A Mechanism:To Organize Spam File using Python

Amit Shukla
Galgotias University
Greater Noida,India
amitshukla@galgotiasuniversity.edu.in

Dr Kirti Shukla
Galgotias University
Greater Noida,India
kirti.shukla@galgotiasuniversity.edu.in

Harshita Deep
Galgotias University
Greater Noida,India
dharshii26@gmail.com

Abstract— "**The organizing of your computer documents is vital to your success in your professional and personal life. The average worker now loses over two hours looking for misplaced documents and emails on their computer. You can avoid this huge loss of time with organization through the use of folders and subfolders. Here's when our Junk File Organizer comes into play. Due to the large number of files, it is a daunting task to sit and organize each file. To make that task easy the below Python script comes handy and all the files are organized in a well-manner way within seconds. It is easier to locate a file if it is placed in a well-named subfolder. It is often difficult to find files that are not organized in any way. If you organize your documents into subfolders, you will be able to find all of the similar files together. To obtain this goal, we go through some steps which include Creating Dictionaries which creates defined directories, Mapping which maps the file formats with directory and; Organizing which does the main job.**"

Keywords—Junk Files, Software, Python, Organizer, Time Management.

I. INTRODUCTION

A. Problem Statement

The business world is fast-paced and always forward bound. To keep up with industrial growth and your own personal advancement, it's essential to practice good organizational skills. And, among this mess, there are tons of programmers and being lazy is a trait of a human being whether he belongs to any profession. If someone is in the IT field, he must have a desktop or laptop full of files. It can be said that being disorganized can hinder your ability to get where you want to be professionally, decrease your productivity, cost your employer time and money, and can even be bad for your physical and mental health. As the location of files of different types is set in a particular computer in different folders, and as the number of files increase, the folder starts to become messy. Due to this, the task of organising each file becomes very daunting.

B. Problem Solution

Many people wish to advance their careers and climb those corporate ladders, but they won't get far unless they are organized. To make this task easy, Python script comes in handy to organize all the files in a well-manner way within seconds. There are two ways to solve this issue. First, you can create different folders for different files manually and then place the files according to their types, for example, video files need to be placed in a folder named Video, audio files need to move to the Audio folder, photos to the Picture folder, and so

on. It's a very time-consuming and boring procedure if there are a huge number of files that need to be sorted. But as we all know, we, humans don't have that much patience and few has lack of time as well.

Coming to the Second way to solve this problem, we can create a program that can do similar work for us but automatically. Being a python programmer, the idea came to my mind that if a python program could be created that will organize a large number of files, then it would make our task easy.

II. LITERATURE SURVEY

A. Background Information

According to Texas A&M University, some types of jobs--such as engineering--require good organizational skills to get the job done. When a position calls for an employee to be orderly, but they cannot exhibit signs of good organizational skills, it can result in the employee being dismissed from a project or terminated from their position.

Poor organizational skills are a hindrance to productivity. Instead of being able to get work done, poor organizational skills will leave you rummaging through documents and attempting to carve out a clutter-free area on your computer. Spending all of this time being non-productive causes your efficiency to suffer, which will not look good to your boss.[1]

B. How Poor Organization affect people

There are series of effects on people with this problem:

- Irregular sleep schedules

- Increase in stress

- Poor organization can stifle productivity and decrease employee engagement.

- It can easily elongate the time of a given task.

- Disorganization can cause employees to confuse dates, mix up assignments, miss deadlines and, in turn, this makes them unreliable and undependable.

C. Statistical Diagnosis

Twenty-five students of the University of Missouri–St. Louis volunteered to participate. The study was approved by the institutional review board of the University of Missouri St.

Louis (approval number 921551–2) and participants provided written informed consent according to the Declaration of Helsinki at the start of the testing session.
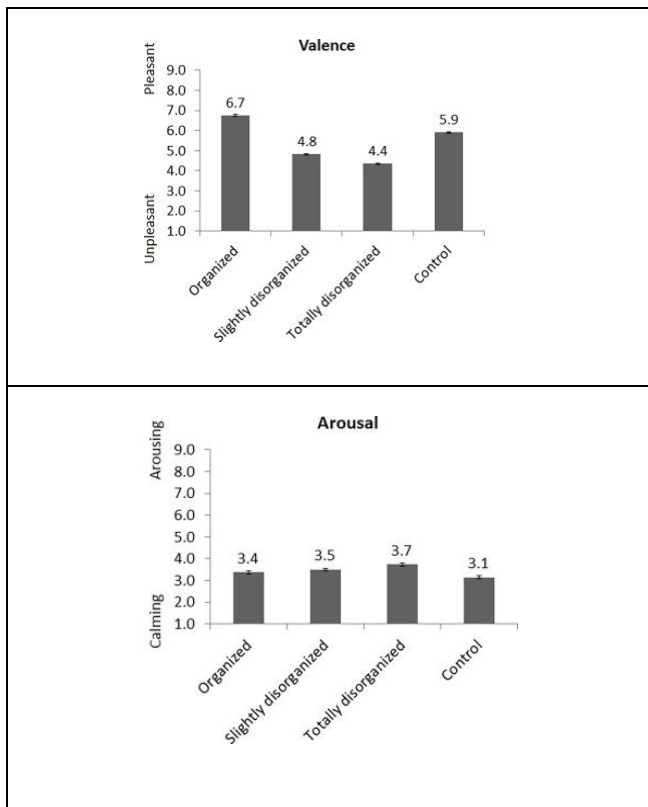

