Industrial ICT Solutions

## SATLYS Toshiba Analytics AI Services Accelerating Digital Transformation

● Launching of SATLYSKATA

To offer SATLYS high-grade analytics artificial intelligence (AI) services to more customers, Toshiba Digital Solutions Corporation has released SATLYSKATA AI analytics services customized for specific purposes, aggregating and standardizing its AI analytics expertise. Its first release comprises two cloud analytics services: SATLYSKATA Maintenance Parts Inventory Optimization and SATLYSKATA Work Activity Estimation. These services incorporate an application programming interface (API), allowing them to collaborate with customers' systems.

The SATLYSKATA Maintenance Parts Inventory Optimization service predicts the number of parts failures based on failure history data, making it possible to reduce the risk of running out of inventory and optimize the inventory of maintenance parts.

The SATLYSKATA Work Activity Estimation service estimates work activities based on acceleration data from wearable devices. Visualization of work activities helps to grasp work data, improve work procedures, and reduce work time.

We will continue to expand the lineup of our AI analytics services.
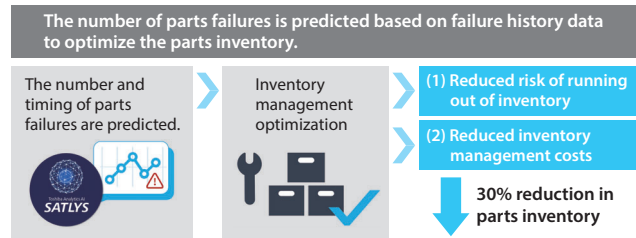
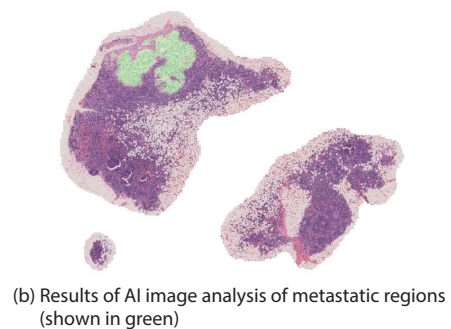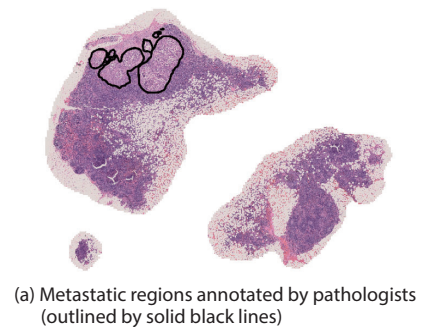Software as a service (SaaS) for Maintenance Parts Inventory Optimization
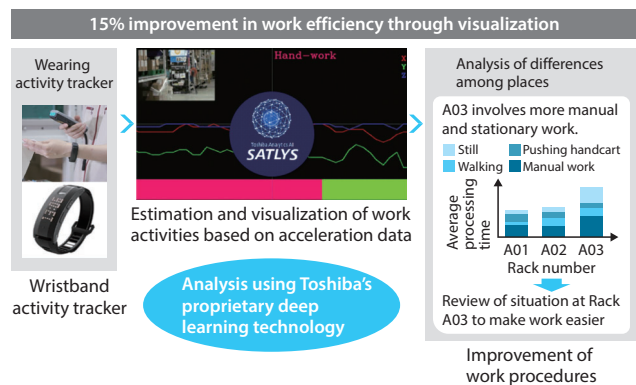
SaaS for Work Activity Estimation

● AI Image Analysis to Support Pathological Diagnosis

Cancer is the leading cause of death in Japan. In particular, Japan has one of the world's highest rates of incidence of gastric cancer. On the other hand, the number of pathologists per capita in Japan is becoming seriously insufficient–as low as one-third that of the United States.

As image recognition using AI is exceeding human capabilities, there are high expectations for its application in the medical field. Against this background, Toshiba Digital Solutions Corporation has developed a pathological image analysis technology for the diagnosis of lymph node metastasis of gastric cancer in collaboration with Chiba University.

