Analysis and Research of Administrative Approval Form
Data Standardization

Jianqin Xie*, Weigang Guo and Zhen Zhu
School of Electronic and Information Engineering, Foshan University, Guangdong
Foshan, China
*Corresponding author e-mail: 34603770@qq.com

Abstract. With the rapid development of Internet and big data, "Internet government
service" has been valued by People of Republic China governments at all levels. Since
2008, Foshan City has begun to build an "administrative approval electronic network
integration system", which has achieved remarkable results. However, in terms of data
sharing and standardization, there are rooms of improvement. This paper takes the
Foshan Natural Person Administrative Approval Form as an example for analysis and
research, in order to highlight insight for standardization, and demonstrate
standardized results by examples. It is expected to promote the digitization of e-forms
and the sharing of e-government information in Foshan City.

1. Introduction
In recent years, with the rapid development of the Internet, "Internet-based Government service" has
developed rapidly, and it has also received the attention of the State Council and governments at all
levels. In September 2016, the State Council issued the "China to step up efforts to promote
internet-based governance " Opinions (Guo Fa [2016] No. 55) [1]; In order to promote the
construction of the government affairs platform, in July 2018, the State Council issued the "Govt to
speed up integrated online service platform " (Guo Fa [2018] No. 27) [2], comprehensively improve
the standardization and facilitation level of government affairs, and better provide full-process
integrated online services for enterprises and the masses.

To achieve an integrated service platform, data sharing is key. However, e-government is still in
isolation because of the barriers between inter-departmental e-government information resources, the
backwardness of the original government information infrastructure, the limited access to resources,
and the difficulty in maximizing the effectiveness of resource integration and utilization[3]. Some data
can not be shared, the existing system can only realize the informationization of the government office
process, and does not effectively analyze a large amount of data [4]. With the introduction of big data
concept and rapid development, the application of big data has received widespread attention. Guizhou
has taken the lead and achieved remarkable results [5, 6]. Hangzhou also made great efforts in big data
government management and services[7]. The focus of big data reform is to break the “data island” [8],
but for the form that the user fills in the administrative approval, it is still based on Word, Excel or
other formats. After the user fills in the submission, the field data in the table cannot be extracted,
resulting in the data can not be reused and shared. This paper will analyze and study the administrative
approval form of natural persons in Foshan City, propose standardization rules and practice, and solve the problem of data isolation.

2. Form analysis

2.1. Total analysis
Before standardizing a form, you first need to analyze and understand the meaning of the form, understand the application context of the form and the corresponding administrative approval items. Only by deeply understanding the application scenarios of the approval items and forms, the meaning of the field can be accurately positioned and correctly understood in the form.

2.2. Carding additional materials
According to the instructions in the form, extract the additional materials needed. For example, if the designated representative or the agent can be a natural person or another organization; if the designated representative or the agent is another organization, a copy of the other organization's license should be submitted separately. And the documents of the specific manager and the identity document of the specific manager." If the agent is a natural person, an additional identity document must be submitted; if the agent is an organization, a copy of the organization’s license and its assignment must be attached. And the document of the manager and the identity document of the specific manager.

2.3. Structure split
The content of the form is divided into two parts by nature: the form skeleton and the content. The form skeleton refers to the row, column structure and label text of the form, which serves as the basis for the field splitting and the basis of the field naming; the filling content refers to the specific content that the applicant needs to fill in, which serves as the basis for the field definition. For data with multiple records in the form, such as a resume, split the form in the form of a sub-table, and then divide the sub-table into two parts: the form skeleton and the filling content.

3. Form standardization rules
Based on the relevant national data standards, we propose the following standardization rules:

3.1. Serial number
For the content to be filled in the form (including the check items), all the information (data) that needs to be digitized is numbered sequentially from top to bottom and left to right, starting from 1 and represented by continuous natural numbers.

3.2. Internal identifier
That is, the internal identifier of the field, the encoding rule is "3 department (business) number + 3 digit form number + 3 digits of the code number". For a form, the first 6 bits of all fields are the same, and the last three bits are 0 to 3 bits for the sequence high.

3.3. Chinese name
Fill in the content for the form, and digitize the Chinese name of the corresponding field in the inventory storage time. Make sure that the field names in the form are unique to distinguish them from other fields, and you can briefly indicate the specific point (meaning) of the corresponding content in the field.

3.4. Data type and value domain
Data can be processed manually or automatically. Data types and value domain are used to represent the type of symbols, characters, or other representations of field data and a set of allowed values, including characters, numbers, logical, and enumeration, dates, binary, combination, and so on. The enumeration type and the combination type are determined according to specific circumstances, such
as a degree that can be represented by an enumeration, including a bachelor's degree, a master's degree, a doctor's degree; a time period can be represented by a combination, including two fields of a start time and an end time.

