Survey on Establishing Evaluation Model for Blockchain Technology-Based System

Executive Summary

March 2017

株式会社三菱総合研究所
（Mitsubishi Research Institute, Inc.）

1. Background and purpose

Background

- The blockchain technology used for virtual currencies, such as Bitcoin, is *extremely difficult to falsify* compared with conventional centrally managed systems. Also, it is possible to construct a *real zero-downtime* system as *low cost*. Therefore, the technology is expected to be applied to a very wide field such as IoT.
- However, it was made clear that there was emerging problems to install the technology. One of these problems is the lack of indicators and criteria for precisely evaluating the features of the technology and comparing it with existing technologies and systems.
- Because there is no evaluation model, there is anxiety, excessive expectations, and misunderstandings about the technology and its system. As a result, a situation where appropriate introduction does not proceed has been created.

Purpose

- To create evaluation model that can appropriately evaluate the characteristics of the Blockchain Technology-based System (BC system), covering all items necessary for evaluation.
- To make a meaningful evaluation model that can be used as evaluation criteria by system vendors when they actually compare the BC System and the existing system.

2. Scope of study and points to be noted

Scope

- In principle, the evaluation model was considered based on the situation as of November 2016.For the blockchain platform, Bitcoin, Ethereum, Hyperledger Fabric etc. were keeping in mind. The use cases were targeted for those whose concepts were announced or where demonstration tests were conducted.
- From now on, the development of various technologies and the creation of use cases that cause social change will also be assumed. The kind of change that would occur was considered as much as possible based on the insight of the committee members*. However, it is difficult to predict the evolution of technology, so it is necessary to pay attention to the fact that the future assumption is limited.

*: See page 3

Points to be noted

Please understand the following points before using this study result

- Blockchain technology is being improved day by day. This study had been implemented for about half a year (from September 2016 to March 2017), and in this short time there had emerged new technologies or a blockchain platform, and an evaluation model supporting all of them could not be studied.
- In this issue, we examined the evaluation model for the technology still under development, but evolving rapidly, as a snapshot of the present time. In other words, we could not consider a universal evaluation model for the future. Therefore, it is necessary to review the evaluation model from time to time as technology evolves and use cases are accumulated.

3. Survey flow and expert opinion committee

- A survey of domestic and international trends on evaluation of BC system was conducted to put in order the points at issue. Based on the result of the discussion, a study of the evaluation model was conducted at an expert opinion committee meeting.
- The committee met five times from November 2016 to March 2017. The chairperson was served as Dr. Takagi Souichiro of the Global Communication Center of the International University of Japan. Committee members included academic experts, blockchain development companies, domestic system vendors, and international consortium participation companies.

<<Survey flow>>

(1) Survey of domestic and international trends on evaluation of BC system

- ✓ Desktop research on characteristics of blockchain technology and evaluation of BC system
- ✓ Survey of current status of evaluation in domestic and international use cases
- ✓ Organizing points at issue through advanced interviews with parties concerned and experts

(2) Establishment of evaluation model for BC system

- ✓ Consideration through five expert opinion committee meetings

<<List of committee members>>

Name	Title, Organization (March, 2016)
Soichiro Takagi (Chairperson)	Research Manager, The International University of Japan, Global Communication Center,
Edmund Edgar	CEO, Social Minds
Hiroshi Oiwa	Group Leader, Research Group of Cyber Physical Ware, National Institute of Advanced Industrial Science and Technology
Yuzo Kano	CEO, bitFlyer,
Masanori Kusunoki	CISO Board, Yahoo
Koichi Shibata	General Manager, IoT Business Development Office, SKEED
Yasunori Sugii	CEO, Currency Port
Masanobu Takagi	blockchain Architect, IBM Japan
Toshiya Cho	General Manager, Financial Innovation Center, Hitachi
Shinichi Toriyama	Manager, Fintech Business Development Office, NEC
Masayuki Hatta	Instructor, Suruga University
Kazumi Hirose	Cloud Solution Architect, Microsoft Japan,
Shigeichiro Yamazaki	Professor, Department of Information and Computer Science, Faculty of Humanity-oriented Science and Engineering, Kindai University

