5.1 MLOps Platform for Industrial IoT Services to Automate Artificial Intelligence Operation Processes and Enable Artificial Intelligence Utilization



MLOps platform AI service operation processes and functions

Artificial intelligence (AI) application data for industrial Internet-of-Things (IoT) services vary depending on various factors. To operate such applications, it is essential to continuously execute highly specialized processes such as performance monitoring, retraining, and model management, which incur substantial costs that must be reduced.

With this in mind, the Toshiba Group has developed a machine learning operations (MLOps) platform that automates these processes, enabling smooth implementation on-site. The new MLOps platform provides the following functions: (1) pipeline management for automated AI model training, (2) model serving to offer AI models as inference services, (3) monitoring capabilities to track data changes and AI model performance, (4) training data management, and (5) continuous integration (CI) and continuous delivery (CD) to automate source code and AI model testing and deployment. This MLOps platform can be easily constructed using automated build scripts.

Applying the new platform to industrial IoT services confirmed that it allows even non-AI specialists to update AI models and handle data changes. We aim to expand the applications of this platform going forward.

# 5.2 Practical Application of Remote OS Update Technology for Linux<sup>®</sup>-Based Embedded Devices



Overview of remote OS update technology and configuration changes necessary for practical applications

Embedded devices require regular operating system (OS) updates to fix security vulnerabilities. To do so, Toshiba Corporation has developed a remote OS update technology for Linux<sup>®</sup>-based embedded devices based on open-source software (OSS) and conducted a proofof-concept (POC) demonstration.

We incorporated the technology into the cloud-connected interface of the ELCLOUD digital transformation (DX) controller, a new service offered by Toshiba Elevator and Building Systems Corporation. In the POC demonstration, embedded devices downloaded OS update data from a server. However, to achieve practical application, we reconfigured the DX controller to receive OS update data from the cloud services it utilizes. We expect that this will help provide timely OS updates and reduce on-site support costs.

Also, to ensure secure OS updates, the new remote OS update technology provides the following features: (1) OS update data integrity verification, (2) OS update data encryption and decryption, and (3) rollback to the previous OS version in the event of an update failure.

The POC demonstration confirmed that the technology can also be applied to other social infrastructure systems.

Linux<sup>®</sup> is the registered trademark of Linus Torvalds in the U.S. and other countries.

### 5.3 GridDB Cloud Multi-Tenancy Technology



App: application API: application programming interface

Reduction in resources achievable with multi-tenancy



Method of setting resource usage limits

GridDB Cloud is a database for IoT and big data which are frequently generated in large volumes. Toshiba Digital Solutions Corporation previously released a single-tenant version of GridDB Cloud, which allows users to monopolize resources such as CPUs and disks. We have now developed a multi-tenant version, which allows multiple users to share systems, databases, and other resources after access rights are granted to individual users.

However, a given user should not consume large proportions of shared resources such as CPUs and disks in a multi-tenant environment because doing so can degrade application performance for other users or hinder data registration.

To prevent this and achieve a low-cost multi-tenant environment, we have developed a function to set an upper limit on resource usage for each user. It denies resource access when the number of requests sent by a given user reaches the upper limit within unit time and places databases in read-only mode to prevent further data registration when the amount of disk space used by a given user reaches the upper limit.

# 5.4 Tenant-by-Tenant Deployment and Scaling with Automated Environment Setup Based on Kubernetes™



Delivery infrastructure architecture and building service environment process

Time required to build each tenant unit	Time required to open new tenant unit	Running cost per tenant unit
40 h 📫 12.8 h 68% reduction	2-4 weeks 1 week 50-75% reduction	\$173 <b>\$149</b> \$149 <b>14% cost</b> reduction

Estimating effectiveness of using AI chatbot service

Toshiba Digital Solutions Corporation is currently developing the Amazon Web Services<sup>™</sup> (AWS) version of a delivery infrastructure as an architecture that combines scalability and service updating flexibility required for recurring businesses. It uses Kubernetes<sup>™</sup> clusters and can be deployed independently on a tenant-by-tenant basis, enabling efficient establishment and operational support for each customer environment.

