

White paper

Trustable Internet

Towards a trustable Internet in which people can use information securely

Version 1.0

October 13, 2022

TIAL : Trusted Internet Architecture Lab.

TABLE OF CONTENTS

	Executive Summary	3
	Assumed reader and prerequisite knowledge.....	4
1	Background.....	4
	1.1 Progress and reliability of a data-based society	4
	1.2 Use case and issues	4
	1.3 Requirements analysis.....	5
2	Trustable Internet.....	6
	2.1 Trustable Internet architecture.....	6
	2.2 Endorsement layer	7
3	Approach for achieving the Trustable Internet	8
	3.1 Global movement toward a credible society	8
	3.2 Endorsement data/graphs that provide evidence for data credibility	8
	3.3 Endorsement graph reflecting the physical space	10
	3.4 Method for supporting viewer’s judgement of data credibility.....	12
	3.5 Overall behavior of the Trustable Internet	13
4	Future work.....	14
5	References.....	15

EXECUTIVE SUMMARY

The Internet has become an essential part of social and economic activities. The widespread use of smartphones and social networking services has made it possible for individuals and organizations to send and view data throughout the world anytime, anywhere. Some data are disinformation containing false information that can cause social problems. In many cases, viewers cannot confirm the credibility of the data on the Internet, so even if there is suspicion of disinformation, they have no choice but to make their own judgments about the credibility. Decisions made on the basis of such data may not be accurate if the data contains uncertain content.

Data authenticity has traditionally been ensured by incorporating trails for each individual system, such as document approval management systems and package tracking systems, or by building trust between data senders and receivers in advance in ways other than the Internet. To make use of such a method in general, the system which can judge the credibility of the data on the Internet is desired. However, it is difficult to achieve this without affecting the existing structure of the Internet.

Therefore, we propose an architecture called "Trustable Internet" that overlays a layer on the Internet, which has a mechanism to confirm the credibility of the data, so that the existing Internet is not affected, and Internet users can use the Web/applications as before.

The architecture has an endorsement layer that provides an interface for storing and sharing "endorsed" information as additional information, which is the basis for data credibility. In addition, it has an endorsement graph with a data structure expressing the connection between additional information linked to the data.

The additional information as evidence for credibility judgment includes the information on the Internet as well as the information from the physical space added by a person or thing, and is given by the person or the device when the data is generated or after the generation. When browsing data on the Internet, the additional information can be confirmed from various viewpoints as necessary, and the information is overlaid on a browser in consideration of the subjectivity of the individual. It also makes data on the Internet more reliable by pointing out the lack of information in the endorsement graph to judge credibility, and allowing new information to be added.

Overlay to the Internet enables a general-purpose approach that ensures data credibility without affecting conventional systems. This makes it possible to add information that can be confirmed to various data on the Internet. The viewer can also obtain additional information capable of confirming the credibility of the data and make a subjective judgment on the basis of the information.

Currently, people have no choice but to judge the credibility of data on the internet by individuals, but with the support of data in the physical space and the related additional information, it is possible to confirm the credibility of data from various viewpoints and make decisions more accurately. This is expected to make it possible to prevent the use of unreliable data and the re-spread of disinformation and fake news. Furthermore, by expanding the scope of judgment, the scope of economic and social activities will be expanded, and the optimal method will be realized in society. For example, it is considered that necessary information can be quickly provided when quick judgment is required for disasters, etc.

ASSUMED READER AND PREREQUISITE KNOWLEDGE

This whitepaper is intended for IT business and IT engineers, as well as for policy makers.

- Target reader: Chapters 1 and 2 are intended to be understood by business stakeholders and policymakers who do not necessarily have a technical background. Chapter 3 covers technical aspects and is intended for IT engineers.
- Prerequisite knowledge: Chapter 3 also assumes knowledge of Web technologies standardized by the W3C, etc.

Note: In this white paper, three terms used in this white paper are used with the following meanings.

- information: endorsed/related data
- data: data that obtained from the Internet or from physical space and that is non-endorsed or to be endorsed.
- data credibility: degree of correctness of data

1 BACKGROUND

1.1 Progress and reliability of data-based society

The Internet has become an essential part of modern social and economic activities. The spread of smartphones and social networking services (SNSs) has made it possible for individuals and organizations to send and view data throughout the world anytime, anywhere. Data obtained from sensors and systems have been used by companies, governments and other organizations for various services. Thus, the credibility of data exchanged over the Internet has become more important than ever.

Some data are disinformation containing false information that can cause social problems. In many cases, viewers cannot confirm the credibility of the data on the Internet, so even if there is suspicion of disinformation, they have no choice but to make their own judgments about the credibility. Decisions made on the basis of such data may not be accurate if the data contains uncertain content.

To improve the credibility of data, there is a growing movement to develop standards and legal systems to identify individuals and organizations sending data over the Internet [1]. In addition to the identification of individuals and organizations, methods are being developed for confirming the qualifications and abilities of individuals and the registration and credit information of organizations [3][4]. A system to confirm the sender of data on the Internet, as well as information necessary for judging the credibility of the data, such as whether the person is an expert in the field, is under consideration, but the system has yet to be put into practical use.

