Software Technologies in the Data-centric Era

Seiji Kihara and Hiroyuki Tanaka

Abstract

With the arrival of a *data-centric society*, where various types of value are created from real-world data, demand for system infrastructure and software-development methods as business evolves has risen to an even greater level. In this article, we introduce various technological challenges and the efforts of NTT laboratories, particularly the NTT Software Innovation Center, in developing an information-processing infrastructure for the data-centric era to meet the requirements of business evolution and support a data-centric society.

Keywords: information-processing infrastructure, data sharing, software

1. Situation surrounding information-processing infrastructure

Processes for solving real-world issues by using a variety of real-world data are taking on greater importance. There is no shortage of examples of the application of such processes. For example, in the manufacturing industry, factory operational data can be used for forecast management and predicting equipment failure. In the retail industry, weather data can be used to forecast supply and demand and customer behavior data can be used to plan marketing activities and design store traffic flow. We call this process of creating value from data *data valorization*. Besides the intensification of competition between companies and the worsening of environmental and social issues, COVID-19 has forced society to transform into a remote-work society. This state of affairs is forcing society as a whole to undergo digital transformation, and even greater expectations are being placed on data valorization. We believe that such an environment will bring about a *data-centric society* in which value is created from data for all aspects of life in society, such as making decisions and predicting the future. In a data-centric society, the role of the information-processing infrastructure, which provides the foundation for sharing and analyzing vast amounts of diverse data, will become even more critical.

Reviewing the history of technological progress to date, we see that the information-processing infrastructure has evolved steadily. In terms of system infrastructure, processor performance has advanced as a result of first boosting single-core central-processing-unit performance through semiconductor miniaturization then evolving to a multi-core processor architecture. The speed and sophistication of the entire system has advanced due to performance speed-up and scaling as a result of the evolution from single-server processing to distributed and parallel processing. NTT has proposed the concept of the Innovative Optical and Wireless Network (IOWN) [1, 2], which, in addition to hardware and network advancements, calls for improvements in software supporting these technologies—in other words, improvements in system-infrastructure technology. In terms of software-development methods, in addition to the waterfall development model, which is well-suited when large-scale, high-quality software such as mission-critical software is required, other development methodologies have become widely adopted. These approaches include agile development, a method well-suited for web applications, which require rapid development for immediate availability and reiterations for improving usability. New software-development technologies that can support accelerating business transformation are thus