3.5. Data Format
The format requirements for field values specified from a business perspective. Field value data format definitions include: A = alphabetic characters; N = numeric characters; AN = alphanumeric characters; .. n = characters with a maximum length of n; .. ul = text of indeterminate length; .. p, q (p, q are natural numbers) = the longest p digits, q digits after the decimal point; YYYYMMDD = "YYYY" for the year, "MM" for the month, "DD" for the date, can be used in combination according to the actual situation. For example, N..17, 2 means the longest 17 numeric characters, two decimal places; A..20 means the longest 20 alphabetic characters.

3.6. Data Sources
An authoritative source channel (reusable channel) that specifies field data. For example, the education comes from the education department, the name comes from the public security department.

3.7. Object class
The boundaries and meanings can be clearly identified, and the characteristics and behaviors follow the same rules of the rules, abstract concepts, or collections of things in the real world. In short, the object class indicates the object class that the field data describes. Such as natural person, organization, item, etc.

3.8. Characteristic word
A feature common to all members of an object class. Such as name, organization name, permission, etc.

3.9. Representation
A domain of values, a combination of data types, and, if necessary, a unit of measure or a character set. Such as name, picture, enumeration, description, etc.

3.10. Definition
The detailed definition of the Chinese name of the field, if the form is marked, should be as faithful as possible to the original form to mark the text content, and can be used as the annotation text and prompt text output by the future form generation tool.

3.11 Remarks
Make necessary descriptions of the object, content, or field value of the field description. Such as electronic signatures, signature pictures.

4. Example
The following is an example of the “Approval Form of post-doctoral living allowance in Gaoming District” of the Foshan Human Resources and Social Security Bureau No. 85, which standardizes and digitizes the form. The original form is shown in Figure 1.
This form is used for post-doctoral applications for living allowances provided by the Gaoming District Government. The information includes basic personal information, resumes, performance and departmental opinions, and so on. According to the description in the table, the ancillary materials mainly include the original and copy of the ID card, academic certificate, agreement or contract, outbound approval form, transfer introduction letter, and the recent one-inch photo.

The resume in the form has multiple pieces of data, which are split into sub-tables. The skeletons in the original table and sub-tables are split according to the row, column structure and label text of the form, and the content is blank. According to the aforementioned standardization rules, the results of standardizing the form data are shown in Table 1. (To save space, only some representative fields are selected in the results).
Table 1. Results of the standardized application for approval form of the post-doctoral living allowance in Gaoming District.

| SN | Internal identifier | Chinese name | Data type and value domain | Data Format | Data Sources | Object class | Characteristic word | Represen-tation | Definition | Remarks |
|----|---------------------|--------------|-----------------------------|-------------|--------------|---------------|---------------------|----------------|------------|---------|
| 1  | 012085 001          | Work unit    | Charac-ters                | AN..120     | personal fill | Organ-ization | work unit           | name            | work unit  | Work unit of applicant |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 2  | 012085 002          | Name         | Charac-ters                | A..30       | police       | natura-l person | name               | name            |            | Name of applicant       |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 3  | 012085 003          | Date of birth | Dates                      | YYYYMMDD    | police       | natura-l person | date               | date            |            | Date of birth certificate and formal registration in household registration management department |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 4  | 012085 004          | Gender       | Enumeration_gender         | A..6        | police       | natura-l person | gender             | Enumeration_gender | Basic physiological characteristics of person |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 5  | 012085 005          | Registered residence | Combination_Register Address | Combina-tion Address | police       | natura-l person | registered residence | Combination_Register Address | registered residence |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 6  | 012085 006          | Photograph   | Binary Photo               | personal fill | Photo       | natura-l person | Photo              | photo           | photograph | Photograph of applicant, scanning acquisition |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 7  | 012085 007          | ID number    | Charac-ters                | AN..18      | police       | natura-l person | ID number          | Only life-long unique identifier | ID number of applicant |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 8  | 012085 008          | Standard of living allowance | Numbers | N..8,2       | personal fill | item           | amount of money    | Measure          | Standard of living allowance |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 9  | 012085 009          | Resume       | Sub-table_resume           | Sub-table   | item         | resume          | sub-table          | Resume of applicant | 012085ZB01 sub-table_resume |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 10 | 012085 010          | Applicant's opinion | Charac-ters                | AN..ul      | personal fill | item           | opinion            | describe | Applicant's opinion |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 11 | 012085 011          | Person in charge of application unit | Binary | Photo | personal fill | item           | autograph          | photo             | Autograph of person in charge |
|    |                     |              |                             |             |              |               |                     |                 |            |         |
| 12 | 012085 012          | Stamp of application unit | Binary | Photo | personal fill | item           | seal               | photo             | Stamp of application unit |

The sub-tables are shown in Table 2 (selected parts), the Combination_Register Address is shown in Table 3, and the Enumeration_gender is shown in Table 4.
Table 2. Sub-table standardization results.

