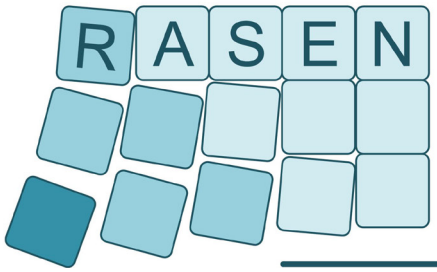


Compositional Risk Assessment and Security Testing of Networked Systems

# Deliverable D2.3.2

# Use case evaluation v.2

| | |
|---|---|
| **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:** | D2.3.2 |
| **Nature of deliverable:** | Report |
| **Dissemination level:** | PU |
| **Internal version number:** | 1.0 |
| **Contractual delivery date:** | 2015-09-30 |
| **Actual delivery date:** | 2015-09-30 |
| **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) | Fredrik Seehusen (SINTEF), Johannes Viehmann (FOKUS) |

## Version history

| Version | Date | Description |
|---|---|---|
| 0.1 | 15-06-08 | First version of TOC. |
| 0.2 | 15-08-20 | First IW contribution |
| 0.3 | 15-08-31 | Integrated SAG contribution, IW contribution finalized |
| 0.4 | 15-09-09 | Finalized conclusion, restructured Section 6 |
| 0.5 | 15-09-14 | Integrated EVRY evaluation |
| 0.6 | 15-09-23 | Updated according to internal review. Ready for final quality check. |
| 1.0 | 15-09-30 | Final quality check done. |

## 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 second phase of the evaluation process taking place within the project's third and final year as well as to evaluate the project progress with regards to partner established criteria.

## 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 completes the work within WP2 that started with the definition of the project's industrial use cases by providing the final evaluation of the RASEN tool-supported methodology using the three use case systems: Software AG's Command Central, EVRY's Net Bank software and Info World's Medipedia eHealth portal. The current document uses the user requirements that were first defined in deliverable *D2.2.1 - Use case requirements definition* as well as an updated evaluation template that was first used within the first evaluation phase, as detailed within *D2.3.1 - Use case evaluation v.1*.

As the technical results of the RASEN project were disseminated as three main innovations and several artefacts, the present document provides a unified, innovation-centric result of the evaluation process. This is achieved by aggregating the evaluation rating of each individual requirement into aggregate ratings for each artefact at first, and then to each project objective. To enable this, a common evaluation scheme as well as a common set of defined user roles are provided. As the final evaluation phase, the current document also provides a comparison between the evaluation results obtained at the M24 mark with those obtained at the present M36 mark.

Furthermore, the document completes the circle by tracing back from the evaluation of the individual requirements to the project's main innovations and artefacts, as well as to the project's objectives, and shows to what degree each objective was fulfilled, using the evaluation of the use case providers. From this, a number of lessons learned from deploying the technical artefacts within the industrial organizations as well as best practices that can be adopted by other industry players from the RASEN tool-supported methodology are provided.

# Table of contents

# 1 Introduction

WP2 consisted of three tasks (cf. Figure 1) which are tightly linked among each other, resulting in the case study evaluation (Task 2.3), within which the present document is the final deliverable.

The first activity of WP2 consisted of identifying relevant case studies originating from different industrial sectors that were 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 undertook the analysis of the partner use cases and identified similarities and differences between existing processes in each organization.

Task 2.2 aimed 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 fell within the purview of Task 2.2. The scope of this task also included 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 was Task 2.3, grouped in two evaluation rounds: the first evaluation in year 2, subject of deliverable *D2.3.1 - Use case evaluation v.1* and the final evaluation that is detailed within the present document.

The present evaluation assesses the RASEN tool-supported methodology and technical implementation against the defined the use-case study requirements. For this, research and technology partners provided the results to the case study partners and assisted in implementing the new tools and methodologies within their processes. As the final evaluation, the current document also links back to the defined project objectives and success criteria, showing how they are covered by the technical work achieved within the project.

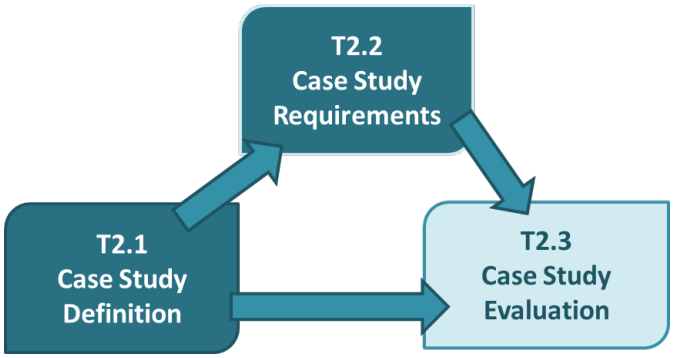


**Figure 1 – Overview and dependability of tasks within WP2**

# 2 Use Case Systems under Evaluation

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

## 2.1 Software AG

The software constituting Software AG's use case is called Command Central (CCE), 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.

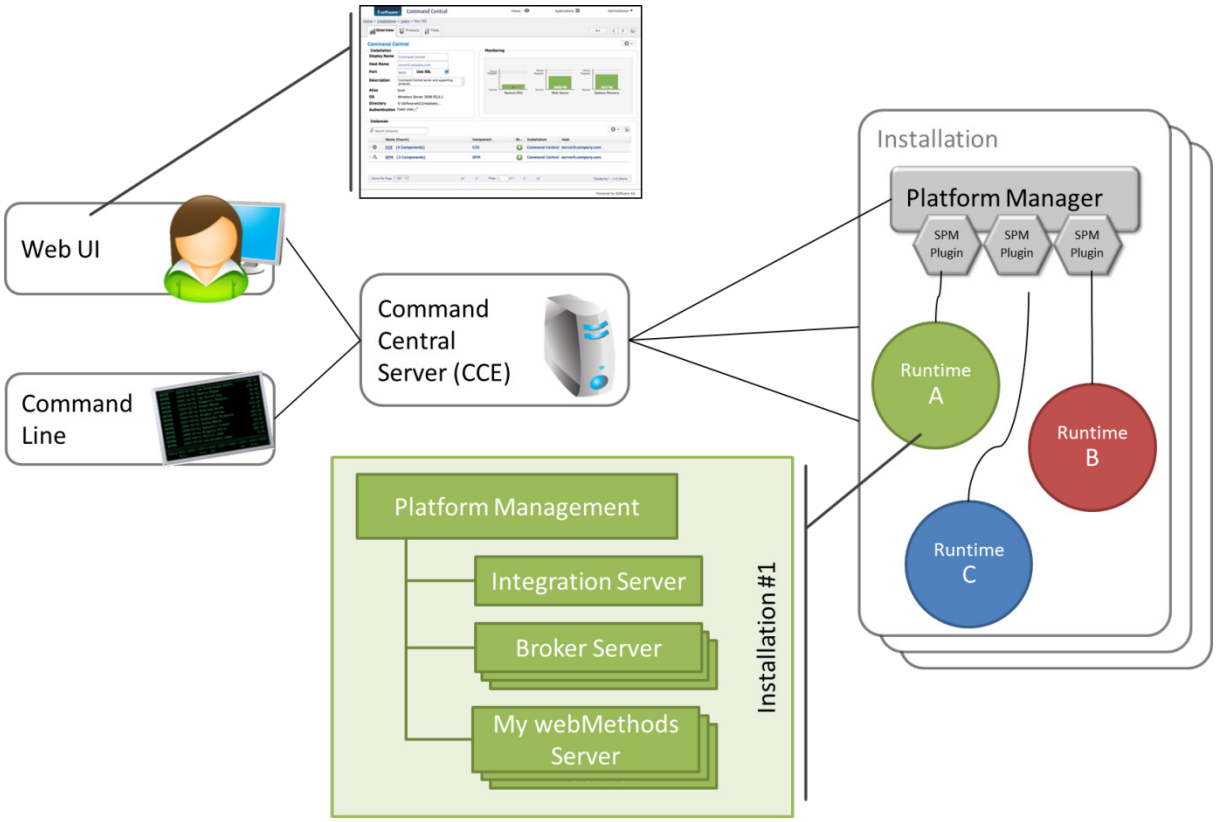


**Figure 2 – Command Central Architecture**

Command Central users can communicate with Command Central Server 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 2.

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 section.

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.

## 2.2 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 3.

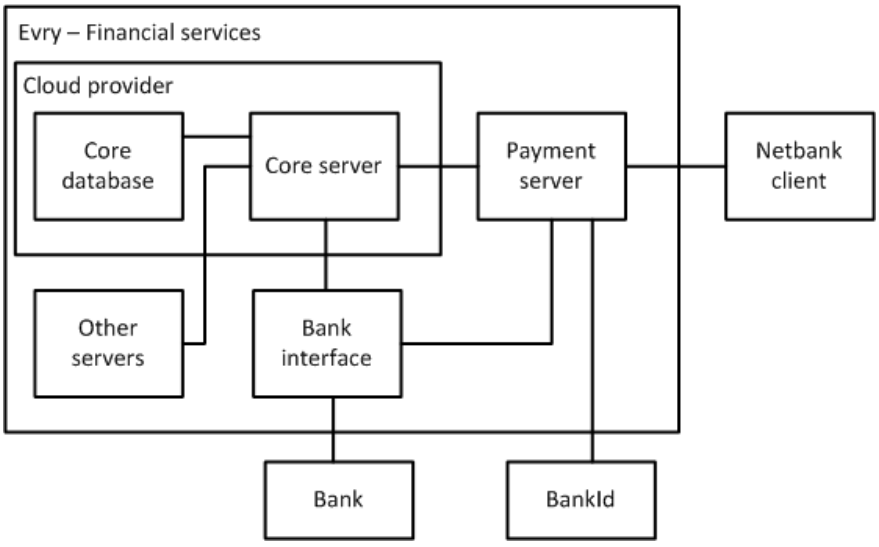


**Figure 3 – Architecture of EVRY Netbank system**

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.

