Classifying information and its implications to business practices

The role of humans and systems in the process

Master's Thesis

Mariel Saretsalo
41186

Master thesis in the International
Information and Knowledge Management
Master's Programme
Supervisors: Gunilla Widén and Shahrokh Nikou
Faculty of Social Sciences, Business and Economics
Åbo Akademi University
Åbo 2020

Abstract

Over the past three decades, the volume as well as the diverse nature of information has increased exponentially. This is both an opportunity and a threat. The opportunity is that information enables us to make decisions smarter, faster, and to trade it even as it was a commodity. The threat is that without a proper system to store information, it becomes difficult, or even impossible to utilize the needed information. If information is irretrievable, all the effort used for storing it becomes wasted time. At the same time business processes are being influenced by the nature of the work, tools being used and an information flow between different stakeholders. All these factors have forced companies to seek efficient, reliable and cost-effective means to organize, manage, share, and dispose digital information in a logical fashion, and in a way that it can be accessed by employees in any location, time or information communications technology that is being used.

This thesis addresses the importance of file management, methods, best practices and technologies to classify information in the business context. The theoretical framework provides a basic understanding of file management operating parameters, such as how information can be classified and labelled for the users to access and retrieve data more quickly and more conveniently, and what is the role of human and systems in that process. The possible implications of the file management from a business perspective are also examined, such as employee performance and productivity.

The empirical research investigated existing file management practices in a case company. The overall aim was to provide recommendations as to how to improve existing file management practices from the current state. Descriptive research was used to examining the current state and maturity of the file management. The results received from the survey indicate that the current folders and files used by the employees in the case company are disorganized and poorly managed. There are no standardized practices to manage files, i.e. how they should be classified, named, or shared with relevant stakeholders. Furthermore, working conditions and work devices has caused additional challenges in terms of accessing files. Having multiple systems to store information, together with a lack of clear rules and guidelines regarding the file management activities, is interfering with the employees' ability to perform and focus on work task.

Although the empirical research was specific to the case company, both the research results and the theoretical framework support the fact that investing in well-designed information classification approaches and systems can result in greater productivity and process efficiency. Thus, impacting companies' overall ability to reach its business objectives.

Table of contents

List of Figures

1 Introduction

The transformation of an industrial society into an information society has changed business processes, especially the ways work tasks are being performed. Work is done independently and/or remotely by utilizing different technologies and systems. Thus, employees currently must manage a higher and more complex amount of information and understand interdependencies between data contents (Lindén, 2015). An industry report, titled "Dealing with Document Deluge and Danger", estimates that more than 2.5 trillion PDFs are created annually, and that this number will continue to grow and multiply in the future (BPI Network and Foxit, 2016). These factors have forced companies to seek efficient, reliable and cost-effective means to organize, manage, share, and dispose digital information in a logical fashion, and in a way that it can be accessed by employees in any location, time or information communications technology (ICT) that is being used.

There is no business without information. In order for the employees to perform at their best, which is also directly linked to business success, they must be able to find information easily. Most importantly they need to be able to trust it. “The high cost of not finding information”, an ICD report indicates that employees use nearly 15-25% of their working time for finding documents from their company’s information databases, and that only half of those searches are successful (as cited in Smallwood, 2013, p. 79). Poor taxonomies and lack of common procedures to classify content complicates tracking and protecting documents. It also increases the risk of duplications. All the above can lead to a low level of employee performance and difficulties to follow laws, rules and regulations related to information management. Thus, impacting companies’ overall ability to reach their business objectives.

Business practitioners suggest that file management practices can mitigate the challenges mentioned above. Therefore, this thesis aims to explore more in details the importance of file management, its methods, best practices and technologies in the business context. The empirical research will focus on investigating existing file management practices in a case company. The overall aim of the thesis is to provide

recommendations how to improve the existing file management practices from the current state.

1.1 Business context

The case company in the research has requested to stay anonymous, thus it will be referred as Case Company X (hereafter CCX). CCX is a local entity, part of a multinational corporation, with a focus on providing innovative solutions in the medical and health industry. Although, it has its own staff and business functions such as marketing, sales, and medical, CCX is part of a Nordic cluster. Thus, many business activities, including IT services**,** are shared across the Nordic countries. All the country business units (CBUs) in the Nordics, share a collective file management system called ***File Explorer*** which is an application, a part of Windows' operating system (hereafter OS), that can be used to manage files and folders. All files are being saved and stored in a shared drive called ***AShare***, in which each CBU have their own parent and subfolders targeted specifically to their use. New Technology File System (hereafter NTFS) permissions are being used to grant access rights to each parent or sub folder separately. In general, each employee has at least access rights to the functions in relation to the country-specific folders that they operate at. For instance, employees working in sales have access to their country's local sales folder. Employees also have the possibility to use OneDrive to store files, but it is only intended for personal file storage. Meaning, the CCX does not have OneDrive for Business.

Operating parameters, such as file management systems, belong to the IT department whereas the actual file management practices, *i.e.* organizing and managing information from the beginning of its creation to disposal, is each CBU's responsibility. There are no standardized file management practices in place, which has caused many challenges in terms of controlling, managing, tracking and protecting files in both repositories, File Explorer and OneDrive. Although, there is a common need across the CBUs, this thesis will focus on investigating the current state of the file management in the CCX.

1.2 Research questions

This thesis addresses the importance of file management, methods, best practices and technologies, and aim to provide recommendations as to how to improve existing file management practices from the current state in the CCX. Although the research is specific to CCX, this thesis aims to provide frameworks that could be utilized by other CBUs as well. To this end, the current thesis aims to answer the following questions:

- **RQ1: How can information be classified, and what is the role of systems and users in the process?**

This question aims to provide a basic understanding of file management operating parameters, such as how information can be organized, labelled and classified for the users to access and retrieve data more quickly and more conveniently. Also, it aims to investigate the possible implications of the file management from system, user, and business perspective, such as, how information can be classified, where the organized information will be stored and who uses the information.

- **RQ2: How is file management linked to business practices?**

This question aims to provide examples of how the frameworks in the RQ1 are linked to file management practices in the business context and what are the possible implications, such as employee performance. The empirical research will focus on investigating existing file management practices in the CCX. The overall aim is to provide further improvement recommendations towards sustainable file management practices which can be done by first understanding the current state and maturity of the file management practices in the CCX. This will be done by a survey in which the questions will be formulated based on the theoretical framework and sent to the whole staff in the CCX.

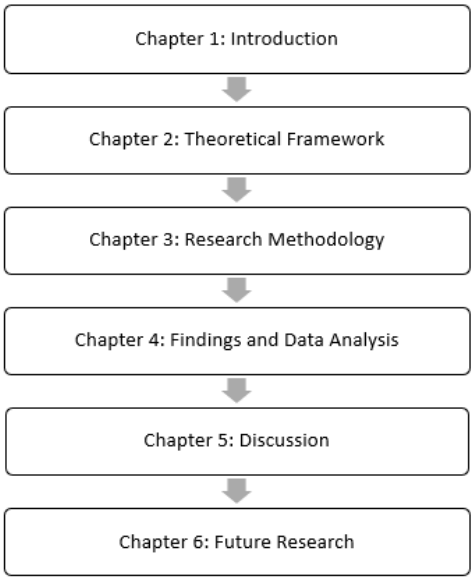
1.3 Context and terminology

Within the scope of information sciences there are many different definitions used to describe information stored in OSs, such as document, file, content, and record.

Each term has its own definition and methods to manage them. There are also different OSs depending on the information that is being processed or tool that is being used. In fact, the world of information management is filled with technical acronyms, such as ECMS (Enterprise Content Management System), DMS (Document Management System), CMS (Content Management System), and FMS (File Management System). It is critical to have a general comprehension of how these systems differ when choosing, for example, the best system for one's needs. However, there are many similarities in the way information is managed and how the systems function. From a holistic perspective, they all aim to aid and streamline the information management processes. Many theories are therefore applicable when elaborating on information in any format or in any system. (Lindén, 2015; Smallwood, 2013.)

To ease the reading process, the term 'file', 'file management' (hereafter FM) and 'file management system' (hereafter FMS) will be used throughout the work (see Figure 1 for the outline of the thesis). The concepts and theories presented in the literature review are applicable to digital files and folders in common computing environments, which is the scope of the thesis.

Figure 1 Outline of the Thesis structure



2 File Management

Throughout the history, in every culture, in every era, the two following questions have remained the same: i) which medium to use to collect or record information, and ii) which location to store the collected information. Over the past three decades, the volume as well as the diverse nature of information has increased exponentially. This is both an opportunity and a threat. (Kipngetich, 2014.) The opportunity is that information enables us to make decisions smarter, faster, and to trade it even as it was a commodity (Lindén, 2015). The threat is that without a proper system to store information, it becomes difficult, or even impossible to utilize the needed information. If information is irretrievable, all the effort used for storing it becomes wasted time (Lindén, 2015; M-Files, 2019).

According to an online encyclopedia Reference (2019): *"File management is the art of storing, naming, sorting and handling computer files. It is the process of maintaining folders, documents and multimedia into categories and subcategories as desired by a user".* In other words, FM seeks to organize, label and classify information in the computer so that a user can access and retrieve data more quickly and more conveniently. To ensure quality, safety and efficiency in the process, different protocols should be taken into consideration. For instance, concepts and principles related to information management, from creation to preservation, are described in the certificate ISO 15489 (International Standard Organization, 2016). Although the focus is on records, approaches and standards are very much applicable to files. Most importantly, this highlights the importance of managing information on a global scale. Many companies utilize this as a competitive advantage. For instance, M- Files (2019) uses ISO certificate as a proof that their system is reliable in terms of managing files systematically yet safely.

Various tools and systems have been developed to store and manage electronical files. A file management system (hereafter FMS) is an application that is used to store, arrange, and access computer files stored on a disk or other storage location. For example, File Explorer is a built-in file management application in Windows' OS (Moran, 2016.) FMS is an enabler that can fulfill the needs and requirements of the

user across the entire lifecycle of a file. Therefore, FMS does not only provide a functionality to store files and folders but also a collection of functions that can be performed with files. According to Stallings (2012, p. 524), the typical requirements and functions of the FMS from a user perspective are the following:

- **Create:** The lifecycle of a new file begins with the creation or receipt of a document that does not exist in the file system. During the production, all the necessary and desired information is added to the file.
- **Store:** Created file or folder is being positioned within the structures of files.
- **Open and use:** An existing file is being opened and a user can perform functions depending on a file's access control restrictions such as, read and write. 'Read' is the process in which the user can access the content of a file whereas 'write' allows user to add new information to opened file. Once user is done using the file, it is being closed until it is being opened again. File size change in case user has done modifications to it. File can also be moved (relocated) or copied to another folder. User can also include the following action
- **Share:** file or folder can be distributed to other users; for instance, as an email attachment or an URL link to the locations of the file (also known as file path).
- **Archive and dispose:** Once a file has reached the end of its lifecycle, the file should be either disposed (deleted permanently) or archived accordingly (e.g. to separate location)

Mokhtar (2017, p. 1), however, argues that in general, various tools have been designed by IT professionals and implemented by organizations with little understanding of the FM principles. IT professionals tend to focus on system functionalities and infrastructure while excluding FM aspects. As mentioned, there is a great number of FMS solutions available on the market, and the distinction between them may be difficult to understand. It is not surprising that systems are usually the starting point for companies when they wish to establish or develop their FM processes (Cadence Group, 2006). However, systems alone are not enough to guarantee improved FM practices. Even the system-oriented perspective in the information sciences addresses the role of a user. FM is a combination of a human and technology. Thus, the needs of knowledge workers and their commitment and motivation to utilize systems as part of their work should also be addressed in the framework (Aujirapongpanin et al., 2010).

The following sections aim to understand how information can be organized and categorized, and what is the role of systems and users in the process. Lastly, the business implications of FM will be discussed.

2.1 Organizing business information

Classification is one of the key foundational elements of FM. *"It is a holistic concept and activity for an information-related organization to organize and manage information from the beginning of its creation to disposal"* (Mokhtar, 2017, p. 19). In other words, classification is used to organize information and facilitate its retrieval in a systematic way. Taxonomy is a system of describing an object (Arthur, 2005) that can exist in different forms (e.g. lists, trees, hierarchies, system maps) (Arthur, 2005; Reinout, 2008; Smallwood, 2013). Classification and taxonomy are both methods used to organize and categorize information in a consistent way without the need to analyze each piece of information separately. The fundamental difference between the concepts, is that taxonomies describe relationship (maps) between objects while classification group objects that are alike (Reinout, 2008). To put it simply, taxonomy is the science and practice of classification. Universally known example of a taxonomy classification could be the presentation of animal kingdom. In the taxonomy animals are organized into smaller and smaller groups. The animal species are defined according to their relationship with the other species in the hierarchy.

In the context of FM, classification is one of the most fundamental activities when organizing files and constructing taxonomies. It can be used as a process or a system of grouping similar contents together that are being placed into a visual hierarchical presentation (taxonomy). Meaning that files are displayed and organized hierarchically in FMS. Furthermore, according to Smallwood (2013, p. 81) *"Taxonomies are the heart of the solution to harnessing and governing information".* Sam Goldman, a business data scientist, also shares this idea by stating that: *"Investing in content management is a valuable undertaking, but without creating a taxonomy as part of the process, companies could find themselves saddled with a next-*

to-useless solution...Put it another way, a company's taxonomy is the skeleton from which all content hangs" (as cited in Cadence Group, 2006).

From the users' perspective, the purpose of a taxonomy is to ease and guide users' behavior. It helps users to navigate and find information in a logical and familiar way, even if they are not sure what they are looking for. As Smallwood (2013, p. 82) states: *"Good taxonomy design makes it easier and more comfortable for users to browse topics and drill down into more narrow searches to find the documents and records they need".* As mentioned, taxonomies can exist in different forms, but they all have one thing in common, which is visualization of the information. Abstract data, that is being classified and presented visually, allows users to make discoveries, decisions, and explanations based on the internal structure and/or causal relationships in it (Smallwood, 2013, p. 83-84). For instance, a file name called 'certificate' stored in a folder called 'education' communicates to the user that this certificate is most likely related to learning rather than it being for example a product certificate. Navigation is a graphical interface, a part of OS, that illustrates objects and their relationship to other objects visually (e.g. files and folders). Parts of OS are examined further in section 2.4.

2.2 Building a taxonomy

The most underlying aim of the taxonomy structure is that it is easy to use by the end-users (Arthur, 2005). Therefore, building a taxonomy from scratch, or improving existing, requires a careful planning. Arthur (2005) suggest to start with a piece of paper and to think questions, such as: what the purpose of the taxonomy is and who are the users and their needs. Smallwood (2013, p. 96-97) supports this view and recommends to involve users, such as subject matter experts, testers and stakeholders, to the project at the very early phase since their knowledge can be used to gain better understanding of how each business unit function and interact. The operating parameters needs also to be considered – *i.e.* where will the taxonomy be applied. OSs have different technical components and functions which will be furtherly presented in section 2.4.

According to Wang, Chaundry and Khoo's study (2007), classification schemes strengthen the structure of the taxonomy and reduces time and effort in the development process. However, there is no single answer to questions; who is responsible or what is the best method to classify information. Classification in the business context usually includes managers from key business units. IT and legal units are also closely involved since one of the rationales of FM is to comply with standards and legislations (Smallwood, 2013.) Each employee should also have a basic knowledge about classification since they are very much involved in the process (*i.e.* end users create and search files on a daily basis).

There are different options available for the taxonomy building process, they can be built from the scratch either manually (Arthur, 2005) or by utilizing current technologies (Mokhtar, 2017). There are also prebuilt taxonomy templates available on the market. Prebuilt taxonomies can constitute industry specific best practices, and therefore allow faster start and implementation than developing a taxonomy from a scratch. However, if the prebuilt taxonomy does not fit well, tailoring it according to business' needs may end up taking longer than building one from scratch. (Smallwood, 2013, p. 88.) System generated taxonomies on the other hand utilizes metadata from existing files. All the files include metadata, that are descriptive data fields, such as author, size of the file, creation date, etc. In this approach, all the files with similar data fields are grouped together. Human contribution is also needed in this approach since system alone cannot assess the real value of the information and how it is used by the user (Mokhtar, 2017; Smallwood, 2013). These factors highlight the fact that there is no "quick or one size fits all" - solution available. However, the resource investment is worth it. Implications of organizing information and well-designed taxonomies for businesses are furtherly reviewed in the section 2.5.

2.3 Classifying information

As stated, classification helps us to organize and make sense of things. People responsible for classifying objects may, however, face problems since there are many ways to classify same concepts and/or objects (Batley, 2005). According to

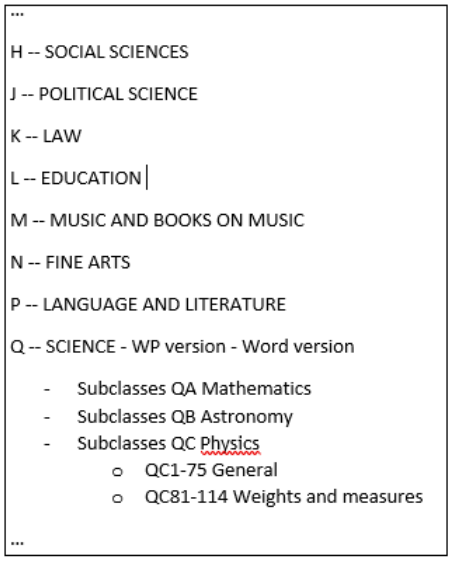
Brooks (2017) there are three typical approaches to classify business contents: subject, organizational and functional. Each type will be shortly explained together with some limitations.

