BONSEYES

ARTIFICIAL INTELLIGENCE MARKETPLACE

Deliverable D5.1

Demonstrator Proof-of-Concept

Point of Contact	Gordana Velikic
Institution	RT-RK
E-mail	Gordana.Velikic@rt-rk.com
Phone	+381695060934

Project Acronym	**BONSEYES**
Project Title	**Platform for Open Development of Systems of Artificial Intelligence**
Grant Agreement No.	**732204**
Topic	**H2020-ICT-2016 Smart Cyber-Physical Systems**
Project start date	**1 December 2016**
Nature	**Other**
Dissemination level	**Confidential**
Due date	**M18**
Date of delivery	**M18**
Lead partner	**RT-RK**
Contributing partners	**NVISO, TCD, FHNW, SYNYO, HES-SO, ARM, ZF**
Authors	**Gordana Velikic (RT-RK), Samuel A. Fricker (FHNW), Tim Llewellynn (NVISO), Lei Wang (ZF), Daniel Ostler (TUM)**
Reviewers	**Samuel A. Fricker (FHNW), Milena Milosevic (RT-RK), Nuria Pazos (HES-SO), Miguel de Prado (HES-SO), Noelia Vallez (UCLM), M. del Milagro Fernandez (UCLM), Oscar Deniz (UCLM), Georgios Goumas (ICCS), Itsik Arbel (NVISO)**

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732204 (Bonseyes).

This work is supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0159.

The opinions expressed and arguments employed herein do not necessarily reflect the official views of these funding bodies.

Revision History

Version	Date	Author	Comment
0.1	April 2018	Gordana Velikic (RT-RK)	First draft
0.2	May 2018	Samuel Fricker (FHNW)	Second draft
0.3	May 2018	Samuel Fricker (FHNW)	Quality assurance
0.4	May 2018	Milena Milosevic (RT-RK), Noelia Vallez (UCLM), M. del Millagro Fernandez (UCLM), Oscar Deniz (UCLM), Georgios Goumas (ICCS), Miguel de Prado (HES-SO), Nuria Pazos (HES-SO)	Integration of all review contributions
0.5	May 2018	Samuel Fricker (FHNW)	Protocol for validating Bosneyes
1.0	May 2018	Gordana Velikic (RT-RK), Samuel Fricker (FHNW)	Final version

Contents

Abbreviations, Participant short names and Glossary

Abbreviations

AI	**Artificial Intelligence**
CPS	**Cyber-Physical Systems**
IoT	**Internet of Things**
DL	**Deep Leaning**
KPI	**Key Performance Indicator**
LPDNN	**Low Power Deep Neural Network**
SDK	**Software Development Kit**

Participant short names

NVISO	**nViso SA**
UCLM	**Universidad de Castilla – La Mancha**
TCD	**Trinity College Dublin**
UEDIN	**University of Edinburgh**
FHNW	**Fachhochschule Nordwestschweiz**
TUM-Med	**Klinikum rechts der Isar der Technischen Universität München**
ICCS	**Institute of Communication and Computer Systems**
SYNYO	**Synyo GmbH**
HES-SO	**Haute Ecole Spécialisée de Suisse Occidentale**
ARM	**ARM Limited**
ZFFNAG	**ZF Friedrichshafen AG**
RTRK	**Institut RT-RK**
SCIPROM	**SCIPROM Sàrl**
BTH	**Blekinge Tekniska Högskola**

Glossary

The Bonseyes glossary is available on https://www.bonseyes.com/glossary/.

List of Tables

List of Figures

Summary

The main objective of this deliverable is to describe how the achievement of the Bonseyes goals will be evaluated, identify risks, and suggest risk elimination actions at an early stage of development via feedback to work packages 2 and 4.

The document describes the role of the demonstrators and use cases for validating Bonseyes (section 1), describes the protocols intended to be used to validate Bonseyes (section 2), and reports the current status and risks of the work on the demonstrators and use cases (section 3).

This document is a mid-term report of Demonstrator Proof-of-Concept, reflecting the status of the task T5.1 at month M18. T5.1 has started in month M13 and will end at month M24. D5.2 will offer the final report of Demonstrator Proof-of-Concept at month M24. The work on the final use cases during months M25-M36 will be reported with the deliverables D5.3-D5.7. D5.7 will include the results of the validation according to the validation protocol.

