Project Number 318736

D6.5 – External Industry Evaluation Results

Version 1.3
27 July 2015
Final

Public Distribution

The Open Group

Project Partners: **Centrum Wiskunde & Informatica, SOFTEAM, Tecnalia Research and Innovation, The Open Group, University of L'Aquila, UNINOVA, University of Manchester, University of York, Unparallel Innovation**

PROJECT PARTNER CONTACT INFORMATION

Centrum Wiskunde & Informatica Jurgen J. Vinju Science Park 123 1098 XG Amsterdam, Netherlands Tel: +31 20 592 4102 E-mail: jurgen.vinju@cwi.nl	**Softeam** Alessandra Bagnato Avenue Victor Hugo 21 75016 Paris, France Tel: +33 1 30 12 16 60 E-mail: alessandra.bagnato@softeam.fr
Tecnalia Research and Innovation Jason Mansell Parque Tecnologico de Bizkaia 202 48170 Zamudio, Spain Tel: +34 946 440 400 E-mail: jason.mansell@tecnalia.com	**The Open Group** Scott Hansen Avenue du Parc de Woluwe 56 1160 Brussels, Belgium Tel: +32 2 675 1136 E-mail: s.hansen@opengroup.org
University of L'Aquila Davide Di Ruscio Piazza Vincenzo Rivera 1 67100 L'Aquila, Italy Tel: +39 0862 433735 E-mail: davide.diruscio@univaq.it	**UNINOVA** Pedro Maló Campus da FCT/UNL, Monte de Caparica 2829-516 Caparica, Portugal Tel: +351 212 947883 E-mail: pmm@uninova.pt
University of Manchester Sophia Ananiadou Oxford Road Manchester M13 9PL, United Kingdom Tel: +44 161 3063098 E-mail: sophia.ananiadou@manchester.ac.uk	**University of York** Dimitris Kolovos Deramore Lane York YO10 5GH, United Kingdom Tel: +44 1904 325167 E-mail: dimitris.kolovos@york.ac.uk
Unparallel Innovation Nuno Santana Rua das Lendas Algarvias, Lote 123 8500-794 Portimão, Portugal Tel: +351 282 485052 E-mail: nuno.santana@unparallel.pt	

DOCUMENT CONTROL

Version	Status	Date
0.1	First outline and preparation of evaluation categories	3 March 2015
0.8	Initial inputs from evaluators	30 March 2015
1.0	Full draft for partner review	16 April 2015
1.1	Additional clarifications and WTP inputs from evaluators	11 May 2015
1.2	Final version after QA	22 May 2015
1.3	Minor revisions to address final EC review comments	27 July 2015

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This deliverable summarises the evaluation of the OSSMETER platform by software development experts from organisations that are external to the project. A total of 12 experts were selected from different industries and types of organisations. These experts evaluated the OSSMETER platform within the context of their own software development activities to assess its ability to analyse and support the selection of open source project alternatives. The results were favourable across all of the evaluation criteria and useful recommendations were also contributed by the external experts. The experts also provided guidance concerning the economic aspects of the platform as inputs to a sustainable strategy for exploitation.

1. INTRODUCTION

This deliverable describes the results of a set of evaluations of the OSSMETER platform carried out by external software development experts in Europe who are typical target users for the OSSMETER technologies. A total of 12 experts were selected from different industries and organisation types (e.g. vendor, software developer, research) to provide differing views on key aspects of the OSSMETER project results. These organisations have evaluated the OSSMETER platform within the context of their own software development activities to assess the ability of the platform to analyse and support the selection of open source project alternatives. The results are described in the following sections of this deliverable:

- Section 2 provides brief profiles of the organisations of each of the experts

- Section 3 describes the evaluation approach that was utilised in collecting the inputs from the external experts

- Section 4 summarises the outcome of the external evaluations and includes the recommendations and suggestions from the experts

- Section 5 summarises the economic assessment provided by the external experts as guidance in the development of a sustainable economic model for exploitation

- Section 6 provides some overall conclusions

The evaluation results have provided important recommendations both for the further development of the platform and have been used in the preparation of deliverable D7.6 – Final Dissemination and Exploitation Plan.

