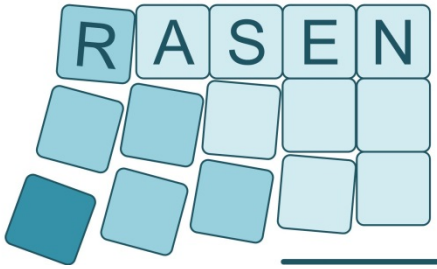

Compositional Risk Assessment and Security Testing of Networked Systems

Deliverable D2.3.1

Use case evaluation v.1

Project title:	RASEN
Project number:	316853
Call identifier:	FP7-ICT-2011-8
Objective:	ICT-8-1.4 Trustworthy ICT
Funding scheme:	STREP – Small or medium scale focused research project

Work package:	WP2
Deliverable number:	D2.3.1
Nature of deliverable:	Report
Dissemination level:	PU
Internal version number:	1.0
Contractual delivery date:	2014-09-30
Actual delivery date:	2014-09-30
Responsible partner:	Info World

Contributors

Editor(s)	Arthur Molnar (Info World)
Contributor(s)	Erlend Eilertsen (EVRY), Arthur Molnar (Info World), Frank Werner (Software AG), Albert Zenkoff (Software AG)
Quality assuror(s)	Jürgen Großmann (Fraunhofer FOKUS)

Version history

Version	Date	Description
0.1	14-07-15	First version of document TOC
0.2	14-08-01	Refined document TOC
0.3	14-08-07	Integrated first contributions
0.4	14-08-08	Integrated SAG contribution
0.5	14-09-16	Integrated EVRY contribution
0.6	14-09-18	Restructured IW contribution
0.7	14-09-22	Refined EVRY and IW contribution, added CRSTIP section. Deliverable ready for internal review.
0.8	14-09-26	Addressed internal review. Ready for final quality check
1.0	14-09-29	Final quality check

Abstract

The overall objective of RASEN WP2 is to identify use case scenarios contributed by the partners in the project, analyze them regarding their requirements and finally evaluate the case studies on software developed within the project.

The purpose of the current document is to detail the evaluation process that took place within the project's second year, to evaluate the project progress with regards to partner established criteria and to provide the roadmap towards third year evaluation activities.

Keywords

case study, requirement definition, requirement evaluation, security risk assessment, legal requirement, business software, medical information systems, financial sector

Executive Summary

The overall objective of RASEN WP2 is to provide use cases in which the R&D results of the RASEN project can be evaluated and exploited. The tasks for WP2 are closely related to WP3, 4 and 5. WP2 is split into three tasks: T2.1, T2.2 and T2.3.

- T2.1: Use case scenario definition – identification and description of use case scenarios from use case providers that are of relevance to the RASEN project.
- T2.2: Use case requirements definition – Extraction of requirements from use cases to the R&D work packages.
- T2.3: Use case evaluation – Evaluation of the R&D results of the RASEN project in light of the use case requirements.

This document builds on the previous deliverables of WP2 by providing the first evaluation of the RASEN methodology and tooling using three complex networked systems: Software AG's Command Central, EVRY's Net Bank software and Info World's Medipedia eHealth portal. To ensure a streamlined evaluation process a template was agreed upon by partners. This incorporates the requirement definitions established during the previous tasks of this work package and includes a section for detailing the results of the evaluation as well as a four-step scale to measure the use case provider's satisfaction with the work achieved so far.

The present document also introduces a new evaluation scheme that allows stakeholders to assess the maturity level of the organization in four key areas targeted by the RASEN project. First developed within previous work and extended during RASEN the CRSTIP (Compliance, Risk Assessment and Security Testing Improvement Profiling) scheme will be further used within project exploitation to highlight results that can be obtained by deploying RASEN artefacts.

As this document was prepared after the 2nd research and development phase of the project, not all requirements were eligible for evaluation; these will be evaluated after the completion of the upcoming research and development phase at the end of the project.

The results of this first evaluation showcase that many of the use case partners' requirements were already addressed by the technical work, with a positive assessment and expectations provided for most technical artefacts.

The current evaluation also shows that project objectives are well covered by the use case partners' requirements, allowing for a thorough assessment of the project work.

Table of contents

1 Introduction

WP2 consists of three tasks (cf. Figure 1) which are tightly linked among each other, resulting in the case study evaluation (Task 2.3) accomplished within this deliverable.

The first activity of WP2 consists of identifying relevant case studies originating from different industrial sectors that will be used to guide and evaluate the results of the RASEN project. The three case study providing partners develop highly-complex networked systems that are widely used and have stringent security and privacy requirements. Therefore, Task 2.1 undertakes the analysis of the partner use cases and identifies similarities and differences between existing processes in each organization.

Task 2.2 aims to extract use case requirements for the RASEN project starting from the case study scenarios that were detailed within task T2.1. Furthermore, the effort of defining a common template and its use in clearly stating identified requirements falls within the purview of the current task. The scope of this task also includes taking the first required steps regarding the evaluation of the RASEN approach by clearly linking identified requirements with RASEN objectives and success criteria.

The final task of WP2 is Task 2.3 is grouped in two evaluation rounds: the first evaluation in year 2 which is subject in this deliverable and the final evaluation in year 3.

The 1st evaluation which is reported in this deliverable assesses the RASEN methodology and technical implementation against the defined the use-case study requirements. For this, research and technology partners provide the results to the case study providers and will assist in implementing the new tools and methodologies within their processes. The 2nd evaluation will be similar to one accomplished in this deliverable and will provide results in the 3rd year of the project.

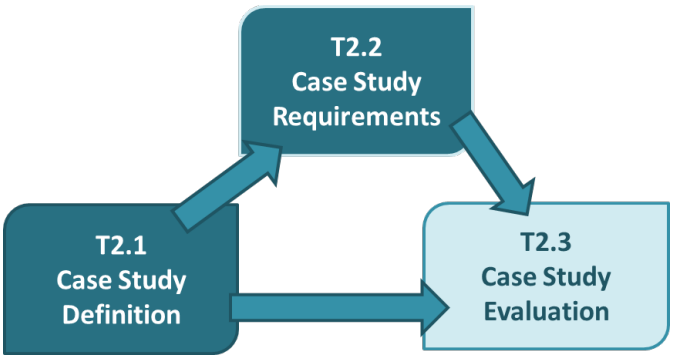

Figure 1 – Overview and dependability of tasks within WP2

2 The CRSTIP Assessment Scheme

The CRSTIP (Compliance and Risk Security Testing Improvement Profiling) assessment scheme was developed in order to provide a simple, straightforward assessment with regards to the organization's current positioning together with providing guidelines regarding what is required to further advance its standing [6]. The approach is based on previous work undertaken within the ITEA2 – Diamonds[1] project, where it was used to assess the progress that could be achieved in selected key areas of the security-testing domain. It was further refined within our project in order to serve as a liaison between our project efforts and organizations that would like to improve their standing within key areas addressed within our project. These areas describe major aspects or activities in a security testing process and are chosen in that way that they cover the most relevant innovations within RASEN.

CRSTIP can be used to assess the readiness level of an organization with regards to four key areas targeted by research in RASEN. Each area consists of four hierarchically organized levels, as shown within Figure 2.

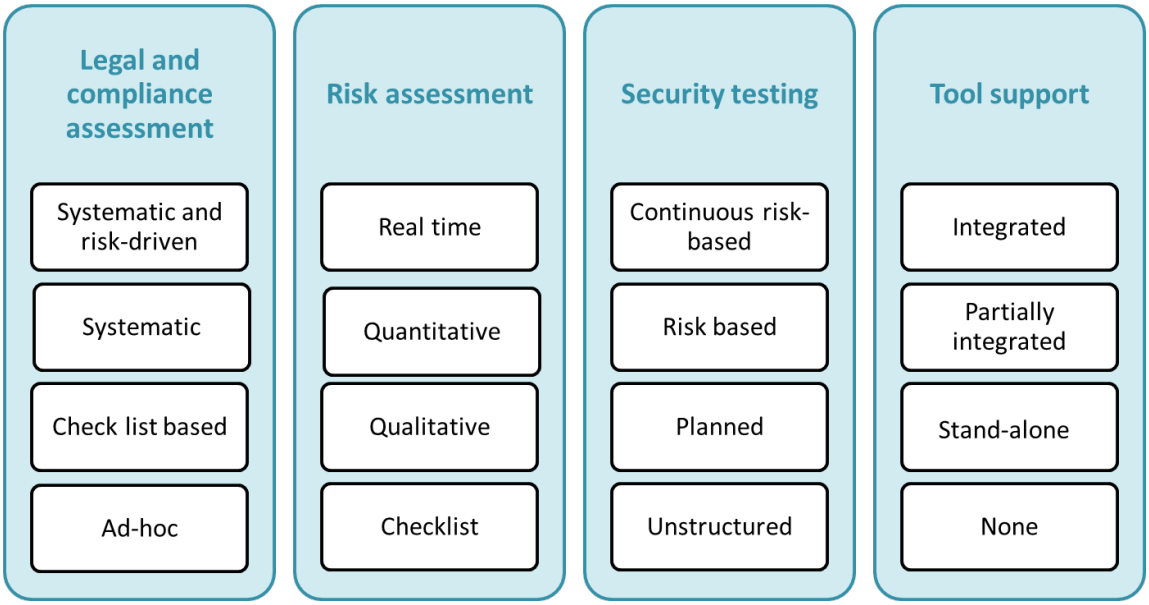

Figure 2 – CRSTIP key areas and levels

The four levels within each of the key areas provide a straightforward description in order to make it easy for stakeholders to evaluate their own organization. These levels are detailed as follows:

Legal and compliance assessment

This refers to the overall process employed with the objective of adhering to the requirements of laws, industry and organizational standards and codes, principles of good governance and accepted community and ethical standards. The overall process should support, to the extent possible, the documentation of compliance.

Key Area	Description
Ad-hoc	The compliance assessment is unstructured, does not use a defined compliance process, and compliance decisions are made primarily on an event-driven basis.
Check list based	The checklist-based compliance assessment uses a checklist to answer a set of standard questions or to tick checkboxes.
Systematic	A systematic compliance assessment follows a structured and planned approach where there is a defined process and structured documentation of compliance. Generally, the process involves the identification of compliance requirements,

[1] ITEA2 Diamonds project http://www.itea2-diamonds.org/evaluation/stip/index.html

	evaluation of the compliance issues and taking measures to ensure compliance.
Systematic and risk-driven	A systematic and risk-driven compliance assessment involves a defined process for risk-driven compliance where compliance requirements are prioritized based on their risks. This approach is supported by a systematic documentation that enables the mapping of different risks and controls to relevant compliance requirements.

Table 1 – Levels in legal and compliance assessment

Risk assessment

Risk assessment is the overall process of risk identification, risk estimation and risk evaluation. Risk identification is the process of finding, recognizing and describing risks. This involves identifying sources of risk, areas of impacts, events (including changes in circumstances), their causes and their potential consequences. Risk identification can involve historical data, theoretical analysis, informed and expert opinions, and stakeholders' needs. Risk estimation is the process of comprehending the nature of risk and determining the level of risk. This involves developing an understanding of the risk. Risk estimation provides the basis for risk evaluation and decisions on whether risks need to be treated, and on the most appropriate risk treatment strategies and methods. Risk evaluation is the process of comparing the results of risk estimation with risk criteria to determine whether the risk and/or its magnitude is acceptable or tolerable. Risk evaluation assists in the decision about risk treatment.

Key Area	Description
Checklist	Risk assessment mainly consisting in answering a sequence of questions or filling in a form.
Qualitative	Risk assessment based on qualitative risk values. Value descriptions or distinctions based on some quality or characteristic rather than on some quantity or measured value.
Quantitative	Risk assessment based on quantitative values. Values based on some quantity or number, e.g. a measurement, rather than on some quality.
Real time	Risk assessment in real-time based on underlying, computerized monitoring-infrastructure.