1.2 Use case and issues

Data, the credibility of which is uncertain, are generated and spread, in the event of a disaster is discussed as a use case.

Disaster Information Transmission (FIG 1. Disaster Information Transmission): In response to the news of a river flooding, a resident posted the news on an SNS with a message saying, "flood is likely to occur." He only wanted to inform people the possible occurrence of a flood. However, uneasy readers who could not confirm the news saw the message and thought that a flood is occurring in their neighborhood, so they told people around them that a flood is occurring. In the transmission process, information on various assumptions was added to the original message that referred to the possibility of flooding. Eventually, some residents began to believe that a flood was occurring, and the message

changed to mention the fact that a flood has occurred and is spreading. Therefore, many residents became anxious, and their arbitrary actions, such as buying up goods and triggering riots, caused social and economic confusion.

When there is a lack of accurate information, as during a disaster, information based on speculation is added to a small amount of information, and the information is spread as inaccurate information in an instant. The content often changes during the transfer because it spreads without people knowing which information is accurate and which is speculation. The Internet is a place where a wide variety of people can freely transmit data, but in many cases, the credibility of the transmitted data cannot be judged; thus, the data must be used with its credibility uncertain. When the decision on the credibility of the data is left to the viewer, it can lead to misunderstandings and human errors.

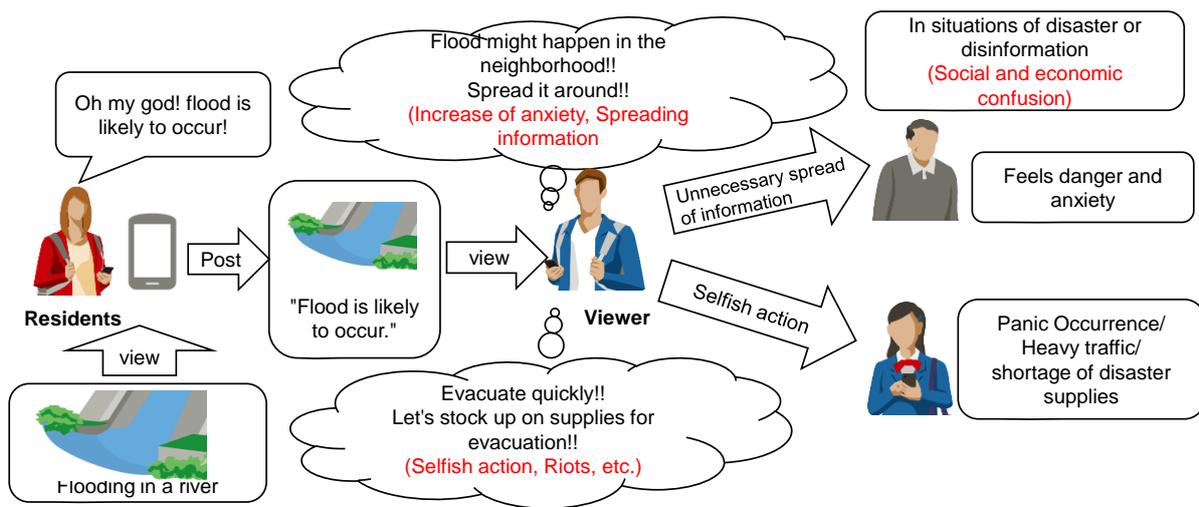


FIG 1. Disaster Information Transmission

1.3 Requirements analysis

There are two types of data for the use cases described in the previous section, i.e., That is, initially generated data and secondarily transmitted data.

Initially generated data: Data initially generated by people transcribing or photographing what people see or hear. The data may also be generated by a device or system. These data do not represent all the situations in the physical space. Also, the data generated by humans may be supplemented by imagination or speculation, or be difficult to express in words, so the receiver of the data may not understand the intention of the original data.

Secondarily transmitted data: Data that a person first receives and then re-transmits. When a message is received by a person, the message may be retransmitted as is, but may be retransmitted by cutting out a part of the message, interpreting the message and correcting the expression. This can involve the imagination and speculation of the receiver. Similarly, such data may be further transmitted after people have received the data.

When the initial data are generated or re-transmitted, the data that differ from the fact are generated in cyberspace because the content changes due to the speculation and interpretation of the person or thing.

To prevent a viewer from unintentionally taking part in the process of generating data that differ from the facts, if the viewer knows that the data they have are not credible, they will be less likely to spread the data or take wrong actions on the basis of the data. By having a support system that shows the credibility of such data, the viewer will not believe the data as is and mistakes will be reduced.

This white paper focuses on the information aspect and aims to make more accurate judgments in cyberspace by accumulating useful information and enhancing the relationship between information. This can be achieved by considering information from a third-party perspective, information from the real world, and the viewpoint of the viewer, as follows.

Requirement(i): Acquisition of evidence for the credibility of data in cyberspace

It is not always possible to judge the credibility of data such as posted messages. However, if the sender of the data can be confirmed as information related to the data, credibility judgment can be supported. It is also possible to find trends such as the existence of contradictions among a multifaceted and large number of information by strengthening the relationship between information. Furthermore, if the information held by the third party who has confirmed the credibility of the generated data, such as an expert generating or judging the generated data can be known later, it can also be evidence for the viewer to judge data credibility.

