Project brief description	EC1 integrates, in a Trustworthy environment prototype, the revisit of classical engineering in view of the effects induced by the usage of AI within critical systems. It federates the integration of technological components developed by the other projects of Confiance within the Trustworthy environment by specifying interoperability, deployment and usage constraints, in consistency with recommendations delivered by projects EC2 and EC6.

IRT Accountable

Last Name, First Name	Function	Email	Phone

IRT Approvers (Internal)

Last Name, First Name	Function	Email	Phone

Modifications history

Version	Date	Comments
1.0	2021 April	First version
2.0	2022 September	Update with Q1->Q3 batch 2 use-cases

Authors

Nom Prénom	Entité
BOSCA Amélie	Soprasteria
FOULLIARON Josquin	IRT SystemX
POUJOL Michel	SopraSteria
GARDET Caroline	SopraSteria
TSCHIRHART Fabien	IRT SystemX

Confiance AI - Project EC1
Industrial Use-Cases

Deliverable 1.3.1.2 : Industrial use-cases handling specification
Deliverable 1.3.2.2 : Industrial use-cases required tools and report
Deliverable 1.3.3.2 : Shared industrial UC data, models and report

TABLE OF CONTENT

1. INTRODUCTION 5

1.1. PROGRAM CONTEXT 5
1.2. PROJECT EC1 5
1.2.1. WORK PACKAGES AND TASKING 5
1.3. DELIVERABLE DESCRIPTION 6

2. DEFINITIONS AND OBJECTIVES 7

3. INDUSTRIAL USE-CASES IN THE PROGRAM 8

3.1. USE-CASES THEMATICS 8
3.1.1. 2D VISION 8
3.1.2. ANOMALY DETECTION 8
3.1.3. PREDICTION 8
3.1.4. NATURAL LANGUAGE PROCESSING 8
3.1.5. OBJECT TRACKING 9
3.1.6. REIDENTIFICATION 9
3.1.7. SURROGATE MODEL 9
3.1.8. VISUAL INSPECTION 9
3.2. USE-CASES SUB-THEMATICS 9
3.2.1. HYBRID 9
3.2.2. IMAGE 9
3.2.3. SYMBOLIC 9
3.2.4. TABULAR 9

3.2.5. TEXT 9
3.2.6. TIMES SERIES 9

4. USE-CASES INTEGRATION AND PROVISIONING PROCESS 10

4.1. REVIEW OF THE SUBMISSION BY THE USE-CASE WORKING GROUP 10
4.2. USE-CASES INSTRUCTION PROCESS 11
4.3. QUESTIONARY TEMPLATE PRESENTATION 11
4.4. IDENTITY 11
4.5. DATA 11
4.6. MODELS 12
4.7. VALIDATION 12
4.8. DEPLOYMENT AND SUPERVISION 13
4.9. ORGANIZATION 13

5. INTEGRATED USE-CASES 14

5.1. PRIMARY USE-CASES 14
5.1.1. AIR LIQUIDE: DEMAND FORECASTING 14
5.1.2. RENAULT: WELDING INSPECTION 14
5.1.3. VALEO: SCENE UNDERSTANDING WITH 2D PERCEPTION 16
5.1.4. AIRBUS: ACAS XU 16
5.1.5. AIR LIQUIDE: CYLINDER COUNTING 18
5.1.6. ATOS: VISUAL SIMILARITY 18
5.1.7. RENAULT: OPINION MINING 18
5.2. SECONDARY USE-CASES 19
5.2.1. THALES LAS: AERIAL PHOTOGRAPH INTERPRETATION 19
5.2.2. SAFRAN: VISUAL INDUSTRIAL CONTROL 19
5.2.3. NAVAL GROUP: ANOMALY DETECTION 20
5.3. SUPPLEMENTARY DATASET 20
5.3.1. EMS DATA QUALITY DATASET 20
5.4. SELECTED FOR MATURATION 20

6. USE-CASES MATURATION 22

7. USE-CASES MANIPULATION TOOLS & ENVIRONMENT 23

7.1. DATA AND DESCRIPTION 23
7.2. DATA & MODEL MANAGEMENT TOOLS 23
7.3. RESULTS COMPARISON AND ANALYSIS 23

8. REFERENCES 24

1. INTRODUCTION

1.1. Program context

The goal of Confiance.AI program is to address the multiples challenges that arises from the integration and the use of artificial intelligence processes in critical systems. The program will deliver an environment, the Trustworthy environment; a set of interoperable technological components, each one covering one or several of the specific challenges previously outlined. This environment will be articulated as instantiable tool chains whose use will be defined in a set of methodological guidelines. It will allow the design, the development, the validation of artificial intelligence modules prior to their integration and support into critical systems.

1.2. Project EC1

Among the seven projects that make up the Confiance.AI program, project EC1 “Federative Environment” holds the specific task of federating the work and results of the other projects. This implies, not only the consolidation and the implementation of the technological components in line with the recommendations and guidelines built by EC2 “Processes & Methods”, but also its end-to-end execution for the design of AI modules and their application on selected industrial use cases.

To this end, EC1 will implement an interoperable **Trustworthy environment** prototype. Made available to the entire program, this environment will be used for different purposes:

- Support the continuous development and integration of technological components,
- Assess the added value of components and tools on representative industrial use cases,
- Maturate shared industrial use cases by the means of selected components and tools,
- Allow the revisit of classical engineering in view of the effects induced by the usage of artificial intelligence module within critical systems.