Table 2 – Levels in risk assessment

Security testing

Security testing is used to experimentally check software implementations with respect to their security properties and their resistance to attacks. For security testing we can distinguish functional security testing and security vulnerability testing. Functional security testing checks if the software security functions are implemented correctly and consistent with the security functional requirements. It is used to check the functionality, efficiency and availability of the specified security features of a test item. Security vulnerability testing directly addresses the identification and discovery of yet undiscovered system vulnerabilities. This kind of security testing targets the identification of design and implementation faults that lead to vulnerabilities that may harm the availability, confidentiality and integrity of the test item.

Key Area	Description
Unstructured	Unstructured security testing is performed, either by the development team or by the testing team, without planning and documentation. The tests are intended to be run only once, unless a defect is discovered. The testing is neither systematic nor planned. Defects found using this method may be harder to reproduce.
Planned	Planned security testing is performed, either by the development team or by the testing team, after a structured test plan has been elaborated. A test plan documents the scope, approach, and resources that will be used for testing.

Risk based	Security tests are planned and executed, either by the development team or by the testing team and planning of security testing is done on the basis of the security risk assessment (i.e. impact estimations or likelihood values are used to focus the security testing and optimize the resource planning).
Continuous risk-based	Continuous risk based security testing is a process of continuously monitoring and testing a system with respect to potential vulnerabilities. Security risk analysis results are still used to focus the security testing and optimize the resource planning. Any evolution of the system, of the environment of the system or of the identified threats, leads to update the security testing so that vulnerabilities could be detected throughout the whole life cycle of the software product.

Table 3 – Levels in security testing

Tool support

This key area specifies the degree of tool support that is available for the above mentioned key areas. Typically, tools work on their own data structures that are well suited to the task, which needs to be performed with or by the tool. Tool integration is the ability of tools to cooperate with other tools by exchanging data or sharing a common user interface.

Key Area	Description
None	No tool support in any of the above mentioned key areas is available.
Stand-alone	Tools are available for some of the above mentioned key areas. However, the tools are not integrated thus, they do not exchange data with other tools nor do they share the same user interface.
Partially integrated	Tools are available for some of the above mentioned key areas. Tool integration is based on point-to-point coalitions between tools. Point-to-point coalitions are often used in small and ad-hoc environments but have problems when it comes to more tools and larger environments (no scalability).
Integrated	Tools are available for nearly all of the above mentioned key areas. Tool integration is based on central integration platforms and repositories (e.g. EMF store, Model Bus, Jazz etc.) that provides a common set of data to be exchanged and respective interfaces. Tool federations better fit to larger tool environments because the existence of a common set of interfaces eases the integration of new tools. However, the definition of a common data set and common interfaces is more complex as defining bilateral point-to-point coalitions.

Table 4 – Levels in tool support

2.1 Assessment of use cases

The CRSTIP assessment scheme was first used in order to provide a baseline for the three RASEN case studies by assessing the level of each use case providing organization before having deployed any of the project artefacts. In addition, each use case provider also expressed their high-level expectations from the RASEN project by identifying targeted levels within each of the key areas. These are the levels expected to be reached once RASEN is fully implemented within the organization. Besides being employed as a high-level evaluation tool, CRSTIP will also be used within the project's dissemination and exploitation activities, which are detailed within deliverable *D6.1.2 - Periodic Standardization, Dissemination and Exploitation Plan v.2*. The following Sections detail the CRSTIP assessment of the project's three use cases.

2.2 Software AG

Figure 3 illustrates the baseline of the Software AG use case (SAG) as well as the partner's expectation once RASEN project artefacts have been deployed within the organization (SAG after RASEN). The main expected benefits of implementing RASEN are expected in the area of security

testing with the implementation of a risk-based process within the company's software development process.

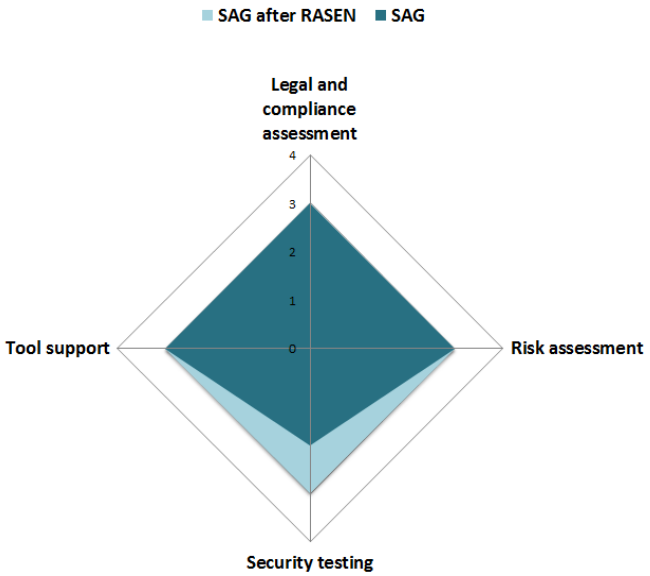

Figure 3 – CRSTIP assessment of the SAG use case

2.3 EVRY

Figure 4 illustrates the CRSTIP evaluation of the EVRY use case. As a player in the financial software market, EVRY stands to benefit greatly from deploying RASEN artefacts. EVRY expects significant process improvements by adapting the security testing methodology that will enable undertaking continuous risk-based testing. Furthermore, RASEN is expected to improve legal compliance assessment processes as well as introduce quantitative risk assessment based on the CORAS method.

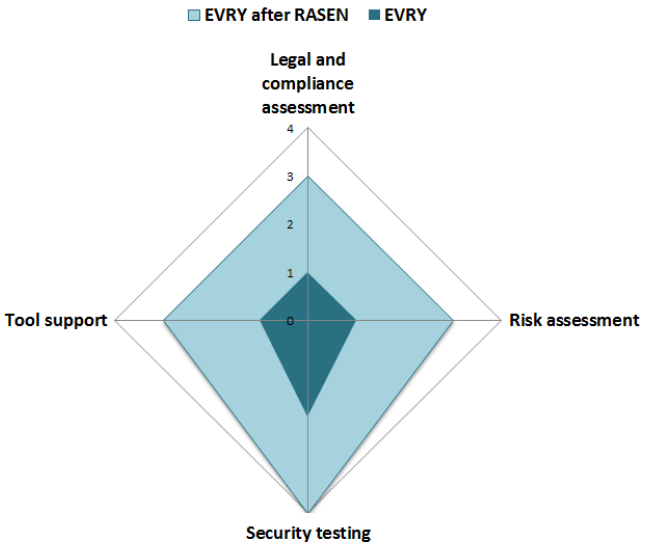

Figure 4–- CRSTIP assessment of the EVRY use case

2.4 Info World

As the development methodology of Medipedia is illustrative for most Info World systems, this initial assessment serves to provide a baseline with regards to key areas addressed by RASEN as well as

highlight the organization's expectation from the project by assessing the impact of implementing RASEN artefacts within key Info World processes. Furthermore, this evaluation will be used external to the project in dissemination and exploitation activities in order to highlight the industrial benefits of the project benefits and encourage its adoption.

Figure 5 showcases the CRSTIP evaluation for Info World [6]. The company currently employs an internal assessment of compliance that is checklist based that we believe can be improved via RASEN artefacts to a systematic approach. The current risk assessment process is qualitative as there is no structured prioritization of risk and no structured methodology. With regards to security testing, as detailed within the Info World use case description in *D2.1.1 - Use case scenarios definition* the process does not depend on any tool support and is not integrated with compliance and risk assessment activities.

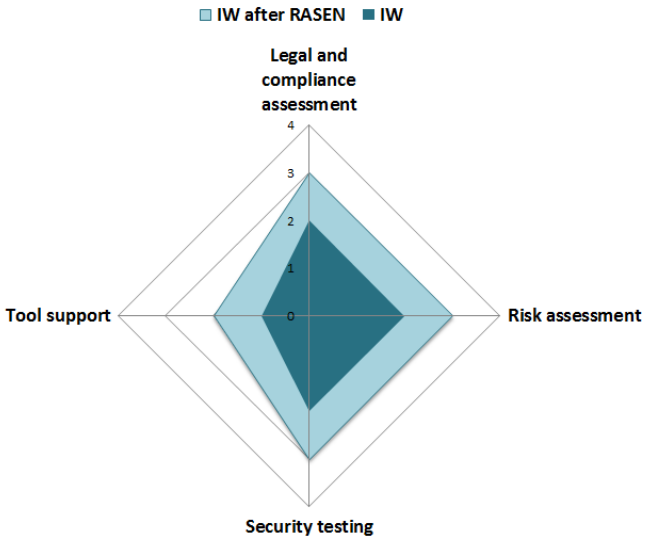

Figure 5 – CRSTIP assessment of the IW use case

3 Use Case Systems under Evaluation

This Section is dedicated to detailing the software systems that are employed in the evaluation of the RASEN methodology and tooling.

3.1 Software AG

The software constituting Software AG's use case is called Command Central [1], a tool from the *webMethods* tool suite allowing release managers, infrastructure engineers, system administrators, and operators to perform administrative tasks from a single location. Command Central assist the configuration, management, and monitoring by supporting the following tasks:

- Infrastructure engineers can see at a glance which products and fixes are installed, where they are installed, and compare installations to find discrepancies.
- System administrators can configure environments by using a single web user interface or command-line tool. Maintenance involves minimum effort and risk.
- Release managers can prepare and deploy changes to multiple servers using command-line scripting for simpler, safer lifecycle management.
- Operators can monitor server status and health, as well as start and stop servers from a single location. They can also configure alerts to be sent to them in case of unplanned outages.

Command Central is built on top of Software AG Common Platform, which uses the OSGi (Open Services Gateway Initiative) framework. Product-specific features are in the form of plug-ins.

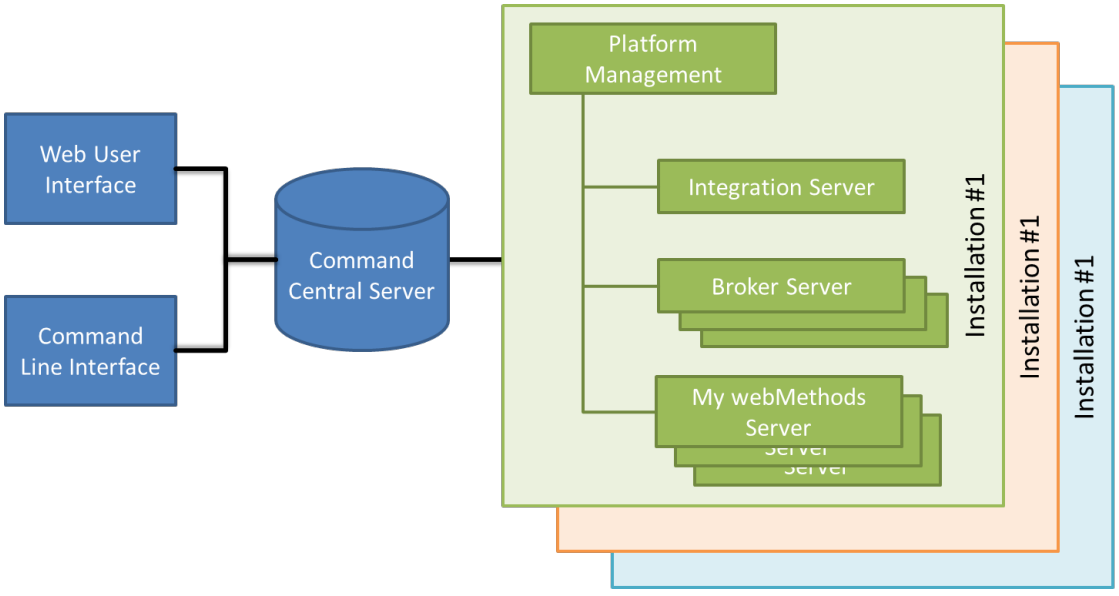

Figure 6 – Command Central Architecture

Command Central users can communicate with Command Central Server [1] using either the Graphical web user interface for administering products using the web, or the Command line interface for automating administrative operations. An architecture overview of the Command Central software is provided in Figure 6.

The Command Central Server accepts administrative commands that users submit through one of the user interfaces and directs the commands to the respective Platform Manager for subsequent execution. An installation in Command Central means one or more instances of the products that Command Central can manage. Products that Command Central manages are referred to as managed products throughout this help.

