MODEL FOR THE EVALUATION OF USABILITY OF SCIENTIFIC WORKFLOWS

AMELIA BINTI ABD RAHMAN

BACHELOR OF COMPUTER SCIENCE
(COMPUTER SYSTEMS & NETWORKING)

UNIVERSITI MALAYSIA PAHANG

UNIVERSITI MALAYSIA PAHANG

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name	: Amelia binti Abd Rahman
Date of Birth	: 12 June 1996
Title	: Model for the evaluation of usability of scientific workflows
Academic Session	: 2018 / 2019

I declare that this thesis is classified as:

☐	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997)*
☐	RESTRICTED	(Contains restricted information as specified by the organization where research was done)*
☑	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Universiti Malaysia Pahang reserves the following rights:

1. The Thesis is the Property of Universiti Malaysia Pahang
2. The Library of Universiti Malaysia Pahang has the right to make copies of the thesis for the purpose of research only.
3. The Library has the right to make copies of the thesis for academic exchange.

Certified by:

(Student's Signature)

960612-10-5814
Date: 8 January 2019

(Supervisor's Signature)

PROF DR. VITALIY MEZHUYEV
Date: 8 January 2019

NOTE : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach a thesis declaration letter.

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Computer Science (Computer Systems & Networking)

____

(Supervisor's Signature)

Full Name : PROF DR. VITALIY MEZHUYEV

Position : Professor

Date : 8 January 2019

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang or any other institutions.

____

(Student's Signature)

Full Name : AMELIA BINTI ABD RAHMAN

ID Number : CA15105

Date : 8 January 2019

MODEL FOR THE EVALUATION OF USABILITY OF SCIENTIFIC WORKFLOWS

AMELIA BINTI ABD RAHMAN

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Computer Science (Computer Systems & Networking) with Honours

Faculty of Computer Systems & Software Engineering

UNIVERSITI MALAYSIA PAHANG

JANUARY 2019

ACKNOWLEDGEMENTS

First of all, thanks to Allah s.w.t Who enabled me to conduct this research. I want to offer this endeavour to Allah for giving me strength, health and peace of mind in order to finish the research.

I would like to thank my supervisor, Prof Dr. Vitaliy Mezhuyev, who helped me in completion of the research. I am extremely grateful to him for his expert, valuable guidance and constant support to appreciate my works. I am very lucky to have a supervisor who constantly checking my works, responding to all my questions and giving advices and good comments to improve my works.

Next, I would like to express my special appreciation to my parents and my younger brother for their constant moral support extended with love and encouragement throughout the years which made it possible for me to complete the research.

Lastly, I would like to thank to my friends and those who helped me throughout the journey of the research.

ABSTRAK

Aliran kerja saintifik adalah sangat penting di kalangan komuniti saintifik, yang menyatukan data saintifik, analisis, tugas simulasi untuk menjalankan sebarang eksperimen pengiraan. Sistem aliran kerja saintifik sedia ada, yang terdiri daripada berpuluh-puluh langkah terperinci dan sistematik, membolehkan para saintis membuat dan memvisualisasikan proses saintifik tanpa banyak kesukaran. Dalam kemajuan sains, masih ada peningkatan yang dapat dibuat terutama dalam penyatuan sistem aliran kerja saintifik kerana masyarakat saintifik yang berbeza memerlukan notasi yang berbeza untuk model kebolehgunaan. Semasa aliran kerja saintifik berkembang dari masa ke masa, sukar bagi komuniti saintifik untuk mencari aliran kerja saintifik yang sesuai untuk tugas mereka. Dalam kajian ini, penilaian kegunaan aliran kerja saintifik sedang dilakukan kerana ia merupakan ciri utama untuk menentukan kejayaan sistem perisian. Kajian empirikal digunakan untuk penilaian kebolehgunaan, berdasarkan penciptaan soal selidik untuk membangunkan kaji selidik dan pemeriksaan hipotesis statistik. Pakej Statistik untuk Sains Sosial (SPSS) digunakan untuk menganalisis data selepas respon dikumpulkan. Untuk pengesahan kaji selidik, pengumpulan data daripada 101 responden yang terdiri daripada komuniti saintifik dan pemaju system aliran kerja saintifik dikumpulkan kemudian analisis kebolehpercayaan (Cronbach's Alpha), One Sample t Test, analisis faktor (KMO dan Bartlett's Test), ujian ANOVA digunakan. MATLAB digunakan untuk pembangunan kod bagi Alpha Cronbach. Hasil penyelidikan ini, yang merupakan model kegunaan aliran kerja saintifik yang dicadangkan, akan membantu menaik taraf sifat-sifat kualiti dalam sistem aliran kerja saintifik.

ABSTRACT

Scientific workflows are extensively important among scientific communities, which unified scientific data, analysis, simulation tasks to conduct any computational experiments. Existing scientific workflow systems, which consist of dozens of detailed and systematic steps, enable scientists to create and visualize scientific processes without much difficulty. Regardless of their potential for advancement in science, there are still improvement that can be made especially in unification of scientific workflow systems since different scientific communities required different notations to model. As scientific workflows evolved over time, it is also difficult for scientific communities to discover suitable scientific workflows for their task. In this research, evaluation of usability of scientific workflows is being done since it is a key characteristic of determining the success of any software system. Empirical study was applied for the usability evaluation, based on creation of questionnaire to develop survey and statistic hypotheses checking. Statistical Package for the Social Sciences (SPSS) was used to analyse the data after responses were collected. For validation of survey, the collection of data from 101 respondents of scientific community and developers of scientific workflow systems was collected then the reliability analysis (Cronbach's Alpha), One Sample t Test, factor analysis (KMO and Bartlett's Test), ANOVA test were used. MATLAB was use for development of coding for Cronbach's Alpha. Results of this research, which is the proposed model of usability of scientific workflows, will help to upgrade the quality attributes in scientific workflow systems.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA	Analysis of variance
CITC	Corrected Item-Total Correlation
RQ	Research question
SAS	Statistical Analysis System
SPSS	Statistical Package for the Social Sciences

CHAPTER 1

INTRODUCTION

1.1 Background of Study

Workflow management system is a specially designed scientific workflow system to construct and run a list of workflows in a scientific operation. Nowadays, scientific community are relying on scientific workflows to conduct analysis and discovery of experiment using distributed computing platforms. Scientific workflows system usually provides a visual programming frontend, which makes it easier to conduct experiments. Other than that, scientific workflows systems offer to users a friendly environment to build their own workflows as well as to execute them and view the results in the real-time.

Scientific workflows execution requires corresponding software tools and high-level mechanism as thousands of parallel tasks including large datasets of input. This may lead to longer duration of the time taken for data to be processed.

The research identifies the challenges to evaluate the usability of scientific workflows systems. Based on analysis of existing research, the model of usability will be proposed. The research collects data from survey and check it using methods of mathematical statistics through Statistical Package for the Social Sciences (SPSS). MATLAB is used for development of coding for checking validity (Cronbach's Alpha). The research is a way to find out the most important usability heuristics from scientific community. Consequently, the research aims to improve quality attributes in existing scientific workflows systems. Based on the proposed model, existing front-ends of scientific workflow systems may be improved.

1.2 Problem Statement

1. Current scientific workflows systems become more and more complex and their usability changes correspondingly.

2. To the best of our knowledge, there are no researches devoted to the evaluation of usability of scientific workflows.

3. Different scientific communities required different notations to model scientific workflows. Thus, unification of approaches is needed. Such unification and development of new workflow systems should follow the criteria of usability of scientific workflows.

1.3 Objectives

i. To analyse existing approaches and models for the evaluation of usability of software systems.

ii. To develop a model for evaluation of usability of scientific workflows.

iii. To evaluate proposed model on a case study with application of methods mathematical statistics.

1.4 Scope

i. The focus of this research are the models for evaluation of usability of scientific workflow systems.

ii. To collect data, the survey was developed and shared to scientific community through Web-based application (Google Docs).

iii. Questions in the survey use Likert scale to measure respondent’s opinion and point of view.

iv. The analysis of collected data includes the methods of mathematical statistics method using SPSS.

v. MATLAB is used for development of coding to check validity (Cronbach's Alpha).

vi. Users of proposed model are scientific community and developers of scientific workflow systems.

1.5 Significance

i. The research proposed a new model for the evaluation of usability of scientific workflows.

ii. Proposed model will improve the quality attributes in new scientific workflow systems.

1.6 Thesis Organization

The research aims to help scientific community to develop better scientific workflows based on quality attributes that were collected. This thesis consists of five chapters.

Chapter 1 presents the introduction on scientific workflows, problem statement, objectives for conducting the research, scope regarding the research, significance of this research and thesis organization.

Chapter 2 consists of systematic literature review on existing models of usability of software systems. This section also considers in detail the problems and solutions in current research.

Chapter 3 illustrates methodology of the research. It also shows approaches used in the research such as questionnaire development, survey sharing, and statistical data analysis obtained from respondent.

Chapter 4 contains results and findings of the research. This chapter discusses the result that have been collected based on survey conducted.

Chapter 5 discusses research results and gives brief conclusion. It also gives limitations as well as future work planning.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In recent years, the request for usable scientific workflow systems has increased particularly. For the most part, it is due to the revolution in users' perception of scientific workflow software systems. Over the last few years, the idea of usability has been interpreted in different quality models and research shows that the key component in the general quality of a software system is a usability. This chapter describes existing works related to this research. The literature review is based on collected information from scientific papers.

2.2 Models Used for the Evaluations of Usability of Scientific Workflows

These are several existing usability models proposed by experts.

2.2.1 Jakob Nielsen

Nielsen (1993) presented usability heuristics for the evaluation of usability attributes to check through the usability principles of the software system. The usability principles of the heuristics should be next utilized by the evaluators for examining a software interface. According to his classification, usability has five sub attributes, which are, learnability (easy to learn), efficiency (efficient to use), memorability (easy to remember), absence of errors (the relevance of catastrophic errors for applications, and satisfaction (pleasant to use) (Gupta, Ahlawat, & Sagar, 2015).

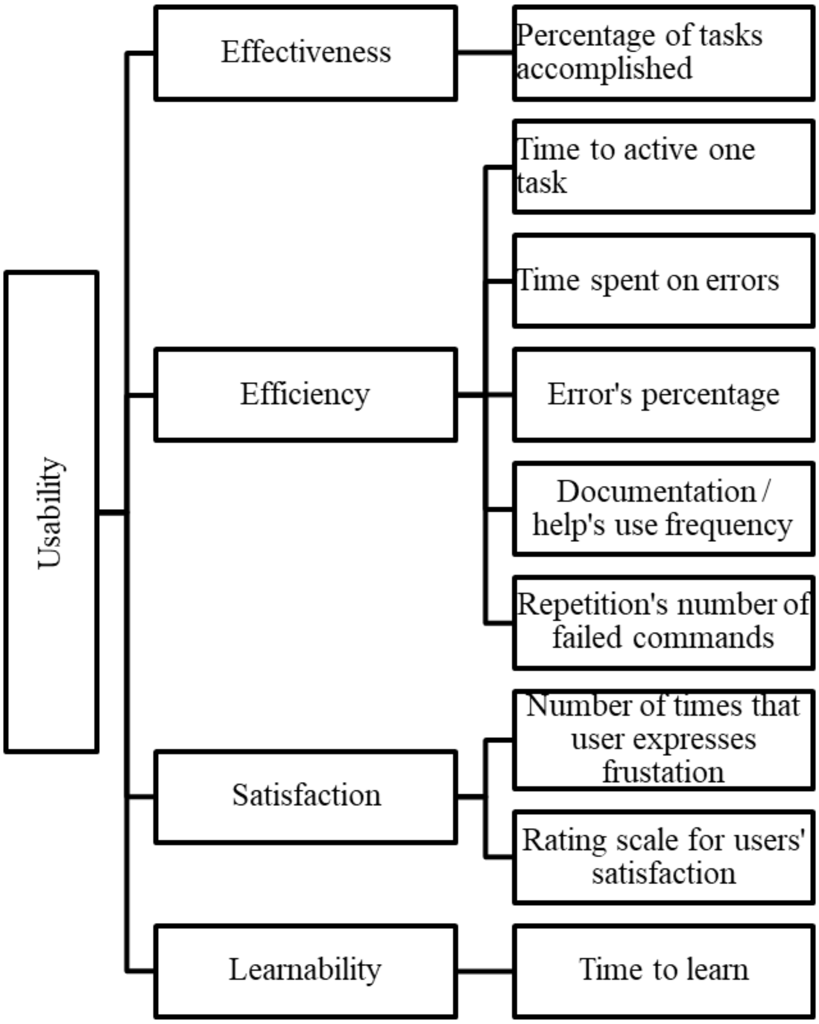


Figure 2.1 Nielsen's Usability Model

2.2.2 ISO 9126 – 1

ISO 9126 – 1 defines usability as "The capability of the software product to be understood, learned, used and be attractive to the user, when used under specified conditions." The ISO 9126 series of the standards (ISO 9126, 2001 -2003) divides software quality into six general categories of characteristics, which are usability, effectiveness, functionalities, reliability, maintainability and portability. The ISO 9126 objective is to offer a framework of software quality evaluation. ISO/IEC 9126, defines a quality model rather than specify quality requirements for software, which can be applied to every kind of software. This ISO standard consists of user's view and includes the concept of quality in use (Abran, Khelifi, Suryn, & Seffah, 2003).

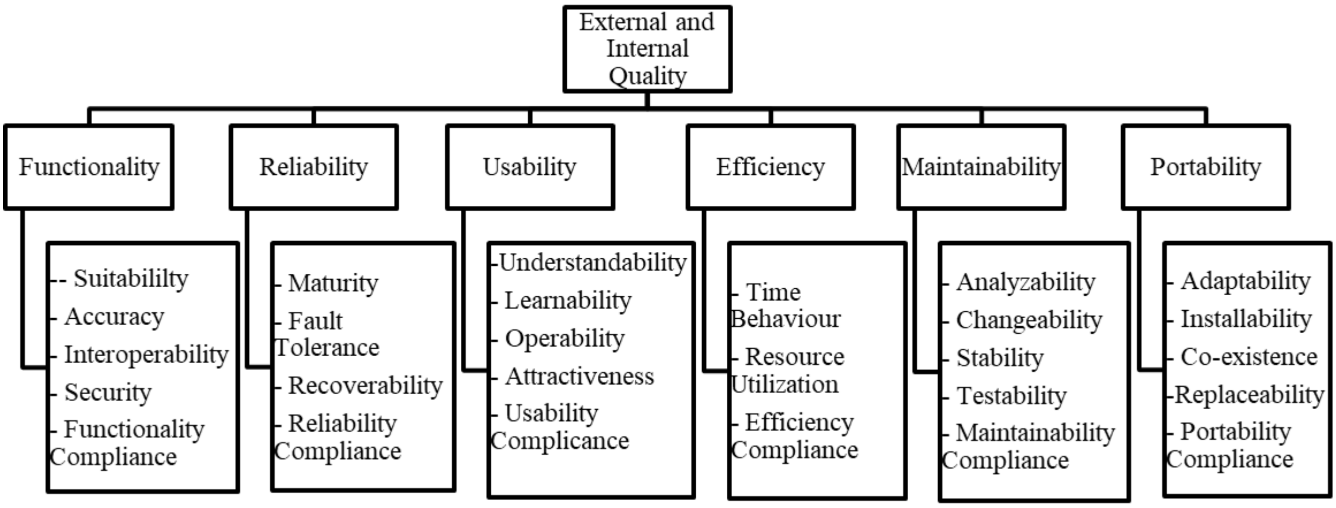


Figure 2.2 ISO 9126 – Software Quality

2.2.3 FURPS (Functionality, Usability, Reliability, Performance and Supportability)

FURPS model was presented at first by Robert Grady and extended by IBM Rational Software (Abiud & Mbugua, 2016). FURPS which stands for functionality, usability, reliability, performance and supportability.

Functionality in FURPS model consists of feature sets, capabilities and security, while for usability it consists of various human factors such as consistency in user interface, aesthetics, online and context sensitive help, wizards and agents, user documentation and training materials. Other than that, reliability in FURPS includes recoverability, predictability, accuracy, frequency and severity of failure and mean time between failures (MTBF). Performance in FURPS focuses on functional requirements such as speed, efficiency, availability, accuracy, throughput, response time, recovery time, and resource usage. Lastly, supportability contain testability, extensibility, adaptability, maintainability, compatibility, configurability and serviceability (Gupta et al., 2015).

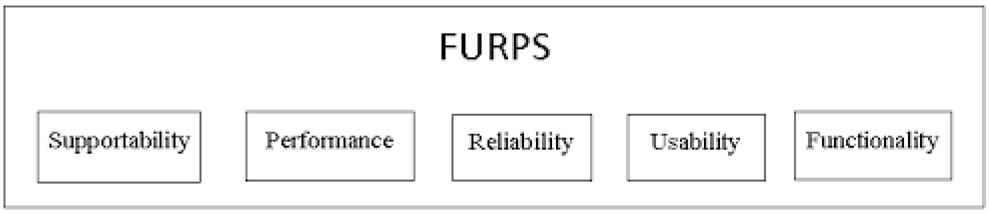


Figure 2.3 FURPS

2.3 Comparison of Usability Attributes in Various Models

The table 2.3.1 shows the comparison of usability attributes based on different standards and models.

