MADES

Model-based methods and tools for Avionics and surveillance embeddeD SystEmS

FP7-ICT-2007 248864

D5.1 Specification of the evaluation criteria of the technical WPs

Work package: WP5
Lead participant: TOG
Editor: Scott Hansen
Authors: Chris Parnell, Scott Hansen, Ian Dobson, Alessandra Bagnato, Gundula Blohm, Neil Audsley
Reviewers: Dimitris Kolovos, Matteo Rossi
Date: 30-11-2010
Version: 4.3

Dissemination level		
PU	Public	
PP	Restricted to other programme participants (including Commission Services)	
RE	Restricted by a group specified by the consortium (including Commission Services)	
CO	Confidential, only for members of the consortium (including Commission Services)	X

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 248864.

The MADES Consortium

TXT e-solutions
Via Frigia 27
20126 Milano
Italy

Contact: Alessandra Bagnato
E-mail: alessandra.bagnato@txt.it
Phone: +39 2 25771.725
Address: Via al Ponte Reale 5
16124 Genova, Italy

SOFTEAM
21 avenue Victor Hugo
75016 Paris
France

Contact: Andrey Sadovykh
E-mail: andrey.sadovykh@softeam.fr
Phone: +33 1 3012 1857

University of York
Heslington
York YO10 5DD
United Kingdom

Contact: Richard Paige
E-mail: paige@cs.york.ac.uk
Phone: +44 1904437222

Politecnico di Milano
via Golgi, 42
20133 Milano
Italy

Contact: Luciano Baresi
E-mail: baresi@elet.polimi.it
Phone: +39 02 2399 3638

The Open Group
Avenue du Parc de Woluwe 56
B-1160 Brussels
Belgium

Contact: Scott Hansen
E-mail: s.hansen@opengroup.org
Phone: +32 2 6751136

CASSIDIAN
Woerthstrasse 85
89077 Ulm
Germany

Contact: Gundula Blohm
E-mail: gundula.o.blohm@cassidian.com
Phone: +49 731 392 3757

Executive Summary

This deliverable describes the evaluation criteria that the MADES project will use to assess the performance of the project towards achieving its objectives. The measures identified will play an important role in effective monitoring of the scientific and technical achievements and provide key indicators that quantify the expected industry impact of the project. The measures have been derived using an industry accepted methodology for defining measurement programmes, which have been grouped in four categories: Research and Development, Industry impact, Dissemination, and Management. The measures defined in this document will be utilised at the later stages of the project to evaluate the scientific results through the use of Case Studies and other actions, which will be reported in follow-on deliverables.

Table of contents

1 Introduction

This deliverable provides a set of evaluation criteria that the MADES project will use to measure and assess the performance of the project in achieving the overarching objectives outlined in the MADES Description of Work. The measures identified will play a vital role in effective project management and are derived using an industry-accepted methodology, known as Goal-Question-Metrics (GQM) method developed by Basili[1], which is well established and recognised to be a valuable approach for helping project and development managers define and align measurement programmes to project goals.

A measurement and monitoring programme needs to be balanced. On the one hand, too little evaluation and assessment can lead to projects that don't achieve expected results, as performance problems are not identified and corrective actions not taken. On the other hand, too many metrics and measures can be time consuming and divert important resources that would be better utilised developing new technology innovations or undertaking actions that create greater industry impact. It is also important to note that some processes and results lend well to measurement, while others like software research and development are better managed using more agile methods for evaluation as over management with too many metrics has been shown to lead to less than optimal software project results.

Overall, the measurements that will be applied within the MADES project are intended to provide four fundamental benefits.

1. **Characterise** - metrics allow partners to better understand processes, resources and project/work status. They also establish baseline measures for future comparisons.
2. **Evaluate** - metrics enable the project partners to see how they are performing against plan or goals. The data shows when and where the project and processes drift from plan and allow for course corrections.
3. **Predict** - metrics can help predict future performance by providing understanding of the relationships between processes and projects. Through reliable data, project partners can better estimate costs, schedules and work effort.
4. **Improve** - A thorough metrics program isolates roadblocks, inefficiencies and root causes in order to improve efficiency, productivity and quality. Once metrics have been established, they can be used to demonstrate alignment with overall project goals and can assist in communication with others outside the project.

These four benefits result in increased project efficiency, greater collaboration, enhanced productivity, larger industry impact and ultimately better use of resources (time, money and people).

The MADES consortium has identified a balanced set of metrics that will support the evaluation and assessment of work within the project, while not diverting substantial resources to

[1] Basili, Victor R.; Caldiera, Gianluigi; Rombach, H. Dieter, "The Goal Question Metric Paradigm", Encyclopaedia of Software Engineering (Marciniak, J.J., editor), Volume 1, John Wiley & Sons

measurement collection and data analysis. We have selected the set of measures and evaluation criteria with a desire to have simple metrics whenever possible so that the achievements of the project are more easily conveyed to the broadest audience.

1.1 Role of this deliverable

This document constitutes MADES project deliverable *D5.1 Specification of the evaluation criteria of the technical WPs* and describes the measures that will be used to assess the level of success achieved by the research and development within the MADES project. The role of this deliverable is to establish a set of quantifiable metrics that can be used to determine the extent to which the project results satisfy the industrial user requirements, provide improvements to the embedded systems development community, and are likely to have a substantial impact on the target industries. These measures will help the project partners to monitor performance and take additional actions when needed to ensure the project achieves its objectives.

1.2 Structure of this document

This deliverable is structured in five parts starting with a description of the methodology used to define the metrics in Section 2, followed by the evaluation criteria categorised into the following sections:

- Section 3: Research and Development measures
- Section 4: Industrial Impact measures
- Section 5: Dissemination measures
- Section 6: Management measures

For each measure the justification of measure selection and the specification of how the measure will be quantified are provided.

1.3 Relationship to other MADES deliverables

This deliverable is closely related to all of the project technical deliverables as each of these provides the technological results that will be evaluated using the metrics defined in this document. The actual data collected for each metric will be reported in deliverables:

- D5.2 Results of first evaluation of the technical WPs.
- D5.3 Final evaluation report

The results of this document will also help shape the final results for work to be carried out in each of the development workpackages as these will be the subject of the many of the evaluations and the results will help identify any final improvements needed for the following:

- D1.7 MADES Final Approach Guide
- D2.1 MADES Modelling Language Specification
- D2.5 MADES Modelling Tools Final Version
- D3.6 Complete validation framework
- D4.1 Model Transformation and Code Generation Specification
- D4.5 Model Transformation and Code Generation Tools Final Version

Other intermediate versions of the above deliverables will also be shaped by the evaluation plans described in this document as awareness of the evaluation criteria will help focus the research and development work.

