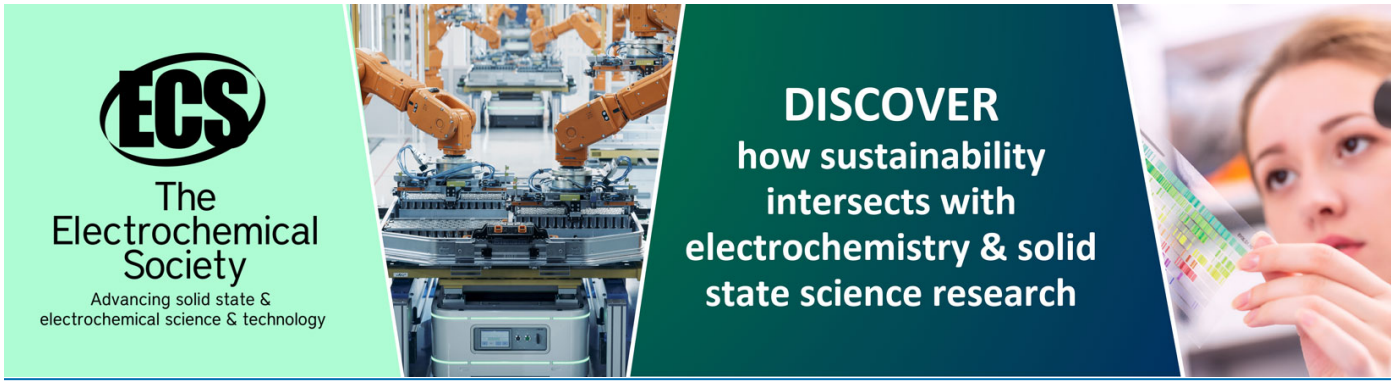
Innovative Research on the Improvement of Visual Quality of Data Journalism in China: Visual Language and Interaction Design



Innovative Research on the Improvement of Visual Quality of Data Journalism in China: Visual Language and Interaction Design

Zhirui Chen[1,a] and Wenchen Guo[2,b]

[1]Communication University of China, No. 1 Dingfuzhuang East Street, Chaoyang District, Beijing, China

[a]czr@cuc.edu.cn; [b]beihuanfanchen@sina.com

Abstract. As the last node of data journalism creation, visual design plays a crucial role in the transmission of data information. Exploring the routes to improve the quality of data journalism at the visual language and interaction design, this paper analyzes the works won data journalism awards and produced by well-known media from 2012 to 2019, and discusses the misunderstandings and shortcomings in the current data journalism visualization practice. This paper presents the improvement routes of data mapping, symbolic expression, interaction level, and interactive narration.

1. Introduction

In the era of big data, the influence of digital media and information dissemination technology is becoming more and more significant. In order to adapt to the ever-changing information environment, journalism explores in practice a new paradigm of data journalism that is different from other news reports. Since Adrian Holovaty, the founder of Every Block website in 2006, put forward the concept of data journalism, many experts, scholars, and media at home and abroad have explained from the perspectives of content form, production process, and development trend. Although the definition of data journalism is not yet unified, from the standpoint of existing viewpoints and research, it can be roughly summarized as the way of data-driven, cleaning, and analysis as intermediate links, and visual content in the form of news content.

The visual design uses visual language and interactive technology to construct a user-centered visual and cognitive interactive experience. Its quality is directly related to whether the data acquired in the initial stage and processed in the intermediate step can be accurately, clearly and profoundly transmitted to the audience, and ultimately affects the quality of the news report. Therefore, this paper is based on the works won data journalism awards and produced by well-known media from 2012 to 2019, and combines journalism, design, psychology and other related theories to explore the data journalism cases in the visual and interactive design problems and successful experience, so as to provide some routes for the development of data journalism in China.

