J. Inf. Technol. Appl. Manag. 31(3): 165~179, June 2024
https://doi.org/10.21219/jitam.2024.31.3.165
ISSN 1598-6284 (Print)
ISSN 2508-1209 (Online)

Why Data Capability is Important to become an AI Matured Organization?

Gyeung-min Kim*

Abstract

Although firms with advanced analytics and machine learning (which is often called AI) capabilities are considered to be highly successful in the market by making decisions and actions based on quantitative analysis using data, the scarcity of historical data and the lack of right data infrastructure are the problems for the organizations to perform such projects. The objective of this study, is to identify a road map for the organization to reach data capability maturity to become AI matured organizations. First, this study defines the terms, AI capability, data capability and AI matured organization. Then using content analyses, organizations' data practices performed for AI system development and operation are analyzed to infer a data capability roadmap to become an AI matured organization.

Keywords : Ai Capability, Data Capability, AI Matured Organization

Received : 2024. 06. 20. Revised : 2024. 06. 27. Final Acceptance : 2024. 06. 29.

※ This research is supported by National Standard Technology Development and Distribution Fund, Korea Planning and Evaluation Institute of Industrial Technology, Ministry of Trade, Industry and Energy

* Professor, MIS, School of Business, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul, 03760, Korea, Tel: +82-2-3277-3581, e-mail: gkim@ewha.ac.kr

1. Introduction

Recently, as computing power permits complex calculations using big data, analytics has advanced from legacy analytics to advanced analytics and machine learning (which is often called AI) using big data ranged structured, semi-structured and unstructured data. Despite the promises of AI, many incumbents are struggling with AI projects due to lack of data. It is estimated that "about 50% of AI projects world-wide are halted or delayed…data quality is among the biggest AI project challenges… companies pursuing such projects generally lack an expert understanding of what data is needed for machine-learning models and struggle with preparing data in a way that's beneficial to those systems [Council, 2019]".

Since the success with AI is dependent on access to data sources, the scarcity of historical data and the lack of right data infrastructure are the problems for the organizations to perform AI projects [Ransbotham et al., 2017]. AI systems are different from legacy analytics in that they revolve around massive data. For example, AI requires not only input datasets to generate the model for prediction but also test datasets to validate the model [Agrawal et al., 2018]. Further, the model's performance is improved with real-world data as it operates. The performance of AI improves with the amount of data. Thus, the biggest challenge that organizations face in deploying these AI systems is to establish data infrastructure and collect data for AI.

To resolve these issues. an increasing numbers of firms are making efforts to have sufficient data for AI systems in terms of volume, variety and velocity by persuading data owners [Def. a senior level employee accountable for the quality of one or more datasets] to share the data if the data owners are within the organization; through data consolidation if the data are fragmented across data sources; through purchase from the market or contract for data supply [Spiekermann, 2019]; and through investment in data to bring new sources when it is not possible to obtain data due to market competition [Daugherty and Wilson, 2018].

In this study, we initially review the characteristics and key capabilities of AI matured organizations through literature review. Using content analysis methodology, data capabilities required to successfully implement AI initiatives are identified. Finally, this research identifies a road map for the organizations to reach data capability maturity to become AI matured organizations. In particular, the following research questions are put forth:

1) What are the characteristics and key capabilities of prominent AI matured organizations?
2) What different data capabilities are needed to become an AI matured organization?
3) What is the road map to reach data capability maturity to become an AI matured organization?

Addressing the lack of study on data capability road map, this study is conducted using content analyses in media and literature to observe organizations' data practices regarding AI system development and infer a roadmap for organizations to improve their data capability.

2. Research Methodology

Content analysis is often used in media

studies to analyze media content as the sample of the study [Neuendorf and Kumar, 2015; Macnamara, 2005]. In this study, the data are collected from diverse sources such as organizational publications, research articles, web site documents and trade publications. Written documents about organizations' data practices regarding AI development and operation are analyzed and the road map to reach data capability maturity are generated inductively.

In designing inductive research, a priori specification of constructs based on extant literature helps link data collection to the study questions [Eisenhardt, 1989]. Based on previous research on AI strategy and strategic alignment, key elements (such as data) required to implement the AI strategy are identified. Using the requirements, content analyses are performed to observe organizations' data practices performed to implement the AI strategy and infer a roadmap for organizations to improve their data capability. The media/written documentation samples are chosen based on theoretical sampling [Glaser and Strauss, 1999]. The sample involved data practices associated with AI, which matches with the objective of this study.