Table 2.1 Comparison table of usability properties of various models

Models / Comparison	**Nielsen**	**ISO 9126 - 1**	**FURPS**
Definition	Usability consists of quality attributes that provides easy to use interface to user and consists of methods to improve ease of use during design	Usability is a product quality for evaluation of a software quality	Representing a model of software quality which includes functional and non-functional requirements
Usability Attributes	Learnability, efficiency, memorability, errors and satisfaction	Functionality, reliability, usability, efficiency, maintainability and portability	Functionality, usability, reliability, performance and supportability
Advantages	Provide quick feedback to designers by assigning correct heuristics	Develop a common understanding of project's priorities based on goals and objectives	Provide technique to validate prioritised requirements based on user's needs.

2.4 Conclusion

This chapter analyses existing approaches and models for evaluation of software usability. These approached will be used for development of the model of usability of scientific workflow systems in chapter 4.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The purpose of this chapter is to discuss methodology, which is used as a development guideline to find out whether the research is done accordingly. Methodology is the theoretical framework of the research that consists of development stages. It is an important phase to discover and evaluate current knowledge repeatedly until problems and solutions are finally defined. Any related information or data were collected, analyzed and presented to verify the objectives in the research. In this research, the methodology focuses on how to evaluate the usability model of scientific workflows. There are four main phases included in the research project, which are systematic literature review, theoretical study, empirical study and lastly statistical study. In addition, there are also details of methods and tools, which are the software and hardware used, related to the research project.

3.2 Methodology Flowchart

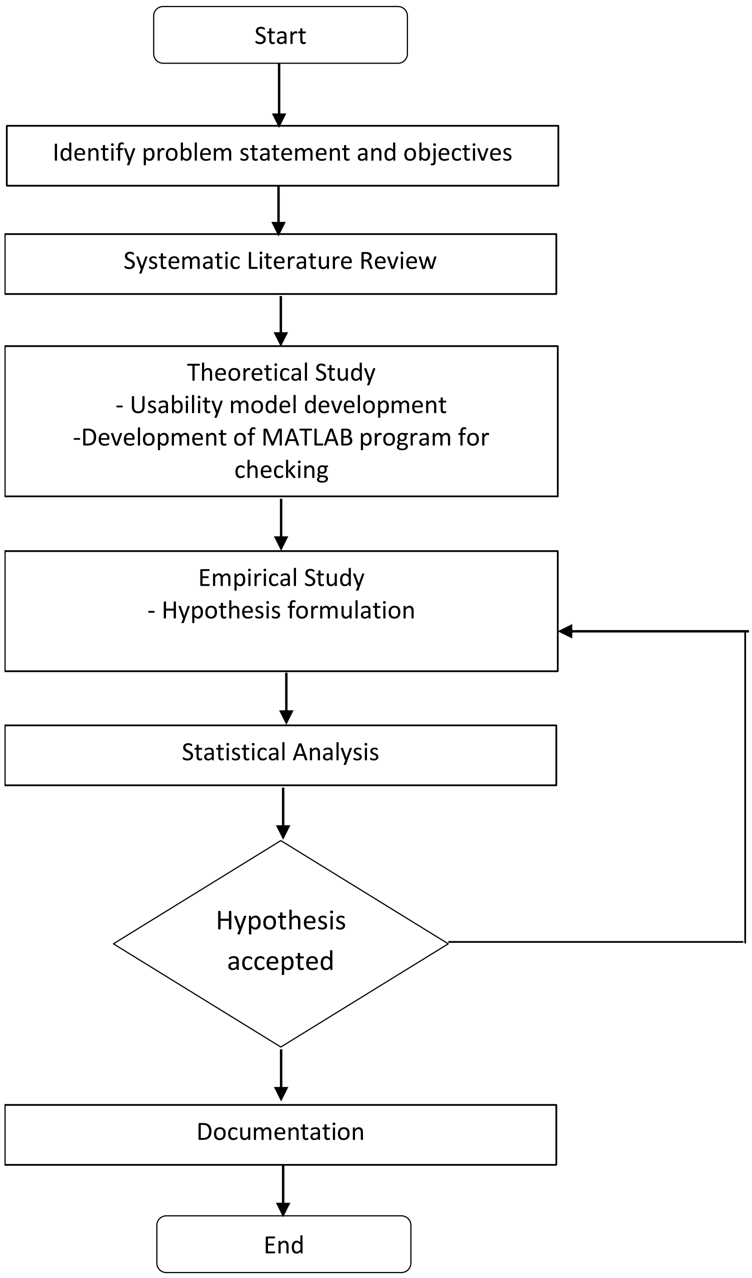


Figure 3.1 Flowchart of the methodology framework

3.3 Research Activities

Below are the steps in the research, which are systematic literature review, theoretical study, empirical study, statistical analysis.

3.3.1 Systematic Literature Review

Systematic literature review is the first action need to be done to identify research background, problem statement, objectives and scope of the research project. Information related to existing models for the evaluation of usability of scientific workflows were analyzed and comparisons have been made of existing models. The aim of systematic literature review is to address problems by determining, classifying and link all important data findings from the related research sources. Systematic literature review includes establishing problem statements from existing research, finding out similar approaches and their drawbacks(Siddaway, 2014). Then, it will allow to develop the model for the evaluation of usability of scientific workflows.

3.3.2 Theoretical Study

Theoretical study includes the analysis of information based on compilation or observation. Other than that, theoretical study also consists of short summary on new or existing concept at the first part of review. In this research, theoretical study is an adaptation of usability of existing models and evaluation methods, which is usability inquiry. Usability inquiry which is usability methods will collect data from respondents' opinions through questionnaire or survey which then contribute to the development of usability model, which will be developed based on existing models of scientific workflows usability.

3.2.1.1 Usability Model Development

In usability model development, there are methods used for usability evaluation. The aim of the usability evaluation is to gain access of the tool's functionality, to find out the effect on users and to identify any problems on the applications. There are three main types of usability evaluation methods, which are testing, inspection and inquiry.

However, for the research, usability inquiry information is the most related to scientific community users' preferences, understanding and requirements in which it is collected in a written form of questions (Yeltayeva, 2012). Questionnaire is the method of usability inquiry where it does not consist of study of user interface but gathers users' opinions. Questionnaire is also considered as a tool used to develop survey for the research. There are series of questions prepared which are divided into demographic, objectives and usability predictors for models of evaluations of scientific workflows based on Likert scale.

3.2.1.2 Development of coding

Development of MATLAB coding will be used to check hypothesis. Below is the interface of MATLAB coding that will be used.

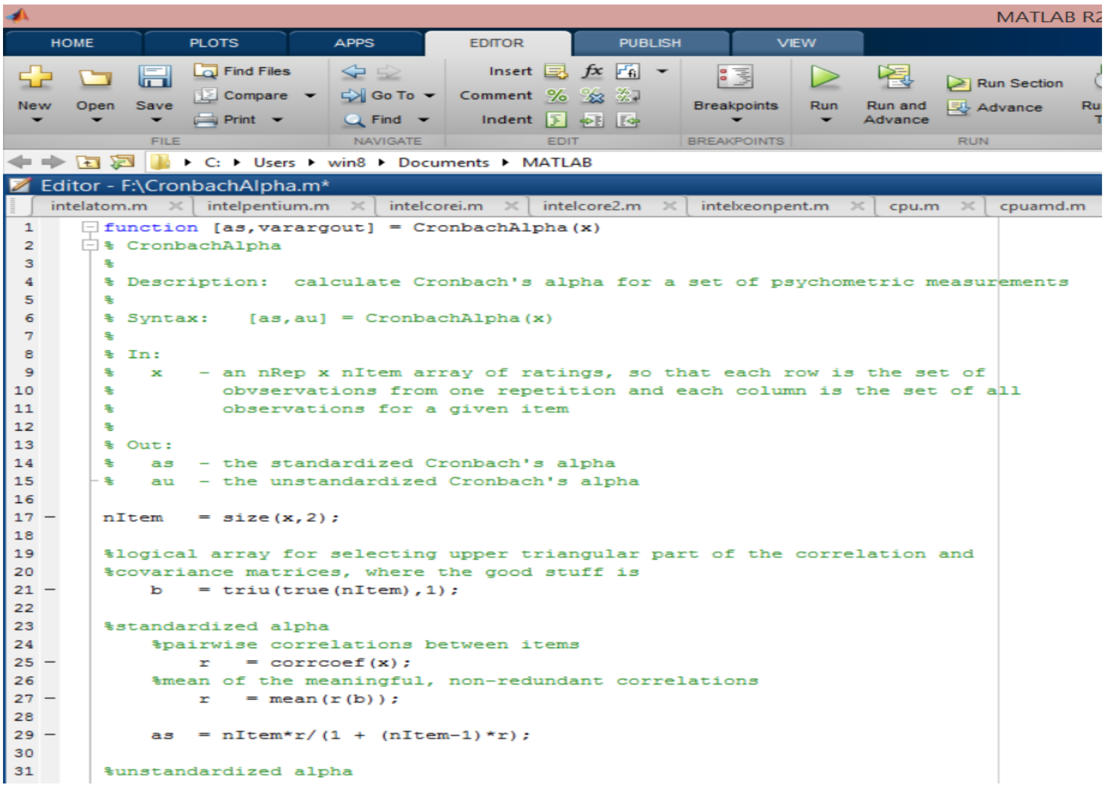


Figure 3.2 MATLAB Coding

```
%unstandardized alpha
if nargout>1
    %variance/covariance matrix
        vc   = cov(x);
    %mean variance (variances are along the diagonal)
        v    = mean(diag(vc));
    %mean covariance, not including variances
        c    = mean(vc(b));

    varargout{1}     = nItem*c/(v + (nItem-1)*c);
end
```

Figure 3.3 MATLAB Coding

3.3.3 Empirical Study

Empirical study describes knowledge, which are observed and experienced using an empirical evidence. Empirical evidence can be divided into two forms, which are quantitative and qualitative form. Qualitative study usually aims to provide what is happening inside the data, point out the features of a new aspect and define status of existing situation. Hypotheses are created over research problem. Quantitative study includes methods of collecting data such as research questions, which are proved by statistical experiments. After that, a list of indicators is prepared to align with distinct research questions. Then, a survey containing set of questionnaires will be prepared. There will be three sections in the questionnaire, which are demographic, objective, and usability predictors for usability models of scientific workflows based on Likert scale. In order to confirm the validation of the questionnaire, pilot study is conducted under the supervision of lecturer. Pilot testers will give any recommendations that are needing to be modified in the questionnaire. Finally, the questionnaire is distributed to scientific community who are interested on filling the questionnaire. The questionnaire is one of the methods of usability evaluation.

3.3.4 Statistical Analysis

Data that are collected in the research project will be analysed through the method of statistical analysis. The collected data will be analysed using reliability analysis, factor analysis test, one sample t test and ANOVA test. SPSS is the mathematical software used to perform analysis process. The generated data from the test will be used to form a conclusion, which later it will be decided whether the hypothesis is accepted or not accepted.

3.3 Hardware and Software Requirement

The hardware required in the research project is any physical devices or hardware components used in system development. To perform all documentation jobs for the research, a laptop is used along with internet access function.

Software specification is needed to perform research documentation and development. Microsoft Word 2013 is used to perform documentation work and Project Plan 365 is used to construct Gantt chart. Other software specification used in the research project are Google Form which used to develop survey form for data collection and MATLAB which is used for coding and SPSS to analyze collected data results through reliability analysis and ANOVA test.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

The aim of this chapter is to develop the model and discuss the results for this study. This chapter evaluates and describes the results that found after data was collected from survey and check it using methods of mathematical statistics through Statistical Package for the Social Sciences (SPSS). In this study, 101 responses of survey form were collected, while 1000 survey form were distributed to the scientific community. The goal of this chapter is to prove or validate the proposed model of usability of scientific workflow systems and to analyse data using reliability analysis, factor analysis test and Analysis of Variance (ANOVA) test. Reliability test is done by using Cronbach's alpha through SPSS. Factor analysis was done which filtered or simplified the data. One Sample T Test is used to test the most significance usability predictors. ANOVA test is used to test the relationship between the independent and dependent variables. Lastly, the data analysis is supported by MATLAB code.

4.2 Proposed Model

Many international standards and usability models can be found through literature review, which describe usability with different attributes in nonhomogeneous manner. This can cause confusion among experts for its application and usage. The inconsistent and varies of usability model is creating major challenge for evaluation of usability of scientific workflows. Researchers do not decide yet which usability model can present scientific workflow systems since there are many usability's definition and lack of information on usability factors. So, it is important for usability model to be generic. This proposed model gives hierarchical based usability model which can unify various existing models for scientific workflow systems.

There are 20 usability predictors being proposed after reviewing existing models of usability. The usability attributes that are included in our proposed model:

a. **Simplicity**, it is to make a software interface simple, common tasks simple to do, communicating in specially developed language and providing good shortcuts that are meaningfully related to complex procedures.

b. **Aesthetics**, it contains pleasing visual design such as color, shape, symmetry and proportion.

c. **User Satisfaction**, it is focuses on user's responses, feelings of users such as comfort and freedom of using the software.

d. **Learnability**, it focusses on how easy the system is for novice users, but usability in general might also include advanced features for experienced users (Informatics, 2008).

e. **Consistency in User Interface**, it is defined as if the system is memorable for user as they should not have to wonder whether different words, situations, or actions is different thing.

f. **Operability**, it is a measure about the capability of the system to operate well.

g. **Efficiency**, it is usability attribute that enables user to produce desired results according to investment of resources (first of all, time).

h. **Performance**, it is the measurement such as, task performance time.

i. **Usefulness**, it is defined as useful result (output) that is produced from interaction of a user with the software.

j. **Reliability**, it is the usability quality of being trustworthy and be able to provide consistency.

k. **Flexibility**, it is about allowing users to tailor actions and interface.

l. **Supportability**, it is the capability of software product to support any operations of a user.

m. **Universality**, it is a measure of software product which can support a user to use the system from different domain of application.

n. **Reusability,** means software product can provide workflows be reused in another application.

o. **Support of Data Transformation**, means software product can convert data to different type of values.

p. **Support of Different Formats of the Output**, it is defined as software product can transform output data to from a source to another format.

q. **Support of Data Analysis**, it is defined as the process of inspecting, analyzing, modeling data helping decision making.

r. **Backup and Recovery of Data**, in case of any error, the software product should provide support for recovery of data which allows the system be reliable

s. **Standard Language or Format**, it is the standard scientific workflow language or standards which any users from different background of domain of application can use.

t. **Context Sensitive Help, Documentation and Training Materials**, it is an online help or information resources provided for a user in any associate situation.

Figure 4.1 Proposed Model

4.3 Demographic Background

Table 4.1 shows the analysis on respondents' age, gender, country, type of employment, educational level, type of project and domain of application of scientific workflows. The respondents for the survey are scientific community and developers of scientific workflow systems. The total number of respondents participated in this survey is 101 respondents, which includes 85 males (84.2%) and 16 females (15.8%). The respondents' age varied from 18 to above 45 years old with majority of 42 respondents' age (41.6%) range from 36 to 45 years old. Most of the respondents were from United States followed by Brazil, India, United Kingdom. The type of employment was checked for academia and industry. Majority of respondents were from academia which is 99 respondents (98.0%) while only two respondents (2.0%) were from industry. Educational level of respondents comprises of 83 Doctorates (82.2%), 13 Masters (12.9%), four Bachelors (4.0%) and one Undergraduate (1.0%). Mostly, the type of project consists of 66 Other Research Projects (65.3%), 31 PhD Projects (30.7%), three Undergraduate Projects (3.0%) and one Industrial Projects (1.0%). The main domains of application of scientific workflows is a Computer Science (30 responders, 31.31%), Physics (20 responders, 20.2%), Biology (18, 18.18%), Chemistry and Medicine (12, 12.12%), Engineering (7, 7.07%) and one in Business (1.01%).

Table 4.1 Demographic Characteristics of Respondents

Demographic Characteristics		**No. of Respondents, n**	**Percentage,%**
Age	18-25	3	3.0
	26-35	31	30.7
	36-45	42	41.6
	Above 45	25	24.8
Gender	Male	85	84.2
	Female	16	15.8
Country	United States	19	18.8
	Brazil	14	13.9
	India	12	11.9
	United Kingdom	10	9.9
	China	7	6.9
	France	6	5.9
	Spain	6	5.9
	Italy	4	4.0
	Germany	3	3.0
	Saudi Arabia	3	3.0
	Netherlands	3	3.0
	Australia	2	2.0
	Iran	2	2.0
	Russia	2	2.0
	Austria	1	1.0

	Czech Republic	1	1.0
	Egypt	1	1.0
	Norway	1	1.0
	Poland	1	1.0
	Tunisia	1	1.0
	Ukraine	1	1.0
	Venezuela	1	1.0
Type of Employment	Academia	99	98.0
	Industry	2	2.0
Education Level	Undergraduate	1	1.0
	Bachelor	4	4.0
	Master	13	12.9
	Doctorate	83	82.2
Type of Project	Undergraduate Project	3	3.0
	Master Project	0	0
	PhD Project	31	30.7
	Other Research Project	66	65.3
	Industrial Project	1	1.0
Domain of Application	Computer Science	31	31.31
	Physics	20	20.2
	Biology	18	18.18
	Chemistry	12	12.12
	Medicine	12	12.12
	Engineering	7	7.07
	Business	1	1.01

4.4 Data on Usability of Scientific Workflow Systems

Figure 4.2 to Figure 4.21 shows the data collection result of each usability predictors of scientific workflow systems. The number of 1,2,3,4 and 5 from Likert scale indicates the levels of agreement (Strongly Disagree, Disagree, Neutral, Agree, Strongly

Agree).