2. Visualization of data journalism

Information visualization is a visual presentation of abstract data. The presented data include both pure numerical information as well as non-numeric information [1]. The generalized information visualization has a long history. In the Paleolithic age, there have been simple expressions and means

of implementation. The drawing of the Southern Changsha State garrison map in the Western Han Dynasty was completely accurate. In the 18th century, information visualization began to develop a more standardized and scientific system. William Playfair pioneered graphic statistics and put forward linear, strip and pie statistics. Since contemporary times, in order to represent more data types and relationships, visual graphics and chart styles have been ever more abundant. The enabling of related theories and technologies in modern graphics and computer science has made information visualization change from static to dynamic, from single dimension to multi-dimension.

As the final step of the data journalism production process, visualization of data journalism is directly oriented to the audience. Therefore, it should not only undertake the functional essence of information visualization and presents the data visually, but effectively integrate and convey information in the graphics. In the vision and interaction, the data are more clearly and kindly to tell the news story, and let the audience understand the content, restore the truth of the news for the audience, then reflect the news value. From the current creation practice, the visualization mainly consists of visual design and interaction design.

3. The visual language in the context of data journalism: data mapping and symbolic expression

3.1. Data mapping

Mapping in mathematics refers to the relationship between elements of a set of two elements that are corresponding to each other. In the visual design of data journalism, the transformation relationship between data and visual results is also a mapping rule. The mapping relationship is the basis of visualization, such as the correspondence between numerical size and length, area, volume, correspondence between color and political tendency Figure 1.

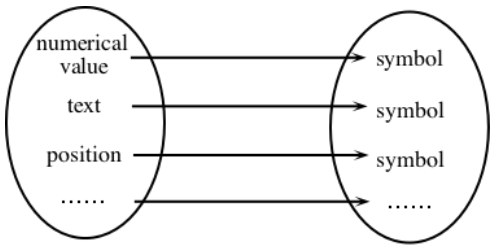


Figure 1. Symbol mapping

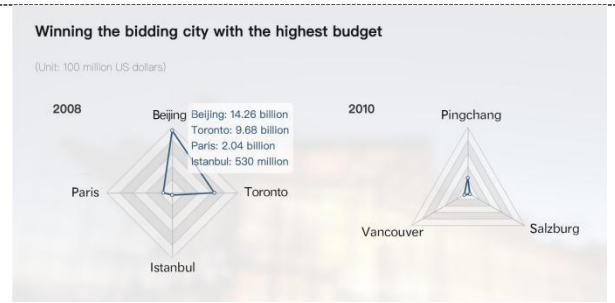


Figure 2. Radar maps are unsuited in *Is the 2022 Beijing Winter Olympics a worthy investment?*

The choice of mapping rules is to comply with the information content and data logic. Mapping data into inappropriate graphical charts can lead to logical confusion. In the *Is the 2022 Beijing Winter Olympics a* worthy *investment?* of the China Data Visualization Competition, in order to express the relationship between the budget of the bidding city and the winning result, the author used the radar chart Figure 2. The radar chart is mainly used to present different aspects of the same subject. Although it is pleasing, it does not conform to the data type. The correspondence between the data points and the variables in the chart is also misplaced. The radar chart here is not optimal mapping rules. In addition to the well-known histograms, line graphs, and ring graphs, the relationship between some infographics and the rendered data is as follows Table 1:

Table 1. Relationship between infographics and data

Chart type	Description
Scatter plot	Judge whether there is some association between variables or the distribution mode of summary coordinate points.

K-line chart	Generally use to show the changes of stock market and futures market.
Chord diagram	Show relationship in display matrix.
Funnel plot	It is applicable to process analysis with long cycle and multiple links, and problems are found through data of each link of funnel.
Contour map	Useg the equal value of the numerical point line to represent the continuous distribution and change of data.
Heat map	Use color to express the two-dimensional numerical size of position correlation.
Venn diagram	Use the closed graph on the plane to express the relationship between data sets.

The human brain has a long-term storage cognitive mode for visual symbols, people form fixed visual cues for the mutual association and influence of corresponding symbols in daily life, the principle of establishing mapping rules should base on cognitive law. By coding the visual symbols, the data are matched with the cognitive patterns in the user's brain. Under the visual cues, the information carried by the symbols is recognized and understood [2]. Besides, the mapping process should be clear and standardized. Data source and unit description are indispensable, which are the basis for users to interpret coded data. The scale division and ruler in the visual component should be scientific and unified to avoid ambiguity.