1.4 Contributors

The Open Group has been a major contributor to this deliverable as well as acting as editor in the preparation of the various versions of the document. TXT e-solutions and EADS as the industrial user partners that will carry out many of the evaluations later in the project have provided insight and guidance concerning the measures that are of highest priority and the feasibility of various measures considered in preparing this document. Politecnico di Milano and University of York have acted as reviewers for the deliverable providing valuable suggestions for final editing and refinements.

2 Assessment criteria methodology

2.1 Introduction

The MADES project has used the approach of goal-oriented measurement as the basis for developing the monitoring and assessment criteria specified in this document. In particular, the project has adapted the GQM method developed by Basili and colleagues during the 1980s, refined during the 1990s, which now serves as the foundation framework for many measurement initiatives.

The GQM method is used to define measurement in such a way that:

- Resulting metrics are tailored to the project and its goals.
- Resulting measurement data play a constructive and instructive role in the project.
- Metrics and their interpretation reflect the values and the viewpoints of the different groups affected (e.g., developers, users, researchers).

As illustrated in the Figure 1 below, GQM begins by identifying measurement goals (conceptual level) that support (are aligned with) overall project goals. The consortium then poses questions (operational level) to further clarify and refine the goals as well as capture the variation of understanding of the goals that exists among the partners. The consortium then identifies metrics that will provide answers to the questions (Quantitative level).

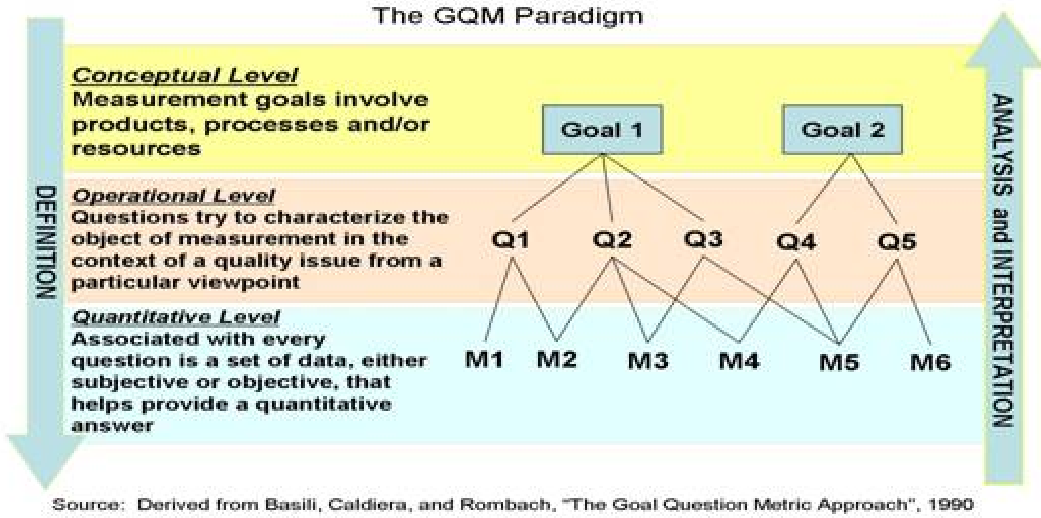

Figure 1 – Overview of GQM methodology

What distinguishes GQM from other measurement paradigms is the hierarchical tree structure used to maintain the relationships amongst goals, questions and metrics.

2.2 Refinement process

GQM uses a six-step process where the first three steps are about using the overarching project goals and desired industry impact to drive the identification of the right metrics, and the last three steps are about gathering the measurement data and making effective use of the measurement results to drive decision making and improvements. The adaptation of Basili's six-step GQM process being utilised by the MADES project is as follows:

1. Identify a set of industry and project goals
2. Generate questions that define those goals in a quantifiable way
3. Specify the measures needed to be collected to answer those questions and track process and project conformance to the goals
4. Develop mechanisms for data collection
5. Collect, validate and analyse the data to provide feedback for corrective action
6. Analyse the data periodically to assess conformance to the goals and to make recommendations for future improvements

The MADES work programme is centred on achieving five primary scientific and technological objectives:

- Develop an avionics/surveillance systems dedicated modelling language based on existing generic RTES modelling languages
- Develop dedicated tools for RTES modelling in the avionics/surveillance domain
- Develop advanced model-based verification and simulation methods and tools for RTES
- Develop advanced model-based code generation methods and tools for RTES

- Integrate MDE methods and tools for avionics/surveillance RTES in a seamless solution

Coupled with these objectives are the higher-level goals of the FP7 ICT workprogramme under the heading ICT-2009.3.4: Embedded System Design, the project also seeks to:

- Develop an integrated design environment for embedded systems that can be extended and customised. This covers software, hardware/software and system design tools for holistic design, from applications down to component and platform level.
- Address challenges encompassing flexibility of the platform to support different applications, increased interoperability of tools from SME vendors and openness in order to facilitate the entry of new industry players, support associated standardisation, easily import existing components and/or handle upgrades.
- Investigate key issues of: (i) technology for efficient resource management, (ii) tools supporting design space exploration, in particular trade-offs when co-developing hardware and software; and (iii) advanced model-driven development.

For each of these objectives there are specific tasks to be undertaken and results in the form of deliverables identified. The focus of this deliverable is to provide specific measures that allow the partners and the European Commission to assess the degree to which these objectives have been achieved, and the impact they have had on the partners, industry and Europe's position in the global embedded systems market place.

2.3 Refinement example

The MADES partners have applied the GQM methodology to establish a set of measures that will be used to evaluate and assess the performance of each of the project workpackages. The Conceptual Level goals were described in the Description of Work for the MADES project and are aligned with Objective ICT-2009.3.4: Embedded system Design, from the ICT Work programme 2009-2010. The project goals were then taken and considered at an operational level in terms of how the goal could be achieved and the implications each goal has on the outcomes of each workpackage. These were stated in the form of questions that should be answered in the affirmative if the higher-level goal is being achieved.

Once the partners had a set of questions to consider at the Operational Level, then the challenge was to identify quantitative measures that could be used to verify that each individual question could be satisfied. As the GQM methodology anticipates, some quantitative measures identified by the partners addressed more than one question at the Operational Level. Furthermore, in some cases the Operational Level questions had multiple dimensions to consider, which in turn required the use of more than one quantitative measure.

An example of how GQM was applied in establishing the evaluation criteria and measures is shown in Figure 2 for one of the overarching goals of the MADES project: provide an integrated design environment for embedded systems that can be extended and customised covering software, hardware/software and system design tools for holistic design, from applications down to component and platform level.

