INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2026, VOL. 72, NO. 3, PP. 1-8
Manuscript received March 23,2026; revised June, 2026. doi: 10.24425/ijet.2026.157956

A cloud platform for the development of resilient and intelligent services for the public sector

Henryk Krawczyk, Piotr Orzechowski, and Bogdan Wiszniewski

***Abstract*—The ongoing progress of artificial intelligence enables the creation of advanced IT services in various areas of human activity. This paper examines the construction of intelligent cloud services for the analysis of digital documents used in the public sector. Examples of such services are provided. The use of a universal platform CAISE is proposed, which allows effective implementation of resilient and intelligent services. The general architecture of such a platform is presented, and their functionality is commented. Verification and validation elements of such platform are described in order to achieve the highest technological readiness level.**

***Keywords*—intelligent services; public sector; PaaS cloud architecture; resiliency metrics; evaluation procedures**

I. Introduction

PAPER documents, along with digital ones, as digitization progresses, form the basis for making appropriate decisions and performing necessary actions. Content analysis of digital documents is widely used in various human activities, particularly in the public sector. The latter involves institutions and organizations managed directly or indirectly by the state or local governments. It encompasses key areas of social and economic development that directly impact the quality of life of citizens. Limited financial and human resources require effective planning and allocation to maximize value for the society. Ensuring high efficiency and quality of public services is an additional challenge. Population aging, migration, and other changes in social structure significantly impact the demand for public services and require the public sector to adapt to changing realities. Innovative solutions must be utilized to meet growing public expectations. Successful digitalization of this sector translates into the sustainable development of public services, including in areas such as healthcare, education, and the justice system. This type of progress is essential for the smooth functioning of society and improved quality of life.

The CAISE (Cloud Artificial Intelligence Service Engineering) project [1], implemented by Gdansk University of Technology, within a framework of the IPCEI-CIS Program [2], is a unique form of support, exclusively for projects approved by the European Commission. The program aims to develop cloud infrastructure and services, which should play a key role in supporting Europe's continued digitalization and digital sovereignty. The technology developed in this program also addresses the growth of the edge computing market, where key issues include data privacy, cybersecurity, and interoperability. The research and development work will lead to the development of open-source software that will enable real-time, low-latency applications using distributed computing resources located close to users. Additionally, the solutions developed should help to effectively reduce the overall carbon footprint associated with the rapidly evolving global cloud infrastructure. As a result, a secure and easy-to-use, vendor-neutral alternative can be offered, democratizing access to edge computing and stimulating competition among companies operating in the EU cloud computing market. The CAISE project addresses these goals well by focusing both on dependability and energy efficiency.

The platform is based on the TASKcloud computing cloud, systematically expanded by CI TASK for several recent years. TASKcloud is gradually enhanced with additional open-source software components, creating advanced environments for the production and execution of intelligent services. In our solution cloud computing [3] and artificial Intelligence (AI) [4] components are largely used. Cloud computing is a data processing model that provides users with access to resources and services belonging to the provider, including computing power, memory and disk space, via the network. This solution corresponds to the Infrastructure as a Service (IaaS) model level. At a higher model level, namely the Platform as a Service (PaaS), tools and components of various kinds are offered to build specific platforms satisfying various client requirements. Both the TASKcloud and the CAISE platform are being developed using the DevOps methodology [5]. Besides, the Continuous Integration and Continuous Delivery (CI/CD) techniques [6] are included in this process, enabling ongoing development throughout the entire project and beyond. Due to such solution each development software process is a more resilient one [7]. It means that expert-level guidance through all phases of the process development could be provided and many best practices, including suitable testing scenarios, and design methodologies should be supplied [8]. One of well-known examples is machine learning Technology Readiness Level (TRL) framework [9], which defines a such principled development process to ensure robust, reliable, and responsible

This work was supported by project “Cloud Artificial Intelligence Service Engineering (CAISE) platform to create universal and smart services for various application areas”, No. KPOD.05.10-IW.10-0005/24, as part of the European IPCEI-CIS program, financed by NRRP (National Recovery and Resilience Plan) funds.

First and third Author are with Gdansk University of Technology, Faculty of ETI, Poland(e-mail: {hkrawk, bogwiszn}@pg.edu.pl).

Second Author is with Gdansk University of Technology, CI-TASK, Poland (e-mail: porzechowski@task.gda.pl).

products. It allows to create new transdisciplinary software engineering models [10].

AI makes traditional IT services replaceable by intelligent services using various Natural Language Processing (NLP) methods [11] and Machine Learning (ML) models [12]. Development of such intelligent services is a rather complex research and development undertaking because it requires extensive knowledge of AI and at the same time a powerful environment in which such a service could be deployed. The CAISE platform facilitates this task by delivering productive environments for building and executing such services. In general, SMEs specializing in specific IT solutions for the public sector use different production environments focused on a given and rather narrow application area. This means that their production effort in terms of investment costs and development time is often unnecessarily duplicated, especially when setting-up dedicated platforms. Our proposal to cope with that is the construction of a universal platform that would allow for designing intelligent services in various sectors of human activity. A key enabler in that regard has been the European IPCEI-CIS initiative, the most important undertaking, which the ultimate goal is to create a common European digital space, based on the cloud-edge continuum [21].