In addition, as the running and orchestration of the Trustworthy environment requires a specific framework to correctly execute the successive steps of the model generation process, EC1 is also responsible for the deployment and provision of an **execution environment** that will ensure the proper execution of every component of the Trustworthy environment.

Finally, EC1 also ensures that Confiance.AI program has the proper tools to collaboratively work at a distance: which implies not only a dedicated information system and tools but also an operational infrastructure running the tools on the one hand and the execution environment on the other. This set belongs to the so-called **program environment.**

1.2.1. Work packages and tasking

Project EC1 is structured around six work packages, each work package is carried out through tasks ranging from two to five.

- WP1: Trustworthy environment engineering tools
  - T1.1. Positioning in relation to industrial environments
  - T1.2. Implementation of the necessary engineering tools

- WP2: Program environment
  - T2.1. Provision of infrastructure & information system
  - T2.2. Provision of the collaboration, project management and development environment

- **WP3: Provision of use-cases in the Trustworthy environment**
  - **T3.1 State of play of industrial use cases in the Trustworthy environment**
  - **T3.2 Implementation of shared industrial use case interfaces**
  - **T3.3 Maturation of shared use cases**

- WP4: Integration and federation of components
  - T4.1 Implementation of an interoperable Trustworthy environment
  - T4.2 Integration of tools and methods from existing initiatives
  - T4.3 End2End evaluation on representative use cases

- WP5: Sustainability and Trustworthy environment support
  - T5.1 Coordination of sustainability strategy for the components of the Trustworthy environment
  - T5.2 Release and documentation of the Trustworthy environment
  - T5.3 Support and training
  - T5.4 Valorisation and choice of business model
  - T5.5 Community building around the Trustworthy environment

- WP6: Development of integrative components
  - T6.1 Identification of trust property aggregation methods
  - T6.2 Construction of trust properties on use cases

1.3. Deliverable description

This document summarizes the activities of provision and maturation of use-case within the program:

It first presents the role of use cases before describing the process used for their selection (identification of their relevance to the program), their instruction and the preparation for their integration into the Trustworthy Environment. Finally, it lists and describes the tools needed to handle the industrial use-cases :

- means of access to use case data and information,
- management of data sets and models,
- tools for comparison and analysis of results.

2. DEFINITIONS AND OBJECTIVES

A use-case is a set containing:

- One problematic given by an industrial
- One or many datasets proposed to deal with this problematic
- A methodology that has been already tested to begin to solve the problematic and that will be used as starting point for works of program confiance.ai

The first objective is to instruct each selected use-cases in order to:

- Know the technical properties of the use-case to prepare the trustworthy environment to receive artefacts.
- Ensure that the use-case has enough artefacts to be usable by the different project of program.ai and estimate the integration workload

Let us recall that the objective of the program confiance.ai is not to solve the use-case starting from nil but to focus on works on "trustworthy". In that context, analyze the completeness of a use-case consists in checking if it its artefacts contains:

- A dataset to learn the models
- Some code to learn a model, or at least some pretrained models
- Some code to perform inference computation on the models

If too many elements are missing or if there is a significant modification of the use-case content from what have been initially proposed, and the final decision come to the steering committee to keep the use-case integration in the ongoing batch, report it to another batch or simply exclude it, depending on the estimated integration workload.

3. INDUSTRIAL USE-CASES IN THE PROGRAM

The industrial use cases correspond to concrete situations in which program partners have raised problematics where trust is needed. They represent shared needs by the industrial partners and thus are used as a guideline for the program to guide the works. They provide a way of gaining a better understanding of the needs of the industrial partners by giving illustrations of these needs, to identify constraints and to choose adapted tools. They govern the choices, whether technological or functional. During the development phase, they ensure the tests of the components developed by all the collaborators participating to the program. Finally, they demonstrate the benefit of the Confiance.ai program by giving concrete examples of the use of integrated components as well as the end-to-end use of the platform.

Use-cases are divided into two categories: primary and secondary.

- Each primary use case has been selected to represent a thematic, a specific typology of problem assuming that it gathers trust problematics of a majority of industrial partners. Its approach must also be generalizable so that it can be easily transferable to another field.
- Secondary use cases have been introduced to complete primary use cases. They are facing the same thematic but the purposes are different, particularly regarding program expectations: a component or a method tested or realized with a primary use case must then be tested on a secondary use case of the same theme.

The ideal situation is of course to have, for each theme, a primary use case used to perform the work, and a secondary use case that allows the validation of the work.

3.1. Use-Cases Thematics

3.1.1. 2D Vision

2D Vision is about locating and identifying things in 2D images, and eventually tracking them in videos which can be interpreted as sequences of 2D images. The corresponding use cases often deals with data recorded from cameras and aim to perform a task which is similar to what our brain is doing when it receives data from our eyes.