| S N | Internal identifier | Chinese name | Data type and value domain | Data Format | Data Sources | Object class | Characteristic word | Representation Definition | Remarks |
|-----|---------------------|--------------|-----------------------------|-------------|--------------|--------------|---------------------|--------------------------|---------|
| 1   | 012085009-01        | Date of beginning and ending | Combination times | Combin -ation personal fill | natural person | times | Combin -ation | Resume time | Combination times |
| 2   | 012085009-02        | Post Characters | AN.30 | personal fill | natural person | post | describe | Applicant’ s post in work unit |         |

Table 3. Combination_Register Address.

| Serial number | Combination name | Data type and value domain | Data Format | Characteristic word | Representation |
|---------------|-----------------|-----------------------------|-------------|--------------------|----------------|
| 1             | province        | enumeration_province        | A..4        | province name      |                |
| 2             | city            | enumeration_city            | A..20       | city name          |                |
| 3             | county(area)    | enumeration_county(area)    | A..20       | county(area) name  |                |
| 4             | town(street)    | characters                  | A..20       | town(street) name  |                |
| 5             | village         | characters                  | A..20       | village name       |                |

Table 4. Enumeration_gender.

| Enumeration name | Code | Enumeration value |
|------------------|------|-------------------|
| Enumeration_gender | 0    | unknown gender    |
| Enumeration_gender | 1    | male              |
| Enumeration_gender | 2    | female            |
| Enumeration_gender | 9    | unexplained gender |
5. The next step of the job
The results of standardization in this paper are mainly used for the integrated platform for administrative approval. In order to continuously improve the data and maximize the sharing of data, the data needs to be further improved. The next step will continue to deepen the following work.

5.1. Combing enumerations and combinations
In the natural person approval form, there are many repeated enumeration and combination type fields. The next step is to extract all the public enumerations and common combinations in the combing process to achieve maximum reuse.

5.2. In-depth investigation and research
For complex forms with strong professionalism, it is necessary to conduct in-depth investigation and research on relevant business departments in order to accurately understand the administrative approval items and application scenarios corresponding to the forms, and correctly understand the meaning of each field. In addition, the current data source data is mainly based on personal understanding of the data, there may be certain errors, only through extensive investigation and research can correctly understand the authoritative source of data, to ensure the accuracy and real-time data.

5.3. Research on the relationship between data
In the administrative examination and approval, the number of forms is high and the complexity is high. The relationship between the fields is complicated. In the context of big data, the traditional government mode has been difficult to meet people's basic needs and requirements. Combining the network to give full play to the interactivity and openness of e-government[9], clear the relationship between data can really promote the extensive application of big data in Foshan government services. The next step will be to build clear hierarchical relationships between fields and use Knowledge Graph to describe concepts and their relationships in the physical world in symbolic form [10], and use visualization techniques to display them. Mine, analyze, construct, draw and display data and the interrelationship between them.

6. Conclusion
Since the establishment of the “Administrative Approval Electronic Network Integration System” in 2008, Foshan has been continuously reforming and innovating, and is in the forefront of the development of national e-government. In 2018, the “Digital Government” construction plan issued by the Foshan Municipal Government (2018-2020) proposed to establish a unified data standard and data resource sharing platform. The implementation of this specification will help realize the implementation of data standards and sharing, and further promote the “Internet-based Government service” construction, and play a leading role in the country.

Acknowledgements
This paper was completed with the support of the “Guangdong Province One-Door Government Affairs Big Data Engineering Technology Research Center” project and the Foshan City Social Science Project “Foshan City Natural Person Administrative Approval Form Data Standardization Research”.

References
[1] Information on http://www.gov.cn/zxxw/content/2016-09/29/content_5113369.htm
[2] Information on http://www.gov.cn/zhengce/content/2018-07/31/content_5310797.htm
[3] Meng Yuyu, Research on E-government Information Resources Sharing Model under the Background of Big Data [J], Henan Science and Technology, 2017(11): 29-30.
[4] MA Bai-yu, Leveraging e-government big data to promote the innovation and development of E-government [J], Electronic components and information technology, 2017, 1(2):10-12+51.
[5] N. Wang, Policy implications of big data development in developed countries and major regions of China, China Management Informationization, Vol.20(2017), No.4,159-160.

[6] Z.N. Mu, S.B. Du, On the Development Situation of Big Data in Guizhou Rural E-commerce and Its Measures, Value Engineering, Vol.35, (2016), No.8,38-39.

[7] N.G. Deng, Research on Government Management Service Innovation in Big Data Era, Journal of Tianjin Administration Institute, Vol.18(2016),No5,18-25.

[8] Y.X. Chen, The Second Information Revolution in Nanhai District of Foshan City, Decision-Making, Vol.23(2016)No.8,52-53.

[9] Zhang Xi, The Impact of Big Data on E-government and the Corresponding Tactics [J], China Computer&Communication, 2017(19):158-159.

[10] Liu Qiao, Li Yang, Duan Hong, Liu Yao, Qin Zhiguang, Knowledge Graph Construction Techniques[J], Journal of Computer Research and Development, DOI:10.7544/issn1000-1239.2016.20148228.53(3)582-600,2016.