4. Issues for establishing evaluation model

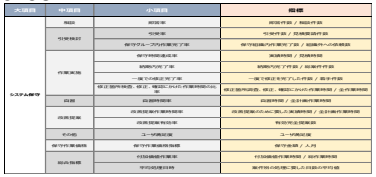
Examination items		Examination policy for evaluation model
Purpose of evaluation model		• The characteristics of a BC system should be appropriately expressed. • As many necessary and sufficient evaluation items as possible should be covered in expected use scenes. • It was decided that an actual evaluation would not be conducted.
Evaluation object	Object of comparison	• Evaluation items that can comprehensively compare existing systems with a BC system and between BC systems were selected. (This is because, although it is assumed at present that existing systems are compared with a BC system in many cases, comparison between BC systems will be necessary in the future.)
	Scope of comparison	• The entire system including the blockchain platform and the subsystem was set as the evaluation model. (Each blockchain platforms have different functions, so a simple comparison is difficult, and user needs are in the evaluation of the whole system.)
	Treatment of business transformation	• Even when the existing business flow changes by utilizing the blockchain technology, it is targeted if the purpose and requirements of the business do not change. • For systems that realize unprecedented services, they were excluded from the evaluation targets because they do not have comparison objects.
	Platform classification	• Based on the types (public and private types) of platforms and consensus algorithm, it was the evaluation model targeting all patterns. • Evaluation points to note due to differences in platform classifications are specified in the remarks.
	Coverage of use cases	• We assumed a comprehensive evaluation model assuming differences in the requirements by use case. (There is a need to judge what kind of blockchain platform is suitable for the case.)
Evaluation item/index/ method		• We extracted the evaluation items highly relevant to the characteristics of a BC system. • Quality, maintenance / operation, and cost as major items. • We decided not to set the evaluation index, but to enrich the description such as points to pay attention to when actually evaluating.

5. Structure of evaluation model

- The evaluation model consist of three major items: quality, maintenance / operation, and cost.
- The items deeply related to a BC system were extracted from ISO / IEC 25010 (system and software quality model) for quality and from IPA (System Reference Manual, Chapter 4 “Maintenance and Operation,” 2005) for maintenance / operation.
- Cost items system vendors recognize as costs (≒ expenses the customers will be charged with) are organized.

Quality model of system & software
Source: ISO/IEC 25010

Evaluation index of system maintenance & operation
Source: IPA



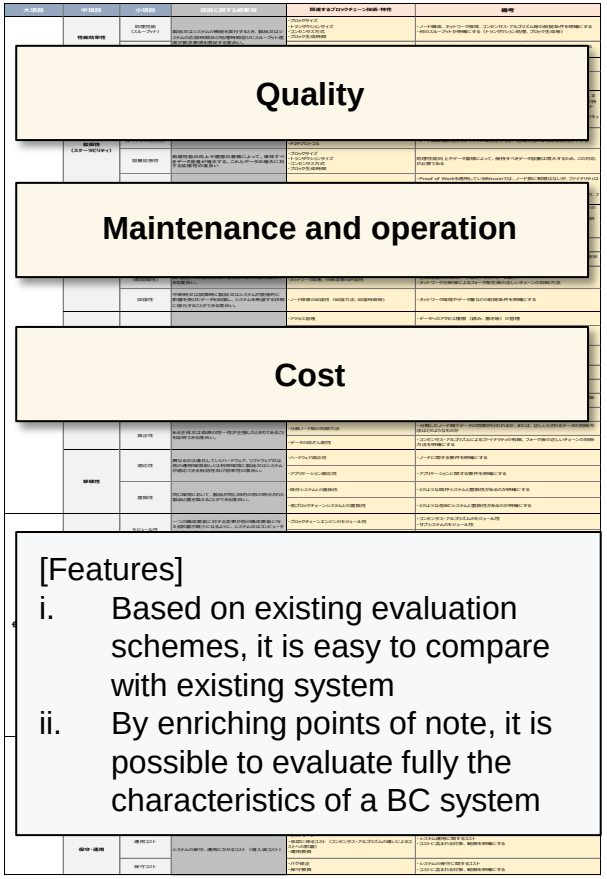
Cost perspective:
- *R&D cost*
- *Implementation cost*
- *Maintenance & operation cost*