2.3.1 Subject classification

In subject classification, objects are generally arranged in an alphabetical order – from the broadest subjects to more precise subjects. Yellow Pages, universally known, is an example that uses subject classification. Classification is widely researched and practiced especially in library science in which subjects are categorized by using classification schemes (Mokhtar, 2017). The International Federation of Library Associations and Institutions (IFLA, 2017) recommends libraries to use subject classifications to arrange books in libraries. Library classification includes schemes, such as Dewey Decimal Classification (referred as DDC), Universal Decimal Classification (referred as UDC) and Library of Congress Classification (referred as LCC). The latter, LCC, is the most commonly used in academic libraries across the world. LCC has a subject heading list that contains 21 main subject classes represented with alphabetical letters and titles (see Figure 2 below). Since the subject headings can often cover somewhat broad concepts, each class has its own subclasses (also known as subheadings and subdivisions). They are used to narrow the subject to a particular aspect. Subject headings are standardized and officially approved schemes, but the subclasses can be added as new information emerges. (Library of Congress, 2015.) To conclude, the focus of classification is coding and organizing library material – from identifying similarities between subjects to physically organizing books on shelves.

The purpose and theories of library classification has received some critique in terms of being retrieval oriented, and therefore the classification may be done from system application perspective rather than the context of the books. For instance, Mokhtar (2017) argues that lack of research has impacted the development of classification systems negatively. Meaning that library scientists' focus has been on how fast the object can be found while excluding other relevant classification principles.

Figure 2 Subject headings in library of congress (Library of Congress, 2015.)



Subject classification can be applied in FM practices, but it is quite limited in use. Most of the times, files are not produced on a subject basis and a single file can relate to more than one activity (Kipngetich, 2014). Figuratively speaking, imagine a situation in which the same book could simultaneously be on different shelves. An example from a business perspective could be a marketing department in which a new logo file for a client A is being produced. All the logos, including this one, are being stored in a folder called 'Client Logos'.

Figure 3 Duplication within classification (Arthur, 2005).

However, there is also a folder called 'Client A' that is a subfolder for a parent folder called 'Clients' that contains all client folders. The complexity with this approach is that the logo needs to be saved in both folders, resulting to file duplication as illustrated in Figure 3 above. (Arthur, 2005.) Subject classification works well for

short-term purposes that are related to a particular entity, person or project but should be avoided in general terms (Smallwood 2013, p. 92.)

2.3.2 Organizational classification

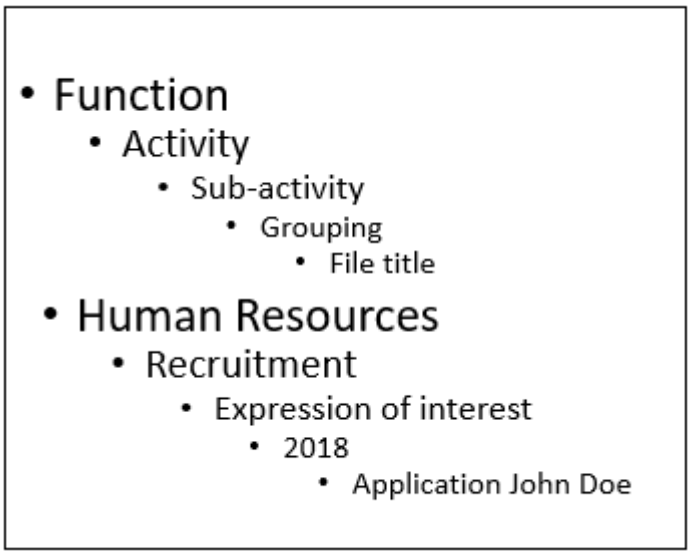
In an organizational classification, files are categorized based on the business units, such as a department or a division (Smallwood 2013, p. 92). Taxonomy; therefore, reflects the structure of an organization. From a user perspective, this makes it easy to decide where to store or find files since the structure often mimics the "original" paper-based filing schemas (Brooks, 2017). However, as mentioned business processes have evolved substantially from the past, hence organization structures are not as stable as they used to be. Different external or internal factors (e.g. mergers and acquisitions) can cause changes in organization structures (Lindén, 2015). This means that the created structure must be revised each time there is an organizational change. Another disadvantage with organizational classification arises when files are shared or managed among multiple business units (Smallwood 2013, p. 92.) For instance, a folder called 'offers' can consist of very different files depending on the business unit (e.g. sales, marketing and HR). The problem usually arises during the storing phase. The receiver needs to evaluate where to store this folder called 'offers' in the structure since another folder with the exact same name already exists. Thus, the risk of duplications is high with this approach as well (Brooks, 2017).

2.3.3 Functional classification

In a functional-based approach, files are classified based on the functions, activities and transactions carried out by the organization. Furthermore, classification of the file is done based on the context rather than the content. Meaning that the file will be classified according to why it exists (e.g. function) rather than what it is about (e.g. subject) (Kipngetich, 2014.) Organization's business processes can be used in the process in which all the folders in the computer are arranged hierarchically. As presented in the Figure 4 below, the highest level represents business function. Second level down constitutes activities performed by that function, and all the files

created as result of the activity are stored in the lowest level (Smallwood 2013, p. 92.) This approach is widely applied in the context of FM, especially when creating a file plan. It is an outline that describes every file in the system, the storage location, rules, such as retention schedule applying to them, and a person responsible for their management. Hence it goes much further than file classification (Kipngetich, 2014.)

Figure 4 Illustration of functional based classification in file directory



Even though this approach is less subjective than a subject-based, it is more enduring than an organizational-based approach. Meaning that functional classification is better suited for organizational shifts and changes. Organization structures may change and evolve but the functions and activities performed usually remain much the same over the time. Created structure will not; therefore, be affected when an employee exits the company (Brooks, 2017.) Functional classification is also flexible in terms of adding new transactions or activities (Smallwood 2013, p. 92). However, if executed to extremes, the structure can become over-complex (Kipngetich, 2014). Each businesses and functions have their own language (also known as jargon). Meaning that people working in the same function and/or activities are familiar with used terms, but others may not (Ravitch, 2007). Thus, a newcomer or an employee from a different function may face challenges finding needed information. Especially, if the folder names contain

context specific terminology and/or codes, or the structure is based on function activities that the user has no further knowledge about. Another disadvantage with functional approach is its inability to handle project or case files and folders. Placing a project file in the hierarchy may be troublesome since these files and folders are a collected, as mentioned, for a certain short-term purpose related to a particular entity, person or project (Smallwood 2013, p. 92).

As illustrated above, each classification type has its pros and cons. There is also a possibility to use a hybrid taxonomy in which two or more approaches are combined or used in parallel. For instance, human resources activities are fairly constant, and therefore a business unit taxonomy could be applied even though the rest of the units would use a functional structure (Smallwood 2013, p. 92).

2.3.4 Thesaurus and controlled vocabularies

It is difficult to establish universally recognized schemes. Meaning that users that are unfamiliar with the topic may not know the appropriate term when searching an item. Classification schemes are therefore supported by a thesaurus which is a controlled vocabulary that shows relationships between terms such as, synonyms or related/preferred terms (*i.e.* suggestion or similar terms), and hierarchical terms (*i.e.* broader and narrower terms) (Brooks, 2017). These terms are usually displayed in the index function. In the Yellow Page example (subject classification), the index function in the back of the book would be a basic thesaurus. A person looking for a pastry shop would most likely begin the search by reviewing letter 'P' (pastry). If the query is unsuccessful, the user could use an index with preferred terms, in which "see also" would direct the person to review letter 'B' (bakery). To conclude, a thesaurus helps users to find the searched subject, create new search terms, and understand the vocabulary of the topic, and possible give suggestions of how to expand or refine the search (Library of Congress, 2015).

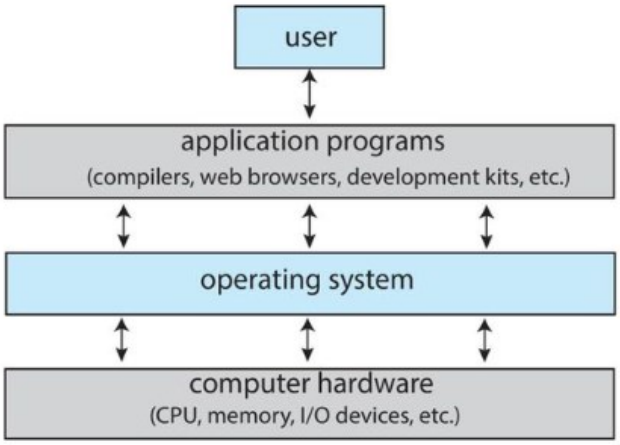
2.4 FM system architecture

One of the considerations when choosing a classification scheme is to consider user needs and operating parameters – where will the taxonomy be applied. OSs have different technical components and functions. Since FM seeks to manage computer files, the following sections aims to understand basic system functionalities (operating parameters). Typical user actions related to managing files and folders will also be reviewed. According to Stallings (2012, p. 524) these are; create, open and use, share, achieve and dispose.

2.4.1 Operating system

A computer is a programmable electronic set of resources that controls movement, storage and processing of data (Stallings, 2012, p. 50). OS is a software that is responsible for managing these resources on a computer's hardware. Each system has basic application programs and manages all the communication between the user and the hardware as illustrated in Figure 5 (Silberschatz, Gagne and Galvin, 2018). Hardware and OS walk hand in hand, thus the hardware defines which OS user can run and install**.** For instance, Windows and Linux can be installed on standard PCs whereas OS X is designed for Apple systems. It is important to understand that all devices, such as smart phone, smart watch, router, and tablet, in addition to computers require an OS. As mentioned, applications are typically created for specific OS since each OS communicates differently and has specific program interface. Many applications and programs are however compatible with multiple OSes (Christensson, 2016). While each OS is different, their functions in the computer systems are the same – providing a graphical user interface and repository to manage files and folders (Moran, 2015; Stallings, 2012). According to Stat Counter (2019), the most common desktop OS is Windows with a market share of around 74%. It is also used by the CCX, and therefore the following chapters will focus on features of the Windows OS.

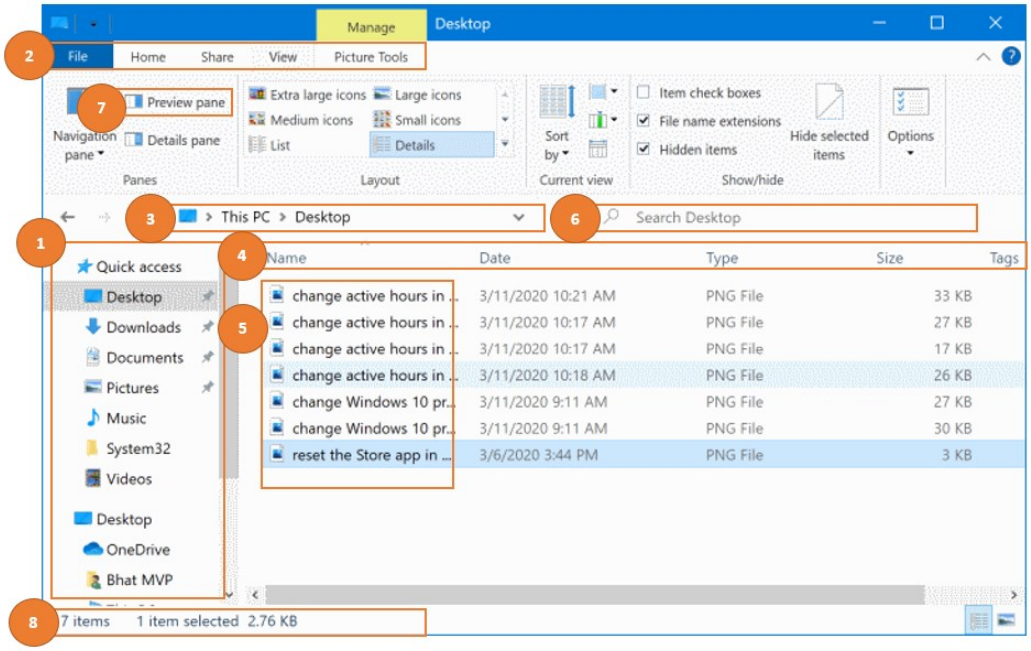
Figure 5 Abstract view of the computer component (Silberschatz et al., 2018)



2.4.2 File Explorer

File Explorer, also known as Windows Explorer, is a file manager application program developed by Windows. It has different versions (e.g. Windows 7 and Windows 10) and it can be installed on multiple brands of hardware (Christensson, 2016). Each version includes a graphical interface with a desktop that allows users to manage files and folders. Graphical interface of the File Explorer with the main parts and functions in Windows 10 are displayed in Figure 6.

Figure 6 Parts of a File Explorer window (Moran 2015, p. 2)



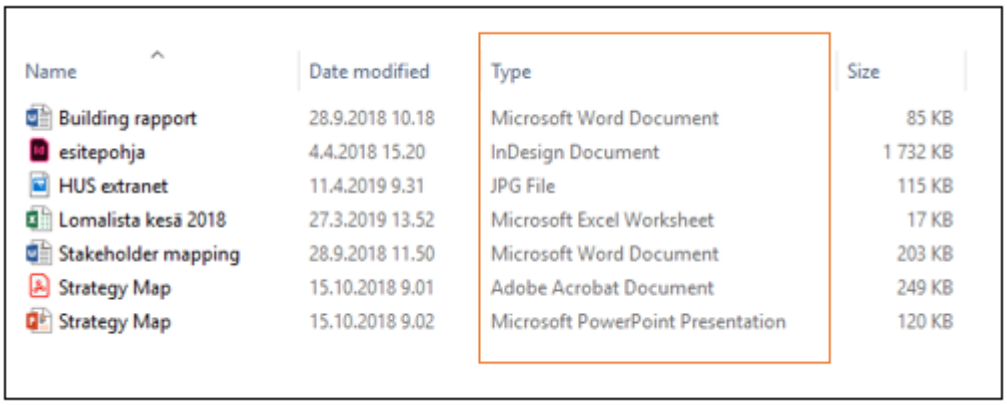
1. Navigation pane allows the user to choose which location to access. It displays all the folders, shortcuts to different storage spaces.
2. Ribbon with a set of tabs labeled ‘File’, ‘Home’, ‘Share’, and ‘View’ provides an access to perform different task or common operations on a selected item. View may differ depending on the location, file or folder that is being selected.
3. Address bar displays the path of the user's current location.
4. Column headings display details that is being that is being displayed in each column. Columns can be modified (e.g. remove or add more informative details) and the details can be sorted (e.g. date modified).
5. File and folder listing display all the files and folders in that specific location.
6. Search box allows user to search subfolders, documents, images, etc. in the current folder.
7. Preview/Details pane allows user to preview an item that is being selected without the need of opening it with a program. This is not a default feature but can be activated in the ‘View’ task bar.
8. Status bar displays information about the folder, such as number of files in the folder.

As stated, the navigation pane displays all the folders, shortcuts to different storage spaces. Roughly categorized, this section has two main parts; i) access to folders and ii) storage spaces. There is also a section called 'quick access', which allows user to access most frequently used folders quickly and conveniently. Windows automatically pins some folders, but users can add their own shortcuts as well. Older Windows' versions have a similar section, but it is called "favorites". Second section displays different storage spaces, such as external storage devices or network folders which are briefly explained further in section 2.4.5. The user can choose preferred location, but some programs will typically use the following default locations to store files once a file is being created (Moran, 2015, p. 5). For example, files created with programs such as Microsoft Word and Microsoft Excel are being stored to documents, files downloaded from the Internet are being stored to documents, music to music folder, and so forth.

2.4.3 File types

There are multiple different definitions to files as well as how they are categorized in the systems. In computer science, a file is a container that holds particular type of information; for example, a document, an image, a program or an audio. In File Explorer, files are most commonly categorized with extension, three- or four-letter -letter abbreviation, that signifies the type of file, file format and the attributes associated with the file. Each extension has its unique icon, which indicates to the user and to the OS, which application can be used with the file (Computer Hope, 2019). For example, in a text file the extension '.doc' indicates that the file can be opened with the Word or another compatible application program. File Explorer typically hides the file extensions from the file name view since the system generally shows in the 'type' column what type of files are being displayed (Moran, 2015, p.13). See figure 7 below.

