VRE4EIC

A Europe-wide Interoperable Virtual Research Environment to Empower Multidisciplinary Research Communities and Accelerate Innovation and Collaboration

Deliverable D2.2

Evaluation methodology

Document version: 1.0

VRE4EIC DELIVERABLE

Name, title and organisation of the scientific representative of the project's coordinator:

Mr Philippe Rohou t: +33 4 97 15 53 06 f: +33 4 92 38 78 22 e: philippe.rohou@ercim.eu

GEIE ERCIM, 2004, route des Lucioles, Sophia Antipolis, F-06410 Biot, France

Project website address: http://www.vre4eic.eu/

Project	
Grant Agreement number	676247
Project acronym:	VRE4EIC
Project title:	A Europe-wide Interoperable Virtual Research Environment to Empower Multidisciplinary Research Communities and Accelerate Innovation and Collaboration
Funding Scheme:	Research & Innovation Action (RIA)
Date of latest version of DoW against which the assessment will be made:	14.01.2015
Document	
Period covered:	M1-M6
Deliverable number:	D2.2
Deliverable title:	Evaluation Methodology
Contractual Date of Delivery:	31.03.2016
Actual Date of Delivery:	31.03.2016
Editor (s):	-
Author (s):	Anneke Zuiderwijk, Yi Yin, Marijn Janssen (TU Delft)
Reviewer (s):	Jacco van Ossenbruggen (CWI), Zhiming Zhao (UvA)
Participant(s):	All
Work package no.:	2
Work package title:	Requirements, use cases and usability, and evaluation
Work package leader:	Yi Yin (TU Delft)
Distribution:	PU
Version/Revision:	1.0
Draft/Final:	Final
Total number of pages (including cover):	27

What is VRE4EIC?

VRE4EIC develops a reference architecture and software components for VREs (Virtual Research Environments). This e-VRE bridges across existing e-RIs (e-Research Infrastructures) such as EPOS and ENVRI^{PLUS}, both represented in the project, themselves supported by e-Is (e-Infrastructures) such as GEANT, EUDAT, PRACE, EGI, OpenAIRE. The e-VRE provides a comfortable homogeneous interface for users by virtualising access to the heterogeneous datasets, software services, resources of the e-RIs and also provides collaboration/communication facilities for users to improve research communication. Finally it provides access to research management /administrative facilities so that the end-user has a complete research environment.

Disclaimer

VRE4EIC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 676247.

Table of Contents

1 Introduction

The VRE4EIC project aims at making it easier for researchers to reuse heterogeneous scientific datasets from multiple disciplines. The project will deliver a VRE reference architecture, a reference implementation, and prototypes for e-Research Infrastructures (e-RIs) (EPOS and 20 others in the cluster of ENVRI^{PLUS}) removing barriers of existing e-RIs and providing a single point of homogeneous access to heterogeneous data and tools that support data reuse.

In order to continuously identify user requirements and to maximize its impact, VRE4EIC is setting up a systematic methodology for evaluating the project results and for assessing their impact. Evaluations can be used to assess what is meaningful (Patton, 2015) and to determine the reliability and usefulness of findings (OECD, 1998). Evaluation can take place for obtaining insight, optimization, controlling and monitoring, decision-making and legitimation (see section 2) (Alexopoulos, Charalabidis, & Loukis, 2012). The objectives of the VRE4EIC evaluation are described in Task 2.3. They include:

- Define in detail the conditions for the evaluation of the VREs;
- Determine the criteria that will be used to assess the VRE architecture, prototypes and use cases.
- Develop and use questionnaires to analyse users' satisfaction while using the VRE4EIC prototypes;
- Apply a method for measuring system, service and project success;
- Develop an analytical conceptual evaluation method, which will be used in task 2.4.

An evaluation methodology is proposed based on technology acceptance theories, on Deliverable 6.6.1 of the ENGAGE project, produced by Alexopoulos et al. (2012), and on papers about project evaluation resulting from the ENGAGE project, such as: Alexopoulos, Loukis, Charalabidis, and Zuiderwijk (2013); Alexopoulos, Zuiderwijk, Charalabidis, Loukis, and Janssen (2014).

This deliverable describes the details of the methodology for evaluating VRE architectures, prototypes and use cases, and for evaluating the impact of the VRE4EIC project after M6. The approach of the evaluation will be described in different chapters of the deliverable as depicted in Figure 1.

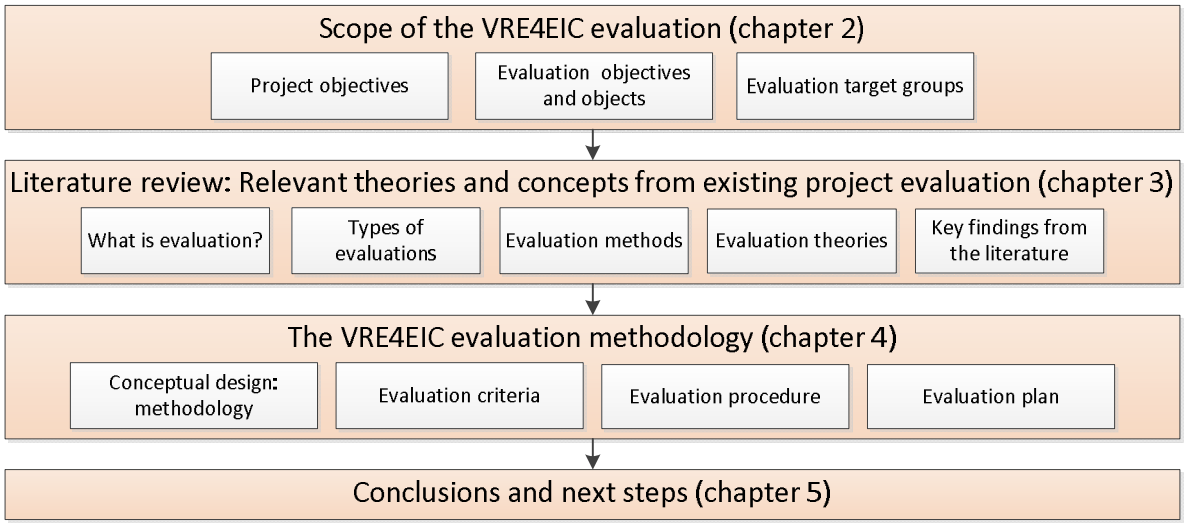

Figure 1: Evaluation approach.

2 Scope of VRE4EIC evaluation

2.1 VRE4EIC project objectives

The objectives of the VRE4EIC project that need to be evaluated are:

1. Increase VRE **usability** in different interdisciplinary domains by closely involving user communities and real-world use cases in the VRE development.
2. Increase the **quality of VRE user experiences** by providing user centred, secure, privacy compliant, sustainable environments on searching data, composing workflows and tracking data publications.
3. Increase the **deployment** of the VRE on different clusters of research infrastructures by abstracting and reusing building blocks and workflows from existing VREs, infrastructures and projects.
4. Improve the **contextual awareness and interoperability** of the metadata across all layers of the resources in the VRE.
5. Promote the **exploitation and standardisation** of the VRE4EIC solution to different research domains and communities.

