Master Thesis Business & Information Technology

UNIVERSITY OF TWENTE.

Exploring Organizational Network Analysis: A Case Study

Author: David Haemers
Student number: s1733540
Track: Enterprise Architecture & IT management
First Supervisor: Prof.Dr. T. Bondarouk (Tanya)
Second Supervisor: Dr. D. Bucur (Doina)
Date:

Abstract

This research evaluates the use of organizational network analysis in a large network with roughly 1100 participants by using a technical solution (Microsoft workplace analytics) to collect and process the data. Previous works have used questionnaires and surveys on smaller networks to assess its purpose. This research took place at a large Dutch telecommunications company and lasted five months. The company launched a pilot-project to experiment with organizational network analysis. The project team created use-cases that were analyzed and discussed with the managers of the business. The three use cases that are discussed in this research are workload balance, collaboration overload, and organizational rigidity and silos. The use cases are evaluated based on several metrics. The metrics are measured based on data that ranges from the beginning to the end of 2020. Since a global pandemic (Covid-19) has forced several measures on way that people work, organizational network analysis is used as a diagnostic tool to evaluate the way that work life has changed. In addition, organizational network analysis is used as an explorative tool to provide noticeable results to the business' managers. After presenting and discussing the results to the business, an assessment is made on the value of organizational network analysis for the company. The results showed that organizational network analysis is a valuable tool for detecting changes in the behavior of employees after new measures are implemented. The method also showed promising results for finding anomalies in the network based on the three evaluated use cases. The managers of the organization were satisfied with the outcome but expect a more targeted approach with analyses based on their own input before they can develop and implement actions. Conclusively, organizational network analysis proved to be a valuable tool for large organizations that use a technical solution as their tool to collect and process data.

Table of content

1. Introduction

This study investigates the application of organizational network analysis in large organizations. Organizational network analysis is a method that is used for studying people and their relationships in an organization. Companies have little insight into collaboration of teams and individuals across the organization. Understanding collaboration throughout the organization could lead to better business outcomes. The tools that provide the data already exist, but companies have an urgency to know whether using those tools will eventually lead to profit. The goal of this research is to determine whether using tools that analyze network data can lead to an improvement on the business goals that are formulated by the organization. This research could subsequently prove to organizations whether network data is a worthwhile investment.

The case study is conducted at one of the largest telecommunication companies in the Netherlands. This organization is in the process of evaluating the concept of organizational network analytics through a set of experiments based on eight different use cases. To extract and analyze organizational network data a technical solution is used. The network data in scope consists of anonymized mail activity, document sharing, instant messaging and meeting details extracted from the internal software systems of the organization. In addition to this, the company added survey outcomes and general organizational data to the data sources used by the manufacturer of the technical tool.

Throughout the project, various use cases are evaluated with HR leaders, people leaders and executive directors relevant to the teams in scope of the experiment. Based on their interpretation and feedback on the use cases, the value of organizational network analytics will be formulated. The final use cases that are analyzed are: workload balance, collaboration overload, and organizational rigidity and silos. Communication data is analyzed to map the behavior of the employees on how they spend their time, whom they collaborate with, and how much they collaborate.

The analyses on the three use cases are executed to investigate in which areas the organization can improve the most. In addition, the results reveal the areas which contribute the most to the overall collaboration of the network. This information is discussed with the managers in the organization to determine whether organizational network analysis will be used full-time in the future.

Currently, the bulk of studies that research organizational network analysis investigate a small network of ten to fifty people. The tools that they use in these studies consist mostly of questionnaires and interviews. This study differentiates by using a technical solution and a much larger sample group. The organization at which the research is conducted is one of the first companies in Europe that experiment with organizational network data. Over the course of 5 months this organization carried out a pilot project that evaluates the potential value of organizational network analysis for the company.

The aim of this study is to provide scientific proof of the value of organizational network analysis in large companies. The study aims to achieve these goals by participating in the pilot

project of the telecommunications company. From the 1st of November 2020 until the 1st of March 2021 the researcher joined the project team that worked on this project.

The main research question (RQ) of this research is:

RQ: What is a potential contribution of organizational network analysis for business value co-creation?

Answering this question with the necessary scientific proof will be an important discovery in the field of network analysis. It is therefore important to answer several sub questions (SQ) that provide accurate information that legitimate the answer to the main research question.

SQ1: How do the employees in the organization spend their time and which business units spent the most time on collaboration hours and on focus hours?

The first use case looks into the time that employees spend while working. The business units are assessed based on the number of hours employees work, and how many of these hours are spent in collaboration or focus hours.

SQ2: How does the organization collaborate and who contributes the most to the collaboration?

The second use case investigates the collaboration in the organization. The purpose of this use case is to find where collaboration comes from in the organization. The goal is to find which parts of the organization contribute the most to the collaboration in the organization and which parts contribute the least.

SQ3: How does the organization connect, who is at the center and who lies at the peripheral parts of the network?

The final use case aims to find which parts of the organization fulfill a prominent position in the network and which parts are at the edge, one of the use cases investigates how the organization connects.

SQ4: *Considering the analyzed use cases, what are the effects of Covid-19 measures on the behavior of the employees?*

One of the purposes of the project is to investigate the effects of the Covid-19 measures on the organization. Some of the analyses that are performed show data for the period before and after the measures are implemented. The before and after situation are compared to see how the organization has changed.

SQ5: How is the value of organizational network analysis assessed in the organization?

The answer to SQ5 will help to determine when the project will be regarded as successful. To decide whether the tool can add value there must be an answer to how value is assessed. During the research it is therefore crucial to look at how the organization interprets the results of the analysis.

The thesis is structured as follows:

- Chapter 2: This chapter introduces the theoretical background information of organizational network analysis. It discusses the current applications and

functionalities and ultimately highlights some of the most important lessons that are learned.

- Chapter 3: This chapter introduces the research design and the methodology. This chapter further discusses the context of the project and investigates the organization, sample size, and instruments that are used.
- Chapter 4: In this chapter the results of the project are examined.
- Chapter 5: The final chapter of this study will answer the research questions and conclude the project.

2. Background information and related works

2.1 Background information: network theory

Network theory is a way of describing the world in terms of a model, called a network, that allows us to capture the information about the relationship between things [29]. Since connectivity is at an all-time-high with internet and globalization, people are getting more interested in capturing information about relationships rather than the components of an entity. The networks are visualized through sociograms in which each entity is represented as a *node* and all connections between them as *ties*. By visualizing a network in a sociogram it becomes much easier to see how a network is connected. This chapter will discuss background information on organizational network analysis and some of its characteristics.

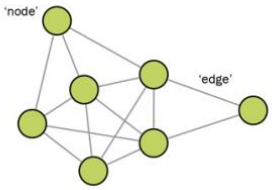


Figure 1 - Edges and Nodes

2.1.1 Organizational network analysis

Organizational network analysis is an application of social network analysis to an organizational entity [8]. Social network analysis is defined as a strategy for investigating social structures [9]. Social network analysis can be applied to many different types of networks such as social media networks [11], terrorist networks [10], sports teams [6], and disease transmission [12].

Social network analysis differs from traditional statistics and data analysis methods [16]. The fundamental discrepancy between social network analysis and traditional methods is the inclusion of *relational* information. Network theorists assume that the behavior of one specific entity influences the others. Patterns in these interactions elicit structure to a network. Structures can be behavioral, economic, political, or social, which grants social network analysis to have many purposes and thus a broad interdisciplinary appeal.

Social network analysis has become interesting for organizations as informal networks are progressively important contributors to performance and job satisfaction of employees [1]. SNA in an organization makes *invisible* interaction patterns of employees visible, which makes it possible to facilitate effective collaboration through group creation. This application of social network analysis in organizations, referred to as organizational network analysis, uses behavioral data that is used to understand how people work and change business strategies [14]. The information that is retrieved from network analyses within organizations can help answer crucial business questions such as: [18]

- Who are our most influential employees?
- What are good leadership characteristics?
- Where in the organization are we collaborating well, and where are we not?
- How immersed are employees in the organization?

2.2.1.1 Passive and active data collection

There are two different ways to collect data for organizational network analysis. The first way is a survey approach in which the data on people's collaboration habits is obtained by asking the employees directly. The second approach uses data that already exists within the organization, for example Email logs, chat logs, or phone records. These two techniques are respectively referred to as active and passive organizational network analysis.

Active ONA

The foremost advantage to active organizational network analysis is that surveys provide the ability to understand nuanced dimensions of relationships between employees by asking people directly what they get from their relationships. The per contra is that a large body of work is required to ensure that people complete the survey. This is crucial to understand as many of the relations as possible.

Passive ONA

The biggest advantage to passive ONA is that the data already exists from sources like email logs and chat logs. This data can be accessed at any time which ensures that there is no reliance on individuals sharing information. However, the employees need to approve the way that the data is used, which can be easier once all data is anonymized. Besides that, an organization requires to have a big infrastructure to process the data. The bigger the number of employees, the more data needs to be processed, the more resources are needed.

The advantages and disadvantages of active and passive ONA are highlighted in the table below.

	Advantage	Disadvantage
Passive	- Data already exists - No reliance on individuals sharing information	- Need to gain consent from employees - Amount of infrastructure that is required
Active	- Understand nuanced dimensions of relationships	- Large body of work that is required

Table 1 - Active and Passive ONA

2.2 Related works: applications of organizational network analysis

This chapter will discuss the applications of organizational network analysis in the current literature. One of the pioneers in the research field of organizational network analysis is Rob Cross. Cross [1] found that network analysis is a powerful tool for a promoting effective collaboration in a group, supporting junctures in networks that cross boundaries, and ensuring integration following restructuring.

2.2.1 Promoting effective collaboration

There are multiple ways to promote the collaboration in an organization with the use of an organizational network analysis. The most obvious improvements would be to increase the participation of underused employees and reduce the participation of overused employees in the network. Cross [1] mentions that network analysis is an excellent tool for finding people

who are central in the organization and identifying those who are peripheral. The case studies on organizational network analysis reveals those applications.

2.2.1.1 Collaboration overload

Cross [38] is one of the first to address collaboration overload in organizations. Collaboration overload occurs whenever employees collaborate to an extent that they do not have enough time to finish their own tasks. Using organizational network analysis, it is possible to identify overloaded employees in the network and relief them of pressure. This is done by finding the employees that others reach out to for information. By highlighting all the information exchanges in the network and mapping those it is very apparent which nodes are the most overloaded. In the case study [22] organizational network analysis is used as a tool to find the overloaded employees. By looking at the number of relations of each employee in the network they can predict whether they are overloaded. While this study uses questionnaires to determine the interactions of an employee, it demonstrates the techniques that are used for mapping the results.

In [23] the authors propose the development of a web application to identify overloaded points in the network. While this application was not used the possibilities of using network analysis look promising, looking at the number of interactions and completed tasks of an individual could potentially highlight those that are overloaded.

2.2.1.2 Peripheral nodes and silos

The nodes at the edge of a network have the lowest number of connections and are identified as peripheral nodes or silos. These nodes can add to the effective collaboration in the network by increasing their participation. In [19], Cross describes how organizational network analysis is used in an organization to identify those employees that are located at the edge of a network. By mapping the information sharing network of the organization the organization quickly identified those individuals and could take measures to integrate them more.

In [17] the authors use questionnaire data of a project team to map a network that can identify the silos and leaders. By asking the members about their communication links to the other members the authors were able to find the nodes that are at the edges or at the core of the network. While the authors use questionnaire data to map their network, this study does reveal the potential of using network techniques for identifying silos in a network.

[20] analyzes the influencers in an organization with different groups. To improve the innovation levels of the organization the organizational network analysis was used to remove the silos. By studying the relations between each group, the authors quickly identified the groups that were not actively participating in the network. The communication data was gathered by asking the participants who they turn to for decision making or problem-solving activities, which is an effective but time-consuming way of collecting data.

2.2.1.3 Key roles and influencers

Since the research by Cross in 2004, the applications of organizational network analysis have been exploited and more ways have been found to promote the collaboration in organizations. Identifying key roles is now a popular application of organizational network analysis. The key roles or influencers in an organization can be used among other things to spread information quicker in the network[2].

[32] Uses both email logs and questionnaires to determine the information exchange in a network. In a trade-off between time and effectiveness, email logs seemed to be the best way

to determine key roles, with questionnaires taking too long to complete. The results showed that mapping the network of around 50 individuals in a university allowed the researchers to find those individuals in the organization that played a key role but where not necessarily high ranked in the organization.
To evaluate and predict future knowledge flows, the authors of [34] identify the influencers of an organization. In this study the authors invite the participants to join their social network. The social network facilitates a place for the employees to discuss and communicate. The data that is collected from the communication streams is used to predict the future knowledge flows and determine the influencers of the network. While this highlights the effectiveness of organizational network analysis as a tool to define key roles, launching a social network to collect data is very time consuming. A quicker collection method is used in [7]. Organizational network analysis finds the individuals that hold the most power in the network by asking the participants who would help to resolve a conflict in the network. With this information they can map a network that have the most influential people at the core. This demonstrates the further application of using network analysis as a tool to identify the most influential people using active network analysis.

2.2.2 Supporting critical junctures in networks

The second application of organizational network analysis that Cross mentions is the support of critical junctures in the network. Networks can cross functional, hierarchical, geographic, or organizational boundaries. In these situations, people that do not work in the same location or have different operational goals must work together in the same network. A network analysis in such cases can help to identify and support the critical junctures. Since this research paper does not focus on geographical boundaries, they are excluded from this literature review, the same goes for organizational boundaries since the sample excludes external individuals. The functional and hierarchical boundaries are discussed and found in other studies.

2.2.2.1 Collaboration across functional boundaries

Functional boundaries happen when multiple groups coexist in the same network. Organizational network analysis can be used to assess the levels of collaboration between these groups. Cross [1] determines this by looking at the collaborative relationships that exist within and between each group in the network. In [4] the authors describe a case in which organizational network analysis helped to improve the connections between divisions in a large organization. The authors used data collected from surveys on the frequency of the communication of the employees. Mapping the findings allowed the researchers and the organization to see that the different divisions in the organization had little collaboration. [4] correlates cross-divisional collaboration positively with innovation. While it shows the effectiveness of network analysis in displaying the relations between departments, the environment in case study [4] was the research & development (R&D) department in a pharmaceutical company.
[13] applies the same techniques in the R&D department of a multinational high-technology company, focusing on the cross-unit knowledge transfer in the network. The authors use three network measures, tie strength, network cohesion, and network range and correlate the strength of each measure to the amount of knowledge acquired in cross-unit transfers. Using survey data, they conclude that each measure is positively correlated with the knowledge that

is transferred between the units. Indicating that building strong relationships would improve the level of cross-unit sharing.
These cases show that using passive data collection methods it is possible to use organizational network analysis in networks with multiple groups to find opportunities to increase to collaboration. While the techniques are already proven in R&D departments, they need to be further explored in other environments.

2.2.2.2 Collaboration across hierarchical boundaries

Besides the functional boundaries that can exist in a network there are hierarchical boundaries that can be analyzed with organizational network analysis. Each organization has a formal hierarchy with managers and individual contributors. The informal network of an organization can be very similar to the formal hierarchy, which constrains the employees [1]. In fluid organizations the employees do not follow the chain of command that strictly to obtain information. Besides that, the network analysis shows how the positions of the leaders are embedded in the informal network. By looking at the relationship patterns of the members of a group it is possible to see how information enters and leaves. The decision making could be improved when the members of the groups only reach out to specific functional areas.
[24] uses network analysis for highlighting the change leaders in an informal network. By the means of questionnaires, the analysts were able to map a communication and information network to find the places of the most influential leaders. The results showed that the archetypical leaders do not always have the most central position in the mapped networks, indicating that for effective change management managers should be aware of the informal networks that exist beneath the hierarchal structure.
In [21] the same techniques are used to improve the enterprise architecture of a small organization. In this case the authors used email logs instead of questionnaires to map the informal network of the organization. By comparing the organizational structure to the social network, the authors saw opportunities to improve the communication in the organization. While [21] shows promising results for the use of passive network analysis as a tool for looking into the informal network beneath the structural network the sample size is low (50).
[36] uses network analysis on multiple networks of different organizations to see how the network of hierarchical organizations compared to fluid organizations. The authors used surveys to collect data and map networks for the organizations and found that the formal structure of organizations is dependent on the size of the organization.
These studies show that network analysis is an effective tool for analyzing and comparing the formal and informal structure of an organization. Creating a map for the formal structure by putting the leaders at the core and the individual contributor on the edges allows analysts to compare it to an informal network created with a network analysis. While the techniques are proven, most studies use an active data collection method, whereas passive methods have only been used in studies with small sample sizes.

2.2.3 Ensuring integration within groups following restructuring initiatives

The final use for organizational network analysis that is mentioned by Cross [1] is in assessing the health of an informal network after a restructuring or an acquisition. Cross [1] uses network analysis in a case to assess the impact of a significant restructuring initiative of an organization. In this case the network analysis was used to see the effects of combining smaller practices into one global network. A network was mapped that showed how the leaders were integrated by looking at the communication of the individuals in the network.

In [31] the authors use network analysis to examine the results of a coaching intervention in an organization. The authors collected pre- and post-coaching data through questionnaires. The findings showed that the participants had an increase in relationships and more ideas were shared within the organization. This shows that in a small sample (18) and using active network analysis methods, it is possible to detect the results of restructuring initiatives in a network.
In [33] organizational network analysis is used to see the effects of implementing new technology in an organization. In this case network analysis was used to analyze the number and efficiency of the interactions of employees. The authors used questionnaires to collect data before and after the implementation of the new technology. The study showed that network analysis is an effective tool for determining the efficiency of interactions of employees.
The studies on organizational network analysis that examine the integration following restructuring initiatives show that it can be used effectively as a diagnostic tool. In these cases, data is collected before and after a restructuring to demonstrate the effects. However, the cases in this chapter use active data collection methods and are focused on small sample sizes.

2.3 Conclusions of literature

The goal of reviewing the literature on organizational network analysis was to examine the current applications of the tool in cases comparable to this one. The cases that are discussed can all be placed under the three applications that are described by Cross [1]. In all the cases network analysis is used as a tool to either support effective collaboration, investigate boundaries, or look at the effects of restructuring initiatives.
While all the cases investigate different networks of separate organizations, the steps that are taken in all studies are familiar:

(optional) 1. Look into a problem that exists within the network

Most studies start their analysis with a problem that has been in the network for a while. The problems usually revolve around inefficiencies that occur within the network. However, not all studies investigate a specific problem, some studies start at step 2.

2. Create a use-case

The problems are attached to a use-case, which usually are determinants of a real-life business goal of the organization. Some popular use cases are influencers, cross-unit collaboration, or innovation.

3. Collect data

There are two ways of collecting data, i.e., passive, and active collection methods. Active methods include data from questionnaires and surveys, while passive methods use data that is collected through logs of email or other electronic communication means.

4. Map a network

After collecting data, a network is mapped that shows the relations between the individuals. Network measures are used to measure the influence of a node in a network.

5. Analyze results

The visualizations of the network are used to analyze the network and find causes for the problems and actions that will improve the use case.

The steps are found in almost all studies that concern organizational network analysis and reveal some sort of framework that exists within the research field. However, each step has variables that may be different for each case. The next chapter will discuss the differences and similarities of this case study compared to the literature.

2.3.1 Similarities and differences

This case study will roughly follow the steps that have been described in the previous chapter. To determine the uniqueness of this study this chapter discusses how the steps in this study are similar or different to the steps of the studies that are discussed in chapter 2.1

1. Look into an existing problem in the network.
This is an optional step for case studies that use network analysis. Not necessarily all studies investigate an issue but are just looking to find improvements in the network. This case study does not investigate existing issues in the network but will start at step 2.

2. Create a use-case
Most of the studies select one or two use cases to analyze. This case study is an exception since eight use cases are selected, with four of them being used for analyses. While most case studies investigate how organizational network analysis can be used for one pre-selected use-case, this case study starts off by manually selecting eight use cases based on their relevance to the network.
While the use cases that are selected are similar to the use cases that are mentioned in the related works, previously unexplored areas are also investigated. All use cases that are previously studied are related to communication relations in networks. However, this case study also investigates the network load by looking into the number of hours a person works in a week and how they are distributed. So not only does this study investigates more subjects at once, but it also evaluates the application of network analysis for previously unexplored areas.

3. Collect data
The vast majority of studies use active data collection methods. Questionnaires are the most popular method to gather data on the relations of individuals in the network. This study will exclusively use data from passive data collection methods. In addition to that, the sample population of this study is substantially larger than the studies that have previously used network analysis. Most studies use sample sizes of 10-50, while some studies use larger samples of 50-250. This is explainable, since a large sample size and active collection methods are hard to combine. Using active data collection methods on a sample population larger than 250 would require a huge amount of time. Since most studies do not have passive data collection methods available to them, they must use a smaller sample size. This study will use both a large sample size (1000+) and a passive data collection method, differentiating itself from most studies.

4. Map a network
Mapping the network is a mandatory step for all network analysis studies. Creating visualizations makes it possible to easily detect the characteristics of the network. Using software packages such as UCINET [34, 57, 58] allows the researchers to create the analyses for social network data. In this case study the workplace analytics software of Microsoft [59]

is used to create analysis. This software uses a pre-determined set of measures that can be selected based on the analysis that is going to be executed. The visualizations are made in Microsoft PowerBI. The methods of creating visualizations and mapping networks is not exclusive to the research field of network analysis. However, there are no studies that explicitly mention the use of workplace analytics.

5. Analyze the data
During the fifth step the researcher decides which metrics to measure on the network. The metrics belong to a use case and go into detail for a specific characteristic. These metrics are sometimes analyzed once, when the researcher wants to explore the network and determine its status [31, 57]. However, the metrics can also be measured multiple times when the researcher wants to determine the effects of a phenomenon that occurred in the network. In this case the analyses will use the metrics for both purposes: explorative and diagnostic.

6. Evaluate the results
There are multiple ways to interpret the results that appear through a network analysis. Some studies merely describe the phenomenon that occur in the network, while others actively look for areas to improve and formulate action, there are also studies that look at the results of actions that have been implemented in the network. Since this case study is part of a pilot project it only describes the phenomenon that occur in the network and hypothesize on actions that can be formulated to improve in certain areas. Because this study does not have a pre-determined goal which declares it as a success, the effectiveness of the project is based on the feedback that is provided by the managers of the business. Similar to other case studies that closely work with the business they perform the analysis in, the success of this project is determined by the perceived useless of the findings by the business for the future.

While this study in many cases is similar to the existing literature on network analysis, there are three main areas in which it differentiates from the bulk of the studies:

1. The studies on network analysis have a clear goal in mind before starting their analysis. They are looking to find the key influencers or examine the communication of a specific group of people. In this case study there are no clear goals, the results of the data analysis will reveal whether interesting results are found in various use-cases.