Figure 7 Type column in File Explorer (Moran 2015, p. 14)

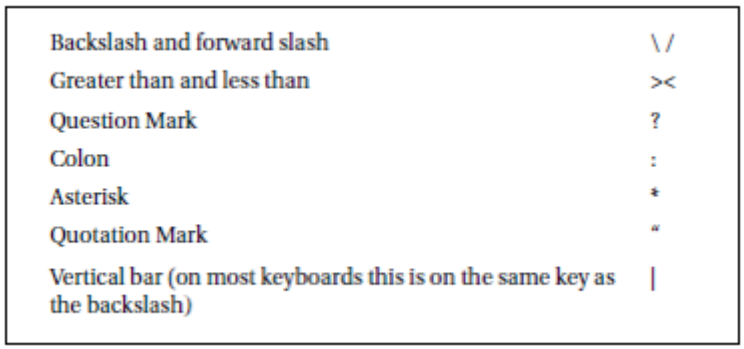


Name	Date modified	Type	Size
Building rapport	28.9.2018 10.18	Microsoft Word Document	85 KB
esitepohja	4.4.2018 15.20	InDesign Document	1 732 KB
HUS extranet	11.4.2019 9.31	JPG File	115 KB
Lomalista kesä 2018	27.3.2019 13.52	Microsoft Excel Worksheet	17 KB
Stakeholder mapping	28.9.2018 11.50	Microsoft Word Document	203 KB
Strategy Map	15.10.2018 9.01	Adobe Acrobat Document	249 KB
Strategy Map	15.10.2018 9.02	Microsoft PowerPoint Presentation	120 KB

2.4.4 File names

Files and folders can be named as preferred but there are a few limitations with the characters in the File Explorer. First, the name itself can contain maximum of 255 characters. Secondly, the complete path to a file cannot be more than 260 characters. File path indicates the specific location of the file in the File Explorer and comprises the file name plus the names of all the folders and subfolders that leads to it (see Figure 9). Lastly, the name cannot contain certain characteristics as presented in Figure 8.

Figure 8 Invalid file name characters (Moran 2015, p. 20)



Backslash and forward slash	\ /
Greater than and less than	><
Question Mark	?
Colon	:
Asterisk	*
Quotation Mark	"
Vertical bar (on most keyboards this is on the same key as the backslash)	|

The majority of scholars in data management seems to focus on using metadata to generate file and directory names. File names serve two important functions. First, they enable uniquely to identify a file over time. Second, description eases the process of recognizing file and its content to the user. (Parker-Wood, 2014, p. 73.). According to Smallwood (2013, p.86) metadata is a vital part of taxonomies and should be leveraged in the designing process. It enables the user to have multiple

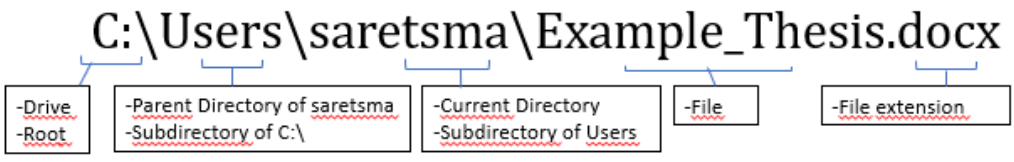
ways to retrieve information. Most importantly, it reduces time used to find a file since metadata used in categorizing increases the accuracy of the search. So rather than getting a large pool of results the user will more likely find the needed piece of information more easily whether it is a specific file or folder.

Naming convention on the other hand is a systematic method for naming files. One can use dates (e.g. YYYY, MM-DD), sequential numbers, or versioning (e.g. V1, V2, V3, etc.). Single files are not in the scope of the study, but the naming convention from a user perspective servers two purposes: i) it illustrates the file content without the need to open it; and ii) it enables users to retrieve and filter files. (Antin, 2016.) The search box in the File explorer, for instance, uses file contents as well as its name when user is searching for a file (Moran, 2015, p.76).

2.4.5 File storage

A directory is a location used for storing files and folders on a computer. All the folders in the File Explorer are arranged hierarchically, in which the highest level is a root folder. The location of an individual file within a directory can be represented with a directory path as illustrated in the figure 9. This path can be copied and shared to other users, which allows other users to access the shared folder path directly from the link. As mentioned, shortcuts to files, folders or programs can be also be added to the navigation pane which can enhance and quicken the information retrieval process.

Figure 9 Example of a directory path and Word document file extension in the File Explorer



Folders do not have extensions like files. Technically, they are the visual representation of a directory and a location to store files (Christensson, 2016).

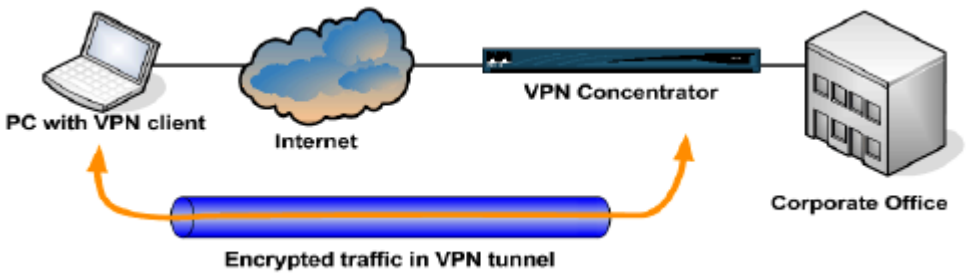
Folders help users to keep files organized, and without the folders all the files would be stored in one place. Folders also enable users to have multiple files with the same name whereas each file would require a unique file name if they were all stored in one folder (Computer Hope, 2019). From an information architecture point of view, one must consider depth and breadth when designing taxonomies (*i.e.* how many parent and subfolders to have). According to Brooks' findings (2017) scholars in the information management fields recommends having seven to twelve top level parent folders with no more than three sub folders below them. Morville and Rosenfeld (2007) supports this and recommends being even more conservative when considering the depth. If users are forced to click through too many levels to find the needed information, they will give up or at the very least, become frustrated. Similar consideration applies with the breadth as well. Therefore, adding folders to the File Explorer should be controlled. (Morville and Rosenfeld, 2007.)

Files and folders can be stored in multiple locations. PC's internal storage is typically labelled as 'OS' or 'Local Disk C' or 'Windows C'. Windows saves the main files and software programs by default to this location. In addition, each computer has a desktop, into which the user can save files and folders. The common factor to these locations is that the user can access this location even without the Internet connection. Despite computer, both laptop and desktop, hard drive technology has evolved, internal storage capacity is still limited. The user can see the amount of used disc space but controlling it is time consuming. Freeing up more disc space can only be done by cleaning up dispensable data whereas with other solutions one may easily expand the storage size (*i.e.* cloud storage). Also, one of the disadvantages with this type of internal disc storage solution is that information is not secured properly (e.g. backup function). In case the computer gets lost/stolen or the hard drive damages, restoring the files is troublesome. (Moran, 2015.)

In addition to internal storage, there are different kinds of external storage solutions available. Factors such as number of users, flexibility, expandability and accessibility determines which solution to choose. One of the most commonly found solution in business settings is a server. According to Christensson (2016): "*A server is a computer that provides data to other computers. It may serve data to systems on a*

local area network (LAN) or a wide area network (WAN) over the Internet." To put it simply, the server is a computer that shares information with other computers (also known as client computer). There are different types of servers available (*i.e.* web servers, mail servers and file servers), and depending on the type, different devices can be used to access the files. Therefore, different factors need to be considered when choosing a server. For instance, the server type determines which devices it can be accessed with, *i.e.* accessing File explorer located in a server that has Window OS may be troublesome with an iPad since Apple OS does not fully support applications designed to Windows. Also, accessing files in certain servers require the Internet connection. In addition, information access to the server should be safe. Companies' overall objective is to protect the integrity, confidentiality and the availability of any information (Sobh and Aly, 2010). In general, the server in the client computers can be accessed through a corporate's local area network (LAN). Meaning that all devices used by the company are connected to the same network. A virtual private network (VPN) is used to secure remote connections to a LAN (*i.e.* an employee that needs to access server from home). VPN can be illustrated as a tunnel in which information can move in a secure manner as illustrated in figure 10 below. Furthermore, it uses security mechanism to prevent unauthorized users from accessing company's files and ensures that they cannot be modified without detection. (Ismail and Ismail, 2010.) For instance, companies can use access control in which a person needs to be identified with a password before they can access the VPN, and thus access the server. In fact, VPN can be used to secure a private access and connection to all internal applications used by the company. In other words, the VPN allows you to work remotely, as if you were at the office.

Figure 10 Sample of VPN technology implementation (Ismail and Ismail, 2010)



There are also multiple cloud storage possibilities, for instance OneDrive, Dropbox, and Google Drive. OneDrive is a cloud file hosting service developed by Microsoft that allows users to: i) access and edit files with all type of devices; ii) protect files in the cloud; iii) share files with others; and iv) conveniently organize and find used files. (Moran, 2015; Microsoft, 2019.). OneDrive is a default application integrated into Windows OS. Furthermore, there are different OneDrive alternatives available depending on certain factors, such as number of users and purpose of the usage. Meaning that companies can, for instance, purchase OneDrive Business that has functions designed to support business processes whereas the default OneDrive in Microsoft solution, also known as consumer basic model, is only for personal use. Hence, it is limited in wider use (Microsoft, 2019).

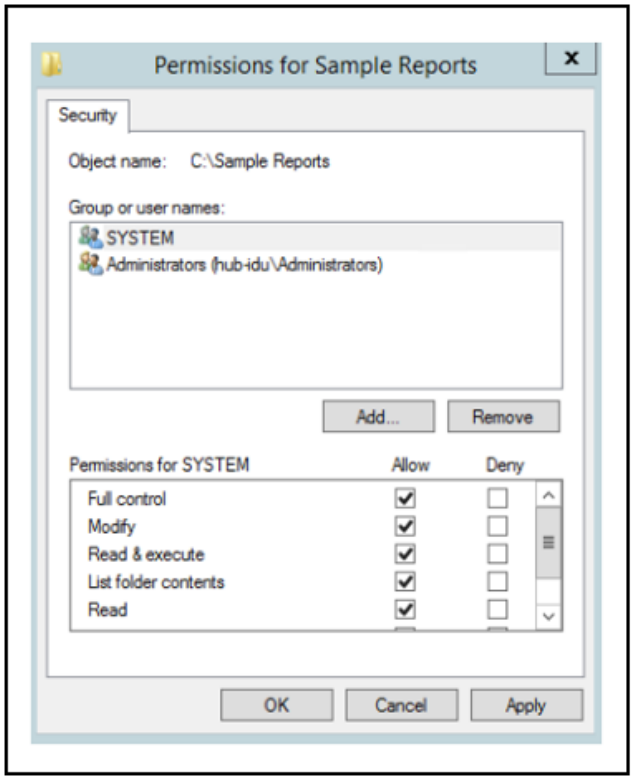
In general, saving and accessing files fall into one of two categories; online-only and available offline. The online-only refers to files that exist in cloud storages. Thus, accessing files requires the Internet connection. Files and folders located online-only do not take disc space from the computer. Files and folders can also be made available offline. Typically, this function can be controlled. For instance, user can manually choose which files and folders can be accessed offline without the Internet connection. Done changes are updated (also known as sync function) once connected to the Internet. This latter option, however, takes up PC's storage space (Moran, 2015, p. 93).

2.4.6 File usage and access rights

As Sobh and Aly (2010) stated, companies' overall objective is to protect the integrity, confidentiality and the availability of any information. For instance, a New Technology File System (NTFS) is another way to reach such objective. It serves two important functions. First, it allows companies to share permissions to drives and folders located in the Windows' network (see Figure 11). Permissions, such as read, write and execute, can be set for individual files and folders, and they are being granted on an individual level (*i.e.* tied to Windows username).

A second advantage of NTFS is, that it includes features to improve reliability. For example, it includes fault tolerance, which automatically repairs hard drive errors without displaying error messages. NTFS can also be used to track hard drive errors with a detailed transaction log. Thus, it can be used to prevent hard disk failures in the future (Christensson, 2008). In business settings, permission rights are generally managed by administrators or folder (directory) owners (e.g. manager of the department). Permissions can either be given or requested (Moran, 2015; Smallwood, 2013).

Figure 11 Example of a permission and ownership view in File Explorer



2.5 FM and business practices

FM practices and systems are vital for the organizations to be competitive in today's knowledge-oriented markets in which knowledge can be considered as a commodity (Akhavan and Pezeshkan, 2014; Lindén, 2015). FM is a continuous process, not a project with an ending period. Therefore, companies should adjust and monitor their resources (assets) and processes (capabilities) on continuous

manner. In order to understand how FM can be improved and/or implemented, one must understand the current situation in which company needs and resources are evaluated (Aujirapongpanin et al., 2010; Seo et al., 2010). The current situation is evaluated from macro to micro level questions, such as i) what is our business goal?; ii) what are the needs of our employees or consumers?; iii) what types of systems we are currently using and are there some limitations?; iv) where are the files stored?; v) who has access to which files?; and so forth (Lindén, 2015, p. 33-43).

In addition to planned FM improvement processes, businesses also need to be agile and have an ability to adapt their FM practices in cases of unexpected or uncontrollable events, such as natural disasters or geopolitical events. Agility is all about reacting to changes whether they are expected or unexpected, and to which businesses should have an ability to adapt in a timely and adequate manner. However, changing, the whole information system infrastructure from financial and operational perspective is costly and takes time, whether the need is planned or unexpected. (Seo et al., 2010). Therefore, Verstraete (2004) suggests that business agility from system perspective could be improved with minimum effort, cost and risk if companies would evaluate the impact of smaller system components rather than trying to replace the whole system at once.

Developing or improving FM is not a straightforward process and there are many different process flow descriptions in the literature. Lindén (2015, p. 18-19) suggest the following: *"Although the fluency and sensibleness of work for knowledge-workers is admirable, business performance is usually the main driver in information management initiatives. Performance in businesses are measured in numbers".* However, it is important to understand that the financial benefits of FM are not always clearly measurable since the development initiatives are linked to different processes *i.e.* people and system (Lindén, 2015; Seo et al., 2010).

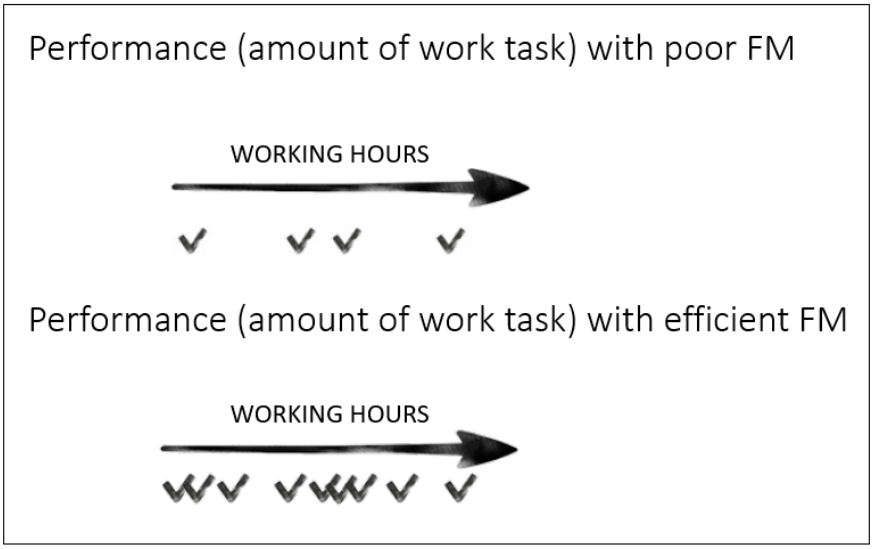
2.5.1 File usage and activities

While FM encompasses so much more than working with files or FMS, the core functionality of this research is not only the FMS' ability to store files and folders but

also a collection of functions that user can be performed with the files. As stated, the typical requirements and functions of the FMS is the ability of users to create, store, open and use, share, archive and dispose files (Stallings 2012, p. 524).

M-Files' study (2019), conducted with small to large businesses, global to national level, with responses from 1,500 office workers found that nearly half of the employees reported experiencing challenges in finding needed information. In addition, 83% reported creating a new file completely because they were unable to find it on the corporate network. As the Candence Group (2006) outlines: *"Well-designed classification system allows users to have better and quicker access to information; resulting in greater productivity and process efficiency."* According to Lindén (2015) productivity, and possible even job satisfaction, increases when knowledge-worker is able to perform a greater number of work task per day due to an ability to access relevant information quickly. Of course, other external (organizational) factors, such as atmosphere and leadership, influence knowledge-workers' job satisfaction and performance, but employees who are satisfied with their work are 25% more productive than those who are not satisfied. Lindén (2015) argues that a well designed and implemented FM practices are an important part of these external factors that influence employees job satisfaction. When employees feel that they have control over their work tasks, productivity increases, thus it increases the amount/number of work tasks finalized (see Figure 12), whereas poor FM practices reduces the productivity and decrease the amount/number of work tasks completed. It also increases the risk of using and sharing false/outdated information which can cause additional unfavorable costs to the company.