1. Role of Demonstrators and Use Cases in the Bonseyes Project

The artificial intelligence (AI) integration into modern technology is faster than it had been imagined few years ago. The level of technology adoption and its maturity to perform smoothly are the main benefits, or obstacles, for the rapid development.

Deep learning (DL) is transforming AI applications to the levels of high consumer acceptance, however, the cost of computation to develop technology based on DL AI models influences the deployment of direct development on the end devices. In addition, lack of general access to training data, may restrict the developmental momentum to companies that have access to specific customer data. Encapsulated development may prevent diversity of applications, encourage monolithic markets, or restrain the competitiveness of small sized companies.

There are many business model scenarios in which the platform may bring improvements in efficiency, performance, reliability, security, and productivity. For illustration purposes, let’s assume that a company developing a product has valuable data, but possesses limited research power, measured in square feet of facilities with necessary equipment and trained personnel. For such company it would be hard to enter the circle of end-to-end product development or services. Also, change in trends of developed technology may turn to “unplanned waste” on invested resources to develop models that are optimized for platform A, while the product constraints change and another kind of platform is in demand to stay competitive. Additional constrains may result in longer time to market, not to mention higher costs. In addition, a vision of a product brings decisions such as the choice of the platform that would be the best fit based on available experts, which may not necessarily be the best choice.

The **Bonseyes** project addresses all these aspects of developmental challenges and envisions the **platform** for **open development** of **AI systems**. The three key platform elements are: Data Marketplace, Deep Learning Toolbox, and Developer Reference Platforms for optional “edge computing”, embedded computing systems, or “cloud computing”. Potential users can choose from many options or create offers to other users. Thus, the Bonseyes platform may help overcome lack in expertise, time, or resources during any phase of the product development. In addition, the platform enables the complete development end-to-end cycle.

The variety of elements that represent the Bonseyes's concepts are developed in related work packages. The openness of a shared developing platform enables collaboration, modularity, and scalability. Such platform is envisioned to be user friendly, such that experts of different levels, from fundamental knowledge to high level expertise, can use the Bonseyes platform for their development processes.

To illustrate the Bonseyes platform in real scenarios, WP5 is planned as a set of demonstrators from top industries, such as automotive, healthcare, and consumer, so as to prove and validate the concepts developed in previous work packages. The use-case scenarios that will be finalized in Demonstrators are initially described in WP1 deliverable D1.1. The demonstrators will help to:

- Detect potential weaknesses and highlight unseen strengths in the project approach
- Eliminate detected risks during the early stage of development of the Bonseyes platform
- Showcase the capabilities of each element of the platform and the platform as a whole
- Ensure that the project progress is in alignment with the project requirements
- Ensure that accompanied documentation and interfaces are in the alignment with the vision that the project platform is easy to use

2. Protocol for Validating Bonseyes

Outlined in D1.1, the Bonseyes framework will be validated by answering the validation questions shown in Table 1.

Table 1 – Validation questions (D1.1)

Validation Question	Motivation
VQ1: Does Bonseyes meet the stakeholders' needs?	Bonseyes aims at delivering a minimally viable data marketplace at TRL4, i.e. a marketplace that is validated in the laboratory and, as a next step, will be ready to be validated in a relevant environment. In Bonseyes, the laboratory is the distributed AI systems development environment that spans the Bonseyes consortium partners.
VQ1.a: Does Bonseyes meet the Bonseyes use case partners' needs?	Bonseyes will primarily be validated with three proof-of-concept demonstrators and use cases. The demonstrators and use cases offer transparency and insights to understand the mechanisms of how industrial system owners and data scientists want to collaborate.
VQ1.b: Does Bonseyes meet the needs of the wider developer community?	To address external validity threats, Bonseyes will be validated in a series of Hackathons. In these Hackathons, use cases, system owners, and data scientists collaborate and utilize Bonseyes that have not been involved in the project definition and execution.
VQ2: Does Bonseyes meet the Bonseyes goals?	Bonseyes is being created for achieving the following four goals: G01: 80% reduction in the cost of ownership. G02: 50% reduction in development time. G03: 90% reduction in computation memory requirements. G04: Adoption by at least 100 developers.

