



INTEROPERABILITY SOLUTIONS FOR
EUROPEAN PUBLIC ADMINISTRATIONS
MONITORING AND EVALUATION

D03.04 Perceived Utility Monitoring Report

ISA action 4.2.6 Interoperable test bed

Framework Contract n° DI/07173-00

16 August 2016

DISCLAIMER

The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this document. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.

© European Commission, 2016

EXECUTIVE SUMMARY

The purpose of this section is to provide an overview of the key findings of the Perceived Utility monitoring of **the ISA Action 4.2.6 – Interoperable test bed**. The objective of the survey was to measure the action's Perceived Utility which is defined as the extent to which the effects (impact) of an ISA action correspond with the needs, problems and issues to be addressed by the ISA programme¹ and the actions' specific objectives.

The survey of the Action 4.2.6 included the evaluation of the Interoperable test bed including all its outputs (documentation, test registry and repository and demo installation package). The survey was designed in the EUSurvey tool and distributed by e-mail to 36 contacts. Over the duration of more than one month², 18 stakeholders have responded, however 7 responses were excluded from further analysis due to the respondents' unfamiliarity with this action (i.e., they have never heard or have just heard but never worked with any of the action's outputs).

Table 1 gives an overview of the main results of the survey. The detailed score calculation process is described in section 5.4.4.

TABLE 1 – ACTION 4.2.6 SURVEY MAIN RESULTS

	Score	Mode	StDev	StErr	Explanation of the score scale
Usefulness Score	5.73	5	1.01	0.31	Average value on a scale from 1 (Not useful at All) to 7 (Very Useful).
Value Score	4.40	4	1.06	0.15	Average value of all the statement means in the range from 1 (Disagree) to 5 (Agree).
User Satisfaction Score	80.75	Not applicable for this score			User Satisfaction Score from 0 (none of the respondents are satisfied) to 100 (all respondents are satisfied with the work performed by the Action).
Net Promoter Score	27	Not applicable for this score			Net Promoter Score from -100 (every customer is a Detractor) to 100 (every customer is a Promoter).
OVERALL PERCEIVED UTILITY SCORE	4.08	Not applicable for this score			The Overall Perceived Utility Score is the average value of the Usefulness Score, the Value Score, the User Satisfaction Score and the Net Promoter Score to a five point scale in range from 1 – the lowest score to 5 – the highest score.

Main findings:

- The survey results demonstrate that **Action 4.2.6 – Interoperable test bed complies with both the ISA programme and the actions' specific objectives.**

¹ Papadomichelaki, X. and Mentzas, G. (2012), "e-GovQual: A multiple-item scale for assessing e-government service quality"

² The survey was launched on the 4th of December 2015 and was active until the 11th of January 2016.

- Respondents of the Interoperable test bed survey are satisfied with the action's outputs – documentation, test registry and repository and/or demo installation package.
- The majority of respondents (54.5%) would recommend the outputs of the Interoperable test bed package to colleagues or other public administrations.
- The weakest aspect of the Interoperable test bed is the documentation output.
- Improvements in the documentation such as making it more readable and user friendly would be of benefit to the action.

REVISION HISTORY

Date	Version	Description	Authors	Approved by
08/02/2016	0.10	Initial version	CGI - Accenture	
02/03/2016	0.20	Updated version	CGI - Accenture	
08/04/2016	0.30	Updated version	CGI - Accenture	
13/04/2016	0.40	Updated version	CGI - Accenture	
19/04/2016	0.50	Updated version	CGI - Accenture	
06/06/2016	0.60	Updated version	CGI - Accenture	
10/06/2016	1.00	Final version	CGI - Accenture	
16/08/2016	2.00			Approved by HVA on 29/7/2016.

TABLE OF CONTENTS

1	INTRODUCTION	8
2	OVERVIEW OF THE ACTION 4.2.6 - INTEROPERABLE TEST BED	9
3	SURVEY METHODOLOGY	10
3.1	PERCEIVED UTILITY	10
3.2	SURVEY MEASUREMENTS	11
3.3	SURVEY ARCHITECTURE	12
4	SURVEY DATA SUMMARY	13
5	SURVEY RESULTS AND ANALYSIS.....	14
5.1	DEMOGRAPHIC PROFILE OF RESPONDENTS	14
5.2	USAGE OF THE ACTION	15
5.3	USEFULNESS SCORE	16
5.4	PERCEIVED UTILITY MEASUREMENTS	17
5.4.1	<i>Perceived Utility Value Score</i>	17
5.4.1.1	PERCEIVED UTILITY DIMENSIONS IMPORTANCE	17
5.4.1.2	PERCEIVED UTILITY DIMENSIONS CONFORMITY	18
5.4.1.2.1	PERCEIVED UTILITY STATEMENT MAPPING TO DIMENSIONS.....	19
5.4.1.2.2	PERCEIVED UTILITY DIMENSIONS CONFORMITY RESULTS	19
5.4.1.2.3	PERCEIVED UTILITY CRITERION SCORE AGGREGATION	21
5.4.2	<i>Perceived Utility User Satisfaction Score</i>	22
5.4.3	<i>Perceived Utility Net Promoter Score</i>	23
5.4.4	<i>Overall Perceived Utility Score</i>	24
5.5	ACTIONS STRENGTHS, WEAKNESSES, OPPORTUNITIES AND THREATS	26
5.6	STATEMENTS BASED ON ACTION OBJECTIVES.....	28
5.7	RESPONDENT RECOMMENDATIONS AND OPINIONS	30
6	SURVEY CONCLUSION AND RECOMMENDATIONS.....	31
7	APPENDIX	32
7.1	RAW DATA EXPORT	32
7.2	GLOSSARY	33

TABLE OF FIGURES

FIGURE 1 – ACTION 4.2.6 USEFULNESS SCORE	16
FIGURE 2 – ACTION 4.2.6 PERCEIVED UTILITY DIMENSIONS IMPORTANCE RESULTS	18
FIGURE 3 – ACTION 4.2.6 PERCEIVED UTILITY DIMENSIONS CONFORMITY RESULTS	20
FIGURE 4 – ACTION 4.2.6 PERCEIVED UTILITY CRITERION SCORE AGGREGATION	22
FIGURE 5 – ACTION 4.2.6 PERCEIVED UTILITY USER SATISFACTION SCORE	22
FIGURE 6 – ACTION 4.2.6 PERCEIVED UTILITY NET PROMOTER SCORE	23
FIGURE 7 – ACTION 4.2.6 ACTIONS STRENGTHS, WEAKNESSES, OPPORTUNITIES AND THREATS	27
FIGURE 8 – ACTION 4.2.6 STATEMENTS BASED ON ACTION OBJECTIVES	29

TABLE OF TABLES

TABLE 1 – ACTION 4.2.6 SURVEY MAIN RESULTS	3
TABLE 2 – ACTION 4.2.6 SURVEY TECHNICAL INFORMATION ABOUT THE FIELDWORK	13
TABLE 3 – ACTION 4.2.6 DEMOGRAPHIC PROFILE OF RESPONDENTS.....	14
TABLE 4 – ACTION 4.2.6 USAGE OF INTEROPERABLE TEST BED	15
TABLE 5 – ACTION 4.2.6 STATEMENT MAPPING TO PERCEIVED UTILITY DIMENSIONS	19
TABLE 6 – ACTION 4.2.6 ADDITIONAL STATISTICAL CALCULATIONS FOR PERCEIVED UTILITY DIMENSIONS.....	21
TABLE 7 – ACTION 4.2.6 OVERALL PERCEIVED UTILITY SCORE CALCULATION	25
TABLE 8 – ACTION 4.2.6 RECOMMENDATIONS AND OPINIONS	30

1 INTRODUCTION

CGI-Accenture has been requested to deliver a Perceived Quality and Perceived Utility Monitoring and Evaluation Report as part of the execution of the ISA programme monitoring (Technical Annex for Specific Contract SC 193 under Framework contract n° DI/07173-00).

