Mastering Data-Intensive Collaboration and Decision Making

FP7 - Information and Communication Technologies

Grant Agreement no: 257184
Collaborative Project
Project start: 1 September 2010, Duration: 36 months

D6.1 – The Dicode Evaluation Framework

Due date of deliverable:	31 July 2011
Actual submission date:	28 July 2011
Responsible Partner:	BRF
Contributing Partners:	UOL, BRF, IMA, PUB

Nature: ☒ Report ☐ Prototype ☐ Demonstrator ☐ Other

Dissemination Level:

- ☒ PU : Public
- ☐ PP : Restricted to other programme participants (including the Commission Services)
- ☐ RE : Restricted to a group specified by the consortium (including the Commission Services)
- ☐ CO : Confidential, only for members of the consortium (including the Commission Services)

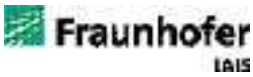
Keyword List: Evaluation, key success indicators, evaluation of Dicode services

The Dicode project (dicode-project.eu) is funded by the European Commission, Information Society and Media Directorate General, under the FP7 Cooperation programme (ICT/SO 4.3: Intelligent Information Management).

The Dicode Consortium

Research Academic Computer Technology Institute (CTI) (coordinator), Greece

University of Leeds (UOL), UK



Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (FHG), Germany

Universidad Politecnica de Madrid (UPM), Spain

Neofonie GmbH (NEO), Germany

Image Analysis Limited (IMA), UK

Biomedical Research Foundation, Academy of Athens (BRF), Greece

Publicis Frankfurt Zweigniederlassung der PWW GmbH (PUB), Germany

Document history			
Version	Date	Status	Modifications made by
1	20-06-2011	First Draft	Anastasia Kastania Sophia Kossida
2	26-06-2011	Second Draft	Anastasia Kastania
3	30-06-2011	Version sent to internal reviewers	Anastasia Kastania Sophia Kossida
4	13-07-2011	Reviewers' comments incorporated (version sent to SC)	Anastasia Kastania
5	28-07-2011	Final (approved by SC, sent to the Project Officer)	Anastasia Kastania Sophia Kossida

Deliverable managers

- Anastasia Kastania (BRF)
- Sophia Kossida (BRF)

List of Contributors

- Anastasia Kastania (BRF)
- Sophia Kossida (BRF)
- Fan Yang-Turner (UOL)
- Lydia Lau (UOL)
- Nikos Karacapilidis (CTI)

List of Evaluators

- Nikos Karacapilidis (CTI)
- Lydia Lau (UOL)

Summary

This deliverable describes the proposed Dicode Evaluation Framework. Herein, we focus on the development of an evaluation framework for the Dicode services, which aim to master data intensive collaboration and decision making in data intensive environments. The deliverable illustrates the main issues to be considered while planning the evaluation of the foreseen services. The first part concerns the evaluation of Dicode Key Success Indicators, which ensure that the overall Dicode objectives would be met by the services. The second part is devoted to general measurements of the Dicode services, which are developed and integrated in WP3-WP5. We also discuss the related data collection and analysis issues.

Table of Contents

1 Introduction

1.1 Context

As mentioned in Dicode's DoW, through WP6 the project pays much attention to the validation and assessment of the services being developed and integrated in WP3-WP5 through three real use cases. Dedicated metrics and instruments will be designed and exploited to evaluate the overall solution and assess the performance of the associated trials. WP6 provides valuable feedback for the refinement and improvement of the work being performed in WP3-WP5.

Work to be done while setting up the Dicode evaluation framework concerns: (a) the identification of indicators (metrics) on which the evaluation will focus (such as quality of services offered, improvement of productivity and creativity, Dicode solutions' usefulness and ease-of-use, as well as adaptability, accessibility and acceptability of the Dicode services), and (b) the design of instruments to elicit data and evidence, thus speeding up the gathering of user feedback.

This deliverable summarizes the outcome of Task 6.1 and proposes the Dicode evaluation framework to be adopted in the project. It examines in detail aspects which will provide measurements for (i) Key Success Indicators for the Dicode project, and (ii) the technical quality of DICODE services and integrated platform.

1.2 Objectives

The evaluation framework should provide measurable evidence on whether the project is being successfully accomplished. Furthermore, it should prove (using appropriate metrics) the successful accomplishment of the services being developed (regarding quality of services offered, improvement of productivity and creativity, Dicode solutions' usefulness and ease-of-use, as well as adaptability, accessibility and acceptability of the Dicode services).

The approach adopted by Dicode project, as described in deliverable D2.2 (entitled "The Dicode approach: User requirements, conceptual integrative architecture, agile methodology and functional specifications"), aims to:

- produce a rich set of usage scenarios which highlight the complex settings for decision making and collaboration in the use case studies (*which will provide context for subsequent evaluation)*;
- develop an agile methodology which brings together users input and application pilot development in an incremental manner *(where periodic evaluation is built in for further refinement and adaptation as necessary)*;
- produce a conceptual framework for a scalable architecture that can evolve with increasing functionalities *(which provides the big picture for the timing of integration evaluation)*.

It is also noted that Figure 16 of D2.2 outlines the role of evaluation in the Dicode Methodology.