3.1.2. Anomaly Detection

The detection of anomalies [1] is declined into several approaches, according to the nature of the anomaly sought:

- Misclassified samples: Objects that are likely to be misclassified and that fall near the decision boundary where the classifier is uncertain.
- Novelty detection: Test points that could be new observations, i.e., the equivalent of an outlier for the test data. An example would be a new rare breed of dogs for a dataset of dog breeds.
- Outlier detection: Data points from the training set that are far from the others, i.e., an unusual or noisy training sample.
- Out-of-distribution detection: Objects that are drawn from a distribution different from the training distribution. - high-

3.1.3. Prediction

As its name indicates, problems related to prediction involve making statements about the state or evolution of a value or a system at a given moment in time. In the framework of a system whose objective is the realization of prediction, it refers to the output of model which has been trained on data sets consisting of historical data and then applied on new or live data so that it can estimate the potential outputs;

3.1.4. Natural Language Processing

Natural Language Processing (NLP) [1], as the main task of communication, refers to the machine's ability to identify, process, understand and/or generate information in written and spoken human communications. It is considered as an AI subdomain from several national strategies and AI experts, encompassing applications such as text generation, text mining, classification, and machine translation

3.1.5. Object Tracking

Object tracking refers to the specific vision task which correspond to be able to localize an object and its movements through time in a video.

3.1.6. Reidentification

Reidentification is the specific vision task which consist in recognize an object or a person in two photos or videos (frames).

3.1.7. Surrogate Model

Surrogate modelling in the confiance.ai program deals with replacing models by another one, under the hypothesis that it presents more advantages for the task and/or the context of use. In particular, to embedded an algorithm with resources constraints, compression can be done which is a form of surrogate modelling.

3.1.8. Visual Inspection

Visual Inspection is about detecting anomalies on object using only images of these objects. It's a non-destructive method to analyse for instance quality on production chains.

3.2. Use-Cases sub-thematics

3.2.1. Hybrid

Hybrid refers to methos used to tackle a use case, which is in this case a mix between a machine learning method and a symbolic method which are mixed together. The goal can be either to insert knowledge in a machine learning method, either to use a machine learning method to allow the symbolic one to perform well.

3.2.2. Image

An image is a visual representation of something, informatically stored with pixels with spatial consistency. Tasks like object detection, object segmentation, tracking, reidentification, anomaly detection and a others can be considered.

3.2.3. Symbolic

Symbolic refers to method which are used to answer a problematic (regression, classification). It gathers all methods which are rule-based, so completely analytical. The term can be opposed to machine learning. It's often use when business rules are weel-known, as one can use a priori values.

3.2.4. Tabular

Tabular data refers to numerical or categorical data which can be ordered in a table. Tasks like regression or anomaly detection are classical ones with this type of data.

3.2.5. Text

Text is a particular type of data which is composed of words to produce a semantic. It can refer to various languages and have many different syntaxes. Tasks like sentiment analysis, translation, text generation or automatic summary are examples of what one can do with text.

3.2.6. Times Series

This sub-thematic gather all use cases for which numerical data with a temporal dimension is at stake (so videos are excluded and match 2D Vision thematic instead). Various task can be considered, as forecasting, anomaly detection or classification.

4. USE-CASES INTEGRATION AND PROVISIONING PROCESS

The integration and provision of a use case within the program is based on a dedicated process. This process guarantees not only the functional interest of the use case pushed into the program but also that the elements that will be made available to the teams are effectively usable and contain the technical elements necessary for their proper exploitation by the teams. Following figure 1 show the different steps of this process.

4.1. Review of the submission by the use-case working group

The process begins with the presentation of the use case to the Use-Case Working Group (GTT[1] Use-Case) by the industrial partner who wishes to push it into the program. The industrial partner is then represented by a use case holder who first presents the use case from an operational point of view, the typology of the problem and finally the technical challenges that emerge from the integration of artificial intelligence within the system.

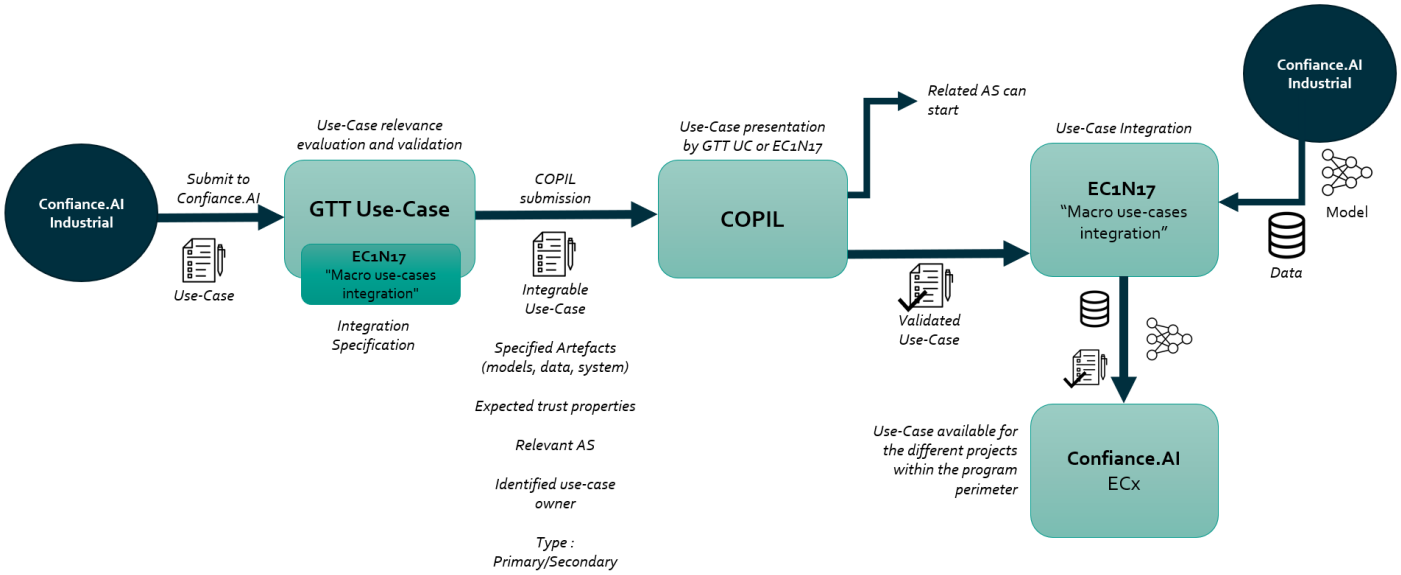

Fig1. Use-Case review and integration process (Batch 2)

