Tentative
Translation

The Conference toward AI Network Society
The Committee on AI Economy
2020 Report (Outline)

July 2020

Chapter 1 Background (History from the establishment of “the Committee on AI Economy” to the present)

Establishment of “the Committee on AI Economy”

- It was established with the aim of considering what kind of social economy should be pursued and how the basic policy and medium-to-long term strategies should be through the promotion of the implementation of AI in society in January 2019.

Outline of “The Committee on AI Economy Report Outline in the Conference toward AI Network Society” (May 2019)

- In the report published in May 2019, the directivity was pointed out in terms of the following 5 key points.

(1) The directivity to be pursued in AI economy (Realization of ”Inclusive AI Economy Society")
(2) Ideal AI investment for sustainable economic growth and the improvement of productivity
(3) How the industrial infrastructure that supports AI economy (Labor/R&D) should be
(4) Japanese enterprises’ global competitiveness over the utilization of AI
(5) Ideal basic policies and strategies on AI economy

The scope of consideration in this report

- ”The Committee on AI Economy” that resumed its consideration in December 2019 pointed out the following 3 items as items to be further considered. In addition, the Special-Interest Group on Data was established with the aim of considering professional/technical items on data economy policy.

(1) Consideration on how the utilization of data requested for the implementation of AI in society
We will consider how the policies should be for the promotion of the utilization of AI/data by grasping how the data should be utilized on the premise that all entities are involved in the implementation of AI in society.

(2) Consideration on the data economy policy in the AI era
In the report published in May 2019, we positioned data as one of production factors, and showed how significant it is to estimate its impact on the improvement of productivity and the necessity of deepening the discussion on a fair remuneration about data. We will consider the estimation method of the value of data and how fair remuneration should be depending on the effects/value of data in order to launch these items.

(3) Consideration on the image of ”Inclusive AI Economy Society"
We look at the future image of the ”Inclusive AI Economy Society“ that can be reached after overcoming the issues shown in (1) and (2).

Chapter 2 The ideal way to utilize data required for the Implementation of AI in Society

The current status of the utilization of data in Japan

- The global economy has changed from a finance-led economy to the one led by digital companies that utilize ICT and data. Japan's competitiveness in the field of digital ICT has been stagnant.
- Regarding the utilization of data in enterprises, there are big disparities between large-scale enterprises and SMEs, and enterprises in large city areas and ones in rural areas.
- Many Japanese enterprises realize that their initiatives for the utilization of data including data collection and analysis are lagging behind because of the lack of human resources for the utilization of data, issues of expenses for the utilization of data and the lack of data collection/storage, even though they are eager to utilize data.
- Many Japanese companies consider that utilizing data in their company was effective.
- Consumers are careful about providing their personal information, behavior history, etc.

Based on **the current status on the utilization of data in Japan**, consider **the ideal way to utilize data required for the implementation of AI in society**

The ideal way to utilize data required for the implementation of AI in society

- In enterprises, ①Position data/AI in management strategies and business architectures, ②Build physical data collection system based on the utilization of data, organizational frameworks for the utilization of data and data infrastructure in organizations, and ③Eliminate or alleviate the anxiety of consumers who provide information as well as respond to the current data protection regulation for customers.
- In administration, accelerate the initiatives to become a "digital government." It's important to promote opportunities to utilize public data such as open data catalog sites, DATA.GO.JP and e-Stat.
- Through the COVID-19 pandemic, the importance of communication network and ICT tools was realized, and it is expected that the awareness of enterprises and individuals will be transformed, going forward. In order to fully exert the potential of communication network and ICT tools, it is essential to realize the ideal way to utilize data.

The functions/roles of data as a "new property" and estimation methods of the effects/value of data

- It's difficult to define or describe the value of data uniquely, because data have various functions/roles and characteristics. Data (or data-related things). The approaches of estimation methods of the value of data can be roughly divided into three categories.