The CAISE platform will be available for use by various domestic SMEs providing IT solutions in areas such as public safety, public administration, transportation, energy, environmental protection, among others. It will offer cloud computing resource management services as well as intelligent text and image content analysis services. The implementation of intelligent cloud services represents in our view a significant product innovation. Integrating DevOps and CI/CD techniques and incorporating quality management procedures into these processes will ensure process innovation as well.

In this paper we will limit ourselves to presenting only a method for creating intelligent and resilient services based on the CAISE platform. Section II provides representative categories and examples of describing intelligent services which were developed with it so far. The next Section III gives the general functionality of the CAISE platform where some implementation problems are also included. Section IV defines main assumptions regarding platform validation and concrete examples of the validation procedures; this aspect of the project is particularly important due to the assumption of achieving by the CAISE platform the TRL-9 technology readiness level and its roll-out on the domestic cloud services market. Questionnaires were used to get feedback from platform users to confirm its benefits, and to consequently ensure its proper development course. In Section V we emphasize two novel concepts introduced by the paper relative to the current, literature on the subject whereas in Section VI we conclude the paper by indicating ways of improving the platform further, which is related to future R&D challenges identified by us during the project.

II. CATEGORIES OF CAISE SERVICES

Our main goal is to develop and implement the CAISE cloud platform on the domestic market, enabling efficient production of intelligent cloud services and applications by domestic SMEs – serving public sector entities in the field of IT solutions related to intelligent processing of digital document content in the Polish language. In this context, three selected IT companies were invited to participate in the project as subcontractors. They represent three public activity sectors mentioned before, i.e. justice, healthcare, and education.

The CAISE platform aims to develop intelligent document content analysis services utilizing the latest advances in deep machine learning, NLP, and image recognition, including methods for ensuring the quality of large-scale training data (Big Data), building and training large language models (LLM foundation models) [13], optimizing their hyperparameters and fine-tuning them using specialized data (LLM fine-tuning), and smoothing the model-generated content (LLM alignment) by end users. This will leverage high-performance computing. The implemented service pipelines enable development of solutions in both – the comprehensive AI, including text summary and abstract generation, document consistency verification, decision identification in descriptive documents, and historical analysis of actions described in a batch of documents, as well as the generative AI, comprising generation of formal (legal) letters, medical notes, or image content descriptions for the visually impaired persons.

Target users of the platform can develop specialized AI/ML services of two categories: core services trained on publicly available datasets to analyse a general digital document content analysis, and sector-specific ones, which are focused on the specificity of a given application sector and tailored for the particular customer needs. They can be implemented, deployed, and published on the CAISE platform, as well as ported as separate new products to other customer-owned environments or platforms.

For illustration of the capabilities of the CAISE platform, we briefly outline four sample core services already developed and tested:

(1) *Phrase location and identification service* [14]; such a service is crucial for tasks involving extraction of textual information about activities and events described in the analysed document. The concept of a "phrase" in the project is understood as fragments of text describing an activity or event with its arguments, such as the performer, the object of the activity, the beneficiary, etc. They correspond to verb phrases, but also to noun phrases, in which the noun is a gerund denoting an action. Automatic extraction of these phrases in the text in the form of "main element: arguments" is crucial for building intelligent applications for analysing activities described in the document text (e.g., to classify the status of the described activity or to examine its compliance with some relevant legal act), examining their chronology (e.g., to identify activities that should be performed in the required order), or generating a description of the required activity (e.g., generating a letter requesting the addressee to perform a given activity).

(2) *Anonymization/pseudo-anonymization service* [15]; it addresses a need to fully anonymize or pseudonymize values of attributes considered sensitive in medical data stored and/or exchanged in accordance with the DICOM standard. Medical imaging data are increasingly being stored in various cloud repositories: to conduct scientific research, plan national and regional health care policies, estimate and optimize health care expenditures, analyse the effectiveness of specific therapies, etc. The introduction of the European Health Data Space should further strengthen and organize this trend [22]. For example, medical data analysis including development of AI/ML models

for scientific research should comprise de-identification of sensitive data, as well as other not essential data. As part of this service, the developed method of anonymizing data in DICOM files (loss and permanent loss of sensitive data) is important for retrospective studies. In the case of prospective studies, the collected data accumulates over time. For example, this may involve collecting data from the same patient before and after therapy. Therefore, pseudo-anonymization removes sensitive data, which can be reconstructed based on a securely stored dictionary, which is stored in the encrypted form, for example, at the data provider's premises. In the project the CT Image IOD modality is investigated as a representative case study. However, the service is implemented in such a way that it can be scaled to other modalities.