Statistical Science allows the collection, classification, presentation and analysis of data using the statistical methods which are objective and have mathematical background and formulation (Kastania & Kossida, 2011). A continuous machine based data collection scheme (using log files) regarding all aspects of services described in Deliverables D2.2, D3.1.1 (entitled "The Dicode Data Mining Framework (initial version)"), D5.2 (entitled "An ontological framework for the capture and representation of stakeholder perspectives to augment collaboration and decision making") and their subsequent updates will allow us extract statistical reports to justify the successful accomplishment of the Dicode's innovative solutions with respect to the following issues (expected benefits and/or results):

- Ensuring of a flexible, adaptable and scalable information and computation infrastructure.
- A semi-automatic, adaptive approach that makes use of both semantic metadata and pre-structured data patterns to provide plausible recommendations, while also learning from the users' feedback to better target their information interests.
- Fully supporting the data mining process (including pre-processing, modelling, validation and deployment), integrating, adapting and extending the Mahout machine learning library, e.g. by enabling distributed data pre-processing on Hadoop, as well as by developing advanced machine learning algorithms for large scale data.
- Exploration, delivery and visualization of the pertinent information in the settings under consideration (Dicode Use Cases). This is also connected with the exploitation of particular user and group characteristics to properly direct or adapt data.
- Provision of a suite of innovative, adaptive and interoperable services (both at a conceptual and a technical level) that satisfies the full range of the requirements reported in deliverable D2.2. Measurement evidence should exist related to the adaptability of the foreseen services with respect to changes in user requirements and operating conditions
- Evaluated collaboration support services, facilitating the synchronous and asynchronous collaboration of stakeholders through adaptive workspaces and serving alternative data visualization schemas (also accommodating the functionalities of WP3 services).
- Proven evidence that the related Dicode services "support the navigation, manipulation and consumption of digital information by means of adaptive user-information interactions".
- Evidence for high quality adaptation of state-of-the-art paradigms for high-performance computing, and their exploitation for complex data mining tasks.
- Evaluation of the Dicode semantic annotation technologies to reduce the overhead of adapting the platform to a new problem domain. Dicode base its data-processing on the cloud computing paradigm and related state-of-the-art enabling technology (MapReduce, Mahout) that makes it more adaptable to various domains (as indicated by the heterogeneity of targeted Dicode use cases).
- Evidence of successful provision to the user with additional relevant and useful information in order to get an overview over his data based on what the data mining services rated as similar or interesting. More generally, we foresee a semi-automatic, adaptive approach that makes use of both semantic metadata and pre-structured data patterns to provide plausible recommendations and then learns from the user's feedback, to more specifically target his information interests.
- Evidence of accomplishment of the adaptation of the user interface to different user communities and types of activities.

- Evidence that a flexible and generic development approach is adopted, thus making the adaptation of the overall framework to other sectors easy.
- Evidence that the foreseen platform improves the quality and quantity of the collaboration process. Since needs and user types evolve over time, that the platform can be customized and adapted to address these quickly changing needs and user types.
- Evidence regarding the web-based development and the global applicability of the platform for a diverse spectrum of industrial sectors that will allow the Dicode platform to be quickly adapted for use internationally (i.e. no cultural barriers to entry).
- Evidence that through open standards, faster adaptation and integration of existing language modules can be achieved (successful modules exist for the handling and processing of multilingual text corpora).
- Evidence that the foreseen Dicode framework will be adaptable and adaptive to be able to "grow" with business needs in the relevant areas on European level.

2 Dicode Key Success Indicators

2.1 Introduction

The first part of the Dicode Evaluation Framework is set up to evaluate the outcomes of the project internally through a series of Key Success Indicators (KSIs) derived from Dicode Description of Work (specifically, from sections of Part B: B 1.1.3 - Scientific and Technical Objectives; B 3.1.2 - Contribution to the expected impacts, and B 3.2.1 - Dissemination and Exploitation). This section elaborates the principles, process and evaluation forms for this evaluation.

2.2 Principles

Internal evaluation is essential in order to improve the quality of Dicode services. To reach its goal, the evaluation process should be designed for speedy implementation in order to provide constructive feedbacks and help the development team work towards increasing service usability, flexibility and acceptability. Three basic principles that were considered for the foreseen evaluation process are:

- ***Simple***
  KSIs should be clearly presented and easily understood by all team members.
  Evaluation instruments should be easily followed.
- ***Measurable***
  KSIs should be designed as quantitative measurable metrics.
  Subjective feedbacks should be also included to support quantitative judgements.
- ***Practical***
  Feedbacks/evaluations instruments should be easily conducted.
  Feedbacks/evaluations results should be easily followed for service improvement.

2.3 Dicode Key Success Indicators (KSIs) Overview

KSIs aim to evaluate the expected impact of Dicode project in terms of both research and technological development. Consequently, KSIs include:

- maturity of technology through Technology Readiness Level (TRL);
- dissemination and exploitation activities;

- usability and acceptability of Dicode services through three use cases.

KSIs are evaluating each service (or a suite of services) when they are instantiated for each use case.

TRL is evaluated by internal reviewers, who estimate the maturity level of the service related technology categorized in nine different areas shown in Figure 1. Appendix A – Technology Readiness Level Evaluation presents the evaluation form to be completed by internal reviewers for each service.

Dissemination and exploitation is evaluated through self-report by technical partners. Appendix B – Dissemination and Exploitation Report presents the evaluation form to be completed by each partner.

Use case evaluations are conducted by use case partners. The evaluation form (Appendix C – Use Cases Validation/Service Evaluation Instrument) has two parts: Part I is completed by service providers, while Part II is completed by use case partners.

