Delivering Personalized Information to Individuals in Super Smart Society

Kentaro Noda[1(✉)], Yoshihiro Wada[1], Sachio Saiki[1], Masahide Nakamura[1], and Kiyoshi Yasuda[2]

[1] Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada, Kobe, Japan
{noda,wada}@ws.cs.kobe-u.ac.jp, sachio@carp.kobe-u.ac.jp, masa-n@cs.kobe-u.ac.jp

[2] Chiba Rosai Hospital, 2-16 Tatsumidai-higashi, Ichihara, Japan
fwkk5911@mb.infoweb.ne.jp

Abstract. In the emerging super smart society, the flood of large-scale and heterogeneous information makes the digital divide in information reception a more serious problem. In this paper, we present a service, called Tales of Familiar (ToF), which autonomously delivers personalized information to individual end users in the super smart society. In ToF, every user is associated with a *familiar*, which is an agent working as an exclusive partner of the user. ToF first generates tales, which are narratives delivered by the familiar, from various information sources, such as direct messages, Web information, sensors, and SNS. The generated tales are sifted for individual users based on personal preferences of the users. Finally, the familiar delivers the selected tales using voice, text or images in an appropriate timing. In this paper, we particularly study the concept, the overall architecture, and data schema of ToF.

1 Introduction

With the rapid advancement of IoT (Internet of Things) [6,7] and cloud computing [2] technologies, the Japanese government has announced the *realization of a super smart society* [4] in 2020 as a part of the 5^{th} Science and Technology Basic Plan. The super smart society is the next-generation society, where heterogeneous systems are inter-connected over cyber and physical spaces, to provide advance and sophisticated services for individual citizens. In the announcement of the Japanese government, the super smart society is expected as a future society, which uses ever more emerging technologies such as robots, AI, bigdata, IoT and novel network devices. Thus, the plan creates the greatest impact on past economy and society.

It can be imagined easily, in the super smart society, that there would be a great flood of large-scale and heterogeneous information, which is incomparable with the current society. Without proper treatment, *digital divide* would be more serious between people who can use the information well and ones who cannot.

V.G. Duffy (Ed.): DHM 2017, Part I, LNCS 10286, pp. 336–347, 2017.
DOI: 10.1007/978-3-319-58463-8_28

Thus, to survive the super smart society, technologies with which one can appropriately collect, analyze and utilize the information become more important.

Towards the oncoming super smart society where large-scale and heterogeneous information are produced, we especially focus on the following three problems with respect to the *information reception by human users*.

Problem P1 (Barrier of Active Access to Information)
Currently, when a user wants to get information, the user has to access the data source *actively by himself*. To find desired information, the user must perform information search with a PC or a smartphone. However, it is difficult for an elderly person or a physically challenged person to search the information, and even to use such high-tech devices with complicated operations.

Problem P2 (Complexity of Identifying Relevant Information)
The super smart society produces and delivers a massive amount of information. However, the actually useful information varies among users, time, contexts, and locations. With the increase of irrelevant information, it becomes more difficult for every user to identify relevant information.

Problem P3 (Discomfort for Mass-Oriented Expression)
Generally, information is expressed in a popularized style, so that all information receivers (including human users and machines) can recognize it easily. For this, the information sender does not consider how *individual* receivers feel with respect to the way of writing. Thus, there is a possibility that a receiver feels uncomfortable by its expression, even if the receiver reaches desired information.

To cope with Problems P1, P2, and P3, this paper proposes a new service, called *Tales of Familiar* (hereinafter, called *ToF* for short), which autonomously provides information exclusively relevant to a given user. Intuitively, a *tale* refers to a narrative (or story) generated from an information source, so that the original information is easily received by the user. A *familiar* originally means a fairy that belongs exclusively to a human master, or a very close friend. Thus, ToF is a service where an agent (as a familiar) always stays by a user and provides exclusively relevant information (as tales) for the user.

The proposed service ToF mainly consists of three parts. First, as a solution to Problem P3, we propose the *tale generation service*. This service produces tales automatically or manually from various information sources in the super smart society. The information sources include direct messages, Web pages, sensors, smart systems, and SNS. A tale is a narrative derived from an information source and is expressed in familiar's spoken language. The tale generation service translates the original information in the mass-oriented expression into the one with more friendly and personalized expression.

