TOOLBOX

Measuring dashboard performance

Optimizing the view on data

Due to the recent technological advancements in data collection, transmission and storage, the amount of data that is available in private or publically accessible databases is growing exponentially. In principle this data may enable individuals and organizations to make well-informed decisions and timely adapt to changing conditions. However, as datasets increase in size and complexity, it becomes more and more difficult to explore the data, select the relevant information, perceive patterns and interpret the data correctly to make the right decisions. Efficient and effective information visualization tools that allow a user to explore and understand the data in an intuitive manner may serve to achieve this goal. Dashboards are promising candidates for this purpose. Dashboards are graphical user interfaces consisting of different components, that organize and present information in a way that is supposedly easy to read and comprehend. The overall quality of dashboards depends on the quality of their components and the synergy between them. Because of their inherent complexity, determining the overall quality of dashboards is difficult. We are currently developing a framework to evaluate and optimize the performance of dashboards. Such a framework will enable the design of efficient and effective dashboards that provide users with an intuitive view on data.

Ward Venrooij, Alexander Toet and Jan van Erp

Dashboards as windows on data

Dashboards are defined as 'a visual display of the most important information needed to achieve one or more objectives that has been consolidated on a single computer screen so that it can be monitored and understood at a glance' (Few, 2004; Figure 1). Hence, to meet the definition of a dashboard given above, a visual information display should:

1. fit on a single screen;
2. contain the most important information needed to achieve one or more objectives, and
3. enable the user to monitor and understand the displayed information at a glance.

Criterion 1 can be regarded as a basic dashboard requirement and a prerequisite for the other two criteria. Visualizations that either require scrolling or divide the information over a set of pages do not enable a user to monitor and understand displayed information at a glance. Links provided by the dashboard to specific details in other documents can be considered as an exception as long as these specific details do not play an essential role achieving the main user objectives. Tabs are generally allowed when they provide additional functionality and when they contain secondary (background) information that is not essential for obtaining an initial situational awareness.

Criterion 2 can be regarded as the potential value the dashboard can offer to its users. The functionality and presented information on a dashboard should support the objectives of the users. The potential of the dashboard is expressed as the degree to which the dashboard truly supports the user's goals. For example: users who need to make decisions on team level, require information and functionality related to teams instead of information available on individual level. This degree can be expressed in a percentage between 0-100%. If a dashboard contains all the functionality and information which is required to achieve all the user's objectives, its potential is 100%. In that case, the user can theoretically achieve all the goals by using the functionality and information offered by the dashboard. In practice, this percentage can be judged by the users or anyone who has access to the dashboard itself and its functional specification document. However,

Figure 1. Example dashboard representing a conceptual model of cognitive workload of soldiers during missions.