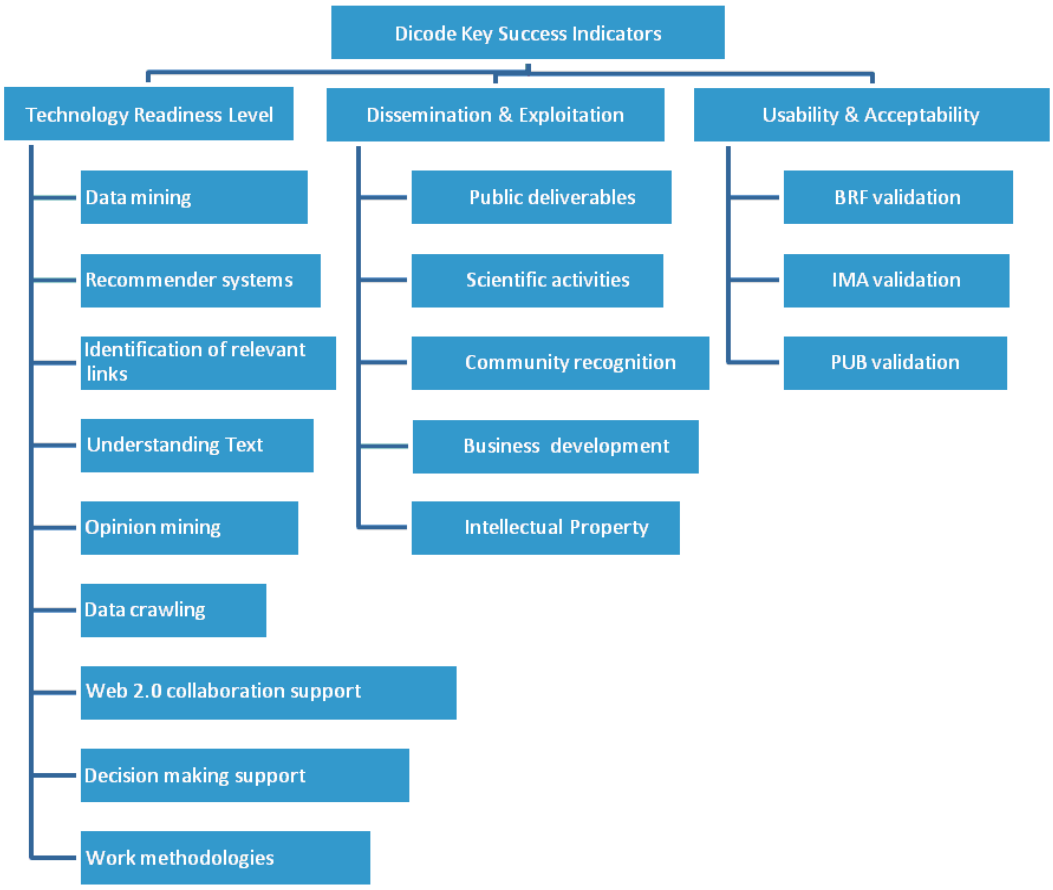


Figure 1. Dicode Key Success Indicators (KSIs)

2.4 Process

Figure 2 illustrates how Dicode services are evaluated through Dicode KSIs. According to Dicode Methodology (D2.2), the Dicode services will be evaluated every six months by both internal reviewers and use case partners. Internal reviewers and use case partners evaluate

KSIs by filling evaluation forms (see Appendices A, B, and C for sample questionnaires) on three main areas: Technology Readiness Level; Dissemination & Exploitation; and Usability & Acceptability. All evaluation result and feedback are accessible through the Dicode Wiki. Service providers then update the service and the specification to improve the quality of services based on the results/feedback.

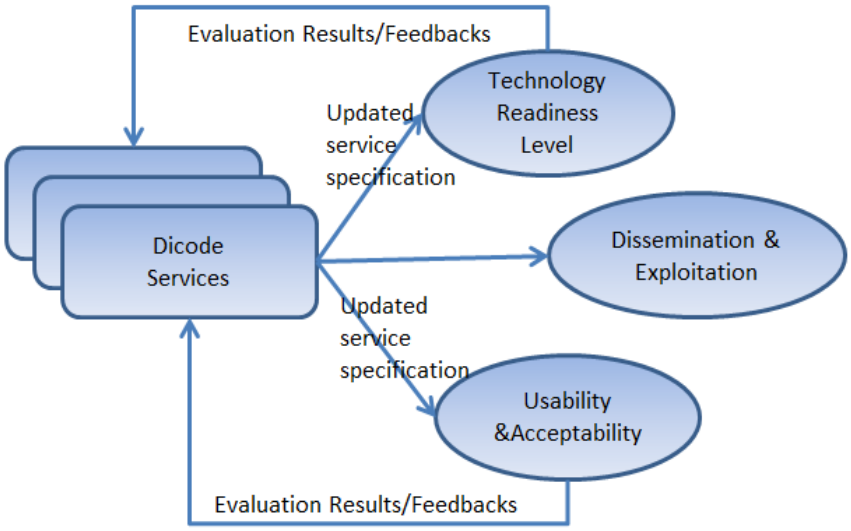


Figure 2. Dicode evaluation cycle

3 General Measurement Guidelines for the Dicode services

This section concerns the identification of indicators (metrics) on which the evaluation will focus (such as quality of services offered, improvement of productivity and creativity, Dicode solutions' usefulness and ease-of-use, as well as adaptability, accessibility and acceptability of the Dicode services).

Various metrics are already available in the scientific literature. Herein, we present a series of metrics extracted taking into account that the Dicode project should assure high quality provision of high-performance scalable data mining in the cloud computing initiative and related open source community which is related to proven added value of the Dicode framework and successful accomplishment of the Key Success Indicator Technology Readiness Level.

Regarding the evaluation of open source development tools for Cloud Computing, a Cloud Technology Capability Matrix has been mentioned in (Rodriguez-Martinez et al., 2010) with the following criteria: Security, Support, Learning Curve, Reliability, Node Extensibility, Cost of Ownership, Use of Standards and Scalability. All these issues will be included in the statistical reports of the Dicode Evaluation Framework. Furthermore, an overview of advantages and inconveniences of Open Source VM-based Cloud Management Platforms has been provided in (Cerbelaud et al., 2009).

3.1 Quality of service for Web Services

Measurement and analysis of quality attributes is necessary for the service design components (Skar et al., 2009). Provision of the desired QoS for web service users is a

difficult task (W3C Working Group Note, 2003). Web service applications also compete for system resources (bandwidth and processing time) as well as for network resources (W3C Working Group Note, 2003). Service quality evaluation aims to reveal to what degree implementation of the desired quality is successful and delivered to users. The QoS requirements have to be ensured with the minimal amount of computational resources.