Figure 12 Relationship between performance and FM (Lindén 2015, p. 22)



Lindén (2015, p. 25) reminds us that companies that are successful in terms of profit, do not necessarily manage information efficiently. Some companies just can afford to pay for non-productive work in which people use comparatively high amount of time looking for an information that they need. In the worst-case scenario, they interrupt a colleague by asking for help to find this piece of information, resulting in two employees being less productive (Lindén 2015, p. 25).

2.5.2 Operating environment

In the digital society new ICT (information communication technology) tools are constantly being developed in addition to FM systems. Hence, the requirements for information management has changed from the past. As Mokhtar's (2017, p. 4) states: *"Communications have become less centralized and workplaces frequently virtual."* Work has therefore become increasingly a thing you do rather than a place you go, and business processes are being influenced by the nature of the work, tools being used and an information flow between relevant stakeholders.

According to Lindén (2015, p. 73) work in the modern businesses is performed with the following terms:

- Information needs to be accessed with different devices

- Information needs to be available around the clock
- Information needs to be available no matter of end-user's location
- Same information needs to be reviewed with different systems
- Information needs to be accesses no matter the language of the content

M-Files' (2019) study also supports the needs of the modern workers. The vast majority (81%) of the respondents reported that they need to access corporate files with mobile devices and only 38% of these respondents reported that it is quick and easy to find needed information with a mobile device. One of the challenges with mobile devices was also the ability to edit or share files (M-Files 2019). Businesses; therefore, do not only need to be proactive on acquiring and keeping their information sources up-to-date but also have nimble operating structures that allows them to react to these changed needs of modern workers (Seo et al. 2010). Furthermore, Seo et al. (2010) see information systems as one of the pillars supporting organization practices while generating competitive advantage. Thus, information systems should not be treated as independent objects but as part of the operational processes just like employees.

2.5.3 File management systems

In addition to systems functionalities, FM gives better premises for companies to respond compliance request, maintain information security and comply with regulations. For instance, retention and disposition scheduling is easier when similar files are grouped together (Smallwood, 2013, p. 11). For instance, General Data Privacy Regulation (GDPR) states that companies need to protect all the personal data, and they are obligated to communicate to which purposes and for how long the personal data will be used (European Commission, 2018 Article 25, Data protection by design and by default). Depending on the OS, files can be scheduled for destruction after a certain period. Either they are disposed automatically, or the destruction needs to be approved or denied separately. Who has access to the personal information is also covered by the regulation, which can be controlled by using, for example, NTFS permissions.

In addition to dealing with growing amount of information and the dilemma of classifying it, the additional issue is the growing number of different systems and repositories. Even though from a holistic perspective they all aim to aid and streamline the information management processes each system is specialized to manage a specific type of information. Examples such as, CRM (customer relationship management), ERP (enterprise resource planning) or CLM (contract lifecycle management) (Lindén, 2015 p. 44). In addition to that companies have traditional FMS such as File Explorer but also cloud based systems such as OneDrive or Dropbox (Moran, 2015). List of systems is long, and despite of the integration possibilities, having information scattered across a variety of systems has several implications.

According to M-Files study (2019):

- on average each organization has 4 different repositories to store and manage information
- 69% store and manage documents in their email inbox, 55% rely on shared network drives, and 24% use document management system.

In addition, 91% agreed that their job would be easier if they could quickly find and access the most current version of a document without having to worry about which system or repository it resides in. Although the user experience of the systems is admirable, security and risk minimization should be considered as vital part of FM approaches. All in all, part of FM initiative is to give better premises businesses to have trusted and reliable information available, which consequently allows end-users to make decisions with more confidence, and to comply with compliance and regulation aspects.

3 Research Method

Broadly speaking, there are two most frequent research methods referred in the literature, quantitative research and qualitative research (Adams et al., 2014, p. 6). These methods do not conflict with each other. In fact, they complement each other. For instance, numbers can add more insights, texture and context to the qualitative data. Qualitative research, on the one hand, attempts to understand motives and problems of human behavior through analyzing participants' perspectives about lived experiences of the event. A quantitative research, on the other hand, is useful for developing and employing numerical models and confirming theories and/or hypotheses. Thus, quantitative research is expressed in numbers and graphs whereas qualitative research is primarily done in non-quantitative characters and expressed in words (Adams et al., 2014; Saldana, 2011; SurveyMonkey, 2019). Same data collection methods (*i.e.* surveys and interviews) can be used for both, quantitative and qualitative approaches. The importance is to understand which method allows researcher to answer designed researched questions (Saldana, 2011).

This thesis has addressed the importance of file management, methods, best practices and technologies. The primary objective of the research is to provide recommendations as to how to improve existing file management practices from the current state in the CCX. This can be done by examining the current state and maturity of the FM in the CCX. Descriptive research was chosen to this thesis, which is a quantitative research method that attempts to collect quantifiable information to be used for statistical analysis of the population sample. It is primarily concerned with finding out "what is" rather than understanding "why" a certain phenomenon occurs (Adams et al., 2014; Bhat, 2019). There are several advantages of using descriptive research. For example, i) there are three distinctive methods that can be used to conduct descriptive research: observational method, case study method and survey research, ii) data can be collected in both qualitative and quantitative manner which gives better premises to have a holistic understanding of the topic, iii) data can be collected in the natural environment of the respondents which can increase the quality of the data, iv) data collection methods are quick and cheap to

conduct, and lastly v) if the sample represents a larger population, it is easy to make decisions on the basis of the statistical analysis of that data. The study group and data collection method are further explained in the following sections.

3.1. Study group

As stated, the case company in this thesis, i.e. CCX has requested to stay anonymous. The scope of the study focused on the CBU located in Finland which has 20 employees. Half of the employees work as a Product Specialists and are part of the sales function. Sales people mostly operate on the field, having face-to-face (F2F) meetings with their customers around the Finland. On average, they have one office day per week for administrative and sales preparation related tasks that they do remotely at home. The other half of the employees are office-based colleagues, which consist of marketing, sales and other administrative support functions. Office based employees mainly operate at the office, but like with sales, they have a possibility to work remotely from home. CCX has provided tablets to the whole sales team whereas the office-based employees use laptops, but tablets can be requested if needed for the job. Based on the background information, the main reason for having tablet as a main tool for sales people is that sales materials are more convenient to be presented via tablet when visiting customers. Based on sales people's needs, materials (*i.e.* sales presentations, leaf behinds, etc.) displayed and shared during customer visits are mainly provided by marketeers.

The CCX uses collective file management (File Explorer), in which they have their own country specific folder FI with subfolders dedicated to specific business functions such as sales, marketing, commercial leadership, and so forth. Access rights to the folders have been given based on employee's function, thus sales people can access sales folder and marketing people has access to marketing folder. In addition, the CCX has one jointly shared folder called 'FIUSERS' which has files that are relevant to the whole organization, meaning all the employees in the CBU has access to the folder. As mentioned, CCX is a part of the Nordic cluster. Thus, some folders and files are being shared across the CBUs. For example, marketing materials used by all Nordic countries are stored in a dedicated marketing folder (*i.e.*

marketing Nordics) whereas local marketing materials are stored under the local folder (*i.e.* marketing Finland). Country specific marketing folder is accessible only for marketeers operating in that CBU whereas the 'Nordic folders' is jointly used by all the marketeers across the CBUs. In addition to File Explorer, everyone has their own personal OneDrive that has been appointed to be used for personal files.

It is important to note that this research is very specific to the CCX. Also, the study group is relatively small. Thus, findings will not be possible to be generalized which was one of the possible advantages of using descriptive research method. However, as Saldana et al. (2011, p. 34) highlights: *"the amount of the participants in the study can depend on many factors, but as long as researchers have sufficient interview data, whether it is from one person or twenty, one should have sufficient corpus for analysis"*. Considering the number of employees (N = 20) working at the case company, the chosen methodological approach deems to be appropriate. The chosen descriptive research method and study group should be sufficient to examine and analyze the current state of FM practices in the CCX.

3.2. Data collection method

The data collection has been conducted via an e-survey which is one of the data collection methods in descriptive research. According to Adams et al. (2014, p. 127) the survey process begins with the design planning in which purpose, delivery method and a sample selection is done. The actual layout of the survey includes writing the questions. The general design principle of the survey is to keep it clear and short, ask only relevant unambiguous question and to use scales that are all going one way. In addition to these, the return mechanism should be simple.

The survey had total of 16 questions which can be found in Appendix 1. Questions were formulated based on the theoretical framework that comprised different activities that users can performed with files. As Lindén (2015) suggests, the current situation in the company can be evaluated from macro to micro level questions such as who is using the files, what are the needs of users, are there some limitations, and so forth.

The survey consisted of both, one open and closed-ended questions, and 15 of the questions had pre-defined options that are measurable with Nominal and Ordinal Scales. As Lindén (2015, p. 18-19) stated, business performance is usually measured in numbers, and it is the main driver of FM initiatives. However, the financial benefits of FM are not always clearly measurable since the development initiatives are linked to different processes i.e. people and system. Numbers can add more insights, texture and context to the qualitative data (Adams al., 2014, p. 6), and therefore the motivation to use questions such as "how much" were used to receive measurements to specific problem or opportunity. One open ended question in the end allowed respondents to describe issues and/or state their feelings related to existing FM practices within the CCX.

Adams et al. (2014, p. 127) suggest piloting the survey before administering the actual survey. This is done to ensure that the survey is clear and free of ambiguous expressions to respondents and to see if something needs to be adjusted in order it to be completed the way intended. Wyatt (2000) also pinpoints this by stating that one downside of conducting survey electronically is that it does not leave much room for asking questions in case the research questions are not clear to the respondent. To avoid such problem, the survey questions were piloted and sent to two mangers to review. Feedback such as, was the questions easy to understand, were they in a logical order and how long did the completion of the survey take were asked. Based on the feedback from the Sales and Marketing Manager, two answer options were adjusted.

The survey was done via a tool called Lyyti. The link to the survey was sent via email to the whole staff in CCX, consisting total of 20 people. The email contained details, such as, responses will be given in anonymous manner, estimated time for the completion, device that can be used, and wished date for the completion. Wyatt (2000) pinpoints additional advantages and disadvantages with such approach. On one hand, data collection in electronic format, makes analyzing faster and cheaper than; for instance, with traditional mailed paper-based surveys. Also, it allows respondents to choose time and place when it is suitable to answer the survey. On the other hand, the results may be threatened in case the survey link is open to

public or if the system has an error. The link to the questionnaire was open, meaning that other people could have responded to the survey in case they would have received the link. The risk for this to happen in this research was relatively small since the scope of the study was very small but also the questions were tailored to this specific study scope. Hence, the risk of receiving answers from other than study group was relatively small. In regard to technical issues, one respondent did report the survey to freeze right before sending the results. Reason for the error was not investigated further, and the respondent completed the survey again successfully.

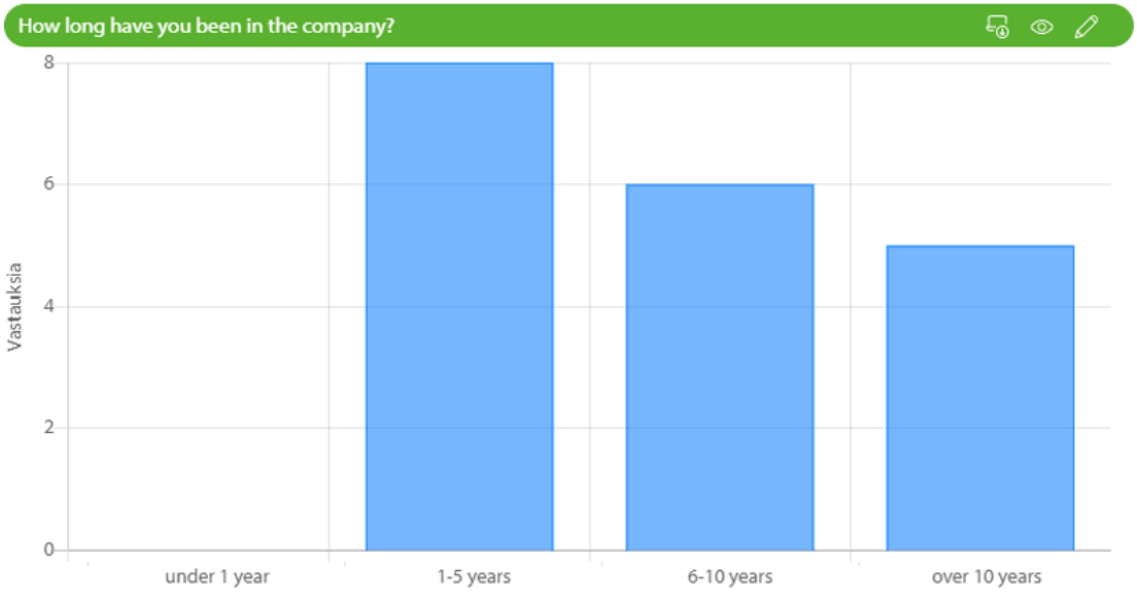
4 Data analysis — current state

In this section, the collected data through the survey will be examined. Data will be presented and analyzed in different themes. Meaning that questions that address the same topic, and/or the intention, are grouped together. Each question and the theme of the grouped questions will be presented. Graphics will be displayed, which allow the reader to visualize captured data easily. As stated, all the data examined in this section is anonymous. Collected data is based on the subjective experience of the respondents. Thus, even the quantitative results are estimations rather than numerical data derived, for example from the FM system.

4.1. Users and file usage

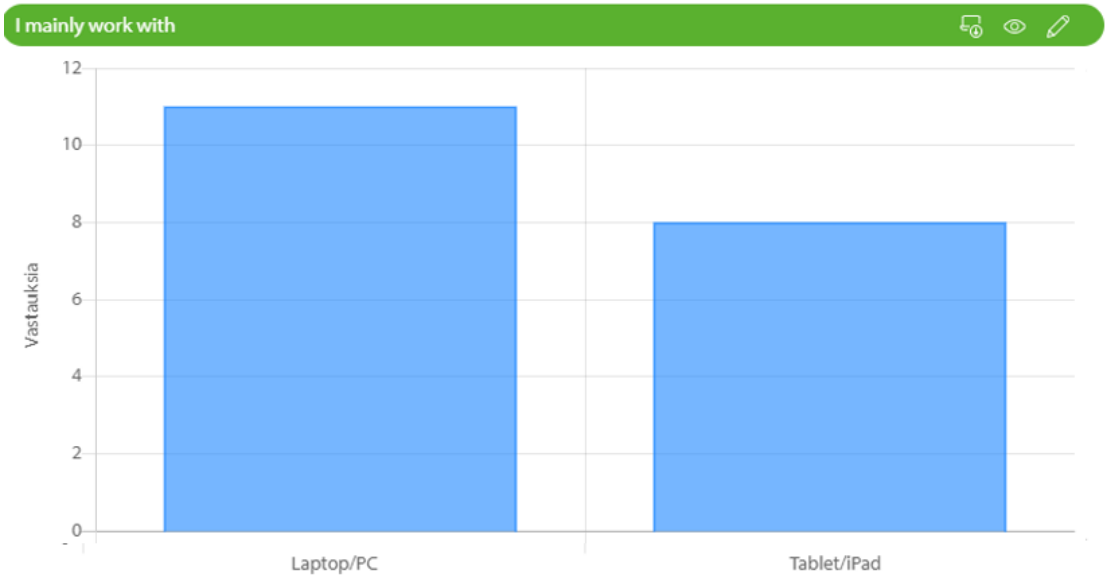
As stated, the study group consisted total of 20 respondents. The response rate was 95%, meaning that total of 19 responses were received. Since the CCX employees are roughly categorized into two groups: sales and office, the intention of the first two basic questions were to find out if there will be deviations in some of the answers depending on one's employment period and/or main device being used.

Figure 13 Employment period



As seen in Figure 13, 8 employees (42%) have been in the company for 1-5 years, 6 (32%) have been 6-10 years, and 5 (26%) have been in the company for over 10 years.

Figure 14 Main work device

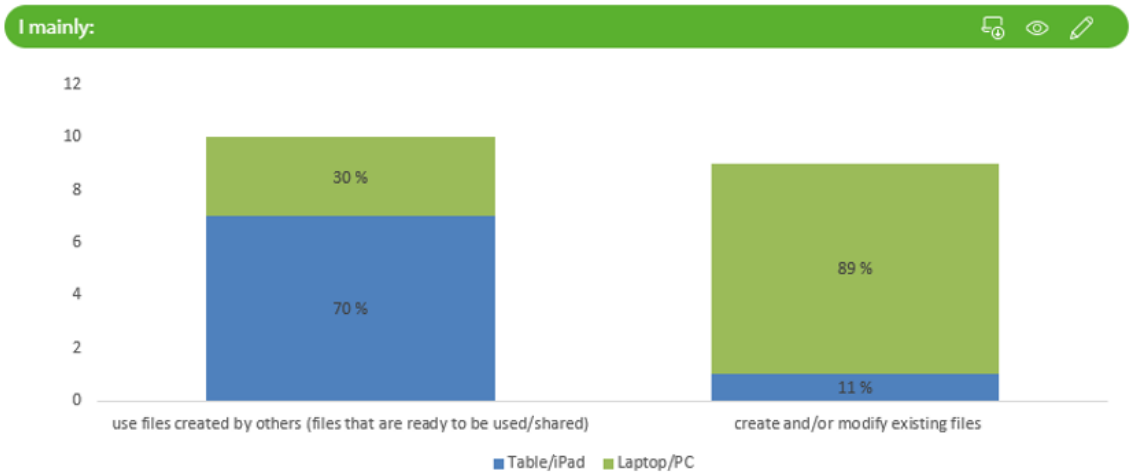


The second question shows that 8 employees (42%) reported to use tablet/iPad as a main device, whereas 11 employees (58%) reported to use laptop/PC, see Figure 14.