2.2 Evaluation objectives and objects

The objectives of the VRE4EIC evaluation are described in Task 2.3. They include:

- Define in detail the conditions for the evaluation of the VREs;
- Determine the criteria that will be used to assess the VRE architecture, prototypes and use cases. Evaluation criteria will include system criteria (e.g. speed and availability) and usage criteria (e.g. usefulness).
- Develop and use questionnaires to analyse users' satisfaction while using the VRE4EIC prototypes;
- Apply a method for measuring system, service and project success;
- Develop an analytical conceptual evaluation method, which will be used in task 2.4.

The VRE4EIC evaluation encompasses the evaluation of the VRE architecture, the three VRE prototypes (including the canonical prototype of the e-VRE and the two domain-specific prototypes for each of the improved existing e-RIs from the EPOS and ENVRI projects building on the canonical prototype) and the 25 use cases developed in the project.

2.3 Evaluation target groups

The viewpoints of the following target groups on the VRE architecture, VRE prototypes and use cases are central to the VRE4EIC evaluation:

- Researchers as *VRE users*, including academic and governmental researchers, research managers, educators, students, innovators, entrepreneurs and the interested citizen;
- *VRE developers*, including commercial (large IT companies, SMEs, entrepreneurs) and non-commercial (universities, not-for-profit organisations, foundations, VRE related projects) developers;
- *Scientific VRE researchers*, including academics who conduct research on VREs, for instance on VRE components and VRE communities;

- *VRE data publishers*, i.e. publishers who wish their data to be available to VRE users, including research institutions and archives, universities, governmental organisations, various researchers and other data publishers.
- *Other*. At the same time, we envision other potential target groups, such as journalists, educators and students, although these groups are not key to the project.

These target groups may overlap. For example, data publishers can also be VRE users. The target groups will be targeted especially in the domains of earth and environmental sciences related to other sciences (e.g. social sciences, humanities, life sciences, physics and other domains), as well as in the other domains mentioned in the Description of Work.

3 Relevant theories and concepts from existing project evaluation

In this chapter we review theories and concepts from existing project evaluations that are relevant in the context of the VRE4EIC evaluation methodology. We start by explaining what evaluation is (section 3.1) and for which purposes one can evaluate. The purpose of evaluation influences the (combination of the) type of evaluation that is needed. Different types of evaluation are explained in section 3.2. Each type of evaluation needs to be related to an appropriate evaluation method. For instance, quantitative evaluations (the type) can take place using quantitative surveys (the method), while qualitative evaluations (the type) can be done using case studies and observations (the methods). In section 3.3, we explain which methods can be used for which evaluation types. Nevertheless, the method does not refer to the variables that are evaluated. The literature and several theories provide insight in which variables might be included in evaluation activities. In section 3.4, various evaluation theories and their key variables are explained. For instance, those theories include variables like 'perceived ease-of-use' and 'perceived usefulness' and provide directions for the variables that may be included in quantitative surveys. Finally, section 3.5 summarizes the key findings from the literature.

Building on the relevant theories and concepts from existing project evaluation, as described in this chapter, the evaluation methodology for the VRE4EIC project will be developed (see chapter 4).

3.1 Evaluation definition

Evaluation refers to making judgements of what is meaningful (Patton, 2015). It is an analytical assessment identifying the reliability and usefulness of findings (OECD, 1998). For instance, it concerns "assessing something against certain standards or criteria, determining its usefulness or quality, comparing it against other similar programs or products" (McCain, 2005, p. 9). As mentioned by Alexopoulos et al. (2012), evaluation research can generally be seen as the application of empirical research methods for a particular objective. Evaluation can take place for various purposes, including:

- Insight: obtaining insight in attitudes and intervention impacts;
- Optimization: the strengths, weakness and opportunities of interventions can be elicited;
- Controlling and monitoring: evaluating the effects and efficiency of interventions;
- Decision-making: making decisions about whether a certain intervention should be promoted, implemented, developed or used;
- Legitimation: support the development or implementation of an intervention to an external stakeholder (Alexopoulos et al., 2012).

There is no standard way of evaluating project results such as the results from VRE4EIC, and a customized methodology focused on the particular objectives of our project is therefore needed. In the context of the VRE4EIC project, the evaluation definition of Hevner and Chatterjee (2010) suits our purposes well: evaluation can be defined as "the systematic determination of merit, worth, and significance of something […] or someone" (p. 109).

The VRE4EIC evaluation methodology combines the above-mentioned types of evaluation of Alexopoulos et al. (2012). These types of evaluation are combined so that the outcomes of the VRE4EIC project are evaluated from a variety of perspectives, which may provide different insights.

3.2 Types of evaluations

The literature describes different types of evaluations that are relevant within the scope of the VRE4EIC project.

- Quantitative and qualitative research methods of evaluation. Simply said, quantitative research involves the collection of numerical data, while qualitative research concerns the collection or analysis of words, pictures and artefacts (Field, 2009; Mertens, 2015). Examples of quantitative methods are questionnaires, mathematical modelling and laboratory experiments, while examples of qualitative methods are case studies, participant observations, document analysis and action research. Quantitative and qualitative research methods may also be combined (Campbell & Fiske, 1959; Jick, 1979), so that the limitations of each method can be compensated by the counter-balancing strengths of another method (Jick, 1979).
- Program/project evaluation and practice evaluation. Alexopoulos et al. (2012) writes that program or project evaluation concerns assessing the effectiveness, efficiency and sustainability of programs. It might include the evaluation of an IT-system as such, where only the evaluator and the IT-system are involved (Cronholm & Goldkuhl, 2003). Practice evaluation, on the other hand, concerns judging the improvement of work carried out by individuals and groups, and on how people act in certain situations (Alexopoulos et al., 2012). It is more related to evaluating an IT-system in use, where a user interacts with an IT-system. This is a richer but also more complex type of evaluation (Cronholm & Goldkuhl, 2003).
- Goal-based, goal-free and criteria-based evaluation. Cronholm and Goldkuhl (2003) explain that goal-based evaluation refers to using particular goals from the organizational context as a yardstick in the evaluation of an IT-system, which is a type of formal-rational approach. Goal-free evaluation is evaluation without a specification of such explicit goals, and can be seen as an interpretative approach. Criteria-based evaluation means that specified general criteria are examined and assessed, which is an action-oriented approach where the interaction between users and IT-systems and/or the IT-systems interface plays an important role.
- Formative/project and summative/tool evaluation. Formative evaluation refers to the evaluation by offering systematic feedback to the designers and implementers (Cronholm & Goldkuhl, 2003; Scriven, 1967; Walsham, 1993). It aims to understand and evaluate program acceptance, whether impact and objectives match, and whether efforts are justified (Alexopoulos et al., 2012). Formative evaluation is conducted at a relatively early stage (Rogers, 2010). Project evaluation is a type of formative evaluation, and refers to criteria including project duration, costs and achievements of a project (Alexopoulos et al., 2012). In contrast, summative evaluation seeks to assess the occurrence or absence of expected internal program or project-specific impacts, as well as external and mandatory impacts (Alexopoulos et al., 2012). It examines and determines the worth of a programme after it has been completed in terms of initially specified success criteria (Cronholm & Goldkuhl, 2003; Scriven, 1967; Walsham, 1993). Summative evaluation is often conducted at a later stage than formative evaluation. Tool evaluation is a type of summative evaluation, and refers to criteria such as user friendliness of the system, transparency and user acceptance (Alexopoulos et al., 2012).

