P-256 *IT新技術がもたらすGLPの将来像*

奥村リョウ（キッセイ薬品工業株式会社）
日本QA研究会GLP部会第3分科会

The future of GLP brought about by new IT technologies

Ryo Okumura (Kissei Pharmaceutical Co., Ltd.)
Japan Society of Quality Assurance, GLP Division, Study Group 3.

ryo_okumura@pharm.kissei.co.jp

Outline

The computerized system has led to increases in efficiency and reliability in GLP facilities. On the other hand, digitalized data and records produced new problems in quality assurance such as difficulties in detecting the falsified signs, distinguishing between originals and duplications, and insufficient duration of electronic recording media; however, such problems have been solved by proper CSV implementation and operational procedure setting. The computerized system, which allows more objective and rigid management of data and records than paper, has become a central core in GLP quality assurance.

Today's innovative IT technologies have made remarkable progress and have been creating the products and services we could not have previously imagined. New IT technologies have permeated our daily lives as common services within a short period and are also expected to bring more efficiency within GLP facilities. However, the introduction of innovative IT technologies has been hesitant in GLP facilities and they are hardly pervasive. This is likely to be caused by difficulty in establishing the new method for quality assurance because there is a case in which those IT technologies are hard to apply to the conventional CSV approaches. In order to overcome this, it is important that the industry and regulatory authorities share the GLP's future image brought about by innovative IT technologies to seek problems and the counter-measures.

Therefore, the study group of Japan Society of Quality Assurance GLP Division 3 started the investigation by envisioning the pictures of near-future GLP facilities which had introduced the latest IT technologies. From considerations, we outlined "Prospects of innovation for GLP data storage based on blockchain technology" and "Basic approach about quality assurance of AI products" on the poster.

Prospects of innovation for GLP data storage based on blockchain technology

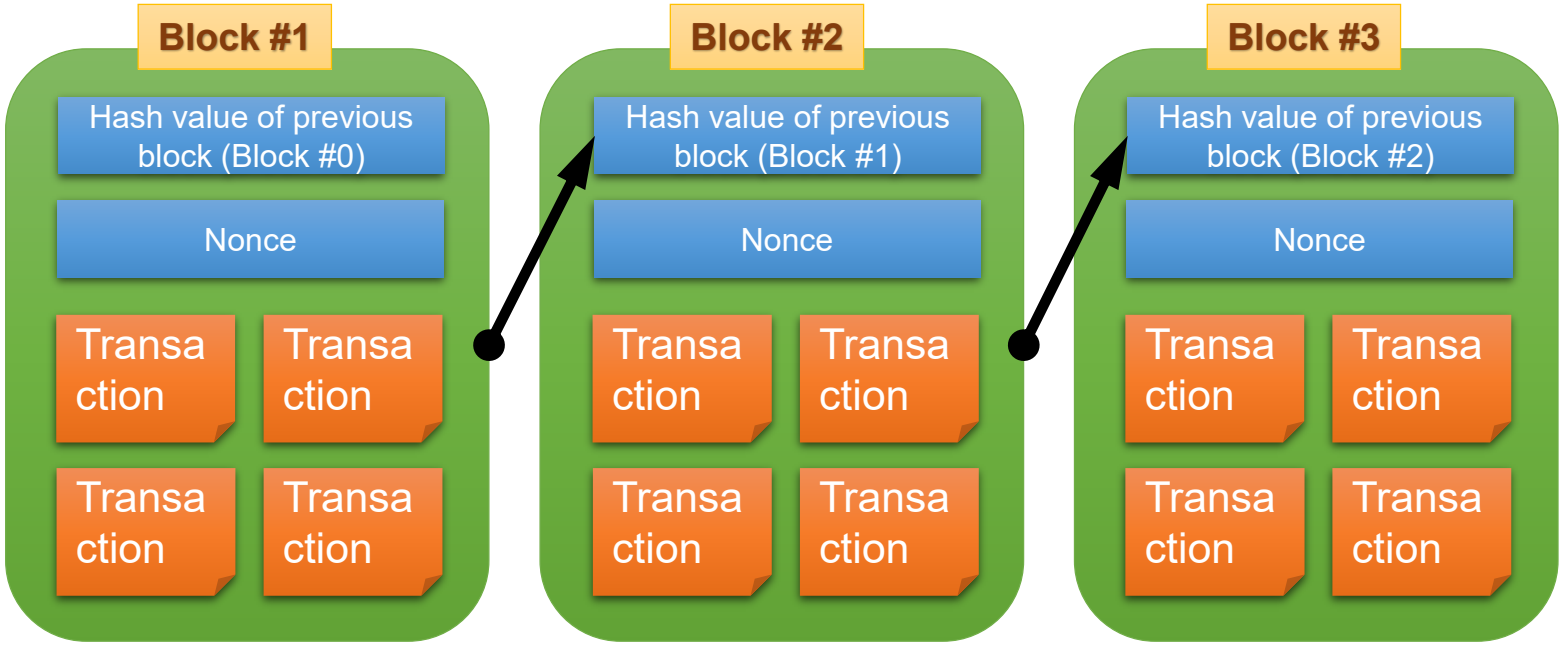
Introduction

Blockchain technology is a method of recording data by combining encryption technology on multiple computers that make up a distributed network. It is widely known as the basic technology for virtual currencies such as Bitcoin, but it is expected to be applied to various fields.