In the following table, the design for this study is described and compared to the recommendations by Eisenhardt [1989].

Although content analysis has disadvantage in terms of methodological rigor, this study used multiple approaches to resolve this issue. To improve construct validity, this study first identifies potentially important constructs from extant literature. To build construct (data capability) measures, several sources of evidence were used [Yin, 1994; Eisenhardt, 1989]. To improve reliability, research steps were detailed and documented as shown in ⟨Table 1⟩.

3. Literature Review

3.1 IT/IS capability

Being defined as "an ability to perform a coordinated tasks utilizing organizational resources [Helfat and Peteraf, 2003: 999].", the capability is regarded as an important constituent of an organization in resource-based theory (RBT) [Byrd and Turner, 2000;

⟨Table 1⟩ Research Design

Eisenhardt's step [1989]	Study design of this research
1. Definition of research question	Three research questions were defined
2. A priori specification of constructs (3. Literature review).	Potentially important constructs such as data capability and capability maturity were identified from the literature on AI strategy and strategic alignment. Those constructs were defined during the literature review.
3. Selection of contents (4. Data practices associated with AI)	Well-known written material about organizations' data practices and capability associated with AI, are analyzed to measure the defined constructs.
4. Multiple data collection methods (4. Data practices associated with AI)	Organizational publications, research articles, trade publications and web site documents are analyzed to provide different perspectives and richness in data.
5. Analyses of Data (5. Road map to improve data capability)	A road map to reach data capability maturity is generated from the analyzed data. A teaching case for data capability was selected to give researcher an acquaintance to the material to be analyzed. This is recommended approach for theoretical study [Quinn, 1980].

Broadbent and Weill, 1997]. IT resources along with organizational processes and other resources such skills and knowledge of the human resources are combined to build an organizational capability such as R&D [Byrd and Turner, 2000; Broadbent and Weill, 1997]. Those RBT scholars assert that inimitable and non-substitutable resources and capabilities determines value creation and sustainability of an organization.

Rooted in RBT, IS/IT capability is the ability of deploying IS/IT resources such as network, application and technology configurations and the knowledge and skills of the IT personnel to support firm's strategies [Peppard and Ward, 2004; Bharadwaj, 2000; Byrd and Turner, 2000; Broadbent et al., 1999; Broadbent and Weill, 1997; Henderson and Venkatraman, 1994]. The alignment of IS/IT strategy with the business strategy are emphasized to obtain strategic benefits from IS/IT [Ward and Peppard, 2002; Peppard and Ward, 2004]. By analyzing business problems and changes in business environment, organizations can not only effectively capture business opportunities and improve business performance but also maximize the return of IT assets [Earl, 1992].

With lack of alignment, business objectives can be ignored; business objectives and strategies cannot be translated into action plans; and applications could be solely dependent on the user' wishes [Teo and Ang, 2001]. Misalignment between business and IT planning can be resolved or minimized through shared domain knowledge and improved communication between IT and business executives [Sauer and Willcocks, 2002; Pepper and Ward, 1999; Reich and Benbasat, 2000; Luftman et al., 2007; Huang and Hu, 2007].

3.2 AI Capability and Value Generation

Based on IS/IT capability definition above, AI capability can be defined as the ability of deploying AI resources (such as internal and external data, algorithm, digital network to connect the firm's ecosystem, and human resources) to support organization's strategy. The AI capability is considered to be critical for organizations to generate value through AI-based predictions, decision makings and offering choices without relying on workers as in traditional business processes. To build AI capability, Ransbotham et al. [2017] suggests the followings to organizations: hiring and educating AI related personnel for training AI as well as continuous learning after AI deployment; executives making a trip to Silicon Valley to better know about the strategic implication of AI for their business and to build AI capability; agile process for AI model development and improvement after deployment; investment in data and data infrastructure; prioritizing AI projects and fostering a fail-fast culture to embrace high uncertainty in AI project success; and learn about AI human teaming. They further elaborate human resource requirements -- adopting AI broadly across an organization requires three types of personnel: 1) technical people, 2) technical people who have business domain knowledge, and 3) people with project management and consulting skills, who can network and bring all those people on board [Ransbotham et al., 2017].

