

An Abductive Process of Developing Interactive Data Visualization: A Case Study of Market Attractiveness Analysis

Qi Li, Kecheng Liu

► **To cite this version:**

Qi Li, Kecheng Liu. An Abductive Process of Developing Interactive Data Visualization: A Case Study of Market Attractiveness Analysis. 18th International Conference on Informatics and Semiotics in Organisations (ICISO), Jul 2018, Reading, United Kingdom. pp.289-298, 10.1007/978-3-319-94541-5_29 . hal-01920741

HAL Id: hal-01920741

<https://hal.inria.fr/hal-01920741>

Submitted on 13 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



An Abductive Process of Developing Interactive Data Visualization: A Case Study of Market Attractiveness Analysis

Qi Li and Kecheng Liu

Informatics Research Center, Henley Business School, University of Reading, Reading, RG6 6UD, UK.

q.li5@pgr.reading.ac.uk, k.liu@henley.ac.uk

Abstract. Data visualization has been widely utilized in various scenarios in data analytics for business purposes, especially helping novice readers make sense of complex dataset with interactive functions. However, due to an insufficient theoretical understanding of the process of developing interactive functions and visual presentations, interactive data visualization tools often display all available automatic graphing functions in front of users, instead of guiding them to choose a visualization based on their demands. Thus, this paper is intended to construct a process of developing interactive visualization with a specific focus on enabling the interoperation between design and interpretation. Stemmed from organizational semiotics, an abductive process will be portrayed in this paper to interpret the process of developing interactive data visualization. Especially the interactive functions will be employed in an iterative process, where producers can be aware of and respond to readers' information demands on semantic, pragmatic and social levels.

Keywords: Interactive Data Visualization, Organizational Semiotics, Abductive Reasoning Process.

1 Introduction

Data visualization plays an important role in business analytic intelligence, in terms of helping users make sense of a large amount of data, by using visual aids in order to convey a message to its readers. [1]. Other scholars further extend the scope of contributions of data visualization to communication between readers and producers, exploration of the complex dataset and making sense of the information demanded decision-making [2, 3]. With the development in-memory computing and cloud techniques, data visualization can be more agile to adapt to users' demands, in other words, quickly responding to users' requests by the embedded interactive functions. In addition, the collaboration among different users will be promoted based on web-based visualization application, where different views from multiple users, producers and experts can be incorporated into the process of developing data visualization[4]. In short, other than

supporting individual's exploration and sense-making of the dataset, data visualization facilitates the communication by interactive functions.

However, according to the observation on the leading visualization tools, such as Tableau, QlikView, and PowerBI, although 'interactive visualization exploration' has been listed as a critical capacity of business intelligence, the focus remains on the generation of various visual representations with automatic graphing and enabling users to analyze and manipulate data by interacting with the visual representations [5]. This observation is also echoed by the prior research, which points out that the concept of interactive data visualization remains unclearly defined and its development process is still vaguely portrayed, even though the diverse technique is available and able to support users to interact with data [6, 7].

This paper will construct a process for developing interactive data visualization with a specific focus on understanding readers' multi-levels of information demands and guiding the producers to fulfil the demands by using interactive functions. Organizational semiotics, the doctrine of sign research, which has been applied to various prior research for understanding the process of information transfer among different parties, will be utilized at the theoretical foundation to the understanding of levels of interpretation. Also, the logical reasoning process will be referred to, to explain how a visualization will be interpreted and to discover the key interactions demanded during this process. In the illustrative case study, the abduction process will be applied to help design a data visualization of market attractiveness analysis.

2 Abduction in Organizational Semiotics

Data visualization can be articulated as a process of communication with graphics means [1, 8, 9]. Semiotic is a theoretical ground of signs and signification. It can help interpret the process where a sign as a carrier to deliver information among different parties. It also guides the discovery of implicit and explicit factors impact the efficacy of information transferring [10]. By having an in-depth understanding of the process and the significant influencing factors, the producers can further work on improving the efficacy of communication, e.g. the right information can be communicated at the right time, by the right method and to the right people. Supported by organizational semiotics, the research can focus on application and usefulness of signs in a business context, where the communication among individuals and business objects are driven by business purposes, serving for business objectives and influenced by organizational environments [11].

2.1 Semiotics

Semiosis reveals the process of sense-making, where an individual understands a sign by interpreting it based on the link with a certain object [12]. It is a universal mechanism which can be utilized to all sign-processing activities, which helps people recognize the importance of creating and using signs. Interactive data visualization can be regarded

as a typical sign-based communication, where visual representations act as signs to facilitate the communication between producers and readers.

