The Innovative Path of Social News Reporting Under the Perspective of Big Data

Xinlan Li(✉)

School of Journalism, Communication University of China, Beijing 100000, China
lxl1239801939@163.com

Abstract. The rapid development of big data technology is driving its deep integration and innovative applications in different industry sectors, and more and more enterprises and organizations are releasing the potential value of data with the data storage and processing capabilities provided by big data platforms. As companies and organizations collect and use more and more data, big data platforms, as data carriers, are exposed to high data security risks. While data journalism has an inherent scientific "character" due to its "quantitative analysis", visualization has advantages in terms of visual impact, intuitive communication and visual presentation that cannot be matched by text or simple photo journalism. The visualization of data journalism achieved the effect of accurate, intuitive, and efficient news reporting during the most serious period of the new pneumonia epidemic in China in 2020.

Keywords: Big Data · Press Release · New media · Data News Visualization

1 Introduction

With the continuous development and advancement of computer and network technology, a variety of online and offline applications have captured a wide range of human data "imprints", and news and information are also included in them [1, 2]. Data journalism is a new form of reporting that has emerged with the advent of the big data era, and its emergence has changed the traditional news production process to a certain extent [3].

2 Graph Matching Based Data Compliance Verification Method

2.1 Methodology Overview

As shown in Fig. 1. The methodology involves three main aspects: the rule layer, the data processing layer and the compliance analysis layer, as detailed below:

P. Qi and Z. Chen (Eds.): ICBDIE 2023, AISR 178, pp. 402–407, 2024.
https://doi.org/10.2991/978-94-6463-238-5_56

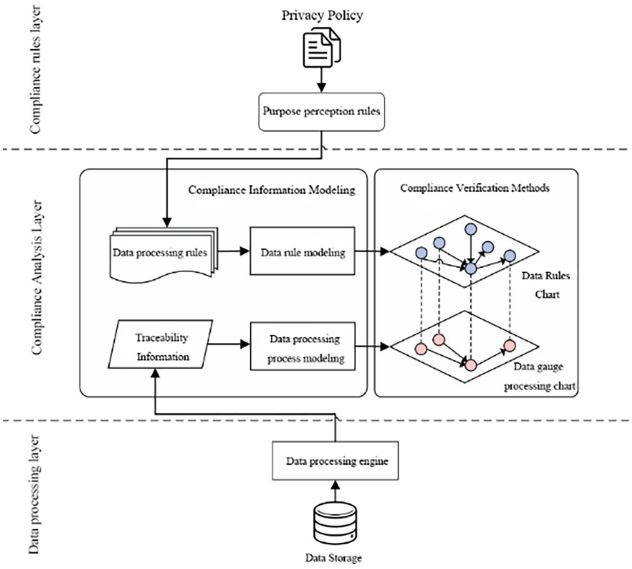

Fig. 1. Data Compliance Analysis Methodology

2.2 Data Rule Modeling

A data processing rule corresponds to a data processing purpose [4], and the roles, operations, and data in the rule are driven by the data processing purpose [5]. The data processing rules are expressed as follows:

$$DPR = \{Role, Process, Resource, Purpose\}$$

Data Rule Graph (DRG) [6]. The data rule graph corresponding to a data processing rule can be expressed as $G_R = (V_R, E_R, Rel_R)$, V_R is the set of vertices of the graph, E_R is the set of edges of the graph, Rel_R is a collection of side labels:

$$V_R \subseteq P_R \cup R_R \cup D_R \cup O_R$$

$$E_R \subseteq (P_R \times R_R \times allow) \cup (R_R \times D_R \times use) \cup (D_R \times O_R \times processedBy)$$

$$Rel_R \subseteq (allow, use, processedBy)$$

where P_R is the set of destination vertices, R_R is the set of role vertices, D_R is the set of data vertices, and O_R is the set of operation vertices [6].

3 Compliance Verification Methods

3.1 Concretized Hierarchical Model of Data Rules

The respective fleshed-out hierarchy for each domain to realize the deduction of the high-level data rules to the low-level implementation and the generalization of the low-level data rules to the high-level policy design [7].

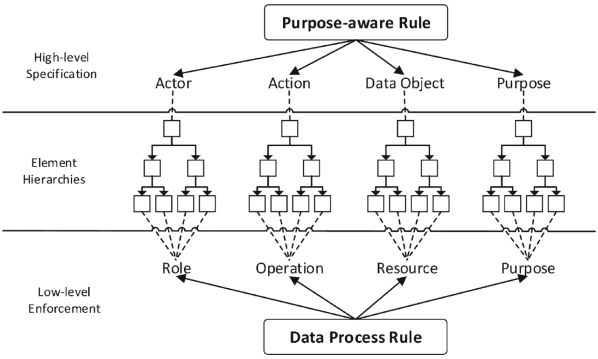

Fig. 2. Data rule hierarchy model

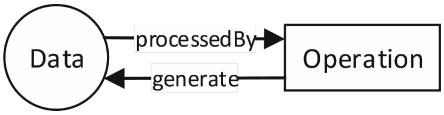

Fig. 3. Data Traceability Graph Model.

The data rule hierarchy model in Fig. 2 shows the mapping relationship between the purpose sensing rules and the data processing rules [8].

3.2 Data Processing Graph Generation

The data traceability graph is represented by a directed acyclic graph $G_T = (V_T, E_T, Rel_T)$, where V_P is the set of vertices of the graph, E_P is the set of edges of the graph, and Rel_P is the set of edge labels.

$$V_T \subseteq D_T \cup O_T$$

$$E_T \subseteq (D_T \times O_T \times usedBy) \cup (O_T \times D_T \times generate)$$

$$Rel_T \subseteq (usedBy, generate)$$

where D_T is the set of data vertices and O_T is the set of operation vertices. There are two relationships between data and operation, when data is used as the input of operation, data and operation are used By relationship, when data is used as the output of operation, operation and data are generated relationship, as shown in Fig. 3.