## 2.3 Info World

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

- **Representativeness**. The chosen system had to be one of the more complex systems delivered by the company, so that it was representative of Info World's product stack. This ensures that successfully applying the RASEN process to it will be later transferrable to other systems developed within the company.
- **Requirements coverage**. The system had to 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 focused 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 included 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 4, 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)

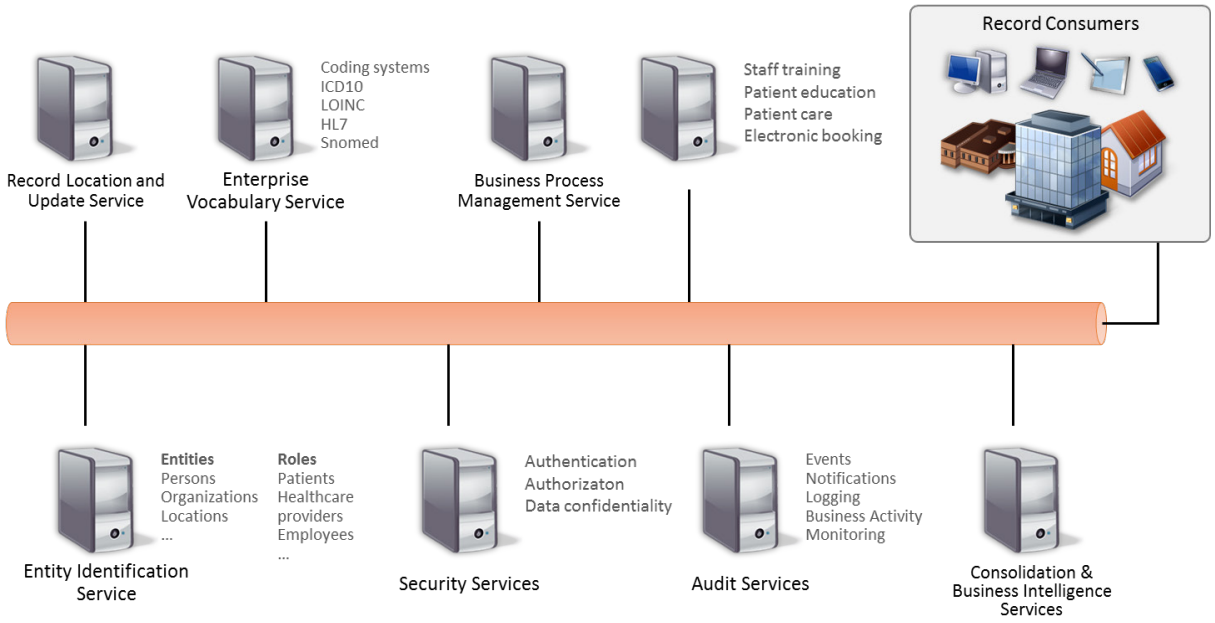


**Figure 4 – Medipedia software architecture**

# 3 Initial Assessment using the CRSTIP Scheme

The CRSTIP assessment scheme allows stakeholders to rank the security and risk assessment processes taking place within their organization using a simple, four-level scale that cover four key areas. The CRSTIP scheme was first introduced in the *D2.3.1 - Use case evaluation v.1* deliverable and was published in [1]. Complete information regarding the assessment scheme, key areas as well as levels within each area are available within Section 9.1 of the present document.

In the RASEN project, the CRSTIP scheme was used 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 post project dissemination and exploitation activities, which are detailed within deliverable *D6.1.3 - Periodic Standardization, Dissemination and Exploitation Plan v.3*. The following Sections detail the CRSTIP assessment of the project's three use cases.

## 3.1 Software AG

Figure 5 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.

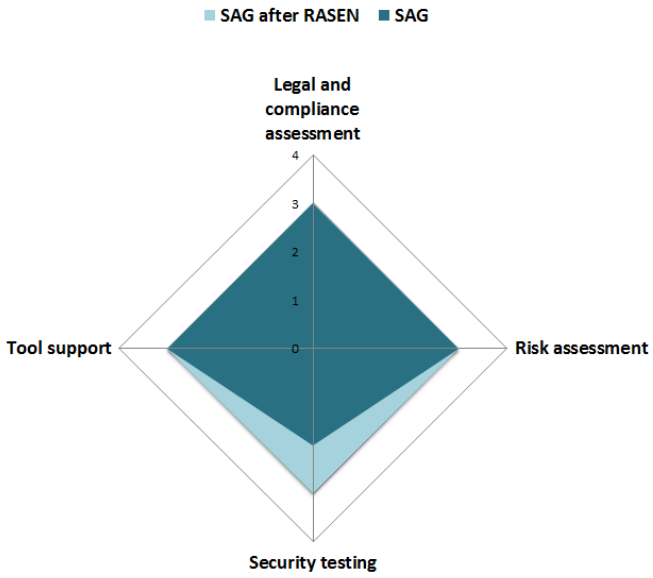


**Figure 5 – CRSTIP assessment of the SAG use case**

## 3.2 EVRY

Figure 6 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.

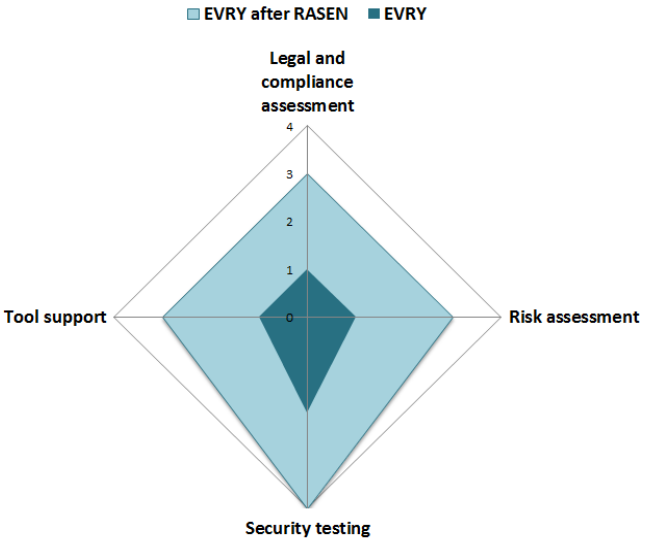


**Figure 6 - CRSTIP assessment of the EVRY use case**

## 3.3 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 7 showcases the CRSTIP evaluation for Info World. 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.

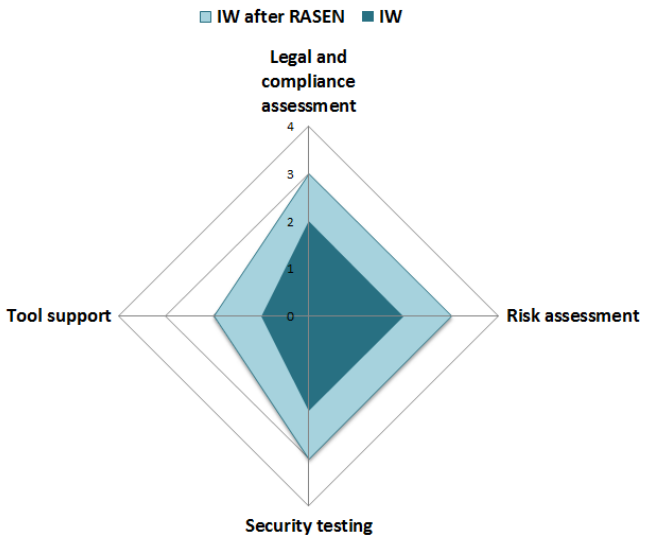


**Figure 7 – CRSTIP assessment of the IW use case**

# 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 at project completion. The current template is an updated version of the one used within the first evaluation at the M24 mark. Table 1 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.<br><br>**E.g.** REQ-SAG-F-010, REQ-SAG-F-010 |
| Requirement | Provides a textual description of the requirements that are evaluated using this template.<br><br>**E.g.** A methodology providing automated security risk assessment. |
| Objective | Provides the project objectives that are linked with the present requirements evaluation.<br><br>**E.g.** O5 |
| Description | A full description of this requirement, as seen from the use case provider's perspective is provided here.<br><br>**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.<br><br>**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.<br><br>**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.<br><br>**E.g.**<br>*F1*: The RASEN test procedure technique is more rigorous than the current test prioritization process. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 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 2.<br><br>**E.g.** Good (3) |
| Evaluation Phase 2 Result | This section details the evaluation results at the M24 mark.<br><br>**E.g.**<br>*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. |
| Involved Role(s) | Here the generic roles (c.f. Section "*5.2.5 Common roles across use case evaluation"*) involved in the evaluation process are listed. Expected roles are Business Analyst, Product Manager, Software Architect, Software Developer, Risk Manager, or Security Manager. |

**Table 1 – 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 |
| N/A | The requirement is not met. |

**Table 2 – Description of evaluation ratings**

# 5 Evaluation

## 5.1 Third Year Evaluation Process

The initial evaluation plan was created as part of Task 2.2. Figure 8 illustrates the RASEN timeline with regards to agreed-upon technical phases and milestones.

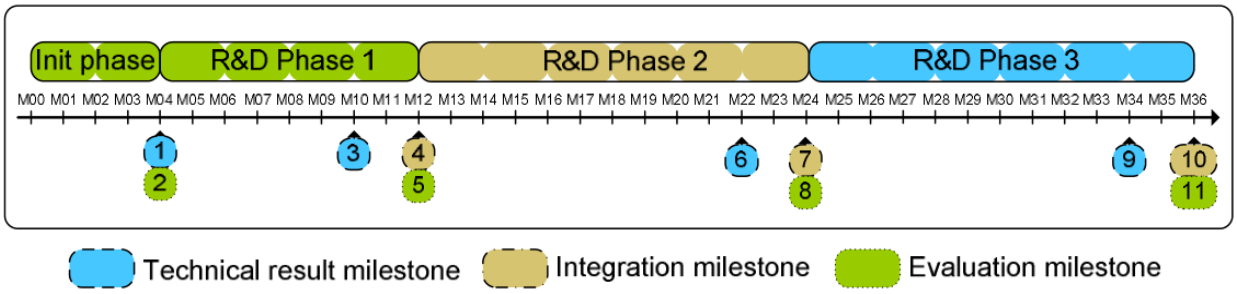


**Figure 8 – Phases, timeline and milestones**

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[1] 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 represented 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 11 is where the project currently stands. During R&D Phase 2, technical and scientific partners have collaborated 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.

The tasks within WP2 represent the middle column within Figure 9. 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*.

The first evaluation of the project's technical results was undertaken within task T2.3 and resulted in deliverable *D2.3.1 - Use case evaluation v.1*, within R&D Phase 2 of the project and consisted of the following steps:

**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

[1] Research and Development

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

As first outlined within deliverable *D2.2.1 – Use Case Requirements Definition*, the current evaluation is the phase 2 of the process. The current phase follows the first one and was 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 consisted of two stages:

- Evaluation Stage 1 – Research and technical partners delivered the RASEN tool-supported methodologies to the use case partners and will helped them implement some of the desired changes in their workflows. The retrospective time-frame of this stage included the first six months of Evaluation Phase 2, therefore the M24 – M30 period of the project implementation.
- Evaluation Stage 2 – Represented the final stage in evaluating the technical results of the project. As preparation of this stage, use case partners received the latest artifacts from the scientific and technical partners and used them with minimal support from the technology providers, where possible. The retrospective time-frame of this stage was represented by the last six months of the project implementation, namely M30 - M36.