2.1 VQ1: Does Bonseyes meet the stakeholders' needs?

To answer VQ1.a), Bonseyes will be prototyped and brought into use incrementally by the Bonseyes use case partners according to the use case descriptions. This process has started in M13 in T5.1 Demonstrator Proof-of-Concepts and continues until M24. Starting with M25, the evaluation shifts to the Bonseyes use cases. The resulting experiences will be used to evaluate the Bonseyes components against the needs described in Table 1. The Bonseyes consortium will investigate qualitatively how each need is addressed by the Bonseyes components, quantitatively with the stakeholders' subjective satisfaction rating, and quantitatively by estimating the extent of benefit in comparison to alternatives to the Bonseyes components that were known to the stakeholders at the start of the Bonseyes project. The qualitative analysis will offer an understanding of the perceived benefits of Bonseyes. The quantitative analyses will offer an understanding of the extent of the benefits.

To answer VQ1.b), prototypes that are mature enough will be brought into use in the workshops and Hackathons. AI tasks will be proposed that may be addressed by letting data providers, annotators, data scientists, and application developers collaborate intensively during a short time. The experiences will again be used to evaluate the Bonseyes components against the needs described in Table 1: qualitatively how each need is addressed and quantitatively in terms of subjective satisfaction rating and estimating the extent of benefit in comparison to alternatives.

Table 2 outlines the stakeholders' needs identified in D1.1 and how the fulfilment of these needs will be evaluated. As the table shows, the following types of evaluation studies will be performed:

- Type I: Design reviews with in-depth interviews.
- Type II: Lessons-learned from dynamic validation.

Table 2 – Stakeholder needs (D1.1 Table 6) and evaluation of need fulfilment

Need	Evaluation of Need Fulfilment
Needs of System Owners	
Need1.1: Minimize time-to-market	Type I: Identify factors that lead to minimized time-to-market and benchmark the Bonseyes implementation against these factors.
Need1.2: Minimize cost-of-ownership	Type I: Identify factors that lead to minimized cost-of-ownership and benchmark the Bonseyes implementation against these factors.
Need1.3: Control the dissemination of AI artefacts	Type I: Identify the threats for unauthorized data or model access and benchmark the Bonseyes implementation against these threats.
Need1.4: Control the dissemination of knowledge about AI systems development	Type I: Identify the threats for unauthorized access of AI challenges and benchmark the Bonseyes implementation against these threats.
Needs of Data Scientists	
Need2.1: Train models that optimize inference accuracy while offering acceptable time and resource consumption	Type II: Utilize the Bonseyes optimization capabilities in the practical use cases and analyse the feedback.
Need2.2: Benchmark trained models against other models	Type II: Utilize the Bonseyes benchmarking capabilities in the practical use cases and analyse the feedback.
Need2.3: Exploit models commercially	Type II: Utilize the Bonseyes licensing and license enforcement capabilities for models in the practical use cases and act on the feedback.
Need2.4: Minimize the training duration of training jobs	Type I: Identify factors that lead to minimized time-to-market and benchmark the Bonseyes implementation against these factors.
Need2.5: Control where the training jobs are performed	Type II: Utilize the Bonseyes federation capabilities in the practical use cases and analyse the feedback.
Needs of Application Developers	
Need3.1: Optimize model inference for target platform	Type I: Identify factors that lead to optimized model inference for each selected target platform and benchmark the Bonseyes Type II: implementation against these factors.
Need3.2: Find a good trade-off between inference accuracy and time and energy consumption	Type II: Utilize the Bonseyes benchmarking capabilities in the practical use cases and analyse the feedback.
Needs of Device Manufacturers	
Need4.1: Minimize the trade-off between model accuracy and execution performance	Type II: Utilize the Bonseyes benchmarking capabilities in the practical use cases and analyse the feedback.
Need4.2: Provide access to the AI systems market	Type II: Utilize the Bonseyes market capabilities in the practical use cases and analyse the feedback.
Needs of Integrators	
Need5.1: Offer data integration for deploying AI capabilities over multiple devices	Type II: Utilize the Bonseyes universal development platform tools in the practical use cases and analyse the feedback.

2.1.1 Generic Design for Studies of Type I: Design Reviews with In-depth Interview Studies.

The studies of type I aim at evaluating whether the design chosen for implementing Bonseyes are appropriate to meet the stakeholder need. This type of study is chosen when functional suitability[1] is important and the goal-fulfilment of utilizing Bonseyes is difficult to measure reliably. For example, the minimization of time-to-market (Need1.1) is difficult to measure because time-to-market may span much calendar time and may be influenced by many confounding factors. In such circumstances, other researchers have interviewed

[1] c.f. ISO/IEC FDIS 25010 taxonomy of quality model framework outlining standard types of quality characteristics.

experienced practitioners to understand whether all relevant means have been considered to achieve the practitioner's goals.