Requirement (ii): Provision of evidence for the data to be consistent with the state of physical space

As shown in FIG. 1, even though no flooding occurred in the neighborhood, the text data included the term flood, which was different from reality. However, by using information in the physical space and by indicating more credible fact information acquired in that physical space, for example, water level information and the location of a person, it is possible to understand whether the word "flood" created in cyberspace and the fact differ.

Requirement (iii): Mechanism to actively provide evidence information.

It is also necessary to construct an ecosystem in which the above mechanism is operated continuously in the real world. Viewers may feel that information from the sender's or a third party's point of view alone is not sufficient to judge credibility. In such a case, enabling the viewer to request further evidence will lead to resolving questions about the credibility of the data. By evaluating the provider of information that is valuable to the viewer as highly reliable, it is possible to give an incentive for the provider to provide evidence.

2 TRUSTABLE INTERNET

2.1 Trustable Internet architecture

The credibility of data has traditionally been ensured by incorporating trails for each individual system, such as document approval management systems and package tracking systems, or by building trust with partners in ways other than the Internet. To use these methods for general purpose, a mechanism for judging the credibility of the data on the Internet is needed. However, it is difficult to implement such a mechanism without affecting the current Internet.

We propose an architecture that overlays a layer on the Internet with a mechanism to enhance the credibility of data. This does not affect the current Internet, enables Internet users to use the Web and applications in the same manner as before, and enables them to acquire additional information that serves as evidence for the credibility of data. We call this architecture as the *Trustable Internet*.

On the Trustable Internet, additional information related to target data and endorsed by a human and/or thing is defined as *endorsement data*. A layer having a mechanism for storing and sharing the data is defined as an *endorsement layer* (FIG 2).



FIG 2 .Trustable Internet architecture

2.2 Endorsement layer

An endorsement layer manages endorsement data added to target data. And endorsement layer has a mechanism for expressing the relationship of the endorsement data linked to the target data. This structure with the target data as a starting point is defined as an *endorsement graph* (FIG 3). The details of the endorsement data and implementation of an endorsement graph are described in the next chapter.

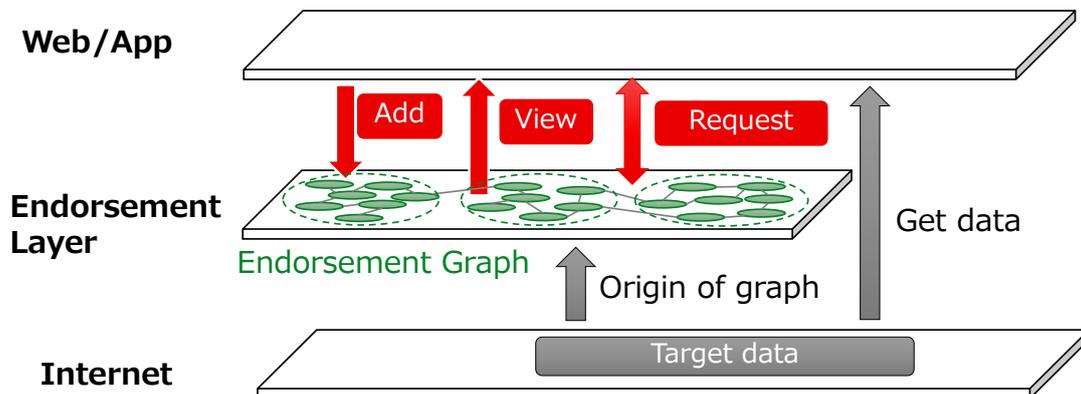


FIG 3. Endorsement Graph

An endorsement layer includes the following interfaces for accessing an endorsement graph.

1. **Add:** An information sender (person or system) can add endorsement data to improve the credibility of the target data acquired from the Internet or the physical space. The endorsement data are added to the endorsement graph at arbitrary time after applying tampering prevention processing.
2. **View:** To judge the credibility of the target data, viewers can view the endorsement data linked to the target data via an endorsement graph.
3. **Request:** If the viewer determines that they cannot judge the credibility of the data due to a lack of additional information, they can request to add more information to the target data. The information provider (person or system) can confirm the number of requests for the target data.

3 APPROACH FOR ACHIEVING TRUSTABLE INTERNET

3.1 Global movement toward a credible society

There has been an active global movement toward a credible society.

In Europe, the European Commission is making progress in applying and revising eIDAS, which is an EU regulation, to provide trust services through the implementation of digital identities in cyberspace that are paired with entities such as individuals and organizations [1]. For example, the specification of identity and attribution in electronic transactions between EU citizens and companies has been unified, and the active introduction of digital identity wallets has been promoted.

In Japan, the “Trusted Web Promotion Council” under the Cabinet Secretariat is formulating “Trusted Web” which is a vision and architecture that uses digital identities of individuals, organizations, and devices to ensure trustability in data exchange.