2. EXTERNAL EVALUATOR ORGANISATION PROFILES

A total of 12 external experts participated in the evaluations of the OSSMETER platform. Figures that characterise the composition of the set of organisations involved in the evaluations are as follows:

- Small to medium size enterprises: 3

- Large multinationals: 2

- Large industrial: 2

- Industrial and academic research centres: 3

- Universities: 2

The two industrial research centres each carry out software development projects funded by industrial organisations in their regions. The third research centre is heavily involved in using OSS project technologies as well as contributing to OSS projects.

In carrying out the evaluations the external experts were asked to consider the OSSMETER platform from the perspective of their own organisation's procedures for adopting open source projects and in relation to other tools and technologies that they may already use. A description of each organisation represented in the evaluations is provided in the following sections.

2.1 SMALL TO MEDIUM SIZE ENTERPRISE

2.1.1 Impronova

Impronova (SE) delivers assessment, training, consultancy, mentoring and support services in the areas of software quality, software process improvement and safety of software systems. Impronova is authorised to lead software assessments of automotive suppliers worldwide on behalf of leading OEMs and also works with other clients to implement processes to meet the requirements of international standard best practice models.

2.1.2 Pragmasis

Pragmasis (PT) is a developer of banking applications used by most banks in Portugal for managing online transactions. The suite of software addresses high volume transactions from the network of ATM machines across Portugal while providing connectivity to transaction networks abroad. The software also handles interbank transactions amongst affiliated financial institutions.

2.1.3 Rheon Media

Rheon Media (UK) are specialists in new media applications providing services within the TV, Broadcast and IPTV arena. Rheon Media clients range from broadcasters to multimedia platform owners to new start-up businesses. The business provides bespoke technical consulting and development services to aid producers of media systems in their tactical and strategic goals.

2.2 LARGE MULTI-NATIONALS

2.2.1 Airbus

Airbus Innovation Works (FR) is the corporate research facility of Airbus with operations in France, Germany, Spain, UK, Singapore and Russia. Its workforce is more is covering all the skills and technology fields that are of critical importance to Airbus product lines including software development technologies and deployment platforms utilised across the company. It is an operational and strategic entity feeding the innovation pipeline from the emergence of new technologies to their maturity and transfer into products.

2.2.2 Thales

Thales Communications and Security (FR) is a company that is part of the Thales Group having 67,000 employees in 56 countries, and addresses activities related to defence, security, and ground transportation. The company develops since many years its activities in civil markets such as civil aviation, radio spectrum monitoring and information technology security.

2.3 LARGE INDUSTRIALS

2.3.1 Intecs

Intecs (IT) is a software company providing leading-edge technological support and products to major European organizations in the design and implementation and verification of complex electronic systems. Intecs has extensive experience in the production of software systems as well as software engineering and quality, and has well-established co-operation with most of the major Italian and European electronic industries for the development of proprietary products. Intecs operates in the Space, Automotive, Telecommunications, ATC and Defence domains.

2.3.2 ISQ

ISQ (PT) provides solutions for clients in all sectors to ensure that their assets and products comply with environmental, quality, health and safety standards and regulations. Founded in 1965, ISQ is a private and independent company providing inspection, testing, training and technical consultancy. ISQ offers high quality and innovative services and solutions, helping customers with infrastructures, equipment, processes and products to meet the standards and regulations for quality, security, environment and social responsibility.

2.4 INDUSTRIAL AND ACADEMIC RESEARCH CENTRES

2.4.1 ATB

ATB (DE) is an RTD institute funded by industry providing core competencies in development of custom and modular software solutions based on latest technologies and implementation of ICT solutions with strong interaction and active customer cooperation. Application areas encompass knowledge management, software systems engineering, systems analysis and design within various domains. Each project foundation stems from a client partnership to address specific technology advances.

