

Evolution of Automation Usage for the Data Visualization User Experience

Marc-Olivier Tricot
Human-IST Institute
University of Fribourg, Switzerland
marc-olivier.tricot@unifr.ch
April 2018

ABSTRACT

This paper describes how the 2007 article *Show Me* [1] uses automation to enhance the user experience for the creation of data visualizations. It then compares the usage done in *Show Me* to three other papers at different periods in time. Using the review, it aims to discuss the evolution of the usage of automation for the improvement of the user experience in data visualization through time.

1 INTRODUCTION

Data visualization is a very important field nowadays, and the ever-increasing prevalence of Big Data makes it necessary to constantly improve and refine the way of presenting the data in the fields of business and multimedia [2].

As more and more ways of presenting information are discovered, the complexity of creating relevant graphs keeps increasing too, with increasingly more professionalism being required to make them. To offer the possibility to more people to make interesting graphs, we think there should be an effort on the technical side to let the user exploit a lot of variety without requiring years of studying the given software.

In this paper, we will study the efforts made to enhance user experience for data visualization software by the means of automation. We were inspired by the paper *Show Me: Automatic Presentation for Visual Analysis*, published in 2007 by Jock D. Mackinlay, Pat Hanrahan and Chris Stolte [1]. We will afterwards do a comparative review of three other publications on the same subject. And finally, we will comment on the evolution of the usage of automation to enhance user experience through the years.

2 BACKGROUND

Tableau and Data Visualization in Business

Tableau, sold by Tableau Software [3], is an interactive tool the aim of which is to aid its users to create graphs from data inside relational databases, OLAP cubes, cloud databases and spreadsheets. It is present in different forms, such as Desktop, Server, Online, Vizable, Reader and Public. It is now a big name in the market of data visualization, next to competition

like IBM Corporation, SAP SE, Microsoft Corporation, Oracle Corporation, Qlik Software, etc.

Data visualization is a growing field, as we can see with the estimated market capitalization of Tableau Software being worth over 7 billion USD early 2018 [4]. With Big Data and the ever-growing access to more information that companies tend to have, supported by the growth of artificial intelligence and even more different kinds of ways of tracking information, data visualization's importance keeps growing. All the data in the world is useless, or even a liability, if you can not understand and make use of it.

Show Me and Automation of Data Visualization

With higher amounts of information comes not only more ways of presenting it, but also more work to create higher quality graphs. Deciding how to create the graph is as important as the data itself. Knowing what kind of graph is most appropriate, but also the shapes, colours and other elements of formatting are what make the difference between a clear or confusing graph [5].

Show Me is an integrated set of user interface commands and defaults that incorporates automatic presentation into the commercial visual analysis system Tableau [1]. It was created as a way to improve user experience when creating graphs, both from an ergonomic point of view and from a time-based one.

Show Me applies the features of VizQL, an algebraic specification language, to automatically present data as a table of views in a way that takes into account both new and skilled users. The *Show Me* paper expands into the research made by Mackinlay et al. to use automatic presentation and defaults to offer the best type of experience to both new and skilled users.

3 SHOW ME'S ADVANTAGES FOR THE PREPARATION OF DATA VISUALIZATION

User Benefits of Show Me

User experience, as defined by ISO 9241-210 [6], is "a person's perceptions and responses that result from the use or anticipated use of a product, system or service".

The work done for *Show Me* expands upon the requirements that the users have for data visualization programs by adding two key contributions: the design of an interactive user experience through automatic presentation functionality and the generation of small multiple views.

Small multiple views are useful for the users representation of how the final product will look, and an automatic generation of those views as examples is a very important way for the user to have faster access to a representative view.

In this way, *Show Me* participated in the current success of Tableau software.

Using Automation to Improve the User Experience

Work about using automation to create better graphs has already been done before *Show Me*. However, using automation as a way of enhancing user experience has not. The automation is now used as a form of support to the user choices, instead of replacing it completely. It moves the aim of the artificial assistant to supporting the flow of visual analysis. A previous work on the subject is that of Roth et al. [7], which only partially addresses the subject.

Different automatic elements of *Show Me* play a role in the **Add to Sheet** command, which adds a single field of data to a view. One such automatic element is *affinity*, which helps to decide what kind of graphs will be suggested with the creation of the new small view. The affinity heuristic is also applied in other cases, such as when adding a new category to small views that already have category fields (for example adding a Country dimension in one that already has a Continent one).