2. While there are some studies that use email logs as a data collection method, by far the most common technique to collect data is through surveys and questionnaires. This study uses email logs and logs of all other electronic communication methods that are used on the internal communication software of the organization.

3. The sample size of most studies is not larger than 50. Network analysis is a popular method for small organizations or separate project teams. While there are studies that analyze the network of 50-250 individuals, it is hard to find studies that use a sample population that is representative of a large organization. In this study the sample size is roughly 1100, making it significantly larger than the average study.

3. Method

3.1 Research design and context

A case study is selected as the preferred research design to study organizational network analysis. A case study is a valuable design when concrete, contextual, and in-depth knowledge about a real-world subject is the objective of the research [41]. Since the goal is to find out how organizational network analysis can be a successful tool for organizations the best way to research is to learn by observing how it is implemented in a real organization. A case where network analysis is used as a pilot in a large company is an ideal place to learn how to start the process and what the main advantages and disadvantages are. In this chapter the tool, data, participants, and other elements will be explained in detail.

3.2 Time of research

The research took place in the period of the 1st of October 2020 until the 31st of March. 2021. Due to various delays regarding software licenses that occurred during the start of the pilot the project at the organization was initiated in the first week January. The time period before the start of the project was used to gather information about the subject and read into various studies that looked into the same field of research. Excluding holidays and unforeseen circumstances the time spent at the organization actively participating in organizational network analysis was around two and a half months.

3.3 Participants

The experimental setting involves 1400 of the roughly 7000 employees within the organization. These 1400 employees are functional in strategic functions related to commercial activities, consisting entirely of knowledge workers. The teams have been selected based on their end-to-end dependencies for product & service delivery in the B2B and B2C markets. Prior and during the pilot, employees from the selected teams had the opportunity to opt-out of the pilot, removing their data from the database that is used in the project. Before the start of the project 4.5% of the sample had decided to opt out leaving a total of 1312 unique members in the sample.

3.4 Instruments

The tool that is used for this project is created for organizational network analysis in large sized companies. The manufacturer of the tool is Microsoft, which also supplies organizational software systems which includes all communication streams that employees require (Microsoft Office 365). All of the communication that takes place on the software systems is stored in the cloud of Microsoft. This makes it convenient to analyze using the organizational network analysis tool of the same supplier. Since it is enormous it is impossible to make an analysis using the entire database. This is why parts of the database are picked out one at a time. Using queries, it is possible to select only the metrics and attributes that are required for the particular analysis. After a query is created the data can be uploaded into a data visualization program that makes it possible to generate different kinds of charts by selecting parts of the data and adding different kind of filters. The goal of the tool is to give insight into how employees collaborate, this starts with metadata that is obtained by the manufacturer of the software. This is simple transactional data that describes when and what kind of transactions took place between employees. This data is then processed and mapped to the organizational data.

In this chapter the tool, along with the employee data that is used in the project will be discussed in detail. Besides the tool that processes the organizational network data, a second tool is used in this project that facilitates the creation of data visualizations. Both tools are explained in the following chapter.

3.4.1 Data processing tool

The software that the organization uses during the project is Microsoft workplace analytics. This software is provided by the manufacturer of the internal communication software (Microsoft Office 365). Since the communication software is stored in the cloud of the manufacturer, the data is readily available. By adding external employee data, a database is formed that contains all communication of the organization and can be categorized by the attributes of the employees. The software uses employee data controlled with privacy standards, this data is merged with data from the organization and eventually used to generate insights and change.

The tool takes a holistic view of collaboration in an organization by analyzing multiple levels of an individual relative to the company. First of all, individuals are analyzed based on their own behavior, for example the hours an employee works overtime. Second, the relationships of an individual are looked into. This level of analysis determines how an individual spends their time and who they spent it with. Time spend on meetings with managers or colleagues is the type of data that is retrieved from this level of analysis. Third, an individual's place in his/her business unit's network is evaluated. This step uses an organizational network analysis to map the network and find influencers or bottlenecks. Finally, the analytics can provide an overview of the organization's network as a system, this can be used to find silos or the most efficient or inefficient parts of the company. All four levels of analysis use data that is retrieved from an organizational network analysis.

3.4.1.1 Data queries

Microsoft workplace analytics allows the analyst to generate queries to answer specific questions. Since there is too much data to analyze the entire database, selections of the database need to be picked out in order to interpret it. There are five query types that can be used for in-depth insights: person, meeting, group-to-group, person-to-group, and network metrics. In order to understand the possibilities of the data these query types will be explained in detail below. The full list of metrics per query type is attached in Appendix B.

Person queries

The person query allows the analyst to look at the data from the perspective of individual employees. This query offers a lot of different metrics that make it convenient to analyze each individual on a wide level. Within the person queries there are four broad categories of analysis: emails, meetings, network, and work.

- *Emails* refer to the emails that each individual sent and receives. Besides the total number of emails that a person it also looks at the hours that a person spends on their emails and the number of emails that are sent during meetings or in after hours.
- *Meeting metrics* specify how many meetings each individual attends. There is a distinction between the total number of meetings and the sum of hours of those

meetings. In addition, it also observes the meetings in which a manager of the individual is present and how many low-quality meetings are attended. The tool marks a meeting as low quality when employees are multitasking, attending a conflicting meeting, or when a meeting is redundant.

- *Network metrics* show the network size of each individual. The network is determined by the number and strength of relations of each employee. Besides the internal network size external network for each employee is also provided.
- *Work metrics* refer to how employees spend their time at work. It tells how many hours are spent as focus hours and how many hours are used for collaboration. Furthermore, it is used to see the workweek span for each individual and how many hours of collaboration they generate for the organization.

Meeting queries

Meeting queries can be used to analyze individual meeting trends and understand the relationship between multiple meeting attributes. Since meetings take a considerable share of the time of current employees [25] it is crucial to look for opportunities where their productivity can be increased.

Meeting queries can be used for general information such as the total hours that employees spent in meetings for a period of time. Using the organizational data this can also be applied to look at meeting hours per business unit or per team which makes it possible to find the biggest and smallest contributors. Furthermore, the meeting queries allow the analyst to look at the number of attendees that are multitasking during a meeting. Since multitasking reduces the meetings productivity [26] the goal could be to find and reduce the number of meetings in which attendees multitask. In addition, the number of people with conflicting meeting hours can be analyzed.

Since the total number of meeting hours each month is such a high number and translates to a great sum of money for the organization, this query will be used to find opportunities where the company can reduce the number of ineffective meetings and ultimately increase the productivity.

Group-to-group queries

Group to group queries are used to understand how a team invests their time across the organization. This query filters out groups of employees and lists the time that the people allocated to other groups. By looking at all the outgoing and incoming communication from a group it is possible to see how well they are embedded into the network of the whole organization. In this case group-to-group metrics use meeting and email metrics that look at all the hours and total number of meetings and emails with the investor group A and collaborator group B. This allows the analysts to examine how and with which other groups a group communicates making it convenient to find groups that are at risk of being silos and groups that are at the center of the network.

Person-to-group queries

Similar to group-to-group queries, this query can be used to compare data from one entity to another collaborator group. In this case the entity is one employee instead of a group of employees. Certain individuals might have different patterns of behavior with one group than

others. This can be used for analysis that look into the involvement of individuals in the organization

Network queries

Network queries use some metrics that are familiar in the research field of organizational network analysis. Using measures such as tie score and influence score, these queries can be used to find out who are the best-connected people in the organization.

Each query type is used for solving different problems. In order to be able to understand what sort of analyses are possible with the tool that is going to be utilized in the project it is crucial to understand these query types. Analyses that look into the number of emails sent during a meeting require a different query than analyses that look into the total number of hours an individual works each week. Since this project looks into eight different use cases that each contain different metrics to analyze, many different queries are created that each consist of different data and attributes. However, before the use cases are introduced it is important to note how a query is going to be analyzed, which is done in the data visualization tool.

3.4.2 Data visualization tool

After generating a query, a specific piece of the database is ready for analysis. To do this the query is uploaded in a separate tool: Microsoft PowerBI. This is a business analytics tool that allows analysists to visualize the data. The visualizations are convenient to display the data in a preferred order or category. The visualizations are made in Microsoft PowerBI using the available functionalities. Besides that, it is also possible to use python scrips that can be ran to make visualizations.

3.5 Dataset

As explained in the beginning of this chapter the raw data that is analyzed in this project consists of two components: *internal communication data* and *external employee data*. The first component consists of all the internal communication data that is stored in the cloud of the internal software provider that the company operates on. The second component regards the employee data that is sent from the organization to the manufacturer of the organizational network analysis tool that is used for this project.

3.5.1 Internal communication data

The internal communication data is stored in the cloud servers of the software provider (Microsoft of the organization. This data consists of all the phone calls, emails, meetings, and instant messages that are sent through Microsoft office 365. An important note to make is that this data does not make up 100% of the communication streams of the organization since phone calls can be made on an employee's private mobile phone. Furthermore, messages can be sent through other services such as WhatsApp. While the vast majority of the communication streams are captured it is beyond the bounds of possibility to get ahold of all communication data of the organization.

3.5.2. Organizational data

The organizational data that was send to the provider of the consists mostly of data that helps categorize the internal communication data and link it to specific results. Table 2 shows a list

of all the attributes that are used. Appendix A describes all the attributes in more detail and highlights what sort of analyses can be made with each attribute.

Attribute	Description
PersonID	Email address of person
EffectiveDate	First day of employee at the organization
LevelDesignation	Level of employee in the hierarchy of the company
Organization	Business unit within the organization
ManagerID	Email address of the manager of the person
SupervisorIndicator	Management level
Gender	
Fte	Hours worked equivalent of fulltime position
Hours	Hours worked by person in a week
EmployeeType	Internal, specialists, temps, others
ContractType	Type of contract of the person
Department	Department of the person
Team	Team of the person
SubTeam	When available, lowest level team, sometimes subdivision per region
Office	Location of office of person
Origin	Hired by legacy organizations or merged organization
HourlyRate	Salary
TimeZone	TimeZone that person works in
Performance Metrics	Manager Appreciation

Table 2 - Organizational data

3.5.3 Combining the components

Both components are combined to create a database that not only has data on all the communication of the organization but can also link this communication data to specific groups and make multiple categorizations. Having the ability to filter the data based on the attributes that are added by the organization makes the data substantially more interesting for analysis. The organization is made up of several business units, understanding each statistic per business unit is valuable information for the managers of each unit. Besides these categorizations the combination of the two different components also allows analyses that show statistics per instance of the attribute, for example male vs female, or individual contributor vs manager. Both components are required for the analyses during this project.

Some examples of the raw data from a meeting query and a person query are shown in Figures 4 and 5 respectively.

MeetingId	StartDate	StartTimeUTC	EndDate	EndTimeUTC	Invitees	Total_redundant_hours	Redundant_attendees
040000008200E00074C5B7101A82E00800000000405A13	10-9-2020 00:00:00	08:00:00	10-9-2020 00:00:00	08:30:00	2	0	
040000008200E00074C5B7101A82E0080000000040EA73	11-12-2020 00:00:00	08:00:00	11-12-2020 00:00:00	08:30:00	2	0	
040000008200E00074C5B7101A82E0080000000010624D	19-10-2020 00:00:00	13:30:00	19-10-2020 00:00:00	14:00:00	2	0	
040000008200E00074C5B7101A82E00800000000B03DB7	11-12-2020 00:00:00	08:30:00	11-12-2020 00:00:00	09:00:00	2	0	
040000008200E00074C5B7101A82E0080000000037F8E5	9-1-2020 00:00:00	10:30:00	9-1-2020 00:00:00	11:00:00	2	0	
040000008200E00074C5B7101A82E0080000000060FDF6	11-12-2020 00:00:00	09:00:00	11-12-2020 00:00:00	09:30:00	2	0	
040000008200E00074C5B7101A82E00800000000D09C37	12-3-2020 00:00:00	09:30:00	12-3-2020 00:00:00	10:00:00	2	0	
040000008200E00074C5B7101A82E0080000000020C387	3-1-2020 00:00:00	10:00:00	3-1-2020 00:00:00	10:30:00	2	0	
040000008200E00074C5B7101A82E00800000000205BCE	3-1-2020 00:00:00	13:00:00	3-1-2020 00:00:00	13:30:00	2	0	
040000008200E00074C5B7101A82E00800000000703966	10-9-2020 00:00:00	11:30:00	10-9-2020 00:00:00	12:00:00	2	0	
040000008200E00074C5B7101A82E0080000000010BBDC	27-7-2020 00:00:00	12:00:00	27-7-2020 00:00:00	12:30:00	2	0	
040000008200E00074C5B7101A82E00800000000901CC5	6-1-2020 00:00:00	10:30:00	6-1-2020 00:00:00	11:00:00	2	0	
040000008200E00074C5B7101A82E00800000000E00718	6-1-2020 00:00:00	12:30:00	6-1-2020 00:00:00	13:00:00	2	0	
040000008200E00074C5B7101A82E0080000000080039E	12-3-2020 00:00:00	15:30:00	12-3-2020 00:00:00	16:00:00	2	0	
040000008200E00074C5B7101A82E00800000000E052BF	6-1-2020 00:00:00	13:30:00	6-1-2020 00:00:00	14:00:00	2	0	
040000008200E00074C5B7101A82E0080000000090E648	11-9-2020 00:00:00	07:30:00	11-9-2020 00:00:00	08:00:00	2	0	
040000008200E00074C5B7101A82E00800000000408108	28-7-2020 00:00:00	08:30:00	28-7-2020 00:00:00	09:00:00	2	0	
040000008200E00074C5B7101A82E00800000000E05F8A	28-7-2020 00:00:00	08:30:00	28-7-2020 00:00:00	09:00:00	2	0	
040000008200E00074C5B7101A82E00800000000B0D740	28-7-2020 00:00:00	09:00:00	28-7-2020 00:00:00	09:30:00	2	0	
040000008200E00074C5B7101A82E00800000000D0A5B5	20-10-2020 00:00:00	12:30:00	20-10-2020 00:00:00	13:00:00	2	0	

Figure 2 – Raw meeting data

Figure 2 shows a small part of the database originating from a meeting query. In this Figure the *meetingid* refers to any unique meeting that took place in the organization. This query uses mostly internal communication data. However, more attributes are available that allow the analyst to see for example when a supervisor has joined the meeting. This information can only be obtained from the organizational data.

PersonId	LevelDesignation	Organization	SupervisorIndicator	Gender	Hours	EmployeeType	ContractType	Departme
E2AE50322E42D822B4A32BDFF097584FD70A2DFF577371	Office & Technology TW12	B2B	IC	Male	40	Internal	ONB	B2B MARKE
E2AE50322E42D822B4A32BDFF097584FD70A2DFF577371	Office & Technology TW12	B2B	IC	Male	40	Internal	ONB	B2B MARKE
E2AE50322E42D822B4A32BDFF097584FD70A2DFF577371	Office & Technology TW12	B2B	IC	Male	40	Internal	ONB	B2B MARKE
E2AE50322E42D822B4A32BDFF097584FD70A2DFF577371	Office & Technology TW12	B2B	IC	Male	40	Internal	ONB	B2B MARKE
E2AE50322E42D822B4A32BDFF097584FD70A2DFF577371	Office & Technology TW12	B2B	IC	Male	40	Internal	ONB	B2B MARKE
E2AE50322E42D822B4A32BDFF097584FD70A2DFF577371	Office & Technology TW12	B2B	IC	Male	40	Internal	ONB	B2B MARKE
E2AE50322E42D822B4A32BDFF097584FD70A2DFF577371	Office & Technology TW12	B2B	IC	Male	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE
6E5EE43B6482E784E6FE91164F850C2648C26CC56B9460	Office & Technology TW13	B2B	IC	Female	40	Internal	ONB	B2B MARKE

Figure 3 – Raw person data

Figure 3 shows a small part of the database originating from a person query. This query showcases the organizational data that was send to the manufacturer of the tool. In this Figure the attributes explained in chapter 3.5 are clearly visible. The *personid* refers to a unique employee. One unique *personid* can appear multiple times in the database since it refers to an instance of one month per employee.

3.6 Examples

This chapter looks at some examples of visualizations that were created in the data visualization tool. These examples use different queries and techniques to analyze the data. Some of the data is anonymized to protect the privacy of the organization. This chapter will

highlight some of the functionalities of the data visualization tool by explaining the respective Figures and what their purposes are.

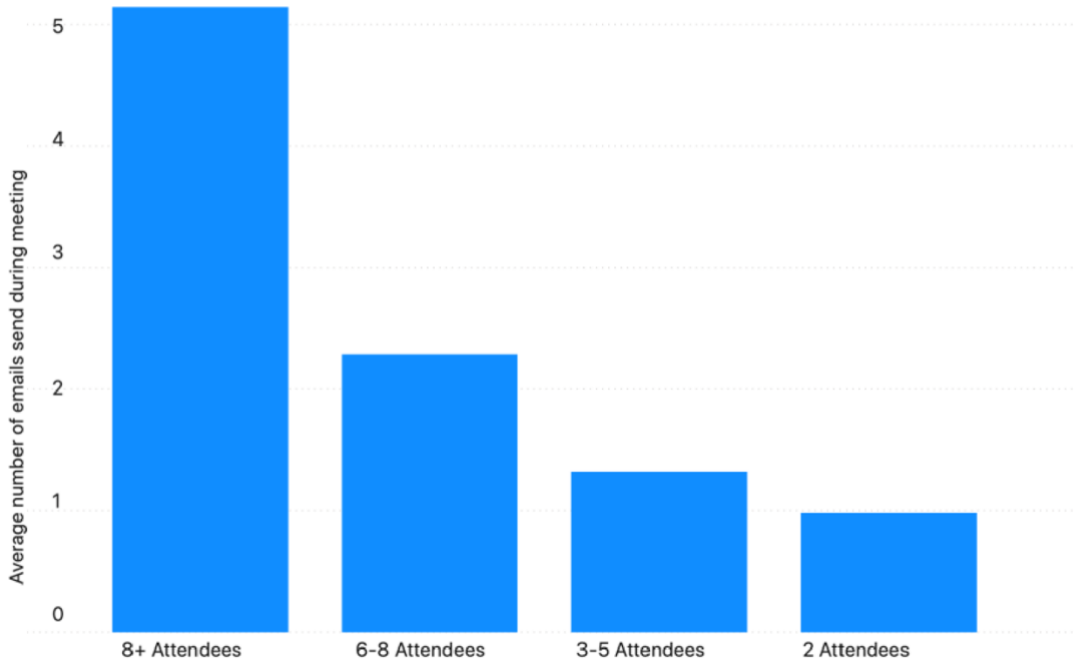


Figure 4 – Average number of emails sent per number of attendees of a meeting. *Basic example to show the functionalities of Microsoft PowerBI.*

Figure 4 is created by using a meeting query and investigates the number of emails that are sent during a meeting by categorizing meetings in number of attendees. On the X-axis the meetings are split into meetings with 2 attendees, 3 to 5 attendees, 6 to 8 attendees, and more than 8 attendees. Since the meetings in the database only show the exact number of attendees, the attendee buckets must be created manually. This is done in the data visualization tool by writing a DAX function. This function generates a new column in the database that categorizes a meeting based on their attendee bucket with the mentioned numbers of attendees. The Y-Axis of this Figure shows the average number of emails that are sent during a meeting. This is one of the metrics of the meeting query functionality of the data processing tool. This Figure was made to look at the different behaviors of individuals in certain types of meetings. These types of analyses help in determining what can be classified as an effective meeting and a non-effective meeting.

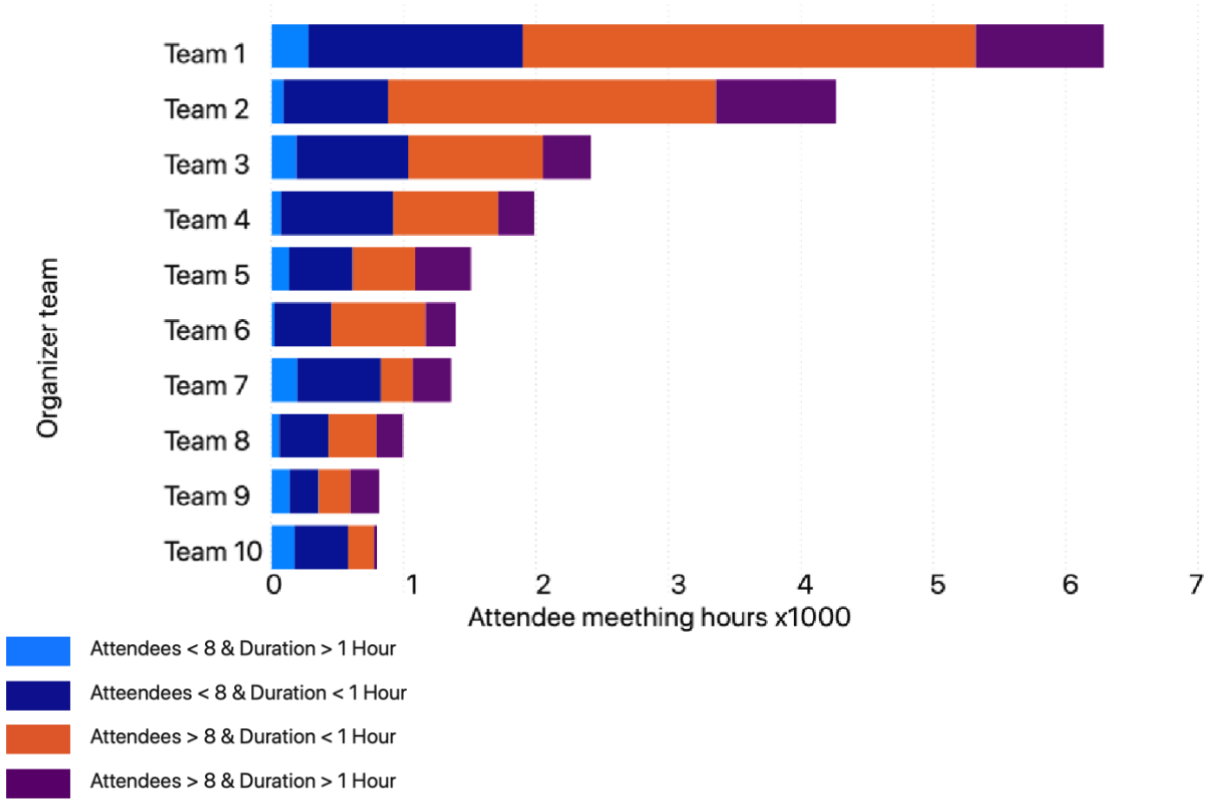


Figure 5 – Attendee meeting hours for top 10 teams most generated meeting hours. Basic example to show the functionalities of Microsoft PowerBI.