Table 1: Estimation Methods of the value of data

Approaches	Methods	Tools	Limitations	Issues
Cost-based Approach	• Calculation method which pays attention to costs related to the creation, management, and utilization of data (personnel, and time, etc.) It is based on the assumption that the value of data is worth the cost.	• Calculating personnel expenses • Calculating time costs • Calculating security measure costs	• This approach can't take into account quality and productivity. • Since prices or quality of ICT equipment are diverse, the estimated value of data may be affected by other factors. • It is difficult to intuitively envision a relationship between cost and value.	• Which aspect of cost can be appropriately ascertained in order to define the value of data?
Market-based Approach	• Calculation method which is based on the market prices of similar products or of willingness of users to pay for those products	• Market prices (data transactions, etc.) • Representation preference method • Case studies of M&A, negative effects • Experimental works	• Lack of data for estimates • In the case of stated preference method, the result can be affected by the user's bias. • The value could be perceived differently between the business side and the consumer side.	• How is data in general measured (not individual data)? • Which point of view is more appropriate for data measurement, the business side or the consumer side?
Impact-based Approach	• Calculation method where the effects (productivity, and revenue, etc.) of utilizing data is estimated	• Econometric analysis (production function analysis, and regression analysis, etc.) • Experimental works	• Statistical estimates are needed. • It is difficult to differentiate between the effects of one type of data and the effects of other types of data.	• It is necessary to clarify subjects for which the value of data is to be estimated (the definition and the value of data) • How can a good analysis model be established based on consideration of complementary elements which are necessary for data to create value?

Source : Prepared based on a published document, etc.

The functions/roles of data as a "new property" and estimation methods of the effects/value of data

- The study conducted empirical analysis with the Production Function that is an impact-based approach as a feasible method to estimate the value of data, noting organizing the functions/roles of data and general purposes/reproductivities applicable to industries and nations, etc.
- As a result, it turned out that the utilized data amount/number has a positive relationship with the added value as well as other production factors (capital and labor).
- Regarding the above mentioned, the status of utilizing data in the following actual business scenes is assumed as follows.

※Please note that this is just an “assumption”, and it doesn’t show each company’s initiative learned through a hearing survey, etc.

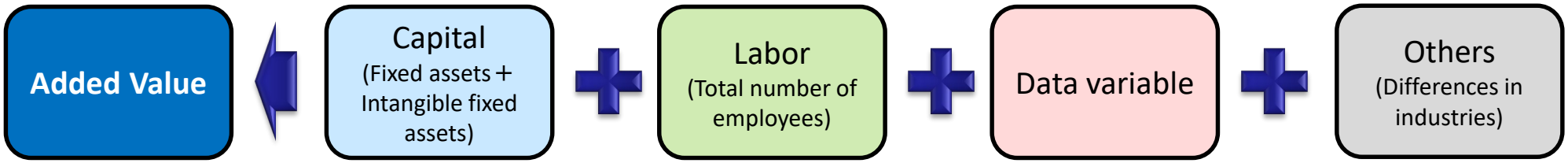
- ① Manufacturing: The companies that grasp the operational status of facilities and detect the abnormality of facilities tend to have realized stable production by visualizing sensor data, etc.

② Retailing: The sales of the companies that finely classify the types of customers from customer attributes and past purchase data, etc. and take optimum marketing measures tend to have increased.

③ EC sites: The sales of EC sites that recommend optimum products for customers based on their customers’ favors and past behavior pattern data tend to have increased.

Table 2: Image of estimation of the value of data through empirical analysis

Analyze the relationship between the owning/utilization of data and output.
As a result, it turned out that the utilized data has a positive relationship with added value.



The functions/roles of data as a "new property" and estimation methods of the effects/value of data

- The study positioned data as a production factor, as with other production factors (capital and labor), and conducted empirical analysis with the Cobb-Douglas Production Function modified to include data as a factor (linear homogeneity was not assumed).

$$V = A_o K^{\alpha} L^{\beta} Data^{\gamma}, \qquad \log(V) = \log A_o + \alpha \log(K) + \beta \log(L) + \gamma \log(Data) + Industry\ Dummy^{*}$$

*Classified into manufacturing and non-manufacturing

Results from an empirical analysis showed that the utilized data amount/number has a positive relationship with the added value as well as other production factors (capital and labor), and the added value increased by 0.05% when the utilized data amount/number increased by 1%. However, it should be noted that this reflects the current status of initiatives of the utilization of data, and that increasing the utilized data amount/number does not necessarily lead to an increase in added value, and that the results show the average effects between enterprises who were successful by utilizing data and those who were not.