Second, to cope with Problem P2, we propose the *tale sifting service*. From massively generated tales, the tale sifting service sifts (i.e., selects) only relevant tales appropriate for the user at that time. The criteria of the sift is based on the preference and context of the user.

Finally, we develop the *familiar* to address Problem P1. A familiar is an agent that autonomously speaks selected tales to the user. It is implemented in a form

of a robot, a virtual agent, or an IoT-embedded stuffed doll. A familiar tells a selected tale to a user in an appropriate timing. Thus, relevant information is automatically delivered in a preferred expression. So the user does not have to search desired information actively any more.

In this paper, we especially present the concept of ToF, the architecture with essential components, and features of each component. More specifically, we first discuss definitions of the tale, the tale generation service, the tale sifting service, and the familiar. Then, we design the system architecture of ToF, and the data schema to realize the service. Finally, based on the data schema and a use case scenario, we explain the workflow of the proposed service.

2 Preliminaries

2.1 Super Smart Society

The super smart society [4] is defined by a society that is capable of providing the necessary goods and services to the people who need them at the required time and in just the right amount; a society that is able to respond precisely to a wide variety of social needs; a society in which all kinds of people can readily obtain high-quality services, can overcome differences of age, gender, region, and language, and can live vigorous and comfortable lives. The Japanese government aims for its realization by 2020. Emerging technologies such as IoT, AI, bigdata and Robot are key components to achieve the society.

We have been using various information in a cyber world, such as e-mail, Web pages, and SNS. In addition to that, with the emerging IoT and cloud computing technologies, now we can accumulate a massive amount of information in a real world, such as health information, car running data and operational status of devices. It is expected that collecting such heterogeneous data over cyber/physical worlds and using the data horizontally across various application fields will create new values. Such continuous activities build the foundation of the super smart society.

2.2 Expansion and Diversification of Information Sources

Due to the spread of the Internet, a massive amount of data are published by human beings every day, from various information sources, including e-mail, Web sites, and SNS. In the super smart society, more data are generated by machines including IoT, sensors and smart systems. Thus, information sources become much more diverse and heterogeneous. Such machine-generated data include activity measures obtained by a smart watch, operation logs of a smart appliance, and environment sensor readings [9] within a smart city [5]. These new information sources yield much more realistic and timely information. On the other hand, we must understand characteristics of each information source, and deal appropriately with larger amount of information more than ever.

As an effective means to handle a huge and wide variety of the information, *curation service* [8] (e.g. Digg [1]) has received a lot of attention in recent years.

The curation is an intellectual activity that collects information on the Internet from a certain viewpoint of a curator, combines these information to create a value, and shares the value within the community. Since the curation service provides well-organized information from an interesting perspective, consumers can obtain relevant information easily. However, the current curation service mainly deals with the Web information only. Also, the curation is conducted subjectively based on the service operator. In this regard, the curation service just summarizes information for the mass but does not personalize for the individuals.

2.3 Challenges in Receiving Massive and Diverse Information

Considering how an end user receives information within the super smart society, we have identified the following three challenges.

First, assuming that the super smart society is developed as an extension of the existing ICT [3] over the Internet, a user has to access the information source *proactively by themselves* to obtain necessary information. The way to accessing the information relies on digital devices, including PCs and smartphones. The use of the digital devices becomes too complex especially for information illiterates. Thus, the digital divide becomes more serious. This challenge corresponds to Problem P1 in Sect. 1.

Second, when a massive amount of diverse information are delivered, the amount of information that are irrelevant or inaccurate for the user increases accordingly. The more information a user can see, the more time the user spends for treating the information. Thus, it would be more difficult for the user to reach relevant information. This challenge corresponds to Problem P2 in Sect. 1.

Finally, the information on the Internet is generally described in a popularized expression, so that the information can be recognized by the mass receivers. The information producer does not necessarily consider how individual receivers feel for the expression. Hence, the receiver might feel uncomfortable by the mass-oriented expression. Assuming that the super smart society is the society for individuals, every user should be able to feel comfortable when he/she receives the information. This challenge corresponds to Problem P3 in Sect. 1.