Figure 5 shows an example of the combination of a meeting query and a group-to-group query. On the Y-axis of this example the teams are mentioned. In this specific example these teams are the top ten teams in the organization that produce the most meeting hours. These are filtered by first creating a list with all the teams and their respective generated meeting hours in the organization. By sorting them descending the top ten teams are retrieved. On the X-Axis of this Figure the total number of generated meeting hours is shown. This is calculated by calculating the sum of all the generated meeting hours of the individuals in the team. The different colors in each bar represent the category of the meeting. As explained in Figure 4, creating a DAX function in the data visualization tool allows the analyst to categorize the data in a customized way. In this DAX function, the meeting is categorized based on the number of attendees and the number of hours of a meeting. For example, the orange part of each bar represents the total number of generated meeting hours of meetings that have more than 8 attendees and last for 1 hour or less. This shows the potential of combining attributes to generate a new category for the analysis. This visualization was made to look at the teams that generate the most meeting hours for the company. Later this visualization changed to also show the number of generated meeting hours per meeting category. This helps to highlight the teams that bring the most collaboration to the organization and in addition to that see to which types of meetings the collaboration extended.

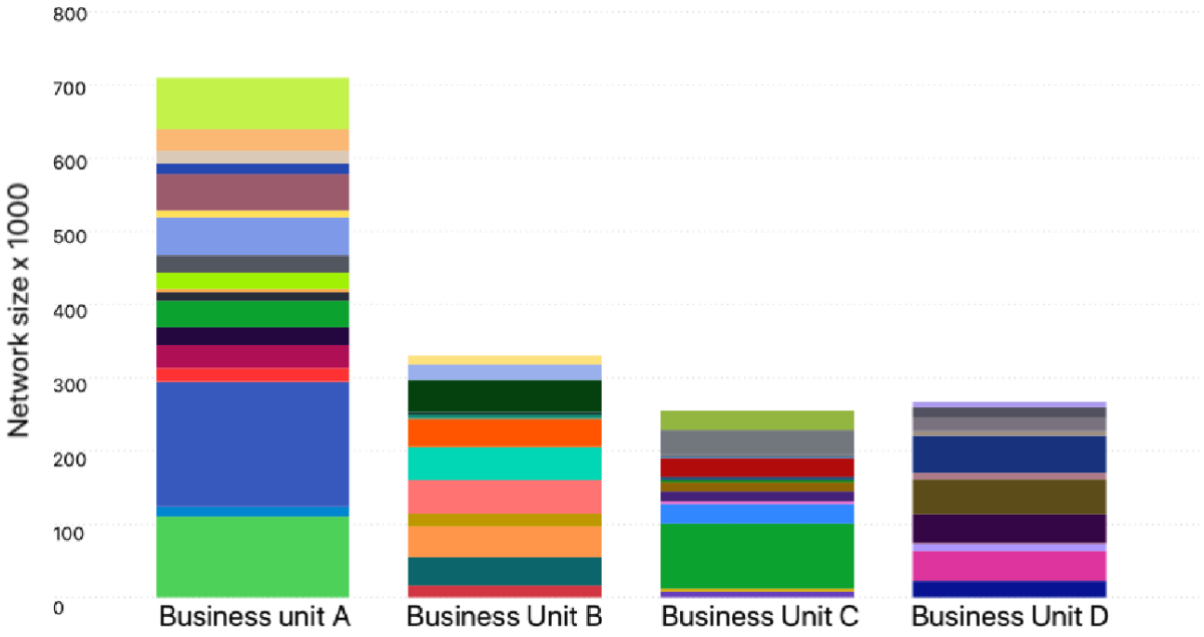


Figure 6 – Network size per business unit. The different colors in the bars refer to the teams in the business unit. Basic example to show the functionalities of Microsoft PowerBI.

Figure 6 shows an example of a combination of network query and person-to-group query. On X-axis it shows the name of the business unit. There are four business units in this data analysis, most of the analyses that were made show the statistics per business unit. A business unit is made up of multiple teams. In the bar these teams are represented by a color. On the Y-axis it shows the network size of each business unit. A network of a person is calculated by the number of people they had at least two meaningful interactions with. A meaningful interaction can be an email, phone call, meeting, or three instant messages. In this Figure the network size of each individual is added to that of their teammates to show the network size of the whole team. The network size of the whole business unit is then calculated by the sum of the network size of the teams. The sum of the network size in Figure 6 is high because the data in this picture is not averaged for network size per month, which is common in the analysis in chapter 4. Instead, the network sizes for each employee per month are added and since there are 12 months of data the network size is substantial. This Figure was created to get more insight in the network sizes of each team in the organization. In addition to that is clear that business unit A has the largest network size by a large margin compared to the other business units. Since all business units have roughly the same number of employees this information helps in determining the most influential parts of the business.

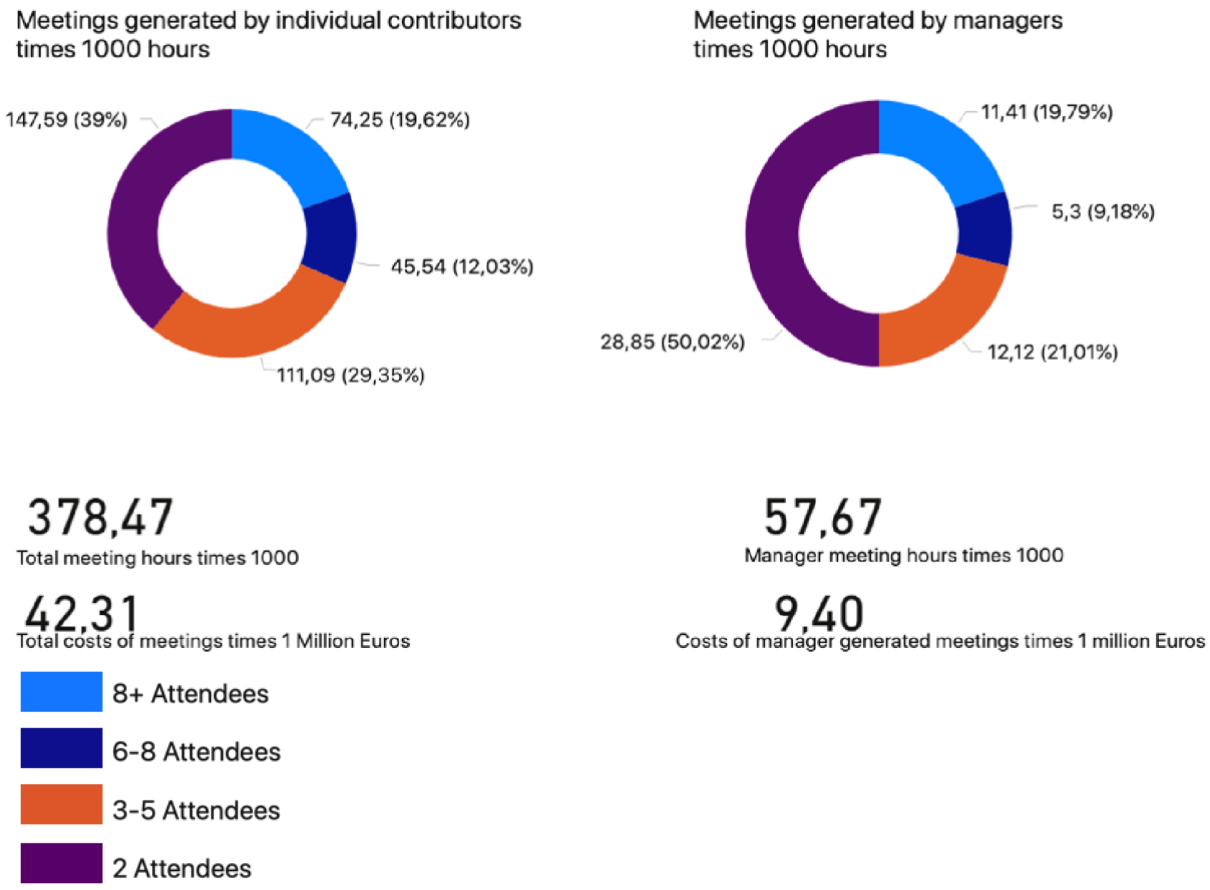


Figure 7- The number of meetings generated by individual contributors and managers. Categorized by number of attendees of the meeting. Includes the costs of the generated meeting hours. Basic example to show the functionalities of Microsoft PowerBI.

Figure 7 shows an example of the combination of a meeting query and a person query. This Figure uses the same categorization as Figure 5. The meetings are split into four categories that split the meetings based on the number of attendees and the duration of the meeting. The chart on the left shows the total number of generated meeting hours for individual contributors (IC). An IC is an employee that is not registered as a manager, this holds true for the large majority of the sample size. The chart on the right shows the generated meetings by managers and managers+. These refer to all the employees that are registered as a manager (manager), or manager of a manager (manager+). On the bottom of both charts the total number of hours of the meetings highlighted. In addition to that the total costs of all the meeting hours are shown as well. This is calculated by adding the hourly wages of all attendees in the meeting. This Figure is helpful to determine where the collaboration in the organizations coming from. It also helps to determine whether the organization is operating with a top-down or bottom-up management.

3.7 Procedure of the Study

Even though the organizational network analysis project at the organization is just a pilot it is quite an extensive process that requires a mandatory series of steps. In this chapter all of the steps are mentioned briefly before being explained in more detail in the following chapters.

1. Use cases

The first order of business for the study will be to evaluate the eight use cases that have been selected by the organizational network analysis team. Use cases provide a direction in which analysts can go. Since the project that the organization has created is merely a test case there are no pre-set guidelines or goals. The goal of the pilot is to determine whether organizational network analysis is a valuable tool for the organization. Projects like these usually include a goal or problem that has roamed the organization for some time but has not been addressed yet. In this specific project that is not the case, and it is the goal of the project team to show the value of organizational network analysis to the leadership team of the organization. At the end of the project the leadership team can then decide to continue with network analysis or shut it down. In order to prove the value of network analysis the analyst will have to dig into the data and show areas where the organization can improve. To make this an organized task the use cases help to steer direction instead of diving into the data without a clear task.

These use cases were picked before the start of the study but have to be evaluated, nonetheless. Knowledge from prior research in organizational network analysis will tell whether these use cases are suitable for analysis and what methods have been used previously to analyze them. The first step of the study will therefore also involve a literature study that will help to evaluate the potential of each specific use case. All of the eight use cases will be analyzed in chapter 4.

2. Attach metrics

Following the evaluation there will be discussions on the analyzability of each particular use case. The use case by itself will just highlight a direction in which the analysts can study the data. In order to make calculated analyses metrics will have to be added to each use case. A metric will allow the analysts to attach statistics to the use cases. Where a use case provides a category and a direction for the analysts, the metrics provide the opportunity to determine how the organization performs in each category.

The metrics, similar to the use cases, have to be created by the project team themselves. In order to determine the best metrics for a use case the project is divided into four sprints. Each sprint takes two weeks and will handle two use cases. This way it is possible to focus on two subjects at a time, allowing the analysts to have more attention for a use case and to create solid conditions for each use case. For each metric it is important that they can potentially highlight some discrepancy in the organization. The goal is to analyze a metric and determine whether parts of the organization are under- or overperforming on that criterium. Another component of a metric is that it should be used for long-term analysis as well. During this pilot the analysts are looking for anomalies in the data in the present. However, when the organization does decide to continue with organizational network analysis the metrics should be able to continue being used. In the future the organization would also prefer to keep track of the metrics and see whether they are improving on these specific parts.

3. Find literature

The metrics are selected by the project team and do not include boundaries or thresholds that determine whether a statistic is an anomaly or normal behavior. However, categorizing the results and establishing the peculiarity of a statistic is required in order to decide in which areas the organization can improve. In order to create boundaries for a metric the literature on each specific subject has to be explored. For example, when the analysts want to categorize the employees based on the hours they work in overtime, literature is required to determine what constitutes low, medium, or high number of overtime hours.

Prior to analyzing the data, it is therefore important to have these boundaries set. This will help to see whether a metric provides interesting results and whether they should be altered. Not necessarily all the metrics require boundaries, as some of the results will be categorized based on the top or bottom percentage performances. For the results that will be categorized in clear brackets it is however crucial to have boundaries that make sense and based on facts.

4. Analyze data using the tools

After evaluating the use cases and attaching metrics with clear boundaries the data can be analyzed. As mentioned before the project is divided into four sprints in which two use cases will be analyzed. Since the database is enormous and it is impossible to carry out analyses using the full database, little pieces have to be selected that can be analyzed separately. For this reason, queries must be made. A data query selects a piece of the database based on the requirements that the analysts select. The queries are made using the data processing tool and afterwards uploaded into the data visualization tool.

When the data query is loaded into the data visualization tool the analysts can display the data on various graphs. This process will be trial-and-error to find the optimal way to display the data and highlight the anomalies. The data will be analyzed using the functionalities of the data visualization tool. All of the results that are created in this step will be put together in a collection file. This file can then be used in the future to select the most important findings that will be shown to the leadership teams of the organization. During the project, organizational network analysis is used as an explorative and diagnostic tool. The results that come out of the analysis are used to answer SQ4, SQ1, SQ2, and SQ3.

5. Review results with business

The results from the analysis are collected in a file for each sprint. These results will be presented to the leadership team of each business unit. There are four different business units that participate in the pilot project. Each of these business units is made up of multiple teams that each have a team manager. All the managers of the teams in the business unit form the business unit leadership team. The results will be presented to the four business unit leadership teams to validate them and see whether there is enough interest to continue with organizational network analysis in the future. Since the leadership teams are interested in the performance of their own business unit all the results will be separately presented. The results will be presented during a digital presentation in which the leadership team can ask questions and share their general thoughts. Once the presentation is given to all units the project team will know what sort of analysis has the most value for the organization and what they need to investigate in the future. The discussions with the leadership teams therefor play an important

part in the decision to continue with organizational network analysis. The results of the reviews with the business are used to answer SQ5.

6 Decision
In the end the project team does not decide whether the organization continues with organizational network analysis in the future. The human resources manager of the company ultimately holds the power to allocate the funds to realize a future for organizational network analysis in the company. This decision will however be based on the results that the project team can present to the leadership teams and the HR manager. The analysis and the business case will have to convince the business of the value of network analysis. Furthermore, the meetings with the business unit leadership teams need to spark some sort of interest with the organizations' most influential people. Conclusively, organizational network analysis needs to show its value by making the business realize that there are serious improvements to be made in certain areas in the organization.

3.7.1. Delay
Unfortunately, some delay occurred whilst working on the project. Due to some problems with the licensing of the data processing tool the research team was unable to initiate the project at the scheduled date of the 1st of October. Eventually the project team had all the tools available to them in the last week of December. This cut the timespan significantly and some changes to the schedule had to be made to reach the deadlines of the 31st of march. Some family related misfortunes also caused the available time of the research to be shortened. Due to these circumstances the project team had to decide that four sprints would be unrealistic and made the decision to cut this down to two sprints instead. Each of the sprints would still contain the original number of two use cases.

3.8 Variable Covid
An unexpected variable entered the project as in March of 2020 the Covid-19 pandemic hit the Netherlands and multiple measurements had to be taken by the government to repress the effects as much as possible. One of those measurements forced people in the whole country to start working from home as much as possible. This has caused all traditional office jobs to move from their office buildings to employees' homes. Whilst it still possible to execute most jobs without significant issues from home, the work dynamics changed since small informal meetings or quick back-and-forth sessions on the work floor are impossible. The idea of organizational network analysis in the organization came about before the start of the covid pandemic. However, the pandemic and in particular the working from home brings interesting opportunities for analyzing. Since the data was available months before the start of the pandemic it is possible to compare the data from the situation to the situation during the Covid-19 pandemic.

In most cases the data will be analyzed based on a timeline that compares the statistics for each month which allows the analysts to see exactly what the effect of the Covid-19 pandemic has on the employee's behavior. This was of course not an intended angle of approach when the organization started with the idea of organizational network analysis, but it turned out to be an interesting variable that arguably made the project more interesting in some parts.

3.9 Privacy

The privacy of the employees is ensured by the guidelines of the general data protection regulation (GDPR) [60]. The GDPR protects the privacy by making it impossible to trace the data back to an individual. In addition to the GDPR guidelines the employees all received an opt-out before the start of the project. The opt-out informed the participants what their data was used for and what the purposes of the project are. Furthermore, it provided them with an opportunity to decline the project and remove their personal data from the project.

To ensure that the data of the people in a team cannot be connected to individuals the teams that have less than 5 individuals are excluded from the analysis. This is only the case for analyses that include information of the team of an individual. In the analyses where team information is not used as an attribute the individuals in such teams are included.

4. Results

The results of this study are split into two components: results from the analysis, and results from the feedback of the organization. The results from the analysis involve all the findings from the creation of the use cases and the data analyses and are used to answer SQ4, SQ1, SQ2, and SQ3. The feedback from the organization relates to all the information that is gathered during the feedback sessions with the organization and is used to answer SQ5. This will respectively demonstrate what the findings of organizational network analysis in the organization are and how the organization interprets this information.

As explained in chapter 3.6 the first step of the project is analyzing the use-cases that will be investigated during the project. All the use cases are mentioned and explained in the chapter 4.1.1. Following the explanation of the use cases will be the report of the sprints that were executed during the project. The results of the sprint include a description of the steps and the results from the data analysis.

4.1 Use cases

Big projects that involve huge databases with endless opportunities to explore require a plan of action before diving head-first in the data [27]. The organization needs to think of several use cases that are worth investing time into. In previous literature organizational network analysis was used as a method to find answers for a specific issue . Zhylinska, (2020) [3] used organizational network analysis as a tool to assess leadership qualities in a software development team. The entire project is focused on leadership qualities which makes for a very clear use case. Gusmerotti (2020) [5] researched the embeddedness of corporate social responsibility in an organization through organizational network analysis. In this case the researcher again revolves their study around one use case. Contrary to this study, the researchers have a clear goal in mind before they start their analysis. This holds true for the majority of studies that look into organizational network analysis. It is common to use organizational network analysis as a tool for looking into specific issues that exist in a company.

In this case there was no apparent issue or clear-cut objective to research which is why the use cases had to be created manually prior to the launch of the project. The use cases where generated based on the input of the provider of the network analysis tool, other research that studies organizational network analysis, and the characteristics of the company itself. The final eight use cases that were established were created by the lead of the organizational network analysis project and his manager. The following paragraph will mention and discuss the eight use cases that were constructed for this project.

Influencers

One of the most popular use cases that is researched in the field of organizational network analysis is influencers. Currently influencers is a widely used term associated with popular online personalities that hold the power to create trends and encourage their followers to buy the products that they promote [42]. Organizational influencers hold a similar position in a company because they have the ability to influence other employees. According to studies by Rob Cross [28] “It is very common for 3-5% of employees to account for 20-35% percent of the value-added collaborations in most workplaces”. Identifying and understanding the

influencers in a company can be very valuable. Influencers can spread information in a company and ensure that organizational changes are enforced [30].

In organizational network analysis influencers are found in two different ways. The first method can only work in active organizational network analysis because all the employees are asked in questionnaires who they regard as the most influential employee. The second method is used in passive organizational network analysis which involves finding the teams or individuals with the most and strongest connections in the organization.

Workload balance

This use case revolves around the hours that employees spent working at the company. The data that was available for this project allowed the analysts to see how many hours were spent on emails, phone calls, meetings, and focus hours. This use case provides insights into the working hours of the different business units in the organization. Highlighting whether there are anomalies between different units and finding possible dangers in cases where individuals are working towards a burn-out. Since the data is anonymized in such a way that one node cannot be linked to a specific individual, this use case is analyzed on business unit level and team level.

Since workload balance has not been researched in organizational network analysis studies there is only one way to analyze this specific use case. The analysis tool calculates the time spent working by an individual by highlighting the first and last interaction on the internal software system of the company. This also allows the analysts to see how many hours are spent in overtime and how many hours are spent in focus hours.

Organization rigidity and silos

The most common association with silos is storage containers for grain. However, in organizational context it is used as a metaphor for parts of the organization that stockpile information and are not actively sharing it with other parts of the organization [43]. Silos are detrimental to the company since they negatively affect the innovative capacity of an organization [35]. This use case is useful in highlighting those areas in the organization that can improve their productivity by involving themselves in the company to a greater extent. This use case also looks into the areas of the business that have high levels of interaction and similar to influencers find key parts of the organization that function as a ‘bridge’ between teams.

Organizational silos can be found in an organization by looking at the flow of communication between the business units and teams. Teams that have hardly any communication streams have a high chance of becoming a silo in the organization while teams that show high levels of communication can function as a ‘bridge’ in the company.

Low and top performers behavior

A common practice for organizational network analysis is analyzing best practices for employees [5],[37]. Best practices are usually deducted by looking at high performing individuals or teams. These behavioral practices can then be taught to other employees. This use case is popular with other studies that look into increasing the overall productivity of a

company. These studies usually identify high performing individuals and compare their behavior to lower scoring employees.

In this case best practices can be deducted by looking at teams that have high performance levels. By observing these teams and comparing their behavior to other teams one could find discrepancies in for example the way that a team communicates or spend their working hours. If found that high performing teams share some characteristics that lower performing teams lack, then the organization can train other teams to improve on certain attributes of their work behavior.

Collaboration overload

Time spent in collaborative activities such as email, meetings, and phone calls have increased by 50% or more over the last two decades [44]. Nowadays at many companies' employees spend roughly 80% of their time on such collaborative endeavors. There is a huge pressure on employees to give input or advice which can cause performance to suffer. To prevent a decrease in performance and even avoid possible burnout it is important to look at the collaborative behavior of an organization. Collaboration overload is a use case that looks into the amount of time that people collaborate, and which teams generate and receive the most collaboration.

Since a lot of meetings, phone calls, and emails are generated every day it is an apparent task to find which parts of the organization contribute most to the total pool of collaboration hours. Firstly, finding the greatest contributors can help to understand which part of the organization requires the most input from other teams or business units. This can lead to questions whether this part of the business is sufficiently equipped for the tasks they are supposed to execute. The second part of this use case is aimed at identifying unnecessary touchpoints and meetings and ultimately reducing the number of collaboration hours for high collaborating individuals and teams.

Onboarding experience

This use case looks is aimed at analyzing the behavior of new employees and finding what causes some employees to better integrate than others during the onboarding process. Ideally this will lead to an increase in productivity by new employees and reducing the first-year attrition in the company.

Similar to best practices, this use case can be analyzed by identifying behavioral characteristics of employees. For example, the involvement of a manager might affect the success rate of new employees. By analyzing these specific characteristics for new employees that have left the company and new employees that have a high-performance rating, best and worst practices for the onboarding experience can be deducted.

Formal vs informal spans and layers

This use case is created to compare the formal structure of the company to the informal organizational design. The formal structure of the company is embedded in the design of the organization and usually regarded as the pattern that employees should follow[45]. However, each organization also has their own informal structure as some teams or individuals might communicate in a way that does not necessarily follow the lines of the formal structure.