3.2. Symbolic expression

Symbols are the result of data mapping and are a full expression of graphic elements such as color, graphic images, texture, and animation.

Compared with the shape of the figure, the color has a higher visual impact and appeal. Some scholars believe that the specific form of color language is formal language, rhetorical language, aesthetic language. In other words, color enables the audience to understand the content through visual perception, reveals the thinking intention of the creators, and gives the audience an aesthetic experience. Color attributes (lightness, hue, and saturation) are often used to classify the levels and categories of data. Lightness and saturation are suitable for encoding quantitative data and reflect the order of data. Hue (such as similar color, complementary color) is suitable for encoding qualitative data, reflecting the data hierarchy and structure. The use of a single and intense color (or dominant color) can create a positive atmosphere and convey strong emotions. For example, the award-winning work of the China Data Journalism Competition *The Status of China's Lost Overseas Cultural Relics* adopts dark red dominant tone on the whole, which is beautiful and elegant and has a warning metaphor, which conforms to the theme. Color harmony and personality collocation can enhance the style of the whole work, give the audience aesthetic experience. Conventional color matching systems include Color Brewer, Color Scheme Designer, and Kuler.

The graphic design of symbols should first have precise readability and recognition. In the Global Data Journalism Awards 2019 *Every Time Ford and Kavanaugh Dodged a Question*, the author hides the text of the hearing behind the line, and the text can appear after clicking the line Figure 3. The graphic is relatively abstract, which is not conducive to user interpretation. Besides, graphic design should adapt to the mobile display. For instance, in the mobile interface design, the minimum font size is not less than 20px [3].

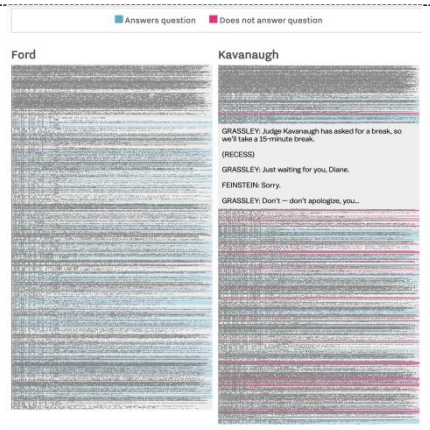


Figure 3. The first page of *Every Time Ford and Kavanaugh Dodged a Question.*

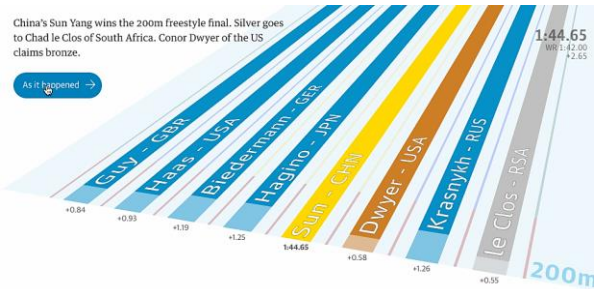


Figure 4. This work, *How Sun Yang Captures the 200m Freestyle Gold* Medal*,* designs a histogram into a swim lane.

Symbol design should focus on variations and innovations. In recent years, data visualization tools and templates have become ever more mature. Users can make visual charts more conveniently by using Google Chart API, d3.js, Echarts, Highcharts, and other tools. However, homogenization and templates of design have also appeared. The innovation of charts needs to be strengthened urgently. Designers can use the basic graphics to construct new shapes through composition, visual association, and imagination through bone, approximation, repetition, gradation, contrast, and space. For example, Xinhuanet's *Home Banquet* uses force-oriented maps to form a new "plate" style Figure 4, which is creative. The Guardian's visualization work *How Sun Yang Captures the 200m Freestyle Gold Medal* has turned the bar chart into a swimming lane [Figure 4], showing the athlete's status from start to finish. Besides, the use of images can bring a sense of realism and identification to the audience. The work of the China Data Visualization Competition *Children Go Home--70,000 Data Interpretation of Chinese Children's Trafficking and Abandonment* combines children's photos into a photo wall. Visual impact evokes a concern for this group.