Command Central can manage one or more installations of the following products:

- Platform Manager
- Command Central
- *webMethods* Broker
- *webMethods* Integration Server
- My *webMethods* Server
- CentraSite
- Universal Messaging

Command Central provides a common location for configuring managed products installed in different environments.

webMethods Platform Manager manages Software AG products. Platform Manager enables Command Central to centrally administer the lifecycle of managed products. In a host machine, you might have multiple Software AG product installations. For each Software AG product installation, you need a separate Platform Manager to manage the installed products.

3.2 EVRY

The software systems that will be targeted in the EVRY case study are so-called Netbank systems which are provided and developed by EVRY on behalf of banks. The Netbank system enable bank customers to perform day to day bank transactions such as paying bills, moving money between accounts, viewing transaction history etc. from their PC, mobile phone, or tablet. An overview of the architecture of the EVRY Netbank system is shown in Figure 7.

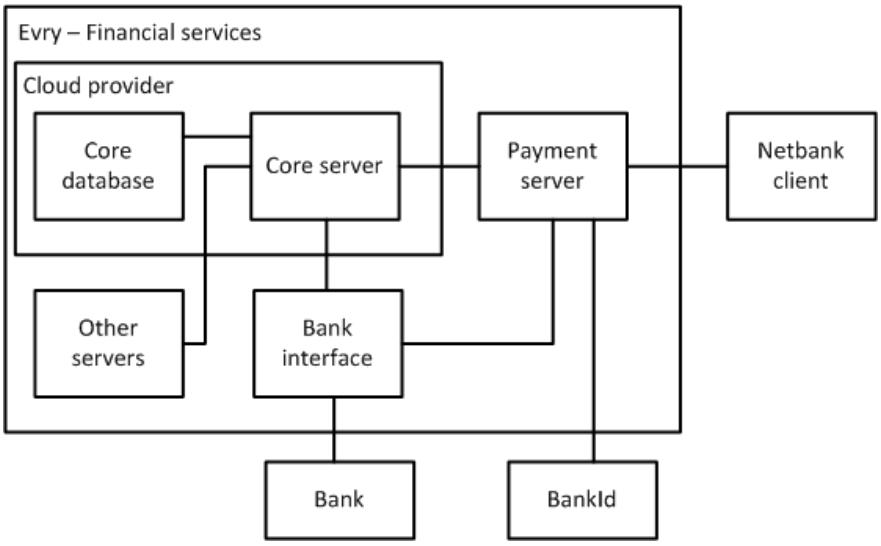

Figure 7 – Architecture of EVERY Netbank system

The Netbank application that is offered by EVRY to the banks can be customized by the banks. However, the standard functionality of the Netbank system (from the client side) are:

- Personal info – a bank customer can read and update personal information such as address, telephone number, etc.
- View account balance – the customer can see the balance for their own accounts
- Internal transfer – transferal of funds between own accounts, e.g. from salary account to savings account.

- Payment – transferal of funds to external accounts, e.g. pay a bill.
- Budget – a customer can set up a personal budget.
- Loan – a bank customer can calculate and apply for a personal loan.
- Transaction – overview of all transactions, both made within the net bank and transactions made with debit/credit cards.

Three different client solutions are provided by EVRY: mobile client for mobile phones, table client for tablets, and web-client for PC's.

3.3 Info World

The system that Info World will employ in order to evaluate the methodological and tooling results of RASEN was selected based on two major criteria:

- **Representativeness**. The chosen system must be one of the more complex systems delivered by the company, so that it is representative of Info World's product stack. This will ensure that successfully applying the RASEN process to it will be later transferrable to other systems developed within the company.
- **Requirements coverage**. The system must present challenges in all the areas addressed by the RASEN project, in order to ensure a full and complete evaluation.

Taking into account the two principles outlined above, Info World's evaluation will focus on the Medipedia system. Medipedia is a complex eHealth web portal that has over 125.000 weekly visitors and enables users to store, share and view their medical history. As the system deals with healthcare data - considered highly sensitive according to personal data protection legislation, the reliability and security of the system are of prime importance. As such, Info World's case study includes aspects of risk assessment and management, deployment and execution of security tests and legal compliance issues. Like all Info World end-user systems, Medipedia is built on the same foundation of standards-compatible software components that were outlined in the previous deliverables of this Work Package and as such we believe it is the most representative system within the company's portfolio. Medipedia provides its users a large selection of features relating to healthcare:

- Users can build, access and share their electronic health record in a safe, reliable environment without incurring any costs.
- Integrated with the nation-wide Medcenter clinical analyses laboratories, Medipedia allows users to receive analyses results directly within their Medipedia account as soon as they become available.
- Healthcare data can be shared by users with trusted physicians, family members and friends.
- Users can schedule appointments within the system.
- Users can interact with peers and healthcare specialists within the active forum system.
- The portal also provides a wealth of healthcare-related information such as descriptions for various medical conditions, analyses results, medications and more.

As shown within Figure 8, the Medipedia system employs the software components that were detailed within deliverable "*D2.1.1 - Use Case Scenarios Definition*":

- Admission, Discharge, Transfer Service (ADT) (section 4.2.2.1)
- Entity Identification Service (EIS) (section 4.2.2.2)
- Retrieve Locate and Update Service (RLUS) (section 4.2.2.3)
- Enterprise Vocabulary Service (EVS) (section 4.2.2.4)
- Security Services (section 4.2.2.5)

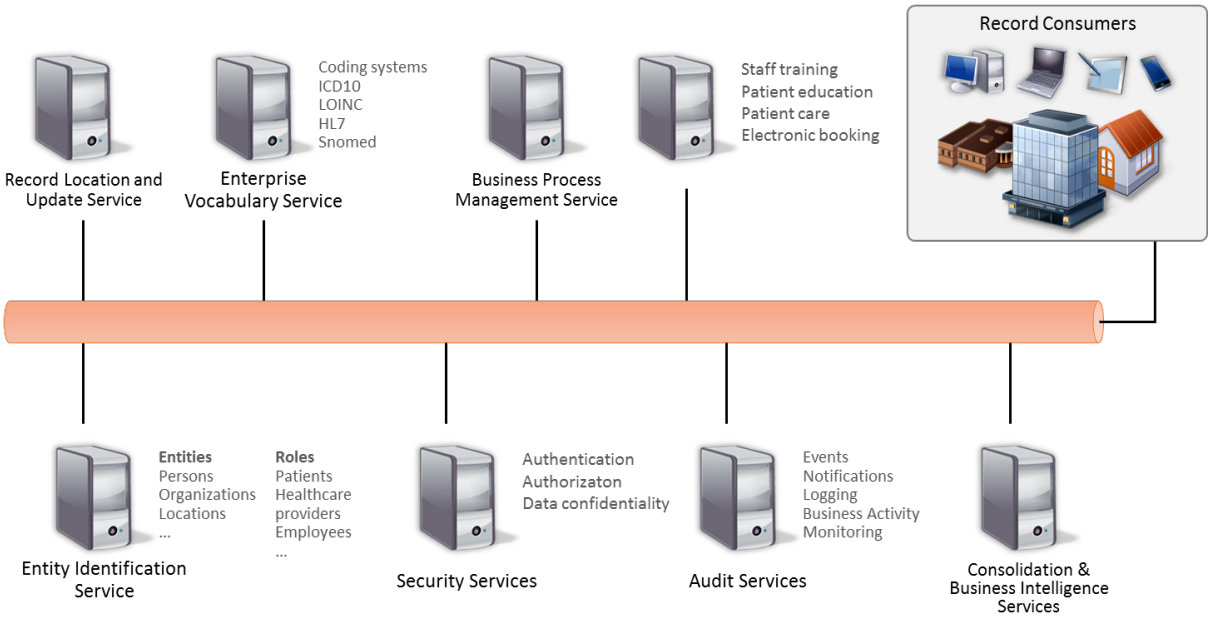

Figure 8 – Medipedia software architecture

4 Template for Requirements Evaluation

In this section we detail the template that use case partners have mutually agreed upon to use for presenting the evaluation criteria of functional requirements and results after the project's second year. Table 5 below illustrates the template used for evaluation. The right-hand side details the meaning for each of the fields.

Requirement Evaluation	
Name	Description
Code(s)	Represents the code or codes of those use case requirements that are evaluated using this template instance. These codes can be found within Section 4 of the "*D2.2.1 - Use Case Requirements Definition*" document. **E.g.** REQ-SAG-F-010, REQ-SAG-F-010
Requirement	Provides a textual description of the requirements that are evaluated using this template. **E.g.** A methodology providing automated security risk assessment.
Objective	Provides the project objectives that are linked with the present requirements evaluation. **E.g.** O5
Description	A full description of this requirement, as seen from the use case provider's perspective is provided here. **E.g.** This requirement identifies Software AG's need for an automated process of security risk assessment. The company provides large software systems that are prohibitively expensive to evaluate manually due to the amount of effort required. Therefore, in order to protect customers, Software AG is looking for new automated methods of risk assessment for these large software systems. The project is expected to deliver (define, create or select) a risk assessment methodology that can be applied in an automated way and provide repeatable and reliable assessment results. In particular, this requirement addresses the systematic approach and clearly defined methodology.
Use Case Provider Satisfaction	The importance attached to this requirement by the use case provider. Represented by an integer between 1 and 5 that denotes the importance that meeting this use case requirement has for the use case provider. A score of 1 denotes very low importance, while a score of 5 represents very high importance. **E.g.** 5
Success Criterion	The project has defined several success criteria within its Description of Work document. Additional success criteria may be defined by use case partners here. **E.g.** *SC-A1*: RASEN specifies a well-defined method to perform risk analysis of a large system in a way that is clearly understandable, systematic and repeatable.
Evaluation Criterion	This section details how the use case partner will evaluate the criteria. As specified in previous documents of this Work Package, evaluation will be undertaken in two phases, at the end of the project's second (M24 mark)

	and third (M36 mark) year. The results of the evaluation undertaken at M24 and presented within this deliverable will be employed in the last R&D Phase that will run throughout the project's final year. **E.g.** *F1*: The RASEN test procedure technique is more rigorous than the current test prioritization process.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	A rating that illustrates how well the requirement is fulfilled at this point. The rating is provided from the use case partner's perspective and detailed within the next field, "*Evaluation Phase 1 Result*". The description of these rating levels is found in Table 6. **E.g.** Good (3)
Evaluation Phase 1 Result	This section details the evaluation results at the M24 mark. **E.g.** *F1*: The current method of prioritization is unstructured and based on expert judgment. The manner of prioritization may also vary from case to case. Adapting artifact A1 would therefore provide rigor to this process.

Table 5 – Evaluation template

Name	Description
Excellent	The requirements are fully met
Good	The requirements are mostly met although there are some deficiencies detected
Fair	The requirements are partly met although there are plenty of improvements needed
Poor	Most of the requirements are not met

Table 6 – Description of evaluation ratings

5 Evaluation

This Section details the evaluation plan and process of the RASEN project.

5.1 Second Year Evaluation Process

The initial evaluation plan was created as part of Task 2.2. Figure 9 illustrates the RASEN timeline with regards to agreed-upon technical phases and milestones.