Some people are able to obtain necessary data rapidly and accurately from a massive amount of information. However, the situation is not easy for information illiterates such as elderly people, physically challenged people, people who are not good at machines, and people who are very busy for work. Without proper treatment, the digital divide with respect to receiving information would be more further serious in the super smart society.

3 Proposed Service

In this study, we propose a new information delivery service, *Tales of Familiar* (ToF, for short) to cope with Problems P1, P2 and P3 indicated in Sect. 1.

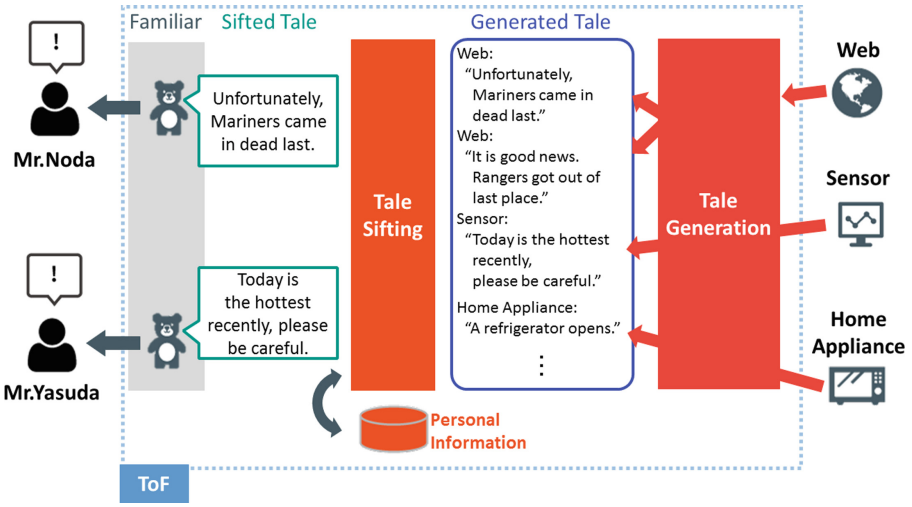

Fig. 1. Use case scenario

3.1 Concept

ToF is an information delivery service that collects information from the heterogeneous sources in the super smart society, and delivers the relevant information to individual users. The information is automatically delivered to a user in such a way that an agent called *familiar* speaks *tales* to the user. Intuitively, a tale refers to a narrative (or story) generated from an information source, so that the original information is easily received by the user. A familiar originally means a fairy that belongs exclusively to a human master, or a very close friend. Thus, the concept of ToF is that an agent (as a familiar) always stays by a user, and provides exclusively relevant information (as tales) for the user.

We explain a use case scenario of ToF using Fig. 1. First, ToF produces tales in the *tale generation* service. Each tale is created from information which collected from an information source such as a Web page, a sensor, or a smart appliance. ToF then sifts the tales in the *tale sifting* service. From massively generated tales, ToF selects only relevant tales appropriate for the user based on personal preference and actions of the user. Finally, a familiar associate to each user speaks the selected tales to the user.

ToF comprises the following key components: (C1) Tale, (C2) Tale Generation, (C3) Tale Sifting, (C4) Familiar. We describe these components in the following sections.

3.2 Tale

A tale refers to information expressed as a narrative (or a story) in familiar's spoken language. The original information source in a popularized style is converted into a tale in a more friendly and personalized style, so that a user can

easily receive and understand the information. We suppose that every tale is created automatically or manually using some appropriate methods, from various information sources in super smart society. For example, let us consider a tale that "Mariners came in dead last!" We suppose that this tale has been created from a sports news Web site for baseball fans of Seattle Mariners.

Each tale has meta-information including an information source, a provider, a category, and a delivery method. In the above example, the source is "the sports news Web site", the provider is the author of the article (or application), the category is "sport/baseball", and the delivery method is "voice" or "text". The meta-information attached to each tale will be used in the tale sifting and the tale provision by the familiar.

We can see in Fig. 1 that there are two pools of tales, where one is a pool for generated tales, and another is a pool for sifted tales. The former pool contains a lot of tales that do not assume specific receivers, and the latter contains tales exclusively selected for individual users.

