

Yash Prakash  
yash2002109@gmail.com  
Varanasi, Uttar Pradesh, India  
Timezone: Indian Standard Time (UTC +5:30)

# Automate Metrics Release and Process Improvement

GSoC Project Proposal for CHAOSS

## Personal Details and Contact Information

- ❖ **Github username:** [yash2002109](#)
- ❖ **Email:** yash2002109@gmail.com
- ❖ **University:** [Indian Institute of Technology BHU \(Varanasi\)](#)
- ❖ **Time-zone:** IST (UTC +5:30)
- ❖ **Address:** C-21/4 H-1, Station Road, Varanasi, Uttar Pradesh - 221002, India
- ❖ **IRC nick:** yash2002109
- ❖ **Microtasks:** <https://github.com/yash2002109/GSoC-21-Microtasks>
- ❖ **Project Link:** <https://github.com/chaoss/website/issues/537>

## Synopsis

CHAOSS metrics have been defined to provide an in-depth view into the various features of an open-source project. The metrics are very useful as with the growing popularity of open-source projects, it is becoming increasingly difficult to keep track of the needs of the project, its implementation and its requirements for the future. The metrics are a key input to help organizations strategically invest their resources.

As of now, CHAOSS operates with five working groups which contribute to the development of metrics. These working groups are:

- ❖ [Common Metrics](#)
- ❖ [Diversity, Equity and Inclusion](#)
- ❖ [Evolution](#)
- ❖ [Risk](#)
- ❖ [Value](#)

I would be working closely with the repositories of these working groups for the project.

Each working group has a set of [focus areas](#) with predefined metrics but they also try to develop more focus areas and metrics in the focus areas.

Currently, the metrics are present in the repositories of these working groups. Each metric is recorded as a markdown file in the repositories. The metrics try to answer a unique question related to an open-source project. The metrics also follow a template to maintain uniformity. These metrics are released biannually in the form of reports.

In addition to the working group repositories, I would also be working with the [translations](#) repository. The translations repository houses the translated version of these metrics. The translated languages are Chinese and Spanish.

The main aim of the project is to understand the metrics release process, propose process improvements and automate this release process.

The basic idea would be to

- ❖ Extract the metrics markdown files and sort them
- ❖ Convert these files into the LaTeX file format.
- ❖ Replicate any other required elements of the report in LaTeX.
- ❖ Create a master file and incorporate all the required files in it.
- ❖ Convert the master file to PDF with formatting to get the final report.

Kevin Lombard, Matt Germonprez and Georg Link are the mentors for this project.

## Benefits to the Community

The manual process of the release can be found [here](#). The process of generating these reports is quite tedious and requires a lot of effort. The project will help automate this

release process allowing the user to generate the required document in just a few simple steps.

Automation is important as it will help provide improved reports, fastrack the release process and also allow the user to format the reports by making some simple modifications in the code.

In the near future, the number of metrics is also expected to increase, handling a large number of markdown files and converting them by hand can be a very frustrating task. Automation, achieved through this project greatly reduces this burden.

The automated system for the release of reports will not only help CHAOSS but can also be implemented by other organizations which store and manage data in a similar manner (after a few customizations, of course) to enhance their quality of report generation.

Currently, there is no report being released for the translations of the metrics. The automated generation of reports for the translations would help people from other countries also to recognize the work CHAOSS is performing.

## **Current Status of the Project**

The majority of the work in this project focuses on the extraction of raw data in the form of markdowns, processing it, and then outputting a single PDF file containing all the required information.

Since there isn't any pre-existing automation in this project, I would be starting fresh and am open to any suggestions or requirements which the mentors may want to incorporate in the reports.

I have already completed the given set of microtasks which have been very helpful to identify the major requirements of this project. It has enabled me to effectively select the ideal tools required for the project.

I have already set up a local development environment which was also required for the microtasks. The microtasks have also helped me in doing a Proof of Concept for this project.

# Goals

## Goal 1:

Implement process improvement ideas that can help enhance the quality of reports.

## Goal 2:

Automate the metrics release process to create PDF from Github markdown files.

## Goal 3:

Automate the process of report generation for the translated versions of the metric markdowns.

Stretch Goals: Implement a GitHub action to check PRs for compliance in terms of structuring and syntax.

# Project Phases

## Phase 1: Shortlisting of process improvement ideas & Organization of the required files

The ideas for process improvements have been listed later in this document. The mentors can provide their feedback on these ideas, with their productive suggestions and recommendation, the ideas will be shortlisted. The shortlisted ideas will then be executed by suitable means.

The metric markdown files are currently in the repositories of the working groups. The required markdown files will be identified along with their dependent files. Any extra or unneeded file will be blacklisted so that the process of report generation is not hampered.

The same process will be followed for files in the translations repository. Since there is a separate repository for translations and it has a different structure, this process would be done by a second automation system to allow easier management.

## Phase 2: Ensuring that all elements of markdown files are displayed correctly and conversion of these files to LaTeX format.

The markdown files support the use of HTML syntax for displaying metrics. Sometimes this syntax can cause problems during the PDF conversion.

Some common errors encountered during the process of conversion are an incorrect table of contents, images not displaying, overflow of lines from the document. These errors will need to be rectified beforehand in the markdown files themselves to ensure that a quality report gets generated.