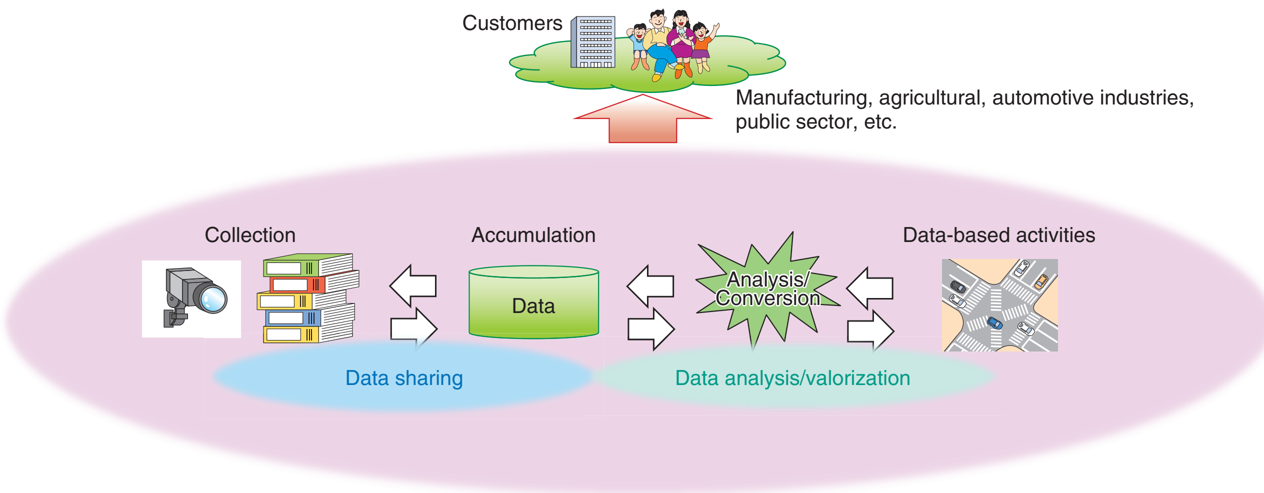

Fig. 1. Value chain in a data-centric society.

in demand.

Against this background, the NTT Software Innovation Center (SIC) is engaged in the research and development of system-infrastructure and software-development technologies that can support a data-centric society and business evolution. Through collaboration with each NTT laboratory, which specializes in related technologies, we are working to develop a future information-processing infrastructure based on IOWN.

2. Information-processing-infrastructure technologies for supporting a data-centric society

The broad flow of data valorization in a data-centric society is shown in **Fig. 1**. First, data related to real-world things and events in various real-world places, such as sensing data and log data, are generated, collected, and accumulated. Next, activities to apply these data to real-world problems are carried out by analyzing and converting these data. In many cases, the data about the real world changed by such activities are repeatedly collected. By reiterating the cycle of data collection, the real world is gradually improved and real-world changes are continuously tracked. With this process of applying feedback to the real world using data as evidence, it is possible to steadily change the real world in a positive direction and provide value to customers in a variety of domains.

We are engaged in the research and development of the following information-processing technologies to support the value chain of such a data-centric society.

(1) Disaggregated computing technologies

In a data-centric society, high-speed processes are required for all aspects of data handling, from data generation to transport, processing, and delivery, to valorize the vast amounts of data generated and used in a data-centric society. NTT has proposed IOWN and is also advancing research activities to achieve a smart society. These goals share in common the demand for high-speed, broadband networks and high-performance computers. To develop such a computer, NTT has proposed the concept of an architecture called *disaggregated computing*, which makes maximum use of photonics-electronics convergence technology (please see "Disaggregated Computing, the Basis of IOWN," NTT Technical Review, July 2021 [3]). We are cooperating with the NTT Network Innovation Center and NTT Device Innovation Center, which operate as part of the NTT IOWN Integrated Innovation Center (established in July 2021 [4, 5]), especially in the area of system software (an example of this effort is introduced in the NTT Technical Review article "Memory-centric Architecture for Disaggregated Computers" [6] in the same issue above).

(2) Data-sharing technologies

To truly use data produced in a data-centric society, it will be necessary to combine data sets separated by geography and networks as well as cross-multiply data sets across organizational and state boundaries. This cross-multiplication requires not only high-speed networks and high-performance computers but also data management and processing in line with the data's generation and accumulation conditions, lifecycle, and permissions based on data owners' agreements. Our effort to solve these challenges is introduced in the feature article "Next-generation Data Hub for Secure and Convenient Data Utilization across Organizational Boundaries" [7] in this issue. This effort is being carried out in partnership with NTT Social Informatics Laboratories (established as part of the reorganization of NTT laboratories in July 2021 [8]).

(3) Data analysis/valorization technologies

To create value from data, computers and information-sharing systems are not enough. We must confront specific problems in each value-producing domain and find appropriate combinations of data and analytic methods for the diverse data sets and methods that can be procured—in short, to formulate the problem. This undertaking requires grounding in statistics and machine learning and understanding of each domain, and, along with these capabilities, support from tools and platforms. Because these capabilities and technologies are needed in every type of field, many NTT laboratories are engaged in their development, with efforts especially centered at NTT Computer and Data Science Laboratories (established as part of the reorganization of NTT laboratories in July 2021 [9]). At SIC, we are working on a processing infrastructure that carries out high-speed conversion of data to value on the basis of problem formulation. A specific example of this effort is introduced in the feature article "High-resolution Multi-camera Analysis Infrastructure to Support Future Smart Cities" in this issue [10].