Combinations of these types of evaluation will be used in the creation of the VRE4EIC evaluation methodology. Since quantitative and qualitative research methods each have advantages and disadvantages, we combine them, so that the limitations of each method can be compensated by the counter-balancing strengths of another method (Jick, 1979). Furthermore, the VRE4EIC evaluation methodology combines program/project evaluation and practice evaluation, so that it covers both the evaluation of the IT-system itself, as well as the improvement of work carried out by individuals and groups and their interaction with the IT-system. Regarding the use of criteria, the VRE4EIC

evaluation methodology uses a so-called goal-based evaluation of IT systems in use (Cronholm & Goldkuhl, 2003), since a clearly focused evaluation is desired for VRE4EIC. Formative/project and summative/tool evaluation are combined in the VRE4EIC evaluation approach, meaning that both systematic feedback to the designers and implementers is provided (formative evaluation) and that the impact and worth of the project are examined in terms of specified success criteria.

3.3 Methods for performing evaluations

A variety of methods for performing project evaluations exist. In this section we describe a number of methods that are relevant for evaluating Virtual Research Environments in general, and for the VRE4EIC project in particular.

- Quantitative surveys. Quantitative surveys can be used to collect quantitative data, for instance, data about the percentage of respondents that agrees with a certain statement (Walliman, 2011). It allows for obtaining standardized data through structured closed questions. The answers are fixed, and respondents can only choose those answers that are provided in the survey.
- Semi-structured surveys and interviews. In semi-structured surveys and interviews, the interviewer presents a number of pre-defined topics to the respondent, yet the respondent has the chance to explore issues that he or she believes are important (Longhurst, 2010). Semi-structured surveys and interviews use a more interpretative approach instead of a very formal approach. Semi-structured data is neither raw data nor strictly typed or table-oriented (Abiteboul, 1997).
- Participant observations and usability testing. Observations can be used to collect data about events and activities (Walliman, 2011). Users of a system can be observed while interacting with the IT-system and while testing it, and one can examine their satisfaction, as well as the performance of the system. To reduce the risk of having a situation in which the measure process is unsystematic (Riley, 1963) or the situation in which the human observer "may selectively expose himself to the data, or selectively perceive them, and, worse yet, shift over time the calibration of his observation measures" (Webb, Campbell, Schwartz, & Sechrest, 1973, p. 114), an observation protocol and a semi-structured observer survey can be used. Usability testing may also concern testing particular scenarios for using the system, or for testing use cases.
- Case studies. Case studies can be used to answer questions that handle operational links rather than frequencies or incidence (Yin, 2003), and are therefore appropriate for qualitative research methods of evaluation. A case study can be defined as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin, 2003, p. 13). Case studies can be used to examine the dynamics of single settings (Eisenhardt, 1989) in their natural environment (Benbasat, Goldstein, & Mead, 1987).
- Web analytics and log data. Web analytics provide information about the number of visitors of a website and their behaviour by tracking them according to specified criteria (Alexopoulos et al., 2012). Advanced web analytics can be used to understand the relation between an individual and a web site (Phippen, Sheppard, & Furnell, 2004). Two popular approaches of web analytics are recording the transactions made by an individual (i.e. log file analysis) and recording which server individual comes from and goes to (i.e. page tagging) (Alexopoulos et al., 2012). Web analytics can be used to examine the impact of an IT-system or project, and to find out whether the target groups have been reached. When collecting log data, due account should be given to privacy aspects.
- SWOT Analysis. A SWOT analysis can be used to identify the strengths, weaknesses, opportunities and threats of a project as a whole. As mentioned by Hill and Westbrook

(1997), a SWOT analysis in itself can make use of various methods, such as document analysis, site visits and user interviews.

- Focus groups. In a focus group, a group of people meet informally to discuss a topic raised by the researcher (Longhurst, 2010). This can also be in the form of a question, and participants get the chance to respond to each other's answers. The group of participants can explore the topics they believe are interesting and important, while the facilitator keep the group on the topic (idem).

Table 1: Classification of research methods, including their advantages and disadvantages

Method	Main advantages	Main disadvantages
Quantitative surveys	Much standardization, collection of large number of responses and generalization of findings	Fixed answers (not flexible), lacks flexibility to pursue certain issues in-depth, sensitive to flaws in question interpretation
Semi-structured surveys and interviews	Provide context information, allow for going in-depth	Sensitive to flaws in question interpretation, resource-intensive for researcher
Participant observations and usability testing	Provide context information, allow for going in-depth, observe people in their 'own' environment while working with a system	Limited steering from the side of the researcher, resource-intensive for researcher, standardization is more difficult
Case studies	Provide context information, allow for going in-depth, flexible, observe people in their 'own' environment	Limited steering from the side of the researcher, resource-intensive for researcher, standardization is more difficult
Web analytics and log data	Detailed data of actual behaviour, limited influence of the respondent by the researcher	No steering from the side of the researcher, lack of context-information and explanation by respondent
SWOT Analysis	Easy and quick, useful for brainstorming and idea generation	Non-standardized, lacks flexibility to pursue certain issues in-depth
Focus groups	Provide context information, allow for going in-depth, flexible	Resource-intensive for researcher

Table 1 provides a short summary to classify the above-mentioned methods and to show their main advantages and disadvantages. The table shows that some methods (e.g. case studies, interviews and focus groups) are more appropriate for collecting in-depth context information, while other methods (e.g. quantitative surveys and log data) are more appropriate for collecting data and responses from large groups of people. Each method has its own advantages and disadvantages. In the VRE4EIC project, different research methods are combined, so that the limitations of each method can be compensated by the counter-balancing strengths of another method (Jick, 1979). The exact evaluation methods used in the VRE4EIC project are explained in section 4.3.

3.4 Evaluation theories

Several existing theories are useful in the context of our VRE4EIC evaluation efforts.

Information Systems (IS) evaluation

For decades, the acceptance and use of Information Technology (IT) has been of substantial importance for Information Systems (IS) research and practice (Lancelot Miltgen, Popovič, & Oliveira,

2013). Much research has already been conducted on the evaluation of Information Systems, for instance, by Hirschheim and Smithson (1988a), Smithson and Hirschheim (1998), Irani (2002) , Irani and Love (2008), Willcocks (1994, 2013) and Farbey, Land, and Targett (1999).

Irani and Love (2008) write that IS evaluation is challenging in nature because of the human and organizational involvement, high risks, erratic funding timing, various portfolio benefits and considerable intangible costs. In their opinion, both hard technological and soft social factors need to be evaluated. Each type of IS requires a different evaluation methodology that considers the specific characteristics, objectives, expected benefits and expected costs.

