Requirements of Incentive Design for Promoting Data Utilization: Case Study of GAIA-X

Jun Oya *, Naoshi Uchihira *

Abstract

This paper provides a case study of GAIA-X to clarify requirements of incentive design for promoting data utilization. In recent years, there are growing expectations for the creation of new values by analyzing large amounts of data accumulated in cyberspace using artificial intelligence (AI). GAIA-X, where Europe aims to build a data ecosystem, has advantageous and unprecedented rules and new systems for the data owner, which sets data sovereignty, assigns ID to data, and monitors the data usage based on Self-Description as defined by the data owner. However, it is unclear whether these rules and systems are sufficient conditions for promoting data utilization. Therefore, this study sorted out common and different points between the data that will be protected and utilized by GAIA-X and the patent right that is one of the intellectual properties and has been utilized from an earlier stage. Next, promoting data utilization by GAIA-X was analyzed focusing on the transaction cost theory promoted by Williamson, Maslow's needs theory, and the acceleration mechanism of Open Source Software (OSS) development. As a result, it was suggested that the requirements of incentive design for promoting data utilization are to organize the altruistic environment, sharing and feedback mechanism.

Keywords: GAIA-X, Data Utilization, Patent Right, OSS, Incentive Design

1 Introduction

In recent years, there are growing expectations for the creation of new values by analyzing the large amounts of data accumulated in cyberspace using artificial intelligence (AI). For example, in Germany, Industrie 4.0 was advocated to advance the manufacturing industry, and in Japan, Society 5.0, which aims to build a human-centered digital society, was advocated. As for the protection of data, which is the source of such national projects, if some data are edited or accumulated as a database that has creative value, they can be protected by copyright, and if they are managed as a trade secret, they can be protected by the Unfair Competition Prevention Act. Moreover, in Europe, regarding the protection of databases, even if the database does not have creative value as required by the copyright law but has a substantial investment, the database can be protected by Sui generis, defined with copyright in Directive 96/9/EC. This prevents the extraction and/or re-utilization of the whole or a substantial part of the database [1]. Recent discussions have raised the issue of the existence of a large amount of raw machine-generated data that cannot be protected by current intellectual property rights [2].

---

* Japan Advanced Institute of Science and Technology

Also, to promote the utilization of data, discussions are held regarding the establishment of rules to ensure the transparency and fairness of data transactions and the elimination of possible risks in advance as much as possible [3][4].

Under these circumstances, discussions on building a data ecosystem are actively held in Europe. The reason is that the cloud market is mostly dominated by US cloud providers, Europe has to depend on US cloud providers, and therefore Europe faces the risk that the data generated in Europe will be accessed by US law enforcement agencies under the US CLOUD Act [3]. Europe admitted that although Europe strengthened the protection of personal data under the General Data Protection Regulation, which in principle prohibits transferring the personal data acquired within the European Economic Area (EEA) to outside the EEA, Europe has been defeated by the US companies in the competition of personal data. However, regarding industrial data, which is the source of European industrial competitiveness, GAIA-X is positioned as a unique and principal European cloud, organizing a data ecosystem "Catena-X" for the automobile industry, which is a major European industry, is being accelerated [5].