Figure 3 presents the QoS requirements for web services (W3C Working Group Note, 2003). The metrics and measurement methods to assure them in the Service Oriented Architecture have been reported in (Yu et al., 2007).

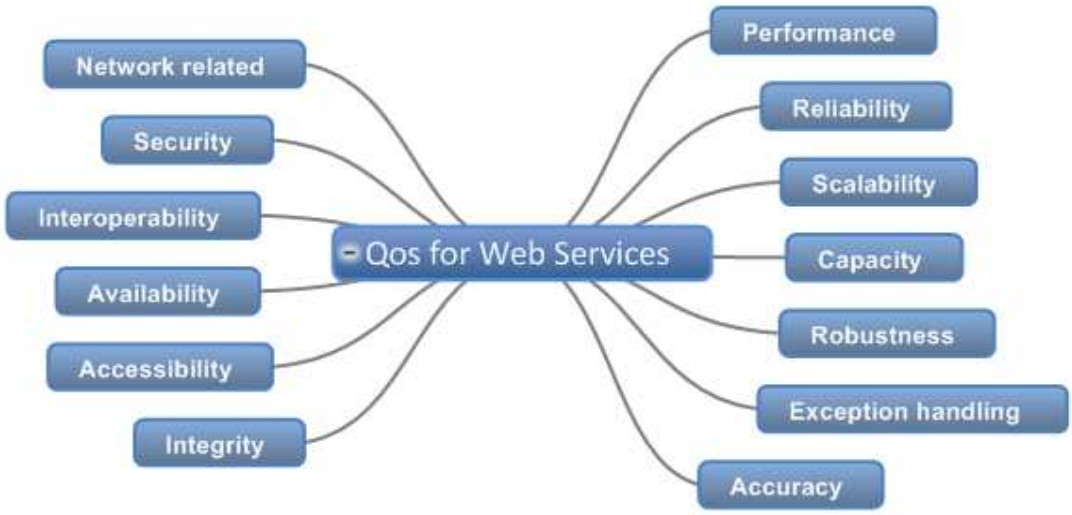


Figure 3 – QoS Requirements for Web Services (W3C Working Group Note, 2003)

It is of interest to examine the possible measurement of some of these aspects in the Dicode services (e.g. accessibility, availability, scalability and interoperability). Apart from the formulas reported for SOA (Yu et al., 2007), exploitation of the metrics described in Appendix D – Metrics involved in assuring QoS Requirements for Web Services in Cloud Computing worth to be considered.

3.2 Improvement of Productivity & Creativity

As mentioned in the Dicode proposal, the acceptance by users involved in Dicode's use cases towards increased users' productivity and creativity is a key issue for the success of the project. The proposed action is to ensure that is the early and continuous involvement of end users in the development and evaluation of the Dicode platform.

In the foreseen Dicode platform, machine-tractable knowledge concerning the full lifecycle of collaboration and decision making is accumulated and maintained. Consequently, the foreseen platform will augment the productivity of stakeholders, e.g. by enabling them to easily locate and meaningfully reuse existing content. This will certainly affect both individuals and the workgroups they belong to.

The instrument to elicit data and evidence, speeding up the gathering of user feedback is the Dicode Productivity & Creativity Assessment User Questionnaire (Appendix E). This instrument will take into account (and be accordingly adapted to) the specific needs and other particularities of each Dicode use case.

3.3 Dicode Solution's Usefulness & Ease of Use

3.3.1 Exploiting metrics extraction for Usefulness

As mentioned in the proposal "Dicode will enable both human as well as machine understandable argumentative discourses in order to support ease-of-use and expressiveness for users and advanced reasoning by the machine. Dicode will achieve this by supporting emergent semantics and the incremental formalization of argumentation discourses." Furthermore, it is written in the proposal regarding utility**.** "To aid and advance the seamless extraction, sharing and analysis of large quantities of heterogeneous data and information sources are of paramount importance for the production of meaningful conclusions and the formation of reliable and useful knowledge".

The instrument to elicit data and evidence, speeding up the gathering of user feedback is the Dicode Solution's Usefulness Assessment User Questionnaire (Appendix F). This instrument will take into account (and be accordingly adapted to) the specific needs and other particularities of each Dicode use case.

3.3.2. Exploiting metrics extraction for Ease of Use

Regarding the user oriented design various methods exist to exploit the ease of use. These are: (i) exploratory methods (record the ease of use user requirements through: interview, observation, questionnaires or recorded sessions), (ii) analytical methods (measure the design quality in usability labs); Experimental methods (record user attitudes and system performance in a usability lab).

Elements that contribute to user satisfaction are the characteristics of the Internet environment, user characteristics, the information retrieval strategy, etc. In user-centered design the following issues should be taken into account (Wang R. Y., & Strong D. M. (1996): (i) the quality categories of interface information (*internal:* the information itself must be "quality", i.e. valid, otherwise it loses its credibility the broadcasting source), (ii) *content:* The information must be relevant to contents of each project. Also the information must be complete, timely and in correct analogy, because the over-information impairs the quality, (iii) Representative: the presentation strategy of the information should be representative and proportional to the content, consistent, concise and easily understood. For this purpose the Dicode Solution's Ease of Use Assessment User Questionnaire (Appendix G) was constructed based on the concepts and notions of Norman (1998), Nielsen (1999) and Lambrakis Foundation (2005). This instrument will take into account (and be accordingly adapted to) the specific needs and other particularities of each Dicode use case.