The metrics file will be converted to LaTeX format along with other elements. The LaTeX format has numerous ways in which formatting can be done to these files. These can be implemented during the process of conversion to generate the final report.

The markdown files for the translations have the syntax to incorporate its elements in English. Hence, the translations can also be edited to fix any of the errors described. The above-described process would be followed for the translated version of the metric markdowns too.

## Phase 3: Release of the final report and removal of unnecessary files.

The report will be generated by combining the LaTeX files and the required elements. The files will then be converted to the PDF format along with the other required elements of the report all at once.

This will generate the table of contents automatically and cuts down on the task of manually creating one and editing it every time new metrics are released.

The report generation process will create folders and intermediate files. Once the report is generated, these files will no longer be needed. Depending on the requirements, any fitting method can be chosen to ensure the removal of these garbage files.

The report for the translated metrics would also be generated in the same manner by a different automation system.

## **Deliverables**

I have listed below concise deliverables that can help the mentors evaluate the status of the project.

### **Deliverable 1: Implementation of Process Improvements, Automation of the Metrics Retrieval Process & their Conversion to LaTeX.**

Expected by Evaluation 1

I will present the mentors, ideas for process improvement. After having careful discussions with mentors, the chosen ideas will be implemented in the required manner.

Automate the process of retrieval of the required markdown files, along with the images and any other file which may be necessary for proper display of markdowns from different repositories in GitHub and store them in an organized manner.

Look for any errors in the metrics markdown files that could lead to erroneous display in the final report. Convert the extracted metric markdown files to the LaTeX file format.

### **Deliverable 2: Automated System for the generation of the English report of released metrics.**

Expected by week 6-7 of the coding period

After completing the 1<sup>st</sup> Deliverable, I will work on replicating/converting the other required elements of the report to LaTeX format. These can be the cover page, list of contributors, release notes, license, etc.

I would then be designing the structure to incorporate these files. The final step to achieve this deliverable would be to convert the master file to PDF which results in the generation of the required report.

## **Deliverable 3: Automated system for generation of final reports for the translated version of the metrics.**

Expected by the end of the coding period

After the completion of the above two deliverables, I would be working on creating an automated system for the translation of the metrics markdowns.

This automated system would also roughly perform the same set of tasks as its counterpart. Since the translations repository has a different structure and workflow, having a separate automation system specifically for this purpose would be preferred.

By this time, I would have completely automated the process for the report generation, the required report and their translations would be generated in just a few input commands by the user.

## **Expected Results**

By the end of this project, the expected result is to have fully functional automation systems in place, which can generate the required report and translations.

Handling the metrics markdown files manually opens up a possibility of human error. With the automation of this project, this potential error will get minimized. In times to come, the number of released metrics is also expected to increase. Hence, it will be best if the release process is automated.

Including the English version of the metrics, there are a total of three languages in which the metrics markdown files currently exist. Hence, a total of three reports would be generated, one for each language. There would be two automation systems in place to achieve this.

The first system will be for the English version of these metrics. Since these metrics are scattered across different working group repositories, this system would extract the required files from each repository and then process them with the other required elements to generate the required report.

The second system would be for the translated version of these metrics. Since a single repository contains all the required metrics of the different languages. This system would walk through and extract the files for a particular language. It would then process them with other elements of the same language to generate the required report. In the future, this automation system can be modified to allow reports for more languages to be generated too.

The automated systems will be expected to perform a set of tasks that when combined will help meet the goals of the project. The set of tasks that the automated systems will be performing have been listed below.

- ❖ Cloning of the required working groups/translations repositories.
- ❖ Identification of the markdown files from the focus-areas/metrics folder for each repository.
- ❖ Organization of all the markdown files along with their necessary elements (e.g – Images)
- ❖ Conversion of the markdown files to LaTeX format.
- ❖ Replication and incorporation of other necessary elements such as cover page, list of contributors, etc in LaTeX format.
- ❖ Inclusion of the required files in a LaTeX master file and subsidiary files.
- ❖ Styling of the LaTeX file to ensure that the generated reports have an easy to follow and elegant design.
- ❖ Converting the master file (which includes the subsidiary files) to generate a single PDF. This would be the final release document.
- ❖ Deletion of all the other unnecessary files and repositories after the report has been generated.

In addition to this, the script will also include checks to make sure that the reports are generated in the required manner.

## **Suggested Process Improvements**

Currently, the release version of the report is quite basic. I have listed below some possible process improvements that can be implemented in the report. Depending on the feedback of mentors, the appealing ideas can be selected and executed during the coding period.



- ❖ While listing down the focus areas and their respective metrics for each working group, there is no specific description provided as to what the working group aims to achieve. A brief summary about the working group can be added and would look good.
- ❖ A little 'participate' section could be appended to the reports, this would help any willing contributor outside the organization to easily get started.
- ❖ The major tools which are responsible for providing the metrics are Augur and Grimoire Lab. A page can be included at the end of the report providing a quick introduction to the reader about Augur and Grimoire Lab. Since these softwares are developed by CHAOSS itself, adding a page will also help popularize them.
- ❖ The contents at the beginning of the report are listed without any heading. A simple heading would give the report better readability.
- ❖ I was unable to find the [front matter](#) (cover page, list of contributors, etc.) in the translated languages. For report generation of translated metric markdowns, these pages will have to be created.
- ❖ Presently under the subheading 'Tools Providing the Metric'. Some metric files have 'Augur' and 'Grimoire Lab' as hyperlinks to their respective websites and some do not. This can be standardized to follow a single format.