Blockchain is resistant to tampering and all the participants have the same chains, therefore, it is easy to restore lost data with it. For that reason, blockchain can be a measure to solve many problems in saving electronic raw data. Our team will introduce three models using blockchain in GLP data storage.

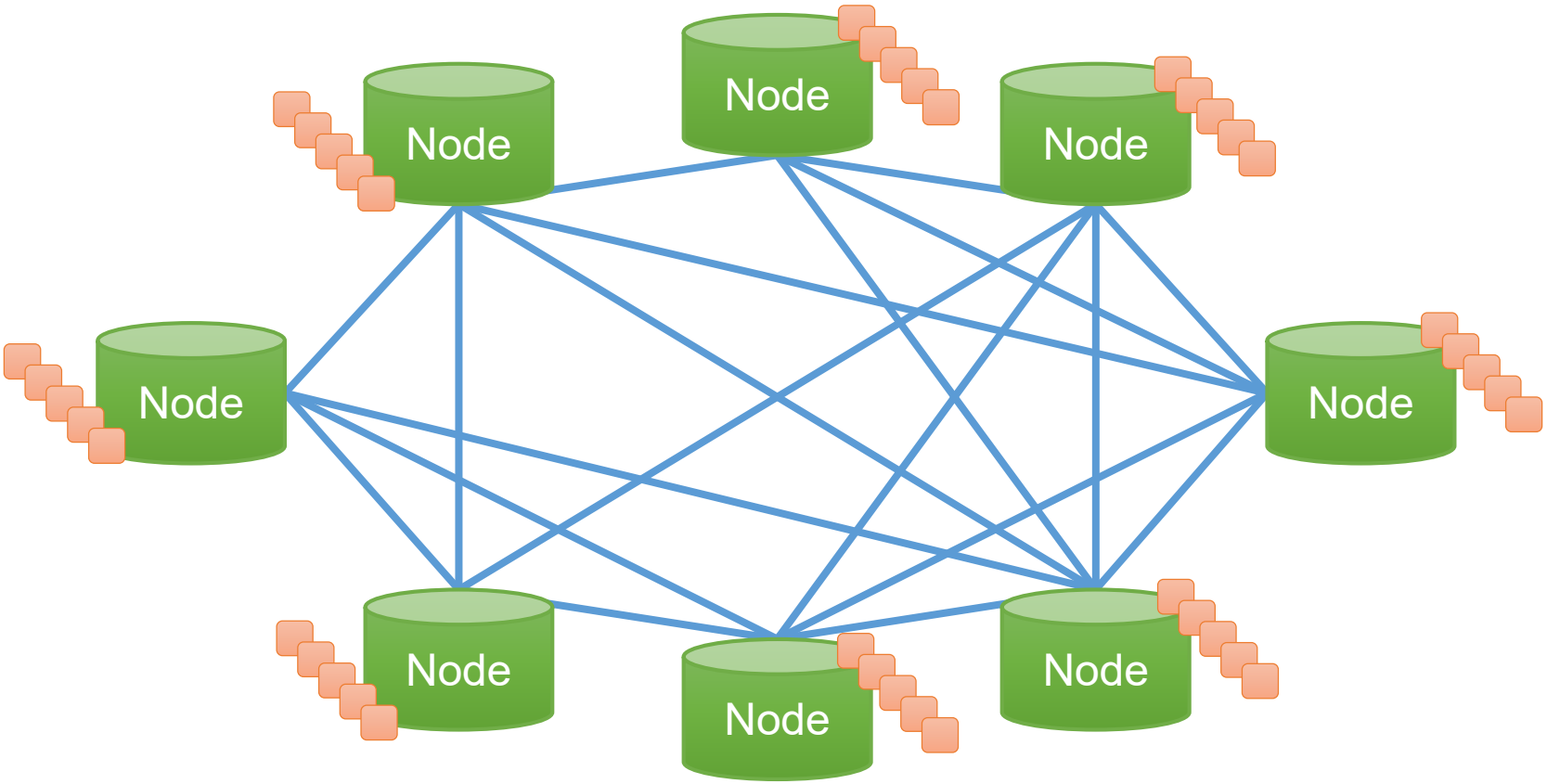
Basic of blockchain

1. One block stores transactions, hash value of previous block, and the nonce.
2. The blocks are associated with each other by including the hash value of the previous block in the new block.
3. Since the hash value is a unique character string that can be obtained by performing a calculation on the data in a fixed procedure, even if one bit changes the data will be a completely different value.
4. Blockchain have high tamper resistance by connecting blocks that are related to each other. (If you tamper with a block, you also need to tamper with all the blocks that follow it)