(3) *Document Image Analysis (DIA) service* [16]; it enables the extraction of blocks of information stored in a raster image of a document, such as a school textbook page, followed by their classification and conversion. Extracted objects are classified into predefined categories, such as text, header text, images, image captions, page numbers, tasks, tables, footnotes, etc. Automatically marking these units in a digital image and assigning them appropriate labels/categories is crucial for implementing intelligent applications such as e-document converters (e.g., e-textbooks), automated test/classification generators, etc. The DIA service has been developed in the project for documents in Polish. It enables conversion of extracted graphic blocks into editable text strings, interpretive textual image descriptions, and more advanced table decomposition for further analysis. Thanks to the cloud technology, the service is scalable, fast, and efficient in processing large volumes of document page raster images, e.g. multi-page textbooks by multiple users simultaneously. Depending on the training sets used, it is possible to extract the content of scanned document images for various education and other purposes, e.g. images of tax forms or invoices. Its main functions include:

- *Layout analysis* examining arrangement of various elements on a page, allowing for better recognition of the document's template structure;
- *Document structure* recognition classifying various document elements, such as headings, tables, charts, captions, and sections – to reconstruct the logical structure of the document;
- *Optical Character Recognition* (OCR) extracting text from raster images. This is a fundamental element of DIA, allowing the conversion of text contained in an image into editable (and interpretable) digital text;
- *Image segmentation* dividing a document raster image into smaller elements, such as text blocks, images, tables, etc., essential for further analysis and content processing;
- *Data export* to a variety of formats, such as PDF, TEX, CSV, XML or JSON, for further processing in other systems.

(4) *Named Entity Recognition (NER) service* [17]; it allows for automatic identification and classification of text fragments relating to given categories, such as people, places, organizations, dates, amounts, products, events, etc. Marking these entities in the text and assigning them appropriate labels is crucial for implementing intelligent applications such as search engines, chatbots, data analytics services, extraction of all kinds of information, etc. The developed NER service can be easily adapted to diverse fields in law (e.g. law firms and bailiffs' offices, common courts, prosecutor's offices, etc.) or medicine (tomography, radiology, general medicine, pharmacology, etc.) – by training it on the respective domain datasets and dictionaries. Users can upload their own data so that the particular NER model can learn to recognize specific entities and terms, as well as define new categories of named entities. The service can provide great precision in recognizing industry-specific terminology and flexibility in adapting to changing user needs, making it a versatile text processing automation tool. Cloud-based technology makes the service scalable, fast, and efficient in processing large amounts of text data, which is extremely useful in a variety of business applications. Its key features include:

- Entity Recognition; the service automatically identifies predefined entities such as people, organizations, places, products, monetary amounts, dates, and other categories that can be customized to the needs of a specific application;
- Classification; each entity in the text is assigned to one of the predefined categories, allowing for more effective data management and analysis;
- Scalability; thanks to its cloud-based architecture, the service can handle both small and large volumes of data without the need for infrastructure management;
- Application Integration; the service can be easily integrated with existing applications such as search engines, business intelligence systems, chatbots, and CRMs.

III. Architecture and Capabilities of the CAISE Platform

We propose a cloud-based solution that will provide users with improved usability, performance, and reliability. This will extend the existing cloud PaaS (Platform as a Service) layer to support development processes for building a representative set of intelligent services. Furthermore, this solution will provide a suitable computing environment for experimenting with such services. Automatic analysis of digital documents and extraction of the knowledge contained within them will enable rational decision-making and the implementation of required action scenarios for many repetitive tasks previously performed manually by humans.

Infrastructure services are related to the basic layers of the TASKcloud cloud computing infrastructure (OpenStack, IaaS), the infrastructure supporting the creation of services and applications on this cloud (Kubernetes, PaaS), and the infrastructure for deploying core and sectoral services (the SaaS layer). Such services provide both access to the developed applications and their proper execution in the cloud environment.

System services of the CAISE platform include platform functions such as management, monitoring, quality assessment, and reporting. These services enable its users to use basic and sectoral services in a friendly and efficient manner, as well as to perform validation procedures regarding the assessment of service quality and platform usability. They support the management of users, data, models, and services, as well as operations such as monitoring and reporting the progress of production processes. They also allow for conducting platform validation experiments, including testing scalability, performance and energy optimization.