Importance of firm's AI capabilities on value generation, organizational resource requirements to build such capabilities has been iterated by many scholars. While Mikalef et al. [2019] emphasize data driven culture and organizational learning, Gupta et al. [2019] em-

〈Table 2〉 Organizational Resources and Routines Needed to Build AI Capability

AI capability	Human resource	Ransbotham et al. [2017]. Tarafdar et al. [2019]
	Leadership/Management role/Change management	Ransbotham et al. [2017], Gupta et al.[2019]
	Agile process	Ransbotham et al. [2017]
	Data and operating platform	Ransbotham et al. [2017], Tarafdar et al. [2019], Iansti and Lakhani [2020]
	Portfolio management	Ransbotham et al. [2017]
	Culture	Ransbotham et al. [2017], Mikalef et al. [2019]
	Organizational flexibility	Ransbotham et al. [2017], Fountaine et al. [2021]

phasize the role of management. Management roles include anticipating business needs for applying big data, coordinating big data-related activities, understanding implications of the output generated from big data, and working with relevant stakeholders such as functional managers, suppliers and customers. Tarafdar et al. [2019] emphasize "data science competence, business domain proficiency, enterprise architecture expertise, operational IT backbone, and digital inquisitiveness (p. 38)" to build AI capability.

Organizational resources and routines needed to build AI capability are summarized in 〈Table 2〉.

3.3 Maturity of AI capability

Coined by Fountaine et al. [2021], the AI-powered organization refers to an organization that excel at creating value by leveraging AI when opportunity arises. Not only that AI maturity is impacted by organizational resources and routines described above, AI maturity can be affected by agility of the organization to find AI opportunities and capture the opportunities by rearranging organizational resources and routines. Positive effects of IT capability (human, technology and relationship capability) on organizational performance are fully mediated by organization's agility based on survey of executives of manufacturing firms, ICT and finance etc. [Kwahk and Hong, 2011]. In Kim [2010], value creation using IT is determined not only by the extent of alignment between IT strategy and organizational strategy but also by flexibility of organizational architecture such as organizational structure and incentive systems. The ability of the organization to respond rapidly to the changing needs of the business environment and reconfigure their strategy and associated organizational architecture is called dynamic capability [Kim, 2010]. Similarly, the effectiveness of change management in AI adoption is associated with the flexibility of organizational structure, work practices (including human-AI teaming) and culture [Ransbotham et al., 2017].

Value creation using AI is also affected by how AI skills and assets are organized. Organizing for AI relates to where AI skills, resources, activities and authority are located. Placing AI specialists on the hub can establish firm's "AI assets and capabilities such as common analytical tools, data processes and delivery methodologies [Fountaine et al., 2021, p. 16]." This centralization of innovation assets avoids duplication, coherence of firm's development efforts, greater division of labor among specialists, maximizes economies of scale [Shilling, 2013]. In addition, the centralization fosters firm-wide AI deployment using a common framework in idea generation, development and diffusion. However, it might not be as agile as decentralization to fit to divisional needs. A hybrid model consisting of a center of excellence with a number of decentralized groups is another alternative of organizing AI assets [Ransbotham et al., 2017]. The center of excellence provides directions and guidance to other internal units in deploying AI.

As time passes and the maturity of AI capabilities increased, AI processes become standardized and can be reside in spokes [Fountaine et al., 2021]. They assert that the deployability of standardized AI processes/teams and responsible oversight of the AI processes with effective control points are indicators of the maturity of AI capability. When firms "need innovation rapidly some companies put more gray-area strategy and capability building in the hub, so that they can monitor industry and technology changes better and quickly deploy AI resources to head off competitive challenges [Fountaine et al., 2021, p. 16]."

Based on the above discussion, an organization with agility that excel at creating value by leveraging AI capability when opportunity arises can be said to be an AI matured organization.

4. Data Practices Associated with AI

The strategic approach to AI emphasizes business strategies and goals that can be advanced through firm's AI capability. Since organizations with AI capabilities can deliver profoundly improved value propositions to customers, suppliers and employees, increasing number of organizations are taking a strategic approach to capture and analyze data by adopting digital technology and leveraging digital opportunities [Ross et al., 2017]. To support the strategic approach, establishing data supply chains is key. Based on the definition of AI capability in the previous section, data capability for AI is defined as the ability of supplying data resources to support organization's AI strategy.