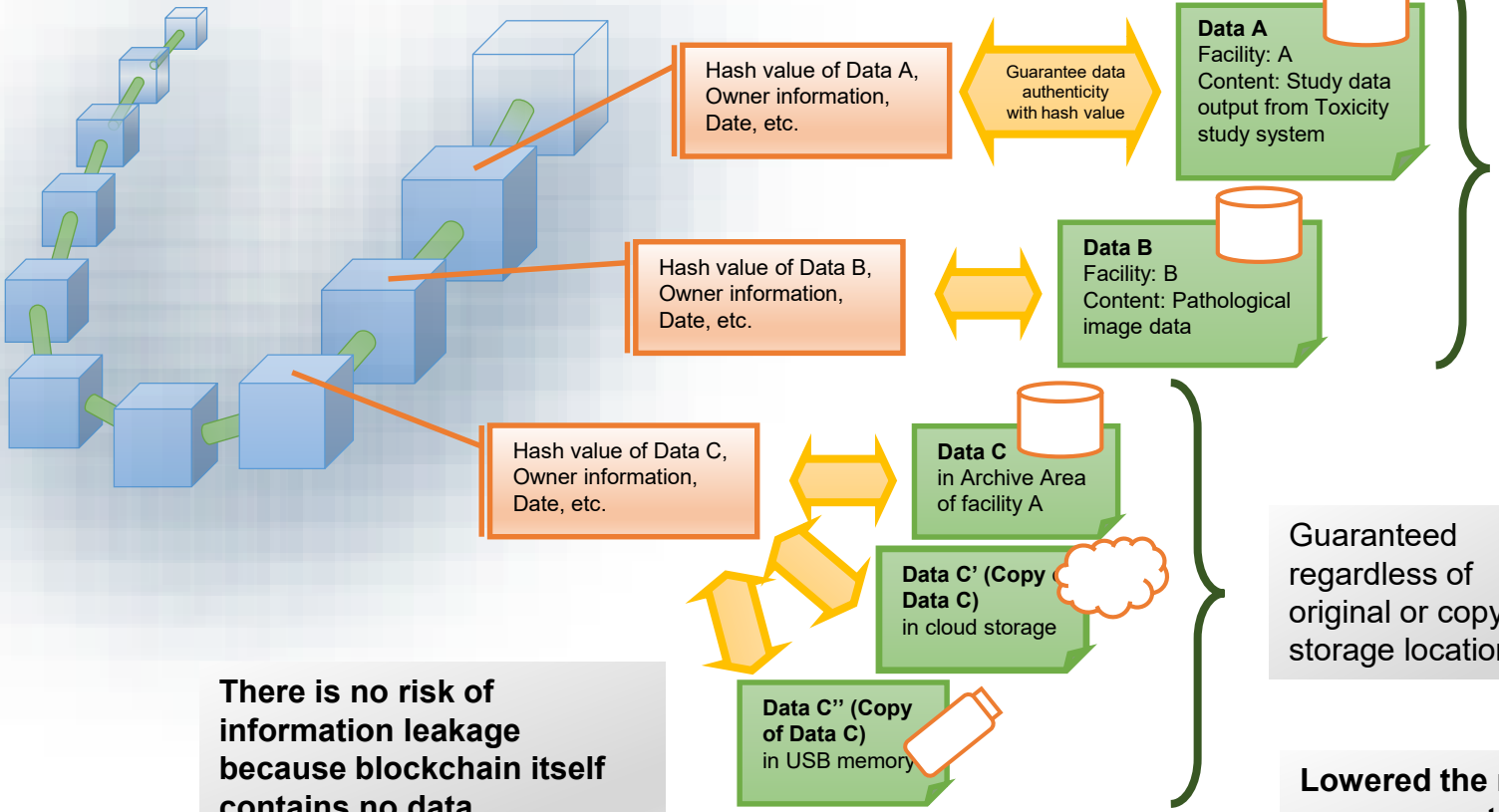
Basic of blockchain

5. When adding a new block to the blockchain, it must be approved by a consensus algorithm such as Proof of Work.
6. Blockchain has high fault tolerance because it is synchronized between nodes. (Backups are held equally on all nodes)



Model 1

Store the hash value of various data in a blockchain.
A hash value protected by the blockchain guarantee that data is a certified copy.