Other automatic elements mentioned in *Show Me* are encoding requirements. From the **Add to Sheet** command, one can find requirements such as Color, Size and Shape, but manually choosing the best options is difficult. Tableau lets the user choose between 20 distinguishable colors, and 10 different shapes which can overlap without losing too much clarity.

By automating the best practices, the user does not need to worry about the visibility of his information. In addition, the calculations are faster and let the user see in minimal times the effects of his choices, so that they are easily changeable if not to his liking. These are the main ways automation improves user experience within *Show Me*.

Using Defaults to Improve the User Experience

The two most important commands are **Show Me**, and **Show Me Alternatives**. These commands display the choices for creating a new chart. These choices are exposed in a grid menu filled in a specific order.

Show Me is a command aimed towards beginner users that brings them a menu filled solely with the six default

choices, while **Show Me Alternatives** offers all possible charts for the more experienced or adventurous users.

The displayed options are chosen according to the disposition and content of the data itself, with higher priority defaults taking precedence over lower ranking ones. Each type of table has some requirements in terms of data needed. The six default choices are as follows, in increasing order of priority:

- At priority level 1, Text Tables are the lowest rank because their primary value is to look up specific values. However, it is very easy to meet the prerequisite; any one field is enough. They can handle a large number of fields and will always be present in the defaults.
- At priority level 2, Aligned Bars require only one Quantitative field to be chosen, because of their effectiveness at conveying information when comparing bar length.
- At priority level 3, Stacked Bars require at least two Category fields and a Quantitative field. This type of graph boasts high legibility and easy-to-fill conditions, and also appears often.
- At priority level 4, Discrete Lines requires at least one Quantitative field, and at least one Date type Category. For showing trends through time, line views are seen to be more efficient.
- At priority level 5, Scatter Plots are chosen if there are between 2 and 4 Quantitative fields, as their strong point is to compare large amounts of values.
- At priority level 6, Gantt Charts are very specialized and therefore perform better for appropriate data, but are rarely available, with at least one Categorical field, 1 to 2 Quantitative fields, and at least one Independent Quantitative field being required.

With **Show Me Alternatives**, the user can access other charts that usually never appear in the defaults because of their limited uses, but can bring interesting ways of presenting data. These charts are Heat Maps, Highlight Tables, Side-by-side Bars, Measure Bars, Continuous Lines, Circle Charts, Scatter Matrices and Histograms. Usually reserved for more advanced users, these charts are all specialized for certain cases, with some of them being only small modifications of default charts.

We can see here how the defaults improve the user experience, by offering an automated standard choice that not only covers most usual cases, but is ranked and filtered depending on the given data. This reduces the choice for the users, and helps them get a better result by being able to choose relevant representations according to their data. Moreover, skilled users can easily gain access to specialized choices they know best fit their data .

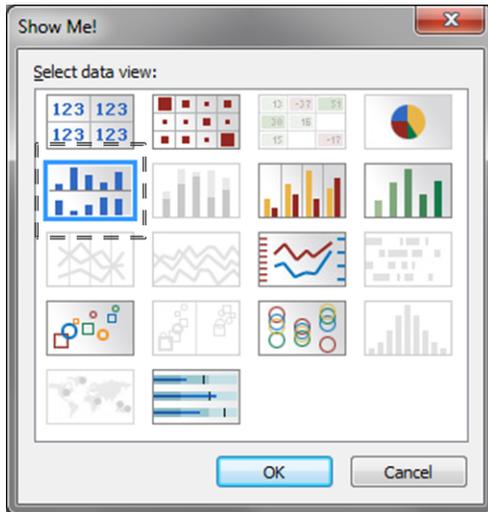


Figure 1: Show Me Button (2011) [8]

Show Me : Evaluation

In the time subsequent to the introduction of *Show Me*, Mackinlay et al. used the feedback both directly from users and from the logs to review the efficiency of their product. From the general positives results, it was decided not to immediately change their work much.

Using the interface logs, there was some research done on identified skilled users, to check if it indeed answered to the level of professionalism required. Some analysis of the log creation file supported the idea that **Add to Sheet** didn't have much impact on skilled users. Yet both **Show Me** and **Show Me Alternatives** had a reasonable amount of results in the logs, and their spread in time suggested regular use by those skilled users.

However, in addition to more types of charts, it was later decided to fuse **Show Me Alternatives** into **Show Me**. In this new version, as it can be seen in Figure 1, all the charts are displayed in a certain order. But those who do not fit the data are faded out, and one kind gets recommended using the automation system previously detailed.

