Technical Report

ISO/IEC TR 24030

Second edition
2024-04

Information technology — Artificial intelligence (AI) — Use cases

Technologies de l'information — Intelligence artificielle (IA) — Cas pratiques

Reference number
ISO/IEC TR 24030:2024(en)

Contents

Page

Foreword vi
Introduction vii
1 Scope 1
2 Normative references 1
3 Terms and definitions 1
4 Abbreviated terms 1
5 Applications 3
5.1 General 3
5.2 Application domains 3
5.3 Deployment models 4
5.4 Examples of AI applications 4
6 Use cases 8
6.1 General 8
6.2 Sources of use cases 8
6.3 Guidance for submitting use cases 8
6.4 Fields of the use case template 8
6.4.1 General 8
6.4.2 ID 9
6.4.3 Use case name 9
6.4.4 Application domain 9
6.4.5 Deployment model 9
6.4.6 Objective(s) 9
6.4.7 Narrative 9
6.4.8 Stakeholders and stakeholder perspectives 9
6.4.9 Data characteristics 10
6.4.10 Key performance indicators (KPIs) 10
6.4.11 Features of use case 10
6.4.12 Threats and vulnerabilities 11
6.4.13 Challenges and issues 11
6.4.14 Trustworthiness considerations 11
6.4.15 Use of standards; opportunities for future standardization 13
6.4.16 SDGs to be achieved 13
6.5 Basic statistics 13
6.5.1 Use cases by application domain 13
6.5.2 Use cases by task 14
6.5.3 Use cases by relevant SDGs 14
6.6 Societal concerns 15
7 Use cases summaries 15
7.1 General 15
7.2 Agriculture 21
7.2.1 Real-time segmentation and prediction of plant growth dynamics using low-power embedded systems equipped with AI (use case 126) 21
7.2.2 Smart agriculture (use case 156) 23
7.2.3 Forecasting of crop yield using decision support system (use case 173) 24
7.3 Digital marketing 26
7.3.1 Improving conversion rates and return on investment (RoI) with AI technologies (use case 53) 26
7.3.2 AI system for digital marketing in retail services (use case 185) 27
7.4 e-commerce/e-business 29
7.4.1 Emotion-sensitive AI customer service (use case 42) 29
7.4.2 Deep learning-based user intent recognition (use case 43) 31
7.4.3 AI virtual assistant for customer support and service (use case 106) 33

7.4.4 Customer relation management (CRM) (use case 157) 34
7.5 Education 36
7.5.1 A recommendation system for industrial training (use case 23) 36
7.5.2 An intelligent marking system (use case 83) 38
7.5.3 Intelligent educational robot (use case 84) 40
7.5.4 AI system to intelligent campus (use case 85) 41
7.5.5 AI adaptive learning platform for personalized learning (use case 102) 43
7.5.6 AI adaptive learning mobile app (use case 124) 45
7.6 Energy 46
7.6.1 Smart energy grid (use case 166) 46
7.7 Fintech 48
7.7.1 Detection of fraud based on collusion (use case 20) 48
7.7.2 Virtual bank assistant (use case 57) 49
7.7.3 Forecasting prices of commodities (use case 91) 51
7.7.4 Financial advice and asset management with AI (use case 114) 52
7.7.5 Loan in 7 minutes (use case 119) 54
7.7.6 Predictive risk intelligence (use case 164) 55
7.8 Healthcare 57
7.8.1 AI system to predict post-operative visual acuity for LASIK surgeries (use case 24) 57
7.8.2 AI system to quality control of electronic medical record (EMR) in real time (use case 50) 59
7.8.3 Discharge summary classifier (use case 79) 60
7.8.4 Generation of clinical pathways (use case 80) 62
7.8.5 Hospital management tools (use case 81) 63
7.8.6 Predicting relapse of a dialysis patient during treatment (use case 87) 64
7.8.7 Instant triaging of wounds (use case 89) 66
7.8.8 Detection of fraudulent medical claims (use case 90) 68
7.8.9 AI platform for chest CT-scan analysis (early stage lung cancer detection) (use case 105) 69
7.8.10 Neural network formation of 3D-models orthopaedic insoles (use case 121) 71
7.8.11 Search of undiagnosed patients (use case 127) 72
7.8.12 A clinical decision support system (use case 131) 73
7.8.13 Symptom assessment (hypothetical) (use case 134) 76
7.8.14 Making using evidence-based medicine and AI (use case 167) 77
7.8.15 AI-service for blood cells and bone marrow scans analysis (use case 168) 79
7.8.16 AI-service for chest X-ray and chest CT (use case 169) 81
7.8.17 Intelligent video analytics system (use case 170) 83
7.8.18 Retrospective analysis (use case 172) 84
7.8.19 Robotization of the federal hotlines on COVID-19 issues (use case 174) 85
7.8.20 Use of computer vision innovative technologies for analysis of medical images and further application (use case 175) 87
7.9 Home and service robotics 89
7.9.1 Device control using AI consisting of cloud computing and embedded system (use case 132) 89
7.10 ICT 91
7.10.1 AI system to help mobile phones to have better picture effect (use case 32) 91
7.10.2 Product failure prediction for critical IT infrastructure (use case 86) 93
7.10.3 AI-based optimized field dispatch (use case 151) 94
7.10.4 Wireless network failure prediction (use case 152) 96
7.10.5 AI performance evaluation of AI-powered messaging bots (use case 176) 97
7.11 Insurance 100
7.11.1 AI services for health insurance companies (use case 161) 100
7.12 Knowledge management 102
7.12.1 Water crystal mapping (use case 77) 102
7.12.2 AI system with a digital knowledge centre for utilizing the knowledge in the organization (use case 187) 104
7.13 Legal 106
7.13.1 AI contract management (use case 120) 106

7.14 Manufacturing ... 108
7.14.1 Quality assurance solution based on AI, to detect defects on wind turbines blades (use case 4) ... 108
7.14.2 Generative design of mechanical parts (use case 15) ... 110
7.14.3 Powering remote drilling command centre (use case 36) ... 111
7.14.4 Quality improvement of adhesive products, based on AI (use case 37) ... 113
7.14.5 Empowering autonomous flow meter control- reducing time taken to “proving of meters” (use case 40) ... 114
7.14.6 Improvement of productivity of semiconductor manufacturing (use case 82) ... 116
7.14.7 AI decryption of magnetograms (use case 104) ... 119
7.14.8 Analysing and predicting acid treatment effectiveness of bottom hole zone (use case 110) ... 120
7.14.9 Automatic classification tool for full size core (use case 112) ... 122
7.14.10 Collaborative AI to assist workers with production and assembly in factories (use case 179) ... 124
7.15 Media and entertainment ... 126
7.15.1 Video on demand publishing intelligence platform (use case 58) ... 126
7.15.2 AI system for promoting DX in customer attraction services at a museum (use case 178) ... 128
7.16 Public sector ... 129
7.16.1 AI ideally matches children to daycare centres (use case 7) ... 129
7.16.2 Open spatial data set for developing AI algorithms based on remote sensing (satellite, drone, aerial imagery) data (use case 122) ... 131
7.16.3 Smart city (use case 165) ... 133
7.16.4 AI tool for species categorization for wildlife population monitoring (use case 177) ... 135
7.17 Security ... 137
7.17.1 Non-intrusive detection of malware (use case 93) ... 137
7.17.2 Detect pickpockets in a crowd - training with privacy-sensitive data (use case 181) ... 138
7.18 Transportation ... 140
7.18.1 Enhancing traffic management efficiency and infraction detection accuracy with AI technologies (use case 29) ... 140
7.18.2 AI system for traffic signal optimization based on multi-source data fusion (use case 49) ... 142
7.18.3 Dynamic routing software as a service (SaaS) based on artificial intelligence (use case 92) ... 144
7.18.4 Misbehaviour detection (MBD) for V2X (use case 182) ... 145
7.18.5 AI system to estimate or predict congestion length for traffic signal control (use case 183) ... 147
7.18.6 Traffic signal control using artificial intelligence (use case 184) ... 148
7.18.7 AI system for predicting rivers’ water levels during flooding (use case 186) ... 149
7.19 Work and life ... 152
7.19.1 Recommendation algorithm for improving member experience and discoverability of resorts in the booking portal of a hotel chain (use case 28) ... 152
7.19.2 Improving the quality of online interaction (use case 88) ... 153
7.19.3 Business use of IoT for surveillance (use case 162) ... 155
7.19.4 Video surveillance (use case 180) ... 157

Annex A (informative) **Use case template** ... **159**

Annex B (informative) **Use cases list of ISO/IEC TR 24030:2021 and ISO/IEC TR 24030:—** ... **162**

Bibliography ... **168**

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 42, *Artificial intelligence*.

This second edition cancels and replaces the first edition (ISO/IEC TR 24030:2021), which has been technically revised.

The main changes are as follows:

— selection of 51 “in operation” use cases from Annex A (informative), Collected use cases of ISO/IEC TR 24030:2021;

— collection and selection of 30 additional use cases;

— enhanced the use case submission form and the structure of use case description in Clause 7 to describe the desirable information of use cases;

— updated the statistics in 6.5 to reflect the use cases in this document;

— removed the subclauses that are no longer suitable for the use cases in this document (e.g. 6.6.3, Annex A and Annex C in the first edition);

— removed most of the terms from Clause 3 to leave two definitions in this document.

Any feedback or questions on this document should be directed to the user’s national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

This document provides a collection of artificial intelligence (AI) use cases in a variety of domains.

In total, 187 AI use cases were submitted by experts between July 2018 and the end of June 2022. In this document, the term “use cases” means “use cases selected from those submitted”. This document selected 81 in-operation use cases from all submissions.

The rationale for this document is as follows:

— illustrating the applicability of the AI standardization work across a variety of application domains;

— input to and reference for AI standardization work;

— sharing the collected use cases in support of AI standardization work with external organizations and internal entities to foster collaboration;

— reach out to new stakeholders interested in AI applicability;

— liaising with organizations to collect requirements for AI through use cases;

— by investigating use cases, it is possible to find new technical requirements (standardized demands) in the market, which can accelerate the pace of transformation of scientific and technological achievements.

While a bottom-up approach was used to collect use cases, a top-down approach is used in this document to identify AI applications, their deployment models and their application domains, as shown in 5.2

The first step taken to collect use cases was to identify application domains of AI systems (described in Clause 5) and to provide a use case template (described in 6.4 and Annex A). Contributors were requested to submit use cases using the provided template.

To improve the quality of use cases, guidance has been provided to contributors. This guidance includes acceptable sources (described in 6.3) and the characteristics of the AI systems (described in 6.4) that are used to develop use cases.

In this document, 6.5 includes basic statistics of use cases. Subclause 6.6 introduces societal concerns that affect many use cases.

The use cases were grouped and categorized according to the identified application domains. In this document, use cases are grouped, categorized and summarized according to the identified application domains in Clause 7. Use cases of specific application domains and their original submissions can be found at https://standards.iso.org/iso-iec/tr/24030/ed-2/en.

The perspectives of security and privacy in the AI use cases can be found in ISO/IEC TR 27563[6]. ISO/IEC TR 27563[6] includes a security and privacy analysis of the use cases in ISO/IEC TR 24030:2021. It is mentioned that the analysis was carried out independently from the use cases in ISO/IEC TR 24030:2021 contributors and therefore that it does not necessarily reflect their views.

AI is an emerging field with use cases and solutions with a wide range of maturity and success. The descriptions are given for the convenience of users of this document and does not constitute an endorsement by ISO.

Information technology — Artificial intelligence (AI) — Use cases

1 Scope

This document provides a collection of representative use cases of AI applications in a variety of domains.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
artificial intelligence
AI
<discipline> research and development of mechanisms and applications of *AI systems* (3.2)

Note 1 to entry: Research and development can take place across any number of fields such as computer science, data science, natural sciences, humanities, mathematics and natural sciences.

[SOURCE: ISO/IEC 22989:2022, 3.1.3]

3.2
artificial intelligence system
AI system
engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives

Note 1 to entry: The engineered system can use various techniques and approaches related to artificial intelligence to develop a model to represent data, knowledge, processes, etc. which can be used to conduct tasks.

Note 2 to entry: AI systems are designed to operate with varying levels of automation.

[SOURCE: ISO/IEC 22989:2022, 3.1.4]

4 Abbreviated terms

For the purposes of this document, the following abbreviated terms apply. The abbreviated terms are extracted from use cases.

AUC	area under the curve
BERT	bidirectional encoder representations from transformers
CNN	convolutional neural network
COBIT	control objective for information and related technology
CRISP-DM	cross-industry standard process for data mining
CRM	customer relations management
CSV	comma separated values
CT	computed tomography
CV	computer vision
DICOM	digital imaging and communications in medicine
DL	deep learning
EHR	electronic health record
GDPR	general data protection regulation
GPU	graphics processing unit
ICT	information and communication technology
ISP	internet service provider
ITIL	information technology infrastructure library
KPIs	key performance indicators
LSTM	long-short-term memory network
ML	machine learning
NLP	natural language processing
NLU	natural language understanding
PACS	picture archiving and communication system
RMSE	root mean square error
RNN	recurrent neural network
ROC	receiver operating characteristic
SaaS	software as a service
SIS	smart information systems
SVM	support vector machine
UT	ultrasonic testing
XGBoost	extreme gradient boosting

5 Applications

5.1 General

This clause identifies AI applications from the perspectives of their application domains and deployment models.

5.2 Application domains

Eighteen application domains were considered as target domains for the use cases. The classifications of the application domains are based on the categories in References [16] and [17].

— agriculture: this domain refers to the science or practice of farming, including cultivation of the soil for the growing of crops and the rearing of animals to provide food or other products (see ISO 20670:2018, 3.2[7]);

— digital marketing: this domain refers to the applications of marketing that uses the Internet and online based digital technologies such as desktop computers, mobile phones and other digital media and platforms to promote products and services;

— e-commerce/e-business (electronic commerce / electronic business): this domain is a category of business transactions, involving two or more Persons, enacted through electronic data interchange, based on a monetary and for-profit basis. Persons can be individuals, organizations or public administrations. The underlying principles and characteristics of e-commerce and e-business include: 1) being business transaction based (of both a financial and non-financial nature); 2) using information technology (computers and telecommunications); 3) interchanging electronic data involving establishment of commitments among persons;

— education: this domain refers to the applications that can provide processes by which an individual or group of people conveys, transfers or obtains knowledge about a subject or concept (see ISO 30422:2022, 3.9[8]);

— energy: this domain refers to the industry that is the totality of all of the industries involved in the production and sale of energy, including fuel extraction, manufacturing, refining and distribution;

— fintech: this domain refers to the companies whose line of business combines software and technology to deliver financial services. the emergence of fintech companies can reshape and improve finance by cutting costs and expanding access to financial services[20];

— healthcare: this domain refers to the applications that provide activities, services, or supplies related to the health of an individual (see ISO/TR 14639-2:2014, 2.31[9]);

— home and service robotics: this domain refers to the science and practice of designing, manufacturing and applying robots that performs useful tasks for humans or equipment excluding industrial automation applications (see ISO 8373:2021, 3.7, 3.10[10]);

— ICT (Information and Communications Technology): this domain refers to group of applications using information and communications (telecommunications) technologies for gathering, storing, retrieving, processing, analysing and transmitting information (see ISO/IEC TR 24704:2004, 3.1.5[11] and ISO/IEC 29138-1:2018, 3.3[12]);

— insurance: this domain provides an effective mechanism for protection and risk management and limits or relieves the financial burden on the insured by mitigating the effects of unpredictable events such as illness, accident, death and natural disasters. Insurance companies pool different types of risk and use statistical analysis to project losses within a given class;

— knowledge management: this domain refers to the applications that provide combination of processes, actions, methodologies, and solutions that enable the creation, maintenance, distribution and access to knowledge (see ISO/IEC 30145-2:2020, 3.7[13]);

— legal: this domain refers to the applications that are used in the legal services industry provide expert advice in all aspects of the law, including contract, corporate, criminal, family and estate, tax and tort law[21];

— manufacturing: this domain refers to the industry that is industries transforming goods, that is, mainly manufacturing industries in their own right, but they also concern the repair and installation of industrial equipment and subcontracting operations for third parties[22];

— media and entertainment: this domain comprises businesses that produce, distribute and offer ancillary digital services and products for motion pictures, television programs and commercials along with streaming content, music, video and audio recordings, broadcast, radio, text and book publishing, eSports and video games sectors[23];

— public sector: this domain refers to businesses and industries that are owned or controlled by the government;

— security: this domain refers to the industry that is made up of companies that manufacture and sell security products. The industry also includes licensed security agents, as well as associations that regulate security agencies, services and products[24];

— transportation: this domain encompasses the movement of humans, animals and goods from one place to another. "Transportation" can be subdivided into "transportation infrastructure", "transportation vehicles" and "transportation operations";

— work and life: this domain refers to the industries in which digital technologies have had profound impacts, good and bad, and other sectors in which automation will likely experience major changes in the near future. Many of these changes have been driven strongly by "routine" digital technologies, including enterprise resource planning, networking, information processing and search[17].

5.3 Deployment models

This document considers the use of AI applications and lists the following possible deployment models of AI applications.

— cloud services;

— cyber-physical systems;

— embedded systems;

— hybrid (embedded systems and cloud services, or on-premise systems and cloud services);

— on-premise systems;

— social networks.

5.4 Examples of AI applications

Examples of AI applications are listed in Table 1. These application examples were derived from the "Artificial Intelligence White Paper"[16]. Each example in Table 1 has an application domain, deployment mode and short description.

The applications in Table 1 are the result of a top-down approach and can be considered to be indicative for collecting use cases. Not all the applications are necessarily addressed by the collected use cases.

Table 1 — Examples of AI applications

Application domain	Application	Deployment model	Short description
Agriculture	Agricultural automation	Cloud services	Monitor and manage field conditions.
		On-premise systems	Accumulate weed or insect patterns and eliminate them.

Table 1 *(continued)*

Application domain	Application	Deployment model	Short description
Agriculture	Craftsmanship skill transfer	Cloud services	Learn about the best practices from craftsmen and provide feedback to others.
Agriculture	Cultivation management	On-premise systems	Monitor the field condition and manage the irrigation condition.
Manufacturing	Construction planning	Cloud services	Learn about the best practices and apply them to future planning.
Manufacturing	Robot construction	Cloud services On-premise systems	Provide autonomous construction robots to construction sites.
Manufacturing	Abnormality or malfunction prediction	Cloud services	Accumulate normal signal patterns to learn normal signals.
		On-premise systems	Find out abnormal signal patterns on the premises.
Security	Cyber security	Cloud services	Monitor cyber transactions against important defence assets and find out attack patterns and prevent their intrusion.
Security	Electronic warfare	Cloud services Embedded systems On-premise systems	Autonomous pilot with cloud support to enable electronic warfare.
Digital marketing	Online campaign performance optimization	Cloud services	Targeted advertising through data analysis.
Education	Adaptive learning	On-premise systems	Provide personalized learning materials via a learning model to achieve efficient learning results.
Education	Scoring	On-premise systems	Provide the most effective feedback to learners via the cognitive learning model to achieve the most effective learning results.
Fintech	Asset management	Cloud services	Accumulate and learn about the best practices and apply them to realize customer satisfaction.
Fintech	Fraud identification	Cloud services	Identify fraud transactions and warn managers.
Fintech	Loan screening	Cloud services	Learn about customers' backgrounds to find out abnormal loan patterns.
Fintech	Security assurance against cyber attacks	Cloud services	Learn and detect known fraud patterns or discover unknown fraud patterns.
Fintech	Stock exchange and trading	Cloud services	Accumulate the best practices and enable 24/7 trading.
Healthcare	Diagnosis support	Cloud services On-premise systems	Provide diagnostic or treatment information and find out abnormal condition compared with normal condition.
Healthcare	Electronic health record system	Cloud services Embedded systems Hybrid On-premise systems	Accumulate and disseminate learning disease patterns and healthy patterns, and assistants as an integrated medical support system.

Table 1 *(continued)*

Application domain	Application	Deployment model	Short description
Healthcare	New drug development	Cloud services	Curation: Find out the correlation among submitted papers.
			Molecular pattern: Find out the effective coordination of target molecule.
Legal	Early case assessment	Cloud services	AI support of work that preps had been doing.
Legal	Judicial recommendation	Cloud services	Judge support by using previous judicial judgment cases.
Transportation	Logistics in the base	Cloud services	Coordinate the best logistic move in the local procurement base warehouse.
		On-premise systems	
Transportation	Procurement logistics	Cloud services	Analyse the procurement context and propose the best procurement actions.
		On-premise systems	
Transportation	Sales logistics	Cloud services	Analyse and learn about the best practices of sales logistics and provide the most effective routes to move sales.
		On-premise systems	
Transportation	Automatic cruise control	Cloud services	Update cruise control software dynamically. Accumulate road condition data and disseminate them to autonomous agents. Mainly enabled on wheelchairs, ships and autonomous robots.
		Embedded systems	Enable autonomous driving without any help from connected devices. Mainly enabled on wheelchairs, ships and autonomous robots.
Transportation	Autonomous driving	Cloud services	Update cruise control software dynamically. Accumulate road condition data and disseminate them to autonomous agents. This application is mainly implemented on vehicles.
		Embedded systems	Enable autonomous driving without any help from connected devices. This application is mainly implemented on vehicles.
Transportation	City-wide traffic control	Cyber-physical systems	Optimize city-wide traffic flow by inspecting real-time traffic images and controlling traffic signals.
Transportation	Dynamic map for autonomous cruise control	Cloud services	Create, maintain, and disseminate map information with semantic tags with real-time communication with mobile agents such as cars, wheelchairs, robots, and human beings.
		On-premise systems	Accumulate actual road situation data and recognize objects that are not on the shared map.

Table 1 *(continued)*

Application domain	Application	Deployment model	Short description
Transportation	Robot taxi	Cloud services	Pick-up arrangement system controls robot taxis effectively.
		Embedded systems	Autonomously drive on the road with dynamic control of steering, acceleration and braking.
		On-premise systems	Autonomously drive on the road using a road map.
Manufacturing	Development design	Cloud services	Accumulate design patterns to help designers.
		On-premise systems	Check design patterns under real constraints on the premises.
Manufacturing	Product quality inspection	On-premise systems	Inspect products by image recognition.
Manufacturing	Production process	Cloud services	Accumulate production quality actuation patterns and estimate quality performance.
Manufacturing	Production process	On-premise systems	Accumulate production throughput-related parameters and estimate the output throughput.
Public sector	Online service support	Cloud services Social networks	Provide residents with support for online services.
Public sector	Public service matching	Cloud services On-premise systems	Optimize matching between residents and public services.
Digital marketing	Autonomous driving store	On-premise systems	Provide autonomous driving sales robots.
Digital marketing	Register-less store	On-premise systems	Monitor all customer movements to realize cash register-less retail shops.
Security	Cyber security	Cloud services	Monitor transactions in cyber space and find out attacks through finding abnormal transaction patterns.
Security	Personal information management	Cloud services	Monitor operations for GDPR conformance assurance.
Security	Video surveillance and crime risk prediction	Cloud services	Monitor behavioural patterns in urban areas, predict crime risk and find out criminal patterns.
Public sector	Equipment operation	Cloud services	Accumulate operational parameters and learn normal operations.
		On-premise systems	Monitor operations and find out abnormal operation patterns.
Public sector	Improvement of operational efficiency	Cloud services	Learn about the correlation among significant parameters and realize the most efficient operations. Traffic control, electricity supply control, etc.
Public sector	Landslide, flood prediction	Cloud services On-premise systems	Monitor weather and ground conditions in real time and predict disasters such as landslides, floods, etc.
Energy	Power demand forecasting	Cloud services	Learn about demand patterns with other significant parameters and forecast future demand.

Table 1 *(continued)*

Application domain	Application	Deployment model	Short description
Work and life	Smart home appliances	Embedded systems	Equip robot vacuums, refrigerators and air conditioners with sophisticated control.
Work and life	Smart personal agent	Social networks	Smart agents assist individual users.

6 Use cases

6.1 General

This document is based on a collection of use cases available at https://standards.iso.org/iso-iec/tr/24030/ed-2/en. Use cases from the first edition of ISO/IEC TR 24030 were included only when they were updated to meet the selection criteria for the second editions. New submissions to the second edition that did not meet the selection criteria (in-operation and of sufficient quality) are omitted from the electronic attachment. This document contains 81 “in-operation” use cases that were selected from those submitted. Annex B gives the use case titles and the subclause numbers in the first edition and this document. Some use cases refer to trademarks or trade names such as company names, product names or service names. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these products or services.

The terms of the template used for the collection of use cases, which is given in Annex A, are defined in 6.4. Clause A.1 describes a blank template and Clause A.2 describes the comparison of the use case templates between ISO/IEC TR 24030:2021 and this document. Subclause 6.2 provides sources of use cases. Subclause 6.3 provides guidance for the submission of use cases. Subclause 6.5 presents basic statistics on the selected use cases. Societal concerns relating to the use cases are described in 6.6.

6.2 Sources of use cases

To improve the quality of use case descriptions, potential sources are:

— peer-reviewed scientific or technical publications on AI applications;

— patent documents describing AI systems;

— technical reports or conference presentations;

— high quality company whitepapers and presentations;

— publicly accessible sources in sufficient detail.

6.3 Guidance for submitting use cases

Use case submissions were considered for inclusion in this document based on the following factors:

— diversity and representativeness: this document is meant to represent a range of AI use cases in terms of application domain, deployment model, objectives, stakeholders, autonomy, data characteristics, architecture and other factors;

— amount of information: to populate as many fields as possible;

— positive impact: demonstration of positive outcomes for society, the environment and stakeholders.

6.4 Fields of the use case template

6.4.1 General

The terms in the template used for the collection of use cases are defined in 6.4.2 to 6.4.16.

6.4.2 ID

The unique number that identifies the use cases in ISO/IEC TR 24030:2021[4] and in this document.

6.4.3 Use case name

Generally, the name of the use case provided by the use case contributor. However, some use cases were renamed for clarity.

6.4.4 Application domain

Generally, the application domain selected by the use case contributor from the list in 5.2. However, the classification of some use cases was changed for consistency reasons.

6.4.5 Deployment model

Deployment model selected by the use case contributor from the list in 5.3.

6.4.6 Objective(s)

Generally, the objectives of the AI system defined by the use case contributor, including the following:

— what is to be accomplished;

— intended beneficiaries of the AI system;

— scope, boundaries and limitations.

The objectives of some use cases were amended to improve their readability.

6.4.7 Narrative

This field asks the contributor to provide description of optimization or inferences being made with the AI model and what decisions, predictions, recommendations are being applied; capabilities and features that are unique to the domain, decision environment; whose decision is being augmented? the dynamics of the decision environment, etc.

6.4.8 Stakeholders and stakeholder perspectives

ISO/IEC 22989[1] defines a stakeholder as any individual, group, or organization that can affect, be affected by, or perceive itself to be affected by a decision or activity. This can include organizations, customers, third parties, end users, the community, the environment, developers and other entities. The use case can describe AI system stakeholders and how their perspectives are taken into consideration. Stakeholder perspectives include stakeholder assets, values and effects.

AI stakeholder roles includes:

— AI provider: provides products or services that use the AI system;

— AI producer: designs, develops, tests and deploys products or services that use the AI system;

— AI Customer: uses an AI product or service either directly or by its provision to AI users;

— AI partner: provides services in the context of AI;

— AI subject: impacted by an AI system, service, or product.

Details of AI stakeholder roles can be found in ISO/IEC 22989:2022, 5.19[1].

6.4.9 Data characteristics

6.4.9.1 General

This field describes the data characteristics that are defined in 6.4.9.2 to 6.4.9.7.

6.4.9.2 Source

Origin of data processed by the AI system, e.g. customers, instruments, IoT, web, surveys, commercial activity, simulations or other sources.

6.4.9.3 Variety

Types of data processed by the AI system, e.g. structured or unstructured text, images, voices, gene sequences, numbers, composite: time-series, graph-structures. This field can also briefly discuss formats, logical models, timescales and semantics.

6.4.9.4 Velocity

The rate of flow at which the data in the AI system is created, stored, analysed or visualized. That can be in real time.

6.4.9.5 Variability

Changes in data rate, format, structure, semantics or quality.

6.4.9.6 Quality

Completeness and accuracy of the data with respect to semantic content as well as syntax of the data (e.g. presence of missing fields or incorrect values).

6.4.9.7 Protected attributes

An attribute for equal treatment across groups (e.g. gender, race, religion, legally regulated attribute).

6.4.10 Key performance indicators (KPIs)

This field describes the KPIs for evaluating the performance or usefulness of the AI system.

6.4.11 Features of use case

6.4.11.1 General

This field describes the features and AI characteristics of the use case.

6.4.11.2 Task(s)

The main task of the use case. A pull-down list includes recognition, natural language processing, knowledge processing and discovery, inference, planning, prediction, optimization, interactivity, classification, recommendation and others.

NOTE NLP includes the tasks such as natural language understanding (NLU), natural language generation (NLG), part of speech (POS), question answering, machine translations, relationship extraction sentiment analysis and automatic summarization. The details of NLP tasks are described in ISO/IEC 22989:2022, 9.2.2[1] and ISO/IEC 22989:2022, 3.6[1].

6.4.11.3 Level of automation

The level of automation of AI systems used in this use case.

AI systems can be compared based on the degree of automation and whether they are subject to external control. Autonomy is at one end of a spectrum and a fully human controlled system on the other, with degrees of heteronomy in between.

The level of automation includes the following options:

— full automation: the system can modify its operating domain or its goals without external intervention, control or oversight;

— high automation: the system can perform its entire mission without external intervention;

— conditional automation: the system performs parts of its mission without external intervention;

— partial automation: sustained and specific performance by a system, with an external agent being ready to take over when necessary;

— assistance: some sub-functions of the system are fully automated while the system remains under the control of an external agent;

— no automation: the system assists an operator.

See ISO/IEC 22989:2022, 5.12[1] for more details on the levels of automation.

6.4.11.4 Method(s)

AI method or AI methods, model or models, or framework or frameworks used in development.

6.4.11.5 Platform

Platform (includes hardware system) used in development and deployment.

6.4.11.6 Topology

Topology of the deployment network architecture.

6.4.12 Threats and vulnerabilities

This field describes threats and vulnerabilities relevant to the use case, such as unwanted bias, incorrect AI system use, security threats, challenges to accountability and privacy threats (hidden patterns).

6.4.13 Challenges and issues

Descriptions of challenges and issues of the use case.

6.4.14 Trustworthiness considerations

6.4.14.1 General

The trustworthiness of AI system is defined as ability to meet stakeholder expectations in a verifiable way (see ISO/IEC 22989:2022, 3.5.16[1]). This field is used to describe how the use case addresses the characteristics of trustworthiness that include, for instance, reliability, availability, resilience, security, privacy, safety, accountability, transparency, integrity, authenticity, quality and usability.

6.4.14.2 Bias mitigation

ISO/IEC TR 24027[3] defines bias as systematic difference in treatment of certain objects, people or groups in comparison to others. In this part of the trustworthiness field, the use case can describe how biases such as human cognitive bias, confirmation, data bias and statistical bias are detected and mitigated in the AI system. The use case can also discuss how the organization has approached bias goals and challenges.

See ISO/IEC TR 24027[3] for further information.

6.4.14.3 Ethical and societal concerns

In this part of the trustworthiness field, the use case can describe how societal and ethical concerns related to the AI system are understood, identified, controlled and mitigated. Current or future measures to address potential ethical and societal risks can also be described, along with protected attributes.

Societal concerns can be a factor when an organization is choosing or recommending an AI technology. Taking context, scope, nature and risks into consideration can mitigate undesirable societal outcomes. In the absence of such considerations, the technology itself can perform flawlessly from a technical perspective but have undesirable social or ethical impacts.

AI ethics is one important aspect of societal concerns that addresses the ethical issues arising from the use of AI systems. AI ethics are being considered in various countries and organizations in the form of principles, guidelines or regulations that ethical AI can follow.

AI ethics and ethical risks is based on the four ethical principles of trustworthy AI of EU high-level expert group (HLEG)[19]:

— respect for human autonomy;

— prevention of harm;

— fairness;

— explainability.

The four ethical principles encompass additional ethical principles described by other AI principles or guidelines. Ethical risk can be defined as a problem caused by an AI system that does not follow the above ethical principles.

6.4.14.4 Explainability

ISO/IEC 22989:2022, 5.14.6[1] defines explainability as a property of an AI system to express important factors influencing the AI system results in a way that humans can understand. In this part of the trustworthiness field, the use case can describe the degree to which the AI system results are explainable and discuss how the organization has approached explainability goals and challenges.

6.4.14.5 Controllability

ISO/IEC 22989:2022, 5.14.5[1] defines controllability as a property of an AI system that a human or other external agent can intervene in the system's functioning. A key aspect of controllability is the determination of which agent(s) can control which components of the AI system (e.g. the service provider or product vendors, the provider of the constituent AI, the user or an entity with regulatory authority). In this part of the trustworthiness field, the use case can describe the degree to which the AI system is controllable and discuss how the organization has approached controllability goals and challenges.

6.4.14.6 Predictability

ISO/IEC 22989:2022, 5.14.7[1] defines predictability as a property of an AI system that enables reliable assumptions by stakeholders about the output. In this part of the trustworthiness field, the use case can describe the degree to which the AI system results are predictable and discuss how the organization has approached predictability goals and challenges.

6.4.14.7 Transparency

ISO/IEC 22989:2022, 5.14.8[1] defines transparency as a property of a system where appropriate information about the system is communicated to relevant stakeholders in system domain. In this part of the trustworthiness field, the use case can describe the degree to which the AI system results are transparent and discuss how the organization has approached transparency goals and challenges.