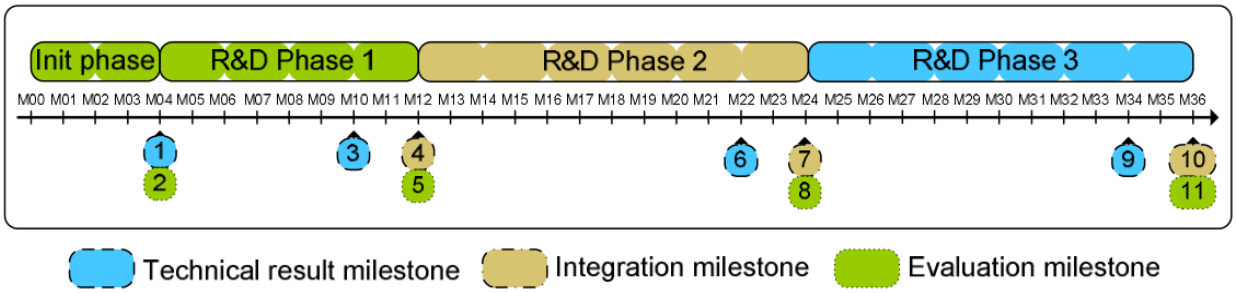

Figure 9 – Phases, timeline and milestones

The project timeline has been divided into the following phases:

- **Initialization**: This phase consisted of two major tasks, the elaboration of the technical baseline and the identification of use case scenarios. Also, at its end this phase the project contained the first evaluation milestone (Milestone 2) evaluating the proposed use case scenarios.
- **R&D[2] Phase 1**: The first technical results of the project were delivered as part of this phase, together with structured requirement definitions and an initial evaluation plan.
- **R&D Phase 2**: This phase represents the second phase of scientific and technical development within the project and the result of its activities are the target of the present document's evaluation. Evaluation Milestone 8 is where the project currently stands. During R&D Phase 2, technical and scientific partners have collaborates with the use case providers to ensure transfer of knowledge and available methodologies and tooling in order to facilitate the use case providers' evaluation of the results obtained thus far. The present deliverable is the documentation of the use case partner's initial evaluation of the suitability of the RASEN methodologies and tools together with obtained results, highlighting existing advantages and drawbacks.
- **R&D Phase 3**: The last R&D phase of the project will use the feedback obtained from the use case partners within the last technical stage of the project. The final evaluation milestone, Milestone 11 is scheduled for the end of the project at month 36 of its implementation.

[2] Research and Development

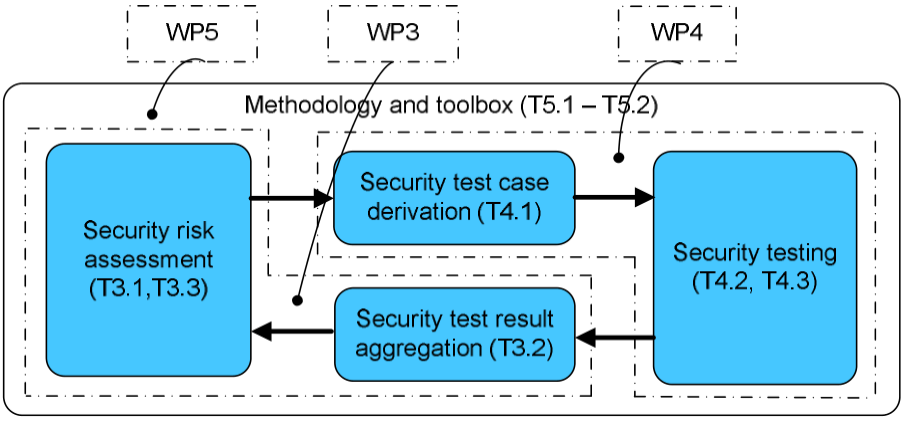

Figure 10 – Mapping work packages to technical results

The proposed evaluation plan is tightly linked with the project's work. Figure 10 shows the mapping of the project's expected main technical results to technical work packages, while Figure 11 details the relationships between the technical results across the project's proposed phases.

The tasks within WP2 represent the middle column within Figure 11. The first task, T2.1 resulted in deliverable *D2.1.1 – Use Case Scenarios Definition* that was elaborated as part of the *Initialization* phase, while task T2.2 resulted in deliverable *D2.2.1 - Use Case Requirements Definition*. As shown in Figure 11, the current task is a direct continuation of already started work.

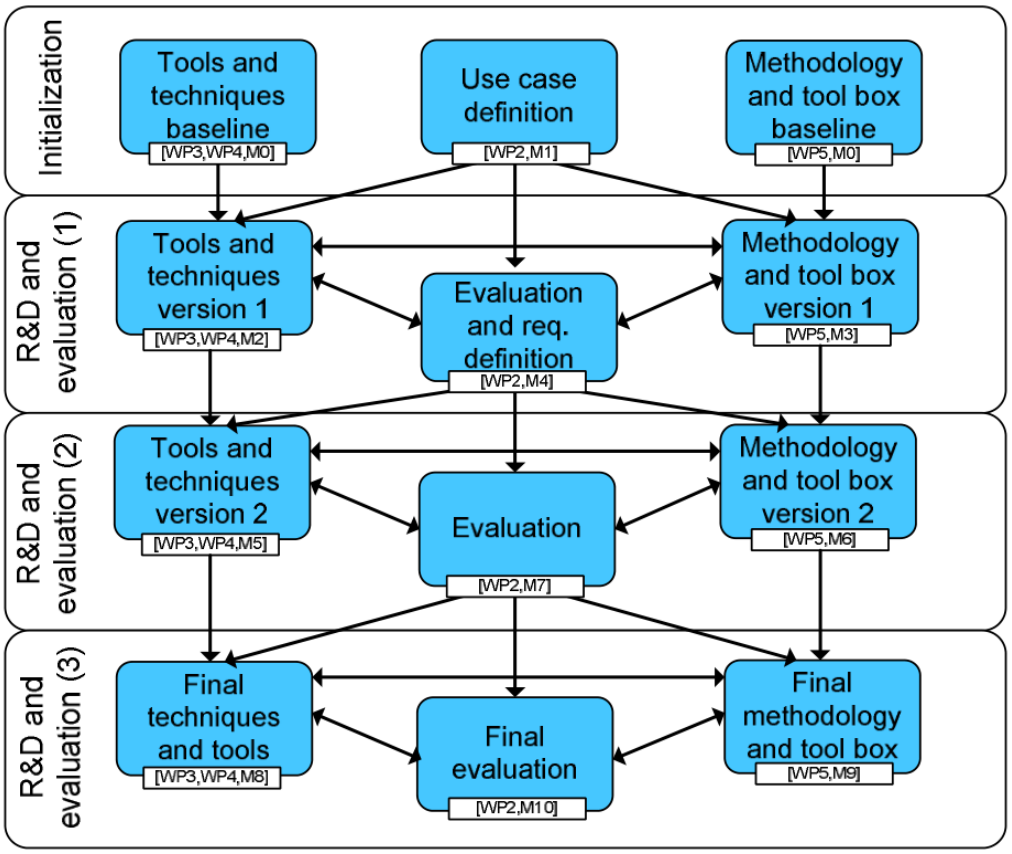

Figure 11 – Relationship of technical results over time

The project evaluation will be undertaken in two phases, concurrently with *R&D Phase 2* (Evaluation Phase 1, which has now completed) and *R&D Phase 3* (Evaluation Phase 2, planned for the upcoming project year), respectively.

The initial evaluation undertaken by use case partners is in accordance with previous plans and consisted of the following stages:

Start-up phase – Contains the first activities undertaken as part of the evaluation by the use case providers. These actions include:

- Identification of relevant tools and methodologies applicable for each use case providing partner.
- Determining the complexity of the evaluation and the length of one evaluation iteration.
- Determine how to best measure the fulfillment of stated requirements

Learning phase – This first evaluation phase represents the use case providers' contact with tools and methods developed within the RASEN project. As such, as part of the current phase use case providers will employ delivered tools with assistance from the Consortium's research and technical partners.

The present deliverable details use case partner's first contact and evaluation with delivered RASEN methodologies and tools. Their feedback will guide the last R&D Phase of the project. Research and technical partners will use the supplied document as a starting point of the last phase of research in order to address outstanding issues and to ensure the suitability and success of the final implementation.

5.2 Evaluation from Use Case Partners

5.2.1 Software AG

5.2.1.1 Evaluation Process

The evaluation of the use case scenario was organized as a sequence of evaluation steps (cf. Figure 12) which individually cover well defined parts. The different evaluation parts are described in the following:

The "*Risk Assessment*" phase was the first part where the product under investigation has been modelled in the ARIS RASEN framework. This has been achieved in a joint workshop with a software engineer as a representative from the product development (Command Central Product Development), a security expert overviewing and ensuring the compliance to security standards, and the RASEN project development team in charge of the implementation. As a result of the workshop the software under consideration has been modelled and weaknesses and risks from the CWE database have been assigned to the product and its components.

In "*Security Test Preparation*" the RASEN project representatives at Software AG conducted an assessment with the security expert to evaluate the model export which provides the artifacts for testing. These test goals define a list of components that need to be tested for the assigned weaknesses.

The "*Test Specification and Execution*" phase will consider the components along with the assigned weaknesses and execute them against a live system. This live system will be provided in terms of a virtual machine applying a black-box testing strategy. It needs to be stressed, that due to the missing implementation of the required testing interfaces – which will be implemented in the next development cycle – an evaluation of this phase was only feasible to a certain degree.

With the test results from the previous phase, the "*Security Risk Integration*" phase will receive the test results and convert them into an appropriate format, suitable for integration into the ARIS RASEN framework. In this evaluation step, all confirmed weaknesses on actual product become visible. This assessment has been conducted together with a security expert and the developers of the ARIS RASEN framework.

Eventually, the evaluation of the "*Risk Valuation & Mitigation*" step will highlight the feasibility of how confirmed risks are summarized on the level of components, but also including the calculation of the confirmed risks on the product level, exhibiting the product riskiness. This evaluation step relies on

interfaces which are only available in the last implementation phase, as such we will only consider them as part of the 2nd evaluation phase.

The following artefacts are subject to the evaluation:

- **A1**: The RASEN method & technique for automated security risk assessment
- **A2**: The RASEN method for compositional security risk assessment
- **A3**: A RASEN method & technique for automated test result aggregation
- **A4**: A RASEN method & tool-chain for automated test case execution
- **A5**: A RASEN method & tool which to provide backward traceability between tests and risks

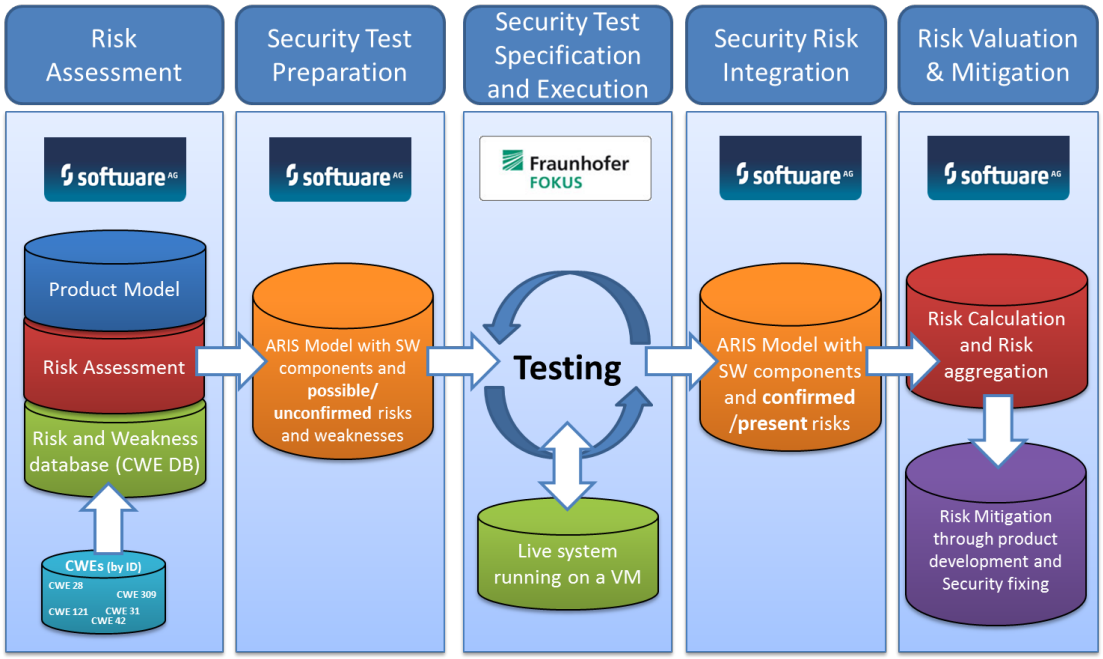

Figure 12 – RASEN tool chain in the SAG use case scenario

The results of the evaluation are discussed in the following sections.