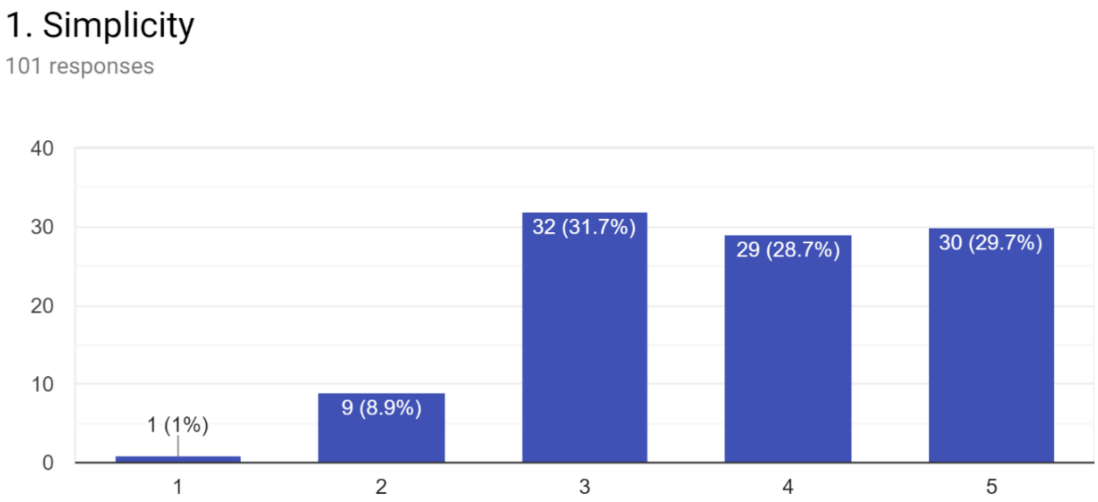


Figure 4.2 Evaluation of simplicity criteria

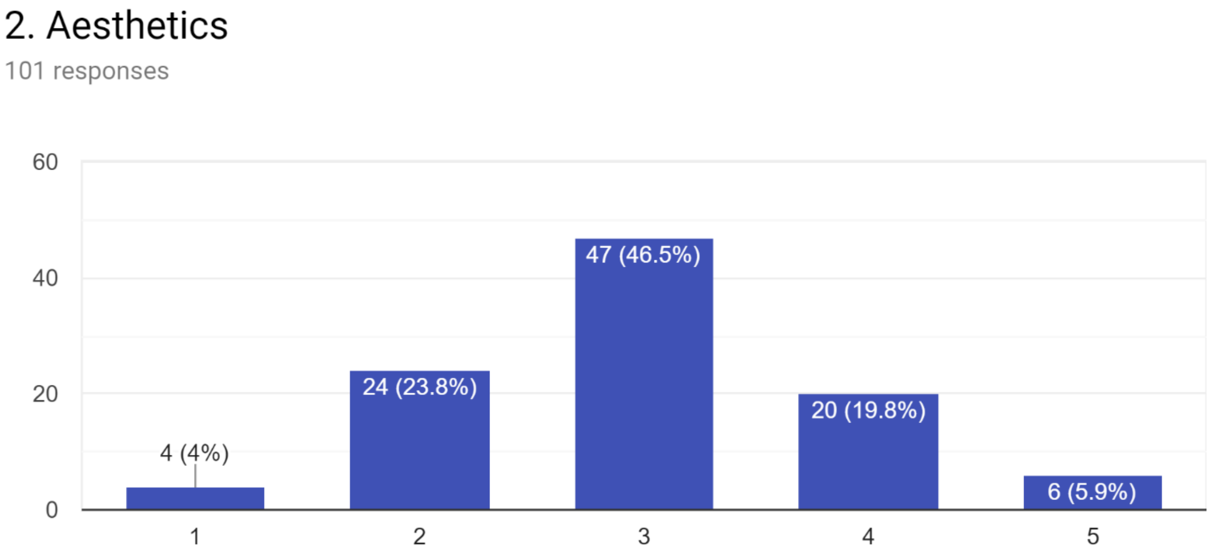


Figure 4.3 Evaluation of aesthetic criteria

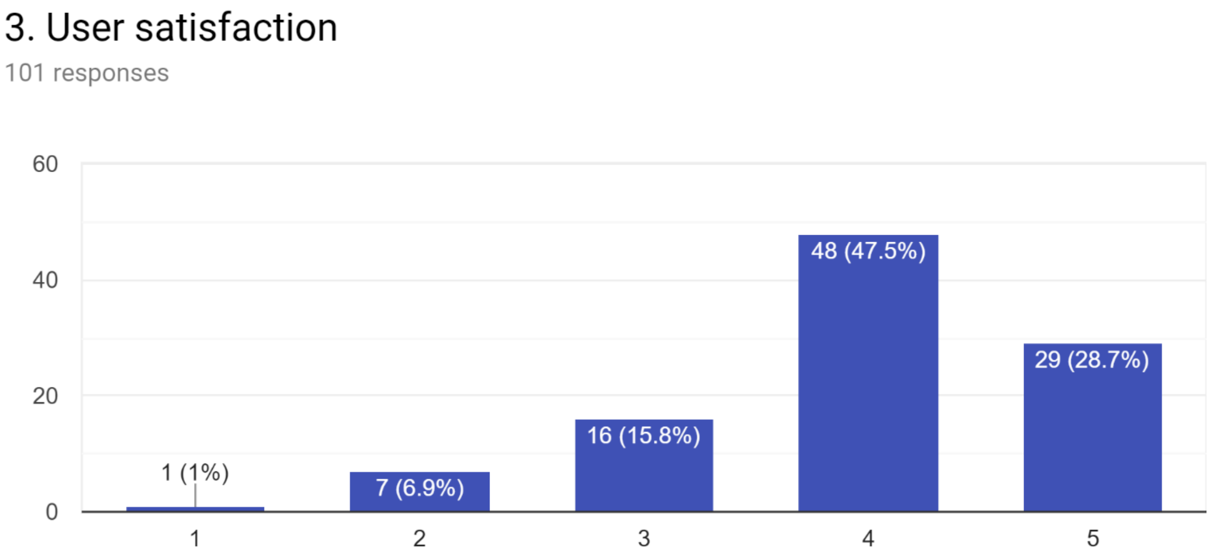


Figure 4.4 Evaluation of user satisfaction criteria

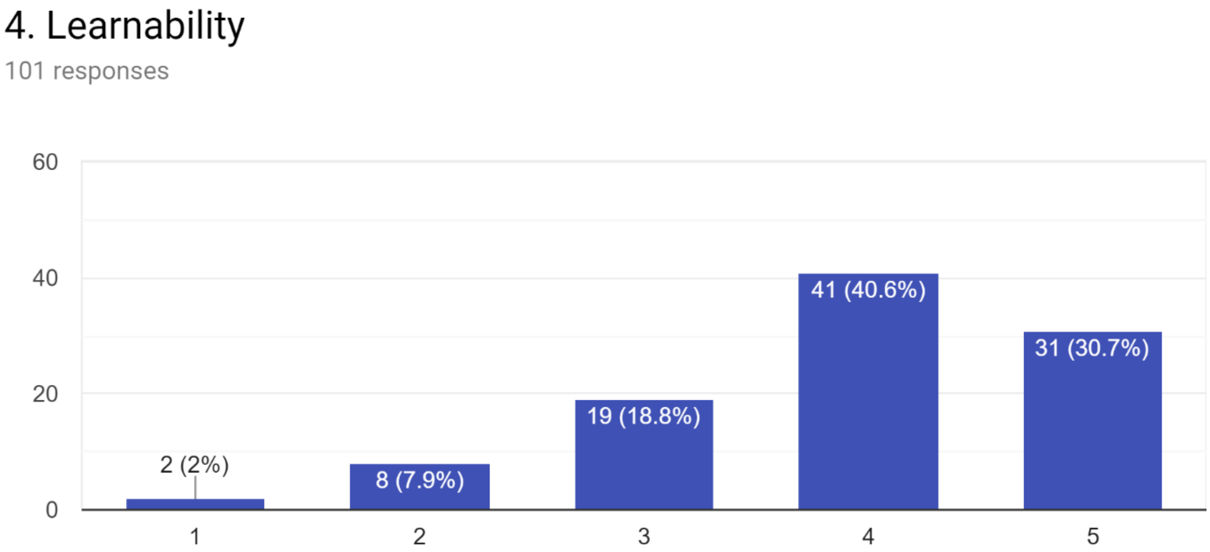


Figure 4.5 Evaluation of learnability criteria

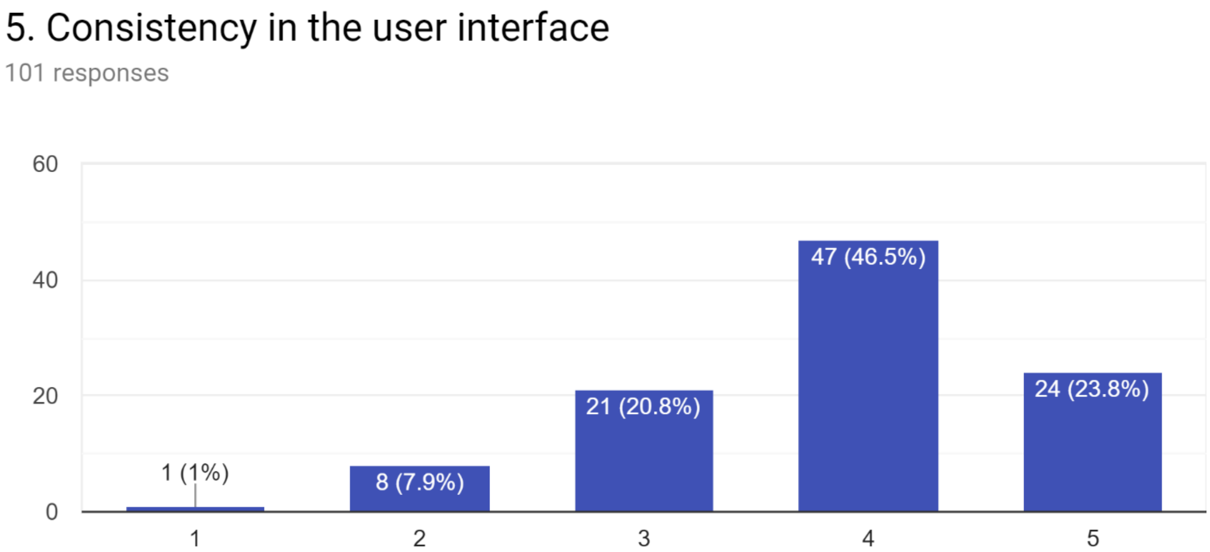


Figure 4.6 Evaluation of consistency in the user interface criteria

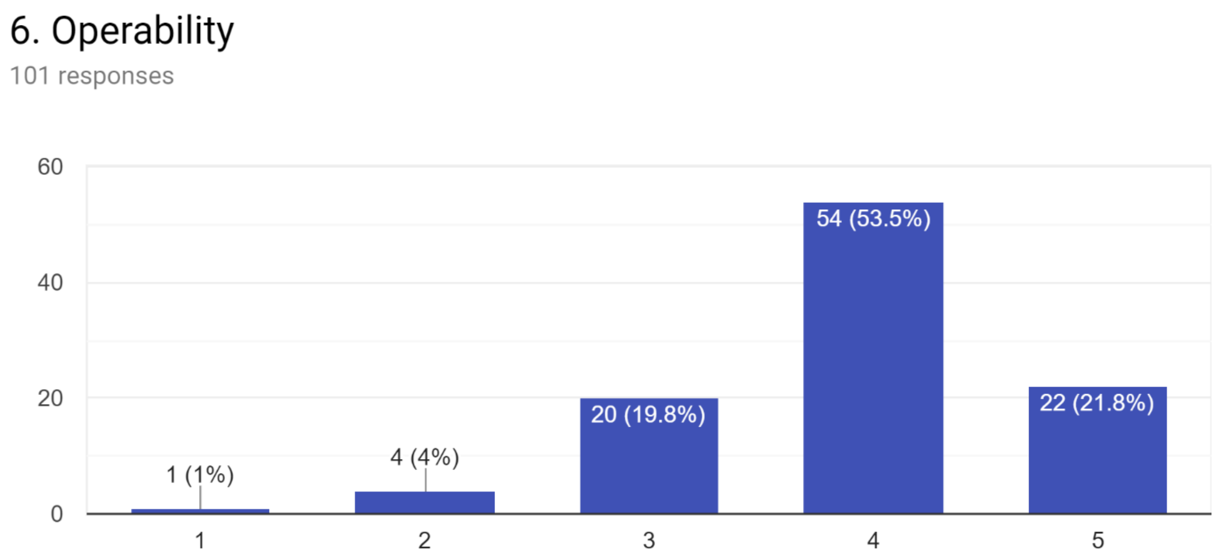


Figure 4.7 Evaluation of operability criteria

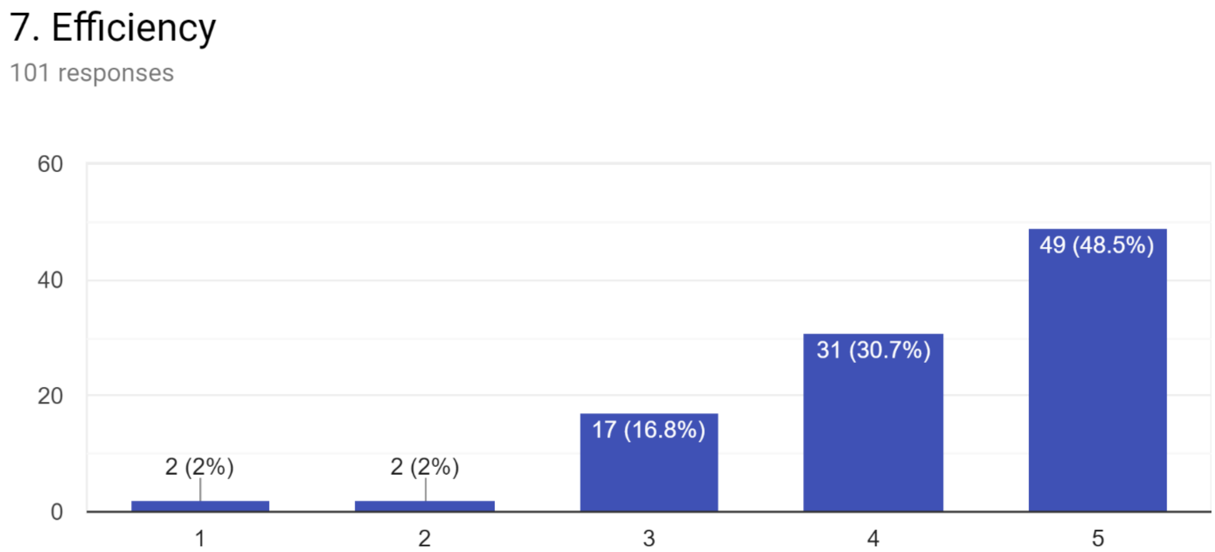


Figure 4.8 Evaluation of efficiency criteria

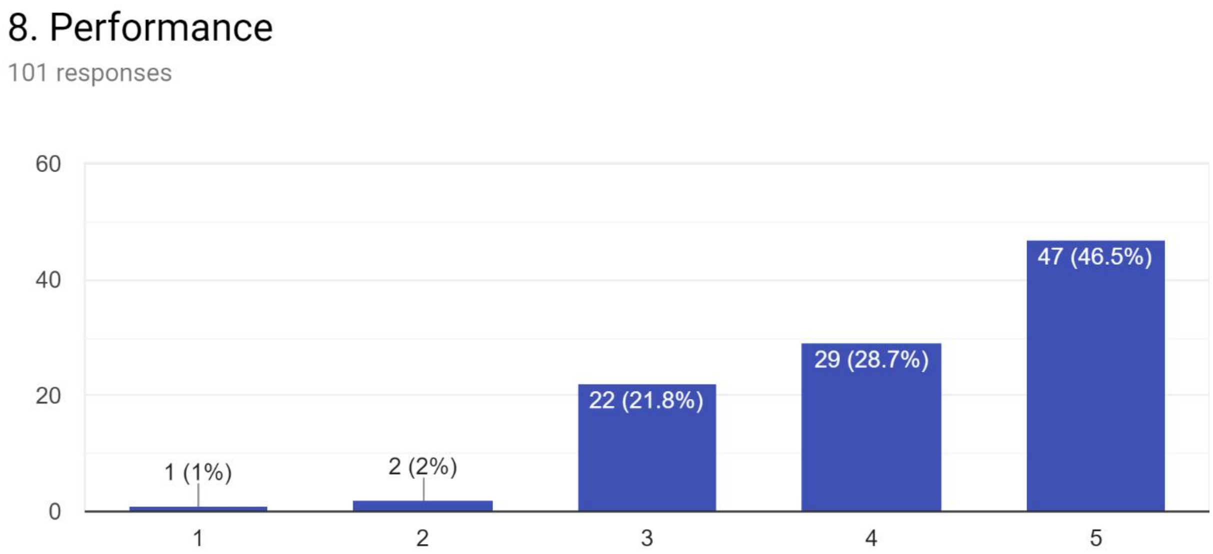


Figure 4.9 Evaluation of performance criteria

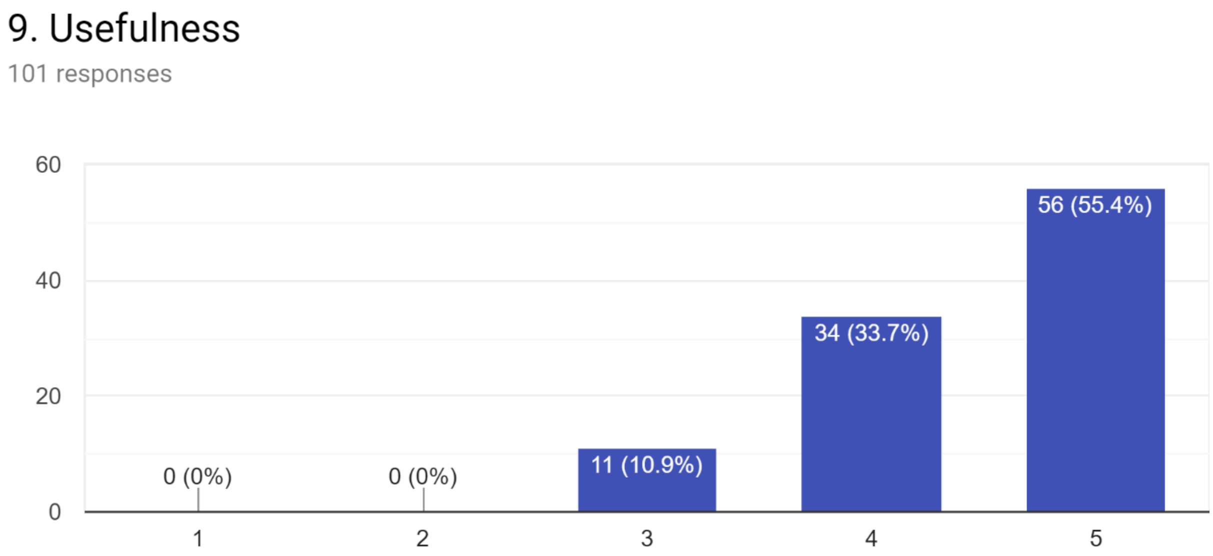


Figure 4.10 Evaluation of usefulness criteria

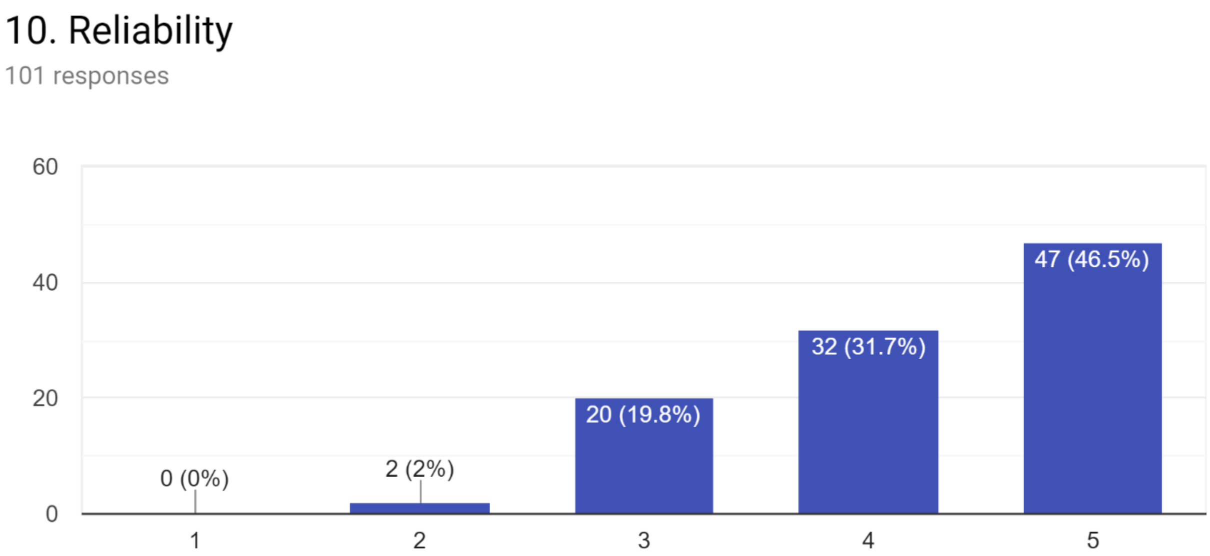


Figure 4.11 Evaluation of reliability criteria

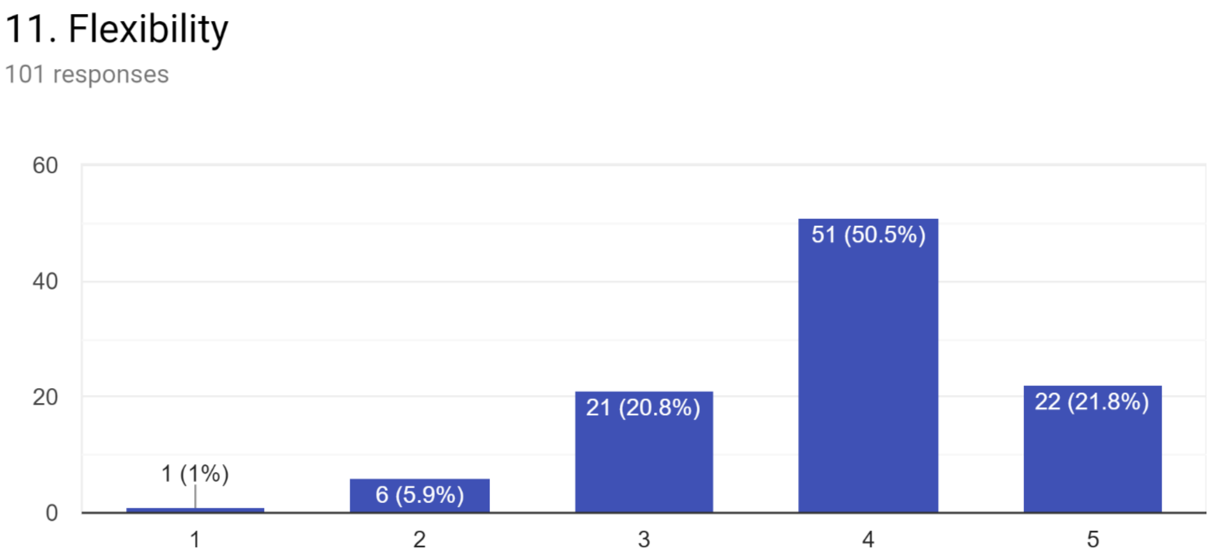


Figure 4.12 Evaluation of flexibility criteria

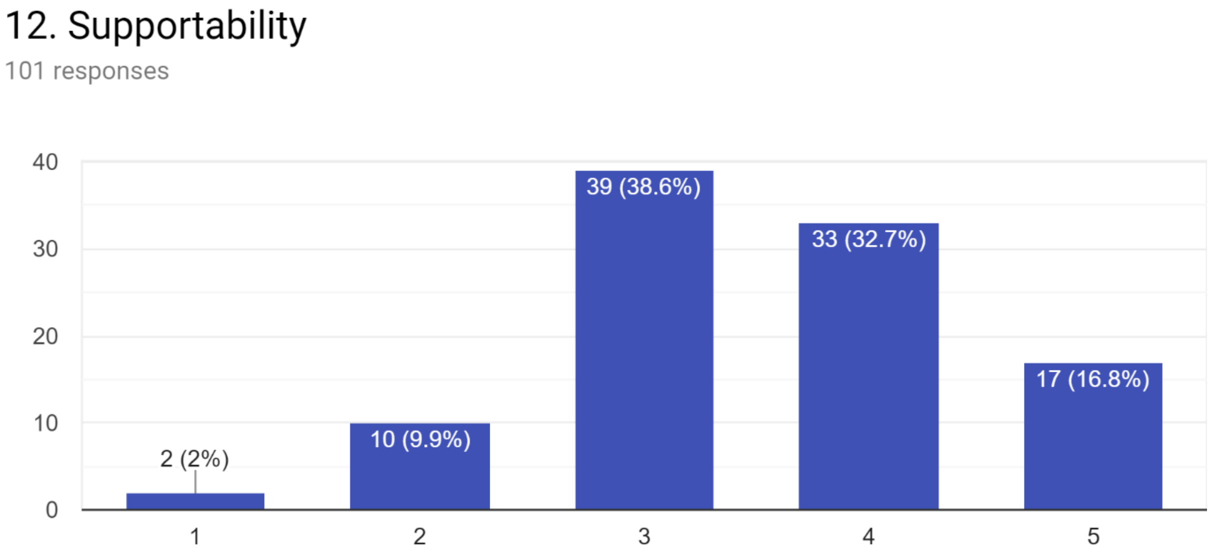


Figure 4.13 Evaluation of supportability criteria

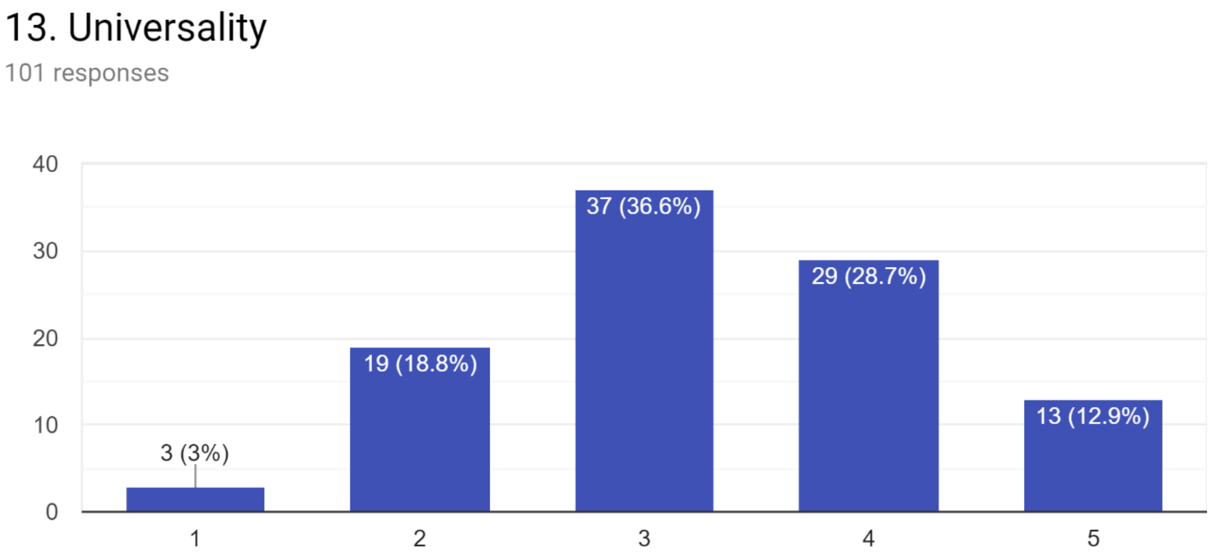


Figure 4.14 Evaluation of universality criteria

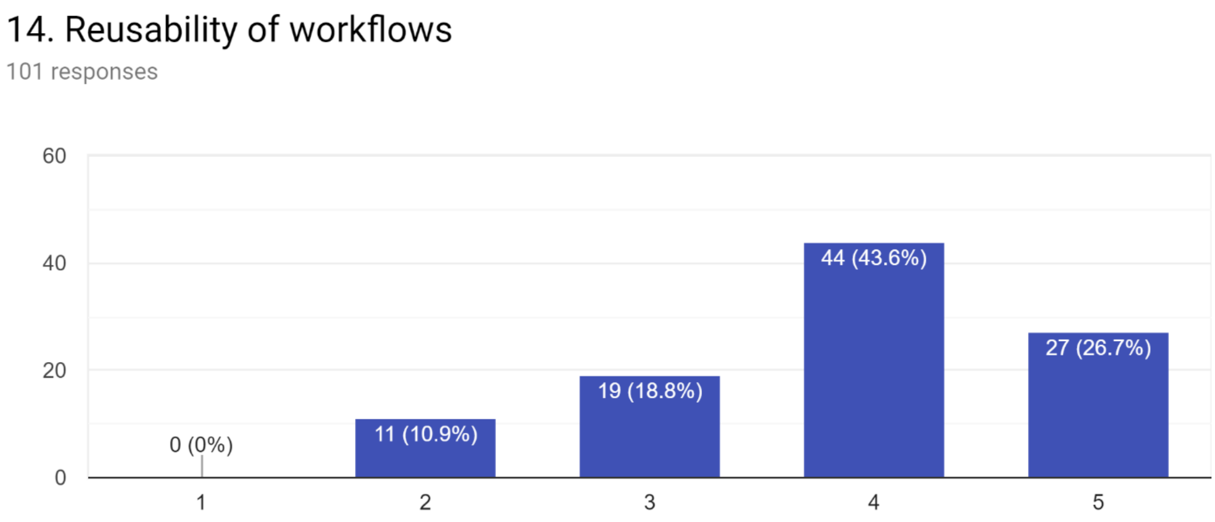


Figure 4.15 Evaluation of reusability of workflows criteria

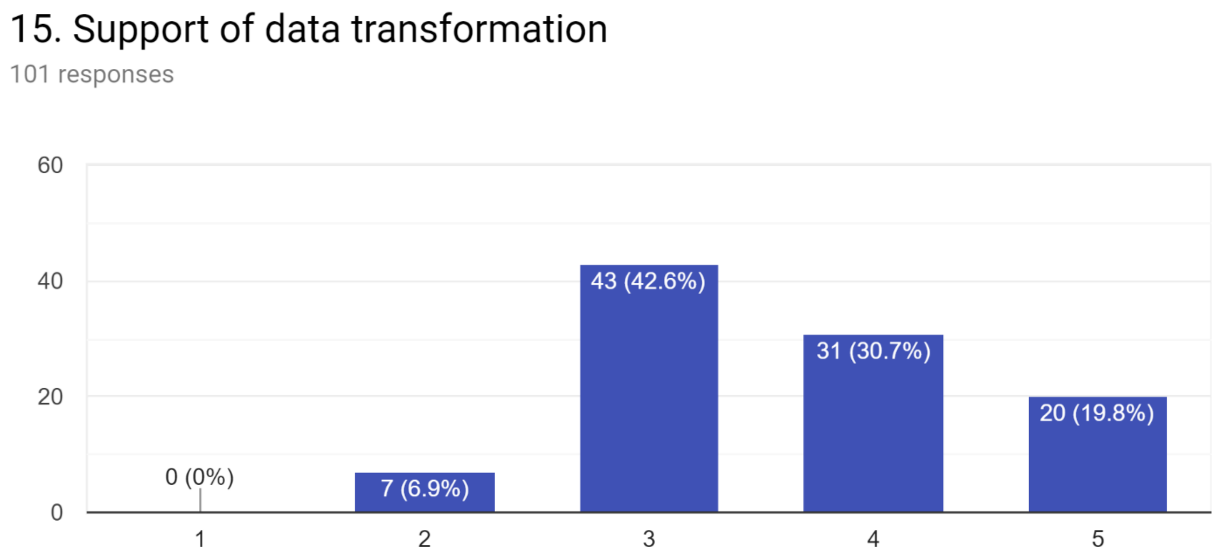


Figure 4.16 Evaluation of support of data transformation criteria

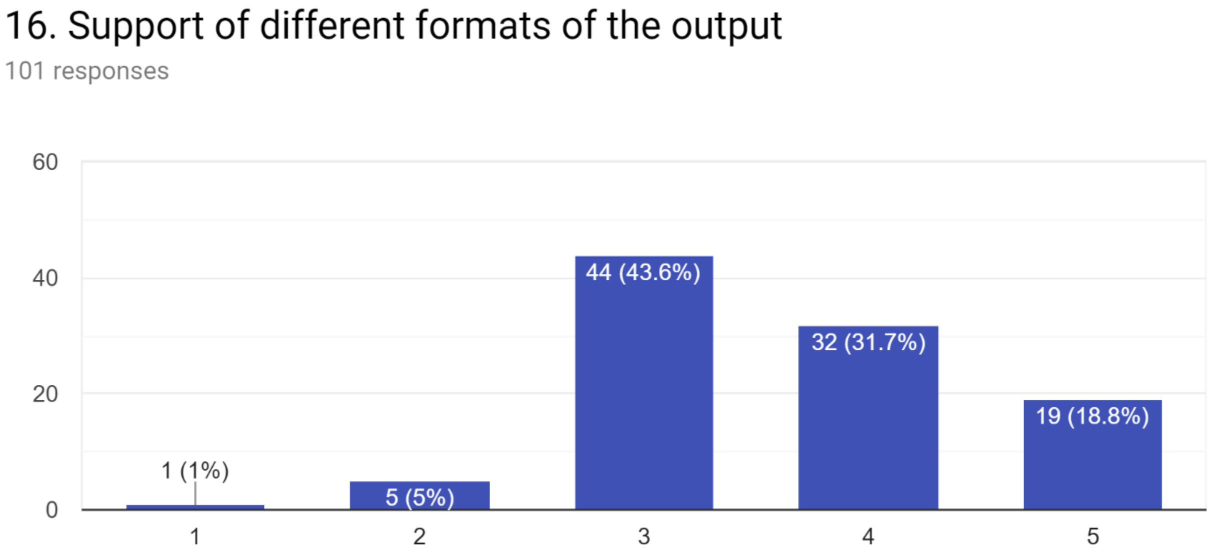


Figure 4.17 Evaluation of support of different formats of the output criteria

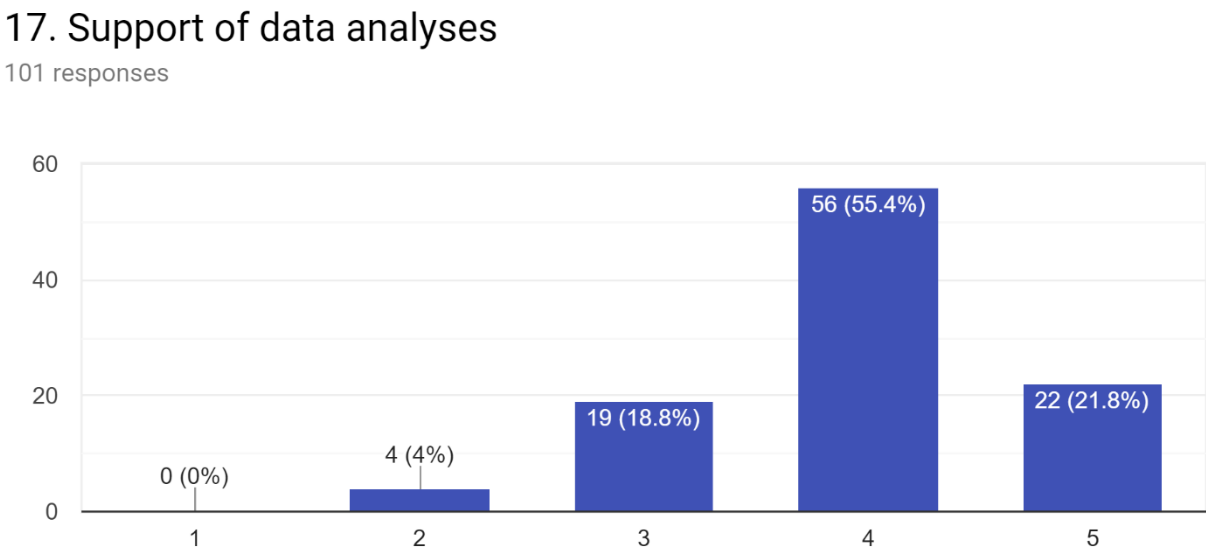


Figure 4.18 Evaluation of support of data analysis criteria

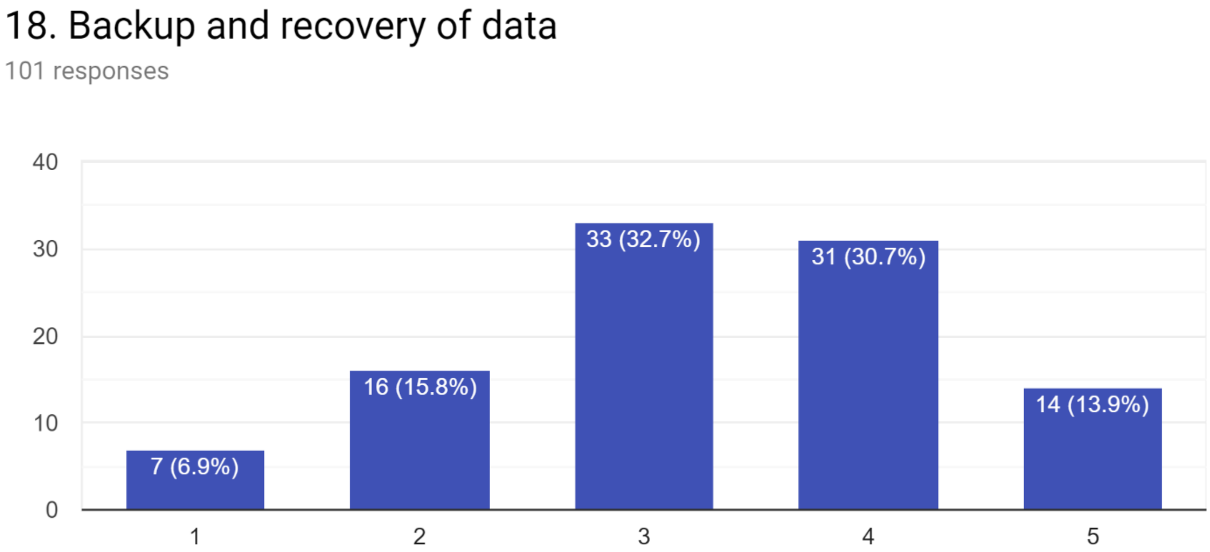


Figure 4.19 Evaluation of backup and recovery of data criteria

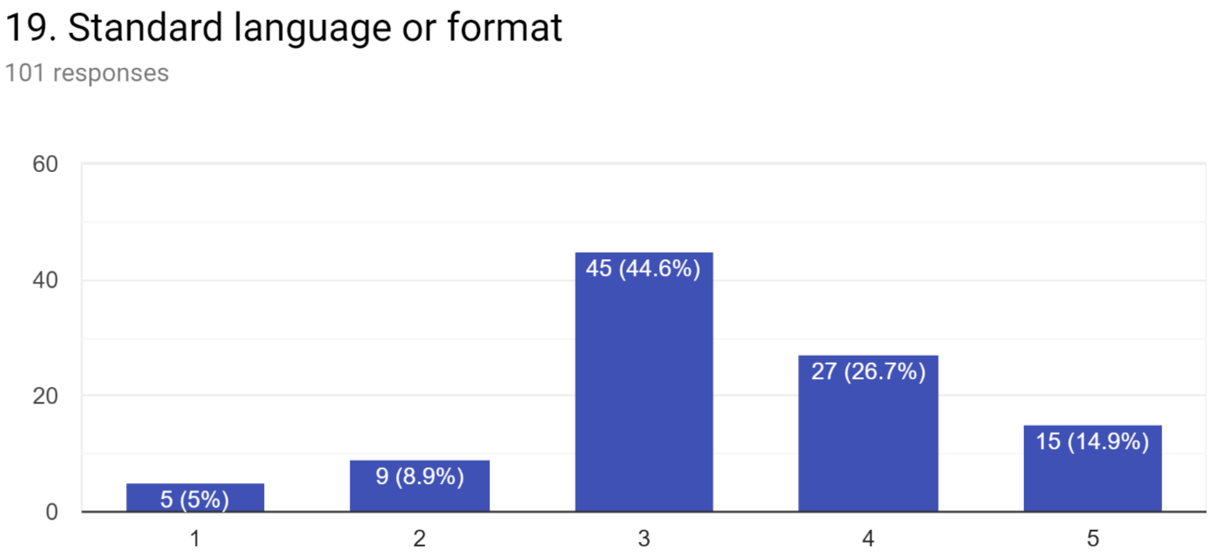


Figure 4.20 Evaluation of standard language or format criteria

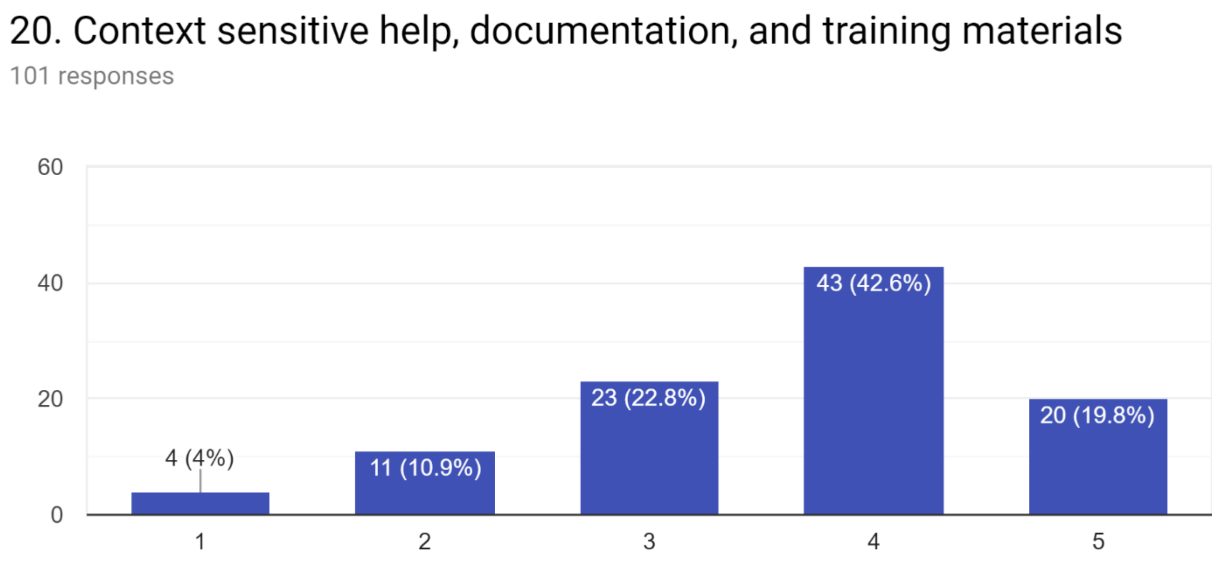


Figure 4.21 Evaluation of context sensitive help, documentation and training materials criteria

4.5 Reliability Analysis

The reliability test was done by Cronbach's alpha which is a measure of internal consistency on how closely related a set of items in a group. It used to evaluate reliability which is commonly used in a survey or questionnaire that form a scale such as, Likert scale. According to the general principle of Cronbach's alpha is that the reliability coefficient value of 0.70 or higher is recognized as acceptable in most scientific research situations while 0.80 or higher is better and 0.90 or higher is best. Other than that, if Corrected Item-Total Correlation (CITC) exceeded 0.3, this suggests that the constructs had adequate reliability. If the obtained Cronbach's alpha value is lower 0.70, the column labelled "Cronbach's alpha if item deleted" is required to be checked in the output table. The column labelled "Corrected Item-Total Correlations" is to be examined to identify on which variable is needed to be excluded to increase the alpha value. The variable with CITC value lower than 0.3 is the variable that would increase the alpha value by deleting them.