Analyzing the communication streams in the company can be an excellent way of deducting the informal structure of the company. To see how teams connect with other teams and comparing this to the formal structure can provide insight into how well the organization design supports this. Based on this the organization design can be adjusted based on the informal communicates that exist within the workplace. The workplace allocation can even be adjusted according to this. Ultimately the goal is to improve the collaboration within critical communities.

Management best and worst practices
This use case is similar to the use case that looks into best and worst practices of employees. The difference between them is that for this case the behavior of management personnel is analyzed exclusively. Identifying what the top percentage of managers does differently compared to the rest can help to establish the behavior that leads to better engagement score. Ultimately this will help the organization to improve on their overall leadership capabilities.

Refining the use cases
Once the use cases are created and the data is available for analysis there is one more step that needs to be executed before the data can be delved into. Since the organization and the business leaders want to be informed during the process of the pilot there are several touch points at which progress will be discussed. Because of this reason the project was divided into four sprints in which two use cases will be analyzed at a time. As explained in chapter 3.7.1 the number of sprints had to be reduced to two sprints instead. Before the start of each sprint two use cases are picked and further elaborated on before analyzing. Since each use case can be attached to a great number of metrics and key performance indicators (KPIs) a selection of them needs to be formed. For each use case KPIs needed to be created to get the best representation possible. The KPIs are created based on the data that is available and the metrics that the tool offers. To find the best KPIs it is therefore important to understand the database and the possibilities of the tool as explained in the previous chapters.

The metrics that were chosen were mostly generated from scratch and did not come with thresholds or guidelines that pointed towards an ideal number or goal to achieve. Therefore, most of the metrics had to be researched extensively to find which results of each metric are considered outliers or normal behavior.

The KPIs were selected in brainstorm sessions in which the lead operated worked together with the researcher. These brainstorm sessions proved to be an advantageous method since the back-and-forth discussions allowed both parties to come up with new ideas and multiple ways of thinking. These sessions lasted about an hour to an hour and a half for each use case and were further elaborated on during other meetings that took place throughout the workweek.

4.2 Sprint 1 – Organizational network load

The two use cases that were selected for the first sprint are *Workload Balance* and *Collaboration Overload*. These were chosen based on the ability to find compelling results that

were unavailable to the organization before quickly. The main question that is answered by these two use cases is how the workload is spread in the organizational network. While *Workload Balance* focusses on how employees spend their time within the organization, *Collaboration Overload* aims attention at how much collaboration time employees generate and receive. Each of the use cases is divided into several metrics that will be explained in more detail below.

4.2.1 Sprint 1 – Preparation phase

- *Workload Balance*

Workload balance is a use case that is generated to examine how employees shape their workweek. For this use case the data is used to analyze three metrics: Collaboration Hours, Focus Hours, and Workweek Span.

- *Collaboration Hours*

Collaboration hours are made up of the time that an employee spends in meetings and on reading or sending emails.

- *Focus Hours*

In this research focus hours are defined as two or more consecutive hours in a person's calendar that do not contain a meeting in between.

- *Workweek Span*

The workweek span is defined as the sum of hours between the first meeting or email in a day and the last email or meeting in a day.

For this use case it is crucial to collect a lot of research that studies each metric. Workweek Span is a metric that looks at the total number of hours that each employee works each week. Since there are no boundaries available that determine whether our results are noticeable or not, the literature on this subject is needed to determine when a workweek is considered long or short.

Workweek Span

Most of the employees in the dataset are hired on a forty-hour contract, which is considered a regular workweek in the Netherlands. The expected average of workweek span is therefore estimated at around forty hours. For this metric the employees are going to be categorized as latent capacity, having normal-hours workweek, or having long-hours workweeks. To establish the boundaries of each category the literature on work week hours must be studied. In this literature review we use studies that are looking into the working hours that employees are expected to be effective and studies that investigate the effects on mental and physical health.

Sparks et al. (2013) studies the effects of working hours on physical health[46]. This study is a meta-analysis in which they compare the results of 12 studies that investigate the effects of working hours on physical health. These studies use different types of samples which include among other white-collar workers that are studied in this project. The study found that people who work more than 48 hours per week have a higher chance to suffer from heart complains than those who work less than 48 hours per week. In another study conducted by Britain's national regulator for workplace health and safety (Health and Safety Executive) some evidence was provided that relate working long hours to stress or mental health issues[47]. In addition to that it was found that there is strong evidence that people perceive that working

long hours leads to poor work-life balance. The study by Bell et al (2012) tends to agree with this finding as they mention that overemployment (when actual workhours exceed desired workhours) has a significantly negative effect on health[48]. This finding is followed by the conclusion that overemployment occurs significantly more often when people work 40 hours or more.

Studies that investigate the effects of working hours on effectiveness contribute to the evidence that working long hours may have bad effects. Research into call center employees in the Netherlands has found that employees work on average 4.6 effective hours per day which would translate into 23 hours per week[49]. The final study that is used to determine the boundaries mentions that output of employees rises at a decreasing rate as hours increase, in addition they conclude that working more than 60 hours per week does not yield any extra productivity for factory employees [50].

Since the literature mostly associates long workhours with loss of productivity and negative health effects this metric is interesting for the business. However, the literature does not provide decisive boundaries. The boundaries are therefore decided based on averages of the aforementioned studies. Since the average workweek in the sample group is 40 hours, a long-hour workweek will be set at 50 or more. This number is positively correlated with loss of productivity and higher chance of mental and physical health issues. Since there are no clear boundaries that suggest the hours of a short workweek, this number is set at 35. This number implies that on average the employee misses out on one hour of work per day.

Collaboration and Focus Hours

The workweek of employees is made up of blocks of focus hours and collaboration hours. In this project the total number of collaboration hours is calculated by the sum of meeting hours and time spent reading and sending emails. Focus hours are defined as blocks of two hours or more without meetings in between. Since blocks of focus hours are decreasing [39] and collaboration is increasingly rising in organizations [40] it is crucial to find a balance between the two. Literature on collaboration and focus hours has not formulated decisive boundaries when it comes to either of the two terms. However, research in collaborative working did conclude that the harder people work collaboratively, the more important it becomes to have time alone[51]. The same research continues to report that people who spend up to 20 percent of their time working remotely are the most engaged at work out of all the workers that were surveyed. In a different survey almost 90 percent of the participants agree to being more productive at work when their focus time is increased[52]. Since it does not state how many hours are spent in meetings or in focus time it is hard to interpret the data, nonetheless it is an indication that workers in general long for more focus hours. In addition, this study states that ideal blocks of focus hours last from one and a half to two hours. One block of focus hours per day would indicate that people need 8 to 10 focus hours per week, at the very minimum. Literature on meetings indicate that 71 percent of employees the causes of too high levels of collaboration[53].

On the other hand, low collaboration by employees could be caused because teams are operating in silos and the company culture does not support collaborative behavior [53]. Other studies on collaboration in organizations indicate that companies where workers collaborate are five times more likely to perform well[54]. This does suggest that employees

that do not participate in collaboration can increase their productivity by increasing their collaboration hours.

All in all, the literature on focus- and collaboration hours suggest that while both are important, employees work most efficient when they do not put excessive hours into either. It is hard to attach boundaries to the metrics since research on the topics mostly suggests a healthy mix of both. However, ideal blocks of focus hours are reasonably believed to be around two hours. When assuming at least one block of focus hours per day, this finding, along with the research that pointed out that employees that spent at minimum 20 percent of their time are found to be most engaged in the company provides enough reason to set the upper boundary of collaboration hours at 30 hours. In this case it is argued that any employee that spent more than 30 hours on collaboration each week is regarded as an excessive collaborator. Assuming an average workweek of 40 hours this means that these employees spend close to 75 percent of their time collaborating. The lower boundary is set at 10 hours, indicating that any employee that spends on average 2 or fewer hours per day on collaboration is regarded as a low collaborating unit. The boundaries for focus hours are set at the same numbers. This means that any employee that spends 30 or more hours on focus time is regarded as an isolated member of the network. On the other hand, employees that spend 10 or fewer hours on focus hours are seen as overused units that struggle to finish their own work.

Collaboration Overload

This use case is created to see who in the organization generates the most collaboration for others in the organization. This is a good indication to see how much workload an individual adds to the network and where the collaboration time of everyone is coming from. This use case is split into two metrics: generated collaboration and received collaboration.

Generated collaboration

The total generated collaboration equates to the sum of hours a person generated for others in terms of meeting and email hours.

Received collaboration

The total received collaboration is calculated by adding the hours a person spent on meetings and emails received by others.

For these metrics it was decided that group-to-group analysis would be optimal. This means that the generation for each team or business unit is analyzed and compared to others. What this provides is an overview of all the business units or teams in which it is clearly visible which teams generate and receive the most collaboration. A list of the top and bottom teams can then be presented on which the business leaders can debate whether these teams should be on there.

Looking into collaboration overload research did not provide any form of boundaries. This type of literature is similar to research on collaboration and focus hours and suggests a balance in collaboration. Some literature does state that too much collaboration can lead to stress and even burn-out [55]. This does imply that teams that are on the list of high receiving collaborators are prone to these negative effects and could be further investigated based on

the data. Similar research has found that excessive collaboration demands can damage employees' productivity and even their health [56]. The same study by collaboration overload expert Rob Cross has found that reducing excessive collaboration can help organizations realize time savings equivalent to 12 to 16 percent of the total hours put in by the workforce. Furthermore, helping the least-efficient collaborators in an organization towards average collaborating can produce time savings equivalent to 8 to 14 percent of the total person-hours worked in that unit. This does indicate that teams on either side of the collaborating spectrum can increase their productivity by either improving or slowing down on their total number of collaboration hours. While not providing exact numbers of what ideal collaboration hours look like, this is a good indication that the top and bottom teams are likely able to improve on their productivity. People who get too many requests can experience stress and can ultimately burn out, whilst people that do not collaborate miss out on productivity.

Objectives Sprint 1

After finding suitable metrics for each of the use cases that will be analyzed in the first sprint the objectives are set by analyzing research and understanding which data will be available. These steps are both mandatory as it makes it possible to make targeted analyses in the database. The final objectives for sprint 1 are as follows:

Workload Balance

Workweek Span

Analyze the hours that people work each week by looking at the first and last activity of each workday. Categorize the employees as low workweek span (<35 hours per week), normal workweek span (between 35 and 50 hours per week), and high workweek span (>50 hours per week).

Focus Hours

Analyze the total numbers of hours in the week in which employees are not disturbed by meetings. Categorize the employees as low focus hours (<10 focus hours per week), normal focus hours (10 to 30 focus hours per week), and high focus hours (>30 focus hours per week.

Collaboration Hours

Analyze the total hours in the week in which employees spent time in meetings or on emails. Categorize the employees as low collaboration hours (<10 focus hours per week), normal collaboration hours (10 to 30 focus hours per week), and high collaboration hours (>30 focus hours per week.

Collaboration overload

Generated collaboration

Analyze the number of collaboration hours that each team generates for the business. Find the top and bottom teams that produce the most and fewest hours of collaboration. Analyze where the generated collaboration ends up in the organization.

Received collaboration

Analyze the number of collaboration hours that each team receives from the business. Find the top and bottom teams that receive the most and fewest. Hours of collaboration. Analyze from which sources these teams receive their collaboration hours.

4.2.2 Sprint 1 – Results

The results of sprint 1 that are presented in the following chapter are the results as they were presented to the leadership teams of the organization. The results of these use cases will help to answer the research questions that are mentioned in the introduction of the report. The workload balance use case focusses on the working hours of employees and how they spend their time, which will answer SQ1. Collaboration overload investigates where the collaboration comes from in the organization and which parts collaborate the most, which will help to answer SQ2. Both use cases compare data from before and after the implementation of Covid-19 measures, the results of these analyses will help to answer SQ4.

4.2.2.1 Workweek span

The first analysis on the workload balance use case concerns the workweek span KPI. This metric is calculated by the data processing tool by taking the sum of hours between the first interaction and the last interaction of an employee. An interaction is defined as sending an email, attending a meeting, or a call or chat. As explained this KPI is categorized into three different divisions, employees' whose workweek span is less than 35 hours, between 35 and 50 hours, and more than 50 hours. As mentioned in the examples in chapter 3.6 it is possible to create buckets in the data visualization tool. For this analysis the employees are put in one of three mentioned buckets. The next step of the analysis is to separate the employees by their business unit. The final variable that is applied to this analysis is the time of the measured period. The data in the database starts from 9th of January, which means there are two months of data in which Covid-19 is no variable yet. The period between the 9th of January 2020 and 13th of March 2020 is therefore used as a baseline, which is shown in Figure 8. The baseline shows data of the company before the Covid-19 measures were put in place. This data is compared to the measure period, which starts on the 4th of October 2020 and ends on the 4th of December 2020, which can be seen in Figure 9. The measure period was selected because the months before October are not as similar to the baseline due to holidays. Once the data is made into graphs they are interpreted and discussed with the project team.

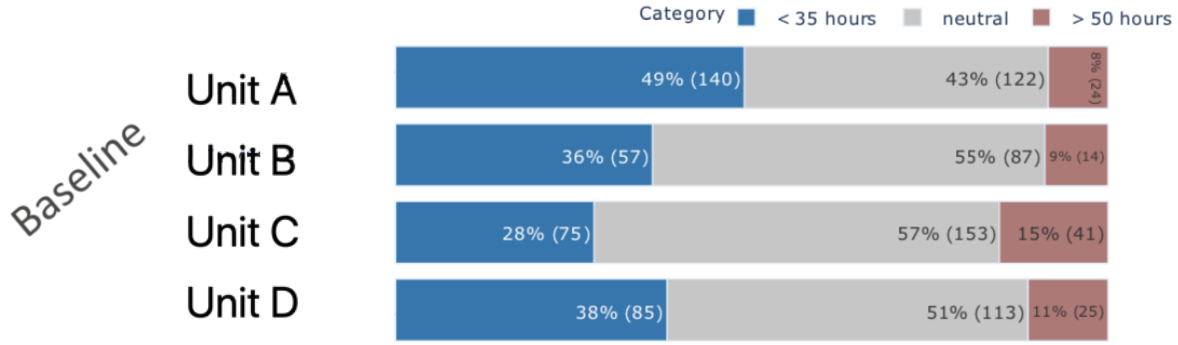


Figure 8 - Workweek span per business unit during the baseline period 09/01/2020 – 13/03/2020

Figure 8 shows the situation before Covid-19 measures and demonstrates that the number of people working less than 35 hours each week is significantly larger than the number of employees working more than 50 hours. Business unit A has the largest group of employees that work in the lowest category and smallest group in the highest category. Business unit C shows the opposite as they have the largest group in the highest category and smallest in the lowest category compared to the other business units.

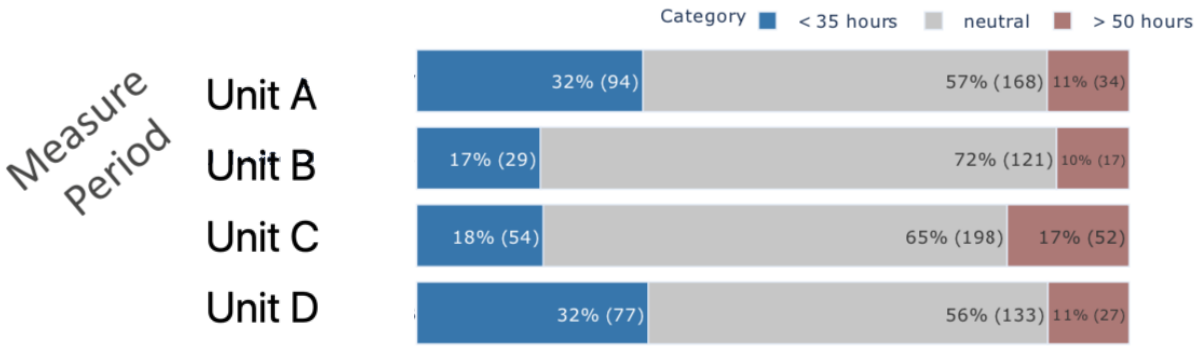


Figure 9 - Workweek span per business unit during the measure period 04/10/2020 – 04/12/2020

The difference between the baseline and measure period is clearly visible. For all business units the percentage of people with a workweek span of less than 35 hours has decreased. Business units A and B have experienced the most significant shift in workweek span dynamics. In both units the percentage of people that work less than 35 hours has dropped by circa 20 percentage points. The business unit with that has endured the least amount of change is business unit D with a reduction of 6 percentage points.
The number of people that work more than 50 hours each week, which is positively correlated with loss of productivity and health issues, has increased for each business unit. Albeit with small margins of absolute numbers, the percentages are compelling. For business unit A the number has increased by 40% and business unit B and C saw an increase of around 25%. The only business unit that did not see a significant increase is business unit D.
Overall, the workweek span has increased for the organization. The number of people that work less than 35 hours has decreased while the number of people working more than 50 hours each week has increased. The number of people working 35 to 50 hours has also increased significantly. While all business units have experienced the same phenomenon, it appears that business unit D is the least affected business unit.

4.2.2.2. Collaboration time

The second KPI that was selected for workload balance is collaboration time. For this KPI the same techniques that were used for workweek span are applied. The employees are put into buckets based on the number of hours they have collaborated in the organization. For collaboration time there are two neutral buckets instead of one. To increase the detail of the information of the graph there are buckets of employees that collaborate 10-20 hours and 20-30 hours each week. The collaboration hours are calculated by the sum of hours spent in meetings or sending and reading emails. The time of the baseline and measure period are the same as the analysis for workweek span. Figure 10 shows the data that was collected before the Covid-19 measures while Figure 11 contains data with active Covid-19 measures.

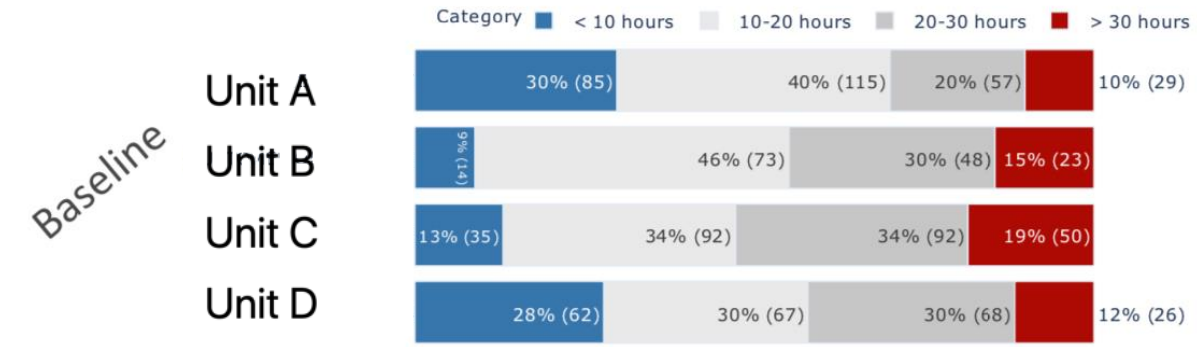


Figure 10 – Collaboration hours per business unit during the baseline period 09/01/2020 – 13/03/2020

Figure 10 shows the situation before the Covid-19 measures and reveal that business unit A and D have the largest groups of employees in the category of people that collaborate less than 10 hours each week. Business units B and C have significantly less employees that collaborate less than 10 hours and show larger groups of employees that collaborate 30 hours or more each week.

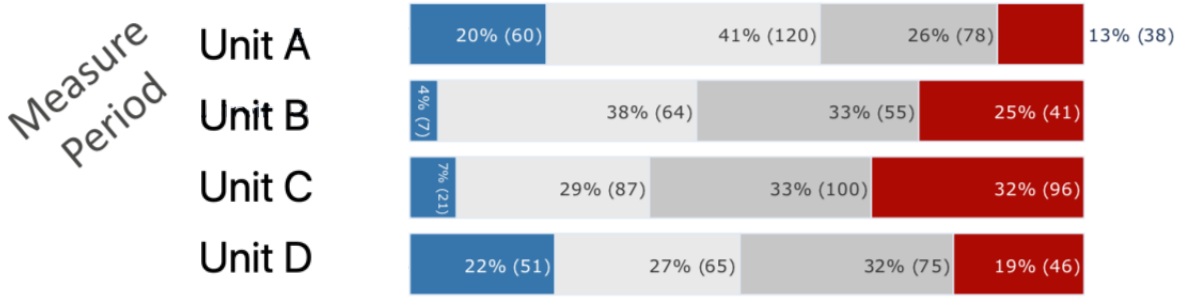


Figure 11 – Collaboration hours per business unit during the measure period 04/10/2020 – 04/12/2020

Comparing Figure 11 to Figure 10 reveals that the number of people that collaborate more than 30 hours each week has significantly increased. For business units B and C, the number of people working in this category has almost doubled, whilst units A and D also saw an increase of about 10 percentage points. The neutral buckets have remained roughly the same for each business unit, indicating that the growth of the 30+ category has come at the expense of the category of employees collaborating less than 10 hours each week. In business units B and C the percentage of employees collaborating less than 10 hours has decreased to 4 and 7 percent respectively. These are the units that saw the highest increase in the 30+ category.
All in all, the graphs reveal a shift in the number of hours that employees collaborate. Across the organization the groups of people that collaborate less than 10 hours each week has reduced and the group of people that collaborates more 30 hours each week has increased. The neutral buckets present the same insights as the 10-20 hour bucket sees a decrease in all units except unit A and an increase for all units except unit C in the 20-30 hour. The increase in collaboration hours could be due to the increase in total workweek span. To determine the exact causes of these analyses the results would have to be discussed with the business in more detail in the future.

4.2.2.3 Focus time

The final KPI that is analyzed for the workload balance use case is focus time. For this KPI the same techniques that have been used for collaboration time are applied. The employees are divided into four different buckets, employees with less than 10 hours of focus time, 10-20 hours, 20-30 hours, and lastly employees with more than 30 hours of focus time. The period

of the analyzed data is the same as the analyses for workweek span and collaboration hours. Figure 12 shows the baseline period before the Covid-19 measures and Figure 13 shows the measure period where Covid-19 measures were in place.