The working group then assesses the relevance of the use case to current or future activities, to this end it is associated with the EC1 project which will conduct an instruction to analyses whether the prerequisites are respected or not, in other words whether the use case will be effectively integrated and usable:

- What are the artefacts (identity, models, data, validation criteria, target system) specifications?
- Which trust attributes are expected?
- Are there any ongoing or upcoming activities for which this use case is interesting?
- Is there a use-case holder and is he able (availability and role) to represent the operational stakeholders and to cover the technical aspects and constraints of the system?
- Is the use-case eligible as a primary or a secondary use-case?

Once this first step has been successfully completed, the use case is presented to the steering committee which must validate its integration and availability in the program. Once this first step has been successfully completed, the use case is presented to the steering committee which must validate its integration and provision within the program.

Finally, the industrial partner will provide the elements presented above to the EC1 project which will verify that the use case works before making it available to the program teams.

[1] Groupe de Travail Transverse, french for Transversal Working Group

4.2. Use-Cases instruction process

Instruction has been done by interviewing the use-cases owners. For each use-case, goal of the interview is to

- *Understand the use-case problematic*
- *Understand the use-case owner expectation about the work produced in the program confiance.ai*
- *List all available artefacts (datasets, code) and their technical properties to prepare the integration*
- *Know the main properties of the methodology that have already been tested to deal with the use-case and that will be the starting point for the program works.*

It was decided that each interview will be conducted by at least two people from the team. To conduct the interviews, a unique questionary template has been made.

4.3. Questionary template presentation

The aim of this questionary is to list in an Excel spreadsheet every question that could be potentially useful for the integration of the use-cases to finally have a synthetic vision of all its significant properties.

The set of questions is divided into five mains thematic:

- Identity
- Data
- Models
- Validation
- Deployment and supervision

Following sections describe the content of each of them.

4.4. Identity

Purpose of this section is to get general information about the use-case

			examples of answer
Identity	UC name		
	UC type		*Object detection, predictive mainetnance, decision support, …*
	Type of problematic		*Computer Vision, NLP, Speech Recognition, Time Series processing,*
	Type of solution		*Classification, Regression, Feature Learning, Representation, Exploration,*
	Type of algorithm		*Machine Learning, Deep Learning, Symbolic AI, expert systems, multi-agents systems, Optimization, …*
	Training		*Supervised, unsupervised, Semi-supervised, Renforcment, transfert learning,*
	TRL		
	Problematic to solve ?		
	Safety criticity ?		
	Operationnal criticity ?		
	Partner expectations ?		

Table 1. Use-case identity description

4.5. Data

This section is divided into two parts:

The first one dealt with properties of the available datasets. Second one is about the properties code for datasets reading and pretreatments.

				Commentary	
Data	Data type (ST, Pictures, Tables,..)				
	Data size				
	Data labelling (metadata)				
	Data availability - how to get it?				
	Condition to get the data ?				
	Can data be stored in our internal server ?				
	Data quality (risk of error)				
	Are there censored data ?				
	Data typology (real, simulated)?				
	Data representativity in future ?				
	CODES	**Couverture**	**Used technologies (languages, frameworks, functions, ...)**	**Etat courant**	**Commentaries current state and/or other useful informations**
	Data reading				
	Storage in memory of read files				
	Data reading in batch mode				
	Data reading in streaming mode				
	Pretreatment				
	Cleaning/ correction				
	Formatting/ datasets making				
	Data augmentation				
	Other useful information on data				

Table 2. Use-case data description

4.6. Models

This section has two parts:

The first one focus on information related to the methodology used to learning the ML model. The second one interests to the code used to learn the ML model and its properties.

				Commentaries	
Model	Chosen model ? Which formatt?				
	Environnment/ package used ?				
	Modelization hypothesis?				
	Data split train/test ?				
	Chosen features ?				
	CPU/GPU required for training ?				
	RAM required for training?				
	Training duration?				
	CODES	**Couverture**	**Used technologies (languages, frameworks, functions, ...)**	**Etat courant**	**Commentaries current state and/or other useful informations**
	Pre trained model reading				
	Traning by transfert learning				
	Active learning				
	Search and optimization of hyperparameters				
	Display of metrics evolution				
	Models saving				
	Other informations related to model training				

Table 3. Use-case models description

4.7. Validation

Questions in this section concerns the existing processes to validate the models and the methodology used to solve the use-case

						Commentaries
Validation	Requirment for "trust" ?					
	Identified limitations , and challenges?					
	Current level of "trust" ?					
	CODES	**Couverture**	**Used technologies (languages, frameworks, functions, ...)**	**Etat courant**	**Commentaries current state and/or other useful informations**	
	Validation « classique » (testset ou holdout)					
	Available scenary for testing ? Execution environment ?					
	Validation croisée					
	Interpretability / explicability					
	logical/formal demonstrations ?					
	Performances (recall, précision, RMSE, ...)					
	Test validity in the future?					