2.4.2 FORTH

The Foundation for Research and Technology-Hellas (GR), established in 1983, is one of the largest research centers in Greece with well-organized facilities, highly qualified personnel and a reputation as a top-level research foundation worldwide. The research and technological directions of FORTH focus on areas of major scientific, social, and economic interest, including: Microelectronics, Lasers, Biotechnology, Computer Sciences, Bioinformatics, Robotics, Telecommunications, Computational Mathematics, Chemical Engineering and Human and Social Sciences.

2.4.3 Ikerlan

Ikerlan (ES) is an R&D technology centre specialising in product, process and service innovation helping companies enhance their competitiveness. Ikerlan offers comprehensive solutions combining the six areas where a high level of specialisation is available: embedded systems, power electronics, microtechnologies, energy, mechatronics and advanced manufacturing.

2.5 UNIVERSITIES

2.5.1 HLRS

The High Performance Computing Center Stuttgart (DE) of the University of Stuttgart supports researchers from Germany and Europe as well as industry with leading edge supercomputing technology. It offers state-of-the-art supercomputing infrastructure, which also includes a PRACE-tier-0 Cray XE6 system, providing computational performance of the petaflop performance range.

2.5.2 Brno University of Technology

Brno University of Technology (CZ) is the second largest technical university in the Czech Republic. It comprises 8 faculties with more than 20,000 students and 2,000 staff members. The Faculty of Information Technology (FIT) research activities include FPGA design and development, hardware-accelerated computing, safety and security, large-scale verification systems, high-performance and cloud technologies, big data, machine learning, and knowledge technologies.

3. EVALUATION APPROACH

3.1 MEASUREMENT CRITERIA

The external evaluators were asked to carry out their assessment of the OSSMETER platform on a *pro bono* basis, which required the project to balance the desire for data from other industrial representatives, with not overwhelming the external experts with too much data to collect and possibly getting no inputs due to other time commitments of the external experts.

The partners identified five key evaluations measures that would be of most value to the project and likely that external evaluators with little knowledge of OSSMETER would feel comfortable investing their time to address. The evaluation criteria used were the following:

- Ease of use of the platform
- Performance of the platform
- Usefulness of the code metrics
- Usefulness of the communications metrics
- Overall value of the platform

Each of the evaluators was provided a unique login / password to access the OSSMETER platform, which was the same instantiation hosted by the University of York for the internal industrial evaluations by project partners allowing external evaluators to quickly access open source project data that had already been collected.

3.2 EVALUATION SCALE

The external experts were asked to use a four point Likert[1] scale that also employed the "forced choice" method when providing their assessments. The forced choice meant there was no middle position in the rating and evaluators were required to take a position either favourable or unfavourable towards the different capabilities of the OSSMETER platform as well as an overall view. The reason for using this approach is studies show evaluators often favour the centre rating in a five or seven point scale as a perceived neutral or non-controversial position. The following four-point rating scale was used:

- Poor = 1

- Acceptable = 2

- Good = 3

- Excellent = 4

The evaluators were informed that the project's intended use of the evaluation measures would be in determining the readiness of the platform for industrial use and also the expected interest in the platform capabilities by typical organisations involved with open source software projects.

In addition to the rating of each measure, evaluators were invited to comment on both the reasoning behind the rating selected and also any features or enhancements that were needed. These have been captured and listed under the results for each of the evaluation criteria in no particular order.

4. OUTCOME OF EVALUATIONS

The following sections summarises the ratings that were received from the external experts for each of the five evaluation criteria. The average rating for each evaluation criteria is provided, along with the standard deviation as an indication of the degree of consensus amongst the reviews (a smaller standard deviation indicates higher consensus). The comments and recommendations received from the experts accompanying their ratings are also provided.

4.1 EASE OF USE

The average rating amongst all of the experts for ease of use of the OSSMETER platform was 3.5, which is between Good and Excellent. The standard deviation amongst responses was 0.5. The distribution of ratings amongst the experts is shown in Figure 1.

