Questionnaire Survey on Corporate Data Utilization

This survey intends to clarify the actual state of data utilization by Japanese enterprises and to study its impact on the Japanese economy. In concrete, the purpose of this survey is to show the effectiveness of data utilization and to contribute to the examination of measures for the promotion of data utilization by companies.

The results of this survey will be presented at a review meeting held by Ministry of Internal Affairs and Communications. The content of the responses may be used for academic purposes (research purposes such as of members of the study group held by Ministry of Internal Affairs and Communications), but individual responses will not be disclosed or specified to the public when the results are published.

We appreciate your cooperation in your busiest moment in advance.

January 2022

Institute for Information and Communications Policy, Ministry of Internal Affairs and Communications

[Precautions when answering questions]

Holding company system

Regarding the companies that adopted the holding company system, please answer about the status of the entire companies in the group, not on a non-consolidated basis (the status of main business companies in the group, if it's difficult).

About the time of reply

Please indicate the status as of the end of the fiscal year (FY) 2020 (March 31, 2021), unless otherwise specified.

About the person who answers

We would like to ask those who are familiar with the analysis and utilization of your company data to answer questions. However, if it is difficult to do so, we would like you to ask other types of personnel of your company to answer questions as appropriate.

[Explanation of terms in the survey]

Data

Data in this survey means the one currently digitalized and processable by computers, and it is limited to the one that can be utilized for your own company for your product development, expansion of services, and streamlining of business. Data provided to others in the course of business and operations, map data, weather data, and various data sets solely for the purpose of providing to others are not included.

Data Analysis

Data analysis in this survey means the act that extracts information by processing data such as characters, values, images, videos, and sounds. Data analysis includes data analysis using data analysis tools and services provided by others, but does not include data browsing activities, data collection and determination, or alarms (for example, the alarm sounds when the temperature exceeds a certain level).

Data Utilization

Data utilization refers to activities aimed at developing and expanding our own products and services, as well as improving operational efficiency. Data utilization includes data analysis, data browsing, automatic data collection and determination, and alarm generation.

Q. Please fill in your company name. [Mandatory]

Your company name

Q. Please fill in the name of the department of your company that answers the questions of this survey. [Mandatory]

Department name

Q1. Does your company perform data analysis in each business domain? Which <u>one</u> of the following best describes your business domain that perform data analysis? **[Mandatory]**

Business domain	Doing	Not doing, under consideration	Not doing, not considering	Business area does not exist
Management planning and back- office operations (general affairs, human resource management, accounting, finance, legal, investor relations, etc.)	۲	0	0	0
Planning and development of products and services (including research activities)	۲	0	0	0
Marketing (including advertising and publicity activities)	۲	0	0	0
Production, manufacturing and service provision (including sales activities at stores and on the Internet)	۲	0	0	0
Logistics and inventory control (including those pertaining to transport of products)	۲	0	0	0
Maintenance and support (including remote monitoring and customer service)	۲	0	0	0

Q2. How does your company process data and use it in each business domain? Please select all applicable items for each

business domain. [Mandatory]

*This question is displayed only in the business domains for which you select "Doing" in Q1. If you don't select any "Doing"

in Q1, this question will be skipped.

Business domain	Aggregation	Analysis	AI using vendor service	AI developed in- house
Management planning and back- office operations (general affairs, human resource management, accounting, finance, legal, investor relations, etc.)	>			
Planning and development of products and services (including research activities)	Y			
Marketing (including advertising and publicity activities)	Y			
Production, manufacturing and service provision (including sales activities at stores and on the Internet)	V			
Logistics and inventory control (including those pertaining to transport of products)	Y			
Maintenance and support (including remote monitoring and customer service)	Y			

[Explanation of terms]

AI

Artificial intelligence (AI) is a technology that reproduces human and other intellectual activities using computers. It is used for prediction using machine learning and deep learning, and includes not only analysis of relationships and optimal behavior, but also automation, determination of missing/illegal items, and monitoring based on data.