We obtained a large number of high-resolution microscopic images of perigastric lymph nodes in cooperation with Chiba University Hospital, and expert pathologists annotated metastatic regions in those nodes. We then developed a neural network algorithm consisting of more than 100 layers to detect metastasis based on these image datasets. Finally, we confirmed that AI can identify metastatic cells as accurately as expert pathologists.

(a) Metastatic regions annotated by pathologists (outlined by solid black lines)

(b) Results of AI image analysis of metastatic regions (shown in green)

Examples of pathological images of perigastric lymph nodes stained with hematoxylin and eosin

# SATLYS Toshiba Analytics AI Services Accelerating Digital Transformation

● Demand Forecasting Technology for On-Demand Ridesharing Services

Toshiba Digital Solutions Corporation has developed a demand forecasting technology to optimize operation plans for on-demand ridesharing services. This technology forecasts both embarkation and disembarkation of passengers.

On-demand ridesharing systems create operation plans based on the locations, dates, and times of advance embarkation and disembarkation reservations. More efficient operation plans can be created if demand for embarkation and disembarkation points is known. However, demand forecasting technology for taxis provides forecasts only for embarkation points.

To address this requirement, we have developed a new technology to forecast the demand for both embarkation and disembarkations points based on the analysis of learned operation data and external environmental data including weather data accumulated by Junpuzi Co., Ltd. over a period of about nine years.

In addition, we have visualized the demand forecast results for a huge number of combinations of embarkation and disembarkation points so as to facilitate reviews by on-demand transport operators, thereby allowing operation schedules to be optimized.
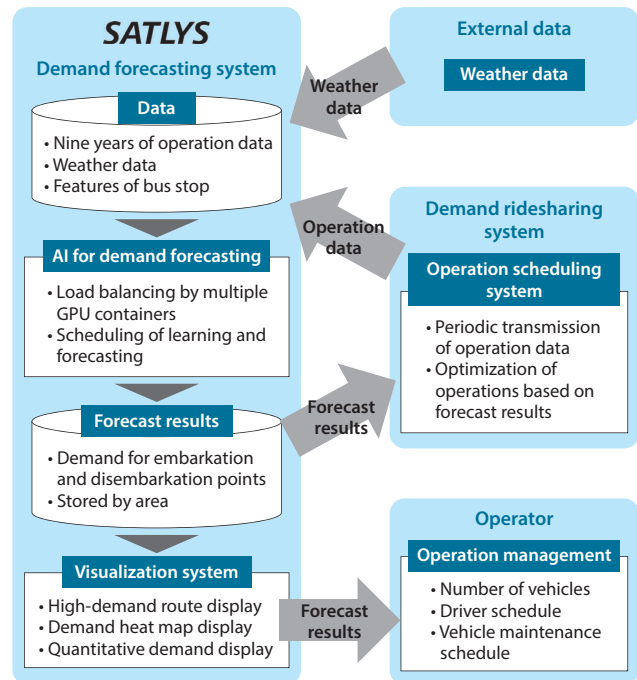
Demand forecasting system for on-demand ridesharing services

● Deep Learning Model Development Platform for High-Quality AI Analytics

In recent years, efforts to develop a wide variety of technologies related to the Internet of Things (IoT), big data, and AI have been actively promoted in the industrial field. However, this trend is now encountering various issues including the lack of an established standard methodology for the development of AI models, differences in experience and skills among AI engineers, and diversity of development tools.