Advanced analytics and ML (Machine Learning) require a large amount of data in terms of volume, variety, and velocity. Volume means that the organization has to have sufficient data in order for algorithms to extract patterns for predictions. Variety means a heterogeneous data sets from multiple reliable sources to address any inadvertent biases in datasets. For example, the diagnosis done with radiology image can be triangulated via other sources such as a sample of blood. Velocity means that the organization needs real-world data to improve model's performance as it operates. That is, the performance of learning based analytical algorithm is improved with experience as it operates [Agrawal et al., 2018; Schmelzer, 2020].

As an effort to provides variety of data with volume for analytical needs, organizations have long been used data warehouse solution. Data warehouse refers to an analytical database and consists of data from the company's operational database and external databases. An external data refers to data purchased from an external company at a certain cost. When data is loaded into data warehouse, ETL (Extract, Transformation and Loading) process is performed. The first function is data extraction and cleaning function. ETL refers to the function of extracting the necessary data for decision support from various operational and external databases, and transforming conflicting data to fit the predefined schema before being loaded into data warehouse. ETL is done periodically in batch mode. Data warehouse is an essential element of organization's OLAP (On-line analytical processing) environment that provides decision makers with advanced data analysis capabilities in real time.

As an effort to increase velocity of data, organizations develop data supply chain that is "dynamic, constantly evolving and constantly fueled by real-time data [Daugherty and Wilson, 2018, p. 175]". This relates to organizations' implementing digital strategy using digital technology such as IoT devices to capture and fuel real time data from disparate sources. As increasing number of firms implement digital strategy a, new technological solutions such as data lake has been implemented to store unstructured data and to provide some structure for analysis. A data lake is a data repository that stores diverse forms of raw data. The data lake allows users to view raw data holistically.

Additionally, data can be bought from the market or contracted for data supply. The data marketplace goes beyond a data sharing community and is operated by companies monetizes data distribution [Spiekermann, 2019]. Some marketplaces go beyond simple brokering between data supply and data purchasers. They provide specialized data services such as consulting and analytical services. With such value-added services, they want to lock-in customers within a specific environment. In other words, as the value and importance of data grows, a data marketplace has emerged that makes it easy for all companies to access data, regardless of size, without being limited to a specific industry.

For another type of data supply, partnership with other companies, the public domain, data-as-service providers, and external data analysis providers can be made. While some partnerships are on a commercial basis, others are based on agreement. For both cases, defining the terms and conditions under which the data may be used is necessary. During this data supply process, many issues such as privacy, security, liability, and market competition can occur [Harrison et al., 2019]. Thus, policies and agreement among involved parties must be established. When two or more parties are entering a data sharing agreement, the agreement must be based on trust and shared goals among stakeholders. Based on their shared goals, the parties should collaboratively set up agreements and policies on data access, data protection, pricing, copyright, and liability [Allen et al., 2014]. Such policies and agreements are codified in data sharing agreements.

Lastly, data can be crowd sourced and outsourced. Crowd sourcing for AI system is defined as the practice of obtaining data for AI tasks from a large group of people who submit their work, typically via internet. Because AI system requires a large amount of data for training and testing, the crowd sourcing has been around for some time. Not only data itself but also metadata are crowd-sourced. Labeling MRI scans in healthcare and labeling attributes on videos are examples of crowd-sourcing the metadata. Reported problems in crowdsourcing include unreliable labels due to negligence and lack of motivation in crowd workers [Abhigna et al., 2108]. The organizations should review the crowd sourcing results (i.e., tagging results) to see how relevant the results are for the organization's purposes. When organizations have multiple workers doing the same task independently perhaps around the globe for example, one way to checking the relevancy of tagging is to examine if the tags appear multiple times in three to five independent workers' results per task.

5. Road maps to improve data capability

5.1 Data Platform Stage

Collecting, storing and cataloging data from internal and external sources are pre-requisite for an organization to analyze data to innovate products, services, processes and business models. Since advanced analytics and ML involves real-time streaming and rapid ingestion of data, data lake and cloud technology fit to the needs of advanced analytics and ML. The data lake enables organizations to perform data analysis using data from a variety of sources (<Figure 1>).

In the data lake, the data pass through three zones: raw, curation and model zones.