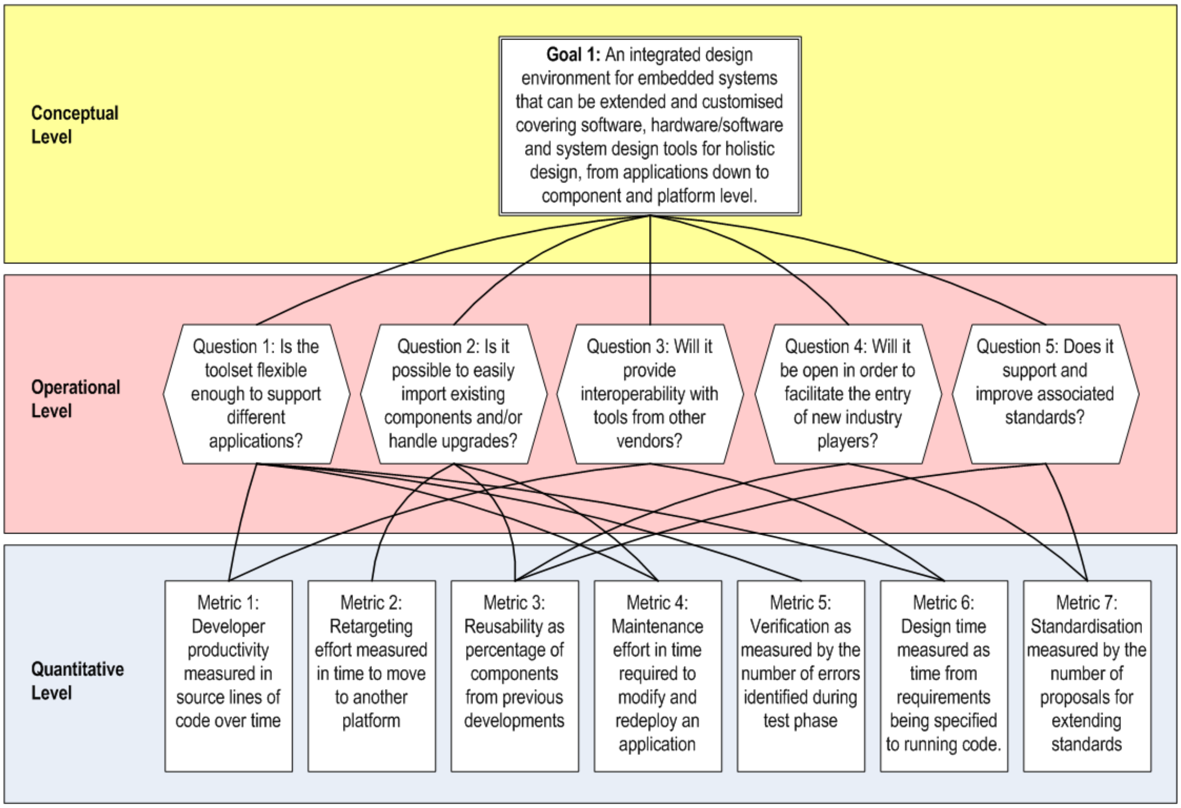

Figure 2 – Applying GQM to determine MADES evaluation criteria

Figure 2 illustrates the application of the GQM methodology for one of several overarching goals that the project partners expect to achieve through the research and development within the project. Other technical objectives as well as market impact objectives have also been addressed and many more measures established using GQM. These measures for evaluating the performance of the project in meeting each of the MADES project goals are specified below in terms of the rationale and the measurement process that will be utilised.

The MADES partners recognise that the success achieved by a project of the size and complexity of MADES will be multifaceted as the project itself starts with several broad primary objectives which require multiple measures in order to sufficiently evaluate whether the project has attained expected levels of achievement.

Using the GQM method the project has identified a set of quantifiable measures that will enable the project partners and the European Commission to evaluate and monitor the progress the partners make towards achieving the overarching goals of the project. As the measures sometimes relate to multiple goals, they have been grouped into four categories:

- **Research and Development** – measures that identify the degree to which project innovations are meeting the current and future needs of industrial organisations and in particular, how well development work is progressing towards achieving a new tool set for embedded systems development suitable for commercial exploitation.
- **Industrial Impact** – measures that indicate the degree to which MADES has had an effect on improving and supporting industrial organisations within the Aerospace and Defence industries, as well as effects on companies that supply information technologies to organisations in these industries.

- **Dissemination** – measures to identify the degree to which information concerning the project and project results has been communicated to the target constituencies including industry, academia, developers, and research.
- **Management** – measures that monitor the performance and resource utilisation of the project against contractual budgets, schedules and the Description of Work.

The measures for assessment and monitoring in each of the above categories are described in detail below.

3 Research and Development Measures

3.1 Introduction

The measures for evaluation of research and development are intended to identify the degree to which the project is progressing towards implementing expected technology innovations for model driven development of embedded systems for the Aerospace and Defence industries, and the extent to which the innovations will contribute to improvements for European embedded systems developers. These measures are intended to be aggregate measures utilised for determining how well the project development work is progressing towards achieving a new set of tools suitable for commercial exploitation.

The research and development work within the MADES project is expected to deliver important improvements in each phase of embedded systems development by providing new tools and technologies that support design, validation, simulation, and code generation, while providing better support for component reuse. As shown in Figure 3 for one of the industrial project partners (TES), the MADES project will provide new tools that add additional capabilities to already established development processes.

