DESTINI

SMART DATA PROCESSING AND SYSTEMS OF DEEP INSIGHT

http://www.destini2020.eu

Deliverable 4.5

Report on Real-World Problems Identified for Future Piloting

Document details:

Editor :	Sapienza
Contributors :	Sapienza, CUT
Date:	30/09/2022
Version:	V3.0

Document history:

Version	Date	Contributor	Comments
1.0	15.07.2022	Sapienza	Initial document, structure and content
2.0	15.09.2022	CUT, Sapienza	Additions of problems and pilots
3.0	30.09.2022	Sapienza	Final review & approval

Contents

1. Introduction

1.1 Purpose

This document presents ideas and proposals on real world problems which have been identified for future piloting, beyond the duration of the project. The piloting activities will be based on research outcomes of the project, and have been identified during the joint activities (online meetings, site visits) which have been carried out by the consortium of DESTINI. In particular, all the travels and site visits which happened in 2022, after the end of the COVID-19 pandemic which stopped travels for most of the project duration, have strongly influenced and contributed to this deliverable.

This deliverable is part of Work Package 4 (WP4). WP4 suggests methods and actions followed by the consortium to extend the community and networks of strategic research collaborators, and support sustainability by attracting EU and national funding in the future.

1.2 Overview

The rest of the document is structured as follows: Sections 2 to 5 describe the different real world problems identified for future piloting. Specifically, we envision four main activities, which will be carried out also on the basis of interested stakeholders (cf. deliverable 4.1). Finally, Section 6 concludes this deliverable by outlining a possible timeline.

2. Piloting #1 - Port analytics via IoT-cloud continuum

1. Project Identification	
Name	Port analytics
Acronym	Smart ports
Funding opportunities	EU Horizon Europe, National projects in Cyprus and/or Italy

2. Summary

Sea ports are involved in 80% of world trade movements, where the real time exchange of information, and operational automation in particular, are considered key enablers for building sustainable, resilient, and high quality seaport navigation infrastructures. To aggregate such real-time information and its sharing between the shipboard and the port digital environments requires the ship to self-implement the manoeuvring path to approach the dock. This necessitates a resilient and highly responsive IoT-edge-cloud infrastructure to build a secured e-navigation software based on IMO[1] standards.

The pilot aims at investigating an autonomous navigation for assisted and real-time docking services. To achieve this, we will use a digitally ready port (e.g. Livorno Sea Port, where Sapienza has some collaboration ongoing) including the commitment of all port authorities. A common scenario represents a ship approaching the port basin for docking. As soon as the ship enters the 5G coverage of the port, it is identified by the 5G-IoT infrastructure (Observation). Next, the port pilots board the ship with a Portable Pilot Unit (PPU)) computing device, where a low latency bi-directional stream of data is shared with the shore side. Further, a "Vessel Companion" is rendered (with meteor-marine conditions, coastline and bathymetric chart) at the Port Authority Control Centre, and the position and the optimal track are calculated (Reasoning and optimization) aggregating 5G, GNSS, and sensor signals (including HR infrared-cameras, distance metres, and radars). The "Vessel Companion" information stream is shared in real-time with the ship. The pilots dynamically and constantly check the performance of the manoeuvre using their PPU. This enables the generation of alerts and employ route corrections. Finally, the ship's navigation system combined with the ship companion stream from the Port Control Centre takes action (Intelligent adaptation) to automatically adapt to the correct route.

The pilot provides e-navigation actuation services for the safety operators with a real-time 3d

[1] https://www.register-iri.com/wp-content/uploads/MSC.1-Circ.1604.pdf

rendering service required to be deployed on PPUs, including the shipping basin, berths and freight terminals (Port Digital Twin). Hence, we utilises advanced cloud tools for automated deployment of the ship navigation software services (e.g. MONICA[2] and TPCS[3] service stack). Initially, we trigger a decentralised runtime system for reasoning about how to dynamically adapt the e-navigation manoeuvre, when the ship is in close proximity to the port's 5G network. Further, the envisioned tool orchestrates the e-navigation serverless stack instantiated on the PPUs and the on-board desk devices. Further, the tool provides AI-guided resource recommendations owing to multiple port entries and mooring manoeuvres to changing edge device constraints for lowering the deployment cost while maintaining scalability and availability of navigation services. Essentially, we enables the smart port to manage compute resources and automate deployment of e-navigation service with diverse operational constraints and manoeuvre.