GAIA-X approves the ownership of data, including data that is difficult to protect under current law, by attaching ID to the data under data sovereignty. Moreover, the data owner can define the range of data usage by Self-Description that can be used by third parties. GAIA-X has a data usage monitoring system that monitors whether the data is used within the range of Self-Description [6][7]. Therefore, in the field of intellectual property, an unprecedented and epoch-making mechanism for data protection and utilization that gives data owners an advantage is being built. Figure 1 shows the overview of GAIA-X architecture [7].
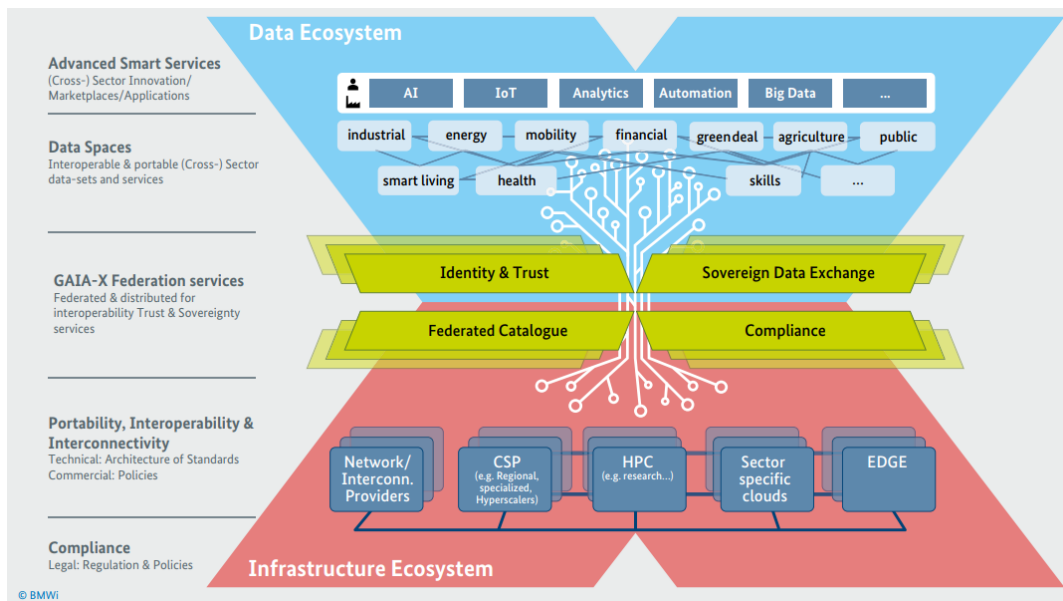

Figure 1: GAIA-X architecture overview [7]

Regarding GAIA-X, white papers have been already published [8][9], and by including some use cases, the understanding of those involved in promoting data utilization has been fostered. However, this discussion is based on the premise that the data will be provided to some transaction entities. It is unclear whether these rules and systems are sufficient conditions for promoting data utilization.

Therefore, this study sorted out common and different points at the stage of utilization between the data which is protected and utilized by GAIA-X and the patent right which is one of the intellectual properties and has been utilized from an earlier stage compared with the data to analyze insufficient points in GAIA-X for the promotion of data utilization. Next, some problems regarding promoting data utilization by GAIA-X were analyzed focusing on the transaction cost theory promoted by Williamson, Maslow's needs theory, and the acceleration mechanism of Open Source Software (OSS) development. Then, the requirements of incentive design for promoting data utilization were clarified.

This paper is presented as follows. Chapter 2 reviews studies on incentives for providing data. Chapter 3 summarizes the comparison between patent right and data utilized by GAIA-X. Chapter 4 summarizes the comparison between OSS and data utilized by GAIA-X. Chapter 5 analyzes the requirements of incentive design for promoting data. Finally, Chapter 6 provides a conclusion.

2   Literature Review

Regarding incentives for providing data, discussions are held to work incentives by strengthening the governance of the data transaction market, such as going through the process of examination and approval of the participation qualifications [10]. However, it is intended to improve extrinsic incentives for transaction entities, not to improve intrinsic and voluntary incentives. Therefore, it is unclear whether this governance is sufficient for promoting data utilization.

Also, providing data can be promoted in the environment of altruism. Chapter IV of the Data Governance Act [11], which Europe is going to establish to promote data utilization, defines Data Altruism, and shows a framework for individuals and companies to provide data altruistically for the public interest. The Data Governance Act states that an independent organization called recognized data altruism organizations, will take on the function of recording information, such as collected data, data processing, income, and expenditure through data transactions involved among multi-stakeholders. The Data Governance Act aims to create an altruistic environment that promotes providing data, but it is unclear whether the framework is sufficient as an incentive to provide data.

In addition, based on the recognition that blockchain, game theory, and AI are incentive techniques for building Internet of Things ecosystems, it is pointed that the incentive techniques facilitate data sharing [12]. Although the use of these techniques to ensure the safety and transparency of data sharing has been discussed, it is not clear how the incentives for data sharing can be increased from the viewpoint of the data provider.