Guaranteed regardless of file size, file type, etc.

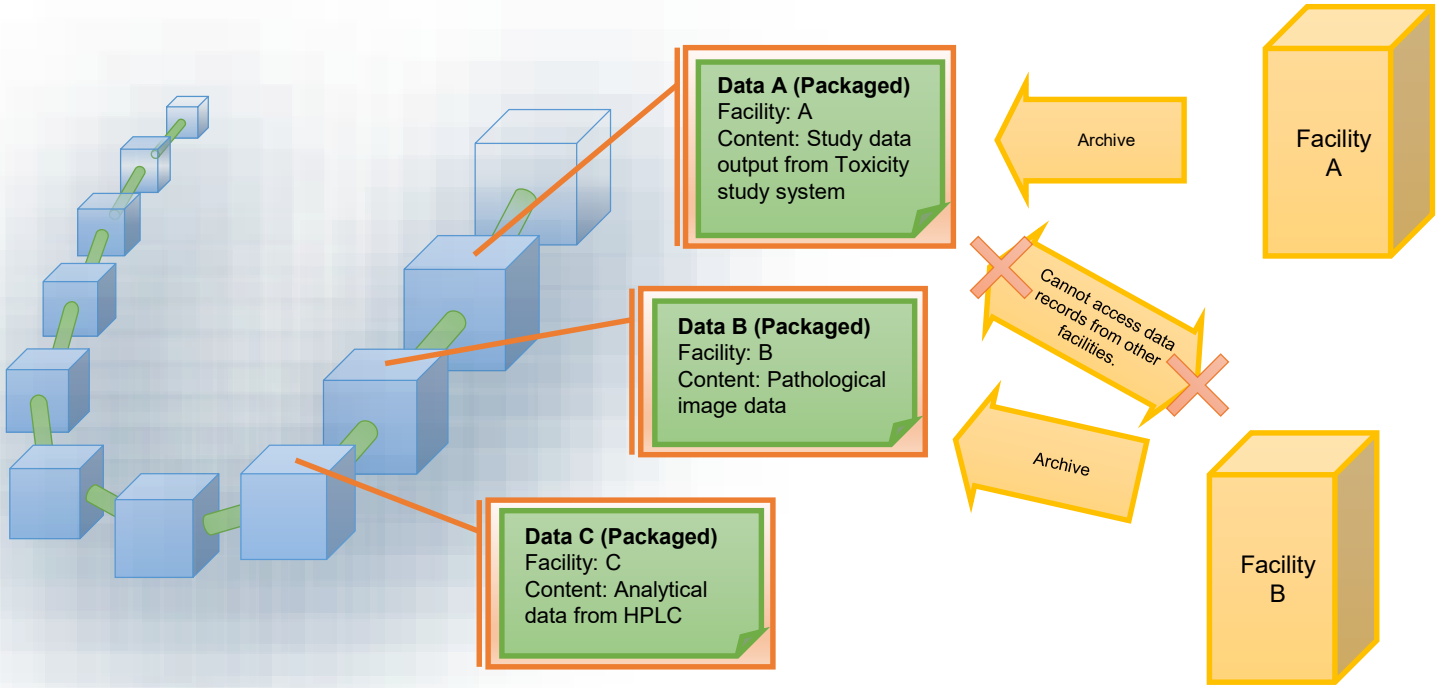
Guaranteed regardless of original or copy, storage location, etc.

There is no risk of information leakage because blockchain itself contains no data.

Lowered the required level of management of stored data allows the management cost to be reduced.