Figure 15 Main activity with files

Answers addressing the main activity with the files are fairly distributed. A total of 10 employees (53%) have stated to mainly use files created by others (files that are ready to be used/shared) whereas 9 employees (47%) mainly create and/or modify existing files, see Figure 15.

Figure 16 Activity with files vs. device

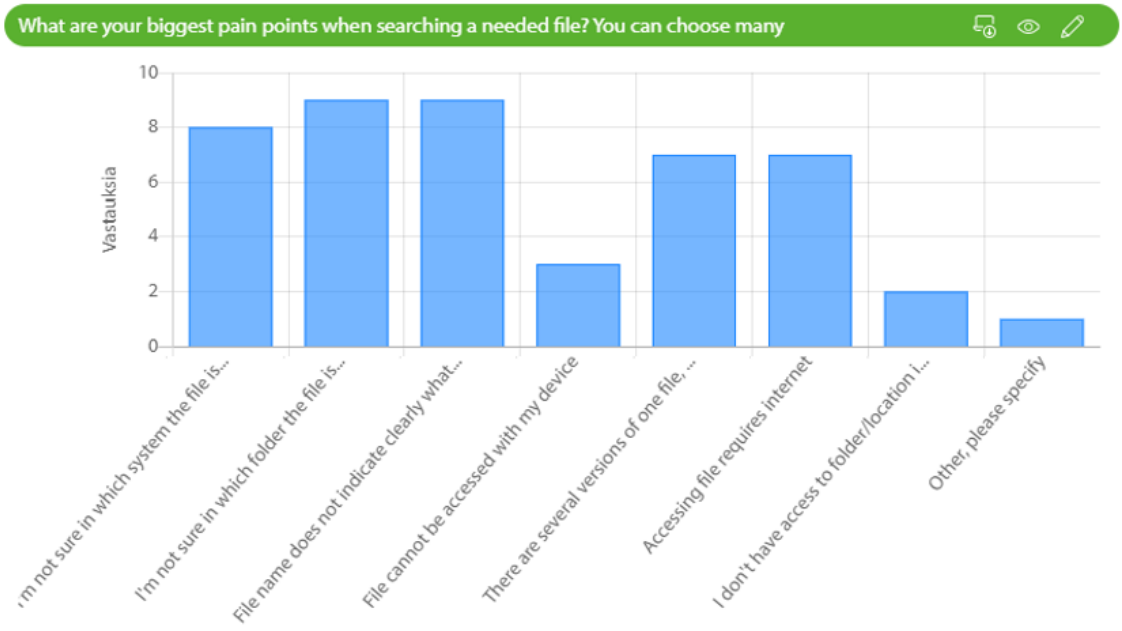


More relevant information can be discovered by comparing the device being used and the main activity. Figure 16 illustrates that 70% of total responses for the option "use files created by others" were given by tablet/iPad users, whereas laptop/PC users stated to mainly "create and/or modify existing files" with a response rate of 89%.

4.2. File search

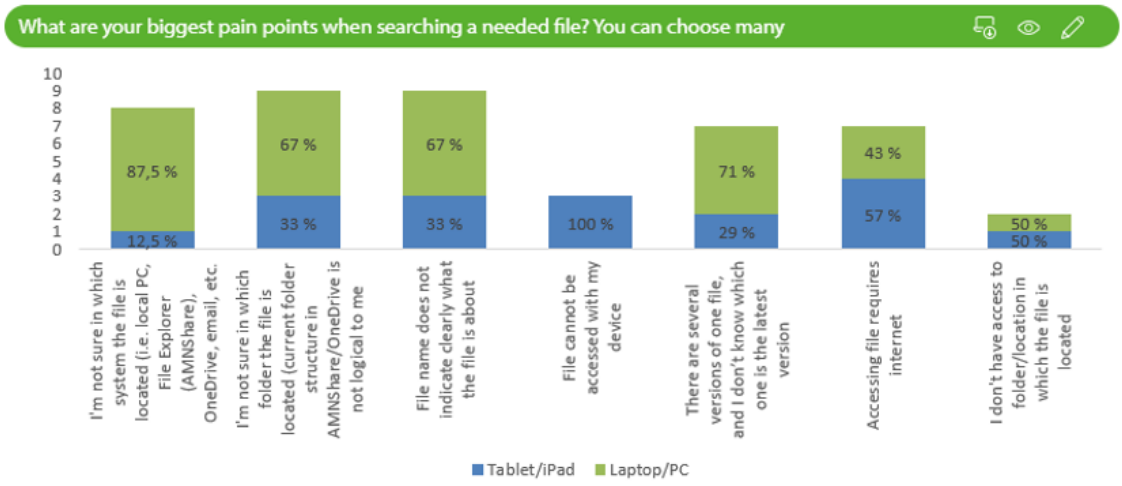
Questions and answers illustrated in this section were asked to find out factors related to a file search process.

Figure 17 Biggest challenges when searching a needed file



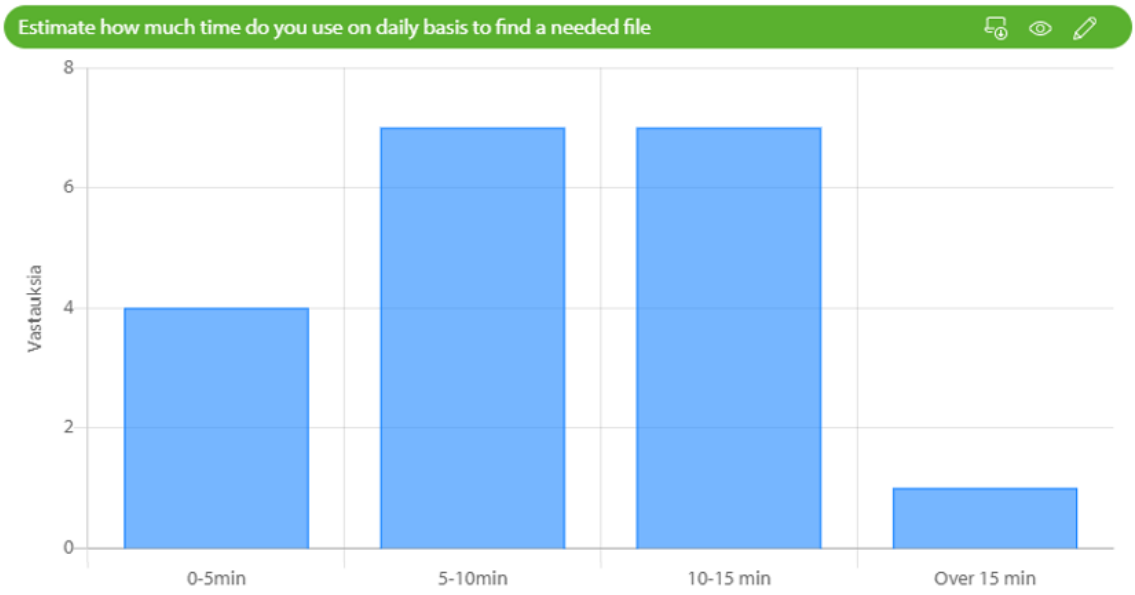
The intention of this question was to allow respondents to define the biggest pain points when it comes to searching a needed file. Respondents were allowed to choose several replies (see Figure 17). The two biggest pain points with a total of 9 responses were: "I am not sure in which folder the file is located (current folder structure in AShare/OneDrive is not logical to me)" and "File name does not indicate clearly what the file is about". The third biggest pain point with total of 8 responses was related to the system: "I am not sure in which system the file is located (*i.e.* local PC, File Explorer (AShare), OneDrive, email, etc.)". Lastly, total of 7 responses were given to the uncertainty of the file (I don't know which one is the latest version) and challenges in accessing file (Accessing file requires the Internet). There was an option where respondents could specify other pain points and consisted the following answer: "I don't use File Explorer (AShare), only OneDrive and Showell". This latter system called Showell is a sales presentation content management system that contains all the sales related materials.

Figure 18 Pain points when searching a needed file vs. device



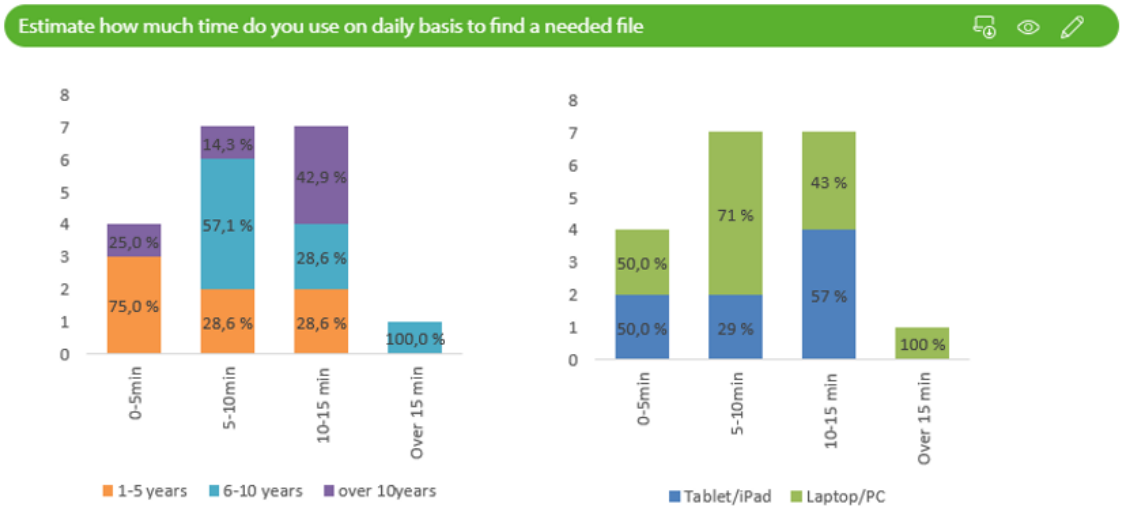
A comparison between the device being used and respondent's perception of what are the biggest pain points when searching a needed file is illustrated in Figure 18. The biggest deviations seemed to be in "I'm not sure in which system the file is located", in which 87,5% of the all answer were given by laptop/PC user. Over half of the responses, 67% related to unclear folder structure or file name were given by laptop/PC users as well. Uncertainty about having the latest file has the ratio of 71% laptop/PC versus 29% table/iPad. The majority of the respondents that identified uncertainty about the file as a pain point are actually employees that mainly create and/or modify existing files (laptop/PC users). The two biggest challenges for tablet/iPad users compared to laptop/PC users were; "Not being able to access file with the device" and "Accessing file requires the Internet", see Figure 18.

Figure 19 Time used on a daily basis to find a needed file



Estimations between daily time spend to find a needed file is distributed equally between 5-10min and 10-15min, with total of 7 responses in each (37%). A total of 4 employees stated (21%) to spend less than 5 minutes per day, and lastly, one employee (5%) estimated to spend over 15 minutes per day to find a needed file, see Figure 19.

Figure 20 Time used on a daily basis to find a needed file vs. employment period and device

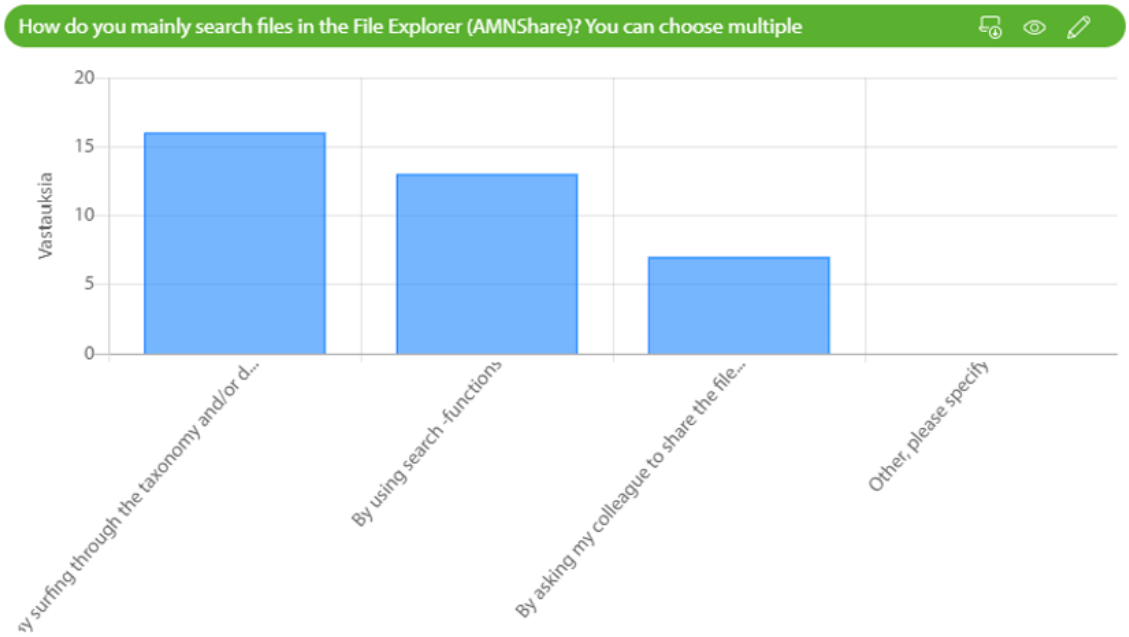


In Figure 20, the left columns illustrate the deviations in answers between employment period and the right column the device being used. The employment

period comparison indicates that the longer employee has been in the company the more time they spend on finding a needed file. For instance, total of 75% of the answers stating to spent 0-5 minutes per day to find a needed file was given by employees that have been in the company less than 6 years whereas 42,9% of total answer to spent 10-15 minutes was given by employees that have been in the company over 10 years.

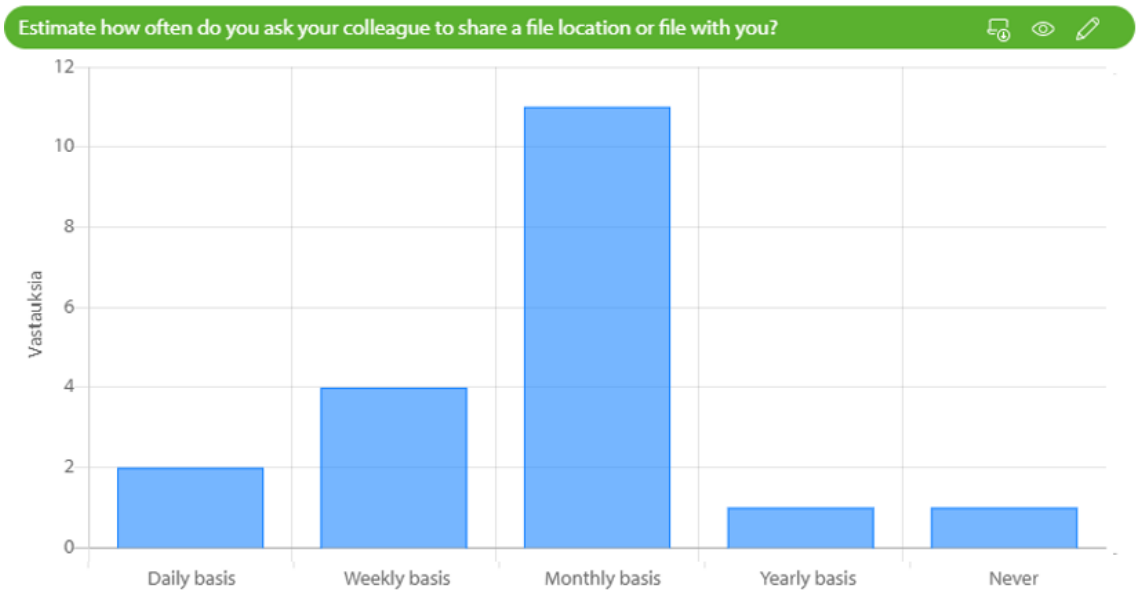
The trend between the device comparison is that 71% of the total answers stating to spent 5-10 minutes per day were given by the laptop/PC users whereas 57% of the total answers stating to spent 10-15 minutes were given the tablet/iPad users. The highest estimation of daily spend (over 15 minutes) was given by an employee who has been in the company between 6-10 years and who uses laptop/PC.