Table 4. Use-case validation criteria description

4.8. Deployment and supervision

This section has questions about the technical properties of the production environment

						Commentaries
Deployement and supervision	Is the UC executed in real time					
	Computation frequency?					
	Use frequency?					
	Execution time constraints ? Result computation time constraints ?					
	CODES	**Couverture**	**Used technologies (languages, frameworks, functions, ...)**	**Etat courant**	**Commentaries current state and/or other useful informations**	
	System integration					
	System deployment					
	Supervision					

Table 5. Use-case deployment and supervision modality description

4.9. Organization

Organization	People available from partner compagny to help for the usecase integration ?	
	Artefact transfert system ?	

Table 6. Use-case organization

5. INTEGRATED USE-CASES

5.1. Primary use-cases

Each primary use case has been selected to represent a thematic, a specific typology of problem assuming that it gathers trust problematics of a majority of industrial partners. Its approach must also be generalizable so that it can be easily transferable to another field.

5.1.1. Air Liquide: Demand Forecasting

Description

This industrial use case belongs to the “Time Series” thematic, with a prediction typology.

The goal of this use case is to make a 1-month prediction of the Air Liquide customer in order to anticipate and organize the production and the distribution. To do so, Air Liquide wants to predict this demand with a confidence interval on the prediction. The main objective in the Confinace.ai program is to guaranty reliability of theses prediction intervals.

To work with this problematic, Air Liquide has provided time series data as a single “.parquet” file. It contains all the interest variables of the problem including the customer demand to predict. Because this data are sensitives, they have been anonymized. Variables could be numerical but also categorical.

In order to use this data within a machine learning process, codes are disponible to train models and test them. A Jupyter Notebook allows to launch any part of the code. The training produces two different models. The first one is a gradient boosting implemented with the xgboost python library which is capable to produce a single demand prediction. The second one is made of three sub-models which predict customer demand and confidence interval bounds (low and high) respectively. These three models are implemented with the xgboost.sklearn python library.

The expectation about confiance.ai work is to compute a confidence interval around the predictions.

Artefacts

The artefacts available for this use-case are:

- Anonymized dataset file containing demand history and context data.
- A code to read the dataset, learn a model with gradient boosting, and perform predictions from the learnt model. A first attempt of prediction with interval bound is done.

The code environment is python using Jupyter-notebook, Xgboost, and Sklearn. Size of the dataset is 350 Ko

5.1.2. Renault: Welding Inspection

Description

This industrial use case represents the “Visual Inspection” thematic.

The problematic of this use case is to qualify the state of welding (OK or KO) from photo taken by cameras on vehicles production chains. The good quality of welding in vehicles on production chain output has a **legal dimension** at the level of the factory direction. Thus, the industrial wants to be sure to have enough "trust" about AI process used for welding defect detection, in order to industrialize the method. This is why Renault is interested through the program in results related to monitoring, robustness under perturbations, explicability and definition of quality metrics.

To work with this problematic, Renault has provided a dataset of about 15000 pictures of welding in full HD (approximatively 2 To of data). These images are labelled to indicate if the welding is OK or KO according to experts. The dataset contains different acquisition conditions, on many chains and different welding such as some of this welding are more critical than others and

cameras are not all the time in the same position. Vibration on a chain could also be different than on another chain. The context data linked to the properties of production installation are specified. It is interesting to note that the dataset is unbalance, as only about 5% of images are characteristic of a defect. Renault also made available about 60000 non-labelled images.

Renault provided exported TensorFlow trained models from Auto ML function of Google Cloud Platform. One model is available by welding beam. A code has been added by the program in order to perform inference on those models. Sklearn SVM models are also available with a Jupyter Notebook provided by Renault Digital in order to infer on this others models.

Significative properties of this use-case are as it follows:

- The dataset is unbalanced (there less than 5% of welding with defects)
- Pictures are requalified with 3 types of labels (OK, NOK, UNKNOWN).
- There are many types of welding with variable criticity.
- Welding position in picture can be different according to such vibration on production chain
- The use-case environment is going to slightly change (new weldings, more cameras, new axles)
- “Trustworhy” has a legal dimension at the level of factory direction

Expectation from confiance.ai works are:

- Work on robustness of methods under perturbations (impact of photos resolution reduction, vibration effects)
- Find high quality metrics (acceptability threshold, deal with indecision cases)
- Work on explicability
- Definition of “trustworthy”

Other subject of research:

- There can be errors in picture labeling, how to detect them?
- There are symmetry properties between some welding, how to use them (transfer learning)
- Study data representativity, and the effect of the use-case evolution on trustworthy.
- Monitoring to detect drifts on new pieces.
- Use different performances metrics depending on welding criticality.

This use-case has an operational constraint, the computation time to qualify a welding has to be less than 1/12 second.

Artefacts

The artefacts available for this use-case are:

- A dataset of 100 000 pictures of welding in full HD (approximatively 2 To of data)
- Context data linked to production installation
- Some pretrained models on the provided dataset

A convolutional neural network is used. The learning and the prediction are done using google cloud platform.

5.1.3. Valeo: Scene understanding with 2D perception

Description

This industrial use case represents the “2D Vision” thematic.