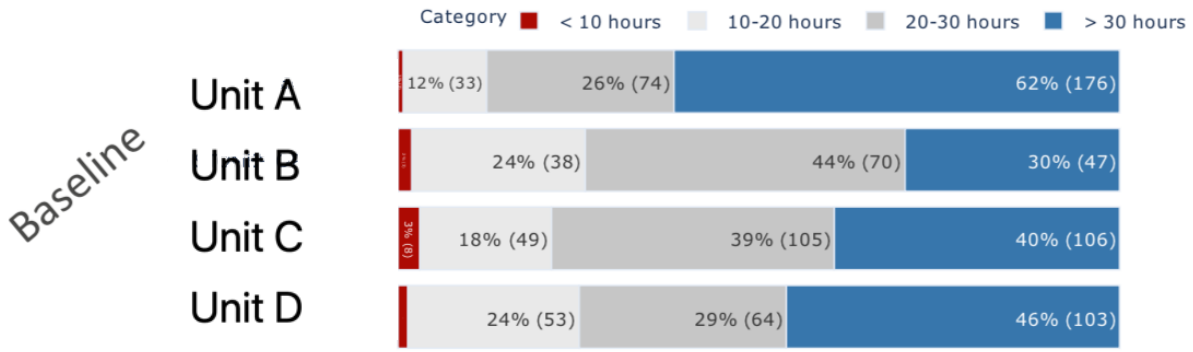


Figure 12 - Focus hours per business unit during the baseline period 09/01/2020 – 13/03/2020

Figure 12 shows that the number of focus hours across the four different business units is quite high for most employees. Around 75 percent of the organization has 20 or more focus hours each week. The number of people that are working with less than 10 focus hours each week is very marginal. Business unit A has the largest group of people in the 30+ hours category, while business unit B has the smallest group.

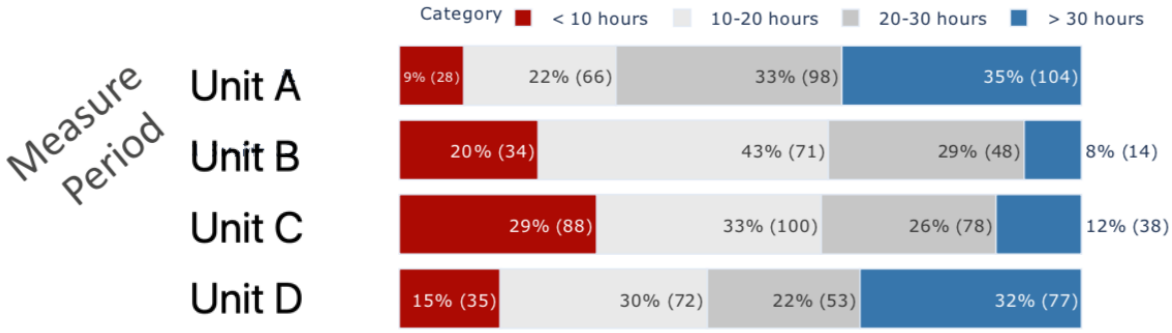


Figure 13 - Focus hours per business unit during the measure period 04/10/2020 – 04/12/2020

As shown in Figure 13, the focus hours have shifted substantially in the organization. The number of people that have less than 10 focus hours each week has seen a huge increase, while the number of people with more than 30 focus hours have decreased by a large margin. For business units B and C, the groups of people with less than 10 focus hours have now become larger than the groups with more than 30 focus hours. Except for business unit D, which has seen a decrease of 14 percentage points, all the business units experienced a 20-30 percentage point drop in the category with 30+ hours. Besides that, the neutral categories have seen a universal change over the organization where the 10–20 hour groups all increased and the 20-30 hour groups all saw a decrease.

Conclusively, the focus hours across the organization have decreased substantially. Employees in all business units have less focus hours. The number of people with less than 10 focus hours, which was barely visible in the baseline graph, is very apparent in measure period. This is noteworthy because according to the earlier literature review into focus hours this means these employees struggle to finish their own work. Whether this is the case for the people in this sample is uncertain, but this could be investigated in the future.

4.2.2.4.Burnout risk

To display the potential danger of having a workweek span that is too high and collaborating too much the results of both KPIs are merged in a burnout risk graph. Based on the literature that was selected for determining the boundaries of each KPI, a potential burnout risk area is identified. The employees in this area have a workweek span of more than 50 hours and collaborate for more than 30 hours each week. Figure 14 shows the percentage of employees in the burnout area for each business unit. The employees are categorized as neutral if their workweek span is below 50 and collaboration hours are below 30. If an employee surpasses the upper boundary for only one of the KPIs they are classified as either high collab or long span. When an employee works more than 50 hours per week and collaborates more than 30 hours they are identified as a potential burnout risk.

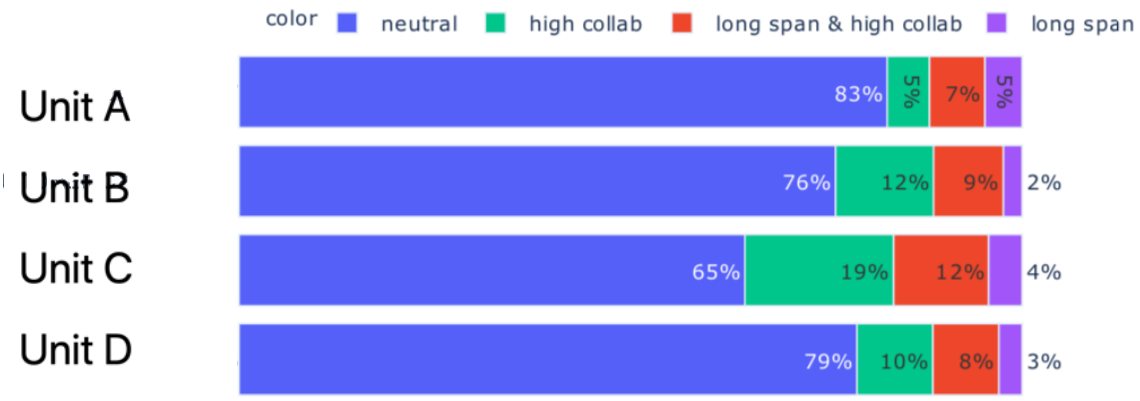


Figure 14 - Burnout risk graph, shows the percentage of people per business unit that have high collaborating hours (30+) and a long workweek span (50+ hours)

Figure 14 shows that around 90% of the employees in the organization are in the neutral zone or surpass 1 of the boundaries. Approximately 10 percent of the employees are categorized as potential burnout risk. The percentage of people with a long workweek span and low collaboration hours is relatively low. The group of employees that are intensive collaborators but do not have a long workweek span is significantly larger. Figure 15 disregards the business units and shows a scatter plot of all the employees in the organization categorized in the same way as in Figure 14. In addition, Figure 15 provides more detail on the number of hours people work and collaborate.

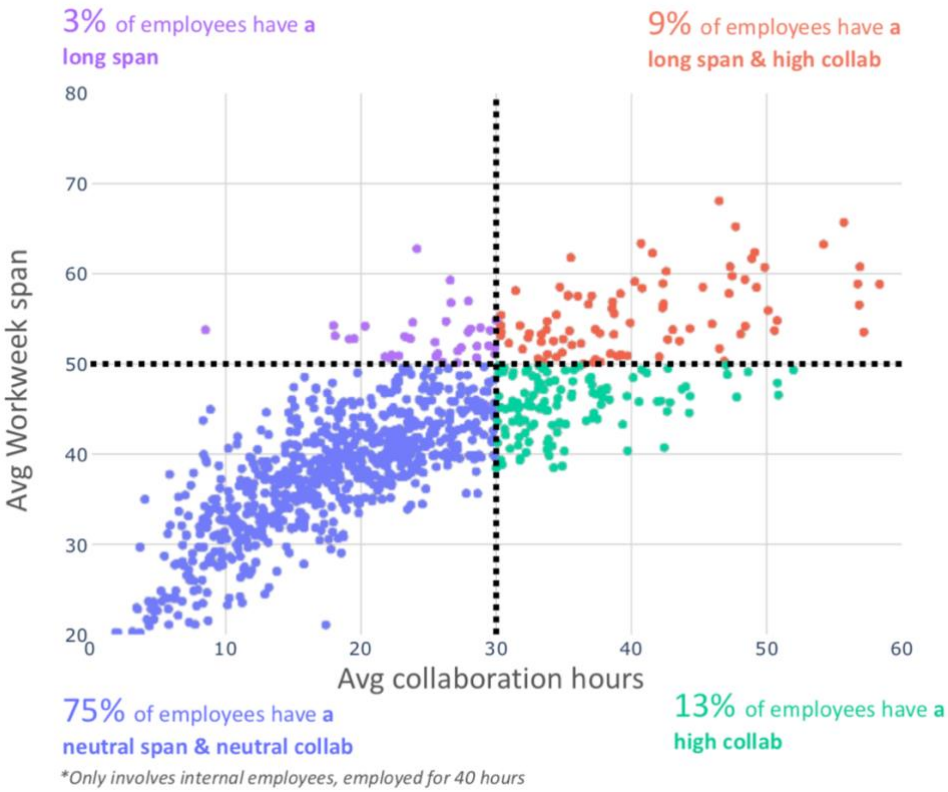


Figure 15 - Burnout risk scatter plot, each node represents an individual and is mapped on the graph according to their workweek span and collaboration hours.

4.2.2.5 Generated collaboration

For the use case collaboration overload the first metric is generated collaboration. This metric investigates where the collaboration in the organization is coming from and how it is spread. This metric calculated by the sum of hours a person generated for other people in terms of email and meetings. The averages of the baseline period and the measure period are shown in Figure 16. As a whole, the organization generated 3 more collaboration hours per week after the Covid-19 measures.

Figure 16 – Average of generated collaboration per week for the organization in the baseline and measure period. Generated collaboration includes the sum of collaboration hours that are generated for other people through emails and meetings.

Figure 16 shows a very basic statistic that is retrieved by averaging the number of generated email and generated meeting hours for all employees in the organization. To understand better where the generation is coming from, the top 20% generators are selected for analysis.

Figure 17 highlights the percentage of people per layer that are in the top 20% generator group. IC are individual contributors who do not have anyone working anyone beneath them. Mngr is a layer of managers, who have people working for them. Mngr+ is one layer above the

manager layer, these are managers who have other managers working beneath them. In Figure 17 it is easily recognizable that the higher an individual operates in the chain, the higher the chance that they belong to the group of people who are in the top 20% collaborators.

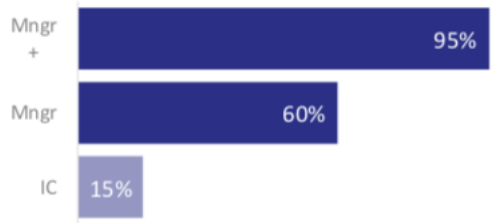
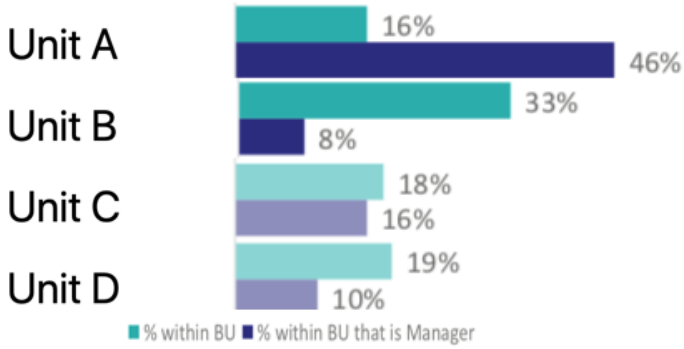


Figure 17 - Collaboration per layer of the organization. IC = individual contributor, Mngr = manager of an IC, Mngr+ = manager of a Mngr. Graph shows the percentage of the group that is in the top 20% of collaborators of the organization.

The next step of the analysis is to look at the generated collaboration per business unit. For this analysis the percentage of employees per business unit that are in the top 20% collaborators are filtered in Figure 18. In addition to that, the Figure shows the percentage of managers in the group of top generated collaborators per business unit. For clarity: 46% of



the group of top collaborators in business unit A (16% of the business unit) are managers.

Figure 18 - Collaboration per layer and business unit. Graph shows the percentage of the business unit that is in the top 20% collaborators and the percentage of that selection that is a manager.

Figure 18 shows that business units A, C, and D all have a similar percentage of people who are in the top 20% of collaborators within the organization. The outlier is business unit B, in which 33% of people are in the 20% top generators of the organization. Business unit B also has the lowest percentage of managers in the generator group, indicating that this business unit has a lot of individual contributors who generate collaboration. In business unit A the percentage of managers in the top contributors is almost 50%, indicating that in this business unit there are little individual contributors in the top 20% of collaborators. Business unit C and D have around 20% of individuals in the top 20% generators, while managers account for 16 and 10% of that group respectively.

To get a broader understanding of the source of collaboration the top 10 teams that have the highest number of employees in the top 20% collaborators are shown in Figure 19. This shows that the top generator team in the organization has 43% of their employees in the top 20% of generators. The other teams in the list show lower percentages but all have more than 20% of employees in the top 20% organization wide. This analysis is made by looking at the

percentage of people in the top 20% for all 25 teams in the database and filtering the top 10 teams.

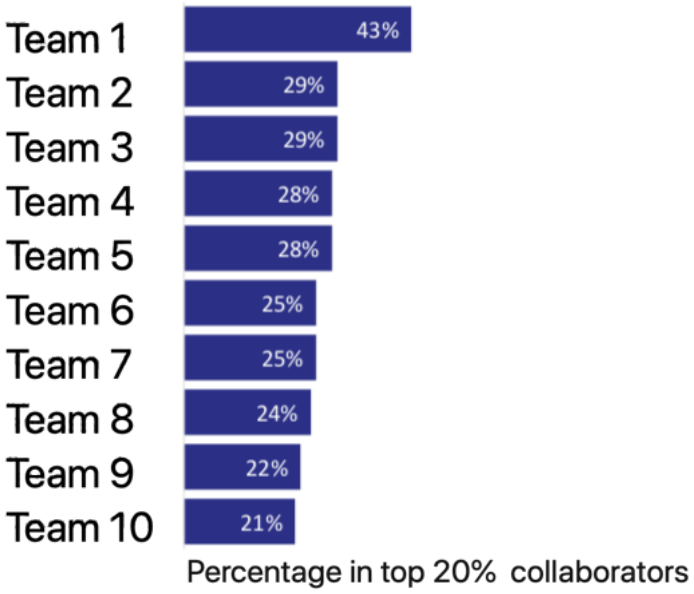


Figure 19 - Top 10 generator teams, graph shows that the 10 teams that generate the most collaboration for the organization have been identified.

Figure 20 shows the cumulative generated workload of the percentage of employees. This analysis is created automatically by Microsoft Workplace Analytics and shows the percentage of employees that account for the percentage of cumulative generated workload. In this case it shows that 10% of the employees create 51% of the collaboration, and 20% of the employees account for 70% of all the generated collaboration in the organization.

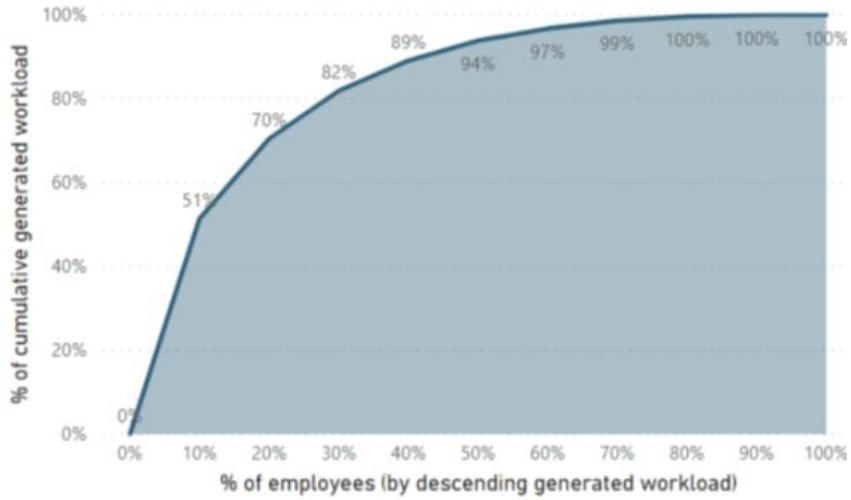


Figure 20 - Cumulative generated workload. Graph shows percentage of people that are responsible for a given percentage of generated workload.

The final analysis that was made for generated workload looks at the meeting hours that are generated in the organization. Figure 21 shows the percentage of generated meeting hours per meeting category. A meeting with more than 8 participants is considered a large meeting, while a meeting that lasts for more than one hour is considered a long meeting. For this analysis the meetings before and after Covid-19 measures are compared by taking the baseline and measure periods. The percentage of large meetings has increased by 15 percentage points while the long meetings saw a decrease in 5 percentage points. The category of long and large meetings has also decreased by 11 percentage points while the decision-making meetings has remained the same.

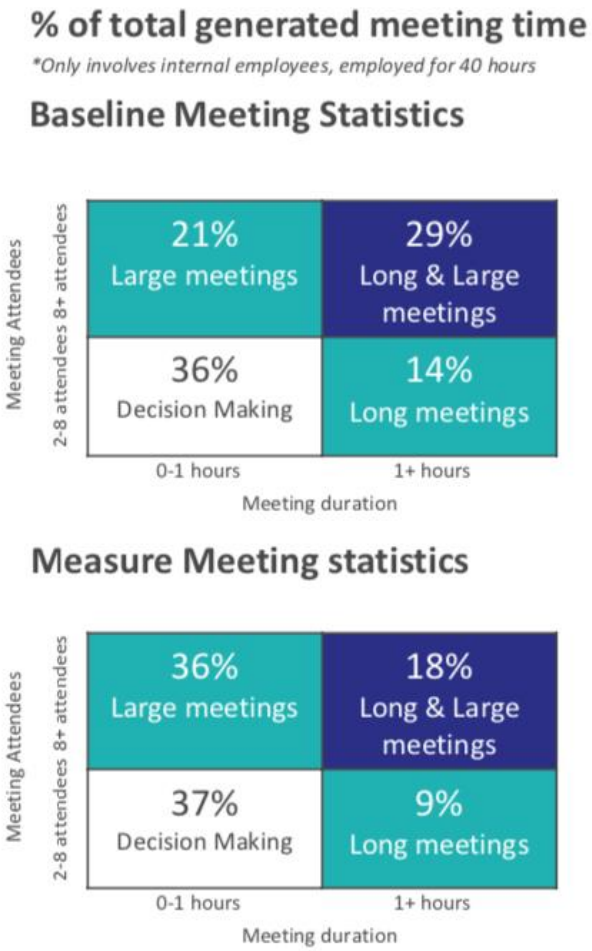


Figure 21 - Generated meeting hours. Shows the percentage of time spent on four different types of meetings. Compares the Baseline period (09/01/2020 – 13/03/2020) to the measure period (04/10/2020 – 04/12/2020).

4.2.2.6 Received collaboration

The received collaboration is calculated by the sum of hours a person spent on meetings and emails received from others. This KPI indicates how much collaboration time a person receives from the internal network. Figure 22 shows that the average received collaboration of individuals has increased from 11 hours in the baseline period to 13 hours in the measure period. This is calculated by dividing the total number of received collaboration hours by the number of employees in the network.

11h -> 13h

Figure 22 - Average of received collaboration per week for the organization in the baseline (09/01/2020 – 13/03/2020) and measure (04/10/2020 – 04/12/2020) period. Received collaboration includes the sum of collaboration hours that people receive through emails and meetings

To understand which areas in the organization, receive the most collaboration the top 10 receivers are filtered out of all the teams. Since the teams differ in size the graph on the right in Figure 23 shows the average received collaboration per person per week. In this graph it shows there are several teams that receive considerably more than the average number of collaboration hours. Other teams that have average scores indicate that they have more employees working in them.

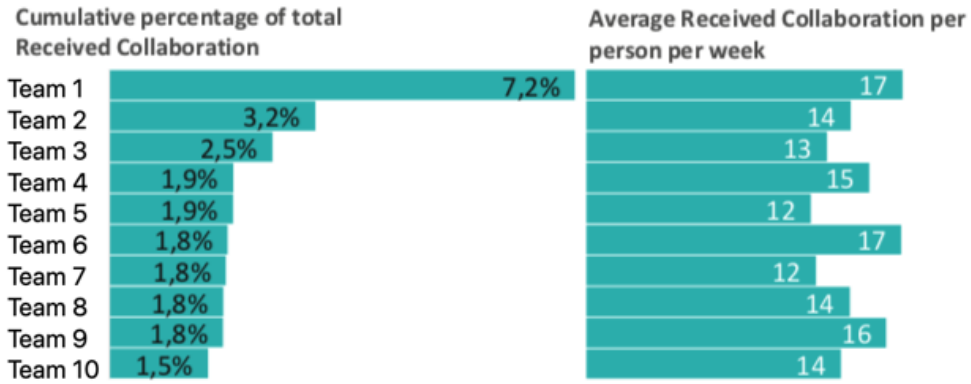


Figure 23 - Top collaboration receivers. Graph shows that the ten teams that received the biggest portion of the total collaboration hours have been identified. Includes the average number of hours that each person in the team received.

4.3 Sprint 2 – Organizational network flow

The use case that is analyzed in sprint 2 is organizational rigidity and silos. This use case is analyzed to see how the organizational network connects. The analysis for organizational rigidity and silos will reveal the areas in the business that see high levels of interaction and the parts of the organization that do not interact much or are at risk of being silos.

4.3.1 Sprint 2 – Preparation phase

This sprint aims to investigates the connections in the organization and how teams communicate. For this use case the focus is on the areas in the organization that receive and send a lot of information, and areas that see little collaboration: silos. The metrics that are attached to this use case are outside team collaboration and within team collaboration. These use cases will reveal how much and with whom a team collaborates. The metrics that are selected for this use case are: within team interaction, outside team interaction, network size, and network breadth. The metrics are explained briefly before showing the results of the analyses.

Within team interaction
This metric investigates how many hours the unit collaborates within the unit. This is calculated by the average time that is spent collaborating with team members.

Outside team interaction
This metric is opposite to within team interaction and looks into the hours a unit collaborates with other units. This is calculated by the average time that is spent collaborating with other teams.

Network size
Network size highlights the total number of connections with other business units. A connection is defined as two individuals sharing at least 2 emails, phone calls, or meetings in the last four weeks.

Network breadth
Network breadth indicates the number of teams that a team connects with. This is calculated by the number of teams a person has a made a connection with. The definition of connection in this case is the same as for network size.

4.3.3 Sprint 2 – Results
The analyses that were made during sprint 2 are presented in the following chapter. These are the results as they were presented to the leadership teams of the organization. The use case organizational rigidity and silos investigates how the organization connects and which parts are the most and least connected, this information is used to answer SQ3. In addition, this sprint aims to analyze the effects of Covid-19 on network breadth and network size, which helps to answer SQ4.

4.3.3.1 Within team collaboration
The within team collaboration hours are calculated by the number of hours a team spends collaborating within their own team per week. This metric is selected to see how many of the hours are spent within a team's own unit and to compare this to the outside team collaboration hours. The first analysis that was made is shown in Figure 24. This analysis shows the collaboration within team for each business unit. The period of the timeline 9th of June 2020 to 13th of December 2020.