In order to provide high-quality AI analytics services, Toshiba Digital Solutions Corporation has constructed a platform to store and share knowledge and proprietary technologies accumulated through the development of deep learning models. This platform consists of a development framework that standardizes the software structure for model learning and inference, development components designed for reuse of model know-how and proprietary technologies, a development environment using virtualization infrastructure, and development standards that define development processes and model deliverables. This platform helps to reduce the costs of constructing development environments and improve model accuracy, making it possible to create high-quality and high-efficiency services in response to customers' requirements.
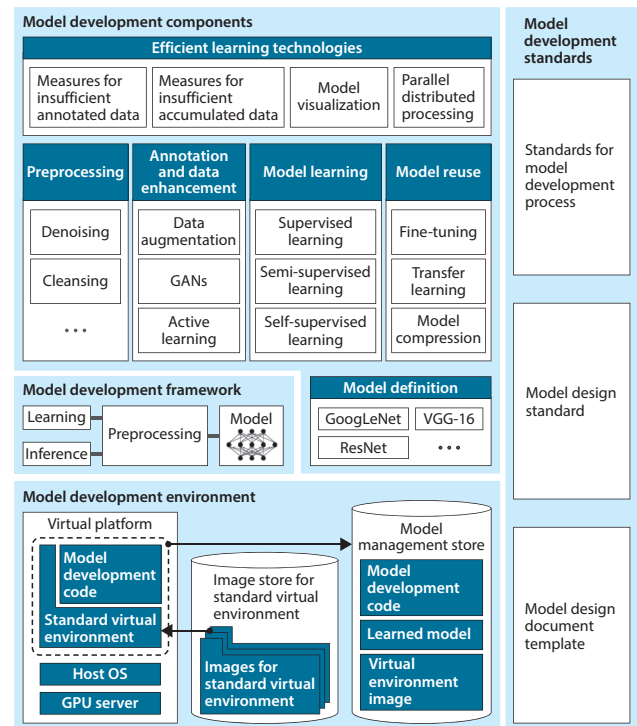
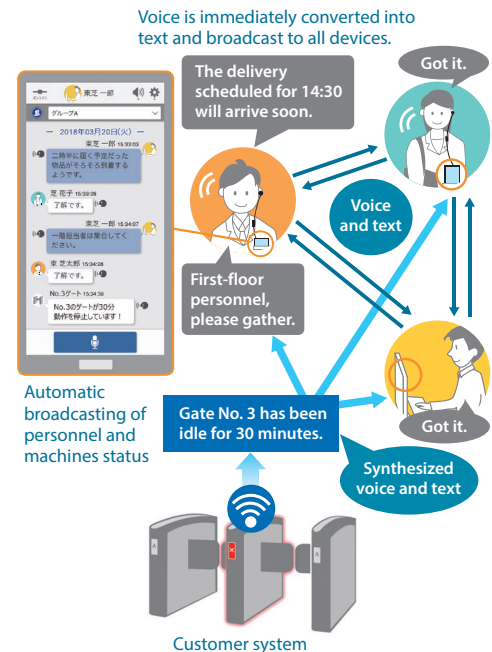Configuration of deep learning model development platform

RECAIUS Toshiba Communication AI Services for Workstyle Innovation
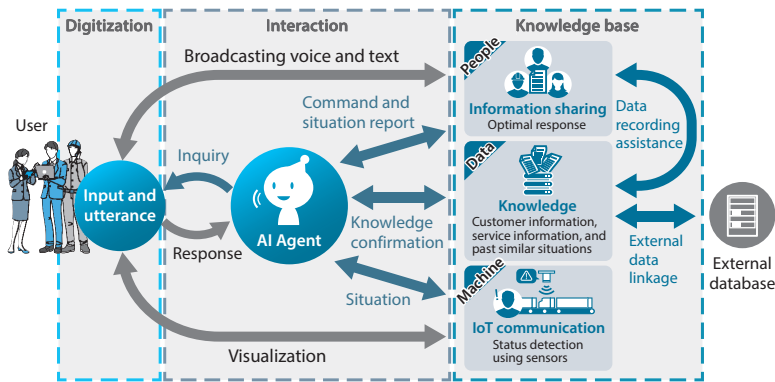
Overview of RECAIUS Toshiba Communication AI

CRM: customer relationship management
＊Screenshots of this product applied to Japanese market

