat je toch, en dat gebeurde bij mij, dat ik dacht je gaat eerst nog een hele tijd krijgen dat je hem mee aan de hand neemt. Het is nog niet meteen dat hij ons gaat leiden.	After	NBC
het algoritme leid ons nog niet meteen, kan ons wel verrassende inzichten geven, opmerkelijke dingen doen, waardoor we correcties gaan aanbrengen of denken hé interessant, maar het is niet dat jij mij nu straks iets toestuurt, een linkje roep ik maar even, en we gaan lekker aan de gang en dat was het dan,	After	NBC
We moeten echt nog wel, ja dan begint het trainen en dat drong wel heel erg tot mij door, volgens mij, toen je die vraag stelde over dat vertrouwen	After	NBC
dat is best logisch natuurlijk, dus ja, dat is helemaal niet gek, maar het is wel iets wat dan ineens doordringt en het is niet het doel wat je gewoon besteld en het doet het.	After	NBC
Je bent hem ook op maat aan het kneden, dus ik denk dat het eigenlijk me dit soort tools misschien wel gebruikelijk is en misschien is dat wel een inzicht wat voor mijn vakgenoten ook echt heel belangrijk is om te weten, omdat wij zo gewend zijn om gewoon iets te bestellen en je trekt het aan en loopt er mee weg of je gaat het gebruiken, maar dat is hierbij dus niet.	After Before	NBC
hoe meer je hem zelf vormt, hoe meer vertrouwen je er natuurlijk in krijgt, want dan weet je wat hij doet en dan krijg je herkenning en als ik dan die lijn trek met een pup dan is het precies hetzelfde	After	NBC
mogen het ook zeven goede zijn en drie niet zo goed? Ja, kan? Ja, dat is precies waar het om gaat, wanneer vind jet het goed genoeg? Wanneer heb je vertrouwen? Nou ik zelf als ik dat zeg dan denk als de helft goed is dan begint het vertrouwen al behoorlijk de goede kant op te gaan.	After	NBC
Ik realiseer me wel heel erg dat dit product bestaat bij de gratie van jou en vervolgens ons wat we er in gaan stoppen en niet dat dit ding zelf iets is, om het zo maar te zeggen.	After	NBC
Ik probeer altijd vanuit de gebruiker te denken. Wij kunnen wel met zijn allen wel begrijpen wat het doet, of waarvoor het is, maar er is ook nog een eindgebruiker die dit proces niet heeft meegekregen en die wel een bol.com stukje ziet, zeg maar, en dat gewoon wil gaan gebruiken en niet een handleiding van honderd pagina's door wil lezen, die wil gewoon plug and play en als ik hier op druk dan is dit het resultaat en als ik daar op druk dan krijg ik een ander resultaat om die reden.	After	NBC
in het begin, niet persé onzeker, maar gewoon nog een beetje proevend was, een beetje van 'nou misschien?' en dat je op een gegeven moment gedurende dat proces kom je steeds zekerder over, ja nee dit moet het zijn?	After Before	NBC

Table C.3 continued from previous page

Quotation Content	Codes	Dept.
het is wel leuk om er een naam aan te geven omdat je dan weet waar je het over hebt en dat moet dan een naam zijn die je aan een pup wil geven, maar het moet ook een onzijdige naam zijn. Het moet niet een hij of een zij zijn.	After	NBC
Maar een Teddy is ook een soort hulpknuffel, neemt je mee, leidraad	After	NBC
Nee, het is echt een langdurig proces, dus je moet een lange adem hebben.	After	NBC
In het begin vond ik het makkelijker om er tijd voor vrij te maken, want dan heb je die energie maar nu moet je wel lang doorgaan.	After Before	NBC
wat zou helpen? Misschien toch een compacter proces, met een duidelijk beeld van dàn, dàn en dàn hebben we afspraken. Ja, misschien is dat het wel? Als je zegt wat maakt het makkelijk, nou 'live'-ontmoetingen, maar dat was (niet mogelijk) door omstandigheden, en de rest is dan denk ik een heel duidelijk proces wat moet er gebeuren, wat verwacht ik van jullie en wanneer dan ongeveer	After	NBC
nu wordt het wel weer concreter, nu is het weer leuker, maar daartussen was even wel zo'n (maakt gebaar van neerwaartse lijn) momentje.	After During	NBC
je moet er zelf energie in steken en dat is natuurlijk met veel dingen van als je er zelf geen energie in steekt dan gaat het je uiteindelijk ook niks opleveren,	After	NBC
dan kan je iets vastpakken en is dan net alsof je een nieuwe collega een hand geeft. Hij is er.	After	NBC

Table C.3: All excerpts from the stimulated recall sessions coded After.