Based on Table 4.2, "N of Items" indicated that the total number of items which is being tested consist of 20 usability predictors and it is revealed that the result of the Cronbach's Alpha value (α) is 0.916 and based on Table 4.3, the Cronbach's Alpha if Item Deleted (CITC) of all the variables are higher than 0.3. So, we can that the reliability coefficient is considered acceptable.

Table 4.2 Reliability Statistics of Usability Predictors

Reliability Statistics

Cronbach's Alpha	N of Items
.916	20

Table 4.3 Item - Total Statistics of Usability Predictors

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Simplicity	71.73	116.338	.637	.910
Aesthetics	72.50	121.032	.463	.914
User satisfaction	71.54	119.410	.556	.912
Learnability	71.60	115.662	.680	.909
Consistency in the user interface	71.66	116.666	.695	.909
Operability	71.59	117.804	.722	.909
Efficiency	71.29	121.887	.410	.915
Performance	71.33	121.142	.461	.914
Usefulness	71.06	123.656	.465	.914
Reliability	71.28	119.842	.584	.912
Flexibility	71.64	120.552	.525	.913
Supportability	71.98	115.560	.718	.908
Universality	72.21	117.106	.595	.911
Reusability of workflows	71.64	118.812	.563	.912
Support of data transformation	71.87	119.973	.543	.912
Support of different formats of the output	71.88	119.826	.550	.912
Support of data analyses	71.55	122.350	.498	.913
Backup and recovery of data	72.22	117.512	.519	.913
Standard language or format	72.13	117.293	.591	.911
Context sensitive help, documentation, and training materials	71.87	116.453	.605	.911

4.5.1 Factor Analysis

The factor analysis is an important method to reduce measured variables into much smaller number of constructs. Factor analysis provides measurements which observed variables are strongly correspond to each other and then grouped it together. In this step, KMO & Bartlett's Test of Sphericity is done as it plays an important role for accepting predictors adequacy (Li, 2013). The predictors are adequate if the value of sampling adequacy is more than 0.6 and the significance value (p – value) is less than 0.05.

Based on Table 4.4, the value of the sampling adequacy is 0.875 which is higher than 0.6 and the significance value is 0.000 which is less than 0.05. So, we can deduct that the predictors are adequate.

Table 4.4 KMO and Bartlett's Test of Usability Predictors

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.875
Bartlett's Test of Sphericity	Approx. Chi-Square	1026.482
	df	190
	Sig.	.000

4.6 One Sample t Test

One sample t test is used to find out whether a sample comes from a population with a specific mean of all usability predictors. To find out which usability predictors are significant; this one sample t test is used along with research question and formulation of hypothesis. In this test, P-value (Sig. 2-tailed) is observed since it is the probability of the sample statistic to assess the null hypothesis to find out whether any of P-value is less than the significance level. One sample t test presents if P-value > 0.05 then the usability predictors is considered as not significance.