Model 2

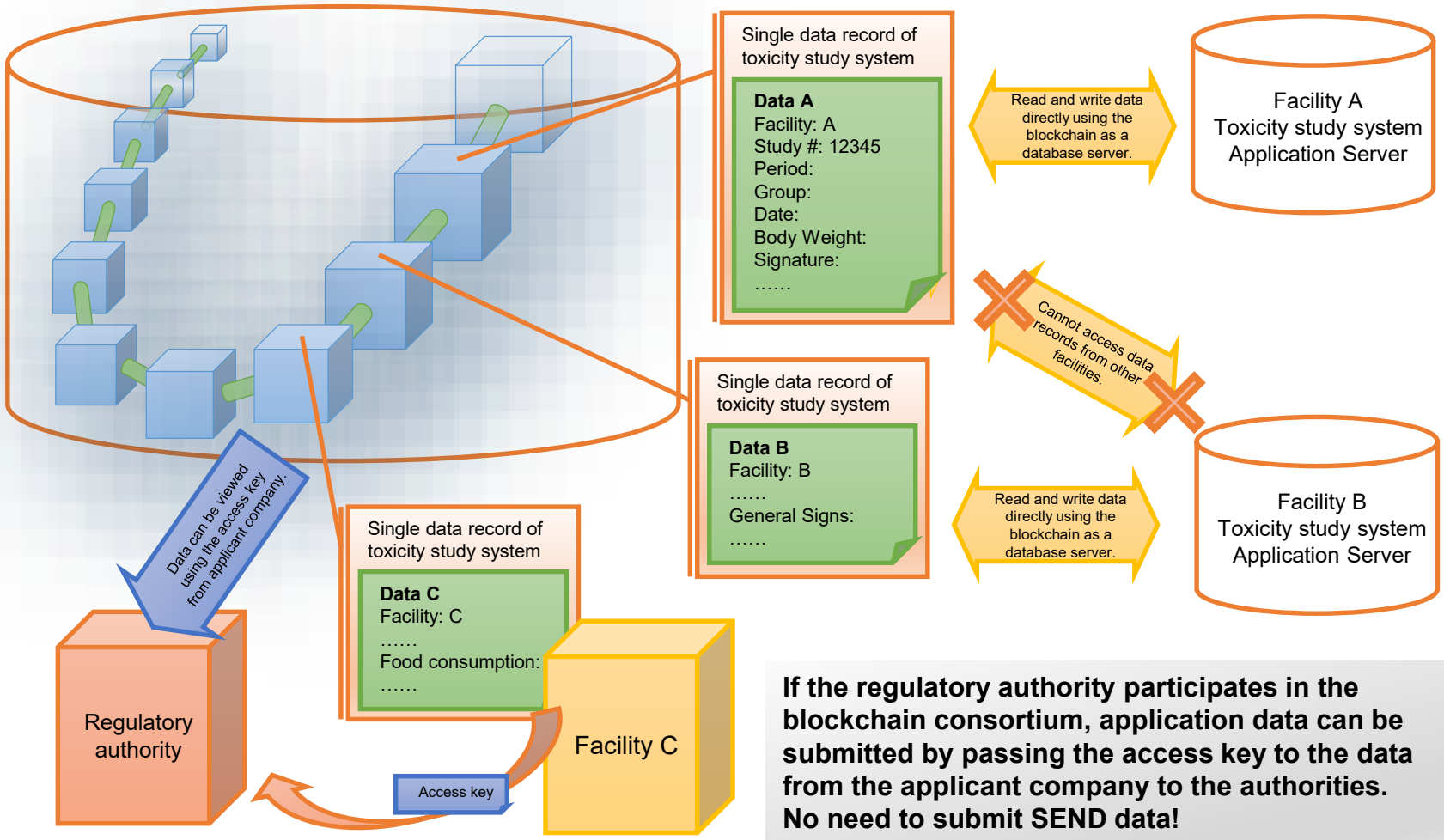
Store various data itself in the blockchain as a archive storage.
Original data is protected by blockchain.
Huge storage is required to store data for all facilities that use the blockchain consortium.



Since the blockchain is held in each participating facility's server, it is also effective for disaster recovery.

Model 3

Store each record of standardized raw data in a blockchain.
Use the blockchain as a database server for toxicity study system.



If the regulatory authority participates in the blockchain consortium, application data can be submitted by passing the access key to the data from the applicant company to the authorities. No need to submit SEND data!

Conclusion

Although each of three models has its own problems, it could have a major impact on GLP data storage if realized. Model 1 may solve problems in authenticity assurance of the data each facility possesses at low cost. In addition, it has a high feasibility because its form has been already used commercially. Model 2 is generally considered improper; however, it will be able to be optimized if operated by a consortium. Model 3 will realize an unprecedented platform which unifies a series of the processes from data recording to the submission to the authority, which would set a higher hurdle. Nevertheless, blockchain has been expanded in application in various fields and would make a paradigm shift also in GLP.