Figure 24 shows that within team collaboration has increased over the year. The first peak in collaboration occurs during the time of Covid-19 measures. The peaks and throughs that follow shortly after are explained by holidays during which many employees are not collaborating. On average, people are collaborating 1,5 hours more within their team in the measure period (4th of October to 13th of December), compared to the baseline period (9th of January to 13th of march. In this Figure business unit A B and D all show similar levels of within team collaboration hours, while business unit C averages circa two and a half more hours.

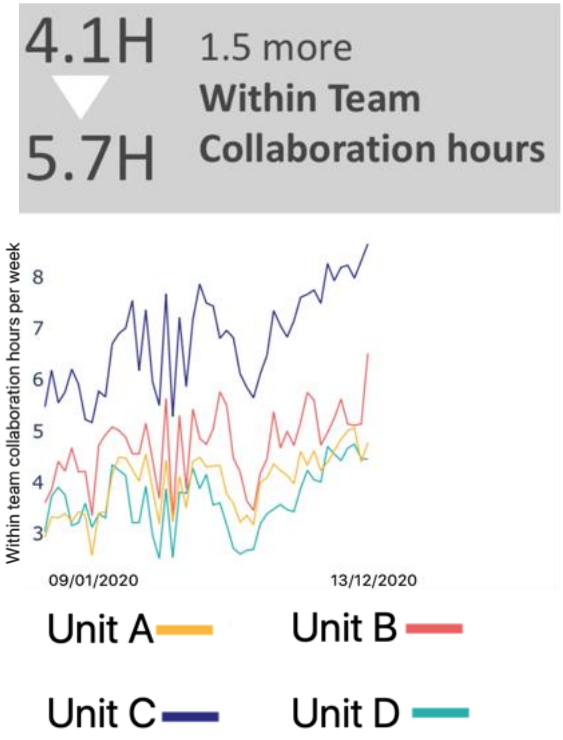


Figure 24 – Within team collaboration hours. Shows the hours that each business unit spends collaborating with people in the same team per week.

Whilst Figure 24 provides some general data on the collaborative behavior of the whole organization, the following Figures show only the top and bottom scoring teams. Figure 25 shows the bottom ten within team interactors. This graph shows the percentage of collaboration that each team spent within their own team. This is calculated by comparing the total number of collaboration hours for each team to the time they spent collaborating within their own team. Figure 25 highlights the teams that spent the least percentage of collaboration within their own team. These teams are filtered because they potentially might not effectively share knowledge within their own team. The low percentage of within team interaction might be due to the responsibilities and the nature of the tasks of the team. Since the results have not been discussed with the separate teams this will have to be investigated in the future.

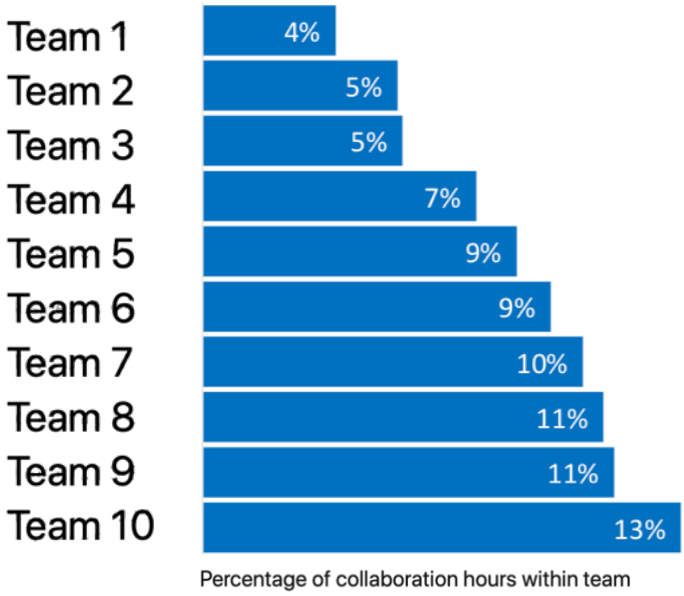


Figure 25 - Bottom ten within team interactors. Graph shows that the ten teams that have the lowest percentage of collaboration within the team have been identified.

Opposite to the analysis in Figure 25, Figure 26 shows the top ten teams that spent the highest percentage of their collaboration time within their own team. All these teams spent more than 40% of their collaboration time within their own team and potentially miss out on sharing knowledge with other parts of the organization.

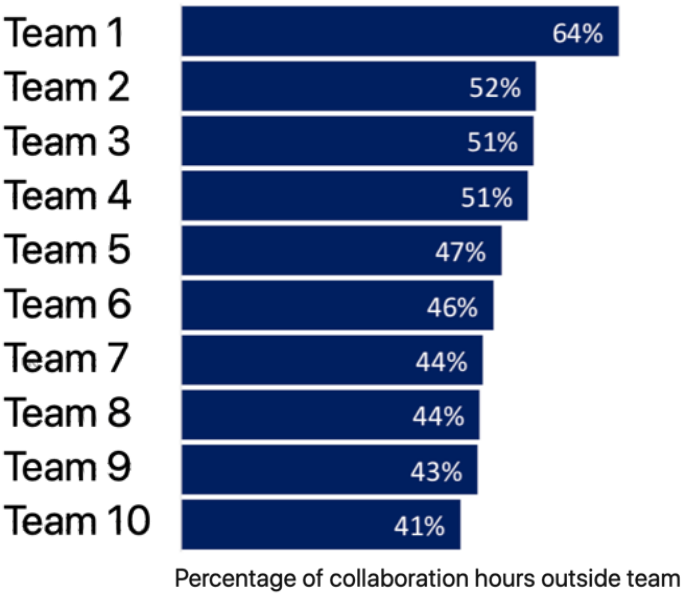


Figure 26 - Top ten within team interactors. Graph shows that the ten teams that spend the largest percentage of collaboration hours outside their own team have been identified.

4.3.3.2 Outside team collaboration

The outside team collaboration is calculated by the number of hours a team spends collaborating with people outside of their team. This metric is selected to see how the number of outside collaboration hours compare to the within team collaboration hours and see which teams in the organization connect the most.

Figure 27 shows the general graph that was made to indicate the hours of outside team collaboration for each business unit. The time-period is the same as in Figure 24. Business unit C, which had the most within team collaboration hours, shows average results for outside team collaboration. Business Unit B has two more hours than unit C, which has the second most outside team collaboration hours. The unit that lacks in outside collaboration hours the most, compared to the other units, is business unit A which has three hours less than business unit C.

On average the organization has collaborated more outside the team in the measure period compared to the baseline period. While the growth is less than within team collaboration the organization has seen an increase in collaboration hours in the network.

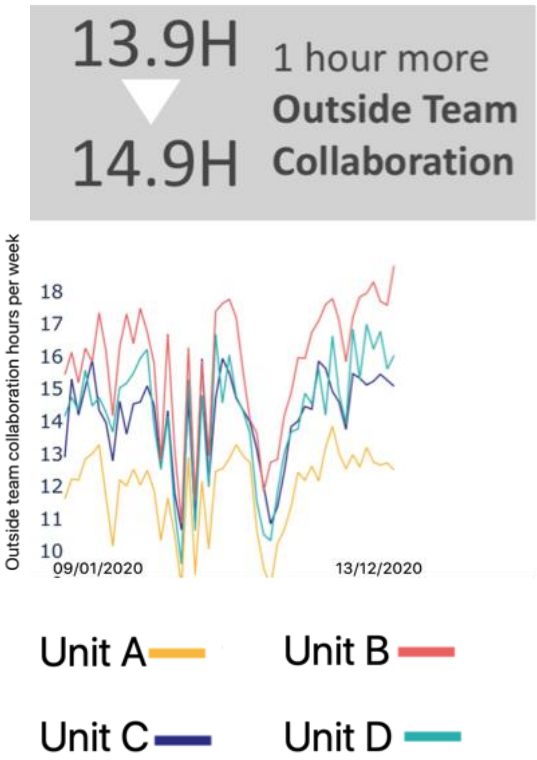


Figure 27 - Outside team collaboration graph. Shows the hours that the people in each business unit spend collaborating with people outside their own team per week.

The final analysis that was created for the team collaboration metrics compares the percentage of total collaboration time of organization. Figure 29 shows the averages of the whole organization for the period 9th of January 2020 to 13th of December 2020. The unclassified collaboration hours are due to connections to employees that are not in the pilot group or external people. The most notable statistic from this graph is that the collaboration time spent within team has grown by 38%. This shows that people in the organization have seen a huge increase of within team collaboration since the Covid-19 measures were put in place.

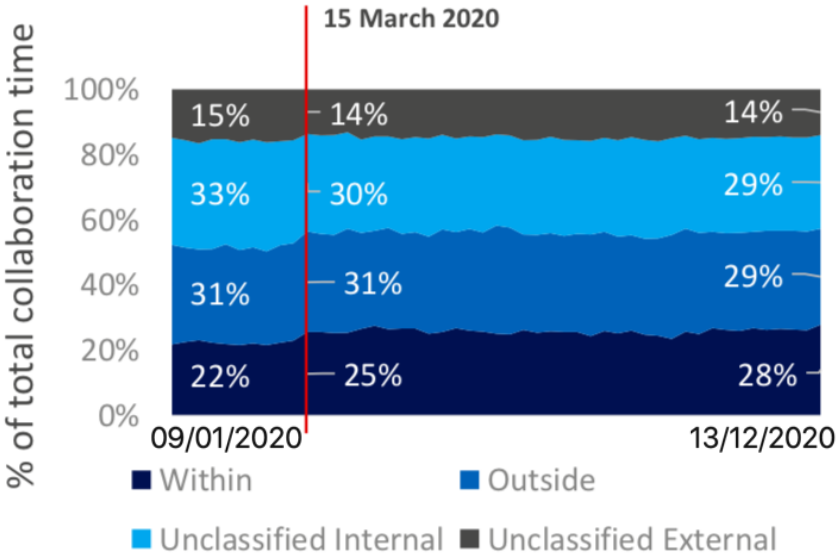


Figure 28 - Percentage of total collaboration time. Shows the percentage of time that the organization spends per collaboration category for 12 months. The red line indicates the start of the Covid-19 measures.

4.3.3.3 Network breadth

The network breadth metric indicates the broadness of a network. In this project there are four business units are analyzed. Each business unit can therefore interact with three different business units. The network breath metric highlights the meaningful connections with other business units. The definition for meaningful connection is set by Microsoft workplace analytics as at least two emails, phone calls, or meetings in the last 28 days.

The first analysis shows the network breadth for each of the business units in the same period that was used for the previous analysis. Figure 30 shows that the average network breadth did not change significantly in the measure period compared to the baseline period. However, the results do show that all four business units show different network breadths, with all of them being divided by circa 0.2.

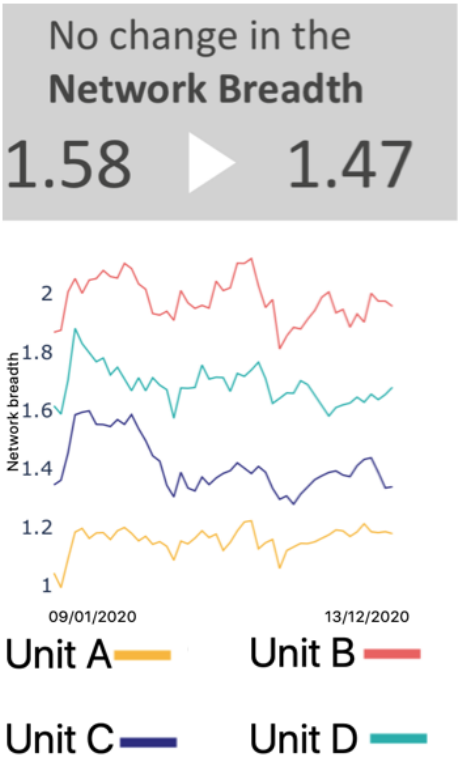


Figure 29 – Network breadth. Shows the network breadth for each business unit for 12 months.

4.3.3.4 Network size

The final metric is network size which indicates how many people in other business unit a business unit has a meaningful connection with. The definition for meaningful connection is the same as for the network breadth metric in Chapter 4.3.3.3.

Figure 31 shows the average network size for employees in each business unit in the period used in the other analysis. The peaks and troughs in this graph are due to holidays. The graph shows that there is no significant change in network size in the measure period compared to the baseline period. This indicates that the Covid-19 measures did not affect the number of people that a person connects with.

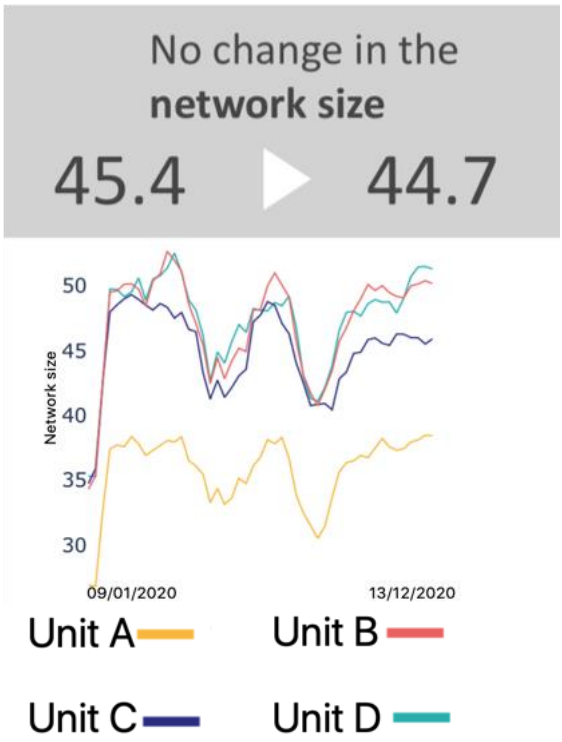


Figure 30 - Network size. Shows the network size of each business unit for 12 months.

Figure 31 shows the network size of each business unit. While this indicates roughly how individuals in the network connect, other analysis are created for adding more detail to the network. Figure 32 shows a network map that shows the connections of everyone. The map shows that business unit A, which has the lowest network size and breadth out of all the business units. Is located on the sides of the network and are the least intertwined with the other business units. Business unit B, which has the highest network breadth and a high network size, is in the middle of the network and connects with the three other business units. Business units C and D have average scores and are both on the side of the network, but show higher levels of intertwinement than business unit A.

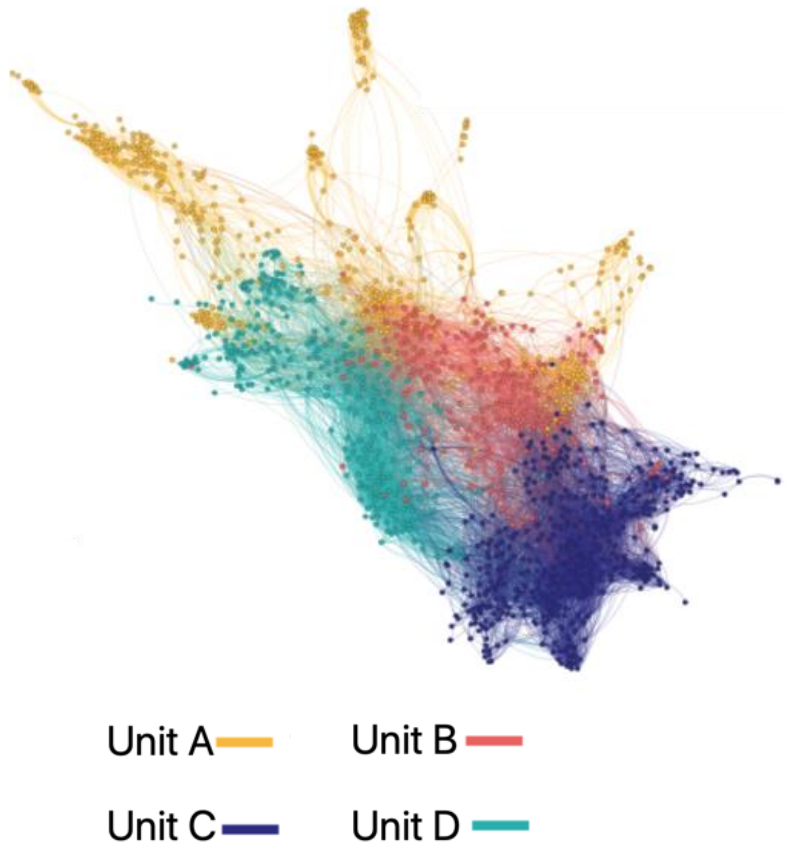


Figure 31 – Network mapped, categorized per business unit. Each node represents a person in the organization and each line represents a meaningful interaction of that person.

4.3.3.5 Network size and breadth

Looking at the network size and network breadth metrics separately provides some interesting insights. However, combining them allows for more analyses, which can be used to detect possible silos in the organization. Figure 32 shows a scatter plot that includes the network size and breadth of all teams (that have more than 5 employees). The boundaries are set by Microsoft workplace analytics. The graph indicates that 42.5% of the teams have a sub-optimal network breadth and network size.

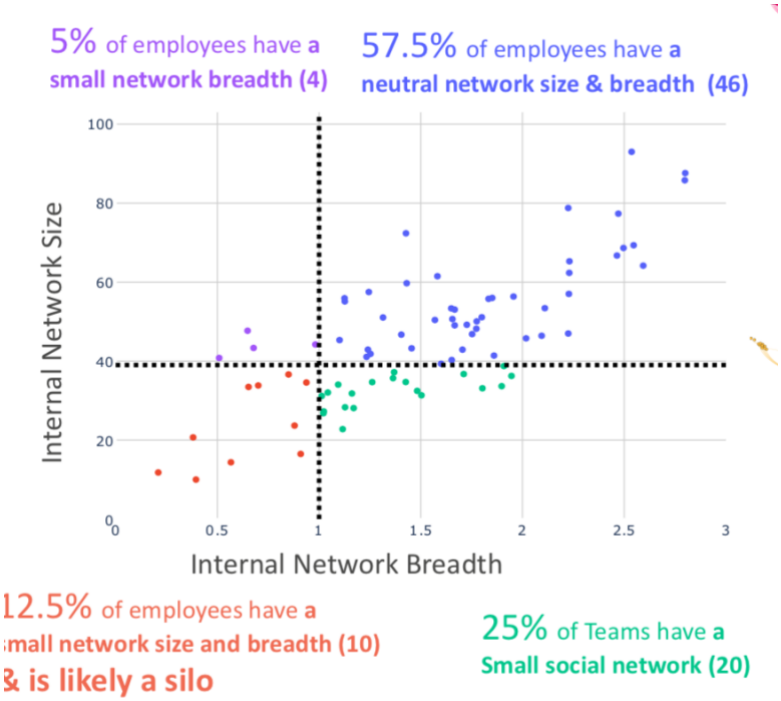


Figure 32 - Network breadth/size scatter plot. Each node is a team and plotted according to their network size and breadth.

The same network graph that is mapped in Figure 32 is also mapped in Figure 34. However, in this network the colors represent the teams as they are categorized in Figure 33. In this network the teams that have a small network size and breadth are visible at the edges. The blue nodes that represent individuals with a neutral network size and breadth are located in the middle of the network.

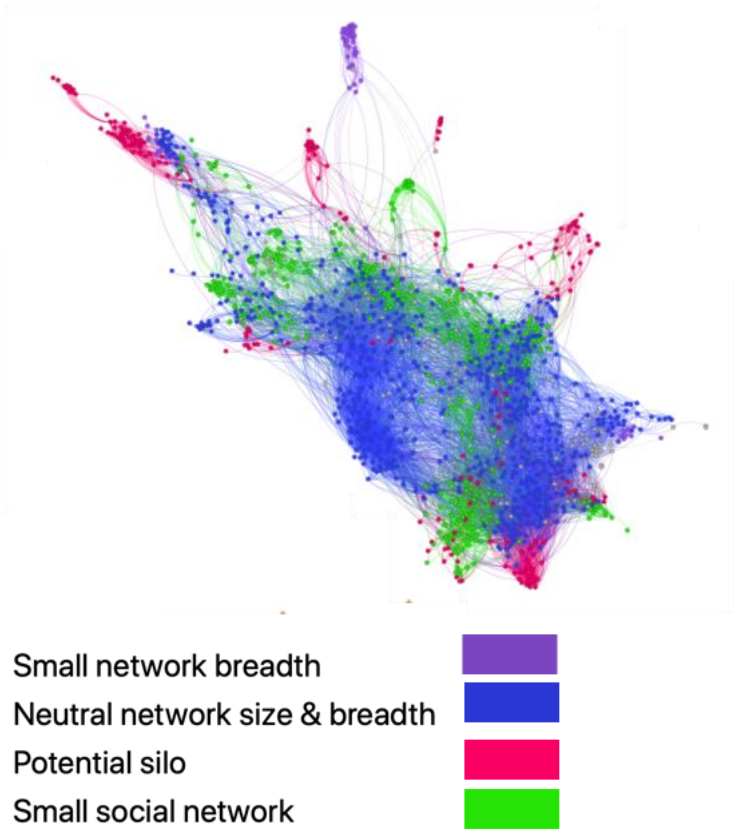


Figure 33 - Network mapped, categorized by network size and breadth . Each node represents a person in the organization and each line represents a meaningful interaction of that person.

The top ten teams that have the lowest network breadth and size are filtered and identified as potential silos. These teams are shown in Figure 35. The identified silos all have a network breadth that is smaller than 1 and a below average network size.

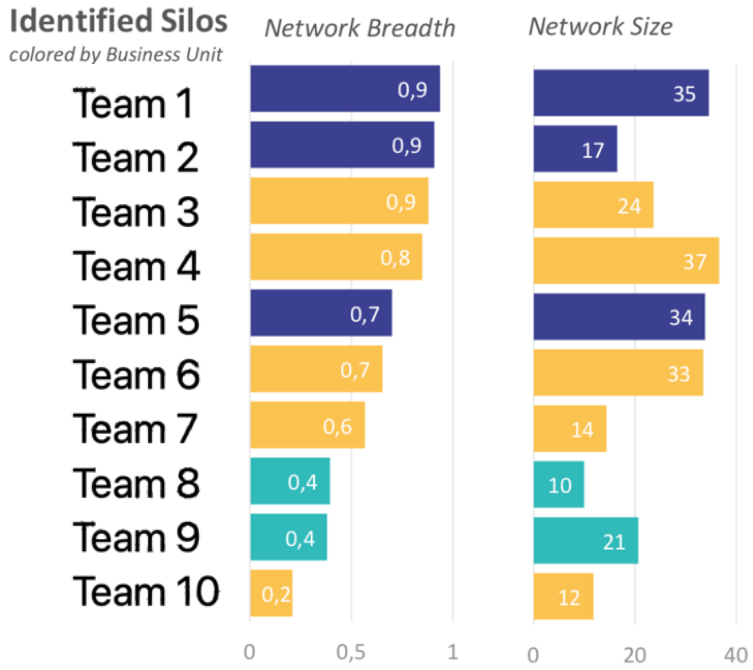


Figure 34 - Silos in the organization. This graph shows that the 10 teams with the lowest network breadth and size have been identified.