(4) Software development technologies

In a data-centric society, because businesses and services will need to evolve at an even faster tempo, conventional methods of software development will eventually hit a limit. We are thus engaged in the research and development of technologies to achieve new rapid software development. These technologies include artificial intelligence (AI) that replaces or transcends a portion of human software-development tasks and allows AI and humans to work together. A specific example of our effort is introduced in the feature article "Test-activity Analysis for Efficient Iterative Testing" [11], which discusses technologies to improve the efficiency of regression testing as software functions are reiteratively added or improved.

3. Going forward

To achieve a future information-processing infrastructure that meets the requirements of business evolution and supports a data-centric society, we must address the variety of challenges described in this article. On the basis of IOWN, we are aiming to establish each technology as early as possible in collaboration with partners in various industries and experts in academic and industrial fields.

References

[1] Press release issued by NTT, "Publication of NTT Technology Report for Smart World: What's IOWN?", May 9, 2019 (in Japanese). https://group.ntt/jp/newsrelease/2019/05/09/190509b.html

[2] NTT R&D website, IOWN, https://www.rd.ntt/e/iown/

[3] A. Okada, S. Kihara, and Y. Okazaki, "Disaggregated Computing, the Basis of IOWN," NTT Technical Review, Vol. 19, No. 7, pp. 52–57, July 2021. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202107fa7.html

[4] Press release issued by NTT, "Reorganization of NTT Laboratories," May 12, 2021. https://group.ntt/en/newsrelease/2021/05/12/210512d.html

[5] NTT IOWN Integrated Innovation Center, https://www.rd.ntt/e/iic/

[6] T. Ishizaki and Y. Yamabe, "Memory-centric Architecture for Disaggregated Computers," NTT Technical Review, Vol. 19, No. 7, pp. 65–69, July 2021. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202107fa9.html

[7] K. Ohmura, H. Zhai, S. Katayama, S. Kawai, K. Kashiwagi, K. Umakoshi, Y. Yosuke, and T. Kimura, "Next-generation Data Hub for Secure and Convenient Data Utilization across Organizational Boundaries," NTT Technical Review, Vol. 20, No. 4, pp. 14–20, Apr. 2022. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202204fa2.html

[8] NTT Social Informatics Laboratories, https://www.rd.ntt/e/sil/

[9] NTT Computer and Data Science Laboratories, https://www.rd.ntt/e/cds/

[10] K. Mikami, R. Kurebayashi, X. Shi, N. Inoue, Y. Matsuo, and I. Yamasaki, "High-resolution Multi-camera Analysis Infrastructure to Support Future Smart Cities," NTT Technical Review, Vol. 20, No. 4, pp. 21–25, Apr. 2022. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202204fa3.html

[11] H. Tanno, H. Kirinuki, T. Kawaguchi, M. Tajima, M. Oinuma, and T. Muramoto, "Test-activity Analysis for Efficient Iterative Testing," NTT Technical Review, Vol. 20, No. 4, pp. 26–31, Apr. 2022. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202204fa4.html

Seiji Kihara
Vice President, Head of NTT Software Innovation Center.
He received a B.S. and M.S. in information science from Tokyo Institute of Technology in 1990 and 1992. He joined NTT in 1992. His current research interests include operating systems, computer networks, and open source software. He is a member of the Information Processing Society of Japan (IPSJ), the Institute of Electronics, Information and Communication Engineers (IEICE), and the Association for Computing Machinery (ACM).

Hiroyuki Tanaka
Executive Research Engineer, Director, NTT Software Innovation Center.
He received a B.E. from Kyushu University, Fukuoka, in 1993 and M.E. from Nara Institute of Science and Technology in 1995. He joined NTT in 1995. His current interests include design and implementation of future computing systems.