A Study on Building of Sports Data Sharing Platform from the Perspective of the Personal Information Protection Act

Kai-Li Wang
Graduate Institute of International Sport Affairs, National Taiwan Sport University, Taoyuan, Taiwan
Email: elverz@ms8.hinet.net

Abstract—The evolution of wearable devices, big data and Internet of Things (IoT) has given rise to today’s emerging apps. Among them, sports- and health-related ones are drawing the most attention. Many giants have thus launched cloud-based platforms that combine hardware products with software services and even social networking functionality. This way, users of sensor-equipped wearable devices can monitor and track their sports data. They can further share the data for more fun with their sports activities. In this study, such platform is termed as “sports data sharing platform” (SDSP). However, the increasing concern over privacy rights has urged more governments to establish new laws on personal data protection worldwide. The move will affect the said developing service models. In this study, we will start from the perspective of the recently enacted Personal Information Protection Act (PIPA) of Taiwan R.O.C. to analyze several important SDSP-related legal issues. We will then discuss several models used to establish the SDSP, and their differences in legal risks. SDSP developers are advised to refer to the conclusion of this study for future improvements. As for the government, suggestions mentioned here may also serve to help handle legal practices about the PIPA. The purpose is to avoid making legal interpretations that might undermine thriving services.

Index Terms—wearable device (wearable technology), big data, Internet of Things (IoT), sports data, personal information

I. INTRODUCTION

Emerging tech-based services such as wearable devices, big data and Internet of Things (IoT) have been taking the technology world by storm in recent years. Among them, wearable devices (or wearable technology) refer to those devices that combine electronic sensors or micro-computers with traditional apparel or accessories. The purpose of this is to sense, track or compute the wearer’s physiological data and to monitor his/her fitness activities [1]. Widely known products include GPS-enabled watches, smart wristbands, the Google glass, as well as sensor-mounted smart apparel and rings. Data acquired through these wearable devices may be linked to the database on a wired or wireless basis through Wifi, Bluetooth or other protocols. This is done to compute the data and send feedback to the wearer in the way it is acquired. The feedback shows the wearer’s data about health status or activities, such as daily step counts and sleep quality. The wearer may examine the data to make improvements. These services are realized through the IoT or database that supports the big data storage and analytics [2] and [3].

As these technologies gain momentum, a variety of apps are also taking center stage. This is particularly evident in the fields of health and sport, where Nike, Inc. pioneered these developments. Examples of its innovations are sensor-equipped sports watches, and the Nike+ FuelBand wristband. The sports giant also built a network-based database that helps track the sports data, as well as dedicated apps under the name of “Nike +” (e.g. Nike + Running, and Nike + Plus Coach). Garmin, the leader in the GPS navigation industry, is making aggressive moves to shift its focus from car navigation to wearable sports technology. The case in point is Garmin Connect, its latest cloud-based platform where cyclists of the same online community can share recorded data about cycling routes. At the time this paper is being prepared, the world is awaiting Apple Watch, for which many expect the smartphone leader to roll out better sports and health management apps. Taken together, the services launched by these giants are architecturally a sports database. This database allows usable data to be stored in it for computation and analysis by its developer. Processed data can then be sent back to its users [4] and [5]. In this study, it is termed as sports data sharing platform (SDSP).

The extremely close relationship of the SDSP to big data stimulates wider interest in such risk issues as information safety and misuse of private information.

The earliest mention of the term can date back to 2001, when Doug Laney [6], the analyst of META Group (now Gartner, Inc.), indicated the so-called 3Vs of data, i.e. volume, velocity and variety, will be greatly challenged and taken advantage of in the future growth of data processing. However, those ideas advanced at that time were no more than a managerial concept. That being said, technological progress has driven the prevalence of sensors of various types, and such tech giants as Google and Amazon have been making great strides in the cloud computing in recent years. These reasons combined...
Today’s wide use of SNSs (Facebook, Twitter, etc.) has built unbelievably close connections between and among people. Just one decade ago, we had to take part in social meet-ups in order to know a friend of our friends. Now only a click on Facebook can create bonds between him/her, let alone the People You May Know function. By using the function, we get to know one of our elementary school classmates with whom we have lost contact for a long time. With everyone now owning at least a mobile device (e.g. smartphones, tablets), location-based services (LBS) have caught on significantly. As one LBS, Foursquare awards users using check-ins with digital prizes or digital badges that can be redeemed for online-to-offline (O2O) marketing of physical products. Another example is SCVNGR. Its QR code-enabled feature allows stores to use the QR code-based payment via the customers’ smartphones to share customer-specific information. Stores may utilize the information to offer exactly the product or service that customers want in a more precise, instant and local manner [8].