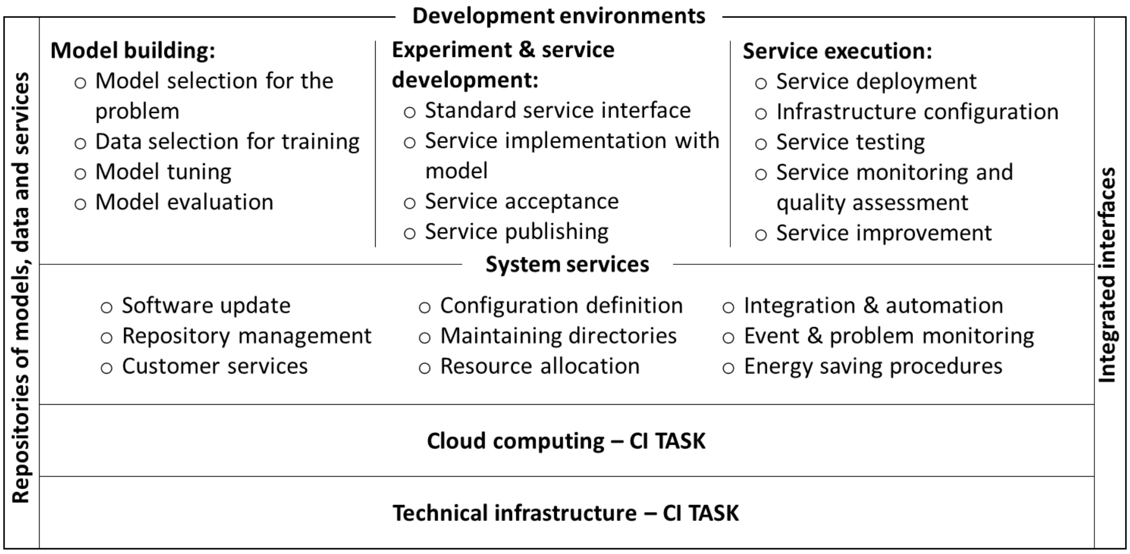

Fig. 1. Architecture of the CAISE platform

The platform also offers three important user environments: model building and training, creating a service based on that model, and deploying the service to clients. The model training environment requires understanding the problem the service is intended to solve and collecting the data necessary to build or find a suitable artificial intelligence (AI) model. Next, appropriate data sources can be identified to select and describe the dataset for the training process. Based on thus prepared data, the model is trained and tested. To create this learning environment, open-source components such as PyTorch, Terraflow (ML/NLP), Argo Workflows were implemented and integrated.

The main task of the User Service Development Environment is to transform the prepared model into the code for the appropriate service. Additionally, a user-friendly service interface must be built and the model packaged, for example, as a Python program. Appropriate testing must then be conducted to validate and upload the implemented service program to the repository. Further service quality improvement occurs in the Client Service Implementation Environment. The service is implemented as a Kubernetes image, and appropriate computing infrastructure is allocated for quality testing. Service behaviour is monitored and evaluated for four types of test data:

- valid – falling within the range expected by the service,
- borderline – close to the limits, but still within the acceptable range,
- incorrect – not compliant with the specification, which should trigger error handling, and
- unusual – rarely used, allowing for the detection of unpredictable behaviours.

Such testing allows for the service to be adapted to the actual cloud environment. In specialized environments, from Keycloak for user management through Prometheus for monitoring and Zammad for error reporting up to GitLab for CI/CD and DevOps support, open-source platform components are integrated and deployed. In consequence one obtains a layer-oriented architecture supporting service development and service execution (see Fig. 1). Thanks to the whole set of system services, we can gain access to computing networks and servers, data stores or repositories, as well as to the entire portfolio of services and applications. There exists also possibility of conducting research experiments regarding the selection of the optimal computing or energy-saving infrastructure.

A specific product of the CAISE platform are use cases (dedicated applications) that define action scenarios in specific areas of public sector activity with a significant market impact. These are implemented by participating subcontractor SME companies, who possess the necessary knowledge regarding marketable applications. The latter can be implemented as core and sector-specific service pipelines and can also be used to validate the CAISE platform. Then the entire development process is monitored, and data regarding platform behaviour is recorded. This allows for the creation of appropriate validation procedures for platform development over a specific time period. A report on the development of the CAISE platform is prepared for each period. The report indicates what changes are needed in the system and support services. These changes are incorporated by the platform developers in the next technology cycle. This approach enables the use of DevOps methodology, enhanced with AI and CI/CD solutions, which lead to continuous improvement of the CAISE platform. This, in turn, creates the opportunity for faster product implementation and lower failure rates [23]. Moreover, ensuring high quality of the services produced and the highest level of technological readiness of the platform is also our basic requirement.

Access to the platform's services is possible via the Services Portal. Depending on their needs, platform users can create a dedicated space for their individual tasks. Within this space, they will be able to select required cloud infrastructure and use core and industry-specific services offered by the platform. The portal also enables communication with the created environments. This communication through the Services Portal is handled by dedicated integrators, allowing users to effectively utilize the capabilities of the offered environments. Among other things, platform users will also be able to efficiently create new services and new applications. Appropriate documentation (user manual) describing how to use the platform and services will be available on the CAISE Infrastructure and Services Documentation website.