We have implemented the delivery infrastructure in a scenario-less AI chatbot service, confirming its effectiveness. The delivery infrastructure provides security function enhancements, a new application update application programming interface (API), and automated compatibility and availability tests to allow flexible tracking of frequent version upgrades. We estimate that incorporating these enhancements into a chatbot service reduces the time required to build each tenant unit by 68%, the time required to open a new tenant unit by up to 75%, and running costs per tenant unit by 14%.

Next, we will implement further enhancements and prepare the Microsoft Azure<sup>™</sup> version of the delivery infrastructure.

Amazon Web Services and other AWS trademarks are trademarks or registered trademarks of Amazon.com, Inc. or its affiliates in the United States and other countries.

Kubernetes is a trademark or registered trademark of the Linux Foundation in the United States and other countries. Windows, Microsoft, and Azure are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

SCIENCE AND TECHNOLOGY HIGHLIGHTS 2024

# 5.5 NOVICEGATE Automated Machine Learning System for Initial AI Evaluation



Proposal for AI installation based on initial evaluation using NOVICEGATE

When integrating AI into business systems or on-site environments, it is essential to use on-site data to evaluate the expected benefits. This includes assessing whether relevant data are being collected and whether the problem setting is appropriate for an AI-based solution. Selecting the right algorithm is crucial for success.

The Toshiba Group has developed an automated machine learning system called NOVICEGATE, which is designed to accelerate the initial AI installation evaluation process. NOVICEGATE makes it possible to receive data from customers, select the best AI algorithm rapidly, assess expected benefits, and present them to the customer. This enables customers to efficiently decide whether to install AI. Furthermore, NOVICEGATE can propose high-accuracy AI-based systems, including our proprietary AI technologies.

# 5.6 AI Test Template to Understand AI Testing and Improve Test Efficiency

Before introduction	n't know how to test Al models or systems.
Al test template	
89 + 米 0 0 > m C >> Code ~	● Python 3 (lyg/same). 〇
4. テストケース一覧と合否	
テスト%。大項目	中国目 デスト方法 合体基準 デスト結果 担当 検査日付 合西
1 デストデータ地址	AFEでんの学校開始にないていてきまた。 10時データ Manach An Lander U. Volich 学校的になった。 18日ジータ Manach - 一品な良したのな物能していない。 18日ジータ Manach - 二品な良したのな物化していない。
2 デストデータ構成	A4学が必要認定になってはできました。 製ビーク いー ・
5. テストコードと実行結果	
<ul> <li>5.1. テストケース1~2</li> <li>ingent : テストに必要なライブラリの歳み込み</li> <li>ant data_gath and read data.テストケース第に対象となるデー</li> </ul>	Organization of AI test template
<ul> <li>setting data学習回該を包括する</li> <li>learning and evaluate Model学習回該だけ学習させ、学習回該</li> </ul>	ددهسته 1. Purpose of testing
<ul> <li>[3]: eccentressessessessessessessessessesee</li> <li>e. 25.6 d. import</li> <li>e. MAR</li> </ul>	2. What is tested
eeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeee	3. Test method
import sys sysjektik-append(*/*) from argarse import ArgumentParser	4. List of test cases and test results
from src.stati import root. from src.stati import Trainor from src.statiser import Trainor	5 Test code and execution results
Supert pandas as pd Supert nampy as ny	S. Test code dila execution results
	Sample test code (for image classification and tabular data regression problems) is reusable.
	Developers can gain insights into AI quality attributes and test methods by testing samples.
After introduction	in a short time by replacing data and models.

Overview of template for AI testing

Because an AI model learns from training data and works on inductive reasoning, AI systems exhibit behavior distinct from traditional software. Therefore, AI-based software quality attributes and test methodologies differ considerably from those of non-AI-based software. To test AI models and data, extensive knowledge is required, including that of AI concepts and statistical techniques, resulting in a heavy burden on developers.