The World Wide Web Consortium (W3C) is developing Decentralized Identifiers (DIDs) [2], which is a standard for decentralized identities, and the Data Model for Verifiable Credentials (VC) [3], which is a digital certificate that can be verified by a third party for digital identities such as personal acquisition credentials. Based on these standards, various systems providing decentralized identities have been developed [4] [5] [6]. The world is approaching an era where people can freely disclose various digitized qualification information on the Internet by their own will.

In general, these activities have been carried out to ensure the credibility of data by using digital identity through ensuring the authenticity of the data source on the basis of Web technology and public key cryptography. However, the Trustable Internet aims for a data society based on a credible Internet that can eliminate the adverse effects of disinformation by focusing on the verification area of data in the physical space while taking into account these new standards and criteria in addition to standard Web technology and public key cryptography.

The following sections describe the approach to implementing the endorsement layer and endorsement graph. The relations between the above approaches and requirements mentioned in 1.3 are as follows.

- 3.2 Endorsement data/graphs that provide evidence for data credibility: requirement (i)
- 3.3 Endorsement graphs reflecting physical space: requirement (ii)
- 3.4 Method for supporting viewer's judgment of data credibility: requirement (iii)

3.2 Endorsement data/graph that provide evidence for data credibility

In this section, we describe endorsement data and endorsement graph that are managed on the endorsement layer and used by viewers to satisfy requirement(i): Acquisition of evidence for the credibility of data in cyberspace.

[Endorsement Data]

Endorsement Data are data in which information relating to data on the Internet, an individual, organization, device, etc., is recorded in a form that cannot be forged using a digital signature. By confirming the sender and generation location, evaluation of the data by another person, etc., from the endorsement data, the viewer can judge the credibility of the data. In addition, identities, or functions of an individual, organization, device, etc. that issued endorsement data is expressed by an endorsement data so that viewers can confirm them.

In the use case mentioned in section 1.2, it is necessary for viewers to confirm information that can be evidence data credibility, such as who is the sender of the data of “flood is likely to occur.” and where the attached image was taken. It is also necessary to know whether the sender was a resident living near the river where the flood occurred, as information for evidence of data credibility.

Endorsement data treat the Information that is evidence of such credibility as *Property*. Property represents attributes or a feature of the data on the Internet or of an individual, organization, device, etc. For example, for the text data on the Internet such as “flood is likely to occur.”, the sender, generation date, time, etc. correspond to properties. In the case of an individual or organization, the name, trade name, address, and qualifications are properties, and in a case of a device, the manufacturer who made it is a property. These properties can be evidence for viewers to judge the credibility of the data. However, the data that are the subject of such judgment, such as the data of “flood is likely to occur.” and the attached image, are called *target data*. Individuals, organizations, devices, etc. that generate the data’s properties and issue endorsement data to the target data are called *endorser*. In the above use case, the endorser is the sender. The endorser who confirmed the property of the address of the sender would be the government or municipality. The target data and endorser are collectively referred to as an *object*. An object is an entity to which endorsement data are issued.

The Trusted Internet defines endorsement data as the following four items linked to the object.

1. Properties
2. Endorser
3. Endorsement data of the endorser
4. Digital signature created on the object's properties to indicate that the endorser confirms the object's properties as true.

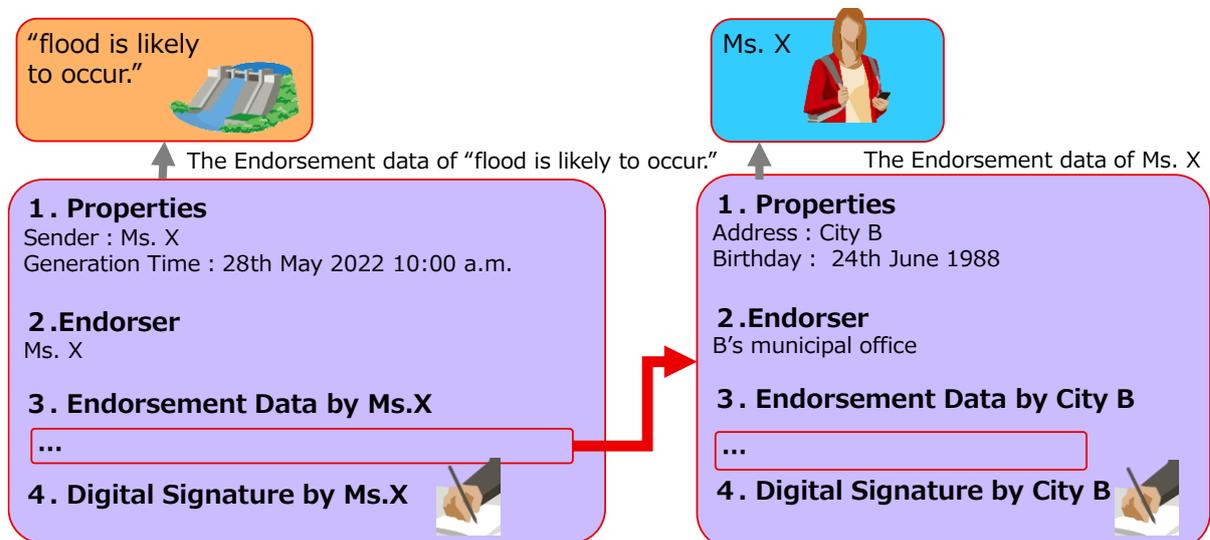


FIG 4. Example of Endorsement Data

The endorsement data of the endorser who is an individual, for example, the data that Ms. X is a resident of City B, could be the digital certificate that can be verified by a third party, as described in Section 3.1. We aim to make trust information in the real world available on the Internet by importing digital certificates issued from various systems to the Trustable Internet as endorsement data.