Figure 21 Main method to search a file in File Explorer



In the question illustrated in the Figure 21, the respondents were allowed to choose multiple answers. Surfing through the taxonomy and/or different folders received the most responses, total of 16, whereas the second most often used method to use search function received total of 13 responses. A total of 7 responses were also given to the option of asking a colleague to share the file/file location. This question was specific to the search method in File Explorer, thus comparison between device or employment period is not relevant.

Figure 22 Estimation of frequency to ask colleague to share a file location or a file

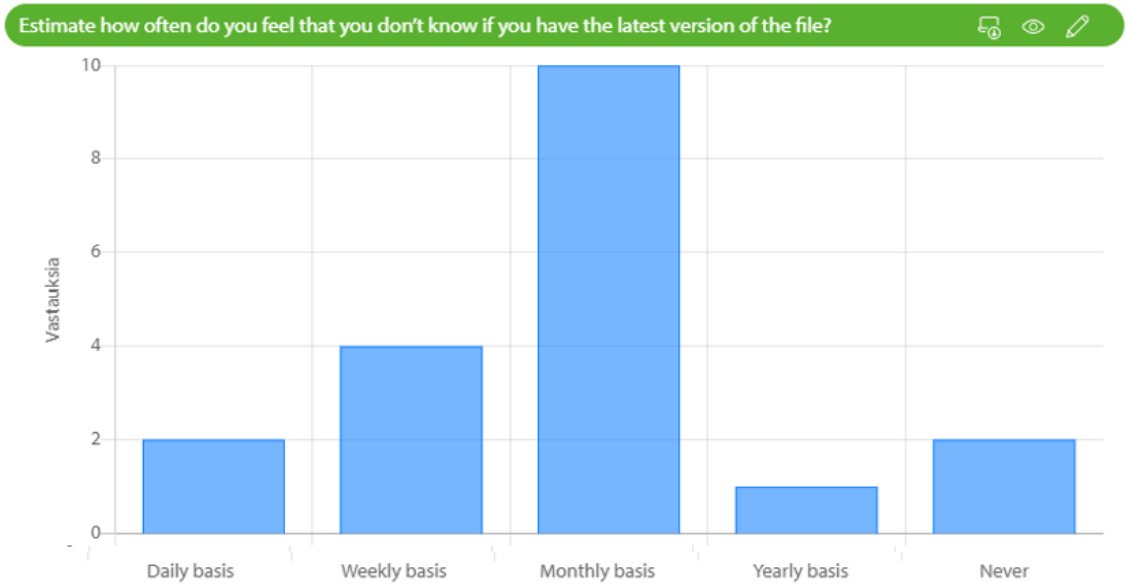


Considering the frequency to ask a colleague to share a file location or a file was asked separately. A total of 6 employees (32%) estimated to practice this on a daily or a weekly basis. The most common estimate, with a total of 11 answers (58%), was to practice this on a monthly basis. Only two employees out of the total group (10%) estimated to practice this rarely (yearly) or never, see Figure 22.

4.3. File retention

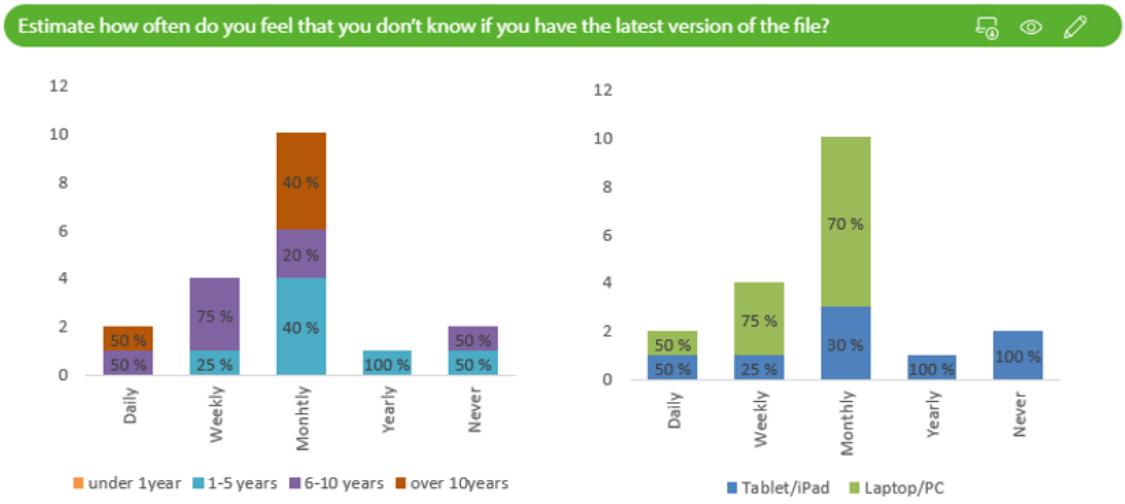
The following questions aimed to gain an understanding of actions related to storing, archiving and preserving files.

Figure 23 Uncertainty about the file



Only three employees out of the total group (15,5%) estimated to experience uncertainty about the file version rarely (yearly) or never. The most common estimate with a total of 10 answers (52,5%) was to practice this on a monthly basis, and 6 employees (32%) estimated to experience uncertainty about the file version on a daily or a weekly basis. Answers between the device being used or employment period did not bring any deviation, see Figure 23.

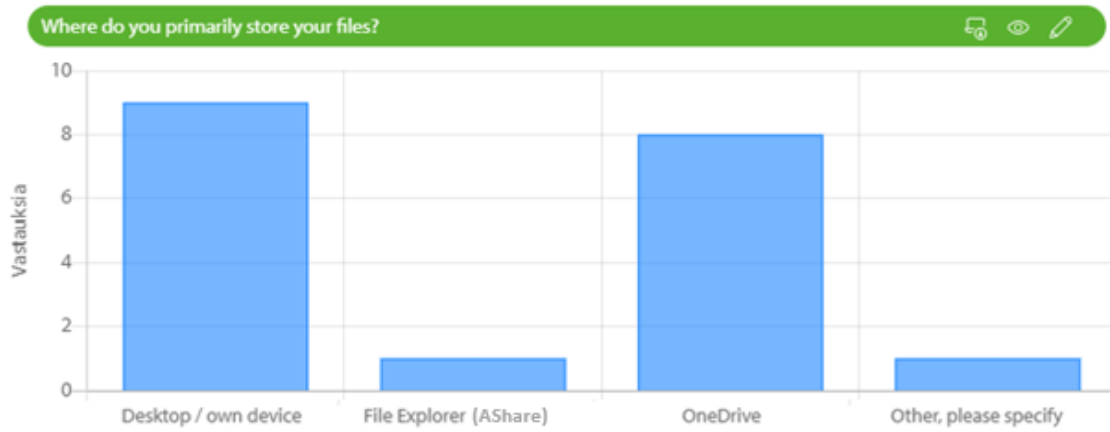
Figure 24 Uncertainty about the file vs. employment period and device



The left column in Figure 24 illustrates the deviations in the answers between employee period and the right column the device being used. The employment

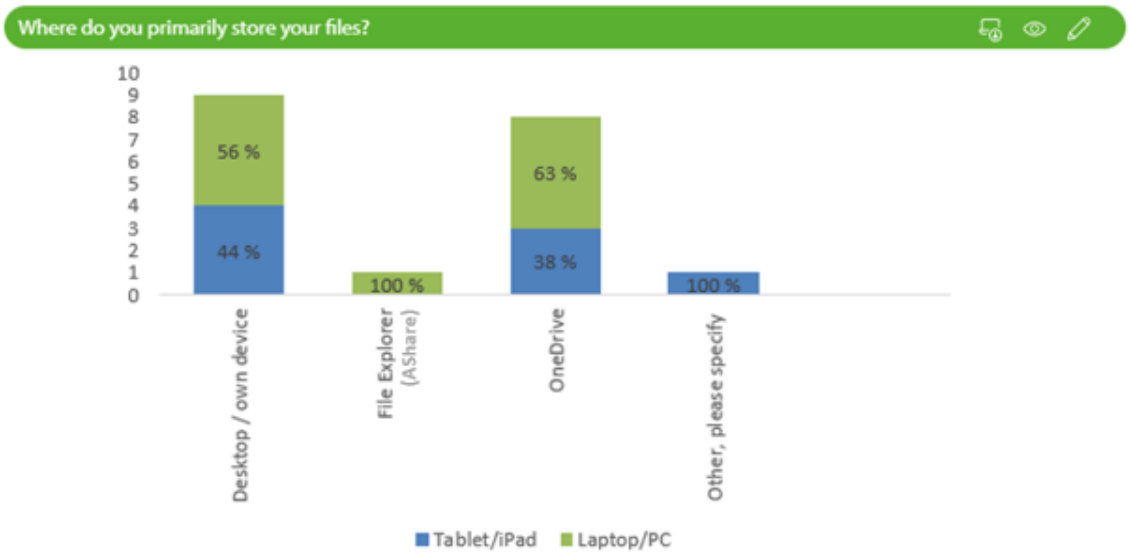
period comparison does not indicate any meaningful deviations. Over half of the answers in feeling uncertainty weekly or monthly was given by laptop/PC users. The answers of the tablet/iPad users were in general fairly equally distributed.

Figure 25 Main location to store files



The most common locations to store files are desktop/own device with a total of 9 responses (47%) or OneDrive with a total of 8 responses (42%). Only one employee stated to use File Explorer and one other, please specify option, consisted the following answer: *"email and/or notes in the iPad"* see Figure 25.

Figure 26 Main location to store files vs. device



As Figure 26 illustrates, there were no meaningful deviations between the device being used and its perception of main location to store files.

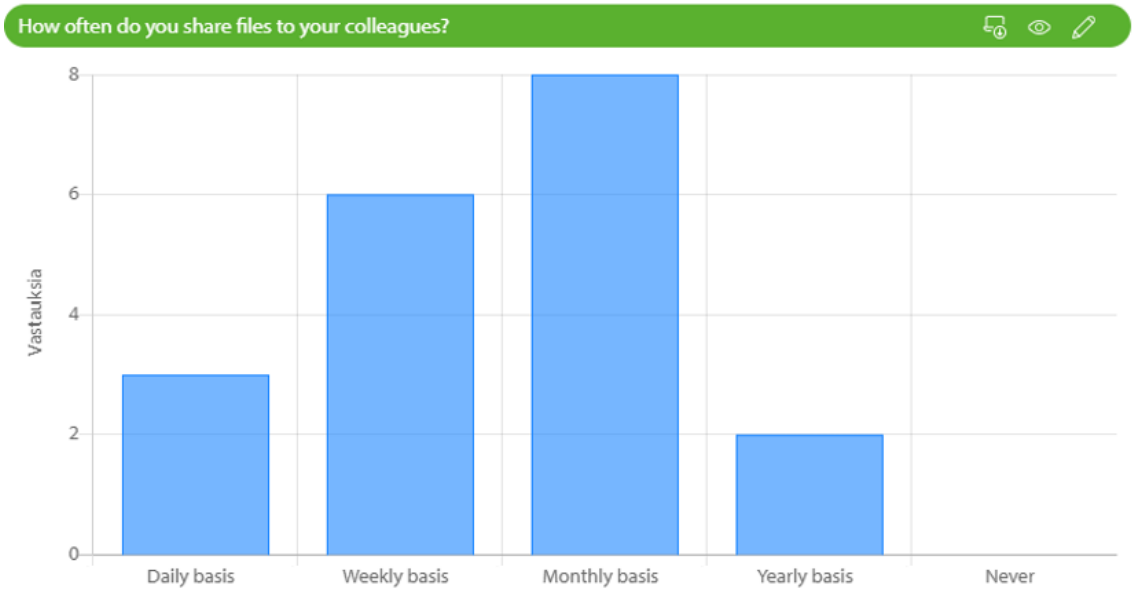
Figure 27 Estimation of frequency to delete created/used files

Based on the answers illustrated in Figure 27, deleting created/used files, is fairly distributed among the employees. Only three employees stated that they delete files on a weekly basis. The most common frequency among the respondents was to delete files monthly (total of 6 employees) or yearly (total of 7 employees). Lastly, three employees reported not to practice deleting files at all. This question was specific to file retention, thus comparison between device or employment period is not relevant.

4.4. File sharing

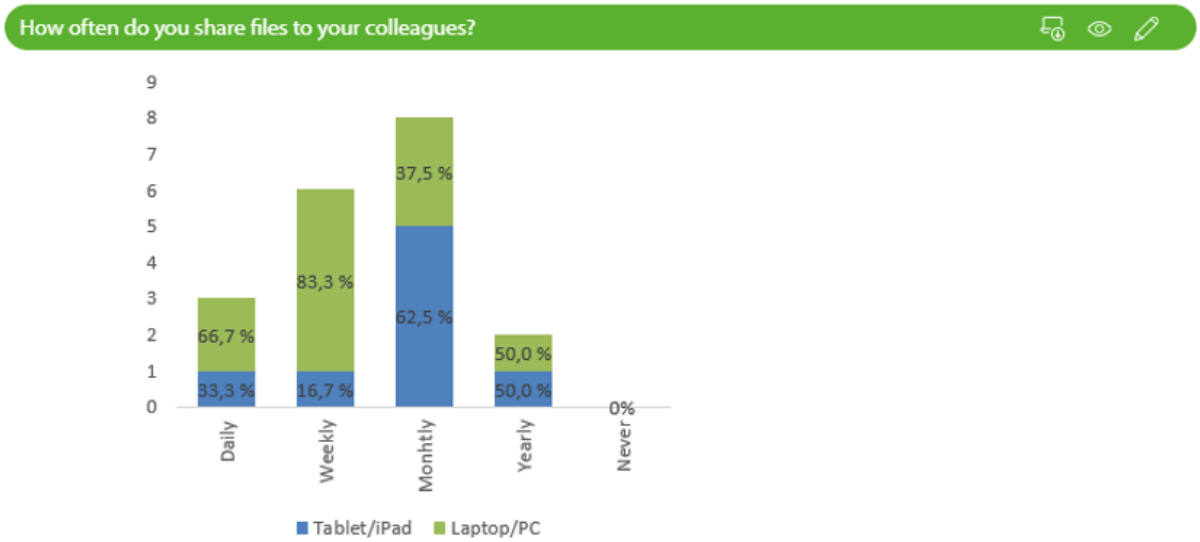
The questions in this section intended to understand how files are being shared within the CCX, why files are being shared, what are the preferred methods and how often files are being shared.

Figure 28 Estimation of frequency to share files with others



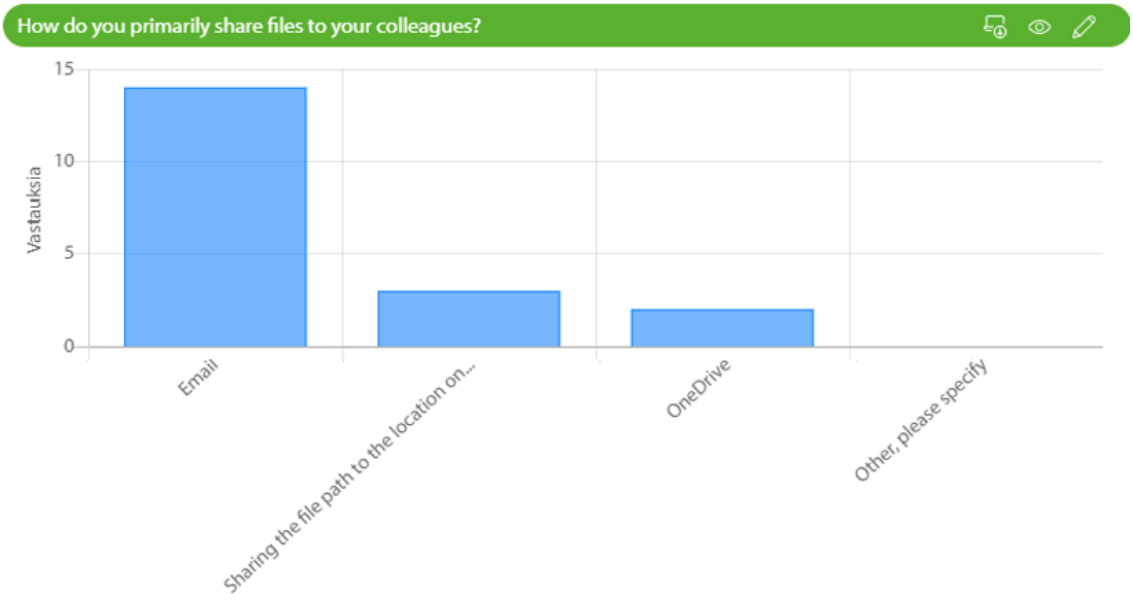
A total of 3 employees (16%) stated to practice file sharing on a daily basis. The majority of the answers were given to share files on a weekly (total of 6) or a monthly basis (total of 8), see Figure 28.