With this in mind, Toshiba Corporation has developed an AI test template that incorporates best practices and techniques to eliminate the knowledge gap among developers. Specifically adapted to the quality attributes of AI, this template provides typical test cases, a concept of expected values, and a simplified test environment. In addition, we offer sample test codes for image classification and tabular data regression problems, which allow developers to gain insights into AI quality attributes and effective test approaches. For AI systems like the samples, the template eliminates the need to create test codes manually as developers only need to replace sample data and expected values in the template.

We will continue with efforts to enhance the quality of AI systems and streamline AI model and data testing.

# 5.7 Inspection Information Management AI Technology for Efficient Plant Infrastructure Inspection



**Overview of inspection information management AI** 

Inspection rounds and maintenance activities for plant infrastructure place a heavy burden on operators, who have conventionally had to record inspection data using pen and paper while referring to images they have captured.

To lighten the workload, the Toshiba Group has developed an inspection information management AI capable of identifying the camera position and equipment anomalies such as cracks and rust from images captured with standard cameras. The AI system can estimate the camera position using deep learning techniques, even for images captured indoors where the global positioning system (GPS) does not work. Furthermore, it can recognize general equipment anomalies by comparing deep feature differences between images captured during normal conditions and those captured during inspections. This feature allows for recognition of subject conditions and the camera position from a single image. These pieces of information are then linked and managed together to considerably streamline the on-site inspection process. The new AI system also standardizes the tasks that were previously dependent on individual expertise.

We have also developed a cloud-based demonstration environment to make it easy to evaluate the effectiveness of the new AI system from various internal and external sites.

# 5.8 Advancements in AI OCR Service Breaking New Ground for Data Businesses

	Group A		Grou	Group B	
Section	Type A	Type B	Type A, B	Others	
	10	20	5	10	
ole withou	ut some ho	rizontal and	vertical lines	;	
Assets			2020/3	2021/3	
Current	assets		2,038,099	2,130,908	
Investm	Investments and long-term loans			534,028	
Tangibl	e fixed assets		575,810	575,442	
Other a	ssets		256,822	260,258	
Total ass	ets		3,383,433	3,500,636	
ble with shading		First con Medical Home m	nsultation fee management nedical care	13	
			ation		
		Examina			
		Examina Diagnos	tic imaging		
		Examina Diagnos Medicat	ion	14	
		Examina Diagnos Medicat Injection	itic imaging ion	14	

**Examples of improved OCR** 

Toshiba Digital Solutions Corporation offers an AI-based optical character recognition (OCR) service to extract specified items of information such as the amount, issue date, and contact information from client-specific unstructured documents, including invoices and quotes. Characters in such documents are sometimes difficult to recognize if they contain complicated tables or tables with shading.

We have improved table recognition by applying our original table analysis and image processing technologies while enhancing robustness against various embellishments by expanding the coverage of training data for deep learning-based OCR. This enables accurate recognition of business forms that were difficult to recognize with conventional OCR systems such as complex tables with merged cells and shading and those without some horizontal or vertical lines. We confirmed that the new OCR system provides a 67% higher recognition rate than the conventional system<sup>(\*)</sup>.

Next, we plan on developing a data business to create new customer value by extracting and utilizing information from unstructured documents that have been difficult to use so far.

(\*) Results in the specified item extraction mode using our sample business forms including invoices

# 5.9 Deployment of Common AI Asset Services for Enterprises Applying Generative AI



Overview of business efficiency improvement and service value enhancement using common assets for generative AI

The use of generative AI is spreading rapidly both in Japan and abroad as it is expected to improve work efficiency, reduce development times, and provide other DX benefits. In the case of business inquiries, for example, users can extract relevant information from related documents to provide as input to a generative AI system for obtaining better answers.

To facilitate the use of generative AI for business, Toshiba Digital Solutions Corporation has developed common assets by combining enterprise-oriented generative AI services<sup>(\*)</sup> with conventional natural language processing technologies. These common assets simplify adding a series of processes to new or existing services, including those for understanding the intent of input questions, high-precision document searches, and generating answers based on search results.

We have commenced a proof-of-concept demonstration of the new generative AI chatbot service called "Commendry with Generative AI" using the new common assets.