6.4.14.8 Verification

ISO/IEC 22989:2022, 5.14.9[1] defines verification as confirmation, through the provision of objective evidence, that specified requirements have been fulfilled. In this part of the trustworthiness field, the use case can discuss how the organization has approached verification goals and challenges.

6.4.14.9 Robustness, reliability and resilience

ISO/IEC 22989[1] defines robustness, reliability and resilience as follows:

— robustness: ability of a system to maintain its level of performance under any circumstances;

— reliability: property of consistent intended behaviour and results;

— resilience: ability of a system to recover operational condition quickly following an incident.

In this part of the trustworthiness field, the use case can describe the degree to which the AI system is robust, reliable and resilient, and discuss how the organization has approached goals and challenges of these characteristics.

6.4.15 Use of standards; opportunities for future standardization

Descriptions of standardization opportunities or requirements associated with use case.

6.4.16 SDGs to be achieved

The Sustainable Development Goals (SDGs)[1], otherwise known as the Global Goals, are a collection of 17 global goals set by the United Nations General Assembly. SDGs are a universal call to action to end poverty, protect the planet and ensure that all people enjoy peace and prosperity.

6.5 Basic statistics

6.5.1 Use cases by application domain

Figure 1 shows the percentage of use cases per application domain. Figure 1 only takes into account the 81 use cases described in Clause 7.

NOTE The classification of some use cases was changed for consistency reasons.

1) https://sdgs.un.org/goals

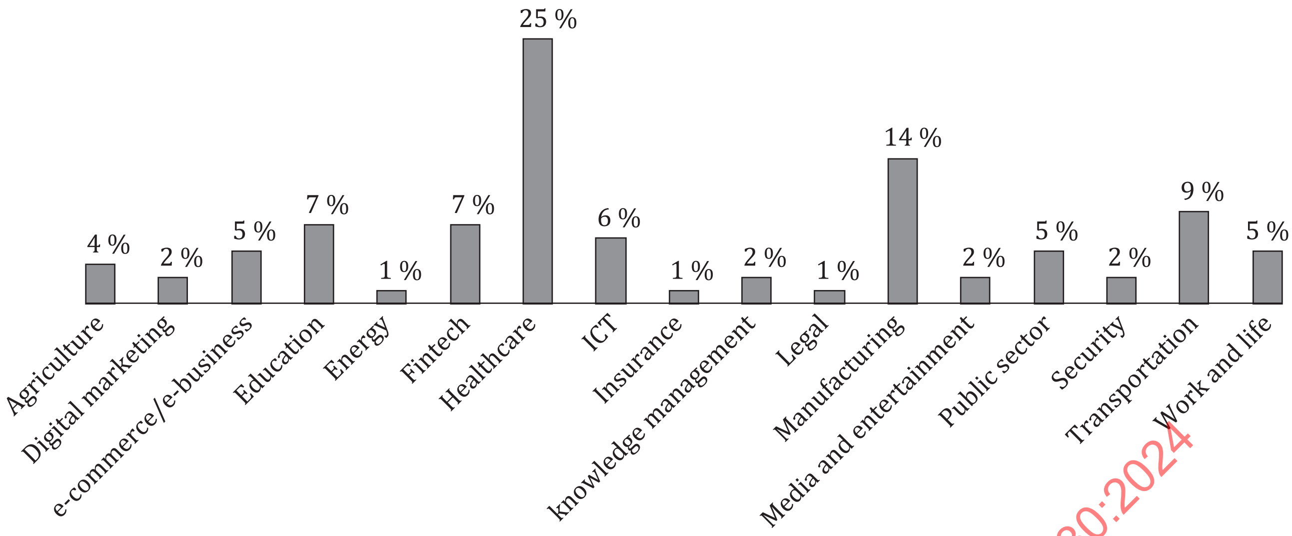

Figure 1 — Distribution of use cases by application domain

6.5.2 Use cases by task

Figure 2 shows the distribution of the 81 use cases according to the tasks listed in 6.4.11.2. Some use cases that include multi-tasks are counted multiple times in Figure 2.

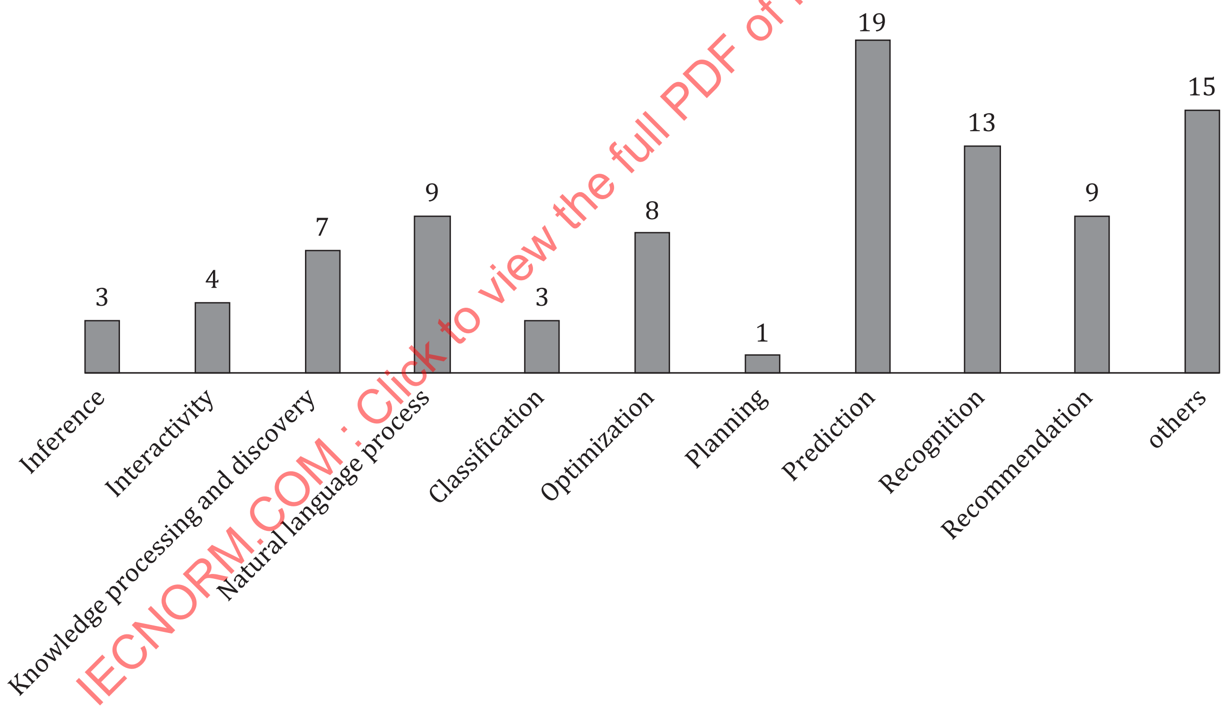

Figure 2 — Distribution of use cases by task

6.5.3 Use cases by relevant SDGs

Figure 3 shows the distribution of the 81 use cases according to relevant SDGs, as identified in Table 2. Some use cases that contribute to several SDGs are counted several times in Figure 3.

NOTE The list of SDGs for certain use cases was modified for consistency reasons.

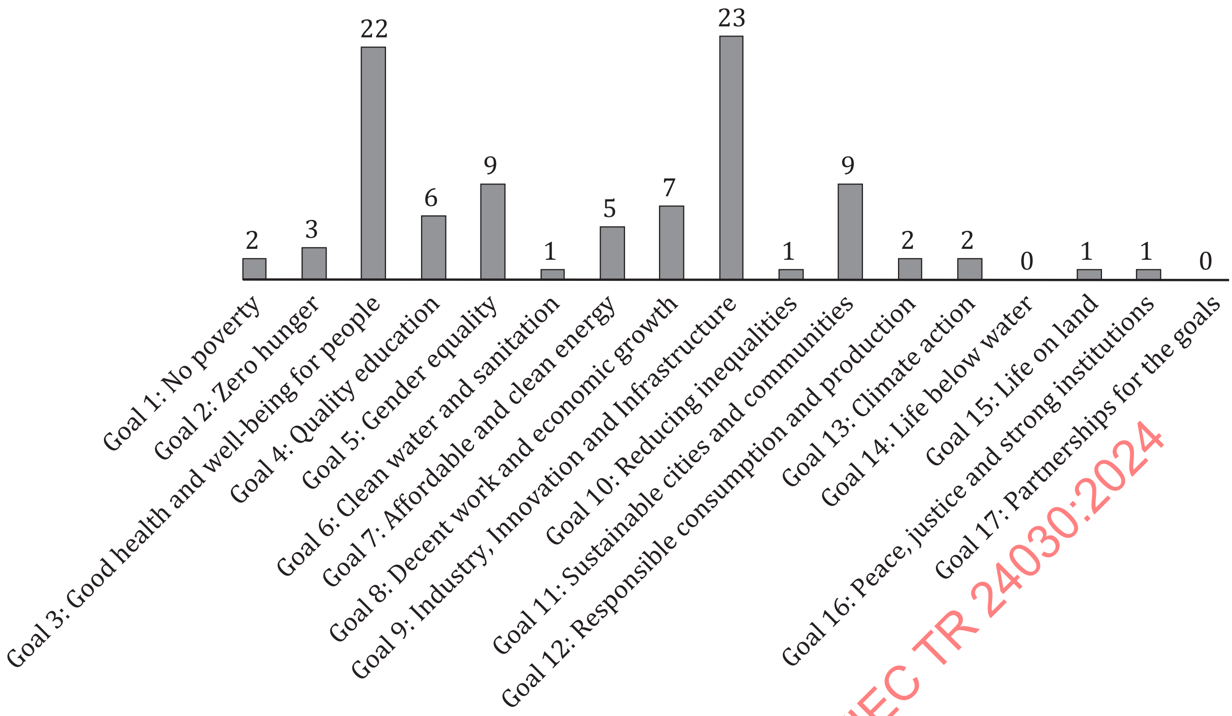

Figure 3 — Distribution of use cases by relevant SDGs

6.6 Societal concerns

ISO/IEC TR 24368[5] identifies a number of ethical and societal concerns. The major topics in this document - accountability, fairness and non-discrimination, transparency and accountability, privacy, security, human control of technology and human-centred design - remain relevant. Since its publication additional trustworthiness characteristics have been identified: most notably the associated impact on environmental sustainability and the degree of societal risk resulting from the use of AI systems and applications. Broadly speaking, the use cases in this second edition reflect the additional attention these characteristics are receiving from consumers, producers, legislators and other AI stakeholders.

7 Use cases summaries

7.1 General

Table 2 provides a summary of information for each of the 81 selected use cases, including the selected application domain, the name of the use case, the subclause in which it is described, its deployment model and relevant SDGs. Subclauses 7.2 to 7.19 provide details of the selected use cases, grouped by application domain

AI is an emerging field, with use cases and solutions that have a wide range of maturity and success. Use case descriptions written by practitioners in their respective fields can include field specific wording. The descriptions are given for the convenience of users of this document and do not constitute an endorsement by ISO. Hyperlinks in the use cases were understood to be active at the time of submission. Not all fields were populated by the use case submitters. Therefore, some fields are blank.

Table 2 — List of use cases

Application domain	Subclause	Use case name	Deployment model	Relevant SDGs
Agriculture	7.2.1	Real-time segmentation and prediction of plant growth dynamics using low-power embedded systems equipped with AI	Embedded system	Good health and well-being for people; zero hunger
	7.2.2	Smart agriculture	Embedded system	Zero hunger
	7.2.3	Forecasting of crop yield using decision support system	Cloud services	Zero hunger
Digital marketing	7.3.1	Improving conversion rates and return on investment (RoI) with AI technologies	On-premise systems	Sustainable cities and communities
	7.3.2	AI system for digital marketing in retail services	Hybrid	Responsible consumption and production
e-commerce/ e-business	7.4.1	Emotion-sensitive AI customer service	On-premise systems	Industry, innovation, and infrastructure
	7.4.2	Deep learning-based user intent recognition	Cloud services	Decent work and economic growth
	7.4.3	AI virtual assistant for customer support and service	On-premise systems	Affordable and clean energy
	7.4.4	Customer relation management (CRM)	Hybrid	Responsible consumption and production; gender equality
Education	7.5.1	A recommendation system for industrial training	On-premise systems	Decent work and economic growth
	7.5.2	An intelligent marking system	On-premise systems	Quality education
	7.5.3	Intelligent educational robot	On-premise systems	Quality education
	7.5.4	AI system to intelligent campus	Cloud services	Quality education
	7.5.5	AI adaptive learning platform for personalized learning	Cloud services	Quality education; gender equality
	7.5.6	AI adaptive learning mobile app	Hybrid	Quality education; gender equality
Energy	7.6.1	Smart energy grid	Embedded systems	Affordable and clean energy

Table 2 *(continued)*

Application domain	Subclause	Use case name	Deployment model	Relevant SDGs
Fintech	7.7.1	Detection of frauds based on collusions	On-premise systems	Decent work and economic growth
	7.7.2	Virtual bank assistant	Cloud services	Industry, innovation and infrastructure
	7.7.3	Forecasting prices of commodities	On-premise systems	Reducing inequalities
	7.7.4	Finance advising and asset management with AI	Cloud service	No poverty
	7.7.5	Loan in 7 minutes	On-premise systems	Industry, innovation, and infrastructure; No poverty; Decent work and economic growth; Gender equality
	7.7.6	Predictive risk intelligence	Hybrid	Industry, innovation, and infrastructure; Gender equality

Table 2 *(continued)*

Application domain	Subclause	Use case name	Deployment model	Relevant SDGs
Healthcare	7.8.1	AI system to predict post-operative visual acuity for LASIK surgeries	Cloud services	Good health and well-being for people; Gender equality
	7.8.2	AI system for quality control of electronic medical records (EMR) in real time	Cloud services	Good health and well-being for people
	7.8.3	Discharge summary classifier	On-premise systems	Good health and well-being for people
	7.8.4	Generation of clinical pathways	On-premise systems	Good health and well-being for people
	7.8.5	Hospital management tools	On-premise systems	Good health and well-being for people
	7.8.6	Predicting relapse of a dialysis patient during treatment	Cloud services	Good health and well-being for people
	7.8.7	Instant triaging of wounds	Cloud services	Good health and well-being for people
	7.8.8	Detection of fraudulent medical claims	Embedded systems	Sustainable cities and communities
	7.8.9	AI platform for chest CT-scan analysis (early stage lung cancer detection)	Cloud services	Good health and well-being for people
	7.8.10	Neural network formation of 3D-model orthopaedic insoles	Hybrid	Good health and well-being for people
	7.8.11	Search for undiagnosed patients	Social networks	Good health and well-being for people
	7.8.12	A clinical decision support system	Cloud services	Good health and well-being for people; Gender equality
	7.8.13	Symptom assessment (hypothetical)	Mobile applications accessing AI backends running on a cloud platform. Stand alone offline mobile applications.	Good health and well-being for people; Gender equality
	7.8.14	Making using evidence-based medicine and AI	Cloud services	Good health and well-being for people
	7.8.15	AI-service for blood cells and bone marrow scans analysis	Cloud services	Good health and well-being for people

Table 2 *(continued)*

Application domain	Subclause	Use case name	Deployment model	Relevant SDGs
Healthcare	7.8.16	AI-service for chest X-ray and chest CT	On-premise systems	Good health and well-being for people
	7.8.17	Intelligent video analytics system	On-premise systems	Good health and well-being for people
	7.8.18	Retrospective analysis	Cloud services	Good health and well-being for people
	7.8.19	Robotization of the federal hotlines on COVID-19 issues	Cloud services	Good health and well-being for people
	7.8.20	Use of computer vision innovative technologies for analysis of medical images and further application	Hybrid	Good health and well-being for people
Home and service robotics	7.9.1	Device control using AI consisting of cloud computing and embedded system	Hybrid	Affordable and clean energy
ICT	7.10.1	AI system to help mobile phones to have better picture effect	On-premise systems	Industry, innovation, and infrastructure
	7.10.2	Product failure prediction for critical IT infrastructure	On-premise systems	Industry, innovation, and infrastructure
	7.10.3	AI-based optimized field dispatch	On-premise systems	Industry, innovation and infrastructure
	7.10.4	Wireless network failure prediction	Cloud services	Industry, innovation and infrastructure
	7.10.5	AI performance evaluation of AI-powered messaging bots	cloud services	Decent work and economic growth
Insurance	7.11.1	AI services for health insurance companies	Embedded systems	Good health and well-being for people
knowledge management	7.12.1	Water crystal mapping	Cloud services	Clean water and sanitation
	7.12.2	AI system with a digital knowledge centre for utilizing the knowledge in the organization	Cloud services	Industry, innovation, and infrastructure
Legal	7.13.1	AI contract management	On-premise systems	Industry, innovation, and infrastructure

Table 2 *(continued)*

Application domain	Subclause	Use case name	Deployment model	Relevant SDGs
Manufacturing	7.14.1	Quality assurance solution based on AI, to detect defects on wind turbines blades	On-premise systems	Affordable and clean energy
	7.14.2	Generative design of mechanical parts	Cloud services	Industry, innovation, and infrastructure
	7.14.3	Powering remote drilling command centre	On-premise systems	Industry, innovation, and infrastructure
	7.14.4	Quality improvement of adhesive products, based on AI	Cloud services	Industry, innovation, and infrastructure
	7.14.5	Empowering autonomous flow meter control – reducing time taken for “proving of meters”	On-premise systems	Industry, innovation, and infrastructure
	7.14.6	Improvement of productivity of semiconductor manufacturing	Hybrid	Industry, innovation, and infrastructure
	7.14.7	AI decryption of magnetograms	Hybrid	Industry, innovation and infrastructure; Affordable and clean energy
	7.14.8	Analysing and predicting acid treatment effectiveness on bottom hole zone	Hybrid	Industry, innovation, and infrastructure
	7.14.9	Automatic classification tool for full size core	Hybrid	Industry, innovation, and infrastructure
	7.14.10	Collaborative AI to assist workers with production and assembly in factories	On-premise systems	Industry, innovation and infrastructure
Media and entertainment	7.15.1	Video on demand publishing intelligence platform	Hybrid	Industry, innovation and infrastructure
	7.15.2	AI system for promoting DX in customer attraction services at a museum	Hybrid	Quality education
Public sector	7.16.1	AI ideally matches children to day-care centres	On-premise systems	Decent work and economic growth; Gender equality
	7.16.2	Open spatial data set for developing AI algorithms based on remote sensing (satellite, drone, aerial imagery) data	Hybrid	Sustainable cities and communities
	7.16.3	Smart city	On-premise systems	Sustainable cities and communities
	7.16.4	AI tool for species categorization for wildlife population monitoring	On-premise systems	Life on land
Security	7.17.1	Non-intrusive detection of malware	Others	Sustainable cities and communities
	7.17.2	Detect pickpockets in a crowd - training with privacy-sensitive data	Social networks	Good health and well-being; Peace, justice and strong institutions

Table 2 *(continued)*

Application domain	Subclause	Use case name	Deployment model	Relevant SDGs
Transportation	7.18.1	Enhancing traffic management efficiency and infraction detection accuracy with AI technologies	Cloud services	Sustainable cities and communities
	7.18.2	AI system for traffic signal optimization based on multi-source data fusion	Hybrid	Sustainable cities and communities
	7.18.3	Dynamic routing software as a service (SaaS) based on artificial intelligence	On-premise systems	Climate action
	7.18.4	Misbehaviour detection (MBD) for V2X	Cloud services	Industry, innovation and infrastructure
	7.18.5	AI system to estimate or predict congestion length for traffic signal control	Cloud services and on-premise system	Sustainable cities and communities
	7.18.6	Traffic signal control using artificial intelligence	Traffic signals, distributed networks and LTE	Industry, innovation and infrastructure
	7.18.7	AI system for predicting rivers' water levels during flooding	On-premise systems or cloud services	Climate action
Work and life	7.19.1	Recommendation algorithm for improving member experience and discoverability of resorts in the booking portal of a hotel chain	Cloud services	Industry, innovation and infrastructure
	7.19.2	Improving the quality of online interaction	Hybrid	Good health and well-being for people
	7.19.3	Business use of IoT for surveillance	Hybrid	Decent work and economic growth
	7.19.4	Video surveillance	Hybrid	Sustainable cities and communities

7.2 Agriculture

7.2.1 Real-time segmentation and prediction of plant growth dynamics using low-power embedded systems equipped with AI (use case 126)

7.2.1.1 Objectives

The project is devoted to the development of a low-power embedded system and AI algorithm for real-time plant segmentation and prediction of its growth. The proposed distributed system is aimed for use in greenhouses and remote areas, where edge-computing autonomous systems are in demand. A branch of this project also aims to develop the payload for drones for the segmentation of harmful plants in real-time.

Prediction of harvest, biomass or leaf area dynamics, leaf index, parameters describing the quality of produced food, consumption of resources from sequences of images of plant growth (including multispectral), data from sensors that describe environmental conditions and artificial growing system parameters representing the state of the growing system.

7.2.1.2 Narrative

Research efforts towards low-power sensing devices with fully-functional AI on board are still fragmented. This use case presents an embedded system enriched with the AI that ensures the continuous analysis and in situ prediction of plant leaf growth dynamics and other important growth parameters. The embedded solutions, grounded on a low-power embedded sensing system with a graphics processing unit (GPU), are able to run the neural networks-based AI on board. Advantages of the proposed system include portability and ease of deployment. This use case uses a sequence of convolutional neural network (CNN) and a recurrent neural network (RNN) called the long-short-term memory network (LSTM) as the core of the AI in our system. The proposed approach guarantees the system autonomous operation for 180 d using a standard Li-ion battery. This use case relies on state-of-the-art mobile graphic chips for smart analysis and control of autonomous devices. This use case used 5 514 images as a source for automated leaf area calculation and follow the training of AI algorithms. Over one thousand records from sensors provide additional information about environmental conditions. All the data were used for training and testing the recurrent neural network, convolutional neural network algorithms, and the segmentation algorithms. This solution provides a root mean square error (RMSE) close to 4 cm^2 in a 3 h prediction horizon. All this allows for high performance in situ optimization of plant growth dynamics and resource consumption.

7.2.1.3 Stakeholders and stakeholder considerations

Stakeholders: agriculture, ecology management, sanitary services.

7.2.1.4 Data characteristics

Data from sensors that describe environmental conditions and artificial growing system parameters representing the state of the growing system.

The AI system used 5 514 images as a source for automated leaf area calculation and follow the training of AI algorithms. Over one thousand records from sensors provide additional information about environmental conditions. All those data were used for training and testing the recurrent neural network, convolutional neural network algorithms, and the segmentation algorithms.

7.2.1.5 KPIs

Accuracy of the predictions of harvest, biomass or leaf area dynamics, leaf index, parameters describing the quality of produced food, consumption of resources.

7.2.1.6 Features of use case

Task(s): prediction

Level of automation: conditional automation

Platform: GPU

7.2.1.7 Threats and vulnerabilities

Hidden patterns, incorrect AI system use.

7.2.1.8 Challenges and issues

The plant growth data significantly depends on multiple factors, including used solutions, illumination characteristics (for greenhouses), weather and seasonal conditions (for outdoors).

The architecture of the neural network is expected to have both high accuracy and high framerate, but a low number of layers and trained parameters for further inference on low-power embedded systems. These controversial factors are expected to be met since embedded systems have limited processing capabilities.

High diversity of data types and no standardization of data obtained by farmers.

7.2.1.9 Trustworthiness considerations

Improving the quality of data by adjusting the data for illumination characteristics (for greenhouses) and weather and seasonal conditions (for outdoors).

7.2.1.10 Use of standards and standardization opportunities

None identified.

7.2.2 Smart agriculture (use case 156)

7.2.2.1 Objectives

This platform is intended to provide personalised information to farmers on crop management.

Food production needs to increase its production levels by 70 % to feed the world's growing population by 2050. In addition, the current ecological footprint is twice the level that it is desired to be; leaving the agricultural sector with the colossal challenge of producing more food, while reducing their ecological impact. Agricultural big data, data analytics, and machine-learning algorithms are the catalysts that are expected to underpin the realization of the world's agricultural goals.

7.2.2.2 Narrative

This platform provides farmers with a range of crop management options within one comprehensive platform. It is intended to complement the company's other programmes, by personalising the exact purchase needs of farmers. The company collaborates with tech companies that can retrieve weather conditions, wind speeds, crop protection, and pest and agronomic data. The company then analyses these data to produce effective agricultural solutions to allow farmers to use their land more effectively.

The project consists of three fundamental components:

— information on local weather predictions, plant disease in situ detection, and recommendation of tools to minimize risk;

— crop and yield previews, farm efficiency and sustainability metrics, and early detection systems for weed, pests and disease;

— support system for farming consultants who are working with farmers to find ideal agricultural planning solutions.

7.2.2.3 Stakeholders and stakeholder considerations

Stakeholders: farmers, suppliers of agricultural products and services, farming consultants, environmental organisations and farmers representative groups.

Stakeholder considerations: overlap with systems threats and vulnerabilities.

7.2.2.4 Data characteristics

The data used in this use case include weather conditions, wind speeds, crop protection, and pest and agronomic data.

7.2.2.5 KPIs

Accuracy of data and recommendations.

7.2.2.6 Features of use case

Task(s): recommendation

Level of automation: high automation

7.2.2.7 Threats and vulnerabilities

Accuracy of data and recommendations: Algorithms can be fallible. There is a wide array of disparities with variables that measure weather, soil and plant disease, training of the image recognition software with accurate and wide variety of images of plant diseases. Natural variations from country-to-country and region-to-region, such as differences in pest management, climatic conditions, and crop types. The system will not replace an agronomist. The platform cannot establish a one-size-fits-all approach for their project and changes their algorithms for different regions. Precise record keeping by farmers is a challenge. Another constraint lies in the actual labour power required to maintain many farm management systems, requiring a large department to work on it. Particularly, when it comes to ensuring the safety and security of the farmer's data.

Privacy and security: The development team are aware of issues surrounding reverse engineering to access their algorithms, so they put great emphasis into securing all of their platforms. If their servers are secure and their software is fully encrypted and secured, then there is little chance that farmers' data will be breached, according to the interviewees.

7.2.2.8 Challenges and issues

Use: creating simple and effective user interfaces.

Data ownership and intellectual property securely protect this data from misuse, hacking, and misappropriation for economic or marketing purposes. It is stated that the company does not use the individual farmer's data to make comparisons between farms, and if this is done in the future, it would not be done without explicit informed consent from the farmers involved. Farmers own their data and can move it to another organization.

7.2.2.9 Trustworthiness considerations

One key factor in the creation of this project was ensuring that it was affordable and easy to use for farmers. The tool is free of charge. It is also for supporting farming consultants, which means that both explanations for farmers and those for farming consultants are to be provided.

Environmental protection: While concerns related to environmental protection were not requested by customers, the project does align with sustainability certification schemes.

7.2.2.10 Use of standards and standardization opportunities

How to deal with conflicting ethical issues and its transparency?

The researchers of this use case have drafted a proposal for guidelines for ethical AI, based on ITIL, COBIT, ISO/IEC 20000 series and CRISP-DM.

7.2.3 Forecasting of crop yield using decision support system (use case 173)

7.2.3.1 Objectives

Increasing the productivity of farms.

Operational adjustment of agricultural works.

7.2.3.2 Narrative

Technology has redefined farming over the years and technological advances have affected the agriculture industry in more ways than one. Agriculture is the mainstay occupation in many countries worldwide and with rising population, which as per UN projections will increase from 7,5 billion to 9,7 billion in 2050, there will be more pressure on land as there will be only an extra 4 % of land, which will come under cultivation by 2050.

This is driving farmers and agro companies to find newer ways to increase production and reduce waste. As a result, AI is steadily emerging as part of the agriculture industry's technological evolution. The challenge is to increase the global food production by 50 % by 2050 to feed an additional two billion people. AI-powered solutions will not only enable farmers to improve efficiencies but they will also improve quantity, quality and ensure faster goto-market for crops.

7.2.3.3 Stakeholders and stakeholder considerations

Agricultural holdings (agronomy, sales department)

7.2.3.4 Data characteristics

The outline of the fields.

Crop rotation history in: field id, sowing year, crop name or indication of land was left uncultivated or was re-cultivated, yield (centner/ha), date of sowing, date of harvest.

Weather data from the fields: date and time of observation, coordinates of the weather station, air temperature, relative humidity, wind speed, atmospheric pressure, precipitation, snow cover height, dew point

7.2.3.5 KPIs

The average error in estimating the yield on the farm is 5,7 %, for a different field it is 15 %.

7.2.3.6 Features of use case

Task(s): decision support system

Level of automation: high automation

Method(s): machine learning (embedding, gradient boosting (XGBoost)

Platform: yield forecast (centner/ha) for each field / farm gets displayed on the web page of the service with the possibility of uploading in tabular formatted CSV files.

Topology decision support system

7.2.3.7 Threats and vulnerabilities

Accuracy of data and recommendations: Algorithms can be fallible. There is a wide array of disparities with variables that measure weather, soil and plant disease, training of the image recognition software with accurate and wide variety of images of plant diseases. Natural variations from country-to-country and region-to-region, such as differences in pest management, limatic conditions and crop types.

7.2.3.8 Challenges and issues

AI solves the scarcity of resources and labour to a large extent and it will be a powerful tool that can help organizations cope with the increasing amount of complexity in modern agriculture.

7.2.3.9 Trustworthiness considerations

The abovementioned accuracy is achieved for the territories where the ML model has been trained, in other regions the accuracy can be slightly lower.

7.2.3.10 Use of standards and standardization opportunities

None identified.

7.3 Digital marketing

7.3.1 Improving conversion rates and return on investment (RoI) with AI technologies (use case 53)

7.3.1.1 Objectives

Utilizing AI technologies in digital marketing.

Help the operation team identify new business scenarios and seize more market opportunities.

Increase conversion rate and marketing effectiveness.

Improve user experience by providing individually customized services.

7.3.1.2 Narrative

Personalized digital marketing has become increasingly important in response to the needs of providing different services to different consumers. The combination of big data and AI algorithms is the core of personalized digital marketing. By modelling user preferences, the AI system can predict the services that users can be interested in, improve marketing effectiveness and enhance user experience.

With economic development, consumers become more emphatic about self-personality. Digital marketing has also begun to focus more on the consumer's personality instead of commonality. Personalized digital marketing has become increasingly important in response to the need to provide different services to different consumers.

The combination of big data and AI algorithms is the core of personalized digital marketing. By modelling user preferences, it can be predicted the services that users can be interested in, improve marketing effectiveness and enhance user experience. There are three main parts of personalized marketing technology:

— audience targeting: forecasting the people who can be interested in the marketing activities, focusing on high-conversion probability populations to increase conversion rates;

— smart subsidy: different marketing subsidies for different users to achieve higher conversion rates at lower cost;

— personalized recommendation: predict user preferences for services or items and recommend to users what they are most likely to be interested in, to increase conversion rates.

Through the application of AI technology, personalized digital marketing has achieved very significant results: the predicted population's conversion rates have shown an improvement of more than 30 %; in the subsidy scenario, personalized digital marketing has achieved a cost reduction of more than 10 % while achieving a 2 % increase in the conversion rate; in the coupon recommendation scenario, the conversion rate has improved by more than 70 %.

7.3.1.3 Stakeholders and stakeholder considerations

Stakeholders: Third-party payment companies, end users (customers), merchants

Stakeholders considerations: User experience, digital marketing RoI, conversion rate, marketing cost.

7.3.1.4 Data characteristics

Description: sample and feature data of marketing campaign.

Source: customers.

Type: log text.

Volume (size): ~500 GB/day.

Velocity: stream and batch.

Variety: device information, location information, conversion information (clicks, transactions), active level.

Variability (rate of change): subject to digital marketing effort (festival, on sale).

Quality: varies (depending on position of data collection and data reflow mechanism).

7.3.1.5 KPIs

Conversion rate: The percentage of users who are responsive to the marketing (e.g. clicks) out of the total number of visitors. The reference of this KPI to mentioned use case objectives is to increase the conversion rate.

RoI: RoI = conversion_rate * (1 - k * cost), k is the cost impact factor and can be adjusted to get a higher conversion rate or lower cost. The reference of this KPI to mentioned use case objectives is to increase marketing effectiveness.

7.3.1.6 Features of use case

Task(s): audience targeting, smart pricing, personalized recommendation

Level of automation: assistance

Method(s): machine learning, deep learning

7.3.1.7 Threats and vulnerabilities

Abuse of personal information, falsified or dirty data.

7.3.1.8 Challenges and issues

How to collect, utilize and protect user information within the scope of what is permitted by relevant national and regional legislation and regulations.

How to let the system evolve and improve continuously by applying new AI models and algorithms.

7.3.1.9 Trustworthiness considerations

For users: enjoy better service at a lower cost

For merchants: Increase profits and decrease costs

For cities and communities: Promote economic prosperity and develop a green economy

7.3.1.10 Use of standards and standardization opportunities

Technical framework of AI-enabled digital marketing system.

Guidelines for collecting, storing and handling of digital marketing data.

Guidelines for applying AI technology to digital marketing.

7.3.2 AI system for digital marketing in retail services (use case 185)

7.3.2.1 Objectives

To improve matching probability and added value of customer satisfaction in retail services by digital marketing introducing AI technologies.