Functional overview of RECAIUS Report Agent

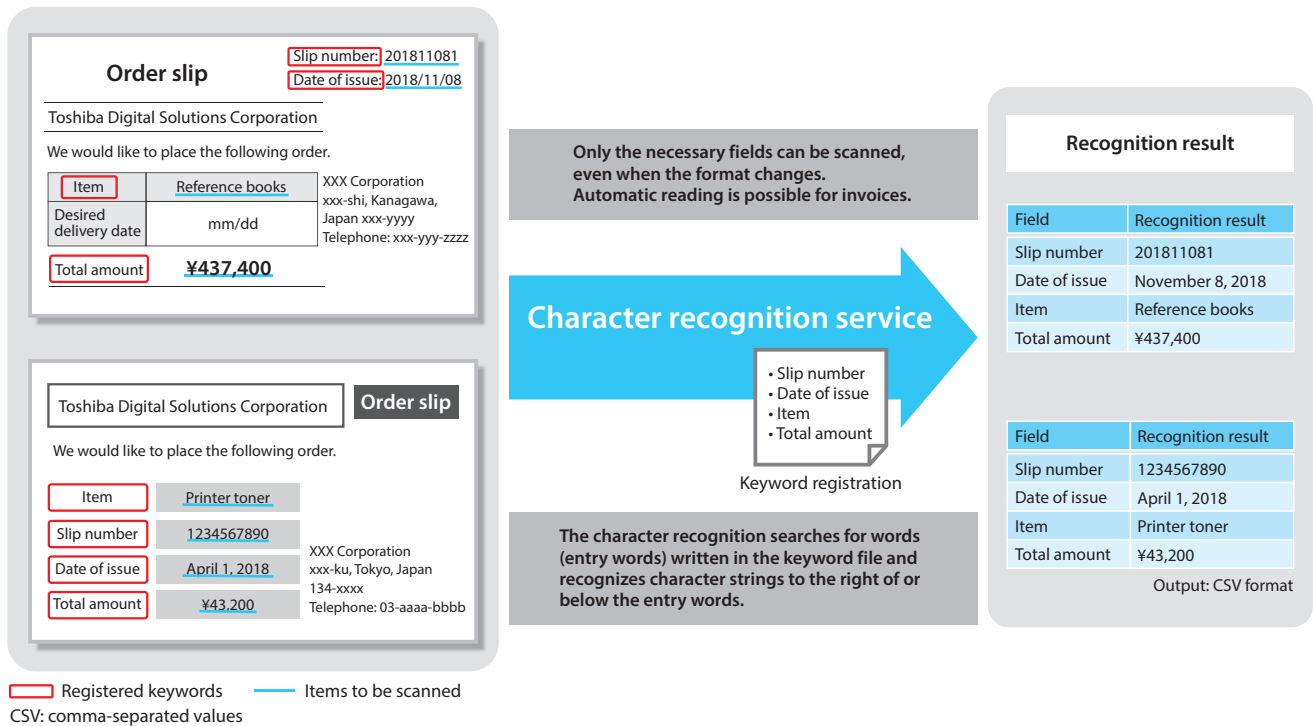Functional overview of RECAIUS Field Voice Intercom Express

With the aim of improving the efficiency and comfort of workplace environments, Toshiba Digital Solutions Corporation has released RECAIUS Report Agent and RECAIUS Field Voice Intercom Express to support sales, field work, and customer engagement tasks.

RECAIUS Report Agent is a service that allows voice-based work reporting through conversations with AI. Users do not need to be concerned with the order, or omission, of report items. If any items are missing, ActiveAI automatically prompts the user to enter them, enabling stress-free reporting in a ubiquitous environment. RECAIUS Report Agent has many patterns of conversation logic necessary for typical professional reporting. These features combine to reduce the time required to create interactive AI conversation scenarios. Building an AI conversation scenario is normally a costly task. However, since programming knowledge is unnecessary for the use of RECAIUS Report Agent, users can build AI conversation scenarios simply by selecting conversation patterns according to the contents of reports. After deployment of a scenario, team members can benefit from the assistance of RECAIUS Report Agent.

RECAIUS Field Voice Intercom Express allows users to share the contents of conversations in real time by means of both text and voice. With conventional wireless intercoms, users can miss announcements if they are not constantly on guard or might find it difficult to identify who is speaking. In contrast, RECAIUS Field Voice Intercom Express incorporates our voice recognition technology to record all utterances together with the utterers' names. Professionals can use this feature to review the communication timeline so as to catch up on missed information, thereby improving customer satisfaction. In fields where communication and information sharing are important, including customer service, security, and nursing, RECAIUS Field Voice Intercom Express helps to improve work efficiency because it allows users to view a log of important messages on their devices and trace their communications for analysis and quality improvement purposes.

We will continue to develop RECAIUS Toshiba Communication AI services to improve people-to-people and people-to-machine communications as a means of enhancing people's lives and workstyles.