Basic approach about quality assurance of AI products

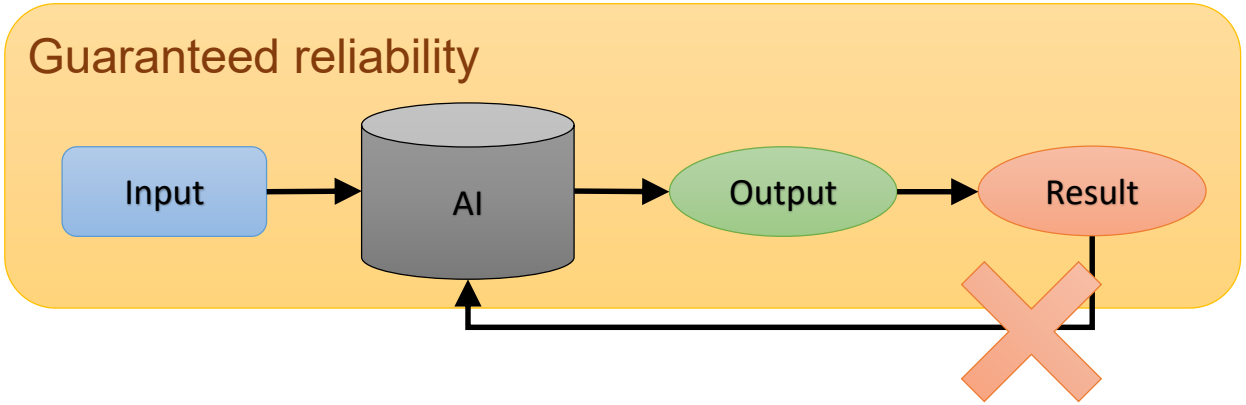
Introduction

The internal process of AI products is a black box from the human perspective, but since the reliability can be guaranteed by verifying the input / output, the black box does not pose a problem in quality assurance. The issue of quality assurance of AI products is that the internal process that derives the result is changed by learning, and the process that is not guaranteed the reliability affects the result.

As a method to solve this problem, we examined a model that separates internal processes that change due to learning and the results of internal processes with guaranteed quality.

Model 1 Stopped learning model

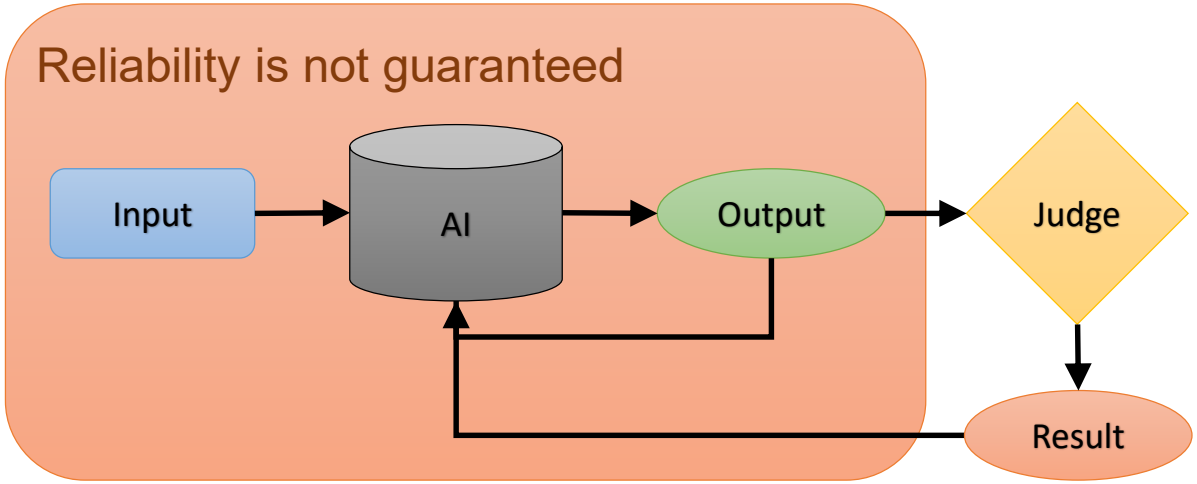
Avoiding the risk of AI products learning during operation by not learning during operation.



As a concern, the difference in the learning data up to the introduction and the data input during operation may affect the output, but there will be no problem with Input within the range where reliability is guaranteed at the time of introduction. However, since there is a possibility that it cannot handle Input that exceeds the expected range, the applicable range will be limited to a certain extent, and versatility will be low. For example, what happens when the staining method is changed in the pathological image diagnosis can be considered.