RQ1: Are all usability predictors are statistically significant?

Hypothesis:

Ho: There is no statistical significance difference between means of usability predictors

Ha: There is a statistical significance difference between means of usability predictors

Based on Table 4.5, test value = 3 indicates that based on Likert scale from the questionnaire, 3.0 is deemed to have normal levels of an agreement which is neutral. *t* shows the comparison of t-distribution between usability predictors. The P-value of usability predictors of aesthetics is higher than 0.05 which shows that difference between the overall population mean and the comparison population mean of usability predictors would be statistically significantly different. Therefore, aesthetics is considered as not significance. Overall, all usability predictors except for aesthetics were accepted as significant usability predictors. However, for usability predictor; backup and recovery of data is near to be considered as not significance because the P-value which is 0.011 is near to 0.05. So, the study rejected Ha and accepted Ho which means that there is usability predictor which is not significance.

Table 4.5 One-Sample Test of Usability Predictors

One-Sample Test

	Test Value = 3					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Simplicity	7.694	100	.000	.772	.57	.97
Aesthetics	.000	100	1.000	.000	-.18	.18
User satisfaction	10.669	100	.000	.960	.78	1.14
Learnability	9.100	100	.000	.901	.70	1.10
Consistency in the user interface	9.258	100	.000	.842	.66	1.02
Operability	11.251	100	.000	.911	.75	1.07
Efficiency	13.106	100	.000	1.218	1.03	1.40
Performance	13.013	100	.000	1.178	1.00	1.36
Usefulness	21.202	100	.000	1.446	1.31	1.58
Reliability	14.772	100	.000	1.228	1.06	1.39
Flexibility	10.059	100	.000	.861	.69	1.03
Supportability	5.523	100	.000	.525	.34	.71
Universality	2.940	100	.004	.297	.10	.50
Reusability of workflows	9.225	100	.000	.861	.68	1.05
Support of data transformation	7.236	100	.000	.634	.46	.81
Support of different formats of the output	7.111	100	.000	.624	.45	.80
Support of data analyses	12.680	100	.000	.950	.80	1.10
Backup and recovery of data	2.605	100	.011	.287	.07	.51
Standard language or format	3.749	100	.000	.376	.18	.58
Context sensitive help, documentation, and training materials	6.087	100	.000	.634	.43	.84

4.7 ANOVA Test (Two – Way ANOVA)

ANOVA analysis is used to test differences between two or more means. The main objective to do ANOVA test is to find out if there is any interaction between two or more independent variables on the dependant variable. ANOVA test provides the decision to reject the null hypothesis or accept the other hypothesis. P-value is compared to the significance level to assess the null hypothesis to find out whether any of the differences between the means are statistically significant. Usually, a significance level or known as α or alpha of 0.05 works well.