Character Recognition of Both Fixed- and Flexible-Form Documents

Order slip

Slip number: 201811081
Date of issue: 2018/11/08

Toshiba Digital Solutions Corporation

We would like to place the following order.

Item	Reference books
Desired delivery date	mm/dd
Total amount	¥437,400

XXX Corporation
xxx-shi, Kanagawa,
Japan xxx-yyyy
Telephone: xxx-yyy-zzzz

Toshiba Digital Solutions Corporation   **Order slip**

We would like to place the following order.

Item	Printer toner
Slip number	1234567890
Date of issue	April 1, 2018
Total amount	¥43,200

XXX Corporation
xxx-ku, Tokyo, Japan
134-xxxx
Telephone: 03-aaaa-bbbb

Only the necessary fields can be scanned, even when the format changes.
Automatic reading is possible for invoices.

Character recognition service

- Slip number
- Date of issue
- Item
- Total amount

Keyword registration

The character recognition searches for words (entry words) written in the keyword file and recognizes character strings to the right of or below the entry words.

Recognition result

Field	Recognition result
Slip number	201811081
Date of issue	November 8, 2018
Item	Reference books
Total amount	¥437,400

Field	Recognition result
Slip number	1234567890
Date of issue	April 1, 2018
Item	Printer toner
Total amount	¥43,200

Output: CSV format

☐ Registered keywords   — Items to be scanned
CSV: comma-separated values

Field reading and searching functions for flexible-form documents

In recent years, optical character reader (OCR) technology has been attracting renewed attention because of its improved character recognition accuracy and high expectations that it will contribute to improved business efficiency when combined with robotic process automation (RPA); that is, business automation using software robots.

With conventional OCRs, however, it is necessary to specify formats as well as the fields to be scanned and their positions for all target documents. The existence of diverse business formats such as invoices, etc. make it difficult to use an OCR to scan business documents. Therefore, the use of OCRs for business applications has been limited to specific tasks.

To address this problem, Toshiba Digital Solutions Corporation has developed a new character recognition function that is capable of recognizing images in printed paper documents. Users only need to define the areas to be scanned via simple mouse operations and register keywords for field searches. This character recognition function simplifies digitization of paper-based documents, making them usable for various office tasks including invoice payments and the registration of applications.

The newly developed character recognition function is characterized by its ability to recognize characters in vouchers, slips, and other nonstandard formats by searching for keywords designated by operators in advance. In particular, characters in invoices can be automatically recognized without keyword registration or any other preprocesses. Furthermore, this function makes it possible to scan handwritten documents and reproduce them as text files. Users can also use a batch processing function and the scanned text can be corrected by registered user.

This character recognition function is available in three deployment options according to the user's business environment: cloud, on-premises software subscription, and an application programming interface (API).

We will apply deep learning-based technology to further improve form identification and image processing performance. We will also enhance our technologies and services to enable high-precision character recognition of a wide variety of forms and paper documents.
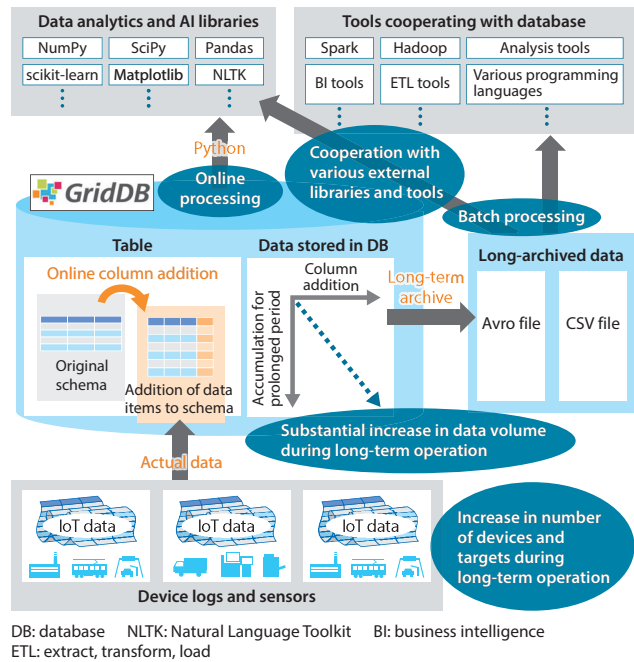
Data Management Technology for GridDB Facilitating Utilization of Long-Stored IoT Data

Toshiba Digital Solutions Corporation provides the GridDB scale-out database to enable users to handle growing volumes of IoT data for prolonged periods.

