Research on the Right Identification and Right Description Standard of Service Data Based On ODRL

Jing Su¹*, Lu Jiang¹, Yaqing Si¹, Guangkai Li², and Qingjun Xiao¹

¹School of Economics and Management, BUPT, 100876, Haidian District, Beijing, China
²Electric Power Research Institute, CSG, 510663, Huangpu District, Guangzhou, China

Abstract. This article mainly solves the problem of standardization of ownership and rights description of the main stakeholders in the field of data circulation in the field of data governance: combining the status quo of data protection and ownership identification, starting with the logical relationship between data and information, deconstructs the service data for subject ownership identification; combs the relevant laws and regulations of data protection, and defines the rights of each subject. In order to realize the structure and machine-readable of the right language, we should refer to the ODRL information model of W3C to build the right description standard of service data, describe the application of the standard in the actual transaction scenario, and build the hierarchical right transfer model in the process of data transaction.

1 Research background

Data has become the new oil in the era of digital economy [1], and there is infinite value behind the cleaned data. The development and utilization of big data has become an inevitable trend. Both the government and enterprises take the rational use of data assets as an important part of strategic planning.

Service data is a kind of data type closely related to economy and personal life. It is produced in all service processes, generated and gradually accumulated under the joint action of consumers and service providers, and controlled by service providers in practice, which has become its unique business competitive advantage. Because of its multi-agent, multi-dimensional, dynamic uncertainty and other characteristics, service data has not yet been defined clearly and generally, lack of legal framework to regulate, especially lack of the definition of the scope of the rights of stakeholders, coupled with the complex composition of service data, involving the identification of personal information rights, making the use and circulation of data a very difficult legal problem. Although personal data transactions have been legalized in the United States, Japan and other countries [2], up to now, there has not been a unified standard for data confirmation, and the flow of data compliance has been hindered.

Data will create more value only if it is circulated, and data innovation is the key factor for market growth and job creation. How to find a balance between the protection of personal data and the free flow of data, better and more secure mining the value behind big data, has
become a problem that countries must solve in the next stage, and the important entry point to solve this problem is the research on the right identification and right description standard of service data.

2 Service data deconstruction and right definition

In a broad sense, the service data resource refers to the data set with a certain scale that records the network service process and is collected and generated by the service provider. The service provider actually controls the definition, formation, collection, storage, disposition, transmission and use of data, so that it can bring commercial value. The most common service data are order data and comment information of e-commerce platform, bill data of communication operators, bank deposit and withdrawal information, etc.

The service data resources show the complex characteristics of multi-agent and multi empowerment, and the identification of its ownership needs to be clarified from the legal level. This study proposes solutions to this problem from the perspectives of law, technology and business.

2.1 Deconstruction of service data and complexity of ownership determination

GDPR (General Regulations for Data Protection) is the integrator of data protection in the EU in recent years, and is also an important reference widely used in the industry. This paper also focuses on the division of the right subjects by GDPR, and defines the main stakeholders in the data circulation link (see Table 1):

<table>
<thead>
<tr>
<th>Right subjects</th>
<th>Describe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data subject</td>
<td>Identified or identifiable natural person pointed by data</td>
</tr>
<tr>
<td>Data controller</td>
<td>A natural person, legal person, or organization that actually owns the data and can determine the purpose and method of personal data processing</td>
</tr>
<tr>
<td>Data processor</td>
<td>The natural person, legal person, or organization entrusted to the controller for processing data</td>
</tr>
<tr>
<td>Data trading platform</td>
<td>A platform to provide relevant services for data and derivatives transactions</td>
</tr>
</tbody>
</table>

After sorting out the service data with wide traffic (as shown in Figure 1), this study deconstructs the single service data:

- Data of data subject: all kinds of data recorded electronically, independently or in combination with other data, which can identify natural person (user) or reflect the activity information of natural person (user).
- Data of service provider: business behavior data left by the service provider in the process of providing services or data containing its intelligence and labor achievements, such as the recommendation information or customer service information provided by the server for the user.
● Data of service object: service object is the purpose for users to use the network service. It can be a specific commodity or a service. It is usually used to describe the commodity or service, such as product specifications, product prices. Although this part of the data is usually public, it still needs to fully consider the wishes of the brand.