Figure 29 Estimation of frequency to share files with others vs. device



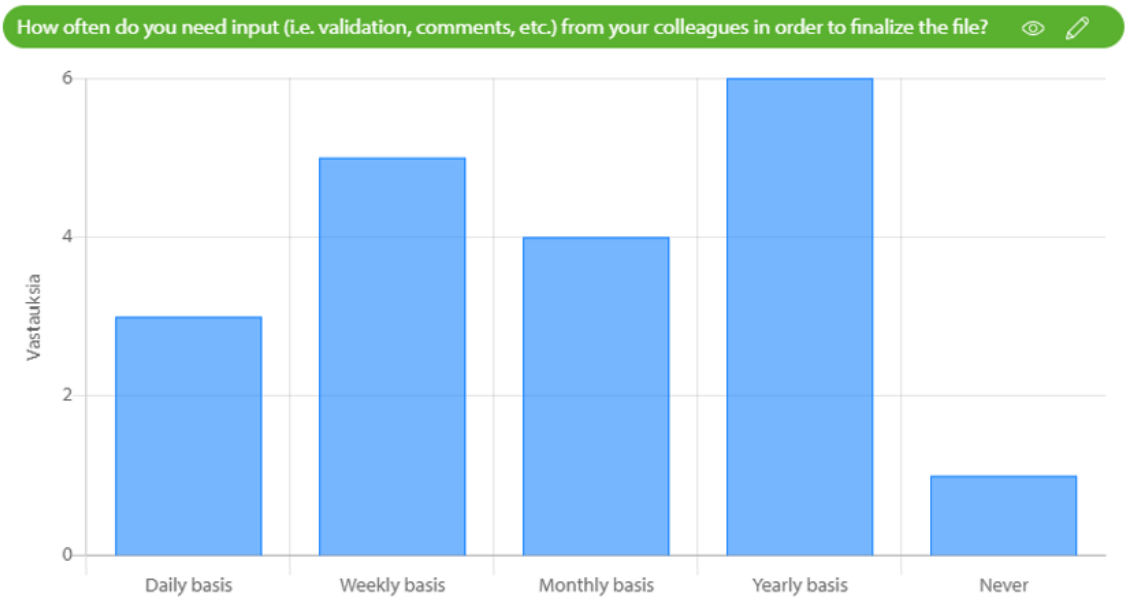
Over 80% out of all the answers 'sharing files on a weekly basis', were given by laptop/PC users. Tablet/iPad users on the other hand, had given over 60% of the answers to the option 'sharing files on a monthly basis', see Figure 29.

Figure 30 Main method to share files with others



The most popular sharing method among all employees is to share files with a colleague via email, with total of 14 responses (74%), see Figure 30. Sharing the file path to the location on File Explorer (AShare) or via OneDrive received a total of 5 responses (26%).

Figure 31 Estimation of frequency to need input from others



Only one employee stated: "I never need input from others" whereas the majority of the respondents reported that they either need input from others daily (total of 3), weekly (total of 5) or monthly (total of 4). Six employees stated they need input from

others on a yearly basis, see Figure 31. In case the respondent stated to need input from others, they received an additional question below.

Figure 32 Main method to collect the input from others

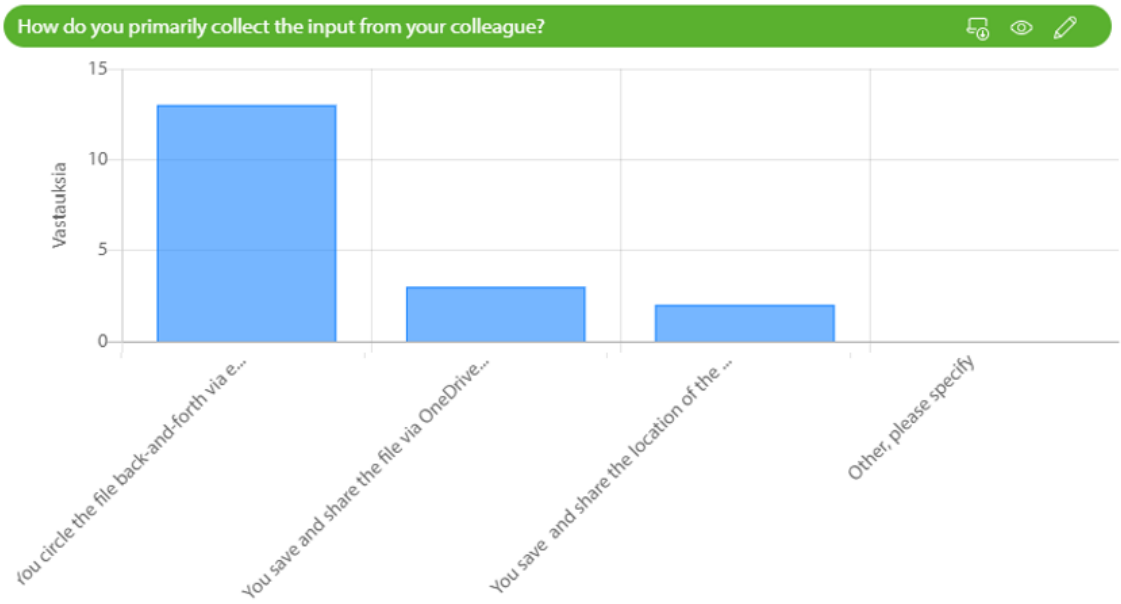
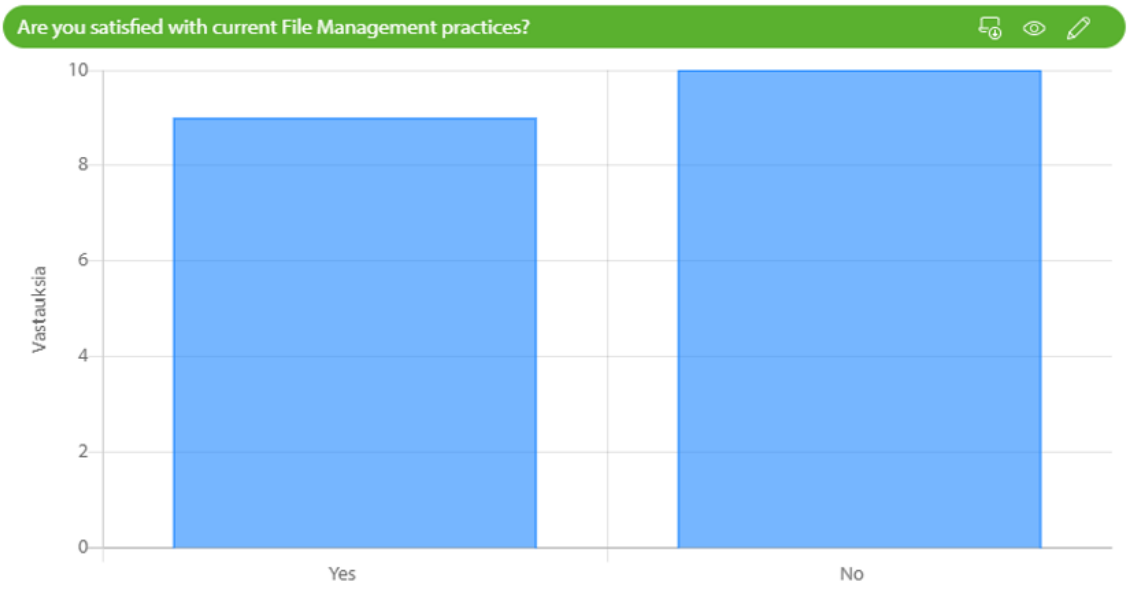


Figure 32 illustrates that that majority of the respondents, 13 out of 18, favor email as the main method to collect input from others. The rest of the respondents (total of 5) stated to use File Explorer (AShare) or OneDrive.

4.5 FM practices

The last two questions gave the possibility to the respondent to express their current satisfaction with the FM practices within the company and provide wishes or development ideas for future FM practices.

Figure 33 Satisfaction with current File Management practices



Overall satisfaction towards current FM practices was fairly distributed between 'yes' and 'no'. A total of 9 employees (47%) stated to be happy whereas 10 employees (53%) expressed dissatisfaction. In case the respondents expressed to be dissatisfied, they were asked to give a further explanation, see Figure 33.

All the 10 responses can be found in Appendix 2. There were two explicit topics that rose from the answers. Lack of common rules and guidelines *i.e.* where to store files and how files should be named was mentioned by several employees. For instance, one respondent had written: *"Clear and commonly agreed procedures are missing, there are no rules to store only the most recent files on the server, or how often old files should be cleaned-up. Clear communication to employees regarding responsibilities, as well as policies, is lacking".* Another common dissatisfaction was related to the structure, search and access which was mentioned by For instance, iPad cannot be used to access files in the File Explorer or File Explorer is disorganized and it is time consuming to find needed files.

The last open-ended question in the survey was: "What type of wishes or development ideas do you have for future File Management practices?". All responses can be found in Appendix 3. A total of 17 employees gave their feedback. Their wishes were quite aligned with the dissatisfaction factors. Meaning that many development ideas were related to establishing clear rules and guidelines, as stated

in the previous paragraph. One of the responses were as followed: *"Clear and aligned structure across the categories, discipline on only saving final files in the shared folders and clean up in old files occasionally. Clear communication to the team for WOWs and roles & responsibilities"*. The word 'clear' was visible in majority of the responses.

5 Discussion

The overall objective of the research was to investigate FM methods, best practices and technologies, and aim to give recommendations for sustainable FM practices in the CCX. The following two sections discusses further defined research questions: RQ1: How can information be classified, and what is the role of systems and users in the process?; and RQ2: How is file management linked to business practices?

The overall concluding statements for the actual data analysis are as follows. The results received from the survey indicate that the current folders and files used by the employees in the CCX are disorganized and poorly managed. There are no standardized practices to manage files – *i.e.* how they should be classified, named, or shared with relevant stakeholders. The taxonomy in the File Explorer is unclear, thus employees are struggling on daily basis to find the needed files. Furthermore, working conditions and work devices cause additional challenges. The employees working in the field, mainly use tablets/iPads whereas office workers use laptops/PCs. Accessing folders and files in the File Explorer with other device than PC is troublesome due to different OSs. Also, the Internet and/or the virtual private network (VPN) is required since the files are either located in the company's local area network (LAN) (server) or in OneDrive. Hence, accessing files is not a straightforward task, especially to those working in the field. The overall results indicate that employees do not use File Explorer as their main storage or sharing application. Other alternatives such as OneDrive, email, and personal desktop are preferred. Increasing number of various systems has led to uncertainty about in which system to search for needed information or whether the found file is the latest version, *i.e.* file can be located in multiple location. Having multiple systems to store information, together with a lack of clear rules and guidelines regarding the FM activities, is interfering with the employees' ability to perform and focus on work

task. This evident when we observed that over half of the employee expressed dissatisfaction towards the current FM practices, and future development ideas and wishes were expressed by a total of 17 employees.

5.1 Classifying information

The first research question: "How can information be classified, and what is the role of systems and users in the process" was presented in the theoretical framework. As stated, there are no single answers to the questions; who is responsible or what is the best method to classify information. Recommendations for improving existing taxonomies, thus the classification schemes, in a detailed level is not possible or even intended since this study did not investigate current classification methods and/or taxonomies in the CCX. However, based on the theoretical framework and the empirical research, the high-level recommendation for the CCX would be to invest, investigate and improve their existing FM practices related to files and folders. As Arthur (2005) suggests, the information classification process should start with the evaluation of: i) what is the purpose of the taxonomy; and ii) who are the users and what are their needs. Based on the research, File Explorer is not supported by the whole staff due to connectivity issues and different OSs. Even the laptop/PC users stated to prefer other applications. Thus, one recommendation would be looking into a cloud-based server which could be accessed with both devices (laptop and tablet) where the files and folders are available offline also. Since the CCX already has OneDrive for the employees to use for their personal files, it could be extended to business use, and this way all the files from company's local area network (server)could be centralized to OneDrive Business.

No matter which system(s) is being used, information classification is an ongoing process, not a project with an ending period. The CCX does not have a standardized protocol to manage files, *i.e*. how they should be classified, named, or shared with relevant stakeholders. In addition to Arthur's view above, Smallwood (2013) suggests, the starting point for the classification practices is to investigate how each business unit operate and interact. Managers from key business units, IT and legal and also end-users should be consulted. The end-users create and search files on

daily basis, and therefore should be involved in the process. Based on the theoretical framework and the empirical research, a functional or a hybrid approach would be ideal for the CCX. In general, functional classification is widely applied in the context of FM (Smallwood, 2013). This approach can be taken beyond classification by creating a file plan which allows companies, including the CCX, to recognize every file in the system, its location and rules. This would not only help to organize the files and folder, but also to maintain FM practices (Kipngetich, 2014). For instance, applying NTFS permissions would be easier and situations of not having access to the file could be avoided. Also, a retention schedule could be easier to organize. All in all, there is no "one size fits all" – taxonomy solution available, thus each company needs to evaluate their needs individually in order to determine best approach to classify files.

As highlighted in the theoretical framework, FM is the combination of a human and technology. Therefore, the operating parameters (*i.e.* systems and applications) needs to be considered in the classification scheme. For instance, File Explorer in Windows OS does not only provide a functionality to store files and folders but also a collection of functions that user can perform with files (Stallings, 2012). There are some limitations, for instance, for file naming convention. However, if used properly, naming convention can enable users and the system itself to recognize files, and consequently ease certain activities in the process; for instance, information retrieval. As stated, it is also important to balance the role of system and human in the process. Meaning that neither of the aspects should be over-emphasized. This thesis did not investigate the role of employees or different units in the actual classification process, but the theoretical framework suggests that the nature of responsibilities in the classification scheme is multifaceted. The main users have the best overview of the business activities (Smallwood, 2013), and therefore the overall responsibility of the FM practices should not; for instance, be appointed to IT or one employee. Based on the theoretical framework and background information about CCX, a high-level recommendation would be that each manager from key business units, is responsible for the classification practices within his/her team, and in which IT, legal or other relevant departments would be considered as support functions. For instance, IT could support technical matters whereas the

legal unit could advice on compliance and regulation aspects. Common ways of working, such as roles and responsibilities, is also important to be communicated across the functions since FM practices is an ongoing process and not a project with an ending period. All in all, in the context of FM, classification is one of the most fundamental activities when organizing files and constructing taxonomies, and in which, both the role of human and systems needs to be evaluated.

5.2 Link between FM and business practices

The second research question, "How is file management linked to business practices", aimed to investigate how frameworks in RQ1 are linked to FM practices in the business context and what are some possible implications, such as employee performance. Research results regarding the file search, illustrate that the majority of the employees (total of 14 out of 19) spend 5-15minutes on a daily basis to find a needed file. Surprisingly, employees with a longer employment period spend more time on a daily basis than those with fewer years of employment. The trend in the answers when comparing devices being used, was that half of the tablet/iPad users spend 10-15 minutes on a day to find a needed file. This could be related to the results found in other sections in which tablet/iPad users mainly use files created by others, and which according to results are not being stored to File Explorer. Thus, the results in the device comparison could be an indication that table/iPad users are struggling to find the needed files from their own devices and/or OneDrive which were stated to be the main locations to store files. Similar results were presented in M-Files' (2019) study in the theoretical part, in which nearly half of the employees reported experiencing challenges in finding the needed information, and in cases where the query is unsuccessful, 83% reported creating a new file. A direct link to duplications and issues in finding the document cannot be done. However, the risk should be minimized since file duplications have several implications. For instance, i) they increase the search time while reducing the accuracy of the search (Smallwood, 2013) ii) they increase the need of storage space, hence influence system operation cost (Moran, 2015) and iii) they can complicate information security management *i.e*. retention and disposition scheduling (Smallwood, 2013).

Over half of the employees in the CCX reported experiencing uncertainty about the file on a monthly basis and over 30% reported experiencing this even on a daily or a weekly basis. At the same time, only three employees stated to delete created/used files on a weekly basis whereas the majority of the employees reported to practice this mainly on a yearly basis or never. The research results about the uncertainty of a file and its accuracy are not directly related to duplications or number of files. Despite the information governance was not in the scope of the study, these findings should support the company's motive to invest in FM initiatives which would allow CCX's employees to have trusted and reliable information available and consequently allow them to make decisions with more confidence while also complying with applicable laws and regulations. All in all, as previously mentioned factors support the hypothesis that all the efforts used to storing information becomes wasted time if information is not retrievable (Lindén, 2015; M-Files, 2019).

Furthermore, over the half of the employees expressed dissatisfaction toward the FM practices in the CCX which is in line with Lindén's (2015) suggestion between the relationship of job satisfaction, productivity and ability to find relevant information quickly. As the theoretical framework discusses, one of the benefits to classify information and/or build taxonomies is the possibility to access and retrieve information more quickly and more conveniently (Brooks, 2017). Furthermore, classifying information, *i.e*. file taxonomies, should help end-users to navigate and find information in a logical and familiar way, even if they are not sure what they are looking for (Reinout, 2008; Smallwood, 2013). Although the fluency and sensibleness of work for knowledge-workers is admirable, Lindén (2015) argues that quantifiable business performance is usually the main driver for FM initiatives. Pre-defined time frames in the survey options (0-5 min, 5-10min, etc.) are too vague to make accurate calculations of how much time in average one person spend on a daily basis to find a needed file. Also, the numerical data received from the survey is based on the employees' own subjective estimation rather than the numerical data derived; for example, from the FM system. However, if looking at the average time in which the majority of the answers were equally distributed between 5-10min and 10-15min, it could be estimated that one person spends 10 minutes daily on average to find a needed file. If calculating further, it means 50 minutes per week, and 3

hours 20minutes per month. The actual cost of spent time in euros depends on different factors, but the point is that over 3 hours on a monthly basis is being spent looking for a file rather than actual work tasks.