3.4 Adaptability

Exploiting adaptability of the Dicode Collaboration and Decision Making Services requires the use of various metrics:

- Metrics for the quality of search: *precision* and *recall* (implement and understand qualitatively) (Marmanis & Babenko, 2009);
- Distance and similarity metrics between users and items: Regarding similarity the Jaccard metric, the Pearson correlation, and variants of these metrics will be used (Marmanis & Babenko, 2009);
- Creating suggestions & recommendations (Marmanis & Babenko, 2009; Herlocker et al., 2004)
    - large-scale comparisons of similarity metrics used (Marmanis & Babenko, 2009)

  - Searching the related literature, there are dozens of quantitative metrics and several qualitative methods for evaluating the results of recommendation systems (Marmanis & Babenko, 2009). The plethos of metrics and methods reflects the challenges of conducting a meaningful, fair, and accurate evaluation for recommendations (Herlocker et al., 2004).
  - Furthermore computational complexity in space and time (Marmanis & Babenko, 2009)
  - Credibility of classifiers such as the precision, the accuracy, the recall, and the specificity. Combinations of these metrics can yield new metrics, such as the F-score (Marmanis & Babenko, 2009)
- metrics to define information content (Marmanis & Babenko, 2009);
- metrics that are based on the confusion matrices (Marmanis & Babenko, 2009);
- overview of the functionality: relevance (ranking) metrics, distance metrics (Marmanis & Babenko, 2009).

3.5 Accessibility & Acceptability

Regarding accessibility issues various metrics are available in the scientific literature: failure rate (Vigo et al., 2007), Web Accessibility Barrier score (Vigo et al., 2007; Freire et al., 2008), Unified Web Evaluation Methodology (Freire et al., 2008), Web Accessibility Quality Metric (Freire et al., 2008), WCAG 1.0 checkpoints (Freire et al., 2008).

These metrics will be used in Dicode Accessibility & Acceptability evaluation.

The instrument to elicit data and evidence, speeding up the gathering of user feedback is the Dicode Solution's Accessibility & Acceptability User Assessment Questionnaire (Appendix H). This instrument will take into account (and be accordingly adapted to) the specific needs and other particularities of each Dicode use case.

3.6 User feedback gathering

The instruments designed to gather user feedback are presented in Appendices E-H. These instruments will take into account (and be accordingly adapted to) the specific needs and other particularities of each Dicode use case.

4 Data collection and analysis issues

The data regarding the provision of evidence about the quality of web services of Dicode Approach will be collected in logs or databases provided by the Technical Partners of the Consortium and according to agreement of the Consortium regarding the criteria that will be used.

A continuous data collection scheme regarding all aspects of services described in Deliverables D2.2, D3.1.1 and D5.2 will allow us to extract metrics and statistics justifying the successful accomplishment of the services developed (regarding quality of services offered, improvement of productivity and creativity, Dicode solutions' usefulness and ease-of-use, as well as adaptability, accessibility and acceptability of the Dicode services).

During the evaluation phases of use cases the data will be statistically analyzed to provide supporting evidence quantifying successful and innovative project outputs and outcomes.

During the data collection and analysis processes, the Commission Regulation (EC) No 753/2004 of 22 April 2004 implementing Decision No 1608/2003/EC of the European Parliament and of the Council as regards statistics on science and technology (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32004R0753:EN:NOT) will be applied.

A series of data analysis techniques will be used (Kastania & Kossida, 2011), including Ethnography (Fetterman, 1998) and Statistical Quality Control Techniques (Montgomery, 2004; Kastania & Kossida, 2011). Special emphasis will be also given on issues of validity and reliability in data analysis (Kastania & Kossida, 2011).

5 Conclusions

In this document we have described the Dicode Evaluation Framework. We have focused on the development of an evaluation framework for open source services provision for mastering data intensive collaboration and decision making in data intensive environments. This description illustrated the main aspects to be considered while planning the evaluation. The first aspect presented is the Evaluation of Dicode Key Success Indicators. The second aspect presented is General Measurement guidelines of the services to be developed and integrated in WP3-WP5 through three real use cases. Furthermore, data collection issues and data analysis issues were presented.

References

Dicode Proposal

Dicode Deliverable D2.1 - Dealing with data intensiveness and cognitive complexity in contemporary collaboration and decision making settings: A state-of-the-art survey

Dicode Deliverable D2.2 - The Dicode approach: User requirements, conceptual integrative architecture, agile methodology and functional specifications

Dicode Deliverable D3.1.1 - The Dicode Data Mining Framework (initial version)

Dicode Deliverable D5.1.1 - Standards and guidelines for development (initial version)

Dicode Deliverable D5.2 - An ontological framework for the capture and representation of stakeholder perspectives to augment collaboration and decision making

Babcock, C. (2011). *IEEE Targets Cloud Interoperability Standards*. www.informationweek.com/cloud-computing/article/229400890?pgno=2.

Cerbelaud, D., Garg, S., & Huylebroeck, J. (2009). Opening The Clouds: Qualitative Overview of the State-of-the-art Open Source VM-based Cloud Management Platform. In *Proceedings of the 10th ACM/IFIP/USENIX International Conference on Middleware* .

Chieu, T. C., Mohindra, A., Karve, A. A., & Segal, A. (2009). Dynamic Scaling of Web Applications in a Virtualized Cloud Computing Environment. In *the Proceedings of 2009 IEEE International Conference on e-Business Engineering* (pp. 281-286). IEEE Computer Society.

Cloud Security Alliance SM. (2010). In Cloud Security Alliance (Ed.), *Domain 10: Guidance for Application Security V2.1*. https://cloudsecurityalliance.org/guidance/csaguide-dom10-v2.10.pdf.

Fetterman, D.M. (1998). Ethnography, Second Edition, Step by Step, SAGE Publications.

Freire, A. P., Bittar, T. J., & R. P. M. Fortes. (2008). An approach based on metrics for monitoring web accessibility in Brazilian municipalities web sites. In *the Proceedings of the 2008 ACM Symposium on Applied Computing* ACM Digital Library.

Fu, C., Ryder, B. G., Milanova, A., & Wonnacott, D. (2004). Testing of java web services for robustness. In *the Proceedings of the 2004 ACM SIGSOFT international symposium on Software testing and analysis* .