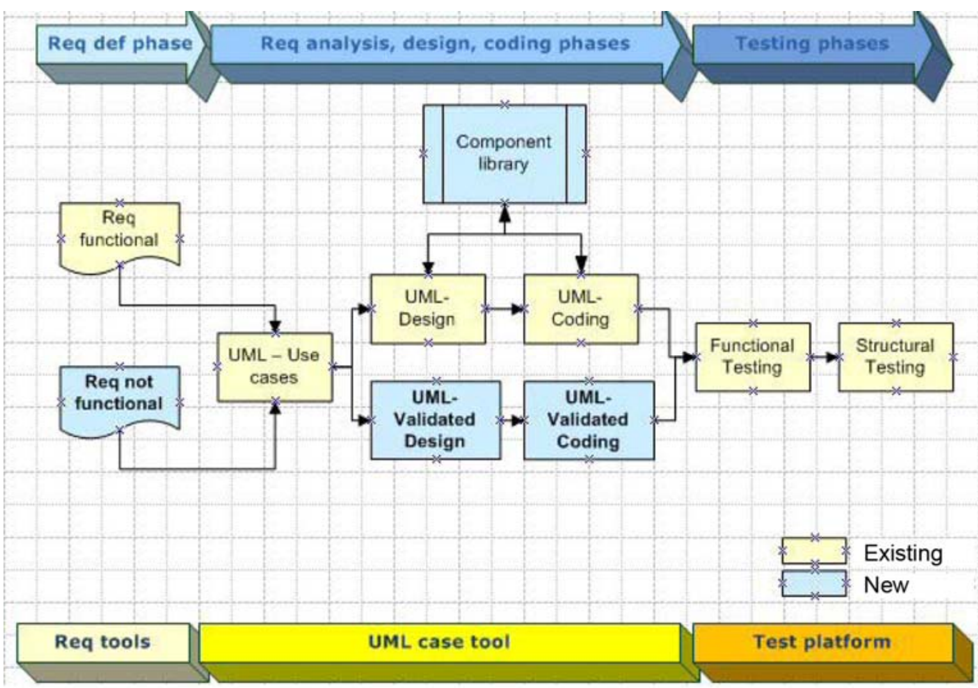

Figure 3 – Example of new innovations from MADES tool set

Several of the evaluation criteria that are driven from the higher-level project objectives are focused on improvements achieved for typical embedded systems developers as compared with current development procedures. The test environment for measuring these improvements will be a representative software development project, or Case Study, that will be carried out by each of the two industrial use partners (TES and EDG) where the MADES tools will be used for industrial application development. As we are concerned with the improvements achieved over existing procedures, it is important to have a clear approach for establishing the baseline used for comparison in determining the extent of any improvements seen in the Case Study evaluations.

3.2 Comparative baseline project

The two industrial user partners each have many years of experience in software development having undertaken numerous projects for different types of embedded systems applications. Each partner has selected a representative application development project as a Case Study for use in determining requirements for the new MADES tool set (see deliverable *D1.1 – Requirements Specification*), and later in the project, the same or a similar application development project will be used to evaluate the MADES tools in carrying out actual application development.

As the impact of many of the research and development tasks within the MADES project represent an important evolutionary improvement in embedded systems software development for the Aerospace and Defence industries, many of the evaluation criteria the project has specified need to be considered with regard to the relative improvements that were achieved when compared with previous software development practices and tools. This provides for a more reliable measurement approach as the normal variations that might occur between software development entities are largely removed, and we can more easily express the impact of MADES research development results in a side by side “before and after” expression.

Not all evaluation criteria are stated in terms of relative improvement as there are new innovative capabilities being provided by MADES tools and methods for which no historical comparative data would be available. However, for some evaluation criteria such as the number of errors discovered during test, the time and resources utilised to complete a specific development phase, a comparative baseline is needed so that the relative improvements can be highlighted.

In order to provide a comparative baseline for measuring the improvements provided by the MADES tools, each of the industrial user partners will select between one to three previous application development projects (or subprojects) that are similar to their Case Study. The exact number will be dependent on the number of projects that are similar and for which sufficient historical data is available. The selection of the projects will comply with the following criteria:

- Completed within the last two years
- Utilised current practices for model driven development (use of UML, etc.)
- Similar size as measured by number of source lines of code (± 20%)

- Similar complexity as measured by number of function points[2] (± 10%)
- Developed in the same language
- Same or similar target hardware platform (e.g. Intel, single or multicore, etc.)
- Empirical data available concerning schedule, test reports, person months, etc.

The data from the selected projects satisfying the above criteria will form the reference for comparison for evaluation. If more than one similar project is identified, the historical data will be averaged so that it is representative of typical projects similar to the chosen Case Study project for each industrial user partner. The empirical data from the previous project or the averaged empirical data when more than one historical project is selected will be referred to as the Comparative Baseline Project (CBP) and will be the baseline measure used when quantifying the subset of the measures specified below that are expressed as comparative improvements.

3.3 Developer productivity

Justification

The MADES project will develop dedicated modelling tools, automated validation technologies, transformation and generation capabilities that will enable modelling by composition with automated verification, generation of code and HDL for critical applications in the Aerospace and Defence industries. As these technologies require development organisations to move away from traditional development methods, an important measure of the return on investment will be improvements in developer productivity. It is expected that the tools developed within the project will automate many functions carried out by developers today. It will also enable developers to more easily develop complex embedded system applications by working at higher levels and using composition and transformations to better manage complexity. A key measure of success for the MADES tools will be the degree to which productivity is improved for a development team.

Specification

Attempting to measure individual productivity is unlikely to be feasible as software development is non-deterministic and there are large variations in individual performance using common software engineering measures. If productivity is measured as source lines of code per person month, it implicitly suggests that the programmer who writes 10 times the amount of code to solve a particular problem is more productive than the programmer who writes 1 times the amount of code, which experience would clearly indicate is not correct.

The measure of most interest for the MADES project is the productivity of the software development team where we are able to average out the expected variations in individual developer productivity. The specification of the measure for developer productivity will be the total elapsed time from the end of the requirements specification phase to delivery of tested code, divided by the number of person months expended.

---

[2] Function points were defined in 1979 in *A New Way of Looking at Tools* by Allan Albrecht at IBM. The functional user requirements of the software are identified and each one is categorized into one of five types: outputs, inquiries, inputs, internal files, and external interfaces. Once the function is identified and categorised into a type, it is then assessed for complexity and assigned a number of function points.

$$\text{Developer Productivity} = \frac{\text{Elapsed time from requirements complete to testing complete}}{\text{Person months expended}}$$

The measure will be compared against the same calculation for the CBP to determine the change, and will be stated as a percentage.

As the Case Studies may involve a learning curve for developers using new tools and development procedures, it may be necessary to normalise the Case Study data using a factor that represents the time to become familiar with the MADES tools. For example, the denominator in the above formula could be reduced by a factor equal to *n months x m staff* to reflect the time required for development personnel to get familiar with the MADES tools. The use of a factor will be decided when the prototype tools are available and the learning curve for the new tools is better understood.

3.4 Retargeting effort

Justification

The use of a model driven architecture approach that relies on modelling and different layers of abstraction, and in particular the use of automatically generated code and HDL, should substantially reduce the effort required to retarget an application to an alternate platform. Retargeting effort is important for embedded systems as the competitive drive for lower cost and lower power consumption for systems, along with increasing computational requirements, leads to periodic retargeting to new lower cost and/or more powerful platforms. The MADES project tools should make it possible to preserve system models while abstracting the process architectures and automating the migration procedure. In addition the virtualisation of hardware techniques should provide the means to better cope with platform portability issues.