Aggregation

Aggregating by period and by company size, etc.

Analysis

Statistical analysis, etc.

Al using vendor service

AI that is not developed or tuned in-house.

Al developed in-house

AI developed and tuned in-house.

Q3. In your company, when you conduct data analysis, how far back does the data you accumulated go? Which one of the following

best describes each type of data? [Mandatory]

*This question will be skipped if you do not select any "Doing" in Q1.

Data type	Almost daily data only	Up to about a week ago	Up to about a month ago	Up to 3 months ago	Up to about half a year ago	Up to about 1 to 2 years ago	Up to about 2 to 3 years ago	Up to about 3 to 5 years ago	Up to 5 years ago	I do not know	No data is accumulated
Customers' (individuals') basic data (name, address, sex, etc.)	۲	0	0	0	0	0	0	0	0	0	0
Customers' (corporations') basic data (name, address, capital, etc.)	۲	0	0	0	0	0	0	0	0	0	0
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)	۲	0	0	0	0	0	0	0	0	0	0
Action data on the Internet (search and access logs, word-of-mouth, posted data, etc.)	۲	0	0	0	0	0	0	0	0	0	0
Data on human activities (biological information, location information, camera image, etc.)	۲	0	0	0	0	0	0	0	0	0	0
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)	۲	0	0	0	0	0	0	0	0	0	0
Data on nature/public (map information, meteorological information, etc.)	۲	0	0	0	0	0	0	0	0	0	0

Q4. What type of data does your company use to analyze data in each business domain? For each business domain, select **all** applicable data types. **[Mandatory]**

*This question is displayed only in the business domain for which you selected "Doing" in Q1. If you don't select any "Doing," this question will be skipped.

*In this question, the types of data for which you select "No data is accumulated" in Q3 do not appear in the choices. If you select "No data is accumulated" for all items, this question will be skipped.

Q4-1. What types of data does your company analyze in the business domain of "Management planning and back-office operations (general affairs, human resource management, accounting, finance, legal, investor relations, etc.)"?

Data type	Answer
Customers' (individuals') basic data (name, address, sex, etc.)	✓
Customers' (corporations') basic data (name, address, capital, etc.)	
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)	
Action data on the Internet (search and access logs, word-of-mouth, posted data, etc.)	
Data on human activities (biological information, location information, camera image, etc.)	
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)	
Data on nature/public (map information, meteorological information, etc.)	

Q4-2. What types of data does your company analyze in the business domain of "Planning and development of products and services (including research activities)"?

Data type	Answer
Customers' (individuals') basic data (name, address, sex, etc.)	V
Customers' (corporations') basic data (name, address, capital, etc.)	
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)	
Action data on the Internet (search and access logs, word-of-mouth, posted data, etc.)	
Data on human activities (biological information, location information, camera image, etc.)	
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)	
Data on nature/public (map information, meteorological information, etc.)	

Q4-3. What types of data does your company analyze in the business domain of "Marketing (including advertising and publicity activities)"?

Data type	Answer
Customers' (individuals') basic data (name, address, sex, etc.)	
Customers' (corporations') basic data (name, address, capital, etc.)	
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)	
Action data on the Internet (search and access logs, word-of-mouth, posted data, etc.)	
Data on human activities (biological information, location information, camera image, etc.)	
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)	
Data on nature/public (map information, meteorological information, etc.)	

Q4-4. What types of data does your company analyze in the business domain of "Production, manufacturing and service provision (including sales activities at stores and on the Internet"?

Data type	Answer
Customers' (individuals') basic data (name, address, sex, etc.)	
Customers' (corporations') basic data (name, address, capital, etc.)	
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)	
Action data on the Internet (search and access logs, word-of-mouth, posted data, etc.)	
Data on human activities (biological information, location information, camera image, etc.)	
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)	
Data on nature/public (map information, meteorological information, etc.)	