While our tendency of sharing online where we go, whom we hang out with or what we like can be an issue of concern, it is far more intimidating that the service providers mentioned above (i.e. Nike, Garmin or even Apple) can obtain and further distribute our physiological or sports data via the SDSP. In this regard, the Federal Trade Commission (FTC) made the earliest move in March 2010 by issuing the report entitled ‘Protecting Consumer Privacy in an Era of Rapid Changes.’ The FTC called on the industry to give the “Do not Track” option to consumers, who could then have the ‘opt-out’ right.

Similarly, United States Department of Commerce (DOC) presented a report ‘Commercial Data Privacy and Innovation in the Internet Economy: A Dynamic Policy Framework’ in December 2010. The DOC suggested that the U.S. government introduce comprehensive legislation to tackle the privacy issues. This ended up with U.S. President Barack Obama declaring the Consumer Privacy Bill of Rights (CPBR) in February 2012. The CPBR holds personal data holders, analyzers, or commercial users liable for 1. individual control, 2. transparency, 3. respect for context, 4. security, 5. access and accuracy, 6. focused collection, and 7. accountability in favor of consumers (i.e. the said personal data holders). In the wake of the Ed Snowden case in 2013, Obama further declared a technological report on the right to privacy in May 2014 in a bid to eliminate public concern over the use of technology. Entitled ‘Big Data and Privacy: A Technological Perspective,’ the report stated five suggestions that the U.S. federal governments should take for policies governing big data and the right to privacy.

In Taiwan, the Personal Information Protection Act (PIPA) was enacted in 2012. The PIPA was based on the laws of EU, Germany, Austria and other countries, under which medical and physical information is defined as specific, or confidential, information that may not be collected without consent. As the SDSP uses the product users’ physiological and sports data directly or indirectly, and actively or passively, one may wonder whether this business model is in violation of the PIPA. If this is the case, what measures should managers take against legal risks? What are the issues to be confronted by the government and the court? To address these questions, we will discuss in Section II how the topic of this study relates to the PIPA. The author will then set out some feasible sports data sharing models in Section III, where the legal measures mentioned in the previous section will be reviewed. Lastly, Section IV will deal with managerial suggestions for SDSP developers.

II. PIPA-RELATED ISSUES FOR SDSP USERS AND DEVELOPERS

Given Taiwan’s developed information industry, both the government and the private sector recognize the importance of personal information that needs full protection in this digital era. In 2010, Taiwan’s lawmakers amended the Computer-Processed Personal Data Protection Act, and renamed it as Personal Information Protection Act (PIPA). The amendment reflects many contemporary issues of grave concern [9]. As such, this study chose the PIPA as the material to spark discussions over issues that may be confronted not only by the island country but also by the whole world.

Regarding the 2010 amendment to the act, it is very interesting to note Article 6 [10], which is as follows:

“Personal information of medical treatment, genetic information, sexual life, health examination and criminal record should not be collected, processed or used. However, the following situations are not subject to the limits set in the preceding sentence:

1. When in accordance with law;
2. When it is necessary for the government agency to perform its duties or for the non-government agency to fulfill the legal obligation, and when there are proper security measures.
3. When the Party has disclosed such information by himself, or when the information concerned has been publicized legally;
4. When the personal information is collected, processed or used under certain methods by a government agency or an academic research institution based on the purpose of medical treatment, personal hygiene or crime prevention statistics and/or study.”

Personal information of medical treatment, genetic information, sexual life, health examination and criminal record may not be collected simply by obtaining the party’s written consent as is allowed under Articles 15
and 19. Therefore, this limitation may deal a blow to how the SDSP—the topic this study handles—can be built.

However, a close look at this article reveals two problems:

- Does the so-called “sports data” fall under the personal (or confidential) information of medical treatment or health examination as defined under the article?

What exactly is sports data? The answer lies in the types of sports. Examples may include offense/defense stats for basketball and baseball players; routes and time for runners and cyclists; as well as data about muscle strength and cardiopulmonary endurance for a body builder. While the first two examples are about non-medical and non-physical objective facts, the third one has conceptually more to do with medical or physical checks.

The following italicized passage is excerpted from the Enforcement Rules of the Personal Information Protection Act, promulgated by the same Taiwanese governing authority that enacted the PIPA.