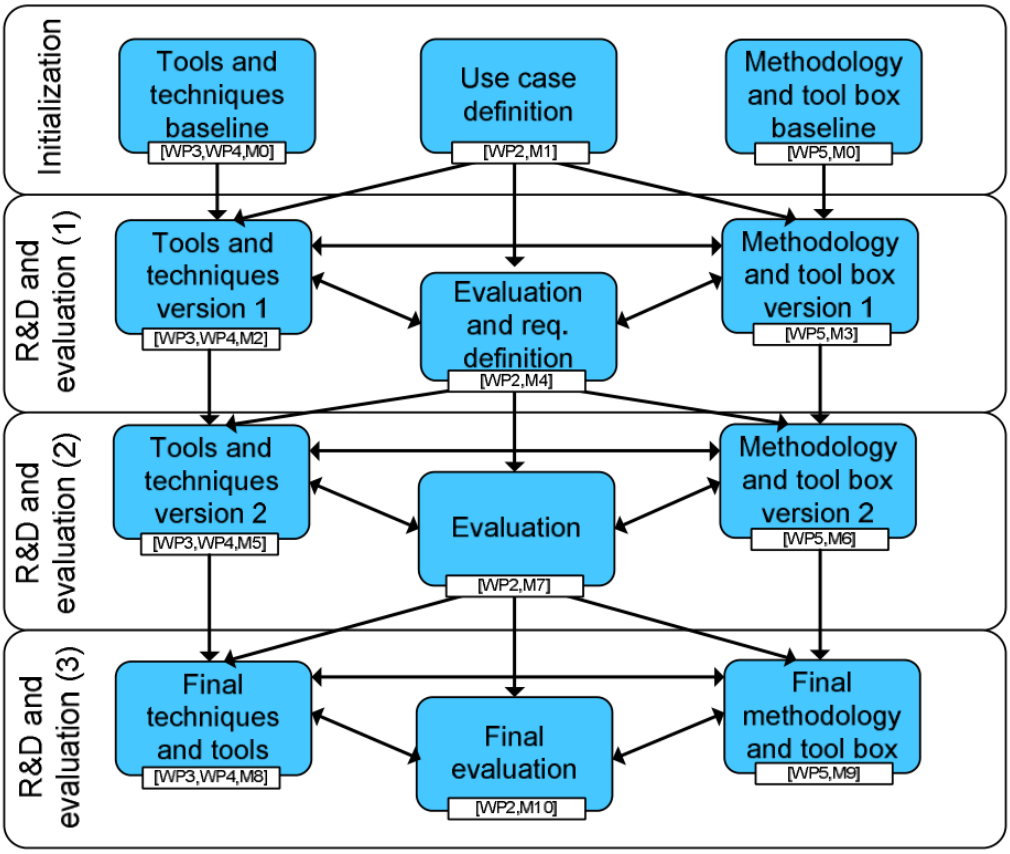


**Figure 9 – Relationship of technical results over time**

The present deliverable details the final evaluation of the project's technical artefacts using the industrial expertise and perspective of the use case partners. Their evaluation forms the basis for post-project exploitation as well as the continuation of work on the existing artefacts.

## 5.2 Evaluating technical innovation within RASEN

The purpose of this section is to present the project artefacts undergoing evaluation. We start by briefly presenting the main three innovations of the RASEN project, which is followed by the extraction

of relevant project artefacts and user roles. The project artefacts and roles are then used throughout the following section, which presents the results of the use case partners' evaluation.

### 5.2.1 Innovation 1: The PMVT approach for security pattern and model-based vulnerability testing

The Pattern-driven and Model-based Vulnerability Testing process (PMVT for short) was developed by the RASEN project for deriving test cases from risk assessment results, aiming to make interrelated and synergetic the activities of risk management and security testing. In the first step, a risk model is created using the CORAS method that identifies threat scenarios and potential vulnerabilities. The risk model is then used for the identification and prioritization of appropriate security test patterns. Based on the security test patterns, test cases are generated by combining information from the risk model, security test patterns, a test model and test generation techniques. The latter are composed of test purposes generated from Smartesting CertifyIt and fuzzing techniques from Fraunhofer FOKUS's fuzzing library Fuzzino. In the last step, test scripts are generated, compiled and executed against the application under test. Hence, the PMVT approach integrates the tools of the project partners: CORAS from SINTEF ICT (to address risk assessment), CertifyIt from Smartesting (to perform risk and model-based test generation) and Fuzzino from Fraunhofer FOKUS (to apply data fuzzing techniques).

***RASEN artefact under evaluation resulting from this innovation:***

***A3***: The RASEN technique for security test automation

### 5.2.2 Innovation 2: The RACOMAT tool – risk assessment combined with automated testing

The RACOMAT tool allows users to combine component based security risk assessment with security testing. Testing can be integrated seamlessly into the incident simulations the tool uses for its compositional risk analysis. Taking benefit from libraries containing risk analysis artefacts like attack patterns and of libraries containing testing artefacts like security test patterns, the RACOMAT tool offers a high level of reusability. Using the assistance the tool offers, many steps of the analytical RACOMAT process form risk modelling to testing and updating the risk picture based on test results can be done automatically.

RACOMAT can use different kinds of risk assessment methods, including fault tree analysis (FTA), event tree analysis (ETA) and the CORAS method, as proposed within RASEN. In general, RACOMAT supports component based risk analysis and compositionality. The tool uses intuitive risk graphs to represent and to visualize the risk picture. For enabling automation of risk based testing, the risk assessment must be made on a low level. The RACOMAT tool allows risk analysts to model close relations to parts and components of the systems that are analyzed. Therefore, the RACOMAT tool introduces the concept of threat interfaces representing entire components and threat ports representing parts of the input / output interface. In order to reduce the manual effort of low level system analysis, RACOMAT integrates techniques for analyzing components automatically. Given (X)HTML pages, source code, compiled programs or listening to common network protocols, it tries to identify the public interfaces of any components and especially the functions as well as ports that could be used for interaction with other components or users. Thereby, an initial system model can be generated without requiring a lot of manual actions.

***RASEN artefact under evaluation resulting from this innovation:***

***A6***: The RASEN tool-supported method for risk assessment combined with automated testing

### 5.2.3 Innovation 3: The RASEN method for risk-based security testing and legal compliance assessment

The RASEN method for risk-based security testing and legal compliance assessment is derived from ISO 31000 and slightly extended to highlight the identification and evaluation of compliance or security issues as one of the major tasks that need to be carefully aligned with typical risk assessment activities.

The method starts with establishing the context and supports additional activities meant to set up and support its management. The process is generic and can be instantiated towards particular instances of integration. We consider three such integrations.

1. A test-based risk assessment starts like a typical risk assessment process and uses test results to guide and improve the risk assessment. Security testing is used to confirm the presence of potential vulnerabilities identified during risk assessment, or to detect new vulnerabilities that have not been identified during risk assessment. This in turn provides a basis for risk values to be verified and adjusted based of tangible test result measurements.

2. A risk-based testing process will start like a typical testing process and uses risk assessment results to guide and focus the testing. Such a process involves identifying the areas of risk within the target's business processes and building and prioritizing the testing program around these risks. In this setting risks help focusing the testing resources on the areas that are most likely to cause concern or supporting the selection of test techniques dedicated to already identified threat scenarios.

3. A risk-based compliance assessment process will start with the identification of compliance issues, and use risk assessment to identify, estimate, and evaluate compliance related risks.

***RASEN artefacts under evaluation resulting from this innovation:***

***A1***: The RASEN tool-supported method for risk-based security testing

***A2***: The RASEN method for compliance risk assessment

***A4***: The RASEN method for compositional security risk assessment

***A5***: The RASEN tool-supported method for test-based security risk assessment and test result aggregation

### 5.2.4 Map of RASEN Artefacts under Evaluation

Table 3 below illustrates the relation between the main innovations of the project, the resulting artefacts and the use cases where those artefacts are evaluated. To ensure a thorough evaluation of all artefacts resulting from the project, each one was evaluated within at least one of the industrial use cases.

| RASEN Artefact | | Evaluation Use Case |
|---|---|---|
| **Innovation 1**: The PMVT approach for security pattern and model-based vulnerability testing | | |
| **A3** | The RASEN technique for security test automation | SAG Command Central, EVRY NetBank and IW Medipedia |
| **Innovation 2**: The RACOMAT tool – risk assessment combined with automated testing | | |
| **A6** | The RASEN tool-supported method for risk assessment combined with automated testing | SAG Command Central |
| **Innovation 3**: The RASEN method for risk-based security testing and legal compliance assessment | | |
| **A1** | The RASEN tool-supported method for risk-based security testing | SAG Command Central, EVRY NetBank and IW Medipedia |
| **A2** | The RASEN method for compliance risk assessment | EVRY NetBank and IW Medipedia |
| **A4** | The RASEN method for compositional security risk assessment | SAG Command Central and IW Medipedia |
| **A5** | The RASEN tool-supported method for test-based security risk assessment and test result aggregation | SAG Command Central, EVRY NetBank and IW Medipedia |

**Table 3 - Artefact evaluation within the use cases**

### 5.2.5 Common roles across use case evaluation

In this section we discuss the all key roles involved in both, the actual software development and the evaluation process as shown in Figure 10 and Figure 11. The overall process consists for all three RASEN pilots of five consecutive steps, namely the Risk Assessment, the Security Test Preparation, the Test Execution, the Security Risk Integration, and the Risk Valuation & Migration.

In the following key roles are generalized consisting of Business Analyst, the Product Manager, the Software Architect and Developer as well as the Risk and Security Manager:.

- **Business Analyst** plays an important role at the beginning and at the end of the overall process. He or she fills the gap between customers and Product Managers. The Business Analyst's main task is to specify the actual software demands in terms of requirements. Additionally requirements are collected in workshops with customers and are merged and aggregated with the prevailing customer expectations and eventually translated into so called features requests.
- **Product Manager's** job is to prioritize the features given from the Business Analyst and align them with the overall product development. He or she reflects the needs of the end users over the complete process and also incorporates market trends, technological advances and the company vision. He monitors the integration with the product into the company's software portfolio and ensures integration. Business Analysts also defines the acceptance criteria in the phase of Security Testing and keep track of general quality assurance issues.
- **Software Architect** defines a technical description out of the abstract features Product Manager and Business Analysts provide. In the Security Test Preparation step the Software Architect plays a leading role by supervising the work of the Software Developers. In the Test Execution phase a Software Architect monitors the alignment with the other software/libraries and ensures the overall compliance to standards as coding guidelines.
- **Software Developers** implement the requests coming from the Software Architect where tasks are usually divided into the implementation of new features coming from feature requests, bug fixes, or change requests to reflect customer's demands and deliver highly customized software,
- **Risk Managers** give recommendations concerning the risk level of a feature or a general changes in the software product. The Risk Manager plays a key role in the Security Risk Integration and the Risk Valuation & Mitigation step. He or she works highly interlaced with the Security Manager, having an eye on the overall risk on software component level, on product level, as well as on company level.
- **Legal Counsel** contributes to the legal risk picture of the company's products. Their job is to ensure that the company's product stack is legally compliant as well as to provide the required measures that are required to be taken to ensure products meet current as well as future legal requirements. This is a complex task as it braces both legal work as well as technical effort. It is also crucially important for many businesses that operate using valuable customer data that must be protected and kept confidential. Given the current time frame – before major changes to data protection legislation at EU level, this role is particularly important.
- **Security Manager** defines standards related to security issues that affect the company. He ensures compliance to security standards, defined especially coding guidelines, and best coding/security practices that Software Developers and Software Architects have to obey. A security manager plays a leading role in the overall process by monitoring security standards, activity with a high security impact, and gives recommendations of how to deal with security related development on company level.

## 5.3 Evaluation from use case partners

### 5.3.1 Software AG

#### 5.3.1.1 Evaluation process

The evaluation of the use case scenario was organized as a sequence of evaluation steps (cf. Figure 10) 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 Architect as a representative from the product development (Command Central Product Development), a Security Manager overviewing and ensuring the compliance to company security standards, and the RASEN project development team in charge of the RASEN trainings, knowledgeable in the RASEN methodology and the function of the ARIS and RACOMAT toolbox. As a result of the workshops 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 one case a Software Developer has been consulted as only he was aware of the current implementation and technical details.

In "*Security Test Preparation*" the RASEN project representatives at Software AG conducted an assessment with the Security Manager and Risk Manager 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 considers 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 is worth to mention, that in the current implementation there is no fully automated interface between – which is only a minor development task – and hence the exchange formats were manually copied, i.e. from the export of the ARIS tools to RACOMAT and vice versa. However we assume that a working and fully automated exchange exists, as the import and respective exports in ARIS and RACOMAT are working.