Smithson and Hirschheim (1998) classify the literature on IS evaluation based on two dimensions, namely (1) the origin of the various approaches which have been imported into the IS field from other disciplines, such as software quality, financial approaches and user satisfaction, and (2) the underlying assumptions of the evaluation approaches, ranging between 'hard' objective/rational engineering assumptions and 'soft' subjective/political social science assumptions. Based on these dimensions, Smithson and Hirschheim (1998) identify three types of evaluation:

- Efficiency, which is concerned with objective/rational evaluation assumptions, making use of relatively detailed specifications or benchmarks.
- Effectiveness, which is characterized by less determined evaluation criteria, and includes studies in the area of utilisation, cost-benefit analysis, comparison with objectives and user satisfaction.
- Understanding, which regards evaluation within a particular organizational context, to examine how people evaluate artefacts and situations.

Willcocks (1994) concludes that "there is a strong correlation between control and measurement of IS and higher effectiveness with IS, however measured" (p. 4-5). While Willcocks (1994) refers to the importance of conducting IS evaluation during all stages of the IS life cycle, in his book he emphasizes the front end of IS evaluation, including strategy, investment appraisal, prioritizing and feasibility evaluation.

Irani (2002) draws on the literature concerning the evaluation of IS and Information Technology (IT) and identifies and tests various conjectures. Based on this study, Irani (2002) concludes upon various lessons learned, including: 1) specific evaluation criteria can make the evaluation process more manageable, and 2) concept justification of IS to operational stakeholders and increased commitment to project success are related.

Farbey et al. (1999) state that IT evaluation includes one or multiple processes of finding and assessing impacts of IT projects and the programme or strategy that they are part of. They describe that such evaluation can take place at different points in time or continuously, and that it includes both quantitative and qualitative approaches. The characteristics of the IT project determine the suitability of a particular evaluation strategy. Farbey et al. (1999) also emphasize the importance of involving external stakeholders in the evaluation process.

Technology Acceptance Models

Various models have been developed to understand which factors influence a person's decision to use a new technology. The key models are as follows.

- **Technology Acceptance Model (TAM)**. Considerable theoretical and empirical support has been found for the Technology Acceptance Model (TAM) (F.D. Davis, 1989; F.D Davis, Bagozzi, & Warshaw, 1989). TAM explains a substantial proportion of the variance in usage intentions and behaviour (Venkatesh & Davis, 2000). According to TAM, a person's intention to use a system is determined by perceived usefulness and by perceived ease-of-use (Venkatesh & Bala, 2008; Venkatesh & Davis, 2000). Perceived usefulness refers to the degree to which people believe that using the system will enhance their job performance. Perceived ease of use is the degree to which people believe that the use of a system will be

free of effort. The perceived ease-of-use also influences the perceived usefulness, since the easier the system is to use, the more useful it can be. According to TAM, external variables that influence the intention to use a system, such as system characteristics, development processes and training, are mediated by perceived usefulness and perceived ease of use.

- **Unified Theory of Acceptance and Use of Technology (UTAUT)**. Based on a review of the Technology Acceptance Model and other theoretical models and literature concerning the acceptance of technology and the predictors of technology acceptance, Venkatesh et al. (2003) proposed the Unified Theory of Acceptance and Use of Technology (UTAUT). This theory can be seen as a unified model for the investigation of the acceptance and use of technology. It is a well-established theory, which has been tested considerably in many different contexts. The key idea of the UTAUT is that a number of factors lead to the behavioral intention to accept and use a system or technology, while this behavioral intention in combination with facilitating conditions leads to the actual use of this system or technology (Sykes, Venkatesh, & Gosain, 2009). In the UTAUT model four constructs directly predict the behavioral intention to use Information Technologies (IT), namely Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI) and Facilitating Conditions (FC). Additionally, four key moderators are defined, including Gender (G), Age (A), Experience (E) and Voluntariness of Use (VU).
- **Integrated model combining UTAUT and the Expectation Confirmation Theory of Information Systems (IS) continuance (ECT).** This model of Venkatesh, Thong, Chan, Hu, and Brown (2011) integrates the Unified Theory of Acceptance and Use of Technology (UTAUT) and the two-stage Expectation Confirmation Theory of Information Systems (IS) continuance. The model is focused on understanding the acceptance and use of a certain technology during the course of its usage. It theorizes that pre-usage variables, usage variables, and variables concerning the intention to continue using the technology influence the acceptance and use of technology. Pre-usable variables are pre-usage beliefs (including perceived usefulness, effort expectancy, social influence, facilitating conditions and trust) and pre-usage attitude. Usage variables include disconfirmation (of perceived usefulness, effort expectancy, social influence, facilitating conditions and trust) and satisfaction. Post-usage variables are post-usage beliefs (including perceived usefulness, effort expectancy, social influence, facilitating conditions and trust) and post-usage attitude. The integrated model of UTAUT and ECT theorizes that these variables influence the intention to use a system and to continue using it.

Information Systems Success Models

The Information Systems Success Model (DeLone & McLean, 1992, 2002, 2003) synthesizes previous research involving IS success and provides a comprehensive view on various interpretations of success. According to this model, a system can be evaluated in terms of technical level, semantic level and information level. “In the D&M IS Success Model, “systems quality” measures technical success; “information quality” measures semantic success; and “use, user satisfaction, individual impacts,” and “organizational impacts” measure effectiveness success.” (DeLone & McLean, 2003, p. 10). The IS Success Model theorizes that information, system, and service quality characteristics affect the subsequent ‘actual use’ and ‘user satisfaction’, which then influence the ‘individual impact’ and the ‘organizational impact’ of the information system. Building on the IS Success Model of DeLone & McLean, Seddon (1997) proposed to further specify the ‘actual IS use’ by introducing four new variables: expectations, consequences, perceived usefulness and net benefits to society. As mentioned by Alexopoulos et al. (2012), we may conclude that IS evaluation should adopt a layered approach based on the interrelated IS success measures and their relations.

e-Services evaluation

Parasuraman, Zeithaml, and Berry (1988) present a model entitled 'SERVQUAL' to determine service quality perception of customers of service and retailing companies. The model consists of 22 items in 5 categories: tangibles (e.g. physical facilities, equipment and personnel), reliability (the ability to deliver the promised service accurately), responsiveness (the willingness to help customers and provide prompt service), assurance (e.g. the ability to assure trust and confidence) and empathy (offering personalized attention to customers).

Zeithaml, Parasuraman, and Malhotra (2000, 2002a) built on this model and developed e-SERVQUAL to assess levels of e-service quality. The e-SERVQUAL model comprises seven dimensions. The first four – efficiency, reliability, fulfilment and privacy - are at the core of the model. These four dimensions are used to assess the customers' perceptions of routine service quality offered by online retailers (the core service scale). The last three dimensions – responsiveness, compensation and contact – become significant only when customers are using service quality outside routine processes, for instance when they run into problems (the recovery service scale).

Zeithaml, Parasuraman, and Malhotra (2002b) posit that "electronic service quality is not uni-dimensional but instead is multifaceted, including dimensions such as ease of use, privacy/confidentiality, reliability, and site design" (p. 371). Moreover, they state that the core service scale differs from the recovery service scale, and that personal service can mostly be found in the category of recovery service rather than core service. In addition, the authors write that electronic service quality influences satisfaction, the intention to purchase a good, and the purchase itself. Finally, they conclude that perceptions of electronic service quality are affected by technology readiness.