Fig. 1. Valence and arousal ratings for each of the four conditions

Follow-up tests showed that the ERP was more negative for totally disorganized than organized and slightly disorganized pictures at frontal and central electrodes.
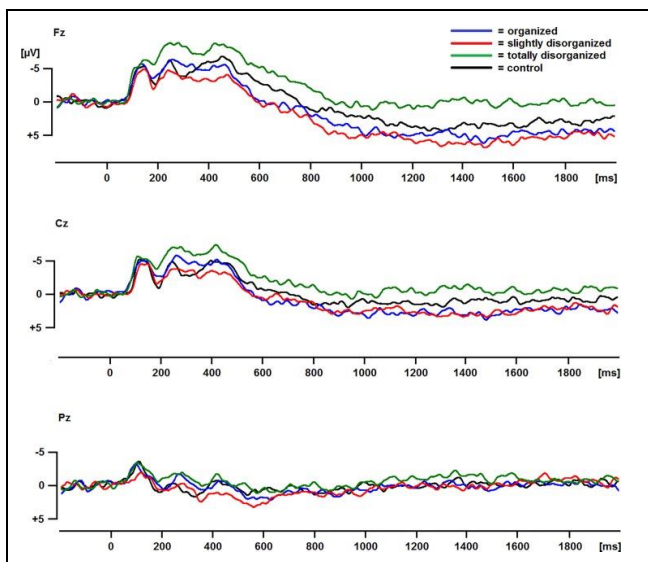

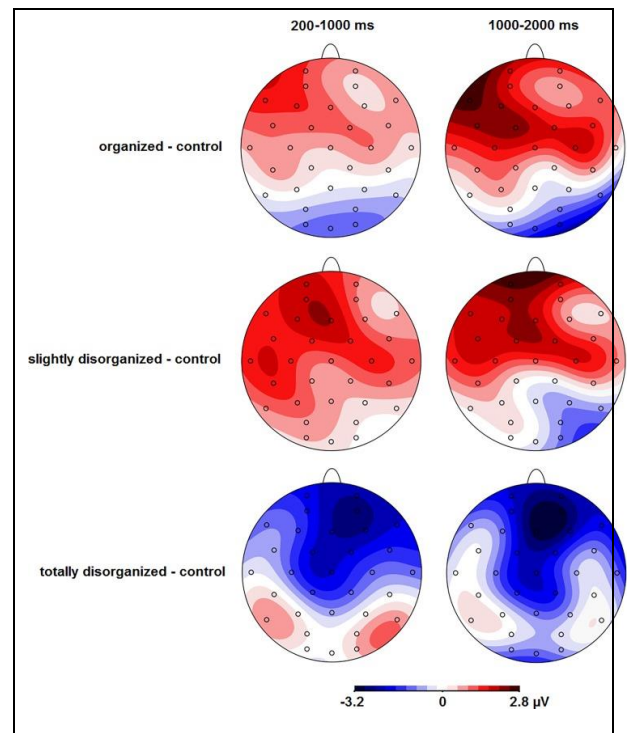Fig. 2. ERPs of each of the four conditions

Fig. 3. Scalp topgraphies of the difference waves in the two time windows.

The goal of this study was to determine how organization, slight disorganization, and total disorganization affect valence (i.e., how pleasant or unpleasant someone feels), arousal (i.e., how calm or aroused someone feels), and the frontal negativity in the ERP in healthy participants. To this end, participants viewed organized, slightly disorganized, totally disorganized, and control pictures that displayed items in their typical arrangement, while their EEG was recorded.

III. DATASET

People who live in crowded cities are also more inclined to keep a cluttered desktop (42 percent), while those that keep it clean are 9 percent more likely to live in the suburbs and 13 percent more likely to live in a rural area.