Based on the scope of the Specific Contract, the Perceived Quality is to be measured for 15 actions and the Perceived Utility is to be measured for 17 actions. This report covers the Perceived Utility measurement for the Action 4.2.6 – Interoperable test bed.

This document is divided into the following sections:

- **Section 1:** provides an overview of the structure of the report;
- **Section 2:** provides an overview of the action and its objectives;
- **Section 3:** explains the methodology used to measure the Perceived Utility;
- **Section 4:** summarises the collected data;
- **Section 5:** focuses on the survey results and the data analysis:
 - The demographic profile of respondents;
 - Usage frequency of the action's outputs;
 - Usefulness Score;
 - Perceived Utility measurements;
 - Actions strengths, weaknesses, opportunities and threats;
 - Statements based on action objectives;
 - Respondent recommendations and opinions.
- **Section 6:** provides the survey conclusion and recommendations;
- **Section 7:** Appendix includes:
 - Raw data export;
 - Glossary.

2 OVERVIEW OF THE ACTION 4.2.6 - INTEROPERABLE TEST BED

Under the ISA programmes and other funding mechanisms of the European Commission, a variety of interoperability solutions is being developed. To avoid compromising the productive system an extensive testing is necessary before connecting new components to these systems (e.g. new partners to a communication network or new clients to a service). Usually these tests require connecting the system to an instance of the service or the communication partner; consequently, there is a need for a reference implementation of this service that is separate from the production instance. In a situation where the compatibility of different systems relies on a conformance to a standard or specification, this conformance can also be assured through testing – either simply by connecting to the reference implementation (which is assumed to implement the specification correctly) or, more reliably, through the execution of detailed test cases to separately test each clause of the specification, or both.

The "Interoperability test bed" action was conceived under the ISA programme to provide an environment where reference implementations of different systems/services could be hosted; studies conducted during the previous phases showed that a dedicated test bed software can support this by providing a user interface, a standardised way to execute tests and access test results and some test automation. In addition, the use of a test bed also enables formal conformance testing against a specification.

The scope of the action has therefore been enlarged to encompass both hosting of reference implementations and provision of a test bed. Under the ISA programme some case studies/pilots were executed to demonstrate this with the help of a test bed software that was developed under the CEN GITB workshop³.

Action's Objectives:

- ITB's (Interoperable test bed) primary objective is to provide a platform for hosting reference implementations of cross-border services, coupled with a test bed that provides a user interface as well as some degree of automation.
- This platform would enable Member States' public administrations (PAs) and their potential vendors to test existing systems or products against a neutral, reliable and responsive test environment of reference.
- The long-term vision is for ITB to become a test center that deploys reference implementations on demand, cooperating with other test centers. Previous studies have shown that using a test bed

³ <http://www.cen.eu/news/workshops/Pages/WS-2015-008.aspx>, accessed on 31/08/2015

conforming to the GITB specifications supports this cooperation well, both for the exchange of test artefacts and for the joint execution of tests.

3 SURVEY METHODOLOGY

A common methodology was developed by the CGI/Accenture team for all the surveys included in the Perceived Quality and Perceived Utility Monitoring and Evaluation Reports. The common methodology enables a comparison between the different action results. This methodology was also applied to perform an evaluation of this action. The first section explains how the Perceived Utility is measured and what dimensions are covered. The next section gives an overview of the main survey measurements. The last section describes the architecture of the survey.

3.1 PERCEIVED UTILITY

Perceived Utility is defined as the extent to which the effects (impact) of an ISA action correspond with the needs, problems and issues to be addressed by the ISA programme⁴ and the actions' specific objectives.

Regarding the Perceived Utility measurement several statements are derived from the objectives of the ISA programme. These statements are grouped in three dimensions which are defined as criteria for measuring Perceived Utility:

- **Potential Re-usability:** The degree to which the action's outcome(s) can be reused by PAs;
- **Sustainability:** To what extent is the financial, technical and operational sustainability of solutions ensured⁵;
- **Collaboration:** The degree to which the action promotes/facilitates collaboration/cooperation between PA's⁶.

The survey statements for the dimensions listed above were developed according to:

- The ISA programme's main objectives: "To support cooperation between European public administrations by facilitating the efficient and effective electronic cross-border and cross-sectorial interaction between such administrations, including bodies performing public functions on their behalf, enabling the delivery of electronic public services supporting the implementation of Community policies and activities"⁷ and actions' specific objectives. The Perceived Utility statements were tailored to reflect these objectives and were based on the ESOMAR⁸ (World Association of Opinion and Marketing Research Professionals) standards.

⁴ Papadomichelaki, X. and Mentzas, G. (2012), "e-GovQual: A multiple-item scale for assessing e-government service quality"

⁵ European Commission (2013), Interim evaluation of the ISA programme, "Report from the Commission to the European Parliament and Council COM (2013) 5 final".

⁶ CRN (2015), Collaboration http://research.crn.com/technology/knowledge_management/collaboration

⁷ Decision No 922/2009/EC of the European Parliament and of the Council of 16 September 2009 on interoperability solutions for European public administrations (ISA) (2009)

⁸ ESOMAR, edited by Hamersveld. M., Bont C. (2007), Market Research, Handbook, 5th Edition

The developed Perceived Utility dimensions will allow to perform a comparison between different actions and also will provide the opportunity to see if the ISA programme objectives have been met (from the user point of view).

3.2 SURVEY MEASUREMENTS

In the data analysis, the core types of measurements which are performed include the Usefulness Score, the Value Score, the User Satisfaction Score, the Net Promoter Score and the Overall Score for Perceived Utility. The survey measurements are divided into two groups: action level measurements and Perceived Utility level measurements.

Action level measurements:

- The Usefulness Score indicates the respondents' evaluation of how useful the action is. The Usefulness Score is calculated taking into account a mean value from a single question: *"In overall, how useful is the "Interoperable test bed" and its documentation/test registry repository and/or demo installation package in your work?"*
- Actions strengths, weaknesses, opportunities and threats: Statements are located in quadrants, based on the dimensions' conformity and dimensions' importance calculated mean values. The quadrants highlight the weak and strong aspects of the action, as well as threats and opportunities.
- Statements based on action objectives shows the respondents' evaluation to what extent the action's objectives have been achieved.

Perceived Utility level measurements:

- The Value Score shows the action's compliance to the dimensions. Two aspects are considered for each dimension. On one side, the importance of the dimension for the users is assessed. On the other side we measure if the action is compliant with the dimension. This section includes statement mapping to dimensions, dimensions conformity results, criterion score and aggregation.
- The User Satisfaction Score shows how satisfied the respondents are with the action. The User Satisfaction Score is assessed with the reference to the results of the dimension importance and conformity evaluation. The User Satisfaction Score is measured at the individual level for each of the survey respondents via the identification of the important dimensions for that particular respondent.
- The Net Promoter Score® (NPS) is a widely used management tool that helps evaluate the loyalty of a customer relationship. In order to evaluate the NPS, the question *"how likely the respondent would recommend the particular action's output to others"* is asked.
- The Overall Score is used to get a single score that would describe the overall Perceived Utility of the action. In order to determine the Overall Score, the average value of the Value Score, the Net

Promoter Score, the Usefulness Score and the User Satisfaction Score is calculated. To calculate the Overall Score, all measurements are reduced to a five point scale.