5.2.1.2 Evaluation Result

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-010**
Requirement	A methodology providing automated security risk assessment.
Objective	O5
Description	This requirement identifies Software AG's need for an automated process of security risk assessment. The company provides large software systems that are prohibitively expensive to evaluate manually due to the amount of effort required. Therefore, in order to protect customers, Software AG is looking for new automated methods of risk assessment for these large software systems. The project is expected to deliver (define, create or select) a risk assessment methodology that can be applied in an automated way and provide repeatable and reliable assessment results. In particular, this requirement addresses the systematic approach and clearly

	defined methodology.
Use Case Provider Satisfaction	5
Success Criterion	SC-A1: RASEN specifies a well-defined method to perform risk analysis of a large system in a way that is clearly understandable, systematic and repeatable. Additionally, the following may be relevant here (not sure if we need or want them to be mentioned here): SC1.1: The approach must ensure traceability between risks and test results. SC1.2: The approach must clearly define how security results can impact the risk assessment picture. SC3.1: The approach should precisely define the rules/conditions for valid composition of security assessment and security testing results.
Evaluation Criterion	Evaluation criteria related to artifact **A1**: **Phase 1 [M24]:** *E1*: SC-A1 may be evaluated by a joint evaluation with the RASEN Project members/developers with a product architect perform a risk analysis of a chosen product with clearly defined scope and environment in accordance with the RASEN specified method. *E2*: The methodology should be clear to the person performing evaluation. **Phase 2 [M36]:** *E1*: SC-A1 may be evaluated by letting a product architect (with no prior experience with respect to this particular work) perform a risk analysis of a chosen large system with clearly defined scope and environment in accordance with the RASEN specified method. *E2*: The methodology should be clear to the person performing evaluation.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Good
Evaluation Phase 1 Results	*F1*: A product developer completed an architectural analysis of a product with the assistance of RASEN Project member. The analysis is deemed complete and satisfactory. *F2*: In the current model there is still knowledge about the modeling system needed. The final system will have additional interfaces to support unassisted modelling by providing a user interface description and wizards. Integration with other security tools (security database, test tools, etc.) is still missing and will be accomplished in the 2nd development phase.

Table 7 – Evaluation for requirement REQ-SAG-F-010

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-020**
Requirement	Tools providing automated security risk assessment.
Objective	O5
Description	The system scale does not allow for manual analysis and therefore we require additional tooling that helps us to perform automated security risk assessment of the company's products. The company provides large

	software systems that are prohibitively expensive to evaluate manually due to the amount of effort required. Therefore, in order to protect customers, Software AG is looking for new automated methods of risk assessment for these large software systems. The project is expected to deliver (define, create or select) a risk assessment methodology that can be applied in an automated way and provide repeatable and reliable assessment results. In particular, this requirement addresses the need for automation support to make the risk analysis feasible and economically viable.
Use Case Provider Satisfaction	5
Success Criterion	The requirement speaks of evaluating a large scale system in an automated way through the use of additional tools. Basically, we need to check that (a) the tools have been selected or provided and (b) the tools do allow us to perform a risk analysis of a large system in an automated fashion. SC-A2: The RASEN project has resulted in selection or creation of automated tools for security risk analysis. The tools are available and ready for deployment into production. SC-A3: The tools provided by the RASEN project facilitate automated security risk analysis of large systems to make the analysis economically viable.
Evaluation Criterion	Evaluation criteria related to artifact **A1**: **Phase 2 [M36]:** *E1*: (SC-A2) Does the project provide a set of tools for automated risk analysis? *E2*: (SC-A3) Using the toolset, a single product is evaluated using the provided methodology. The amount of effort is analyzed and extrapolated to the whole company.
Evaluation Result	
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	This criterion cannot be evaluated in the 1st evaluation phase.

Table 8 – Evaluation for requirement REQ-SAG-F-020

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-030**
Requirement	A methodology providing compositional security risk assessment.
Objective	O1
Description	This requirement identifies our need to have a clear-cut methodology for compositional risk assessment due to the modular architecture of the software.
Use Case Provider Satisfaction	5

Success Criterion	SC3.1: The approach provided by RASEN defines clearly and precisely the rules for valid composition of risk assessment and security testing results.
Evaluation Criterion	Evaluation criteria related to artifact **A2**: **Phase 1 [M24]/Phase 2 [M36]:** *E1*: There is a methodology available which allows for compositional security risk assessment. This methodology can be used to evaluate the risk to Software AG's software suit based on the evaluations at lower levels.
Evaluation Result	Succeeds if Yes.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Poor
Evaluation Phase 1 Results	*F1*: In the current methodology there is no composition of risk ratings from component to product level due to the lack of an aggregation function. If an aggregation function is available (apart from the simple sum of risk ratings), this requirement will receive a good rating.

Table 9 – Evaluation for requirement REQ-SAG-F-030

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-040**
Requirement	Tools supporting automated compositional security risk assessment.
Objective	O1
Description	The requirement captures our need to have state of the art tools supporting automation of the compositional security risk assessment of software. This requirement is for automation. Basically, we cannot perform any manual composition of risk analysis, e.g. through expert valuations. We need a method where changes at lower levels are automatically and completely reflected at the top level without manual intervention.
Use Case Provider Satisfaction	5
Success Criterion	SC-A4: The method of security assessment composition provided by RASEN allows for a fully automated implementation of such composition provided that the "bottom-of-the-graph" evaluations are available.
Evaluation Criterion	Evaluation criteria related to artifact **A2**: **Phase 2 [M36]:** *E1*: Perform a risk evaluation through the proposed methods with the supplied tools. *E2*: If we have some results at the bottom, applying an automated tool that implements the method should give us the results at the top. This should be automatic including the testing interfaces.
Evaluation Result	The method is: *F1*: Implementable as a tool *F2*: When run against the assessment results it shall provide a higher-level composite assessment.

Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	This criterion cannot be evaluated in the 1st evaluation phase.

Table 10 – Evaluation for requirement REQ-SAG-F-040

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-050**
Requirement	Tools providing generation of test cases guided by security risk assessment.
Objective	O2
Description	The requirement captures the importance of translating semi-formal security analyses into automatically generated executable tests that complement tests provided by security testing teams.
Use Case Provider Satisfaction	5
Success Criterion	SC-A5: are there tools for test case generation based on risk assessment? SC-A6: do these tools automatically generate suitable and usable test cases?
Evaluation Criterion	Evaluation criteria related to artifact **A4**: **Phase 2 [M36]**: *E1*: There must be a tool to support the generation of test cases *E2*: With the analysis of the risk experts are generated which allow the testing suite a generation of test cases guided by the input from the security risk assessment.
Evaluation Result	*F1*: the tools are available and running *F2*: we get usable test cases that lower the risk
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	This criterion cannot be evaluated in the 1st evaluation phase.

Table 11 – Evaluation for requirement REQ-SAG-F-050

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-060**
Requirement	Executable test cases providing adequate security coverage relative to the supplied risk picture.
Objective	O2
Description	This requirement relates to the quality of the automatically generated test

	cases and identifies the need for generating high-coverage test sets.
Use Case Provider Satisfaction	5
Success Criterion	SC2.2: The approach should help uncover more relevant security vulnerabilities than traditional security testing approaches (which are not guided by risk assessment). Frankly, that is fine but we still want to have the coverage as well, not limited to finding some more vulnerabilities. SC-A7: The tests generated by the RASEN tools must provide the coverage suitable for the level of risk evaluated.
Evaluation Criterion	Evaluation criteria related to artifact **A4**: **Phase 2 [M36]**: *E1*: The tool must provide the coverage data that will be evaluated by an expert versus the provided risk assessment. *E2*: The coverage should correlate to the level of risk as assessed by the methods of this project.
Evaluation Result	*F1*: The tools provide coverage analysis. *F2*: The coverage is deemed adequate to the risk level by an expert in the field.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	This criterion cannot be evaluated in the 1st evaluation phase.

Table 12 – Evaluation for requirement REQ-SAG-F-060

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-070**
Requirement	Tools providing execution of generated test cases.
Objective	O2
Description	This requirement identifies the need for toolbox components that enable running the generated security test cases.
Use Case Provider Satisfaction	5
Success Criterion	SC-A8: The RASEN project selects or creates tools suitable for running the generated test cases against large software systems.
Evaluation Criterion	Evaluation criteria related to artifact **A4**: **Phase 2 [M36]:** The evaluation basically boils down to: *E1*: Are there tools to run the test cases? *E2*: Do these tools function automatically with the RASEN generated test cases? *E3*: Are we able to run them against a large software system?

Evaluation Result	Answers Yes to the set questions are a pass.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	This criterion cannot be evaluated in the 1st evaluation phase.

Table 13 – Evaluation for requirement REQ-SAG-F-070

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-080**
Requirement	A methodology and toolset that supports automated aggregation of obtained test results into the risk picture.
Objective	O1
Description	This requirement identifies the need of a supporting methodology and toolset that enables the aggregation of security test results back into the high-level risk picture in an automated fashion.
Use Case Provider Satisfaction	5
Success Criterion	SC-A9: The results generated by running the RASEN test tool chain with the generated test cases are automatically imported back into the risk analysis and the risk analysis picture is updated to take into account the imported results. SC3.2: Composition at the risk assessment level should be well behaved with regards to composition at the testing level, e.g. the order in which risk assessment results are composed and transformed to the testing level should be irrelevant.
Evaluation Criterion	Evaluation criteria related to artifact **A3**: *E1*: Executing the tool chain, testing results will be imported, resulting in an update of the risk analysis. *E2*: In order to check the second requirement, we should be able to change the order of result creation/import, and then we can re-run the analysis and see whether the result is still the same. **Phase 1 [M24]**: *E3*: A simple aggregation function (like the mathematical sum or union) is used which ensures the aggregation of test results into the risk picture of the product, resulting a product risk score. **Phase 2 [M36]**: *E4*: A more sophisticated and reliable aggregation function is implemented, delivering more meaningful results of the individual risk sources to the product level risk picture.
Evaluation Result	*F1*: This criterion is currently difficult to assess due to the missing integration of the test-execution into the risk aggregation. However, test results are easy to import which fulfill this criterion. *F2*: This criterion is a simple modification of the input and easy to be fulfilled with an working test execution integration *F3*: A simple aggregation function is currently present which fulfills this criterion.

Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Good
Evaluation Phase 1 Results	A simple aggregation function has been implemented which allows the import of test results into the risk picture by dropping unconfirmed weaknesses from the analysis.

Table 14 – Evaluation for requirement REQ-SAG-F-080

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-090**
Requirement	A methodology and a toolset that supports automatic import and aggregation of secondary risk evaluation sources at component and aggregate level.
Objective	O5
Description	This requirement identifies the need of a supporting methodology and toolset that enables the aggregation of security risk relevant information obtained from external sources back into the high-level risk picture.
Use Case Provider Satisfaction	3
Success Criterion	SC-A10: The risk analysis tool chain provides a clear definition and an implementation of a communication interface that allows influencing the risk analysis by supplementing information.
Evaluation Criterion	Evaluation criteria related to artifact **A3**: **Phase 2 [M36]**: *E1*: The interface is tested by importing externally available sources containing security risk information and adding this information to the corresponding place in the risk assessment model. Is requires a naming convention in place.
Evaluation Result	*F1*: Yes – succeeds.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	This criterion cannot be evaluated in the 1st evaluation phase.

Table 15 – Evaluation for requirement REQ-SAG-F-090

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-F-100**
Requirement	A methodology and toolset that supports reverse analysis of the impact of risk evaluation sources at component and aggregate level.
Objective	O5

Description	This requirement identifies the need of a supporting methodology and toolset that enables us to analyze the impact of changes and tracing them back to the evaluation sources from the high-level risk picture. The visibility of the risk impact of different sources of the security risk is important in tracing the impact back to its origin.
Use Case Provider Satisfaction	5
Success Criterion	SC-A11: The tool chain must provide clear traceability between the top-level risk assessment and the influencing factors.
Evaluation Criterion	Evaluation criteria related to artifact **A5**: **Phase 1 [M24]:** *E1*: After a complete risk assessment is done, we change components (e.g., by changing the set of applicable CWEs) on the bottom of the pile, i.e. the leaves in the product tree, and see the resulting risk assessment (risk rating) changes. **Phase 2 [M36]**: *E2*: After a complete risk assessment is done, we change something at the bottom of the pile (what?), start the testing tool chain and see the resulting assessment change according to the test results. Now, can we trace the change all the way back to where the original change was made unambiguously?
Evaluation Result	*F1*: the composition is currently missing.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	The evaluation criteria cannot be checked as the composition is missing (REQ-SAG-F-030: A methodology providing compositional security risk assessment)

Table 16 – Evaluation for requirement REQ-SAG-F-100

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-N-020**
Requirement	Provided tools must support large systems and enable the compositional security risk analysis of large software products within an economically viable level of investment.
Description	This requirement ensures the applicability of the results to the Software AG infrastructure.
Use Case Provider Satisfaction	4
Success Criterion	SC-A12: The project provides tools developed or selected for the purposes of compositional risk analysis of large software systems with multiple hierarchy levels of components. SC-A13: The tools must support automated and semi-automated processes and integration with other tool chains, allowing for a commercially viable analysis of large software products in the development process.
Evaluation Criterion	Evaluation criteria related to artifact **A2:** **Phase 2 [M36]**: *E1*: The tools are available for integration

	E2: The tools can be integrated with the development process at a reasonable cost for automated analysis, the success is guaranteed.
Evaluation Result	There are currently no evaluation results available in M24.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	This criterion cannot be evaluated in the 1st evaluation phase.