The whole process of semiosis can be articulated into the following triangular model (**Error! Reference source not found.**). The firstness is a sign or representation which is utilized as a sign vehicle linking to a secondness. The secondness is an object, which should be reflected by the sign in the firstness. However, the reflection might not be generic and spontaneous, where readers cannot perfectly receive the information sent by producers without any deviation. Instead, the reflection will be impacted by the readers' interpretation based on prior knowledge, various purposes of interpretation and pressures from the organisational and social environment.

In the context of interactive data visualization, the meaning of the three elements in semiosis framework can be further expanded (**Error! Reference source not found.**):

Table 1. Elements in Semiosis in the Context of Visualization

| Elements [11] | Explanations in the context of visualization |
|---|--|
| The sign, which is considered as a signifier | Visual representations, including a diagram, chart, map and table [13] |
| The object, which is considered as the signified | Business reality reflected or implied by the visual representations, e.g. market size; sales trend |
| The interpretant, which is considered as the effect of signs on readers' action (incl. reading, interpreting and behaving upon) | A process and result of interpreting signs and identifying their reflection based on readers' subjective elements [14] e.g. knowledge, experience and perception of environmental pressures e.g. driven by the sales-oriented strategy, managers will focus more on the information related to current and potential sales |

Even though the semiosis portrays a general framework for discovering the visualization process where readers make sense of visual representation, the interpretant can be explained further, especially identifying the factors influencing interpretant on both technical and social aspects.

2.2 Semiotic Ladder

Interpretant has a broader scope than interpretation, which covers not only signifying a sign and identifying the meaning associated with the sign, but also involving readers' background knowledge, intentions and influences (incl. support and restrain) from social norms [15]. Thus, the semiotic ladder offers a framework of taxonomy to categorize the various influencing factors towards interpretant to 6 levels. By understanding the concepts and characteristics of different levels, visualization producers can have an in-depth understanding in terms of the barriers which hinder readers from making sense of visual representations.

Stemmed from the theory of organizational semiotics, which suggests understanding the barriers hindering the communication in the context of business through the lens of semiotics, Stamper [10] suggests analyzing the sign effect through 6 levels, consisting of two aspects of human information function and IT platform. IT platform is closely related to the infrastructure, physical quality and structure of sign. Different from the traditional semiotic framework which mostly focuses on the meaning and interpretation, Stamper points out that the physical quality and construction of sign will impact on humans understanding of sign. On the aspect of human information function, the semiotic framework is intended to address the challenges on signifying signs in terms of transferring their meaning, fulfilling readers' intentions and responding to the social norms [12].

When it comes to interactive data visualization, the lower three layers encourages producers to incorporate the Gestalt Law and pre-attentive attributes into visualization design, in order to assist human brain perceptible system to visually identify the patterns e.g. size, proximity and colours [16]. On the upper three layers of the semiotic framework, the focus shifts from visual representation (signs) to interpretant of visual representation (sign effect). As it is implied from the comment 'featureless data is equivalent to noise' [17], there is a big challenge on the cognition aspect of interactive data visualization: to enable users to capture the pattern of dataset, to make sense of them based on their background knowledge, intentions and to cope with social pressure. Since this paper mainly focuses on the sense-making aspect of interactive data visualization, the process framework will focus more on the key questions and norms on the upper three layers (Table 2). However, the semiotic framework might have offered a comprehensive guideline for producers to recognize a series of social and technical factors which might affect the interpretation of signs – making sense of visual representations, but it does not offer a set of tangible methods to elicit and document the elements and produce a practical solution.

Table 2. Upper Three Levels of Semiotic Ladder

| Factors | Explanation [12] | In the context of interactive data visualization |
|--------------|--|--|
| Semantics | Meaning indicted by signs: the relationship between signs and objects | Do readers have a statistic or mathematics background to understand the algorithms for data analytics? What factors will readers mainly consider for measuring market? |
| Pragmatics | Intentions of readers to make sense of the dataset | What is the motivation(s) for readers to interpret the visual representations? What is the business purposes? |
| Social World | Context or environment where some factors might impact readers' focus and interpretation of visual representations | Based on what a reader can recognise, what are the major social and environmental factors which might impact on readers' opinions or focus? E.g. corporate strategy, tones from the top. |

2.3 Abduction

Tan and Liu [3] state that the process of developing data visualization can be depicted as a shared semiosis where the visual representation is used as a carrier to facilitate the communication between the producers and readers. Not only is it focusing on the artefact which carries the visual representation in the final stage, but also focusing on the process where a reader interprets the visualization. Also, norm centric activities where norm can be used as a powerful tool to help make producers and document readers' aware of the explicit and implicit demands of the various levels of interpreting.