Value measurement models

Drawing from the technology acceptance model, the IS success model and e-services, Loukis, Pazalos, and Salagara (2012) propose an approach for transforming user evaluation data into useful business analytics. Their e-service value model contains three layers, namely:

- efficiency measures, determining the quality of basic resources and capabilities provide by the e-service to the user (e.g. information quality, service quality and technical performance);
- usage and effectiveness measures, assessing the level of e-service use and its results (e.g. the extent to which user objectives are attained, users' fun and enjoyment and their satisfaction);
- users' future behaviour measures (e.g. the intention of users to use the e-service again in the future or the measure of whether they would recommend the e-service to friends and colleagues).

These layers are expected to provide insight into the users' ratings to identify strengths and weaknesses of e-services. Moreover, these measures impact value generation mechanisms and provide information on improvement priorities.

Although the above-mentioned theories provide useful variables for our VRE4EIC evaluation methodology, these variables and theories are relatively generic and abstract. They have not been created specifically for the evaluation of VREs, and we therefore need to specify the variables from the theories and literature in the context of our project. The specific evaluation criteria will be described in section 4.2.

3.5 Key findings from the literature

Drawing from the described literature and interpreting it in the context of the VRE4EIC project, we conclude that:

- Both objective/rational engineering assumptions and 'soft' subjective/political social science assumptions can be included in IS evaluation (Hirschheim & Smithson, 1988b; Smithson & Hirschheim, 1998).
- Specific evaluation criteria should be developed (Irani, 2002).
- It is important to justify concepts of IS to operational stakeholders, since this may lead to increased commitment to project success are related (Irani, 2002).
- IT evaluation can take place at different points in time or continuously (Farbey et al., 1999), and it is important to evaluate at different development stages (Willcocks, 1994).
- IT evaluation can include both quantitative and qualitative approaches (Farbey et al., 1999).
- The characteristics, objectives, expected benefits and expected costs of the IS or the IT project determine the suitability of a particular evaluation strategy (Farbey et al., 1999; Irani & Love, 2008).
- The Technology Acceptance Model (TAM) (e.g., F.D. Davis, 1989; F.D Davis et al., 1989), the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003), and the Integrated model combining UTAUT and ECT (Venkatesh et al., 2011) can be used to understand which factors influence a person's decision to use a new technology.
- The Information Systems Success Model (DeLone & McLean, 1992, 2002, 2003) offers measures of success, including system quality, information quality, use, user satisfaction, individual impacts and organizational impacts.
- The SERVQUAL model and its associated service quality dimensions (Parasuraman et al., 1988; Zeithaml et al., 2000, 2002a) can be used to evaluate IS service quality.
- Specific individual characteristics and skills should be taken into account in measuring the intention to use a new technology, in measuring success and in measuring electronic service quality.

These key findings from the literature will be used in the development of the VRE4EIC evaluation methodology in the following chapters.

4 VRE4EIC evaluation methodology

As mentioned by Cronholm and Goldkuhl (2003), setting up a methodology for Information Technology (IT) evaluation requires determining *how* the evaluator should act, and *what* should be evaluated. In this section we describe these aspects for the VRE4EIC evaluation methodology. In addition, we describe *when* which types of evaluation will take place.

4.1 Conceptual design: VRE4EIC evaluation methodology

The VRE4EIC evaluation methodology combines the different types of evaluation mentioned by Alexopoulos et al. (2012) (see section 2.1), namely obtaining insight in attitudes and project impacts, identifying strengths, weakness and opportunities of the project, evaluating the effects and efficiency of the project, making decisions about whether a certain aspect of the project should be promoted, implemented, developed or used, and supporting the development or implementation of project components to an external stakeholder. These types of evaluation are combined so that the outcomes of the VRE4EIC project are evaluated from a variety of perspectives. The different types of perspectives make it possible to evaluate both the project and the project products.

Since quantitative and qualitative research methods each have advantages and disadvantages, we combine them, so that the limitations of each method can be compensated by the counter-balancing strengths of another method (Jick, 1979). Furthermore, the VRE4EIC evaluation methodology combines program/project evaluation and practice evaluation, so that it covers both the evaluation of the IT-system itself, as well as the improvement of work carried out by individuals and groups and their interaction with the IT-system.

Regarding the use of criteria, the VRE4EIC evaluation methodology uses a so-called goal-based evaluation of IT systems in use (Cronholm & Goldkuhl, 2003). This type of evaluation has been selected, since a clearly focused evaluation is desired for VRE4EIC. Goal-based evaluation of IT systems in use aims to derive knowledge about whether the IT-system has attained its desired objectives, which positive and negative effects the IT-system has and what its contributions are (Cronholm & Goldkuhl, 2003). Data sources are derived from the "IT-system, goal descriptions, requirement specifications, descriptions of the IT-system, the interaction between users and the IT-system, the users perceptions of the IT-system" (Cronholm & Goldkuhl, 2003, p. 72).

Formative/project and summative/tool evaluation are combined in the VRE4EIC evaluation approach, meaning that both systematic feedback to the designers and implementers is provided (formative evaluation) and that the impact and worth of the project are examined in terms of specified success criteria.

4.2 Evaluation criteria: variables, measurements and methods

4.2.1 Evaluation against project objectives

Below we describe the variables that we will evaluate for the architecture, use cases and prototype, and for the VRE4EIC project as a whole (see Figure 2). The variables have been derived from the objectives of the project and from the literature overview provided in chapter 2. Each of the variables shown in Figure 2 is thereafter described in Table 2, and is then related to the evaluation measurements, the cumulative evaluation targets in each project year and the used evaluation methods. Table 2 includes targets and measures related to different aspects of the VRE4EIC outcomes, including the performance of the system and the building blocks, its interoperability and scalability, and the quality of the system.

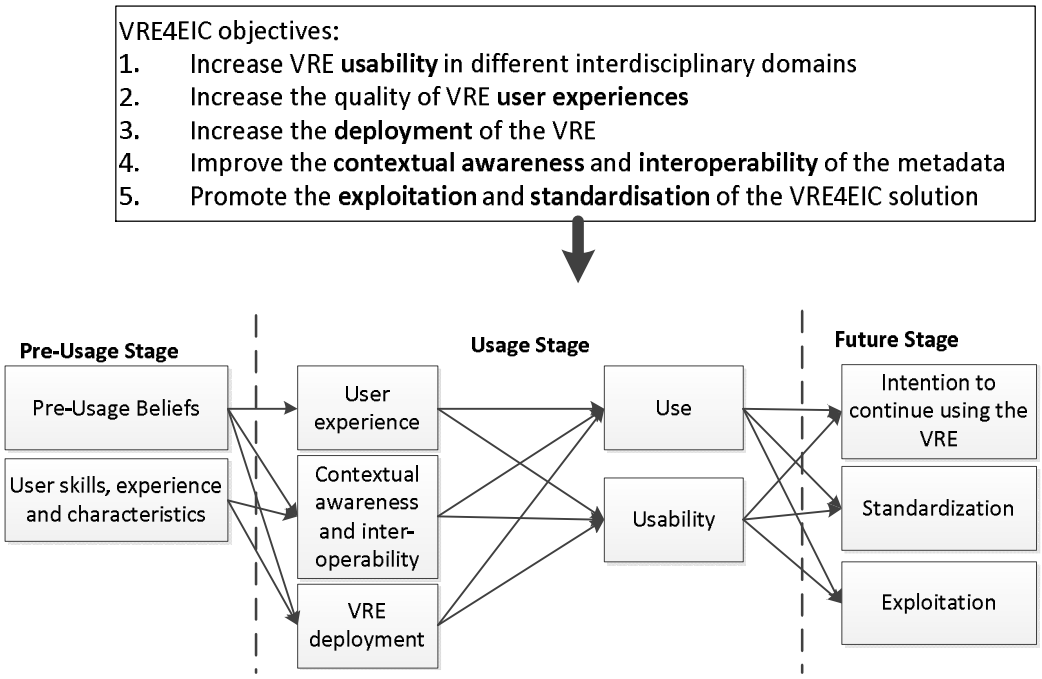

Figure 2: VRE4EIC project evaluation criteria

The measurements described in Table 2 are derived from the measurements described in Tables 4-8 of the Description of Work.