## Approach

[Python](#) is an easy to understand, portable, object-oriented, high-level language. It also has a vast collection of libraries that can simplify and help in the automation of many processes. Hence, I would be doing most of the coding required for this project in Python, I will also be using its various libraries to achieve the goals of the project.

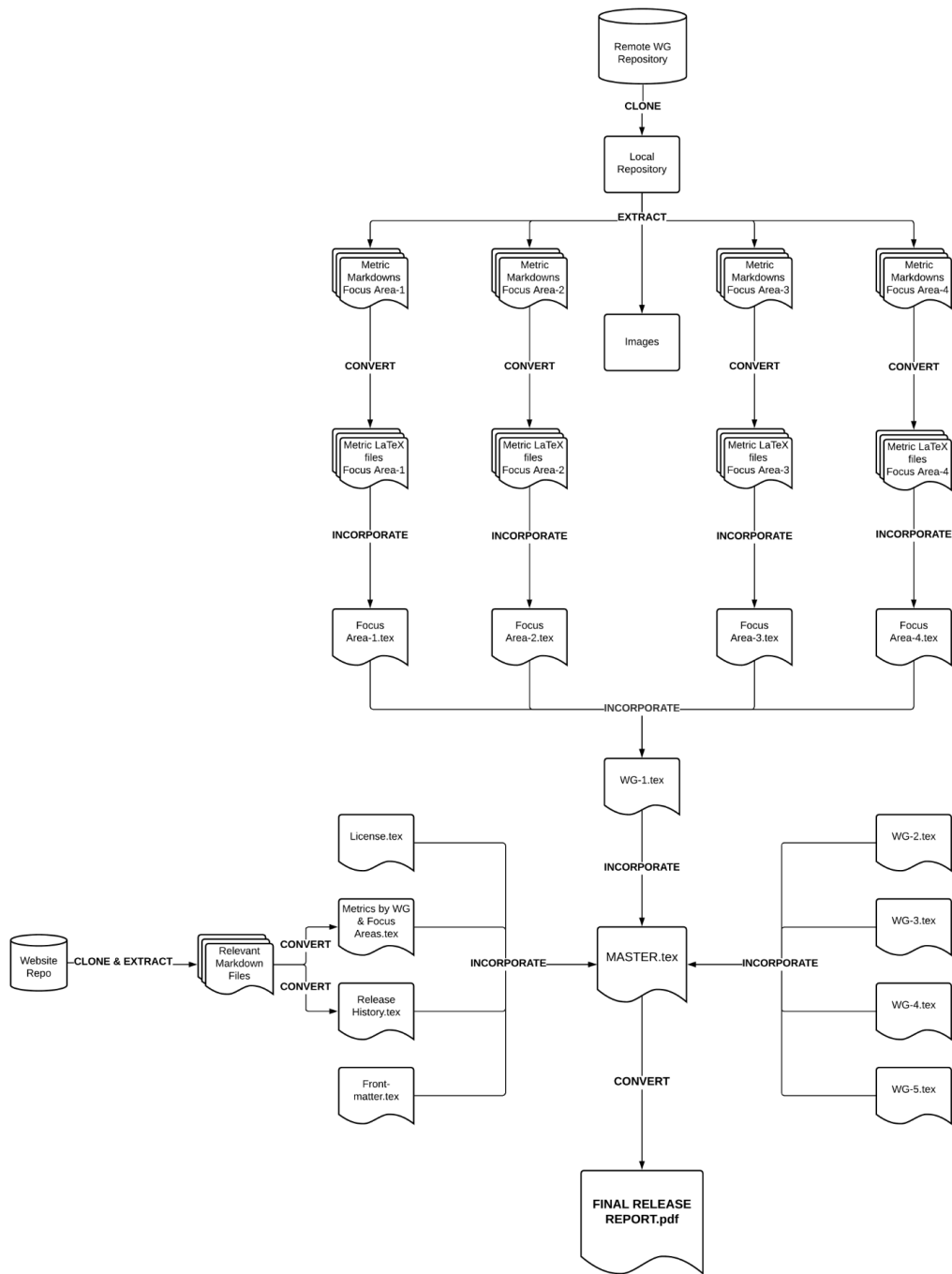
Since most of the projects and softwares which are designed in CHAOSS are developed using Python, it makes even greater sense to use Python. This will allow anyone from the organization to easily understand and edit the code, if required in the future.

I plan to create two automation systems in the course of this project.

1. Automation system for English version of the metric report.
2. Automation system for translated versions of the metric reports.

Also in the future should the structure of either the working groups or the translations repository change, the other automation system for the unchanged repository can continue to function as required.

The automation systems will mostly be following the same set of processes but on different structures and workflow of the repositories. These processes along with their implementation have been discussed in this section.