IV. VALIDATION PROCEDURES AND EVALUATION RESULTS

It was assumed that completion of each CAISE project task culminates in milestones, MS = 1, …, 6. The first three tasks were completed in the first 12 months of the project. The next two tasks are to be completed in the second year, each lasting six months, and the final one will also last six months. The results presented in this article concern the project's 18th month, i.e., after reaching milestone MS = 4. More frequent monitoring (more milestones) in the first year of the project was required due to the assumption of delivering a larger range of functionalities, which required frequent verification of the work carried out.

A key element of the project is the timely completion of the validation process, the purpose of which is to document and confirm that the CAISE platform meets all its intended objectives. As mentioned before, the goal of the project has been to develop and roll-out a cloud platform on the domestic market, enabling efficient creation of intelligent cloud services portfolio, along with a representative set of applications, by domestic SMEs serving the public sector units with IT solutions involving intelligent analysis of textual and image content of digital documents in Polish.

The evaluation method should consider the product being developed properly, according to the principles of software engineering (verification), and whether it meets the requirements set by the client (validation). Teams that are not associated with either designers or users of the system are used for the evaluation. These teams work in parallel with the system designers and carry out validation. They decide whether the platform demonstrator is adequate for the actual system. The validation team has knowledge of both the system and the goals and applications of the platform. Validation results are collected during the execution of each design task and summarized after the completion of each such task.

Generally, the goals of the CAISE platform validation are the following:

- Assurance of the required product functionality by taking advantage of the best practices in software engineering as well as quality standards.
- Analysing the stages of platform development and improving the development process (DevOps), as well as ensuring efficiency of its operation and productivity of the services being developed.
- Continuous monitoring of task execution and documentation of work progress (key project milestones), as well as identifying and eliminating emerging irregularities to ensure the reliability of the platform's operation.
- Checking adequacy and appropriateness of selected solutions from the users' perspective to ensure high usability and, more generally, the maturity of the platform.

Assessment of the above quality attributes forms the basis of validation processes. The scope of these procedures results from their detailed definition. Functionality is described through a list of offered features, including lists of smart services available to users, as well as the scope of their operation. Productivity is determined both by the efficiency of the platform's operation, expressed through the times of creation and execution of basic services on it, and also considers the assessment of completeness and integrity, as well as the degree of automation of activities. Dependability is expressed through the availability and certainty of the implemented functionalities, which is associated with the ability to eliminate all emerging anomalies or problems throughout the service and platform lifecycle, as well as protection against intruders. Maturity is determined by customer acceptance, its universality, i.e., use in various application sectors, and the possibility of transferring the platform to other cloud environments.

Evaluation procedures may concern the entire platform or its basic components [18]. In this discussion, we will limit ourselves to the first option and focus on the above quality attributes, which will be assessed through two questionnaires: Post-Study System Usability Questionnaire (PSSUQ), a modified version of [19], and a TRL Questionnaire (TRLQ) [20], both filled-out regularly by platform users, which take into account the aforementioned quality attributes. PSSUQ measures qualitative aspects related to the platform use. The survey is repeated in the same form for subsequent releases of the platform. TRL questionnaires investigate principles to ensure robust, reliable, and responsible product. We distinguish six milestones in the project. We start construction of the platform from TRL = 3, which finally reaches TRL = 9 in subsequent milestones. Answers to each respective question $a(q_i)$ in these surveys belong to one of the following options:

$$a(q_i) = \begin{cases} \text{It does not meet the requirements} & Q \in [0.0, 0.3) \\ \text{It meets the requirements partially} & Q \in [0.3, 0.5) \\ \text{It meets the requirements satisfactory} & Q \in [0.5, 0.7) \\ \text{It meets the requirement completly} & Q \in [0.7, 0.9) \\ \text{It meets the requirements with a surplus} & Q \in [0.9, 1.0] \end{cases} \quad (1)$$

In the case of PSSUQ the overall survey result represents the average value of numerical responses to each question, i.e.

$$\tilde{a}(PSSUQ) = \frac{1}{n}\sum_{i=1}^{n} a(q_i) \quad (2)$$

where n is the number of PSSWQ questions. Values of $\tilde{a}(PSSUQ)$ collected at milestone MS=4 are presented in Fig.2.

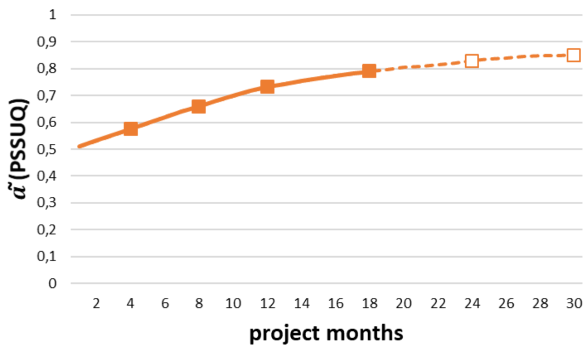

Fig. 2. Values of $\tilde{a}(PSSUQ)$ upon reaching milestone MS=4 in month 18

It may be readily seen that values of $\tilde{a}(PSSUQ)$ increase with the development of the platform, while reaching each subsequent milestone. The current value for milestone MS = 4 in month 18 of the project is 0.75. It is expected that by the end of the project (in month 30) it will increase by almost 0.1.