Table 3: Analysis Results

Data variable	Sample size	Adj R^2	*K* (Capital)	*L* (Labor)	*Data* (Data)
Utilized data amount (= total data amount x ratio of data used for analysis)	258	0.8343	0.44 ◎	0.50 ◎	0.05 ◎
Utilized data number (= total number of data x ratio of data used for analysis)	135	0.8157	0.34 ◎	0.55 ◎	0.07 ◎
Utilized data amount (FY2015)	258	0.8332	0.44 ◎	0.51 ◎	0.05 ◎
Utilized data number (FY2015)	135	0.8136	0.34 ◎	0.56 ◎	0.06 ○
Externally acquired data amount (= total data amount x ratio of data acquired from outside)	267	0.8401	0.47 ◎	0.47 ◎	0.05 ◎
Externally acquired data number (= total number of data x ratio of data acquired from outside)	140	0.8193	0.35 ◎	0.51 ◎	0.07 ◎
Internally owned data amount (= total data amount - externally acquired data amount)	267	0.8379	0.47 ◎	0.48 ◎	0.05 ○
Internally owned data number (= total number of data - externally acquired data number)	140	0.8134	0.35 ◎	0.55 ◎	0.06 ○
Total data amount x data utilization rate (use area/type/processing method)	261	0.8334	0.44 ◎	0.49 ◎	0.05 ◎
Total number of data x data utilization rate	137	0.8150	0.32 ◎	0.58 ◎	0.05 ○
Total data amount x data diversity (diversity of data acquisition and provision)	173	0.8537	0.53 ◎	0.42 ◎	0.04 ○
Total number of data x data diversity	85	0.8576	0.41 ◎	0.49 ◎	0.06 ◎

※Variables not described as FY2015 are values for FY2018.
The ratio of externally obtained data was corrected based on the situation of external data acquisition.

(Note) ◎:significance level 1%、○: 5%、△: 10%

The ideal fair remuneration according to the effects/value of data

- In many cases, data are treated as capital, but there are some opinions that data should be treated as labor and individuals who can produce data should be rewarded financially.
- As for paying remuneration for the users who produce data, while it's positively seen from the standpoint of "data as labor", it's seen both positively and negatively from the standpoint of "data as capital." There are some expected issues in both standpoints.

Table 4: Major characteristics of "Data as Capital" and "Data as Labor"

	Data as Capital	Data as Labor
Contributors to creating the value of data	Value utilizers of data	Value the creators of data
Incentive	Entrepreneurialism	Usual contribution (equal to labor)
Future of Labor	Universal Basic Income	Data Labor
Expected concerns	・AI's contribution to improving productivity might be hindered because AI depends on active involvement of individuals who produce related data. ・Market dynamism might become stagnant if the data produced by individuals are concentrated on certain companies.	・Tragedy of the anticommons (possibility that assets to be shared are subdivided and privatized, then utilization of resources useful in society can be hindered) might arise ・The original motivation for the activity might be lowered due to the monetary consideration.

Source：Prepared referring to Imanol Arrieta-Ibarra, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier and E. Glen Weyl.(2018) "Should We Treat Data as Labor? Moving Beyond 'Free'," *American Economic Association Papers & Proceedings*, Vol.108, pp.38-42. and other referential information, and discussion in Special-Interest Group

Source: Prepared through published documents

Table 5: Issues on the approach of "paying remuneration for the users who produced data"

	Data as Capital		Data as Labor
Paying remuneration for the users who produced data	Positive	Negative	Positive
Reasons for paying remuneration/not for paying remuneration	・With the advance of AI, data will be more valued, so the incentive to provide data will be needed. ・Existence of enterprises that take most of the value produced through data as private profits	・Free data sets and machine learning algorism are freely available ・Data can produce value only when it's analyzed, so data analysts should be paid. ・Giving monetary consideration might lower the original motivation for activities.	・"Data as labor" can be an important opportunity to complement the income of citizens who suffer from the expansion of economic disparity, and bring about awareness of contribution to society.
Process for paying remuneration	Market mechanism (Micropayment/information trust function (Information bank)) ・Regulation (competition policy/tax system/social security, etc.)	—	・(Limitations) Build an appropriate technology system to estimate the contribution and retroactively track the value of data individual users produced.
Expected issues	・Market mechanism: Establishment of business models ・Regulation: Objective correctness of introduction (quantitative index)	・Possibility that uneven distribution of wealth between enterprises/data analysts and other individuals (further decline of labor distribution rates) might arise	・(Limitations) Estimation of contribution and feasibility of technology ・Possibility that estimation of contribution might bring about surveillance society and further disparity/ discrimination

The current status and issues toward marketing and visualization of data transaction

The current status of data transaction in Japan

- While initiatives of marketing/visualization of data transaction are proceeding, the current status of data transaction in Japan is as follows.
  - ① Data transaction with other companies is inactive.
  - ② Organizational data have different statuses of providing/receiving data, depending on the type of data
  - ③ The use of data on individuals is asymmetric between the sides of providers and receivers
  - ④ Individuals are concerned about their data being collected and used.
  - ⑤ Enterprises tend to be afraid of being criticized for their data transaction by society.
  - ⑥ Individuals are expecting enterprises to ensure the purpose of using data and data security.