The generic design of such a study involves the answering of two validation questions: VQI.1: *what are the potentially relevant options*? VQI.2: *have the relevant options been chosen*? VQI.1 may be answered with a literature review that results in an enumeration of design options. VQI.2 may be answered with an in-depth interview study in which the design options chosen for implementation are evaluated and ranked (example: Berntsson Svensson[2]). A study of type I will result in a set of confirmed, missed, and unnecessary design options.

2.1.2 Generic Design for Studies of Type II: Lessons-Learned from Dynamic Validation.

The studies of type II aim at evaluating whether the implementation of Bonseyes is appropriate to meet the stakeholder needs that emerge in relevant or operational environments[3]. This type of study is chosen when functional suitability and context coverage are important and fully-fledged use of the Bonseyes feature is feasible within given calendar time constraints. For example, it will be easy to let practitioners use the licensing and license enforcement capabilities and offer feedback about the experience of using these capabilities in the practitioners' relevant or operational contexts. In such circumstances, other researchers have interviewed practitioners about their user experience.

The generic design of such a study involves the answering of the following two validation question: VQII.1: *does Bonseyes allow the practitioner to meet the stated purpose of Bonseyes use*? VGII.2: *does Bonseyes meet the needs that emerge in the relevant or operational environment*? VQII.1 and VQII.2 may be answered with an in-depth interview study that follows the use of Bonseyes in which the lessons learned are elicited and evaluated. VQII.1 involves quantitative questions in which the interviewee estimates the extent of met purpose. A study of type II will result in a description of how Bonseyes is used, a set of Bonseyes strengths and weaknesses, and recommendations for improving Bonseyes.

2.1.3 Generic Protocol for Studies involving Interviews

The interview protocol will be based on a questionnaire prepared with the answer for VQI.1. The interview will capture the respondent's profile and background. It will review the design options, let the respondent rank and select the options, and identify the rationales for the practitioner's choices.

The respondents should be chosen to represent the environments in which Bonseyes is expected to be used. For example, each Bonseyes use case represents one such environment: the medical use case represents open academic collaboration, the automotive use case secret collaboration with a large-scale company, and the consumer use case secret collaboration with a start-up.

Each interview may be performed with individuals or with a group. The capture of a group opinion is important if multiple stakeholders have conflicting opinions and succeed to reach a consensus. Each interview should be short and include the elicitation of the "hows" and "whys" that indicate the respondent's thinking and rationales. The interviews should be recorded, or in-depth notes be taken.

The analysis of the interview results involves quantitative and qualitative elements. The ranking of options may be evaluated quantitatively for significant differences with a non-parametric Wilcoxon rank sum test, thus reducing the number of differences that would otherwise be implied. The discussions may be

[2] R. Berntsson Svensson, T. Gorschek, B. Regnell, R. Torkar, Al. Shahrokni, R. Feldt: "Quality Requirements in Industrial Practice – an extended interview study at eleven companies." IEEE Transactions on Software Engineering 38(4):923-935, 2012.

[3] T. Gorschek, C. Wohlin, P. Garre, S. Larsson: "Model for Technology Transfer in Practice." IEEE Software 23(6):88-95, 2006.

qualitatively analysed with conventional content analysis where themes are identified through coding thus allow grouping of important experiences and opinions.

2.2 VQ2: Does Bonseyes meet the Bonseyes goals?

To answer VQ2), the following approaches are pursued:

- Goal G01, 80% reduction in the cost of ownership (study type III): The Bonseyes consortium will take the innovator and system owner's perspective and compare a) development based on AI Marketplace-enabled outsourcing with b) development based on monolithic in-house development. The monetary cost of each task, data generation, annotation, model training, AI application development, and AI system development, will be measured. The minimal reduction in cost will be calculated by assuming no reuse of readily available AI artifacts and by comparing the lead-time and speed of in-house development vs. outsourced development. The maximal reduction in development time will be calculated by assuming availability and reuse of all AI artifacts and comparing development without reuse vs. development with reuse according to the pricing the AI artifacts would be offered.
- Goal G02, 50% reduction in development time (study type III): The Bonseyes consortium will take the innovator and system owner's perspective and compare a) development based on AI Marketplace-enabled outsourcing with b) development based on monolithic in-house development. The effort invested in each task, data generation, annotation, model training, AI application development, and AI system development, will be measured. The minimal reduction in development time will be calculated by assuming no reuse of readily available AI artifacts and by comparing the lead-time and speed of in-house development vs. outsourced development. The maximal reduction in development time will be calculated by assuming availability and reuse of all AI artifacts and comparing development without reuse vs. development with reuse.
- Goal G03, 90% reduction in computation and memory requirements (study type IV): The Bonseyes consortium measures the computation and memory resource usage during inference on a) the Bonseyes developer platform with the model optimized with Bonseyes and b) the standard developer platform with the model optimized with standard tooling. The difference will show the extent of reduction in computation and memory requirements.
- Goal G04, adoption of Bonseyes by at least 100 developers: the Bonseyes consortium counts the participants in the workshops and hackathons that were using the Bonseyes components.