3   Comparison between Patent Right and Data Utilized by GAIA-X

Common and different points at the stage of the promotion of utilization between the data protected and utilized by GAIA-X and the patent right, which is one of the intellectual properties and has been utilized from an earlier stage compared with the data, were sorted out to analyze insufficient points in GAIA-X for the promotion of data utilization. Table 1 shows the protection

of technical ideas under the patent right and the protection of data under GAIA-X, evaluation, availability of search, and the method of transaction.

3.1 Difference in Procedure

The major difference is that GAIA-X has the same effect as a patent application in that the owner of the data is approved by attaching ID to the data, but GAIA-X does not have a procedure equivalent to substantive examination like the patent right. It is difficult to directly evaluate the value of the data itself because the value of the data is usually generated through the stages of preprocessing and analysis, and various values are generated depending on the point of view of the subject of analysis. In other words, it is difficult to design procedures equivalent to substantive examination.

3.2 Availability of Search

There is also a major difference in the availability of search. Patent rights can be searched by narrowing down the technical idea to a certain range using the International Patent Classification. In addition, the technical idea is described in the patent specification, in which it is possible to perform a full-text search using technical terms. Therefore, it is possible to directly reach the technical ideas in which the transaction entities are highly interested. However, data requires metadata for the search that is not attached to the original data. Even if data is narrowed down using metadata, an analysis may be difficult due to the differences in data formats, and the value of data cannot be recognized without going through the analysis stage.

Table 1: Comparison between patent right and data utilized by GAIA-X

	Patent Right	Data (GAIA-X)
Filing, Examination	Formality and substantive examination	ID is attached to the data. No substantive examination
Evaluation	As a substantive examination, the examiner makes an objective determination of patentability. Granted patents are novel and inventive over the prior art.	It is difficult to evaluate the value from the data itself. Various values can be generated through data preprocessing, processing, and analysis methods.
Availability of Search	International Patent Classification are attached to all patent publiacions. Full-text search is also available using keywords.	Metadata is attached to the data. International Data Space proposed Broker to manage the metadata.
Method of Transaction	The sales conditions or license conditions are determined through the negotiation. In recent yeas, package license (cross-license, patent pools), patent pledge, and so on are facilitated.	The data provider defines the range of data usage by Self-Description as a condition of the transaction. Data usage are monitored based on the Self-Description.

3.3 Transaction

Next, the study examined whether the market transaction of data utilized by GAIA-X would work. In the case of the transaction of patent rights, through substantive examination, they are narrowed down by patent search. The technical idea that is the subject of the transaction is specifically recognized by reading the description of the patent right. Then, the sales conditions or license conditions are determined through the negotiation. Therefore, in principle, market transactions can be established. However, they are not necessarily activated for the following reason.

In the past, INPIT (National Center for Industrial Property Information and Training), an independent administrative agency under Japan Patent Office, established the Licensable Patent Information Database for patent utilization, but the key to the activation of patent utilization was that about 100 patent utilization advisors visited companies and matched the needs of each company to the patent right [13].

This kind of intermediary has also been introduced in private matching businesses. For example, Ninesigma collects technical needs from clients and then calls for proposals that fulfill technical needs to increase the matching rate of the patent transaction [14]. In other words, patent rights transactions cannot be activated by only developing a transaction market infrastructure equipped with a search function. It is necessary to design a system that supports technology connoisseurs and negotiations, namely, a system that complements the capabilities of transaction entities and reduces transaction costs.

In the case of the transaction of the data utilized by GAIA-X, candidate data for transactions are narrowed down using metadata, the conditions of data usage are shared between data providers and data users based on Self-Description, and the data usage is monitored. Therefore, the transaction cost can be reduced by GAIA-X. However, since neither the data provider nor the data user can recognize the value of the data without data analysis, the data provider will be afraid of lower, unreasonable evaluation of the provided data even if GAIA-X works. As a result, incentives to provide the data do not work well.

4 Comparison between OSS and Data Utilized by GAIA-X

In addition to reducing transaction costs, what other strategy can be taken to improve incentives for providing data?