3.3 Tale Generation

In the tale generation, ToF translates the information into tales in familiar's spoken language. The original information is collected from various information sources in the super smart society. In Fig. 1, tales are generated from three information sources of Web, a sensor and, a home appliance. Moreover, two tales "Unfortunately, Mariners came in dead last!" and "It is good news! Rangers got out of last place!" are produced from the Web.

The tale "Today is the hottest recently, please be careful!" is generated using a temperature sensor of the smart home. This tale is generated from the current room temperature value and the average value of the past several days. On the other hand, the tale "A refrigerator opens" is generated from the status of a smart refrigerator, which alerts the user to save the energy.

Multiple tales can be generated from the same information source, since there are many ways to express the information as a tale. For example, let us consider a sports Web news that Seattle Mariners and Texas Rangers have switched positions. For this news, we can generate two tales "Unfortunately, Mariners came in dead last!" for Mariner's fans, and "It is good news. Rangers got out of last place!" for Ranger's fans. In this way, we can choose more personalized expressions for individual users with different preferences.

In the tale generation, the original information expressed in the popularized style is translated into a tale expressed in a more friendly and personalized style. Thus, we aim to solve Problem P3 in Sect. 1.

3.4 Tale Sifting

In the tale sifting, ToF sifts all the generated tales, and selects only tales relevant for a user at that time. The tale sifting is performed for every user based on *personal information* of the user. In this paper, the personal information involves

user's profile, age, sex, hobby, preference, taste, dislike, and so on. As mentioned before, every generated tale has a set of meta-data. For every tale, ToF evaluates whether or not the user is interested in the tale, by collating the meta-information and the personal information. Thus, all generated tales are sifted for each user.

In Fig. 1, we can see that the tales are sifted for Mr. Noda and Mr. Yasuda, according to their personal information. In this example, we assume that Mr. Noda is a Mariners fan. Thus, the tale related to Mariners is picked up, as the tale is characterized by the "sports/baseball" category by the meta-data. Also, we assume that Mr. Yasuda lives in a smart home. Thus, the tale generated from his smart home is selected. Note that this tale is relevant to Mr. Yasuda only, since the tale is generated from Yasuda's home. So, the tale is not selected for Mr. Noda.

In the tale sifting, a lot of generated tales are sifted so that individual users can focus on only valuable and interesting tales. Thus, we aim to solve Problem P2 in Sect. 1.

3.5 Familiar

The familiar delivers selected tales to the user. We suppose that a familiar is implemented in a form of a robot, a virtual agent, or an IoT-embedded stuffed doll. According to the concept, every familiar belongs exclusively to a user. There are various types of familiars, and different familiars have different capabilities for delivering the tales. For example, the stuffed doll as a familiar provides tales by voice, while the virtual agent as a familiar delivers tales by an avatar, text and images. Also, a familiar is deployed in various places within a house, such as entrance, living room and others.

A familiar monitors the sifted tales, and delivers them to the user by an appropriate trigger. The trigger of the tale delivery can be user's request, user's proximity detected by a sensor, arrival of the emergency tale, fixed time and so on. When there are multiple tales in a familiar, the familiar delivers a tale with higher priority first. We suppose that the priority value is determined by the urgency, the elapsed time from the tale generation, the degree of matching to user's interest.

If the user wants to know more details of the tale, the user can access to information sources, as long as the familiar has a capability. For example, if a tale was generated from a Web site, and the familiar type is a virtual agent, then the user can access to the Web page displayed on a screen of the virtual agent. Also, a familiar can receive a feedback of the tale from the user. The user's feedback is used to update the user personal information, to grasp user's preference more accurately. In Fig. 1, there are two familiars whose types are both stuffed doll. Each of the two familiars is respectively assigned to Mr. Noda and Mr. Yasuda. They deliver sifted tales relevant to them.

The familiar delivers sifted tales to the user autonomously, considering appropriate time, place, and situation. Therefore, the user can obtain relevant information passively, without operating PC or smartphone by themselves. Thus, we aim to solve Problem P1 in Sect. 1.