Thus, this research, inspired by the three principles from Liu [14], is intended to construct a framework for producing data visualization, especially empowering readers to implement abductive reasoning, guiding producers to place interactive functions based on norms and specifying the process of developing data visualization to steps.

The concept of abduction can be traced back to Peirce, 1930s, which can be demonstrated in semiosis where people explore signs with their prior knowledge, spots new (unmatched patterns with their prior knowledge) and refine the prior knowledge by proposing new propositions and hypothesis which might result in further actions, e.g. further discovery by other means [18]. Dubois and Gadde [19] claim that abduction can be used as an approach to push creativity and help the reader form a proposition by making sense of what They have observed and how it differs from their own understanding. Other than abduction, induction and deduction are two other main-stream reasoning processes.

In contrast, deduction encourages people to extract the logical conclusion from the prior theories, to form new hypotheses and propositions, and to test them in the form of empirical study. Induction follows the opposite way compared with deduction [20]. Instead of obtaining knowledge from prior literature, induction will guide people to generalize a theoretical form based on an observation [21].

Different from induction and deduction, the method of abduction supports humans to develop or refine their knowledge by systematizing the creativity and intuition into their logic reasoning process [22]. The factors, such as prior knowledge and context, is also recognized to be influenced by the people's understanding, instead of purely relying on what people can observe in the empirical study [23]. Therefore, the aim of abduction is more than to spot the difference between empirical studies and prior understanding, but also to understand the new phenomenon and re-frame the prior knowledge based on the fresh inputs.

3 Constructing a Process for Developing Interactive Data Visualization

In this research the logical reasoning process of abduction can be depicted as follows, consisting of 5 steps.

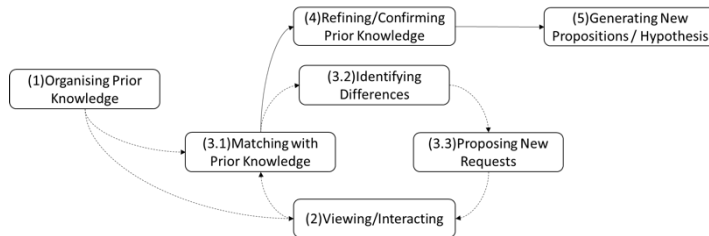


Fig. 1. Abductive Process of Developing Data Visualization

3.1 Step One: Capturing and Organizing Readers' Prior Knowledge

Stemmed from the framework of the semiotic ladder, readers' interpretant of signs can be influenced by their prior knowledge (semantic level), intentions (pragmatic level) and social context (social world level). Also, based on the concept of abduction, readers always intend to observe any phenomenon with a series of hypotheses generated from their prior knowledge. Thus, during the primary stages of developing the concept of data visualization, producers are encouraged to grasp readers' requirements and prior understanding by incorporating six interrogatives [24]. It is only when after these interrogatives have been considered that further investigations can take place as a way of guiding the design of interactive functions.

3.2 Step Two: Viewing the Initial Data Visualization

Based on the information obtained from the previous step and dataset available on hand, the producer can draft the initial version of data representation and present it to the user. The user can then have the initial view of the data representation and try to extract the demanded information. The design of interactive function at this stage is based on the initial input of users' demand and data availability. Thus, the user can explore the dataset based on its initial understanding.

3.3 Step three entering iteration loop

Once readers have found that the information revealed from the initial data visualization is different from their prior understanding, the readers might enter into an iterative process (s3.1, s3.2 and s3.3), in which they can request further questions based on their information demands to interpret the differences.

Based on the observation in step two, readers would compare the information derived from viewing the data visualization with their prior knowledge (s3.1). In other words, they will compare what they have seen from the visualization with what they have understood from their prior experience and identify the differences (s3.2), where they can further address new questions related to data visualization by its interactive functions (s3.3).

3.4 Step Four: Refining the Prior Knowledge

Through continuously addressing different information demands, readers will eventually be able to gain an in-depth understanding of the domain question(s). They can then add the information learnt via the interaction with data visualization into the prior knowledge, and generate a new understanding for the specific domain questions, which can then guide their behaviors.

3.5 Step Five: Generating New Hypothesis

After refining the prior understanding, the readers can further generate a new proposition for their domain questions, which can be further articulated to be a solution for the domain questions they raised up at the very beginning. Also, the readers might generate a further hypothesis which can be tested in the reality or in different scenarios. By this way, they might enter another abductive process by other means to further refine their knowledge.