4.1 Altruism

O.E. Williamson pointed out that trying to make decisions based on insufficient information leads to higher transaction costs, bounded rational, and opportunistic (selfish) behavior [15]. In addition, H.A. Simon pointed out that an altruistic society has potential for development, but a selfish society does not have such potential [16].

GAIA-X discloses information to some extent about transaction entities, transaction data, and data usage through ID, Self-Description, usage control, and so on. Then, it may be expected that the information asymmetry condition is improved, and selfish behavior is to some extent suppressed. However, altruistic behavior that expects that the provided data will be utilized by many transaction entities cannot be necessarily activated. Some kind of additional incentive design is required.

4.2 Maslow's Needs Theory

We examined this problem from the viewpoint of psychology. According to Maslow's needs theory, there are five categories of needs from the lowest level to the highest level: physiological, safety, belongingness and love, esteem, and self-actualization. The lower-level needs derived

from extrinsic motivation and the higher-level needs derived from intrinsic motivation. It is proposed that the higher-level needs can be satisfied when the lower-level needs are satisfied to some extent [17].

As a correspondence relationship of the GAIA-X mechanism, GAIA-X enhances security functions and monitors the data usage of the provided data, which ensure safety and transparency of data transaction. Then, the safety need, which is a lower-level need derived from extrinsic motivation in Maslow's needs theory, is satisfied, but the higher-level needs derived from intrinsic motivation still are not satisfied. Furthermore, the incentives to provide data, namely, the incentives to promote altruistic behavior, cannot be explained.

4.3 Mechanism of OSS Development

To solve this problem, we focused on analysis of OSS development. OSS development is one of the examples where altruistic behavior is promoted. In addition, the source code of OSS can be protected by the copyright law as well as the database, OSS is based on the premise that the source code is shared among developer and third party. Although data provision is not exactly data sharing between the data provider and the data user, it can be expected that the mechanism of OSS development provides hints for promoting data provision.

According to a previous study on the motivation of OSS development, the motivations for participating in OSS development are "gift culture" and "craftsmanship." They do not have trouble making a living, and they seek a reputation from their colleagues by giving something to their colleagues [18]. OSS development takes advantage of the developer's personal profit motivation [19]. In addition, OSS developers share improved source code with each other free of charge, because the sharing of improved source code is ruled. Therefore, there is no information asymmetry within the developer community, and altruistic behavior is a common perception. Unlike GAIA-X, OSS development does not have a problem of satisfying the lower-level needs in

Table 2: The difference between OSS development and GAIA-X

	OSS	Data (GAIA-X)
Stakeholder	Developer, Third party	Data provider, Data user
Description of Copyright, Owner	Included	Included (ID)
Outcome	Improved source code is shared among developer and third party.	Not ruled
Information Assymmetry	No information assymmetry (Improved source code is shared.)	Information asymmetry can be improved by ID, Self-Description, usage control, meta data. However, there is no rule about outcome of analyzed data.
Incentive	Intrinsic and voluntary motivation can work, because it is possible to seek personal profit. Improved source code is evaluated among the community. Therefore, developer can get feedback, motivation can be increased.	Extrinsic incentives are designed by ensuring security and transperency of data transaction. However, intrinsic incentives are not designed.
Altruism	Altruistic behavior is a common perception. In additon, intrinsic motivation works in an environment of altruism.	GAIA-X aims to build altruistic environment. However, it is unclear whether altruistic environment facilitate to provide the data.

Maslow's needs theory. In such an environment, the motivation to seek evaluation among their community and the motivation to seek personal profit, namely the intrinsic and voluntary motivation work.

Compared seeking evaluation among the community and seeking personal profit with Maslow's needs theory, it can be considered that the former corresponds to the esteem need and the latter corresponds to the self-actualization need, which is the highest need in Maslow's needs theory. In other words, the activation of OSS development is the result of satisfying the self-actualization need over the esteem need in an environment where there is no information asymmetry among developers. Altruistic behavior is a common perception; namely, this is the result of being piled up of personal profits based on the intrinsic and voluntary motivation of each developer in an environment of altruism.