Texture and animation are also often used in the expression of visual symbols. The texture is a combination of various visual element attributes, generally including shape, color, direction, etc, which can be used to indicate information such as category and height; however, because its constituent elements are slightly complicated, it is easy to cause visual fatigue when used. The animation is an association of various motion attributes, generally including the direction, speed, frequency, etc. The direction attribute can encode qualitative data attributes, while the latter two are often used to encode quantitative data attributes. However, because the animation effect is visually attractive to the user, it is easy to interfere with the user's interpretation of other data. Therefore, designers must weigh the pros and cons when using the above two visual elements and consider them carefully.

The visual design of the data journalism is the unit of functionality and aesthetics. To sum up the methodology of design, combined with the specific context of journalism, the following methodology can be used for reference. First, it is intuitive and straightforward. On the basis of ensuring the accuracy of information transmission, symbols with relatively few elements are selected for expression. Too cool symbols will interfere with information transmission and affect the visual experience. Secondly, the emphasis should be highlighted. Designers should order the importance of information on the limited interface, so that the essential information content can be paid attention to by users. Finally, it is the balance, such as the balance of color usage, the balance of picture composition, the balance of text and chart, etc.

4. The interaction design in the context of data journalism: interaction level and interactive narrative

Interaction is the most seriously misunderstood and ruthlessly abused computer term [4]. From a technical point of view, all the user's input behaviors on the digital media will interact with the interface and generate output, which can be called interaction. The interaction in the context of data journalism can be understood as the process in which users operate and understand data and news information through interaction with the visual system.

4.1. The interaction level

Referring to Weber and Rall's research paper on interactive information graph in 2012 [5], this paper classifies data journalism interactions of the winners of the data journalism awards at home and abroad in 2012 and 2019.

Through the analysis, based on the time dimension, it can be seen that from the first Global Data Journalism Award in 2012 to the latest in 2019, the proportion of intermediate interactions and high-level interactions has increased significantly. Based on the award-winning works of China and the world, there is no significant difference in the level of interaction. The data journalism' interaction level at home and abroad is still at an intermediate level. Although the form of interaction ultimately serves the content, interaction can effectively engage the user, enhance the understanding and analysis of the data, and alleviate the contradiction between insufficient visual space and data overload [6].

The improvement of the interaction level mainly depends on interaction technology. The technologies commonly used in data journalism works (pages) to control behaviors or dynamic interaction effects are CSS3, JavaScript, jQuery, Ajax, etc. Animation in CSS3 can replace Flash to achieve some simple dynamic effects, but the interaction effect is challenging to achieve. JavaScript achieves dynamic and interactive effects by changing images and events over time. The principle of JavaScript is to change the attributes of elements (such as size, location, resources) when some elementary events occur so that to achieve dynamic effects. These changes are usually designed in the form of custom functions in the scripting language, which is a challenge for people with no programming background. JQuery is a special subset of JavaScript on dynamic interaction design. It has a unique chained syntax and a short and clear multi-function interface. What is more, it has a convenient plug-in extension mechanism. Designers can make use of already wrapped functions to achieve effects easily. For example, the following program code implements a sliding display of hidden paragraph tag elements:

```
$ (′ ′ .btn2 ′ ′ ).click (function ( ){ });
$ (′ ′ .p ′ ′ ).slideDown ( );
```

4.2. Interactive narrative

Interactive narrative refers to narrative behavior using interactive techniques. On the basis of interaction form and level, Chinese and foreign scholars have explored the interactive narrative mode of data journalism from different angles. Segel and Hear summarized it as author-oriented, reader-oriented, and mixed-type from the perspective of the degree of narrative subject bias [7]. Mary Lynn Young summarized interpretive and exploratory narratives based on the degree of interaction [8]. Fang Yihua, a scholar at the Communication University of China, summarized it as reproduction, composition, and interaction from the perspective of grammatical function [9]. Zhang Chao, a scholar at Shandong University, summed up the "database-based exploration narrative" and "game-based experience narrative" from the technical and formal aspects [10]. Among them, the model concluded from the perspective of technology and form is more suitable for the needs of actual design and creation, so this paper will discuss from the viewpoint of scholar Zhang Chao.