For each technology readiness level $k = 3, \ldots, 9$ considered in the project a relevant set of questions Q_k is asked in the TRLQ questionnaires. The mean value obtained for answers for each set Q_k, $k = 3, \ldots, 9$, is denoted as $\tilde{a}(Q_k)$. In general, the CAISE platform is considered to achieve the l-th technology readiness level when the following formula is satisfied:

$$l(TRLQ) = \max_{k=3,\ldots,9} k \mid \tilde{a}(Q_k) > 0.80 \quad (3)$$

Two curves in Fig.3 represent values of $\tilde{a}$(TRLQ) obtained at two milestones, MS = 3 in month = 12 and MS = 4 in month = 18. In may be seen that for the first 12 months of the project their values essentially overlap and are above 0.8. According to Formula (3), this corresponds to achieving level TRL-6. Lower values than 0.8 indicated by these curves identify the need for further improvement of the platform. Moreover, these curves show that actions taken in the period between MS = 3 and MS = 4 were successful. Indeed, from month 12 to month 18 the TRL7 level was achieved, what is indicated by the upper curve for MS = 4. Clearly, the top TRL-9 level can only be achieved at the completion of the project.

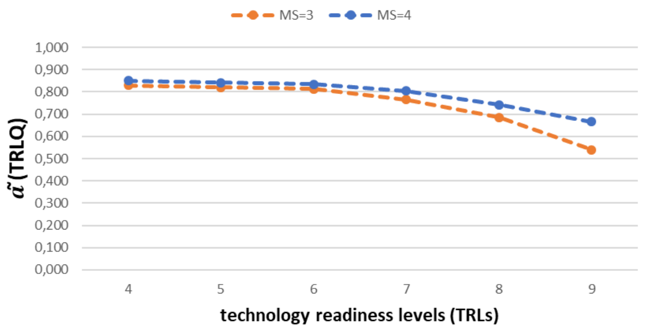

Fig. 3. Values of $\tilde{a}(TRLQ)$ collected at two different milestones MS = 3 (month = 12) and MS = 4 (month = 18) of the project.

Results of our two kinds of surveys plotted in Fig. 2 and 3 are closely related. Apparently, a higher level of quality represented by PSSUQ corresponds to the higher level of TRL. This means that the platform integration and deployment process is being run correctly and gradually becomes resistant to various kinds of anomalies or occurring problems.

Results of the above survey is, to some extent, a subjective evaluation of the CAISE platform – it is carried out by human teams. Therefore, the validation procedure should consider quality measures resulting from the measurement of the production processes. One of these is the integrity degree of open-source software components used to build the platform. As shown in Fig. 4 the project assumed the use of 14 different types of open-source components. Their integration, indicated by blue edges, is the process of combining various components and their resources to enable both their cooperation and coordination of activities within a single integrated system. In other words, integration ensures free exchange of information and effective interaction between pairs of such components, e.g., through various types of data exports from one component to another.

The aforementioned task is not easy, if possible at all; it most often requires modification of the integrated software components. Nevertheless, the main benefits of such a software integration may be the following:

- Improved operational efficiency,
- Automation of processes carried out on the platform,
- Efficient data exchange,
- Reduction of operational costs.

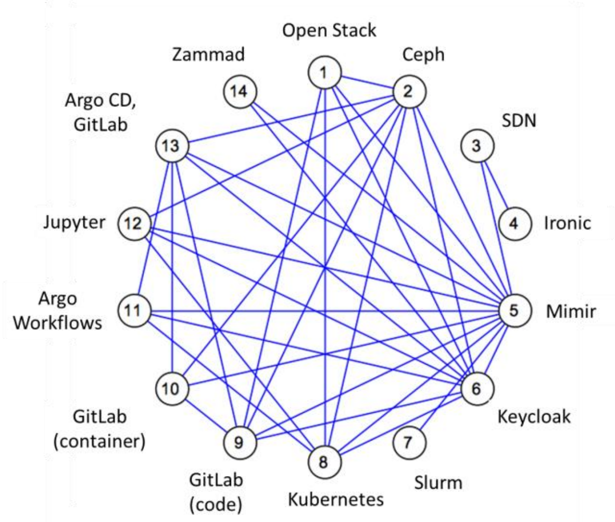

Fig. 4. Integration of open-source components of the CAISE platform

Ultimately, the degree of integration of the CAISE platform can be represented as an adjacency matrix of a graph, where the set of vertices represents the open software components running on the CAISE platform, and the edges between pairs of vertices represent the integration of the specified environments. The metric for the degree of integration is taken as the ratio of the number of integrated component pairs (N_i) at a given stage of platform development (where i denotes a milestone number) to the number of required integrations (N):

$$degree(i) = \frac{N_i}{N} \quad (4)$$

Table I presents the progress in integrating components of the CAISE platform.