4 Service Design

4.1 System Architecture

Figure 2 shows the system architecture of ToF. In this architecture, first, *Writer* collects information from information sources, and generates tales. The generated tales are stored once in a global repository of tales, called *GTale*. Then, *Picker* picks up tales relevant for every user from tales stored in GTale. The sifted tales are stored in a local repository of tales, called *LTale*. There is also a case that the tale which Writer generates for a particular person is directly stored in LTale. Finally, *Familiar* delivers the tales stored in LTale to the user. In proposed architecture, ToF is composed of the following five components.

S1:GTale. Gtale is a repository that stores global tales. It is a data pool deployed in the cloud. Tales in Gtale are generated by the original information with mass-oriented expression, extracted from Web, API, sensors and so on. GTale stores tales just after the tale generation by Writer and before the tale sifting by Picker.

S2:LTale. LTale is a repository that stores local tale. It is a data pool independently associated with every user. LTales stores tales just after the tale sifting by Picker. The tales here are the ones sifted from GTale, or the ones directly imported from Writer.

S3:Writer. Writer is a module that generates tales from various information sources. Writer also puts the generated tales in GTale or LTale. There are various types of Writer for different information sources, such as Web, sensor, direct mail, and others. Tales dedicated for a particular user can be forwarded to LTale directly. Such tales are generated typically from private services such as direct message and smart home. Writer is responsible for (C2) *Tale Generation* in Sect. 3.

S4:Picker. Picker is a module that picks up tales relevant to a user from GTale, and stores the selected tales in LTale. For each tale, Picker considers in which context the tale should be delivered, and determines an appropriate familiar to deliver the tale. Picker is responsible for (C3) *Tale Sifting* in Sect. 3.

S5:Familiar. Familiar is an agent that delivers the tales in LTale to the user. It also receives feedback from the user. Familiar is responsible for (C4) *Familiar* in Sect. 3.

4.2 Data Schema

We here design data schema of ToF, in order to achieve efficient management of tales. Figure 3(1) shows the whole data schema represented by an ER diagram. This diagram follows a notation proposed in [10], where a square shows an entity, (+—∈) shows a parent-child relationship, (+—· · ·) shows a reference relationship, and (+—○+) shows a sub-type relationship. Figure 3(2) shows the data schema around GTale including such as GTale and GTag. Figure 3(3) shows

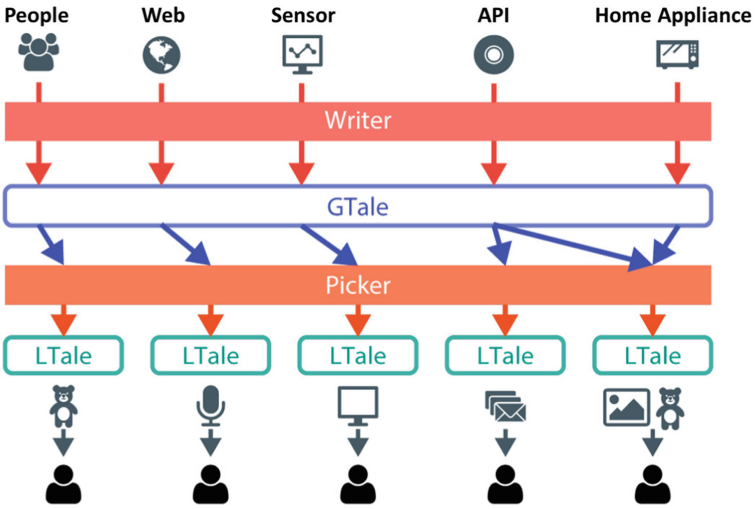

Fig. 2. Whole architecture

the data schema of master information including Provider, User, Content Type, Category, Favorite and Interest. Figure 3(4) shows the data schema around LTale including LTale, LTag, Delivery, Familiar and Capability. We describe the details of each table (i.e., entity) as follows:

GTale: This table stores tales in Gtale. GlobalID is a primary key to identify each tale. Body represents a body text of the tale. Source describes the information source of the tale. CreatedAt denotes date when the tale is created. ExpiredAt denotes an expiry date when the tale is nullified. For example, if the expired date is 2016-11-23 21:11:11, Picker does not pick up that tale after expired date. CategoryID is a reference to the category of the tale. Content-typeID is a reference to the content-type of the tale. When Writer generates a tale, Writer has to generate meta-data to insert the tale into the GTale table.