7.3.2.2 Narrative

Through the smartphone application, a unique participant ID is issued to the participants of the campaign and questionnaires are conducted. Using participant's past purchase history data in retail services, probabilistic latent semantic analysis (PLSA) finds customer latent segments with similar purchasing behaviour. To characterize the customer segment Bayesian networks that represents the conditional probability, P (Seg | Q1, ..., Qn), the relationship between the questionnaire data and the customer segments are constructed. At the same time, conditional probabilities P (Product | Seg) are obtained from the result of PLSA. For new campaign participants and participants with no purchase history, an AI system can calculate the purchase prediction probability for particular participant, P (Product | Seg) P (Seg | Q1, ..., Qn) from the results of the questionnaire, and recommends the product with the highest value to the participant. When the products or services are recommended, discount coupons for real retail services can be offered. The fact that the recommendation was accepted can be verified by history of using the coupons. The AI system constantly collects data generated from the smartphone application and utilizes it to improve the matching probability and added value of customer satisfaction in retail services. P (Product | Segment) and P(Segment| Q1,...,Qn) will be re-learned and the recommendation model will be updated. The AI system also forecasts demand and makes it possible to reduce inventory and waste loss. At the same time, looking at the distribution of customer segments, service providers and content providers optimize information content and merchandise.

7.3.2.3 Stakeholders and stakeholder considerations

Participants as both customers of retail services and AI customers.

Service providers as both providers of the retail service and AI customers.

AI research centre of a national institute as both AI producer and AI provider.

7.3.2.4 Data characteristics

Source: interactions with the AI system (“smartphone application”) and past purchase history data in retail services.

Variety: structured.

Velocity: real-time.

Variability: what data to collect and its format can be changed when the questionnaires, products and services are reviewed.

Quality: high enough.

Protected attributes: collected data are anonymous.

7.3.2.5 KPIs

Effectiveness of recommendation: how effective the AI system’s recommendation and coupons can be used in retail services.

Customer satisfaction: customer experience with the AI system’s interaction with visitors.

7.3.2.6 Features of use case

Task(s): interactivity, recommendation and probabilistic inference.

Level of automation: assistance.

Method(s): Bayesian network, probabilistic latent semantic analysis. Human computer interaction, information recommendation using Bayesian network models.

Platform: Servers for running AI programs and storing data. Smartphone application system for interaction with participants.

Topology: Network of servers and smartphone application system.

7.3.2.7 Threats and vulnerabilities

None identified.

7.3.2.8 Challenges and issues

Since the smartphone application and server system are continual improvement, what data to collect and analyse and its format are under periodic review and subject to change.

7.3.2.9 Trustworthiness considerations

There are some concerns on the part of interaction with visitors, such as inappropriate recommendation due to poor user estimation.

7.3.2.10 Use of standards and standardization opportunities

Standard master tables for recommended products

Standard database of questionnaires

Standard database of coupons in retail services

7.4 e-commerce/e-business

7.4.1 Emotion-sensitive AI customer service (use case 42)

7.4.1.1 Objectives

Extracting sentiment and its intensity from customers' input, and responding with an appropriate attitude in order to improve the quality of customers' inquiry experience.

To design an efficient solution for detecting customers' sentiment and its intensity, especially in situations for which there is a limited training data set.

7.4.1.2 Narrative

Some customer service representatives have to deal with a large number of requests every day. Regular AI customer service systems, which are online 24/7, are capable of offering instant assistance, which frees up labour resources to a large extent. However, it is quite challenging, if not impossible, for those systems to interpret emotions from customer input and respond in as friendly a manner as a human being.

Against this background, based on a huge data set of customer comments and the rich experience of NLP, this AI system can automatically detect sentiments like "happy, angry, anxious," etc. Moreover, this system can also detect the intensity of the customer's sentiment. Furthermore, this use case adapts convolutional neural networks, a widely used technique in visual computing, to interpret the semantic meaning of the customer's expression. It can improve the system's performance for sentiment classification and intensity detection. Moreover, with the adoption of transfer learning, the system can also be applied to various types of data. To overcome the difficulty of limited training data, the AI system also use data augmentation methods such as reverse translation and data noise to increase the variability of training data.

Up to now, the system has reached 90 % recall and 74 % accuracy for sentiment classification over seven categories. The overall recall and accuracy for sentiment intensity are also around 85 %. The system has increased customer satisfaction by 57 %.

7.4.1.3 Stakeholders and stakeholder considerations

Stakeholders: Customer service and customers.

7.4.1.4 Data characteristics

Description: conversation data from after-sales customer services for sentiment classification. These data are annotated by professional annotators into seven categories of sentiments. For sentiment intensity detection: Consists of sentiment data for “angry” and “anxious” only; these data are annotated into three degrees of intensity: “low, medium, high”.

Source: conversation data from real-time customer services.

Type: text

Volume (size): around 60 000 sentences for sentiment classification and 20 000 for sentiment intensity detection.

Velocity: batch processing

Variety: real-time data from various categories of products.

Variability (rate of change): static

Quality: high

7.4.1.5 KPIs

Customer satisfaction: the rate of customer satisfaction when using this system for requests. The expectation is 100 %. The reference of this KPI to mentioned use case objectives is increasing the rate as much as possible.

Accuracy: the current overall rate of accurate prediction across all customer sentiment classifications is 76,4 %. The reference of this KPI to mentioned use case objectives is increasing the rate to 90 %.

Recall: the current overall rate of accurate prediction across all customer sentiment intensity detections is 90 %. The reference of this KPI to mentioned use case objectives is increasing the rate to 90 %.

Accuracy: the current overall rate of accurate prediction across all customer sentiment intensity detections is 85 %. The reference of this KPI to mentioned use case objectives is increasing the rate to 90 %.

Recall: the current overall rate of accurate prediction across all customer sentiment intensity detections is 85 %. The reference of this KPI to mentioned use case objectives is increasing the rate to 90 %.

7.4.1.6 Features of use case

Task(s): natural language processing

Level of automation: conditional automation

Method(s): deep learning, transfer learning, data augmentation

7.4.1.7 Threats and vulnerabilities

The low degree of humanization, and lack of semantic diversity for response.

Reducing the number of human customer service staff.

Identified privacy risks include specific data collected from customers.

7.4.1.8 Challenges and issues

Challenge: the system's performance is expected to be as good as a human customer server.

Issues: 1) limited training data; 2) sentiment classification among seven categories.

Adequately addressing identified privacy risks.

7.4.1.9 Trustworthiness considerations

Improving the efficiency of customer service, improving the customer service experience.

Reducing labour costs and reducing operating costs.

This use case raises serious privacy concerns, especially with respect to specific customer data.

7.4.1.10 Use of standards and standardization opportunities

The system can be promoted to as many customer services companies as possible once provided with enough training data for the specific application scenario.

7.4.2 Deep learning-based user intent recognition (use case 43)

7.4.2.1 Objectives

Recognizing users' intent in order to solve their problems in e-commerce fields.

To recognize and understand users' intent using AI and deep learning technologies and apply such technologies to build chat bot systems to further reduce labour costs and for use in various fields.

7.4.2.2 Narrative

This intelligent customer service chat bot is mainly used to categorize users' questions, recognize users' intent and intelligently answer users' questions related to various business jobs. Currently, this chatbot is used to handle 90 % of online customer service and has saved the company on labour costs each year.

The AI provider is committed to using technology to drive business growth and improve the user experience in all areas of customer service.

Based on improved customer advice experience and the evolution of artificial intelligence technology, the company decided back in 2012 to develop intelligent chatbots to meet the needs of its growing business, to save on customer service costs and to increase service capability. Intent recognition is a key and core technology for building this type of intelligent customer service chat bot. By applying natural language processing technologies, deep learning technologies, and traditional machine learning algorithms, intent recognition accuracy has reached 95 %. Based on accurate intent recognition and a series of solution-finding algorithms, our chat bot can solve the user's problems in most cases and give the user a high-quality consulting experience. Finally, to provide diversified and personalized customer services, the AI system has continuously improved the accuracy of intent recognition, personalized solution generation, sentiment recognition, and image recognition. So far, intelligent customer service has revolutionized the traditional customer service consulting business.

7.4.2.3 Stakeholders and stakeholder considerations

Stakeholders: Customer service and customers

Stakeholder Considerations: Customer service expectations and customer experience

7.4.2.4 Data characteristics

Description: question response data from the online dialogue log

Source: customer dialogue log

Type: text

Volume (size): millions

Velocity: real time

Variety: various scenarios, various businesses, various categories of products

Variability (rate of change): nonlinear

Quality: good

7.4.2.5 KPIs

Accuracy: the percentage of users whose intent is correctly recognized over the total number of users. Currently, accuracy is 95 %. The reference of this KPI to mentioned use case objectives is to improve accuracy of recognizing users' intent.

Resolution: the percentage of questions solved over the total number of questions asked. The reference of this KPI to mentioned use case objectives is to improve resolution of questions from users.

Satisfaction: the percentage of users who are satisfied with customer service over the total number of users. The reference of this KPI to mentioned use case objectives is to improve user experience.

7.4.2.6 Features of use case

Task(s): natural language processing

Level of automation: conditional automation

Method(s): machine learning and deep learning

Platform: GPU and CPU

Topology: TensorFlow

7.4.2.7 Threats and vulnerabilities

High semantic ambiguity, multiple language expressions in one sentence.

Identified privacy risks include risks related to knowledge of user intent and training data.

7.4.2.8 Challenges and issues

Current challenges of deep learning and intent recognition:

— high semantic ambiguity, similar sentences can convey different meanings;

— unclear classification rules caused by complicated business logics;

— hard to answer reasoning questions;

— adequately addressing identified privacy risks.

7.4.2.9 Trustworthiness considerations

Solve problems intelligently to increase efficiency.

Free up labour from repetitive work to save a large amount of resources for society.

This use case raises serious privacy concerns, particularly with respect to the knowledge and training data related to user intent recognition.

7.4.2.10 Use of standards and standardization opportunities

Process standardization would improve quality and productivity.

7.4.3 AI virtual assistant for customer support and service (use case 106)

7.4.3.1 Objectives

Customer support service, product and service consulting.

Limitations - support for dialogues exclusively within the AI provider's products.

Target audience - b2b, b2c clients of the AI provider.

Optimization of company resources for support and customer service by automating the customer service process. As a result of the implementation of the system, the company was able to cover a greater volume of customer requests without necessity to increase its staff of operators. This prevented an increase in the company's operating expenses.

7.4.3.2 Narrative

The system automatically answers customer questions within the application and on the company website. At its peak, service automation reaches 85 %.

Chatbot assists the client in the selection of tariffs and services, and advises on the financial condition of the account. Chatbot promotes new products without the need for an operator. The client can ask a question in free form; the system would understand the request. If necessary, the system can ask additional questions before delivering its answer to the client.

Chatbot is integrated with internal billing systems, CRM, with a product catalogue and many other key services of the company. This allows each client to be provided with an individualized service. If chatbot is unable to help the client, or if the service procedure requires an operator, the dialogue is transferred to the operator. Currently, chatbot serves more than 1 million requests per month, working 24/7 to serve customers.

7.4.3.3 Stakeholders and stakeholder considerations

Stakeholders: Customer service department and customers.

Stakeholder considerations: Customer service department - maintaining / increasing customer loyalty, saving resources.

Stakeholder considerations: Customers - meeting expectations

7.4.3.4 Data characteristics

Source: customer profiles and a history of questions

Type: text, voice

Volume (size): Millions of hits (historical data)

Velocity: in real time

Variety: collected data sets

Variability (rate of change): the system is updated daily with new scenarios.

Quality: high

7.4.3.5 KPIs

Automation: solving a customer issue with a chatbot without operator intervention. The reference of this KPI to mentioned use case objectives is the optimization of customer service costs.

Quality: customer satisfaction rating. The reference of this KPI to mentioned use case objectives is to ensure high customer loyalty to the AI provider’s brand.

7.4.3.6 Features of use case

Task(s): optimization, natural language understanding, dialogue management

Level of automation: assistance

Method(s): deep learning, NLP

7.4.3.7 Threats and vulnerabilities

Information security, communication secrecy and the safety of personal data

Identified security and privacy risks include security and privacy risks related to client profiles and question history.

7.4.3.8 Challenges and issues

The readiness of external systems’ API for integration with the bot platform.

Biased customer attitudes towards chatbots.

Adequately addressing identified security and privacy risks.

7.4.3.9 Trustworthiness considerations

This use case raises serious security and privacy concerns, particularly with respect to client profiles and question history.

7.4.3.10 Use of standards and standardization opportunities

None identified.

7.4.4 Customer relation management (CRM) (use case 157)

7.4.4.1 Objectives

This use case focuses on determining which ethical issues arise in the use of AI in CRM and how can they best be addressed.

Smart information systems (SIS - big data and artificial intelligence) are used in customer relations management (CRM) to help manage large customer databases and improve customer interaction by companies.

The aim of the use case is to identify what ethical issues arise from the use of SIS in CRM, whether companies that use SIS for CRM have policies and procedures in place for addressing these concerns, and if practitioners are facing additional issues not addressed in the current literature.

The aim of the company's model is to further increase customer orientation and cost-efficiency.

7.4.4.2 Narrative

Customer relationship management deals with the processes and systems that support business strategies to build long-term and profitable relations with customers.

Access to customers' online data (through social networks, search engine history, cookies and other tracking systems) allow companies to gather a variety of information about customers. Such access also allows companies to create cloud systems with data, and strategize and automate CRM practices.

Data engineering blends data from all sources and builds ML models that will accurately predict the propensity of a customer to churn (cease to be a customer) as well as predict the propensity to buy a new product.

The use case identified the following issues: autonomy, control, manipulation, privacy, knowledge/lack of knowledge, informed consent, bias, responsibility and trust.

7.4.4.3 Stakeholders and stakeholder considerations

Stakeholders: Companies, marketers, customers and public.

Stakeholder considerations:

Autonomy, control and manipulation: there is a concern that the autonomy of customers is being undermined through the employment of data analytics. Data analysis techniques are not a neutral tool or measurement: they can expand, constrain or alter people's choices and behaviours, each of which has an influence over the user. Data analytics treats individuals as already having made or being inevitably about to make certain choices, often with a moral component. The rise in consumer debt occurred at the same time as the development of complex marketing profile methods that included neural networks and predictive models to target consumers.

Privacy and knowledge: accessing the internet using big tech companies' platforms results in the disclosure of personal information to the platforms and other websites, even if customers do not want to reveal this information. Therefore, privacy issues are a significant concern for CRM practices: people unknowingly and unintentionally are communicating personal data. For marketers, big data are powerful weapons for capturing consumer data directly, indirectly, unobtrusively, with and without permission and participation.

Bias: concerns of bias entering into AI and ML have been growing. Despite surface assumptions that computers are unbiased (as, for example, they do not recognize skin colour or gender) increasing research has been conducted evidencing the potential for the outputs of automated systems to be prejudiced against certain groups of society.

Responsibility and trust: the unequal power relationship between companies and consumers creates concerns of accountability, and data ownership. Thus, responsibility for the data are a key concern.

Transparency and the company's vulnerability: it is usually difficult to make algorithms public, and most customers would not have the ability to understand those algorithms or the input data needed to run them. Furthermore, the algorithms change quickly and models are constantly being updated. From this, one can see how the notion of public "acceptance" of certain algorithms does not really work.

7.4.4.4 Data characteristics

Description: customers' online data

Source: social networks, search engine history, cookies and other tracking systems

7.4.4.5 KPIs

To help manage large customer databases and improve customer interaction.

To predict the propensity of a customer to stop being a customer as well as the propensity to buy a new product.

7.4.4.6 Features of use case

Task(s): prediction

Level of automation: assistance

7.4.4.7 Threats and vulnerabilities

Informed Consent: A challenge to the collection and use of data for CRM and elsewhere is the possible lack of informed consent given by the consumer for data to be used. Some data is given initially for a particular reason, but can be replaced by other interests at a later date. In such cases the company can hold the data but not have received informed consent for its use to the latter end.

In the case of CRM, the data are collected by the company as customers subscribe to services, but customers do not always consent to having their data analysed for insights or profiling. Even when consent is given the apparent means of gaining consent can consist of merely ticking a box to say that it is agreed to the terms and conditions. However, it is widely accepted that very few people ever read these terms and conditions, and as such the validity of that consent is questionable.

GDPR and different regulations outside EU: It would be comparatively easy for a company based and operating outside of the EU jurisdiction to develop a data set that would be unethical or illegal to produce in Europe. Based on this data set, the company can then develop a highly effective AI which can then be sold in Europe. This would be economically advantageous to that non-EU company, provided no personal data collected in this manner ever entered the EU.

7.4.4.8 Challenges and issues

Lack of knowledge on behalf of the customers about what is happening with their data.

Guidelines to make informed consent more realistic (rather than clicking away the pop-up and unknowingly giving consent).

Adequately addressing identified privacy risks.

7.4.4.9 Trustworthiness considerations

There is a need to protect customers from being exploited and misguided through data manipulation and analysis. New advances in CRM practices and data technologies bring social benefits, but also ethical issues to be examined.

7.4.4.10 Use of standards and standardization opportunities

How to deal with conflicting ethical issues and its transparency?

The researchers of this use case have drafted a proposal for guidelines for ethical AI, based on ITIL, COBIT, ISO/IEC 20000 series and CRISP-DM.

7.5 Education

7.5.1 A recommendation system for industrial training (use case 23)

7.5.1.1 Objectives

Find skill requirements and relevant training based on an employee's career objectives.

Recommend a personalized list of "best" training courses to an employee, which can help him/her meet his/her career objectives.

7.5.1.2 Narrative

The recommendation system helps employees improve their skills by recommending appropriate training courses from a given list and historical data.

Continuous training is crucial for creating and maintaining the right skill profile for an industrial organization's workforce. There is tremendous variety in the available training within an organization: technical, project management, quality, leadership, domain-specific, soft-skills, etc. Hence it is important to assist the employee in choosing the best training that perfectly suits his/her background, project needs and career goals. In this use case, the AI system is built on algorithms for recommending training in an industrial setting. The AI system provides training recommendations, taking into account the employee's training and work history. The AI producer has developed several new unsupervised sequence mining algorithms to mine past training data from the organization for making the next personalized training recommendation. Using real-life data about training 118 587 employees over 5 019 distinct training courses from a large multi-national IT organization, this use case show that these algorithms outperform several standard recommendation engine algorithms as well as those based on standard sequence mining algorithms.

7.5.1.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the recommendation system;

— data provider: the organization that provides the training course;

— AI users: the organization that adopts this recommendation system and the employees that use this recommendation system for industrial training.

The stakeholder considerations include:

— the AI users consider job requirements, skill profile, job description requirements and training requirements by using this recommendation system.

7.5.1.4 Data characteristics

The data used in this use case are the employee's training and work history from the organization which includes real-life data about training 118 587 employees over 5 019 distinct training courses.

7.5.1.5 KPIs

Prediction accuracy: number of employees undertaking courses from the recommended list. The reference of this KPI to mentioned use case objectives is to improve accuracy.

7.5.1.6 Features of use case

Task(s): recommendation

Level of automation: partial automation

Method(s): deep learning

Platform: GPU-enabled servers

Topology: GPU-enabled servers

7.5.1.7 Threats and vulnerabilities

The threats include that the different sources of bias based on model training, incorrect AI system use can cause stress in employees.

7.5.1.8 Challenges and issues

The challenges include that the large amounts of training data are desired and predicting human behaviour is tricky.

7.5.1.9 Trustworthiness considerations

This recommendation system is desired to consider the impact on the motivation of the AI user because the employees can feel challenged or demoralized by using recommendations.

7.5.1.10 Use of standards and standardization opportunities

The standards for the unsupervised sequence mining algorithms to mine the past data can be used for this system.

7.5.2 An intelligent marking system (use case 83)

7.5.2.1 Objectives

The system can realize intelligent detection and grading of all subjective questions to significantly reduce labour and organizational costs.

7.5.2.2 Narrative

This use case is an intelligent marking system based on core technology design research, including independent intellectual property rights of the AI producer, handwriting recognition, natural language understanding, intelligent evaluation and other artificial intelligence. It can realize the detection of blank questions for all question types except multiple choice questions, and computer intelligent evaluation of Chinese/English composition, English translation, literature synthesis category short answer questions and English blank questions. At the same time, for Chinese/English composition, it can also effectively detect abnormal answer papers which are highly similar to the dry content of the test paper or the content of the external model text.

The intelligent marking system can provide a new generation of intelligent scanning network evaluation solutions for large-scale pen and paper examinations combined with the mature scanning network evaluation technology. In the process of scanning, the detection and screening of similar volume and blank volume questions and the intelligent evaluation of subjective questions are carried out in real time. Taking the data outputted from the scanning link as the objective third party quality evaluation standard, the online or offline quality monitoring of the marking paper is carried out to improve the quality of the marking paper. At the same time, the computer intelligent evaluation of subjective questions can assist manual marking to a certain extent and effectively reduce the workload of manual marking of subjective questions.

The intelligent marking system has many advantages. First, it has a scientific and unified scoring standard, which can avoid the differences in the scoring scale and subjective interference among different reviewers, and ensure the fairness of the marking results. Second, the investment requires only a small number of technical personnel and servers, which can reduce the organizational cost of existing manual marking by about 50 %. Third, it can detect the abnormalities in the answers, such as blank questions and similar volumes. At the same time, through the real-time comparison with the manual marking data, it achieves quality monitoring.

7.5.2.3 Stakeholders and stakeholder considerations

Stakeholders: marking teachers and technicists

Stakeholder considerations: efficiency

The stakeholders include:

— AI users: marking teachers and technicists;

— data provider: the school that adopts this system and provides scanning student papers.

The stakeholder considerations include:

— AI users consider the efficiency of this system.

7.5.2.4 Data characteristics

Description: scanning student papers

Source: the data from scanning student papers

Type: text, picture

Velocity: batch processing

Variety: single source

Variability (rate of change): static

Quality: high

7.5.2.5 KPIs

The KPIs of this system include:

— cost: reduce the cost of existing manual marking;

— efficiency: improve the efficiency of existing manual marking;

— accuracy: improve the accuracy of existing manual marking.

7.5.2.6 Features of use case

Task(s): natural language processing

Level of automation: high automation

Method(s): deep learning, semantic recognition

7.5.2.7 Threats and vulnerabilities

The inadequate accuracy of the system is the threat to be considered.

7.5.2.8 Challenges and issues

The accuracy of marking papers is necessary to be further improved.

7.5.2.9 Trustworthiness considerations

There is a scientific and unified scoring standard, which can ensure the fairness of the marking results.

This system can significantly reduce labour and organizational costs. This use case can have impacts on teacher job descriptions and roles.

7.5.2.10 Use of standards and standardization opportunities

None identified.

7.5.3 Intelligent educational robot (use case 84)

7.5.3.1 Objectives

The robot is designed to support children's education and to improve the pleasure of learning.

7.5.3.2 Narrative

The educational robot is a new teaching tool to cultivate students' comprehensive ability. It mainly uses artificial intelligence technology, speech recognition technology and bionic technology to cultivate students' various abilities. Educational robots have hearing, vision, oral skills, recognition, emotional detection and the ability to interact for a long time.

An educational robot is designed for interaction with children, primary school students, and junior high school students. It can be used for study or entertainment.

It has teaching materials and lectures by famous teachers for all grades and disciplines, and students can receive high-quality teaching without leaving home. It can also present knowledge forgotten by students and solve students' learning problems in time. Moreover, it can correct students' bad habits such as overdependence on others, hating to get out of bed, and inattentiveness by giving instructions and intelligent reminders so as to cultivate students' good learning behaviour and living habits. For learning English, the robot can train students' oral English ability by practicing dialogue with them, and can also bring students' pronunciation more in line with the standard and improve students' communication ability.

7.5.3.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI users: students, parents and teachers;

— data provider: the learner provides the learner inputs.

The stakeholder considerations include:

— the teachers (AI users) consider the students' grades and learning interest.

7.5.3.4 Data characteristics

Description: learner input, including pronunciation, visual information, keystrokes, etc.

Source: data from the learner

Type: voices, visual information, keystrokes, etc.

Velocity: batch processing

Variety: multiple source

Variability (rate of change): static

Quality: high

7.5.3.5 KPIs

The KPIs of this system include:

— interest: improve students' interest in learning;

— grades: improve students' academic performance.

7.5.3.6 Features of use case

Task(s): recognition

Level of automation: high automation

Method(s): deep learning, automatic speech recognition, bionics techniques

7.5.3.7 Threats and vulnerabilities

The inadequate teaching effect of intelligent robot is the threat to be considered.

7.5.3.8 Challenges and issues

The challenges include:

— be able to sense students' emotions like teachers;

— accurately capture students' gestures, postures, face information, etc.

7.5.3.9 Trustworthiness considerations

This system is to consider the impacts on the students, such as the system can give students emotional support and stimulate students' interest in learning.

After repeated training, the intelligent educational robots can assist students’ study like teachers. It is to consider the transparency, explainability and predictability of the actions of the robots.

7.5.3.10 Use of standards and standardization opportunities

None identified.

7.5.4 AI system to intelligent campus (use case 85)

7.5.4.1 Objectives

This use case is a full range of products and integrated solutions for teaching, examination, evaluation, management and learning.

This use case provides a comprehensive intelligent sensing environment and comprehensive information service platform for teachers and students, so as to realize the integration of human and business information.

7.5.4.2 Narrative

Based on big data and artificial intelligence technology, the use case brings teaching, examination, learning and management into the integrated system of mutual cooperation, based on accompanying data acquisition and dynamic big data analysis, combined with process evaluation, to help teachers and students to realize teaching according to their aptitude and individualized learning, to help managers to supervise and assist decision-making, and to greatly promote the transformation of education, learning and management to intelligence.

In teaching, the AI system includes an intelligent and efficient classroom based on the cloud network end. Through docking the resource cloud platform and school-based resource library, it can realize synchronous push of high-quality resources and help teachers prepare classes efficiently.

In the area of examinations, the oral evaluation technology used in this AI system is the only technology to have passed the certification of the National Language Commission, widely used in the national Chinese Mandarin online test and used in classroom teaching. The AI provider applies its the industry-exclusive artificial intelligence core technology to the examination and automatic approval of traditional offline homework, which greatly reduces the burden on teachers and digitalizes the daily examination process. Big data analysis technology can be used to promote personalized teaching and learning.

In learning, this AI system supports students' online adaptive learning by building a question bank system, evaluation system and online learning system. Through analysis of the students' examination results, the system can evaluate the students' mastery of knowledge points and the stability of their grades, and then combine the key points of the teaching materials with the high-frequency test points. Through intelligent analysis, the scheme can recommend the optimal learning path.

In the area of management, this AI system includes a smart campus solution that covers more than 10 departments such as the academic affairs office, student office and school office. The system provides more than 60 applications to meet the needs of normal campus management. It is worth mentioning that in order to cope with the challenges of educational administration brought by the new curriculum and the new college entrance examination, this AI system, based on the classification algorithm of a deep neural network, puts forward an intelligent course arrangement system, effectively avoids the conflict of course selection, and realizes the optimal voluntary satisfaction rate under the premise of uniform teaching and classroom resources, so that every student can attend classes according to his own volition.

7.5.4.3 Stakeholders and stakeholder considerations

Stakeholders: students, teachers, schools, governments

Stakeholder considerations: privacy

The stakeholders include:

— AI provider: the organization brings teaching, examination, learning and management into the integrated system of mutual cooperation;

— AI users: students, teachers, schools;

— data provider: the schools that adopt this AI system;

— AI subject: governments.

The stakeholder considerations include:

— AI provider considers the students' privacy.

7.5.4.4 Data characteristics

Description: the data comes from students and teachers as well as from their learning and administration processes.

Source: intelligent education products or platforms

Type: structured or unstructured data

Velocity: in real time

Variety: students information, teachers' information, information generated during the course of teaching, learning and management.

Variability (rate of change): in real time

7.5.4.5 KPIs

The KPIs of this system include:

— efficiency: improve students' learning and teachers' administration efficiency. The reference of this KPI to mentioned use case objectives is to improve efficiency.

7.5.4.6 Features of use case

Task(s): knowledge processing and discovery

Level of automation: high automation

Method(s): knowledge processing and discovery

7.5.4.7 Threats and vulnerabilities

Disclosure of privacy data for teachers and students is the threat of this AI system.

7.5.4.8 Challenges and issues

The implementation of an intelligent campus requires the collection and processing of large quantities of data on students and teachers, which is likely to lead to the disclosure of private data. Therefore, the data privacy protection mechanism is expected to be strengthened in the intelligent platform.

7.5.4.9 Trustworthiness considerations

This AI system has not identified the trustworthiness considerations on bias mitigation, explainability, controllability, predictability, transparency, verification, robustness, reliability and resilience.

Regarding ethical and societal concerns, this AI system of intelligent campus is desired to consider the impacts on the AI users (teachers and students), because it brings artificial intelligence technology onto the campus and into the classroom, promotes students' learning and teachers' teaching, and facilitates teaching management.

7.5.4.10 Use of standards and standardization opportunities

None identified.

7.5.5 AI adaptive learning platform for personalized learning (use case 102)

7.5.5.1 Objectives

This use case is an AI system for equal access to high-quality education.

Properties of the system are: open access, interactive tasks, personalization, user-generated content and learning graph.

7.5.5.2 Narrative

The adaptive learning platform (AiEd platform) is an e-learning platform and course-builder that uses AI for forming adaptive learning paths.

The adaptive learning platform is a cloud-based platform designed to create and distribute interactive educational content, enhanced by various types of automatically graded assignments with a real-time feedback. The platform is suitable for any kind of e-learning activity, from private on-campus classes to massive open online courses (MOOCs)s. The platform is designed keeping the needs of computer science education in mind.

The platform aims to apply data mining techniques to make education more efficient and to improve the way people learn and teach. Adaptive and personalized learning are one of the key priorities of our platform.

7.5.5.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI users: students;

— data provider: teachers (content providers);

— AI subjects: third-party services (via experience API), academic researchers (sets of eduDATA).

The stakeholder considerations include:

— AI provider and AI users consider the personal data concerning interests and preferences, safety, privacy of students;

— data providers (teachers, content providers) consider the reputation, trustworthiness, high quality content;

— AI subjects (third-party actors) consider the safety of the system.

7.5.5.4 Data characteristics

The data used in this use case are the student's education and learning history.

7.5.5.5 KPIs

The KPIs include:

— performance: increases educational results through personalized learning process. The reference of this KPI to mentioned use case objectives is personalization;

— variability: educational content makes up the learning graph. The AI adaptive learning engine based on the learning graph allows automatically creation of a huge number of education programs, acceptable for everyone. The reference of this KPI to mentioned use case objectives is learning graph.

7.5.5.6 Features of use case

Task(s): optimization

Level of automation: high automation

Method(s): recommendation-based approach that uses item response theory and the ELO rating system.

Platform: none, cloud-based solution is used

7.5.5.7 Threats and vulnerabilities

Verification of new content is the threat of this AI system.

7.5.5.8 Challenges and issues

Edstories (micro-learning video stories) are desired to be included to satisfy the pedagogical model of movement-based learning.

7.5.5.9 Trustworthiness considerations

The system is expected to be integrated into secondary and tertiary school systems that still face legal boundaries and limitations for scaling, it is desired to consider the trustworthiness issues related to the impact of the education of the students (e.g. the fairness of education opportunities).

In addition, after repeated training, the adaptive learning system would be highly efficient. It is desired to consider the transparency and the predictability of the adaptive learning system.

7.5.5.10 Use of standards and standardization opportunities

None identified.

7.5.6 AI adaptive learning mobile app (use case 124)

7.5.6.1 Objectives

A foreign language learning mobile program for all age groups that adapts to the student and builds an individual learning track based on artificial intelligence.

Providing easy, convenient and adaptive learning of English with the help of a virtual teacher based on artificial intelligence.

7.5.6.2 Narrative

This use case describes a mobile application for learning English, which is based on a program that adapts content to the student and learns with them. During registration, the program analyses the user's account on a social network and draws up an individual training plan based on the student's interests.

The application analyses successes and develops a curriculum adapted for each user. The user is required to first indicate their level of knowledge of the language and follow the instructions of the virtual teacher.

The program pays more attention to developing vocabulary and learning grammar rules. Notably, the program collects various information while the student interacts with it, including the user's training rate, the percentage of correct and erroneous answers, how well the user knows and understands various grammar rules, etc. By collecting this information, the application can tailor different activities to meet goals that have already been achieved and those toward which the student still wants to strive.

The virtual teacher suggests choosing a level of difficulty, and then monitors the execution of tests and tasks, analysing errors. If the student cannot cope, it offers to repeat the material. If the student solves the tests and tasks without errors, it skips.

7.5.6.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI provider: the organization that develops the AI system and provides the service.

— AI users: all age groups with the goal of learning a foreign language.

The stakeholder considerations include:

— AI provider considers the learning interest, effectiveness of acquiring new knowledge, the involvement in the educational process through gamification.

7.5.6.4 Data characteristics

The data used in this use case are the information while the student interacts with the AI system, including the user's training rate, the percentage of correct and erroneous answers, how well the user knows and understands various grammar rules, etc.

7.5.6.5 KPIs

The KPIs include:

— efficiency: improve students' learning through an adaptive learning format. The reference of this KPI to mentioned use case objectives is to improve efficiency;

— interest: improve students' interest in learning. The reference of this KPI to mentioned use case objectives is to improve involvement.

7.5.6.6 Features of use case

Task(s): optimization

Level of automation: partial automation

Method(s): in-depth study of user actions and user information

7.5.6.7 Threats and vulnerabilities

Teaching effect of virtual teacher is the threat of using this AI system.

7.5.6.8 Challenges and issues

The development of a personalized approach to learning is the challenge of this AI system.

7.5.6.9 Trustworthiness considerations

This case of the use of artificial intelligence in the educational process can complement teachers as knowledge transmitters and make education accessible to everyone. It is desired to consider the fairness of the accessibility of the students.