Automation System for English version of metrics release report  
 (Click [here](#) to view the enlarged image)

## Phase 1: Automation of the metrics retrieval process

### ❖ Cloning of the required working groups/translations repositories.

- The process of cloning repositories to create a local copy can be done by the following two methods.
  1. Using the Github API ([PyGithub](#) library).
  2. Using the [Git-Python](#) library.
- Of the two methods described above, I have concluded, the second option would be most suitable for the purpose.
- Since the project calls for only the repositories to be cloned locally and no changes need to be made in the existing repository, there should be no need to use the API.
- Using the Git-Python library is simple and performs the required task in minimalistic commands.

```
# code to clone repository

repo_url = "https://github.com/chaoss/wg-common.git"
repo_name = "wg-common"

# absolute path of directory in which script is being run
script_dir_path = pathlib.Path(__file__).parent.absolute()

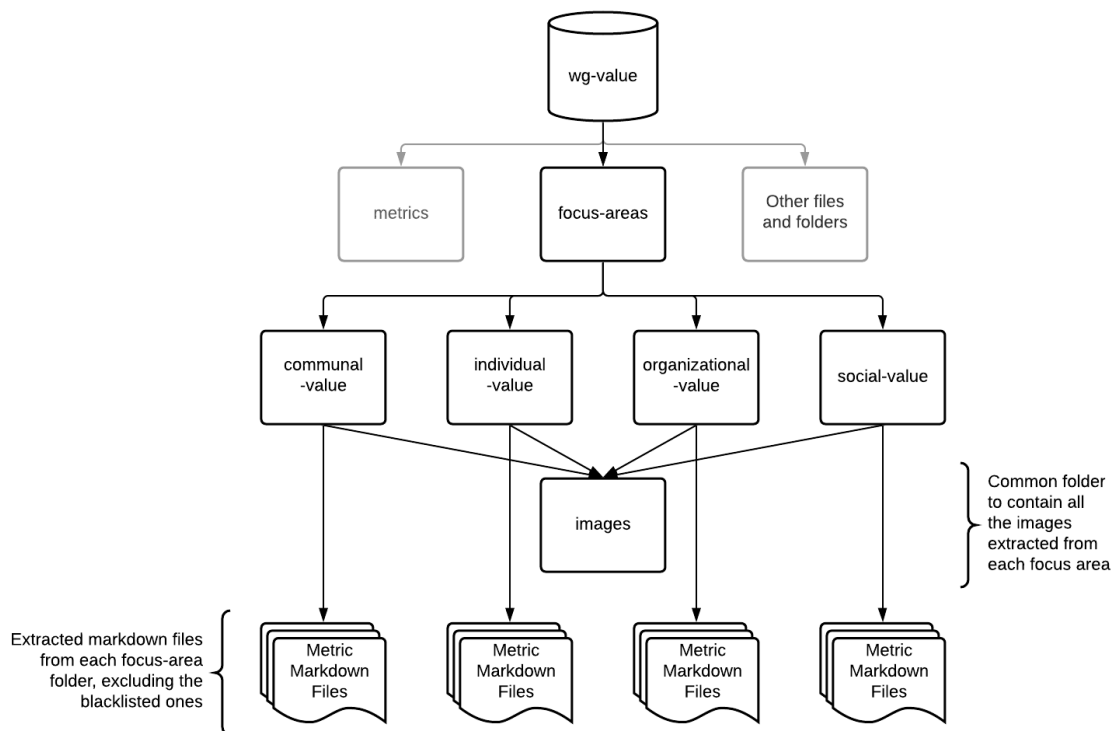
clone_to_path = os.path.join(script_dir_path, repo_name)

try:
    print(f"Cloning '{repo_name}' from {repo_url}")
    git.Repo.clone_from(repo_url, clone_to_path)
    print("Cloned Successfully")
except:
    print(f"Error: Unable to clone {repo_url} repository")
```

### ❖ Identification of the required markdown files from each repository.

- Currently, the required markdown files for the English version of the release are stored in the '*focus-areas*' folder of each repository.
- Some repositories do not yet follow the standard [template](#). The process of creating a uniform structure for these working group repositories is underway.

- Any of the markdown files that are present in the metrics folders that do not have to be included in the final report can be blacklisted. These files can be placeholders or README markdowns.
- The translations of these markdown files are stored in another repository. The repository presently contains two support for two languages (Chinese and Spanish). Each folder for these languages is divided further based on working groups. The metric markdowns files are located in these folders and will be extracted.



(Structure of [wg-value](#) - Click [here](#) to view the enlarged image)

- The [os](#) library can be used to create a directory tree of all the folders containing these metrics. The `os.walk()` function can then be used to walk through each of these files and copy them to a separate folder for conversion.
- The process of copying the required files can be done easily with the help of [shutil](#) library.

```
# following code is used to copy metrics markdown file
# present in each focus area to a common folder (root = "focus-areas")
for folder, sub_folders, files in os.walk(root):
```

```

#check for images folder or if all elements of files are blacklisted
if folder.endswith("images") or check_subarray(files, blacklisted_files_list):
    continue

copied_metric_md_list = []

for file in files:

    if file not in blacklisted_files_list:
        file_path = os.path.join(script_dir_path, folder, file)
        dest_path = os.path.join(script_dir_path, file)

        if os.path.isfile(dest_path):
            print(f"File with same name already exists at destination:
{file_path}")
            copied_metric_md_list.append(file)
        else:
            shutil.copy(file_path, dest_path)
            copied_metric_md_list.append(file)
            print(f"Copied Successfully: {file_path}")

# At this point we will have the markdown files for a specific
# focus group in the common folder