3.3 SURVEY ARCHITECTURE

The survey is divided into several sections which are outlined below:

- The demographic profile: for the purpose of identifying the respondent's demographic profile, they are asked to answer several questions. The demographic profile illustrates the diversity of the respondents who participated in the survey.
- Usage of the action outputs: for the purpose of identifying the usage rate of the action outputs, the respondents are asked to answer several questions regarding the usage of every action output. These questions also work as filters, selecting respondents who should evaluate the statements regarding the specific action output.
- The action's Usefulness: for the measurement of the action's usefulness, the respondents are asked to evaluate a single question using a 7-point Likert grading scale⁹.
- The Perceived Utility Measurement: in order to measure the Perceived Utility, the respondents are asked to grade dimensions and statements based on their level of importance and agreement. A 5-point Likert scale⁹ is used as a grading scale. Responses to these questions are used to determine the Value Score, action strengths, weaknesses threats and opportunities, and User Satisfaction Score.
- The Net Promoter Score: there is a single question that measures the Net Promoter Score. By answering this question, the respondents indicate their likelihood of recommending the action's outputs to colleagues or other public administrations.
- Actions strengths, weaknesses, opportunities and threats shows the location of the action statements based on dimension conformity and importance results.
- Statements based on action objectives: in order to evaluate the extent to which these statements conform to the particular action, the respondents are asked to grade statements based on their level of agreement. A 5-point Likert scale⁹ is used as a grading scale.
- The recommendations: the last section includes three open questions for recommendations and opinions regarding the action and the survey.

⁹ A Likert Scale is a widely used scaling method developed by Rensis Likert. Likert scale refers to the use of an ordinal 4- or 5-point rating scale with each point anchored or labeled.

4 SURVEY DATA SUMMARY

This section aims to provide detailed information about the data gathering fieldwork. Table 2 gives an overview of the survey start and end dates, the number of persons to whom the survey was proposed, the amount of responses collected, as well as the survey launching method.

TABLE 2 – ACTION 4.2.6 SURVEY TECHNICAL INFORMATION ABOUT THE FIELDWORK

Start date:	04/12/2015
End date:	11/01/2016
The survey launch method:	E-mail notification
Reminders:	E-mail reminders sent out on 10/12/2015, 19/12/2015 and 05/01/2016
Target Population:	36
Number of reached respondents:	18
Number of suitable respondents for the survey:	11

5 SURVEY RESULTS AND ANALYSIS

This section aims to provide the detailed survey analysis and to present the results.

5.1 DEMOGRAPHIC PROFILE OF RESPONDENTS

The respondents' demographic profile tends to describe the action respondents from the demographical point of view. It illustrates the diversity of the respondents.

Table 3 gives an overview of the demographic profile of the respondents. It is important to take into account that only 11 out of 18 respondents who participated in the survey were identified as suitable (7 responses were excluded from the further analysis due to the respondents' unfamiliarity with this action), thus the percentage value of one respondent is 9.1%.

TABLE 3 - ACTION 4.2.6 DEMOGRAPHIC PROFILE OF RESPONDENTS

RESPONDENT PROFILE			
		Amount	Col %
ALL SUITABLE RESPONDENTS		11	100.0
RESPONDENT GROUP*	System owner requiring testing facilities	1	9.1
	Test bed owner	7	63.6
	Test developer	5	45.5
	Tester	2	18.2
	Other (1 respondent - LSP eSENS)	1	9.1
ORGANISATION	Public administration at national level	3	27.3
	Public administration at regional level	1	9.1
	EU institution	2	18.2
	Academic	2	18.2
	Private sector	3	27.3
LOCATION	Belgium	1	9.1
	Bulgaria	1	9.1
	Germany	2	18.2
	Italy	2	18.2
	Portugal	2	18.2
	Spain	1	9.1
	Sweden	1	9.1
	Other (1 respondent - Turkey)	1	9.1
POSITION LEVEL	Management level	5	45.5
	Technical level	6	54.5

Base: all respondents, n=11

* There were multiple choices possible for these questions. This explains why the percentage of responses can exceed 100%.

5.2 USAGE OF THE ACTION

The usage profile provides an overview of the usage rate of the action. Table 4 illustrates the diversity of the action's output usage. The respondents are provided with a set of questions which are asked to determine the usage rate also filters the respondents, thus leading them to evaluate the outputs of the action which they have used.

Table 4 gives an overview of the frequency of using the Interoperable test bed. It is worth taking into account that only 11 out of 18 respondents who participated in the survey were identified as suitable for completing it, thus the percentage value of one respondent is 9.1%.

TABLE 4 – ACTION 4.2.6 USAGE OF INTEROPERABLE TEST BED

USAGE PROFILE			
		Amount	Col %
ALL SUITABLE RESPONDENTS		11	100.0
INTENSITY WORKING WITH THE INTEROPERABLE TEST BED	Just looked at it	3	27.3
	Once	1	9.1
	Occasionally	6	54.5
	Regularly	1	9.1
INTEROPERABLE TEST BED DOCUMENTATION USAGE*	Assessment of GITB service compliance for Minder	3	27.3
	General applicability of GITB to ITB	8	72.7
	Other (1 respondent - Technical and hardware requirements documentation)	1	9.1
INTEROPERABLE TEST BED DATA IMPLEMENTATION*	None	7	63.6
	Assessment of GITB service compliance for Minder	2	18.2
	General applicability of GITB to ITB	3	27.3
DEMO INSTALLATION PACKAGE USAGE	Never heard about it before	4	36.4
	Just heard, but don't use/work with it	2	18.2
	Just looked at it	2	18.2
	Installed it	1	9.1
	Installed it and worked with it occasionally	2	18.2
TEST REGISTRY AND REPOSITORY USAGE	Never heard about it before	2	18.2
	Just heard, but don't use/work with it	4	36.4
	Just looked at it	1	9.1
	Once	1	9.1
	Occasionally	3	27.3

Base: all respondents, n=11

* There were multiple choices possible for these questions. This explains why the percentage of responses can exceed 100%.

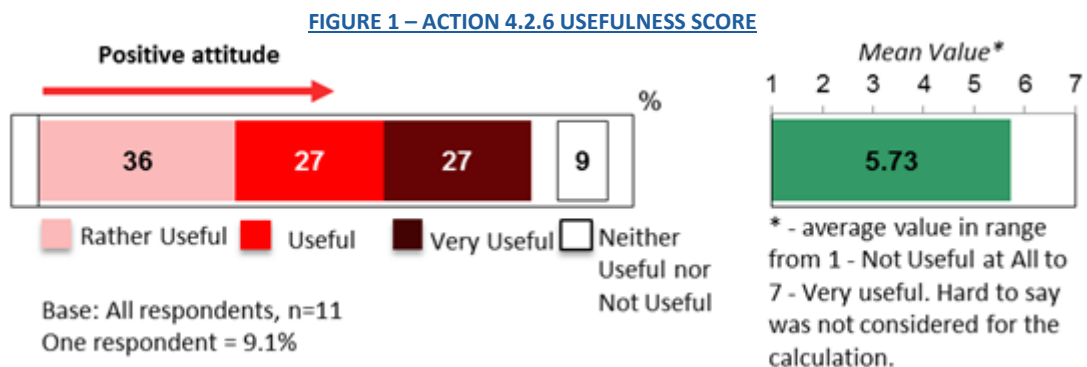
5.3 USEFULNESS SCORE

The Usefulness Score is calculated taking into account a single question: “In overall, how useful is the “Interoperable test bed” and its documentation/test registry repository and/or demo installation package in your work?”

Each respondent is requested to provide his/her opinion using the 7-point Likert grading scale. For evaluation of the usefulness, a grading scale is used with values ranging from “Very Useful” to “Not Useful at All”. An additional “Hard to Say” option is provided, however this score is excluded from the score calculations. Before performing the survey data calculations, the 7-point Likert scale values are interpreted as numeric values:

- 7 – Very Useful;
- 6 – Useful;
- 5 – Rather Useful;
- 4 – Neither Useful nor Not Useful;
- 3 – Rather Not Useful;
- 2 – Not Useful;
- 1 – Not Useful at All;
- 0 – Hard to Say (*is not considered for the calculation*).