Table 17 – Evaluation for requirement REQ-SAG-N-020

Requirement Evaluation	
Name	Description
Code	**REQ-SAG-N-030**
Requirement	Provided tools must enable the compositional security risk analysis of large software products within a linear or better time relative to the number of components (number of classes, lines of code, number of tests etc.) analyzed.
Description	This requirement ensures the applicability of the results to the Software AG infrastructure.
Use Case Provider Satisfaction	5
Success Criterion	SC-A14: The RASEN method and tool chain must operate in linear or better time relative to the complexity of the system (number of components or classes, lines of code, number of tests etc.)
Evaluation Criterion	Evaluation criteria related to artifact **A2**: **Phase 2 [M36]:** *E1*: There are tools – in particular the testing tool chain – available *E2*: Tools operate in linear time when tested on our products vs. the LOC and number of components?
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Not Available
Evaluation Phase 1 Results	This criterion cannot be evaluated in the 1st evaluation phase.

Table 18 – Evaluation for requirement REQ-SAG-N-030

5.2.2 EVRY

5.2.2.1 Evaluation Process

The evaluation related to the EVRY case study is mainly conducted in collaboration between EVRY on the one hand side and SINTEF, UiO, and Smartesting on the other hand. The latter partner's main interest is to evaluate the following artifacts in the EVRY case study:

- **A1**: The RASEN method and technique for risk-based test identification and prioritization.
- **A2**: The RASEN method for compliance risk assessment

- **A3**: The RASEN techniques for security test automation

Note that we will refer to these artifacts in the next evaluation section.

The evaluation process in the EVRY case study involves applying the above mentioned artifacts to security assess EVRY's Netbank system, and to compare this assessment with the way the Netbank system is currently assessed by the process currently in place at EVRY.

During the last year, the evaluation has also involved two workshops as part of the collaboration between EVRY, SINTEF, and UiO for evaluating artifacts A1 and A2. The first workshop (a two-day meeting in Trondheim) primarily addressed artifact A1, and its use for the purpose of security test prioritization. The second workshop (a one day meeting in Oslo) primarily addressed artifact A2 for the purpose of legal compliance assessment, and involved legal compliance experts from EVRY.

A third workshop is planned for October 9th and will mainly address artifacts A3 for the purpose of security test automation. This workshop will involve EVRY, Smartesting, and SINTEF. After this third workshop, the case study will reiterate, starting again with risk assessment and legal compliance, followed up by security testing.

During the case study, the artefacts will be continuously evaluated according the evaluation criteria (as summarized in the next section).

5.2.2.2 Evaluation Result

Requirement Evaluation	
Name	Description
Code	**REQ-EVRY-F-010**
Requirement	The RASEN artifacts must improve EVRY's security test prioritization process if adapted.
Objective	O5
Description	This requirement refers to the improvement of the "Test Requirements Gathering" and the "Test Planning and Prioritization" activity of the EVRY security testing process. The kinds of improvements possible are: time/effort reduction and efficiency (roughly corresponding to the number of security issues uncovered w.r.t. effort).
Use Case Provider Satisfaction	5
Success Criterion	SC5.1
Evaluation Criterion	Evaluation criteria related to artifact **A1**: *E1*: All relevant security test cases can be seen as a refinement of a test procedure derived from CAPEC according to artifact **A1**. *E2***:** The level of abstraction of the CAPEC derived risk model (by artifact **A1**) is appropriate for security test identification. *E3*: The prioritization of the test procedures generated by artifact **A1** is according to intuition. *E4*: The risk visualization of security test related risks by **A1** is according to intuition. *E5*: The likelihoods are defined appropriately by **A1** *E6*: Estimating attack success likelihood according to **A1** is easy. *E7*: Estimating technical impact likelihood according to **A1** is easy. *E8*: The effort spent on test prioritization according to **A1** will be saved in the testing phase. Evaluation criteria related improvement of EVRY's testing process through

	A1: *F1*: The RASEN test procedure technique (**A1**) is more rigorous than the current EVRY test prioritization process. *F2:* Artifact **A1** helps prioritize test procedures more accurately than EVRY's current process for doing this. *F3:* Test prioritization according to artifact **A1** may help save time during the testing phase of the EVRY testing process. *F4:* Taking the test procedures derived according to **A**1 as starting point for testing is better than current starting point at EVRY (this is the security requirements)
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Fair
Evaluation Phase 1 Results	*E1*: CAPEC seems to be a sufficient starting point for the kind of security testing which is performed at EVRY. However, it is important that the CAPEC catalogue be kept up to date so that the latest security vulnerabilities and attacks can be addressed in the testing process. *E2*: The level of abstraction in which test procedures are described seems ok, and is similar to the level EVRY is currently using to describe security requirements (which are EVRYS's starting point for test identification). The test procedures should not be described in more detail so as to not limit explorative security tests. *E3 – E5*: These criteria are currently difficult to assess. However, we cannot say that the criteria are obviously fulfilled or not fulfilled at this point. *E6 - E7*: During the case study estimating likelihoods according to artifact **A1** has been fast and the participants of the case study have been able to get a quick intuitive feeling about the estimates. This suggests that criteria E6 and E7 are both fulfilled. *E8*: This criterion is probably true/fulfilled. However, for critical systems (such as the one addressed in the EVRY case study), cutting/not considering certain kinds of security tests might not be an option. However, in that case, one could consider spending less time on lower priority tests. *F1*: The current method of prioritization at EVRY is unstructured and based on expert judgment. The manner of prioritization may also vary from case to case. Adapting artifact A1 would therefore provide rigor to this process. *F2*: This criterion in the sense true that EVRY does not perform any structured/documented prioritization of test cases. However, during the testing, a prioritization is performed implicitly, and it is currently hard to assess whether the prioritization obtained through artifact A1 is more accurate than this implicit prioritization. , *F3*: This evaluation criterion is probably true/fulfilled. For critical systems, cutting certain tests might not be an option, but less time could be used. Currently, it might be the case that too much time is spent on tests that are not worth it. *F4*: If security requirements are used as input in the process for deriving

	test procedures according to artifact **A1,** then is will probably be a better basis for testing than using security requirements alone. *Summary:* Adaptation of artifact **A1** into the EVRY testing process will likely provide rigor to the manner in which test cases are prioritized. In addition, the level of abstraction in which test procedures are described in artifact **A1** corresponds well with the way this is currently done at EVRY. A this point, we cannot say that test procedure prioritization of artifact **A1** is wrong/unintuitive. However, we cannot say that is completely correct either. In addition, the time saving benefit of using the prioritization of artifact **A1** needs to be assessed more carefully. Therefore we overall evaluation rating is currently *Fair*.

Table 19 – Evaluation for requirement REQ-EVRY-F-010

Requirement Evaluation	
Name	Description
Code	**REQ-EVRY-F-020**
Requirement	The RASEN artifacts must improve EVRY's test execution process if adapted.
Objective	O5
Description	This requirement is primarily related to the need of automation parts of the security execution which is currently performed manually at EVRY. The requirement is that the automation will save time and that it will at least result in equal or better test results.
Use Case Provider Satisfaction	5
Success Criterion	SC5.1
Evaluation Criterion	Evaluation criteria related to artifacts **A3**: Evaluation criteria related to improvement of EVRY's security testing process through **A3** *F1*: Adapting artifact **A3** into EVRY's process will automate parts of the security testing process which is currently performed manually. *F2*: The adaptation of artifact **A3** will result in more security vulnerabilities being uncovered trough testing than what is currently being uncovered through EVRY's testing process. *F3*: Adapting artifact **A3** will save time during the test execution phase of the EVRY security testing process.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Fair
Evaluation Phase 1 Results	*F1*: It is clear that the adaption of artifact **A3** will result in the automation of some of the test execution tasks which are currently performed manually at EVRY. *F2*: This criterion is currently being evaluated. *F3*: This criterion is true provided that the artifact **A3** has been properly configured/set up prior to the test execution phase. The quantification of the criterion F3 is currently under evaluation. *Summary:* The adaption of artifact **A3** will result in the automation of part

	of EVRY's security test execution phase. The quantification of the benefits/drawbacks of this is currently under evaluation.

Table 20 – Evaluation for requirement REQ-EVRY-F-020

Requirement Evaluation	
Name	Description
Code	**REQ-EVRY-F-030**
Requirement	The RASEN artifacts must enable better decision making related to security test and compliance assessment if adapted.
Objective	O3, O4
Description	This requirement is related how the security test results are communicated and used as basis for decision making The requirement also relates to EVRY's compliance process.
Use Case Provider Satisfaction	5
Success Criterion	SC5.1
Evaluation Criterion	Evaluation criteria related to artifacts **A1** and **A2**: *E1*: The costs of using the method (of artifact **A2**) is, in the long run, lower than the value of the benefits from its use. *E2*: The risk matrix (as obtained by artifact **A1**) is a useful way of communicating the security test results. Evaluation criteria related to improvement of EVRY's compliance process through **A2**: *F1*: The method of artifact **A2** provides an increased level of confidence on the compliance of the organization, compared to EVRYS current method. *F2*: The method of artifact **A2** provides better input to decision making than EVRYS current method. *F3*: The technique for compliance risk identification (artifact **A2**) enables a better structuring in identifying compliance risks than EVRYS current method.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Fair
Evaluation Phase 1 Results	*E1*:It is currently too early to tell whether this criterion is fulfilled or not. However, we have learned that the identification of legal and compliance risks involves too much analytical activity, which can sometimes be frustrating. Therefore, we have developed a technique for structuring the identification of compliance risk. This technique enables a mechanization of compliance risk identification process, therefore potentially saving time. *E2*: This criterion is currently under evaluation. *F1*: This criterion is currently under evaluation. *F2*: This criterion is currently under evaluation. *F3*: We believe that this criterion is fulfilled. Currently, EVRY does not have a structured method for identifying compliance risks. Therefore the adaption of artefact **A2** would provide such a method. *Summary:* We currently believe that the adaptation of artefact **A3** into

	EVRY's process for compliance assessment will provide more structure and rigor to the process. This could potentially save time and increase/strengthen the accuracy of the compliance assessment. The latter point is however currently under evaluation.

Table 21 – Evaluation for requirement REQ-EVRY-F-030

5.2.3 Info World

5.2.3.1 Evaluation Process

The evaluation of the Info World case study was conducted with Info World on one side, and SINTEF, Smartesting and UiO on the other. The research and technical partners' have identified the following RASEN artifacts that can be evaluated using the Info World use case:

- **A1**: The RASEN method and technique for risk-assessment
- **A2**: The RASEN method for compliance risk assessment
- **A3**: The RASEN methodology for security risk assessment compositionality
- **A4**: The RASEN techniques and tools for risk-based testing and test-based risk assessment.

The artifacts mentioned above will be reused within the use case evaluation within the following section.

The evaluation process was organized as three interactive online workshops attended by RASEN scientists, on one hand, and Info World decision makers, system engineers and developers on the other. The first of these workshops focused on presenting and evaluating the current status of the RASEN methodology and toolset for risk assessment, the second workshop was focused on the automation of security testing and its integration with risk assessment while the final workshop addressed the issue of legal compliance. The present section details these activities, while evaluation criteria and results per se are detailed within the following section.