This use case is based on perception and understanding of a scene from a camera sensor which is a fisheye camera on cars. It is a computer vision problem using deep learning models to detect and identify objects. For now, the goal is to perform identification on pictures but the final purpose is to do that task on videos, or in other words, on temporally coherent sequences of pictures. The criticality of the problematic is high since this model will be part of a system that will manage emergency braking for the autonomous car. The Confiance program is expected to accompany Valeo on subjects such as data annotations, interpretability, model construction and integration pipeline and robustness against adversarial attacks.

The dataset used to deal with this use case is the WoodScape dataset by Valeo. WoodScape comprises four surround-view cameras and nine tasks, including segmentation, depth estimation, 3D bounding box detection, and a novel soiling detection. Semantic annotation of 40+ classes at the instance level is provided for over 10,000 images. Through the program, only annotations of 6 classes are used. Valeo provided also another dataset to deal with the image annotation problematic, including images with annotation errors.

Two baseline models are provided by Valeo to perform detection and identification of 6 types of objects, based on classical architectures ResNet18 and ResNet50 implemented in Pytorch. The DeepLab has made available another model they have implemented, also in Pytorch. Full codes are shared to annotate data, train and infer on these models and visualize examples of results.

The expectation about confiance.ai work is to improve metrics and score function to evaluate the “trust”, work on interpretability, robustness particularly in case of adversarial attack, real and unseen data.

Artefacts

The artefacts available for this use-case are:

- The public dataset Woodscape created by Valeo containing fish-eyes photos in urban scenes with annotations
- Valeo Deep Perception dataset and a corresponding code to perform semantic segmentation
- The open dataset European persons provided by the university of Delft.
- A private dataset from Valeo that will be available only in 2022 because it requires some pretreatment (blur some parts of pictures) in order to satisfy the GDPR:
- A code provided by Valeo for the dataset reading, the learning of models, and the inference process (should be provided at the middle of May 2021

The size of the complete dataset will be at least of 40 Go and the provided code environment is python and Pyorch.

The type of models used are convolutional networks and their learning require GPUs, 12go of RAM ant takes approximatively between 2 days and 1 week.

5.1.4. Airbus: ACAS XU

Description

This industrial use case represents the “Surrogate Model” thematic.

ACAS Xu is an Airborne Collision Avoidance System for UAV (Unmanned Aerial Vehicle). It is dedicated to decision support: ACAS Xu is solving conflicts with cooperative traffic using vertical motions and with non-cooperative traffic using horizontal motions. Given the current status (distance, bearing angle, current headings, …) of a UAV (called the "ownship") with respect to another aircraft (called the "intruder") with which it could potentially collide, the ACAS indicates the operation (new heading) it must

perform to avoid this collision. To do that, ACAS tables defines the score policy by mapping possible states to actions. These tables are too large to be implemented in many embedded systems. The use case aims to replace these tables with a much lighter, on-board model for which the problem of certification then arises as the ACAS tables corresponding in operations which have been precalculated and validated (certified). Airbus is expected within the program to progress on this validation and certification aspect.

The dataset in use contains 25 tables, which represent 4 GB of data. They represent the policy Valeo wants to encode in neural networks. These tables are delivered by the RTCA (Radio Technical Commission for Aeronautics) and are named “Lookup tables”.

There have been many attempts to replace the ACAS tables by a model. Some works were done at Stanford which has published certain models that are therefore publicly available. Stanford models are multi-layers perceptrons, with 6 fully connected layers and 50 neurons per layers. A code allows to infer and test these models. In addition, some new models have been developed by the IRT Saint Exupery as part of the ANITI/DEEL project. They are multi-layers perceptrons with 7 fully connected layers and 50 neurons per hidden layers, for which the full code is available to train and test.

The expectations from confiance.ai is to be able to transpose the certification of the ACAS table to the surrogate model.

Artefacts

The artefacts available for this use-case are:

- The ACAS XU Tables which require a subscription to RTCA (Approximatively 4 go of data),
- Some trained models provided by Stanford university in Nnet format. Python package to read the model are provided.
- Some code done inside DEEL project by Irt Saint-Exupery (format and environment are unknown)

The already used models are multi-layers perceptron's with 6 layers and 50 neurons by layers. The training requires GPU and the training duration is about 2 of 3 days.

About the validation process, Stanford, proposes ten properties to check to ensure that the AI model is correct.

5.1.5. Air Liquide: Cylinder Counting

Description

This use case is about object identification and tracking. The goal is to count gas cylinders getting in and out a warehouse using video taken by RGB cameras.

Confiance will deals with embarcability, robustness and monitoring purposes.

Artefacts

Air Liquide has provided:

- 200 videos splitted in Full HD frames (with restricted access to the persons who have made a specific demand)
- Pytorch code for detection, and counting tracking
- Trained Pytorch model

5.1.6. Atos: Visual Similarity

Description

The problematic tackled here is the visual similarity learning for person re-identification. It's about the capacity to identify someone (or an object in a similar way) detected in one camera in another camera's field. Although it's a computer vision problem, it implies different evaluation metrics and learning objectives than classical images machine leaning issues. Data it's also critical because of ethical and privacy purposes, that's why synthetic data are particularly relevant in that context. Domain adaptation is de facto a very linked subject. Confiance will also deals with explicability and robustness issues.