In order to have an overview of the positive (Rather Useful, Useful and Very Useful) and negative (Rather Not Useful, Not Useful and Not Useful at All) attitude proportions, the bars in blue represent the negative attitude, whereas the bars in pink and red represent the positive one. In addition, a neutral opinion (the bars in white) is presented separately on the right. An explanatory legend with colour codes represents the data which is available. The average mean value is presented on the right side of the figure.



The survey results show that the Interoperable test bed is useful to the respondents in their work; none of the respondents provided a negative response. To more than one fourth of the respondents, the Interoperable test bed appears to be very useful. The mean value is **5.73** points and it is between 5 - 'Rather Useful' and 6 - 'Useful' values.

5.4 PERCEIVED UTILITY MEASUREMENTS

This section aims to provide a detailed Perceived Utility measurement analysis and to present the results.

5.4.1 Perceived Utility Value Score

This section includes the analysis and results of the Value Score and is structured into two main sections: the dimensions importance and conformity via statements.

5.4.1.1 PERCEIVED UTILITY DIMENSIONS IMPORTANCE

Prior to the evaluation of the dimensions' conformity to the outputs of the action, it is essential to initially ascertain whether these dimensions are important to the respondents while working with the action. If a specific dimension is important to respondents, then it is essential that its conformity assessment is positive. However, if a dimension is not important to respondents, then it should not be considered as the action's weakness because of the non-compliance with the outputs of the action.

Three Perceived Utility dimensions are included in the survey: Collaboration, Potential Re-usability and Sustainability. This section describes the respondent's answers regarding the importance of the dimensions.

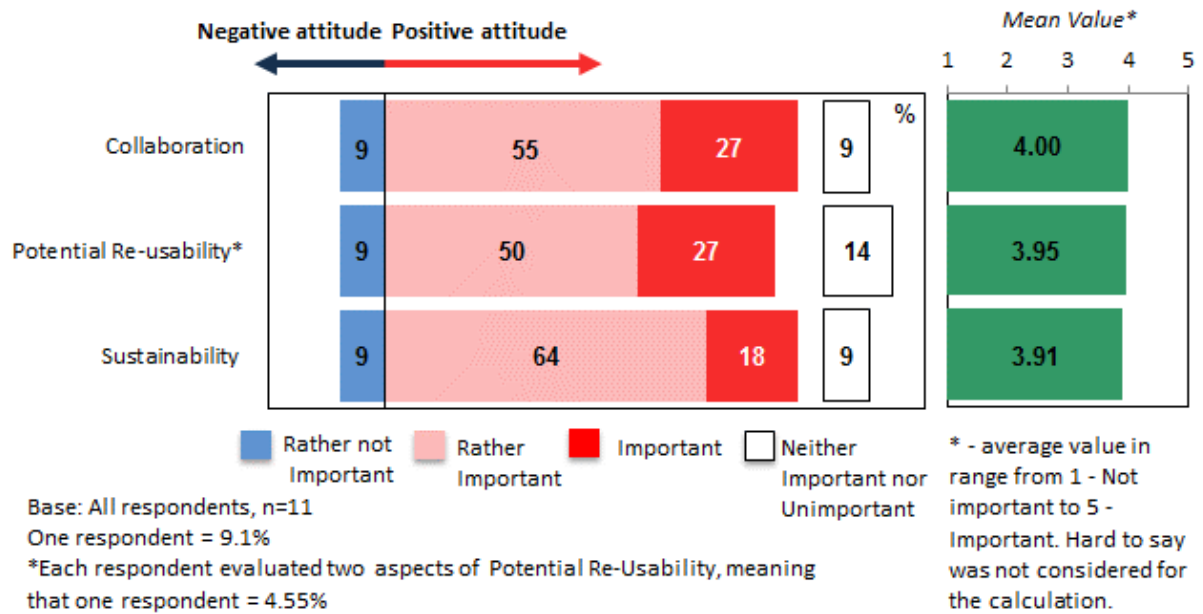
Each respondent is requested to provide his/her opinion using the 5-point Likert grading scale. For the dimension importance evaluation, a grading scale with values ranging from 'Important' to 'Not important' is used. An additional 'Hard to Say/Not Applicable' option is provided, however this score is excluded from the score calculations. Before performing the survey data calculations, the 5-point Likert scale values are interpreted as numeric values:

- 5 – Important;
- 4 – Rather Important;
- 3 – Neither Important nor Unimportant;
- 2 – Rather not Important;
- 1 – Not Important;
- 0 – Hard to Say/ Not Applicable (*is not considered for the calculation*).

In order to have an overview of the positive and negative attitude proportions, the bars in blue represent the negative attitude (answers 'Not Important' and 'Rather not Important'), whereas the bars in pink/red represent the positive one (answers 'Rather important' and 'Important'). In addition, a neutral opinion (the bars in white) is presented separately on the right. An explanatory legend with colour codes represents the available data. The average mean value for each of the dimensions is presented on the right side of the figure.

FIGURE 2 – ACTION 4.2.6 PERCEIVED UTILITY DIMENSIONS IMPORTANCE RESULTS

"How important to you are these factors when using "Interoperable test bed", taking into consideration all the outputs – documentation/test registry repository and/or demo installation package?"



The survey results indicate that the Collaboration, Potential Re-usability and Sustainability dimensions are equally important Interoperable test bed dimensions. None of the respondents evaluated these dimensions as not important, and at least three quarters of the respondents consider them as rather important or important. The mean value for Collaboration is 4.00, for Potential Re-Usability it is 3.95 and for Sustainability it is 3.91. Also the mean value of the lowest evaluated dimension is still higher than the average value/neutral value (3 - 'Neither Important nor Unimportant'), meaning that all dimensions are important.

5.4.1.2 PERCEIVED UTILITY DIMENSIONS CONFORMITY

In order to measure the Perceived Utility dimensions' (Collaboration, Potential Re-usability and Sustainability) conformity to the action, a set of descriptive statements was developed for each dimension. By evaluating the statement conformity to the action, the extent to which the dimensions correspond to the ISA programme's objectives is measured.

This section provides an analysis of the statements. It starts with statement mapping to dimensions, which is followed by the analysis of the Perceived Utility dimension conformity statements. Finally, the last section provides an overview of the statement conformity scores, which are summarised in groups according to the dimensions.

5.4.1.2.1 PERCEIVED UTILITY STATEMENT MAPPING TO DIMENSIONS

The survey statements are linked to each of the action’s outputs: documentation, demo installation package and test registry and repository.

Table 5 gives an overview of the statements representing each dimension. Potential Re-usability and Sustainability dimensions are represented by 5 statements each, while the Collaboration dimension is represented by 3 statements.

TABLE 5 – ACTION 4.2.6 STATEMENT MAPPING TO PERCEIVED UTILITY DIMENSIONS

	Statement	Dimension
1	Overall, the action activities help save costs	Potential Re-usability
2	Overall, the action activities help save time	Potential Re-usability
3	Overall, the Demo Installation Package supports effective reuse of services	Potential Re-usability
4	You are planning to use the Test Registry and Repository in future	Potential Re-usability
5	Overall, the Test Registry and Repository supports effective reuse of tools	Potential Re-usability
6	You are planning to use the documentation in future	Sustainability
7	Overall, the documentation is easily (re)usable	Sustainability
8	You are planning to use the Demo Installation Package in future	Sustainability
9	The Demo Installation Package represents a sustainable solution that will also be relevant in future	Sustainability
10	The Test Registry and Repository provides sustainable solutions that will also be relevant in future	Sustainability
11	The documentation helps successfully cooperate with other public administrations/departments	Collaboration
12	Overall, the Test Registry and Repository supports effective electronic cross-border and cross sector interaction	Collaboration
13	The Test Registry and Repository supports the implementation of European community policies and activities	Collaboration

5.4.1.2.2 PERCEIVED UTILITY DIMENSIONS CONFORMITY RESULTS

This section shows the evaluation of the statements on a scale of the following values: 'Disagree' (note - none of the respondents evaluated any of the statements with this value; this explains its absence in the Figure 3), 'Rather Disagree' (in light blue colour), 'Neither Agree nor Disagree' (note - none of the respondents evaluated any of the statements with this value as well), 'Rather Agree' (in pink colour) and 'Agree' (in red colour).