3. Involved partners from DESTINI	
Partner	Country
Cyprus University of Technology (CUT)	Cyprus
Sapienza Università di Roma	Italy

[2] https://www.monicapmslivorno.eu/

[3] https://tpcs.tpcs.eu/

3. Piloting #2 - Process mining in juridical systems

1. Project Identification	
Name	Process mining in juridical systems
Acronym	PM@Justice
Funding opportunities	Italian funding on NRRP

2. Summary

E-government emerged during the ’90s and early 2000s as the usage of ICT to provide public services to citizens and enterprises. Methods and approaches were developed, to define novel digital interactions between citizens and their government (C2G), between governments and other government agencies (G2G), between government and citizens (G2C), between government and employees (G2E), and between government and businesses/commerce (G2B).

The counter-side of this digitalization has been the explosion of many different applications and complex information systems, which represent nowadays a huge bunch of legacy systems which, per-se, cannot exploit novel techniques (as those ones proposed by Artificial Intelligence - AI), in order to be further improved. They are not extensible and pluggable, and investments should be preserved in the long term. As an example, in the Italian judicial system, which has received significant ICT investments in the last decade, we have more than 55 different information systems, with an average age of 20+ years, and the raw transitions of the civil judicial processes account to more than 110 000 ones.

Now Italian courts offices can rely on information systems able to manage the processes driving the trials, but are challenged in exploiting the full potential of the data generated by such systems so as to analyze and possibly improve the processes and ultimately complete more trials in shorter times. This calls for exploiting and improving data and process methods and tools, with the objective of making this data actionable and of measuring relevant KPIs, yielding to improved and effective resource usage some possible research directions for the application of AI-driven data and process science in the judicial domain.

When originally digitizing, the emphasis has been first on representing data and processes with the purpose of designing information systems, starting from their execution traces. Only very recently, the emphasis has shifted towards analyzing data and process executions. Here AI (in particular

methods from knowledge representation and machine learning) could play a major role, provided that they could somehow be plugged over existing legacy information systems, especially considering the following challenges.

- Process mining offers an innovative approach to analyze (judicial) data from a process-oriented perspective, in particular considering the execution of several trials as a whole. In a nutshell, process mining exploits the evidences associated to events occurring during the execution of a process, and stored in so-called event logs, to provide additional insights in the classical business process management phases: i.e., process discovery, analysis, enactment, and improvement. However, events logs are not easily available in legacy systems, which were not designed to record and store events at the proper abstraction levels (which is higher than that adopted in system logs and lower than the one provided in logs of citizens-/employees-visible actions).
  Can we derive process logs, suitable for process mining, without severe technical interventions on the legacy information systems?
  Can we derive events on the "surface" of the system, through the observation of data and events, without deeply understanding the internals of the systems?
- Process analysis requires process modeling, which is typically performed by means of simple specification languages, including statecharts, dataflow, or more recent ones as BPMN. Due to the multiplicity of applications in which data are employed, a more general purpose specification language, equipped with a powerful logic, is needed. In the judicial domain, properties of data about trials are always strictly connected with temporal constraints, therefore the specification language must be able to describe temporal behaviours and be declarative more than control-flow oriented ones: it must be able to describe ordering constraints among events and specific and detailed temporal constraints.
  Can we devise novel approaches to process modeling, which better support process analysis about the temporal facet?
- Conformance checking techniques can be useful especially for a temporal analysis of the processes. When analyzing trials, two main KPIs are of interest: i) Disposition Time (DT), to enable the measurement of the duration of administrative procedures in courthouses and to estimate the impact of delays in terms of costs of procedures; ii) the Clearance Rate (CR), defined as the ratio between adjudged trials and incoming new cases. While such indexes are easy to compute at an aggregate level (for instance, all cases in a period), the challenge is in analyzing details of trials to identify patterns that may be associated with strong deviations of the above KPIs. Concerning the compliance, it is also fundamental to tackle the problem of checking whether the process verifies the correctness of the procedures. This entails checking if the process is following the steps required by the legislation (which can be described with some formalism) and if the software specifications guarantee the alignment with specifications of all paths that terminate, while paths that do not terminate are not

produced.