With the test results from the previous phase, the "*Security Risk Integration*" phase receives test results and converts them into an appropriate format, suitable for integration into the ARIS RASEN framework. Apart from this, additional artefacts are generated by RACOMAT along with the security test describing test coverage, risks, test conditions, a risk aggregation up to product level and a risk tree. These are also exported as they contain valuable information regarding the system under test and the applied testing strategy to be considered further in the overall risk assessment. In evaluation step, all confirmed weaknesses on actual product become visible. This assessment has been conducted together with a Security Manager, the developers of the ARIS RASEN framework, and the developers from Fraunhofer FOKUS in charge of the RACOMAT tool.

Eventually, the evaluation of the "*Risk Evaluation & Mitigation*" step highlights 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. The risk picture is created by using both, direct inputs from RACOMAT to compute a simple risk aggregation along the product tree based on ARIS's internal aggregation function as well as a more sophisticated aggregation already computed by RACOMAT.

An overview of artefacts used within the Software AG pilot are – which have been subject to the evaluation process – are already denoted in Table 3.

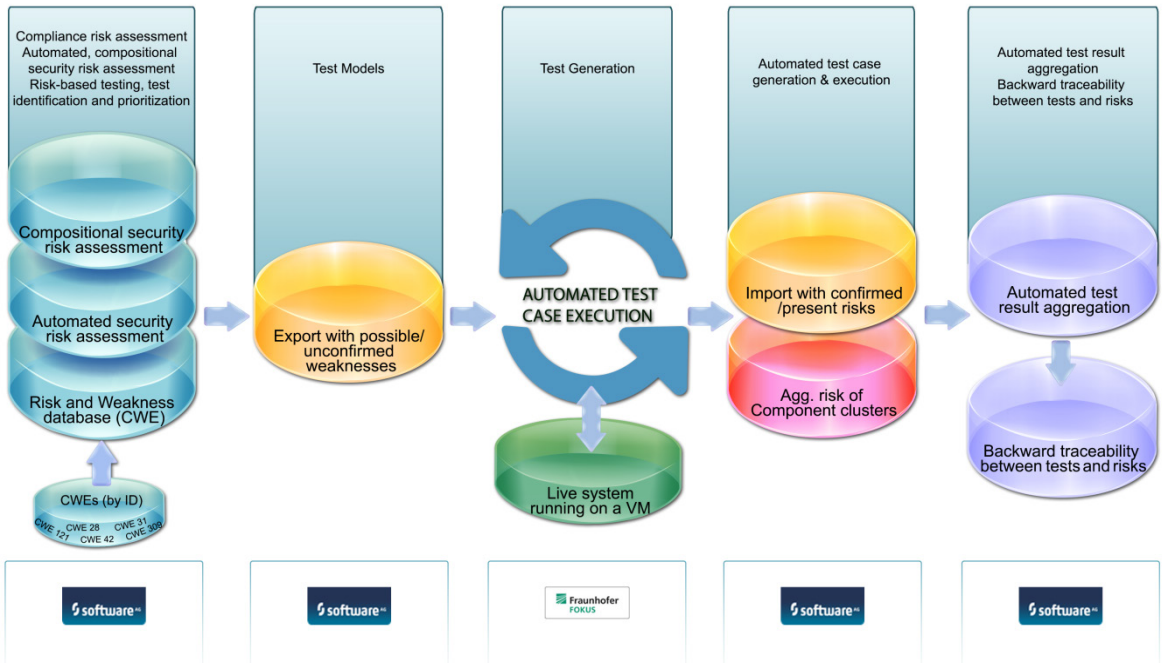


**Figure 10 – RASEN tool chain in the Software AG use case scenario**

The results of the Software AG evaluation are discussed in the following subsection.

### 5.3.1.2 Evaluation Results

| 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.<br>Additionally, the following may be relevant here:<br>SC1.1: The approach must ensure traceability between risks and test results.<br>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**:<br><br>*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.<br>*E2*: The methodology should be clear to the person performing evaluation. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 Result | *E1* An architectural analysis was been accomplished by a Product Manager and Software Architect after a short introduction to the RASEN methodology through the RASEN expects. The analysis has been complete and all the modelling elements were available.<br><br>*E2*: A short training session provided sufficient knowledge about the modeling system, the new web-based user interface, and the RASEN methodology such that the final system could be modelled. Through the support of reports and macros (wizards) all major steps are automated that enable the creation of a new product, new components, new generic components, and support in creating a vignette. Integration with other security tools is still considered as an asset but it turned out that for the actual assessment these are not essential. |
| Involved Role(s) | Product Manager, Software Architect, Software Developer |

**Table 4 – 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.<br>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.<br>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**:<br><br>*E1*: (SC-A2) Does the project provide a set of tools for automated risk analysis?<br><br>*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 Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Excellent |
| Evaluation Phase 2 Result | *E1*: The RASEN project provides a toolset consisting of RASEN Extension based on the ARIS Business Architect and the RACOMAT tool developed by Fraunhofer FOKUS. Both tools are integrated and models can be exported from ARIS to RACOMAT and in turn test results can be exported from RACOMAT to ARIS Business Architect.<br><br>*E2*: The above mentioned toolset has been applied to a Software AG product called Command Central (CCE). During the design of the ARIS Extension focus has been on effectiveness and usability e.g., by introducing generic component types. Using these generic types a new product model can be instantiated quickly, using existing and frequently used models which speeds up the modeling process. Due to training effect and the easy to apply methodology we believe that also other components can be quickly modelled and a risk assessment becomes feasible with moderate efforts. |
| Involved Role(s) | Security Manager, Risk Manager |

**Table 5 – 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**: |

| | |
|---|---|
| | *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 Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 Result | *E1*: The current RASEN methodology has concepts of compositions of risk ratings from the component to the product level. Considering the aggregation function currently two solutions exist:<br>• An aggregation which is purely based on the known weaknesses is feasible although the aggregation function is simple and limited in its meaningfulness. Aggregation functions that we considered directly in ARIS are arithmetic mean, a probabilistic distribution of risks, and the sum.<br>• An aggregation which is based on the risk tree from RACOMAT gives meaningful results which are imported in ARIS. These results are reliable as they are based on the risk graph and in addition provide testing details like testing coverage. |
| Involved Role(s) | Risk Manager, Security Expert |

**Table 6 – 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**:<br><br>*E1*: Perform a risk evaluation through the proposed methods with the supplied tools.<br>*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 Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Fair |

| | |
|---|---|
| Evaluation Phase 2 Result | *E1*: The required method is implemented in RACOMAT from Fraunhofer FOKUS, delivering the composition and aggregation functionality required when computing the composite assessment.<br><br>*E2*: The testing can be automated although some issues naturally exist when using testing on web interfaces. Those interfaces prevent the derivation of data structures or variable types used. Hence the test automation requires minor manual intervention as an automation is per se not applicable here.<br>In addition, when facing login screen the testing system cannot derive valid credentials to obtain access. Hence it is impossible to automate the testing of a backend if it is protected by login screens. |
| Involved Role(s) | Security Manager, Software Architect |

**Table 7 – 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?<br>SC-A6: do these tools automatically generate suitable and usable test cases? |
| Evaluation Criterion | Evaluation criteria related to artifact **A4**:<br><br>*E1*: There must be a tool to support the generation of test cases<br>*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 Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Excellent |
| Evaluation Phase 2 Result | *E1*: Through RACOMAT we obtained a solution which delivers the security test automation for Software AG. Based on this the automated generation of test cases is feasible. Tools are available and running<br><br>*E2*: Applying security testing to our Command Central (CCE) instance obtains test cases that substantially lower the risk. |
| Involved Role(s) | Software Engineer, Software Architect |

**Table 8 – 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).<br>Frankly, that is fine but we still want to have the coverage as well, not limited to finding some more vulnerabilities.<br>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**:<br><br>*E1*: The tool must provide the coverage data that will be evaluated by an expert versus the provided risk assessment.<br>*E2*: The coverage should correlate to the level of risk as assessed by the methods of this project. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 Result | *E1*: The applied tool chain provides apart from the test results test coverage.<br><br>*E2*: This test coverage seemed on first glance being adequate for the considered vulnerability. However as the present evaluation does only consider a small software product from Software AG we want to further analysis the level of risk and the coverage criteria and compare it with results from our code analysis tools. |
| Involved Role(s) | Security Manager |

**Table 9 – 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 | 5 |

| | |
|---|---|
| Satisfaction | |
| 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**:<br><br>The evaluation basically boils down to:<br>*E1*: Are there tools to run the test cases?<br>*E2*: Do these tools function automatically with the RASEN generated test cases?<br>*E3*: Are we able to run them against a large software system? |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 Result | *E1*: RACOMAT comes with an internal code generation which is based on test pattern. As most attacks can be executed by only a few test cases, the current version of RACOMAT does not have test patterns implemented for all 600 existing CWEs, but only for the most often used ones.<br><br>*E2*: The test cases from RASEN did automatically work. When using test cases based on web interfaces we denoted that login credentials must be provided by the user as all parameters handled as strings which cannot be derived by the software.<br><br>*E3*: The security tests were executed on Command Central (CCE) which is a considerable small software product. However experiences made are promising that the solution will also work with large software product albeit additional issues appear due to complexity. As such we expect that very large risk graphs (which are displayed over 20 monitor screens) are difficult to check. However it is beyond the RASEN project to account for this. |
| Involved Role(s) | Security Manager, Software Architect |

**Table 10 – 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**:<br>*E1*: Executing the tool chain, testing results will be imported, resulting in an update of the risk analysis.<br>*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.<br>*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 Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Excellent |
| Evaluation Phase 2 Result | *E1*: The supported tool-chain and implemented interfaces allow for security testing, import of the test results along with detailed testing details and are used to update the risk analysis.<br><br>*E2*: This criterion is a simple modification of the input and easy to be fulfilled with fully functional test execution integration. So when following the evaluation criterion twice the same results are obtained by the testing framework.<br><br>*E4*: A sophisticated aggregation function exists in two flavors, one it is implemented in the RASEN Extension of ARIS which account for simple risk aggregations and a more sophisticated version is implemented in RACOMAT which is based on the risk graph. |
| Involved Role(s) | Security Manager |

**Table 11 – 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**:<br><br>*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 | *E1*: Yes – succeeds. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Fair |
| Evaluation Phase 2 Result | *E1*: The import of externally available sources containing security risk information is only possible if the external component has been previously modelled and tested with RACOMAT. Otherwise the risk model is unknown and there can be no assumption made about the security of externally available sources. |
| Involved Role(s) | Security Manager |

**Table 12 – 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**:<br><br>*E2*: After a complete risk assessment is done, we change something at the bottom of the pile, 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 Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 Result | *E1*: In the current implementation it is only possible to run a risk rating and risk aggregation based on security test results. Without such a testing, when changing the risk ratings on components only a simplistic aggregation function can be applied which does not consider the risk graph.<br><br>*E2*: When using the testing tool chain the resulting tests and confirmed risks are reflected. Yet the implementation does not account for making changes w.r.t. testing in the component tree visible, .i.e. the delta of |

| | |
|---|---|
| | previous and current risks are not displayed. This could be overcome by manually taking screen-shots and comparing previous and current results. On the other hand this becomes too complex when considering large scale networked systems but such a feature could be implemented. |
| Involved Role(s) | Security Manager, Risk Manager |

**Table 13 – 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.<br>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:**<br><br>*E1*: The tools are available for integration<br>*E2*: The tools can be integrated with the development process at a reasonable cost for automated analysis, the success is guaranteed. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 Result | *E1*: The RASEN Extension of ARIS along with the RACOMAT tool from Fraunhofer FOKUS are such tools which are integrated and allow compositional security risk analysis with an economically level of investment.<br><br>*E2*: The tools are – as of today – integrated using a common import and export format which needs to be manually triggered. |
| Involved Role(s) | Security Manager, Software Architect |

**Table 14 – 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**:<br><br>*E1*: There are tools – in particular the testing tool chain – available<br>*E2*: Tools operate in linear time when tested on our products vs. the LOC and number of components? |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Fair |
| Evaluation Phase 2 Result | *E1*: The RACOMAT tool is available and integrated into the testing chain within the RASEN scenario.<br><br>*E2*: Tools operate in linear time. However, then testing large scale networked systems there are further tests needed on the system under tests to ensure that the system is still functional and has not been killed through the tests. Due to the complexity of software products, consisting of several hundred thousand lines of code, it is advisable to further integrate code analysis tools. This helps to tremendously reduce the number of considered test cases and leads to a more focused testing strategy. This has however not part of the RASEN project. |
| Involved Role(s) | Software Architect, Security Manager, Risk Manager |

**Table 15 – Evaluation for requirement REQ-SAG-N-030**

## 5.3.2 EVRY

### 5.3.2.1 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
- **A5**: The RASEN tool-supported method for test-based security risk assessment and test result aggregation

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 several workshops between EVRY, SINTEF and UiO, in collaboration with security testers as well as legal counsel and risk managers.