Based on **the current status of data transaction in Japan**, consider **the issues to be addressed**

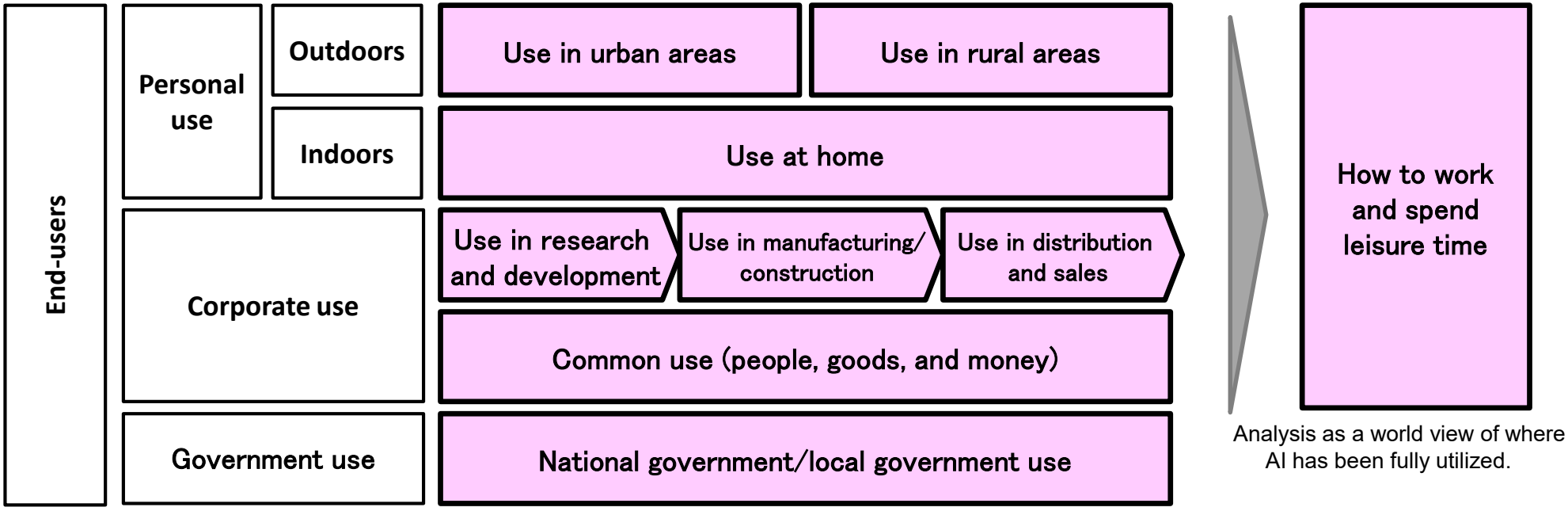
Issues to be addressed for marketing and visualization of data transaction

- Standardization of contract formats concerning data transaction for the utilization of data
  - ① Contractual coverage according to sales channels/price negotiation/manufacturers' responsibility
  - ② Necessity of grasping the distribution/quality/ traceability/ freshness, etc. of data
  - ③ Identification and fidelity guarantee of persons in charge of data labeling/annotation
- The necessity of management of data owned by enterprises by a third party
- The necessity of management of personal data by a third party (Personal Data Trust Bank's potential)
- Finding 'hidden needs' in data transaction

The image of “Inclusive AI Economy Society"

- In ”Inclusive AI Economy Society“, it is desirable that all the entities including individuals, small and medium-sized enterprises, and local enterprises actively engage in social and economic activities in utilizing AI and data in ingenious ways based on their intentions and judgment, contribute to the improvement of productivity, acquire the distribution according to the contribution, fulfilment, and expansion of leisure time, and share the affluence in the entire society.
- In the analysis of the perspective of the utilization of AI, the utilization scenes of AI in individuals, enterprises and government are classified into 9 scenes (Use in urban areas, Use in rural areas, Use at home, Use in research and development, Use in manufacturing/construction, Use in distribution and sales, Common use (people, goods, and money, National government/local government use and How to work and spend leisure time.