Can we devise novel approaches to conformance checking and compliance, specifically tailored to the analysis of temporal KPIs, but also suitable to identify anomalies and patterns originating such anomalies? Is explainable conformance checking a reality?

AI techniques can help in addressing the above challenges, in particular those fields commonly referred to as Knowledge Representation and Reasoning (KRR), Machine Learning (ML) and Natural Language Processing (NLP).

1. Document-based event logs generation. The problem of extracting key concepts or relevant data from unstructured documents has recently captured the attention of researchers, seeking to distill knowledge for the disparate corpus of natural language content. Automatic retrieval of temporal information from texts such as verdicts and the creation of a link to the judicial process events could allow the identification of critical points of the process. As pointed out, this would allow identifying events without intervening on the information systems. While this activity is often performed manually, through document annotation, a possible research direction for automatic text processing is the discovery of new procedures for the creation of process models from documents. In this context, a further interesting challenge is how to define the quality of the event log for guaranteeing the significance of the results obtained with the process analysis. In addition to the information offered by the textual documentation made processable by natural language processing techniques, an important source of information is offered by the logs of the programs used to manage the process files. These programs, built starting from finite state automata that describe the possible state transitions of the processes, generate logs that describe the series of state transitions of the trial documents (dossiers) and that, in some way, also tell the history of the processes. In this context, data may take the form of infinite data streams, rather than finite stored data sets. Several aspects of data management need to be reconsidered in the presence of data streams. We aim to focus primarily on the problem of defining methods to describe and verify properties on data, capable of alerting the user if some of these properties are violated.
2. Process modeling and temporal analysis. Process modeling of courthouses administrative procedures deserves attention to study the proper representational formalism to include different perspectives, and a combination of knowledge-driven and data-driven techniques in logs analysis, to take into account the semantics of logs due to existing law regulations and limits. Knowledge representation approaches, based on ontologies, temporal logics and declarative process models, still maintaining reasoning tractability, may offer the opportunity to better infer interesting facets. Clearly, in order to ensure seamless adoption of novel approaches, validation studies with justice practitioners is needed. Once processes are modeled, the current approaches in judicial process analysis are mainly addressing the identification of temporal outliers using process mining techniques focusing on single cases. Possible

new research directions focus on the analysis of processes more in general, analysing multiple cases for predicting the effect of improvement actions. Such analysis and predictions can successfully exploit machine learning techniques.

3. Process simulation and digital twins. Techniques are needed in order to assess the impact of possible improvement actions on trials in general, and how to better use resources. For instance, a possible direction that could be of interest and could be adapted to trial process analysis is to combine process mining together with significant events simulation to propose an approach based on digital twins to evaluate the impact of improvement actions, including process model changes and resource allocation.

The role of AI techniques is significant in many of the directions proposed above. First of all, process mining tools can learn the actual structure of processes and support the comparison with modeled processes. Natural Language Processing (NLP) can support the extraction of information about events both from documents and from existing unstructured process logs. Finally, data-aware ML can support the identification of structural anomalies in process execution.

3. Involved partners	
Partner	Country
Cyprus University of Technology (CUT)	Cyprus
Sapienza Università di Roma	Italy

4. Piloting #3 - Electrospindle 4.0

1. Project Identification	
Name	Zero defect manufacturing in spindle production
Acronym	Electrospindle 4.0
Funding opportunities	EU Horizon Europe

2. Summary

Industry 4.0 represents the last evolution of manufacturing. With respect to Industry 3.0, which introduced the digital interconnection of machinery with monitoring and control systems, the fourth industrial revolution extends this concept to sensors, products and any kind of object or actor, i.e., thing, involved in the process. Internet-of-Things (IoT) enters in the manufacturing sector. The tremendous amount of data produced is intended to be analyzed by applying methods from artificial intelligence, machine learning and data mining.