---

[1] A Likert scale is a psychometric scale commonly involved in research that employs questionnaires and is the most widely used approach to scaling responses in survey research. Respondents specify their level of agreement or disagreement on a symmetric agree-disagree scale for a series of statements.
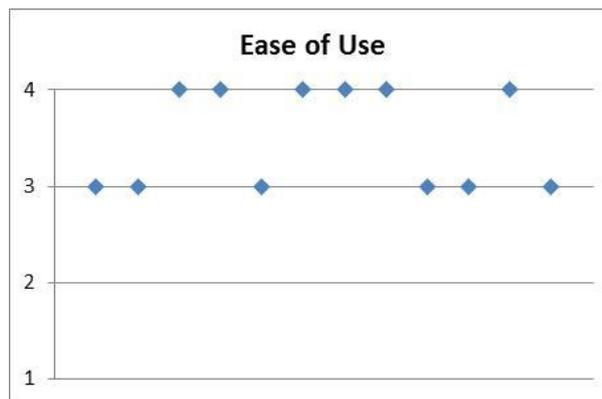
Figure 1: Distribution of evaluations of ease of use of the platform

4.1.1 Comments and recommendations

The following were comments and recommendations from the external experts that accompanied their rating of the ease of use of the platform:

- I've found the OSSMETER interface quite clear and user-friendly. Maybe the home is a little bit crowded, unlike the other pages.

- Images for "other metrics" can be hard to read; moreover a description of the metrics could help: for example, it is not clear to me what Churn means.

- Using OSSMETER was quite straightforward. The GUI is very intuitive.

- The welcoming page is clear. It is easy to understand how to use it. When typing the name of the project, OSSMETER automatically fetches and display information about the project, provided that it is hosted by one of the forges it can handle. I made the test with Acceleo, which is an Eclipse project.

- Nevertheless, one day when OSSMETER will have scanned a lot of projects it will be difficult to find a specific one. It could be nice to a have a category or tag system: it would help people find projects that correspond to what they seek.

- It would help if projects are grouped into types of applications, i.e. CMS, DB, Utilities, Modelling etc. It might be an administrator placing the new projects into a category. Also if by selecting the category you get a ranked list, then you know which is the active open source project for that group. You could also show the best one on the main page.

4.2 PLATFORM PERFORMANCE

The average rating amongst all of the experts for performance of the OSSMETER platform was 3.3, which is between Good and Excellent. The standard deviation amongst responses was 0.6. The distribution of ratings amongst the experts is shown in Figure 2.
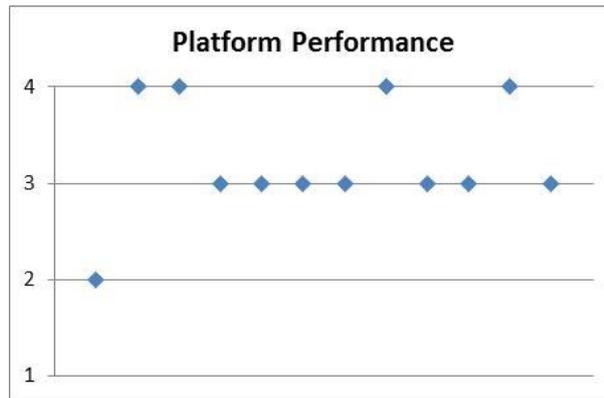
Figure 2: Distribution of evaluations of platform performance

4.2.1 Comments and recommendations

The following were comments and recommendations from the external experts that accompanied their rating of the platform performance:

- When accessing some of the analysed projects (the bigger ones) it seems that the platform is lagging a little bit. However, I'm pretty sure this could be easily scaled.

- I've noticed some delay in displaying the results of metrics, in particular when comparing projects. However, I don't believe these delays will hinder the overall usability of the platform.

- The interface was fast enough. However, the new project analysis which is supposed to take a couple of days, seems somewhat too long.

- Everything is fast. Some sparklines are not displayed, but it may be due to my company's firewall.