Specification

It is unlikely that there will be specific retargeting actions being carried out within the Use Cases of the industrial user partners and selecting a previous project where retargeting was involved for inclusion in the CBP is also problematic. For these reasons it would be difficult to define a direct measurement for retargeting effort that could be used for comparison. Instead, the measurement approach will be to utilise two surrogate measures where data can be collected during the Case Study evaluations and when considered together will provide an indication of the expected improvement in retargeting effort:

Measure	Description	Calculation
Abstraction Completeness	Degree to which the Case Study application can be defined entirely at the modelling level	Number of modifications made of MADES generated code to ensure or correct execution or achieve required performance and determined according to the following scale: • Low: no code modifications • Medium: less than 3% of generated code required modification • High: more than 3% of generated code required modification

Measure	Description	Calculation
Virtualisation complexity	The effort required to define a virtualisation profile and associated mappings required to support an additional hardware platform	The time in person days to specify a new hardware platform to a level sufficient to generate code using the MADES tools using the following scale: • Low: less than 2 person days • Medium: between 3-5 person days • High: more than 5 person days

These two measures have been selected as they characterise key attributes that affect retargeting. For example, if the abstraction of an application including functional and extra-functional properties is less than complete then modifications to code after transformations will be needed to reflect additional functionality or constraints not captured at the higher abstraction levels. These modifications will often require significant re-implementation effort when re-targeting to a different platform. Similarly, if the effort to specify the virtualisation profile and mapping s for a new hardware environment is difficult or complex, then it is likely that the retargeting effort will not be substantially improved over current practices.

3.5 Design time

Justification

The MADES tool set includes an enhanced commercial modelling tool with specific features developed within the project to support model driven embedded systems development for the Aerospace and Defence domains. Also included are tools that support analysis and simulations at the model level that allow design concepts to be verified, as well as a component repository that should reduce development time by automating routine and repetitive design tasks. These tools are expected to shorten the time required to design embedded systems applications.

Specification

The measure that will be used for evaluating design time will be the elapsed time required during the Case Study evaluations to transition from the point of having completed the requirements specification to delivering code for testing, normalised against the total person months that were expended.

$$\text{Design Time} = \frac{\text{Elapsed time from requirements complete to code delivered to testing}}{\text{Person months expended}}$$

The measure will be compared against the same calculation for the CBP to determine the change, and will be stated as a percentage.

As the Case Studies may involve a learning curve for developers using new tools and development procedures, it may be necessary to normalise the Case Study data using a factor that represents the time to become familiar with the MADES tools. An example of how normalisation to address learning curves could be applied has been provided in Section 3.3 above.

3.6 Reusability

Justification

The project intends to improve the reusability of software components by introducing persistent models supported by a component repository, and by integrating a system modelling by component composition approach within the MADES tool set. Components will be better characterised and the ability to identify, select, and integrate existing components within a system will be improved over current practices. Reuse of components saves time and costs in software development but requires that knowledge of reusable components be captured and maintained and that integration of components can be achieved while guaranteeing required system qualities. Reuse of components is not widely practiced for embedded systems so improvements in reusability made possible by the MADES tool set would have important benefits for industry.

Specification

The Use Cases of the industrial user partners do not start from a library of existing components available to be reused so directly measuring reusability of components is not feasible. However, there are certain attributes of the MADES tools that if present would indicate reusability would likely be more widely practiced when MADES tools are deployed. The attributes for measurement are the following:

Measure	Rationale	Calculation
Component Information	If sufficient knowledge about a component and its behaviour is captured and stored then future reuse is more likely	Survey of Use Case developers questioning their view as to whether the information stored concerning components in the repository is sufficient to enable substantial reuse using the following scale: • Low: insufficient information is stored to enable reuse • Medium-Low: Some information required to enable reuse is stored • Medium-High: Most information required to enable component reuse is stored • High: Nearly all information required to enable component reuse is stored
Component Selection	If finding applicable components that have already been developed is an easy process then future reuse is more likely	Survey of Use Case developers questioning their view as to whether facilities for searching within the repository for specific component functionality and features is sufficient to enable substantial reuse using the following scale: • Low: search facilities are

Measure	Rationale	Calculation
		insufficient to enable reuse • Medium-Low: search facilities are partly sufficient to enable reuse • Medium-High: search facilities are largely sufficient to enable reuse • High: search facilities are fully sufficient to enable reuse
Component Assembly	If assembling components into systems (i.e. composition) is straightforward then future reuse is more likely	Survey of Use Case developers questioning their view as to whether MADES composition and verification methods would enable substantial reuse using the following scale: • Low: composition methods are insufficient to enable reuse • Medium-Low: composition methods are partly sufficient to enable reuse • Medium-High: composition methods are largely sufficient to enable reuse • High: composition methods are fully sufficient to enable reuse

The overall survey scoring and their consistency across the three measures will provide an indication as to whether component reuse would be better facilitated by the new MADES tools and therefore greater reuse of components within organisations developing embedded systems would be expected.

3.7 Verification effort

Justification

The MADES project is developing technology innovations in the use of models and transformations for the Aerospace and Defence domains that support advanced verification and simulation methods. These methods are expected to enable developers to verify the correctness of a substantial portion of an embedded system application during the design phase. As a result design times should be improved (see measure defined under Section 3.5) as well as the effort required for testing reduced as errors are discovered and corrected during the design phase leaving fewer errors to be detected during the testing phase.

Specification

The new verification and simulation methods that are being developed in the project indicate three measures are needed to evaluate the impact the MADES tools have on development. This is because the new methods shift some development tasks previously carried out during the testing phase to the design phase, so we need to understand the changes in each phase to fully evaluate the impact the new technologies will have on verification. The measures to be used are the following:

1. Number of errors discovered during the test phase – collected from existing error tracking systems used by the partners when carrying out the Use Case evaluations

2. The time required for system design – as already defined under Section 3.5 for design time above

3. Number of errors discovered during the design phase with the new verification and simulation tools – these will need to be collected as an additional tracking task during each Use Case evaluations

These three measures will be compared against the same calculations for the CBP to determine the changes, and each will be stated as a percentage.

It may be the case that data concerning the number of errors discovered during the design phase (measure 3 above) is not available for the CBP. The measurement should still be carried out during the Use Case evaluations as knowing the extent to which the tools detected errors during the design phase is an important metric for evaluating the benefits provided by the MADES tools.

3.8 Maintenance effort

Justification

Software maintenance costs often comprise more than half of the total software development costs over the life of a system. The MADES tools will allow developers to validate the systems at the design phase with advanced verification and simulation methods, which will lead to significant improvements in the quality of the generated embedded systems applications. In addition, working at the model level provides a consistent and holistic view of the system which helps to locate and understand design level problems should they occur and avoid some of the reverse engineering challenges that often arise when having to maintain systems that have been in service for some time. The transformation and automated generation capabilities of the MADES tool set is also expected to reduce the number of errors while implementing the model for a particular platform, which in turn would mean the effort required for maintenance of applications developed with the MADES tool should be reduced.

Specification

The Use Cases of the industrial user partners involve the development of new applications so direct measurement of maintenance effort is not feasible as the applications will not yet have entered the maintenance phase. To establish a measure that could act as a predictor of the improvement in maintenance effort the project will utilise indicative measures based on data that can be collected during the Case Study evaluations.