In long-running IoT systems, the amount of data stored increases exponentially over time. Therefore, data management is crucial to balance the time required to store IoT data and the time required for long and complex queries to retrieve data for analysis.

To facilitate the use of huge amounts of IoT data, we have developed data archiving methods using CSV and other common file formats to accommodate the growth in data volume as well as interfaces to Python programming language, which is commonly used for data analysis to facilitate communication with open-source software (OSS) tools. Furthermore, to simplify the addition of devices to an IoT system during operation, we have incorporated a method for online data structure reconstruction into GridDB and increased the upper limit on the number of data columns.

The enhanced data management and analysis capabilities simplify the handling and increase the value of IoT data. The newly developed technologies are incorporated in GridDB SE/AE 4.1.


DB: database   NLTK: Natural Language Toolkit   BI: business intelligence
ETL: extract, transform, load

Cooperation with external libraries and tools for easy utilization of huge volumes of IoT data accumulated in GridDB database for prolonged period

Industrial IoT Security Reference Architecture

Toshiba Digital Solutions Corporation has defined unique security reference architecture for IoT systems in line with the security levels required by the ever-evolving digital transformation (DX). This security architecture systematizes various technologies related to the security of IoT systems according to their characteristics and provides three levels of security measures based on the criticality of the IoT systems and the skills of cyber-attackers.

The newly developed security architecture is compliant with the International Electrotechnical Commission (IEC) 62443 standards, a set of industrial control system security standards, as well as other security guidelines. The new security architecture is a cost-effective solution that allows the implementation of necessary and sufficient security measures according to the evolution of DX.

We have determined the technical measures for the three security levels based on the security functional requirements of the IEC 62443 standards, which are being implemented in individual IoT systems according to their respective security requirements.
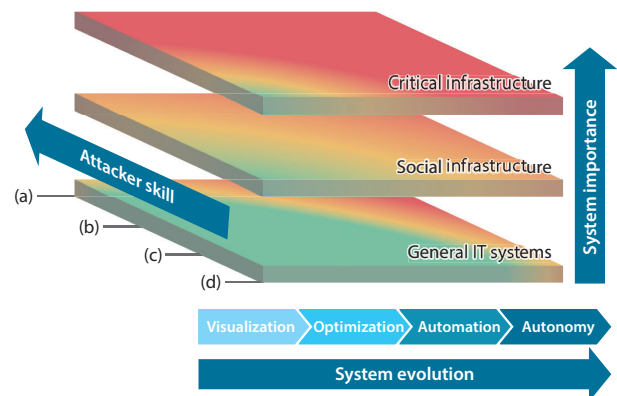
(a) Nation
(b) Specialized organizations (terrorist and crime organizations)
(c) Specialists (black hackers)
(d) Script kiddies

Security level

Low	Medium	High
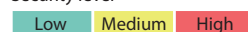
Compliant with IEC 62443 industrial control system security standards

Security reference architecture composed of three security levels

Toshiba Hybrid Crowdsourcing Service

Crowdsourcing is an outsourcing service in which tasks are performed by a large, undefined network of people for such purposes as creating text and images and collecting data in various domains. Crowdsourcing provides an attractive solution for the creation of teaching data for supervised machine learning at low cost, and is attracting the attention of many researchers in the AI field. However, quality and security control remain major concerns.