4.3 USEFULNESS OF CODE METRICS

The average rating amongst all of the experts for usefulness of the code metrics provided by the OSSMETER platform was 3.4, which is between Good and Excellent. The standard deviation amongst responses was 0.6. The distribution of ratings amongst the experts is shown in Figure 3.
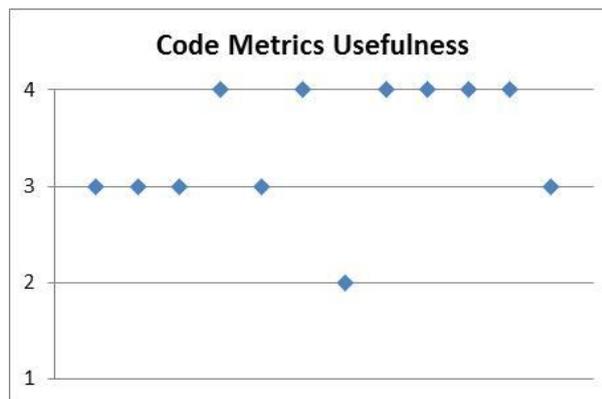
Figure 3: Distribution of evaluations for usefulness of code metrics

4.3.1 Comments and recommendations

The following were comments and recommendations from the external experts that accompanied their rating of the usefulness of the provided code metrics:

- The existing metrics show a good coverage of commonly used metrics. It might be interesting if users of the platform are interested in defining their own metrics, e.g. by combining predefined metrics.

- I've found quite interesting the possibility to have information about the project development and activities. But maybe it can be more appealing for potential sponsors or company participation of a project.

- The code metrics were very useful. However, they might be better structured in groups to enable an easier navigation through the values.

- Code metrics are interesting. However, users may not want to always consider all these metrics.

- Radar chart: There are some axis that could be related with the quality of the project (coupling) whereas others probably could not (code size). I expected to see a correlation between the area covered in the radar chart and the health of the project.

- It would be useful to show a tooltip when on mouse hovering an axis in radar chart (brief explanation of the scale). Then it will enable to make the explanation (on the right) less verbose and avoid repetitions (e.g. explanation of the sentiment analysis).

- It would be useful to explain the facts behind the information provided when on mouse hovering (trend is going down "because" three months was and now is…).

- It would be nice to add support to colour code a specific site Green / Amber / Red based on the analysis of different KPI's. All I want to know as a user is this

a good source to use. Perhaps with a time axis to see how the colour has changed.

4.4 USEFULNESS OF COMMUNICATION METRICS

The average rating amongst all of the experts for usefulness of the communications metrics provided by the OSSMETER platform was 2.8, which is between Acceptable and Good. The standard deviation amongst responses was 0.4. The distribution of ratings amongst the experts is shown in Figure 4.
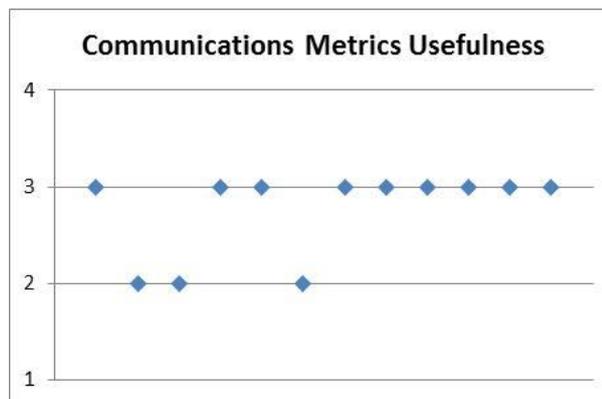

Figure 4: Distribution of evaluations for usefulness of communication metrics

4.4.1 Comments and recommendations

The following were comments and recommendations from the external experts that accompanied their rating of the usefulness of the provided communication metrics:

- To be honest, I would personally look more for code metrics and issue related stuff. This could be because I'm used to the open source community and carrying out my own investigations, but others may find it useful.

- I've found quite interesting the possibility to have information about issue tracking, as it can give a nice view on the project team's support. Communication metrics alone cannot always be trusted to judge the health of the project, but they can give a good feeling together with the other metrics.