Poor user maintenance of device content generates a build-up of digital clutter. Some of the major problems that result from digital data overload are an increase in intrusive and unsolicited ads that often use vulnerabilities to penetrate the device: for computers (55%).

This digital clutter and careless user behavior is leaving devices — and the sensitive information they contain — vulnerable to security threats: for computers (51%).

I've gathered some interesting statistics on the cost of clutter and disorganization in time, money and quality of life. Seriously consider what clutter and chaos is costing you in your home and/or business. It can be a real eye opener.

Take a few minutes to review the Organizing and Time Management Statistics below:

A. *Information Management*

- 96% of office workers are frustrated by their company's information management, Harte-Hanks
- According to Forbes ASAP, the typical executive today wastes 150 hours a year, almost one month, searching for lost information. For someone earning $50,000 a year, that translates to a loss of $3,842
- Typical US worker is interrupted by communications technology every 10 minutes, Institute for Future and Gallup
- 80% of papers and information that we keep, we never use, Agency Sales Magazine
- Email is increasing print volume by 40%, Document Magazine

B. *Stress and Work/Life Balance*

- 80% of our medical expenditures are stress related as stated by The Centers for Disease Control and Prevention, Fast Company 2003
- Stress-related illness costs the nation $300 billion a year in medical costs and lost productivity, American Institute of Stress
- Perfectionism costs 50% or more of the total effort to squeeze out the last 10% or so of quality, Jeff Olson, The Agile Manager's Guide to Getting Organized, 1997

C. *Time Management*

- The National Association of Professional Organizers says we spend one year of our lives looking for lost items

- Office workers waste an average of 40% of their workday. Not because they aren't smart, but because they were never taught organizing skills to cope with the increasing workloads and demands, Wall Street Journal Report
- In a 2008 NAPO survey of 400 consumers nationwide:
  - 27% said they feel disorganized at work, and of those, 91 percent said they would be more effective and efficient if their workspace was better organized
  - 28% said they would save over an hour per day
  - 27% said they would save 31 to 60 minutes each day
- It's estimated that a manager loses 1 hour/day to disorder, costing the business up to $4,000/yr if earning $35,000/yr or $8,125/yr at $65,000\
- 280 hours (7 weeks) per year are lost by workers seeking clarification due to poor organization & communication
- The average American spends one year of their life looking for lost or misplaced items at home and in the office, US News and World Report
- According to the American Demographic Society, Americans waste more than 9 million hours each day looking for lost and misplaced articles
- For every hour of planning, 3 to 4 hours are saved from redundancy, waiting for information, not being prepared and poorly managed tasks
- 40% of adults say if they had more time, they would spend it with family, Day Runner Survey

D. *Organizing Products and Tools*

- Using the correct organizational tools can improve time management by 38%, Mobile Technology Product
- Home storage products have become a $4.36 billion industry, Newsweek 2004