Herlocker, J.L., J.A. Konstan, L.G. Terveen, and J.T. Riedl (2004). Evaluating Collaborative Filtering Recommender Systems. *ACM Transactions on Information Systems,* Vol 22, 5-53. ACM Press, 2004.

Kastania, A.N. & Kossida, S. (2011). Statistics in Telemedicine, chapter in the book "Advances in Telemedicine: Technologies, Enabling Factors and Scenarios", INTECH Open Access Publisher.

Lambrakis Foundation (2005). Quality Guide for Websites of Cultural Agencies – funded by the Operational Programme Information Society under the 3rd Community Support Framework [online]: http://www.infosoc.gr/meletes/

Lee, J. Y., Lee, J. W., Cheun, D. W., & Kim S.D. (2009). A Quality Model for Evaluating Software-as-a-Service in Cloud Computing. In *the Proceedings of 2009 Seventh ACIS International Conference on Software Engineering Research, Management and Applications* (pp. 261-266). IEEE Computer Society.

Marmanis, H. & Babenko, D. (2009). Algorithms of the Intelligent Web, Manning Publications Co.

Montgomery, D.C. (2004). *Introduction to Statistical Quality Control*, Wiley.

Nielsen, J. (1999). Designing Web Usability: The Practice of Simplicity, New Riders Publishing, Indianapolis.

Norman, D. A. (1998). The Design of Everyday Things, The MIT Press.

Rodriguez-Martinez, M., Seguel, J., & Greer, M. (2010). Open Source Cloud Computing Tools: A Case Study with a Weather Application. In *the Proceedings of 2010 IEEE 3rd International Conference on Cloud Computing* (pp. 443-449).

Skilton, M. (2010). *The 8 easiest ways to measure cloud ROI*. www.computerworlduk.com/how-to/infrastructure/3373/the-8-easiest-ways-to-measure-cloud-roi/

Skar, L. A., Lennon, R., Berre, A.J. (Ed.). (2009). T*owards best practices in Cloud Computing*. Orlando, USA.

Oriol, M., Marco., J., Franch., X., & Ameller, D. (2008). Monitoring Adaptable SOA-Systems using SALMon. In *the Proceedings of MoNA+* (pp. 19-28).

Thangarathiman, T. (2003). *Exception Handling in Web Services*. www.developer.com/net/csharp/article.php/3088231/Exception-Handling-in-Web-Services.htm

Vigo, M., Arrue, M., Brajnik, G., Lomuscio, R., & Abascal, J. (2007). Quantitative metrics for measuring web accessibility. In *the Proceedings of the 2007 International Cross-Disciplinary Conference on Web Accessibility* ACM Digital Library.

W3C Working Group Note 25 November 2003. (2003). In Lee, KC., Jeon, JH., Lee, WS., Jeong, & S-W. Park (Eds.), *QoS for Web Services: Requirements and Possible Approaches*. www.w3c.or.kr/kr-office/TR/2003/NOTE-ws-qos-20031125/.

Wang R. Y., & Strong D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers, Journal of Management Information Systems, 12, 5-34.

Wu, J., Liang, Q., & Bertino, E. (2009). Improving Scalability of Software Cloud for Composite Web Services. In *the Proceedings of 2009 IEEE International Conference on Cloud Computing* (pp. 143-146). IEEE Computer Society.

Yu, W. D., Radhakrishna, & Kolluri, V. (2007). Modeling the Measurements of QoS Requirements in Web Service Systems. *Simulation, 83*(1), 75-91.

Appendices

Appendix A - Technology Readiness Level Evaluation

Service Description (to be completed by service provider)

Name:
Location of service documentation/source code:
Version:
Provider (Organization):

Initial Maturity Level: ☐ TRL 1 ☐ TRL 2 ☐ TRL 3 ☐ TRL 4 ☐ TRL 5 ☐ TRL 6 ☐ TRL 7
Targeted Maturity Level: ☐ TRL 1 ☐ TRL 2 ☐ TRL 3 ☐ TRL 4 ☐ TRL 5 ☐ TRL 6 ☐ TRL 7

Service Evaluation (to be completed by internal reviewer)

Internal Reviewer (Organization):
Evaluation Date:
Maturity Level Evaluation: ☐ TRL 1 ☐ TRL 2 ☐ TRL 3 ☐ TRL 4 ☐ TRL 5 ☐ TRL 6 ☐ TRL 7

Reviewer Comments:

Notes
The TRLs used for Dicode estimation are:

- TRL1: Scientific research begins to be translated into applied research and development, basic properties
- TRL2: Practical applications are invented based on basic principles; speculative application, no detailed analysis
- TRL3: Analytical or laboratory studies to validate analytical predictions, not yet integrated or representative
- TRL4: Basic components are integrated to show how they work together, allows ad-hoc models
- TRL5: Components are integrated with reasonably realistic supporting elements, tests can commence
- TRL6: Representative model or prototype system, tested in a relevant environment, restricted conditions
- TRL7: Demonstration of an actual model or a system prototype in a realistic operational environment
- TRL8: Technology/models proven to work in final form and under expected conditions, end of true development
- TRL9: Actual application of the models and technology in its final form and under realistic operational conditions

Appendix B - Dissemination & Exploitation Report

Partner:

Dissemination & Exploitation Activities

Category	Activities	Impact
Deliverable		
Scientific Activities		
Community Recognition		
Business Development		
Intellectual Properties		

Appendix C - Use Cases Validation / Service Evaluation Instrument

This evaluation instrument includes two parts:

Part I is for service provider (to provide a service specification).

Part II is for service evaluator (to evaluate the service in terms of service quality, which includes a scenario walkthrough, post walkthrough questionnaire, acceptability evaluation and overall quality evaluation). For single-choice questions, please also give comments to explain why the choice has been made.

Part I - Service Specification

Provider (Organization):

Note: The most up-to-date service specification should be uploaded to Dicode Wiki.