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

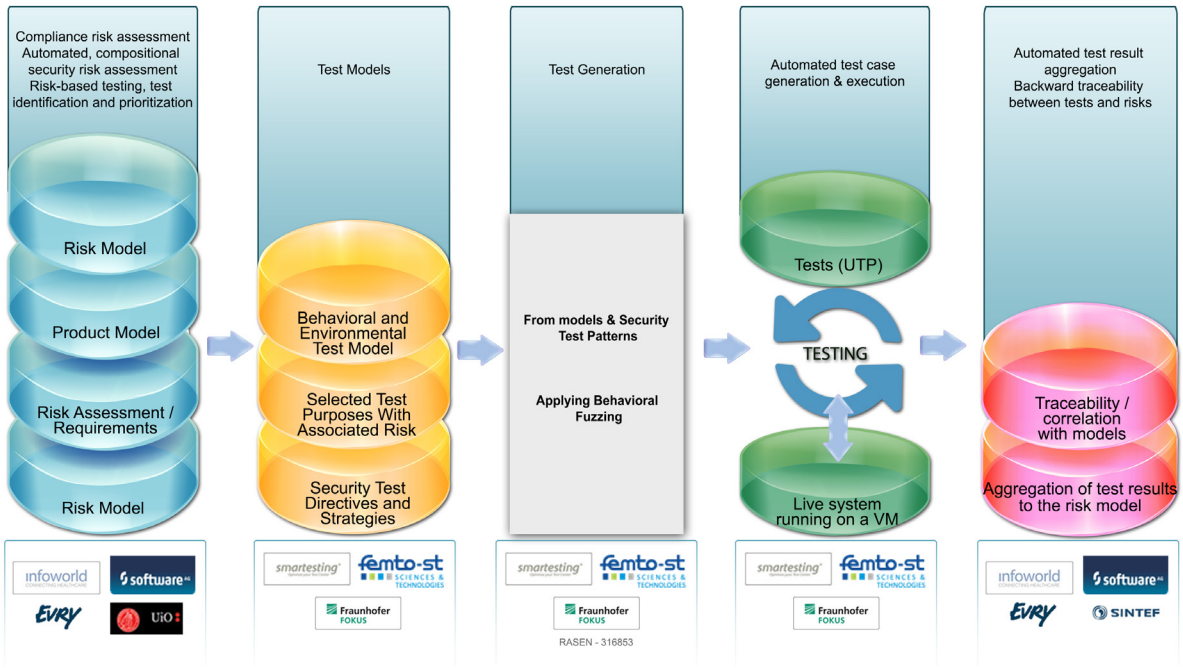


**Figure 11 - RASEN tool chain in the EVRY and Info World use cases**

### 5.3.2.2 Evaluation Results

| 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.<br><br>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**:<br>*E1*: All relevant security test cases can be seen as a refinement of a test |

| | |
|---|---|
| | procedure derived from CAPEC according to artifact **A1**.<br>*E2***:** The level of abstraction of the CAPEC derived risk model (by artifact **A1**) is appropriate for security test identification.<br>*E3*: The prioritization of the test procedures generated by artifact **A1** is according to intuition.<br>*E4*: The risk visualization of security test related risks by **A1** is according to intuition.<br>*E5*: The likelihoods are defined appropriately by **A1**<br>*E6*: Estimating attack success likelihood according to **A1** is easy.<br>*E7*: Estimating technical impact likelihood according to **A1** is easy.<br>*E8*: The effort spent on test prioritization according to **A1** will be saved in the testing phase.<br><br>Evaluation criteria related improvement of EVRY's testing process through **A1**:<br>*F1*: The RASEN test procedure technique (**A1**) is more rigorous than the current EVRY test prioritization process.<br>*F2:* Artifact **A1** helps prioritize test procedures more accurately than EVRY's current process for doing this.<br>*F3:* Test prioritization according to artifact **A1** may help save time during the testing phase of the EVRY testing process.<br>*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 2** [M36] | |
| Evaluation Phase 2 Rating | Fair |
| Evaluation Phase 2 Result | *E1*: Almost all security test cases we are aware of can be captured by a CAPEC attack pattern.<br><br>*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.<br><br>*E3*: The experience of the participants of the case study, is that it is fairly easy to understand the intuition behind the test prioritization technique.<br><br>*E4*: The visualization of the risk values of the risk picture was intuitive and a helpful support to the test prioritization.<br><br>*E5*: The participants of the case study found the process of likelihood definition/estimation to be easy to understand.<br><br>*E6, E7*: Estimating attack success likelihood according to A1 is easy.<br>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. However, this depends on your knowledge of the system under test, and the process may require external input (although it was not necessary in the case study).<br><br>*E8*: The effort spent on test prioritization according to A1 will be saved in the testing phase. Time could be saved since this approach can be used to filter/ignore irrelevant tests. In any case, the time spent in the test prioritization phase i very low compared to the time spent on the testing. This suggests that the criterion is fulfilled. |

| | |
|---|---|
| | *F1*: The RASEN test procedure technique (A1) is more rigorous than the current EVRY test prioritization process. Currently, there is no structured method for doing the prioritization. Artefact A1 would therefore provide a more structured prioritization approach.<br><br>*F2*: Currently, the priority of tests are not documented either quantatively or qualitatively. 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.<br><br>*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.<br><br>*F4*: The security requirements are perhaps more abstract than the test procedures generated from artefact A1. All in all however, they are quite similar to the test procedures (minus the priority values). They would therefore be a good starting point for test design and implementation<br><br>*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. In addition, the test procedure prioritization technique as well as the risk visualization was found to be intuitive and understandable. It is difficult to empirically compare the effectiveness of the current information prioritization with the structure prioritization technique of artefact **A1**. Therefore we overall evaluation rating is *Fair*. |

**Table 16 – 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**:<br><br>Evaluation criteria related to improvement of EVRY's security testing process through **A3**<br>*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.<br>*F3*: Adapting artifact **A3** will save time during the test execution phase of the EVRY security testing process. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | *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<br><br>*F2*: This is probably true, since the artifact A3 would lead to the execution of more tests that is currently possible using manual techniques. However, the criterion has not been through fully validated empirically due to technical difficulties encountered during the evaluation.<br><br>*F3*: This criterion is true provided that the artifact **A3** has been properly configured/set up prior to the test execution phase. |
| Evaluation Phase 2 Result | Fair |

**Table 17 – 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<br>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 **A2** and **A5**:<br>*E1*: The test measurements are adequate for capturing the test results, i.e. the measurements capture everything that is needed.<br>*E2*: Mapping test results to test measurements requires little effort.<br>*E3*: The measurements are adequate for aggregation to the risk model<br>*E4*: Mapping test measurements to the risk model is easy<br>*E5*: The risk matrix provides is a good way of providing input to decision makers.<br>*E6*: The costs of using the method (of artifact **A2**) is, in the long run, lower than the value of the benefits from its use.<br><br>Evaluation criteria related to improvement of EVRY's compliance process through **A2**:<br>*F1*: The method of artifact **A2** provides an increased level of confidence on the compliance of the organization, compared to EVRYS current method.<br>*F2*: The method of artifact **A2** provides better input to decision making than EVRYS current method.<br>*F3*: The technique for compliance risk identification (artifact **A2**) enables a |

| | |
|---|---|
| | better structuring in identifying compliance risks than EVRYS current method. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Fair |
| Evaluation Phase 2 Result | **Artefact A2**<br>*F1*: **A2** contributes to mitigating subjectivity in making compliance decisions. By providing a structure and similar criteria to assess compliance risk, the method is considered to introduce some level of objectivity in assessing consequences.<br><br>F2: The creation of generic risk models facilitates reusability of results and thus contributes to long term cost benefits. The fact that such risks are based-on the requirements mean that there is no need to start from scratch every time there are changes in an organization or system.<br><br>F3: By decomposing compliance norms into different elements through the natural language pattern and structuring these elements in a template, the RASEN method simplifies the transition from normative statements (obligations or prohibitions) to the graphical risk models. It also facilitates a potential for future automated model.<br><br>F3: The visualization of compliance risks in CORAS stimulates and focuses discussions during the risk assessment. The focus is achieved by limiting the discussion to a specific generic threat at a time and providing a clear guidance in terms of the relevant inputs and outputs of each step.<br><br>**Artefact A5**<br>*E1*: Yes, based on these measurements, the new likelihoods of the risk model can be automatically calculated.<br><br>*E2*: Yes, most of the process is automated. The only thing the user needs to do is to estimate the measurements: "Test vulnerability discovery likelihood" and "likelihood of false positive" as documented in deliverable D3.3.2 and D4.3.3.<br><br>*E3*: Yes, bases on the measurements, the new likelihoods of the risk model can be automatically calculated as documented in deliverable D3.2.3.<br><br>*E4*:. In some cases, it can be difficult to estimate the measurements that are needed for automated aggregation.<br>However, it is possible to obtain the required measurements based on historical data and statistics.<br><br>*E5*: The decision makers are familiar with risk matrices already. They should therefore easily be able to understand the risk matrix generated on the basis of the test results. |

**Table 18 – Evaluation for requirement REQ-EVRY-F-030**

### 5.3.3 Info World

#### 5.3.3.1 Evaluation process

The evaluation of the Info World case study was conducted with Info World on one side, and SINTEF, Smartesting, UFC, FOKUS and UiO on the other. The research and technical partners' have identified the relevant RASEN artifacts that can be evaluated using the Info World use case, as illustrated within Table 25.

Within its use case, Info World evaluated the RASEN tool chain that is illustrated within Figure 11, also common to the EVRY use case. The evaluation was undertaken as a series of activities that had representation from both RASEN scientific and technology providing partners as well as Info World representatives. Each evaluation activity focused on the evaluation of one or several pieces of the tool chain shown within Figure 11, and they are as follows.