“Personal information of medical treatment shall mean medical records and other check-ups or treatments made by doctors or other medical personnel for the purpose of treating, correcting or preventing the diseases, harms or disabilities of human body or for other medical due reasons, or shall mean other personal information produced through prescription, medication, operation or disposition based on the above results of check-ups. Personal information of health examination shall mean the information produced by medical examination not for the purpose of diagnosing or treating a specific disease.”

One may reason that, based on the definition above, sports data is not personal information of medical treatment. But can it be produced through health examinations? This depends on how to define the term “medical practice.” In Taiwan, the Ministry of Health and Welfare and the Sports Administration, Ministry of Education have long held opposite opinions about the difference between medical and healthcare practices. In sum, the controversy—whether sports data is personal information generated through check-ups—is to be clarified in real-world practices.

In this study, we assume that sports data is about physical performance rather than confidential information as defined under the article above. This is largely because the sports data referred to here is used to measure a body builder’s physical performance, such as how much weight he/she can lift and how many sit-ups he/she can do. As the measurement of heartbeat, height and weight is physiological, the sports data obtained so is more likely to be confidential.

- What does it mean by “when the Party has disclosed such information by himself, or when the information concerned has been publicized legally?”

If the party publicizes his/her own personal data by himself/herself or the information has been publicized legally, such disclosure is not considered privacy breach. Take, for instance, a post about his/her new weight lift record on an SNS website, or a professional player’s height/weight covered by mass media. It is therefore allowed to collect, process and use this type of data. However, we may have to look at one key issue: how to define this kind of spontaneous disclosure. This can be addressed from two perspectives. One is the “definition of publicizing/disclosure;” the other is its “relation with data collection by SNS providers.”

According to Article 13 of the Enforcement Rules of the Personal Information Protection Act, publicizing/disclosure is defined as follows:

“The personal information disclosed by the Party himself/herself shall mean the personal information disclosed by the Party himself/herself to the general public or specific persons.”

The general public referred to here can be anyone that has access to the information (e.g. learning from television). But here we would want to take a closer look at “who are specific persons” from the practical perspective. Can a five-member online community be one of them? How about a community joined by more than 500 users? What if a cyclist shares his cycling route on his social media page with his 500 online friends? As such, the governing authority is advised to provide further explanations on this still-controversial issue.

Let us now turn to the second issue: how the user’s spontaneous disclosure relates to data collection by SNS providers. The first question to be examined is: “is there a difference or even a conflict between user-publicized data and data collected by SNS providers?” To explain this, we can follow the definition under Article 2 of the PIPA “Collection: To collect personal information in any form and way” to look at the example above. If this is the case, the SNS provider is arguably collecting the user’s personal information when he/she shares his/her cycling route on the SNS website.

Given the explanation above, we can conclude that it is legal for an SNS provider to collect the user-publicized information from a public domain designed by the provider itself. This meets the exclusion clause we discussed earlier about the collection, processing and use of personal confidential information.

To avoid controversy, however, SNS providers are supposed to properly warn their users in advance about the risk of not being protected by the PIPA before they upload their personal data onto the SNS website.

To sum up, we still cannot clearly define the legal scope of sports data. The closer it is to confidential information, the more the development of SDSPs will be hampered. In this regard, we have to consider legal measures against risks, which will be discussed in Section III.

III. SDSP OPERATING MODELS THAT RESPECT THE PIPA

Before our discussion over SDSP operating models from the PIPA perspectives, we have to first examine how the word “operating” and its related legal terms (i.e. collection, processing and use) can be defined. As the previous section has dealt with the “collection,” here we
need to go over the other two by citing the following passages from Article 2 of the PIPA.

“Processing: To record, input, store, compile, correct, duplicate, retrieve, delete, output, connect or internally transmit information for the purpose of establishing or using a personal information file;  
Use: All methods of personal information use other than processing;”

Fig. 1 shows a common service platform where Users A, B and C can upload their sports data. The service provider analyze and sum up all uploaded data to derive a sum, which is sent back to the users along with their personal data. When necessary, the service provider may give advice on sports or health management to them based on the summarized data. An example of this is how the Nike+ Running analyzes a runner’s speed and offer the information about his/her ranking and percentile ranking among all users of the app. A further example is the Nike + Coach, which gives sports advice.