[Endorsement Graph]

An endorsement graph is a graph representing the relationship between the endorsement data and uncertain subjects and the issue relationship of the endorsement data. Representation as a graph enables viewers to trace the evidence for credibility in a chain fashion from the endorsement data of the uncertain object. We describe an endorsement graph below with specific examples.

FIG 5 is an example of an endorsement graph created for a post (= target data) that insists a “flood is likely to occur.” Viewers can conclude that a post is not worthy of belief because the posted image was taken in a location different from the location of sender’s residence and data posted by the Local Government are contrary to the content of the post.

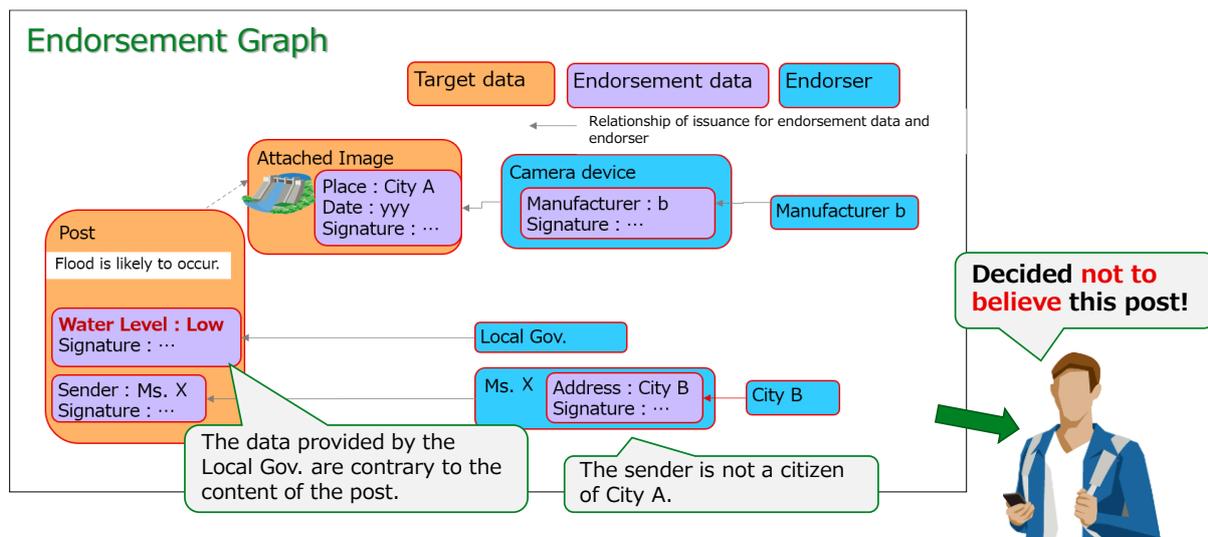


FIG 5. An Example of Endorsement Graph

In this manner, the endorsement layer manages endorsement data and endorsement graph and makes them available to the viewer, so that the evidence for judging data credibility can be provided to the viewer.

3.3 Endorsement graph reflecting the physical space

In this section, the process of providing and using evidence for judging credibility so that it can be confirmed that the data in cyber space and facts match on the basis of the fact information acquired in the physical space will be explained in the order of sensing the situation in the physical space, adding to the endorsement graph, and matching with the data in cyber space on the basis of the use case.

[Sensing conditions in physical space]

It is possible to sense or recognize the state and situation of the physical spaces by acquiring various fact information in the physical spaces such as water-level information and location information through the Internet.

Sensors that acquire various fact information from the physical space are equipped with Internet of Things (IoT) functions, connected to the Internet through edges (e.g., IoT gateways). Fact information are obtained at the edge in accordance with data models that can be used on the Internet. Such sensor device or equipment is normally manufactured in accordance with proper procedures and installed and maintained in accordance with procedures set by

the installer and maintenance personnel. The manufacturer information and maintenance information of the devices can be provided at an edge. It becomes possible to confirm that information indicating that these processes have been carried out is imparted by browsing the endorsement data.

In this manner, since it can be confirmed that the sensor device is manufactured, installed, and maintained correctly, the information acquired from sensor devices can be used as evidence for judging data credibility.

By using location information of a physical space from a system (network equipment, etc.), constituting the Internet, it is possible to obtain endorsement data such as where the data were generated and how the data were traced, thus enhancing the credibility of the data. By using this type of endorsement data, one can see if it is truly about an event that happened in a location.