Characteristics of a BC system
- Discussions at committee meetings
- Interview with experts
- Desktop survey

Based on existing evaluation shemes
- Easy to compare with existing systems
- Evaluation model with high coverage

Extraction & addition of evaluation model focusing on characteristics of BC system
- Evaluation model of BC technology & characteristics
- Focus on notes

Evaluation model of BC system



[Features]
i. Based on existing evaluation schemes, it is easy to compare with existing system
ii. By enriching points of note, it is possible to evaluate fully the characteristics of a BC system

6. Evaluation model (1) Quality （1/5）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
Performance Efficiency	Throughput	This characteristic represents the performance relative to the amount of resources used under stated conditions.	・Block size ・Transaction size ・Consensus algorithm ・Block confirmation time	・ Clarify the conditions for the node configuration, network environment, and consensus algorithm. ・ Clarify the definition of throughput. For example, “The target is transaction processing and theoretical performance.” ・ Clarify the tradeoff relationship between these items and other evaluation items. *For details of the tradeoff, see the item “Scalability - Throughput.”
	Network latency		・Network environment ・Node Distribution	・ Clarify the preconditions for the node configuration and network environment. ・ Clarify the definition of latency. For example, select two nodes randomly and measure the time taken to transmit ● kB data between the nodes. Average the iteration value.
	Block confirmation		・Consensus algorithm ・Network environment ・Node Distribution	・ Clarify the definition of block confirmation time (e.g., time from when a transaction is thrown until when a block is generated). ・ Clarify the characteristics and tradeoff relationship of the consensus algorithm that is used as in the following: “In the case where the block is confirmed completely, it is the time until confirmation. In the case where the block is not confirmed completely, it is the time until confirmation with a probability of ●%.” For example, when PoW is used and the block cannot be confirmed completely, but when the block depth is 6, the probability is ●%. Also, there is no restriction on the number of nodes. ・ Clarify the tradeoff relationship between these items and other evaluation items. *For details of the tradeoff, see the item “Scalability - Throughput.”
	Data reference		・Node Distribution ・Network environment ・Block structure	・ Clarify the preconditions of the node configuration and network environment when referring to a specific block and transaction.

6. Evaluation model (1) Quality （2/5）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
Interoperability	Interoperability with existing systems	Degree to which two or more systems, products, or components can exchange information and use the information that has been exchanged.	・Data structure ・API specification	・ Clarify the conditions for interoperability. ・ Clarify the experience of interoperability with existing systems.
	Interoperability with other BC systems		・Data structure ・Consensus algorithm ・API specification	・ Clarify the conditions for interoperability. ・ Clarify the experience of interoperability with other BC systems.
Scalability	Throughput	Degree of improvement of performance	・Block size ・Transaction size ・Consensus algorithm ・Block confirmation time	・ Clarify the method for improving the throughput and the tradeoff caused by the method. [Clarify the tradeoff relationship] ・ Tradeoff relationship related to reliability: For example, as the throughput performance improves, the data capacity increases. The load of the full nodes, especially where all the data are retained, becomes heavier, and the sites that can bear the full nodes decrease. As a result, the reliability is lowered. If the number of full nodes decrease to ● or less, it is difficult to meet the requirements. ・ Tradeoff relationship resulting from applied consensus algorithm: For example, the throughput is increased by applying high-speed ●● consensus algorithm. This algorithm requires approved nodes that are specifically managed, and the upper limit of such nodes is around 30 for practical operation. If one thirds of these nodes are suspended, the system function cannot be maintained and thus the availability decreases, leading to ▲.
	Network latency		・Node Distribution ・Network environment ・P2P protocol	・ This system operates in a distributed environment, and performance largely depends on the network environment. ・ Clarify the key points of network latency improvement by showing the bottleneck of network latency improvement.
	Capacity	The degree to which capacity can be expanded.	・Block size ・Transaction size ・Consensus algorithm ・Block confirmation time	・ Data continues to increase. Estimate the volume of data accumulated after a certain period of time and clarify countermeasures.