IV. METHOD TO IMPLEMENT

A. Flowchart
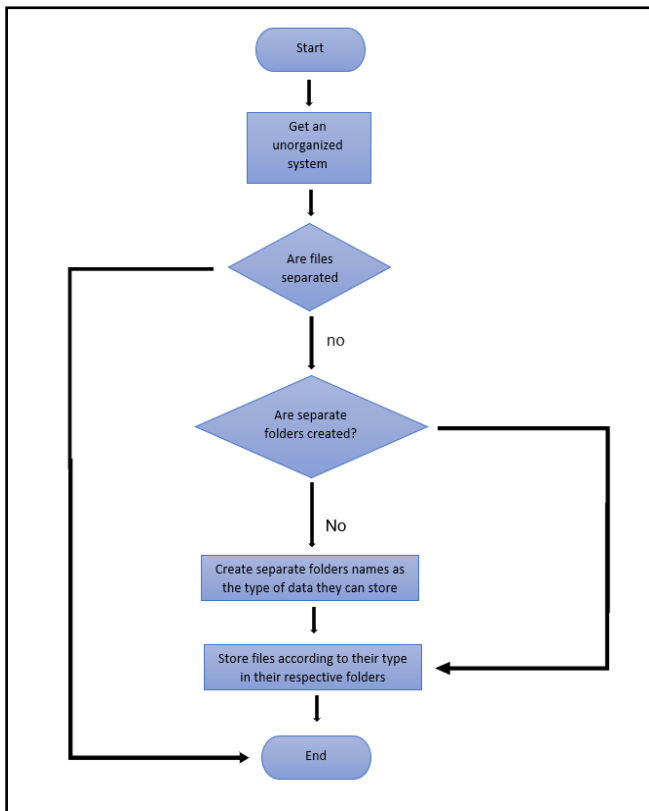
B. Creating Dictionaries

The code below will create the defined Directories.

```
DIRECTORIES = {
    "HTML": [".html5", ".html", ".htm", ".xhtml"],
    "IMAGES": [".jpeg", ".jpg", ".tiff", ".gif", ".bmp",
               ".png", ".bpg", "svg",".heif", ".psd"],

    "VIDEOS": [".avi", ".flv", ".wmv", ".mov",

              ".mp4",".webm",".vob", ".mng",

              ".qt", ".mpg", ".mpeg", ".3gp"],

    "DOCUMENTS": [".oxps", ".epub", ".pages",

                 ".docx", ".doc",".fdf", ".ods",

                 ".odt", ".pwi", ".xsn",".xps",

                 ".dotx", ".docm", ".dox",

                 ".rvg", ".rtf", ".rtfd", ".wpd",

                 ".xls", ".xlsx", ".ppt", "pptx"],

    "ARCHIVES": [".a", ".ar", ".cpio", ".iso", ".tar",

                ".gz", ".rz", ".7z", ".dmg", ".rar",

                ".xar",".zip"],
```

```
    "AUDIO": [".aac", ".aa", ".aac", ".dvf", ".m4a",

             ".m4b", ".m4p", ".mp3", ".msv",

              "ogg", "oga", ".raw", ".vox",

              ".wav", ".wma"],

    "PLAINTEXT": [".txt", ".in", ".out"],

    "PDF": [".pdf"],

    "PYTHON": [".py"],

    "XML": [".xml"],

    "EXE": [".exe"],

    "SHELL": [".sh"]


}
```

C. Mapping

Now we will map the file formats with the directory.

```
FILE_FORMATS = {file_format: directory

    for directory, file_formats in

    DIRECTORIES.items()

    for file_format in file_formats}
```

D. Mapping file extensions to directory

Here, we map file extensions with the directory.

```
def organize_junk():

    for entry in os.scandir():

        if entry.is_dir():

            continue

        file_path = Path(entry)

        file_format = file_path.suffix.lower()

        if file_format in FILE_FORMATS:

          directory_path =
Path(FILE_FORMATS[file_format])

            directory_path.mkdir(exist_ok=True)


file_path.rename(directory_path.joinpath(file_path))
```

```
for dir in os.scandir():
    try:
        os.rmdir(dir)
    except:
        pass
```

E. Organizing

The code for Python Lazy Junk Files Organizer will organize everything in the appropriate folder in a single go and remove empty directories.

F. Information Overload?

These days our devices are used as storage for all aspects of our digital lives. Users tend to save everything, even things that they know, deep down, they'll never need or be able to find. They save files downloaded, apps, emails archived, photos taken and even those funny videos of cats!

This type of data is a major contributor to unnecessary clutter on our devices. But unlike clutter in the real world — where stacks of papers, books, clothing and other assorted junk can be physically seen — there are no obvious outward signs that could indicate a problem.
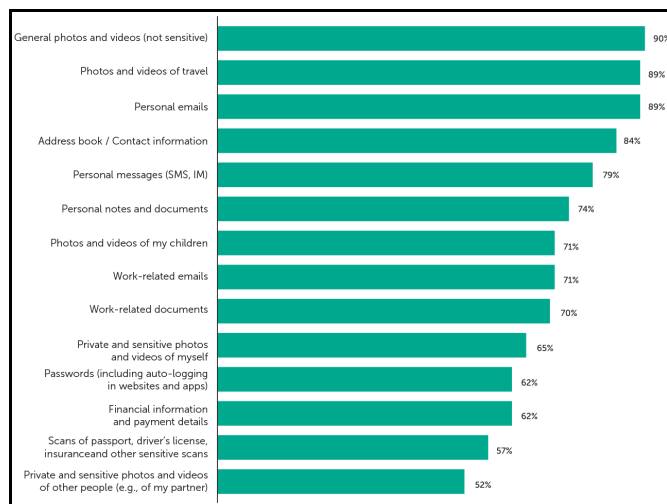
G. Approach

1. Import all the library functions required ..like shutil...os etc..

2. Create different folders based on type of files we are going to segregate into different folders using dictionaries.

3. Each folder will represent the name of the files present inside it.

4. Create the map of the file types into their respective folders.

5. Create a function to filter file types into their respective folders.

6. Use the OS module of python to recursively list out all the files that are present in the folders into newly created folder like their respective name.