(\*) System configuration that allows limited access to the service (available on a closed network) and does not use input/output data for model learning

# 5.10 Next-Generation Speech Synthesis Technique Based on Deep Neural Networks



user-controllable prosody, parameter tuning, incremental generation using fixed memory, and high-fidelity voices comparable to recorded voices of real people

Block diagram of next-generation speech synthesis based on DNN



Five-point MOS of speech naturalness

Toshiba Digital Solutions Corporation has developed a next-generation speech synthesis technique based on deep neural networks (DNNs) for ToSpeak RECAIUS speech synthesis middleware and related text-to-speech products. Our goal was to achieve high-fidelity speech synthesis comparable to real humans while maintaining ToSpeak functionality. The new technique meets customer requirements, including those for compact memory size and parameter tuning options for voice quality and prosody adjustment.

The new speech synthesis technique consists of (1) a framework with a phoneme encoder and acoustic feature decoders and (2) a neural vocoder for waveform generation from acoustic features. For each DNN model, a model compactization technique was employed to reduce memory size and computational requirements. Sound naturalness evaluations by non-expert

subjects using a crowd-sourcing system demonstrated that the synthesized voices closely resembled the recorded voices of real people. Even compact DNN models caused only a slight degradation in voice quality.

To expedite the market launch, we plan to offer an initial version that implements the neural vocoder (DNN Hybrid), followed by a version that incorporates an encoder-decoder DNN model in the framework (DNN Full). We will provide licensing for the technology that can be used for a wide range of applications from embedded systems to advanced content applications.

# 5.11 SQBM+ for AWS<sup>™</sup> Capable of Solving Combinatorial Optimization Problems with Up to 10 Million Variables



Overview of new version of SQBM+ for AWS capable of solving combinatorial optimization problems with up to 10 million variables

In 2021, Toshiba Digital Solutions Corporation launched the SQBM+ service, a quantum-inspired optimization solution. SQBM+ solves large-scale combinatorial optimization problems at high speed and with high accuracy using the simulated bifurcation algorithm invented thanks to research into quantum computing. SQBM+ for AWS<sup>™</sup>, an Amazon Web Services<sup>™</sup> version of SQBM+, succeeded in solving problems with up to 100 000 variables. However, as we worked on real-world problems, we recognized the need to handle optimization problems with even more variables. There were also requests to support new graphics processing unit (GPU) instance types to further improve computing performance.

With this in mind, we have released a new version of SQBM+ for AWS, which supports up to 10 million variables and runs on GPU instances with various architectures and memory configurations, including the NVIDIA<sup>TM</sup> Ampere GPU architecture. The new version allows users to select the optimal GPU instance type, considering calculation cost and performance. In addition, it is capable of efficiently solving small- to large-scale problems by adjusting the GPU resource allocation according to scale.

We have also improved the calculation algorithm resulting in the new version of SQBM+ for AWS succeeding in finding the optimal solution to 60 problems out of 71 G-set benchmark problems in comparison with 50 achieved by the previous version (with a calculation time of 3 000 seconds).

Amazon Web Services, AWS, and Amazon Elastic Compute Cloud (Amazon EC2) are trademarks of Amazon.com, Inc. or its affiliates in the United States and other countries.

NVIDIA is a trademark or a registered trademark of NVIDIA Corporation in the United States and other countries.

# 5.12 Quantum Key Management System (Q-KMS) V2.1



Overview of Quantum Key Management System (Q-KMS) V2.1

Quantum key distribution (QKD) networks deliver quantum keys between QKD systems via optical fibers, which, in turn, generate encryption keys for applications from the quantum keys. Applications then use the encryption keys to convert plaintext data into ciphertext to ensure secure communication.

Toshiba Digital Solutions Corporation has released Quantum Key Management System (Q-KMS) V2.1, which allows encryption keys to be shared with any QKD systems that are not directly connected. Q-KMS V2.1 optimizes the key distribution route according to failure and performance conditions, improving availability. Furthermore, when used by multiple applications, Q-KMS V2.1 adjusts the quantity of keys provided for each application to prevent key shortages due to overuse in a specific application and thereby ensure key provision.