[Adding to endorsement graph]

As described above, the information obtained from the physical space through sensors can be used as endorsement data. Physical-space information can be added to an endorsement graph at any time in the form of post-verification, as well as during data generation. For example, if the sensor is a water level sensor installed in a river, it can be added as evidence for indicating the state of the river. In the endorsement layer, endorsement data are added to the information from the sensor (the water-level from the water-level sensor) as evidence of the credibility of the data to be judged (in this example, SNS data about the flood).

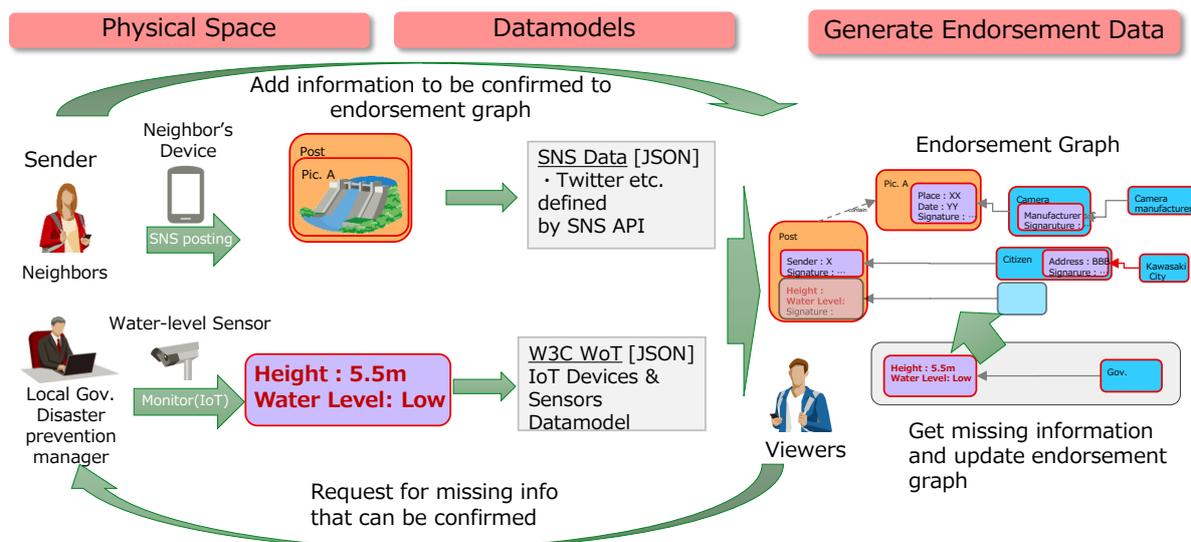


FIG 6. Endorsement Graph reflecting physical space

[Matching data in cyberspace]

To confirm the credibility of data, a technique using big data analysis is often used to analyze the original data then analyze the elements and the associated large amount of data as big data. For example, various related data such as whether there is a post like the original data or whether there is a disaster that occurred in the past at the same location or a location with similar topography are collected from the Internet, and the credibility of the original data is estimated by machine learning.

However, such big data analysis does not increase the credibility of the data, such as the data on the occurrence of a flood or the images attached to it, which are target data; thus,

there are cases in which inconsistencies with reality occur. By using endorsement data obtained from the physical space in addition to cyber space, it becomes clear that there is a discrepancy between the data and facts, for example, that the location where the data was sent and the location where the water-level was obtained are different.

By imparting endorsement data to various data on the Internet, various endorsement graphs are accumulated in the endorsement layer. Through multi-faceted evaluation of these endorsement graphs, it becomes possible to detect coincidence, inconsistency, and inconsistency of information, and to judge the credibility of data to which no endorsement data are attached. This can further increase the importance and necessity of the endorsement layer.

3.4 Method for supporting viewer's judgment of data credibility

On the Trusted Internet, when viewers search the information for a river, they can acquire data such as water level, flood information, and information on endorser which is evidence for the certainty of such information. This means that viewers can obtain information from both the Internet and endorsement graph.

Viewers can judge data credibility by using endorsement data. When too little information can make judgement difficult, viewers can request to add further endorsement data to the target data. If endorsers can add endorsement data to the requested data, viewers will consider that these endorsers provide valuable data (FIG 7).