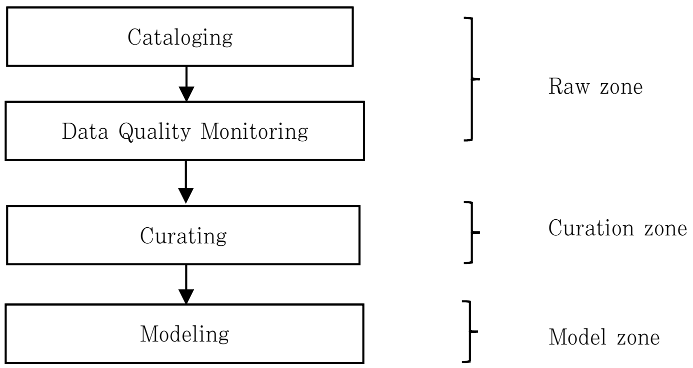
In the Raw zone, data coming from multiple sources must be cataloged; and the data quality has to be monitored before moving to curation.

In the Curation zone, users perform the following activities:

Step 1: Identify data in the archive
Step 2: Get the data; and convince the data owner to share data if necessary.
Step 3: Find and fix any issues with the data; and the data issues include unexpected values, mysterious variables, and data containing too much detail.
Step 4: Make the data findable and usable; and add tags and documentation so that the data can be found and used.
Step 5: Save the data and documentation

In the model zone, the data are fed to the algorithm for training models for prediction.

Among various types of cloud services, DaaS (Data as a Service) providers provide data lake and data warehouse solutions [Asay, 2021], which helps firms collect, store, integrate, query, and analyze data. In addition, DaaS providers are starting or planning to provide various types of data marketplaces that broker data transactions for the firms looking for external data resources to secure analytical competitiveness.



<Figure 1> Data Flow in the Data Lake

One of the most prominent examples of AI-matured organization is Alibaba. Alibaba's expansion to Ant financial was based on data from Alipay's mobile payment platform [Iansti and Lakhanim, 2020]. Alibaba is excelling at connecting businesses through an operating platform, "aggregating the data that flows among them, and extracting its value through analytics and AI [Iansti and Lakhani, 2020, p. 7]." With a common data architecture, data catalog and algorithm libraries, firm can experiment and deploy AI-based digital services and processes across different lines of businesses to capture business opportunities. With "universal set of capabilities [Iansti and Lakhani, 2020, p. 8]" of data sourcing, processing and advanced analytics/ ML model development, organizations can digitally transform their businesses.

With scalable technologies such as IoT, data lake, data warehouse, analytics, ML provided by the cloud service providers, firms can build a data platform to aggregate data from internal and external sources. This also allows the firms to analyze data using advanced analytics and ML to create value when opportunity arises.

5.2 Data Governance Stage

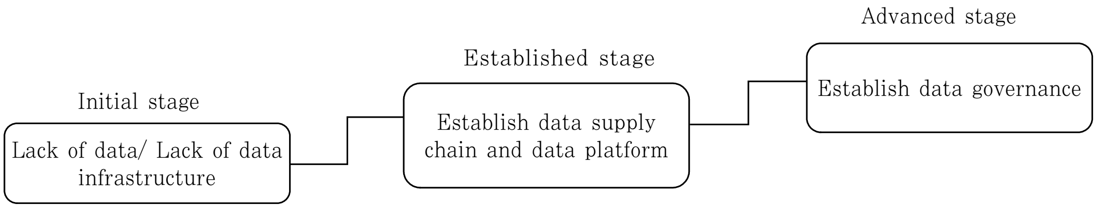
Capturing, cleaning, integrating, curating and storing data could be enterprise-wide activities [Daugherty and Wilson, 2018]. Some firms believe that "democratization of access to AI tools and decision-making power among managers and employees" creates more tangible value from AI and big data. For example, "Airbnb believes that every employee needs to have access to data platforms to make informed decisions; and skills to use and interpret data and tools [Di Fiore, 2018, p.3]." Democratize discovery is also indicated in TMC (Texas Medical Centre) example in [Daugherty and Wilson, 2018].

Enterprise-wide data trust frameworks is essential to ensure organization's activities for data capturing, cleaning, integrating, curating and storing in relation to data lake. Data governance is a system of controlling and directing the use of data which includes "the policies and plans that allocate decision rights and responsibilities to organizational actors involved in managing data [Harrison et al., 2019, p. 175]."

Since datasets for ML are likely to be combined across organizational boundaries, the origin and the context in which the data are created must be recorded. If ML models used datasets taken from original production site without proper records or documentation, decision outcomes produced by the ML model could have distortion.