Fig. 4.Example of data processing chain simplification

Different cases of data processing chain simplification are illustrated in Fig. 4.

4 Experimental Analysis

4.1 Validity Testing and Analysis

The data compliance analysis method proposed in this paper is mainly for the purpose of protecting users' personal information (Table 1).

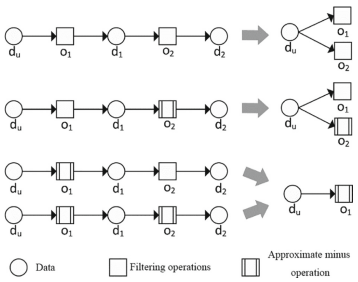

Fig. 4. Example of data processing chain simplification

Table 1. Test rules

Number	Rules
1	Personalized analysis, User1, {Customer, Customer_address}, {{projection, filter, aggregation}, {projection,filter,aggregation}}
2	Statistical Analysis, User2, {Customer, Customer_address}, {{filter,aggregation}, {filter,aggregation}}
3	Sensitive analysis, User3, {Customer, Customer_demographics}, {{projection,filter, aggregation},{projection,filter,aggregation}}

4.2 Performance Testing and Analysis

(1) Additional time overhead for compliance analysis (Fig. 5).

(2) Performance analysis of data processing graph generation algorithm [9]. The experimental results are shown in Fig. 6.

(3) Performance analysis of compliance analysis algorithms. The experimental results are shown in Fig. 7 and (Fig. 8)

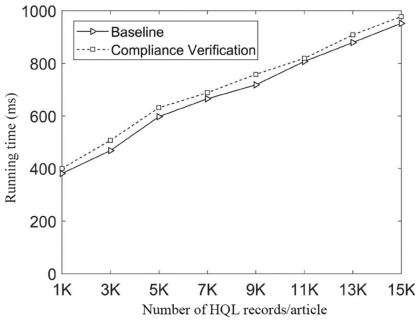

Fig. 5. Comparison of audit information query time before and after adding compliance analysis to Atlas.

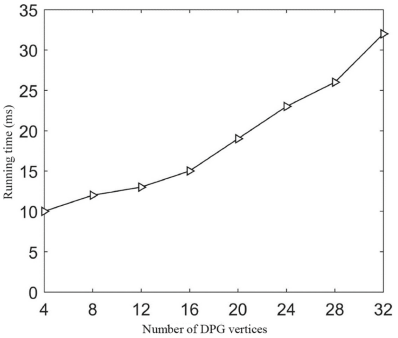

Fig. 6. Graph simplification algorithm running time as influenced by the number of graph vertices.

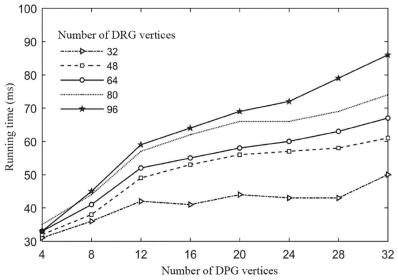

Fig. 7. Compliance analysis time for different data processing graph and data rule graph sizes.

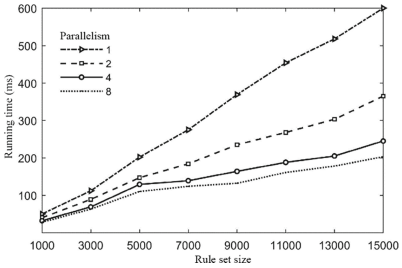

Fig. 8. Compliance analysis time for different rule set sizes

5 Conclusion

This paper conducts research on data privacy compliance issues for big data interactive analytics platforms, and proposes a graph matching-based data compliance verification method. Based on the analysis of the provisions of the standards and norms related to personal information protection, three compliance requirements are proposed for the role of data processing, the use of data and the purpose of processing. The data compliance analysis problem is transformed into a graph matching task by modeling the

data processing rules and the data processing process to construct a data rule graph and a data processing graph based on directed acyclic graphs, respectively.

References

1. Cui Ping. The innovation path of social news reporting in the context of big data[J]. China Press,2022(23):36-37. DOI: https://doi.org/10.13854/j.cnki.cni.2022.23.027.
2. Zhou Xueyan. The innovation path of social news reporting in the context of big data[J]. Journalism Research Guide,2020,11(08):154-155.
3. Feng Xinying. The innovation path of journalism communication talent training mode in the context of big data[J]. Western Radio and Television, 2020(06):83-84.
4. Chen Qiumeng. Research on the cognition and social opinion of news communication in the era of big data--a review of "The era of big data in news communication"[J]. Leadership Science,2019(22):129. doi: https://doi.org/10.19572/j.cnki.ldkx.2019.22.082.
5. Zhang Nan. Exploring the impact of the development of big data on news communication[J]. Media Forum,2019,2(18):85.
6. Yang Ruonan. Research on the content of the column "Illustrated News" of People's Daily Online [D]. Hebei University,2019.
7. Zhang Chao. The current situation and social value of big data news development[J]. Television Research,2018(08):77-79.
8. Bayan-Bahati. Research on Data News Column of Xinjing News New Media Platform[D]. Xinjiang University,2018.
9. Gan Yongkun. Exploring the evolution of social change and news communication concept in the era of big data[J]. Communication Power Research,2018,2(14):52.