Q4-5. What types of data does your company analyze in the business domain of "Logistics and inventory control (including those pertaining to transport of products)"?

Data type	Answer
Customers' (individuals') basic data (name, address, sex, etc.)	V
Customers' (corporations') basic data (name, address, capital, etc.)	
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)	
Action data on the Internet (search and access logs, word-of-mouth, posted data, etc.)	
Data on human activities (biological information, location information, camera image, etc.)	
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)	
Data on nature/public (map information, meteorological information, etc.)	

Q4 -6. What types of data does your company analyze in the business domain of "Maintenance and support (including remote monitoring and customer service)"?

Data type	Answer
Customers' (individuals') basic data (name, address, sex, etc.)	V
Customers' (corporations') basic data (name, address, capital, etc.)	
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)	
Action data on the Internet (search and access logs, word-of-mouth, posted data, etc.)	
Data on human activities (biological information, location information, camera image, etc.)	
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)	
Data on nature/public (map information, meteorological information, etc.)	

Q5. Please indicate the total amount of data owned by your company as of the end of FY 2020 (March 31, 2021) and how much it has increased or decreased since the end of FY 2019 (March 31, 2020). **[Mandatory]**

[Notes]

Please fill in the calculation method and prerequisites if possible. [Optional]

*This question will be skipped if you do not select any "Doing" in Q1.

Amount	of data
Amount	u uala

Aggregate data amount does not include:

- other entities' data simply stored in data centers, etc. managed by your company

- data that have not been accessed, updated, or used for a long period of time

Estimates of the approximate amount of data are as follows:

- 12 MB for 1 minute video (1 TB for about 1,400 hours)

- 2.5 MB for 1 photographic image file (1 TB for about 400,000 photos)

- 10 KB per page for a MS Word document file: (about 1 TB per 100 million pages)

How to answer

If you do not know the exact amount of data owned by your company, please estimate an approximate value

(Example: 100 TB, 1000 TB, etc.).

The total amount of data increased or decreased is + (positive) if increasing, or - (minus) if decreasing.

(Example: + 10 TB, -5 TB, etc.).

*Enter figures with "+" and "-" in half-width character.

Total data capacity as of the end of FY 2020 [Mandator	1	ТВ
Change from the end of FY 2019 [Mandatory]		ТВ

Calculation method, assumptions, etc. [Optional]

Q6. What percentage of each data type is utilized in your company? Please answer so that the total is 100%. If you don't know the exact percentage, please calculate approximate figures (e.g., 25%, 50%, etc.). [Mandatory]

*This question will be skipped if you do not select "Doing" in Q1.

*In this question, the types of data for which you select "No data is accumulated" in Q3 do not appear in the choices. If you select "No data is accumulated" for all the types of data, this question will be skipped.

Data type	Answer	
Customers' (individuals') basic data (name, address, sex, etc.)		%
Customers' (corporations') basic data (name, address, capital, etc.)		%
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)		%
Action data on the Web (search and access logs, word-of-mouth, posted data, etc.)		%
Data on human activities (biological information, location information, camera image, etc.)		%
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)		%
Data on nature/public (map information, meteorological information, etc.)		%

Q7. What types of data were obtained from external sources in FY 2020 among the data used in your company? Choose <u>all</u> that apply. [Mandatory]

that apply. [Mandatory]

*This question will be skipped if you do not select "Doing" in Q1.

How to obtain data	Answer
Using domestic data sales company	✓
Using overseas data sales company	
Using what domestic vendors/partners own	
Using what overseas vendors/partners own	
Using data from domestic public agencies	
Using data from overseas public agencies	
Not using external data	

Q8. How much did your company spend in FY 2020 to obtain the external data you answered in Q7? Please fill in the total amount. If you do not know the exact amount, please calculate an approximate figure (e.g., 1 million yen, 10 million yen,

etc.). [Mandatory]

*This question will be skipped if you do not select any "Doing" in Q1.

*This question will be skipped if "Not using external data" is selected in Q7.