Table 2: VRE4EIC project evaluation objectives, measurements, targets and methods

VRE4EIC project objectives	Measurements	Targets per project Year (Y) (cumulative)			Evaluation method
		Y1	Y2	Y3	
1. Usability	Number of users of the VRE4EIC building blocks[1]	-	25	100	Web analytics & log data
	Percentage of recurring users of the VRE4EIC building blocks[1]	-	5%	10%	
	Number of research collaboration groups on the e-VRE[1]	-	5	10	
	Number of domains involved in the e-VRE	-	5	10	
2. Use	Percentage of end-users involved in the evaluations finding it easy to use the e-VRE, EPOS VRE and ENVRI+-VRE (e.g. for sharing research data and information, searching research data and information, processing and using research data information, collaborating with other researchers, composing workflows and for tracking data publications)[1]	-	20%	40%	User questionnaires, participant observations, usability tests, group discussions
	Number of VRE prototypes developed[1]	-	1	3	Web analytics & log data
	Number of high quality software services offered to developers through the prototype(s)[1]	-	3	15	
	Number of datasets (with related publications) available to users through the prototypes[1]	-	500	10,000	

[1] Since the VRE4EIC building blocks and prototypes will be developed in the second and third year of the project, there are no targets for this measurement in year 1.

	Number of researchers targeted as potential end-users	4,000	25,000	70,000	
2. User experience	Number of VRE4EIC-related training materials developed (e.g. MOOC videos)	5	10	15	Analytics
	Number of languages in which training materials are available to users	1	3	5	
	Number of users (with different skills) accessing the training materials	2,000	4,000	7,000	
	Percentage of developers convinced that the VRE4EIC building blocks enable researchers to work more effectively[1]	-	20%	75%	User questionnaires, participant observations, usability tests, group discussions, semi-structured interviews, log data
	Percentage of developers convinced that the VRE4EIC building blocks enable researchers to work more efficiently[1]	-	20%	75%	
	Percentage of developers convinced that the VRE4EIC building blocks support higher productivity of researchers[1]	-	20%	75%	
	Percentage of end-users satisfied with the VRE prototype (including the VRE services and the VRE learning environment) [1]	-	25%	50%	
	Percentage of end-users satisfied with security and privacy policies of the VRE prototype (describing conditions under which data can be shared and used)[1]	-	25%	50%	
3. VRE deployment	Number of building blocks and workflows from existing VREs, infrastructures and projects reused by VRE4EIC[1]	-	5	15	Analytics
	Number of use cases that the VRE building blocks support[1]	-	7	25	Participant observations, usability tests
	Strategies for handling security, privacy and trust issues developed	Yes	Yes	Yes	SWOT-analysis, analytics
4. Contextual awareness and inter-operability	Number of metadata and architecture standards supported by the VRE4EIC architecture	1	3	5	SWOT-analysis, analytics
	Number of metadata mappings conducted	1	3	5	
5. Exploitation and standardi-zation	Exploitation plan for the VRE architecture and the canonical prototype developed	-	Yes	Yes	(Web) Analytics, SWOT-analysis
	Number of conference papers about the project	1	5	10	
	Number of (open access) journal papers about the project	-	3	10	
	Total number of scientific publications by the project partners	2	10	25	
	Number of commercial associate partners (SMEs) involved in the project	2	4	4	

	Number of services disseminated by the competence centre (open source, for VRE developers)[1]	-	5	15	
	Number of countries where the VRE4EIC building blocks are available to users and developers[1]	-	15	28	
	Number of languages that the e-VRE is available in[1]	-	2	5	
	Number of standardisation group meetings (e.g. W3C, RDA)	2	5	12	
	Number of VRE-related projects (and/or their communities) with which VRE4EIC project partners collaborate	2	20	53	
	Number of press-releases	2	4	6	
	Number of project newsletters	2	4	6	
	Number of public websites developed	1	1	1	
	Number of project stands developed	1	1	1	
	Number of posters developed	-	1	2	
	Number of brochures disseminated	300	600	1,000	
	Number of social media used (e.g. Twitter accounts)	2	3	3	
	Number of workshops organized	5	10	15	
	Number of interviews for analysing user requirements	10	10	10	
	Number of other interviews (excluding those for user requirements)	-	3	5	
	Number of project presentations	5	12	20	
	Number of survey responses of potential end-users	20	50	100	
	Number of potential end-user training sessions	2	10	20	
	Number of lab experiments	-	-	2	
	Number of (university) lectures	1	5	10	
	Number of meetings with the commercial associate partners	1	2	3	

4.2.2 Evaluation of Impact

The quantitative indicators designed to evaluate the project results against the project objectives also serve to evaluate impact since they cover the criteria described for measuring impact in the Description of Work (Section 2.1 Table 4).

4.2.3 Key Performance Indicators

It should be noted that (an appropriate subset of) these quantitative indicators may be used as KPIs (Key Performance Indicators) for the project.

4.3 Evaluation procedure

As shown in Table 2, both internal and external evaluation methods are used in the VRE4EIC evaluation methodology.

4.3.1 Internal evaluation

The internal evaluation includes the following methods:

- **Web Analytics**. Web Analytics will be used to obtain information about the number of users of the e-VRE prototype and other objective numbers of VRE use and provision.
- **SWOT-analysis**. The VRE4EIC project partners will conduct a SWOT analysis. The SWOT analysis is used to identify the strengths, weaknesses, opportunities and threats of the project, and includes document analysis, site visits and an assessment of the project by the project partners. Figure 3 will be used for the SWOT-analysis.

Strengths: • … • … • …	**Weaknesses:** • … • … • …
Opportunities: • … • … • …	**Threats:** • … • … • …

Figure 3: SWOT analysis template.