Model 2 Human judge model

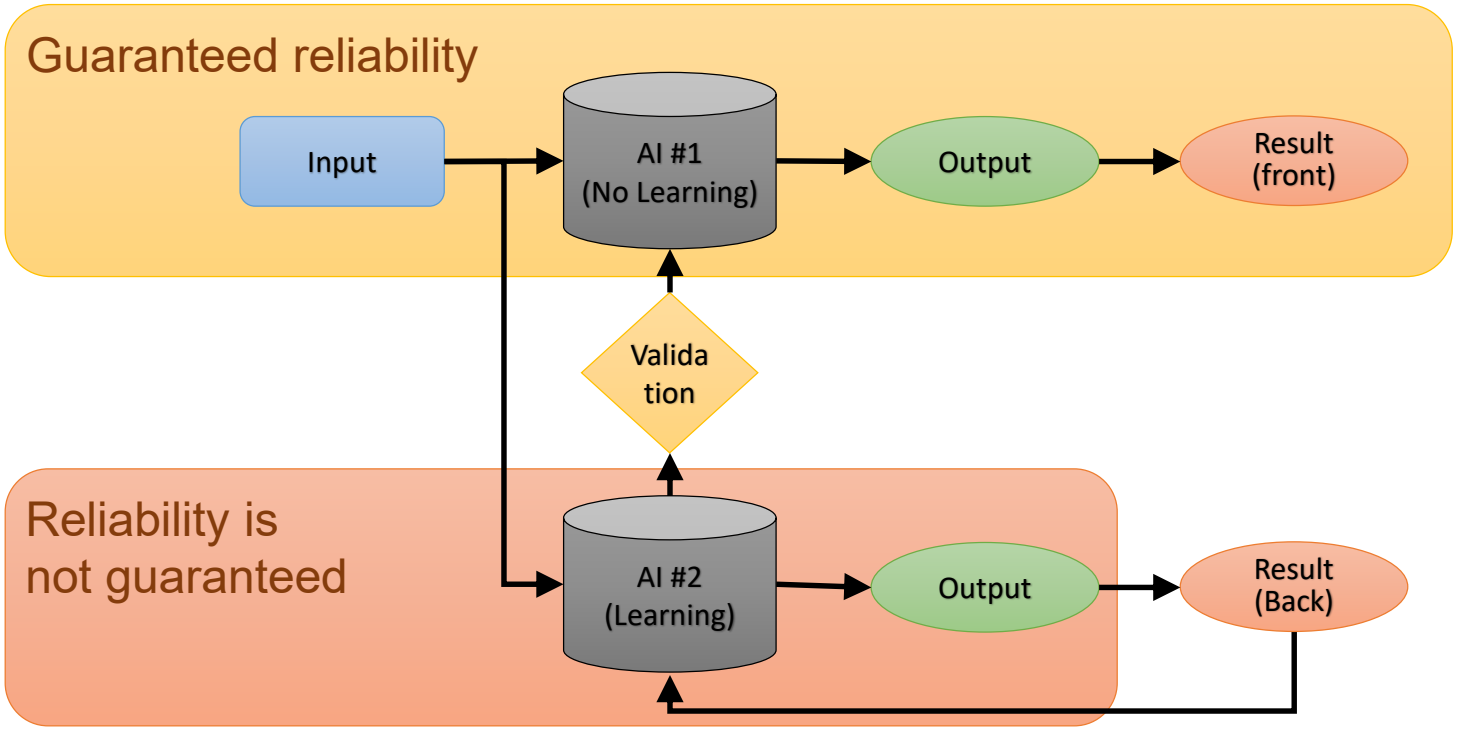
Avoid the risk of AI products learning during operation by incorporating a process of ultimately making human judgments.



The accuracy of output can be expected to be improved by continuing the learning at the operation site, but in the case of machine learning, especially in the case of supervised learning, there is a risk that correct output cannot be performed if human judgment is not appropriate. Alternatively, in the case of deep learning, AI automatically finds the feature amount without human control, so there is a risk that correct output cannot be performed depending on the quality of the input. Therefore, it is considered necessary to have a mechanism that can perform re-learning and roll back to the past learning state if necessary.