Q-KMS V2.1 is utilized in QKD testbeds in Japan and we plan on improving functions based on the insights gained from such testbed verification.

### 5.13 DNCWARE Blockchain+ Managed Service Suitable for Business Co-creation

In 2022, Toshiba Digital Solutions Corporation launched DNCWARE Blockchain+, a highly reliable enterprise blockchain developed based on proven cluster technology. One of the features of a blockchain is its ability to enable transparent data collaboration among multiple enterprises. DNCWARE Blockchain+ leverages this feature with the aim of creating a world where various types of data are connected to expand services. It allows anyone to freely develop services on the blockchain platform and facilitate collaboration among the services to generate new business opportunities. To achieve this vision, it is essential to offer service and sales models that make it easy for anyone to adopt DNCWARE Blockchain+.

With this in mind, we have launched managed services through an electronic commerce (EC) platform. DNCWARE Blockchain+ employs technology that allows it to be used without the need to be aware of complex blockchain mechanisms such as consensus algorithms and digital signatures. It provides a clear separation of authority between application developers and service providers operating blockchain nodes, making it suitable for managed services. In addition, it provides automatic node restart and reconfiguration mechanisms to ensure swift recovery in the event of a node failure. Therefore, application developers can use DNCWARE Blockchain+ without having to consciously build or manage a blockchain platform.

By leveraging our EC platform which integrates with external services, we have enabled usage-based billing and credit card payment, which were previously unattainable through a conventional order-based system integration (SI) mechanism. The setup of application environments and the collection of billing information have also been automated to reduce the time required from usage initiation to billing.

To date, several startup companies that faced challenges with traditional order-based SI models have adopted EC sales to conclude contracts for Blockchain+ services.



DNCWARE Blockchain+ goal



API: Application Programming Interface

Overview of DNCWARE Blockchain+ managed service

# 5.14 Sound Field Control to Create Space Where Sound is Audible Only in Specific Area









Toshiba Digital Solutions Corporation has released sound field control software to create a space where sound is audible only in a specific area. Designed to create a certain listening experience, the software produces a directional sound field with only three general-purpose speakers.

To establish a directional sound field, many speakers are usually necessary to increase the sound pressure using phase differences. The new sound field control software, however, combines acoustic power and phase difference control to reduce sound pressure levels around the speakers according to the acoustic power law while increasing directional sound pressure levels using sound phase differences.

Combining these two principles, the new software succeeded in creating a more than 12 dB difference in sound pressure between audible and inaudible areas with only three speakers. This allows the software to deliver sound to a desired area only, even in places where it is difficult to install multiple sonic or ultrasonic speakers.

Using this software, we aim to offer a sound field control solution suitable for many use cases by achieving various sound pressure distribution profiles.

# 5.15 Weather Data Service to Analyze Weather Observation Data Using Proprietary Method



\*1: Data from parabolic dual-polarization weather radar are currently used.

\*2: Gust detection service is currently in development and scheduled for release in near future.



In May 2023, Toshiba Digital Solutions Corporation launched the Weather Data Service, a solution to analyze observation data received from weather radars in real time and with high accuracy.

The Toshiba Group has a history of providing weather radars and weather observation systems capable of processing observation data. By leveraging our analytics technology and expertise accumulated through many years of experience in manufacturing radars, the Weather Data Service analyzes data received from public weather radars and provides weather forecast data to various users.

The service consists of three elements: Rainfall Forecasting Service, Particle Classification Service, and Gust Detection Service. The Rainfall Forecasting Service analyzes the growth of rain clouds using our proprietary method to accurately forecast the signs and amount of localized heavy rainfall in each 250-meter mesh up to 30 minutes in advance. The Particle Classification Service identifies various types of hydrometeors in real time, including hail, snow, and rain, and provides forecasts.

In recent years, there has been an increase in weather-related disasters caused by rapid weather changes such as localized heavy rain, hail, snowstorms, and gusting winds. We aim to contribute to disaster prevention and mitigation by offering a solution to monitor sky conditions and forecast weather conditions.