One of the objectives of such analysis is Zero Defect Manufacturing, i.e., a manufacturing process where acquired data during the entire life cycle of products is used to continuously improve the product design in order to provide customers with unprecedented quality guarantees.

The achievement of such an objective also requires to rethink the entire manufacturing process and the entire supply chain, in order to consider every single phase of the product life cycle. The Electrospindle 4.0 pilot aims at applying Zero Defect Manufacturing principles to the production of spindles.

A spindle is a rotating motor device employed in manufacturing for machining different types of material, by applying different possible tools. Such machine undergoes severe functioning conditions in terms of vibrations and shocks, thus reliability is a fundamental feature for customers, which do not want to sustain inactivity periods due to maintenance operations.

HSD is an Italian company and one of the international leading companies in the production of spindles for manufacturing customers. The Electrospindle pilot has objectives that are coherent with some of the lines of action proposed by the Italian Cluster “Fabbrica Intelligente” (CFI), which clearly identifies HSD as a Light House Plant for Zero Defects Manufacturing, i.e., a national reference point for what concerns technological innovation. The identified lines of action are: (i) production systems for a personalized production, (ii) efficient production systems and (iii) adaptive and evolutionary production systems; that are also coherent with the European strategies for the development of a resilient country in the manufacturing sector.

The new family of spindles will be equipped with special sensors that will enable them to independently transmit data related to their status and working conditions while performing their functions on the machine tool on which they are installed. Such devices will be the result of the application of several techniques, including cloud computing, machine learning (ML), artificial intelligence (AI), digital twins (DT) and Design for X. The new manufacturing process will promote a continuous improvement of design and manufacturing processes, with the goal of creating more reliable and efficient products.

3. Involved partners	
Partner	Country
Cyprus University of Technology (CUT)	Cyprus
Sapienza Università di Roma	Italy

5. Piloting #4 – Paradisiotis group (poultry meat production)

1. Project Identification	
Name	Smart Data Processing and Innovative Tools for Paradisiotis Group
Acronym	SMART-PARADIS
Funding opportunities	EU Horizon Europe, National Cypriot (Research & Innovation Foundation (RIF))

2. Summary

This idea is an ongoing, working attempt to attract funding for applied research in Industry 4.0. Having a strong local company as collaborator, namely Paradisiotis group, a factory that breeds chicken and then slaughters and packages poultry meat. A proposal has already been prepared and submitted under the call CO-DEVELOP of the Research & Innovation Foundation of Cyprus. The summary is as follows:

SMART-PARADIS aims at establishing a collaboration between the Cyprus University of Technology as the main research partner, and Paradisiotis Group, a poultry meat production enterprise, as the corporate partner and end-beneficiary of the project, supported by Sapienza University of Rome acting as external research partner, for transferring new and innovative knowledge and expertise between the partners, and allowing challenging problems within the specific business domain to be tackled. Collaboration activities include trainings, workshops, brainstorming, staff mobility & exchange, and hands-on experience development, as well as sharing of technological resources and infrastructure. Latest innovative technologies in the areas of smart data processing, process mining and blockchain will enable addressing specific business priority areas in the relevant working environment, and meeting specific needs and challenges, such as reducing chicks' mortality rate, improving waste management, optimizing delivery routes, lowering fuel consumption and maintenance costs, and linking energy consumption with breeding conditions. This project will contribute to bridging the gap between academia and industry in Cyprus, and at the same time produce innovative results that will benefit the participating enterprise, as well as other similar stakeholders. The targeted priority areas are aligned with main pillars of the Smart Specialization Strategy of Cyprus, such as food production, health & quality of living, and energy preservation. The outputs will allow the enterprise to increase its competitive stand and pioneer in technology related to its line of business, which in turn will result in better products with positive effects on public health, increased productivity and lower costs, and energy preservation. The research partners will also be benefitted as they will be able to excel their research portfolios and

perform applied research offering solutions to real-world problems and connect better with the industry and market.