	×ten
Amount expended	thousand
	yen

Q9. What is the percentage of external data used in Q7 for the types of data used in your company that you answered in Q6? Which one of the following best describes each type of data? If you don't know the exact percentage, pick the <u>one</u> that you think most likely applies. [Mandatory]

*This question will be skipped if you do not select any "Doing" in Q1.

*In this question, the types of data for which you select "No data is accumulated" in Q3 do not appear in the choices. If you select "No data is accumulated" for all items in Q3, this question will be skipped.

*This question will be skipped if you select "Not using external data" in Q7.

Data type	0%	Less than 5%	5~10%	10~20%	20~30%	30~40%	40~50%	50~60%	60~70%	70~80%	80~90%	90% or more
Customers' (individuals') basic data (name, address, sex, etc.)	۲	0	0	0	0	0	0	0	0	0	0	0
Customers' (corporations') basic data (name, address, capital, etc.)	۲	0	0	0	0	0	0	0	0	0	0	0
Data of communication with customers, etc. (daily sales reports, purchase history, inquiry history, results of market research, etc.)	۲	0	0	0	0	0	0	0	0	0	0	0
Action data on the Internet (search and access logs, word-of-mouth, posted data, etc.)	۲	0	0	0	0	0	0	0	0	0	0	0
Data on human activities (biological information, location information, camera image, etc.)	۲	0	0	0	0	0	0	0	0	0	0	0
Data on the operation of machines, equipment, vehicles, etc. (operational situation, location information, speed, etc.)	۲	0	0	0	0	0	0	0	0	0	0	0
Data on nature/public (map information, meteorological information, etc.)	۲	0	0	0	0	0	0	0	0	0	0	0

Q10. In your company, is there a person responsible (e.g., CIO, CDO, CDXO) for data utilization? What is the **most** applicable?

[Mandatory]

*This question will be skipped if you do not select "Doing" in Q1.

Whether or not there is a person in charge of managing data utilization	Answer
Full-time employee in charge (could be in charge of all digital work including data utilization)	۲
Full-time employee in charge (could be in charge of wider area including digital work)	0
None	0

Q11. What kind of environment is prepared in your company so that data can be utilized? What is the **most** applicable? [Mandatory]

*This question will be skipped if you do not select any "Doing" in Q1.

Environment for data utilization	Answer
Built an environment enabling data utilization enterprise-wide	۲
Built an environment enabling data utilization among multiple departments	0
Built an environment enabling data utilization in a single department	0
Built an environment enabling data utilization only for specific tasks	0

Q12. What kind of corporate regime (department/person in charge) does your company apply for data analysis? Choose <u>all</u> that apply. If you select "Other," please specify your answer. [Mandatory]

*This question will be skipped if you do not select "Doing" in Q1.

Data analysis system	Answer
Employee in charge of data analysis department	V
Data analysts in each business unit	
Employee in business units are not data analysis experts	
Outsourcing	
Joint analysis involving other enterprises such as alliances and consortia	
Other	

Other

Q13. Please indicate the number of employees of your company whose primary function is data analysis (including parttimers, contract employees, temporary employees, etc., in addition to regular employees). If you do not know the exact number of the employees of this kind, please fill in an approximate figure (e.g., 10 people, 100 people, etc.). [Mandatory] *This question will be skipped if you do not select "Doing" in Q1.

Q14. How much did your company spend on data storage in FY 2020? If you do not know the exact amount, please calculate an approximate figure (e.g., 100,000 yen, 1 million yen, etc.). [Mandatory]

Please fill in the calculation method and prerequisites if possible. [Optional]

*This question will be skipped if you do not select "Doing" in Q1.

The amount that your company spent for the storage of data include expenses for "devices" such as servers and storage for storing data, for "software" for database management, and for "data storage services" using cloud computing, etc.