4 COMPARATIVE REVIEW

Now that we have seen how **Show Me** and **Add to Sheet** are used in Tableau Software to improve the user experience by automation, it is interesting to see how other implementations are done in order to attain the same objective. We have found three relevant examples, an older work by Roth et al. on automatic presentation [7], then an article on a conversational interface by Sun et al. [9], and a newer paper by Sakhet et al. about visualization by demonstration [10].

Interactive Graphic Design

Interactive Graphic Design Using Automatic Presentation Knowledge is a 1994 article by Roth et al. [7] where they present three novel tools (for the time) for creating data graphics.

Those three tools are as following:

- SageBrush, a tool for assembling graphics from separate objects such as axes, lines and bars.
- SageBook, a tool for flipping through previously used graphs relevant to the current work.
- SAGE, a knowledge-based presentation system that systematically interprets the users needs and conveys the information to other tools that automatically create the needed graphs.

Those tools are put together in order to enhance user-assisted design through automation of tasks such as retrieval of previous graphs relevant to the current work according to the data content and appearance, creating novel displays of partial to entire completeness and also showing possible alternatives when asked to by the user.

This paper contains some core elements that were used in *Show Me*, especially on the automation side, such as the differing alternatives or using the relevancy of previous graphs. However, where *Show Me* spent time concentrating on the effects of their work for the user experience, Roth et al. spend more time detailing the technical effects of the automation.

Their work on integrative displays, knowledge-based systems, partial completion and relevant graph comparison shows the great concern given at the time on the use of automation and on a newer form of graph creation. They address themselves this fact in their conclusion, where it is mentioned that there is no ability for users to specify their information-seeking goals or tasks, nor is there an easy and interactive way for users to manipulate data displays, both of which penalize beginner users. This is addressed in *Show Me*, when a shift towards user-friendliness appeared.

Articulate

Articulate: a Conversational Interface for Visual Analytics is a 2009 article by Sun et al. [9] that aims to ease the access to data visualization. Their particular way of doing so is creating *Articulate*, a semi-automated visual analytic model that is guided by a conversational interface. The aim was to make it simple to create graphs by replacing the need to interact with complex user interfaces.

Instead of having to choose from a large amount of options, *Articulate* would understand vocal commands and process them to create a visualization corresponding to the user's needs. Figure 2 details the system model of the software.

The software uses **ConVL**, a high-level conversational language based on first-order predicate logic, with additional

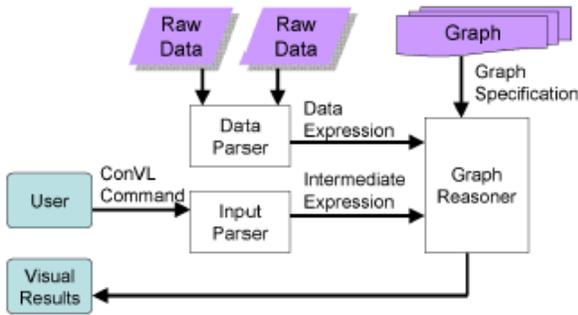


Figure 2: System Model for *Articulate* [9]

constraints added to attain the simplicity and flexibility required for a conversational language.

The input parser translates the **ConVL** command into a smaller instructions which are used as the basis for the end visual product. The data parser then reads the original data file to find its properties and sends the relevant data to the graph reasoner. The reasoner is the core part of the semi-automated generation, as it treats the information sent to it and makes the appropriate links in order to generate the relevant visualization.

From this article, we can see the advance towards the democratization of the visualization creation experience, a domain increasingly worked on [11].

In order to follow the general trend of improving user experience, there is a lot of work done towards creating a more natural way of interacting with software in general, and it affects data visualization too. While the automation takes a different form than the other cases, it remains nonetheless a valid way to improve the user experience, in particular for beginner and casual users.

Visualization by Demonstration

Visualization by Demonstration: An Interaction Paradigm for Visual Data Exploration is a 2017 paper by Sakhét et al.

In this article, the authors present *Visualization by Demonstration*, a user interaction paradigm which advocates for the mapping of visual demonstrations provided by the users, and from them, recommending new, relevant examples or transformations of the provided ones. To make that possible, they implemented a prototype named VisExemplar to show the feasibility of their idea. The three main visual elements of VisExemplar are the **ThinkBoard**, the **Detail View** and the **Recommendation Gallery**.

When using VisExemplar, the given data is distributed on a relevant, automatically decided representation, such as a scatter plot, on the **ThinkBoard**. From there, the user can individually select the data which are interesting, and see the details on the **Detail View**. The user can interact with

the data on the **ThinkBoard**, being able to move, resize or recolor them so that they make more sense for them.