Beyond this proposal, CUT and Sapienza will continue along the same line to apply BPM concepts in the Paradisiotis environment and enhance these with integration with IoT and digital twins. The next plan is to collect information from various sensors installed in modern machinery that Paradisiotis uses (e.g. a system for feeding, watering and controlling climate conditions for chick farms) and process this information for optimizing production and minimizing losses (i.e. reducing chick deaths in the farms).

3. Involved partners	
Partner	Country
Cyprus University of Technology (CUT)	Cyprus
Sapienza Università di Roma	Italy

6. Piloting #5 – Smart Data Processing for Mastitis

1. Project Identification	
Name	Smart Data Processing for Mastitis
Acronym	SMARTMAST
Funding opportunities	EU Horizon Europe, National Cypriot (Research & Innovation Foundation (RIF))

2. Summary
Mastitis is a common disease found in cows which affects the quality of the dairy food produced, and probably may have a negative impact on the human health. As a result of the complexity of the disease, researchers are struggling to create accurate diagnostic tools, so as to identify the health impact on the humans, and identify the source, in order to prevent the disease. In this context, Fuzzy Cognitive Maps will be used in order to find hidden correlations in the data, aiming to ultimately create a monitoring system to be able identify the disease faster, analyze and identify the health impact on the humans from dairy products produced from cows with the disease, and also predict the types of contamination that has been expressed.

3. Involved partners	
Partner	Country
Cyprus University of Technology (CUT)	Cyprus
Jheronymous Academy of Data Science	JADS

7. Piloting #6 – MONITORING OF LOOTING AND TRAFFICKING OF CULTURAL HERITAGE ITEMS

1. Project Identification	
Name	Smart and operational platform and tools enabling collective and actionable Insights to jointly fight illegal trafficking of cultural goods
Acronym	SOTERIA
Funding opportunities	HORIZON EUROPE

2. Summary
SOTERIA's objective is threefold. Firstly, SOTERIA will systematically conduct pan-EU, evidence-driven, multi-disciplinary research to deliver deep insights in illegal cultural heritage trafficking, combining social-sciences, humanities, criminology, AI and engineering. Secondly, SOTERIA will consolidate on H2020 and related R&D projects, and integrate tools and techniques for enabling security experts to conduct network analysis and improve insights in market mechanism. Special emphasis will be placed on prevention of criminal networks, and to more effectively use techniques like counter insurgency to leverage prevention and more effectively deal with the criminal network threats. Thirdly, increasing awareness and training. At the same time, SOTERIA will respect fundamental rights and align with European societal values. SOTERIA will achieve its ambitious gaols by fostering improved strategic and long-lasting collaboration amongst security agencies within EU member states through a pan-European and institutionalized community of security agencies and experts involved in crime-fighting ITCG that are enacted through the SOTERIA Platform. SOTERIA will deliver and promote the uptake of novel and actionable methods, tools and theories for lawful acquisition of data sources, and collectively, pro-actively and analytically perform investigations in such a way that the trail of evidence can be used in court. The SOTERIA platform will initially find its home in the ENLETS, European network of Law Enforcement Agencies.

3. Involved partners	
Partner	Country
Cyprus University of Technology (CUT)	Cyprus

Jheronymous Academy of Data Science	JADS

8. Concluding remarks and timeline

This deliverable is a part of work package 4 and presents envisioned piloting activities from the consortium of DESTINI. Most of these activities are already in place, thanks to previous collaborations among partners of the DESTINI project and relevant stakeholders.

A possible timeline for the pilots can be:

- smart ports, to be started on March 2023 and lasting a couple of years
- process mining in the justice domain is already started and will last till end 2023
- Electrospindle 4.0 is already started and will last till middle 2024
- Smart Paradis is expected to be submitted in the first semester of 2023
- SMARTMAST to be started on started on 2023
- Re-submission of a proposal for monitoring looting and trafficking of cultural artifacts in 2023 Horizon calls