The first measure will be the time, measured in person days[3], required to address Type 1 errors discovered during testing using the MADES tools. The measure should including the time required for analysing the error, making the correction and generating a new release of the application for further testing. This time estimate would correlate closely with a typical

[3] If error tracking logs for the CBP are not detailed enough to determine person days then elapsed time in days may be substituted.

maintenance task that would be carried out when an error is discovered after application deployment. The figure will be compared against the same calculations for the CBP to determine the change stated as a percentage.

A second set of measures related to maintenance effort are a collection of measures already specified but are closely tied to maintenance effort. The additional measures to be used as part of the evaluation of maintenance effort are the following:

Measure	Rationale	Calculation
Developer productivity	If developer productivity has improved due to the MADES tools, the time required to carry out developments as part of maintenance tasks are also likely to be improved	See developer productivity measure specification in Section 3.3
Design time	If design times have improved then the MADES tools are likely to offer similar improvements when design changes for corrective actions are required after deployment	See design time measure specification in Section 3.5
Number of errors discovered during the test phase	If fewer errors are discovered during the test phase it is likely that higher quality applications are being generated using the MADES tools, which would likely require fewer corrections after deployment	See verification measure specification in Section 3.7

These measures also provide a sense of how well complexity of embedded systems applications is being managed through the use of MADES tools. This is also useful for evaluating maintenance effort as application complexity is a well established predictor[4] of maintenance effort. If the MADES tools allow better handling of complex applications it is likely that maintenance effort will also be reduced.

3.9 Usability

Justification

The MADES tools are intended to eventually be used for development of industrial embedded system applications and their take-up by industry will depend in part on how easy they are to use and how stable they are in providing the expected features and functions. Developers' expectations for user interface, integration with other tools, and configurability have steadily increased thanks to IDE platforms such as Eclipse and the consolidation of smaller specialist tool vendors into larger tool technology suppliers. Usability must therefore be considered as a relative measure that depends on the existing tool environment and development practices of an organisation.

[4] "Determinants of software volatility: a field study", Journal of Software Maintenance and Evolution: Research and Practice, Volume 15, Issue 3, pages 191–204, May/June 2003

Specification

The MADES project will use a survey instrument with Use Case developers where their view as to the usability of each component of the MADES tool set is collected. The survey will include the following questions for each MADES tool stated in terms relative to other existing tools already in use in the development organisation:

- How easy was it to use the MADES tool?
- How stable was the MADES tool?
- How well did the MADES tool perform?
- How well did the MADES tool integrate with other MADES tools?
- How well did the MADES tool integrate with existing development tools?

The scale for answering the questions will be the following:

- Low: the attribute in question was insufficient
- Medium-Low: the attribute in question was partly sufficient
- Medium-High: the attribute in question was largely sufficient
- High: the attribute in question was fully sufficient

In addition, there will be an opportunity for Use Case developers to provide comments and clarifications concerning their experiences and why they chose specific ratings in each category. The MADES partners expect that usability will improve between the early versions of the tool set made available to the industrial user partners and the final versions of the tools. The survey instrument will therefore be utilised at both the early and final phase evaluations to provide an accurate representation.

3.10 Coverage of user requirements

Justification

The MADES software framework and the underlying tools architecture are driven by industrial user requirements from the Aerospace and Defence domains. The project will track each of the key requirements (see deliverable D1.1 – Requirements Specification) and address these requirements both from an architectural and development standpoint. While not all requirements may be within the scope of the project, the percentage of user requirements that have been satisfied provides an important measure of the applicability of the project results in solving the industrial challenges targeted by the project.

Specification

The MADES project will use a requirements management tool that will provide tracking of identified requirements and mapping of requirements to specific workpackages and components within the MADES tool set. The percentage of total requirements being implemented in one or more components of the MADES tool set will be available from the requirements management tool for periodic reporting. Priorities of requirements (e.g. mandatory, optional, etc.) have also been specified by the project partners and will be included in the coverage reporting.

3.11 Change requests during Use Case evaluations

Justification

The MADES project includes tasks to carry out two separate Use Case evaluations by industrial user partners in order that the final research and development is shaped by extensive user inputs and validation. The number of change requests that arise from the industrial user partners during the Use Case evaluations provides a measure of the degree to which users are engaged and driving the direction of the research towards industry needs. It also provides a measure of the degree to which research and development work has delivered a tool set suitable for industrial application development.

Specification

Change requests from industrial user partners and any others that gain experience with the MADES framework will be formally logged and tracked through final disposition. The following quantitative figures will be available from the change request tracking system:

- total number of change requests submitted by users
- number of change requests that were successfully addressed
- number of change requests that are still being addressed
- number of change requests that were beyond the research scope of the project

As prototype software, the project partners recognise that it may be the case that not all change requests may be resolved and that some further development work may be required as part of the exploitation of project results after the project is completed.

3.12 Components deployed for Use Case evaluations

Justification

The use of industrial Use Cases within the MADES project will enable new technologies and embedded systems development methods to be driven from user needs and validated during the project. One of the challenges facing the project is making the new tools accessible and stable enough for industrial developers to deploy outside of the laboratory environment of the research and development partners. The number of functional components deployed for Use Case evaluations is an important measure of the extent the technologies are ready for further dissemination and exploitation by the project partners and the European embedded systems development community.

Specification

The MADES tool set will be defined from an architectural standpoint and will have well-defined components and interfaces indicating the relationship and interaction between components. The measure will be a ratio specified as follows:

$$\text{Component Deployment} = \frac{\text{Number of MADES tool chain components deployed in one or more Use Cases}}{}$$

Number of MADES tool chain components
specified in the MADES Tool Chain architecture

The measure will be stated as a percentage and is not related to the CBP.

3.13 Requirements implemented

Justification

The requirements measure counts the number of requirements that have been defined and allocated to software components and successfully developed, keeping in mind that it may not be possible to implement some requirements until late in the development process. Tracking the requirements allows the project to plan resources and to prioritise those requirements that provide the greatest value towards achieving the overall goals of the project.

Specification

The MADES project will use a requirements management tool that will provide tracking of identified requirements and mapping of requirements to specific components within the MADES tool set. The requirements associated with each functional component will be tracked and when a component has been implemented the requirements will be reported as implemented. In some cases requirements will span multiple components in which case all functional components will need to have been completed for a requirement to be considered fully implemented.

4 Industrial Impact Measures

The industrial impact measures are outward looking and relate primarily to the industrial user and technology supplier communities in the Aerospace, Defence and other industries. These measures are intended to indicate the degree to which the MADES project has had an effect or is likely to have an effect towards encouraging, improving and supporting industrial organisations implementing model driven development of embedded systems. The measures are not presented in any particular order.