Artefacts

The artefact for this use-cases are:

- 9 public datasets
- Modified ResNet to tackle Re-identification problem, implemented in Tensorflow: training and test code and a trained model

5.1.7. Renault: Opinion Mining

Description

The problematic of this use case is to qualify the comments left by customers on Renault stores, which means perform sentiment analysis. It's a Natural Language Processing problem.

Confiance is about to study robust sentiment analysis, which means in that context be able to understand irony for example. Explicability is also a perspective.

Artefacts

The artefact for this use-cases are:

- A dataset of 700 anonymized comments
- A pretrained camemBERT model on the provided dataset and the code to reproduce the training and testing. The learning can be done with wandb but also without.

5.2. Secondary use-cases

Secondary use cases have been introduced to complete primary use cases. They are facing the same thematics but the purposes are different, particularly regarding program expectations.

5.2.1. Thales LAS: Aerial Photograph Interpretation

Description

This industrial use case reaches the "2D Vision" thematic.

This use case deals with the detection of objects, such as buildings, vehicles and amongst other, people (if the resolution allows it) on aerial photographs. It is a computer vision and classification problematic using deep learning models and supervised training methods. The task is intended to operator assistance and therefore has a low to medium criticality. However, it has high real time stakes. This is why the main expectations from the program are about on-boarding, in particular maintaining performance when switching to embedded mode. Thales is also interested on works on interpretability of results, domain adaptation and trust characterization.

The dataset used XView which is public. XView is one of the largest datasets of overhead imagery. It contains images from complex scenes around the world, annotated using bounding boxes with 1 million object instances, 60 classes and 0.3-meter resolution. For the use case, Thales has defined a 4-classes ontology ("small vehicle", "large vehicle", "aeronef", "boats") by re-mapping some of the xView classes. Only the train set of xViex is used, as the validation one doesn't contain annotations.

Thales propose to use YoloV5 architecture to detect Vehicules, as a baseline. The code is implemented in Pytorch and remains very close to the Ultralytics original implementation. It allows to train, infer and test the performance of the network and visualize the results.

The expectations from work of program confiance.ai is maintain/improve performance when switching to embedded mode (which implies model reduction), study domain adaptation from XVIEW to WAMI dataset and help to characterize confidence, with relevant metrics.

Artefacts

The artefacts available for this use-case are:

- Public XVIEW dataset, which is one of the largest publicly available datasets of overhead imagery. It contains images from complex scenes around the world, annotated using bounding boxes
- WAMI Thales dataset
- YoloV5 implemented in Pytorch with code to train and test the model
- YoloV5 trained in Pytorch

5.2.2. Safran: Visual Industrial Control

Description

The goal of this use-case is to perform out of domain detection from photo of industrial pieces.

The expectations from work of program confiance.ai is to improve the TRL, benchmark the method used with others and evaluate its reliability. Another important point is to ensure that an alert is raised when the detection algorithm has a dysfunction.

Artefacts

The artefact for this use-cases:

- A set of photos with cracks (496 different types), and another set without cracks (total of 20 000 photos for a total of 1 Go of data volume)
- Some trained models.
- A code to learn a model of type “real NVP” and to perform the inference for the detection. This code is written in python using Pytorch.

The learning step require 3 GPU, 16 go of RAM and 3 days, to train a model on data.

5.2.3. Naval Group: Anomaly Detection

This use case deals with anomaly detection on time series data. All Naval Group ships (surface vessels or submarines) are equipped with thousands of sensors of different types. Among these, vibration sensors are distinguished by the fact that they must meet 2 needs:

- on the one hand, to detect or prevent breakdowns,
- on the other hand, to ensure that the acoustic stealth of the ships has not deteriorated.

The detection of machine malfunctions using vibratory sensors is a very common problem and often one of the most difficult to solve. Indeed, vibratory sensors have a very high sampling frequency and recordings can be instable and fluctuating. Most of the time, useful information is buried and there are a lot of non-linear interactions due to vibratory waves.

Naval Group has provided a dataset with time series data and corresponding anomaly labels through time and a simple detection model as a baseline.

The use case is about robust anomaly detection, monitoring and explicability.

5.3. Supplementary dataset

Some others datasets are available in the program in order to be used by some projects to deal with specific problematics or data properties.

5.3.1. EMS DATA Quality Dataset

This is a large-scale time-serie dataset provided by Air Liquide whose main purpose is to feed works about time-series anomalies detection.

The context of this use case is anomaly detection on multivariate time series data recorded from sensors. These sensors are installed in a factory to look after and optimize performances, using different tools that exploit these sensors. A lot of different anomalies can be found and there are a lot of sensors.

The aim of the use case is to detect unknown anomalies.

To deals with the use case, Air Liquide has provided a 27Go dataset with time series data recorded from sensors during a whole year in a factory.

5.4. Selected for maturation

A use-case selected for maturation is a use-case that will be fully tooled, and implemented to the industrial partner. in order to see if it can evolve in an extended version

The planned use-case for maturation are:

- Renault: Welding inspection
- One other use-case, yet to be defined.

6. USE-CASES MATURATION

Some use-cases are to be selected for maturation. The maturation of a use case means two different things:

- first, this use-case will be used to evaluate the end-to-end AI engineering workflow of the trustworthy environment designed within in Confiance.ai program. This evaluation involves the implementation of all the steps, components and methods identified in the program that should enable the design of trusted AI-based systems. This work shall allow the evaluation of the coverage and maturity of the trust environment: identification of trust attributes, metrics, methods and adapted components, characterization and comparison of results...
- second, the use case will be completed on the basis of the methodological work of defining an AI-based system: specification, requirements, risks, evaluation...