At the same time, artificial intelligence, performing the functions of analytics, packaging and personalization of educational content, is much more effective than a person in the role of an assistant to a teacher and shifts the role of a classical teacher towards mentoring. It is desired to consider the impact of the job descriptions of the teachers.

In addition, the use of a virtual teacher in the educational process enables the analysis of student actions and builds individualized learning tracks based on the data received. It is desired to consider the bias and the verification perspectives regarding the impact on the students.

7.5.6.10 Use of standards and standardization opportunities

None identified.

7.6 Energy

7.6.1 Smart energy grid (use case 166)

7.6.1.1 Objectives

This use case is a smart grid case study that explores the principal ethical issues that occur in the use of smart information systems (SIS) in smart grids and offers suggestions as to how they can be addressed.

7.6.1.2 Narrative

The introduction of smart grids and active demand systems that monitor and incentivise alternative energy consumption habits can enable dynamic consumption of energy by enabling customers to shift their consumption to take advantage of dynamic pricing.

The use of SIS in energy promises to ensure sustainable affordable energy for the ever-increasing demands of smart living without big investments in the energy distribution systems in two ways. First, SIS systems optimize the management of energy demand and energy supply from existing resources. Smart grids also involve a host of intelligent technologies to improve the management of the energy distribution network that connects energy producers with consumers.

7.6.1.3 Stakeholders and stakeholder considerations

Stakeholders: Energy distribution system operators, energy suppliers, energy consumers, politicians and (governmental) policy advisors.

Stakeholder considerations:

— affordability and energy equity: While smart grids are seen as one of the solutions to effecting energy justice or equity, what they in fact try to achieve is energy abundance so that there is enough supply to satisfy the disproportionate increases in energy demand. Poorer socio-economic strata would be the most motivated to save on their energy costs but can find it difficult to benefit from dynamic pricing as their energy use is frugal to begin with. There is also a concern that dynamic pricing will leave consumers who are unable to shift their energy consumption worse off, as companies will try to 'penalise' energy use during peak times by raising prices.

— privacy and informed consent: smart meter acceptance levels are dropping due to ethical tensions in the interaction between customers and organisations. GDPR has raised the public's privacy concerns and suspicion towards companies.

7.6.1.4 Data characteristics

The data used in this use case include energy demand and energy supply from existing resources.

7.6.1.5 KPIs

The accuracy of optimising the management of energy demand and energy supply from existing resources.

7.6.1.6 Features of use case

Level of automation: conditional automation

7.6.1.7 Threats and vulnerabilities

Cyber-risks and Security: do we jeopardize energy security?

Cyber-attacks on the smart grid can do significant damage, yet are still inadequately addressed, due to their low frequency of occurrence.

7.6.1.8 Challenges and issues

Sustainability: doing our bit for climate change.

The more accurate and real-time the modelling matches energy production with consumption, the more responsive the grid will be in managing electricity flows. On the other hand, smart meters and SIS technologies come at an energy cost.

Prioritization of energy distribution and energy justice.

While the goal of smart grids is to avoid energy scarcity, balancing energy provision between competing priorities will become a point of political and societal debate. These concerns relate primarily to issues of distributive justice.

7.6.1.9 Trustworthiness considerations

Privacy and Informed Consent: Does the smart grid give away household privacy?

Privacy concerns relate to the granularity of electricity consumption data that can be collected about a household and the intimate lifestyle information that can be inferred from it. To mitigate such concerns, and in response to the GDPR, utilities have embarked on educating customers regarding their privacy policies, the reasons behind data collection and sharing, accessing of data, and consumer rights.

There is a lack of deliberation of how these ethical issues, arising from the constant surveillance of energy consumption in the energy transition will be resolved, or institutional and political commitments on how they will mitigate. While the discussion in the energy transition debate elaborates on the ability to incentivise behavioural change via economic benefits, the ethical and social implications of dynamic pricing with respect to energy equity or energy justice are neglected.

7.6.1.10 Use of standards and standardization opportunities

— code of conduct for system operators on use of data to build trust with consumers in terms of goals, privacy and informed consent;

— guidance for companies for ethics by design principles to build AI systems;

— guidance for companies to consult the public about innovations and strategic developments;

— guidance on how to deal with conflicting ethical issues and its transparency;

— requirements on a fair costing and pricing system for the different user groups.

7.7 Fintech

7.7.1 Detection of fraud based on collusion (use case 20)

7.7.1.1 Objectives

Validating the predicted collusion set is effort-intensive, investigative and asks legal expertise.

Automatic unsupervised detection of frauds based on collusions.

7.7.1.2 Narrative

This use case describes a set of unsupervised machine learning algorithms to detect collusion-based frauds, particularly circular trading and price manipulation in stock market trading.

Fraud is prevalent across all industries; and is particularly severe in today's computerized, web-connected, mobile-accessible, and cloud-enabled business environments. A Federal Bureau of Investigation (FBI) report states that the insurance industry in the US, which consists of over seven thousand companies and collects over one trillion dollars in premiums, loses about 40 billion dollars annually to fraud in the non-health insurance sector alone. The aggregate size of the 52 regulated stock exchanges across the world (total market capitalization) was $55 trillion as of December 2012. Given the money involved, it is not surprising that the stock market is a target of fraud.

Many malpractices in stock market trading, e.g. circular trading and price manipulation, use the modus operandi of collusion. Informally, a set of traders is a candidate collusion set when they have “heavy trading” among themselves, as compared to their trading with others. This use case formalizes the problem of detection of collusion sets, if any, in a given trading database. This use case show that naïve approaches are inefficient for real-life situations. This use case adapts and apply two well-known graph clustering algorithms for this problem. This use case also proposes a new graph clustering algorithm, specifically tailored for detecting collusion sets; further, this use case establishes a combined collusion set. Treating individual experiments as evidence, this approach allows us to quantify the confidence (or belief) in the candidate collusion sets. This use case has carried out detailed simulation experiments to demonstrate effectiveness of the proposed algorithms. The system is also operational in a government organization. Note that all our collusion detection algorithms are completely unsupervised and do not need any training data.

7.7.1.3 Stakeholders and stakeholder considerations

Stakeholders: stock market regulator, stock traders and stock investors.

Stakeholder considerations: fair price, prevention of collusions and frauds.

7.7.1.4 Data characteristics

The data used in this use case are the trading data consists of records having the following form:

stock ID, timestamp, seller ID, buyer ID, quantity, price, value.

Actual trading data contains many other details, e.g. trader details, client details, (sales and purchase) order details such as IDs, placement times, matching time, quantities and prices quoted in the order etc. Each record in this trading database refers to a single trading transaction.

7.7.1.5 KPIs

Prediction accuracy: How many predicted collusion sets were actually involved in fraudulent activites? The reference of this KPI to mentioned use case objectives is to improve accuracy.

7.7.1.6 Features of use case

Task(s): knowledge processing and discovery

Level of automation: partial automation

Method(s): machine learning

Platform: GPU-enabled servers

Topology: GPU-enabled servers

7.7.1.7 Threats and vulnerabilities

Incorrect fraud detection can lead to unnecessary alerts.

7.7.1.8 Challenges and issues

Actual examples of collusion-based fraud cannot be readily available, even for evaluation and testing.

7.7.1.9 Trustworthiness considerations

Incorrect detection of collusions and frauds can cause unnecessary stress in stock traders.

7.7.1.10 Use of standards and standardization opportunities

Graph-based clustering.

7.7.2 Virtual bank assistant (use case 57)

7.7.2.1 Objectives

Use of advanced chatbots and dialogue systems to automate part of call centre activities.

Provide better quality help desk support to employees.

7.7.2.2 Narrative

A bank's virtual assistant is the first point of contact for branch operators, who receive immediate answers at any time - it allows optimization of the time during which the service desk can be manned by human operators, who are dedicated to activities of greater value.

A bank in Italy has created a virtual consultant to support internal staff in their operations and interaction with customers.

The solution enabled a significant change in the service model of the bank, allowing it to achieve important results in terms of greater contact volumes, extension of service hours and reduction of low-value human-centric activities.

The virtual assistant has been conceived as the first (and only) access point for assistance. It is easy to use and responds with a high level of reliability to the questions of branch colleagues. The virtual assistant has not been designed as a simple “chatbot” trained on a specific topic but as a virtual “colleague” to turn to for any question, completely integrated into the bank knowledge chain. To date, the virtual bank assistant manages all 14 knowledge domains of the bank, receiving thousands of answers.

From the beginning of its use (January 2018), the virtual assistant manages 100 % of requests, partly independently and partly in collaboration with the human operators of the service desk.

The effectiveness of the solution is evidenced by the very high level of satisfaction, with positive feedback from users exceeding 90 % and the reduction in the time spent by service desk operators in providing support to the branches, which today can be quantified as a reduction of 25 %.

7.7.2.3 Stakeholders and stakeholder considerations

Stakeholders: bank support employees, bank customers, bank managers.

Stakeholder considerations: to provide better quality help desk support to employees.

7.7.2.4 Data characteristics

The data used in this use case include a knowledge base which includes 14 knowledge domains of the bank and didn’t include the training data for the chat bot.

7.7.2.5 KPIs

Greater contact volumes with the bank: the objective is to expand the quantity of internal support activities provided by the bank to its employees. The reference of this KPI to mentioned use case objectives is to improve productivity of service desk operators (already measured an improvement of 25 %).

Extension of service hours: expand internal support activities 24/7. The reference of this KPI to mentioned use case objectives is always on.

Reduction of low-value human-centric activities: reduction of low-level labour activities and let employees concentrate on more added-value activities. The reference of this KPI to mentioned use case objectives is to improve the quality of work.

7.7.2.6 Features of use case

Task(s): natural language dialogue systems

Level of automation: conditional automation

Method(s): NLP

Platform: web-based solution

7.7.2.7 Threats and vulnerabilities

None identified.

7.7.2.8 Challenges and issues

Provide natural and consistent interaction with users who have different levels of experience (and thus vocabulary) and different backgrounds.

7.7.2.9 Trustworthiness considerations

None identified.

7.7.2.10 Use of standards and standardization opportunities

None identified.

7.7.3 Forecasting prices of commodities (use case 91)

7.7.3.1 Objectives

Build a neural network to forecast the price of base metal commodities.

Use forecasted prices to interpret trading trends.

7.7.3.2 Narrative

A trading company sought to improve the accuracy of price point forecasting for specific commodities.

The trading company has access to very good data to develop regression models. However, its model was insufficient to differentiate the impact of long term versus short-term externalities. As such, a neural network was developed to ingest both structured market data as well as unstructured aggregate social media data to improve the inference and retraining ability to forecast prices.

7.7.3.3 Stakeholders and stakeholder considerations

Stakeholders: trading company, manufacturers, suppliers.

Stakeholder considerations: loss in spread of trades, market research for clients.

7.7.3.4 Data characteristics

The data used in this use case are the structured market data as well as unstructured aggregate social media data.

7.7.3.5 KPIs

Forecast accuracy: difference between forecasted and actual price. The use case depends on higher and timely accuracy of the price for necessary trades.

Model latency: the latency for the model to retrain and output inferences. As the trading sector becomes more automated, it was important for the model to reduce latency.

Money saved: the loss incurred in poor or negative spreads. The trading company can use better forecasts to save their clients' money and reduce stress on cash flow.

7.7.3.6 Features of use case

Task(s): prediction

Level of automation: assistance

Method(s): neural networks

Platform: on cloud service accessible by secure API.

Topology: star

7.7.3.7 Threats and vulnerabilities

Possible tightening of aggregate data access policies of social media platforms, which can require the neural network to be remodelled.

7.7.3.8 Challenges and issues

Modelling a neural network model that ingests a large and wide array of data, while calibrating for variables that have short-term versus long term impact.

7.7.3.9 Trustworthiness considerations

Unpredictable flow of materials and commodities due to price shocks.

7.7.3.10 Use of standards and standardization opportunities

AI price prediction can be used to predict the price and flow of goods can be used to hedge against unpredictable externalities such as civil unrest or territorial disputes when the accuracy of prices and amounts is critical to the mission.

7.7.4 Financial advice and asset management with AI (use case 114)

7.7.4.1 Objectives

Financial advice and portfolio management for financial institutions and consumers.

Designed to manage exchange-traded securities portfolios of conservative investors in real time, using asset price data and macroeconomic data, and make the most accurate decisions at a given yield and moderate risk. Prediction of significant depreciation of exchange-traded asset prices as a result of a sharp monetary contraction (financial crisis).

7.7.4.2 Narrative

The core of the system carries out a structured collection from open sources and multi-threaded parallel analysis of information; it regulates the application of basic algorithms and rules for changing these algorithms that modify the purpose of the task. (Intermediate goal setting is one of the elements of “strong AI”). One of the tasks is to assess market trends, as well as market and interest rate risk. Changes in the algorithm of actions depend on the macroeconomic information received from the outside. It translates notoriously weakly formalized parameters into specific decisions on the formation of investment portfolios and issues orders to brokers to purchase, rebalance or sell assets in stock exchanges. The macroeconomics unit is an autonomous system that generates indicators of time periods and geographical areas with different weights of investment potential.

For the purpose of realizing efficiency that cannot be achieved by competitors, the project uses more complex technologies than other standard solutions for building neural systems. All algorithms of the basic core of the project are developed by the creators themselves.

The idea that neural systems are absolute, impenetrable “black boxes” is a myth by the developers. Therefore, by understanding exactly what technologies are used to achieve analysis goals, overloaded “boxed” solutions can be optimized. This was done in the project.

The algorithm of simple regression analysis of prices (model William Sharpe/Harry Markowitz) did not lead to the required efficiency. Therefore, the project uses complex model when weighing factors and the algorithms of simple regression analysis of prices change depending on the “field,” formed by the regression assessment of other economic parameters.

The William Sharpe/Harry Markowitz model is a simplification because it is very resource-intensive, particularly when it comes to the hundreds of asset names globally for the diversification expected in this model. Applying a straight-line approach to the assessment of dozens or even hundreds of additional macroeconomic parameters of each of the dozens of different countries (in the interrelated world economy), ask either supercomputers and very expensive neural models, or building a fundamentally new economic model for the AI core.

In this project, the regression evaluation of higher-order macroeconomic indicators “guides” all subsequent lower-order models, thereby resolving issues.

7.7.4.3 Stakeholders and stakeholder considerations

Stakeholders: state regulators.

Stakeholder considerations: here is no difference from human control.

7.7.4.4 Data characteristics

Description: historical and real-time securities price data. Historical and real-time macroeconomic data.

Source: securities prices from exchanges, open source, websites of central banks and the International Monetary Fund (IMF).

Type: structured data

Volume (size): 4 TB

Velocity: real-time data replenishment 100 mbps

Variety: mostly structured

Variability (rate of change): high

Quality: high

7.7.4.5 KPIs

Portfolio yield: The percentage return of the portfolio compared to the benchmark. The reference of this KPI to mentioned use case objectives is long-term, from 10 y to 20 y, where the retention of positive returns is significantly higher than the base.

Sharpe ratio: Risk assessment strategies. The reference of this KPI to the use case objectives is that a higher Sharpe ratio is an indication of a higher level of control reliability (1 to 2 or more).

7.7.4.6 Features of use case

Task(s): prediction, advise and management

Level of automation: partial automation

Method(s): ensemble models

Platform: 64 GB RAM, 2 x Intel Core i7

7.7.4.7 Threats and vulnerabilities

Changes in state regulation of financial activities.

7.7.4.8 Challenges and issues

Data can be noisy, can have several missing values and are desired to have appropriate pre-processing and treatment before feeding to the model algorithm.

Working with financial assets requires high reliability of computing systems and replication systems.

7.7.4.9 Trustworthiness considerations

None identified.

7.7.4.10 Use of standards and standardization opportunities

None identified.

7.7.5 Loan in 7 minutes (use case 119)

7.7.5.1 Objectives

A completely automated solution that analyses customer behaviour and makes the best loan offers for a customer.

Create lending products for clients of medium and large businesses (LMB) with the shortest delivery time possible, taking into account the extremely detailed customer profile.

7.7.5.2 Narrative

Loan in 7 minutes is a solution where a credit decision is made by artificial intelligence in just a few minutes without human participation.

A complex machine learning settlement system was implemented on a largest Hadoop-cluster (tens of petabytes of data) and integrated into the bank's corporate lending business process.

The new project has significantly improved customer experience:

— eliminated the need for the client to contact the bank in person for a loan;

— asks no additional documentation from the client to get a decision;

— bank's automated systems were improved in terms of automatic transaction creation;

— substantially simplified the process of issuing a loan.

If the client applies for a loan, they fill out a short form on the bank's online system to reflect the recent changes in the business. As soon as the client provides the necessary information the solution is activated.

It interacts with the internal and external systems (e.g. transactional data and credit bureaus), collects all detailed information about the client, applies algorithms based on artificial intelligence and machine learning methods, automatically performs risk estimation and calculates appropriate offers for the client.

The client chooses the appropriate lending terms. The solution calculates the interest rate, generates an electronic version of credit documentation and sends it to the client via the web interface.

Along with the terms of the loan, a list of legal documents that are expected to be received from the customer for the deal to succeed is formed. The function of a legal officer is performed automatically by the Robot Lawyer, which does the same set checks on client documents as a human lawyer in a standard credit process would do.

The client signs the documentation using his electronic certificate. The signature applied has full legal force and can be verified automatically by a certificate authority.

The loan conditions chosen by the client are reflected in the bank's internal accounting system.

The speed of the decision making on a loan application in the solution is significantly improved and represents an important step in the development of corporate lending in Russia for the LMB segment.

7.7.5.3 Stakeholders and stakeholder considerations

Stakeholders: customers

Stakeholder considerations: fair treatment

7.7.5.4 Data characteristics

Millions of customer profiles collected from more than 50 systems inside and outside the bank and included previous history of credit offers.

7.7.5.5 KPIs

Non-performing loans ratio: Ratio of the sum of borrowed money upon which the debtor has not made the scheduled payments for a specified period to the total loans. The reference of this KPI to mentioned use case objectives is to improve efficiency.

Time to decision: minutes for generating appropriate loan offers. The reference of this KPI to mentioned use case objectives is to shorten delivery time.

7.7.5.6 Features of use case

Task(s): natural language processing, decision making, graph

Level of automation: high automation

Method(s):

— NLP: neural networks CNN + bi-LSTM, BERT + attention and few-shot learning (Proto- name entity recognition).

— Decision making for loan approval: neural network (NN), XGBoost, LogReg + L1/L2 regularization.

— Graph: investigation of companies' influence on each other to consider it in decision making.

7.7.5.7 Threats and vulnerabilities

Different sources of bias.

7.7.5.8 Challenges and issues

Nonlinear models based on big data need significant computational power during the training phase.

7.7.5.9 Trustworthiness considerations

Data attributes such as gender, race and address are used to verify fairness and take corrective action as necessary. This action will correct unfair lending decisions for a particular group.

The system establishes an inquiry means from the examination of the loan applicant. Responding to inquiries will improve the fairness of loan screening and the reliability of loan screening companies. In technological innovation and infrastructure is a crucial driver of higher levels of productivity and economic growth.

Because of gender and racial bias in training data, loans are less likely to pass for a particular gender or race.

There is no way for loan users to appeal.

7.7.5.10 Use of standards and standardization opportunities

Standardization needs for setting up this use case are currently under further investigation.

7.7.6 Predictive risk intelligence (use case 164)

7.7.6.1 Objectives

Predictive risk intelligence for a wide range of domains including supply chain management, insurance, finance, sustainability and medicine.

Risk management.

7.7.6.2 Narrative

The core technology used is a predication engine. The prediction engine uses smart information systems (SIS) that enable the company to detect risk events on a global scale, days and sometimes even weeks before they happen. Predictive intelligence is used to improve decision-making for various application domains and markets, ranging from non-governmental organisations (NGOs) to supply chain management, corporate sustainability, and the insurance industry.

The AI system is used to analyse media streams with advanced data analytics technologies. This involves data retrieval from a range of social media sites, which is then analysed by ML algorithms to determine predictive outcomes for clients. The algorithms are trained in a number of languages and can evaluate a wide range of risk factors.

7.7.6.3 Stakeholders and stakeholder considerations

Stakeholders:

- for all: policy makers, social media users, social media providers, designers of risk prediction algorithm;
- supply chain management: business analysts, buyers, suppliers;
- insurance: insurance companies, consumers;
- finance: stock traders, business owners, investors;
- medicine: doctors, clinical analysts, patients.

Stakeholder considerations: overlap with systems threats and vulnerabilities.

7.7.6.4 Data characteristics

The data used in this use case are the media streams which is from a range of social media sites.

7.7.6.5 KPIs

The accuracy of detecting the risk events on a global scale, days and sometimes even weeks before they happen.

7.7.6.6 Features of use case

Level of automation: assistance

7.7.6.7 Threats and vulnerabilities

Security and privacy concerns: use of social media data for predictions raises privacy questions, consent from those whose data were collected.

Integrity: a lack of integrity when designing or using algorithms, when making predictions for specific target clients and therefore formulating algorithms that will service those clients a bias can arise.

Transparency, fairness and accountability: when it comes to transparency and fairness in the automated decision-making process, such as predictive risk intelligence, users or clients only get a limited idea of why a decision has been made in a certain way, which does not mean the decision is justified or legitimate. The AI's methodology is confidential to protect intellectual property

Trust and accuracy: the predictions rely on social media input for a specific representation of the world. In certain regions social media can be used less or experience a lot of fake news.

Bias: the company 'consciously' selects its clients so there is a high risk of bias.

7.7.6.8 Challenges and issues

Overlap with systems threats and vulnerabilities.

7.7.6.9 Trustworthiness considerations

Privacy and data governance.

Use of social media data to make risk analysis and predictions.

7.7.6.10 Use of standards and standardization opportunities

How to deal with conflicting ethical issues and its transparency. Researchers are drafting a proposal for guidelines for ethical AI.

Guidelines on how to engage stakeholders and provide transparency (reporting) to stakeholders on ethical code and decisions. CWA 17145 on ethics assessment for research and innovation can provide input here.

7.8 Healthcare

7.8.1 AI system to predict post-operative visual acuity for LASIK surgeries (use case 24)

7.8.1.1 Objectives

Predicting post-operative visual acuity for laser-assisted in situ keratomileusis (LASIK) surgeries from retrospective laser-assisted in situ keratomileusis (LASIK) surgery data with patient follow-ups.

Given: pre-operative examination results and demographic information about a patient.

Predict: post-operative uncorrected visual acuity (UCVA) one day, one week and one month after surgery.

7.8.1.2 Narrative

LASIK surgeries have been quite popular for treatment of myopia, hyperopia and astigmatism over the past two decades. In the past decade, over 10 million LASIK procedures have been performed in the United States alone with an average cost of approximately $2 000 USD/s. While 99 % of such surgeries are successful, the commonest side effect is a residual refractive error and poor UCVA. This system aims at predicting the UCVA post LASIK surgery. The task of this system is a regression problem and use the patient demography and pre-operative examination details as features. This is the work to systematically explore this critical problem using machine learning methods. Further, LASIK surgery settings are often determined by practitioners using manually designed rules. It is explored that the possibility of determining such settings automatically to optimize for the best post-operative UCVA by including such settings as features in the regression model. The experiments of this system on a data set of 791 surgeries provides an RMSE of 0,102, 0,094 and 0,074 for the predicted post-operative UCVA one day, one week and one month after surgery respectively.

7.8.1.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the AI system;

— data provider: the hospitals that provides the data set;

— AI users: patients undergoing LASIK surgeries.

7.8.1.4 Data characteristics

Description: The data set contains information for 404 patients in the age range of 18 y to 47 y. Of these patients, 215 are female, and the rest are male. The 791 LASIK surgeries were performed in 2013 and 2014.

Of the surgeries, 397 were performed on the left eye and the remaining ones on the right eye. Most of the surgeries are either of the Wavefront guided-LASIK type or of the Plano-scan-LASIK type. Orbscan is the most popular topography machine used; Oculyzer being the second most popular. Pre-operative UCVA values vary between 0,15 and 2. Post-operative UCVA values vary between −0,2 and 1 for d 1, −0,3 and 1 for one week and −0,2 and 0,95 for one month after the operation. Although usually large data sets improve the accuracy of the learned machine learning models, it is difficult to obtain large data sets in this domain.

Source: measured using various medical machines.

Type: structured data

Volume (size): 791 instances from 404 patients.

Velocity: batch

Variety: single source. data from multiple centres of the hospital.

Variability (rate of change): static.

Quality: contains some noise. high quality after pre-processing.

7.8.1.5 KPIs

Recommendation: the system can be used to automatically recommend the right LASIK surgery to the patient. The reference of this KPI to mentioned use case objectives is new use-case in healthcare.

Improve accuracy: the accuracy of the model is reasonably good in terms of being practically useful. The reference of this KPI to mentioned use case objectives is to improve accuracy.

7.8.1.6 Features of use case

Task(s): prediction

Level of automation: partial automation

Method(s): machine learning, gradient boosted decision trees-based regression.

Platform: machine with 1 CPU and 2 GB RAM. Any operating system.

Topology: LASIK surgeries, UCVA, uncorrected visual acuity, regression

7.8.1.7 Threats and vulnerabilities

Different sources of bias; incorrect AI system use

7.8.1.8 Challenges and issues

The challenges of this use case include:

— a large amount of data about such surgeries is not easily available;

— there are a lot of pre-operative measurements that can be used as signals;

— data are sparse, i.e. there are a lot of missing values.

7.8.1.9 Trustworthiness considerations

The trustworthiness considerations of this use case include: the explainability of the predication, the bias of the data set of LASIK surgeries and the privacy of the patients' data.

7.8.1.10 Use of standards and standardization opportunities

The standards regarding the format and exchange of LASIK surgeries data can be used in this system.

7.8.2 AI system to quality control of electronic medical record (EMR) in real time (use case 50)

7.8.2.1 Objectives

Detecting defects in EMRs by inspecting unstructured data based on NLP ability.

To ensure the completeness, consistency, punctuality and medical compliance of EMRs written by physicians.

7.8.2.2 Narrative

This AI system is developed to simultaneously detect mistakes while physicians write the electronic medical record (EMR).

Using NLP ability, it can process a large amount of unstructured text and judge its accuracy according to recognized medical references.

It achieved 80 % coverage of all EMR quality control requirements issued by the Chinese government, and reduced human labour for EMR quality control (QC) by 60 %, which translated into cost savings and enhanced physician education.

7.8.2.3 Stakeholders and stakeholder considerations

Stakeholders: doctor, hospital and patient.

The stakeholders include:

— AI producer: the organization that develops this system;

— AI users: the doctors that use this system to create EMR. The hospitals that contain and use the EMR data;

— data provider: the hospitals and the EMR system vendor;

— AI subjects: the hospitals and the patients relevant to the EMR.

The stakeholder considerations include safety, privacy, fair treatment and trustworthiness.

7.8.2.4 Data characteristics

Description: EMR text data

Source: EMR system

Type: text data from EMR system vendor

Velocity: real time

Variety: multiple data sets

Variability (rate of change): static

Quality: high (depending on EMR system)

7.8.2.5 KPIs

Coverage: ratio of EMR QC requirements covered by the solution or all issued EMR QC requirements in China. Ideal target is 100 %. The reference of this KPI to mentioned use case objectives is to improve accuracy.

7.8.2.6 Features of use case

Task(s): natural language processing

Level of automation: conditional automation

Platform: cloud service

7.8.2.7 Threats and vulnerabilities

The threats include new privacy and new security of using the EMR data to develop the AI system.

7.8.2.8 Challenges and issues

Challenges: Achieve all EMR QC requirements in different disease areas.

Issues: 1) lack of medical reference data; 2) lack of medical knowledge graph.

7.8.2.9 Trustworthiness considerations

The quality of the system is an important trustworthiness consideration of this use case. The system is developed to achieve 80 % coverage of all EMR quality control requirements issued by the Chinese government, and reduced human labour for EMR QC by 60 %. The quality of the system can be translated into cost savings and enhanced physician education, which is a value of this system.

7.8.2.10 Use of standards and standardization opportunities

The standards regarding the format and exchange of EMR data and the standards of the EMR quality control requirements can be used in this system.

7.8.3 Discharge summary classifier (use case 79)

7.8.3.1 Objectives

Classification of discharge summaries

7.8.3.2 Narrative

This system proposes a method for construction of classifiers for discharge summaries. First, morphological analysis is applied to a set of summaries and a term matrix is generated. Second, correspond analysis is applied to the classification labels and the term matrix and generates two dimensional coordinates. By measuring the distance between categories and the assigned points, ranking of keywords is generated. Then, keywords are selected as attributes according to the rank, and training examples for classifiers are generated. Finally learning methods are applied to the training examples. Experimental validation shows that random forest achieved the best performance and the second best was the deep learner with a small difference, but decision tree methods with many keywords performed only a little worse than neural network or deep learning methods.

7.8.3.3 Stakeholders and stakeholder considerations

The stakeholders include:

- AI producer: the organization that develops the AI system;
- data provider: the organisations that provide the training data for developing the AI system;
- AI users: the medical staff that use this system to create discharge summaries;
- AI subjects; the medical staff and the patients.

The stakeholder considerations include the stakeholders' assets and values on the quality of medical care.

7.8.3.4 Data characteristics

Source: hospital information system

Type: text, numerical: time-series

Volume: 1 GB text

Velocity: real time

Variety: text, numerical, (time series)

Variability: every hour

Quality:

— records: dependent on medical staff;

— numerical: automatic.

7.8.3.5 KPIs

Accuracy: classification accuracy. Check of decision summaries.

Length of stay: length of stay in inpatient ward. Management of ward.

7.8.3.6 10. Features of use case

Task(s): knowledge processing and discovery

Level of automation: assistance

Method(s): text mining, decision tree, random forest, SVM, BNN, deep learning

Platform: servers for analytics, data servers

Topology: network of data and analytics servers

7.8.3.7 Threats and vulnerabilities

The threats of this system include the bias in hospital texts.

7.8.3.8 Challenges and issues

The challenges of this system include reduction of the computational complexity, the refinement of medical texts and the enhancement of the medical hospital management.

7.8.3.9 Trustworthiness considerations

The trustworthiness considerations of this use case include: the quality of the AI system and the privacy of the patients' data.

7.8.3.10 Use of standards and standardization opportunities

The standards that are relevant to big data analytics can be used in this use case.

7.8.4 Generation of clinical pathways (use case 80)

7.8.4.1 Objectives

Nursing clinical pathway

7.8.4.2 Narrative

This system proposes a temporal data mining method to construct and maintain a clinical pathway used for schedule management of clinical care. Since the log data of clinical actions and plans are stored in a hospital information system, these histories give temporal and procedural information about treatment. The method consists of the following four steps:

— first, histories of nursing orders are extracted from the hospital information system;

— second, orders are classified into several groups by using clustering and multidimensional scaling methods;

— third, by using the information on groups, feature selection is applied to the data and important features for classification are extracted;

— finally, original temporal data are split into several groups and the first step is repeated. After the grouping results are stable, a new pathway is constructed based on the induced results.

7.8.4.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the AI system;

— data provider: the organisations that provide the training data for developing the AI system;

— AI users: the nursing staff that use this system for the schedule management;

— AI subjects; the nursing staff, other medical staff and the patients.

The stakeholder considerations include the stakeholders' assets and values on the quality of medical care.

7.8.4.4 Data characteristics

Source: hospital information system

Type: text, numerical: time-series

Volume: 1 GB text

Velocity: real time

Variety: text, numerical, image (time series)

Variability: every minute/hour

Quality:

— records: dependent on medical staff;

— numerical/image: automatic.

7.8.4.5 KPIs

Pathway complexity: complexity of nursing orders. This KPI impacts on the management of nursing orders.

Length of stay: length of stay in inpatient ward. This KPI impacts on the management of ward.

7.8.4.6 Features of use case

Task(s): knowledge processing and discovery

Level of automation: assistance

Method(s): decision tree, clustering

Platform: servers for analytics, data servers

Topology: network of data and analytics servers

7.8.4.7 Threats and vulnerabilities

The threats of this system include the bias in hospital data.

7.8.4.8 Challenges and issues

The challenges of this system include reduction of the computational complexity and the enhancement of the hospital management.

7.8.4.9 Trustworthiness considerations

The trustworthiness considerations of this use case include: to achieve the good practice of medical services with considering the privacy of the patients' data.

7.8.4.10 Use of standards and standardization opportunities

The standards that are relevant to big data analytics can be used in this use case.

7.8.5 Hospital management tools (use case 81)

7.8.5.1 Objectives

Hospital management

7.8.5.2 Narrative

Temporal data mining methods (multi-scale comparison with clustering and temporal frequent item sets) are applied to hospital data.

A scheme for innovation of hospital services based on data mining. Then, based on this scheme, data mining techniques are applied to data extracted from hospital information systems. The results included several interesting findings, which suggests that the reuse of stored data can provide a powerful tool to improve the quality of hospital services.

7.8.5.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the AI system and hospital services;

— data provider: the organisations that provide the training data for developing the AI system;

— AI users: the hospital administrators that use this system for the hospital management;

— AI subjects; the medical staff and the patients.

The stakeholder considerations include the stakeholders' assets and values on the visualization of medical staff behaviour in hospital.

7.8.5.4 Data characteristics

Source: hospital information system

Type: text, numerical, images: time-series

Volume: 1 GB text data and 4 TB image data

Velocity: real time

Variety: text, numerical and image (time series)

Variability: every second/hour

Quality

— records: dependent on medical staff;

— numerical and image: automatic.

7.8.5.5 KPIs

Waiting time: waiting time of outpatient clinic. This KPI impacts on management of outpatient clinic.