The RASEN methodology for *Compositional Security Risk Assessment* was evaluated using the Medipedia eHealth portal as a system under test. The first step consisted of modelling relevant risks for Medipedia in the CORAS tool developed by SINTEF. Then, using the input from Medipedia Developers and Product Manager the risk model was fine tuned to provide accurate information in the form of a prioritized list of relevant risks taken from the CAPEC database. When compared with Info World's ad-hoc approach to risk assessment, the RASEN methodology facilitates a comprehensive approach by using well-known and updated software vulnerability repositories such as the CAPEC. Furthermore, a prioritized list of risks that are considered relevant for the system under test allows selection and prioritization for existing security test cases. The final year of the project also saw the development of additional tooling that facilitates the automation of these activities. However, while the RASEN methodology can provide the next step in ensuring the security of systems such as Medipedia, what Info World's software team found that in their state tooling was not mature enough for industrial deployment within a complex development process.

The next step of the evaluation concerned the automated generation, execution and interpretation of security test cases, which was achieved in collaboration with partners Smartesting, UFC and FOKUS. The first step was to model a section of the Medipedia platform using an in-house developed DSL. The obtained language, together with test patterns and selected test inputs allowed for the generation of abstract test cases for the previously modelled section of the Medipedia system. These were exported to JUnit test cases that were fuzzed using FOKUS's Fuzzino library. The recorded execution time for all tests shown within Figure 12 was of approximately 45 minutes [2], making it suitable for integration into the overnight build processes.

As shown within Figure 12, a multi-step vulnerability was found in the Medipedia forum. Indeed, on the "*New Forum Post as a Visitor*" page, the name field is vulnerable to XSS because the value of the field is used as output on the "*Display forum topics*" page without proper sanitation. To be effectively detected, this multi-step XSS vulnerability requires a complex verdict assignment process, which is built-it in the PMVT process, and not easy to find using the current practice based on scanner inspections [2].

The second vulnerability detected is a single step XSS in the same forum component. Again, on the "*New Forum Post as a Visitor*" page, the content field is vulnerable to XSS because the value of the field is rendered back raw on the "*Display Post*" page [2]. Initial experimentation showed that some of the executed tests misreported a vulnerability or missed one. For instance, 24 tests for XSS revealed to be false negatives. This is not alarming since these false negatives came in fact from attack vectors whose purpose is to detect XSS in very specific configurations, which was not the case in Medipedia [2]. In addition, 2 tests targeting CSRF attacks came back positive, but we were not able to reproduce the attacks manually [2].

The technical side of test generation and execution was well understood within Info World, as well as the technical limitations of the process. After the initial experimentation phase, the Medipedia development team (software developers, testers, product manager) received all the tooling required for extending the Medipedia model to cover a larger part of the portal application as well as to deploy it within the organization as part of post-project exploitation.

| Vulnerability | Abstract Test Cases | Attack Vectors | Executable Test Cases | Detected Vulnerabilities | False Positives | False Negatives |
|---|---|---|---|---|---|---|
| SQL Injection | 47 | 10 | 470 | 0 | 0 | 0 |
| Single Step XSS | 18 | 105 | 1890 | 1 | 0 | 12 |
| Multi-Step XSS | 9 | 105 | 945 | 1 | 0 | 12 |
| Cros Site Request Forgery | 11 | 1 | 11 | 0 | 2 | 0 |

**Figure 12 – Test Execution Results of Medipedia Use Case**

The final step of the evaluation targeted the legal compliance aspects of eHealth software. The working scenario was based on the expected introduction of the General Data Protection Regulation (GDPR) that introduces several updates to the data protection legislation in high confidentiality areas, including eHealth. The evaluation of the RASEN method for legal compliance was undertaken by UiO researchers within the project on one hand side and by Info World's legal team on the other. The first step was *risk identification.* During a joint UiO – Info World meeting, the following risks were discovered with regards to Info World's situation pertaining to eHealth and the upcoming introduction of the GDPR:

***Risk 1:*** Every incident shall be considered as important irrespective of a prejudice or a non-prejudice. If the provision shall remain in this format, the result will a major responsibility for us (draft article 31, GDPR).

***Risk 2:*** The incorrect classification of the data and the inadequate handling of them. he introduction of some new terms in the GDRP, like: *personal data in large scale filing systems on children*, *genetic data or biometric data*, *data concerning health*, *main establishment*, *binding corporate rules*, *group of undertakings*, *child*. It is important to take into consideration how the data should be classified according to the new definitions from the GDRP project

***Risk 3:*** The necessity of adaptation of the condition for obtaining the explicit consent of the subjects for personal data processing in order not to be reconsidered. According to the GDRP project, the consent can be obtained through any means which allows the subject to exactly express the subject's conditions and to allow the person to express it affirmatively. For example: to tick off an item the moment of the visiting of a website, a declaration or an attitude which explicitly shows the fact that the subject accepted the operations of personal data processing.

***Risk 4:*** The necessity to establish a new strategy in order to erase those data and in the same time those data to do not show up in the searching having as object the name of the person or the subject (draft article 17, GDPR).

***Risk 5*:** The administrative sanctions- their value was increased: staring from 250,000 Euro until 2% from annual worldwide turnover (draft article 79, GDRP) for the enterprises.

The next phase consisted of *risk estimation*, that included well known risk assessment steps such as establishing likelihood and consequence scales as well as modelling the identified risks using the CORAS tool, and structuring them within a risk template developed within the project, according to the methodology presented in [3]. When compared with the existing process, the RASEN legal compliance tool-supported methodology provides a structured approach with more controls provided by the CORAS tool environment as well as facilitating organization and communication within the company by employing a common set of diagrams and templates to document risk.

The results of the Info World evaluation are discussed in the following subsection.

#### 5.3.3.2 Evaluation Results

| 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**<br>*E1*: The artifacts allows for understanding the relevant business and regulatory environment<br>*E2*: Relevant compliance requirements can be identified according to **A2**<br>*E3*: **A2** provides for the identification of compliance risks<br>*E4*: Compliance risks can be modeled in a structured manner according to the **A2** artefact.<br>*E5*: **A2** enables structured estimation of compliance risk<br>*E6*: **A2** enables structured evaluation of compliance risk<br>*E7*: Estimation, evaluation of compliance risks is easier using **A2**.<br>*E8*: **A2** can be applied to systems of various complexity and modularity, from in-house software components to assembled software systems delivered to customers.<br><br>Evaluation criteria improvement of Info World's testing process through **A2**:<br>*F1*: The RASEN artifact **A2** provides increased level of confidence on the compliance of the organization, compared to the alternative.<br>*F2*: The RASEN artefact **A2** enables better input to decision making than the alternative.<br>*F3*: The cost of using **A2** is in the long run lower than the value of the benefits from use. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 Result | *E1*: Using the CORAS-backed tool-supported methodology allows structuring the risk estimation process. Furthermore, by employing structured templates such as detailed within [3] facilitates communication between the organization's levels, as well as facilitating the documentation of risk at all times within the software project's lifetime.<br><br>*E2, E3*: Identification of compliance requirements and risks is currently |

| | |
|---|---|
| | undertaken manually, but supported by the presence of the structured templates [3].<br><br>*E4*: Compliance risks for the Medipedia platform were modelled using a graphical notation within the CORAS tool.<br><br>*E5 – E6*: Estimation and evaluation of compliance risk were achieved in the Medipedia use case using the structured CORAS notation and tool that produces risk values once likelihood estimations were provided.<br><br>*E7*: When compared with the existing ad-hoc approach to addressing compliance risk, the RASEN methodology provides both a structured approach via templates as well as tool-support in the form of CORAS.<br><br>*E8*: The present evaluation was focused on the Medipedia system, one of the most complex systems developed within the company. Its system architecture includes several components that are broadly reused within Info World, making their safety, reliability and legal compliance of paramount importance.<br><br>Given the complexity of the evaluated system we estimate **A2** to be applicable across systems of different scales and complexities.<br><br>*F1 – F2*: **A2** provides a clear and structured approach as well as effective tooling to support decision making.<br>*F3:* Given the scale and scope of the evaluation we cannot provide a long-term perspective on the involved costs. However, both the tool as well as the structured template are easy to use and do not require significant additional training for the company's legal team. |
| Involved Role(s) | Legal Counsel, Compliance Manager, Software Developer |

**Table 19 – 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**:<br>*E1*: A structured, tool-backed methodology that deployable for undertaking security risk assessment of Info World's software components and end products is available.<br><br>*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.<br><br>*F2*: Employing **A1** will bring more confidence in the correctness of the risk model that the current approach.<br><br>*F3*: A large part of the risk model produced following **A1** can be tested using conventional security tools. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 Result | *E1*: The RASEN tool-supported methodology was employed targeting the Medipedia eHealth system that resulted in a structured risk assessment model which Info World considers to be more advanced when compared with the existing process. The system that was modelled is both representative for the company as it is one of the most complex systems developed by Info World as well as representative, by using many of the common libraries that were developed in house.<br><br>As such, while a definitive answer cannot be provided before evaluating the methodology using several systems we believe the tools and methodology are transferable to other system of the same complexity.<br><br>*F1*: The RASEN tool-supported methodology for security risk assessment allowed the creation of a prioritized list of risks for the Medipedia system. While it currently does not allow automated generation of security tests, it facilitates the distribution of testing effort to those areas that are perceived as presenting high-risk.<br><br>*F2*: The evaluated methodology both requires and produces structured, quantifiable input and output, respectively, which facilitate test selection on the low level and decision making on a higher level.<br><br>*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. |
| Involved Role(s) | Software Architect, Software Developer |

**Table 20 – 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 2** [M36] | |
| Evaluation Phase 2 Rating | Fair |
| Evaluation Phase 2 Result | *E1*: This criterion was evaluated together with the security risk assessment evaluation process, where risks pertaining to individual components were aggregated into the product's risk picture.<br><br>While the scope of the evaluation was limited, the research providers showed that the **A1** artifact 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. |
| Involved Role(s) | Security Manager, Software Architect, Software Developer |

**Table 21 – 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 | 5 |

| | |
|---|---|
| Satisfaction | |
| Success Criterion | SC2.1 |
| Evaluation Criterion | Evaluation criteria related to artifact **A5**:<br>*E1*: Security test cases are a refinement of a structured test procedure targeting known types of vulnerabilities<br>*E2*: Test cases for **A5** can be generated using artifact **A1**.<br>*E3*: Effort spent on additional actions for obtaining test cases within **A5** is saved in the testing phase:<br>*E4*: Adapting **A5** results in more security vulnerabilities being uncovered than what is being uncovered using the current process.<br>*E5*: Testing results of **A5** can be used to update **A1**.<br><br>*F1*: The RASEN test technique (**A5**) is more rigorous than Info World's current test prioritization process.<br>*F2*: Artifact **A5** helps prioritize test procedures more accurately than Info World's current process.<br>*F3*: Test prioritization according to artifact **A5** may help save time during the risk assessment and testing phase for Info World. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Good |
| Evaluation Phase 2 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.<br><br>*E2:* Currently there is no automated generation of security test cases using the results of the risk assessment process.<br><br>*E3*: While the effort of obtaining test cases is not changed from the existing approach, the risk assessment methodology that provides the prioritized list of risks allows fine tuning the testing process and selecting most relevant test cases, which we expect will lead to a reduction in cost.<br><br>*E4:* Initial experimentation with the generation and execution of test cases (**A5**) as discussed within [2] has shown a link between the risk estimations resulting from the risk assessment process and results obtained during the testing process. Due to the limited time as well as having only one system within the use we cannot provide a definitive evaluation at this point.<br><br>*E5:* This criterion cannot be evaluated at this point.<br><br>*F1, F2*: The RASEN technique provides a prioritized list of risks that can be used to focus testing on, and as such is superior to current Info World processes.<br><br>*F3*: Test prioritization is expected to contribute to significant time savings once the product risk model and risk estimation are completed. |

| | |
|---|---|
| Involved Role(s) | Software Architect, Software Developer, Software Tester |

**Table 22 – 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. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Excellent |
| Evaluation Phase 2 Result | All the RASEN tools used in the evaluation were mature and reliable. |

**Table 23 – 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 | *E1*: The tools evaluated within the organized workshops were based on |

| | |
|---|---|
| Result | 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. |
| **Evaluation Phase 2** [M36] | |
| Evaluation Phase 2 Rating | Excellent |
| Evaluation Phase 2 Result | RASEN tooling used in the Info World use case is available as a series of plugins for well-known software tools such as Eclipse, and as such as easy to deploy and use. |

**Table 24 – Evaluation for requirement REQ-IW-N-120**

## 5.4 Unified Results of Use Case System Evaluation

The objective of the current section is to provide a unified result of the industrial use case evaluation of the innovations within the RASEN project, as well as to illustrate the progress that was achieved within the final implementation year of the project as seen from an industry perspective.