The app’s in-system computation is actually the legal term we are now discussing: processing (e.g. editing and input) and use (e.g. using algorithm). In the example above, Users A, B and C acquire a sum and their own data, which does not violate their privacy rights. In this study, this kind of platform is termed as “independent feedback sharing platform”

![Image 1](https://via.placeholder.com/150)

**Figure 1.** Independent feedback sharing platform.

The major difference between the two platforms lies in whether the data with specific personal traits can be accessed, as shown in Fig. 2. This model is called “relevance feedback sharing platform” in this study. Take, for instance, the Garmin Connect. The app gives users of the same community the access to every member’s scores (speed, etc.), and allows them to exchange information about cycling routes as a virtual benchmark. This is how the relevance feedback sharing platform works.

The two models operate under the condition that the sports data is not confidential information. When one user considers his/her heart rate or other data confidential, these models would not help. But we can imagine how eager a user is to find the data of an athlete with comparable performance through the big data and use it as the virtual benchmark to improve his/her game or training scores. His/her goal can be more achievable with more careful consideration taken into as well as more accurate physical and fitness stats.

To address these needs, we will examine several solutions. Nevertheless, none of them is flawless.

- Replacing physiological data with athletic scores
- The service provider can replace physiological data, which bears closer relation to check-ups, by athletic scores (speed, time, etc.), which is considered overt. Doing so, however, may cause big data algorithms to be less accurate.
- Defining user’s uploading as spontaneous disclosure
- This approach goes against the protection of personal data. The service may also be of less exclusivity, thereby taking a toll on how the business model operates.
- Having the data collected, processed and used by hospitals or other legally responsible organizations Hospitals may therefore be less concentrated on their core duties. We should also think about whether a hospital is legally obligated to help find a sports partner.
- Separating physiological data from athletic scores for data computation

![](https://via.placeholder.com/150)

**Figure 2.** Relevance feedback sharing platform.

Fig. 2 illustrates another model. Similarly, this model also computes the user-uploaded data. The only difference lies in what is sent back to the app users. In fact, the feedback here includes not only the sum and the user’s own data but also the athletic performance stats of others. This means the data with specific personal traits can be accessed, as shown in Fig. 2. This model is called “relevance feedback sharing platform” in this study. Take, for instance, the Garmin Connect. The app gives users of the same community the access to every member’s scores (speed, etc.), and allows them to exchange information about cycling routes as a virtual benchmark. This is how the relevance feedback sharing platform works.

The major difference between the two platforms lies in whether the data with specific personal traits can be accessed, as shown in Fig. 2. This model is called “relevance feedback sharing platform” in this study. Take, for instance, the Garmin Connect. The app gives users of the same community the access to every member’s scores (speed, etc.), and allows them to exchange information about cycling routes as a virtual benchmark. This is how the relevance feedback sharing platform works.

The major difference between the two platforms lies in whether the data with specific personal traits can be accessed, as shown in Fig. 2. This model is called “relevance feedback sharing platform” in this study. Take, for instance, the Garmin Connect. The app gives users of the same community the access to every member’s scores (speed, etc.), and allows them to exchange information about cycling routes as a virtual benchmark. This is how the relevance feedback sharing platform works.

The major difference between the two platforms lies in whether the data with specific personal traits can be accessed, as shown in Fig. 2. This model is called “relevance feedback sharing platform” in this study. Take, for instance, the Garmin Connect. The app gives users of the same community the access to every member’s scores (speed, etc.), and allows them to exchange information about cycling routes as a virtual benchmark. This is how the relevance feedback sharing platform works.
in order to facilitate the growth of emerging services. Platform developers are also supposed to pay more attention to accompanying legal risks.

- SDSPs can be divided into “independent feedback sharing platform” and “relevance feedback sharing platform,” based on how the feedback is shared. While the former delivers the sum and the user’s own data, the latter may include other users’ specific sports data. Developers of relevance feedback sharing platforms must pay necessary attention to the content of service contracts between and among users. The purpose of this is to avoid violating the PIPA.

- To avoid the legal risk of privacy breach while handling sports data, we have examined four possible solutions in Section 3. Although none of them is flawless, service providers are advised to consider these approaches. We also call on governments all over the world to address this issue from the practical perspective by making amendments or legal interpretations accordingly.

REFERENCES


K. L. Wang works as an assistant professor in Graduate Institute of International Sport Affairs, National Taiwan Sport University. His area of interest is the law issue and public policy in leisure, entertainment and sport field. Recently, he did many researches about the image right, patent, public policy and legislation regarding the sport industry based on the approach of economics of law.