Length of stay: length of stay in inpatient ward. This KPI impacts on management of ward.

7.8.5.6 Features of use case

Task(s): knowledge processing and discovery

Level of automation: assistance

Method(s): temporal data mining, clustering

Platform: servers for analytics, data servers

Topology: network of data and analytics servers

7.8.5.7 Threats and vulnerabilities

The threats of this system include the bias in hospital data.

7.8.5.8 Challenges and issues

The challenges of this system include reduction of the computational complexity and the enhancement of the hospital management.

7.8.5.9 Trustworthiness considerations

The trustworthiness considerations of this use case include: to achieve the good practice of medical services with considering the privacy of the patients' data.

7.8.5.10 Use of standards and standardization opportunities

The standards that are relevant to big data analytics can be used in this use case.

7.8.6 Predicting relapse of a dialysis patient during treatment (use case 87)

7.8.6.1 Objectives

Build an AI system to augment dialysis nurses.

Use AI to predict if a patient is possible to relapse during dialysis to reduce patient trauma.

7.8.6.2 Narrative

A deep learning model to learn from historical and real-time parameters about a patient to identify the probability that patient can relapse during dialysis.

The private dialysis clinic was relying solely on the discretion of trained nurses to make a call whether or not a patient can get started for a dialysis session or is expected to be taken to a hospital ahead of the treatment due to possible relapse. This created inconsistencies in the patient's experience and 10 % of the patients would relapse and suffer trauma in the middle of their sessions. The deep learning model was able to provide a more consistent call about the likelihood of relapse, upon which the trained nurses can decide proactively for or against starting the dialysis session.

Prediction models can improve global quality of care for patients of kidney diseases or failure, and can allow the services to be more federated and standardized.

7.8.6.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the AI system;

— data provider: the organisations that provide the training data for developing the AI system;

— AI users: the dialysis nurses;

— AI subjects; the dialysis patients and the partner hospitals.

The stakeholder considerations include the stakeholders' assets and values on the percentage of relapses as a total of all sessions, cost of incomplete sessions.

7.8.6.4 Data characteristics

Description: dialysis appointment history data

Source: dialysis company database

Type: structured data with Boolean, numerical and alphanumerical data

Velocity: batch

Variety: single

Variability (rate of change): dynamic, updated weekly

Quality: high

7.8.6.5 KPIs

Prediction accuracy: consistency of prediction compared to actual relapse rates. The reference of this KPI to mentioned use case objectives is that the prediction accuracy is expected to be 90 % or more to ensure only the true-positive relapses are proactively sent to hospitals.

Ease of use: ease of interpreting the inference of the models. The reference of this KPI to mentioned use case objectives is that the output of the model is expected to be easily understandable to the nurses.

Money saved: the loss incurred for incomplete sessions. The reference of this KPI to mentioned use case objectives is that the proactive decisions not to commence high-relapse-chance patients' sessions to reduce the cost of incomplete sessions.

7.8.6.6 Features of use case

Task(s): prediction

Level of automation: assistance

Method(s): deep learning

Platform: clinic computers and laptops

Topology: hybrid

7.8.6.7 Threats and vulnerabilities

If the equipment to identify the on-premise vital stats of the patient is incorrect or inaccurate, these would feed incorrect data into the model and the prediction output would also be inaccurate, leading to misguided decisions.

7.8.6.8 Challenges and issues

Challenges in feature engineering the scores of data sets into a logical format that allows the prediction model to retrain without need for high compute.

7.8.6.9 Trustworthiness considerations

The trustworthiness considerations of this use case include: the lack of reliable and accessible healthcare facilities, and the privacy of the patients' data.

7.8.6.10 Use of standards and standardization opportunities

The standardizations that can be used in this use cases include to standardize and federate the service of kidney diseases or failure.

7.8.7 Instant triaging of wounds (use case 89)

7.8.7.1 Objectives

Build an AI system to augment triaging decisions by wound nurses.

Use AI to identify and classify the intensity of wounds.

7.8.7.2 Narrative

A computer vision model able to use RGB and infrared (IR) wavelengths to measure the size, depth and intensity of a wound.

A wound nurse is the first line of medical attention when a patient comes to the hospital suffering from serious external wound injuries. The problem is more chronic in diabetic patients. The wound nurse is expected to spend time to view the wound and decide how to triage the seriousness of the wound before sending the patient to the doctor. A CV model was built that can use a 2-megapixel mobile camera and off-the-shelf IR camera attachments to visualize wounds within seconds, to help the wound nurse make faster and more consistent triaging decisions.

Using computer vision can make medical attention more globally accessible, in particular for poor and remote areas without compromising on the quality of care.

7.8.7.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the AI system;

— data provider: the organisations that provide the training data for developing the AI system;

— AI users: the wound nurses;

— AI subjects; diabetes patients and hospitals.

The stakeholder considerations include the stakeholders' assets and values on the time and accuracy of triaging wounds.

7.8.7.4 Data characteristics

Description: images of wounds in RGB and IR spectrum

Type: image data

Volume (size): 250 GB

Velocity: batch

Variety: single

Variability (rate of change): static

Quality: high

7.8.7.5 KPIs

Visualization accuracy: the visual representation of the wound is close to the actual condition. The reference of this KPI to mentioned use case objectives is to unburden the nurse from the stress of accurately identifying the severity of wounds.

Ease of use: ease of interpreting the visual models of the wound. The reference of this KPI to mentioned use case objectives is that the visualization of the wound is expected to be easily understandable for the wound nurses.

Time saved: the time taken to view, assess and triage each patient. The CV model would create a visualization of the wound within seconds which can otherwise take a wound nurse 10 min to 30 min.

7.8.7.6 Features of use case

Task(s): knowledge processing and discovery

Level of automation: assistance

Method(s): Computer vision

Platform: mobile phones, hospital computers

Topology: bus

7.8.7.7 Threats and vulnerabilities

Externalities like poor lighting or damages in the phone camera can ingest incorrect data into the CV model and output inaccurate visualizations.

7.8.7.8 Challenges and issues

Challenges in integrating RGB models and IR models into a single, interpretable visualization for the nurses.

7.8.7.9 Trustworthiness considerations

The trustworthiness considerations of this use case include: shortfalls in access to trained nurses and medical imaging technology, and the privacy of the patients' data.

7.8.7.10 Use of standards and standardization opportunities

The standards that are relevant to computer vision can be used on this use case.

7.8.8 Detection of fraudulent medical claims (use case 90)

7.8.8.1 Objectives

Build an ML model to classify if a particular claim can be fraudulent.

More effective fraud detection, moving from human-only detection to computer-assisted detection through machine learning.

7.8.8.2 Narrative

The third-party administration (TPA) company has a very good visualization dashboard to observe trends by patient, by physician and by condition of medical claims submitted to the insurance companies it serves.

However, the identification of anomalies in the visual representation was still based on subjective judgment.

A machine learning model was developed to determine whether a particular claim is fraudulent and to identify anomalies in claims due to fraudulent activity by the patient, the doctor or both in collusion.

7.8.8.3 Stakeholders and stakeholder considerations

Stakeholders: TPA, medical insurance companies, doctors and patients.

Stakeholder considerations: percentage of true-positive, false positive and false negative fraudulent claims detected out of the total set of claims.

7.8.8.4 Data characteristics

Description: medical claim data

Source: medical insurance companies

7.8.8.5 KPIs

Inference accuracy: Number of true-positives detected vs false-positives. The reference of this KPI to mentioned use case objectives is that the better the accuracy of the model is, the more surgical the TPA's intervention is in identifying and controlling fraud.

Time of inference: The latency in the model to retrain and generate new inferences. The reference of this KPI to mentioned use case objectives is that the latency in the model is expected to be reasonable to allow the TPA to take faster action against fraudulent activities.

Insurance company client's satisfaction: The reduction in the number of fraudulent claims that the insurance company client is expected to disburse money to. The reference of this KPI to mentioned use case objectives is that the loss to the clients of the TPA (i.e. insurance companies) would reduce if more fraudulent claims are detected.

7.8.8.6 Features of use case

Task(s): inference

Level of automation: partial automation

Method(s): machine learning

Platform: TPA's own devices and servers

Topology: ring and hybrid

7.8.8.7 Threats and vulnerabilities

If the features of the model are not updated every few years, the model is possibly not be able to detect modes of fraud that have never been seen before.

Identified privacy risks include patient privacy risks (e.g. re-identification of patient data).

7.8.8.8 Challenges and issues

A major challenge was to develop separate models for each major source of fraudulent claims.

Adequately addressing identified privacy risks.

7.8.8.9 Trustworthiness considerations

Inadvertent or illegal use of funds intended for essential services to people.

This use case raises serious privacy concerns, particularly with respect to patient data.

7.8.8.10 Use of standards and standardization opportunities

Machine learning models capable of detecting fraud can be used worldwide to protect the integrity of public or private funds intended for essential services such as medical care, housing, education or sanitation.

7.8.9 AI platform for chest CT-scan analysis (early stage lung cancer detection) (use case 105)

7.8.9.1 Objectives

Detecting malignant neoplasms (lungs) on chest CT-scans.

To facilitate early stage oncology chest CT-scans through the application of the platform based on artificial intelligence.

7.8.9.2 Narrative

This AI system is a software platform for the diagnosis and assessment of pathology risks using artificial intelligence technologies. The product supports radiologists and oncologists, facilitating the analysis and recognition of diagnostic images of CT-scans, digital X-rays and mammography. The project aims to reduce costs and improve diagnostic accuracy, while detecting pathologies at early stages.

This AI system implements its own patented technology to create a digital model of the patient. This allows for state-of-the-art results derived from the company's algorithms, confirmed by scientific publications. This AI system platform core goals are improved oncology detection at early stages and prioritization of patient flow. The company provides its own developed digital imaging and communications in medicine (DICOM) viewer. The platform can be integrated into any type of picture archiving and communication systems (PACS)/central archive of medical images such as SaaS solutions, or as part of a medical institution's closed infrastructure. The company is ready to provide customizable integration options to fit the needs of varying customers.

"Hybrid intelligence" technology allows for the combination of the AI platform's sensitivity with the specificity of a skilled radiologist. This AI system can also be used to manage the flow of different radiological studies.

This AI system is designed to increase the efficiency and effectiveness of radiological analysis. This product addresses two main medical issues: an undersupply of radiologists in the workforce and missed malignant neoplasms on chest CT-scans. With the introduction of this technology, thousands of lives can be saved via improved early-stage oncology.

7.8.9.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the AI system;

— data provider: the organisations that provide the training data for developing the AI system;

— AI users: the healthcare authorities;

— AI subjects; the patients and hospitals.

The stakeholder considerations include the stakeholders' assets and values on the reputation, saved lives and cost savings.

7.8.9.4 Data characteristics

The data used in this use case are the diagnostic images of CT-scans, digital X-rays and mammography.

7.8.9.5 KPIs

Accuracy: 93 % detection rate of malignant neoplasm on chest CT-scans (AUC). The reference of this KPI to mentioned use case objectives is to improve accuracy.

Speed: from 4 min to 6 min (depending on Internet speed). The reference of this KPI to mentioned use case objectives is to improve speed.

7.8.9.6 Features of use case

Task(s): recognition

Level of automation: conditional automation

Method(s): deep learning

7.8.9.7 Threats and vulnerabilities

Loss of trust of this AI system.

7.8.9.8 Challenges and issues

Achieving a higher confirmed level than accredited radiologists in the detection of lung cancer.

7.8.9.9 Trustworthiness considerations

The trustworthiness considerations of this use case include: to consider the privacy of the patients' data.

7.8.9.10 Use of standards and standardization opportunities

The standards that are relevant to big data analytics can be used in this use case.

7.8.10 Neural network formation of 3D-models orthopaedic insoles (use case 121)

7.8.10.1 Objectives

Artificial intelligence methods are used to construct individual medical products to reduce the risk of developing diseases of the musculoskeletal system.

Development of comfortable, individualized, anatomically correct orthopaedic 3D insoles for the treatment of flat feet.

7.8.10.2 Narrative

Using artificial intelligence methods, the system converts a pre-scanned footprint into an innovative, medically based 3D-insole. The AI-system would independently make a medical decision based on the collected medical history and anthropometric data.

Initial training of the AI-system would take place together with the doctor. In the future, the system would begin by independently choosing the most suitable location options for a patient's vaults and indentations and plan an anatomically correct and secure 3D-insole.

The system consists of two parts, hardware and software.

The hardware scans 3D/2D foot images of patients and receives a production file format ready for loading into a specialized machine or a 3D printer.

In the software, a local orthopaedic 3D model of the insole is formed according to a unique author’s technique using a local software package based on artificial intelligence.

The received data are stored on a cloud platform.

The 3D-method makes it possible to more accurately orthose complex pathologies and atypical deformations due to the sophisticated equipment used and accurate removal of anatomical physiological parameters of the foot up to 10 000 p/sm^2. The patient’s foot is scanned in the sitting position; it is not exposed to loads; the 3D laser scanning is 6 cm high, which allows for obtaining full-colour 3D models of the patient’s legs with an accuracy of half a millimetre. Further automatic milling is highly accurate for orthopaedic shoes. The process of creating insoles is completely autonomous and personalized and does not require the intervention of an orthopaedic doctor.

Overall, the system is modularized with capabilities to self-learn and for future extensions.

7.8.10.3 Stakeholders and stakeholder considerations

Stakeholders: medicine, public sector.

Stakeholder considerations: Improving the quality of life.

The stakeholders include:

— AI producer: the organization that develops the AI system;

— data provider: the organisations that provide the training data for developing the AI system;

— AI users: the medicine related organizations;

— AI subjects; the patients and hospitals.

The stakeholder considerations include the stakeholders’ assets and values to improve the quality of life of the patients.

7.8.10.4 Data characteristics

The data used in this use case include medical history, anthropometric data and the 3D/2D foot images of patients and receives a production file format ready for loading into a specialized machine or a 3D printer.

7.8.10.5 KPIs

Individualized, anatomically correct orthopaedic 3D insoles: Local orthopaedic 3D model of the insole is formed. The reference of this KPI to mentioned use case objectives is to reduce the risk of developing diseases of the musculoskeletal system.

7.8.10.6 Features of use case

Task(s): construction

Level of automation: high automation

Method(s): neural networks

Platform: 3D printer, scanner, cloud platform

7.8.10.7 Threats and vulnerabilities

Incorrect AI system use.

7.8.10.8 Challenges and issues

None identified.

7.8.10.9 Trustworthiness considerations

The trustworthiness considerations of this use case include: to consider the privacy of the patients' data.

7.8.10.10 Use of standards and standardization opportunities

Tolerance criteria for predicted product characteristics

7.8.11 Search of undiagnosed patients (use case 127)

7.8.11.1 Objectives

Search for undiagnosed patients with orphan diseases, define patients' journey.

Deep semantic analysis of unstructured texts (based on meaning rather than keywords, i.e. using natural language processing technology).

7.8.11.2 Narrative

Knowledge extraction from the massif of user posts in patient forums, physicians' professional networks, health-related portals, etc.

Full-scale crawling of search engines.

Semantic and statistical analysis of found posts related to description of particular symptoms, description of clinical analyses, diagnostic procedures, etc.

Identification of insights and presentation of results.

Semantic artificial intelligence (AI) tools that can read and interpret electronic free text at scale. Real patient journey, patient subgroups, etc. are to be evaluated.

A unified medical and social image of the user (patient) can be created.

7.8.11.3 Stakeholders and stakeholder considerations

Stakeholders: patients, government affairs, physicians and pharma companies.

Stakeholder considerations: personal data of the subjects to be identified, especially patients', i.e. special health information can potentially be at risk.

7.8.11.4 Data characteristics

Source: user posts in patient forums, physicians' professional networks, health-related portals, etc.

Variety: real time

Variability (rate of change): multiple

7.8.11.5 KPIs

Patient journey: real patient journey is to be clarified based on obtained data. Disease guidelines are to be changed accordingly.

Effectiveness: % of all identified patients is expected to be close to the number of patients predicted by prevalence data.

7.8.11.6 Features of use case

Task(s): natural language processing

Level of automation: conditional automation

7.8.11.7 Threats and vulnerabilities

Difficulties with ordering and finding patients.

7.8.11.8 Challenges and issues

Personal data of the subjects to be identified, especially patients', i.e. special health information can potentially be at risk.

7.8.11.9 Trustworthiness considerations

The reliability of the AI system can deteriorate with fluctuations of descriptions of symptoms, clinical analyses, diagnostic procedures, etc.

7.8.11.10 Use of standards and standardization opportunities

None identified.

7.8.12 A clinical decision support system (use case 131)

7.8.12.1 Objectives

Screening for cardiovascular disease risk prediction with machine and deep learning methods.

Advances in precision medicine would require an increasingly individualized prognostic evaluation of patients in order to provide the patient with appropriate therapy.

7.8.12.2 Narrative

Cardiovascular disease (CVD) continues to be the most prelevant health problem of most countries in the world, including the Russian Federation. According to the World Health Organization, more than 17 million people die each year from CVD worldwide, including more than seven million from coronary heart disease (CHD).

The machine learning models outperformed traditional approaches for CVD risk prediction (such as strategies concentrating on risk evaluation (SCORE), prospective cardiovascular Munster (PROCAM), and Framingham equations). This approach was used to create a clinical decision support system (CDSS). It uses both traditional risk scales and models based on neural networks. Of notable importance is the fact that the system can calculate the risk of cardiovascular disease automatically and recalculate immediately after adding new information to the EHR. The results are delivered to the user's personal account.

The CDSS is a ready-made, trained solution to identify high-risk patients and prevent morbidity and mortality. The characteristics of this system include:

— automatic risk stratification of patients;

— a more efficient organization of preventive work aimed at a personal group of patients with a high risk of complications and death;

— the ability to route patients depending on the assessment obtained;

— reduced morbidity and mortality;

— reliable digital assistance, trained on the results of evidence-based medicine and modern clinical guidelines;

— automatic identification of risk factors;

— automatic determination of the likelihood of developing a disease;

— compliance with clinical practice guidelines;

— reduced time of the patient risk assessment;

— powerful artificial intelligence to evaluate medical data and identify risk factors without development costs;

— the addition of medical decision support functions;

— ready service for evaluating EHR and to identify the risk factors;

— reducing the costs of development of the medical information system.

7.8.12.3 Stakeholders and stakeholder considerations

Stakeholders: end-users (physician, nurse, laboratory technologist, pharmacist, patient), sales and marketing team, CDSS product development and maintenance team (system administrator, system developer, system architect, project manager and system maintenance).

Stakeholder considerations: Competitiveness, cost savings

7.8.12.4 Data characteristics

Description: FHS-Cohort

Type: structured/unstructured text: time-series

Volume (size): 84 MB

Velocity: variety

Variability (rate of change): never

Quality: presence of missing fields or incorrect values

7.8.12.5 KPIs

Area under curve receiver operating characteristic (AUC ROC): AUC provides an aggregate measure of performance across all possible classification thresholds. The reference of this KPI to mentioned use case objectives is to determine the quality and correctness of classification models.

Confusion matrix (TP, FP, TN, FN): metrics that can be used to measure the performance of a classifier or predictor. The reference of this KPI to mentioned use case objectives is that some of these people have the disease, and our test correctly says they are positive. They are called true positives (TP). Some have the disease, but the test incorrectly claims they do not. They are called false negatives (FN). Some do not have the disease, and the test says they do not – true negatives (TN). Finally, there can be healthy people who have a positive test result – false positives (FP). These can be arranged into a 2 × 2 contingency table (confusion matrix), conventionally with the test result on the vertical axis and the actual condition on the horizontal axis.

Metrics (accuracy, precision and recall): evaluation metrics for machine learning. The reference of this KPI to mentioned use case objectives is to evaluate the performance of a model in ML.

7.8.12.6 Features of use case

Task(s): natural language processing

Level of automation: partial automation

Method(s): SpaCy, natural language toolkit (NLTK), StanfordNLP, Tensorflow, Keras

Platform: CPU, tensor processing unit (TPU)

Topology: collaboratory Google, web-services

7.8.12.7 Threats and vulnerabilities

Injuries and error: the most obvious risk is that AI systems would sometimes be wrong, and that patient injury or other health-care problems can result.

Data availability: training AI systems requires large amounts of data from sources such as electronic health records, pharmacy records, insurance claims records, or consumer-generated information like fitness trackers or purchasing history. But health data are often problematic. Data are typically fragmented across many different systems.

Privacy concerns: another set of risks arise around privacy. The requirement of large data sets creates incentives for developers to collect data from many patients. Some patients can be concerned that this collection violates their privacy, and lawsuits have been filed based on data-sharing between large health systems and AI developers.

Bias and inequality: there are risks involving bias and inequality in health-care AI. AI systems learn from the data on which they are trained, and they can incorporate biases from those data. For instance, if the data available for AI are principally gathered in academic medical centres, the resulting AI systems would know less about—and therefore would treat less effectively—patients from populations that do not typically frequent academic medical centres.

7.8.12.8 Challenges and issues

To provide physician tools to easily calculate cardiovascular risk anywhere in the world.

7.8.12.9 Trustworthiness considerations

One of the major concerns about AI-assisted CDSS is how the machines reach decisions, and whose decision is expected to prevail when there is disagreement between the CDSS and the medical professional. This lack of transparency is referred to as the 'black box' of AI. In addition to the lack of transparency, the necessary use of large training data sets coupled with mathematical and statistical algorithms and sometimes neural networks, whether with or without full understanding of the internal workings, presents a challenge in educating doctors to use these tools in a clinically relevant way.

7.8.12.10 Use of standards and standardization opportunities

None identified.

7.8.13 Symptom assessment (hypothetical) (use case 134)

7.8.13.1 Objectives

General humans (patients), health workers and health professionals.

Support patients and doctors with (pre)-diagnosing health problems and providing advice on next steps.

7.8.13.2 Narrative

As people are living longer and as the population is growing, there is a global shortage of caregivers. This shortage varies by region. One way to more efficiently use caregivers is to have patients use symptom assessment software that can provide them with a first assessment that informs their next steps.

The World Health Organization estimates the shortage of global health workers to increase from 7,2 million in 2013 to 12,9 million by 2035. This shortage is driven by several factors including growing population, increasing life expectancy and higher health demands. The 2017 Global Monitoring Report by the WHO and the World Bank reported that half of the world's population lacks access to basic essential health services. The growing shortage of health workers is likely to further limit access to proper health care, reduce doctor time, and worsen patient journeys to a correct diagnosis and proper treatment.

While the problem in low- and middle-income countries (LMIC) is worse, in more developed countries health systems face challenges such as increased demand due to increased life expectancy. Additionally, available doctors have to spend considerable amounts of time on patients that do not always need to see a doctor. Up to 90 % of people who seek help from primary care have only minor ailments and injuries. The vast majority (>75 %) attend primary care because they lack an understanding of the risks they face or the knowledge to care for themselves. In the United Kingdom alone, there are 340 million consultations at the general practitioner (GP) every year and the current system is being pushed to do more with less resources.

The gold standard for correct differential diagnosis, next step advice and adequate treatment is the evaluation of a medical doctor who is an expert in the respective medical field, which is based on many years of university education and structured training in hospitals. Depending on context, steps such as triage preceding diagnosis are responsibilities of other health workers. Decision making is often supported by clinical guidelines and protocols or by consulting literature, the internet or other experts.

In recent years, one promising approach to meet the challenging shortage of doctors has been the introduction of AI-based symptom assessment applications that have become widely available. This new class of system provides both consumers and doctors with actionable advice based on symptom constellations, findings and additional contextual information like age, sex and other risk factors. Starting from some general background information about the patient, these systems allow to enter the most relevant presenting complaints. After this step systems proactively collect further relevant evidence - usually in the form of a dialogue inspired by the patient-doctor conversation. In a final step these systems provide some general pre-clinical triage (e.g. to see a doctor on the same day or to try self-care). Most systems also provide the list for most reasonable underlying diseases.

7.8.13.3 Stakeholders and stakeholder considerations

Stakeholders: patients, health workers, health professionals (doctors), clinics or hospitals, health systems, large companies, governments, health related NGOs, WHO, companies developing symptom assessment systems.

Stakeholder considerations: personal or patient health, cost for diagnosis, cost for treatment, time to diagnosis, time to treatment and recovery time.

7.8.13.4 Data characteristics

The data used in this use case include the general background information about the patient such as age, sex and other risk factors as well as the most relevant presenting complaints.

7.8.13.5 KPIs

The accuracy of predication of pre-clinical triage.

The accuracy of the list for most reasonable underlying diseases.

7.8.13.6 Features of use case

Task(s): classification, interactive evidence gathering (chatbots)

Level of automation: partial automation

Method(s) diverse; mostly data or knowledge driven probabilistic, heuristic or deductive expert systems; partially with conversational ML based NLP technology

Platform: diverse; generic cloud platform hardware

Topology: diverse

7.8.13.7 Threats and vulnerabilities

Incorrect (pre)-diagnosis, incorrect triage, insufficient robustness against missing, incorrect or contradicting evidence, insufficient consideration of patient context (e.g. age, gender, region, season, ethnicity)

7.8.13.8 Challenges and issues

None identified.

7.8.13.9 Trustworthiness considerations

Since the symptom assessment system is not 100 % reliable, it escalates some cases to human doctors based on the pre-clinical triage.

7.8.13.10 Use of standards and standardization opportunities

Standardized quality benchmarking.

Standardization of input space (symptoms, findings, etc.) and output space (conditions, ICD10, pre-clinical triage levels etc.).

7.8.14 Making using evidence-based medicine and AI (use case 167)

7.8.14.1 Objectives

Decision support system for doctors, based on knowledge generated by experts.

7.8.14.2 Narrative

Clinical decision support system based on the intellectual analysis and processing of medical texts - medical care standards, clinical recommendations and guidelines, EHR, drug use guide, evidence-based medicine recommendations to ensure that medical decisions can be made in pharmacotherapy.

Thematic information relevant to the user's request:

— national clinical guidelines;

— standards for primary health care;

— standards for specialized health care;

— clinical guidelines and EBM guidelines;

— summary of evidence with reference to sources;

— criteria for assessing the quality of care.

System services:

— machine learning to find similar ones;

— multimodal mathematical models for search;

— intelligent search tools for semantic search;

— processing requests in a natural language;

— voice search engine;

— interfaces for the system from mobile devices and gadgets;

— cross-language support;

— the system works with international ontologies and thesauruses.

7.8.14.3 Stakeholders and stakeholder considerations

Stakeholders: healthcare facilities and healthcare educational.

Stakeholder considerations: using care reduces: medication prescription errors, adverse drug events and other medical errors.

7.8.14.4 Data characteristics

The data used in this use case are the medical experts' knowledge generated from the medical texts (medical care standards, clinical recommendations and guidelines, EHR, drug use guide, evidence-based medicine recommendations).

7.8.14.5 KPIs

AI algorithm accuracy: 94 %

Medical error reduction: 30 %

7.8.14.6 Features of use case

Task(s): semantic search

Level of automation: assistance

Method(s): machine learning, NLP, topic modelling

7.8.14.7 Threats and vulnerabilities

None identified.

7.8.14.8 Challenges and issues

None identified.

7.8.14.9 Trustworthiness considerations

Produced a correct diagnosis in > 90 % of cases, compared to the clinicians' success rate of 80 %.

Prevent medical error.

7.8.14.10 Use of standards and standardization opportunities

None identified.

7.8.15 AI-service for blood cells and bone marrow scans analysis (use case 168)

7.8.15.1 Objectives

To automate the routine process of microscopic blood and bone marrow scans analysis using Third Opinion AI service based on artificial intelligence.

7.8.15.2 Narrative

This AI system aims to automate the routine process of microscopic blood and bone marrow scans analysis and assist doctor on signs of abnormalities detection.

Analysis of images with the tagging next 34 blood and bone marrow cells types:

- agranular myeloblast;
- band neutrophil;
- bare lymphocyte nuclei;
- basophilic normoblast (erythrokaryocyte);
- hypergranular atypical promyelocyte;
- hypogranular atypical promyelocyte;
- large lymphocyte;
- lymphoblast, l1 - l2 subgroups;
- lymphoblast, l3 subgroups;
- lymphocyte, variant form;
- mature basophil;
- megakaryoblast;
- metamyelocyte;
- mitosis;
- monoblast;
- monoblast with granularity and auer rod;
- monocyte;
- myelo and monoblast;

— myeloblast with auer rod;

— myeloblast with azurophilic granules;

— myeloblast with fuzzy-granules;

— myelocyte;

— neuroblastoma;

— no lineage spec blast (unidentified blast cell);

— oxyphilic erythrokaryocyte;

— plasma cell;

— polychromatic normoblast (erythrokaryocyte);

— promonocyte;

— promyelocyte;

— segmental eosinophil;

— segmental neutrophil;

— small lymphocyte;

— smudge cells;

— undifferentiated anaplastic blast cells.

Report with detected and quantified cell types.

This AI system is based on a supervised learning convolutional neural network algorithm.

Application areas:

— processing the flow of images in diagnostic centres and laboratories;

— dynamic studies of one patient: new picture is compared with previously saved in in EHR or laboratory information system (LIS) under the same ID of a particular patient;

— processing the images flow in specialized medical institutions (oncology, tuberculosis, hematology clinics) and general institutions (city and regional hospitals).

The use of this AI system:

— speeds up the decision-making process both for urgent and for routine cases;

— allows to route patients within the medical institution effectively;

— contributes to the transparency and cost of treatment control.

The service can be integrated with EHR / LIS.

This AI system obtains the patent of the Russian Federation with the title “method for isolation and classification of blood cell types using deep convolutional neural networks”.

7.8.15.3 Stakeholders and stakeholder considerations

Stakeholders: morphologists, digital microscopes developers, EHR or LIS developers, healthcare facilities and the ministry of health.

7.8.15.4 Data characteristics

Origin of data processed by the AI system are medical instruments.

Digitized bone marrow smear from users (pathologists).

Variety: images

Quality: accuracy: 92 to 99 % accuracy rate of blood and bone marrow cells recognition (by type)

Protected attributes: all personal data

7.8.15.5 KPIs

Accuracy: 92 to 99 % accuracy rate of blood and bone marrow cells recognition (by type).

Report with cells labelling and interpretation.

7.8.15.6 Features of use case

Task(s): assistant, studies routing, etc.

Level of automation: assistance

Method(s): computer vision

7.8.15.7 Threats and vulnerabilities

Indifference of the doctor to the process of analysis of the study.

7.8.15.8 Challenges and issues

Lack of digital diagnostic equipment in laboratories, lack of digital transformation skills in the domain.

7.8.15.9 Trustworthiness considerations

Ethical concerns are mitigated by limiting the statement of the conclusion by artificial intelligence. Thus, the infliction of harm by AI is prevented.

7.8.15.10 Use of standards and standardization opportunities

To use artificial intelligence, it is necessary to rework the standard ISO 15189:2022[30], Medical laboratories.

7.8.16 AI-service for chest X-ray and chest CT (use case 169)

7.8.16.1 Objectives

To automate the routine process of chest X-ray and chest CT analysis using this service based on CV technologies.

The service can detect signs of pulmonary pathologies (including pneumonia COVID-19) and lung cancer.

7.8.16.2 Narrative

This AI system can be used in 3+ scenarios to support radiology departments and transform the workflow.

7.8.16.3 Stakeholders and stakeholder considerations

Stakeholders: radiologists, healthcare software development companies [e.g. picture archiving and communication system (PACS), electronic health record (EHR)], healthcare facilities and the ministry of health.

7.8.16.4 Data characteristics

Origin of data processed by the AI system are medical instruments.

customers or partners.

Variety: images

Velocity: 10 s - speed of study analysis

Quality:

— area under the ROC curve: 92 %;

— accuracy: 87 %;

— sensitivity: 94 %;

— specificity: 86 %.

Protected attributes: all personal data.

7.8.16.5 KPIs

speed: 10 sec

accuracy: 87 %

sensitivity: 94 %

specificity: 86 %

7.8.16.6 Features of use case

Task(s): recognition for binary triage, radiologist's assistant, studies routing, etc.

Level of automation: assistance

Method(s): computer vision

7.8.16.7 Threats and vulnerabilities

Indifference of the doctor to the process of analysis of the study - the doctor is asked to review the AI-report.

7.8.16.8 Challenges and issues

Scaling challenges with on-premise installation model.

7.8.16.9 Trustworthiness considerations

Ethical concerns are mitigated by limiting the statement of the conclusion by artificial intelligence. Thus, the infliction of harm by AI is prevented.

But normal scans can be identified with AI and that AI-system can be approved for an autonomous AI-medical imaging diagnostic system

7.8.16.10 Use of standards and standardization opportunities

ISO 12052:2017 is considered in this use case.

7.8.17 Intelligent video analytics system (use case 170)

7.8.17.1 Objectives

This AI system is an intelligent monitoring system for increased patient safety. Using computer vision technology, while securely protecting patient information, the system monitors patient's safety and the quality of care provided by medical staff.

7.8.17.2 Narrative

This AI system provides the solution on:

— the incidence of pressure ulcers development: the system calculates the time spent by the patient in each pose. If the bedridden patients are left in one position too long, the pressure cuts off normal blood supply increasing the bedsores development risk. If it exceeds the time limit set by the doctor or nurse, the system sends the notification;

— patient falls: the system detects the patient's position, fall risk and emergency signals;

— safety: the system can authorize the personnel and visitors and control the care quality;

— the lack of regular monitoring: the system monitors the patient's state of health using information from connected wearable devices.

Using this AI system aims to reduce:

— patient neglect due to nursing shortage;

— falls treatment costs;

— bedsores treatment costs.

7.8.17.3 Stakeholders and stakeholder considerations

Stakeholders: nursing facilities, healthcare facilities and the ministry of health.