The user requirements were first defined in deliverable *D2.2.1 - Use case requirements definition* where they were partitioned into functional and non-functional. Table 25 provides a unified view of the use case evaluation results for every artefacts within the principal innovations of the project.

| Artefact | Evaluated using Requirement | Evaluation Result M24 | Evaluation Result M36 |
|---|---|---|---|
| **Innovation 1**: The PMVT approach for security pattern and model-based vulnerability testing | | | |
| **A3** The RASEN technique for security test automation | REQ-SAG-F-060 | N/A | Good |
| | REQ-SAG-F-070 | N/A | Good |
| | REQ-EVRY-F-020 | Fair | Fair |
| | REQ-IW-F-090 | Fair | Good |
| | REQ-IW-F-100 | Fair | Good |
| **Innovation 2**: The RACOMAT tool – risk assessment combined with automated testing | | | |
| **A6** The RASEN tool-supported method for risk assessment combined with automated testing | REQ-SAG-F-080 | Good | Excellent |
| **Innovation 3**: The RASEN method for risk-based security testing and legal compliance assessment | | | |
| **A1** The RASEN tool-supported method for risk-based security testing | REQ-SAG-F-010 | Good | Good |
| | REQ-SAG-F-020 | N/A | Excellent |
| | REQ-EVRY-F-010 | Fair | Fair |
| | REQ-IW-F-050 | Fair | Good |
| | REQ-IW-F-060 | Fair | Good |
| **A2** The RASEN method for compliance risk assessment | REQ-EVRY-F-030 | Fair | Fair |
| | REQ-IW-F-010 | Fair | Good |
| | REQ-IW-F-020 | Fair | Good |

| | | | |
|---|---|---|---|
| | REQ-IW-F-030 | Fair | Good |
| | REQ-IW-F-040 | Fair | Good |
| **A4**<br>The RASEN method for compositional security risk assessment | REQ-SAG-F-030 | Poor | Good |
| | REQ-SAG-F-040 | N/A | Fair |
| | REQ-SAG-F-090 | N/A | Fair |
| | REQ-SAG-F-100 | N/A | Good |
| | REQ-IW-F-070 | Fair | Fair |
| | REQ-IW-F-080 | Fair | Fair |
| **A5**<br>The RASEN tool-supported method for test-based security risk assessment and test result aggregation | REQ-SAG-F-050 | N/A | Excellent |
| | REQ-EVRY-F-030 | Fair | Fair |
| | REQ-IW-F-090 | Fair | Good |
| | REQ-IW-F-100 | Fair | Good |

**Table 25 – Aggregated evaluation result of functional requirements**

Table 26 illustrates the evaluation of the no-functional requirements of the use case providers. As non-functional requirements, they have not been linked with RASEN innovations of artefacts.

| Non-functional Requirement | Evaluation Result M24 | Evaluation Result M36 |
|---|---|---|
| REQ-SAG-N-020 | N/A | Good |
| REQ-SAG-N-030 | N/A | Fair |
| REQ-IW-N-110 | Excellent | Excellent |
| REQ-IW-N-120 | Excellent | Excellent |

**Table 26 – Evaluation of non-functional requirements**

Given the complexity of the RASEN project, as well as of the artefacts that were obtained during the three years of implementation, it becomes important to provide a unified, domain-independent and aggregate result of the evaluation activities within the project. As such, the individual rating obtained within all three project use cases, irrespective of domain, and for each evaluated artefact were averaged in order to provide an aggregated result, as shown within Figure 13.

The results of both evaluation activities (Phase 1, that was finalized at M24 as well as Phase 2, the current and final evaluation) are shown within Figure 13, for each of the project's most important artefacts. The scale that was employed is a five step one which was devised within *D2.3.1 - Use case evaluation v.1*, where it was employed to rate the results of the first evaluation.

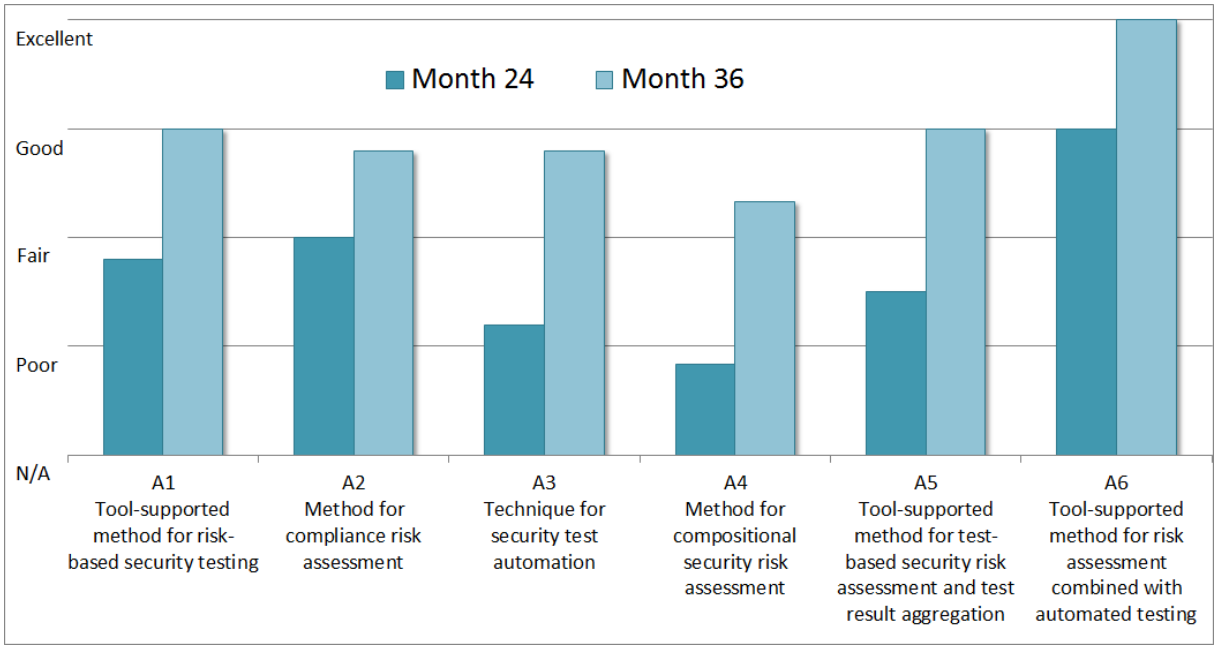


**Figure 13 - Unified rating of evaluated RASEN artefacts**

# 6 Best practices for software system makers and users based on experiences from evaluation

The use case providers feel that the application of the RASEN tool-supported methodology has proven to be applicable for smaller systems, including for limited deployment within the use case systems considered within the project. However, as some of the RASEN pilots also consider the use of large scale networked systems the applied methodology is expected to also show its natural limits. For example, in the case of Info World's Medipedia, this is illustrated by manual intervention required for generating security test cases once the risk estimation is completed; Info World feels that further work to complete the risk assessment – security testing circle would be of great benefit and would greatly ease deploying the methodology in an industrial context.

Out of the experience gained through using the RASEN methodology as well as evaluating it within two stages we hence recommend that the RASEN graphical models and testing strategies are complemented with additional valuable input coming from software code analysis. While such an approach was not part of the work in the RASEN project we believe that code analysis – which is rather cheap in terms of resources -- will clearly highlight major weaknesses and enhance existing testing strategies allowing the RASEN methodology to be applicable to very large systems.

## 6.1 Risk Management & Security Process in Business Industries

In order to deliver reliable and secure software to its customers worldwide, the production process of enterprise software at Software AG includes various stages of testing, starting from component testing, product testing, and testing of combinations of products. In addition the development process is compliant to the ISO 15408 standard which defines a common criteria framework used to specify functional security and assurance requirements of IT products and computer systems. More over customers in particular from the United States require adherence to NIST publications like "Minimum Security Requirements for Federal Information and Information Systems" (FIPS PUB 200), "Security and Privacy Controls for Federal Information Systems and Organizations" (NIST Special Publication 800-52), "Security Requirements for Cryptographic Modules" (FIPS PUB140-2) and many more related to cryptography, security and secure hashing.

The Software AG pilot is used to assess RASENs ability to improve the software production process and relate security testing with risk assessment methodology. As of now it is considered as a great benefit of RASEN to enable the risk analysis of newly implemented features. Thus whenever a new feature is implemented, the requirement analysis contains a careful assessment of needs and conditions, taking into account existing possible conflicts with requirements from various stakeholders. In this sense risk analysis is continuously assuring an exhaustive list of addressed risks which eventually contribute to the risk level on the product level. To mitigate the overall risk to a justifiable level, risky features must be identified which contribute excessively to the overall resulting risk.

Following this risk selection and mitigation approach helps identify risky program code that can consequently be altered to mitigate the risk of the feature. The pilot further strives to demonstrate compositional risk tracking which enables developers to trace the security impact of a newly implemented feature based on the risk analysis and vice versa the risk analysis of the feature is automatically reflected in the risk analysis of the feature is automatically reflected in the risk analysis of the product. As the list of software products is extensive and the implementation of tools, processes, and measures across several hundred developers a challenge, results of the RASEN methodology should reveal a comparative product risk analysis in order to prioritize product prioritization and risk rating.

## 6.2 Test Automation in eHealth

The cornerstone of IW's use case is represented by the protection of our end-users' healthcare data. Current legal regulations are represented by the 95/46/EC Directive. However, the mid 2010's are a time of crossroads with regards to data privacy, as new data protection regulations are expected to come into effect in the form of the General Data Protection Regulation (GDPR) sometime after 2015. The GDPR is expected to provide an updated legal framework accounting for the effect of disruptive

technologies such as rich Internet applications and cloud services. Coming into effect in a time frame where eHealth data breaches are becoming common, its adoption will require companies to recalibrate their efforts in order to ensure compliance with the upcoming regulations.

The latest draft of the GDPR article 31 specifies that all data breaches, regardless of caused prejudice must be reported to the supervisory authority within 24 hours, which changes the impact of data breaches independently of prejudice. This is compounded by increased administrative sanctions, which can now reach 2% of annual worldwide turnover. Furthermore, the draft also introduces the term "explicit consent", with data controllers such as Info World bearing the burden of proof for the data subject's consent.