In this part, hypotheses were formulated along with research questions (RQs) to do analyses. Usability predictors of the dependent factors were allocated which consists of simplicity, aesthetics, user satisfaction, learnability, consistency in the user interface, operability, efficiency, performance, usefulness, reliability, flexibility, supportability, universality, reusability of workflows, support of data transformation, support of different formats of the output, support of data analyses, backup and recovery of data, standard language and format, and lastly, context sensitivity help, documentation and training materials. Meanwhile, type of employment, educational level, type of project and domain of application are allocated as the independent factors.

In this part of research four RQs were formulated: Does different type of employment affect the evaluation of usability factors of scientific workflows? Does different educational levels affect the evaluation of usability factors of scientific workflows? Does different type of project affect the evaluation of usability factors of scientific workflows? and Does different type of domain application affect the evaluation of usability factors of scientific workflows?

RQ1: Does different type of employment affect the evaluation of usability factors of scientific workflows?

Hypothesis:

Ho: Different types of employment do not affect the evaluation of usability factors of scientific workflows

Ha: Different types of employment affect the evaluation of usability factors of scientific workflows

Based on Table 4.6, the significance value (p – value) is 0.702 for simplicity, 0.439 for aesthetics, 0.470 for user satisfaction, 0.567 for learnability, 0.596 for consistency in the user interface, 0.473 for operability, 0.274 for efficiency, 0.289 for performance, 0.048 for usefulness, 0.215 for reliability, 0.551 for flexibility, 0.971 for supportability, 0.777 for universality, 0.038 for reusability of workflows, 0.306 for support of data transformation, 0.315 for support of different formats of the output, 0.005 for support of data analyses, 0.785 for backup and recovery of data, 0.597 for standard language and format, and 0.122 for context sensitivity help, documentation and training materials. All the significance values are greater than 0.05 except for usefulness and reusability of workflows.

Therefore, there is statistical significance difference in the mean of usefulness, support of data analyses and reusability of workflows' factors with different type of employment. So, the study rejected Ho and accepted Ha for usefulness and reusability of workflows' factor. Meanwhile for other usability factors, the study accepted Ho.

It can be concluded that different type of employment does affect the evaluation of usefulness and reusability of workflows while for other usability factors, we can conclude that different type of employment does not affect the evaluation of other factors of usability of scientific workflows.

Table 4.6 Output of the ANOVA analysis for Usability Factors by Type of Employment

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Simplicity	Between Groups	.151	1	.151	.147	.702
	Within Groups	101.611	99	1.026		
	Total	101.762	100			
Aesthetics	Between Groups	.510	1	.510	.605	.439
	Within Groups	83.490	99	.843		
	Total	84.000	100			
User satisfaction	Between Groups	.432	1	.432	.526	.470
	Within Groups	81.409	99	.822		
	Total	81.842	100			
Learnability	Between Groups	.328	1	.328	.329	.567
	Within Groups	98.682	99	.997		
	Total	99.010	100			
Consistency in the user interface	Between Groups	.238	1	.238	.283	.596
	Within Groups	83.227	99	.841		
	Total	83.465	100			
Operability	Between Groups	.344	1	.344	.518	.473
	Within Groups	65.854	99	.665		
	Total	66.198	100			
Efficiency	Between Groups	1.051	1	1.051	1.208	.274
	Within Groups	86.157	99	.870		
	Total	87.208	100			
Performance	Between Groups	.939	1	.939	1.135	.289
	Within Groups	81.854	99	.827		
	Total	82.792	100			
Usefulness	Between Groups	1.824	1	1.824	4.002	.048
	Within Groups	45.126	99	.456		
	Total	46.950	100			
Reliability	Between Groups	1.081	1	1.081	1.558	.215
	Within Groups	68.682	99	.694		
	Total	69.762	100			

Flexibility	Between Groups	.266	1	.266	.358	.551
	Within Groups	73.793	99	.745		
	Total	74.059	100			
Supportability	Between Groups	.001	1	.001	.001	.971
	Within Groups	91.187	99	.921		
	Total	91.188	100			
Universality	Between Groups	.084	1	.084	.081	.777
	Within Groups	103.005	99	1.040		
	Total	103.089	100			
Reusability of workflows	Between Groups	3.782	1	3.782	4.442	.038
	Within Groups	84.278	99	.851		
	Total	88.059	100			
Support of data transformation	Between Groups	.819	1	.819	1.059	.306
	Within Groups	76.626	99	.774		
	Total	77.446	100			
Support of different formats of the output	Between Groups	.794	1	.794	1.022	.315
	Within Groups	76.909	99	.777		
	Total	77.703	100			
Support of data analyses	Between Groups	4.293	1	4.293	8.101	.005
	Within Groups	52.460	99	.530		
	Total	56.752	100			
Backup and recovery of data	Between Groups	.092	1	.092	.075	.785
	Within Groups	122.581	99	1.238		
	Total	122.673	100			
Standard language or format	Between Groups	.289	1	.289	.282	.597
	Within Groups	101.414	99	1.024		
	Total	101.703	100			
Context sensitive help, documentation, and training materials	Between Groups	2.622	1	2.622	2.430	.122
	Within Groups	106.823	99	1.079		
	Total	109.446	100			

RQ2: Does different educational level affect the evaluation of usability factors of scientific workflows?

Hypothesis:

Ho: Different educational levels do not affect the evaluation of usability factors of scientific workflows

Ha: Different educational levels affect the evaluation of usability factors of scientific workflows

Based on Table 4.7, the significance value (p – value) is 0.314 for simplicity, 0.529 for aesthetics, 0.342 for user satisfaction, 0.420 for learnability, 0.572 for consistency in the user interface, 0.854 for operability, 0.207 for efficiency, 0.496 for performance, 0.09 for usefulness, 0.899 for reliability, 0.079 for flexibility, 0.728 for supportability, 0.166 for universality, 0.093 for reusability of workflows, 0.592 for support of data transformation, 0.592 for support of different formats of the output, 0.075 for support of data analyses, 0.376 for backup and recovery of data, 0.482 for standard language and format, and 0.814 for context sensitivity help, documentation and training materials. All the significance values are greater than 0.05.

Therefore, there is no statistically significance difference in the mean of usability factors of scientific workflows and different educational level. So, the study accepted Ho.

It can be concluded that different educational level does not affect the evaluation of usability factors of scientific workflows.

Table 4.7 Output of the ANOVA analysis for Usability Factors by Educational Level

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Simplicity	Between Groups	.151	1	.151	.147	.702
	Within Groups	101.611	99	1.026		
	Total	101.762	100			
Aesthetics	Between Groups	.510	1	.510	.605	.439
	Within Groups	83.490	99	.843		
	Total	84.000	100			
User satisfaction	Between Groups	.432	1	.432	.526	.470
	Within Groups	81.409	99	.822		
	Total	81.842	100			
Learnability	Between Groups	.328	1	.328	.329	.567
	Within Groups	98.682	99	.997		
	Total	99.010	100			
Consistency in the user interface	Between Groups	.238	1	.238	.283	.596
	Within Groups	83.227	99	.841		
	Total	83.465	100			
Operability	Between Groups	.344	1	.344	.518	.473
	Within Groups	65.854	99	.665		
	Total	66.198	100			
Efficiency	Between Groups	1.051	1	1.051	1.208	.274
	Within Groups	86.157	99	.870		
	Total	87.208	100			
Performance	Between Groups	.939	1	.939	1.135	.289
	Within Groups	81.854	99	.827		
	Total	82.792	100			
Usefulness	Between Groups	1.824	1	1.824	4.002	.048
	Within Groups	45.126	99	.456		
	Total	46.950	100			
Reliability	Between Groups	1.081	1	1.081	1.558	.215
	Within Groups	68.682	99	.694		
	Total	69.762	100			
Flexibility	Between Groups	.266	1	.266	.358	.551
	Within Groups	73.793	99	.745		
	Total	74.059	100			

Supportability	Between Groups	.001	1	.001	.001	.971
	Within Groups	91.187	99	.921		
	Total	91.188	100			
Universality	Between Groups	.084	1	.084	.081	.777
	Within Groups	103.005	99	1.040		
	Total	103.089	100			
Reusability of workflows	Between Groups	3.782	1	3.782	4.442	.038
	Within Groups	84.278	99	.851		
	Total	88.059	100			
Support of data transformation	Between Groups	.819	1	.819	1.059	.306
	Within Groups	76.626	99	.774		
	Total	77.446	100			
Support of different formats of the output	Between Groups	.794	1	.794	1.022	.315
	Within Groups	76.909	99	.777		
	Total	77.703	100			
Support of data analyses	Between Groups	4.293	1	4.293	8.101	.005
	Within Groups	52.460	99	.530		
	Total	56.752	100			
Backup and recovery of data	Between Groups	.092	1	.092	.075	.785
	Within Groups	122.581	99	1.238		
	Total	122.673	100			
Standard language or format	Between Groups	.289	1	.289	.282	.597
	Within Groups	101.414	99	1.024		
	Total	101.703	100			
Context sensitive help, documentation, and training materials	Between Groups	2.622	1	2.622	2.430	.122
	Within Groups	106.823	99	1.079		
	Total	109.446	100			

RQ3: Does different type of project affect the evaluation of usability factors of scientific workflows?

Hypothesis:

Ho: Different type of projects do not affect the evaluation of usability factors of scientific workflows

Ha: Different type of projects affect the evaluation of usability factors of scientific workflows

Based on Table 4.8, the significance value (p – value) is 0.825 for simplicity, 0.503 for aesthetics, 0.399 for user satisfaction, 0.730 for learnability, 0.820 for consistency in the user interface, 0.731 for operability, 0.430 for efficiency, 0.566 for performance, 0.123 for usefulness, 0.273 for reliability, 0.466 for flexibility, 0.042 for supportability, 0.189 for universality, 0.827 for reusability of workflows, 0.440 for support of data transformation, 0.485 for support of different formats of the output, 0.327 for support of data analyses, 0.721 for backup and recovery of data, 0.508 for standard language and format, and 0.845 for context sensitivity help, documentation and training materials. All the significance values are greater than 0.05 except for supportability factor.

Therefore, there is statistically significance difference in the mean of supportability factor with different type of project. So, the study rejected Ho and accepted Ha for supportability's factor. Meanwhile for other usability factors, the study accepted Ho.

It can be concluded that different type of project does affect the evaluation of supportability while for other usability factors, we can conclude that different type of employment does not affect the evaluation of other usability factors of usability of scientific workflows.

Table 4.8 Output of the ANOVA analysis for Usability Factors by Type of Project

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Simplicity	Between Groups	.936	3	.312	.300	.825
	Within Groups	100.826	97	1.039		
	Total	101.762	100			
Aesthetics	Between Groups	2.002	3	.667	.790	.503
	Within Groups	81.998	97	.845		
	Total	84.000	100			
User satisfaction	Between Groups	2.440	3	.813	.994	.399
	Within Groups	79.401	97	.819		
	Total	81.842	100			
Learnability	Between Groups	1.308	3	.436	.433	.730
	Within Groups	97.702	97	1.007		
	Total	99.010	100			
Consistency in the user interface	Between Groups	.787	3	.262	.308	.820
	Within Groups	82.678	97	.852		
	Total	83.465	100			
Operability	Between Groups	.873	3	.291	.432	.731
	Within Groups	65.326	97	.673		
	Total	66.198	100			
Efficiency	Between Groups	2.436	3	.812	.929	.430
	Within Groups	84.772	97	.874		
	Total	87.208	100			
Performance	Between Groups	1.705	3	.568	.680	.566
	Within Groups	81.087	97	.836		
	Total	82.792	100			
Usefulness	Between Groups	2.702	3	.901	1.975	.123
	Within Groups	44.248	97	.456		
	Total	46.950	100			

Reliability	Between Groups	2.735	3	.912	1.319	.273
	Within Groups	67.028	97	.691		
	Total	69.762	100			
Flexibility	Between Groups	1.915	3	.638	.858	.466
	Within Groups	72.145	97	.744		
	Total	74.059	100			
Supportability	Between Groups	7.343	3	2.448	2.832	.042
	Within Groups	83.845	97	.864		
	Total	91.188	100			
Universality	Between Groups	4.927	3	1.642	1.623	.189
	Within Groups	98.162	97	1.012		
	Total	103.089	100			
Reusability of workflows	Between Groups	.803	3	.268	.297	.827
	Within Groups	87.257	97	.900		
	Total	88.059	100			
Support of data transformation	Between Groups	2.117	3	.706	.909	.440
	Within Groups	75.329	97	.777		
	Total	77.446	100			
Support of different formats of the output	Between Groups	1.924	3	.641	.821	.485
	Within Groups	75.779	97	.781		
	Total	77.703	100			
Support of data analyses	Between Groups	1.974	3	.658	1.165	.327
	Within Groups	54.778	97	.565		
	Total	56.752	100			
Backup and recovery of data	Between Groups	1.665	3	.555	.445	.721
	Within Groups	121.008	97	1.248		
	Total	122.673	100			
Standard language or format	Between Groups	2.397	3	.799	.780	.508
	Within Groups	99.306	97	1.024		
	Total	101.703	100			
Context sensitive help, documentation, and training materials	Between Groups	.914	3	.305	.272	.845
	Within Groups	108.532	97	1.119		
	Total	109.446	100			

RQ4: Does different domain of application affect the evaluation of usability factors of scientific workflows?

Hypothesis:

Ho: There is no difference between responders by domain of application of scientific workflows towards the evaluation of usability factors of scientific workflows

Ha: There is a difference between responders by domain of application of scientific workflows towards the evaluation of usability factors of scientific workflows

Based on Table 4.9, the significance value (p – value) is 0.216 for simplicity, 0.806 for aesthetics, 0.883 for user satisfaction, 0.664 for learnability, 0.178 for consistency in the user interface, 0.924 for operability, 0.486 for efficiency, 0.545 for performance, 0.560 for usefulness, 0.494 for reliability, 0.662 for flexibility, 0.829 for supportability, 0.585 for universality, 0.845 for reusability of workflows, 0.452 for support of data transformation, 0.332 for support of different formats of the output, 0.394 for support of data analyses, 0.410 for backup and recovery of data, 0.091 for standard language and format, and 0.788 for context sensitivity help, documentation and training materials. All the significance values are greater than 0.05.

Therefore, there is no statistically significance difference in the mean of usability factors of scientific workflows and different domain of application. So, the study accepted Ho.

It can be concluded that different domain of application does not affect the evaluation of usability factors of scientific workflows.

Table 4.9 Output of the ANOVA analysis for Usability Factors by Domain of Application

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Simplicity	Between Groups	8.441	6	1.407	1.417	.216
	Within Groups	93.321	94	.993		
	Total	101.762	100			
Aesthetics	Between Groups	2.602	6	.434	.501	.806
	Within Groups	81.398	94	.866		
	Total	84.000	100			
User satisfaction	Between Groups	1.996	6	.333	.392	.883
	Within Groups	79.846	94	.849		
	Total	81.842	100			
Learnability	Between Groups	4.136	6	.689	.683	.664
	Within Groups	94.874	94	1.009		
	Total	99.010	100			
Consistency in the user interface	Between Groups	7.404	6	1.234	1.525	.178
	Within Groups	76.061	94	.809		
	Total	83.465	100			
Operability	Between Groups	1.331	6	.222	.322	.924
	Within Groups	64.867	94	.690		
	Total	66.198	100			
Efficiency	Between Groups	4.828	6	.805	.918	.486
	Within Groups	82.380	94	.876		
	Total	87.208	100			
Performance	Between Groups	4.196	6	.699	.836	.545
	Within Groups	78.596	94	.836		
	Total	82.792	100			
Usefulness	Between Groups	2.325	6	.388	.816	.560
	Within Groups	44.625	94	.475		
	Total	46.950	100			
Reliability	Between Groups	3.815	6	.636	.906	.494
	Within Groups	65.947	94	.702		
	Total	69.762	100			
Flexibility	Between Groups	3.102	6	.517	.685	.662
	Within Groups	70.958	94	.755		
	Total	74.059	100			
Supportability	Between Groups	2.657	6	.443	.470	.829

	Within Groups	88.531	94	.942		
	Total	91.188	100			
Universality	Between Groups	4.909	6	.818	.783	.585
	Within Groups	98.180	94	1.044		
	Total	103.089	100			
Reusability of workflows	Between Groups	2.443	6	.407	.447	.845
	Within Groups	85.616	94	.911		
	Total	88.059	100			
Support of data transformation	Between Groups	4.504	6	.751	.967	.452
	Within Groups	72.941	94	.776		
	Total	77.446	100			
Support of different formats of the output	Between Groups	5.378	6	.896	1.165	.332
	Within Groups	72.325	94	.769		
	Total	77.703	100			
Support of data analyses	Between Groups	3.585	6	.598	1.057	.394
	Within Groups	53.167	94	.566		
	Total	56.752	100			
Backup and recovery of data	Between Groups	7.578	6	1.263	1.031	.410
	Within Groups	115.096	94	1.224		
	Total	122.673	100			
Standard language or format	Between Groups	10.939	6	1.823	1.888	.091
	Within Groups	90.764	94	.966		
	Total	101.703	100			
Context sensitive help, documentation, and training materials	Between Groups	3.552	6	.592	.525	.788
	Within Groups	105.894	94	1.127		
	Total	109.446	100			