Risk assessment workshop (25th and 26th of June, 2014)

The RASEN security risk assessment methodology, assisted by currently available RASEN tooling was applied to Info World's Medipedia eHealth web platform during a two-day online workshop held on the 25th and 26th of June, 2014. The workshop and technical activities were attended by SINTEF's *Fredrik Seehusen*, specialized in risk assessment, as scientist working on the RASEN project. On the other hand, Info World's researchers involved in the project (*Cornel Botea*, *Arthur Molnar*) together with Info World's technical staff involved in the design, development and maintenance of the Medipedia platform (*Daniel Jianu*, *Vlad Racovita* and *Silvia Dusceac*) attended.

The workshop was organized as such:

- Technical artifacts that model considered security risks found in the CAPEC[3] vulnerabilities database were prepared ahead of time by SINTEF.
- Info World's representatives provided technical information regarding the design, known vulnerabilities and attack patterns possible against the Medipedia platform.
- This information was incorporated into the already prepared models with the risk picture being automatically generated by SINTEF. This activity resulted in a structured risk assessment of the Medipedia platform, which is stored in encrypted form on the project's eRoom.
- Discussion regarding future improvements required, especially with regards to integration of the various tools that will form the final RASEN toolchain.

The results of this activity from an evaluation standpoint are detailed within the following section of this deliverable.

[3] https://capec.mitre.org/

Automated security testing workshop (5th of August, 2014).

The purpose of the second online workshop was to allow RASEN research partners to showcase methodology and tooling achievements that are relevant to the Info World use case, to present the roadmap for further work and to gather feedback that will ensure the usability and utility of the project's resulting artifacts for actual use case partners and future users of the platform. The security testing workshop was coordinated by Smartesting's *Julien Botella* and was organized as follows:

- Brief presentation of the RASEN tool chain between risk assessment and security testing, as illustrated within Figure 13.
- Presentation of the DSL employed to model the Medipedia, together with further planned improvements such as an advanced editor that will facilitate the creation of such artifacts.
- Presentation of how test purposes and patterns, together with risk assessment data are employed in the creation of test cases, together with future work required for the integration of full risk assessment data within the test generation process.
- Presentation regarding the concretization of tests. How abstract test cases are transformed into JUnit tests that can be automated.
- Showcasing automated test execution using a privately-hosted Medipedia clone across the Internet.

The results of this activity from an evaluation standpoint are detailed within the following section of this deliverable.

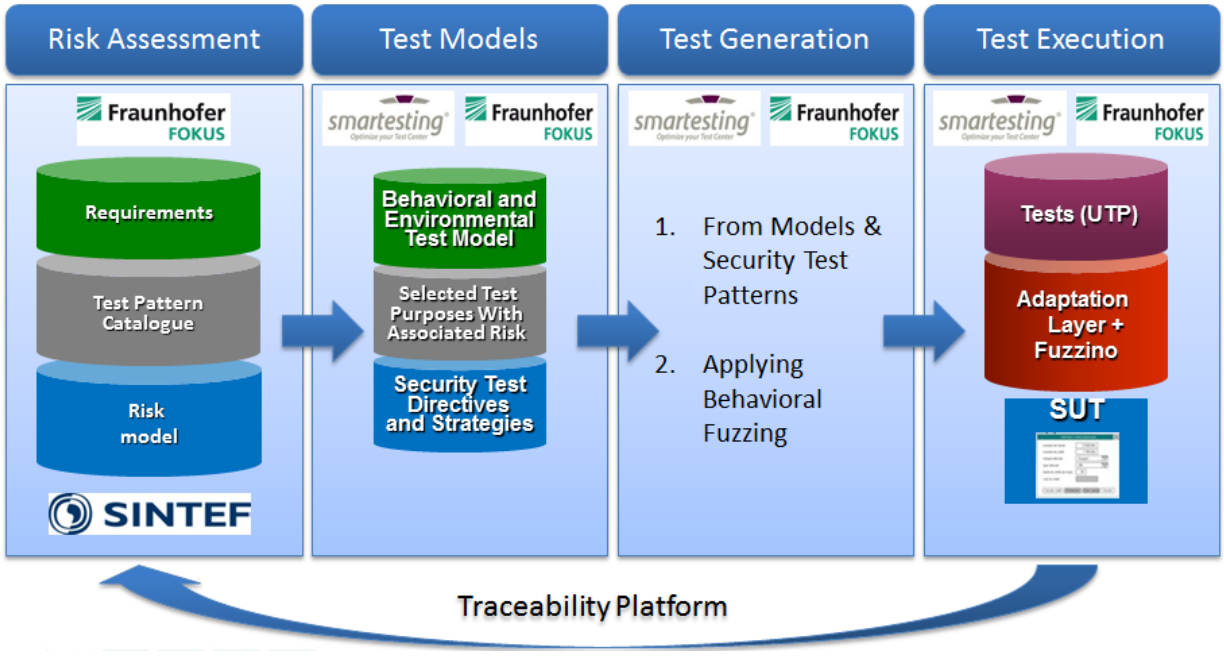

Figure 13 – RASEN tool chain

Legal compliance workshop

The first evaluation activity regarding legal compliance within the Info World use case took place on September 4th and was attended by UiO and Info World researchers. The workshop was organized as follows:

- Understanding the business and regulatory environment for the Info World use case.
- Identification of relevant compliance requirements.
- Identification of possible compliance issues
- Identification of compliance risk

- Compliance risk modelling, estimation and evaluation using CORAS

As this activity was the first evaluation of methodology and tooling for legal compliance risk management, the workshop was not focused on identification and estimation of all risks; the focus was to provide an end-to-end scenario for several risks that can materialize within the Info World business case.

The results of this activity are detailed within the following section.

5.2.3.2 Evaluation Result

Requirement Evaluation	
Name	Description
Codes	**REQ-IW-F-010, REQ-IW-F-020, REQ-IW-F-030, REQ-IW-F-040**
Requirement	A structured methodology and associated software tooling that provides means for ascertaining the legal compliance of Info World developed software components as well as customized software solution deployments to a set of legal norms.
Objective	O4
Description	This requirement identifies the need for developing a new methodology and toolset that support the process of checking for legal compliance of existing software components and software solution deployments against a well-defined body of legislation.
Use Case Provider Satisfaction	4
Success Criterion	SC4.1, SC4.2 and SC4.3
Evaluation Criterion	Evaluation criteria related to artifact **A2** *E1*: The artifacts allows for understanding the relevant business and regulatory environment *E2*: Relevant compliance requirements can be identified according to **A2** *E3*: **A2** provides for the identification of compliance risks *E4*: Compliance risks can be modeled in a structured manner according to the **A2** artefact. *E5*: **A2** enables structured estimation of compliance risk *E6*: **A2** enables structured evaluation of compliance risk *E7*: Estimation, evaluation of compliance risks is easier using **A2**. *E8*: **A2** can be applied to systems of various complexity and modularity, from in-house software components to assembled software systems delivered to customers. Evaluation criteria improvement of Info World's testing process through **A2**: *F1*: The RASEN artifact **A2** provides increased level of confidence on the compliance of the organization, compared to the alternative. *F2*: The RASEN artefact **A2** enables better input to decision making than the alternative. *F3*: The cost of using **A2** is in the long run lower than the value of the benefits from use.
Evaluation Phase 1 [M24]	
Evaluation Phase 1	Fair

Rating	
Evaluation Phase 1 Result	*E1*: Understanding the business and regulatory environment remains still a manual undertaking for legal experts. However, preparatory steps such as introduction of structured likelihood and consequence scales can bring further structure to the process. *E2*: Identification of compliance requirements is currently not affected by the A2 artifact. *E3*: Compliance risks are still identified manually, however this is done according to a consistent template. *E4*: Compliance risks can be modelled using the CORAS tool. *E5 – E6*: Estimation and evaluation of compliance risks can be achieved using structured CORAS notation and the CORAS tool that produces risk values once likelihood estimations are provided. *E7*: While on the studied system and modelled risks estimation and evaluation of risks was easy, at this point we cannot extrapolate this to the entire system. *E8*: The present evaluation was focused on the Medipedia system. Its system architecture includes several components that are broadly reused within Info World, making their safety, reliability and legal compliance of paramount importance. While this first evaluation did not have the scale and complexity that would allow assessing this, given the complexity of the evaluated system we estimate **A2** to be applicable across systems of different scales and complexities. *F1 – F3*: Given the limited time and scale of this first evaluation, we cannot readily provide estimates regarding the long-term feasibility and cost-effectiveness of **A2**. However, with regards to decision making input **A2** appears to provide clear structure and effective tooling. Given the limited time and scale of this evaluation, as well as the observations above we provide the *Fair* rating with regards to applying **A2** to Info World's Medipedia use case.

Table 22 – Evaluation for requirements REQ-IW-F010, REQ-IW-F-020, REQ-IW-F-030 and REQ-IW-F-040

Requirement Evaluation	
Name	Description
Code	**REQ-IW-F-050, REQ-IW-F-060**
Requirement	A methodology and toolset providing structured security risk assessment for Info World developed software components and end products.
Objective	O5
Description	This requirement identifies Info World's need for a structured process of

	security risk assessment. Due to the security implications of dealing with sensitive personal data, such risks must be considered at each step of the development process. However, currently Info World only employs an ad-hoc process that is based on the technical knowledge of its analysts, developers and testers without employing a formalized methodology or specialized tooling.
Use Case Provider Satisfaction	5
Success Criterion	SC5.1
Evaluation Criterion	Evaluation criteria related to artifact **A1**: *E1*: A structured, tool-backed methodology that deployable for undertaking security risk assessment of Info World's software components and end products is available. *F1*: The process enabled by **A1** provides a more correct risk model than the current alternative. More precisely, if the target system is analyzed using both the current and **A1** methods by committing the same resources, the model yielded by **A1** will be equally or more correct. *F2*: Employing **A1** will bring more confidence in the correctness of the risk model that the current approach. *F3*: A large part of the risk model produced following **A1** can be tested using conventional security tools.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Fair
Evaluation Phase 1 Result	*E1*: It is currently too early to provide a definitive evaluation. **A1** was already employed targeting the Medipedia eHealth system that resulted in a structured risk assessment model which Info World considers to be more advanced that what it had available before. However, due to the limited scope of its applicability (only 1 system) and due to the fact that identified risks have not yet undergone testing we believe a current evaluation of Fair would be most adequate. *F1-F2*: These criteria are currently under evaluation. More information will be available once identified risks have undergone testing. *F3*: As identified risks are linked with well-known vulnerability databases, they are geared towards the same end as Info World's existing methods. As such, we believe the risk model is conductive to the deployment of automated testing tools.

Table 23 – Evaluation for requirements REQ-IW-F050, REQ-IW-F-060

Requirement Evaluation	
Name	Description
Code	**REQ-IW-F-070, REQ-IW-F-080**
Requirement	A methodology and toolset providing compositional security risk assessment for Info World's software solutions.
Objective	O3
Description	This requirement identifies Info World's need of a structured methodology and associated tooling that will enable the organization to obtain up to date security risk assessments for its end-products by composing the results of available assessments both for individual software components as well as for its end products.
Use Case Provider Satisfaction	5
Success Criterion	SC1.2, SC1.3
Evaluation Criterion	*E1*: Artifact **A3** allows risk assessments for Info World's software components to carry across to its assembled end products. When an updated risk model is available for a software component, the assembled product risk model and testing prioritization can be updated.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Fair
Evaluation Phase 1 Result	E1: This criterion cannot be evaluated. At this point Info World's Medipedia system has undergone a risk evaluation, however there was no focus on individual software components or their interplay. However, there is reason to believe the A1 artifact that was demonstrated can be employed for the assessment of software components both simple and large, with likelihood and consequence scales that appear feasible for reusing results in the picture of a large-scale, assembled system.