That is why organizations should have such policies that the sources and context in which data are created must be recorded. When data is taken into data lakes, data have to be created with the meta data and para data (i.e., data about the process by which the data were collected). Keeping records about how data is created, curated, moved and used, allows organization to maintain data quality throughout the data life cycle. With these reasons, organizations should establish data trust framework for which metadata-based data lineage is particularly important [Pdpc, 2019]:

- Backward data lineage showing a trail from use of data until the source of data
- Forward data lineage showing a trail from the source of data until the use of data
- End to end data lineage showing the entire trail of the data combining both back-



〈Figure 2〉 Data Capability Road Map

ward and forward lineage

One of the reasons for the data governance being more important than ever is that data should be in compliance with laws and regulations dealing with privacy [Zarkadakis, 2020]. Specifically, certain data treatments legislated are done accordingly and sensitive data should be protected against unauthorized disclosure or modification. This issue can be resolved with help from relevant data governance solutions that provide the data trust to protect both data providers and data customers [Felici et al., 2013]. DaaS providers should also be held responsible for determining an appropriate approach to enable their client to use trustworthy data for analysis.

A road map to reach data capability maturity is derived below in 〈Figure 2〉.

6. Conclusion and Future Research

This research addresses the lack of research in data capabilities of AI matured organization and identifies key data capabilities to become an AI matured organization. Then key milestones that organizations need to build such capabilities are also identified. According to [Council, 2019], "the cost and hassle of collecting and preparing data comes as a shock for some companies." The results of this study indicate the importance of data capability to become AI-matured organization. The data capabilities not only to support data sourcing, integrating, cataloguing, querying and model development but also safeguard those activities through proper data governance help organization experiment with model development and deployment across businesses responsibly; and enable organizations to capture AI opportunities and become AI matured organizations.

For the future research, study results will be compared with management interview and other IS literature. That is, the road map (themes) arise through the content analyses is triangulate via interviews with an IT manager who had experiences in AI development as well as other IS literature.

As future research, more rigorous maturity model for AI capability as well as data capability for AI, can also be developed. Maturity assessment in certain practice (e.g., business-IT alignment) evaluates current status of organization's practice and provides roadmaps to attain the more ideal status (ultimately to reach the most ideal status in those practices). To assess a maturity level, first, maturity criteria need to be identified. In other words, areas (categories) of the organizational practice to be assessed need to identified. For example, to assess business-IT alignment practice, five areas of business-IT alignment practices are identified by [Luftman, 2003]: communication, competence, governance, partnership, technology scope and skills. Then for each practice area (i.e., maturity criteria), detailed practices that needs to be performed are identified. To determine the level of achievement on each practice area, measurement scale (e.g., 5 Likert scale) needs to be prepared. An average score of the evaluation team for each practice area becomes basis of an average score of each practice area (category) as well as the organization's overall maturity score for the assessed practice. The next higher-level score provides the roadmap for the organization to reach.

Similarly, each dimension of 〈Table 2〉 can be elaborated to assess the maturity of organization's AI capability. Data and operating platform dimension of 〈Table 2〉 can be more detailed to assess the maturity of organization's data capability. An organization with the highest score in every dimension can be said to be AI capability matured organization. Additionally, the dimensions of 〈Table 2〉 can further be studied from which agility dimension (e.g. Alibaba creates Ant financial based on data from Alipay's mobile payment platform) can be sorted out. An organization with the highest scores on both AI capability dimensions and agility dimension can be said to be AI matured organization. The AI maturity model provides "a management system with associated improvement roadmaps that guide strategies to continually improve, develop, and manage [Edward and Tuikka, 2022, p. 21]" the AI capability and agility within their organization.

References

[1] Abhigna, B., Soni, N., and Dixit, S., "Crowdsourcing-A step towards advanced machine learning", international conference on computational intelligence and data science (ICCIDS 2018), Procedia Computer Science, Vol.132, 2018, pp. 632-642.

[2] Felici, M., Koulouris, T., and Pearson, S., "Accountability for data governance in cloud ecosystems", IEEE 5th International Conference on Cloud Computing Technology and Science, 2013, pp. 327-332.

[3] Agrawal, A., Gans, J. and Goldfarb, A., Prediction machine: the simple economics of artificial intelligence, 2018, Harvard Business Review Press.