Table 2 shows the difference in stakeholders involved in OSS development and data utilization by GAIA-X, handling of the outcome, incentives, and so on.

5   Requirements of Incentive Design for Promoting Data Utilization

Based on the analysis in Chapters 3 and 4, the requirements of incentive design for promoting data utilization can be clarified as follows.

5.1 Altruistic Environment

Referring to the mechanism of OSS development shown in Chapter 4, the first approach to increase incentives for providing data in GAIA-X is to organize an altruistic environment. R. Axelrod pointed out that true altruism can well develop in kinship, such as "parent" and "child" relationships, because the attribution is clear [20]. However, in business conditions, the attribution of transaction entities is not always clear. He also pointed out that friendly relationships are not necessary to develop the altruistic environment, cooperative relationships can be created by narrowing down transaction entities and deepening individual relationships [20]. This means that intermediary function shown in Chapter 3 is necessary for building an altruistic environment. GAIA-X has systems consisting of ID, Self-Description, usage control, and so on, and if these systems are strengthened as intermediary function, the altruistic environment can be improved.

5.2 Sharing

As shown in Chapter 4, sharing is an important condition for OSS development. Even if it is difficult to create a completely altruistic environment, it is effective to organize the somewhat closed sharing condition of secondary data preprocessed and analyzed by the data user by defining contract templates. Regarding J.R. Commons' institutional economics, well-organized institutions enable future-oriented actions by guaranteeing expectations for the future in the presence of uncertainty [21]. Namely, concerning data governance, if the sharing condition of secondary data is ruled in advance, the uncertainty of data sharing for the data provider can be improved and incentives for providing data can be increased.

Actually, in Japan, according to the Contract Guidelines for AI and Data in the Agricultural Sector published by the Ministry of Agriculture, Forestry and Fisheries, one of the contract templates states that the secondary data generated by the data user must be shared with the data provider because the agricultural data may include the know-how of farming, and the data provider may hesitate to provide data [22][23].

5.3 Feedback

In addition, feedback is also important for altruistic environments. As shown in Chapter 4, feedback is an important mechanism of OSS development. Sharing secondary data is also a kind of feedback. It is important to enable to recognize the results of providing data, such as how the provided data were used for the benefit of society. J.R. Hackman and G.R. Oldham advocated a job characteristics theory, which explains that one of the job characteristics that enhance intrinsic motivation is feedback [24]. E.L. Deci and R. Flaste advocated that the reward of intrinsic motivation is a sense of enjoyment and accomplishment, to build such an environment, it is necessary to create a mechanism that links actions and results [25].

6   Conclusion

This study focused on GAIA-X to clarify the requirements of incentive design for promoting data utilization. We sorted out common and different points between the data, which will be protected and utilized by GAIA-X, and the patent right, which is one of the intellectual properties and has been utilized from an earlier stage compared with the data. Next, promoting data utilization by GAIA-X was analyzed focusing on the transaction cost theory promoted by Williamson, Maslow's needs theory, and the acceleration mechanism of OSS development.

It was clarified that although GAIA-X has advantageous and unprecedented rules and new systems for the data owner, which sets data sovereignty, assigns ID to data, and monitors the data usage based on Self-Description as defined by each data owner, incentives to provide data do not work well, since neither the data provider nor the data user can recognize the value of the data without data analysis, the data provider will be afraid of lower, unreasonable evaluation of the provided data even if GAIA-X works. As a result, incentives to provide data do not work well.

In addition, it was clarified that altruistic behavior that expects that the provided data will be utilized by many transaction entities cannot be necessarily activated in GAIA-X, some kind of additional incentive design is required. The mechanism of OSS development provides hints for promoting data provision. The requirements of incentive design for promoting data utilization are to organize the altruistic environment, sharing and feedback mechanism.

This study is still an initial proposal on requirements of incentive design for promoting data utilization. The data, patent right and OSS which this study focused on are classified in the same category as the intellectual property. However, all mechanisms for promoting the patent rights and OSS utilization are not necessarily suitable for promoting data utilization. Therefore, we are going to add some case studies related to the incentive design for promoting data utilization.