4.1 New or revised standards resulting from the project

Justification

Making the technologies from the MADES project the standard for model driven development used by many within embedded systems community will accelerate the wider take-up of MADES results, increase the opportunities for commercial exploitation, and give industrial users greater confidence when considering investments in new MDD based tools and methods. In this context we define “standard” as the official uptake within a standardisation process of a standards body (e.g. OMG, JCP, W3C, etc.) a specification (or part thereof) developed partially or completely within the MADES project.

It’s important to consider that standardisation is most often a consensus-based process involving many actors within industry, and sometimes government, who have an interest in a particular topic. The consensus process usually requires a significant amount of time (typically 1.5 to 3 years) as actors involved must consider what is being proposed as a potential standard, understand the implications, consider various viewpoints and alternatives, and eventually reach agreement on an often modified version of the original proposal. Within the time-frame of the MADES project it is unlikely that the entire consensus process for a new or revised standard would be completed, however, important and substantial progress can be made and the evaluation criteria focuses on actions that can be achieved within the operation of the project.

Specification

The success of the standardisation activities and the level of involvement of MADES in standards processes will be measured by monitoring the following actions by the project partners:

- MADES documents provided as a specification to a working group of a standardisation body
- MADES documents accepted as a working item of a working group of a standardisation body
- Contributions from MADES to working items of a working group of a standardisation body

In addition, the total number of standards groupings where MADES partners are active will be monitored to provide an indication of the scope of standards initiatives MADES is pursuing.

4.2 Industry links to MADES tool chain architecture

Justification

The MADES tool set architecture provides a common framework that can be extended to support additional tools or address additional industries beyond the Aerospace and Defence industries targeted by the project. The architecture specifications and interfaces will be published and the degree to which these are referenced through online links, citations in technical papers and commercial product and service literature provides an indication of the progress the project is making towards the MADES tools becoming an industry reference for embedded systems developers.

Specification

The number of industry links will be collected primarily online as part of the activities of Workpackage 6 addressing dissemination and exploitation, and in particular Task 6.2 – Market and standardisation watch, established to identify market trends and initiatives. MADES will collect data by undertaking the following actions at regular intervals during the project operation:

- Internet searches for references to the MADES architecture or tool chain on other websites.
- Collection of product literature and services information from known suppliers of related tools and technologies to look for references to the MADES tools or specifications.
- Online searches of relevant research journals, industry newsletters and publications for references to the MADES tools or specifications.

The project anticipates that data will not be available until the later phases of the project when a number of public deliverables have been published along with articles and presentations as planned by the project partners.

4.3 Open source availability

Justification

The MADES tool set will provide a platform for industrial users, technology suppliers and researchers to create more advanced and reliable embedded systems. While the project will undertake two Use Cases, the broader objective is to see that the tool set is utilised in a wide range of industrial applications and new commercial offers, as well as expanded through continued development and contributions of new innovations and technologies. The percentage of the MADES technology components available in open source provides a measure of the degree to which the project results have been made accessible to industry, developers, and researchers.

Specification

The MADES tool chain has been described from an architecture standpoint and has well-defined components and interfaces indicating the relationship and interaction between components. The MADES project will provide a set of reference implementations of each tool along with a methodology guidebook that describes how the tools are used and deployed within a development organisation. The measure of open source availability will be a ratio of the following:

$$\text{Open Source Availability} = \frac{\text{Number of MADES tool chain components made available as open source}}{\text{Number of MADES tool chain components specified in the MADES Tool Chain architecture}}$$

MADES may choose to include for some components de facto standard implementations from proprietary vendors or other suppliers, which are not available in open source format. Should this occur, two ratios will be measured: one which includes all components; and one which excludes components that are industry standards but not available in open source. This will provide a more accurate view of the components developed or adapted by MADES that have been made available in open source.

4.4 Patents resulting from MADES research

Justification

Information about the number and nature of patents generated and being applied for by the MADES partner organisations, with other external partners or with existing or potential customers gives a measure of the importance of innovations and expected exploitation of project results. The number of patents is also a strong indicator of the level of development and commercial readiness for exploitation.

Specification

In order to measure the indicator the following information will be collected:

MADES partners	Patent name	Patent status	Intention of patent

Where information provided under the headings will be the following:

- **Patent status**: indicates whether the patent is accepted and filed by the patent agency or pending.
- **Intention of patent**: describes briefly the purpose of the patent/relevance for exploitation (basis for developing a MADES product or service or for licensing).

4.5 Projects/initiatives using MADES results

Justification

Collaboration with other projects and initiatives leverages additional resources that can help in the take-up of the MADES project results by industry, while providing additional inputs and collaborations in carrying out the research and development in the project. The degree to which other projects and initiatives are using the project results (specifications, technologies, methods, etc.) provides an indication of the success that the MADES project is achieving in becoming the basis for or contributing to new embedded system industry initiatives or projects.

Specification

In order to measure the success of bringing MADES results into related initiatives and projects the following information will be collected:

Project Name	Type of project/ initiative	Partners of the consortium involved	Level of exploitation

For the *Type of Project or Initiative* we have two major groups of projects:

- Research projects that are addressing the market with a longer term perspective of at least a two-year timeframe. Typically the goal is to further enhance the results of MADES in order to be part of an exploitation activity.
- Innovation projects that are closer to the market and have a shorter horizon for the integration into commercial products or services.

The projects will be identified as being one of the following in order to provide a detailed measure for assessment:

- **Internal Research Project -** This is a research project within one or more organisations of the MADES consortium.
- **Collaborative Research Project -** This is a research project of MADES partners that involves at least one partner that's not a member of the MADES consortium.
- **External Research Project -** This is a research project that does not involve any consortium member.
- **Pilot User Innovation Project -** This is a project where MADES results, possibly combined with other results, are used to solve a particular problem for an early adopter.
- **Product or Service Innovation Project -** This is a project where it is planned to integrate MADES results into an existing product or service already offered.
- **External Innovation Project -** This is a project where MADES results are used for business purposes outside the consortium.

The *Level of exploitation* heading provides an understanding of how the results of MADES are or will be used. The following categories will be used to indicate the level of exploitation in these external projects:

- **Concepts:** This means that the other project is e.g. using documents, concepts and ideas of the project but not (yet) design, codes, or prototypes.
- **Architecture and Design:** For this level the external project must not only use documents and concepts but also artefacts from the MADES software framework. E.g. Design Patterns, Subsystems and their relations, dynamic behaviour patterns (subsystem interactions).
- **Software Components:** A project has reached this level of exploitation if components from the MADES software framework are used additionally to the artefacts above.
- **Alternative Direction:** This indicates that a project has previously been in one of the above-mentioned levels but has changed its direction away from the results / solutions from MADES.

4.6 Collaboration with other initiatives

Justification

If the results of MADES have a value for other initiatives and projects it is a clear indicator of the quality of the achieved results. It could also deliver valuable input for the exploitation activities as the collaboration on a technical level could establish the basis for joint activities towards standardisation and exploitation.