2.2.1 Generic Design for Studies of Type III: Comparison of Process Variants with a Multiple Case-Study.

The studies of type III aim at describing differences between variants of engineering processes and evaluating the impact of these differences on economic variables such as cost-of-ownership and development time. This type of study is chosen when a complex phenomenon, such as an engineering process, is studied in its real-world context. Case study research aims at describing how the phenomenon unfolds and why it is unfolding in the observed way[4]. Case studies may be exploratory providing new insights, descriptive portraying a phenomenon, explanatory seeking explanation, or improving trying to improve a certain aspect of the studied phenomenon.

The generic design of a type III study involves the answering of the following two validation questions: VQIII.1 (descriptive): *how do practitioners develop AI systems with/without Bonseyes*? VGIII.2 (explanatory): *how does the use of Bonseyes affect the economic variable of interest*? VQIII.1 may be answered by describing the engineering process steps with and without the use of Bonseyes by following formalisms such as UML Activity

[4] P. Runeson, M. Höst: "Guidelines for conducting and reporting case study research in software engineering." Empirical Software Engineering 14(2):131-164, 2009.

Diagrams or BPMN. VQIII.2 may be answered by measuring or estimating the economic variable of interest on each process step. The instances of the engineering processes studied in the case studies may concern the situations prior and after the adoption of Bonseyes, respectively AI systems development projects that use or not use Bonseyes. A study of type III will result in a description of the engineering process and a description of the Bonseyes impact for the economic variables of interest grounded in the observed cases.

Case study research is a flexible type of research where the protocol is adapted to the observations and the insights that emerge by contrasting the observations with supporting or conflicting theory.

As samples, we will use the 6 Bonseyes AI systems development projects, i.e. the 3 proof-of-concept and the 3 use cases. Each such case involves collaborative development of AI systems. The sampling represents a maximum variation sampling approach because the cases were chosen to reflect as different engineering contexts as possible (academic/open, large-scale company/closed, start-up company/closed).

Data of different modalities will be collected from multiple sources. The diversity allows triangulation, hence to evaluate how robust the findings are or how they vary. The researchers interact with the cases over prolonged time by performing interviews and collecting documents that describe the engineering processes, the economic variable of interest, and related information.

The data will be analysed by creating a map of the engineering process steps and how they vary across companies and cases. Of particular importance will be the position of the Bonseyes marketplace and tools and how they affect the structure of the processes and the economic variable of interest. Where possible, quantitative measurements will be collected and aggregated. Qualitative data collected from the interviews or extracted from the documents will be analysed with conventional content analysis. If patterns about the impact of Bonseyes become evident, we intend to extend the content analysis with grounded theory methodology[5].

2.2.2 Generic Design for Studies of Type IV: Evaluation as a Computer Experiment.

The studies of type IV aim at evaluating whether Bonseyes generates results, such as an AI model or algorithm, of quality superior to results generated with other means. This type of study is chosen when results can be generated and quality measurements applied on the results reliably. For example, with predefined datasets and a specified operational environment, it is possible to generate an AI algorithm and apply accuracy and performance measurements on the execution of the AI algorithm. An example of such a study has been given by Elafrou et al[6], and we refer to the design of that study as a template and example for other computer experiment studies.

2.3 Pragmatic Use of the Study Results

Positive impact and praise will be used for formulating experience reports in support of Bonseyes dissemination and exploitation. Skepticism and criticism will be used as a feedback to adjust the requirements and improve the Bonseyes concept and implementation.

[5] K. Charmaz. Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis. Sage Publications. 2006.

[6] A. Elafrou, G. Goumas, N. Koziris : "Performance Analysis and Optimization of Sparse Matrix-Vector Multiplication on Modern Multi- and Many-Core Processors." 46th International Conference on Parallel Processing (ICPP 2017), Bristol, U.K.