4.2.1. Database-based exploration narrative design. A database is a collection of data that is based on specific rules, stores large amounts of data, and can be shared with multiple users. Database-based interactive narratives give users more autonomy in choosing the story behind the data. The key to

designing such narratives is to filter the data built on journalism value and logical hierarchical layout, because value and level determine the depth of exploration narrative. Finally, through an intuitive and friendly chart, to meet the user's information needs. The China Data Journalism Competition's award-winning work Endangered Animal Protection: The Edge of Extinction showed the endangered species in 2007-2018 Figure 5. Each dot is a species. When the user hovers the mouse over the dot, the name of the species appears; as the user scrolls through the text, the chart is linked with the text and presented to the user. The user can not only grasp the number and change of species endangered level, but also check the specific name of each species.

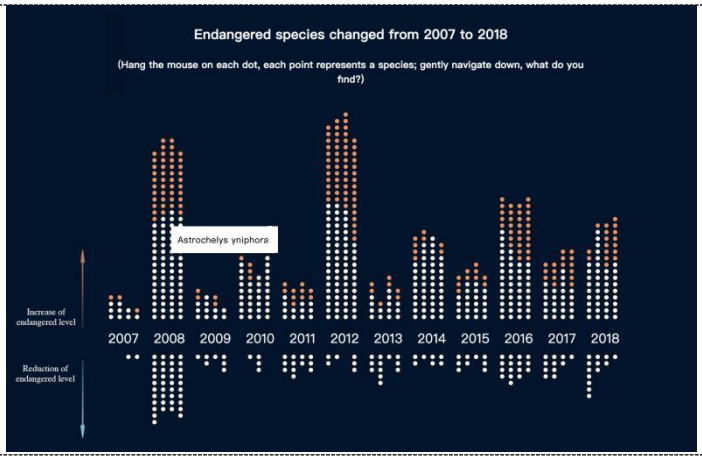


Figure 5. This figure presents the interactive chart of *Endangered Animal Protection: The Edge of Extinction.*

4.2.2. Game-based experience narrative design. News games are gamification of news. Game design elements are used in the context of news, and users can interact and understand the information in a virtual context [11]. News game is an essential form of an interactive narrative of data journalism. The winner of the China Data Journalism Competition, The toilet revolution? Not obeying to "build"! , users can choose to build toilets in rural or urban areas according to their interests Figure 6. After completing the construction by installing accessories by clicking, dragging, the text explanation will also be based on three different results [Figure 7]. This work uses the game to assist the "toilet revolution" reporting. At the same time, it also allows users to participate in the narrative. Users are acting as the news parties or stakeholders, and build a narrative framework of personalized cognition and emotion in the game.

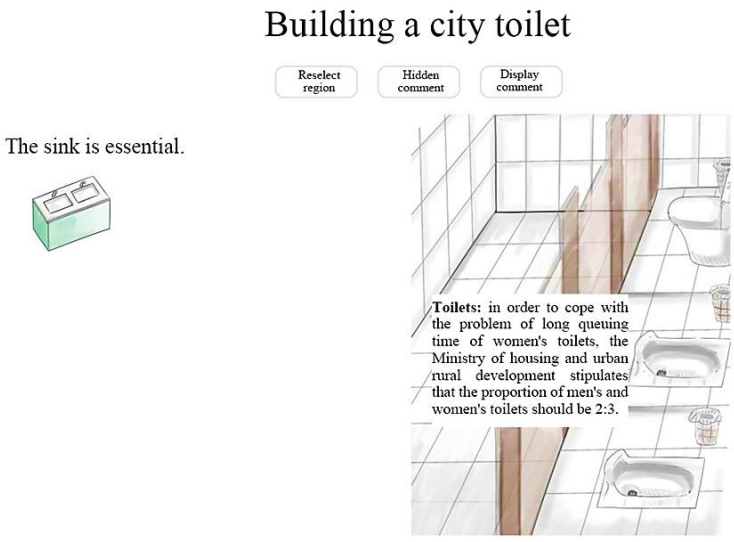


Figure 6. This figure presents the interactive interface of The toilet revolution? Not obeying to "build"!