GTag: This table stores tags of the global tale. Each tag is one of meta-information of global tale. GrobalID and sequence are composite primary key. Tag describes an element of the tale. When Picker picks up tales, Picker will use GTag table.

Familiar: This table stores familiar information. FamiliarID is a primary key to identify each familiar. ProviderID means the owner of the familiar. Also, providerID refers to User table and assigns the user that familiar should provide tale for. Place describes the place of the familiar.

Capability: This table stores familiar's capability for delivering the tales. FamiliarID and sequence are composite primary key. Content-typeID is a reference to the content-type of the tale.

LTale: This table stores tales in LTale. FamiliarID and localID are composite primary key to identify each tale. FamiliarID is a reference to the familiar which provide the tale. Therefore, we manage local tale for each familiar.

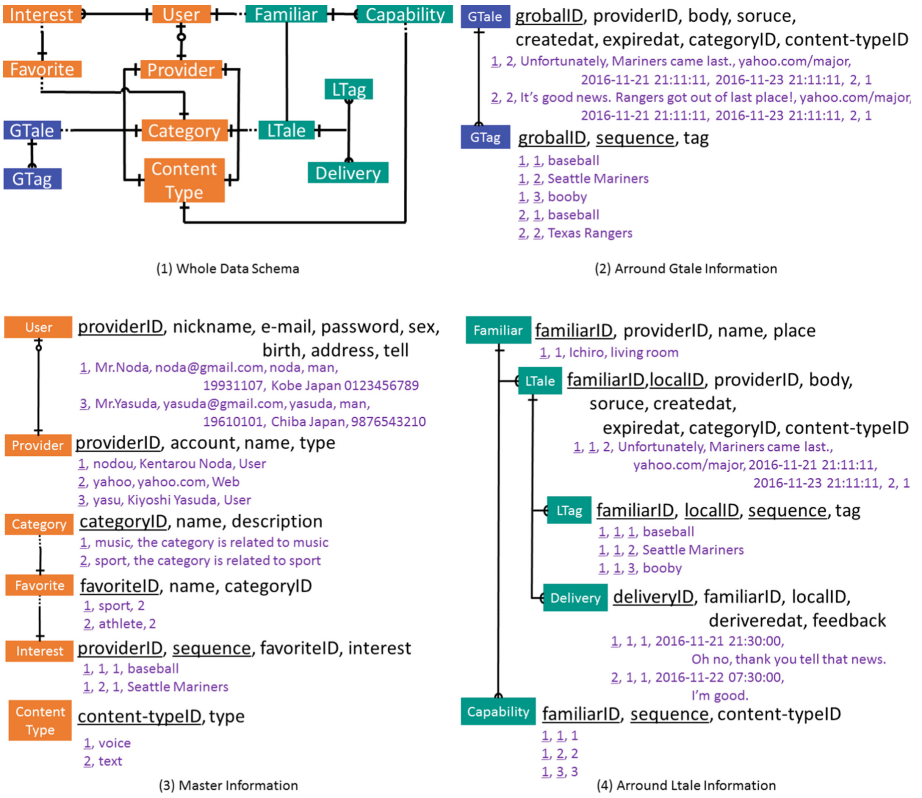

Fig. 3. Data schema

Body, source, categoryID and content-typeID are duplicated from GTale table. Familiar provides tales for the user using this table.

LTag: This table stores tags of the local tale. FamiliarID, localID and sequence are composite primary key. FamiliarID is a reference to the familiar which provide the tale. Tag is duplicated from GTag table.

Delivery: This table stores the log of tale provision. DeliveryID is a primary key. FamiliarID is a reference to tale provider. LocalID is a reference to provided local tale. DeliveredAt represents provided date Feedback describes the user reaction. The user's feedback is used to provide tale again, study the user response and update the user personal information. Familiar stores feedback of provided tale in this table.

Provider: This table stores provider information. ProviderID is a primary key to identify each provider.

User: This table stores user personal information. That include nickname which familiar calls, e-mail, password. For example, the user, providerID is 3 in Fig. 3(3), is called "Mr.Yasuda" from familiar. The user is one of the information provider. The instance which type is User in Provider table refers to User table also.