From the way the data are modified, VisExemplar will provide recommendations of data visualizations in the **Recommendation Gallery**, according to algorithms that calculate relevant comparison elements so the graphs make the most sense. There are also possibilities of changing a high number of elements in the original plots according to those same elements, such as recoloring all elements within a certain proximity of the chosen data, or any of higher or lower value than the selected one.

In this work, we can see a new, modern way of making the data visualization user experience even more user-friendly. By changing the standard method of graph creation and taking advantage of the effectiveness of letting people create spatial representations manually [12], we obtain a better experience for the user.

The automation is still very much used to help the users through their data visualization creation, but it changes form again to let the user make the choices through direct manipulation of the data.

5 CONCLUSION

While not a new field of work, the use of automation to improve the user experience is still a living field with a lot of possibilities. From the first, mostly mechanical works on the subject, the focus has slowly moved from the automation as a replacement of the user to the automation as a complement to the user, as shown in *Show Me*.

At the same time, following a general trend in the field of improving the user experience of software, differing ways of interacting with the users gradually appeared, outside of the original mouse and keyboard experience. Modular interfaces, still supported by automation that assists user choices, can be seen to appear in the late 2000s.

As changes were made to the concept of user-assisted creation, we can see a gradual shift from software centered around skilled users to beginner users, with an intermediate state shown by a number of mixed-initiative systems [13] which balance responsibility equally. As part of a new paradigm seen in the late 2010s, user-inferred design supports letting the user manipulate the data directly as seen in the work of Sakhét et al. [10].

From there, even after all the work done, we can still see that there is more to this field, for example taking inspiration from other industries (such as graphic design [14]), and while automation has proven its usefulness for improving user experience thoroughly, its shape still continues shifting through time, and it will be interesting to see what the newest paradigm will be.

ACKNOWLEDGEMENTS

This paper is a student work, written for Seminar SS2018 "Data Visualization", under the supervision of Prof. Denis Lalanne and Florian Evequoz. Human-IST Institute, University of Fribourg, Switzerland, 2018.

REFERENCES

- [1] Chris Stolte Jock D. Mackinlay, Pat Hanrahan. Show me : Automatic presentation for visual analysis. *IEEE Transactions on Visualization and Computer Graphics*, pages 1137–1144, 2007.
- [2] Data visualization: What it is and why it matters, April 2018. URL: https://www.sas.com/en_id/insights/big-data/data-visualization.html#dmimportance.
- [3] Business intelligence and analytics | tableau software, April 2018. URL: <https://www.tableau.com>.
- [4] Tableau software (nyse:data) - share price, news & analysis, April 2018. URL: https://simplywall.st/stocks/us/software/nyse-data/tableau-software?unique_symbol=NYSE:DATA.
- [5] Seven elements of good data visualization - extension, April 2018. URL: <https://www.extension.org/2017/04/11/7-elements-of-good-data-visualization/>.
- [6] Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. Standard, International Organization for Standardization, 2010.
- [7] Joe Mattis Jade Goldstein Steven F. Roth, John Kolojejchick. Interactive graphic design using automatic presentation knowledge. *Proceedings of the SIGCHI*, pages 112–117, 1994.
- [8] Create dual axis chart in tableau, April 2018. URL: <https://danjharrington.wordpress.com/2011/08/04/create-dual-axis-chart-in-tableau/>.
- [9] Andrew Johnson Dennis Chau Yiwen Sun, Jason Leigh. Articulate: a conversational interface for visual analytics. IEEE Symposium on Visual Analytics Science and Technology, 2009.
- [10] Eli T. Brown Alex Endert Bahador Sakhet, Hannah Kim. Visualization by demonstration: An interaction paradigm for visual data exploration. *IEEE Transactions on Visualization and Computer Graphics*, pages 331–340, 2017.
- [11] C. Gilpin D. Ebert, K. Gaither. Enabling science discoveries through visual exploration. NSF Workshop Report, 2007.
- [12] A. Endert C. Andrews and C. North. Space to think: large high resolution displays for sensemaking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 55–64, 2010.
- [13] E. Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 159–166, 1999.
- [14] Alan Wilson Mira Dontcheva James Delorey Sam Grigg Bernard Kerr Zhicheng Liu, John Thompson and John Stasko. Data illustrator: Augmenting vector design tools with lazy data binding for expressive visualization authoring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2018.