Andrew Tanuwijaya, 2602158216

Literature Review of File Naming Conventions, Habits, and Systems

Introduction

Naming files is something that many people have to do. Students studying in not only technology fields, but any field that has to do with storing information will likely have its students manage files. For example, business students will have to manage organizational files of consumer data, product information, etc. Computer science students manage and store hundreds of files for their programs. Even if someone is not studying, this is still very applicable to a work environment. Most businesses have a certain workflow where data is stored, and employees access it constantly almost every day. Thus, naming files is an important aspect of data storage in general.

Yet, despite that, there are tons upon tons of people who are either too lazy, or not skilled enough to name their files properly. Students simply downloading files from their school's learning management system, or employees downloading those files for a one off and being done with it. Eitherway, this will cause mismanagement in the long run, due to the large amount of files that will eventually accumulate.

The purpose of this review is to investigate why it is important to properly name one's files, the different types of file naming convention that exists, and what more can be done for this seemingly minor issue that causes long-run problems. Being able to gain a higher depth of knowledge for this specific topic will allow a better understanding into the reasoning behind people's minds for their file names, and give an insight on what, possibly, is the most efficient file naming system. So, the main question in this review is a simple one: "What file naming conventions exist, and which exactly would be the most efficient by impacting the productivity of the user?"

Body

This specific literature review would aim to try and gauge six different literary papers that cover file naming conventions and file system management, especially focusing on the section where they all talk about how their users name their files exactly. The choice behind these papers is to find a diverse group of papers that, although covering different main topics, relate closely enough to the main topic at hand for this literary review specifically.

1. A Guide to Developing a File Naming Convention [1].
2. A File System for Information Management, or specifically, a file system that explicitly supports information management [2].
3. The TILDE File Naming Scheme, which is a popular naming method among programmers [3].
4. File Naming in Humanities and Social Sciences [4].
5. File Naming Among Personal Computer Users in 1999 [5].
6. Predicting the Contents of a File through its Name [6].

As to simplify the paper, it will be divided into three primary themes. This will help break up the paper into more digestible sections:

1. The Importance of Having a File Naming Convention [1, 2].
2. The Naming of Files on the Profession Spectrum [1, 3, 4]
3. Innovating File Naming Practices [].

The Importance of Having a File Naming Convention

The first question to answer would be why someone would want to have a file naming convention. Well, simply put, having this convention would definitely help retrieve those files much more easily. It will reduce the amount of time you would have to search for it. Once you are clear, concise, and consistent when it comes to naming files, it will minimize the time it takes to search for these files [1].

The Nebula file system is a different perspective at file structuring. They use a structural approach rather than the traditional navigation method. It is much less conventional as it is super uncommon by using descriptive names for the files. However, with this system, the files are found way faster and more efficiently [2].

The Naming of Files on the Profession Spectrum

Everyone has different file naming conventions. Usually, those who label their files would have it named based on their needs and priorities. For example, lots of programmers use the TILDE method, where the files are organized in a more hierarchical structure known as the "forest of tree-directory tree". This more or less allows the user to keep their personal view of the directory tree, which reduces any conflicts of names and cross-system portability [3]. This way, files will remain organized no matter what system the programmer uses.

On a different side of the spectrum, those studying humanities and social sciences use descriptive filenames with the date, subject, and project identifiers. This way, files are still recognizable when they are away from their original directory [4].

This just goes to show how differently people name their files in different fields. There are lots of different priorities when naming files like how portable it is, the descriptiveness of the metadata, and so on. Each field’s file naming are often defined by their workflow and the type of data they manage.

Innovation in File Naming Practices

The amazing thing is that techniques to name files and manage systems have been changing alongside the growing amount of data as scalability, retrieval, and organization becomes more and more necessary. As mentioned before, the Nebula file system is still one of the more innovative practices [2]. The more universal and consistent method of naming,

however, often includes consistent date formats and version numbers. These are the kinds of simple factors that usually help maintain file organization and decrease confusion over file versions.

Research was even done on people who did name their files, and it shows that their naming conventions are usually reflected by their personalization and organization needs, though differing throughout a large span of individuals [5]. There were so many different methods, from basic abbreviations to extremely descriptive names. This goes to show how important it is when designing a system that caters to both the simple and complex name preferences.

A different study also showed that even the most labels people give for their files are often predictive of their properties and usage. It showed that proper name based information has helped enhance file system performance [6].

Conclusion

In simple terms to conclude this review, file naming is an important practice that will influence how efficiently you retrieve and organize your data for the long term. The key to having excellent organization is that a good structured file naming convention will make sure that your information stays organized, accessible, and easy to manage later on even when the amount of data increases.

Clarity, consistency and your personal approach is important. The Guide to Developing a File Naming Convention and the Nebula File System displayed that without a doubt. Programmers with the TILDE system also showed that even with less conventional methods, as long as you understand what you are doing, any method will work.

After examining all these literary pieces, the most crucial part of file naming is to make sure you understand why you chose that convention. With the digital world continuously

expanding, having your own naming convention for your data will become even more important down the line to make sure you stay organized and keep your files accessible.

References

[1] University of Glasgow, "A Guide to Developing a File Naming Convention," 2020. https://edshare.gla.ac.uk/807/1/File_Naming_v2_20200608.pdf

[2] M. Bowman, et al., "A File System for Information Management," 1999. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=99bcfbe095b591c696809ca21533bcd1d5502587

[3] D. Comer and R. E. Droms, “The tilde naming system,” *Proceedings of the 2nd workshop on Making distributed systems work - EW 2*, pp. 1–4, 1986. doi:10.1145/503956.503972

[4] J. W. Crowder, J. S. Marion, and M. Reilly, “File naming in Digital Media Research: Examples from the Humanities and Social Sciences,” *Journal of Librarianship and Scholarly Communication*, vol. 3, no. 3, Dec. 2015. doi:10.7710/2162-3309.1260

[5] H. Chapman, *The file naming habits of personal computer users*, Mar. 1999. doi:https://doi.org/10.17615/z4xq-2188

[6] Ellard, Daniel, J. Ledlie, and M. Seltzer. 2003. “The Utility of File Names.” Harvard Computer Science Group Technical Report TR-05-03.