4.3.2 External evaluation

The external evaluation includes the following methods:

- **User questionnaires**. The exact content of the online user questionnaire depends on what the e-VRE and the prototypes will look like, which on its turn depends on the elicited requirements. At this stage of the project, it is not possible to predict exactly what the e-VRE and the prototypes will look like. However, even though we cannot create each of the questions of the User Questionnaire yet, we expect the User Questionnaire to contain at least the following components:
  - Introduction
  - Questions related to the e-VRE building blocks and functions
  - Questions related to the VRE prototypes
  - The evaluation criteria as described in Figure 2 (e.g. pre-usage beliefs, usability, use), related to each of the five VRE4EIC project objectives
  - Questions related to users' satisfaction
  - Questions related to user demographics
  - Suggestions and comments

  Responses to questionnaires will preferably be collected electronically.
- **Semi-structured surveys and interviews**. Depending on the need for feedback, in-depth interviews will be conducted with representatives of each of the target groups (see section 3.3 for a target group overview). The topics of the semi-structured survey and interviews depends on the need for feedback, and could include:

  - Questions related to the e-VRE building blocks and functions
  - Questions related to the VRE prototypes
  - Questions related to users' satisfaction
  - Interviewee background questions
- **Participant observations, usability tests and log data**. Users of the e-VRE and of the prototypes will be asked to conduct scenarios (based on the use cases) with the developed system and services. Scenarios can defined as narrative descriptions of interactions between users and proposed systems (Potts, 1995). More specifically, "scenarios highlight goals suggested by the appearance and behaviour of the system, what people try to do with the system, what procedures are adopted, not adopted, carried out successfully or erroneously, and what interpretations people make of what happens to them" (Carroll, 1999, p. 2). When users conduct scenarios, which will be a form of usability tests, they will be observed, and log data (taking due account of privacy aspects) will be collected. This type of evaluation of VRE4EIC will also be integrated in education and students will be involved in the evaluation procedure.
- **Group discussions**. Groups discussions will be organized to discuss the e-VRE and the prototypes. Groups discussion will take place in various forms, including:
  - Group discussion with users. Preferably, these discussions are organized directly after users have conducted scenarios in which they worked with the e-VRE and the prototypes. Group discussions can also be part of workshops with users. For instance, particular aspects of the VRE4EIC building blocks and services can be discussed, and users can be asked for their feedback on these building blocks and services.
  - Group discussions with developers. These group discussions take place with developers of VREs and Research Infrastructures related to VRE4EIC, such as developers involved in the EPOS and ENVRI+ projects. Group discussions can also be part of workshops with developer. For instance, the architecture of the e-VRE can be discussed, and developers can be asked for their feedback on this architecture.
  - Group discussions with scientific VRE researchers. Academics who conduct research on VREs will be consulted for feedback on the e-VRE. For example, they can provide feedback on VRE components and VRE communities.
  - Group discussions with VRE data publishers. Providers of data for the e-VRE will be asked for feedback on the e-VRE and the prototypes.

The data collected through the above-mentioned methods will be used to improve the outcomes of the VRE4EIC project. It will feed directly into Work Package 3 concerning Architecture, VRE development, integration and scalability. The data obtained through the evaluations that can be made available as open data, will be opened (taking due account of privacy aspects).

4.4 Evaluation plan

Since the reviewed literature (e.g. models like UTAUT, TAM and IS Success models) does not provide sufficient depth for the VRE4EIC evaluation, we have set up an ambitious initial evaluation plan to better understand the needs of our target groups through in-depth sessions (see Table 3). This plan still includes a number of uncertainties that will be specified as the project progresses. The evaluation sessions are targeted to specific audiences of specialists. As described in the Description of Work, the evaluation has been divided into two phases. In phase 1 (M6-M24), evaluations are focused mainly on the collected requirements and the use cases, progressively moving more towards requirement refinement and architecture evaluation. These evaluation outcomes are input for e-VRE architecture. In the second evaluation phase (M25-M36), the focus of the evaluation activities moves towards system and prototype evaluation.

Table 3: VRE4EIC evaluation plan

VRE4EIC evaluation phase 1 (April 2016- September 2017; M6-M24)							
Event #	**Evaluation event title/topic**	**Target audience**	**Partner(s) responsible**	**Internal/ external evaluation**	**Event**	**Type of evaluation**	**Approximate date**
1	Evaluation of requirements	VRE developers	All	Internal	VRE4EIC meeting with involved partners	Group discussions, SWOT analysis	April 2016
2	Evaluation of requirements	VRE users (researchers) and data publishers	TU Delft	External	Conference on e-Democracy and Open Government (CeDEM)	Group discussions	May 2016
3	Evaluation of refined requirements	VRE users (researchers) and data publishers	TU Delft	External	Conference on E-Government (EGOV)	Group discussions	September 2016
4	Evaluation of architecture (1st version)	VRE developers	All	Internal	VRE4EIC meeting with involved partners	Group discussions, SWOT analysis	December 2016
5	Evaluation of use cases (1st version) and architecture	VRE developers	ERCIM, INGV	Semi-external	EPOS project meeting	Group discussions, SWOT analysis	January 2017
6	Evaluation of use cases (1st version) and architecture	VRE developers	ERCIM, UvA	Semi-external	ENVRI+ project meeting	Group discussions, SWOT analysis	January 2017
7	Evaluation of use cases (1st version) and architecture	VRE users (researchers)	TU Delft, euroCRIS	External	Not event-related, audience reached through consortium network	Semi-structured interviews/surveys	February 2017
8	Evaluation of use cases (1st version)	Other (students)	TU Delft	External	TU Delft	Group discussions, SWOT analysis	March 2017
9	Architecture and prototype evaluation	VRE users (researchers) and data publishers	TU Delft	External	Conference on e-Democracy and Open Government (CeDEM)	Group discussions	May 2017
10	Architecture and prototype evaluation	VRE users (researchers) and data publishers	TU Delft	External	Conference on E-Government (EGOV)	Group discussions	September 2017

VRE4EIC evaluation phase 2 (October 2017-September 2018; M25-M36)							
Event #	**Evaluation event title/topic**	**Target audience**	**Partner(s) responsible**	**Internal/ external evaluation**	**Event**	**Type of evaluation**	**Approximate date**
11	Evaluation of the use cases and prototypes	VRE developers	ERCIM, INGV	Semi-external	EPOS project meeting	Group discussions, SWOT analysis	October 2017
12	Evaluation of the use cases and prototypes	VRE developers	ERCIM, UvA	Semi-external	ENVRI+ project meeting	Group discussions, SWOT analysis	December 2017
13	Evaluation of the use cases and prototypes	VRE developers	All	Internal	VRE4EIC project meeting	Group discussions, SWOT analysis, questionnaires	February 2018
14	Evaluation of the prototypes	VRE users (researchers), data publishers	ERCIM, TU Delft	External	Not event-related	Web analytics	April 2018
15	Evaluation of the use cases and prototypes	VRE users (researchers), data publishers, VRE developers		External		Semi-structured surveys and interviews	June 2018
16	Evaluation of the use cases and prototypes	VRE users (researchers), data publishers, VRE developers		External		User questionnaires, participant observations, usability tests and log data	May 2018

5 Conclusions and next steps

The VRE4EIC project has now developed a methodology for evaluating the VRE4EIC project outcomes, including the e-VRE architecture and the prototypes with the use cases. Based on the VRE4EIC objectives, the evaluation encompasses the following stages and variables:

- Pre-usage stage: pre-usage beliefs, user skills, experience and characteristics;
- Usage stage: User experience, contextual awareness and interoperability, VRE deployment, use and usability;
- Future stage: intention to continue using the VRE, standardization and exploitation.