V. RESULT AND DISCUSSION
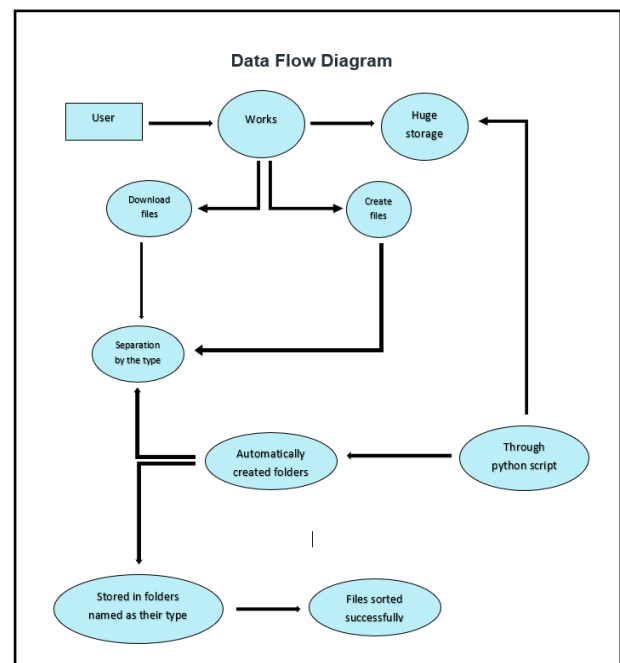
Here is how, our ***Junk File Organizer using Python*** works. With time, we've unknowingly accumulated mountains of unwanted data that could leave us exposed to ever-increasing cybersecurity threats. The digital world is growing and so is our capacity to store this data. But the fact that we have the capability to store vast information doesn't mean we should. The reason that most people are not organized is that it takes time. This is always a matter of personal preference, and is not something that a geek on a website can tell you. You should always choose your own system, based on how your own brain is organized (which makes the assumption that your brain is, in fact, organized).

Figure 1: Before decluttering

Figure 2: After decluttering

VI. CONCLUSION

This report demonstrates the scale of the problem with the data that permeates across all the devices that help us to manage our digital lives. Digital clutter is increasing and so are user problems associated with it.

User behaviour and attitudes to applications are the source of many of these issues. Many users fail to undertake the simple but essential care of their device that cleans and updates software and apps, adjusts settings and uninstalls unused apps. These actions are important to the hygiene of devices and the data that exists on it — from phone through to tablet and computer.

Just as it's become traditional to clean out our closets, attics, and garages each spring, it's good practice to regularly clear out and refresh digital spaces. It keeps them running smoothly and protects security, so spending a little time to get your digital house in order could prevent you from losing important or sensitive information in the future.

REFERENCES

[1]   Barry, L., Holloway, J., & McMahon, J. (2020). ASD Research, 78, 101617.

[2]   Bailey, Christopher. "A database system for organising musique concrete." ICMC. 2010.

[3]   Xin Yang, Paul T. Schrader, and Ning Zhang. [3] Yang, Xin, Paul T. Schrader, and Ning Zhang. "ABIDE Deep neural network study on ASD

[4]   Krieger, Lars, and Jared R. Peacock. "MTpy: A Python toolbox for magnetotellurics." Computers & geosciences 72 (2014): 167-175.

[5]   Röst, Hannes L., et al. "pyOpenMS: a Python‐based interface to the OpenMS mass‐spectrometry algorithm library." Proteomics 14.1 (2014): 74-77.

[6]   V. Subramanian, V. Subramanian, V. Subramanian, V (2018). Deep Learning with PyTorch: A step-by-step guide on using PyTorch to create neural network models. Packt Publishing Ltd. is a company that publishes books.

[7]   Telles, M. A. (2008). Python power!: the comprehensive guide. Course Technology PTR.

[8]   E. Tebenkov and I. Prokhorov (2021). Machine learning techniques are being used to train AI conversation bots. 735-744 in Procedia Computer Science.

[9]   Lutz, Mark. Programming python. " O'Reilly Media, Inc.", 2001.

[10] Feldman, Stuart I. "Make—A program for maintaining computer programs." Software: Practice and experience 9.4 (1979): 255-265.

[11] Guthrie, Ruth, and Paul Gray. "Junk computing: is it bad for an organization?." Information Systems Management 13.1 (1996): 23-28.

[12] D. I. Jon, W. Kim, H. R. Wang, Y. S. Woo, J. S. Seo, Y. M. Park,... & W. M. Bahk (2017). Maintenance treatment according to the Korean pharmaceutical algorithm for depressive illness (KMAP-DD). S530-S530 in European Psychiatry, 41(S1).