7.8.17.4 Data characteristics

Origin of data processed by the AI system, e.g. customers, instruments, IoT, web, surveys, commercial activity, simulations or other sources.

Video from patient's room

Variety: video

Quality: accuracy: 95 %

Protected attributes: all personal data

7.8.17.5 KPIs

Accuracy: 99 %

Alarms with immediate response

7.8.17.6 Features of use case

Task(s): recognition for staff overload, meeting patient expectations etc.

Level of automation: assistance

Method(s): computer vision

7.8.17.7 Threats and vulnerabilities

Saving privacy is a threat in this use case.

7.8.17.8 Challenges and issues

Additional capital expenditure (CAPEX) for healthcare facilities.

7.8.17.9 Trustworthiness considerations

Securely protecting patient information.

7.8.17.10 Use of standards and standardization opportunities

The use of AI in video surveillance systems can be considered to propose a standard.

7.8.18 Retrospective analysis (use case 172)

7.8.18.1 Objectives

Service for the analysis of lung CT archives collected during the COVID-19 pandemic, in the case of nodules.

7.8.18.2 Narrative

AI algorithms given in this system carry out automatic segmentation of developed medical devices for various nosologies.

Rapidly obtaining a preliminary diagnosis using artificial intelligence.

7.8.18.3 Stakeholders and stakeholder considerations

Stakeholders: medical information and analytical centres, outpatient covid centres, and cancer centres.

7.8.18.4 Data characteristics

Series of CT images.

7.8.18.5 KPIs

ROC/AUC > 0.91.

Analysis speed < 5 min/series.

7.8.18.6 Features of use case

Task(s): prediction of nodules in lungs scans

Level of automation: high automation

Method(s): ResNet based AI

Platform: Docker

Topology: clustered

7.8.18.7 Threats and vulnerabilities

Medical data security, standard CT research protocols required and possibility of non-target rare diseases finding.

7.8.18.8 Challenges and issues

As a result of incorrect adjustment of CT devices, deviations in accuracy are possible.

As a result of the lack of integration of multisystem inflammatory syndrome (MIS) and CT devices, the value of the product decreases.

A medical specialist is required for the correct interpretation of the results.

7.8.18.9 Trustworthiness considerations

Implemented data transfer using a secure network.

Implemented data depersonalization before transmission (and depersonalization).

Ease of use lowers the entry threshold for physicians.

7.8.18.10 Use of standards and standardization opportunities

DICOM protocol.

ГОСТ Р МЭК/ТО 62266-2009 for AI.

ГОСТ 28147-89 and ГОСТ Р 34.11-94 for gateway.

7.8.19 Robotization of the federal hotlines on COVID-19 issues (use case 174)

7.8.19.1 Objectives

To increase the availability of information on COVID-19 issues for the population of the Russian Federation. To process 100 % of citizens' appeals without a queue on the line. The ability to provide information 24/7.

7.8.19.2 Narrative

The solution is based on the robot-operator for the federal hotline. It automatically answers subscribers' questions about symptoms, the course of the disease, how to protect yourself and your family without involving contact centre operators. The robot provides statistics on the number of cases and information about the method and procedure of vaccination, PCR tests. The knowledge base contains about 1,500 questions and answers on COVID-19. The customer is provided with a convenient and intuitive format of editing the knowledge base for self-update of the robot-operator.

7.8.19.3 Stakeholders and stakeholder considerations

AI provider: IT company.

AI producer: IT company.

AI Customer: Ministry of Digital Development, Communications and Mass Media of the Russian Federation and Ministry of Health.

AI subject: Population of the Russian Federation.

7.8.19.4 Data characteristics

Source: incoming telephone traffic on the federal hotlines.

Variety: audio, text.

7.8.19.5 KPIs

Call waiting time less than 10 sec.

The delay between the question and the answer of the robot is less than 2 seconds.

Processing 100 % of calls, incl. at peak loads.

7.8.19.6 Features of use case

Task(s): speech recognition, search for question-answer pairs.

Level of automation: full automation

Method(s): implementation of the knowledge base functionality that allows to quickly update with new robot data and search for an answer to the question in real time.

The knowledge base with about 1 500 questions and answers on COVID-19 supports the following algorithms:

— Damerau-Levenshtein distance: A measure of the difference between two strings of characters, defined as the minimum number of operations of insertion, deletion, replacement and transposition (permutation of two adjacent characters) required to transfer one string to another. It is a modification of the Levenshtein distance: the operation of transposition (rearrangement) of symbols has been added to the operations of inserting, deleting and replacing symbols defined in the Levenshtein distance.

— BLEU (bilingual evaluation understudy): An algorithm for calculating points for individual segments of speech - in this case, sentences - by comparing them with a set of high-quality reference copies. BLEU is always a number between 0 and 1. This value indicates how similar the candidate text is to reference instances, with values closer to 1 representing more similar texts.

— The k-nearest neighbours' algorithm (k-NN) is a metric algorithm for automatic classification of objects or regression. In the case of using the method for classification, the object is assigned to the class that is most common among the k neighbours of the given element, the classes of which are already known. In the case of using the method for regression, the object is assigned an average value over the k objects closest to it, the values of which are already known. The algorithm can be applied to samples with a large number of attributes (multidimensional).

— BERT is a language model based on the transformer architecture, designed to pre-train language representations for their subsequent application in a wide range of natural language processing tasks. BERT is a neural network based on a composition of transformer encoders. BERT is an autoencoder. Each layer of the encoder applies two-way attention, which allows the model to take into account the context on both sides of the token in question, and therefore more accurately determine the values of the tokens. When a text is submitted to a network input, it is tokenized first. Tokens are words available in the dictionary or their constituent parts - if a word is absent in the dictionary, it is split into parts that are present in the dictionary. In the neural network itself, tokens are encoded by their vector representations (embeddings), namely, the representations of the token itself (pre-trained), the number of its proposal, and also the position of the token within its proposal are connected. Input data comes to the input and is processed by the network in parallel, and not sequentially, but information about the mutual arrangement of words in the original sentence is saved, being included in the positional part of the embedding of the corresponding token. The output layer of the main network has the following form: the field responsible for the answer in the task of predicting the next sentence, as well as tokens in an amount equal to the input. The reverse transformation of tokens into the probabilistic distribution of words is carried out by a fully connected layer with the number of neurons equal to the number of tokens in the original dictionary.

Platform: common platform

7.8.19.7 Threats and vulnerabilities

Regular AI training through the involvement of journalists and editors who study open sources in order to recognize the reality of the current morbidity situation.

7.8.19.8 Challenges and issues

Low-quality data (background noise, poor phone connections)

7.8.19.9 Trustworthiness considerations

Fault tolerant solution.

Service level agreement (SLA): 99,9

Accommodating new knowledge about COVID-19 is a challenge.

7.8.19.10 Use of standards and standardization opportunities

Customer contact centre complies with GOST R 55540-2013 standards.

The service level indicator is 100 %.

7.8.20 Use of computer vision innovative technologies for analysis of medical images and further application (use case 175)

7.8.20.1 Objectives

To improve the quality of radiology reports by the implementation of AI services.

To reduce the time from a study's completion to the provision of the radiology report by a radiologist.

7.8.20.2 Narrative

Testing and implementation of AI-based services in the unified radiological information service of the united medical information and analytical system (URIS UMIAS) for the use in the city healthcare system in order to improve a quality and efficacy of the diagnostic process carried out by radiologists.

7.8.20.3 Stakeholders and stakeholder considerations

Stakeholder: the city medical facilities, department of information technologies of the city government, AI developers and the patients.

7.8.20.4 Data characteristics

Data source: the unified radiological information service, which collects radiological studies performed in city medical facilities.

Analysed data: DICOM files obtained from diagnostic devices of radiology departments:

— CT scanners – 176 units;

— X-ray and fluorography machines – 730 units;

— mammography units – 123.

Use cases:

— automatic routing of studies (according to specified criteria – procedure code, patient age, type of a medical facility);

— manual sending a specific study for processing by the selected AI service at the radiologist’s request (a marketplace with all available AI services is implemented in the interface of the radiology workstation).

Number of processed studies per day ~ 16,000.

Performance monitoring and quality control:

— automated weekly monitoring of technological defects in the AI performance;

— comprehensive monthly monitoring of technological and clinical defects in the AI performance.

7.8.20.5 KPIs

Experiment directions: 13.

Number of directions in the stream processing: 8.

Number of companies participating in the experiment: 21.

The number of AI services participating in the experiment is 45, of which:

— 22 AI services are at the testing stage in URIS UMIAS;

— 23 AI services are in the stream study processing in URIS UMIAS.

Total number of participating city medical facilities: 103.

Total number of diagnostic devices, from which studies were analysed by AI services: 1 029.

Total number of analysed studies by AI services for 2020 and 2021: 4 000 000.

Satisfaction level with AI results by radiologists participating in the experiment: 65 %.

Number of scientific papers published: 14 scientific papers have been published by researchers.

7.8.20.6 Features of use case

Task(s): recognition by processing and analysing DICOM files received from diagnostic devices of radiology departments in the city by AI services to improve a quality of radiology reports and reduce the time from a study's completion to a provision of the report by a radiologist.

Level of automation: assistance

7.8.20.7 Threats and vulnerabilities

Exceeding a permissible percentage of technical defects in the AI-service operation.

Exceeding a permissible percentage of clinical defects in the AI-service operation.

7.8.20.8 Challenges and issues

AI integration into the national clinical guidelines and a development of financing mechanisms for the Obligatory Medical Insurance system.

7.8.20.9 Trustworthiness considerations

The basic functional and diagnostic requirements are developed for the AI services. It includes methods for step-by-step testing and monitoring technological, diagnostic and clinical quality. More than 250 labelled data sets have been created using original methodology. This includes one of the first and the largest COVID-19 data sets.

7.8.20.10 Use of standards and standardization opportunities

Eleven national standards for using AI in the healthcare came into effect based on the scientific results of the experiment. A proposal for developing a guideline on conducting clinical trials of AI-services in diagnostic radiology was submitted to the international standardization organizations in 2020.

7.9 Home and service robotics

7.9.1 Device control using AI consisting of cloud computing and embedded system (use case 132)

7.9.1.1 Objectives

Learn the user's preferred temperature in each situation for the control of home appliances (air conditioning equipment).

Keep rooms comfortable by running home appliances (air conditioning equipment) at the user's preferred temperature according to the situation.

7.9.1.2 Narrative

Motivation: the temperature at which the user feels comfortable varies depending on the conditions outside the air conditioner, such as the outside temperature, intensity of sunshine, time of day, day of the week, etc. Always maintain a comfortable environment by eliminating the need for this setting change.

Problem statement: though the temperature at which the user feels comfortable depends on the situation, such as the time of day and the day of the week, it is impossible to present these settings at the time of product shipment. Even if the designer of the product provides a method to enable the user to set such a setting, users themselves do not know they are expected to set their preferred temperature preferences. Long-term data cannot be stored in the device, so the model is forced to learn in the cloud. Training the model in the cloud takes longer to be able to cope with the sudden variations in the operating pattern of the user.

Current situation: the temperature is set using the controller every time the user feels uncomfortable.

Solution approach and solution steps: in addition to training the model using long-term historical data in the cloud, the model is also adjusted by frequent training in embedded devices. When the user changes the temperature setting using the controller, not only the setting but also accompanying data, such as the setting time, are stored in the air conditioner. Data about the operating status, such as temperature sensor values installed for the control of the air conditioner, are stored in the air conditioner. Data stored in the air conditioner is periodically uploaded to the cloud instance held by the manufacturer. The latest weather forecast information is kept on the cloud service at all times. A model is created to represent what the set temperature is expected to be depending on the external conditions around the air conditioner (including the forecast) by periodic learning for each air conditioner on the cloud. The model is delivered to the corresponding air conditioner. Online machine learning is performed based on the data stored inside the air conditioner, and the internal parameters of the model are adjusted. This embedded learning is performed frequently, e.g. once an hour, and it is possible to reflect sudden changes in the user's usage pattern to the model. The online machine learning algorithm inside the air conditioner and the batch machine learning algorithm in the cloud are tuned as close as possible to prohibit radical change of the model from the model adjusted by online machine learning when the model is delivered via cloud computing and the adjusted model is overwritten. The air conditioner predicts the preferred temperature by using the model, and the result is used as the set temperature of the air conditioner. As in normal operation, the air conditioner performs control so that the temperature of the room remains at the set temperature.

Results and effects: since the prediction is done by the air conditioner (embedded), it keeps working even if there is a network failure or a cloud failure. The only impact of a failure is the inability to upload data and the inability to update the model as it learns via the cloud. Operation training with a long-term cycle, such as a fixed operation for each day of the week, is effective if the model is trained from the accumulated operation history. A model with this effect is created mainly by learning on the cloud. If there are sudden operation pattern changes, e.g. when the outside temperature rises suddenly and the user reacts to it, high frequency online machine learning inside the air conditioner can adjust the model immediately.

7.9.1.3 Stakeholders and stakeholder considerations

Stakeholders: equipment users, manufacturers, providers

Stakeholders' assets, values:

— equipment users: comfort, unintended (unpleasant) behaviour, riskless behaviour, privacy;

— manufacturer: competitiveness, reputation, reliability, safety;

— distributor: no claims for unintended (unpleasant) behaviour.

7.9.1.4 Data characteristics

Source: the data sources include the air conditioner which stored the controller settings (such as the temperature setting, the setting time) and the operating status (such as temperature sensor values), and the weather forecast information that are provided from the meteorological agency.

Variety: the structured data.

Velocity: the data from air conditioner are uploaded periodically (setting by the manufacture) and the weather forecast information are increased daily.

Quality: high.

7.9.1.5 KPIs

Number of cancel operations: the air conditioner changes the temperature setting based on the prediction. When the user notices an unintended(unpleasant) setting, the user operates the controller to cancel the setting that was based on the prediction.

Distance between models: the difference between the model trained in cloud computing and the one learned in the embedding equipment (air-conditioner).

7.9.1.6 Features of use case

Task(s): prediction

Level of automation: conditional automation

Method(s): machine learning, online machine learning

Platform: PC (pre-validation), cloud, cloud-to-device communication (Internet), embedded equipment

Topology: All air conditioners are connected to one cloud. On the cloud, keep a history of past operations and operating conditions for all air conditioners. Learning for each air conditioner on the cloud, and delivering the created model to the air conditioner. The air conditioner retains the operation history and operation status history data for a certain period of time, and also maintains the delivered model. The model is adjusted regularly by executing online machine learning in the air-conditioner. Change the set temperature based on prediction, which is based on the model in the air conditioner.

7.9.1.7 Threats and vulnerabilities

Creating an incorrect model by machine learning using a child's mischievous operation history.

Creating an incorrect model by machine learning using the history of operations based on user misunderstanding, for example, operations that set the temperature extremely low when the user wants to cool down immediately.

When an air conditioner is resold, the use pattern of the original user leaks to the next purchaser when the air conditioner is used.

Threats to cloud computing in general.

7.9.1.8 Challenges and issues

During actual use, there is a possibility of significant difference between the model that learns by cloud computing and the model adjusted inside the air-conditioner. There is a significant change in the temperature setting if the model in the air conditioner is overridden by the model that learns on the cloud.

How and when to detect whether there has been a significant difference.

How the air-conditioner explains a significant difference when it is detected. Criteria for determining whether to explain this difference.

7.9.1.9 Trustworthiness considerations

The purposed solution address different aspects of the trustworthiness in AI systems, related with the performance efficiency of characteristics like capacity and behaviour time to make the required adjustments of variables such as the temperature taking into account overheating or overcooling parameters. Smart-home applications, also they use to take into account trustworthiness aspects related with accountability, security and responsibility when introduce final products within consumer digital markets.

7.9.1.10 Use of standards and standardization opportunities

This use case is a typical smart-home application, with medium-high degree of conditional automation, with a scenario where the AI systems have more relevance over the external oversight of some internal parameters of the home such as the temperature. In order to distinguish standardization opportunities, it is important to understand that it is a hybrid solution, where ML algorithms are implemented into the embedded AI systems, and also in the cloud to develop the prediction tasks required by the solution. It is possible to distinguish the following areas of improvement through the implementation of standardization processes:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk impact assessment.

7.10 ICT

7.10.1 AI system to help mobile phones to have better picture effect (use case 32)

7.10.1.1 Objectives

Better understanding the image and improving the image effect on a smartphone by using a DL model that is trained on the cloud or offline.

To find an efficient solution to increase camera image quality on a smartphone without significantly increasing the operation and power burden on the mobile phone.

7.10.1.2 Narrative

An AI system was developed that can increase smartphone camera image quality. Using deep learning, the smartphone can identify more scenarios and objects than before. Based on the identified scenarios and objects, the smartphone can better understand the image and improve the image effect.

At present, there are 1,4 billion smart phone shipments in the world every year. Photography is one of the most important functions of smart phones, and the industry has been trying to improve the picture quality of mobile phone photography. The industry goal is to reach the quality of a professional single-lens reflex

(SLR) camera. The traditional image processing algorithm has hit its ceiling in terms of potential. Traditional algorithms cannot be used for photographing many scenes, simply because the effect is very poor.

This deep learning algorithm provides a turning point for solving the above problems. By using the AI system, smartphones can better “understand” the pictures they take. Based on the deep learning algorithm, the smart phone can analyse the scene in real time and intelligently identify a range of scenes during the shooting process, such as blue sky, flowers, green plants, night view, snow scene, etc. The smart phone can also intelligently detect the objects in the scene. Based on scene recognition and object detection, the smartphone can automatically adjust and set parameters for different pictures, so as to get better photo effects.

Mobile phones can now recognize one hundred kinds of scenes, and this number will reach the hundreds in the future. By using the deep learning algorithm, mobile phones can now detect 20 types of objects, and in the future, they will be able to detect hundreds of objects. Object detection can be used for smart digital microscope (auto focus on targets), and portrait segmentation can be used for background blur or light efficiency.

7.10.1.3 Stakeholders and stakeholder considerations

Stakeholders: mobile phone manufacturers, end users, third party testing and evaluation agencies.

Stakeholder considerations: competitiveness

7.10.1.4 Data characteristics

Description: annotated pictures

Source: public picture library, self-collection picture library, web crawling pictures or automatic synthesis of pictures.

Type: picture format supported by a training platform and smart phone.

Variety: single source

7.10.1.5 KPIs

Mean intersection over union (MIoU): the intersection of the prediction area and actual area divided by the union of the predicted area and actual area. The ideal target is 100 %. The reference of this KPI to mentioned use case objectives is to improve accuracy.

False acceptance rate (FAR): negative samples are identified as positive samples / total number of negative samples. The lower the value of FAR is, the more likely it is the smartphone would detect the correct scenes and objects. The reference of this KPI to mentioned use case objectives is to improve accuracy.

7.10.1.6 Features of use case

Task(s): recognition

Level of automation: conditional automation

Method(s): deep learning

Platform: neural network processing unit (NPU), GPU, CPU etc.

Topology: no need

7.10.1.7 Threats and vulnerabilities

New privacy threats (hidden patterns).

7.10.1.8 Challenges and issues

Challenges: achieve the same picture quality as a professional SLR camera.

Issues:

— lack of data for certain scenes;

— lack of computing ability on the device side;

— users can see an improvement in image quality but do not know that it is due to AI.

7.10.1.9 Trustworthiness considerations

Incorrect object detection can lead to racial prejudice or privacy protection problems.

7.10.1.10 Use of standards and standardization opportunities

The standardized content includes:

— the formats of training picture data;

— the formats of deep learning model generated offline or cloud, which would be transplanted to smart phones;

— the platform to support the transplanted model in the smart phone;

— API which can be used by other applications, such as picture classification, security.

7.10.2 Product failure prediction for critical IT infrastructure (use case 86)

7.10.2.1 Objectives

Building an AI system to augment quality assurance (QA) engineers.

Reduce the likelihood of releasing defective batches of hardware.

7.10.2.2 Narrative

A deep learning model to learn from a visual representation of the number of items that failed in a specific batch of hardware as well as the type of defect.

The hardware manufacturing company was using a few QA engineers to make subjective calls on whether a specific batch was good enough to be released into the market. The graphical representation of the shortfalls and defects was also done manually. This led to inconsistent labelling and many unsatisfied customers. To augment the QA engineers, a deep learning AI model was developed to do more accurate and consistent labelling of which batches were likely to be most defective and the major type of defects.

7.10.2.3 Stakeholders and stakeholder considerations

Stakeholders: QA engineers, manufacturing line technicians, technical sales.

Stakeholder considerations: customer satisfaction index, cost of returned merchandise, time spent on QA.

7.10.2.4 Data characteristics

The data used in this use case are the visual representation of the number of items that failed in a specific batch of hardware as well as the type of defect.

7.10.2.5 KPIs

Prediction accuracy: consistency of prediction compared to actual defect rates. The reference of this KPI to mentioned use case objectives is that the prediction accuracy is expected to be 80 % or more to ensure only the true-negative batches are inspected.

Time saved: time for QA engineers to inspect every batch. The reference of this KPI to mentioned use case objectives is that the prediction model highlights the most obvious defective batches and allows the QA engineers to spend time only on high-discretion tasks.

Customer satisfaction: the number of returns from the manufacturer's customers. The reference of this KPI to mentioned use case objectives is that the satisfaction goes up when the number of defects is reduced upfront before the sales process.

7.10.2.6 Features of use case

Task(s): prediction

Level of automation: conditional automation

Method(s): deep learning

Platform: private on-premise servers

Topology: bus and hybrid

7.10.2.7 Threats and vulnerabilities

If the retraining model is compromised due to significant changes in the input data, the prediction model can generate incorrect outcomes and cost the hardware manufacturer serious loss.

7.10.2.8 Challenges and issues

Challenges in identifying which deep learning model gives the best performance output, and challenges in indexing raw flat files into visualization images.

7.10.2.9 Trustworthiness considerations

Address issues of sustainable manufacturing and high-value technical jobs.

7.10.2.10 Use of standards and standardization opportunities

Failure prediction models can improve global standards in manufacturing by reducing the waste of materials used and energy and water consumed.

7.10.3 AI-based optimized field dispatch (use case 151)

7.10.3.1 Objectives

Optimization of field dispatch based on AI.

Save operating expense (OPEX) by avoiding false field dispatching decision.

7.10.3.2 Narrative

The trouble tickets issued by the ISPs when the network problems occur, have been used to guide the field engineers to the right problem spot by informing them the location of the fault site and the reasons as well. However, some of the tickets do not require any field actions, but just give the additional information to the control room operators.

To minimize the number of false dispatch tickets, DL technology has been introduced. As it is natural for any ISPs to store their engineers' work records for a certain period, it is relatively easy to train DL model using them.

After being approved to be superior than human operators, the implemented DL model has been applied on the field.

In order for ISPs to maintain the reliable and secure communication network infrastructures, it is essential to send a field engineer on time to the fault site to recover the failure.

But sometimes it is unclear whether a field engineer can really be sent to the fault site only depending on the logs and symptoms in the control room. It is important to pinpoint which alarms can be recovered by dispatching the field engineer, otherwise false dispatching occurs and OPEX loss follows.

Traditionally ISPs issue the trouble tickets to dispatch the field engineers. Based on the directions on the tickets, the engineers can move to the fault site and fix the problems.

With the rapid development of networking technologies, network equipment has become more sophisticated and the structure configured by that equipment has become extremely complicated, resulting in the change of the nature of trouble tickets. The trouble tickets started to show only the information to be referred by the operators and they do not mean field dispatch anymore. Some of them are just for informational or can be covered from the remote-control room.

As a result, it became important for the operating staff in the control room to tell whether a certain trouble ticket really need to be covered by the field engineers. But it is not easy even for the experienced operators to make a correct decision all the time. It is neither easy to make it a complete rule set to tell the 100 % correct prediction.

Recently, the telecom company has been adapted DL technologies on this matter and successfully applied it to field. It was relatively easy to apply deep learning technology in this matter, because engineers' work records have been kept well in large quantities and the stored records contained the information about whether the site dispatch was necessary.

After investigating couple of different models, CNN-style models have been chosen to be implemented because at that time its performance was already proved especially in the computer vision area. As the trouble tickets can be transformed into the 2D images, any CNN-style models can easily get them into their input images and show the excellent performance.

Since this model shows better performance than human operators in terms of determining whether a ticket can be taken care of by a field engineer or not, it has been applied to the field quickly.

7.10.3.3 Stakeholders and stakeholder considerations

Stakeholders: ISPs dispatching field engineers.

Stakeholder considerations: avoid the unnecessary dispatching based on the wrong decision.

7.10.3.4 Data characteristics

The data used in this use case are the 2D images for CNN-style models. The 2D images are transformed from the trouble tickets which are extracted from engineers' work records contained the information about whether the site dispatch was necessary.

7.10.3.5 KPIs

Classification accuracy: how to accurately make a decision on whether the dispatch is needed. The reference of this KPI to mentioned use case objectives is to save the OPEX by not sending out the field engineers.

7.10.3.6 Features of use case

Task(s): prediction

Level of automation: high automation

Method(s): CNN-style DL model has been applied to classify the trouble ticket records.

7.10.3.7 Threats and vulnerabilities

Lost the opportunity to save the OPEX by sending out their engineer to the field unnecessarily.

7.10.3.8 Challenges and issues

None identified.

7.10.3.9 Trustworthiness considerations

None identified.

7.10.3.10 Use of standards and standardization opportunities

None identified.

7.10.4 Wireless network failure prediction (use case 152)

7.10.4.1 Objectives

Wireless core network failure prediction based on AI.

Avoiding serious wire core network failure before it happens.

7.10.4.2 Narrative

Since almost all communication services are based on the wireless network environment, stable operation and management of wireless network is very important.

In particular, if the wireless core network can be closely monitored and the symptom of abnormality can be detected in advance, the survivability of the network will be greatly improved.

A telecom company has implemented a wireless core network failure prediction system using auto-encoder model and deployed it successfully. The key algorithm is to create an auto-encoder model with LSTM model to train the normal traffic. The trained model can detect the faulty traffic movement and prohibit the serious network fault pre-actively.

As the 5G market expanded in earnest, stable operation of wireless core networks became more important than ever. In the wireless core network, the traffic of users flows in a collective form. By monitoring this in real time, it is possible to quickly diagnose the abnormality of the entire network.

Service providers are monitoring the collective wireless user traffic by tapping key connection points between the wireless core network components for monitoring purposes and for accumulating the monitoring results.

Accumulated data are provided to the human operators through a specific network management system to help determine the abnormality of the traffic and the network itself on which it flows.

However, the prediction of anomalies based on empirical knowledge varies widely depending on the person in terms of prediction accuracy and swiftness of processing. It can be possible for the human operators, especially not experienced ones, not to recognize a particular symptom which can cause large-scale damage on the core.

In order to resolve this problem, the telecom company has developed an AI-based wireless core network failure prediction system. As for the core algorithm, a long-term memory (LSTM) model has been first investigated among the recurrent neural network (RNN) technology that shows excellent performance in detecting patterns occurring in the time axis among deep learning models. It is applied to the legacy management system to implement and commercialize the function to predict wireless network failure in advance.

Inside the core engine, it implemented an auto-encoder model based on LSTM model. The auto-encoder model has been trained by the accumulated wireless core network traffic logs to learn the normal traffic patterns. The trained model can predict the abnormal traffic before the problem actually happens, so that serious failure situation can be avoided.

7.10.4.3 Stakeholders and stakeholder considerations

Stakeholders: ISPs operating wireless core network.

Stakeholder considerations: predict the abnormality of wireless core network in advance.

7.10.4.4 Data characteristics

The data used in this use case are the accumulated wireless core network traffic logs, which are used to learn the normal traffic patterns.

7.10.4.5 KPIs

Prediction accuracy: how accurately predict the anomaly situation? The reference of this KPI to mentioned use case objectives is to improve pre-activeness of abnormality detection.

7.10.4.6 Features of use case

Task(s): prediction

Level of automation: high automation

Method(s): auto-encoder model using LSTM building blocks

7.10.4.7 Threats and vulnerabilities

Missing and ignoring a particular symptom which can lead to a serious damage on the whole wireless network.

7.10.4.8 Challenges and issues

None identified.

7.10.4.9 Trustworthiness considerations

None identified.

7.10.4.10 Use of standards and standardization opportunities

None identified.

7.10.5 AI performance evaluation of AI-powered messaging bots (use case 176)

7.10.5.1 Objectives

Determine the efficacy of AI messaging bots and automations using another AI to provide a numerical rating of conversation quality. Includes a middle AI layer to provide targeted analysis of specific issues in bot automation. Reduces manual time and energy spent monitoring and reading individual automated conversations. Limited to messaging bot conversations only; not applicable in current stages to human-to-human interactions or automations outside the proprietary messaging ecosystem.

7.10.5.2 Narrative

The ubiquity of human-bot interactions is a given and standard part of our lives. Humans are as critical of human-bot interactions as human-human interactions and are prone to negativity bias. However, because humans communicate with bots in a very command-like language (for example, "Siri, play Despacito" or "Directions to the nearest petrol station"), traditional AI conversation quality measurement techniques using NLP/NLU fail to accurately discern the quality of the human-bot interaction.

The meaningful automated conversation score (MACS) replaces the traditional numerical and statistical techniques that measure the performance of messaging bots (commonly referred to as chatbots) and automations.

Additionally reduces the entirety of automation corpora to the subset that contains identifiable and actionable issues for the purpose of bot AI tuning. Allows bot tuners to focus on targeted conversations to more rapidly generate improvements/recommendations.

MACS uses 5 machine learned models, one for each "failure reason" and the score itself.

7.10.5.3 Stakeholders and stakeholder considerations

The stakeholders include:

— bot developers can assess the MACS rating of the bots they have developed and take corrective action directly;

— bot tuners can generate recommendations for usage by bot developers;

— bot managers can assess the MACS rating for all bots deployed by an organization;

— business managers can assess the root cause of customer dissatisfaction with a particular automation (at the intent level);

— end users will benefit from the corrective action undertaken to improve the MACS of individual bots and automations.

7.10.5.4 Data characteristics

General: MACS is measured on a discrete scale of "1", "2" and "3" with "1" denoting a poor experience and "3" denoting a great experience.

Source and variety: the ML model has been trained on roughly 5 000 conversations that were classified by humans. Conversation data are mostly text but can include structured content like menu tables and images. Additionally includes conversation metadata and proprietary features like consumer intent, presence of transfer or escalation.

Velocity: conversation consumed by the model is being generated real-time and historically. Data are currently ingested, used for computation and stored daily. Visualizations are also updated daily.

Quality: conversation data are used in internal stores with full data integrity, with the possible expectation of outages. Not all metadata fields are applicable to all conversations (some do not use intent recognition, for example) but data are not expected to be incorrect or missing.

Protected attributes: model was trained on a subset of data where all personally identifiable information (PII): names, numbers, were masked for security reasons. Gender, race, religion is not identified or separated in the data.

7.10.5.5 KPIs

MACS score:

— main KPI of MACS is correlation: the product decision was to focus on being correlated with the human evaluation of the conversation experience. As a perceived score 1 conversation is a problem conversation.

MACS failure state:

— precision: how accurately do we identify the reason for failure in the conversation?

— recall: from all the automated conversations with a failure, how many failures did we truly identify?

Model runtime:

— model runs batched, nightly across 3 regional servers for all of the previous day's closed conversation data. It scores (and provides all failure reasons) for 600,000 conversations in 45 min.

7.10.5.6 Features of use case

Task(s): natural language processing

Level of automation: high automation

Method(s):

— AI methods: machine learning, dialogue management and measurement, natural language understanding, language modelling, dialogue act classification.

— Models: BERT, XGBoost, RandomForest, gated recurrent unit (GRU).

— Frameworks: PyTorch, TensorFlow.

Platform:

— Development: JupyterHub, Pachyderm.

— Deployment: Pachyderm and MySQL tables, Conversation Builder (proprietary web-based bot building User Interface).

Topology: data are initially extracted in the extract transform load (ETL) layer. It then runs through a featurizer - to clean and prepare the data for as model input. The models are then run on the data input; as a result, models can be easily updated (as long as they continue to use the same features.)

7.10.5.7 Threats and vulnerabilities

Implicit biases of annotators or biases based on training data set (of brands willing to share data) exist and are unidentified. Security threats to the model would represent threats to the company as a whole, as the model, data and infrastructure live behind a firewall on the internal private cloud platform. PII data are masked for security reasons, and the score and reasons do not directly expose any aspect of the conversation data, PII or otherwise. It is possible the model, if acquired by malicious actors, can be used to backwards engineer some masked conversation text.

7.10.5.8 Challenges and issues

Limitations include the fact that the model mostly predicts negative failure states and does not currently track improvements to the conversation flow. It can be considered that there are additional failure states to be discovered or defined; the subset of failure states does not cover all potential failures. Also, the model is dependent on the conversation text. This AI system currently only support English language conversations, as the effort to translate and annotate across languages is resource intensive.

Accuracy of the model is useably passable at the aggregate level; which meets the current use case requirements for pointing out major flaws within bot design.

7.10.5.9 Trustworthiness considerations

Bias mitigation: data for development was sampled from more than 60 accounts to ensure uniformity. Conversations were annotated by multiple annotators, all of whom were exposed to no more than 10 % of the entire data set and were trained on bias identification when annotating. The model itself is not subject to any bias mitigation efforts.

Ethical and societal concerns: It did not identify any specific concerns in this system.

Explainability: MACS produces two interrelated measures; the score is a measure of consumer experience and the reasons provide an explanation for a bad score. MACS is annotated in the way it is applied, this use case has the annotations provide a score and provide an explanation for the score, the data are QA'd by the expert taxonomists and only data that fits with the defined conceptual relationship between scores and failure states is used to develop our models.