TABLE I
DEGREES OF CAISE COMPONENT INTEGRATION

MS =	1	2	3	4	5	6
degree (i)	0.23	0.59	0.74	0.94	1.0*	1.0*

**target values*

As a result, 14 open-source software components can be implemented, with the target number of integrated component pairs being 47. The level of integration increases with each milestone and will reach its maximum value six months before the project completion in month 30. This will enable a comprehensive assessment of the CAISE platform during the *First Industrial Deployment* (FID) phase of the project, planned for the period between milestones MS = 5 (month = 24) and MS = 6 (month = 30).

V. RELATED WORK

There are various types of AI PaaS services available in the computing market, offered by different cloud service providers. They differ in functionality, tools' capabilities and technical support. Professional solutions include Google Cloud Platform, Amazon Web Services, and Heroku, to name a few. AI PaaS'es enable faster development, testing, and deployment – the key activities for Agile and DevOps teams. Furthermore, they directly complement the release flow of CI/CD processes and helps support a full DevOps release cycle. Open-source solutions, such as GitHub, GitLab, Bitbucket or SourceForge, are also available for hosting. What is more, they allow users to build, test, deploy, and manage applications without having to manage hardware infrastructure. The CAISE platform, integrating a range of open source components, constitutes a similar solution, dedicated for supporting development of intelligent services for analysing digital documents. It addresses needs of SMEs suffering shortages of highly qualified AI staff by providing access to cloud resources and tools enabling creation and implemention of such services and assessing their quality. This contribution is twofold.

One is striving for a fully operational, open-source-based, cloud-based solution that spans the entire lifecycle of document intelligence in the public sector – in one coherent productized workflow. By offering a portfolio of composable and containerized services, the platform uniquely enables SMEs to build resilient, legally compliant solutions for public administration units [1]. Individual pieces of workflows that could be deployed with the CAISE platform are treated separately in the current literature. No paper yet advocated for comprehensive, production-grade, end-to-end solutions that would integrate universal intelligent services for formal administrative procedures and decisions, including multi-document process historicization, auditable anonymization, human-in-the-loop LLM refinement or structure-constrained generative outputs [21]. So far, the literature focus is on various modules of classic pipeline architectures, like UIMA [24] or on considering only conceptual aspects of using AI in public administration [25] – but not integrated production pipelines.

Another innovative concept signaled in this paper is linking TRL level to software quality metrics. Direct empirical studies on that are sparse. The literature contains many studies either on software quality metrics or on TRL adaptations, but no paper attempted to empirically correlate an assigned TRL with downstream quality metrics. However, one paper postulated mapping of existing metrics into readiness indices [26]. Authors showed how to map quality and process metrics into a composite readiness index and then studied correlation with operational out-comes, but their work is just a proposal rather than a broad empirical evidence. And here in our view opens a promising new direction of research – can one reasonably expect that projects that meet the artifact thresholds for higher TRLs should show better operational metrics? To verify that claim thorough cross-project statistical studies that control for domain, team size, and criticality are needed; these are mostly absent or limited in scope today.

VI. Conclusion

DevOps methodology, Artificial Intelligence (AI) and Machine Learning (ML) can already be seen as a trend in the development of intelligent cloud services. Their practical exploitation led to a variety of alternative solutions, which are worth considering but may need evaluation. We have introduced one original solution – the CAISE platform, which can support service engineering in various segments of the public sector. Intelligent services developed with it have a great potential to leverage digitalization of many important public institutions, making this sector more resilient and responsive to the modern society needs.

We have discussed the selected validation procedures that can be carried out during the platform's development and can signal the platform's weaknesses that need to be further addressed. Thanks to SME subcontractors, active on the domestic (Polish) market of software suppliers for various segments of the public sector, and who are involved in the CAISE platform validation process, such an assessment is reliable and, consequently, will allow achieving the highest level of technological readiness.

In this context, we presented the CAISE platform and its capabilities, which have already been substantially confirmed by the feedback from its target users. Moreover, during the development of the project, interesting correlations were discovered between the qualitative attributes of the platform and the levels of technological readiness. These can be incorporated into the DevOps methodology and enable the continuous further improvement of the CAISE platform.

References

[1] Project Portfolio, "Cloud Artificial Intelligence Service Engineering (CAISE) platform to create universal and smart services for various application areas", Gdansk University of Technology, August 12, 2023.

[2] IPCEI-CIS: https://digital-strategy.ec.europa.eu/pl/news/ipcei-next-generation-cloud-infrastructure-and-services-boost-europes-digital-decade.

[3] S. Manvi, G.K. Shyam: "Cloud Computing Concepts and Technologies", CRC Press, Taylor & Francis Group, LLC, 2021.