[4] Allen, C., Jardins, T., Heider, A., Lyman, K., McWilliams, L., Rein, A., Schachter, A., Singh, R., Sorondo, B., Topper, J., and Turske, S., "Data governance and data sharing agreements for community-wide health information exchange: Lessons from the beacon communities", The Journal for Electronic Health Data and Method, Vol. 2, No. 1, 2014.

[5] Asay, M., "Forget about database, what we need is data platform", InfoWorld, 2021.03.17 Available at https://www.itworld.co.kr/news/186858 (accessed 02-02-2022).

[6] Bharadwaj, A., "A Resource-based perspective on information technology capability and firm performance: An empirical investigation", MIS Quarterly, Vol. 24, No. 1, 2000, pp. 169-196

[7] Broadbent, M. and Weill, P., "Management by maxim: How business and IT managers can create IT infrastructures", Sloan Management Review, Vol. 38, No. 3, 1997, pp. 72-92.

[8] Broadbent, M., Weill, P., and Neo, B., "Strategic context and patterns of IT infrastructure capability", The Journal of Strategic Information Systems, Vol. 8,

No. 2, 1999, pp. 157-187.

[9] Byrd, T. and Turner, D., "Measuring the flexibility of information technology infrastructure : Exploratory analysis of a construct", Journal of Management Information Systems, Vol. 17, No. 1, 2000, pp. 167-208.

[10] Council, J., "Data challenges are halting AI projects, IBM executive says", CIO Journal, May 28, 2019, Available at https://www.wsj.com/articles/data-challenges-are-halting-ai-projects-ibm-executive-says-11559035800 (accessed 06-08-2026).

[11] Daugherty, P. and Wilson, J., Human + Machine: Reimagining Work in the Age of AI, Harvard Business Review Press, March 20, 2018.

[12] Davenport, T. and Harris, J., Nature of Analytical Competition: Using Analytics to Build a Distinctive Capability, Harvard Business Review Press, 2007.

[13] Di Fiore, A., "Why AI will shift decision making from the c-suite to the front line", Harvard Business Review, August, 2018.

[14] Earl, M., "Putting IT in its place: Polemic for the nineties", Journal of Information Technology, Vol. 7, 1992, pp. 100-108.

[15] Edward, C. and Tuikka, T., "An organizational maturity model for data spaces: A data sharing wheel approach", Data spaces; Design, Deployment and Future Direction, Curry, E., Scerri, S. and Tuikka, T. Eds., 2022, Springer, pp. 21-42.

[16] Eisenhardt, K., "Building theories from case study research", Academy of Management Review, Vol. 14, No. 4, 1989, pp. 532-555.

[17] Fountaine, T., McCarthy, B., and Saleh, T., "Building the AI-matured organization: Technology isn't the biggest challenge. Culture is", Harvard Business Review, Special Issue Summer, 2021, pp. 10-19.

[18] Gartner, Gartner Top Strategic Technology Trends for 2021, available at https://www.Gartner.com (accessed 01-10-2021).

[19] Glaser, B. and Strauss, A., The Discovery of Grounded Theory: Strategies for Qualitative Research, Aldine Transaction, New Brunswick, N.J, 1999.

[20] Gupta, S., Qian, X., Bhushan, B., and Luo, Z., "Role of cloud ERP and Big Data on firm performance: A dynamic capability view theory perspective", Management Decision, Vol. 57, No. 8, 2019, pp. 1857-1882

[21] Harrison, T., Luna-Reyes, L., Pardo, T., DePaula, N., Najafabadi, M., and Palmer, J., "The Data Firehose and AI in Government: Why Data Management is a Key to Value and Ethics", Proceedings of 20th International Conference on Digital Governance Research (DGO2019), June 18-20, 2019, Dubai, United Arab Emirates.

[22] Helfat, C. and Peteraf, M., "The dynamic resource-based view: Capability lifecycles", Strategic Management Journal, Vol. 24, No. 10, 2003, pp. 997-1010.

[23] Henderson, J. and Venkatraman, V., "Strategic alignment: Leveraging information technology for transforming organizations", IBM Systems Journal, Vol. 38, No. 2, 1999.

[24] Huang, C. and Hu, Q., "Achieving IT-business strategic alignment via enterprisewide implementation of balanced scorecards", Information Systems Management, Vol. 24, No. 2, 2007, pp. 173-184.