Human autonomy: the AI provider strictly PII masks all data that are used for development of MACS, the models cannot be tampered with to retrieve any PII information.

7.10.5.10 Use of standards and standardization opportunities

None identified.

7.11 Insurance

7.11.1 AI services for health insurance companies (use case 161)

7.11.1.1 Objectives

This use case analyses two AI technology providers to healthcare insurance companies (German health insurance company and business intelligence centre for healthcare insurers) and evaluates ethical issues.

Smart information systems in the insurance industry are used for:

— fraud detection;

— risk management;

— marketing.

7.11.1.2 Narrative

An estimated 80 % to 85 % of insurance companies use smart information systems in 2018. Data from various sources play a vital role in helping insurance companies:

— improve claims management;

— improve claims tracking;

— identify of fraudulent behaviour;

— conduct finance and accounting;

— decision making;

— predictive modelling.

The use case focuses on two companies providing AI services to health insurance companies. The cases show that the ethical issues identified in the literature (such as bias, reliability, discrimination, data governance, security and privacy, job losses) differ from the issues identified by the two case companies (e.g. trust, informed consent and responsibility). Transparency and accessibility of data as ethical issues appeared to be important in both literature and practice.

7.11.1.3 Stakeholders and stakeholder considerations

Stakeholders: insurance companies, AI service providers, healthcare organisations, health insurance, customers, patients and clients.

Stakeholder considerations:

— Privacy: the most obvious issues when using smart information systems concern privacy.

— Trust: health insurance companies are the main clients and it is very important that health insurance companies trust the AI service provider. Data sharing requests from universities or research are only honoured with the consent of the health insurer. They are less concerned about the trust of insured persons, but that follows from the specifics of their functions.

— Solidarity: solidarity is a core value for insurers. The use of AI allows for detailed distinction between customers cohorts that present greater or lesser risky. Particularly where risks cannot or partially, be influenced by the customer, the risks can undermine solidarity in the sector.

— Data governance: there are many questions regarding data governance when using smart information systems and smart technologies. The insurance industry currently lacks specific regulations, except for the GDPR.

— Bias and reliability: data can be missing, corrupted, inconsistent, recorded in different formats in different data sources and lack a standard vocabulary. Data problems in healthcare are often perceived to be the result of the volume, complexity and heterogeneity of the data, their poor mathematical characterization and their non-canonical form. When data are not statistically sound, they reduce their efficacy for training. As a result, artificial intelligence can make many false assumptions or being discriminatory (e.g. making insurance more expensive for minorities).

— Discrimination: as smart information systems become more sophisticated, discrimination emerges. The predictive modelling capacities of artificial intelligence systems constitute a natural “fit” to the assessment of risk inherent in the processes of insurance ratemaking and pricing. This nature of the AI system can increase the discrimination in this use case.

— Transparency: determining which set of variables are the most relevant for the artificial intelligence model is not always easy, because artificial intelligence chooses those completely by itself in accordance with its experience. But the customer has a legal right to be informed about the use of this personal data. That is why it is very important what methods insurance professionals will use to be able to respond fully to customer requests for explanation of the reasoning that underlies those determinations, given the mystery that cloaks the algorithms by which cognitive systems produce their results. The two cases are not open about their methods to the clients, nor do they think this is relevant.

7.11.1.4 Data characteristics

The data used in this use case are the healthcare data of the health insurance customers, patients and clients.

7.11.1.5 KPIs

The accuracy of fraud detection.

The improvement of claims tracking.

The accuracy of decision making.

The accuracy of predictive modelling.

7.11.1.6 Features of use case

Task(s): optimization

Level of automation: assistance

7.11.1.7 Threats and vulnerabilities

Accessibility of data: within the organizational limits required by regulation and security, data are essential for insurance companies for training and machine learning. The main limitation is accessibility of the data, because data often exists in different settings and systems, such as administrations, clinics, laboratories, registers, public and private companies.

Security: adoption of smart information systems in the insurance sector raises many barriers and challenges. When talking about security, authentication is very important. Authentication is understood as the act of establishing or confirming that claims made by or about a subject are true.

7.11.1.8 Challenges and issues

Job losses: although artificial intelligence and smart information systems have been promised as problem-solvers, they have also sparked concerns about job losses.

7.11.1.9 Trustworthiness considerations

As smart information systems become more sophisticated, discrimination emerges. The predictive modelling capacities of artificial intelligence systems constitute a natural “fit” to the assessment of risk inherent in the processes of insurance ratemaking and pricing. The business strategy of the private company can diverge from the societal or governmental objectives for (health) insurances with respect to equity and justice and solidarity.

7.11.1.10 Use of standards and standardization opportunities

Guidelines and requirements for transparency: to be able to respond fully to customer requests for explanation of the reasoning that underlies those determinations of AI methods used.

7.12 Knowledge management

7.12.1 Water crystal mapping (use case 77)

7.12.1.1 Objectives

Similarity classification of water crystals as a knowledge map.

7.12.1.2 Narrative

Much of the earth’s surface is covered by water. As was pointed out in the 2020 edition of the World Water Development Report[18], climate change challenges the sustainability of global water resources, so it is important to monitor the quality of water to preserve sustainable water resources. Quality of water can be related to the structure of water crystals, the solid-state of water, so methods to understand water crystals can help to improve water quality. The data set provided by an NPO containing 5 007 crystal photos has been created as the first annotated data set of water crystals. Water crystal shapes has been classified into 13 categories. A data set containing 20 K crystal photos has been delivered for service production.

7.12.1.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the AI system for water crystal classification;

— AI customer: the organization that adopts this AI system to analyse the water quality, such as municipality, county, regions, UN, water-related researchers;

— AI subject: citizens;

— data provider: the organizations that provided and annotated 5 K water crystal photos.

The stakeholders' considerations include the following:

— AI producer considers how to enhance the performance of water crystal classification;

— AI customer considers how to utilize the classification results for water quality analysis.

7.12.1.4 Data characteristics

Source: the water crystals have been provided by an NPO. Crystals were produced from water samples collected from many countries and sources, with the help of scientists all around the world.

Type: very high-resolution images (5 472 × 3 648 pixels)

Velocity: real-time in production phase

Variety: image data only

Variability (rate of change): moderate

Quality: high

7.12.1.5 KPIs

The main KPIs of this use case include:

— the water quality (turbidimetry) that is used for water crystal ranking;

— the crystal similarity (crystal category) that is used for crystal classification in the water crystal map.

7.12.1.6 Features of use case

Task(s): water crystal similarity ranking and clustering.

Level of automation: partial automation

Method(s): deep learning approach and crystal structure embeddings

Platform: migration from server to cluster machine.

Topology: distributed

7.12.1.7 Threats and vulnerabilities

None identified.

7.12.1.8 Challenges and issues

The challenges of this use case include to use the classification results to enhance the water quality.

7.12.1.9 Trustworthiness considerations

Trustworthy machine learning to be evaluated in 2022-2023.

7.12.1.10 Use of standards and standardization opportunities

The standardization opportunity in AI model generation can be considered in this use case.

7.12.2 AI system with a digital knowledge centre for utilizing the knowledge in the organization (use case 187)

7.12.2.1 Objectives

The AI system establishes and uses a digital knowledge centre to utilize knowledge for assisting in the operation and maintenance of IT services that depend on human knowledges.

7.12.2.2 Narrative

In recent years, as ICT systems have become increasingly complex due to the use of multiple cloud services, business opportunities can lose due to the longer time it takes to respond to inquiries and conduct troubleshooting from users.

The AI system in this use case establishes and utilizes a "digital knowledge centre" that digitizes the knowledge of skilled staff and utilizes it throughout the organization. This changes the way operations and maintenance work is carried out to ensure business succession and increase the value of IT services.

This AI system includes the following functions:

— In inquiry response work, instead of searching incidents (previous cases) and knowledge (FAQ) by trial and error, simply select from a list of candidates and get an answer right away;

— In trouble shooting work, any staff can follow the procedure by simply following the flow without asking a skilled person for the procedure. This AI system can associate existing documents with procedures by simply registering them without organizing them;

— In knowledge maintenance, this AI system efficiently maintains knowledge without analysing a huge amount of response history, as well as understands the status in real time without manual extraction and analysis.

This AI system has the following characteristics for establishing and utilizing the digital knowledge centre:

— To consolidate knowledge: this AI system uses knowledge networking technology for large documents such as automatic pre-defined tagging and learning to optimize tagging for large internal documents. The AI system can extract the knowledge from the existing documents in the organization knowledgeable by specifying a storage location without organizing the document. Thus, the problem that a beginner does not know the contents of a document and cannot find information scattered in the organization is solved. Even beginners can use the tag information to find information in existing documents from knowledge networks.

— To utilize knowledge: this AI system uses a knowledge inspection technology based on task similarity for automatically displaying the previous cases, similar cases and related documents optimized according to the utilization state of corresponding candidates by navigating a task procedure in a process bar. Thus, even if a beginner does not have the knowledge necessary for performing the work, it is not necessary to ask an expert how to proceed the work, and the work can be performed while viewing a flow in which the work process and knowledge are displayed as long as the flow is followed.

— To increase knowledge: this AI system uses a fit and gap technology based on knowledge enrichment to present knowledge extension points without having to look back at work. As a result, it is not necessary for the skilled staff to manually extract and analyse the enormous records of the inquiry response work to find insufficient knowledge, and a beginner can increase knowledge by the assistance of the AI system. This AI system also provides visibility into how knowledge is being used and optimizes the presentation of extension points.

7.12.2.3 Stakeholders and stakeholder considerations

The stakeholders include:

— AI producer: the organization that develops the AI system and provides the technologies to establish and utilize the digital knowledge centre;

— AI customer: the organization that adopts this AI system to utilize the knowledge for the operation and maintenance support work;

— AI subject: the staff that are related to the operation and maintenance work. The users who contact the organization to require the trouble shootings and other operation and maintenance supports.

— data provider: the organization that adopts this AI system and provides the internal documents to the AI system for establishing and utilizing the digital knowledge centre.

The stakeholders' considerations include the following:

— AI producer considers how to enhance the performance of establishing and maintaining the digital knowledge centre;

— AI customer considers the workload and staffing of the operation and maintenance support work by adopting the AI system, as well as considering the customer satisfaction on the support work;

— Data provider (AI customer) considers how to collect the internal documents and other useful data adequately and effectively (e.g. considering what kinds of the documents can be transited to the digital knowledge centre).

7.12.2.4 Data characteristics

The data used for this use case are the internal documents (includes paper documents) and the existing historical records regarding the operation and maintenance support work.

Source: the internal data storages, includes the on-promise file servers in the organization and the cloud storage services.

Variety: the structured and unstructured text, images of the paper documents.

Velocity: the data and documents are increased daily (e.g. the daily support records and the documents daily created in the organization).

Quality: high.

7.12.2.5 KPIs

The main KPIs of this use case include:

— the accuracy and process time on consolidating the knowledge daily increased internal documents;

— the workload reduction of the operation and maintenance support work.

7.12.2.6 Features of use case

Task(s): knowledge processing and discovery, natural language processing, information retrieval.

Level of automation: conditional automation

Method(s): AI models for the natural language processing technologies

Platform: cloud services

7.12.2.7 Threats and vulnerabilities

No threats and vulnerabilities identified.

7.12.2.8 Challenges and issues

The challenges of this use case include:

— to enhance the performance of establishing and maintaining the digital knowledge centre (e.g. the technologies to process the paper documents);

— to enhance the usability for assisting the operation and maintenance support work.

7.12.2.9 Trustworthiness considerations

This use case did not have the trustworthiness issues on the bias, explainability, controllability, predictability and transparency.

The impact to the support staff are desired to be considered because the AI system can cause unemployment.

7.12.2.10 Use of standards and standardization opportunities

The standards can be used in this use case include:

— the standards regarding the document formats;

— the standards regarding the exchange and storage format of the support records.

7.13 Legal

7.13.1 AI contract management (use case 120)

7.13.1.1 Objectives

Building an AI contract management solution for automating business processes related to documents: data classification, automatic data extraction and contract monitoring.

Creating a solution that is able to standardize the contract management process, improve the quality of work on problematic contracts and claims, and optimize lawyers' working process and relieve them from routine tasks.

7.13.1.2 Narrative

The AI contract management solution is built on an AI legal core, which includes technology that converts different types of documents into a digital format, replicates natural human-like text recognition and extracts data to automate business tasks.

It is a platform for automatic reading and analysis of legal documents and extraction of data with a high level of accuracy.

Based on the extracted data automatic contract monitoring and execution can be performed.

The system contains the following AI contract management features:

— structured digital document archive;

— hierarchical chain and connections of all documents in relation to a primary document, whether it is a contract, order or anything else;

— monitoring and control of key contract terms;

— creation of all necessary documents: notifications, claims, etc;

— auto filling the required enterprise resource planning (ERP) systems with relevant data.

7.13.1.3 Stakeholders and stakeholder considerations

Stakeholders: procurement department, legal department.

Stakeholder considerations: acceleration and rising quality of legal operations and processes.

7.13.1.4 Data characteristics

Description: different types of documents: contracts, agreements, non-disclosure agreements (NDA), etc.

Source: data warehouses (DW)

Type: structured or unstructured text, images

Velocity: real-time in production phase

Variety: different types of source with mostly structured data

Variability (rate of change): moderate

Quality: moderate

7.13.1.5 KPIs

Recall: also known as sensitivity, the fraction of the total amount of relevant instances that were actually retrieved.

Precision: also called positive predictive value, the fraction of relevant instances among the retrieved instances.

Customer satisfaction: the ratio of customer satisfaction when using this system for requests. The expectation is 100 %. The reference of this KPI to mentioned use case objectives is to increase its ratio as high as possible.

Algorithm accuracy: output when compared to human expert analysis of the same data.

Task completion rate: the performance is calculated by dividing the number of cases that have been completed successfully by the total number of assigned tasks. The success or failure of a task is set according to the criteria of each system. The reference of this KPI to mentioned use case objectives is the accurate task completion using the AI system.

Cost: minimize the financial costs and reduce the risk of penalties under the contracts.

Efficiency: improve the efficiency of existing manual document processing.

7.13.1.6 Features of use case

Task(s): contract management

Level of automation: high automation

Method(s): optical character recognition (OCR), NLP and knowledge representation, NLU, neural networks, machine learning, CV

Platform: 40 CPU, 80 GB RAM, SSD ~3,9 TB

7.13.1.7 Threats and vulnerabilities

Security threats, privacy threats.

Usually contracts contain trade secrets, disclosure of which can lead to serious financial losses. For this reason, the solution operates in a closed client protected form.

Bias due to changes in requirements on the customer's end or inappropriate training data.

7.13.1.8 Challenges and issues

Noisy data (variable scan quality).

Working with private data (information security).

Nonlinear models need significant computational power during the training phase.

7.13.1.9 Trustworthiness considerations

A helpful industrial solution that can optimize the current contract management process and assist to make the legal department's job easier.

7.13.1.10 Use of standards and standardization opportunities

The opportunity to bring the working process on a contract to a single standardized format (meta-document process) with ability to extract a key data set.

7.14 Manufacturing

7.14.1 Quality assurance solution based on AI, to detect defects on wind turbines blades (use case 4)

7.14.1.1 Objectives

To find an accurate and efficient solution to detect defects without compromising the detection of in-material damage and risking a loss in the manufacturer's reputation.

7.14.1.2 Narrative

This AI system can automatically detect defects in produced wind turbine blades through an image optimization method, which transforms raw data from the non-destructive ultrasonic testing (UT) scanner into image data based on RGB where deep learning-based image recognition can be applied effectively. It achieved high coverage (see KPIs) of more than 95 % for various defects, e.g. detecting missing back wall echo, foreign objects, dry glass defects and wrinkle defects. Quality controllers can focus their efforts on suspicious regions and disregard all clean data; humans are only expected to examine the regions of blades that are flagged as suspicious by the AI system. Since it achieved low split (see KPIs) of less than 20 %, it reduced evaluation time for UT scanning data by 80 %.

Each blade can be up to 75 m in length and takes a highly skilled professional quality controller up to 6 h to evaluate the UT scanning data in the quality assurance process. With 5,000 blades produced every year, that adds up to a saving of almost 32 000 person-hours, which translates into significant cost savings, reduced production lead times, and increased productivity. Today, there is a shortage of ultrasonic engineers and inspectors. This solution means the same inspector can do four to five blades per day instead of one previously.

7.14.1.3 Stakeholders and stakeholder considerations

The manufacturer is the customer who wants to enhance the efficiency of the quality assurance process on wind turbine blades while maintaining the quality level of the products.

The actual users of the solution are quality controllers and inspectors of produced wind turbine blades.

The producer is responsible for deploying an AI system capable of solving the above problem of efficiency. The provider is responsible for developing an AI system for the solution.

Other stakeholders include utility companies and neighbourhood communities.

7.14.1.4 Data characteristics

Source: UT scanner

Type: ultrasonic data from scanner vendor

Velocity: batch

Variety: single source

Variability: static

Quality: high

7.14.1.5 KPIs

Coverage (for accuracy): proportion of the defects automatically reported without manual inspection. Ideal target is 95 %.

Split (for efficiency): proportion of the regions of the blade that are flagged as suspicious and asking for manual inspection. The less split, the more efficient the total quality assurance process becomes.

7.14.1.6 Features of use case

Task(s): recognition

Level of automation: partial automation

Method(s): deep learning

Platform: deep neural network

Topology: convolutional neural network

7.14.1.7 Threats and vulnerabilities

Changes in defects of in-material damage over time.

7.14.1.8 Challenges and issues

Challenges:

Achieve the same level as ultrasonic accredited engineers for detecting critical defects.

Issues:

— lack of defect data per defect type;

— how to create good images for deep learning from UT raw data;

— back wall detection.

7.14.1.9 Trustworthiness considerations

The solution based on AI, is a mixed technical proposal, human-machine, where the system does not provide a 100 % of reliability. To resolve to possible lacks robustness and resilience, there is needed additional processes of continuous validation (monitoring and retraining) of the system.

Related with the transparency of the system, the explainability of the solution based on AI is achieved across a causal explanation of their functionality, developed by groups of AI stakeholders related with final user of the application, such as the controllers of the oversight of the system and quality inspectors, for the goal of understanding how the AI system of defect recognition, arrives at its results and to explain the chain of causal attributions of mechanisms by which the input features are processed to produce the given result.

7.14.1.10 Use of standards and standardization opportunities

Taking into account the level of autonomy of the AI system, the use case implements a mixed scenario between the AI system and the supporting with oversight of human experts. Both scenarios can be developed, across the stages of acquire, process and apply. In the particular use case of defect detection in wind turbine blades, can be identified the following issues and opportunities to be addressed by standardization:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk-impact assessment.

7.14.2 Generative design of mechanical parts (use case 15)

7.14.2.1 Objectives

To help mechanical engineers design lighter, strong and better parts.

To create optimized parts following precise mechanical constraints while enabling cost savings by reducing the amount of material necessary to achieve goals.

7.14.2.2 Narrative

Generative design is an iterative design process that involves a program that generates a certain number of outputs that meet certain constraints, and a designer that is possible to fine tune the feasible region by changing minimal and maximal values of an interval in which a variable of the program meets the set of constraints, in order to reduce or augment the number of outputs to choose from.

7.14.2.3 Stakeholders and stakeholder considerations

Stakeholders: organizations, designers, customers, end users.

Stakeholder considerations: competitiveness, safety, stability.

7.14.2.4 Data characteristics

Because the generative design is one of the design processes and it can be used in various domains, there is no specific data or data type that can be identified for the AI system in this use case.

7.14.2.5 KPIs

Typical design criterion's:

— weight reduction: to minimize the part weight respect to previous designs, to be improved;

— mechanical constraints metrics: to use mechanical metrics to obtain strong and maximum strength parts to be integrated in the final mechanical design.

7.14.2.6 Features of use case

Task(s): optimization

Level of automation: assistance

Method(s): genetic algorithms, optimization algorithms, generative adversarial networks

Platform: CPU, GPU

7.14.2.7 Threats and vulnerabilities

Highly dependent on engineer input for constraints and requirements.

7.14.2.8 Challenges and issues

Challenges: the engineers using this technology are still desired to know how to define the constraints, start and end points for the piece.

Issues: pieces generated to satisfy a set of constraints can still have design flaws overlooked because of misunderstanding by the user.

7.14.2.9 Trustworthiness considerations

Modular design of mechanical engineering requires the highest reliability ratios in the design of the parts, about aspects related to the performance efficiency and compatibility or re-usability in different parts and sections of final designs. Also is relevant the importance of aspects related to the security and accountability over the commercialized product, taking into account issues related to both trustworthiness characteristics since early stages of the design.

7.14.2.10 Use of standards and standardization opportunities

Depending on the final development, based on a degree of assistance automation, with supervised machine learning algorithm based on a greedy neural architecture search (GNAs), generative adversarial networks, it is possible to distinguish a scenario where AI acquires more relevance regarding the human expertise, providing a useful tool to improve design parameters related to the final performance of the design like the weight and resilience of the final product.

Taking into account aspects previously mentioned related to the use case, the following areas of improvement in the design stages of the product can be defined through the application of standardization process overall product life cycle and pipeline of different ML methods based on AI:

- acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency;
- process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;
- apply: acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk-impact assessment.

7.14.3 Powering remote drilling command centre (use case 36)

7.14.3.1 Objectives

Automatic generation of a daily performance report, reduction in overall drilling time, reduction in invisible loss time and improvement of rig asset management.

7.14.3.2 Narrative

It is important for a drilling contractor to monitor rig parameters in real time to optimize operations. The customer lacked granular insights during drilling and cannot ascertain the root cause of non-productive time, and manual interpretation of signals led to the failure to notice anomalies, further degrading performance.

The AI system extracted and ingested different types of signals from surface and downhole sensors to perform near real-time processing. More than 170 vital signals every second from each oil rig were processed by the AI system to provide near real-time insights into drilling operations. This was achieved by handling data format and data extraction standards, and the AI system provides the flexibility of generating customized asset utilization reports, thus helping the oilfield engineers to understand the root causes of non-productive time and better utilize the assets on field. Rig-specific utilization reports, and weekly and monthly utilization reports, helped to plan drilling operations, improving drilling efficiency.

7.14.3.3 Stakeholders and stakeholder considerations

Stakeholders: customers of oil and gas companies, relevant environmental regulatory and policy making authorities.

Stakeholder considerations: challenges to accountability and security threats.

7.14.3.4 Data characteristics

Description: data from an oil and gas rig

Source: drilling equipment

Type: time-series sensor data

Velocity: 2,5 billion+ data points each day

Variety: machine data

7.14.3.5 KPIs

Downtime: loss time that it is indicated in the lost time of the asset in being idle or off, or unplanned downtime. The report generated by the use case recovers KPIs related with the availability of the asset, with utilized time, lost time and their causes.

Overall drilling time: specific KPI that it indicates the time spent on one drilling job inclusive of all downtimes. The reference of this KPI to mentioned use case objectives is that the real time visibility into operations gives the operators early warnings to take actions immediately.

7.14.3.6 Features of use case

Task(s): knowledge processing and discovery

Level of automation: conditional automation

Method(s): utilization and performance evaluation

Platform: application server: 64 GB RAM/ 16 Core / 500 GB HDD; data server: 128 GB RAM/ 16 Core, 3 TB HDD

7.14.3.7 Threats and vulnerabilities

Challenges to accountability, security threats.

7.14.3.8 Challenges and issues

Compliance of organizations

7.14.3.9 Trustworthiness considerations

The proposed solution addresses different aspects of the trustworthiness in AI systems, related to the performance efficiency in aspects related to the improvement of availability drilling system and effectiveness parameters.

Together with the production process improvements, the provider of the solution remarks important trustworthiness challenges related to the reliability of the system, accountability or aspects related to the security of the final solution.

7.14.3.10 Use of standards and standardization opportunities

Depending on the final development, based on premise systems, with conditional automation, it is possible to distinguish a scenario where AI acquires more relevance regarding the human expertise, providing a useful tool to improve performance characteristics related with a better performance efficiency of temporal and resource management of parameter of the final solution. It is possible to distinguish the following areas of improvement through standardization process:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk-impact assessment.

7.14.4 Quality improvement of adhesive products, based on AI (use case 37)

7.14.4.1 Objectives

Enhance adhesive quality, performance benchmarking.

7.14.4.2 Narrative

The IoT signal intelligence platform provides the ability to have a holistic perspective and understanding of the sensitivity of the key parameters affecting output quality and the ability to monitor and control the process in real-time. This would avoid variations in yields, build-up of inventories and missed customer deadlines.

The IoT signal intelligence platform ingested three plus years of process data and sensor data regarding plant operations from temperature, rpm, torque and pressure sensors that were strapped on to industrial mixers. These are the mandatory sensors for the operations. The AI system provider used its episode detection algorithms (deep learning) to filter signals from noise and specifically identify the contributors to quality (anomaly signatures) that can then be used as signals to predict quality. It used its proprietary N-dimensional Euclidian distance-based scoring algorithms to normalize and present a unified score to the business team. This unified health score provided the process team with a different lens to benchmark, specifically target and radically improve process efficiencies. The AI system provider then leveraged its sophisticated ensemble models to predict potential quality failures, allowing the operations team to take real-time actions to control process deviations. The signals identified in the earlier steps provide model explainability to the end-user for reasons behind quality deviation.

7.14.4.3 Stakeholders and stakeholder considerations

Stakeholders: AI providers, AI customers or final users as manufactures and producers, relevant environmental regulatory and policy making authorities.

Stakeholder considerations: competitiveness (respond to and exceed customers' and consumers' expectations by providing the best value, quality, service and winning innovations, brands and technologies to create sustainable value).

7.14.4.4 Data characteristics

This AI system ingested three plus years of process data and sensor data regarding plant operations from temperature, rpm, torque and pressure sensors that were strapped on to industrial mixers.

7.14.4.5 KPIs

Accuracy: precision about to what extent has the model been able to predict correctly. The reference of this KPI to mentioned use case objectives is that provided ability as to % of times the quality complied.

7.14.4.6 Features of use case

Task(s): prediction

Level of automation: conditional automation

Method(s): N-dimensional Euclidian distance-based scoring algorithms

Platform: application server: 64 GB RAM/16 Core/500 GB HDD; data server: 128 GB RAM/16 Core, 3 TB HDD

7.14.4.7 Threats and vulnerabilities

Challenges to accountability, new security threats.

7.14.4.8 Challenges and issues

Patented process if any, security restrictions.

7.14.4.9 Trustworthiness considerations

The use case addresses different aspects of the trustworthiness in AI systems, remarking by the AI provider aspects related to the sustainability and environmental protection, through the investing in scientific research and innovation as the way to achieve a sustainable model. The purpose to improve the efficiency performance of previous developments, by the provider, through the development of technology based on AI is also relevant.

7.14.4.10 Use of standards and standardization opportunities

Taking into account the development model hybrid as a mixed of a platform of cloud services and IoT devices, joined to a relevant degree of autonomy of the AI system of conditional automation. Accepting the following stages of production process, acquire, process and apply, it is possible to identify the next issues and opportunities to be addressed by standardization:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency.;

— process: general issues such as the criterion for training data, criterion for validation model, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk impact assessment.

7.14.5 Empowering autonomous flow meter control- reducing time taken to “proving of meters” (use case 40)

7.14.5.1 Objectives

Reduce the time taken for trial and error methods to set the variable frequency device (VFD) and flow control valves (FCV) setpoints.

7.14.5.2 Narrative

With the previous system, the customer used to set VFD and FCV percentages manually to achieve the desired flow rates using trial and error methods, which usually can take about 3 h to 4 h, with a relevant lost time. The efficiency of proving the meters was very low, and improvement was necessary to remove any possible outlier in reading, as it was time-consuming. Through of the present use case, Cerebra module is integrated within the system, taking into account considerations of the flow of fluid. The customer can choose between the available options of a high flow rate, low flow rate or multi-viscous flow. Then, with the master meter in the loop of testing, the meter from the field is introduced to analyse and to quantify previous outliers and then prove the system by a more efficiently manner.

Since it takes some time to get the exact VFD and FCV percentage values, to achieve the desired flow rate, Cerebra's prognostics engine is integrated into the system.

Based on machine learning algorithms, the data models for the VFD and FCV percentages are used to predict the values to be chosen, with an accuracy around of 98 %. Since the system is a closed-loop, these predicted-values are automatically registered on the valves' monitors, and at the end only requires small tweaking to complete the operational process, thus reducing human labour.

7.14.5.3 Stakeholders and stakeholder considerations

Stakeholders: AI providers, AI customers or final users as manufacturing industries, relevant environmental regulatory and policy making authorities.

Stakeholder considerations: competitiveness and stability.

7.14.5.4 Data characteristics

This AI system ingested three plus years of process data and sensor data regarding plant operations from temperature, rpm, torque and pressure sensors that are used to set the variable frequency device (VFD) and flow control valves (FCV) setpoints.

7.14.5.5 KPIs

Accuracy: the accuracy of the predictive model is very highest achieving values around 98 %, extending to the setpoints to obtain predicted the values.

Reduction in calibration time: the amount of time saved from manually setting the calibration.

7.14.5.6 Features of use case

Task(s): prediction

Level of automation: conditional automation

Method(s): random forest prediction, one hot encoding, cross validation, normalization

Platform: application server: 64 GB RAM/ 16 Core / 500 GB HDD; data server: 128 GB RAM/ 16 Core, 3 TB HDD

7.14.5.7 Threats and vulnerabilities

Challenges to accountability, security threats

7.14.5.8 Challenges and issues

None identified.

7.14.5.9 Trustworthiness considerations

The use case addresses different aspects of the trustworthiness in AI systems, remarking by the AI provider aspects related to the sustainability and environmental protection, through the investing in scientific research and innovation as the way to achieve a sustainable model.

A relevant purpose is to improve the efficiency performance of previous developments, by the provider, through the development of technology based on AI, and aspects related with a predictive maintainability of the systems and operational calibrations across the use specific modules to achieve flow rates saving lost time and automation of tasks. Taking into account challenges related with the security of the system but also the accountability of the final customers and providers with the responsible use of the technology.

7.14.5.10 Use of standards and standardization opportunities

According to the level of autonomy of the AI system of conditional automation, it is a typical scenario with a high degree of automation based on AI.

In this particular use case of improvement of performance efficiency of the smart-sensor, the following issues and opportunities can be identified to be addressed by standardization, depending of the operational stage:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk impact assessment.

7.14.6 Improvement of productivity of semiconductor manufacturing (use case 82)

7.14.6.1 Objectives

Analysis of data taken from production equipment and improvement of productivity based on the analysis.

Cost reduction of semiconductor manufacturing.

7.14.6.2 Narrative

In modern semiconductor manufacturing, a huge amount of data is gathered and used to improve yields. However, it is difficult even for skilled engineers to promptly achieve the improvements by means of manual analysis because of the complexity of the production process and the scale of the data. In operations in the semiconductor factory, where more than five thousand pieces of equipment are running, and two billion records of data are created daily, it is difficult to secure enough engineers to resolve the problems that arise during production. This use case tackled the issue with AI technology including machine learning. The endeavour resulted in an improvement in productivity through stable quality based on semi-automated data analysis.

This use case consists of the following themes:

— Support for analysis of cause of failure based on wafer map patterns: at the final stage of semiconductor manufacturing, each chip on a wafer is tested and a pattern of the failed chips distributed on the wafer is produced. Analysis of the cause of failure is carried out based on the pattern and the history of usage of manufacturing devices. The analysis is supported by the following four technologies:

  — clustering of wafer map patterns: clustering of the wafer map patterns is carried out in order to grasp an overview of the occurrence of the failure. Because there are 200 thousand wafers per month, a fast-clustering algorithm is required to promptly provide information to engineers. Making use of scalable k-means++, the clustering process is 72,5 times faster than the previous method.

  — cause estimation based on pattern mining: if failure of a particular manufacturing device frequently occurs in the history of a wafer belonging to a wafer map cluster and failure of the device seldom occurs in the history of other wafers, then the device is likely to be the cause of the failure. The candidates for the cause of the failure and their likelihoods are calculated based on the number of occurrences of the combinations of the devices promptly counted by a pattern mining algorithm FP-Growth and ranking through chi-squared test.

  - wafer map classification based on CNN: a wafer map is classified into registered typical wafer maps in order to monitor the recurrence of the failure. The classification accuracy (F1 score) with SVM was 0,898. Making use of CNN, the accuracy is improved to 0,95.
  - web portal for yield analysis: the information provided by the above technologies is shown in a web portal. The portal has improved the automatic classification of scanning electron microscope (verage analysis time from 6 h to 2 h.
- Automatic classification of scanning electron microscope (SEM) images of defects: Tests of wafers are carried out not only at the final stage of the production but also between processes, where the result of the previous processes is checked. One of the tests is classification of images of microscopic aspects of the defects observed by SEM. Thirty thousand images are taken daily. It is an important test because the class of a defect can provide valuable insight for cause estimation. Previously, the classification was carried out semi-automatically by an engineer with a tool with classification function. However, the human workload was relatively high because the tool's ability was quite limited. Making use of CNN, the number of defect categories that are automatically classifiable has dramatically increased. Now the automation ratio is 83 %, improved from 49 %.
- Analysis of cause of variation of quality characteristic value: In this use case, the cause of the variation of a quality characteristic value is identified and the yield is maintained by countermeasures. For quick identification, various data including process parameters and sensor measurements from a manufacturing device are stored in a database. Therefore, the number of attributes becomes huge by the time of the completion of production. It is not uncommon for the number of attributes to be much greater than the number of products to be analysed, sometimes by several orders. Making use of Lasso regression for data with 23 600 attributes and 303 products, a regression model predicting a quality characteristic value has been built, with automatic feature selection. Engineers' cause identification tasks are also supported by a network diagram that visualizes the causal structure of the selected features. As a result, the average analysis time is improved to one day from seven days.