Category: This table stores tale category. Picker picks up tales using this table.
Content Type: This table stores method of delivering tale. Picker decides familiar which provide tale using this table.
Favorite: This table stores favorite thing. This table is used to register what user is interested in. For example, the phrase "sports" in "What your favorite sport?" is stored in this table. FavoriteID a primary key to identify each favorite thing. Also, a FavoriteID corresponds to a CategoryID and connects user interest thing with tale category.
Interest: This table stores the user interest thing. ProvideID and sequence are composite primary key. Also, this table has reference relationship for Favorite table. Picker picks up tales using this table.

4.3 The Entire Flow of the Service

We describe how to execute the use case scenario in Fig. 1, based on proposed data schema.

Assumption: Tale provider is "yahoo.com" on Web site. We assume the tale which Mr.Noda is interested in is produced (refer to Provider and User in Fig. 3(3)). Mr.Noda have the familiar called Ichiro (refer to Familiar in Fig. 3(4)).
STEP1: The information about the rank of American League in Major League Baseball is posted in "yahoo.com". Writer produces tales from that information and stores these tales in GTale. Each tale spotlights each baseball team. When a tale is produced, Writer produces tags of that tale also (refer to GTale and GTag in Fig. 3(2)).
STEP2: Picker gets personal information of Mr.Noda from personal information database. Picker compares user interest in personal information to tag and category of each tale in GTale and evaluates whether each tale is appropriate for the user. Thus, picker picks up the tale such as "Unfortunately, Mariners came last." which is appropriate for Mr.Noda.
STEP3: This tale should be provided by voice, because of the content type of tale (refer to ContentType in Fig. 3(3)). Picker checks Ichiro can manage that content type and stores the tale in LTale. At the same time, sequences and tags in GTale about stored tale are duplicated and stores in LTag (refer to LTale and LTag in Fig. 3(4)).
STEP4: Ichiro is placed on the living room (refer to Familiar in Fig. 3(4)). Ichiro finds Mr.Noda in the living room by the motion sensor and provides tale by voice. "Mr.Noda, unfortunately, Mariners came last". Mr.Noda replies "Oh no, thank you tell that news". Familiar stores that feedback and the log of tale provision in Delivery (refer to Delivery in Fig. 3(4)).

5 Conclusions

In this paper, toward Super Smart Society, we proposed the autonomous information delivery service for information illiterate. In the proposed service, a massive

amount of tales are produced from various sources of information. Produced tale is sifted and only the tale which is appropriate for the user is picked up. Sifted tale is provided for the user autonomously by familiar. Familiar is the IoT device that has sensor and microphone. As the future challenges, we have to develop a prototype of proposed service and test that service.

Acknowledgments. This research was partially supported by the Japan Ministry of Education, Science, Sports, and Culture [Grant-in-Aid for Scientific Research (B) (No. 16H02908, No. 15H02701, No. 26280115), Young Scientists (B) (No. 26730155), and Challenging Exploratory Research (15K12020)].

References

1. Digg. http://digg.com/
2. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M.: A view of cloud computing. Commun. ACM **53**(4), 50–58 (2010)
3. Dutton, W.H.: Information and Communication Technologies: Visions and Realities. Oxford University Press, Inc., Oxford (1996)
4. The Japanese Ministry of Education, Culture, S.S., Technology: The 5th science and technology basic plan (2016)
5. Hollands, R.G.: Will the real smart city please stand up? City: Anal. Urban Trends Cult. Theory Policy Action **12**(3), 303–320 (2008)
6. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gener. Comput. Syst. **29**(7), 1645–1660 (2013)
7. Atzori, L., Iera, A., Morabito, G.: The internet of things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010)
8. Rosenbaum, S.: Curation Nation: How to Win in a World Where Consumers are Creators. McGraw-Hill Education, New York (2011)
9. Sakakibara, S., Saiki, S., Nakamura, M., Matsumoto, S.: Indoor environment sensing service in smart city using autonomous sensor box. In: 15th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2016), Okayama, Japan, pp. 885–890, June 2016
10. Watanabe, K.: Hanbai Kanri System de Manabu Modeling Koza. Shoeisha, Tokyo (2008). (in Japanese)