Service Name	
Overview	
Objectives	
Requirements	
Use cases	
Scenario instruction (step-by-step to use the service. This will be used for usability evaluation, see part II)	

Part II - Service Evaluation by Use Case Partner / Service User

Evaluator Name:
Evaluator Organization:
Evaluation Date:

1. Scenario Walkthrough (follow the scenario instruction described in Part I)

Tasks	User action	System response	Evaluator's comments

2. Post Walkthrough Questionnaire (please circle your answer and comment on your choice)

- **I am satisfied with the** instruction **given to complete the scenario.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I am satisfied with the** time **spent to learn to use this service.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I am satisfied with the** ease **of completing the scenario.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I am satisfied with the** time **of completing the scenario.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I am satisfied with the** information **(system message, system help etc.) provided by the service.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

3. Acceptability Evaluation (please circle your answer and comment on your choice)

- **I found the service very** helpful **to my work.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **The service has all the** functions **I expected.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **The** interface **of the service is pleasant.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I will** use **this service in my work.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I will** recommend **this service to my peers/community.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

4. Overall Quality (please circle your answer and comment on your choice)

- **I think the** scenario **addressed the data-intensive decision making issues in my work.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I think the** objectives **of the service are met.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I think the service is** novel **to my knowledge.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **Overall, I am satisfied with the** performance **of the service (answer and comment).**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **Overall, I am** satisfied **with this service.**
  Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

Finally, I would like to provide comments not covered by above questions.

Appendix D - Metrics assuring QoS Requirements for Web Services in Cloud Computing

QoS requirement for web services	Metrics
Performance	• throughput (W3C Working Group Note, 2003) • response time (W3C Working Group Note, 2003) • current response time in milliseconds to access to a Web Service (Oriol et al., 2008) • minimum response time in milliseconds to access to a Web Service (Oriol et al., 2008) • maximum response time in milliseconds to access to a Web Service (Oriol et al., 2008) • average response time in milliseconds to access to a Web Service (Oriol et al., 2008) • latency (W3C Working Group Note, 2003) • execution time (W3C Working Group Note, 2003) • transaction time (W3C Working Group Note, 2003)
Reliability	• number of failures per day, week, month, or year (W3C Working Group Note, 2003) • assured and ordered delivery for messages being transmitted (W3C Working Group Note, 2003) • assured and ordered delivery for messages being received (W3C Working Group Note, 2003) • Service stability (Lee et al., 2009) • Service accuracy (Lee et al., 2009)
Scalability	• the bandwidth consumption per composite service (Wu et al., 2009) • number of concurrent users (Chieu et al., 2009) • number of active connections (Chieu et al., 2009) • number of requests per second (Chieu et al., 2009) • average response times per request (Chieu et al., 2009)
Capacity	• number of simultaneous requests completed with the guaranteed performance per second (W3C Working Group Note, 2003) • IT capacity per second (Skilton, 2010) • Storage (Skilton, 2010) • CPU cycles (Skilton, 2010) • network bandwidth (Skilton, 2010) • memory workload (Skilton, 2010)
Robustness	• dataflow test coverage metrics (Fu et al., 2004) (e.g. overall fault-catch coverage metric) • control-flow coverage metrics (Fu et al., 2004) (branch, edge, basic block coverage) • proposed in Deliverable 5.1.1 • number of test cases • number of classes covered • number of classes not covered • code coverage

Table 1 – Metrics involved in assuring QoS Requirements for Web Services in Cloud Computing

QoS requirement for web services	Metrics
Exception handling	• message-contents of the exception (Thangarathiman, 2003) • type of fault code (ClientFaultCode, ServerFaultCode) (Thangarathiman, 2003) • actor-URL of the Web service method where the exception has occurred (Thangarathiman, 2003) • detail-detail element to be used to provide more information to the callers about the exception (Thangarathiman, 2003) • number of ClientFaultCode detected per second (Thangarathiman, 2003) • number of ServerFaultCode detected per second (Thangarathiman, 2003)
Accuracy	• current functionality compliance of a Web Service (Oriol et al., 2008) • parameter accuracy factor: the result of the division of accepted correct type parameters passed to a Web Service method with the number of expected parameters for this method (1 is better) (Oriol et al., 2008) • result accuracy factor: the result of the division of the number of correct type results returned by a Web Service method with the number of expected results of this method (1 is better) (Oriol et al., 2008) • fault factor: the result of the division of the number of faults a Web Service method had generated with the total times that this method had been executed (1 is worse) (Oriol et al., 2008)
Integrity	• Integrity = number of successful transactions / total number of transactions (Yu et al., 2007)
Accessibility	• failure rate (Vigo et al., 2007) • Web Accessibility Barrier score (Vigo et al., 2007; Freire et al., 2008) • Unified Web Evaluation Methodology (Freire et al., 2008) • Web Accessibility Quality Metric (Freire et al., 2008) • WCAG 1.0 checkpoints (Freire et al., 2008)
Availability	• current availability: of a Web Service all the time in slots of time (Oriol et al., 2008) • accumulative availability time: the percentage how much time a Web Service has been available since it has been monitored for the first time (Oriol et al., 2008) • accumulative unavailability time: the percentage how much time a Web Service has been unavailable since it has been monitored for the first time (Oriol et al., 2008) • average recovery time: measures in seconds the average time that a Web Service needs to be available again (Oriol et al., 2008)

Table 1 (cont.) – Metrics involved in assuring QoS Requirements for Web Services in Cloud Computing

QoS requirement for web services	Metrics
Interoperability	• external interfaces tests & measurements (to be determined, (Babcock, 2011) • internal context tests & measurements (to be determined, (Babcock, 2011))
Security	• Compliance and Governance Metrics (Cloud Security Alliance SM., 2010) • Identity and Access Metrics (Cloud Security Alliance SM., 2010) • Threat and Virus Metrics (Cloud Security Alliance SM., 2010) • Vulnerability and Patch Metrics (Cloud Security Alliance SM., 2010) • Data Security Metrics (Cloud Security Alliance SM., 2010)
Network related	• network delay per second (W3C Working Group Note, 2003) • delay variation per second (W3C Working Group Note, 2003) • packet loss per second (W3C Working Group Note, 2003)