FIGURE 3 – ACTION 4.2.6 PERCEIVED UTILITY DIMENSIONS CONFORMITY RESULTS

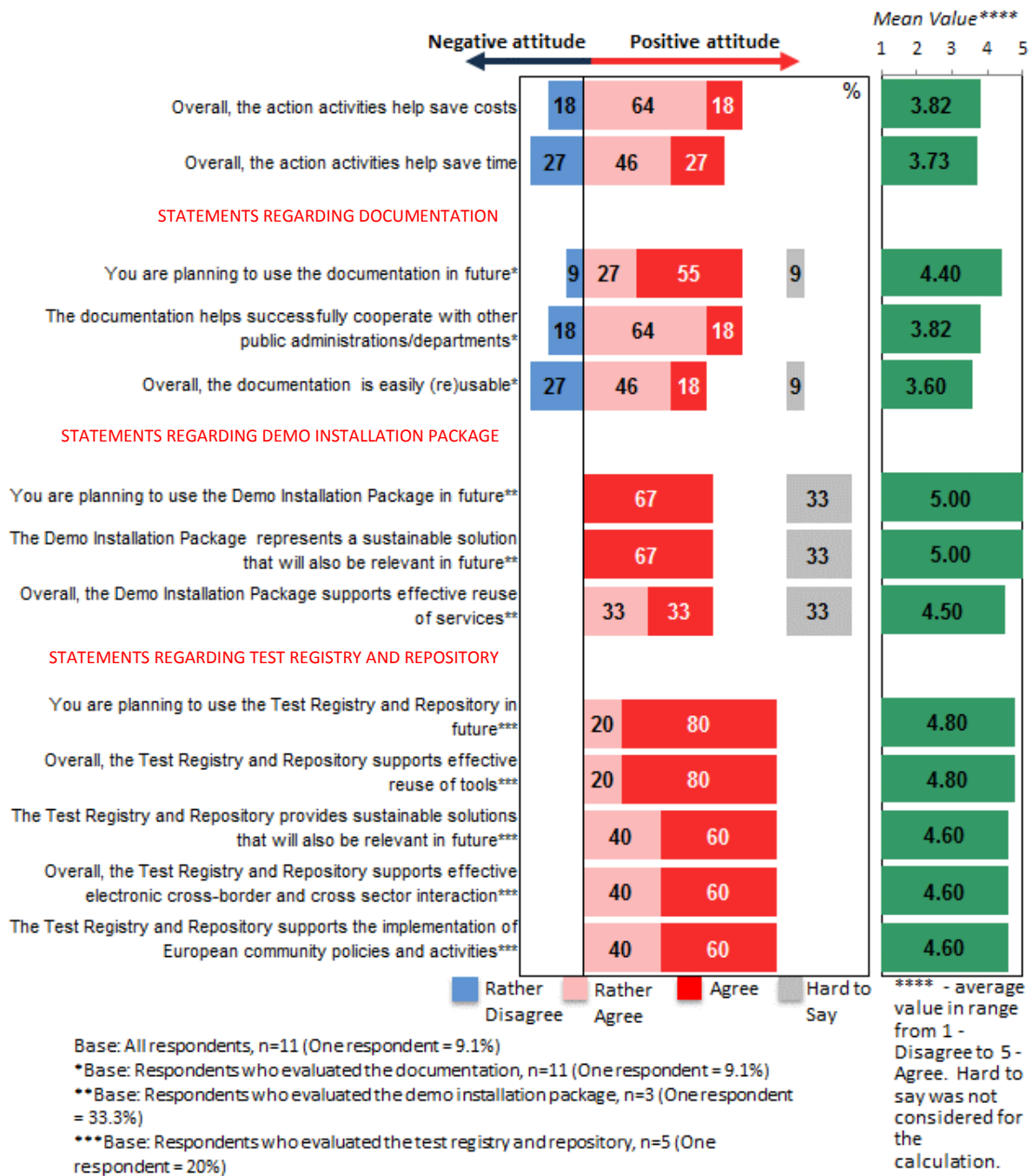


Figure 3 indicates that all of the statements are evaluated as more or less relevant to the Interoperable test bed outputs as the average value is higher than a neutral value of 3 ('Neither Agree nor Disagree'). The most relevant statement to this action for all three outputs is:

- 'You are planning to use [action outputs] in future'.

When comparing the statements, the difference in bases should be taken into account (some statements were evaluated by 3 respondents, others by 5 and 11), as well as the total amount of respondents.

Table 6 provides an overview of the statement conformity scores, which are summarised by dimensions. To calculate these scores, the average values of all the relevant dimension statements are taken into account.

Table 6 also provides an overview of the additional statistical calculations¹⁰ - mode, standard deviation and standard error. With reference to the theory used in business research methods,¹¹ it is concluded that for statistically meaningful calculations, the minimum respondent number is equal to or greater than ten per statement.

TABLE 6 – ACTION 4.2.6 ADDITIONAL STATISTICAL CALCULATIONS FOR PERCEIVED UTILITY DIMENSIONS

	Dimension	MEAN	MODE	StDev	StErr
Per dimension	Collaboration	4.34	4	0.99	0.30
	Sustainability	4.52	4	1.13	0.26
	Potential Re-usability	4.33	4	1.07	0.23
Total Criterion Score		4.40	4	1.06	0.26

The survey results show that the dimension conformity score is evaluated as equally high for all of the dimensions, as they all are in the range of the Standard Error. Also, none of the dimensions is irrelevant, since the value is higher than the neutral value of 3 - 'Neither agree nor disagree'.

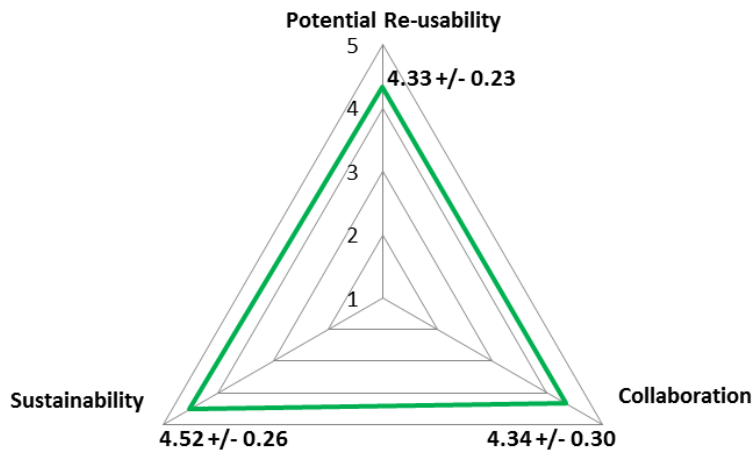
5.4.1.2.3 PERCEIVED UTILITY CRITERION SCORE AGGREGATION

Figure 4 provides a visual overview of the dimension conformity scores.

¹⁰ Dictionary of statistics & methodology: a nontechnical guide for the social sciences (page 226).

¹¹ Cooper D. R., Schindler P. S. (2013), Business Research Methods, 12th Edition

FIGURE 4 – ACTION 4.2.6 PERCEIVED UTILITY CRITERION SCORE AGGREGATION



5.4.2 Perceived Utility User Satisfaction Score

The User Satisfaction Score shows how satisfied and happy the respondents are with the performance of a specific action. The User Satisfaction Score is expressed as a percentage from 0 to 100, where 0 signifies that there are no satisfied and happy respondents, whereas 100 signifies all respondents are satisfied and happy with the work performed by the action.