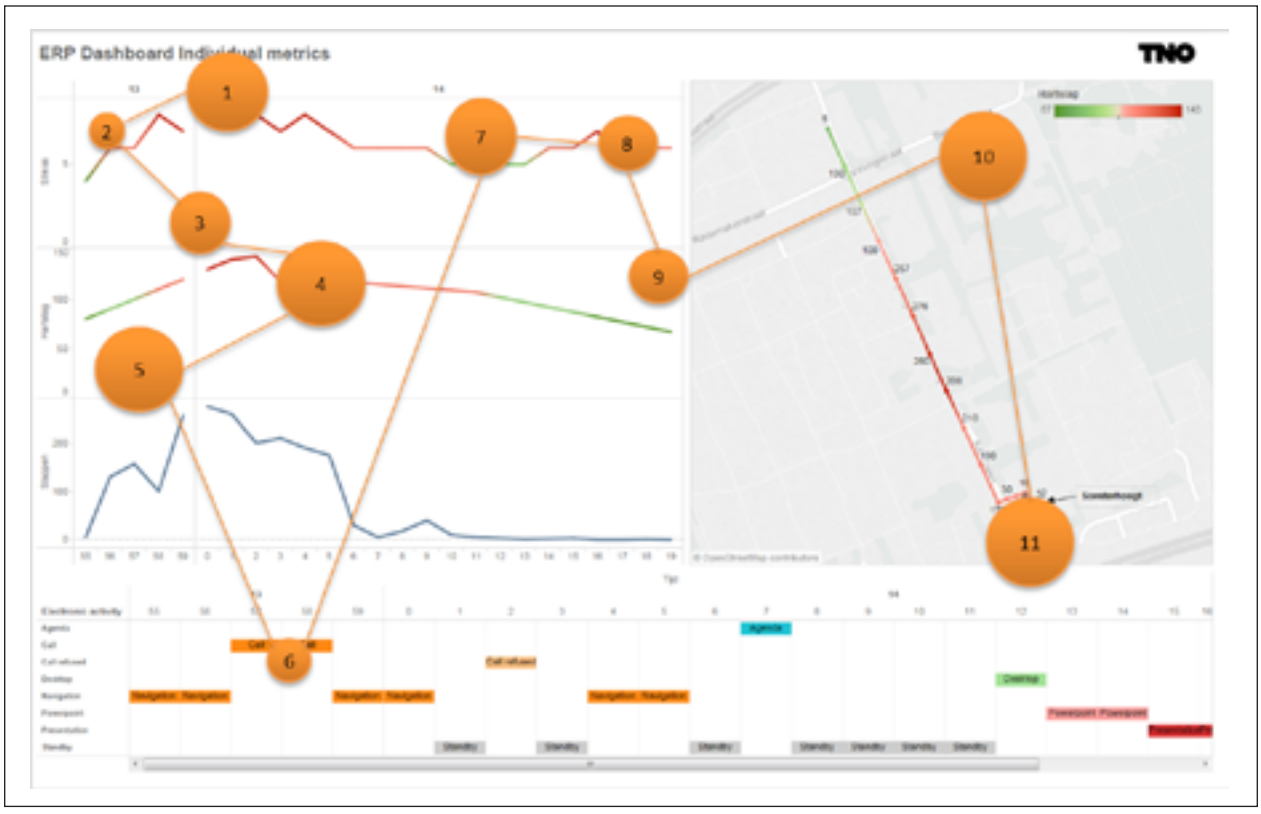

Figure 2. An example of a dashboard with eye-tracking data gathered during an experiment. The dashboard combines several individual health metrics like stress, hart rate, activity and physical location.

Figure 3. Left: an example of the overuse of both color and alerts. Right: an example of a better design (example Few, 2013).

the potential does not provide information on how easy the objectives of the user can be met. This will be the focus of criterion three.

Criterion 3 focusses on the degree to which the potential of the dashboard is realized in practice. A dashboard may have a high potential because it contains all the necessary functionality and information, but still fail to communicate the information to the user due to bad interface design. Criterion 3 deals with the ability of a dashboard to communicate information to the user and to support the user in monitoring and understanding the displayed information at a glance. The realization of a dashboard's potential can be expressed in several performance scores, such as effectiveness and efficiency. Effectiveness expresses the degree to which a dashboard communicates its message correctly. Effectiveness can for example be derived from the number of interpretation mistakes. Efficiency expresses the amount of effort required to perceive (cognitively process) the information presented by a dashboard. Efficiency scores can for instance be computed from observer response times and visual scan paths.

Figure 4. Example dashboard concerning the psychosocial well-being of selected groups during disasters (OPSIC, 2016).

Table 1. Three essential criteria that any dashboard should meet.

Criterion	Measures
Criterion 1 Does it fit on a single screen?	This criterion is regarded as a basic criterion. Visualizations that either require scrolling or divide the information over a set of pages do not enable a user to monitor and understand displayed information at a glance.
Criterion 2 Does it contain the most important information needed to achieve one or more objectives?	This criterion can be regarded as the potential of a dashboard which varies between 0 and 100%. It can be measured by evaluating whether a dashboard contains the required information to enable the user to meet his or her goals.
Criterion 3 Does it enable the user to monitor and understand the displayed information at a glance?	Criterion 3 focusses on the realisation of the potential of the dashboard. This can be expressed in both efficiency and effectiveness measures.

Table 2. Dashboard design criteria and measures that can be used to quantify the extent to which a design meets these criteria.