3. Proof of Concept: Light Version of Bonseyes

The main objective of the first deliverable D5.1 in WP5 is to create a light but complete version of results that integrate the Bonseyes project throughout its work packages. In doing so, many hidden issues related to end-to-end integration can be encountered and solved. These problems will reveal potential risks and help analyze the background to find solutions for the risk mitigation, recommend paths of further development, and potential changes.

The light version is built using initial version of the core Bonseyes components (Figure 1):

- Data Marketplace (D2.2)
- Deep Learning Methods (D3.1)
- Reference Architectures and Models (D3.3. & D3.5)

At this stage no Developer Platforms are used, rather a prototype application is built using PC.

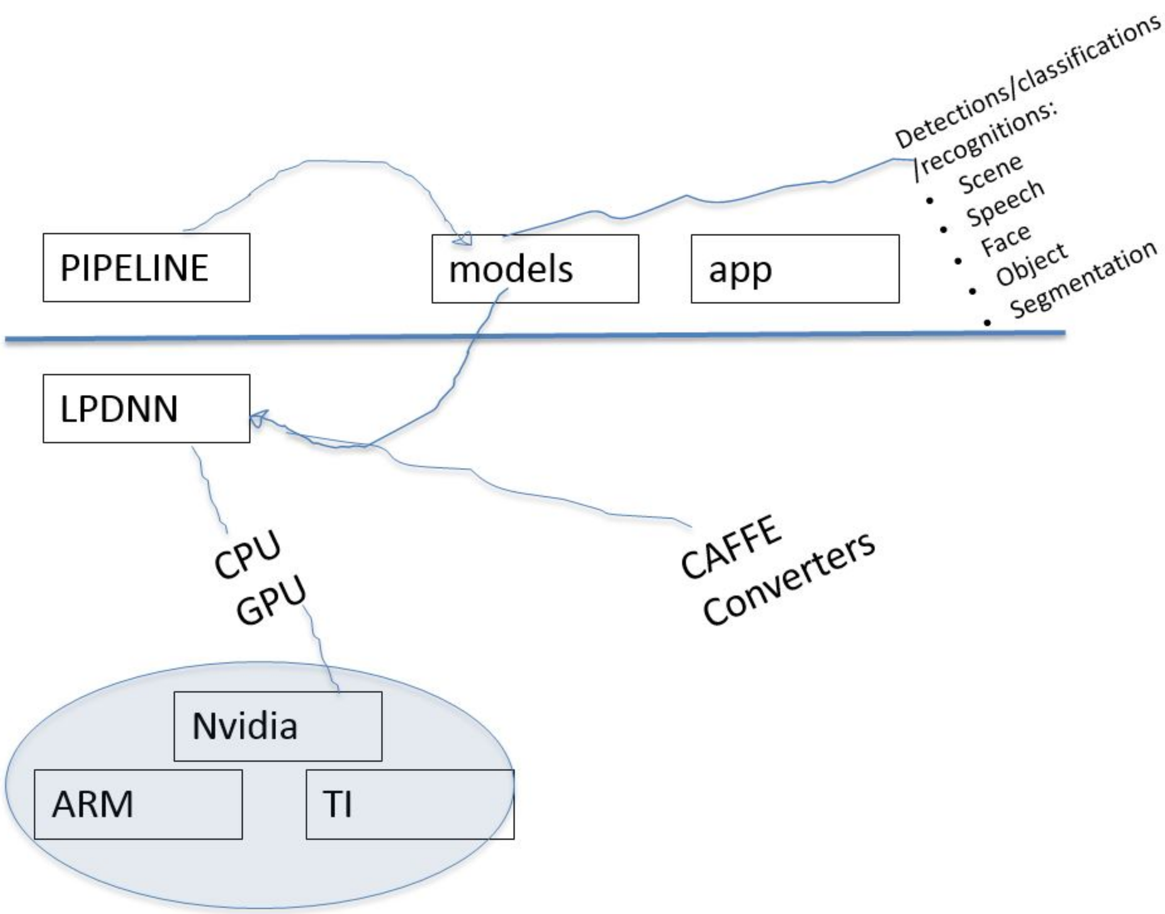

Figure 1. Initial simplified sketch of the Bonseyes proof of concept prototype from one of the WP5 team brainstorming sessions. There are three major steps in the Bonseyes product development: 1. Training model (Market place), 2. Deployment (LPDNN), 3. Embedded platform

4. Proof of concept as a risk detector

During WP5 brainstorming sessions at workshops and conference calls, a number of project risks have been identified. In addition, the measures to eliminate, manage, or reduce impacts of risks have been proposed. This step will ensure to have a properly functioning final System-of-Systems Demonstrators.