Table 6: Perspective of the Utilization Scenes of AI

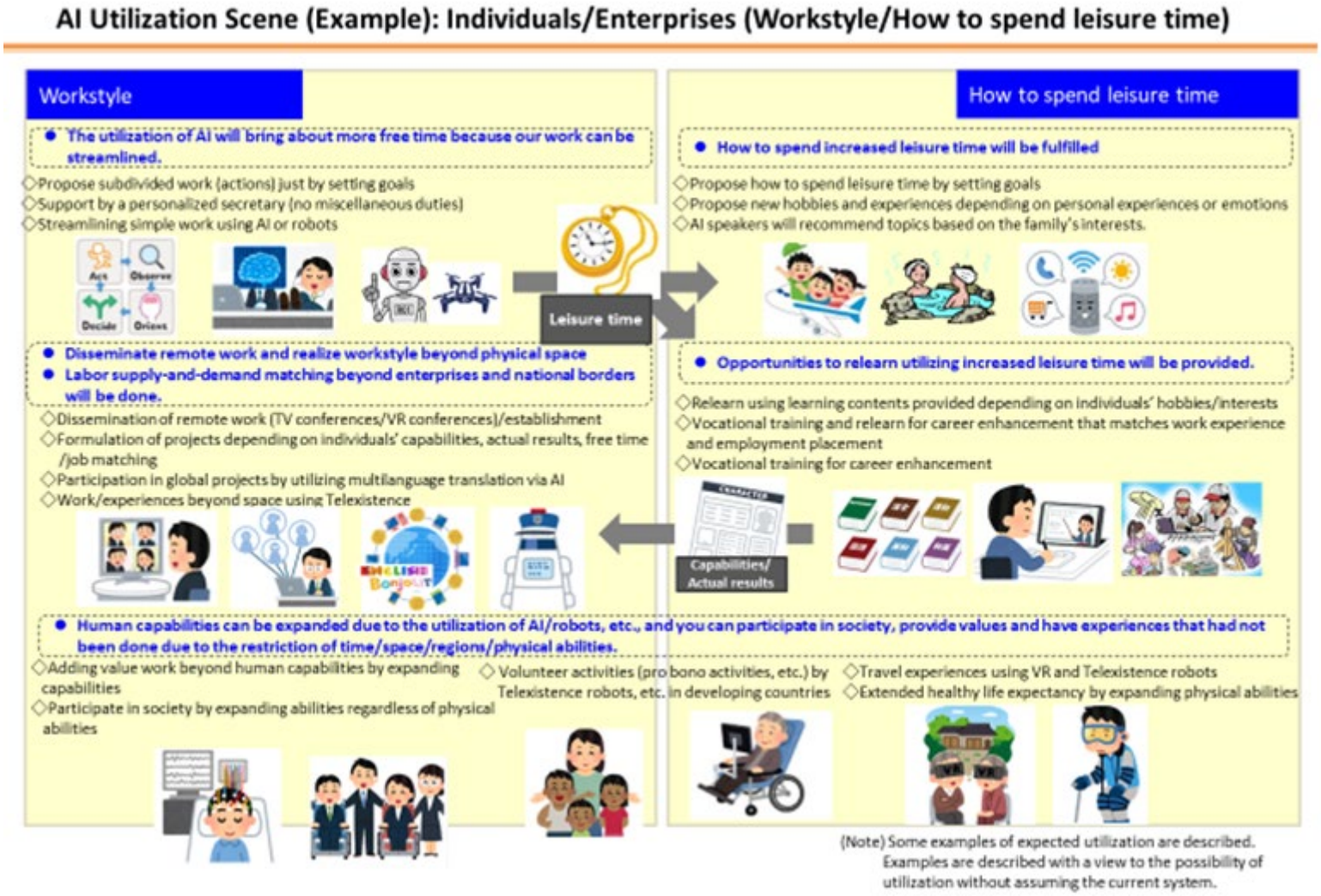


Analysis as a world view of where AI has been fully utilized.

Examples of the Utilization Scenes of AI

- We looked at the utilization scenes of AI in individuals, enterprises, administration and how to spend leisure time.