7.14.6.3 Stakeholders and stakeholder considerations

Stakeholders: AI producers related with semiconductor manufacturing companies.

Stakeholders considerations: competitive edge based on manufacturing cost reduction.

Business continuity based on the fewer number of required data scientists.

7.14.6.4 Data characteristics

This use case uses the various data including process parameters and sensor measurements from a manufacturing device are stored in an internal database. Making use of Lasso regression for data with 23 600 attributes and 303 products, a regression model predicting a quality characteristic value has been built, with automatic feature selection.

Source: detected by manufacturing devices.

Variety: the structured data.

Velocity: the data are increased daily.

Quality: high.

7.14.6.5 KPIs

Accuracy of wafer map classification: how accurately the recurrence of a failure is detected?

Time to identify the cause of failure: how quickly the cause of a failure is identified?

Accuracy of defect classification: how accurately the defect SEM images are classified?

Accuracy of feature selection: how accurately important features to quality characteristic values are selected?

7.14.6.6 Features of use case

Task(s): recognition, prediction, optimization, interactivity, recommendation

Level of automation: partial automation

Method(s): clustering, pattern mining, CNN, web portal, Lasso regression

Platform: PC cluster with GPU

7.14.6.7 Threats and vulnerabilities

Delay of the analysis tasks caused by inaccurate AI outputs.

Delay of countermeasure deployment caused by the fact that the physical model of a failure is unknown.

7.14.6.8 Challenges and issues

Guarantee of the correctness of analysis by AI.

Automatic physical model building for a failure.

7.14.6.9 Trustworthiness considerations

The use case addresses different aspects of the trustworthiness in AI systems related with the improvement of efficiency performance of manufacturing processes based on the implementation of different methods and ML algorithms to analyse images for validate and verified the quality parameters of semiconductors. The main improvement addresses productivity and robustness issues, trying to implement this improvement to save production costs.

Other relevant challenges acquired with the final client include accountability, security or reliability and a responsible use and commercialization of the technology.

7.14.6.10 Use of standards and standardization opportunities

According to the information provider, the use case has a medium degree of automation, with use of ML algorithms like CNN or pattern recognition with clustering methods for image treatment, drawing a mixed scenario with processing based on AI under the expert human oversight. Disguising the operational stages of acquire, process and apply in the manufacturing process, it is possible to identify the following areas of improvement across standardization methods:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency and specific issues as standardization data formats of manufacturing devices;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk-impact assessment. Also specific issues related with the standardization of results and outputs related with manufacturing devices.

7.14.7 AI decryption of magnetograms (use case 104)

7.14.7.1 Objectives

Oil and gas transportation. AI system to quickly identify defects during the quality assurance process on a field pipeline.

Detection of internal defects (e.g. pits, ulcers).

Detection of structural elements (e.g. welds, bends).

7.14.7.2 Narrative

A solution has been developed that allows for the detection of internal defects and structural elements.

In the territory of the Russian Federation, there are tens of thousands of kilometres of small diameter production pipelines under varying degrees of condition facing varying numbers of internal defects (e.g. pits, ulcers) and structural elements (e.g. welds, bends).

There are in-tube flaw detectors that allow the signal from the magnetometer sensors to be read. These robots are not widely used due to the speed of data interpretation.

Automation of the recognition of structural elements and defects would reduce the pipeline diagnostics process by at least 160 ×.

7.14.7.3 Stakeholders and stakeholder considerations

Stakeholders: AI customers of oil and gas manufacturing companies, relevant environmental regulatory and policy making authorities.

Stakeholder considerations: decision speed

7.14.7.4 Data characteristics

Description: data from 64 robot sensors

Source: flaw detector

Type: raw data, transformed into .csv

Volume (size): 60 GB

Velocity: batch

Variety: different source

Variability (rate of change): static

Quality: low

7.14.7.5 KPIs

Coverage for welds detection: detection accuracy’s ideal target is 95 %. The reference of this KPI to mentioned use case objectives is to improve accuracy.

Coverage for defects detection: detection accuracy’s ideal target is 100 % with 50 % defect depth. Target is 90 % with 30 % defect depth. The reference of this KPI to mentioned use case objectives is to improve accuracy.

7.14.7.6 Features of use case

Task(s): recognition

Level of automation: conditional automation

Method(s): machine learning, classic computer vision

Platform: flaw detector

Topology: trees, random forest

7.14.7.7 Threats and vulnerabilities

Condition of the flaw detector

7.14.7.8 Challenges and issues

To achieve high level accuracy in recognizing defects and welds.

To reduce the processing time of magnetograms.

7.14.7.9 Trustworthiness considerations

There are addressed different quality characteristics related with the trustworthiness of the implemented solution. Aspects as the improvement of efficiency performance about the capacity or time behaviour about the maintainability and coverages of the solution are developed by the provider of the solution.

Also are relevant environmental challenges acquired to be proactive to possible environmental risks, taking into account trustworthiness issues as the accountability, security or reliability of the company with a responsible application of the technology to avoid possible environmental disaster related with oil spills.

7.14.7.10 Use of standards and standardization opportunities

According to the information provider, the use case has a medium degree of automation, with use of ML algorithms based on tree decision and random forest, and a hybrid architecture. It is possible to identify the following improvement areas across the standardization:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk-impact assessment.

7.14.8 Analysing and predicting acid treatment effectiveness of bottom hole zone (use case 110)

7.14.8.1 Objectives

Mining of oil and gas; digital assistant for analysing and predicting the effectiveness of acid treatments of the bottom hole zone.

Predict the effectiveness of acid treatments of the bottom hole zone

7.14.8.2 Narrative

Predict the effectiveness of acid treatments on the bottom hole zone.

Currently, a long and subjective selection of candidate wells for acid treatments is being carried out.

An application with mathematical models for automating statistical analyses and predicting the technological and economic efficiency of acid treatments of the bottom hole zone of the well in the form of additional oil and well production.

The ranking of wells according to the degree of effectiveness of acid treatment of the bottom hole zone.

Determining the significance of various factors on the regression model for the field.

The goal is a convergence of the obtained forecast of the mathematical model with historical data of at least 80 %.

7.14.8.3 Stakeholders and stakeholder considerations

Stakeholder: AI customers of oil and gas manufacturing companies, relevant environmental regulatory and policy making authorities.

7.14.8.4 Data characteristics

Description: data from different well sensors

Type: structured data, .csv

Volume (size): 100 MB

Velocity: real time

Variety: different source

Variability (rate of change): static

Quality: position updates can be incomplete

7.14.8.5 KPIs

Coefficient of determination: prediction accuracy's ideal target is 0,9. The reference of this KPI to mentioned use case objectives is the prediction of future outcomes.

7.14.8.6 Features of use case

Task(s): prediction

Level of automation: assistance

Topology: trees, random forest, boosting

7.14.8.7 Threats and vulnerabilities

None identified.

7.14.8.8 Challenges and issues

To achieve high level accuracy of prediction efficiency of acid treatments

7.14.8.9 Trustworthiness considerations

There are addressed different quality characteristics related with the trustworthiness of the implemented solution. Aspects as the improvement of efficiency performance about the capacity or resource utilization with the target to achieve prediction improvements in implemented treatments.

Also are relevant environmental challenges acquired to be proactive to possible environmental risks, taking into account trustworthiness issues as the accountability, security or reliability of the company with a responsible application of the technology to avoid environmental disasters.

7.14.8.10 Use of standards and standardization opportunities

The use case application presents a hybrid architecture model, with an automation degree of assistance, and with an acquisition and processing pipeline where there are recovered lithographic parameters to be processed by ML algorithms based on trees decisions and random forest. This scenario can provide the following areas of improvement through the application of standardization processes:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk-impact assessment.

7.14.9 Automatic classification tool for full size core (use case 112)

7.14.9.1 Objectives

Oil and gas exploration, classification of rock types, oil saturation, carbonate and fracture according to core images.

Classification of rock types.

Classification of oil saturation.

Classification of carbonate.

Classification of fracture according to the core.

7.14.9.2 Narrative

A solution has been developed that allows for the classification of rock types into four classes. This resulted in an 80 % reduction in core image analysis.

To describe the core of an exploratory well, three specialists are allocated for a period of up to two weeks with travel expenses. The results of the description are subjective and can contain conflicting positions of experts. Automation of the process of classifying rock types, saturation, carbonate and rock layer degradation by daylight and ultraviolet images using machine learning mechanisms can reduce the lithotype typing time to three days.

7.14.9.3 Stakeholders and stakeholder considerations

Stakeholders: AI customers of oil and gas manufacturing companies, relevant environmental regulatory and policy making authorities.

Stakeholder considerations: decision speed

7.14.9.4 Data characteristics

Description: DL and UV core photos

Source: UT scanning instrument

Type: photo

Volume (size): 50 GB

Velocity: batch

Variety: single source

Variability (rate of change): static

Quality: middle

7.14.9.5 KPIs

Coverage: rock type accuracy's ideal target is 80 %. The reference of this KPI to mentioned use case objectives is to improve accuracy.

Splits: detection of splits in the rock with an accuracy of 2 cm. The reference of this KPI to mentioned use case objectives is to improve efficiency.

7.14.9.6 Features of use case

Task(s): recognition

Level of automation: assistance

Method(s): machine learning, classic computer vision

Platform: camera

Topology: trees, random forest

7.14.9.7 Threats and vulnerabilities

Quality of images received from special equipment.

7.14.9.8 Challenges and issues

To achieve the same level of accuracy of recognition of rock types as lithologist experts.

To minimize the set of laboratory tests due to visual recognition of rock types and their parameters from core images.

7.14.9.9 Trustworthiness considerations

There are addressed different quality characteristics related with the trustworthiness of the implemented solution. Aspects as the improvement of efficiency performance about the capacity or resource utilization with the target to achieve prediction improvements in implemented treatments.

Also are relevant environmental challenges acquired to be proactive to possible environmental risks, taking into account trustworthiness issues as the accountability, security or reliability of the company with a responsible application of the technology to avoid environmental disasters.

7.14.9.10 Use of standards and standardization opportunities

According to the information provided, the use case presents a low degree of automation with a level of assistance develop in a hybrid architecture, where UT scanning instrumentation, provide information to be acquired, processing and applied, permitting identify the following areas of improvement across standardization processes:

— acquire: general issues such as choice of data frequency, establishing bias in input data, minimum size for training data sufficiency;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk-impact assessment.

7.14.10 Collaborative AI to assist workers with production and assembly in factories (use case 179)

7.14.10.1 Objectives

To create autonomous, dexterous and market-oriented robot prototypes that provide new ways of automating the manufacturing and processing of flexible and deformable materials.

The system can map and recognize the environment, as well as selected key objects.

Objects can be tracked for the duration of their lifetime.

Workers can control the robot using a predefined set of gestures.

7.14.10.2 Narrative

Robotic hands are expected to be used in the production chain to assist workers in routine tasks, performing repetitive processes such as product inspection, assembly, sorting, or distribution in conveyor belts and assembly lines. The robot is necessary to detect key objects in conveyor belts or other types of assembly lines and track each object for the duration of the assembly process. Different use cases are proposed by project partners during development to perform tasks like: packing of poultry pieces, handling pillows during assembly, improving manual bagging of insoles, assisting in the assembly of blowing machines, and handling of flexible circuit components. The system can be controlled by operators using static and dynamic gestures, facilitating collaboration between them. The robots are designed so that they can be easily reprogrammed and repurposed for different scenarios.

7.14.10.3 Stakeholders and stakeholder considerations

Workers at industrial facilities are expected to be directly benefited from the system, by reducing workload as well as the risk of workplace hazards.

Indirectly, everyone in the company benefits as the system is necessary to have an impact on efficiency by cutting down on processing and production times, ultimately improving the company's competitiveness and profitability.

7.14.10.4 Data characteristics

Data source: the different beneficiary partners involved in the project, will provide videos of their facilities, procedures and key objects to train and test the system. They will also provide the system with real-time data of their processes that the system is designed to process, analyse, understand and act upon, as well as control gestures to interact with the robot.

Several use cases have initially been proposed for the design and development of the system:

— packing of poultry pieces;

— handling pillows during assembly;

— improving manual bagging of insoles;

— assisting in the assembly of blowing machines;

— handling of flexible circuit components.

Additional data include output from the system sensors to provide information about the robotic hand's eight degrees of freedom (translation and rotation in x, y and z, as well as pan and tilt coordinates).

Data varieties: unstructured data from day-to-day operations; semi-structured data from control gestures; and structured data from the system sensors.

Data velocity: once deployed in production, the system will receive data and react to it in real time.

7.14.10.5 KPIs

Environment reconstruction: 3D reconstruction with an error < = 0,5 cm.

Object detection and characterization: detection over 95 % for all objects / Surface reconstruction with error < = 0,5 cm.

Predictive tracking: MOTA over 95 %.

Non-verbal communication: Static and dynamic gesture detection rate over 95 %.

7.14.10.6 Features of use case

Task(s): recognition, classification, mapping, tracking, control

Level of automation: assistance

Method(s): Computer vision, deep learning, image processing, control

Platform: robotic hand, sensors and a central processing unit equipped with one or several graphics processing unit(s) (GPU) and a set of algorithm libraries to process the information.

Topology: on-site system consisting of a robotic hand with a set of sensors and actuators connected to a central unit. Cloud-based federated knowledge base.

7.14.10.7 Threats and vulnerabilities

System malfunction

Unforeseen events (e.g. detection of unknown objects, objects or workers crossing the trajectory of the robot)

Security threats

7.14.10.8 Challenges and issues

Operation errors caused by wrong or incomplete definitions of either the assembly (or production) process, or the work environment on premises.

Sensor accuracy errors.

Errors derived from human interactions with the system.

7.14.10.9 Trustworthiness considerations

Robotic hands are one of the most relevant applications of AI in manufacturing sectors. This use case of AI systems highlights significative trustworthiness issues that are desired to be developed, in order to ensure the sustainability of the model development as the user controllability of the systems, but also to achieve the highest degree of autonomy with highest ratios of performance efficiency, in aspects related with time behaviour, capacity and resource utilization.

Scenarios combining machines with highest degree of autonomy together with the oversight of human external agents suppose an important challenge from the point of view of security and safety in order to ensure the robustness of the equipment but also the integrity of line operators. Also there are relevant aspects related with the maintainability of the system such as the repairability and reusability that can be addressed with a modular design to be able to achieve highest levels of trustworthiness AI systems and also assuring the sustainability of the development model.

7.14.10.10 Use of standards and standardization opportunities

The use case is a typical smart-manufacturing application, where it is possible to distinguish a mixed scenario with human external oversight over machines with highest degree of conditional automation. The use case presents a hybrid architecture, where it is possible to distinguish the following standardization opportunities across the life cycle and acquisition pipeline of the application:

— acquire: general issues such as creation of benchmark data set, choice of data frequency and synchronization of acquisition, establishing bias in input data, minimum size for training data sufficiency;

— process: general issues such as the criterion for training, criterion for validation, criterion for retraining or the implementation in existing systems;

— apply: general issues such as explainability of results, acceptable output for commercialization, fail safe mode of operation against biases, safety, health and environment impact or risk-impact assessment, standardization events, fault and error code types.

7.15 Media and entertainment

7.15.1 Video on demand publishing intelligence platform (use case 58)

7.15.1.1 Objectives

Predictive maintenance platform for a video on demand content preparation process.

The goals of the project are:

— process fault comprehension;

— fault prediction;

— fault recovery through a recommendation engine;

— productive interaction between the fault prediction and recovery recommendation engines for proactive process maintenance.

7.15.1.2 Narrative

An E2E platform was developed in order to achieve accurate fault prediction with machine learning and useful recovery action recommendation using Reinforcement Learning.

The fault prediction engine allows to simulation of the outcome of a process instance. The machine learning engine predicts the outcome using:

— the current state of the target applications;

— the current state of the target IT systems;

— the recent state of the target applications (20 min);

— the recent state of the target IT systems (20 min).

The ML models give insights on the most important variables in predicting the outcome. These variables can point directly to the error cause or be related to it.

The recovery recommendation engine is able to:

— use the ML models to find a data-driven optimal action;

— incorporate user feedback to add custom actions;

— incorporate user feedback in order to further improve its recommendation strategy.

Model- and user-defined actions challenge each other in order to provide the current best action. User feedback is incorporated in a reinforcement learning fashion.

7.15.1.3 Stakeholders and stakeholder considerations

Stakeholders: the factory, predictive maintenance platform developers, maintenance staff.

Stakeholders considerations: the maintenance staff consider reducing the cost of the fault recovery, accurate fault prediction, ML engine processing time.

7.15.1.4 Data characteristics

Description: the data used in this use case include current and recent (20 min) state of applications and IT systems, user feedback.

Source: applications and IT systems.

7.15.1.5 KPIs

Error frequency: error frequency to be reduced. The reference of this KPI to mentioned use case objectives is productive interaction between the fault prediction and recovery recommendation engines for proactive process maintenance.

Lateness: number of time-consuming tasks to be reduced. The reference of this KPI to mentioned use case objectives is productive interaction between the fault prediction and recovery recommendation engines for proactive process maintenance.

Model AUC: KPI to monitor the classification quality of the models. The reference of this KPI to mentioned use case objectives is fault prediction.

User feedback: user feedback is used to tune the recommendation engine. The reference of this KPI to mentioned use case objectives is fault recovery through a recommendation engine.

7.15.1.6 Features of use case

Task(s): the fault prediction engine and the fault recovery recommendation engine work in synergy: the first yields a fault probability based on the current and recent state of applications and IT systems, providing the latter with a recommended recovery action. This action is challenged by other user-defined actions in the recommendation engine. The platform suggests the winning action to the user. The user can then give feedback, allowing the recommendation engine to improve in a reinforcement learning fashion.

Level of automation: high automation

Method(s): random forest, variable importance evaluation, reinforcement learning.

Platform: virtual machines

Topology: machine learning, reinforcement learning, recommendation engine, environmental logs, application log, next best action, process mining.

7.15.1.7 Threats and vulnerabilities

None identified.

7.15.1.8 Challenges and issues

The machine learning engine processing time had to be very short.

7.15.1.9 Trustworthiness considerations

None identified.

7.15.1.10 Use of standards and standardization opportunities

None identified.

7.15.2 AI system for promoting DX in customer attraction services at a museum (use case 178)

7.15.2.1 Objectives

To promote digital transformation (DX) with improved productivity and added value in customer attraction services at a museum by introducing AI technologies.

7.15.2.2 Narrative

The museum installed dedicated terminals and distributed IC cards or QR codes to visitors for interactively collecting questionnaires with visitor’s IDs and recommending through visitors touching their card at the entrance, each exhibition booth and the exit. It constantly collects big data generated at the service site and utilizes it to improve the productivity and added value of its customer attraction services.

In addition to the features such as “AI vending machine” or digital signage and interaction with visitors, the AI system can learn from the big data the structure among key goal indicators (KGIs), KPIs and key factors for success (KFSs) in terms of a Bayesian network, which can be used for DX of the customer attraction services. It can also analyse and visualize the big data for science communicators and managers, respectively. For example, it can execute clustering of the visitors based on their actual routes and questionnaires. As a result, science communicators and managers improve their navigation skill and the contents in the museum.

7.15.2.3 Stakeholders and stakeholder considerations

Visitors as both customers of the museum service and AI customers.

Science communicators (museum staff) as both providers of the museum service and AI customers.

Museum managers as both providers of the museum service and AI customers.

AI research centre of a national institute as both AI producer and AI provider.

7.15.2.4 Data characteristics

Source: visitors via their moves inside the museum, interactions with the AI system (“AI vending machine” component) and questionnaires.

Variety: structured (tracking visitor’s feedback data) and unstructured data (evaluation data from stakeholders).

Velocity: data can be real-time when the museum is open.

Variability: what data to collect and its format can be changed when the overall DX process is reviewed.

Quality: high enough

Protected attributes: collected data are anonymous.

7.15.2.5 KPIs

Effectiveness for DX: how effective the AI system’s outputs can be used for promoting DX.

Visitor satisfaction: customer experience with the AI system’s interaction with visitors.

7.15.2.6 Features of use case

Task(s): interactivity, recommendation and probabilistic inference

Level of automation: assistance

Method(s): Bayesian network, probabilistic latent semantic analysis. Human computer interaction, information recommendation using user models.

Platform: servers for running AI programs and storing data. On-site system for taking care of interaction with visitors.

Topology: network of servers and on-site system

7.15.2.7 Threats and vulnerabilities

None identified.

7.15.2.8 Challenges and issues

Since the overall DX process is continual improvement, what data to collect and analyse and its format are under periodic review and subject to change.

7.15.2.9 Trustworthiness considerations

There are some concerns on the part of interaction with visitors, such as inappropriate recommendation due to poor user estimation.

7.15.2.10 Use of standards and standardization opportunities

AI for facilitating multi-stakeholder processes.

7.16 Public sector

7.16.1 AI ideally matches children to daycare centres (use case 7)

7.16.1.1 Objectives

Assignment pattern that satisfies applicants' complex requirements.

To automatically determine the assignment pattern that is expected to fulfil the preferences of as many applicants as possible.

7.16.1.2 Narrative

The number of children on daycare centre waiting lists has becomes a social issue. Matching children to daycare centres while accommodating each family's preferences is time- and labour-intensive for local governments.

The basic goal of daycare admissions screening is to satisfy the preferences of applicants according to the priority ranking of children in consideration of the number of places in each daycare centre. In addition, each local government can incorporate more complex requirements, such as applicants who want their siblings assigned to the same daycare centre and who want siblings assigned in the same period, in order to increase the satisfaction of applicants. Saitama city government has eight requirements concerning sibling admissions as well as the timing of the siblings' admissions. The screening rule thus became more complex, and consequently there are cases where multiple assignment patterns can fulfil the rule or no patterns fulfil the rule. This means the city officials are required to take a long time to carefully determine the assignment of applicants to be absolutely sure that the relevant rules have been correctly fulfilled.

This AI system has made it possible to match children to daycare centres, meeting as many preferences as possible, following a priority ranking. This is done by modelling the dependency relationships of complex requirements, including parents who prioritize siblings going to the same daycare centre, or parents who do not mind if their children go to different daycare centres as long as both children get a seat, using a mathematical model based on game theory, which rationally resolves the relationships between people having differing values. When this technology was evaluated using anonymized data from about eight thousand children in the city of Saitama, Japan, it successfully calculated an optimal assignment result in just a few seconds.

7.16.1.3 Stakeholders and stakeholder considerations

Stakeholder: city officials, daycare centres and applicants.

Stakeholder considerations: maintaining fairness of matching results, reducing the burden of seat assignment tasks, enabling the smooth return of parents to the workplace.

7.16.1.4 Data characteristics

The data used in the AI system is provided by the city official. It includes 8 000 applications of the daycare. The application data are anonymous and include the request of siblings' application, working conditions of parents, childcare conditions, and whether or not parents have nursing care.

Source: city officials.

Variety: the structured text or database.

Velocity: the data can be updated frequently or not be updated.

Quality: high.

7.16.1.5 KPIs

Accuracy: the matching rate of assignment. Automatic assignment.

Time: the computation time to find an optimal assignment. Time reduction.

7.16.1.6 Features of use case

Task(s): optimization

Level of automation: partial automation

Method(s): game theory

7.16.1.7 Threats and vulnerabilities

It is desired to consider the privacy of applications in this use case. The data that make the parents or the family possible to be identified are desired to be removed.

7.16.1.8 Challenges and issues

Challenges: determine an optimal assignment pattern instantly and fairly depending on unique and complex rules in each local government.

Issues: long calculation time is required in the case of a large number of children and siblings.

7.16.1.9 Trustworthiness considerations

Supporting working parents.

Resolving the problem of children waiting for day care.

7.16.1.10 Use of standards and standardization opportunities

Need to consider unique requirements for assignment rules in each local government.

7.16.2 Open spatial data set for developing AI algorithms based on remote sensing (satellite, drone, aerial imagery) data (use case 122)

7.16.2.1 Objectives

Analytical services for automatic detection of changes in the state of ground surface objects for administrative, government and social purposes in different use cases, such as:

urban monitoring: cadastral data, land management, estimation of the living population, etc;

— emergency mapping: estimation of disaster damage;

— security and risk management monitoring of protected zones (powerlines, railroads, pipelines): detection of vegetation growth, control of safety, etc.

The growth of the market for geo-analytical cloud services based on remote sensing data and AI technologies; open benchmark data sets for the research and development community; and bringing the power of AI and the global coverage of remote sensing imagery closer to the people.

7.16.2.2 Narrative

Despite the increasing number of data sets and competitions in remote sensing data science [e.g. the provider of very-small-aperture terminal (VSAT) satellite-based data network services], there is still a lack of geographical diversity, of training classes, and of interoperability of data sets.

The proposed approach is to be extended to different types of remote sensing data and application domains based on classification of the natural and man-made objects that have a clear interpretation either in satellite or aerial imagery.

7.16.2.3 Stakeholders and stakeholder considerations

Stakeholders: community

Stakeholder considerations: trustworthiness, safety and competitiveness.

7.16.2.4 Data characteristics

The data used in this use case are the remote sensing (satellite, drone, aerial imagery) data.

7.16.2.5 KPIs

Georeference: maps (e.g. geographic database) for data labelling require objects' coordinates. The reference of this KPI to mentioned use case objectives is that simply annotated photos are not enough.

Time series: emergency mapping requires the detection of changes in residential infrastructure analysis before and post-event images. The reference of this KPI to mentioned use case objectives is to observe places in dynamic and calculate comparative indicators.

Cartographic styled labelling and classification: competition of network with manual mapping. The reference of this KPI to mentioned use case objectives is that the maps make an abstracted interpretation of earth observation images.

Advanced classification: the help of different bands combination. The reference of this KPI to mentioned use case objectives is thematic interpretation of satellite imagery.

Open API and web tools: integrate both mapping and data science approaches in ways demanded by users. The reference of this KPI to mentioned use case objectives is to access and preview.

7.16.2.6 Features of use case

Task(s): recognition

Level of automation: partial automation

Platform: high-performance computing (HPC)

7.16.2.7 Threats and vulnerabilities

New privacy and security threats, challenges to accountability

7.16.2.8 Challenges and issues

There is no standard or criteria regulating the process of labelling (manual or automatic) remote sensing [satellite, drone or unmanned aerial vehicle (UAV)] images with a geographic reference. Development of such a standard is vital to AI algorithms for guaranteeing the quality of training data and for testing and benchmarking.

The following criteria are considered that the perfect data set collection of electro-optical (EO) imagery is expected to match:

— georeference: annotated photos are not enough. Maps for data labelling (e.g. geographic database) require objects' coordinates;

— time series: to observe places dynamically and calculate comparative indicators. The main application is "emergency mapping" that requires the detection of changes in residential infrastructure analysis of before and post-event images;

— cartographic-styled labelling and classification: maps make an abstracted interpretation of earth observation images; it therefore can be considered that the previous approach of labelling images with boxes does not satisfy the criteria for accurate image segmentation and will not work. For neural networks, it is now necessary to compete with manual mapping, and to calculate its accuracy at least some ground truth that looks like a map. At the same time there are many other sources beyond the EO imagery that can be useful for mapping, such as point of interest (POI), and collecting field works in order to accumulate addresses. At this moment our goal is to compare ML methods with the information that can be extracted by a cartographer using only optical bands of imagery and some geographic information system (GIS) software. For such purposes, it is proposed the basic classifier that is part of training and testing data sets;

— multispectral: next, it is proposed to extend this approach to advanced classification, which is comparable to thematic interpretation of satellite imagery with the help of different bands combination. That is why the proposed classifier includes classes that require even more specific training and non-optical bands for better recognition.

Providing open API and web tools to access and preview data sets. Despite the data set collection representing structured data, it would be much more capable for further and updated use based on the standards for interoperability of geodata. In this use case, it tried to join both mapping and data science approaches in the way seeing new tools and services demanded by users. For many users from the data science community, maps and remote sensing are becoming just one of the sources of information that is necessary to be structured and classified. And for many mappers that are involved in the process of geodata interpretation and classification, the map itself is the perfect tool to interact with the data, no matter whether implemented in a python notebook or loaded in a desktop GIS application.

7.16.2.9 Trustworthiness considerations

Global extension of this technology brings society new possibilities of situational awareness and digital instruments for natural and man-made resource management.

7.16.2.10 Use of standards and standardization opportunities

None identified.

7.16.3 Smart city (use case 165)

7.16.3.1 Objectives

This use case offers insights into the development of smart cities in a regional context, the use and implementation of smart information system (SIS) in urban environments. The use case evaluates the ethical issues in the literature and how they contrast and diverge from those faced by professionals in practice.

This use case offers practitioners, policymakers, smart city organisations, and private ICT companies, interesting observations about ethically responsible approaches towards SIS implantation in smart city projects.

Evaluation of the ethical issues in the literature on the use of SIS in smart cities identifies a number of concerns that need to be addressed. The concerns from literature are evaluated by four organisations working on smart city projects to determine if the ethical issues found in the literature correlate with those in real-life examples.

The use case presents four advanced smart city projects to demonstrate how the cities are implementing SIS in practice.

7.16.3.2 Narrative

This use case unpacks the ethical challenges of AI by looking at four cities: citizens' complaints AI, parking permit chat-bot, platform for data exchange, and a project with an open-source algorithm.

Smart cities are in their infancy, which means that availability and accuracy of data, and as a result the accuracy of recommendations remains an issue.

Data are a potentially helpful tool for citizens and planners alike to regain control and access of information within their respective cities. Consent, transparency and data ownership are prominent ethical considerations, with focus on citizens regaining control over their own data.

Collaboration is at the heart of a smart city. A public-private model facilitates both the business development- and the citizen-engagement sides of the smart city. A bottom-up approach is the most effective way to ensure that smart cities work and are used by citizens.

7.16.3.3 Stakeholders and stakeholder considerations

Stakeholders: policy makers, smart city organisations, ICT companies, city planners and governance, and citizens.

Stakeholder considerations:

Accuracy of SIS and bias: determining the accuracy of information and bias are both crucial components to the case study's validity. The chat-bot is still in its early stages of implementation. The Amsterdam case raises a similar issue. The system is still in the developmental stage and is being used in situations that will have minimal negative impacts on the lives of Amsterdam citizens.

Availability and accuracy of data: the city data exchange was originally conceived as a general platform based on the idea that data are interchangeable, but in practice, the specificity of data-usage undermined this original idea, delivering results was not straightforward.

Data availability is a fundamental requisite for the success of smart city projects, but in practice, this can be difficult. Municipalities shared the need to be able to retrieve and use their own training data for their algorithms, rather than relying on third-party vendors for this.

Economics and inequalities: smart cities are about collaboration between corporations and municipalities and not simply the sale of services according to a standard business model. Apart from the need to overcome bureaucratic hurdles also objectives can differ. Investors are interested in the economics while municipalities can have other objectives such as sustainability. Public entities focus on providing value to citizens whereas this can have been a hindrance in the economic development of projects.

Privacy and data ownership: smart city projects aim to allow citizens greater control over their personal data and to ensure their privacy. Privacy of citizens and GDPR compliance of suppliers is an important requirement for smart city projects. Citizens have to provide informed consent on the use of the data, and the data are anonymized

Transparency and trust: there have to be a symbiotic relationship of trust between corporations, citizens and municipalities working on smart city projects. There is the need for transparency about the accessibility and availability of the city data to achieve trust from the citizens.

7.16.3.4 Data characteristics

The data used in this use case are the data of citizens and city government.

7.16.3.5 KPIs

The accuracy of recommendations of the AI systems.

7.16.3.6 Features of use case

Task(s): interactivity

Level of automation: conditional automation

7.16.3.7 Threats and vulnerabilities

Data ownership and control: there is a strong emphasis on the citizens owning and controlling their data and personal information and consenting to its use. However, it is also important to identify who has control over data in public-private smart city projects. Therefore, there is a need for transparency and involvement between partners in order to ensure a fair and equitable interaction within smart city projects. While private organisations can be best suited to control and manage SIS, there needs to be an understanding between partners and a strong degree of transparency in their relationship.

Inequalities: there is concern that the technology can replace humans in many areas of the smart city. Many people fear that SIS will replace customer service, driving and factory jobs within the coming decade.

Smart cities need an intellectual infrastructure to deploy SIS, becoming hubs for technological innovation and advancements, which can subsequently lead to a ‘brain-drain’ in rural areas. The most educated and prosperous citizens will be located in cities, which can have a dramatic effect on the education, prosperity, and growth of rural areas. SIS can increase digital divides and inequalities.

Privacy: there are many privacy concerns in the context of smart cities, such as the use of technologies that track movement, that scan bodies, and those that record and recognize audio. Protection of one’s property and physical space and overall privacy is very important to individuals.

Since smart cities are based on the collection and use of data through SIS technologies, safeguarding the data are crucial for the maintenance of the goal of a smart city to benefit citizens.

7.16.3.8 Challenges and issues

Economic pressure: most cities are far from reaching the desired benefits outlined in smart city agendas because they are still in the early stages of development. At the same time, monetary benefits are increasingly linked to efforts to become smart.

7.16.3.9 Trustworthiness considerations

Responsible innovation favours the inclusion of all relevant stakeholders in the development of major innovative projects such as smart cities. Many smart city initiatives are devised by SIS technology corporations and city governments, disregarding civic participation and civic input. Smart city initiatives can place a greater emphasis on technical fixes, instead of implementing political and social solutions to try to tackle urban issues