4.4 Findings data analysis

The data analysis has provided the organization with a lot of insights on the way that the employees communicate and how they spend their time. This chapter will conclude the results of the data analysis to determine what organizational network analysis was able to find. The conclusions are divided into three parts, each part will present conclusions for a different use case.

4.4.1 Workload balance

The results showed that there are clear differences for the workweek span metrics in the measure period compared to the baseline period. The literature review in chapter 2 did reveal that organizational network analysis is often used as a diagnostic tool, which is proven with the analyses in this use case. The analyses on workload balance displayed a change in the number of hours in the workweek span, collaboration hours, and focus hours of the employees.

The average employee in the organization works more hours per week, while also spending more time on collaborating. There is a decrease in the number of focus hours. Some of this is explained by the Covid-19 measures, which forced people to start working from home and made it impossible to talk face-to-face with people without using some sort of online connection. The increased number of collaboration hours is therefor an expected result of the analysis. In addition to the workweek span metrics, based on the criteria for burnout risk that were set, the project team was able to demonstrate how many employees were at risk of a burnout.

4.4.2 Collaboration overload

The results of the collaboration overload analysis demonstrated that, using organizational network analysis, it is possible to find the parts of the organization that contribute the most and least to the overall collaboration. The analysis into the total number of generated and received collaboration hours has revealed that employees saw an increase on both sides of the collaboration in the measure period compared to the baseline period. The most common explanation for this in the business is the Covid-19 measures, which forces the employees to collaborate more often with their colleagues, since informal face-to-face meetings are no longer possible.

4.4.3 Organizational rigidity & silos

The outcomes of the analysis on the organizational rigidity and silos use case showed that the organization saw an increase in within- and outside team collaboration. In addition to that, the most and least participating teams were identified. Using the network breadth and network size metrics it was possible to determine which of the business units occupies the most central position in the network and which were on the outside. The teams that show the lowest participation in the network based on their network breadth and network size are now highlighted in the organization.

4.5 Results business meetings

The results of the following chapter help to answer SQ5. The results in chapter 4.2 and 4.3 are presented in separate meetings to the leadership teams of the business units. The leadership teams consist of the senior managers of each business unit. Besides the four business units the results were presented to the human resources leadership team which includes the lead HR business partners and the senior management team of the HR departments.

The results are presented in slides that include the Figures that are discussed in chapter 4.2 and 4.3. The slides were different in the presentations for the four participating business units as they only showed the results that were specific for the business units. The slides included the results from the first two sprints. The HR leadership team was presented with the results from all the business units like they are presented in chapter 4.2 and 4.3.

4.5.1 Business reactions

The first reactions to the data from the business units were promising. All four management teams showed that they want to use organizational network analysis in the future based on the results that were presented to them. However, since the data that was presented is just general information the business units felt that in the future a different approach was needed. The positive reactions to the presentations along with the improvements that need to be made for future network analysis are discussed in the following chapters.

4.5.1.1 Positive reactions

The reactions to the presentations were in positive in general. The managers were excited to see the results and many of them felt like they wanted to start working with the data right away. The reoccurring message that came from the meetings was that the results confirmed some thoughts that they had but never before were looked into. The number of meetings was one of the topics that was never investigated or checked before. Most of the managers were therefore very interested to see the number of long and bloated meetings in the organization.

The results in the presentation showed that the managers produce a lot of the collaboration in the organization. The reactions in the meetings revealed that while this was a suspicion of the management teams, they never had the confirmation. This showed that the management teams recognize the phenomena that were filtered from the data analysis.

4.5.1.2 Improvements for future network analyses

Even though the presentations showed promising results to the managers, they wanted to see more from organizational network analysis in the future. In short, the reactions from the business related to improvements for future network analyses can be summarized into two categories; a more targeted approach to data analysis, and ready-to-go actions that can be implemented to improve the results.

Targeted approach to data analysis

While the management teams were very interested and, in some cases, even shocked to see the results, some managers were left with a feeling that they wanted a more targeted approach for the data analysis. Since some of the results showed the data for the whole business unit it is difficult to formulate actions. The results of the workload balance use case

are shown as an average for the business unit. While the managers showed interest in those results, they wanted more specific data to see in which parts of their unit they had to interfere.

Ready-to-go actions
The presentations included just the results as they are presented in the previous chapters. While the managers reacted enthusiastically, some of them wanted to know what the next steps were. While the managers believe that the results are interesting and, in some cases, should be improved, they wanted to know how the improvements were to be accomplished. The managers appreciated the results of the analysis but did not quite understand how they would add value to the employees and the organization. Since the project team did not prepare this for the pilot, this would have to be altered in the future of organizational network analysis.

4.5.1.3 Recurring remarks
Besides the positive reactions and the improvements for the future there were some other reoccurring remarks. These are discussed in the following chapter.

Missing data
Many of the mangers commented that the database was not incomplete and misses some communication data. WhatsApp was brought up the most as the managers admitted they called and texted via WhatsApp for business purposes. While this is a concern, the project team did think of this before and concluded that using the data that was available through the internal communication software would be sufficient to execute the pilot project.

The future of organizational network analysis
Since this was only a pilot project, and the reactions were mostly positive, some managers were thinking ahead and wondered how this should be implemented in the organization for the future. Since organizational network analysis could be a full-time method for the organization in the future it would require a dedicated team. To approve the team the business needs to understand the costs and benefits of the project. This would require a business case that explains how organizational network analysis would be beneficial to the organization.

4.5.1.4 Conclusion
The aim of the meetings with the business units and the management teams was to determine whether organizational network analysis can be used in the future. All the management teams of the business units agreed that it has the potential to be a useful tool for the organization. Since this was just a pilot project there were no targeted actions to be made based on the results, which is what many of the managers were missing now. However, they believe that actions can be formed based on the results, if they can obtain results that show more specifically were the anomalies in the results are coming from. All in all, the organization was excited about the pilot project and wants to continue with the analysis in the future.

5. Conclusion

5.1 Research questions

The research will be evaluated by answering the research questions. The questions will be answered based on the method and results of the study.

SQ5: How is the value of organizational network analysis assessed in the organization?

The value of organizational network analysis was a necessity to assess in the project since the organization had to decide whether to continue with network analysis in the future. The decision to continue was based on the results that were presented in the meetings with the managers of the organization, which are described in chapter 4.5. Since the results were general and the managers are looking for a targeted approach in the future, the value of organizational network analysis is assessed partly on the possibilities that it can bring in the future. However, the general results that were presented in the meetings were promising, as the managers responded that with a more targeted approach, they could interpret the results better and develop actions.

In this project, the value of the project was assessed by the response of the business to the results of the analysis. The positive response to organizational network analysis is due to the possibilities that it has for the future. The managers foresee potential actions that can be implemented based on the results that organizational network analysis can bring.

SQ4: *Considering the analyzed use cases, what are the effects of Covid-19 measures on the behavior of the employees?*

The results have shown that there are clear changes in the organization due to the Covid-19 measures. The metrics for workload balance show that across all the business units the workweek span has increased, as demonstrated in chapter 4.2.2.1. Additionally, each unit saw a rise in collaboration hours, shown in chapter 4.2.2.2, whilst the number of focus hours has decreased as can be seen in chapter 4.2.2.3. The cause of these changes cannot be explained to their full extent. However, the sudden increase in collaboration hours is related to working from home by the managers of the organization. They have mentioned in the presentation sessions that people are forced to collaborate more often through emails and calls since they people do not meet each other on the work floor.

The results on within team and outside team collaboration reflect the previous results as both increased. Chapter 4.3.3.1 shows that within team collaboration saw a significant percentual increase of almost 40, while outside team collaboration increased by roughly 7%, as highlighted in chapter 4.3.3.2. An explanation of this is that employees spoke with their teammates on the work floor, which they are now forced to do using calls and emails. The marginal increase on outside team collaboration shows that emails and calls where the preferred method to reach outside the team before Covid-19 measures.

Chapters 4.3.3.4 and 4.3.3.5 show that the Covid-19 measures did not have a substantial effect on the network breadth and size of the organization. Changes of 5% and 1% respectively reveal that the employees are not forced to increase their network due to Covid-19 changes. This indicates that the increase in collaboration is due to spending more hours with the same collaborators.

Conclusively, the results have shown that organizational network analysis is a tool that can be used to see the effect of the pandemic on the organization, most notably the increase in collaboration. The changes in behavior of the employees are clearly visible and demonstrate the purpose of organizational network analysis as a diagnostic tool.

SQ1: How do the employees in the organization spend their time and which business units spent the most time on collaboration hours and on focus hours?

The analysis in chapter 4.2.2 of the first use case provided insight into the workweek span, collaboration hours, and focus hours of the four business units. While most of the employees are categorized as neutral in the three metrics, roughly 12% of the organization work for more than 50 hours each week, which is defined as high. In addition to that, approximately 18% of employees spend less than 10 hours on focus hours, which means that they potentially do not have sufficient time to finish their own projects. The collaboration hours have increased but are harder to interpret due to the Covid-19 measures which has forced employees to collaborate more frequently.

The results show that business unit C has the highest workload of all four. The employees in this unit have on average the highest workweek span while also having the most collaboration hours. Unit C also has the highest number of people working with less than 10 focus hours each week. Based on the criteria that were composed for potential burn out risks, this unit has the largest number of employees that are at risk of getting a burnout.

Business unit A and D are the units with the shortest workweek span and the lowest number of collaboration hours. In these units the employees have the lowest risk of getting a burnout, but there is a potential latent capacity that can be unlocked which could increase the productivity out of these units.

Conclusively, the results have shown that with the analyses it is possible to discover how the organization spends their time and whether they prefer collaborating or spending time on focus hours. With the criteria that have been composed for potential burnout risks, an opportunity to identify potential cases is created.

SQ2: How does the organization collaborate and who contributes the most to the collaboration?

The analysis on collaboration in the organization provided insight into the top collaborators of the organization and revealed where the collaboration is generated. The results in chapter 4.2.2.5 show that the managers play an important role in the collaboration of the company. 95% of managers+ and 60% of managers are part of the top 20% collaborators. This indicates that the structure of the organization is based on a top-down design where the managers dictate a high percentage of the work. This is most visible in business unit A, in which 46% of the top collaborators are managers.

Furthermore, business unit A has the lowest percentage of top contributors out of the four business units. The leader in this category is business unit B, in which 33% of the employees are in the top 20% of contributors, indicating that this business unit contributes the most to the collaboration of the organization. In addition to the top collaborators, the analysis has also

revealed the top 10 receivers of collaboration in the organization. The teams producing and receiving the most collaboration are at the heart of the organization.

The analysis into the collaboration within and outside the team in chapter 4.3.3 showed that business unit C is the most active collaborator within teams, and business unit B is the most active outside team collaborator. This is in line with the previous results that showed that business unit B has the highest number of people collaborating in the organization. Business unit A has low scores on both which also confirms the earlier findings that showed they have the lowest number of employees in the top contributors.

All in all, the results show that it is possible to analyze the collaboration in the organization by investigating where the collaboration comes from and who receives it. This information can be used by the organization to detect which parts of the organization are crucial for spreading information and fulfill an important task in the collaborative capacity of the company.

SQ3: How does the organization connect, who is at the center and who lies at the peripheral parts of the network?

By looking at the network breadth and network size of the employees in each business unit, analyzed in chapters 4.3.3.3 and 4.3.3.4, it is possible to determine which unit is at the heart and which unit is at the edge of the network. Business unit B has the highest network breadth while also showing high network size. In the mapped network they are clearly visible at the center of the network. This is in line with the results of the other research questions which showed that business unit B has is the most collaborative unit out of the four business unit. The same goes for business unit A, which scored consistently low and is at the clear edge of the network.

Business unit A has 5 teams that are identified as potential silos, which have below average network breadth and network size. Business unit B has 0 identified potential silos. This information can help the organization to become more connected by integrating the silos better in the organizational network.

These results show that organizational network analysis is a useful tool for identifying the most and least connected part of the organization and by doing so determining which teams are potential silos in the network.

RQ: What is a potential contribution of organizational network analysis for business value co-creation?

The goal of the pilot project was to examine the possibilities of organizational network analysis and decide whether to continue with it on a full-time basis. Since the organization decided to proceed with organizational network analysis and dedicate a team to it, the pilot project is successful. The organization expects that a full-time team devoted to network analysis will bring sufficient value to the organization based on the results that were presented in the meetings.

The results of this research have shown that organizational network analysis can be used a diagnostic- a well as an explorative tool in a large network. The Covid-19 variable allowed the team to use network analysis as tool to compare the behavior of the network before and after

the government installed measures. The analysis further showed that using network analysis as an explorative tool provided the organization with new insights, as well as a method to validate some assumptions that had regarding the behavior of the organization.

The managers of the organization are most excited for a targeted approach that will provide insights into the behavior of specific teams. In the future, organizational network analysis can be a valuable tool for the organization to examine the top or low scoring teams for metrics used in this research or other metrics. Identifying the teams will then assure that the organization can make targeted actions and improve the behavior of the teams.

To execute a successful project in the future there is one important condition. To generate results that the business can act on, the analysts that perform the analysis should discuss with the managers which areas they want to gain insights in. The results that were presented in this project were based on the use cases created by the analyst's team. While these proved to be insightful, the managers indicated that they required more detailed results based on their input to develop and execute actions that would change these results.

Conclusively, organizational network analysis can contribute to the organization in the following ways. It can be used by an organization as a tool that identifies parts of the network that underperform on measures. This information can be used to formulate actions and potentially change the behavior of the low scoring units. Another contribution is the usage of organizational network analysis as a diagnostic tool that is used to detect the effects of transformations on the network. Both applications are appreciated by the managers of the business, who in the future are also looking to see whether organizational network analysis can produce results based on a targeted approach initiated by input by the managers.

5.2 Limitations

The participants of this study are all white-collar workers in a telecommunication company in the Netherlands. The conclusions of this research are therefore limited to organizations that have the same sample group. More research in different types of organizations and on different types of employees is required to assess the use of organizational network analysis.

This project evaluated the use of organizational network analysis with the help of a technical tool that collects and processes the data. The tool in this case is Microsoft workplace analytics. The results and the conclusions of this research are exclusive to this tool. In future studies different tools must be evaluated and tested on large target groups to determine the value of organizational network analysis.

The final limitation of this study regards the use of organizational network analysis as a diagnostic tool. The change that was evaluated in this research concerned the measures that the government enforced due to a global pandemic. These are obviously quite severe and therefor clearly visible in the way that people work. To further explore the use of organizational network analysis as a diagnostic tool, different types of changes must be evaluated.

References

[1] Cross, R., Borgatti, S. P., & Parker, A. (2002). Making Invisible Work Visible: Using Social Network Analysis to Support Strategic Collaboration. *California Management Review*, *44*(2), 25–46. https://doi.org/10.2307/41166121

[2] Zhu, T., Wang, B., Wu, B., & Zhu, C. (2014). Maximizing the spread of influence ranking in social networks. *Information Sciences*, *278*, 535–544. https://doi.org/10.1016/j.ins.2014.03.070

[3] Hatala, J.-P. (2006). Social Network Analysis in Human Resource Development: A New Methodology. *Human Resource Development Review*, *5*(1), 45–71. https://doi.org/10.1177/1534484305284318

[4] Garcia, S. (2015). Improving innovation with organizational network analysis. *OD Practitioner*, *47*(2), p10-16.

[5] Microsoft. (2020). Case study: Transforming global commercial sales with Microsoft Workplace Analytics. Retrieved from https://insights.office.com/sales-excellence/transforming-commercial-sales-with-workplace-analytics/

[6] Lusher, D., Robins, G., & Kremer, P. (2010). The Application of Social Network Analysis to Team Sports. *Measurement in Physical Education and Exercise Science*, *14*(4), 211–224. https://doi.org/10.1080/1091367x.2010.495559

[7] Ramos, V., Franco-Crespo, A., González-Pérez, L., Guerra, Y., Ramos-Galarza, C., Pazmiño, P., & Tejera, E. (2019). Analysis of organizational power networks through a holistic approach using consensus strategies. *Heliyon, 5*(2). doi:10.1016/j.heliyon.2019.e01172

[8] Merrill, J., Bakken, S., Rockoff, M., Gebbie, K., & Carley, K. M. (2007). Description of a method to support public health information management: Organizational network analysis. *Journal of Biomedical Informatics*, *40*(4), 422–428. https://doi.org/10.1016/j.jbi.2006.09.004

[9] Otte, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, *28*(6), 441–453. https://doi.org/10.1177/016555150202800601

[10] Choudhary, P. (2015). A Survey on Social Network Analysis for Counter-Terrorism. *International Journal of Computer Applications, 112*(9), 24-34. doi:10.5120/19695-1455

[11] Grandjean, M. (2016). A social network analysis of Twitter: Mapping the digital humanities community. *Cogent Arts & Humanities, 3(1).* doi:10.1080/23311983.2016.1171458

[12] Emch, M., Root, E. D., Giebultowicz, S., Ali, M., Perez-Heydrich, C., & Yunus, M. (2012). Integration of Spatial and Social Network Analysis in Disease Transmission Studies. *Annals of the Association of American Geographers, 102(5), 1004–1015.* doi:10.1080/00045608.2012.671129

[13] Tortoriello, M., Reagans, R., & McEvily, B. (2012). Bridging the knowledge gap: The influence of strong ties, network cohesion, and network range on the transfer of knowledge between organizational units. *Organization science*, *23*(4), 1024-1039.

[14] Pentland, A. (n.d.). People Analytics: Transforming Management with Behavioral Data. Retrieved from https://professional.mit.edu/course-catalog/people-analytics-transforming-management-behavioral-data

[15] Al-Taie, M. Z., & Kadry, S. (2017). Theoretical Concepts of Network Analysis*. Python for Graph and Network Analysis, 1–32.* doi:10.1007/978-3-319-53004-8_1

[16] Lazega, E., Wasserman, S., & Faust, K. (1995). Social Network Analysis: Methods and Application*s. Revue Française de Sociologie, 36(4), 781.* doi:10.2307/3322457

[17] Farshchi, M. A., & Brown, M. (2011). Social networks and knowledge creation in the built environment: A case study. Structural Survey, 29(3), 221-243. doi:10.1108/02630801111148202

[18] Newman, G. (2017). Why Informal networks are set to revolutionize HR and People Analytics. Retrieved from https://www.linkedin.com/pulse/why-informal-networks-set-revolutionize-hr-people-analytics-newman/

[19] Cross, R. (N/A). Case study: Restructuring. Retrieved from https://www.robcross.org/wp-content/uploads/2020/05/CC-Case Study_Restructuring.pdf

[20] Cross, R. (N/A). Case study: Innovation. Retrieved from https://www.robcross.org/wp-content/uploads/2020/05/CC-Case-Study_Innovation.pdf

[21] Kazienko, P., Michalski, R., & Palus, S. (2011, June). Social network analysis as a tool for improving enterprise architecture. In *KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications* (pp. 651-660). Springer, Berlin, Heidelberg.

[22] Cross, R. (N/A). Case study: Collaboration. Retrieved from https://www.robcross.org/wp-content/uploads/2020/05/CC-Case-Study_Collaboration.pdf

[23] Zhylinska, O., Chornyi, A., Dzhuliy, V., & Yemchuk, L. (2020). Organizational Network Analysis as a Tool for Leadership Assessment in Software Development Team. *2020 10th International Conference on Advanced Computer Information Technologies (ACIT)*. doi:10.1109/acit49673.2020.9208827

[24] Andrzejak, P. (2019). Identifying potential change leaders through social network visualizations. *Nauki o Zarządzaniu*, *24*(1), 37-47.

[25] Mroz, J. E., Allen, J. A., Verhoeven, D. C., & Shuffler, M. L. (2018). Do we really need another meeting? The science of workplace meetings. *Current Directions in Psychological Science*, *27*(6), 484-491.

[26] Kleinman, L. (2010). *Physically present, mentally absent? Technology multitasking in organizational meetings* (Doctoral dissertation).

[27] Phillips, P. P., & Stawarski, C. A. (2008). *Data collection: Planning for and collecting all types of data* (Vol. 175). John Wiley & Sons.

[28] Cross, R., Rebele, R., & Grant, A. (2016, February). Collaborative Overload. Harvard Business Review. https://hbr.org/2016/01/collaborative-overload

[29] Systems Innovation. (2016). Network Theory Overview. Retrieved from https://www.systemsinnovation.io/post/network-theory-overview
[30] Zielinski, D. (2019, August 16). Find Your Influencers with Organizational Network Analysis. SHRM. https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/find-your-influencers-with-organizational-network-analysis.aspx

[31] Terblanche, N. H. D., & Erasmus, E. D. (2018). The use of organisational network analysis as a diagnostic tool during team coaching. *SA Journal of Industrial Psychology, 44.* doi:10.4102/sajip.v44i0.1548

[32] Becheru, A., & Luncean, L. (2015). Organisational Network Analysis within the academic world: Initial steps. *2015 IEEE 19th International Conference on Computer Supported Cooperative Work in Design (CSCWD).* doi:10.1109/cscwd.2015.7230959

[33] Zhu, M., Liu, W., Hu, W., & Fang, Z. (2009). Social Network Analysis in IT Company. *2009 International Conference on E-Learning, E-Business, Enterprise Information Systems, and E-Government.* doi:10.1109/eeee.2009.77

[34] Leon, R. – D., Rodríguez-Rodríguez, R., Gómez-Gasquet, P., & Mula, J. (2017). Social network analysis: A tool for evaluating and predicting future knowledge flows from an insurance organization. *Technological Forecasting and Social Change*, 114, 103–118. doi:10.1016/j.techfore.2016.07.032

[35] Briody, E. K., & Erickson, K. C. (2014). Success despite the Silos: System-Wide Innovation and Collaboration. *International Journal of Business Anthropology*, *5*(1).

[36] Segarra, L., Herrera, R. F., Alarcón, L. F., & Pellicer, E. (2017). Knowledge management and information flow through social networks analysis in Chilean architecture firms. In *Proc. of the 25th Ann. Conf. of the Int'l Group for Lean Construction* (pp. 413-420). Heraklion, Greece.

[37] Wen, Q., Gloor, P. A., Fronzetti Colladon, A., Tickoo, P., & Joshi, T. (2020). Finding top performers through email patterns analysis. *Journal of Information Science*, *46*(4), 508-527.