4.8 MATLAB Coding for CronbachAlpha

MATLAB integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. In the study, MATLAB Cronbach Alpha function was developed to calculate the Cronbach's alpha for data received from responders. The questionnaire's answers must be a 2-dimensional matrix, which is the only function's input. If the result is greater than 0.8 then the answers are considered as reliable. If the score is below 0.2 then the answers are not reliable. Below shows the MATLAB code used to calculate Cronbach's alpha:

```
function [as,varargout] = CronbachAlpha(x)

% CronbachAlpha

% Description:    calculate Cronbach's alpha for a set of psychometric
measurements

% Syntax:    [as,au] = CronbachAlpha(x)

% In:

%    x    - an nRep x nItem array of ratings, so that each row is the set of

%           obvservations from one repetition and each column is the set of
all

%           observations for a given item

% Out:
```

```
%	as	- the standardized Cronbach's alpha

%	au	- the unstandardized Cronbach's alpha

nItem	= size(x,2);

%logical array for selecting upper triangular part of the correlation and

%covariance matrices, where the good stuff is

	b	= triu(true(nItem),1);

%standardized alpha

	%pairwise correlations between items

		r	= corrcoef(x);

	%mean of the meaningful, non-redundant correlations

		r	= mean(r(b));

	as	= nItem*r/(1 + (nItem-1)*r);

%unstandardized alpha

if nargout>1

	%variance/covariance matrix
```

```
        vc          = cov(x);

    %mean variance (variances are along the diagonal)

        v           = mean(diag(vc));

    %mean covariance, not including variances

        c           = mean(vc(b));

    varargout{1}    = nItem*c/(v + (nItem-1)*c);

end
```

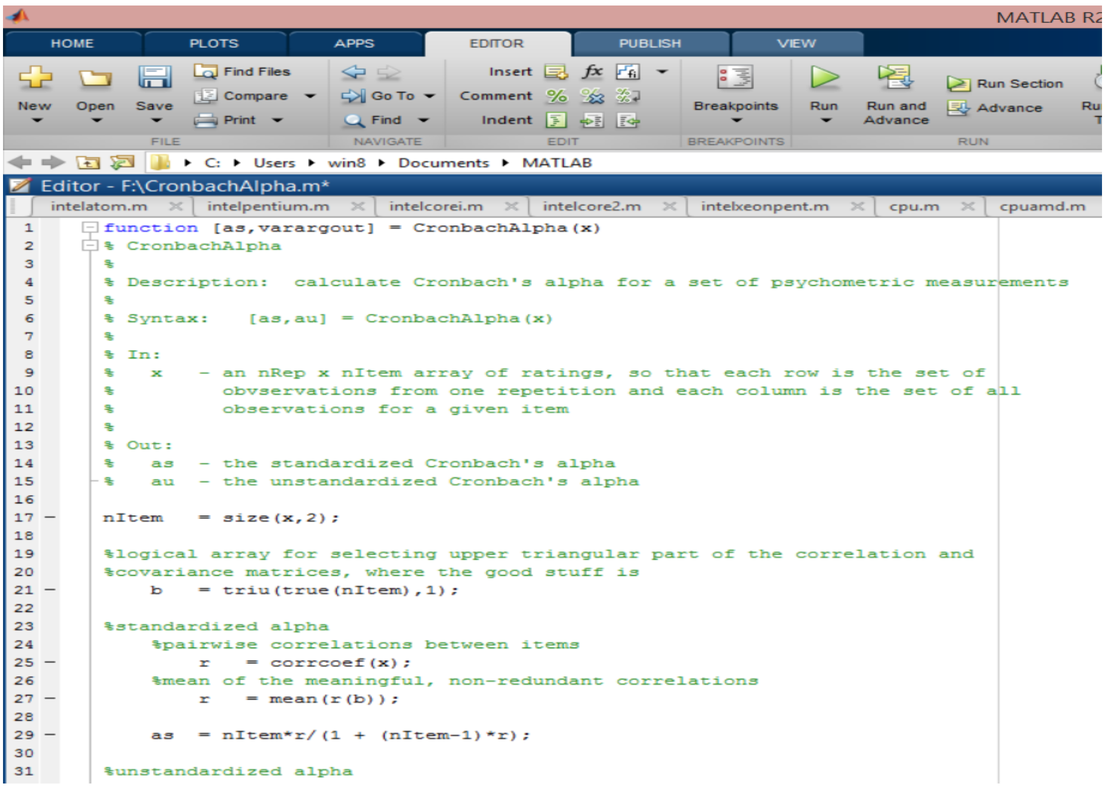


Figure 4.22 MATLAB Coding

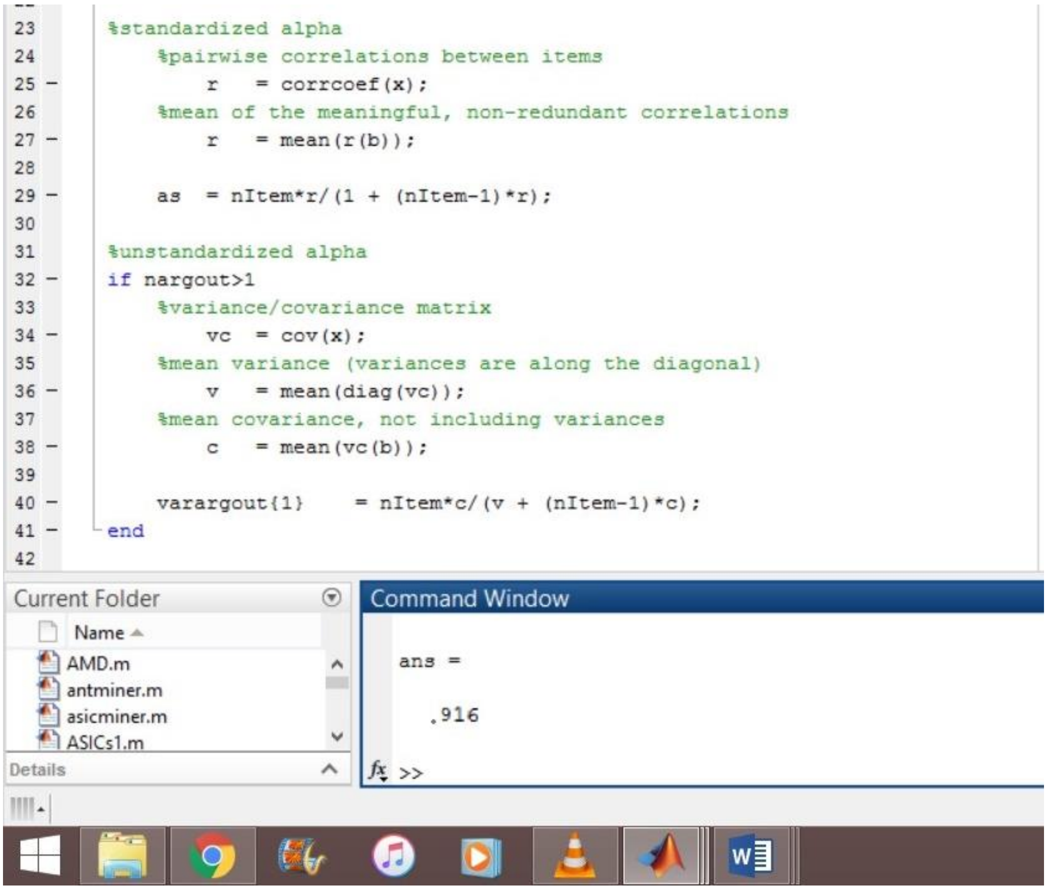


Figure 4.23 MATLAB Coding

Based on Figure 4.22 and Figure 4.23, the interface of MATLAB code, nItem is equal to the number of items, c is the average inter-item covariance among the items and v equals to average variance. From the formula, if a user increases the number of items, the Cronbach's alpha will increase. If the average inter-item correlation is low, alpha will be low. As the average inter-item correlation increases, Cronbach's alpha increases as well. The result of Cronbach's alpha is 0.916.

4.9 Discussion

Validity of the collected data of 101 respondents from scientific communities is checked. Then, proposed model is evaluated based on appropriate case study with application of methods mathematical statistics using SPSS. The first test that was performed is reliability analysis which calculates Cronbach's Alpha. Based on the obtained result, the Cronbach's Alpha (α) is 0.916. So, it is deducted that the reliability coefficient is considered acceptable.

The most important test was, to find out significance of usability predictors by One Sample t Test. P-value of usability predictors were compared and any P-value which is higher than 0.05 was considered as not significant. There were only one usability predictors was found which is considered as not significant - aesthetics.

Then, two- way ANOVA is performed to find out if there is any interaction between independent variables on the dependant variable. Hypotheses were formulated along with research questions (RQs) to do analyses and determine whether to accept or reject the hypothesis. Significant difference between independent factors (type of employment, educational level, type of project and domain of application) and dependent factor(simplicity, aesthetics, user satisfaction, learnability, consistency in the user interface, operability, efficiency, performance, usefulness, reliability, flexibility, supportability, universality, reusability of workflows, support of data transformation, support of different formats of the output, support of data analyses, backup and recovery of data, standard language and format, and context sensitivity help, documentation and training materials) were checked in term of usability predictors. Based on the obtained result, type of employment does affect evaluation of usability factors which are usefulness and reusability of workflows. Besides, type of project also does affect supportability factor. So, the conclusion is that type of employment and type of project are the factors that have significant role on the evaluation of usability of scientific workflows.

CHAPTER 5

CONCLUSION

5.1 Overview

This purpose of this chapter is to conclude the findings of the research. In this chapter, summary of research, findings of the research, research constraints and directions for future work are concluded.

5.2 Introduction

The purpose of this research is to develop the model of evaluation of usability of scientific workflows. The aim of the research was successfully achieved by studying and analysing the existing models of usability. Based in analyses, a model was proposed. Next, to evaluate the proposed model the questionnaire was developed for data collection and analysis. Later, the finalised questionnaire was distributed to scientific community through Web-based application (Google Docs). A sample data from 101 respondents was collected. Then, statistical methods (reliability test, factor analysis, one sample t test and ANOVA test) along with development of MATLAB coding for Cronbach's Alpha were done to analyse the data. Finally, the proposed model of evaluation of usability of scientific workflows was assessed from the findings.

5.3 Research Result

Results of the study show there is changes in the proposed model since aesthetics was considered not significant, so it will not be included in the final model. Based on ANOVA test result, the most affected usability predictors based on the interaction between independent variables on the dependant variable of the usability of scientific workflows are usefulness, supportability, support of data analysis and reusability.

However, all the other usability predictors were accepted, strongly correspond with each other and considered as adequate.

5.4 Research Constraints

Although the results collected from the questionnaire were significant, there are some limitations can be identified and will be solved in future studies. Firstly, time was one of the research constraints. This is because the time taken for data collection of valid responses was longer than predicted compared to the time taken to analyse and test the data. Secondly, the response rate of respondents among scientific community was very low. Only 101 responses were gathered out of 1000 emails being sent.

5.5 Future Works

There are several improvements that can be done in future research. First, increasing the number of respondents that will be involved in the survey. This will allow a better assumption for evaluating the usability of scientific workflows among scientific community.

Furthermore, this research was done by allocating 20 factors of usability predictors in terms of scientific workflows systems. In future, it can be done using other usability factors from various usability models or measurement.

REFERENCES

Abiud, M. M., & Mbugua, P. (2016). An analytical comparative analysis of the software quality models for software quality engineering, *1*(2), 15–24.

Abran, A., Khelifi, A., Suryn, W., & Seffah, A. (2003). Usability meanings and interpretations in ISO standards. *Software Quality Journal*, *11*(4), 325–338. https://doi.org/10.1023/A:1025869312943

Gupta, D., Ahlawat, A., & Sagar, K. (2015). A Critical Analysis of A Hierarchy Based Usability Model, (September 2017). https://doi.org/10.1109/IC3I.2014.7019810

Informatics, A. (2008). A Survey on Usability Evaluation Techniques and an Analysis of their actual Application, (450).

Siddaway, A. (2014). WHAT IS A SYSTEMATIC LITERATURE REVIEW AND HOW DO I DO ONE?, (Ii), 1–13.

Yeltayeva, K. (2012). Usability Study of the Taverna Scientific Workflow Workbench, 99.

https://www.technologynetworks.com/informatics/articles/one-way-vs-two-way-anova-definition-differences-assumptions-and-hypotheses-306553

MatLab (www.mathworks.com)

One Sample t Test (https://statistics.laerd.com/spss-tutorials/one-sample-t-test-using-spss-statistics.php)

Cronbach's Alpha (https://www.mathworks.com/matlabcentral/fileexchange/7829-cronbach-s-alpha)

APPENDIX A

PSM 1 Gantt chart

The figures below show the Gantt chart of the project.

		Task Name	Duration	Start	Finish
1		Briefing session	5 days	Mon 12/02/18 8:00 AM	Fri 16/02/18 5:00 PM
2		Meet up with supervisor to discuss on research title	5 days	Fri 16/02/18 8:00 AM	Thu 22/02/18 5:00 PM
3		Research on problem statements, objective and scope of research	10 days	Tue 20/02/18 8:00 AM	Mon 05/03/18 5:00 PM
4		Correction for introduction and discuss on literature	5 days	Mon 05/03/18 8:00 AM	Fri 09/03/18 5:00 PM
5		Literature review;study on previous related researches	10 days	Fri 09/03/18 8:00 AM	Thu 22/03/18 5:00 PM
6		Correction on literature review and discuss on	8 days	Thu 22/03/18 8:00 AM	Mon 02/04/18 5:00 PM
7		Methodology	10 days	Wed 04/04/18 8:00 AM	Tue 17/04/18 5:00 PM
8		Correction on methodology	8 days	Tue 17/04/18 8:00 AM	Thu 26/04/18 5:00 PM
9		Overall correction on PSM1	10 days	Thu 26/04/18 8:00 AM	Wed 09/05/18 5:00 PM
10		Finalizing PSM1	5 days	Mon 30/04/18 8:00 AM	Fri 04/05/18 5:00 PM
11		Submit and present PSM1	7 days	Fri 04/05/18 8:00 AM	Mon 14/05/18 5:00 PM