An evaluation plan has been created to show when which partner should conduct which activities. The next step is to use this methodology, refine it further throughout the project (e.g. to define the user questionnaire questions), and to apply it to Task 2.4: the actual evaluation of the VRE4EIC architecture and the prototypes with the use cases. The actual evaluation will start from M7 of the project, and the evaluation results will be communicated actively with the VRE4EIC partners. The evaluation results will be used to ensure that the architecture, prototypes and use cases meet the requirements of our target groups, so that uptake of the project results is maximized.

6 References

Abiteboul, S. (1997). *Querying semi-structured data*. Heidelberg: Springer Berlin.

Alexopoulos, C., Charalabidis, Y., & Loukis, E. (2012). Deliverable D6.6.1: ENGAGE evaluation framework.

Alexopoulos, C., Loukis, E., Charalabidis, Y., & Zuiderwijk, A. (2013). *An evaluation framework for traditional and advanced open public data e-infrastructures*. Paper presented at the 13th European Conference of Electronic Government, Como, Italy.

Alexopoulos, C., Zuiderwijk, A., Charalabidis, Y., Loukis, E., & Janssen, M. (2014). *Designing a second generation of open data platforms: integrating open data and social media* Paper presented at the 13th conference on Electronic Government, Dublin, Ireland.

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of Information Systems. *MIS Quarterly, 11*(3), 369-386. doi: 10.2307/248684

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105.

Carroll, J. M. (1999). *Five reasons for scenario-based design*. Paper presented at the 32nd Hawaii International Conference on System Sciences, Hawaii, U.S.A.

Cronholm, S., & Goldkuhl, G. (2003). Strategies for information systems evaluation-six generic types. *Electronic Journal of Information Systems Evaluation, 6*(2), 65-74.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science, 35*(8), 982-1003.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research, 3*(1), 60-95.

DeLone, W. H., & McLean, E. R. (2002). *Information systems success revisited.* Paper presented at the 35th Annual Hawaii International Conference on System Sciences, Hawaii.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of management information systems, 19*(4), 9-30.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532-550.

Farbey, B., Land, F., & Targett, D. (1999). Moving IS evaluation forward: learning themes and research issues. *The Journal of Strategic Information Systems, 8*(2), 189-207.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: SAGE.

Hevner, A. R., & Chatterjee, S. (2010). *Design research in Information Systems*. New York: Springer.

Hill, T., & Westbrook, R. (1997). SWOT analysis: it's time for a product recall. *Long range planning, 30*(1), 46-52.

Hirschheim, R., & Smithson, S. (1988a). *A critical analysis of information systems evaluation*. Paper presented at the Information Systems Assessment: Issues and Challenges, Amsterdam, The Netherlands.

Hirschheim, R., & Smithson, S. (1988b). A critical analysis of information systems evaluation. In N. Bjorn-Andersen & G. Davis (Eds.), *Information Systems Assessment: Issues and Challenges* (pp. 17-37).

Irani, Z. (2002). Information systems evaluation: navigating through the problem domain. *Information & Management, 40*(1), 11-24.

Irani, Z., & Love, P. E. (2008). Information systems evaluation: A crisis of understanding. *Evaluating Information Systems–Public and Private Sector", Butterworth-Heinemann, UK*.

Jick, T. D. (1979). Mixing qualitative and quantitative methods: triangulation in action. *Administrative science quarterly, 24*(4), 602-611.

Lancelot Miltgen, C., Popovič, A., & Oliveira, T. (2013). Determinants of End-User Acceptance of Biometrics: Integrating the "Big 3" of Technology Acceptance with Privacy Context. *Decision Support Systems, 56*, 103-114. doi: http://dx.doi.org/10.1016/j.dss.2013.05.010

Longhurst, R. (2010). Semi-structured interviews and focus groups. In N. Clifford, S. French & G. Valentine (Eds.), *Key methods in Geography* (2 ed., pp. 103-115). Los Angeles: SAGE.

Loukis, E., Pazalos, K., & Salagara, A. (2012). Transforming e-services evaluation data into business analytics using value models. *Electronic Commerce Research and Applications, 11*, 129–141.

McCain, D. V. (2005). *Evaluation Basics*. U.S.A.: ASTD Press.

Mertens, D. M. (2015). *Research and evaluation in education and psychology: integrating diversity with quantitative, qualitative, and mixed methods*. Los Angeles: SAGE Publications.

OECD. (1998). Best practice guidelines for evaluation. *PUMA Policy Brief No. 5, Public Management Service.* Retrieved January 26, 2016, from https://portal.publicpolicy.utoronto.ca/en/StudyReports/OECD/BestPracticeGuidelinesforEvaluation/Documents/Best%20Practice%20Guidelines%20for%20Evaluation.pdf

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). Servqual. *Journal of retailing, 64*(1), 12-40.

Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. U.S.A.: SAGE Publications.

Phippen, A., Sheppard, L., & Furnell, S. (2004). A practical evaluation of web analytics. *Internet Research, 14*(4), 284-293.

Potts, C. (1995). *Using schematic scenarios to understand user needs*. Paper presented at the 1st conference on Designing Interactive Systems: Processes, Practices, Methods, & Techniques, Ann Arbor, U.S.A.

Riley, M. W. (1963). *Sociological research: I. A case approach*. New York: Harcourt, Brace & World.

Rogers, E. M. (2010). *Diffusion of innovations*. New York: Simon and Schuster.

Scriven, M. (1967). *The methodology of evaluation*. Chicago: Rand McNally.

Seddon, P. B. (1997). A respecification and extension of the DeLone and McLean model of IS success. *Information Systems Research, 8*(3), 240-253.

Smithson, S., & Hirschheim, R. (1998). Analysing information systems evaluation: another look at an old problem. *European Journal of Information Systems, 7*(3), 158-174.

Sykes, T. A., Venkatesh, V., & Gosain, S. (2009). Model of Acceptance with Peer Support: a Social Network Perspective to Understand Employees' System Use. *MIS Quarterly, 33*(2), 371-393.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*(2), 273-315.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly, 27*(3), 425-478.

Venkatesh, V., Thong, J. Y. L., Chan, F. K. Y., Hu, P. J.-H., & Brown, S. A. (2011). Extending the two-stage information systems continuance model: incorporating UTAUT predictors and the role of context. *Information Systems Journal, 21*(6), 527–555. doi: 10.1111/j.1365-2575.2011.00373.x

Walliman, N. (2011). *Research methods: The basics*. Oxon: Routledge.

Walsham, G. (1993). *Interpreting information systems in organizations*. Chichester: Wiley.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1973). *Unobtrusive measures. Nonreactive research in the social sciences* Chicago: Rand McNally & Company.

Willcocks, L. (1994). *Information Management: The evaluation of Information Systems investments*. London: Chapman & Hall.

Willcocks, L. (2013). *Information management: the evaluation of information systems investments*: Springer.

Yin, R. K. (2003). *Case study research. Design and methods*. Thoasand Oaks: SAGE publications.

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2000). e-service Quality: Definition, Dimensions and Conceptual Model. *Working Paper*.

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2002a). An Empirical Examination of the Service Quality-Value-Loyalty Chain in an Electronic Channel. *Working Paper*.

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2002b). Service quality delivery through web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science, 30*(4), 362-375.