[Notes]

	imes ten
Amount spent [Mandatory]	thousand
	yen

Calculation method, assumptions, etc. [Optional]

Q15. How effective was the utilization of data at your company in terms of input (e.g., cost reduction through improving operational efficiency) in FY 2020? Which <u>one</u> of the following best describes its impact on each business domain? [Mandatory] *This question is displayed only in the business domains for which you select "Doing" in Q1. If you don't select any "Doing" in Q1, this question will be skipped.

Business domain	Very effective	Somewhat effective	Neutral	Not so effective	Not at all effective
Management planning and back- office operations (general affairs, human resource management, accounting, finance, legal, investor relations, etc.)	۲	0	0	0	0
Planning and development of products and services (including research activities)	۲	0	0	0	0
Marketing (including advertising and publicity activities)	۲	0	0	0	0
Production, manufacturing and service provision (including sales activities at stores and on the Internet)	۲	0	0	0	0
Logistics and inventory control (including those pertaining to transport of products)	۲	0	0	0	0
Maintenance and support (including remote monitoring and customer service)	۲	0	0	0	0

Q16. How effective was the utilization of data in terms of <u>output (increase in sales, etc.)</u> at your company in FY 2020? Which **one** of the following best describes its impact on each business domain? **[Mandatory]**

*This question is displayed only in the business domains for which you select "Doing" in Q1. If you don't select any "Doing" in Q1, this question will be skipped.

Business domain	Very effective	Somewhat effective	Neutral	Not so effective	Not at all effective
Management planning and back- office operations (general affairs, human resource management, accounting, finance, legal, investor relations, etc.)	۲	0	0	0	0
Planning and development of products and services (including research activities)	۲	0	0	0	0
Marketing (including advertising and publicity activities)	۲	0	0	0	0
Production, manufacturing and service provision (including sales activities at stores and on the Internet)	۲	0	0	0	0
Logistics and inventory control (including those pertaining to transport of products)	۲	0	0	0	0
Maintenance and support (including remote monitoring and customer service)	۲	0	0	0	0

Q17. In the market where your company sell your main products and services, how many competitive companies do you have in Japan and the world in total? What is the most applicable? If you don't know the exact number, pick the <u>one</u> that you think most likely applies. [Mandatory]

Number of competitors	Answer
0 company	۲
$1 \sim 2$ companies	0
3 ~ 5 companies	0
$6 \sim 10$ companies	0
11 ~ 20 companies	0
$21 \sim 50$ companies	0
$51 \sim 100$ companies	0
101 + companies	0

Q18. What are the purposes of data utilization in your company? Choose <u>all</u> that apply. If you select "Other," please specify your answer. [Mandatory]

*This question will be skipped if you do not select "Doing" in Q1.

Purpose of data utilization	Answer
Development of new products and services	
Improving the quality of existing products and services	
Improvement of business processes	
Improvement in labor productivity	
Reduction in labor inputs (man-hours or number of workers)	
Increasing added value	
Reduction in costs of raw materials, outsourcing, inventory management	
Other	

Other

Q19. What are your company's internal issues regarding data utilization? Choose <u>all</u> that apply. If you select "Other," please specify your answer. [Mandatory]

Internal issues in data utilization	Answer
Organizational barriers, e.g. lack of management understanding	J
Human resource barriers, e.g. lack of employees with know-how	
Lack of external resources, e.g. appropriate consulting or software	
Lack of business management models or successful use cases	
Concerns about data protection and security	
Financial difficulty with the cost burden	
Other	

Other

Q20. What are your company's external issues regarding data utilization? Choose <u>all</u> that apply. If you select "Other," please specify your answer. [Mandatory]

External issues in data utilization	Answer
Used for purposes other than intended (abusive use)	
Our intellectual resources may be used by our competitors	
Concern about inappropriate handling of personal data	
Unable to receive full revenue generated from the data provided to external parties	
Uncertainty about data governance by counterparties of the transactions	
Absence of a fair trading market	
Other	

Other

Q21. Regarding the utilization of data, some opinions suggest that there are expectations for open API initiatives in the financial sector. Please describe issues and challenges for your company in further promoting business using open APIs. **[Optional]**

*This question appears only for those in the banking industry.