Table 7: AI Utilization Scene (Example): Individuals/Enterprises (Workstyle/How to spend leisure time)



The ideal policy toward the promotion of the utilization of AI/data

The ideal directivity

- Various restrictions brought about by COVID-19 have changed the workstyle and lifestyle of individuals and the way enterprises should be, so there is a possibility that the “new normal” might be produced.
- The transition to the economy society based on the utilization of AI is more likely to be accelerated in the “post COVID-19 era”. Actually, the “social intelligence” that humans should display in the ”Inclusive AI Economy Society“ has been already required.
- In this situation, measures to support enterprises, individuals and administration are required.
  ① In enterprises, supporting activities to encourage SMEs and local enterprises to launch this initiative for the utilization of AI/data is needed.
  ② In individuals, promoting public-relations campaign to rightly understand new technologies such as AI is required.
  ③ In administration, it is necessary to work on further preparation for the systems so that enterprises and individuals can utilize data trustworthily, and it is important to accelerate the initiatives of the digital government that can increase opportunities for private citizens to utilize public data.

Take **specific measures** based upon the ideal directivity

Specific measures

- The support for SMEs and local enterprises to promote the utilization of AI/data
- Enlightenment activities by the government and private sectors for right understanding of new technologies such as AI
- Initiatives for sharing the image of ”Inclusive AI Economy Society"
- Review of the system that promotes that utilization of AI/data.
- Promotion of the utilization of AI/data keeping the “post-COVID-19” society in mind.

The ideal way of data economy policy required for the promotion of data-driven economy in the AI era

The Ideal Way

- It is expected that data-driven economy will grow going forward, but it's hard to say market mechanism in data transaction is fully working, so it is desirable to reach efficient resource allocation based on market mechanism as soon as possible.
- The mechanism where data produce added value is complicated, and the stakeholders in the mechanism are roughly divided into four; resource data producers, data analysts, data owners and data transaction mediators.

	Resource data producers (Net service users)	Data analysts (Data scientists)	Data owners (enterprises/platforms etc.)	Data transaction mediators (Personal Data Trust Bank, etc.)
Reasons for paying/not paying remuneration	Approval of the contribution by resource data producers	Contribution to analysis	Revenue from data	Commission fee from data transaction
Payment methods of remuneration (process)	Free services, information provision、points, cash, etc.	Evaluation for analysis results	Business revenue	Business revenue
Standard of remuneration	Intention of data owners/data transaction mediators	Labor market	Advertising market, etc.	Cross trading

- From the perspective of revitalization of the transaction of data on individuals, it is important to raise individuals' awareness of managing the data independently, complying with the rule for dealing with personal information. However, it should be noted that various costs are needed for the compliance of the said rules.
- The economic characteristics of data should be considered to conduct appropriate data economy policy.

Take **Specific measures** based upon the **ideal directivity**

Specific measures

- Building data infrastructure
- Realization of data cooperation
- Promotion of corporate data management by a third party/initiatives of Information Bank
- Review the legal system to promote the utilization of data

The Ideal way of data economy policy required for the promotion of data-driven economy in the AI era

Issues in the estimation methods of the value of data

- To realize data transaction markets, observation and analysis of the reality surrounding the data are important. The empirical analysis for estimating the value of data in Chapter 3 is the clue to that.
- Data are added to capital equipment that expands the capability of workers, and are positioned as a provider of new capabilities. Therefore, it is necessary to position data as a factor of production (stock) as the beginning of a theoretical study, and organize them theoretically.
- The issues in the estimation methods of the value of data are organized and the directivity for further brush-up is shown in the following.

① Building capital stock based on the economic characteristics of data.

✓ As the economic characteristics of data, temporal changes in the value of data and expansion of value due to data accumulation are cited, so how capital stock in the target year should be built needs to be considered through the status of owning and utilizing data, taking these characteristics into account.

② Further analysis of the relationship of data and the value created by data

✓ Regarding the relationship between quantity or quality of data and the value of data, analysis by industry type and production process is needed. Conducting a multifaceted analysis with more survey data is required, but to do so, building the system where information is continuously collected in a large scale needs to be considered.

③ Social enlightenment toward data economy

✓ To further polish the estimation methods of the value of data, the improvement of social recognition of the value of data and the expansion of initiatives of analysis are needed. To realize that, preparing and publishing data that can be utilized for analysis is important. One of the methods to do that is publishing data on official statistics or financial statements.