References

[1] Directive 96/9/EC of the European Parliament and of the Council, Chapter III, Mar. 1996.

[2] European Commission, "Building a European Data Economy," p. 10, Jan. 2017.

[3] European Commission, "A European Strategy for Data," Chapter 3, Feb. 2020.

[4] Ministry of Internal Affairs and Communications, Ministry of Economy, Trade and Indus-try, "Case Study on New Data Distribution Transactions," Mar. 2022 (in Japanese).

[5] Catena-X Automotive Network e.V., https://catena-x.net/en/, (accessed 20.Aug. 2022).

[6] GAIA-X, "GAIA-X: A Pitch Towards Europe," Federal Ministry for Economic Affairs and Energy, May. 2020.

[7] GAIA-X, "GAIA-X: Driver of Digital Innovation in Europe," Federal Ministry for Economic Affairs and Energy, p. 9, May. 2020.

[8] GAIA-X, "White Paper GAIA-X Secure and Trustworthy Ecosystems with Self Sovereign Identity," Jun. 2022.

[9] GAIA-X, "GAIA-X and European Smart Cities and Communities White Paper V21.09," Oct. 2021.

[10]   World Economic Forum, "Developing a Responsible and Well-designed Governance Structure for Data Marketplaces," Aug. 2021.

[11]   European Commission, "Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European Data Governance and Amending Regulation (EU) 2018/1724 (Data Governance Act)", Jun. 2022.

[12]   P.K.R. Maddikunta, Q.-V. Pham, D.C. Nguyen, T. Huynh, O. Aouedi, G. Yenduri, S. Bhattacharya, T.R. Gadekallu, "Incentive Techniques for the Internet of Things: A Survey," Journal of Network and Computer Applications, vol. 206, Oct. 2022.

[13]   INPIT, "Investigative Research Report on the Outcome of the Patent Utilization Promotion Business," Mar. 2010 (in Japanese).

[14]   NineSigma, https://www.ninesigma.com/, (accessed 21.Aug. 2022).

[15]   O.E. Williamson, "Markets and Hierarchies: Some Elementary Considerations," Organiza-tional Forms and Internal Efficiency, vol. 63, No. 2, pp. 316-325, May. 1973.

[16]   H.A. Simon, "A Mechanism for Social Selection and Successful Altruism," Science, 250(4988), pp. 1665-1668, Dec. 1990.

[17]   A.H. Maslow, "Motivation and Personality (Second Edition)," Prabhat Prakashan, Chapter IV, 1987.

[18]   E.S. Raymond, "Homesteading the Noosphere," https://firstmonday.org/ojs/index.php/fm

/article/download/1474/1389?inline=1, (accessed 21.Aug. 2022).

[19] K.R. Lakhani, E. von Hippel, "How Open Source Software Works: Free User-to-user As-sistance," Research Policy, vol. 32, pp. 923-943, 2003.

[20] R. Axelrod, "The Evolution of Cooperation," Basic Books, Oct. 1985.

[21] T. Nakano, "Enrich the Country and Strengthen the Army, An Introduction to Geographical Economy," Toyo Keizai Inc., Chapter 13, Dec. 2016 (in Japanese).

[22] Ministry of Agriculture, Forestry and Fisheries, "the Contract Template," https://www.maff.go.jp/j/kanbo/tizai/brand/attach/pdf/keiyaku-31.pdf, (accessed 21.Aug. 2022) (in Japanese).

[23] Ministry of Agriculture, Forestry and Fisheries, "Contract Guidelines for AI and Data in the Agricultural Sector ~For Protection of Agricultural Expertise and Promotion of Data Uti-lization~, " https://www.maff.go.jp/j/kanbo/tizai/brand/keiyaku.html, (accessed 21.Aug. 2022) (in Japanese).

[24] J.R. Hackman, G.R. Oldham, "Motivation through the Design of Work: Test of a Theory," Organizational Behavior and Human Performance, vol. 16, pp. 250-279, 1976.

[25] E.L. Deci, R. Flaste, "Why We Do What We Do, The Dynamics of Personal Autonomy," Penguin Books, Jan. 1995.