Criterion	Measures
1. Fit on a single screen.	Can be assessed by simple human judgement. It should be considered which type of screen is being used mostly to view the dashboard. Currently dashboards can be viewed on several devices like mobiles, tablets, computers, televisions and beamers. It should be assessed if the necessary information fits on one screen and is still readable.
2. Contain the most important information needed to achieve one or more objectives.	Digital questionnaires to gather functional requirements and evaluate whether the user objectives are met.
3. Enable the user to monitor and understand the displayed information at a glance.	Digital questionnaires to evaluate user experience. These questionnaires can quantify the effectiveness of a dashboard by measuring how much information users extract from the dashboard and how well users understand the information. Eye tracking data (figure 2) can serve to quantify the efficiency of a dashboard: an efficient dashboard design will result in less eye movements (e.g., when the information can be extracted in a few glances) than an inefficient design (e.g., when users need to go back and forth between individual dashboard elements). Ideally, the dashboard layout should help the user to follow the most efficient scan path.

Towards the optimal dashboard

Today, a steadily increasing amount of literature is available to assist dashboard designers in their quest to develop effective and efficient dashboards. Designers can use insights from various fields like visual analytics, usability and human centered design to increase the quality of their dashboards. In the current body of literature several rules and guidelines have been presented for the design of efficient and effective visual components like charts and tables. Some examples mentioned in Few (2013) are:

- use line charts to show changes over time;
- use bar charts for comparisons between categories;
- use alerts to focus attention on the most important information (Figure 3);
- do not overuse color (Figure 3);
- avoid pie charts.

Applying these guidelines will help designers to improve the performance of single visual means, like a chart or a table. However, dashboards often combine various visual means to deliver the insights required to make decisions. As a result, dashboard performance cannot be considered as a simple function of the performance of their individual components. It also depends on the degree to which information displayed by these subcomponents has to be combined to gain the required insights. So in addition to the quality of the individual elements, it is also the synergy between them that determines the performance of a dashboard. Therefore, we need better guidelines: simply applying the guidelines mentioned above to individual visual means does not guarantee that their combination will result in an effective and efficient dashboard. Hence, based on the three criteria and the initial guidelines given above, we explored a way to measure dashboard performance in a consistent way. The ability to measure dashboard performance will provide us with more insight in both their strong points and their weaknesses, which enables us in turn to create a new set of guidelines that can be applied in designing dashboards. In order to create these guidelines we are currently developing a set of measures to evaluate the performance of dashboards as a whole. Table 1 presents an overview of the different classes of measures and their relation to the dashboard design criteria.

The first criterion deals with the fit on one screen and can be easily to assessed by simple human judgement. The type of screen which is being used mostly to view the dashboard should be taken into account.
To evaluate the second criterion, digital questionnaires can be used to check if all functional requirements and user objectives are met. The dashboard effectiveness mentioned in the third criterion can be evaluated through digital questionnaires. Finally, dashboard efficiency can be deduced from eye movement data. All methods are common in Human Factors research and hence require specialists in this domain to contribute to the development of effective dashboards for (big) data services.
Currently we are developing standardized questionnaires and eye movement metrics to support the Human Factors community in their work in big data dashboard development. Evaluating the performance of dashboards on their most important criteria, taking into account the synergy between their visual components in a consistent and coherent way, will help to create a set of guidelines for improving future dashboard performance.

References

Few, S. (2004). Dashboard Confusion, Intelligent Enterprise.
Few, S. (2013). Information Dashboard Design, Analytics Press.
OPSIC (2016). Available at: http://opsic.eu/.

About the authors

W. Venrooij, Msc
Research Scientist in Data Visualization, Department Training and Performance Innovations
TNO, Soesterberg
ward.venrooij@tno.nl

A. Toet, PhD
Senior Scientist
Department of Perceptual and Cognitive Systems
TNO, Soesterberg

Prof. J.B.F. van Erp, PhD
Principal Scientist
Department of Perceptual and Cognitive Systems
TNO, Soesterberg
Professor of tangible user interaction
University of Twente, Enschede