Figure 7. This figure shows the end interface of The toilet revolution? Not obeying to "build"!, giving feedback to the user's choice.

According to the level of gamification of news games, news games can be divided into three categories: imitation experience, social survey, and text editing. In the design of imitation experience games, we should pay attention to the creation of immersion, such as keeping the story, background, and interface in the design, using real or lifelike images, so that users can integrate into the visual experience as soon as possible, and then immersed in the narrative context. The design of social survey games should strengthen user interaction and data collection and analysis, make full use of questionnaires, role-playing, dialogue, comment, and other means to obtain user information, analyze and filter out valuable content, and obtain news added value. The design of text editing games should focus on the combination of linear and non-linear. For example, the designer can use the bifurcation tree to let the user make judgment choices in the fork. Different selection order will arrange the elements in the news event differently. In turn, there are many possibilities such as sequence, flashback, and narrative. The narrative of the whole news changes from a single fixed linear to a linear and nonlinear combination and the gamified narrative structure becomes more diverse.

5. Conclusion

Data journalism is a news reporting paradigm that uses data to discover facts and express them visually. However, there are few researches on the path and methodology of improving the visualization quality of data journalism. Through case analysis of the award-winning works of domestic and foreign data journalism awards and the works produced by well-known media, it is found that there are still some problems in the data presentation of current data journalism reports, such as the lack of standardization, directness, weak symbol design ability, low-level of interaction, and lack of immersion. Based on relevant theories and practical creation experience, this paper summarizes the promotion path and methodology: first, the establishment of data mapping law should be based on information content, data logic, presentation purpose, and cognitive law, and the mapping process should be standardized and unified. The symbolic expression should bring into full play the tension of visual elements such as color, graphics, and images. On the condition of ensuring the identification, design composition and visual association should be utilized to enhance innovation and give users an aesthetic experience. Third, further enhance the level of interaction, relying on interactive technology to improve human-computer interaction experience; Fourthly, interactive narrative can explore and experience narrative design by means of database and news game, so as to build a narrative framework of user cognition and emotion.

References

[1] Xinyi F, Shixia L and Yingqing X 2017 The development and thinking of information visualization *J. Decoration.* **4** pp 16–19
[2] Donald A N 2016 *The Desig 0n of Everyday Things* (Beijing: CITIC publishing group) p 65
[3] Allen C and Robert L 2015 *About Face 4* (Beijing: Electronic Industry Press) pp 414–452
[4] Chris C 2015 *Chris Crawford on Interactive Storytelling* (Beijing: People's post and Telecommunications Press) p 23

[5] Weber and Rall 2012 Between data visualization and visual story elling:the interactive information graphic as a hybrid form *C. 2012 Annual Conference of International Communication Association* (Phoenix Arizona) pp 1–37
[6] Su Z, Lina Z and Wen W 2016 *Big Data Visualization Technology* (Beijing: Tsinghua University Press) pp 169–172
[7] Segel E and Heer J 2010 Narrative Visualization :Telling Stories with Data J. *IEEE Transactions on Visualization & Computer Graphics* **16** pp 1139–1148
[8] Mary L Y Alfred H and Johanna F 2018 What Makes for Great Data Journalism? *J. Journalism Practice* **12** 115–135
[9] Yihua F and Huihan Y 2018 On the narrative paradigm of data news *J. Modern Communication (Journal of Communication University of China)* **40** pp 45–49
[10] Chao Z 2017 A Preliminary Study of the Interactive Narrative of Data News J. *The Press* **45** pp 10–15
[11] Jingwei C 2015 *Game Psychology* (Beijing: Communication University of China Press) p 82