To illustrate business implications further, time spend to fulfill a colleague's request to share a file location or a file could be evaluated. In fact, the majority of the employees (total of 11) stated to ask a colleague to share a file location or a file with them on a monthly basis. This question does not give the respondent the possibility to state the motive for this action. Hence, the request for support may not be related to investigate pain points in the search. However, the results could fit into Lindén's (2015) theory: *"in case the query does not bring results, employees may interrupt a colleague by asking for help, resulting in two employees being less productive".* Based on the survey results, information in CCX information systems is not only disorganized but there is also paradox regarding FM applications. The most common location to store files in the CCX was the desktop or OneDrive. Moran (2015) highlights several issues of using own devices as a storage, such as i) other employees or the company cannot access these files, thus the overall FM cannot be controlled centrally; ii) adding more disc space is challenging since computer's are generally purchased with a certain disc space capacity; and lastly iii) information is not being secured properly, *i.e*. in case the computer gets lost/stolen or the hard drive gets damaged, restoring the files is troublesome. The question "where do you primarily store your files" did not give the respondents the possibility to share further details (*i.e.* motives) about the chosen storage location. It also leaves out the possibility of using multiple locations in parallel. However, the research results indicate that File Explorer is not generally preferred as a storage location. In fact, File Explorer as a FM application was not found appealing since access requires the Internet and/or to use laptop/PC. These results should be taken into account when considering alternative FMS in the CCX.

Despite there is no "quick or one size fits all" – taxonomy solution available, both the research results and the theoretical framework support the fact that investing in well-designed classification systems is worth it and can result in greater

productivity and process efficiency. Thus, impacting companies' overall ability to reach its business objectives.

5.3 Future research

The FM is a quite largely researched topic which does not make this study novel. Perspectives in the theoretical framework was chosen to meet the overall objective and to give further recommendations for the FM practices in the CCX. A high-level recommendation was possible to be provide. However, in order to execute micro level changes or reach concrete achievements, recommendation would be to investigate theoretical frameworks related to taxonomy project planning and using more sophisticated methodological approach using, e.g. semi-structured interviews. Future research could give indications to different actions that companies, including CCX, should undertake in order to invest and improve existing FM practices related files and folders. Ideally, the taxonomy project planning would investigate different phases in the process, such as: how to conduct an actual file management inventory and how business process analysis is linked with information classification. This could also give insights of how to maintain FM practices in the long run. Recommendations received from this type of study would again be a company specific since there is no one-size-fits all solution available. Recommendations in the theoretical level could be more generalized whereas the actual solution would be more business specific. In case the CCX changes FMS or executes actions (*i.e.* taxonomy project planning), with the aim of improving existing taxonomies and/or classification schemes, it would be interesting to repeat the conducted survey. Comparing results after and before a taxonomy project would provide interesting insights, especially from the return of investment point of view.

The study group was relatively small which does not necessarily infer with the creditability or reliability of the study. However, since the CCX is a part of the Nordic cluster, part of multinational corporation, the study group could be expanded to other CBUs. Expanding the research to other Nordic CBUs would be a good starting point since some processes and services are already shared across the Nordic CBUs. Expanding the study to other Nordic countries could provide wider understanding of the user needs within the Nordic cluster. Also, one could argue that it is unlikely

that CCX would; for instance, change its FMS locally. However, if a broader need could be demonstrated, the need for a change may be more likely.

Lastly, it would also be interesting to investigate the link between the employees' performance and FM further. There is a large amount of studies around this area, but it would be interesting to collect data and compare different studies. Especially focusing on the studies done within the past few years since operating environments (i.e. ways of working and systems) have changed greatly in the past decade. In addition, the focus could be on industry or the company size related. Narrowing down the research area could provide more in-depth understanding of the implications of the FM in the chosen business context. Furthermore, as presented in this thesis too, one could argue that the motives to invest FM is to improve the overall business performance. Having business specific data that can be illustrated and calculated in numbers could be beneficial. Information such as, how much money is spent to maintain different FMS (i.e. yearly fees), or how much time employees use on daily basis to find a needed file in which the actual data would be derived from the FMS could improve the reliability and validity of the research.

6 References

Adams, J., Khan, H. T. A., & Raeside, R. (2014). Research methods for business and social science students. SAGE Publications, New Delhi.

Akhavan, P. and Pezeshkan, A. (2014), "Knowledge management critical failure factors: a multi-case study", VINE, Vol. 44 No. 1, pp. 22-41.

Antin, K. (2016). File naming conventions: why you want them and how to create them. [online] Available: https://www.huridocs.org/2016/07/file-naming-conventions-why-you-want-them-and-how-to-create-them/ [Accessed 25 May 2019].

Arthur, M. (2005) "Expanding the system definitions and configurations (taxonomy and data structure)". Journal of Digital Asset Management Vol 1. No 4, pp. 279-297.

Aujirapongpan, S., Vadhanasindhu, P., Chandrachai, A. and Cooparat, P. (2010), "Indicators of knowledge management capability for KM effectiveness". VINE, Vol. 40 No. 2, pp. 183-203.

Batley, S. (2005). Classification in theory and practice. Oxford, England; Rollinsford, New Hampshire: Chandos Publishing.

Bhat, A. Descriptive research: definition, characteristics, methods, examples and advantages. [online] Available at: https://www.questionpro.com/blog/descriptive-research/ [Accessed 21 May 2020].

Brooks, C. (2017). Introduction to Classification, Taxonomy and File Plans – Foundational Elements of Information Governance: Part 2. [online] Available at:https://imergeconsulting.com/introduction-classification-taxonomy-file-plans-foundational-elements-information-governance-part-2/. [Accessed 13 May 2019].

Cadence Group (2006). Taxonomies: "The Backbone of Enterprise Content Management". Journal of ARMA international, Vol. 20, No 3.

Christensson, P. (2016). Operating System Definition. [online] Available at: https://techterms.com. [Accessed 8 April 2019].

Christensson, P. (2008). NTFS Definition. [online] Available at: https://techterms.com. [Accessed 8 April 2019].

Computer Hope (2019) File extension. [online] Available at: https://www.computerhope.com/jargon/f/fileext.htm. [Accessed 21 May 2019].

IFLA, The International Federation of Library Associations and Institutions. (2017). Subject classification schemes [online]. Available at:

https://www.ifla.org/best-practice-for-national-bibliographic-agencies-in-a-digital-age/node/9042. [Accessed 21.4.2019]

International Standard Organization, ISO 62542 (2016). Information and documentation - Records management. Part 1: Concepts and principles.

Ismail M. N. and Ismail M. T. (2009). "Analyzing of Virtual Private Network over Open Source Application and Hardware Device Performance". European Journal of Scientific Research, Vol. 28, No. 2, pp. 215-226.

Kipngetich, K. (2014). Examine various filing and classification systems that are operationalized in institutions. Research paper [online] Available at: https://www.academia.edu/13735672/EXAMINE_VARIOUS_FILING_AND_CLASSIFICATION_SYSTEMS_THAT_ARE_OPERATIONALIZED_IN_INSTITUTIONS. [Accessed 19 May 2019].

Lindén J-P. (2015). Tiedonhallinta & Yrityksen menestys, 2nd edition. Netera Consulting. Tampere.

M-Files (2019). The Intelligent Information Management Benchmark Report [online] Available at: https://www.m-files.com/fi/whitepaper-2019-intelligent-information-management-benchmark. [Accessed 19 May 2019].

Microsoft (2019). [online]. Available at: https://www.microsoft.com [Accessed 21 May 2019].

Mokhtar, U. A. (2017). Records Classification: Concepts, Principles and Methods. Chandos Publishing.

Moran J. (2015) Keeping Files and Folders Organized. In: File Management Made Simple, Windows Edition. Apress. Berkeley, CA.
Morville, P. and Rosenfeld, L. (2007). Information architecture for the World Wide Web (3. ed.). Sebastopol, CA: O'Reilly.

Moskowitz, J. (2010). Group Policy: Fundamentals, Security, and the Managed Desktop, Wiley.

Parker-Wood, A. (2014). Improving File Management Through Provenance And Rich Metadata.

Patton, M. Q. (2015). Qualitative research & evaluation methods: Integrating theory and practice (Fourth edition.). Thousand Oaks, CA: SAGE.

Ravitch, D. (2007). EdSpeak: A Glossary of Education Terms, Phrases, Buzzwords, and Jargon.

Reference. (2019). What Is the Definition of "file Management". [online] Available at: https://www.reference.com/technology/definition-file-management-b964eb8adb7dfa5a [Accessed 13 May 2019].

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Availabe at: http://data.europa.eu/eli/reg/2016/679/2016-05-04 [Accessed 10 May 2019]

Reinout, R. (2008). "Clarity in the usage of the terms ontology, taxonomy and classification". Civil Eng. 20.

Saldana, J (2011). Fundamentals of Qualitative Research, Oxford University Press, Incorporated, Cary.

Seo, P, La Paz, A. and Miranda, J. (2010). #Role of Information Systems in Organizational Agility: The case of eClasSkeduler at Universidad de Chile". 16th Americas Conference on Information Systems 2010, AMCIS 2010. 6. 578.

Silberschat A., Gagne G, and Galvin P. (2018). Operating System Concepts,10th Edition. John Wiley & Sons.

Smallwood, R. F. (2013). Managing electronic records: Methods, best practices, and technologies. Hoboken, New Jersey: John Wiley & Sons, Inc.

Sobh T. and Aly Y. (2011). Effective and Extensive Virtual Private Network. Journal of Information Security (JIS), Vol.2 No.1, pp: 39-49.

Stallings, W. (2012). Operating Systems: Internals and design principles (7th ed., int.ed.). Upper Saddle River, NJ: Prentice Hall.

StatCounter GlobalStats. (2017). Desktop Operating System Market Share Worldwide [online] Available at: http://gs.statcounter.com/os-market-share/desktop/worldwide/#monthly-201711-201711-bar [Accessed 3.3.2019].

Survey Monkey (n.d). [online]. Available at: https://www.surveymonkey.com [Accessed March 1st 2020].

The Library of Congress. (2015). Classification and Shelflisting [online]. Available at: http://www.loc.gov/aba/cataloging/classification/ [Accessed 21.4.2019].

Verstraete, C. (2004). "Planning for the unexpected [business agility]". Manufacturing Engineer. Vol. 83, Issue: 3, p:18-21.

Wang, Z, Chaudhry, A., and Khoo, C. (2008) "Using classification schemes and thesauri to build an organizational taxonomy for organizing content and aiding navigation", Journal of Documentation, Vol. 64 Issue: 6, pp: 842-876.

Wyatt, J. (2000). When to Use Web-based Surveys, Journal of the American Medical Informatics Association [online]. Available at:

https://academic.oup.com/jamia/article-abstract/7/4/426/714361 [Accessed 21.3.2020].

7 Appendices

APPENDIX 1 – Survey questionnaire

*Compulsory

1. **I mainly work with: ***
   - Laptop/PC
   - Tablet/iPad

2. **How long have you been in the company? ***
   - under 1 years
   - 1-5 years
   - 6-10 years
   - over 10 years

3. **I mainly: ***
   - create and/or modify existing files
   - use files created by others (files that are ready to be used/shared)

4. **What are your biggest pain points when searching a needed file? You can choose many ***
   - I'm not sure in which system the file is located (i.e. local PC, File Explorer (AShare), OneDrive, email, etc.)
   - I'm not sure in which folder the file is located (current folder structure in AShare/OneDrive is not logical to me)
   - File name does not indicate clearly what the file is about
   - File cannot be accessed with my device
   - There are several versions of one file, and I don't know which one is the latest version
   - Accessing file requires the Internet
   - I don't have access to folder/location in which the file is located
   - Other, please specify

5. **Estimate how much time do you use on a daily basis to find a needed file ***
   - 0-5min
   - 5-10min
   - 10-15 min
   - Over 15 min

6. **How do you mainly search files in the File Explorer (AShare)? You can choose multiple***
   - By surfing through the taxonomy and/or different folders
   - By using search -functions
   - By asking my colleague to share the file location/file with me
   - Other, please specify

7. **Estimate how often do you ask your colleague to share a file location or file with you? ***
   - Daily basis
   - Weekly basis
   - Monthly basis
   - Yearly basis
   - Never

8. **Estimate how often do you feel that you don't know if you have the latest version of the file? ***
   - Daily basis
   - Weekly basis
   - Monthly basis
   - Yearly basis
   - Never

9. **Where do you primarily store your files? ***
   - Desktop / own device
   - File Explorer (AShare)
   - OneDrive
   - Other, please specify

10. **How often do you delete files that you have created/used***
    - Daily basis
    - Weekly basis
    - Monthly basis
    - Yearly basis
    - Never

11. **How often do you share files to your colleagues? ***
    - Daily basis
    - Weekly basis
    - Monthly basis
    - Yearly basis
    - Never

12. **How do you primarily share files to your colleagues? ***
    - Email
    - Sharing the file path to the location on File Explorer (AShare)
    - OneDrive
    - Other, please specify

13. **How often do you need input (i.e. validation, comments, etc.) from your colleagues in order to finalize the file? ***
    - Daily basis
    - Weekly basis
    - Monthly basis
    - Yearly basis
    - Never

14. **How do you primarily collect the input from your colleague? (only visible in case respondent needs input from colleagues in question 13)**
    - You circle the file back-and-forth via email
    - You save and share the file via OneDrive, and ask to work with the file in there
    - You save and share the location of the file in the File Explorer (AShare), and ask to work with file in there
    - Other, please specify

15. **Are you satisfied with current File Management practices? ***
    - Yes
    - No (please specify why not*)

16. **What type of wishes or development ideas do you have for future File Management practices?**
    - Free text

APPENDIX 2 – Survey question 15 – additional responses

Answers given to question 15 "Are you satisfied with current File Management practices?" by respondents that had chosen 'No'.

1) All the files and folders are disorganized
2) Clear and commonly agreed procedures are missing, There are no rules to store only the most recent files on the server, or how often old files should be cleaned-up. Clear communication to employees regarding each responsibilities, as well as policies, is lacking
3) Ashare is disorganized
4) Server is too disorganized and time consuming to use
5) Lack of systematic, unclear structure, lack of agreed naming conventions with files and folders. There has been a shift from personal folders to shared folders, but no clear rules have been established
6) The files are scattered on OneDrive and File Explorer. There are also several locations in File Explorer. Users have final file versions in their personal drive.
7) Finding files quickly is too difficult.
8) Many people cannot use files at the same time if using files in the server
9) Clear and commonly agreed procedures are missing. Meaning everyone seems to work with files and folders as they see fit.
10) You cannot access File Explorer with iPad.

APPENDIX 3 – Open-ended answers

Answers given to question 16: “What type of wishes or development ideas do you have for future File Management practices?”

1) Clarify the ways of working
2) Clear and aligned structure across the categories, discipline on only saving final files in the shared folders and clean up in old files once in a while. Clear communication to the team for WOWs and roles & responsibilities
3) It would be nice to have a quick guide of how to store and share files, it is frustrating that everything needs to be learned with “test and learn” approach.
4) Clear naming convention, logic and agreed ways of working.
5) Make clear guidelines:
    a. * where and how files should be saved
    b. * where and how the files should be named
    c. * how files should be shared
    d. * where old files should be archived if you do not want to delete them
    e. * who is responsible for managing the files and in which folder "
6) Final files are always saved File Explorer AShare (including InDesign files)
7) I am pleased with Showell :)
8) Since I mainly use iPad in my work, I would like access server with the iPad
9) Although the structure of AShare folders has been made clearer in recent years, it could certainly still be clarified even further. But I think a lot has happened over the last two years. I could also want tips to clarify my own folders.
10) Clear instructions on where everything related to a specific topic is stored.
11) Clear instructions on where to find anything. Right information would ease the use. The similarity in storing files i.e. storing files related to function. This way everyone would know where to search needed file.
12) In a perfect world, you would also access AShare from your iPad. Thank you so much for cleaning up the AShare. I perhaps has the most recent experience of the former mess in tended files but now it is really neat and logical.
13) I find server/AShare really difficult and therefore I don’t use it. OneDrive and Showell are great places and email works well if the file is not too big. I also share information / files through discussion groups (iPhone). I think company works pretty well for informing about what is available and where to find it. I also use my desktop since it does not require Internet connection.
14) Sometimes we could use more traditional email to share information
15) Clear rules of the ways of working, which storage location to use, how files should be named, shared, etc.
16) There should be clear instructions on which storage location to use. And AShare Drive should be accessible with the iPad.
17) OneDrive is good when you get there, but otherwise it is a little weird. File format is not the same as in PC, especially when people go to view / edit files there. AShare would be good, but you cannot access it with iPad.