● Service process data: data generated by the initiation of service process, used by the service provider to record the service process.

● Data of other subject: data not belonging to the data subject or the service provider, but related to the subject but actually belonging to other data subjects. For example, a group photo uploaded by a user on a social platform, a comment posted under a friend status, etc.

From the structure of service data, it can be seen that it not only contains personal data involving personal sensitive information, but also contains enterprise data related to business information and intellectual achievements of service providers. The special composition of data makes the service data ownership complex, multi-dimensional and multi-agent.

2.2 Definition of service data rights

After clarifying the ownership relationship of each part of the service data, it is necessary to clarify the right boundary owned by each subject. On the basis of extensive literature research, the rights of each right subject of service data are summarized as follows:

Table 2. Rights of each subject.

<table>
<thead>
<tr>
<th>Right subjects</th>
<th>Rights of the subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data subject</td>
<td>The right to know, the right to access, the right to correct, the right to be forgotten, the right to limit processing, the right to carry data, the right to refuse, the right to decide independently, the right to judicial relief, the right to cancel, the right to withdraw consent, the right to restore</td>
</tr>
<tr>
<td>Data controller</td>
<td><strong>Control:</strong> Sharing right, donation right, publication right and marking right <strong>Right to use:</strong> Reproduction right, merger right, anonymity right, derivative right, renewable property right, entrusted processing right, conversion right and index right <strong>Usufruct:</strong> Right to sell, right to transfer <strong>Disposition right:</strong> Modification right, storage right, transmission right, deletion right and extraction right</td>
</tr>
<tr>
<td>Data processor</td>
<td>Generally speaking, the data controller itself is also a data processor, but there are also cases where a third party is entrusted to process the data. The data processor only has the rights (mainly the right to dispose and the right to use) granted by the entrustment agreement. After the completion of the contractual obligations, all data and their backups must be deleted, and the data cannot be occupied or used any more.</td>
</tr>
<tr>
<td>Data buyer</td>
<td>The data buyer can have certain data property rights through the transaction contract. At the same time, as a data product consumer, it also has the rights and interests of consumers, including the right to security, the right to know, the right to choose independently, the right to fair trade, the right to claim, the right to associate, the right to know, the right to be respected for human dignity and the right to supervise.</td>
</tr>
<tr>
<td>Data trading platform</td>
<td>1. <strong>Distribution platform:</strong> Through the transfer clause of the transaction contract rights with the data commodity supplier (data controller), we can obtain the right of return of the data commodity. 2. <strong>Channel platform:</strong> As an information intermediary and trading channel, it provides trading places for both parties, so it has no right to data. 3. <strong>Self-produced and self-sold platform:</strong> The data commodities in the trading platform are independent compliance data or brand-new derivative</td>
</tr>
</tbody>
</table>
3 Right description of service data

In the current data circulation market, the data controller needs to obtain the data product identification issued by the regulatory department or trading platform before putting the anonymized service data product into the data circulation market, as the data owner’s statement of all rights of the data product.

In order to further improve the efficiency of online data circulation and promote the automation of data transactions, the right identification and ownership certification of service data should describe the right content of data goods in a structured and machine readable way. ODRL (Open Digital Rights Language) of W3C is a kind of policy expression language. It provides a flexible and interoperable information rights model. It uses the form of policy to express the behavior that the subject allows or forbids the resources in a certain business scenario, as well as the additional restrictions. It uses a specific vocabulary and coding mechanism to represent the usage of content and service, which is suitable for the ternary semantic structure based on subject, object and right.

Referring to the ODRL information model and its common operation vocabulary, combining with the right content of the new property right of the data controller mentioned above, a four-level classification system is defined according to the right intensity as follows.