As a highly visible, feature-rich system Medipedia provides a large attack surface. Any loss of data confidentiality or integrity can bear multiple financial and legal consequences imposed by the National Authority for Data Protection, partner medical clinics or compromised end-users. In addition, since medical decision are taken every day on the basis of the data stored in the system its corruption can have direct and undesirable medical consequence for its users. This outstanding combination between complex networked systems that handle highly-personal data such as Medipedia, the constantly changing and challenging security climate represented by an increasing number of threats to data security and privacy as well as the introduction of complex new legal requirements such as the GDPR means that companies which desire to remain at the forefront of their field must invest in new methodologies for ensuring the security and confidentiality of their most prized assets.

## 6.3 Advantages for the Finance Industry

EVRY's use case is used to show how RASEN can improve the security test methodology and process by using RASEN's method and technique for risk-based test identification and prioritization. EVRY's financial suite that consist of various financial system, ranging from net banks for private customers (which we have used in EVRY's use case) to management systems for banks, need to have a high level of security since these systems handles huge transfers of funds and handles sensitive and business critical data. The results from RASEN should show that by introducing formal risk assessment and prioritization methods should do lead to more effective and less time consuming security testing.

## 6.4 Lessons learned using the Industrial Pilots

By applying the RASEN methodology **Software AG** sees an increased use of developer-capacities through the prioritization as most vulnerable components and code fragments are revealed through the automated testing. As such the software development cycle substantially improved with respect to facilitating management decisions on investments to achieve the maximum risk mitigation while raising the awareness of confirmed risks software areas that represent vulnerabilities. Though the product management is able to evaluate and mitigate risks on software products both, on product and on company level. As it is infeasible to test for every vulnerability, an automation of risk assessment and security testing is obtained. When considering the benchmark of the CRSTIP Evaluation, SAG already has a mature and exhaustive software development where the overall software quality of software could be further improved through the application automated security testing as proposed by RASEN. Beyond those achievements, there are still theoretically unsolved questions regarding the meaning of risk when considering the composition of components which need to be further investigated.

For **Info World**, the adoption of the RASEN tool-backed methodology is expected to bring several benefits. First of all, by employing a well-defined, structured approach to legal and compliance assessment facilitates communication of new requirements across organizational levels and allows maintaining exact records regarding legal requirements, risks as well as necessary steps that need to be taken. As the upcoming GDPR brings several important changes for eHealth companies, this is one of the key aspects of RASEN adoption for the company. Second of all, automation of the security testing process facilitates faster implementation of required features and a quicker time to market of new products. Even more so, updating the product risk picture from security test results allows maintaining an updated per-product risk picture that facilitates taking any required measures, thus lessening the company's exposure to actions from National Authorities for Protection of Personal Data or the loss of reputation that is to be suffered from the public at large.

In the **EVRY** case RASEN is expecting to deliver a more effective use of the testing time available by introducing risk-based test identification. This goes along with an increased level of confidence that correct test cases with highest impact on overall security level is selected. As future work EVRY plans to generalize the input from RASEN to be used in the general methodology to be used in all security testing, and not only in the system used for the test case.

# 7 Fulfilling project objectives

The present Section is dedicated to showing how the requirements from the use case providers as well as their evaluation 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 was planned to be achieved through the following 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 27 – 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 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 28 – Evaluation coverage of S&T objectives**

As shown within Table 28 above, all stated project objectives were evaluated using at least one requirement defined by the industrial use case providers. However, we believe that in the light of the project's final evaluation it is also worthwhile to examine the degree to which use case providers have evaluated the fulfillment of each project objective.

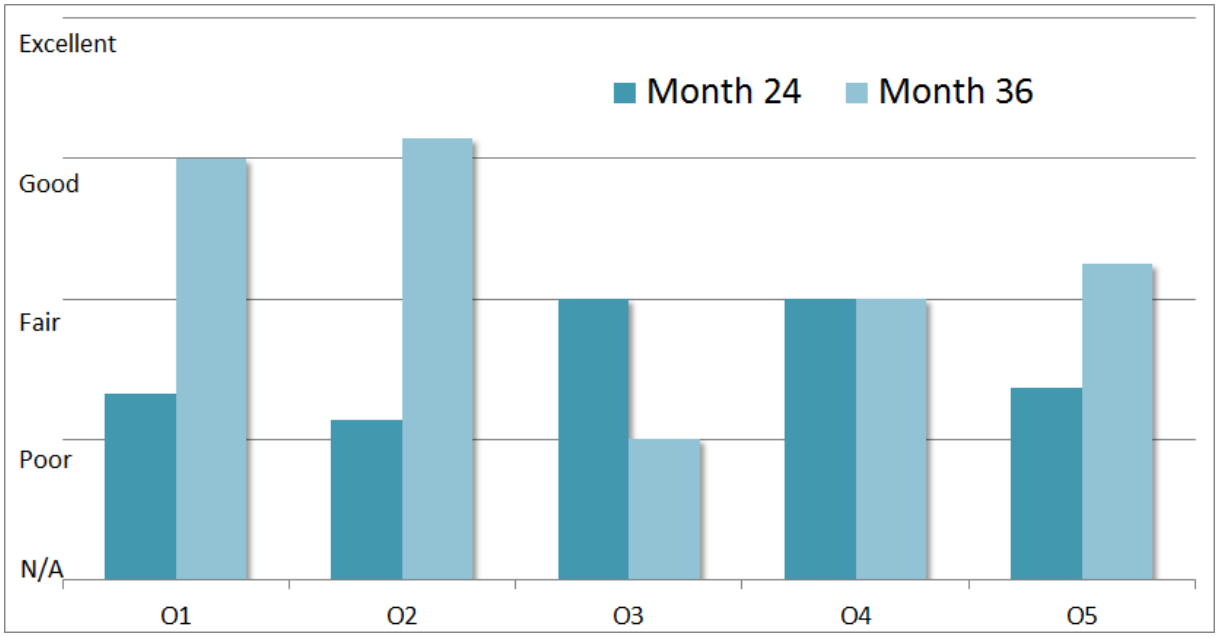


**Figure 14 – Use case evaluation result for each objective**

Figure 14 illustrates the aggregated rating for each of the project objectives. The values shown are calculated as follows: first of all, each use case requirement is linked to a project objective, to enable this kind of traceability. As part of the M24 and M36 evaluations, the fulfillment of each use case requirement was assigned a rating, as illustrated within Table 2. This information is averaged and aggregated within the figure above in order to provide a view that is independent from both use case providers as well as requirements.

# 8 Conclusion

The goal of WP2 was to clearly define the proposed use case scenarios, to extract clearly defined and evaluable 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.

In order to make the evaluation comparable across the use cases, each requirement was assigned an evaluation ranking between N/A (if the requirement could not be evaluated at all) and Excellent. The first evaluation of the project was finalized in M24 and brought the first results from an industry use case perspective: of the 23 defined functional requirements, 7 could not be evaluated and 1 was evaluated at Poor. The rest of the requirements already received ratings of *Fair* or *Good*. The present evaluation illustrates the progress made in the final research and development phase: all requirements were evaluated and given at least a *Fair* rating, with over half of requirements being fulfilled at a *Good* level.

Each use case provider undertook their own independent evaluation based on the requirements that were first defined in deliverable *D2.2.1 – Use Case Requirements Definition.* After the evaluation process, the results were aggregated in order to illustrate project progress since the first evaluation at the M24 mark. Furthermore, as use case requirements were given at a lower level, the evaluation information was aggregated at both artefact as well as project objective level, as shown within Sections 5.4 and 7.

As also illustrated in the publicly-available whitepapers that were produced by the industry partners [4], the evaluation process resulted in a common approach that is aimed towards organizations interested in adoption the RASEN tool-supported methodology. We believe that a common assessment of the evaluated artefacts, such as presented within Section 6 of this deliverable is a valuable tool for both organizations within the industry, as it provides an insight into the strengths and weaknesses of the developed technologies, as well as for researchers and technology providers for planning the next wave of innovation.

# 9 Annexes

## 9.1 Annex I - 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 [1]. The approach is based on previous work undertaken within the ITEA2 – Diamonds[2] 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 15.

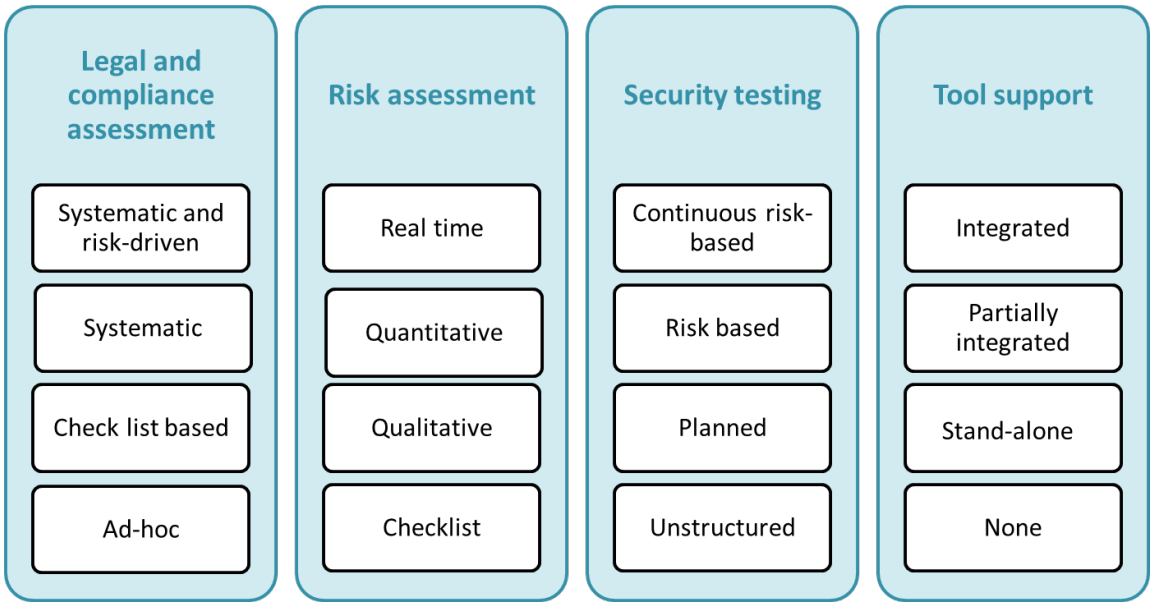


**Figure 15 – CRSTIP key areas and levels**

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. |

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

| | |
|---|---|
| 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, 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 29 – 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 30 – 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 31 – 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 32 – Levels in tool support**

# 10 References

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

[2] Alexandre Vernotte, Bruno Legeard, Fabien Peureux, Cornel Botea and Arthur Molnar - *Risk-Driven Vulnerability Testing: Results from eHealth Experiments using Patterns and Model-Based Approach* (RISK 2015 workshop)

[3] Samson Esayas, Tobias Mahler, Fredrik Seehusen, Frode Bjørnstad and Veda Brubakk - An Integrated Method for Compliance and Risk Assessment (Experiences from a Case Study), to be published within IEEE Conference on Communications and Network Security, Florence 2015.

[4] Frank Werner , Albert Zenkoff, Arthur Molnar, Erlend Eilertsen - An Industrial Perspective on Security Testing, Risk Assessment, and Legal Compliance