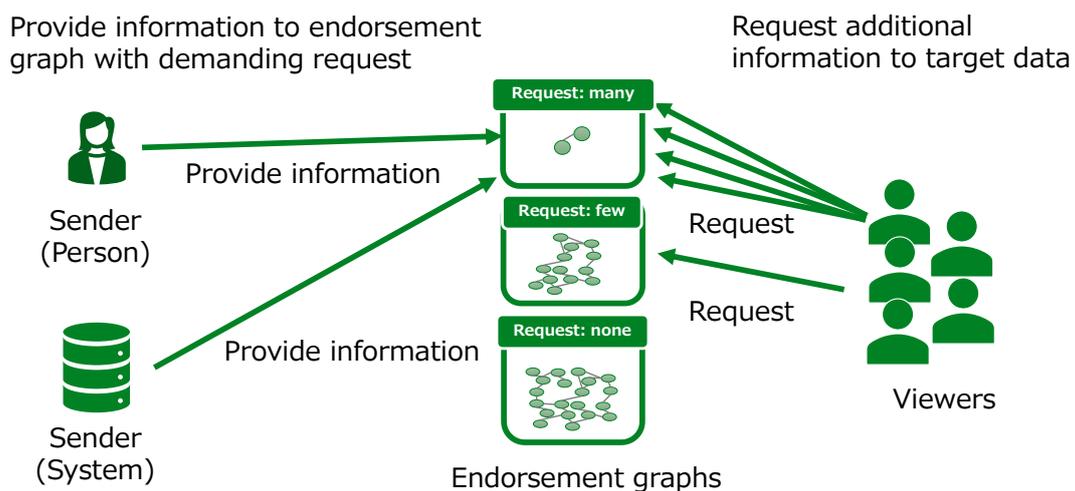


FIG 7. Request/Add information

However, when too much endorsement data or complex relationships can also make judgement difficult, support mechanisms in application and browser extensions could help viewers make judgements. For example, an application that shows viewer only the data according to the predefined criteria can be considered. These criteria include, for example, SNS posting rules and guidelines, and data-usage policies of companies (FIG 8).

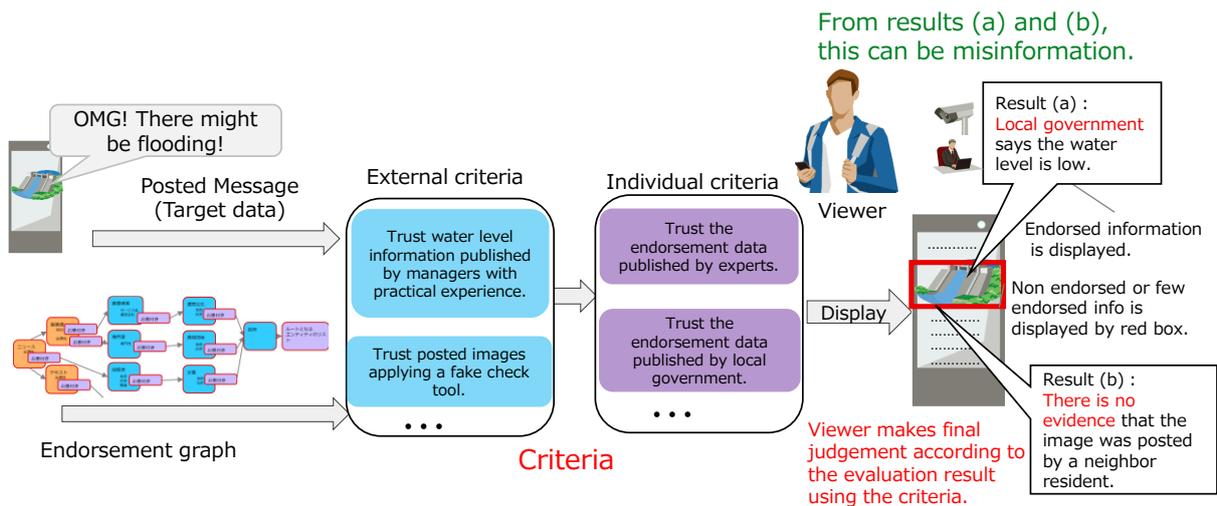


FIG 8. Mechanism to help viewers make judgements

Figure 8 shows an example of analyzing an endorsement graph using external and individual criteria. The external criteria define conditions associated with the data-generation process, and the individual criteria are based on viewer's principles or preferences such as blacklist or whitelist. The result of these criteria can be used to support a viewer by displaying auxiliary information of data or displaying a warning. With this type of support, requesting further endorsement data and adding data to the request are possible.

3.5 Overall behavior of the Trustable Internet

We describe a series of steps to solve the problem of the diffusion of uncertain information in a use case by using an endorsement graph generated from data in the physical space or data from a web app. (FIG 9)