Table 1 (cont.) – Metrics involved in assuring QoS Requirements for Web Services in Cloud Computing

Appendix E - Productivity & Creativity Assessment User Questionnaire

A. Productivity Augmentation Measurement

A1. *Rate the value of* ***scientific interactions*** in the Dicode platform (in a scale from 0 [none] to 10 [excellent]) _____

A2. *User Productivity*

- Rate the **time savings** of the Dicode platform in your working processes (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the capability of the Dicode to easily **locate existing content** (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the **search times** of the Dicode (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the capability of the Dicode to allow you **meaningfully reuse existing content** (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the capability of the Dicode to provide you **easier access to information** (in a scale from 0 [none] to 10 [excellent]) _____

A3. *Network structure impact*

- Rate your overall satisfaction from knowledge sharing using Dicode collaboration tools (services) (in a scale from 0 [none] to 10 [excellent]) _____
- Rate your overall satisfaction from knowledge improvement using Dicode decision making tools (services) (in a scale from 0 [none] to 10 [excellent]) _____

A4. *Data about completed processes*

Data about **USE CASE processes** completed

- Number of times a use case service was used (per service type) _____
- Overall satisfaction from each kind of each use case service offered (in a scale from 0 [none] to 10 [excellent]) _____

Data about **collaboration processes** completed _____

- Number of times a collaboration tool (service) was used per service type
- Overall satisfaction from each kind of collaboration tool (service) offered(in a scale from 0 [none] to 10 [excellent]) _____

Data about **decision making processes** completed

- Number of times a decision making tool (service) was used per service type _____
- Overall satisfaction from each kind of decision making tool (service) offered (in a scale from 0 [none] to 10 [excellent]) _____

B. Creativity Augmentation Measurement

- Rate the overall satisfaction from the platform to assist your long term development (*Incubate*) (in a scale from 0 [none] to 10 [excellent]) _____

- Rate the overall satisfaction from the platform to assist you in the development of breakthrough Ideas (*Imagine*) (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the overall satisfaction from the platform to assist you in Improving your skills and knowledge (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the overall satisfaction from the platform to complete your short term goals (*Invest*) (in a scale from 0 [none] to 10 [excellent]) _____
- Rate your overall satisfaction from creativity improvement using Dicode collaboration tools (services) (in a scale from 0 [none] to 10 [excellent]) _____
- Rate your overall satisfaction from creativity improvement using Dicode decision making tools (services) (in a scale from 0 [none] to 10 [excellent]) _____

Appendix F - Usefulness Assessment User Questionnaire

- Rate the solutions that the service provides to remedy the *rising number of information* (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the solutions that the service provides to remedy *uncertainty of information* (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the solutions that the service provides to remedy *information diversity and increasing number of alternatives* (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the solutions that the service provides to remedy *ambiguity of information:* the situation where information can be interpreted in several ways (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the solutions that the service provides to remedy *complexity of information* (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the solutions that the service provides to deal with *intensity of information*: importance of particular information items (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the provided *information quality and value*: the degree of worth of information (in a scale from 0 [none] to 10 [excellent]) _____
- Rate the solutions that the service provides to remedy *overabundance of irrelevant information* (in a scale from 0 [none] to 10 [excellent]) _____

Appendix G - Ease of Use Assessment User Questionnaire

A. Usability principles (based on concepts and notions of Norman (1998), Nielsen (1999))
In a scale from 0 [none] to 10 [excellent],

- Rate the Visibility: optical information to facilitate the user _____
- Rate the Tolerance: behavior similar to expectations _____
- Rate the Physical mapping: conceptual correspondence between commands and functions _____
- Rate the Restrictions: Incomplete ways in the actions performed and in the design of operations in understandable manner _____
- Rate the Conceptual Models: the operation of the proposed actions according to the perception of user for these actions _____
- Rate the Feedback: Notification regarding the user's position _____
- Rate the Error prevention: Restrict user errors & support for their resolution _____
- Rate the Flexibility: Variety of operation modes _____
- Rate the Ease of Recognition: easy identification of the required actions _____
- Rate the Flexibility of the use efficiency: Shortcuts provision, capabilities for the user to affect the configuration of the system _____
- Rate the Provision of clear error messages: Simple language in error messages and a proposals to resolve them _____
- Rate the Aesthetics of the minimalist design: Messages with the necessary information _____
- Rate the Help facilities and the Documentation facilities: Help facilities related to user actions _____
- Rate the Visibility of the system status: user briefing regarding the work progress ___
- Rate the System matching with the real world: Comprehensive and explanatory language _____
- Rate the User Control capabilities and the freedom of action: understandable and direct processes, as undo & redo _____
- Rate the Consistency and presentation standards: Maintain the same presentation of the interface _____
- Rate the Fault prevention: Mechanism for preventing error occurrence _____
- Rate the Ease of detection: obvious & easy actions required during service use _____

B. Usability indicators (according to Lambrakis Foundation)

Ease of learning how to use the system	□ No □ Yes
Overall, good performance in the system operation	□ No □ Yes
Maintenance of system usability over time	□ No □ Yes
Few bugs in the system operations and ways to resolve them	□ No □ Yes
User satisfaction according to his/her own criteria	□ No □ Yes

Appendix H - Accessibility & Acceptability Assessment User Questionnaire

- **I am satisfied with the** access **given to the Dicode Solution.**

Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I am satisfied with the** instruction **given by the Dicode Solution.**

Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I am satisfied with the** time **spent to learn to use the provided solution.**

Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree

- **I am satisfied with the** ease **of use of the solution to complete my tasks.**

Strongly disagree ☐1 ☐2 ☐3 ☐4 ☐5 strongly agree