# 5.16 Energy and Space Management for Factory Buildings Utilizing IoT Products and Technologies



Configuration of energy and space management system for factory buildings

In 2022, in collaboration with Toshiba Digital Solutions Corporation and Toshiba Lighting & Technology Corporation, the Toshiba Fuchu Complex enhanced the energy and space management of the new manufacturing building using Toshiba Group's latest IoT products and technologies.

For energy management, we introduced the Unified Controller Vm series typeS and the CP30 Model 300 small embedded industrial computer from Toshiba Infrastructure Systems Corporation to enable secure and rapid IoT data collection. We also employed our Asset IoT Cloud Service for Factories to transform the collected data into value-added information, thereby establishing mechanisms for power consumption monitoring and signage to improve area and equipment energy efficiency.

For space management, we introduced the ViewLED camera-equipped LED lighting fixture and the ViewLED Solution cloud AI image analysis service from Toshiba Lighting & Technology Corporation to visualize shop floors and identify available space, enabling effective space utilization.

# 5.17 Distributed Co-Simulation Technology Contributing to Development of Battery Electric Vehicles



Thermal management co-simulation of BEV using VenetDCP connecting domestic and overseas sites

Toshiba Digital Solutions Corporation provides the VenetDCP distributed co-simulation platform to Japanese and foreign automakers for developing thermal management systems for battery electric vehicles (BEVs). Thermal management, which integrates control of heat and electric power from the battery, drivetrain, and air conditioner, is an essential element that determines BEV cruising performance.

At Toyota Motor Corporation, VenetDCP has successfully started to connect simulators at its domestic and overseas locations, contributing to BEV thermal management evaluation using overseas driving environment data. This achievement was presented at the prostep Symposium 2024.

In China, where the spread of BEVs is accelerating, we have concluded a distribution agency agreement with Shanghai ComfortUni Smart Technology, which possesses cutting-edge technology for developing thermal management systems and have begun providing VenetDCP for BEV development companies in the country.

# 5.18 Security Monitoring Technology for Steel Rolling Control Systems That Does Not Interfere with System Operation



Security monitoring technology using IDS for steel rolling control systems

Increasing cyberattacks affect operation at factories and plants. In the steel industry, the need continues to grow for steel rolling control system security.

With this in mind, Toshiba Digital Solutions Corporation has developed security monitoring technology for steel rolling control systems that does not interfere with system operation. It uses an intrusion detection system (IDS) from Claroty Ltd.

The three issues for monitoring steel rolling control systems are: (1) identifying TC-net 1G, Toshiba's control protocol, using the IDS, (2) ensuring that IDS connectivity does not interfere with system operation, and (3) providing long-term support for the stable control system operation.

We have resolved these issues by (1) developing a function to identify TC-net 1G with the IDS, (2) verifying that the IDS connects to the steel rolling control system and does not communicate with it, interfering with system operation, and (3) installing the IDS on the industrial server with long-term support.

The new control system can monitor for unauthorized device connections and abnormal communication over the long term.

# 5.19 ALWAYS V Care Needs Assessment Support System



Features of ALWAYS V care needs assessment support system

Streamlining care needs assessment is a pressing issue in Japan due to an ever-increasing number of care recipients and severe labor shortages. With this in mind, Toshiba Digital Solutions Corporation has developed ALWAYS V, a care needs assessment system, which leverages digital technology to considerably reduce the workloads of social workers while allowing them to carefully assess the health of care applicants.

ALWAYS V allows social workers to perform all care needs assessment tasks on a tablet, ranging from preparatory tasks to creating survey reports. It provides the following benefits:

- (1) Elimination of paper-based surveys: ALWAYS V eliminates traditional paper-based surveys for a paperless workflow.
- (2) Real-time input of handwritten notes: Social workers can enter survey results on a tablet while taking handwritten notes. This allows them to interact more closely with care applicants during surveys.
- (3) Comparison and consistency checks: ALWAYS V makes it easy to compare information from previous visits and check for inconsistencies between survey items, ensuring a fair assessment regardless of social workers' abilities.

These benefits help to maintain the quality and fairness of the care insurance system with the goal of an inclusive society where the dignity of all care recipients is respected.