Table 24 – Evaluation for requirements REQ-IW-F-070 and REQ-IW-F-080

Requirement Evaluation	
Name	Description
Code	**REQ-IW-F-090, REQ-IW-F-100**
Requirement	Tools supporting generation and execution of security test cases guided by security risk assessment and aggregation of test results back into the updated risk picture.
Objective	O2
Description	The requirement captures the importance of translating structured security analyses into automatically generated executable tests that complement Info World's security testing team. The generated tests must allow for obtaining comprehensive coverage of the software systems targeted by the RASEN approach.
Use Case Provider Satisfaction	5
Success Criterion	SC2.1

Evaluation Criterion	Evaluation criteria related to artifact **A4**: *E1*: Security test cases are a refinement of a structured test procedure targeting known types of vulnerabilities *E2*: Test cases for **A4** can be generated using artifact **A1**. *E3*: Effort spent on additional actions for obtaining test cases within **A4** is saved in the testing phase: *E4*: Adapting **A4** results in more security vulnerabilities being uncovered than what is being uncovered using the current process. *E5*: Testing results of **A4** can be used to update **A1**. *F1*: The RASEN test technique (**A4**) is more rigorous than Info World's current test prioritization process. *F2*: Artifact **A4** helps prioritize test procedures more accurately than Info World's current process. *F3*: Test prioritization according to artifact **A4** may help save time during the risk assessment and testing phase for Info World.
Evaluation Phase 1 [M24]	
Evaluation Phase 1 Rating	Fair
Evaluation Phase 1 Result	*E1*: The **A4** artefact is in close interplay with **A1**, which is based on the CAPEC vulnerability database. During the Info World evaluation activities, it was shown to be comprehensive for commercial use and targeting the same types of vulnerabilities and attacks that internal testing at Info World was already focused on. As such, we consider *E1* to be fulfilled satisfactorily at this point. *E2-E3*: These criteria cannot be evaluated at this point due to more integration work required between the required software components. *E4-E5:* This criterion cannot be evaluated at this point as testing using **A4** was not deployed at this point. *F1*: The current process of test prioritization is based on the expertise of the company's development and testing staff. While we have reason to believe a structured methodology could improve our current processes, at this point such a comparison cannot be made. *F2*: The result of the risk assessment targeting the Medipedia system was a prioritized list of risks that must be explored through security testing. The obtained priorities were consistent with the company's previous experience and we believe **A4** could lead to better overall prioritization. However, the **A4** testing process must be deployed to evaluate this. *F3*: This criterion cannot be evaluated at this point.

Table 25 – Evaluation for requirements REQ-IW-F-090 and REQ-IW-F-100

Requirement Evaluation	
Name	Description
Code	**REQ-IW-N-110**
Requirement	Provided tools must work under recent versions of Microsoft Windows (at least XP/Vista/7/8)
Description	This requirement ensures ease of use within Info World's IT infrastructure.
Use Case Provider Satisfaction	3
Success Criterion	-
Evaluation Criterion	*E1*: RASEN tooling is available and offers full functionalities under versions of Microsoft Windows (at least XP/Vista/7/8)
Evaluation Phase 1 Rating	Excellent
Evaluation Phase 1 Result	*E1*: Available tooling was evaluated under Microsoft Windows 7 and found to work without issues. Due to the interoperability of their underlying development platform RASEN tooling is expected to fulfill this requirement.

Table 26 – Evaluation for requirement REQ-IW-N-110

Requirement Evaluation	
Name	Description
Code	**REQ-IW-N-120**
Requirement	Provided tools must come with intuitive graphical user interfaces
Description	This requirement ensures ease of use from the end users' perspective, helping with easy adoption of the toolbox.
Use Case Provider Satisfaction	3
Success Criterion	-
Evaluation Criterion	*E1*: Tooling associated with artefacts **A1**-**A4** must provide an intuitive user interface, with clearly marked controls that present a gentle learning curve for domain specialists.
Evaluation Phase 1 Rating	Excellent
Evaluation Phase 1 Result	*E1*: The tools evaluated within the organized workshops were based on the well-known Eclipse framework ant offered a user-friendly GUI experience. More so, the graphical representation for various concepts used in security risk assessment are taken from the CORAS methodology that already has extensive documentation available and is therefore intuitive and easy to follow.

Table 27 – Evaluation for requirement REQ-IW-N-120

6 Coverage of Project Objectives

The present Section is dedicated to showing how requirements made by the project's use case partners cover the RASEN project objectives. The main objective of the project is to "*strengthen European organizations' ability to conduct security assessments of large scale networked systems through the combination of security risk assessment and security testing, taking into account the context in which the system is used, such as liability, legal and organizational issues as well as technical issues*".

The project's main objective will be achieved through the following scientific and technical objectives:

Objective	Description
O1	Enable organizations (including their non-technical experts) to understand what low-level security test results mean in terms of risks and legal obligations by aggregating security test results to the risk assessment level.
O2	Enable organizations to guide the security testing by high-level technical as well as non-technical considerations through systematic derivation of security test cases from risk assessment results.
O3	Enable organizations to obtain a global view of the security of large scale network systems through compositional assessment.
O4	Make it easier for organizations to show that they are compliant with legal norms of relevance to security.
O5	Enable continuous and rapid security risk assessment of large scale networked systems.

Table 28 – RASEN S&T Objectives

Adequate coverage of project objectives ensures that all aspects addressed by the project are evaluated within at least one of its use cases. As such, the requirements template that was defined within deliverable *D2.2.1 – Use Case Requirements Definition*, Section 3.2 includes the *Objective* section, enabling use case providers to link each requirement to a project objective. This was carried on in the present deliverable, where the evaluation template which is detailed within Section 4 includes the same row.

Objective	Coverage
O1	Meeting this is evaluated within the Software AG use case through the evaluation of requirements REQ-SAG-F-030, REQ-SAG-F-040 and REQ-SAG-F-080.
O2	Meeting O2 is evaluated within both the Software AG and Info World use cases through the evaluation of requirements REQ-SAG-F-050, REQ-SAG-F-060 and REQ-SAG-F-070 for Software AG, as well as REQ-IW-F-090, REQ-IW-F-100 for Info World.
O3	Meeting objective O3 is evaluated within through the EVRY use case via requirement REQ-EVRY-F-030 as well as within the Info World use case, via requirements REQ-IW-F-070 and REQ-IW-F-080.
O4	Whether objective O4 is met is evaluated within the Info World use case using the requirements REQ-IW-F-010, REQ-IW-F-020, REQ-IW-F-030 and REQ-IW-F-040 targeting legal compliance.
O5	This objective is evaluated within all project use cases. Requirements REQ-SAG-F-010, REQ-SAG-F-020, REQ-SAG-F-090 and REQ-SAG-F-100 target O5 from Software AG's perspective. Requirements REQ-EVRY-F-010 and REQ-EVRY-F-020 evaluate O5 via the EVRY use case while REQ-IW-F-050 and REQ-IW-F-060 do so within the Info World use case.

Table 29 – Evaluation coverage of S&T objectives

The pie chart in Figure 14 details how project objectives are covered within the requirements defined by the use case partners.

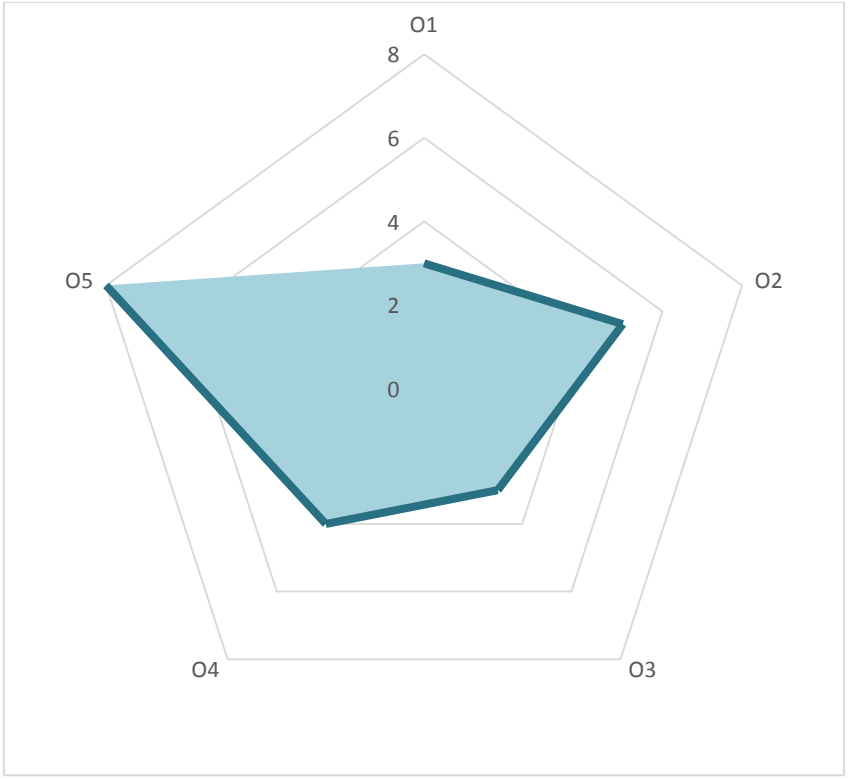

Figure 14 – Objective coverage chart

The presented data shows that all scientific and technical objectives are covered by use case partner requirements, with 4 out of the 5 objectives being covered by more than one use case.

7 Conclusion and Future Work

The goal of WP2 is to clearly define the proposed use case scenarios, to extract clearly defined and measurable use case requirements and to evaluate the technical progress of the project with regards to how the developed methodologies, tools and techniques help use case providers with finding solutions to the proposed requirements. To facilitate a broad-reaching approach, three organizations from three different countries that develop secure complex networked software as a main part of their activities were selected as use case providers.

The current document shows that project objectives are adequately covered by use case partner requirements and that the work from the project's technical work packages already addresses most partner requirements.

As first outlined within deliverable *D2.2.1 – Use Case Requirements Definition*, evaluation phase 2 will be undertaken during the project's final year and it will be finalized at the M36 mark. Evaluation Phase 2 will follow the first phase and will be undertaken within the last year of the project, up until Evaluation Milestone 11, as shown on Figure 9. Like Evaluation Phase 1, it will also consist of two stages:

- Evaluation Stage 1 – Research and technical partners will deliver latest tools and methodologies to the use case partners and will help them implement desired changes in their organizational workflows. Because by this point use case providers should already be experienced in using and implementing RASEN methodologies and tools, the role of the technical partners is expected to be more limited than within the *Learning phase* of Evaluation Phase 1. The prospective time-frame of this stage includes the first six months of Evaluation Phase 2, therefore the M24 – M30 period of the project implementation.
- Evaluation Stage 2 – Represents the final stage in evaluating RASEN methodologies and tools. As preparation of this stage, use case partners will receive the latest artifacts from the scientific and technical partners. As the final stage of evaluation, use case providers will employ the delivered tools without assistance from the technical partners involved. Stage 2 of the evaluation will conclude with a report delivered by the use case partners that provides detailed information regarding the successful use of developed methodologies and tools within each use case provider organization. The produced report will link obtained results with use case requirements stated in this deliverable and will use the RASEN objectives and success criteria for a thorough assessment of the project results. The prospective time-frame of this stage is the last six months of the project implementation, namely M30 - M36.

Current plans for the next research and development phase include work on risk assessment composition, further work on updating security risk assessment using security testing results as well as integration work on components of the RASEN tool chain. These efforts are expected to cover those use case partner requirements that could not have been evaluated at this point.

References

[1] Handbook: webMethods Command Central Help, Version 9.6, April 2014, http://documentation.softwareag.com/webmethods/wmsuites/wmsuite9-6/Command_Central_and_Platform_Manager/9-6_Command_Central_Help.pdf

[2] RASEN Deliverable D5.3.2

[3] R. Marselis, R. van der Ven, TPI NEXT CLUSTERS FOR CMMI, http://www.tmap.net/sites/tmap.net/files/attachments/ TPI NEXT clusters for CMMi 0.pdf, 2009.

[4] SOGETI, Website of SOGETI, http://www.sogeti.nl/, 2009

[5] R. van Veenendaal, Test Maturity Model integration, http://www.tmmi.org/pdf/TMMi.Framework.pdf

[6] Arthur-Jozsef Molnar and Jürgen Grossmann - CRSTIP - An Assessment Scheme for Security Assessment Processes (accepted paper at RISK 2014 workshop within ISSRE14).