4 Illustrative Case Study: Market Attractiveness Analysis

In this research, an analysis of global market attractiveness of energy drink industry will be referred to as an illustrative case study. The key domain question raised up by the target readers is to identify the most attractive market(s) to develop a new brand of energy drink, and they expect the delivery to be able to reveal the answer by graphics that they can easily understand. Thus, this case study illustrates all 5 steps of the abductive process of developing interactive data visualization.

4.1 Step one: Capturing and organizing readers' prior knowledge

A semi-structured interview with the target readers takes place at the very beginning of capturing the initial requests of developing data visualization. Information obtained from the interview will be mapped based on the framework of six interrogations, and then migrate to the 3rd-level of the semiotic framework.

4.2 Step two: Viewing the initial data visualization

Based on the information obtained in step one, two bar charts are drawn-up for fulfilling the initial information requests – showing the market with the highest / lowest sales and volume. The advantage of bar charts, is that it enables readers to easily compare the data by bar length and to identify the highest / lowest data by ordering.

After viewing the initial presentation of data visualization, the target readers confirm some of their initial hypothesis, by focusing on countries with a high population, which as a result will allow for an increase in sales and volume. These countries include China, Japan and the United States. However, they also discovered that the information differed from the prior experience. For example, some markets like Brazil and UAE might

not rank high in terms of sales but ranked in the top tier of sales due to the high unit price.

4.3 Step Three: Entering in Iteration Loop

Based on the questions generated from the initial view of data visualization in step two, the target readers enter the iterative loop where they can address further requests for information based on their new hypothesis. A round of interviews took place to allow the target readers to compare the information from data visualization and prior experience, to articulate the specific gaps and to reveal more details of their new requests and hypotheses. During the interview, the target readers expressed the idea of taking more variables from the non-sales aspects into the measurement of market attractiveness, since an attractive market should not merely be identified by the sales data in a short period. Instead, incorporating non-sales data might help reveal a view of long-term market development.

4.4 Step Four: Refining the Prior Knowledge

Once the target readers find the data visualization provides sufficient information for them to justify the prior hypothesis, they decided to end the iteration and tried to refine the prior understanding. The information grasped from the data visualization would be added to the prior knowledge. For example, the target readers had thought to put the focus on the western European market since it seems to be a mature market. However, with the aid of data visualization, they found that the new brand is likely to be launched in Brazil and the Middle East. The reason for this being is that there tends to less competitiveness, which lends itself to more room for a new brand to set up and grow. Also, instead of merely focusing on a single market, Hub-and-Spoke can be considered to apply in Middle Eastern markets since some similarities can be found on data patterns of their energy drink consumption and market competition, e.g. setting UAE as the Hub and gradually expand the brand influence on its neighbor countries (spokes) via a new fashion.

4.5 Step Five: Generating New Hypotheses

During the final step, a workshop took place to finalize the interactive data visualization based on documents of all hypothesis and information requests provided by the target readers. The format of 'context-content-conclusion' has been used for the final presentation, which can demonstrate the key questions and hypothesis (context), the filtered data in a graphic format with interactive function (content), and summary of data pattern (conclusion). The target readers can take the interactive data visualization as an input to their business strategy formation or a document to provoke a discussion of strategic decision.

5 Discussion and Conclusion

This paper portrays an abductive process of developing interactive data visualization, where different mechanisms have been used to facilitate the communication and interoperation between producers and target readers. The information demands from target readers have been analyzed by different levels based on the semiotic framework, including semantic, pragmatic and social levels. Also, the iteration loop allows target readers to continuously address the requests for more information to justify their hypothesis, which can be documented and analysed in the format of the norm and eventually lead the design of interactive functions. In the end, a case study of global market attractiveness analysis has been used to illustrate the proposed process of developing interactive data visualization.

In terms of contributions of this research, it further develops the statement from [3] - visualization as a process of abduction, by demonstrating a detailed process where visualization producers can capture targets' multi-level information demands and elicit them into the design of interactive function. Also, by further developing the idea of enabling the interoperation between producers and readers, the iteration loop enables both two parties to continuously synergize the understanding of information request, targets', purposes, and potential influence from the corporate strategy (social environment).