- There are too many metrics. I am not sure we need so many details to evaluate the quality of a project. It may be interesting to let users select the criteria they want to consider, and have a unique spiderweb diagram of all these selected criteria instead of a set of spiderweb diagrams.

- The "metrics in depth" provides so much data that it's difficult to decide how to interpret everything. For the moment, OSSMETER demonstrated its ability to extract data from forges; now it must provide synthetic and consolidated data, so that users can actually draw conclusions.

- I find very useful to choose the metrics that are going to be shown (in the account's profile).

- New bugs/features versus fixed etc. shows more clear than individual lines with counts, i.e. the rate of new stuff and closed being resolved over time. Then a mouse overlay can be provided to get the individual metrics.

4.5 OVERALL VALUE

The average rating amongst all of the experts for overall value of the OSSMETER platform was 3.3, which is between Good and Excellent. The standard deviation amongst responses was 0.4. The distribution of ratings amongst the experts is shown in Figure 5.
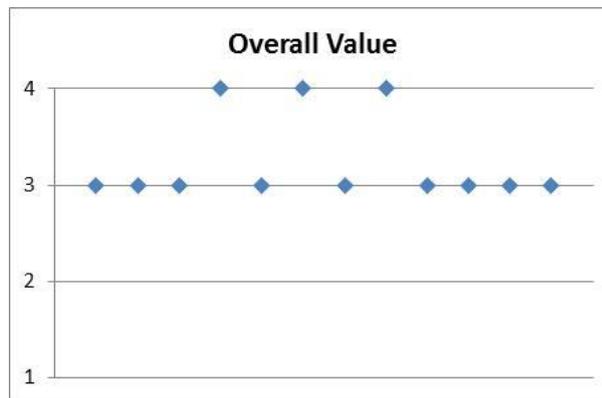

Figure 5: Distribution of evaluations for overall value

4.5.1 Comments and recommendations

The following were comments and recommendations from the external experts that accompanied their rating of the overall value of the platform:

- The project has analysed a number of projects that we use to implement solutions for our industrial partners. If the project/platform increases the number of analysed projects, I think this platform is very interesting for developing partners.

- Overall, I have a good feeling about the presentation and usefulness of the OSSMETER project. I believe that having information about projects before starting to use them could be quite helpful to any individual, team or company.

- The chance to have information about the development status and activity of a project, as well as the support of the project team to user requests, is surely valuable.

- It would be good to have more Eclipse projects available in the list. Moreover it could be great to have the possibility to compare projects which address similar

goals (for example, in the Eclipse environment we can find projects such as ATL and QVTo which both deal with M2M Transformation)

- Interesting platform - provides quite a deep insight into the project statistics. I personally would be interested to use it.

- The comparison is quite useful in order to choose two projects that offer the same functionality.

- There is no doubt OSSMETER provides an interesting service. However, the website provides too large amount of data: users will probably be lost.

- It would be really nice to have projects organized into categories, to ease their selection and comparison.

- OSSMETER makes very relevant statistics to help decide what project to use. It would be nice to display the license of projects, as it is a key criteria for project selection (GPL, LGPL, BSD, possibility of commercial license, etc.) as well as the number of users (this is probably difficult to evaluate) to have an idea of the popularity of a given project.

5. ECONOMIC ASSESSMENT

The economic assessment of the OSSMTER platform was introduced as an additional element for input from the external experts beyond the rating of the platform against the five criteria (e.g. ease of use, performance, usefulness, etc). It's inclusion was motivated by activities under workpackage 7 tasks to identify within the final version of the exploitation plan an economic model that would be sustainable after the project is completed.

The external experts were asked to provide any guidance they might have concerning the amounts that could be charged for the OSSMETER platform as a subscription service, or to suggest other revenue generation approaches that would likely be acceptable to software developers within their industry sector based on the value provided by the platform. Some experts noted they were not comfortable addressing marketing and economic aspects; however most provided inputs or recommendations The recommendations received were the following:

- Regarding price policy I would suggest to use an approach such as GitHub: Relatively small fee for individuals and a higher fee for companies.