The maturation of a use case involves the following activities:

- Specifying the problem following the approaches pushed by the EC2 and EC6 projects.
- Problem analysis, functional constraints, requirements, ODD.
- On this basis, identification of relevant available components.
  - AI characterization and qualification component.
  - Component allowing the robustness, explicability, repeatability of the AI component.
  - Component for the extension of the available data.
- Generalization of the components and adaptation to the use case
- Application of the components for the design of the AI module for the use case system.
- Evaluation of the methods, components and completion of the use case.

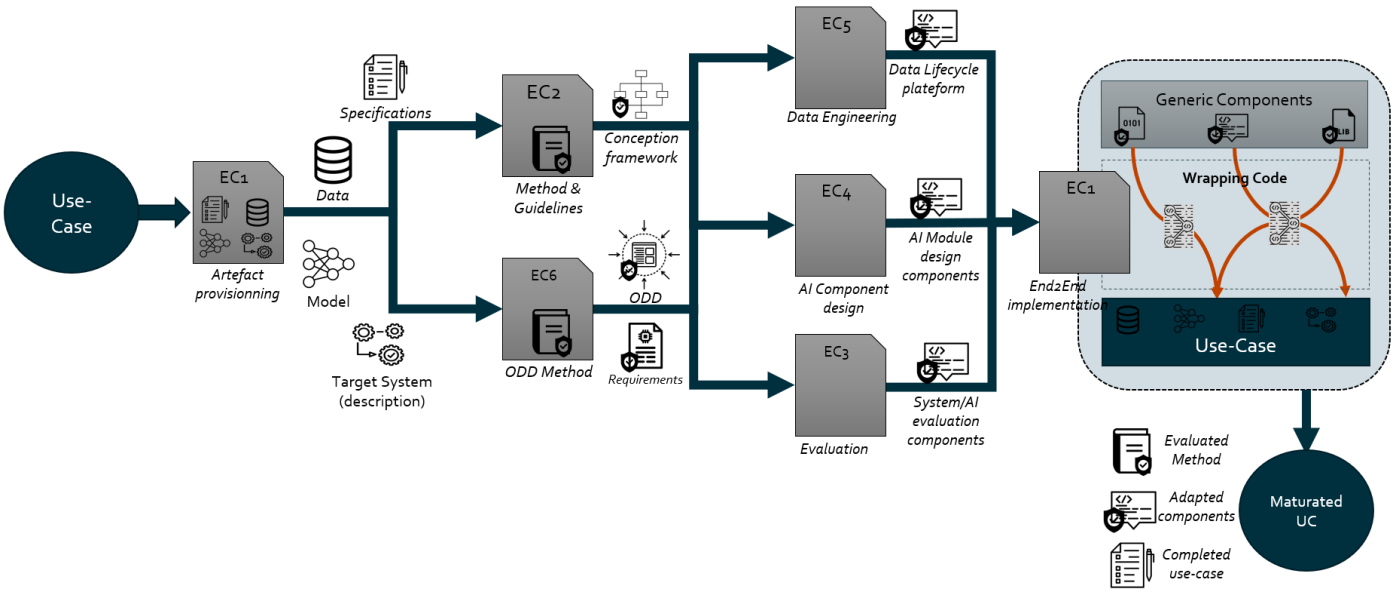

Fig2. Use-Case review and integration process (Batch 2)

7. USE-CASES MANIPULATION TOOLS & ENVIRONMENT

A NFS server on OVH cloud is used to store the datasets and pretrained models. All codes are versioned on Gitlab.

The V1 of the Confiance environment is available. NFS volume is mounted on user environments to allow users to directly access the artefacts of use cases, using JupyterLab.

A virtual machine (sandbox) without GPU for computations is also available and 4 others virtual machines hosted by IRT St-Exupery with GPU can be used.

However, workforce of the program is strongly encouraged to use the V1, in order to test the environment.

7.1. Data and description

In order to work with use cases through the program, data is available to member of confiance.ai program on a Network File System (NFS) server. Already available codes, mainly provided by the bearer if the use case, is shared on a specific git repository, which can be found at https://git.irt-systemx.fr/confianceai/use-cases.

The full description of the data, the codes and how to run and use them is specified on dedicated wiki program pages: https://wiki.confiance.ai/wiki/Use_Cases.

These tools are temporary solutions to allow collaborators to work with use cases, as data and code management solutions have not been integrated yet in the current version of Confiance platform.

7.2. Data & model management tools

The actual and temporary solutions used to manage data (NFS) doesn't include versioning capabilities. Models are also stored on this NFS, so models are not versioned either. The study of tools for configuration management is part of the scope of the program and choices regarding theses aspects will be made to integrated appropriate tools on the environment. However, the code is versioned with git, which remains a temporary solution as the industrial use cases are not dedicated to be integrated in the final product.

7.3. Results comparison and analysis

Batch 1 results on use Renault and Air Liquide primary use cases can be visualized on a Kibana Dashboard which have been produced by the program.

It can be found at opensearch-dashboard.apps.confianceai-public.irtsysx.fr

The DebiAI component has been integrated in the current version of the platform. This tool is able to visualize data and therefore, could help to compare results, for examples performances of many models.

DebiAI is available at DEBIAI (irtsysx.fr)

8. REFERENCES

[1] EC2-Taxonomy, «Taxonomy,» 2022.