Model 3 Learning environment isolation model

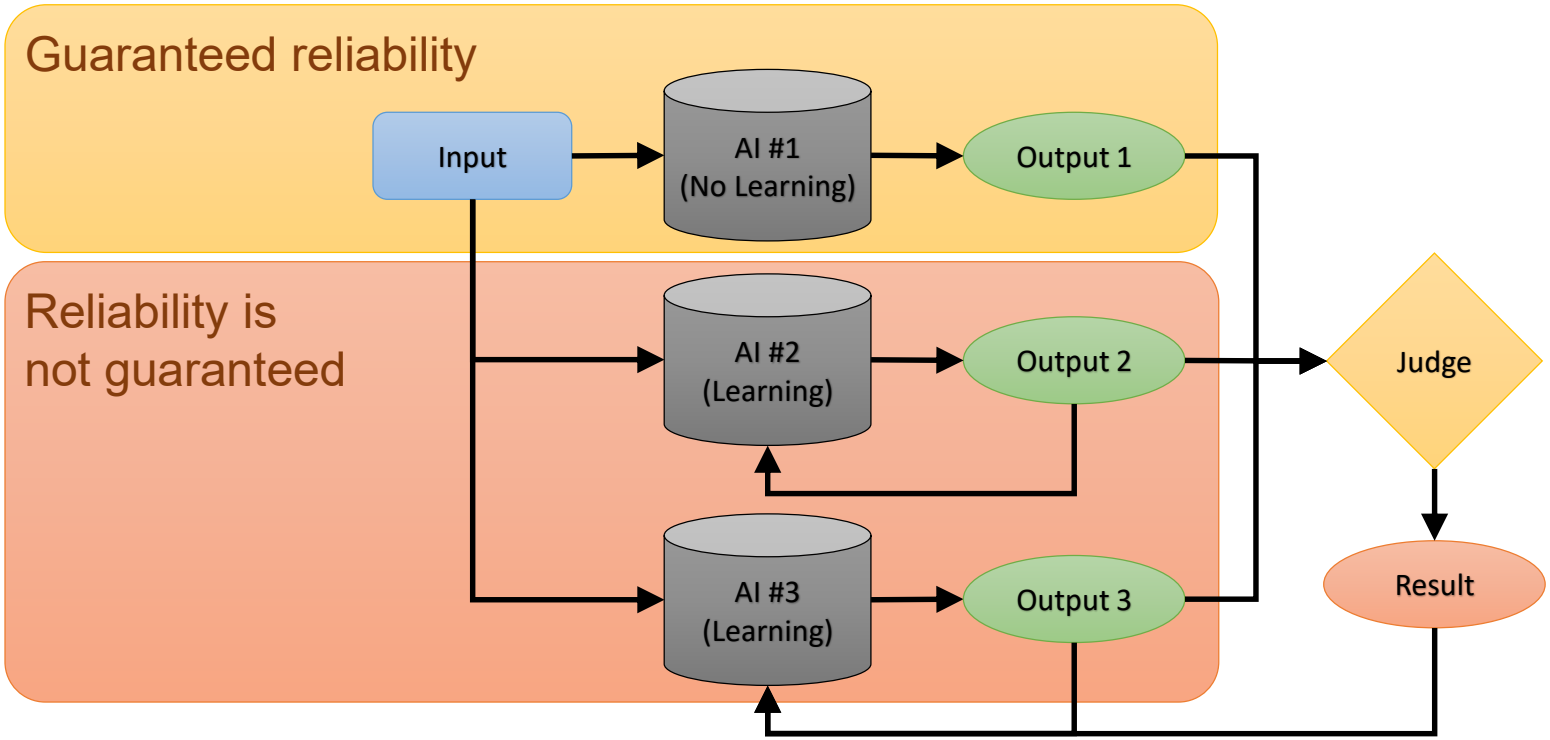
This is a model in which the learning AI runs in parallel in the background, performs validation regularly, and if there are no problems, applies the learning results to the learning stop AI.



Generating raw data while benefiting from learning can only be a validated system. For the validation of learning AI, the method of inputting the audit data prepared in advance and confirming that the output has not changed is realistic. Also, if it is a white-box AI that can visualize the inference process of AI, a method of confirming that there is no change in the inference process can be considered.

Model 4 Multiple concurrent model

In order to complement the human judge model, this is a model in which multiple AIs are operated in parallel and ultimately human judgments are made.



A combination of learning continuation AI, learning stop AI, and AI with different algorithms and parameters. For example, in the case of a pathological image diagnosis system, it is assumed that a person evaluates the findings of each AI to determine raw data. It is possible to envisage a model in which the output is decided by majority without any human judgment, but it is not suitable for situations where reliability assurance is required.

Conclusion

Which model is appropriate depends on the requirements to be applied, but from the viewpoint of guaranteeing the reliability of a dynamically changing black box, a common methodology can be considered. That is, to understand what part is a black box, when and how it can change, and to incorporate that part from the data generation process. Although four models are presented in this paper, the methodologies can be summarized as two: isolation of learning environment and human judge. Given this, reliability assurance of AI products is not very complicated. It is necessary to consider each AI product thoroughly, but it is not impossible. Considering the merits of AI technology, there is no reason to deny its use.

Japan Society of Quality Assurance, GLP Division, Study Group 3.

Author: Ryo Okumura (Kissei Pharmaceutical Co., Ltd.)
Co-Author: Atsuko Sato (Aska Pharmaceutical Co., Ltd.),
Motoyuki Ito (Daiichi Sankyo Co., Ltd.),
Shinichi Horikawa (Ina Research Inc.),
Hiko Fukumura (Maruho Co.,Ltd.),
Masahito Iemura (Senju Pharmaceutical Co., Ltd.),
Megumi Ueda (Sumitomo Dainippon Pharma Co., Ltd.),
Kaori Saito (Taiho Pharmaceutical Co., Ltd.), Yoshiaki Hiraishi (Terumo Corporation)

Contact: ryo_okumura@pharm.kissei.co.jp