To address this problem, Toshiba Digital Solutions Corporation is offering the Toshiba Hybrid Crowdsourcing Service, which provides two quality and security control methods: (1) analyzing a customer's requirements and adapting the customer's existing processes to our crowdsourcing system, and (2) securing crowdsourcing workers consisting of dedicated workers in high security, community workers comprising former employees of companies in the Toshiba Group, and unspecified public workers working outside the Toshiba Group at low cost. Furthermore, our crowdsourcing service manages and combines these workers according to the customer's requirements so as to acquire numerous teaching data for voice recognition, voice synthesis, and image recognition systems.
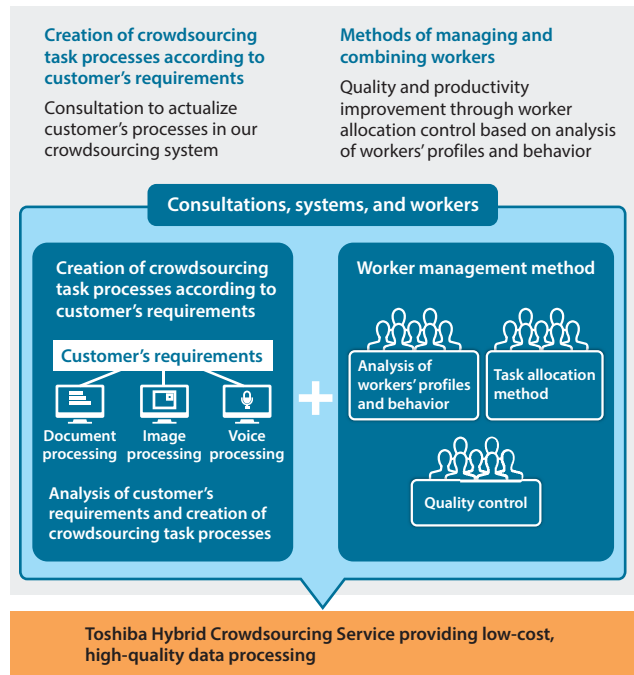
Features of Toshiba Hybrid Crowdsourcing Service

Automated Document Processing Solution Utilizing RPA and OCR for Intellectual Property-Related Businesses

Robotic process automation (RPA), which automates routine business tasks using a PC, is attracting considerable attention. Toshiba Digital Solutions Corporation has been offering an intellectual property (IP) management cloud service to support the patent application and maintenance processes of individual customers. We have now released an automated document processing solution that helps to efficiently digitize necessary data in documents including patent application papers and invoices utilizing RPA and OCR.

Since this solution provides a templated scenario, users can automate routine tasks simply by defining field names, etc. Once these definitions have been made, our newly developed item-spotting OCR technology can automatically detect the positions of fields to be scanned, eliminating the need to perform the time-consuming task of field setting for each form.

Prior to the introduction of the RPA solution, users had to find an item on a search screen while viewing an invoice then switch to another screen to enter information about the item found. This process was necessary for each invoice. Finally, users had to output and check the entries.

In contrast to such manual operations, RPA automates the entire process up to the output of the entered information. Therefore, users only need to put invoice files in a folder, execute a scenario, then check the entered results, making it unnecessary to repeat entries for each invoice.

PDF: Portable Document Format
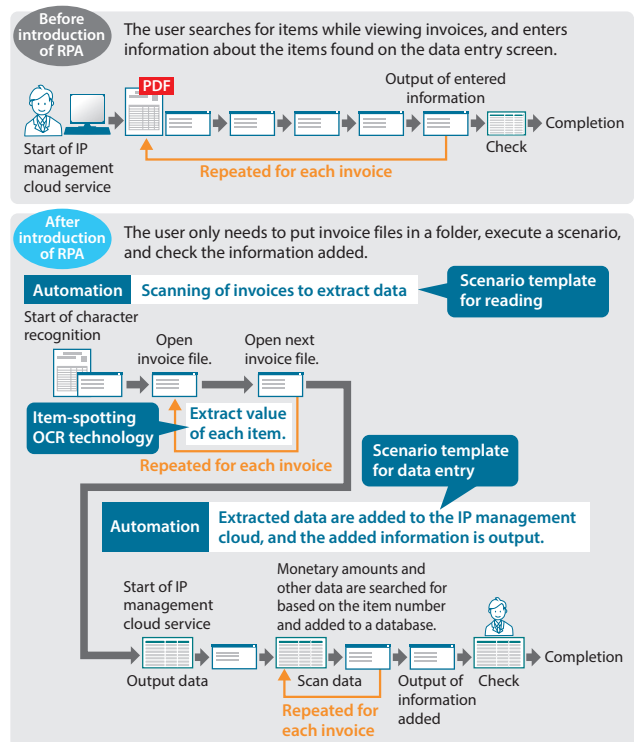
Application of RPA automated invoice processing to IP-related business