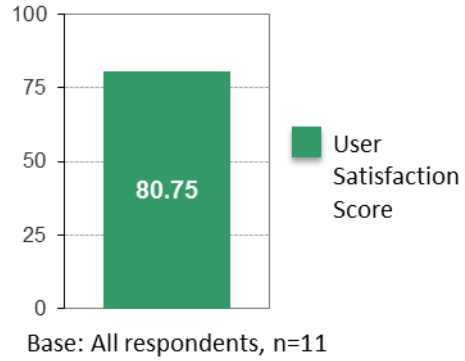
The User Satisfaction Score is assessed with reference to the results of the dimension importance and dimension conformity evaluation. The User Satisfaction Score is measured at the individual level for each of the survey respondents via identification of the important dimensions for that particular respondent.

To increase the accuracy of the calculation, a specific weight coefficient is applied to the dimensions. To those dimensions which respondents evaluated as “Important” a weight coefficient of 1 was applied, while a coefficient of 0.5 is applied to the dimensions which respondents evaluated as “Rather Important”. A coefficient of 0 is applied to all the other dimensions. Finally, all the individual values are summed.

As the next step, an analysis of the statements which represent these identified dimensions is performed. If a respondent claimed that a particular statement fully corresponded to the specific dimension (value 5 – ‘Agree’), then a coefficient of 100 (100% eligibility) is assigned. If evaluated with 4 – ‘Rather Agree’, a coefficient of 75 applies, if evaluated with 3 – ‘Neither Agree nor Disagree’, a coefficient of 50 applies, if evaluated with 2 – ‘Rather Disagree’, a coefficient of 25 applies, and in the case it was evaluated with 1 – ‘Disagree’, the coefficient is 0.

FIGURE 5 – ACTION 4.2.6 PERCEIVED UTILITY USER SATISFACTION SCORE

Figure 5 shows that the User Satisfaction Score is **80.75**. The result indicates a high level of respondent satisfaction with the outputs of Interoperable test bed.



5.4.3 Perceived Utility Net Promoter Score

The Net Promoter Score® (NPS) is a widely used management tool that helps evaluate the loyalty of a customer relationship¹². This management tool has been adapted to suit the ISA programmes’ Evaluation and Monitoring activities and measures the overall respondents’/stakeholders’ experience and loyalty to a specific ISA action.

In order to evaluate the NPS, the question ‘how likely the respondent would recommend the particular action’s output to others’ is asked. The assessment is done on a scale from 0 to 10, where 0 represents the answer ‘Not likely at all’ and 10 – ‘Extremely likely’¹³. After the data analysis, the respondents are classified as follows:

- **Promoters** (numeric values from 9 - 10) - loyal users who will keep using the action’s final outcome and refer others, promoting the usage of the action's outcomes;
- **Passives** (numeric values from 7 - 8) - satisfied but unenthusiastic users who will most probably not recommend the action's outcomes to others;
- **Detractors** (numeric values from 0 - 6) - unhappy users who can damage the image and decrease the usage of the action's outcomes.

The NPS final score calculation is done based on the following formula:

$$\text{NPS} = \% \text{ of Promoters} - \% \text{ of Detractors}^{13}$$

The result can range from a low of -100 (every customer is a Detractor) to a high of +100 (every customer is a Promoter).

FIGURE 6 – ACTION 4.2.6 PERCEIVED UTILITY NET PROMOTER SCORE

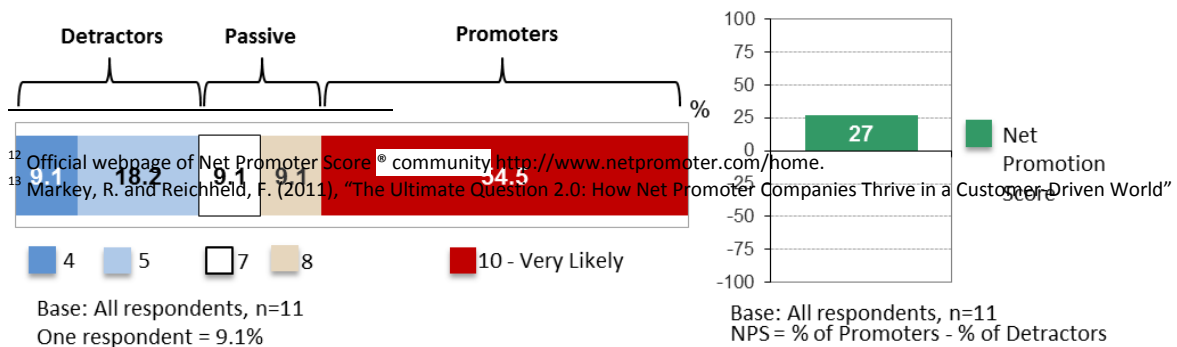


Figure 6 shows that more than half of the respondents (54.5%) are Promoters of the action. Considering the Perceived Utility value of the Interoperable test bed, it is very likely that the respondents would recommend the Interoperable test bed outputs - documentation/test registry repository and/or demo installation package to colleagues or other public administrations. Slightly more than one quarter (27.3%) of the respondents would not recommend the Interoperable test bed's outputs to colleagues or other public administrations. According to the NPS terminology, such respondents can be identified as Detractors. The Net Promoter Score value is calculated as the percentage difference between the Promoters and Detractors. In the case of the Interoperable test bed the NPS is 27 (NPS is expressed in whole numbers). This indicator can be assessed as good because of a positive value (above the neutral value of 0).

5.4.4 Overall Perceived Utility Score

Referring to the performed measurements described earlier, namely, the Usefulness Score, the Value Score, the User Satisfaction Score and the NPS, an Overall Perceived Utility Score is calculated.

To calculate the Overall Perceived Utility Score, all measurements are reduced to a five point scale (the statements used to calculate the Value Score are already expressed using a scale from 1 to 5, the Usefulness Score had values from 1 to 7, NPS - from -100 to +100, and the User Satisfaction Score - from 0 to 100). In order to determine the Overall Perceived Utility score, the average value of these four measurements is calculated. To reduce any linear scale to a different linear scale the following formula¹⁴ is used:

$$Y = (B - A) * (x - a) / (b - a) + A$$

- Y = Value after reducing to a five point scale
- x = Value in the initial scale
- B = The highest value of the new scale (in this case it is 5, as we are reducing other scales to a five point scale)
- A = The lowest value of the new scale (in this case it is 1, as we are reducing other scales to a five point scale)
- b = The highest value of the original scale (for Net Promoter Score and User Satisfaction Score it is +100, for Usefulness Score it is 7)
- a = The lowest value of the original scale (for the Net Promoter Score it is -100, for the User Satisfaction Score it is 0 and for the Usefulness Score it is 1)

Example of reducing Net Promoter Score to a five point scale:

$$(5-1) * (27.20 - (-100)) / (100 - (-100)) + 1 = 4 * 127.20 / 200 + 1 = 508.8 / 200 + 1 = 2.544 + 1 = 3.544$$

¹⁴ Transforming different Likert scales to a common scale. IBM. Retrieved February 04. 2016., from <http://www-01.ibm.com/support/docview.wss?uid=swg21482329>

TABLE 7 – ACTION 4.2.6 OVERALL PERCEIVED UTILITY SCORE CALCULATION

NAME OF THE SCORE	ORIGINAL VALUE	VALUE AFTER REDUCING TO A FIVE POINT SCALE
Usefulness Score	5.73	4.15
Value Score	4.40	4.40
User Satisfaction Score	80.75	4.23
Net Promoter Score	27	3.54
OVERALL PERCEIVED UTILITY SCORE		4.08

The survey results show that on a five point scale the Value Score has the highest score, which indicates the significant dimension conformity to the action. The Net Promoter Score has the lowest score, yet in general it is positive, indicating that a higher proportion of respondents would recommend the Interoperable test bed outputs - documentation/test registry repository and/or demo installation package to colleagues or other public administrations.