```

❖ Organization of all the markdown files along with their necessary elements (e.g. – Images)

- Once the paths of all markdown files and the other required elements have been identified. I will ensure that no element incorporated in the markdown files is missing, that is all the images and if any other files which are required for proper display of markdown have been successfully moved to their respective folders.
- Identify any files that do not follow the metrics display template and fix them.

```

# code to copy all the images from each focus-area in a particular
# repository to a common images folder
for folder, sub_folders, files in os.walk(root):
    for file in files:
        if folder.endswith("images"):
            file_path = os.path.join(script_dir_path, folder, file)
            dest_path = os.path.join(script_dir_path, "images", file)

```

```
if os.path.isfile(dest_path):
    print(f"File with same name already exists at destination:
{file_path}")
else:
    shutil.copy(file_path, dest_path)
    print(f"Copied Successfully: {file_path}")
```

## Phase 2: Ensuring that all elements of markdown files are displayed correctly and conversion of these files to LaTeX format.

For conversion of markdown files into PDF. I have decided to use the [PyPandoc](#) library. [Pandoc](#) is free and open-source software, released under the [GPL](#). It is widely used as a conversion tool and can be used to convert files from markdown to other formats.

PyPandoc is a library that has been created on the foundations of Pandoc. It allows the user to easily implement various utilities provided by Pandoc in a Python script. PyPandoc provides the necessary functions for this task and also saves the unnecessary hassle of creating a Shell script.

The process of conversion would be done in three steps:

1. Conversion of metric-markdown files LaTeX.
2. Conversion of front content, list of contributors, release notes, etc. from their respective format to LaTeX.
3. Conversion of the master LaTeX file to generate the final PDF.

The processes described will be implemented in both the automation systems.

- ❖ Making necessary changes to the markdowns to ensure proper display of all elements and any formatting (if required).
  - The markdown files when opened locally do not display the images which are incorporated using HTML or raw Github URLs. I will be fixing any such inconsistencies and changing it to the markdown syntax with relative paths.
  - Since this process is required to be done only once for all times to come in each markdown, I will be doing this manually and submitting the required PRs ([wg-common/pull/112](#)) to ensure that the above syntax is followed.

- In addition to manually checking for these errors, the script will contain checks to make sure that these elements are not incorporated in the wrong manner.
- If such an error is found, the script would notify the user describing the error.
- I will also be on the lookout to see if any other modifications need to be made to the markdown files.
- I will also make a checklist regarding the changes that will ensure that a markdown file gets displayed fittingly in the created PDF.
- In the future, the working groups can check their newly released metrics with this checklist before releasing it to ensure that the metric file does not display erroneously when added in the final report.
- The translations also follow the English commands and the syntax is the same irrespective of the language, hence they too would be checked and edited the same way.

❖ Conversion of metric-markdowns to LaTeX.

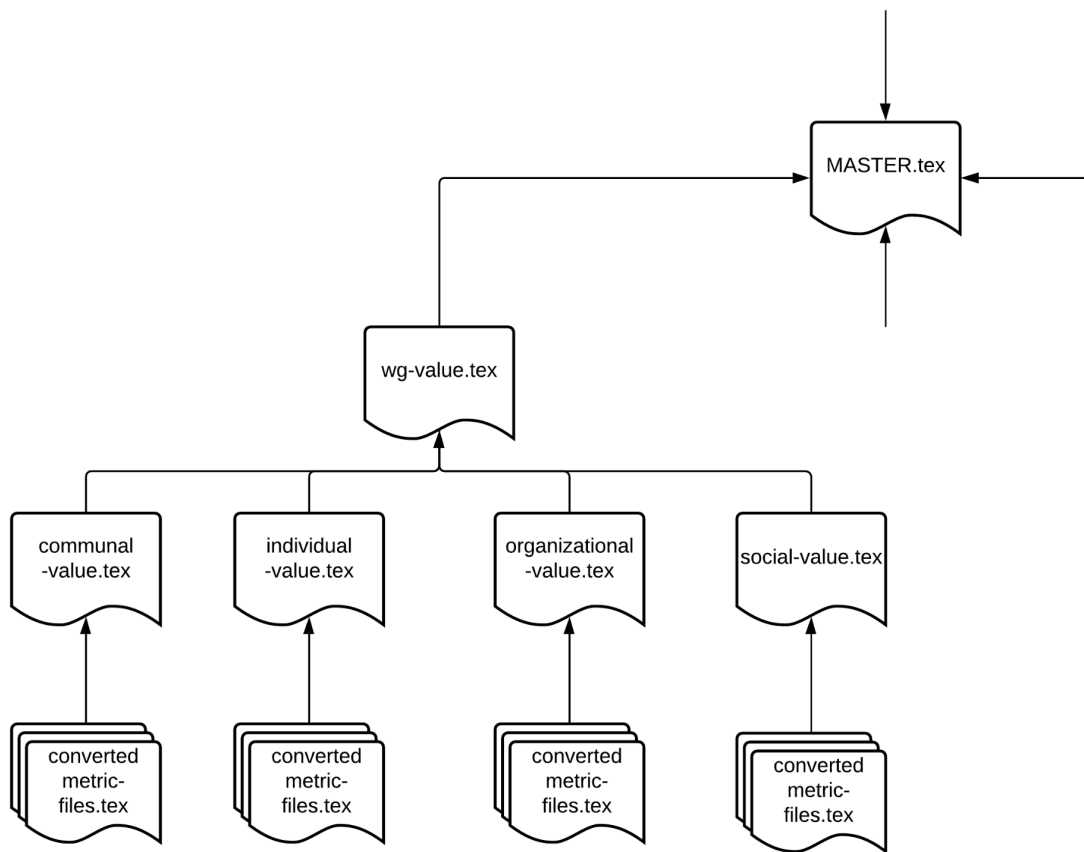
- The process of converting metric markdowns can be done by Pandoc itself.
- The extracted markdown files can be directly converted to LaTeX file format.
- These files will then be further incorporated at their appropriate level in the structure (described below) of the automation system.