[4] M. Kalech, R. Abreu, M. Last, "Artificial Intelligence Methods for Software Engineering", Scientific Publishing Co. Pte. Ltd, 2022.

[5] J. Davis, R. Daniels, "Effective DevOps", O'Reilly Media, Inc., 2016.

[6] H. Merode: "Continuous Integration (CI) and Continuous Delivery (CD): A Practical Guide to Designing and Developing Pipelines", Apres, 2023.

[7] D. Rosenberg, B. Boehm, B. Wang, K. Qi, "Rapid, evolutionary, reliable, scalable system and software development: the resilient agile process", Proceedings of the 2017 International Conference on Software and System Process (ICSSP 2017), doi: 10.1145/3084100.30841

[8] M.S. Merkow, R. Lakshmikanth, "Secure and Resilient Software Development", CRC Press, Taylor & Francis Inc.

[9] A. Lavin, C.M. Gilligan-Lee, A. Visnjic, et al., "Technology readiness levels for machine learning systems", Nat. Commun. 13, 2022, doi: 10.1038/s41467-022-33128-9

[10] B. Combemale, J. Gray, B. Rumpe, "Research software engineering and the importance of scientific models". Journal Software and Systems Modeling (SoSyM), Vol. 22(4), pp. 1081-1083, Jul. 2023

[11] H.M. Elsherif, "Natural Language Processing (NLP). The Complete Guide". Eldon-USA Publishing, (2024)

[12] J. Sen,, S. Mehtab, at al., "Machine Learning: Algorithms, Models, and Applications", doi: 10.48550/arXiv.2201.01943

[13] W.X. Zhao, K.Z. Li, et al., "A Survey of Large Language Models", [2303.18223] A Survey of Large Language Models. https://arxiv.org/abs/2303.18223

[14] M. Godlewska, B. Wiszniewski, "Serverless edge computing with mobile documents agents", TASK Quarterly 28 (1) 2024

[15] D. Tkachenko, M. Sobotka, A. Górecki, J. Kłopotek-Główczewski, T. Neumann, J. Rumiński, "Detection of Electronic Protected Health Information in DICOM Objects: Rule-Based Methods vs. LLMs". In: Ladyzynski, P., Pijanowska, D.G., Liebert, A. (eds) Joint 20th Nordic-Baltic Conference on Biomedical Engineering & 24th Polish Conference on Biocybernetics and Biomedical Engineering. NBC PCBBE 2025 2025. IFMBE Proceedings, vol 131. Springer, Cham. doi: 10.1007/978-3-031-96538-8_12

[16] M. Boczoń, "Automatic service for generating image descriptions in Polish". MSc Thesis, Gdańsk University of Technology, 2025, supervised by T. Dziubich.

[17] J. Majkutewicz, J. Szymański, "An Integrated Approach for Prototyping and Deploying Large Language Models". In: Saeed, K., Dvorský, J., Fukumoto, M., Nishiuchi, N. (eds) Computer Information Systems and Industrial Management. CISIM 2025. Lecture Notes in Computer Science, vol 15927. Springer, Cham. (2025). doi: 10.1007/978-3-032-02406-0_13

[18] H. Krawczyk, B. Wiszniewski, "Resilient Software Development of Intelligent Services in Cloud Environments", Advances in Dependable Systems and Networks - Proceedings of the Twentieth International

Conference on Dependability of Computer Systems DepCoS-RELCOMEX, June 30 - July 3, 2025, Brunów, Poland, Springer.

[19] PSSUQ calculator: PSSUQ Calculator - UIUX Trend

[20] S. Bhattacharya, V. Kumar, S.N. Nishad "Technology Readiness Level: An Assessment of the Usefulness of this Scale for Translational Research", Productivity, 62, January 2022, doi: 10.32381/PROD.2021.62.02.2

[21] 8ra Cloud-Edge Continuum; Europe's Next Generation Cloud Infrastructure and Services, https://www.8ra.com/

[22] European Health Data Space Regulation (EHDS), https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en

[23] DORA 2024. Accelerate State of DevOps.

https://dora.dev/research/2024/dora-report/2024-dora-accelerate-state-of-devops-report.pdf

[24] D. Ferrucci, A. Lally, UIMA: an architectural approach to unstructured information processing in the corporate research environment. Natural Language Engineering. Vol. 10(3-4), pp. 327-348, (2004), doi: 10.1017/S1351324904003523

[25] S. Weerts, Generative AI in public administration in light of the regulatory awakening in the US and EU, Cambridge Forum on AI: Law and Governance, Vol. 1(e3), pp. 1–19, (2025) doi: 10.1017/cfl.2024.10

[26] D.C. Lowe, L. Justham, M.J. Everitt, Multi-index analysis with readiness levels for decision support in product design, Technological Forecasting and Social Change, Vol. 206, (2024),

doi: 10.1016/j.techfore.2024.123559.