Data controller/buyer under the full buyout intensity has exclusive control right, revenue right, use right, disposal right and consumer rights; data controller/buyer under the joint ownership intensity has non-exclusive control right, revenue right, use right, disposal right and consumer rights; data controller/buyer under the resale license intensity has revenue right and use right, the right of disposition and the rights and interests of consumers; the data controller/buyer under the general use license intensity only has the right of use, the right of disposition and the rights and interests of consumers.

3.1 Example of right description standard for data controller

Before registering the data product for online circulation, the data controller needs to output the metadata file of the right description of the data product according to the above-mentioned right description specification (or apply to the regulatory authority for the right identification of the product accordingly), so as to declare the rights of the data controller to the data product.

According to the above, the data controller has control right, disposition right, use right, income right and the sub rights of the four rights subordinate to the service data goods, that is, from the perspective of property right, the data controller’s online transaction behavior can be judged by JSON description of the rights. For example, the data controller who lacks the right of return cannot trade data products online.

The right to control data is the premise and basis for the realization of other data property rights enjoyed by data controllers. Take the control right as an example to describe the control right using JSON based on ODRL:
Table 3. JSON description of control right based on ODRL.

```json
{
  "@context": [
    "http://www.w3.org/ns/odrl.jsonld"
  ],
  "@type": "Set",
  "uid": "http://example.com/policy:1000",
  "profile": "http://example.com/odrl:profile:01",
  "ConflictingTerm": "invalid",
  "permission": [
    {
      "target": {
        "@type": "AssetCollection",
        "uid": "http://example.com/asset:6666.SDR"
      },
      "assignee": {
        "@type": ["Party", "vcard:Organization"],
        "uid": "http://example.com/company/A",
        "vcard:fn": "Company A",
        "vcard:hasEmail": "company-contact@example.com"
      },
      "refinement": {
        "xone": {
          "@list": [
            { "@id": "http://example.com/party:Original" },
            { "@id": "http://example.com/party:Non-Original" }
          ]
        }
      },
      "action": "control"
    }
  ]
}
```

The control right JSON policy of the data controller is a set class policy assigned to the assignee, which means that the data controller (company A) is allowed to control the target assets (asset:6666.SDR). The basic description of the assignee (company A) is given with vCard namespace, and its enterprise name and contact information are introduced. In order to make the identification of control right of data controller play a role of traceability in infringement tracking, we add originality as a constraint in the Party class (refinement).

### 3.2 Example of right description standard for data buyer

Data buyer becomes another controller of data goods to some extent after acquiring data goods. The range of rights and interests of data buyer in data goods depends on the strength of rights he chooses when he purchases data goods. According to the common purchase purpose in the current data circulation market, the right intensity can be classified into four levels: level I—full buyout, level II—joint possession, level III—resale license and level IV—general use license.

Take level I - full buyout as an example: the data seller transfers the ownership of the data goods to the data buyer and deletes the corresponding data owned by it. Therefore, the data buyer becomes the only controller and possessor of the data, and owns the complete ownership of this part of data.
Table 4. JSON description of ‘level 1 - full buyout’ based on ODRL.

The JSON strategy of full buyout is an agreement strategy assigned by the assigner to the assignee, which means that the data buyer (company A) is allowed to perform the operation of full buyout on the target asset (asset:666.SDR), but the data buyer must fulfill the obligation of purchasing level I right intensity. After the data buyer completely buys up the data commodity, the data seller must fulfill the obligation of deleting the corresponding data, otherwise it will bear the consequences of being chased and corrected by the regulatory authorities.
4 Conclusion

In the era of digital economy, data has become an important factor of production with unlimited value, and the demand of data circulation is increasingly strong. This study is devoted to the discussion of defining the ownership of service data and the right description standard. On the basis of the understanding of the basic logical relationship between data and information, the study deconstructs the service data, solves the problem of multi-agent power confirmation, defines the ownership and the right content of the main stakeholders, and structures the right content and the right transfer transaction activities in formal language. It can improve the efficiency and automation level of online data circulation, promote the sound development of data circulation market compliance, and provide the possibility for legal supervision such as infringement tracking.

This work was supported by National Key R&D Program of China (2017YFB1401100).

References