❖ Incorporation of metrics in the master file.

- The converted metrics files now have to be linked to the master file.
- This process can be done by creating a tree structure and incorporating the required file at its respective level
- This incorporation can be done by using LaTeX commands (`\include{}`, `\input{}`, `\includeonly{} etc.`) .
- Some LaTeX commands do not allow nesting, hence their alternatives would be found and used during the project course.
- The master file will contain the required front-matter, license, and incorporate the working groups' file.
- The working group files would contain the description of the working groups and incorporate the focus areas files.
- The focus area files for each working group would incorporate the converted metrics files.



- The flowchart added below provides an example of how the process would be done for each repository.



*(Incorporation of metrics - Click [here](#) to view the enlarged image)*

- The metrics for translations would also be organized in a similar manner where the LaTeX master file will be one for each language.
- ❖ Adding other markdown elements such as metrics by working groups and focus areas, release notes, license, etc.
- The metrics files sorted by working groups and focus areas can be found [here](#). These markdown files contain the required information in a tabularized manner.
  - These markdown files need to be inserted before the individual metric files in the report.
  - The [release notes](#) are also recorded as markdowns which will be added to the report.

- Finally, the [license](#) file which currently exists markdown in each repository will be added to the report.
  - All the files mentioned above will be added to the report by once again converting them first to LaTeX format.
  - These files will then be incorporated at the appropriate level in the hierarchy of structure of the LaTeX files.
- ❖ Incorporation of the cover page, list of contributors, and any other element which needs to be prepended.
- Currently, CHAOSS uses a word document for the [preliminary content](#) of the report, makes the necessary changes to this document, and incorporates it into the final report.
  - This document will have to be converted/replicated to the LaTeX format.
  - Any other element which is required to be added before the table of contents would be a part of this file.
  - The replicated LaTeX file would then be added to the report during its PDF conversion.

### Phase 3: Release of the final report and removal of unnecessary files

The process of report generation would generate some intermediate files. Once these files are no longer required their removal can be done by the automated system itself. The removal of all the inessential files can be done by using `os` and the `shutil` library once again.

All these processes described below would be common to both the automation systems.

- ❖ Formatting the PDF during conversion
- Since the intermediate form of conversion is a LaTeX file, The PDF can be styled by creating a LaTeX file and including it as a header during conversion.
  - The LaTeX file will contain the commands that are required for the proper formatting of the markdown files.
  - These commands can be used to modify font size, font style, list styles, etc.

- Additionally, a JSON file could be added changing the different highlight styles. Currently, Pandoc provides some highlighting styles but the JSON scripts for these files can be modified to customize these styles.
- Pandoc also has support for the YAML file format. The YAML header options can be added to the final markdown file before converting it to PDF for easy formatting.

## ❖ Conversion of the LaTeX master file to PDF.

- After deciding on the proper formatting. I plan to incorporate all these changes by using different formatting flags offered by PyPandoc
- The current version of the report is released by converting the master file to PDF. This converts all the files incorporated to the root level to be converted to PDF.
- I plan to use XeLaTeX as the PDF engine, unlike other PDF engines XeLaTeX supports the use of other languages hence converting PDFs for the translations too should not be a problem.
- The converted document would be a single PDF. This PDF would contain all the required elements of the release version of the report.

[illegible]

```

print(f"Created successfully: '{final_report_name}")
except:
    print("Error: PDF could not be generated successfully")

```

❖ Adding any other required customizations. (Optional)

- This section is optional and the changes would only be done to the converted PDF in case any of the demands of the document are unmet by LaTeX formatting.
- Depending upon the choice and recommendations of mentors, any further customizations required can be done to the PDF itself at this stage.
- The PDF obtained after this process would be the required release version of the final report.

❖ Removal of other unnecessary files and folders

- Depending on the choice of the user, the repositories which were cloned initially and the other intermediate files which were generated in the process can be deleted.
- The 'os' and 'shutil' library once again can be used to perform this task.

```

def delete_files(file_path_arr):
    for file_path in file_path_arr:
        try:
            if os.path.isfile(file_path):
                print(f"Deleting: {file_path}")
                os.remove(file_path)
                print(f"Deleted Successfully")
        except:
            print(f"Unable to delete {file_path}")

def delete_folder(folder_path):
    try:
        if os.path.isdir(folder_path):
            print(f"Deleting folder: {folder_path}")
            shutil.rmtree("images")
            print("Deleted Successfully")
    except:
        print(f"Unable to delete folder: {folder_path}")

```

## Timeline

I plan to implement the process improvement ideas and create the required automation system for English version of the release simultaneously. Once these two goals are achieved, I would be focusing on creating the system for automation of translated versions of metrics release reports. I would be working on each automation one at a time.