6. Evaluation model (1) Quality （3/5）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
Scalability	Number of nodes	The degree to which nodes can be increased.	・Data capacity ・Consensus algorithm	・ Clarify the upper limit of the number of nodes (full node, light node, etc.). As the number of nodes increases, the number of transactions may exceed the processing performance. To prevent this issue, clarify the number of nodes matching the throughput performance. ・ Clarify the tradeoff relationship between these items and other evaluation items. *For details of the tradeoff, see the item “Scalability - Throughput.”
Reliability	Maturity	Degree to which a system, product or component meets needs for reliability under normal operation.	・Existing practical technology (cryptographic technology etc.) ・Newly developed technology (consensus, algorithm, etc.) ・Maturity and operating results as a practical system utilizing blockchain technology	・ Maturity is usually evaluated based on the actual introduction record. However, the introduction record of blockchain technology is very few, and it is difficult to evaluate it briefly. Therefore, since a BC system is made up of existing technologies (cryptographic techniques, etc.) and new technologies (consensus algorithm, throughput, etc.) that have been researched and developed to improve performance and functions, it shows the maturity as a system, based on practical achievements of individual elemental technologies, operation results of similar systems, operation results in a test environment, and the like.
	Availability	Degree to which a system, product or component is operational and accessible when required for use.	・Single point of failure	・ Clarify whether there is a single failure point node. ・ In the case where there is no single point of failure, clarify the rough number of ineffective nodes caused by disconnection or other failures that affects the reliability as a system.
			・Consensus algorithm	・ Clarify the conditions (number of nodes, etc.) for obtaining the correct consensus. ・ Clarify the conditions where the consensus will not function, such as an illegal state (51% attack) caused by the consensus algorithm and a consensus disabled state (disconnection with one-third or more nodes in PBFT).

6. Evaluation model (1) Quality （4/5）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
Reliability	Fault Tolerance	Degree to which a system, product or component operates as intended despite the presence of hardware or software faults.	・Node failure tolerance ・Tolerance of network failure and network attacks	・ Clarify the definition of normal operation. ・ Clarify node conditions and network conditions for normal operation. ・ Clarify the main chain determination method after fork generation due to network divisions.
	Recoverability	Degree to which, in the event of an interruption or a failure, a product or system can recover the data directly affected and re-establish the desired state of the system.	・Recoverability of node failure (recovery method and time, etc.)	・ Clarify preconditions for the network environment and data volume.
Security	Confidentiality	Degree to which a product or system ensures that data are accessible only to those authorized to have access.	・Access control	・ Clarify the management method of the access authority (reading, writing, etc.) to the data, setting level, etc.
			・Data anonymization	・ Clarify the existence of a function to conceal data. ・ Clarify the object and scope of concealment. ・ Clarify concealed data verifying method by third party.
			・Transaction anonymization	・ Clarify the existence of a function to conceal transactions. ・ Clarify the object and scope of concealment. ・ Clarify the verification method by third party.
	Integrity	Degree to which a system, product or component prevents unauthorized access to, or modification of, computer programs or data.	・Membership management	・ Clarify the existence of membership management functions.
			・Access control	・ Clarify the management method of the access authority (reading, writing, etc.) to the data, setting level, etc.
	Non-repudiation	Degree to which actions or events can be proven to have taken place, so that the events or actions cannot be repudiated later.	・Consensus algorithm	・ Clarify the existence of block confirmation by consensus algorithm. Also, clarify how to decide the main chain after fork.
			・Hard fork policy	・ Clarify the rule, method, and the extent of the influence of block rollback.