5.5 ACTIONS STRENGTHS, WEAKNESSES, OPPORTUNITIES AND THREATS

When analysing the data results of the dimensions' conformity versus the dimensions' importance, the action's strengths, weaknesses, opportunities and threats can be identified.

Statements are located in quadrants, based on the dimensions' conformity and dimensions' importance calculated mean values. The quadrants highlight the weak and strong aspects of the action, as well as threats and opportunities.

In general, all the statements that are attributed to the action can be grouped into four categories:

- Strengths – Essential to respondents and relevant to the action (1st quadrant);
- Weaknesses – Essential to respondents but not relevant to the action (2nd quadrant);
- Threats – Not essential to respondents and not relevant to the action (3rd quadrant);
- Opportunities – Not essential to respondents but relevant to the action (4th quadrant).

Three colours are used to identify Perceived Utility dimensions:

- Dark blue: Potential Re-usability;
- Red: Sustainability;
- Brown: Collaboration.

As seen in Figure 7, all 13 statements are evaluated as essential to respondents and relevant to the action - all of them are placed in the 1st quadrant and are identified as strengths of the Interoperable test bed.

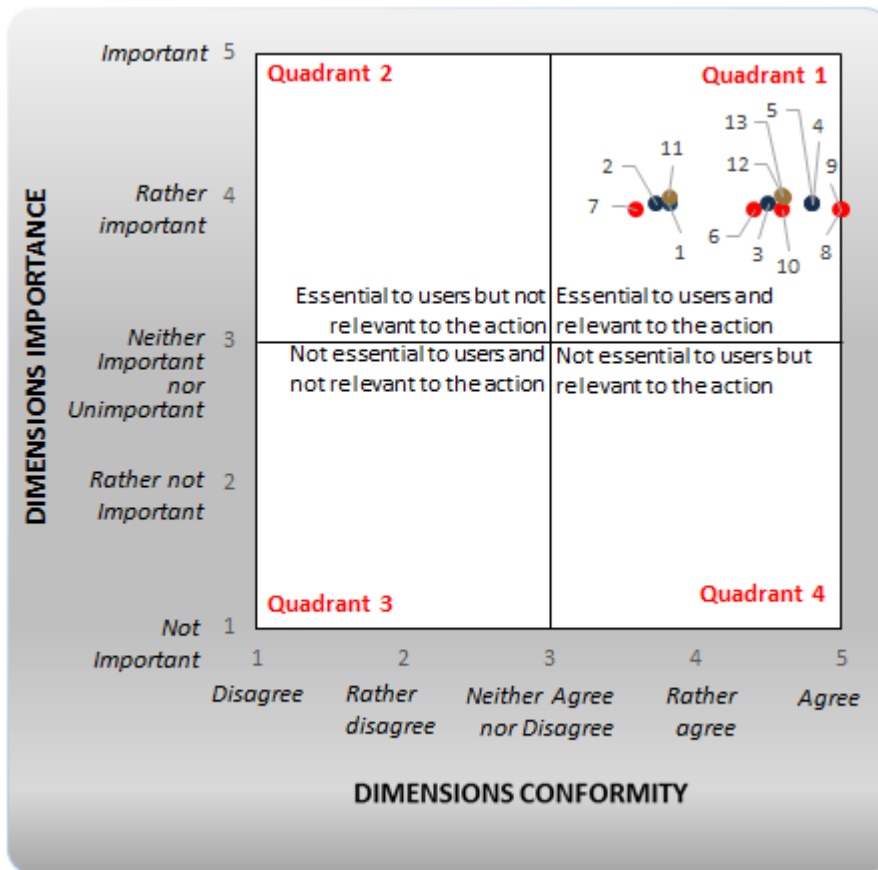
When comparing different statements, it is evident that the following two statements are the action's most important strengths (the most relevant to the action and important to the respondents):

- *'You are planning to use the Demo installation package in future'* (statement 8) and
- *'The Demo Installation Package represents a sustainable solutions that will also be relevant in future'* (statement 9).

As a relatively less relevant (but not irrelevant, because the average score is higher than 3 - point mark) are the following statements:

- *'Overall, the action activities help save time'* (statement 2) and
- *'Overall, the documentation is easily (re)usable'* (statement 7).

FIGURE 7 – ACTION 4.2.6 ACTIONS STRENGTHS, WEAKNESSES, OPPORTUNITIES AND THREATS



I. Potential Re-usability:

- 1 - Overall, the action activities help save costs
- 2 - Overall, the action activities help save time
- 3 - Overall, the Demo Installation Package supports effective reuse of services
- 4 - You are planning to use the Test Registry and Repository in future
- 5 - Overall, the Test Registry and Repository supports effective reuse of tools

II. Sustainability:

- 6 - You are planning to use the documentation in future
- 7 - Overall, the documentation is easily (re)usable
- 8 - You are planning to use the Demo Installation Package in future
- 9 - The Demo Installation Package represents a sustainable solution that will also be relevant in future
- 10 - The Test Registry and Repository provides sustainable solutions that will also be relevant in future

III. Collaboration:

- 11 - The documentation helps successfully cooperate with other public administrations/departments
- 12 - Overall, the Test Registry and Repository supports effective electronic cross-border and cross sector interaction
- 13 - The Test Registry and Repository supports the implementation of European community policies and activities

5.6 STATEMENTS BASED ON ACTION OBJECTIVES

For the purpose of describing the action's objectives, statements based on action objectives are designed for this survey. The respondents are asked to evaluate the extent to which these statements conform to the particular action, namely, if the action's objectives have been achieved.

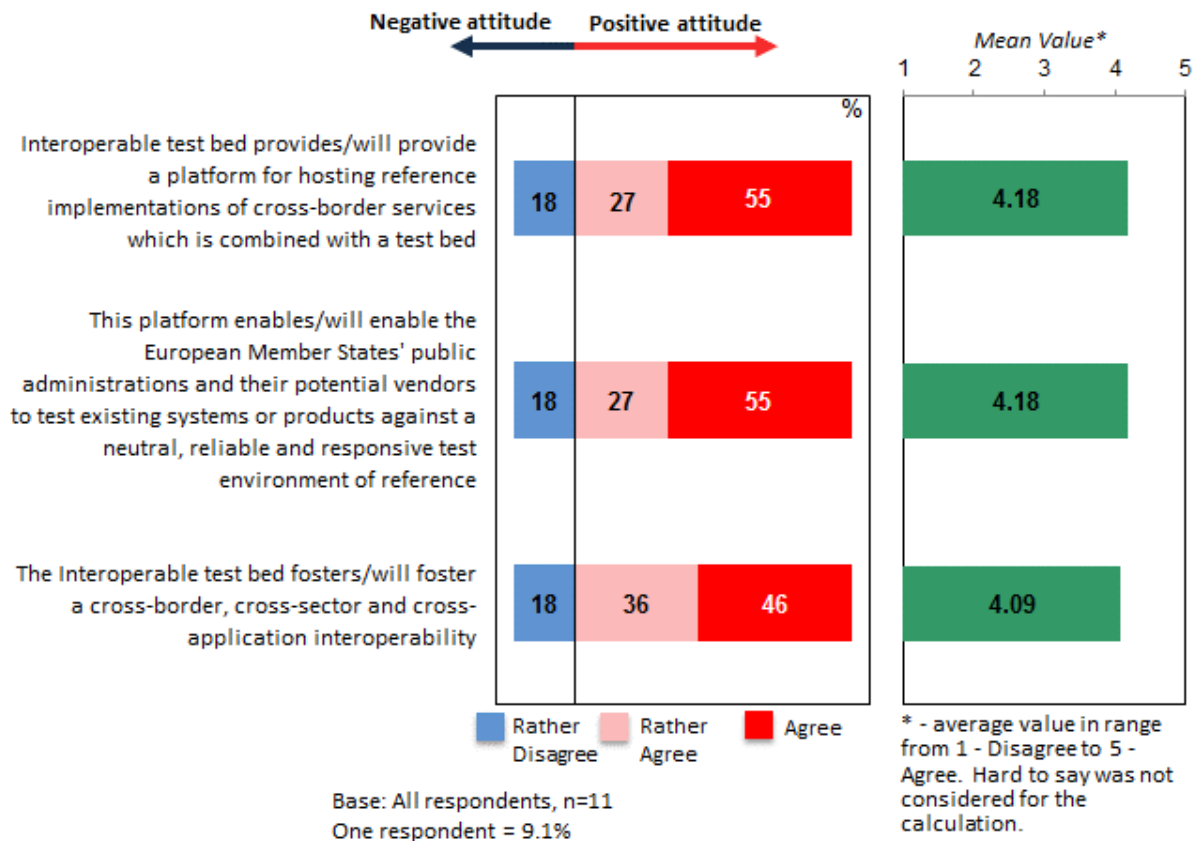
Each respondent is requested to provide his/her opinion using the 5-point Likert grading scale. For the dimension conformity evaluation, a grading scale with values ranging from 'Agree' to 'Disagree' is applied. An additional 'Hard to Say/Not Applicable' option is provided, however this score is excluded from the score calculations. Before performing the survey data calculations, the 5-point Likert scale values are interpreted as numeric values:

- 5 – Agree;
- 4 – Rather Agree;
- 3 – Neither Agree nor Disagree;
- 2 – Rather Disagree;
- 1 – Disagree;
- 0 – Hard to Say/ Not Applicable (*is not considered for the calculation*).

In order to have an overview of the positive and negative attitude proportions, the bars in blue represent the negative attitude (answers 'Disagree' and 'Rather Disagree'), whereas the bars in pink/red represent the positive ones (answers 'Agree' and 'Rather Agree'). In addition, a neutral opinion (the bars in white) is presented separately on the right. An explanatory legend with colour codes represents the available data. The average mean value for each of the dimensions is presented on the right side of the figure.

Figure 8 shows the respondents' answers to the question: *"To what extent do you agree with the following statements, taking into consideration the "Interoperable test bed" and its documentation/test registry repository and/or demo installation package?"*

FIGURE 8 – ACTION 4.2.6 STATEMENTS BASED ON ACTION OBJECTIVES



The survey results indicate that the statements based on action objectives (statements which describe the action's objectives) have been evaluated as relevant to the action. The mean values for these statements are 4 and higher, which is between the 'Rather Agree' and 'Agree' answer options. The statements:

- 'This platform enables/will enable the European Member States' public administrations and their potential vendors to test existing systems or products against a neutral, reliable and responsive test environment of reference' and
- 'Interoperable test bed provides/will provide a platform for hosting reference implementations of cross-border services which is combined with a test bed'

have been evaluated identically - more than half of the respondents (55%) answered that they agree to the fact that these statements are relevant (the mean value for both is **4.18**). The third statement:

- 'The Interoperable test bed fosters/will foster a cross-border, cross-sector and cross-application interoperability'

has been evaluated slightly lower - the mean value is **4.09** (which corresponds to the 'Rather Agree' assessment).

5.7 RESPONDENT RECOMMENDATIONS AND OPINIONS

This section provides an overview of the recommendations and opinions received. It should be noted that each response is given by a single survey respondent, which means that the number of different answers to each question is the same as the number of respondents who had an opinion or a recommendation to the specific question.

TABLE 8 – ACTION 4.2.6 RECOMMENDATIONS AND OPINIONS

<i>"Do you have any recommendations to improve the "Interoperable test bed", taking into consideration the "Interoperable test bed" and its documentation/test registry repository and/or demo installation package?"</i>
Even if the artefacts can be shared and stored in a common repository, there is a growing number of private companies providing testing services that should be leveraged and protected avoiding implementing public services that can compete with these initiatives. Creating a market of private companies offering testing services can help the deployment of electronic systems in Europe.
Some FAQs or a quick guide may help newcomers to learn more about the test bed.
We will provide extensive feedback for GITB specification when we finish Minder service level GITB compliance.
<i>"What are the main benefits or the most valuable things about the "Interoperable test bed"?"</i>
A common repository of artefacts for the projects developed within the EU is a main benefit. Providing a general test bed platform is not beneficial for this sector to take-up.
To facilitate the methods of testing software and provide an easy and sustainable way of communicate between cross border systems and promoting at the same time the cooperation between public and private sector.
Consistent use of metadata about test resources.
Interoperability and facilitation of reusability of assets.
<i>"Do you have any other recommendations to share with us?"</i>
In 2016, we will try to use the frameworks this Action provides in our work, helping to expose our testing frameworks for other sectors to reuse. We have been doing a lot of planning work to now and know that this Action is helping to shape our approach and ensure cross-sector interoperability can be established, where relevant.
You can put definitions of 'interoperable test bed', 'demo installation package' in this and future surveys.

6 SURVEY CONCLUSION AND RECOMMENDATIONS

The objective of this survey was to evaluate the Perceived Utility of Action 4.2.6 – Interoperable test bed. The following conclusions have been drawn based on the analysis performed:

- The ISA Action 4.2.6 - Interoperable test bed received a **positive Perceived Utility assessment with an Overall Perceived Utility Score of - 4.08 out of 5**. The high Overall Perceived Utility Score and individual parameters indicate that respondents consider the Interoperable test bed useful.
- Positive evaluation of the statements based on action objectives indicates that Interoperable test bed corresponds to the action's specific objectives.
- The majority of respondents (54.5%) would recommend the outputs of the Interoperable test bed including the documentation, test registry repository and/or the demo installation package to colleagues or other public administrations.
- The results show that the Interoperable test bed is perceived as equally beneficial in terms of Collaboration, Potential Re-usability and Sustainability.
- The findings represent that the weakest aspect of the Interoperable test bed is the documentation output. 27% percent of the respondents think that the documentation is not easily usable.

Based on the conclusions drawn, CGI-ACN adduces the following recommendations:

- Improvements in the documentation such as making it more readable and user friendly would be of benefit to the action.
- Documentation might benefit from the addition of a short and simple video guide and a frequently asked question section for the newcomers.

7 APPENDIX

7.1 RAW DATA EXPORT

The attached file contains the survey result export.



Raw Data.xls

7.2 GLOSSARY

- A Likert Scale is a widely used scaling method developed by Rensis Likert. Likert scale refers to the use of an ordinal 4- or 5- point rating scale with each point anchored or labelled.
- The mean¹⁰ (average) is the most popular measure of location or central tendency; has the desirable mathematical property of minimizing the variance. To get the mean, you add up the values¹⁰ or each case and divide that sum by the total number of cases;
- Mode¹⁰ refers to the most frequent, repeated or common value in the quantitative or qualitative data. In some cases it is possible that there are several modes or none;
- The Net Promoter Score® (NPS) is a widely used management tool that helps evaluate the loyalty of a customer relationship. Customers are classified as Promoters, Passive and Detractors.
- 'Perceived Quality' is defined as the extent to which the outputs of an ISA action are meeting its direct beneficiaries' expectations;
- Standard deviation¹⁰ shows the spread, variability or dispersion of scores in a distribution of scores. It is a measure of the average amount the scores in a distribution deviate from the mean. The more widely the scores are spread out, the larger the standard deviation;
- Standard error¹⁰ is the standard deviation of the sampling distribution of a statistic. It is a measure of sampling error; it refers to error in estimates due to random fluctuations in samples. It goes down as the number of cases goes up. The smaller the standard error, the better the sample statistic is as an estimate of the population parameter – at least under most conditions;
- 'Perceived Utility' is defined as the extent to which the effects (impact) of an ISA action correspond with the needs, problems and issues to be addressed by the ISA programme;