Figure 0.1 Research phases and estimated time for PSM1

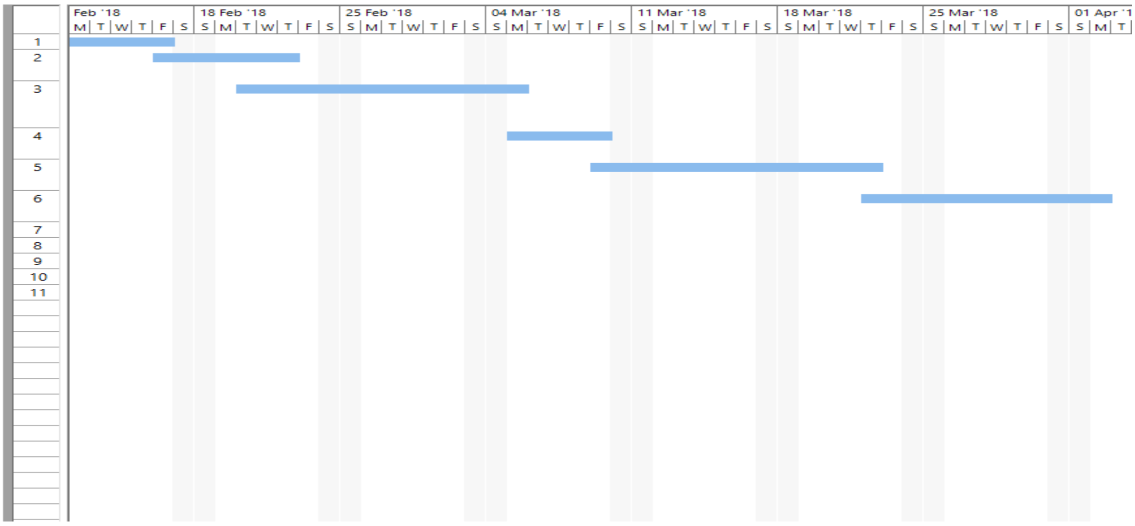


Figure 0.2 Gantt chart from Week 1 until Week 6

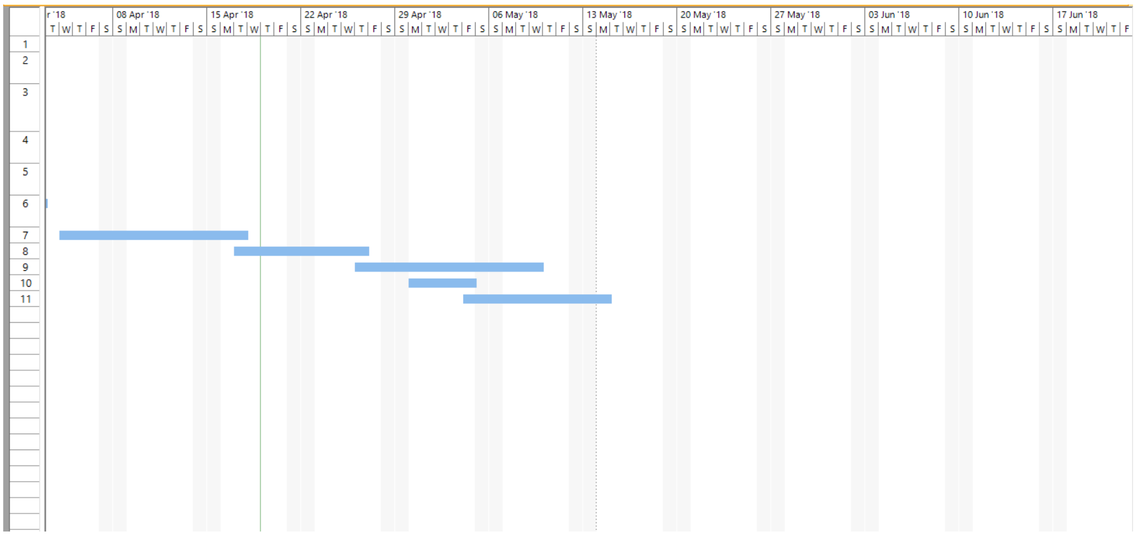


Figure 0.3 Gantt chart from Week 7 until Week 11

PSM 2 Gantt chart

	Task Name	Duration	Start	Finish
1	PSM2 Briefing Session	1 day	Thu 9/13/18	Thu 9/13/18
2	Meet up with SV to discuss Result & Discussion: Implementation & Testing	5 days	Fri 9/14/18	Thu 9/20/18
3	Result & Discussion	12 days	Thu 9/20/18	Fri 10/5/18
4	Submit Result & Discussion	5 days	Mon 10/8/18	Fri 10/12/18
5	Correction of Result & Discussion	8 days	Wed 10/10/18	Fri 10/19/18
6	Submit correction of Result & Discussion	5 days	Sat 10/20/18	Thu 10/25/18
7	Conclusion	6 days	Thu 10/25/18	Thu 11/1/18
8	Submit Conclusion	4 days	Thu 11/1/18	Tue 11/6/18
9	Correction of Conclusion	5 days	Wed 11/7/18	Tue 11/13/18
10	Submit correction of Conclusion	4 days	Tue 11/13/18	Fri 11/16/18
11	Documenting Thesis	11 days	Fri 11/16/18	Fri 11/30/18
12	PSM2 Briefing Session	1 day	Wed 11/21/18	Wed 11/21/18
13	Submit Thesis to SV	6 days	Fri 11/30/18	Fri 12/7/18
14	Correction of Thesis	6 days	Mon 12/3/18	Mon 12/10/18
15	Submit Thesis to Faculty	3 days	Mon 12/10/18	Wed 12/12/18
16	Presentation PSM2	3 days	Tue 12/18/18	Thu 12/20/18

Figure 0.4 Research phases and estimated time for PSM2

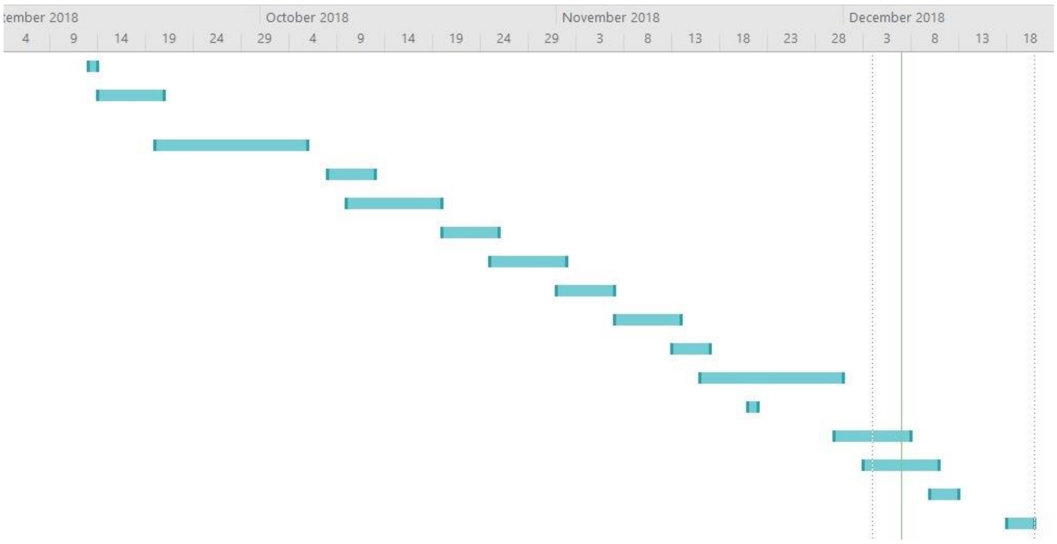
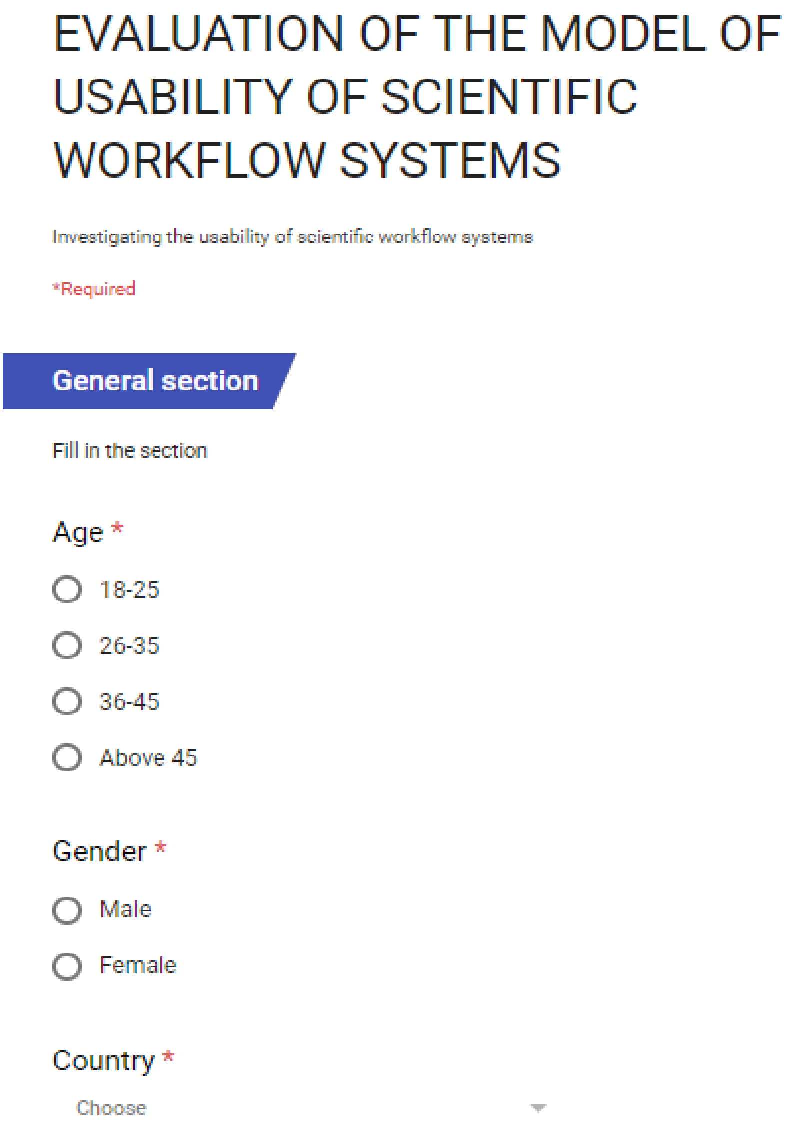


Figure 0.5 Gantt chart from Week 1 until Week 14

Questionnaire Form (Google Docs)



EVALUATION OF THE MODEL OF USABILITY OF SCIENTIFIC WORKFLOW SYSTEMS

Investigating the usability of scientific workflow systems

*Required

General section

Fill in the section

Age *

- 18-25
- 26-35
- 36-45
- Above 45

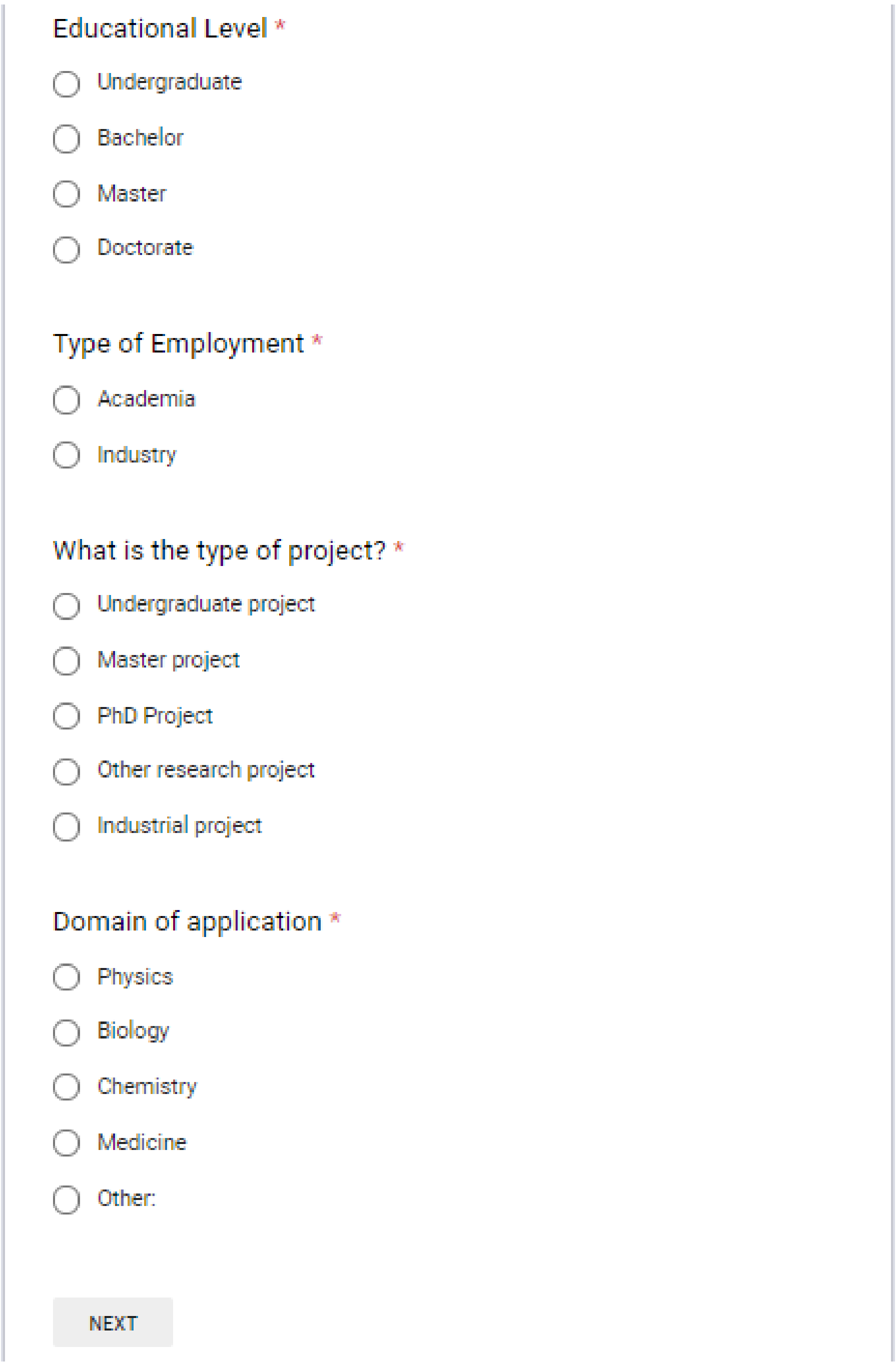
Gender *

- Male
- Female

Country *

Choose

Figure 0.6 Questionnaire Form



Educational Level *

- Undergraduate
- Bachelor
- Master
- Doctorate

Type of Employment *

- Academia
- Industry

What is the type of project? *

- Undergraduate project
- Master project
- PhD Project
- Other research project
- Industrial project

Domain of application *

- Physics
- Biology
- Chemistry
- Medicine
- Other:

NEXT

Figure 0.7 Questionnaire Form

EVALUATION OF THE MODEL OF USABILITY OF SCIENTIFIC WORKFLOW SYSTEMS

*Required

EVALUATE THE FACTORS OF USABILITY OF SCIENTIFIC WORKFLOW SYSTEMS

Rate the criteria using 5 Likert scale

1. Simplicity *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

2. Aesthetics *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

3. User satisfaction *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

4. Learnability *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

Figure 0.8 Questionnaire Form

5. Consistency in the user interface *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

6. Operability *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

7. Efficiency *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

8. Performance *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

9. Usefulness *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

10. Reliability *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

Figure 0.9 Questionnaire Form

11. Flexibility *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

12. Supportability *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

13. Universality *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

14. Reusability of workflows *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

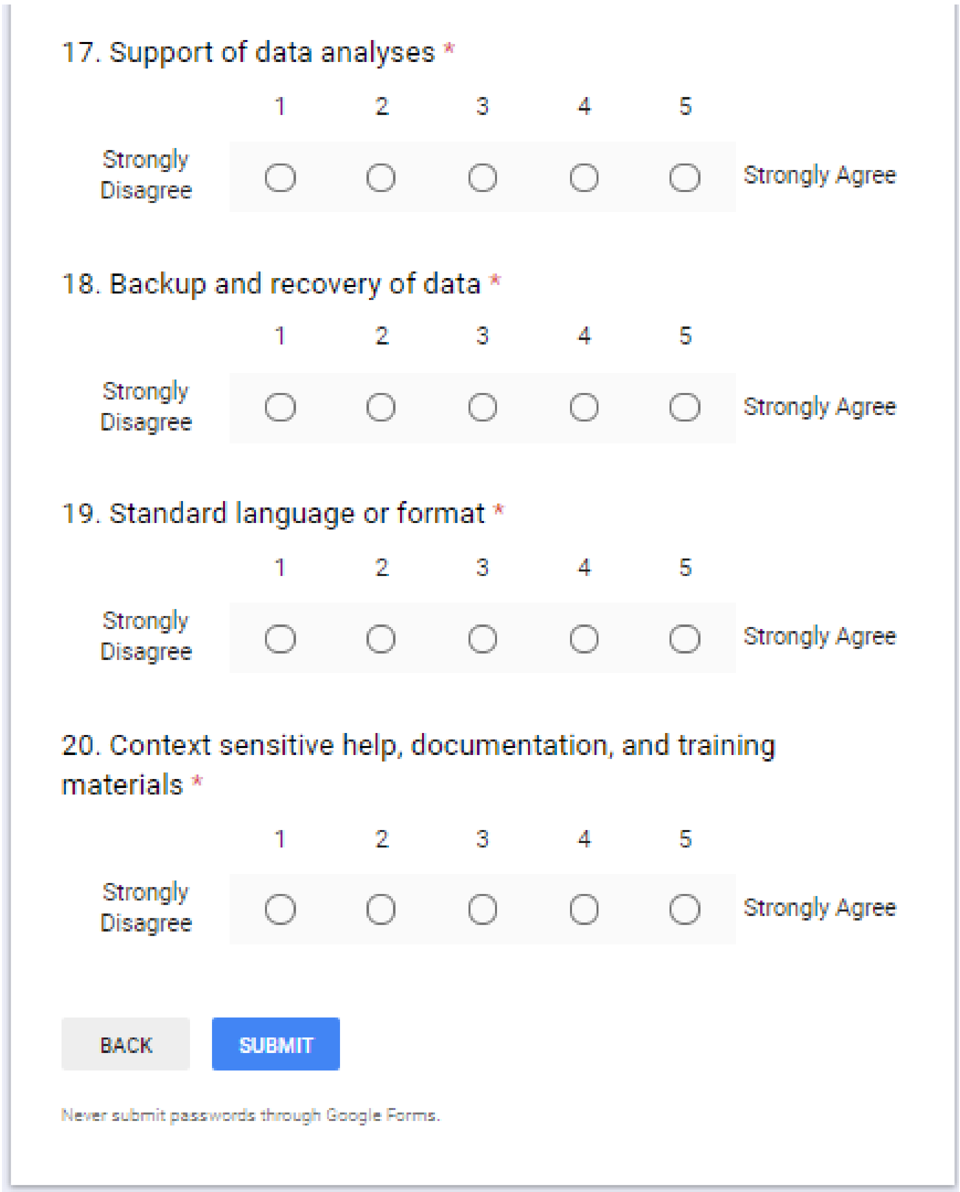
15. Support of data transformation *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

16. Support of different formats of the output *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

Figure 0.10 Questionnaire Form



17. Support of data analyses *

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

18. Backup and recovery of data *

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

19. Standard language or format *

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

20. Context sensitive help, documentation, and training materials *

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

BACK SUBMIT

Never submit passwords through Google Forms.

Figure 0.11 Questionnaire Form