Since I would be having the experience of creating one automation system during the project period, the other one can be created in a shorter duration of time. Some elements of the first automation system such as the required formatting, cover page, etc. can be directly utilized in the second automation system too.

Time Period	Tasks
After proposal submission [April 13 - May 17]	<ul style="list-style-type: none"><li>• Continue contributions to the website and the working groups' repositories, centralized around metric markdowns and their release process.</li><li>• Restructure the translations repository and organize the translated metric markdowns similar to the structure in working groups ( <a href="#">#11</a>).</li><li>• Increase familiarity with LaTeX file format</li><li>• Explore alternative approaches to the project</li></ul>
Community Bonding Period [May 17 – June 7]	<ul style="list-style-type: none"><li>• Interact with the mentors, discuss the proposal and implement their suggestions</li><li>• Try to look for more ways to enhance the quality of reports</li><li>• Learn more about the structure and workflow of the repositories.</li><li>• Shortlist the process improvement ideas for report generation.</li><li>• Decide and finalize approaches to the chosen ideas.</li></ul>

GOAL 1 and 2	Implement the process improvement ideas & Create the automated system for generation of report for English version of metrics
<p>Week 1</p> <p>[June 7 – June 14]</p>	<ul style="list-style-type: none"> <li>• Create the draft version of the required files for the implementation of process improvement ideas.</li> <li>• Work on the data management scripts for the automation system.</li> </ul>
<p>Week 2</p> <p>[June 14 – June 21]</p>	<ul style="list-style-type: none"> <li>• Make necessary changes and finalize the files for process improvement after feedback from mentors.</li> <li>• Identify and organize the required metric markdown files.</li> </ul>
<p>Week 3</p> <p>[June 21 – June 28]</p>	<ul style="list-style-type: none"> <li>• Organize the other required elements for proper display of markdown files.</li> <li>• Scan and fix the markdown files for any errors in their display and fix them in CHAOSS's remote repository.</li> </ul>
<p>Week 4</p> <p>[June 28 – July 5]</p>	<ul style="list-style-type: none"> <li>• Create the script to convert markdown files to LaTeX file format using PyPandoc.</li> <li>• Create the required tree structure containing the master file and subsidiary files to incorporate the metrics files along with other report elements.</li> </ul>
<p>Week 5</p> <p>[July 5 – July 12]</p>	<ul style="list-style-type: none"> <li>• Convert the other required elements of the report such as the front-matter, metrics by working groups and focus areas, and release history, License to LaTeX.</li> <li>• Work on formatting the markdowns with the help of LaTeX commands, YAML headers, formatting flags, etc.</li> </ul>
<p>Week 6</p>	<ul style="list-style-type: none"> <li>• Generate the final release report</li> </ul>

[July 12 – July 19]	<ul style="list-style-type: none"> <li>• Take feedback from mentors and make any necessary changes to the report.</li> <li>• Test the automation system for any issues.</li> </ul>
<b>GOAL 3</b>	Automate the process of report generation for the translated versions of the metric markdowns.
Week 7 [July 19 – July 26]	<ul style="list-style-type: none"> <li>• Work on data management scripts for the 'translations' repository(second automation system).</li> <li>• Identify and organize the necessary elements required for proper display of these translated markdown files using the data management scripts.</li> <li>• Fix the translated metric markdown files for any errors in displaying them and update them in the remote repository.</li> </ul>
Week 8 [July 26 – August 2]	<ul style="list-style-type: none"> <li>• Convert the other required elements of reports in LaTeX. These would be in the translated languages.</li> <li>• Create the tree structure with the master file one for each language of the translations to incorporate their respective elements in the report.</li> </ul>
Week 9 [August 2 – August 9]	<ul style="list-style-type: none"> <li>• Decide on the required formatting of the reports.</li> <li>• Create the required script to convert the master files to PDF to generate the final release version of the reports for translations.</li> <li>• Start documentation of the automated systems.</li> </ul>
	Documentation and Submission of the work done
Final Week [August 9 – August 16]	<ul style="list-style-type: none"> <li>• Test the automation system for any issues.</li> <li>• Finish up the documentation of the automated systems.</li> <li>• Submit the final work and the work report.</li> </ul>

## About Me

I am Yash Prakash, a freshman undergraduate student. I'm pursuing my Bachelor's in the field of Electrical Engineering from the Indian Institute of Technology, BHU (Varanasi). The idea of open-source projects has fascinated me for quite some time now.

I personally enjoy the fact that people all around the world can connect and work together as a team towards a common goal. The open-source communities are fantastic and provide me with a very good environment to expand my skill set and also to contribute to projects which are developed for the good of the community.

Ever since I attended the first workshop on open-source projects, I have actively been on the lookout for opportunities to contribute and interact with people. The best part is not only do I get to learn a lot and make new friends but also the fact that I'm able to collaborate on projects which have a global impact.

I have significant experience writing Python and C. I'm also interested in web development and competitive programming. I have been using Git and Github for quite some time now.