- This is a no trivial point: In my opinion it would be better to avoid subscription fee, to initially encourage the usage of the tool and have feedback; donations could be considered as an alternative approach.

- A part of my work consists in developing tools to manipulate and transform models. We sometimes rely on different open source projects, such as Acceleo, XTend, DokuWiki, etc. A tool like OSSMETER would be definitely relevant to

select the right tool: we usually select the tool we heard of the most; it is difficult to have precise criteria.

- Regarding pricing policies that could be applicable, I think we must consider two things. First, nowadays most people are not ready to pay to get information: we get news from free newspapers, we can freely listen to music on YouTube, etc. Such information is of relatively low quality, but it is sufficient for most people: they will not pay to get better quality. Second, comparison services like OSSMETER are useful to make decisions from time to time, but it is difficult to justify the payment of a monthly fee. So, I would recommend the following funding policy:

  − Put advertisement on the OSSMETER website. As such a website is likely to receive a very large number of visits, advertisement will probably be profitable.

  − Provide metrics publicly. It will help generate traffic and will reinforce the popularity of the website.

  − Develop additional services that are not free, but for a very reasonable amount (e.g. €10), so that many people can afford to:

    ➢ Have access to more detailed information (e.g. number of users of a given project) to make their decision.

    ➢ Have additional open source projects analysed and included in the database.

    ➢ Get a clean and official project comparison chart with different weights for criteria, so that it is possible to get a unique score for each project. Getting an official project comparison chart is interesting for companies when they need to decide what technology to use. I think they would pay for it.

- Getting paid to add projects in the database corresponds to common practices in open source projects: the project has its own roadmap and features are developed according to this roadmap. But if you really want a given feature, you may pay for it in order to speed up the development. Once developed, the feature is available for everybody.

- I am very reluctant to pay for services. If there is the possibility to choose between a less-powerful free service and full-of-features paid service, I will go for the free one. In this way and putting the focus in GitHub, the data regarding to watch/forks and graph is enough.

- In order to "put a price for the tool", it would be interesting to show in the webpage some use scenarios and clarify in which way OSSMETER will improve the customer business. If it is missing, people will find it difficult to quantify the added value and probably they will not subscribe to it.

A common thread from the contributions received from the external experts was the fees charged for individuals to access the information collected on open source projects should be zero or only a modest amount. Something similar to the fees charged by GitHub could be acceptable, which would typically be €5-€20 for individuals and €20-€80 for organisations. Or alternatively, a model where a large amount information is free to access, but specific features or additional content incurs a small fee.

6. CONCLUSIONS

The evaluations carried out by the external experts have shown the OSSMETER platform has great potential to provide important benefits for European software developers and others who make decisions concerning the selection and continued support of open source projects.

The platform was rated Good to Excellent in every criteria except for the communication metrics where it scored between Acceptable to Good. The lower rating of the communication metrics is likely to be related to their newness as they are less familiar metrics for evaluating open source projects as compared to more traditional code metrics. Also, the communication metrics are typically used to complement the code metrics where the quality of the code is first verified, followed by further analyses of the supporting communication channels.

The recommendations also provided important inputs for further development of the OSSMETER platform including:

- Providing facilities for OSS projects are categorised when displayed perhaps *a priori* as part of the OSS analysis process, or by providing support for user defined tags to be assigned to OSS projects for customised groupings.

- Providing additional online guidance to users concerning tailoring of the display of metrics that are of most importance as some users were able to utilise this user profile feature while others were not aware of the capabilities available for customising the display of metrics.

- Supporting hover hints that would pop-up on specific items like radar graph axis or more metrics with further explanations. Also, using of colour to quickly identify the overall quality level of OSS projects.

- From an economic standpoint, a prevalent view was that a low/no fee base service should be used to make the OSSMETER platform attractive and also aligned with other types of online services familiar to software developers (e.g. GitHub), and that high traffic volumes should create opportunities for revenue generation either through advertising, or use of basic/premium service models.

The recommendations provided are useful for incorporating in future revisions of the platform and the economic assessment provides important inputs for finalising the exploitation plans in the project.