6. Evaluation model (1) Quality （5/5）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
Security	Authenticity	Degree to which the identity of a subject or resource can be proved to be the one claimed.	・Synchronization method between distributed nodes	・ Clarify whether the data are synchronized between distributed nodes. Also, clarify how to decide the data that should be considered right at the time of synchronization.
			・Consensus algorithm	・ Clarify the existence of block confirmation by consensus algorithm. Also, clarify how to decide the main chain after fork.
Portability	Adaptability	Degree to which a product or system can effectively and efficiently be adapted for different or evolving hardware, software or other operational or usage environments.	・Hardware adaptability	・ Clarify requirements on nodes.
			・Application adaptability	・ Clarify application requirements.
	Replaceability	Degree to which a product can replace another specified software product for the same purpose in the same environment.	・Replaceability with existing system	・ Clarity what kind of existing system has substitutability.
			・Replaceablity with another BC system	・ Clarify what kind of another BC system has substitutability.

6. Evaluation model (2) Maintenance / Operation（1/2）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
Maintainability & operability	Modularity	Degree to which a system or computer program is composed of discrete components such that a change to one component has minimal impact on other components.	・Blockchain platform	・ Clarify the components of the blockchain platform and the modularity of technical elements. For instance, since the consensus algorithm is a highly modular implementation, it is easy to change to another algorithm.
			・Subsystem	・ Clarify the components of subsystems and the modularity of technical elements. For instance, regarding the ●● function of the subsystem, upgrading the ●● function is easy due to modular design considering high functionality.
			・Contract code	・ Clarify the specifications of the contract code (description language and other code).
	Reusability	Degree to which an asset can be used in more than one system, or in building other assets.	・Blockchain platform	・ Clarify the reusability of a consensus algorithm. For example, the consensus algorithm implemented on the BC platform ●● can be implemented on the BC platform ▲▲.
			・Subsystem	・ Clarify the reusability of a subsystem. For example, the subsystem ●● can be used in the ▲▲ system.
			・Contract code	・ Clarify the specifications of the contract code (description language and other code).
	Analyzability	Degree of effectiveness and efficiency with which it is possible to assess the impact on a product or system of an intended change to one or more of its parts, to diagnose a product for deficiencies or causes of failures, or to identify parts to be modified.	・Fault detection	・ Clarify the presence or absence of a function to detect a failure. ・ Clarify the presence or absence of a function to identify where a failure has occurred (node failure, network failure, etc.). ・ Clarify the presence or absence of a function to identify the range of the influence of a failure.
			・Performance analysis	・ Clarify the presence or absence of performance monitoring function such as throughput, network performance, and scalability.

6. Evaluation model (2) Maintenance / Operation（1/2）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
Maintainability & operability	Modifiability	Degree to which a product or system can be effectively and efficiently modified without introducing defects or degrading existing product quality.	・Bug handling	・ Clarify the method of modifying bugs and the party of responsible.
			・Contract code	・ In BC systems, since it is impossible to modify (falsify) written code, clarify how to respond when a bug is found in the contract code.
			・Hard fork	・ In BC systems, since it is impossible to modify (falsify) written code, clarify the way to rollback a block when a bug or improper data by illegal access is found.
	Testability	Degree of effectiveness and efficiency with which test criteria can be established for a system, product, or component and tests can be performed to determine whether those criteria have been met.	・Blockchain platform	・ Clarify what kind of function and performance test can be done in what environment and how the platform is influenced when the environment changes because test results are influenced by the structure of nodes and networks.
			・Tolerance of node failure and network failure ・Scalability ・Consensus algorithm	・ Since resistance tests against the failure of nodes and networks, capacity and node extensibility test, and testing of consensus algorithm are important in a distributed environment, clarify what kind of test can be carried out.
			・Contract code	・ In BC systems, since it is impossible to modify (falsify) written code, contract code needs adequate testing. Therefore, clarify what kind of tests can be carried out.