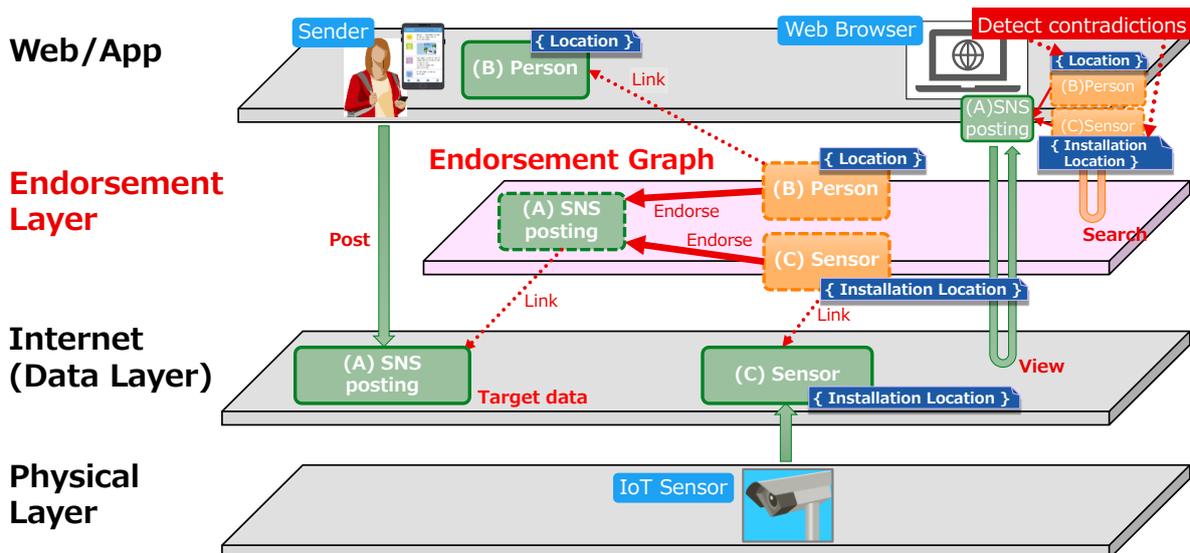


FIG 9. Overall behavior of the Trustable Internet

- (1) A resident post a message to SNS using a smartphone that “there might be flooding”. The posted message is digitized and can be viewed on the Internet (A). The endorsement layer generates a graph of the sender’s ID as evidence for judging the credibility of the SNS post (B). As mentioned above, whether a disaster, such as a flood, has occurred is based on the transmission of information by individuals with expertise or those on site (for example, weather forecasters, specialists in civil engineering and disaster countermeasures, and local residents). In this case, this information is represented by the attributes of the person that are digitized in (B).
- (2) To sense the state of the physical space, information from water level sensors installed in the river is used as evidence to indicate the state of the river. The information from the sensor is added to the endorsement graph as endorsement data to use it as evidence for judging the credibility of the original SNS post (C). The endorsement data can be added later, as well as when the original SNS was posted.
- (3) When viewers search the Internet for the state of the river, they can obtain information from both the Internet and endorsement graph, and simultaneously view the information about the river as well as the endorsement data (A)(B)(C), which is evidence of the state of the river. For example, when a large amount of endorsement data, such as a person and sensor, are applied to the target data, if there is a contradiction between the endorsement data, the credibility of the target data can be judged as lacking in credibility. In this use case, viewers can see that the sender’s location of the SNS post and location of the sensor are different, so they can conclude that they misunderstood a flood was likely to occur near them and prevent spreading the post further. The credibility of the state of the river is judged by the person on the basis of the data (A) and the underlying endorsement data (B) and (C).

As described above, it is possible to link the data on the Internet with the physical space, and judge the credibility of the data by viewers while confirming objective evidence.

4 FUTURE WORK

The Trusted Internet Architecture Lab would work on the following activities in both technology development and ecosystem to realize the Trustable Internet:

(1) Develop protocols for dealing with endorsement graph

the Internet and Web protocols to be used in general for handling endorsement data and endorsement graph should be established.

- Establishment of protocols for decentralized environments
- Building a multidimensional endorsement graph for multiple viewpoints of Viewers

(2) Activities to build an ecosystem

An architecture including an ecosystem should be considered.

- Viewers can benefit just by accessing the information because they can judge the credibility of the data. However, it is also necessary to consider the motivation and benefits of those who add endorsement data.
- To solve this problem, for example, a reputation mechanism is introduced can give an incentive, such as a high evaluation score to the provider who provides information that enhances the certainty of the data. Alternatively, a direct reply such as a token or NFT (non-fungible token), may be given.
- Technology could be used to identify disinformation and fake news. To incorporate various types of information, it is necessary to consider what can be described in endorsement data, for example, a negative property or a stochastic property.

- It is also necessary to consider privacy and safety aspects to prevent the unexpected use and spread of endorsement data.
- It will also require mechanisms to enhance UX (user experience) so that both providers and viewers can easily and repeatedly use the Trustable Internet.

(3) Expansion of use cases

Other use cases should be considered. For example,

- Use of endorsement graph for disinformation and fake news.
- Use of endorsement graph in combination with existing communication tools and information collection mechanisms.
- Improving data credibility in cross-border distribution.

5. REFERENCES

- [1] European Commission, "European Digital Identity Architecture and Reference Framework – Outline | Shaping Europe's digital future," [online]. Available — <https://digital-strategy.ec.europa.eu/en/library/european-digital-identity-architecture-and-reference-framework-outline>.
- [2] The World Wide Web Consortium (W3C), "Decentralized Identifiers (DIDs) v 1.0," [online]. Available — <https://www.w3.org/TR/did-core/>.
- [3] World Wide Web Consortium (W3C), "Verifiable Credentials Data Model v 1.1," [online]. Available — <https://www.w3.org/TR/2022/REC-vc-data-model-20220303/>.
- [4] The Linux Foundation, "Hyperledger Indy – Hyperledger Foundation," [online]. Available — <https://www.hyperledger.org/use/hyperledger-indy>.
- [5] Microsoft Corporation, "Distributed Identity (Identity) Solutions | Microsoft Security," [online]. Available — <https://www.microsoft.com/ja-jp/security/business/identity-access-management/decentralized-identity-solution>.
- [6] TBD, "Web5 | TBD," [online]. Available — <https://developer.tbd.website/projects/web5/>.