However, there are two limitations which should be addressed in order to inspire further studies. Firstly, this research does not set criteria to measure the satisfaction of target readers about fulfilling their information demands. Without the criteria, the target readers might be trapped by 'confirmation bias where they think they might have had sufficient information but actually, this is not the case [25]. Therefore, the following research can further work on specifying the criteria for measuring readers' satisfaction of information fulfilment. Secondly, this research does not compare the proposed process with the traditional way of developing visualization from the readers' perspective about the extent in which the new process helped them understand data better than the traditional approach. Thus, a comparative research between the abductive process and non-abductive process of developing interactive data visualization should be conducted to justify the helpfulness of the abductive process.

References

1. Chen M, Ebert D, Hagen H, et al (2009) Data, Information, and Knowledge in Visualization. *Comput Graph Appl IEEE* 29:12–19 . doi: 10.1109/MCG.2009.6
2. Segel E, Heer J (2010) Narrative Visualization: Telling Stories with data. *Vis Comput Graph* 16(6):1139–1148 . doi: 10.1109/TVCG.2010.179
3. Liu K, Tan C (2015) Semiotics in Visualisation. In: 16th International Conference I 2014 (ed). Enterprise Information Systems, Lisbon, Portugal, pp 5–7
4. Tableau (2014) Visual Analysis Best Practices. 41
5. Gartner (2017) Magic quadrant for business intelligence and analytics platforms. Gartner 1–126 . doi: 10.1017/CBO9781107415324.004

6. Shneiderman B (1996) The eyes have it: a task by data type taxonomy for information visualizations. Proc 1996 IEEE Symp Vis Lang 336–343 . doi: 10.1109/VL.1996.545307
7. Strecker J (2012) Data Visualization in Review: Summary. In: Eval. Off. IDRC. <https://idl-bnc.idrc.ca/dspace/bitstream/10625/49286/1/IDL-49286.pdf>
8. SAS (2012) Data Visualization: Making Big Data Approachable and Valuable. Whitepaper, Source IDG Res Serv 4
9. Wang Y, Liu D, Qu H, et al (2016) A Guided Tour of Literature Review: Facilitating Academic Paper Reading with Narrative Visualization. Proc 9th Int Symp Vis Inf Commun Interact - VINCI '16 17–24 . doi: 10.1145/2968220.2968242
10. Stamper RK (2001) Information, Organisation and Technology: Studies in Organisational Semiotics. Springer US, Boston, MA
11. Liu K, Li W (2014) Organisational Semiotics for Business Informatics. Taylor & Francis, UK
12. Stamper R, Liu K, Hafkamp M, Ades Y (2000) Understanding the roles of signs and norms in organizations - a semiotic approach to information systems design. Behav Inf Technol 19:15–27 . doi: 10.1080/014492900118768
13. Friendly M (2006) A Brief History of Data Visualization. Handb Comput Stat Data Vis 1–46 . doi: 10.1007/978-3-540-33037-0_2
14. Liu K (2014) Semiotics in Visualisation. 5–7 . doi: 10.1007/978-3-319-22348-3_1
15. Stamper RK (1973) Information in Business and Administrative Systems. Wiley
16. Tufte ER (2001) Envisioning Information. Graphics Press
17. Chi EH, Riedl JT An Operator Interaction Framework for Visualization Systems 1 Introduction 2 Related Work
18. Kovács G, Spens KM (2005) Abductive reasoning in logistics research. Int J Phys Distrib Logist Manag 35:132–144 . doi: 10.1108/09600030510590318
19. Dubois A, Gadde LE (2002) Systematic combining: An abductive approach to case research. J Bus Res 55:553–560 . doi: 10.1016/S0148-2963(00)00195-8
20. Ho YC (1994) Abduction ? Deduction ? Induction ? Is there a Logic of Exploratory Data Analysis ? Proc Annu Meet Am Educ Res Assoc 28 . doi: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.135.3507&rep=rep1&type=pdf>
21. Sowa JF (2000) Knowledge Representation: Logical, Philosophical, and Computational Foundations. 594
22. Kolko J (2010) Abductive Thinking and Sensemaking: The Drivers of Design Synthesis. Des Issues 26:15–28 . doi: 10.1162/desi.2010.26.1.15
23. Arrighi C, Ferrario R (2005) Abductive Reasoning , Interpretation and Collaborative Processes by. 1–16
24. Tan C, Liu K (2013) An organisational semiotics inspired information architecture: Pervasive healthcare as a case study. ICISO 2013 - Proc 14th Int Conf Informatics Semiot Organ IFIP WG81 Work Conf 35–44
25. Kodagoda N, Attfield S, Wong BL, et al (2013) Using Interactive Visual Reasoning to Support Sense-Making: Implications for Design. IEEE Trans Vis Comput Graph 19:2217–2226 . doi: <http://dx.doi.org/10.1109/TVCG.2013.211>