[38] Cross, R., & Gray, P. (2013). Where Has the Time Gone? Addressing Collaboration Overload in a Networked Economy. *California Management Review, 56(1), 50–66.* doi:10.1525/cmr.2013.56.1.50

[39] Perlow, L. A., Hadley, C. N., & Eun, E. (2017, July). Stop the meeting madness. Harvard Business Review. https://hbr.org/2017/07/stop-the-meeting-madness

[40] Cross, R., & Taylor, S. (2019, January 23). How to Manage Collaborative Overload. Babson Thought & Action. https://entrepreneurship.babson.edu/manage-collaborative-overload/

[41] McCombes, S. (2020, June 19). How to Do a Case Study: Examples and Methods. Retrieved from https://www.scribbr.com/methodology/case-study/

[42] What is an Influencer? - Social Media Influencers Defined [Updated 2021]. (2021, January 28). Retrieved from https://influencermarketinghub.com/what-is-an-influencer/

[43] Kenton, W. (2021, May 07). Understanding Silo Mentality. Retrieved from https://www.investopedia.com/terms/s/silo-mentality.asp#:~:text=In business, organizational silos refer,shared because of system limitations.

[44] Lansmann, S. Collaborative Overload: A Dark Side of Enterprise Social Software.

[45] APA Dictionary of Psychology. (n.d.). Retrieved from https://dictionary.apa.org/formal-organizational-structure

[46] Sparks, K., Cooper, C. L., Fried, Y., & Shirom, A. (2013). The effects of working hours on health: A meta-analytic review. In From Stress to Wellbeing Volume 1 (pp. 292-314). Palgrave Macmillan, London.

[47] Harrington, J. M. (1994). Working long hours and health.

[48] Bell, D., Otterbach, S., & Sousa-Poza, A. (2012). Work hours constraints and health. *Annals of Economics and Statistics/Annales d'économie et de statistique*, 35-54.

[49] Collewet, M., & Sauermann, J. (2017). Working hours and productivity. *Labour economics*, *47*, 96-106.

[50] Pencavel, J. (2015). The productivity of working hours. *The Economic Journal*, *125*(589), 2052-2076.

[51] The Privacy Crisis - Multi Tasking & Employee Privacy. (2020, October 08). Retrieved from https://www.steelcase.com/research/articles/privacy-crisis/?__hstc=191316853.20fcb526031e867da77e4f5065c0c18e.1608552504042.1608552504042.1608552504042.1&__hssc=191316853.1.1608552504042&__hsfp=3160465627

[52] Cathy Reisenwitz in Future of Work. (n.d.). What is Focus Time and How Does it Work? Retrieved from https://www.getclockwise.com/blog/what-is-focus-time

[53] Madehmer. (n.d.). Week in the life metrics in Workplace Analytics - Workplace Intelligence. Retrieved from https://docs.microsoft.com/en-us/workplace-analytics/use/explore-metrics-week-in-the-life

[54] Živković, M. (2021, February 18). Over-Collaboration vs Not Enough, How to Hit the Right Balance at Work. Retrieved from https://www.chanty.com/blog/balance-at-work/

[55] MacKay, J. (2020, February 5) Overcoming collaboration overload: How to choose (and use) the best collaboration tools for your team. Retrieved from https://plan.io/blog/collaboration-overload/#:~:text=Collaboration%20overload%20is%20when%20a,being%20bombarded%20daily%20with%20emails

[56] Cross, R., & Gray, P. (2013). Where has the time gone? Addressing collaboration overload in a networked economy. *California management review*, *56*(1), 50-66.

[57] Rasskazov, S., Rubtcova, M., Derugin, P., Pruel, N., & Malychev, V. (2016). Social network analysis as an organizational diagnostic tool: the case of small business in Russia. *International Review of Management and Marketing, 6*(1).

[58] Leon, R. D., Rodríguez-Rodríguez, R., Gómez-Gasquet, P., & Mula, J. (2017). Social network analysis: A tool for evaluating and predicting future knowledge flows from an insurance organization. Technological Forecasting and Social Change, 114, 103-118.

[59] Workplace Analytics. (n.d.). Microsoft. https://www.microsoft.com/en-ww/microsoft-365/business/workplace-analytics

[60] *General Data Protection Regulation (GDPR) – Official Legal Text*. (2019). General Data Protection Regulation (GDPR). https://gdpr-info.eu/

[61] *Metrics description.* (2021) Workplace analytics. https://docs.microsoft.com/en-us/workplace-analytics/use/metric-definitions

Appendix A – organizational data attributes

PersonID
The internal communication data needs to be able to be linked to an individual. In this case an email address of the person is sufficient since this the unique ID that employees use for the internal software system, making it the most convenient way to connect an employee to their communication data.

EffectiveDate
EffectiveDate refers to the date at which an employee starts working at the company. In some analyses it is required to distinguish the tenure years of employees. When analyzing networks of people there might be a contrast in those of people that work for a short time as opposed to the veteran employees. In addition to that, some analyses demand the need to filter out solely the new employees.

LevelDesignation
This attribute categorizes employees based on their level in the hierarchy of the company. Besides the obvious manger layers there are several layers for the individual contributors (IC) in the company. The higher the layer the higher the responsibilities and the salary. This attribute could be useful to highlight different levels of involvement for the different layers in which the employees reside.

Organization
There are four business units that participate in the project. Most of the analyses will look at the same metrics for each unique business unit in the company. Since the results are presented to each business unit individually, this is one of the attributes that will be used for a large portion of the analysis.

ManagerID
All employees need to be linked to their manager; therefore, each employee is provided with their managers' email address. In the analyses this can be required in order to relate the behavior of individuals to the behavior of their manager.

SupervisorIndicator
This indicates the management level of the node, they can either be an Individual Contributor (IC), Manager, or Manager+. Manager+ indicates a manager of a manager.

Gender
In some analyses the gender of the employees can be selected as a filter in order to find anomalies in the behavior of males and females. In addition, gender can be used to observe whether the ratio of male to female is conforming to the ideals of the organization.

FTE
FTE is used as an abbreviation of Full Time Equivalent which refers to the hours worked equivalent to a full-time position. In organizational context this is an important term since managers usually express the time of employees in FTE. In order to translate the time it costs

or the time that can be saved into a term that is understood immediately by business partners FTE is a required attribute.

Hours
The hours that a person works in a week. This is a necessary attribute to make a distinction between full-time and part-time employees. For some analyses the part-time employees might have to be filtered out.

EmployeeType
There are three possible types of employees in the database: internal employees, external employees and temporary employees. In most cases the analyses will only look at the internal employees.

ContractType
Employees can have different types of contracts: temporary contracts, fixed-time contracts, or zero-hours contract. In some analyses this distinction might have to be made to determine the involvement of employees that are employed on different types of contract.

Department
Each business unit is divided into several departments. This attribute helps in order to make a distinction in the analyses between those departments. For the mangers of each business unit it is very useful information to know which departments score high or low for certain metrics.

Team
Each department is further divided into teams. Once again this attribute helps to get closer to the outliers. Some metrics can be analyzed in such a way that the top ten or bottom ten scoring teams are picked out of the findings.

SubTeam
One layer below Team is SubTeam. Not necessarily each team has a SubTeam, but when available this can be used to zoom further into the results.

Office
This refers the geographic location of the office in which the person operates. Since most teams operate in the same location this attribute might

Origin
The organization originated from a merger. The Origin attribute informs whether an employee is originally hired by the legacy organization or by the merged organization.

HourlyRate
HourlyRate refers to the hourly salary of an employee. In some analyses monetary reward can be used as a variable to see whether it affects the behavior of an employee.

TimeZone
Some external employees can work outside of the Netherlands, which is why the TimeZone

attribute can be used as a filter. However, most analyses will only include the employees working in one of the main offices of the organization in the Netherlands.

PerformanceMetrics
Managers grade their employees with a numerical rating. These performance metrics can be included in an analysis to relate the behavior of an employee to the way they perform in the organization.

Appendix B – Workplace analytics metrics

The metrics available from Microsoft workplace analytics to generate queries. Retrieved from the workplace analytics website [61].

Person metrics

PERSON METRICS

Metric	Description	Query type	Data type	Customizable
After hours collaboration	Number of hours the person spent in meetings, emails, IMs, and calls with at least one other person, either internal or external, after deduplication of time due to overlapping activities (for example, calls during a meeting), outside of working hours.	Person	Hour	No
After hours email hours	Number of hours the person spent sending and receiving emails outside of working hours.	Person	Hour	Yes
After hours in calls	Number of hours a person spent in scheduled and unscheduled calls through Teams, outside of working hours. For calls that started during working hours, this number only includes the part of the call that occurred outside of that person's work schedule (as set in Outlook).	Person	Hour	Yes
After hours instant messages	Number of hours a person spent in instant messages through Teams, outside of working hours.	Person	Hour	Yes
After hours meeting hours	Number of hours the person spent in meetings outside of working hours.	Person	Hour	Yes
Call hours	The number of hours the person spent in scheduled and unscheduled calls through Teams with at least one other person, during and outside of working hours.	Person	Hours	Yes
Collaboration hours	Number of hours the person spent in meetings, emails, IMs, and calls with at least one other person, either internal or external,	Person	Hour	Yes

Metric	Description	Query type	Data type	Customizable
	after deduplication of time due to overlapping activities (for example, calls during a meeting).			
Collaboration hours external	Number of hours the person spent in meetings, emails, IMs, and calls with at least one other person outside the company, after deduplication of time due to overlapping activities (for example, calls during a meeting).	Person	Hour	No
Conflicting meeting hours	Number of meeting hours where the person had overlapping meetings in their calendar. The count includes the entire duration of all overlapping meetings, not just the amount of time that overlaps. (This number includes all non-declined meeting times, which includes accepted, tentative, or no responses to meeting invitations.)	Person	Hour	Yes
Email hours	Number of hours the person spent sending and receiving emails.	Person	Hour	Yes
Emails sent	Number of emails the person sent.	Person	Count	Yes
External network size	The number of people external to the company with whom the person had at least two meaningful interactions in the last four weeks.	Person	Count	Yes
Generated workload call hours	Number of hours the person spent calling internal recipients through Teams.	Person	Hour	Yes
Generated workload call participants	Number of internal participants of calls organized by the person. (Counts each participant once for each call.)	Person	Count	Yes
Generated workload calls organized	Number of calls organized by the person.	Person	Count	Yes

Metric	Description	Query type	Data type	Customizable
Generated workload email hours	Number of email hours the person created for internal recipients by sending emails.	Person	Hour	Yes
Generated workload email recipients	Number of internal recipients on emails sent by the person. (Counts each recipient once for each email received.)	Person	Count	Yes
Generated workload instant message hours	Number of instant message hours the person created through Teams for internal recipients by sending instant messages.	Person	Hour	Yes
Generated workload instant message recipients	Number of internal participants of calls organized by the person. (Counts each participant once for each call.)	Person	Count	Yes
Generated workload meeting attendees	Number of internal attendees in meetings organized by the person. (Counts each attendee once for each meeting.)	Person	Count	Yes
Generated workload meeting hours	Number of meeting hours the person created for internal attendees by organizing meetings.	Person	Hour	Yes
Generated workload meetings organized	Number of internal meetings organized by the person.	Person	Count	Yes
Instant message hours	Number of hours a person spent in instant messages (IMs) through Teams with at least one other person, during and outside of working hours.	Person	Hours	Yes
Instant messages sent	Total number of instant messages (IMs) sent by a person through Teams, during and outside of working hours.	Person	Count	Yes

Metric	Description	Query type	Data type	Customizable
Internal network size	The number of people within the company with whom the person had at least two meaningful interactions in the last four weeks.	Person	Count	Yes
Low-quality meeting hours	Number of meeting hours in which an attendee multitasked, attended a *conflicting meeting*, or attended a meeting that exhibits *Redundancy (organizational)*. Workplace Analytics admins can set the hourly rate of low-quality meeting time; if this value has not been set, the cost defaults to $75 per person hour. **Note**: Calculations for conflicting meeting hours are affected by meeting exclusion rules and adjustments based on the type of meetings that overlap (non-declined work meetings, focus hours, and out-of-office time).	Person	Hour	Yes
Manager coaching hours 1:1	Total number of hours that a manager spends in one-on-one meetings with *all* of the manager's direct reports.	Person	Hour	Yes
Meeting hours	Number of hours the person spent in meetings with at least one other person during and outside of working hours.	Person	Hour	Yes
Meeting hours during working hours	Number of hours the person spent in meetings, during working hours, with at least one other person.	Person	Hour	Yes
Meeting hours with manager	Number of meeting hours where attendees included at least the person and their manager.	Person	Hour	Yes
Meeting hours with manager 1:1	Number of meeting hours involving only the person and their manager.	Person	Hour	Yes
Meetings hours with skip level	Number of meeting hours that the person attends where their manager's manager also attends the meeting.	Person	Hour	Yes
Meetings	Number of meetings the person attended.	Person	Count	Yes
Meetings with manager	Number of meetings where attendees include at least the person and their manager.	Person	Count	Yes

PERSON METRICS

Metric	Description	Query type	Data type	Customizable
Meetings with manager 1:1	Number of meetings involving only the person and their manager.	Person	Count	Yes
Meetings with skip level	Number of meetings where the manager of the person's manager is an attendee.	Person	Count	Yes
Multitasking meeting hours	Number of meeting hours where the person sent: • Two or more emails sent per meeting hour • Two or more emails sent per meeting for meetings less than one hour	Person	Hour	Yes
Networking outside company	The number of distinct external domains outside the company a person has had at least two meaningful interactions in the last four weeks.	Person	Count	Yes
Networking outside organization	The number of distinct organizational units within the company that the person had at least two meaningful interactions in the last four weeks.	Person	Count	Yes
Open 1 hour block	Number of one-hour blocks in the person's calendar without meetings during the work day.	Person	Count	Yes
Open 2 hour blocks	Number of two-hour blocks in the person's calendar without meetings during the work day.	Person	Count	Yes
Peer average (customer collaboration)	The total amount (in hours) of customer collaboration for all of the participants in the plan divided by the number of participants in the plan.	Person	Hour	No
Peer average (internal collaboration)	The total amount (in hours) of internal collaboration for all of the participants in the plan divided by the number of participants in the plan.	Person	Hour	No
Redundant meeting hours (lower level)	Number of meeting hours a person spent in a meeting with both their manager and their skip-level manager present in the meeting.	Person	Hour	Yes

Metric	Description	Query type	Data type	Customizable
	This metric is *not* used in calculating *Low-quality meeting hours*. Analysts can use this metric only when creating Person queries.			
Redundant meeting hours (organizational)	Number of meeting hours a person spent with attendees from three or more distinct levels within that person's organization. Used in calculating *Low quality meeting hours*.	Person	Hour	Yes
Time in self-organized meetings	Number of hours spent in meetings organized by the person with at least one other person.	Person	Hour	Yes
Total calls	Total number of calls a person joined through Teams, including scheduled and unscheduled calls during and outside of working hours (as set in Outlook).	Person	Count	Yes
Total email sent during meeting	Number of emails the person sent during meetings.	Person	Count	Yes
Total focus hours	Total number of hours with two or more hour blocks of time where the person had no meetings.	Person	Hour	Yes
Working hours collaboration hours	Number of hours the person spent in meetings, emails, IMs, and calls with at least one other person, either internal or external, after deduplication of time due to overlapping activities (for example, calls during a meeting), during working hours.	Person	Hour	No
Working hours email hours	Number of hours the person spent sending and receiving emails during working hours.	Person	Hour	Yes
Working hours in calls	Total number of hours a person spent time in scheduled and unscheduled calls with Teams, during working hours.	Person	Hour	Yes
Working hours instant messages	Total number of hours a person spent time in instant messages through Teams, during working hours.	Person	Hour	Yes

PERSON METRICS

Metric	Description	Query type	Data type	Customizable
Workweek span	The time between the person's first sent email, meeting attended, or Teams call or chat, and the last email, meeting, call, or chat for each day of the work week. The total number of hours are based on the person's work week that is set in Outlook, which the user can change at any time. If a work week is not defined in Outlook (or if Workplace Analytics is unable to access a user's Outlook settings), the totals are based on the default of Monday through Friday, with a minimum of four hours and a maximum of 16 hours per day. If reported for the week, the metric is a sum of the daily values for the week. If reported for the month, the metric is the sum of the daily values for the month.	Person	Hour	No

Meeting metrics

MEETING METRICS

Metric	Description	Query type	Data type	Customizable
Attendee meeting hours	Total number of adjusted meeting hours for all attendees. A *meeting query* focuses on the meeting as the main entity and reports on the various meeting attributes; a *person query* looks from a person's perspective and aggregates multiple meetings for the selected time period. Because the two query types have different purposes, their output also differs.	Meeting	Hour	Yes
Attendees	Number of people who attended the meeting.	Meeting	Count	Yes

MEETING METRICS

Metric	Description	Query type	Data type	Customizable
Attendees multitasking	Number of attendees that sent emails during the meeting. • In meetings of one hour or less, two or more emails. • In meetings longer than one hour, two emails per hour. (Example: Sending four emails during a two-hour meeting would count as multitasking.)	Meeting	Count	Yes
Attendees with conflicting meeting	Number of attendees with meetings that overlap with the meeting (includes all non-declined meetings, which include accepted, tentative, and no responses to meeting invites).	Meeting	Count	Yes
Emails sent during meetings	Number of emails the person sent during all meetings.	Meeting	Count	Yes
Invitees	Number of people invited to the meeting.	Meeting	Count	Yes
Redundant attendees	The number of attendees of a meeting who are redundant, as defined by the *Redundant meeting hours (lower level)* metric. For more information about *Redundant meeting hours (lower level)*, see the table that lists Person metrics.	Meeting	Count	Yes
Total meeting cost	The total cost of all attendees in a meeting. The meeting cost for each attendee is defined as the product of the attendees' meeting hours multiplied by the attendees' hourly rates. If no hourly rate is available for one or more attendees, the default rate of $75/hr (US dollars) is used to calculate the cost of those attendees.	Meeting	Currency	Yes
Total redundant hours	The total number of redundant hours metric for all attendees in a meeting.	Meeting	Hour	Yes

Group-to-group metrics

GROUP-TO-GROUP METRICS

Metric	Description	Query type	Data type	Customizable
Collaboration hours	Number of hours spent in meetings, emails, IMs, and calls, after deduplication of time due to overlapping activities (for example, calls during a meeting) between the time investor and collaborator groups.	Group	Hour	No
Email hours	Number of hours spent sending and receiving emails between the time investor and collaborator groups.	Group	Hour	No
Meeting attendee count	Total number of attendees in all meetings from the time investor and collaborator groups.	Group	Count	No
Meeting hours	Number of meeting hours the time investor group has spent meeting with the collaborator group.	Group	Hour	No
Meeting invitee count	Total number of invitees in all meetings from the time investor and collaborator groups.	Group	Count	No
Meetings	Number of distinct meetings with at least one attendee from the time investor and collaborator groups.	Group	Count	No

Person-to-group metrics

PERSON-TO-GROUP METRICS

Metric	Description	Query type	Data type	Customizable
Collaboration hours	Number of hours that the time investor spent in meetings, emails, IMs, and calls with one or more people in the collaborator group, after deduplication of time due to overlapping activities (for example, calls during a meeting).	Group	Hour	No

PERSON-TO-GROUP METRICS

Metric	Description	Query type	Data type	Customizable
Email count	Count of unique email exchanges (sent and received) that the time investor had with one or more people in the collaborator group.	Group	Count	No
Email hours	Number of hours that the time investor spent sending and receiving emails with one or more people in the collaborator group.	Group	Hour	No
LastTimeContacted	The last date and time that the time investor (a measured employee) emailed or attended a meeting with one or more people in the collaborator group for the specified date range. Note that this metric refers only to interactions that were initiated by the time investor.	Group	DateTime	No
Meeting hours	Total number of hours that the time investor spent in meetings with one or more people in the collaborator group. This metric uses time allocation logic.	Group	Hour	No
Meetings	Number of unique meetings that the time investor attended with one or more people in the collaborator group.	Group	Count	No
Network size	Number of people in the collaborator group who had at least two meaningful interactions in the last 28 days with the time investor. This counts both licensed and unlicensed employees in the collaborator group.	Group	Count	No

Organizational network analysis (ONA) metrics

ORGANIZATIONAL NETWORK ANALYSIS (ONA) METRICS				
Metric	**Description**	**Query type**	**Data type**	**Customizable**
Diverse tie score	A numeric score that indicates how varied and how broad a person's connections are. This is based on both the infrequent direct collaboration between two people and on the differences in the common network they share between themselves. (Collaboration activities consist of emails, meetings, Teams calls, and Teams chats.) A person need not have much direct collaboration with their diverse ties, so it's easy to have more diverse ties than strong ties. Diverse ties present good sources of fresh and varied information from across the company.	ONA	Score	No
Diverse tie type	A value that indicates the relative diversity of the person's diverse ties. 0 means that the tie is not diverse; 1 means that the tie is diverse; 2 is an intermediate value that means more diverse than 0 but less diverse than 1. (The Diverse tie type metric is derived from the Diverse tie score metric, which in turn is based on the thresholds that are described in The last columns give the results.)	ONA	Score	No
Diverse ties	The number of diverse ties that the person has, that is, the number of ties whose Diverse tie type is 1.	ONA	Count	No
Influence	A numeric score that indicates how well connected a person is within the company. A higher score means that the person is better connected and has greater potential to drive change. (A person's connection score is based on the frequency of collaboration activities, which include emails, meetings, Teams calls, and Teams chats with other people within the company.)	ONA	Score	No
Influence rank	One of a sequence of numbers that starts with 1. A rank of 1 represents the person with the greatest Influence score; a rank	ONA	Score	No

Metric	Description	Query type	Data type	Customizable
	of 2 represents the person with the next greatest Influence score, and so on. If two people have the same Influence score, they also have the same influence rank.			
Manager overlapping strong ties	The number of strong ties that are in common between a manager and their direct reports. If a manager shares a significant number of strong ties with their directs, this can indicate that they are on the same page and executing on a well known, well understood, common plan. This metric reflects the manager's ability to ensure that the team is working toward progress and team members are up to speed.	ONA	Count	No
Manager unique strong ties	The number of strong ties that are unique in the manager's network when contrasted with the strong ties of their direct reports. This metric helps answer the question: "What is the potential of this manager to bring fresh connections and fresh ideas to their team?"	ONA	Count	No
Strong tie score	A numeric score that indicates how strong and tight a person's engagements are. It is based on both direct collaboration between two people and on the common network they share. (Collaboration activities consist of emails, meetings, Teams calls, and Teams chats.) For example, a "strong tie" between a manager and a direct report reflects both the amount of direct collaboration they have with each other and the time they both invest in connections that are common to both of them. Typically, a person has only a few strong ties because such ties take more effort to maintain.	ONA	Score	No
Strong tie type	A value that indicates the relative strength of the person's strong ties. 0 means that the tie is not strong; 1 means that the	ONA	Score	No

Metric	Description	Query type	Data type	Customizable
	tie is strong; 2 is an intermediate value that means stronger than 0 but weaker than 1. (The Strong tie type metric is derived from the Strong tie score metric, which in turn is based on the thresholds that are described in The last columns give the results.)			
Strong ties	The number of strong ties that the person has; that is, the number of ties whose Strong tie type is 1.	ONA	Count	No