6. Evaluation model (3)Cost（1/2）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
R&D	R&D of Blockchain platform technical elements	R&D cost (pre-implementation cost)	・New consensus algorithm ・Development of high-speed P2P protocol	・ Estimate the R&D cost for the technical elements of the BC platform for improving processing performance. ・ Organize the function and performance levels of existing technologies and estimate the cost by clarifying the target function and performance.
	R&D of subsystems		・Application development ・Contract development environment	・ Estimate the R&D cost of the subsystem to expand application fields. ・ Estimate the cost by clarifying the target function and performance.
Implementation（Commercialization）	Hardware cost	Cost for system implementation (cost at the time of implementation)	・Node ・Network ・Subsystem	・ Clarify the objects and range included in the cost. ・ Clarify that estimation objects and the range of estimation are different between the types (public and private types) of platforms. For example, in the case of the public type, materials of unspecified participants are not (or cannot be) estimated as a cost, and in the case of the private type, since the number of nodes and participants are clear and equipment that carries a server role is assumed, these costs are included in the estimation.
	Software cost		・OS ・Middleware ・Application	・ Clarify objects and range included in cost. ・ Clarify that estimation objects and the range of estimation are different between the types (public and private types) of platforms. For example, in the case of the public type, materials of unspecified participants are not (or cannot be) estimated as a cost, and in the case of the private type, since the number of nodes and participants are clear and equipment that carries a server role is assumed, the cost of software is included in the estimation.
	System implementation cost		・Assembly, implementation, and test	・ Clarify the objects and range included in cost. ・ Clarify that estimation objects and range of estimation are different between the types (public and private types) of platforms. For example, in the case of the public type, materials of unspecified participants are not (or cannot be) estimated as a cost, and in the case of the private type, since the number of nodes and participants are clear and equipment that carries a server role is assumed, the cost of installation is included in the estimation.

6. Evaluation model (3)Cost（2/2）

Characteristics	Sub-characteristics	Outline of characteristics	Related BC technology & BC characteristics	Points to remember and remarks on using the evaluation model
Maintenance & operation	Operational cost	Cost for maintenance and operation of the system (cost after implementation)	・Node ・Network ・Cost for consensus（Impact on cost by difference of consensus algorithm）	・ Clarify the objects and range included in the cost.
	Maintenance cost		・Bug modification	・ Clarify the objects and range included in the cost. ・ Clarify the modification frequency that it expected to be high due to new technology, and the technical level required for modifications.

7. How to use the evaluation model

- It is assumed that the evaluation model on the BC system created in this survey will mainly be used in the following three scenes:

	<<Scenes of utilization>>	<<Method, effect, remarks on utilization>>
1	**Comparative evaluation when existing system is replaced by BC system**	• A case where a system vender carriers out the evaluation and communicates the result to the system introducer is assumed. • While evaluating the entire system within the framework of evaluation system vendors are currently using (such as ISO/IEC 25010 and IPA), the quality, maintenance and operation, and cost of blockchain technology are evaluated and compared by using the entire evaluation model, and the result of the evaluation is proposed to customers.
2	**Evaluation of verification test results of BC system**	• In addition to a case where the evaluation is based on the entire evaluation model, a case where items to be noted are selected in accordance with the requirement of a use case where the evaluation is also assumed. • It is hoped that cases are accumulated so that the results of plural demonstration tests can be compared.
3	**Comparative evaluation when selecting the blockchain platform to be used**	• When making the comparison, evaluation items of the evaluation model especially closely related to a blockchain platform are the focus, and a decision is made (such as throughput, block confirmation, and reliability). • An additional list showing functions that can be used is necessary because functions that can be supported differs from one platform to another.

8. Future issues and consideration for promoting social implementation of blockchain technology

Future issues to be considered in utilizing the evaluation model

The evaluation model examined this time becomes meaningful in practical use, and in order to continue to be utilized, the following efforts are necessary.

- ✓ Implementation and accumulation of evaluations of the actual system
- ✓ Verification of coverage of evaluation model according to accumulation of use cases
- ✓ Maintenance of evaluation model according to technology development
- ✓ International standardization of evaluation model

Consideration for Promoting Social Implementation of Blockchain Technology

In order for the evaluation model to contribute to the social implementation of blockchain technology, the following efforts are necessary.

- ✓ Grasp and familiarize the characteristics of blockchain technology
- ✓ Build a problem solving consortium with stakeholders
- ✓ Organize the element technologies necessary for constructing BC systems
- ✓ Create an environment for social implementation (revision of regulations and institutions, etc.)