[25] Huggett, J., Dipping in Data Lakes,

Introspective Digital Archaeology, Understanding the computational turn in archaeology, Available at https://introspectivedigitalarchaeology.com/2019/07/15/dipping-in-data-lakes/ (accessed 2020-09-17).

[26] ISO/IEC JTC1 17788 INTERNATIONAL STANDARD Information technology – Cloud computing – Overview and vocabulary, 2014-10-15

[27] Kwahk, K. and Hong, M., "The effects of IT competency and organizational learning on firm performance: With a focus on the role of organizational agility", Korean Management Review, Vol. 40, No. 4, 2011, pp. 1075-1108.

[28] Luftman, J., "Assessing IT/business alignment", Information Systems Management, Vol. 20, No.4, 2003, pp. 9-15.

[29] Macnamara, J., "Media content analysis: Its uses, benefits and best practice methodology", Asia Pacific Public Relations Journal, Vol. 6, No. 1, 2005, pp. 1-34.

[30] Mikalef, P., Krogstie, J., Pappas, I., and Pavlou, P., "Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities", Information and Management, 57, 2020, pp. 1-15.

[31] Neuendorf, K. and Kumar, A., "Content Analysis", The International Encyclopedia of Political Communication, John Wiley & Sons, 2015, pp. 1-10.

[32] Pdpc, "A Proposed Model Artificial Intelligence Governance Framework", Personal Data Protection Commission, Singapore, January, 2019.

[33] Pepper, J. and Ward, J., "Mind the Gap", Diagnosing the relationship between the IT organization and the rest of the business", Journal of Strategic Information Systems, Vol. 8, No. 1, 1999, pp. 29-60.

[34] Quinn, J., "An incremental approach to strategic change", McKinsey Quarterly, Winter 1980, No. 1, pp. 34-52.

[35] Ransbotham, S., Kiron, D. Gerbert, P. and Reeves, M., "Reshaping Business with Artificial Intelligence", MIT Sloan Management Review, Fall, 2017.

[36] Reich, B. and Benbasat, I., "Factors that influence the social dimension of alignment between business and information technology objectives", MIS Quarterly, Vol. 24, No. 1. March 2000, pp. 81-113.

[37] Ross, J., Sebastian, I., and Beath, C., "How to develop a great digital strategy", MIT Sloan Management Review, Vol. 53, No. 2, Winter, 2017, pp. 7-9.

[38] Sauer, C. and Willcocks, L., "The evolution of organizational architect", MIT Sloan Management Review, Vol. 43, No. 3, 2002, pp. 41-49.

[39] Schilling, M., Strategic management of technological innovation, McGraw-Hill, 2013.

[40] Schmelzer, R., "How Do You Test AI Systems?", Forbes, Jan. 3, 2020, Available at https://www.forbes.com/sites/cognitiveworld/2020/01/03/how-do-you-test-ai-systems/#473c9b53afd5 (accessed 2020-07-23).

[41] Spiekermann, M., "Data marketplaces: Trends and monetization of data goods", Intereconomics, Vol. 52, 2019. pp. 208-216.

[42] Tarafdar, M., Beath, C., and Ross, J., "Using AI to enhance business operations", MIT Sloan Management Review, Vol. 50, No. 4, Summer, 2019, pp. 37-44.

[43] Teo, T. and Ang, J., "An examination of major IS planning problems", Interna-

tional Journal of Information Management, Vol. 21, 2001, pp. 457-470

〔44〕 Ward, J. and Pepper, J., Strategic Planning for Information Systems, Wiley, 2002

〔45〕 Yin, R., Case Study Research: Design and Methods. Sage Publications Inc., New York, CA., 1994.

〔46〕 Zarkadakis, G., "Data Trusts Could Be the Key to Better AI", Harvard Business Review, November 10, 2020.

Author Profile

Gyeung-min Kim

Gyeung-min Kim is currently a full professor of Management Information Systems in College of Business Administration at Ewha Womans University, Korea. She received her M.S. and Ph.D. degrees in Management Information Systems from Texas Tech University. She earned her B.S. degree in Computer Science from Ewha Womans University in Korea. Her research interests include platform ecosystem, AI governace and capability, global IS sourcing, and gender issues in IS area. Her publications have appeared in the Journal of Organizational Computing and Electronic Commerce and Journal of Systems and Software among others. She is a member of Association for Information Systems (AIS) and Korea Society of Management Information Systems (KMIS).