The following risks have been identified:

- Use cases not identified well by partners
- Use case partners do not know how to measure Key Performance Indicators (KPI) of the project
- Demonstrator development platforms not ready
- Demonstrator performance is poor
- Tools and methods are too difficult for partners to use
- Tools and methods do not work
- Marketplace does not have content, or content is not sufficient

4.1 Risk: Use cases not identified well by partners

4.1.1 Risk mitigation

One of the risks is that the partners may misunderstand what they need to demonstrate. We have defined five scenarios for Demonstrators (WP1) as illustrative examples of the Bonseyes platform potential:

- two automotive use cases (owner ZF)
- a consumer use case (owner nViso)
- two healthcare use cases (owner TUM)

The consortium partners' collaborative work has singled out a nice set of core capabilities from specific demonstrators. The set is rich enough to have a fair amount of capability to allow the structure to demonstrate a number of different use case applications. The use case applications for the proof of concept stage are built upon specific abstracted tasks for the example scenarios.

For the initial risk elimination phase, the algorithms that are building blocks of system of systems are given in the Table 3.

Table 3. Building blocks of system of systems. Ten algorithms are initial building blocks. ZF contributes with: pedestrian detection, face recognition, distracted driver detection; nViso with face detection, face recognition, emotion recognition, and body pose detection and TUM with keyword spotting, scene classification and instrument detection.

Task/Algorithm vs. Application Demo Scenario	Object Detection	Face Recognition	Scene Classification / static	Scene Classification / temporal (LSTM network)	Keyword Spotting, in Speech Recognition	3D Body Pose (Open pose reference architecture)
Emotional (nViso)	Face Detection	Face Recognition	Emotion Recognition (the face is the scene)	Emotion Recognition (analysing a sequence of images)		3D Body Pose
Healthcare (TUM)	Instrument Detection		Surgery stage Classification	Surgery stage Classification (analysing a sequence of images)	Patient / Physician Conversation	
Automotive (ZF)	Pedestrian Detection	Face Recognition	Driver Distraction	Driver Distraction		

For the proof of concept, ten algorithms have been selected from five scenarios. One of the key benefits of the Bonseyes platform is (re)usability of the developed models that are reflected in the algorithms. These algorithms constitute the frame for 10 challenges. Each partner has provided the resources needed for specific challenge. In addition, each partner is required to define the challenges of each algorithm, such as for example, those challenges that single out the algorithm from the algorithms that exist on the Bonseyes market. For illustration purposes facial recognition can be specific challenge. Facial recognition requires detection of small objects like eyes, nose or lips while other challenges that require object detection may be less demanding because the objects detected are larger. We have algorithms that are developed for the autonomous vehicles and run on fast processors in real time, however, we may be required to develop the algorithms for tractors ECUs that do not have same powerful processors.

The summaries of the technical hurdles, such as challenges in the illustrative examples above, resources (datasets, reference models), evaluation of the results on target platform, and reference models are given in the D3.5. The technical hurdles that algorithms need to overcome are the content for the marketplace, and also add to competitive edge of the marketplace. One of the technical hurdles for face recognition are small objects detection. Adding solutions to these problems may increase number of different options for object detection and thus increase the competitiveness of the marketplace. The challenges also provide the insight of differences between the regular market and Bonseyes marketplace, thus defining the reason to use the Bonseyes marketplace.

4.2 Risk: Use case partners may not know how to measure the KPIs of the project

4.2.1 Risk mitigation

Key Performance Indicator (KPI) is a simple and yet, clear indicator that helps understand the progress of the project. Although the word “simple” is used in this definition, very often the view on KPIs differ considerably. To prevent the risk of inadequate KPIs, the FHNW team, based on their SW Engineering expertise, has provided a list of KPIs in the deliverable D1.1 and protocols for the KPIs evaluation for Bonseyes validation in this document (section 2). The protocols will be effective for all partners in the consortium. In addition, the FNHW will collect measurements from use case partners of “before-Bonseyes” situation and measurement of “after-Bonseyes”. The measurements will help signal where we can improve time and cost of the development of demonstrators for the project, and further, for the novel Bonseyes platform users after the project is over.