Specification

Project/Initiative Name	Type of collaboration

Where information concerning types of collaboration will be one the following:

- Informal exchange of ideas, or initial contact
- Exchange of documents, information
- Exchange of software components
- Joint developments and demonstrations

Other categories may be included as the nature of collaborations with other initiatives evolves over the life of the MADES project.

4.7 New products

Justification

A key exploitation related indicator is the level at which project results will move from research and development prototypes into products (either in source or commercial form) available to the broader community of embedded systems developers. The number of products expected towards the end of the project helps in assessing the impact of dissemination and exploitation actions of the project. It should be noted that the work within MADES is advanced research and that any new products will generally appear towards the end of the project or after the project is completed.

Specification

Products may be tools and other software products and services and the type of product will be classified into three categories:

- Commercial - products developed and sold by commercial organisations.
- Academic - products that are intended for reuse by academic research groups.
- Internal - reuse of project products for internal purposes by consortium members.

For each of the three types of products, MADES will provide the number of new products.

5 Dissemination Measures

The research and development carried out within the technical workpackages is intended to be widely disseminated and create opportunities for exploitation by the project partners as well as external organisations from industry and academia. Successful dissemination is an important indication of the quality and significance of the scientific results as it provides a means to evaluate the interest and acceptance of the project technologies amongst the wider embedded systems community targeted by the project. The measures for evaluating scientific and technical dissemination are described in no particular order.

5.1 Technical and research publications/presentations

Justification

The MADES consortium members will deliver many presentations during the operation of the project and will aim at the publication of research, technical and more industry orientated papers. Different types of dissemination aimed at different levels of awareness, understanding or role have been identified in the project plan. The number and type provide indicators of the fulfilment of the expected level of the dissemination activities.

Specification

In order to measure this indicator, data for each activity (e.g. publication or presentation) will be collected as shown in the following table:

Activity	Date	Audience or Event	Response

5.2 Courses in academic institutions

Justification

The innovations being developed within the project are expected to change the way embedded systems are developed and provide new tools for software engineers to develop complex systems. The academic community is an important target audience that will help advance the MADES technologies after the project has been completed. Actions carried out by university partners within the project that lead to MADES results being presented in new or existing curricula reinforce the exploitation of the project results.

Specification

In order to measure this indicator the information included in the table below will be collected:

Course name	Course date	Number of attendees	Academic institution	Academic level	Mandatory / Optional

Where information provided under the headings is further described as follows:

- **Academic institution**: indicates the name of the academic institution where the course is delivered.
- **Academic level**: describes the academic level the course is introduced (bachelor, master level, etc).
- **Mandatory/Optional**: denotes whether the course that is given is mandatory or optional element of the curricula of the attendees.

The MADES project results are expected to primarily be used to augment existing course materials and model driven design and development curricula.

5.3 Page views of the MADES website

Justification

As all dissemination actions will communicate the URL of the project website, the number of visits will help to identify the interest in general for further information on the project and can be used to understand if the dissemination efforts are successful. As the number of page hits are recorded on a monthly basis in the database of the website management system a temporary increase in page hits might be mapped to a particular event in order to understand the impact (e.g. of the presence of MADES partners at a particular conference). Also registration information will be collected for accessing public downloads of project results and the type of visitors can be analysed to determine the level of interest and from which industries.

Specification

The measure will be the data provided by the website management system used to realise the MADES website. The most important indicators will be:

- Absolute number of page views per month
- Increase/Decrease of pages views compared to the same month last year and previous quarter
- Type of industry of visitors registering to download MADES information

Further demographics of page views will also be collected for information purposes such as originating regions when available.

5.4 Downloads of public project documents

Justification

Most of the MADES deliverables are planned to be made available as public documents to be shared with the industry, academic and research communities interested in embedded systems development. As these documents are published the number of downloads from the project website provides an indicator of the level of interest and the degree of success the project is having in carrying out dissemination activities.

Specification

The measure will be the data provided by the website management system used to realise the MADES website. The following indicators are of most importance:

- Number of downloads per deliverable/ technical paper/ presentation
- Overall number of downloads of deliverables
- Overall number of downloads of technical papers
- Overall number of downloads of presentations

This information will be collected on a monthly basis using the website management system.

6 Management Measures

The project management measures identified by the project are well-established and common to most FP7 ICT research projects. These measures satisfy the contractual reporting requirements defined by the European Commission and provide appropriate measures to assess performance as well as any issues with regard to resource utilisation, schedule or partner performance. They are included here for completeness as they are important measures that will be evaluated in monitoring the progress of the technical workpackages.

6.1 Person month effort vs. budget

Justification

The ambitious nature of the MADES research project means that it is important to monitor human resource utilisation in order to achieve the greatest level of technology innovation and industry impact within the resource budgets available. Regular monitoring of person month effort utilised by each partner and planning of workpackage resources at each phase of the project will ensure resources are effectively used and sufficient to achieve expected project results.

Specification

The project work programme includes planning of expected effort by partner broken down by workpackage and task. The project will collect quarterly from each partner the expended effort and monitor resource usage, with any substantial variations noted.

6.2 Project costs vs. budget

Justification

The project operates under a fixed budget with respect to European Commission contributions. Costs per partner should be monitored on a regular basis to identify and mitigate situations where costs overrun could jeopardise the project achieving expected project results or industry impact.

Specification

The project work programme includes planning of expected costs by partner broken down by workpackage. The project will collect annually from each partner the costs incurred and monitor expenditures, with any substantial variations noted.

6.3 Conformance to project schedules

Justification

The project seeks to achieve on time performance of each of the project deliverables and a good measure of project operations is the degree to which tasks are completed when expected. The partners recognise that due to the nature of the advanced research and development in the project it is impossible to be extremely precise in scheduling and occasionally additional development time and/or resources to implement specific innovations may be required. This measure is still useful nonetheless as several deliverables arriving late would be cause for concern and indicate a need for corrective action either in better planning or closer monitoring.

Specification

The project work programme includes specific dates for achieving defined deliverables and milestones. The project will regularly report on the degree which deliverables and milestones have been completed on schedule, with any substantial variations noted.

7 Conclusion

The measures described above provide a strong basis for the evaluation and assessment the MADES project as it seeks to achieve the major goals set out in the Description of Work and an important area of the FP7 ICT programme. The project partners have used their experience and the GQM methodology to identify at an early stage in the project the details of each measure and how these should be presented. As the project progresses the measures may be adapted based on experience and the activities undertaken in the project and industry. The MADES project remains fully open to including additional measures in order to better manage the project and to ensure the partners and European Commission is able to fully assess the technical work achieved and the intended impact of the MADES project.