During my period of association with CHAOSS, I have extensively worked with the metric markdown files from each of the working group repositories. I'm well familiar with the current structure of each repository and have even sent some PRs focused on standardizing them.

## Post GSoC

- ❖ By the end of the project, I hope to be a part of CHAOSS's family. I will be happy to make regular contributions to CHAOSS post the completion of this project.
- ❖ It is my interest to learn more about the workings of Augur and Grimoire Lab. Further, I would love to take part in more meetings and interact with the community.
- ❖ While being associated with CHAOSS, I have learned a lot of things in a very short duration of time, I'm really thankful to the community who have been very helpful and played an integral role in my growth.



## Contributions

- ❖ (Merged) Sample PR for restructuring the repository - [#14](#)
  - Repository - [chaoss/translations](#)
  - Proposed a structure similar to that of working groups.
  - Related to [#11](#)
- ❖ (Merged) Added Images & Fixed their links- [#12](#)
  - Repository - [chaoss/translations](#)
  - Related to [#11](#)
- ❖ (Merged) Fixed Hyperlink - [#571](#)
  - Repository - [chaoss/website/](#)
- ❖ (Merged) Fixed wrong question for 'contributor location' metric - [#583](#)
  - Repository - [chaoss/website/](#)
- ❖ (Merged) Renamed images and updated links to images - [#112](#)
  - Repository - [chaoss/wg-common](#)
  - Ensured that repository follows the standard template given in [#102](#)
  - Changed syntax of incorporated images
- ❖ (Merged) Updated LICENSE Year - [#110](#)
  - Repository - [chaoss/wg-common](#)
- ❖ (Open) Standardized Repo Structure - [#356](#)
  - Repository - [chaoss/wg-diversity-inclusion](#)
  - Modified metric files and focus area subfolder to follow the template in [#346](#)
  - Updated Links in README.md
- ❖ (Open) Standardized Repo Structure - [#406](#)
  - Repository - [chaoss/wg-evolution](#)
  - Created new directory to classify metric markdown files
  - Added Images required
  - Related to [#405](#) & [#395](#)
- ❖ (Merged) Updated LICENSE Year & Fixed Capitalizations -[#140](#)
  - Repository - [chaoss/wg-value](#)

❖ (Merged) Fixed Typo - [#188](#)

- Repository - [chaoss/metrics](#)

Working with the required working group repositories has allowed me to understand their current structure and their aims. Moreover, I have also regularly attended CHAOSS weekly community call-up meetings and the web-content meeting. This has familiarized me with the community and needless to say I'm very excited to work together.

## Microtasks

### [Link to the Microtasks Repository](#)

I have successfully completed all four microtasks assigned. I was quite successful at replicating the required results.

In addition to completing the microtasks, I have also worked on some mini-projects related to the project idea. The mini-projects are automation systems and help generate the PDF of metrics for a specific working group from its remote repository. These automation systems were created to choose the best approach for this project.

These projects can be found in the microtasks repository itself. The POC for the presented idea in the proposal can be found [here](#).

## Why CHAOSS?

I have been interested in contributing to the open-source community ever since I heard of it. The software developed in CHAOSS is written majorly in Python. Python is also the language I have been most comfortable with while working on my personal projects. The workflow of CHAOSS is well organized and makes it easy for contributors to locate the relevant issues and give their suggestions.

CHAOSS also has a very helpful community that has actively cleared my doubts and taken note of my opinion. The community bonds with everyone like family. They are always willing to connect and discuss any of the issues encountered. I would love to work with them further.

## Work Management and Communication

- ❖ I intend to host a Github repository of the project, during this project I will be regularly updating this repository which will help mentors keep track of the project.

- ❖ I'm quite flexible regarding the working hours since I enjoy working at night time, the time zone difference shouldn't be an issue.
- ❖ I also plan to communicate with mentors via Zoom calls every week or fortnight to update them regarding the status of the project. I will also discuss my doubts and consider their suggestions during that time.
- ❖ I intend to provide an update to the CHAOSS community by giving them a brief summary regarding the status of the project on a fortnightly basis. This can be done through the mailing list.

## Availability and Other Commitments

- ❖ I do **not** have any work-related commitments this summer.
- ❖ I plan to devote around 25 hours of my time to this project, per week, during the coding period. Depending on the project requirements, I'm prepared to devote more time to the project.
- ❖ My end-semester exams are scheduled to take place from 1<sup>st</sup>- 10<sup>th</sup> July. During this period, I would be able to allocate 10-15 hours per week to the project.
- ❖ To ensure that I do not lag, I will commit more time to the project during weekends and try to stay ahead of the timeline.

## Note of Thanks

I would like to thank Georg Link, Kevin Lombard and Matt Germonprez for their valuable suggestions. They have actively helped me in clearing my doubts and even provided their feedback for the improvement of this proposal. I would also like to thank other members of the CHAOSS community for their suggestions and help.

## More Information

For any further doubts or clarifications, please feel free to drop me a mail at [yash2002109@gmail.com](mailto:yash2002109@gmail.com)

## References

- ❖ <https://github.com/chaoss/website/issues/537>
- ❖ <https://handbook.chaoss.community/community-handbook/community-initiatives/metrics/chaoss-metric-release>
- ❖ <https://chaoss.community/metrics/>
- ❖ <https://pandoc.org/MANUAL.html>