Problems and issues with open APIs

Q22. How does your company use AI? Which of the following **best** describes your company's view on each type of AI technologies? **[Mandatory]**

Types of AI technologies	Utilizing Interested but not utilizing		Not interested	
Technology for recognizing, analyzing, and generating text data (sentences & languages)	۲	0	0	
Technology for visual data recognition, analysis & generation	۲	0	0	
Technology for sound/voice data recognition, analysis & generation	۲	0	0	
Technology for machine operation & control	۲	0	0	
Technology for optimization & inferencing	۲	0	0	

Q23. How does your company collect and analyze data for analysis using AI? Which of the following <u>best</u> describes how your company deals with such data related to each type of AI technologies? **[Mandatory]**

*In this question, only the types of AI technologies selected as "Utilizing" in Q 22 are displayed in the choices. If you select no "Utilizing" choice in Q22, this question will be skipped.

Types of AI technologies	Data collection & analysis both in- house	Data collection in- house, leverage results of vendor's analysis		Vendor aggregates and analyzes data; leverage the vendor's results
Technology for recognizing, analyzing, and generating text data (sentences & languages)	۲	0	0	0
Technology for visual data recognition, analysis & generation	۲	0	0	0
Technology for sound/voice data recognition, analysis & generation	۲	0	0	0
Technology for machine operation & control	۲	0	0	0
Technology for optimization & inferencing	۲	0	0	0

Q24. How does your company evaluate the effectiveness of using AI? For each purpose, choose the one that <u>best</u> describes it. If you select "Other," please specify your answer. **[Mandatory]**

Purpose of using AI	Very effective	Somewhat effective	Neutral	Not so effective	Not at all effective	Not used for this purpose
Development of new products and services	۲	0	0	0	0	0
Improving the quality of existing products and services	۲	0	0	0	0	0
Improvement of business processes	۲	0	0	0	0	0
Improvement in labor productivity	۲	0	0	0	0	0
Reduction in labor inputs (man- hours or number of workers)	۲	0	0	0	0	0
Increasing added value	۲	0	0	0	0	0
Reduction in costs of raw materials, outsourcing, inventory management	۲	0	0	0	0	0
Other	۲	0	0	0	0	0

*This question will be skipped if you do not select any "Utilizing" in Q 22.

Other

Q25. How has the use of AI changed the number of employees (including part-timers, contract employees, temporary employees, etc., in addition to regular employees) in your company? What is the most applicable? **[Mandatory]**

*This question will be skipped if you do not select any "Utilizing" in Q 22.

Changes in the number of employees after starting to use AI	Answer	
Utilizing AI led to a cut in staffing or shifting employees to different tasks	۲	
Hired new employees for work created by the use of AI	0	
No change in number of employees due to AI use	0	

Q26. To what extent has the number of employees in your company (including part-timers, contract employees, temporary employees, etc., in addition to regular employees) increased or decreased by the use of AI? If the number of employees has increased, add + (plus) and indicate the change in percentage compared to before using AI. If the number has decreased, add - (minus) before the percentage figure. If you do not accurately count the change in the number of employees, please calculate an approximate figure (for example, -1%, + 5%, etc.). [Mandatory]

Please specify the calculation method and prerequisites if possible. [Optional]

*This question will be skipped if you do not select any "Utilizing" in Q 22.

*This question will be skipped if you select the choice of "No change in number of employees due to AI use" in Q25.

Increase/decrease (percentage) in the number of employees that the use of AI has brought about [Mandatory]

Calculation method, assumptions, etc. [Optional]

%

Q27. Please specify requests or opinions regarding the improvement of this survey's questions or implementation method if

any. After reviewing, we will take advantage of your requests and opinions for planning the next survey. [Optional]

Requests and opinions concerning this survey