4.3 Risk: Demonstrator development platforms not ready

4.3.1 Risk mitigation

In the first task of WP4 several prospective embedded development platforms are evaluated. The goal was to select low power platforms with high performance processors (GPU, DSP, VPU) necessary to meet smart Cyber-Physical Systems (CPS) constraints within the industry demonstrators. Deliverable in form of the report is part of the WP4, however, relevant considerations will be provided in the appendix. The consortium agreed on the following platforms:

- DSP : RT-RK Alpha AMV based on TI
- GPU : NVIDIA Platform : TX2 and TX2i
- CPU : RK3399
- VPU : Renesas V3M Platform

The consortium members are working to ensure that the platforms are available and to get the needed support for the platform according to the Demonstrator specifications.

RT-RK has contacted Renesas to enter the list for the platform delivery. The list is long and may affect the time line. To prevent delays, RT-RK and HES-SO will explore further possibilities. Platforms are selected according to performance and capabilities, experience of consortium partners in using the platform, the partner developed platforms, or consortium members that have good collaboration with platform provider/manufacturer to avoid potential problems of undelivered platforms or lack of adequate support.

4.4 Risk: Demonstrator performance is poor

4.4.1 Risk mitigation

To avoid poor performance of the Demonstrators, e.g. the models that run on the board being slow, conference calls, workshops and hand-on workshops have been organized. The important questions are clarified during these meetings, such as: possible problems in deployment, speed of execution, appropriate satisfaction of the performance on embedded system for the use-case demonstrator. The reports for the first measurements for selected algorithms and scenarios are collected to confirm that we can provide working, useful demonstrators that run in real time. In the process are evaluations of the performance for Object Detection (Face and Pedestrian), Face and Body Posse Recognition.

4.5 Risk: Tools and methods are too difficult for partners to use

4.5.1 Risk mitigation

During evaluation workshops and phone calls, some partners expressed concern that the tools are difficult to use. Since this observation is also important from the aspect of user, not just a team member that collaborates in development, a series of hands-on workshops have been arranged to make sure that the tools are clearly explained or modified. During workshops, the interaction between the Marketplace and Training Pipeline has been explained. The workbench to use the training pipelines has been shown and explain. Further, the logic behind the glossary and specific tools has been explained. The team members tested developed elements of the Bonseyes platform, providing feedback. Moreover, adequate support for SDK, libraries, tools, and drivers for each platform is provided among partners. In case that the support is needed outside the consortium, good connection and collaboration with SoC vendor providers will help to get temporary support for the targeted problem.

4.6 Risk: Tools and methods do not work

4.6.1 Risk mitigation

The ultimate goal is to have tools and methods that can be executed on the development platforms. Having defined the algorithms, the training methods, and the deployment of the model on PC, the question at hand is how they will behave on the real platform. To answer the question, we have to have a clear understanding of the marketplace training pipeline and interaction, and what Bonseyes is doing for the use cases. The Bonseyes platform provides an easy deployment from training down to selected board. We have a concrete set of algorithms, as mentioned in 3.1. of this report, that with initial development of marketplace and pipeline, proves the concept. Low power deep neural network (LPDNN) should be the bridge between the tools, and methods and on the board deployment. With LPDNN we use feedback from the board to provide correctness of the model. The feedback is used by the Pipeline to have an evaluation on the results from the board to confirm that the board runs the model with the same accuracy.

4.7 Risk: Marketplace does not have content

4.7.1 Risk mitigation

The partners will help to populate the content of the marketplace during hands-on workshops, in addition to the content already generated as part of WP2 and WP3. This includes the challenges that are identified as part 3.1 of this document.

5. Conclusion

Major risks are applicable to all elements of the Bonseyes platform and should be eliminated to ensure successful end-to-end integration. Initial analyses have shown that gaps in the project are minor in this phase, and that previous delay in WP3 will not impact the final timeline. To ensure smooth development and to increase the success of the final deployment, hands-on workshops are organized. The workshops will help to train partners to use tools, upgrade them per need, and to test the performance of the developed elements of the platform during deployments, before the complete integration. Instead of testing per se, rather we use team work during workshops to improve the features, performance, and user experience of the platform.

This deliverable will influence the ensuing work and future deliverables as follows. The remaining work of T5.1 will consider the recommended risk mitigation actions and pilot the validation protocols until month M24. D5.2 will offer the final report of Demonstrator Proof-of-Concept reporting about the achieved preliminary validation results at month M24. The work on the final use cases during months M25-M36 will be reported with the deliverables D5.3-D5.7. D5.7 will include the results of the validation according to the validation protocol.