# Collaborative Visualizations of Self-impacted Data

**Pierre Vanhulst** **Denis Lalanne**

Human-IST Research Center
University of Fribourg
Switzerland

**ABSTRACT**

This paper presents a literature review led with two purposes: understanding the impact of collaboration in visualization and analyzing the current lack of engagement of end-users with their personal data. Based on this review, we propose a few trails to follow in order to design efficient collaborative visualization systems aiming for behaviour change. We conclude by highlighting which are the unresolved challenges that will require further researches.

**Keywords**: Collaborative visualization, personal visualization.

**Index Terms**: • Human-centered computing~Visual analytics • Human-centered computing~Information visualization • Human-centered computing~Visualization systems and tools.

## 1 INTRODUCTION

The blooming of the smart meters' market allowed pervasive computing to become reality. We are now facing a crucial question that should have been answered earlier: how will those tools improve our life? There are basically two schools of thought when it comes to this: those who automate changes by machine learning and those who inform users so that they understand their personal behaviours and their consequences. While the first category requires less commitment from the users, the second allows them to reflect and understand how their daily routines affect them [6]. This realization is desirable, as it makes people less dependent on machines: humans learn and adapt their behaviours based on their personal data.

In our study, we defined the type of data gathered by quantified self devices as "**self-impacted**", stressing the fact that they are not only personal, but also directly influenced by the people who produce them. Thus, "Self-impacted data" share the four following properties: they are *personal*, *continuous*, *actionable* and generally *invisible* (for example, end users usually have no idea how much energy they consume for a given household task). By opposition to quantified self data, self-impacted data also include dimensions that are not directly produced or absorbed by humans such as fuel or electricity consumption.

Using visualizations to make sense of large amounts of data is a popular research topic. While most studies focus on single users, some [12] regard "sensemaking" not only as a cognitive process, but as a social one too: under certain circumstances, groups working on a visualization may achieve better results than single individuals. This specific research area is at the crossroads of two others, "Computer Supported Collaborative Work" (CSCW) and "Information Visualization" (InfoVis). At the other side of the spectrum, "Personal Visualization & Personal Visual Analytics" (PV&PVA) have recently been defined as new fields of research [13]. They encompass all the usages of visualizations in personal contexts for various purposes. Finally, "persuasive technologies" and "gamification" were also reviewed in an attempt to understand what motivates users to engage with such systems.

In this paper, we report a literature review of several research papers that will eventually lead to the design of a collaborative visualization system of self-impacted data. Based on this review, we also identify the challenges that stand in the way and provide guidelines to support the design of such systems.

## 2 WORK HYPOTHESES

At the origin of our research lies the assumption that collaborating while analysing a visual representation of self-impacted data would foster more relevant findings and ideas for behaviour change.

Hyp. #1: Collaborating on a shared visualization allows users to make more relevant findings.

Acquiring a relevant new behaviour, and sticking to it, is a difficult task. Studies in psychology field [3] argue that behaviour changes require years before being assimilated, and that extinct behaviours are not erased but might resurge whenever a change of context occurs. On the top of this, we assumed that users would not engage easily with visualization systems. Thus, our second hypothesis stresses the need for advanced incentives to make people use the systems and to sustain wanted behaviour changes.

Hyp. #2: New techniques are required for users to engage with the analysis of their self-impacted data and to evaluate the impact of their behaviour change.

## 3 METHODOLOGY

We reviewed 68 articles[3], each of them being selected for its relevance regarding at least one of our hypotheses, using search engines such a Google Scholars or ScienceDirect. The final step of our research was to review governmental policies in the world and the scientific reports that either advocated, or not, the use of smart meters. While this search was more context-specific, field deployment reports were especially useful when it came to understanding current pitfalls of self-impacted data visualization systems.

## 4 USER ENGAGEMENT

The problem of user engagement (UE) for analysing self-impacted data is dual: first, users do not engage easily with InfoVis. Second, once they engage and find potential solutions to their problems, it is difficult to commit to these solutions. We address both aspects below.

[1] pierre.vanhulst@unifr.ch
[2] denis.lalanne@unifr.ch

Electronic proceedings of the IEEE VIS 2015 workshop
*Personal Visualization: Exploring Data in Everyday Life*
The authors remain the holders of the copyright

[3] This short paper only references 25 of the 68 papers. Those not referenced covers aspects that are more technical. The full list is available at http://human-ist.unifr.ch/?q=team/pierre-vanhulst/cvbc-guidelines-papers

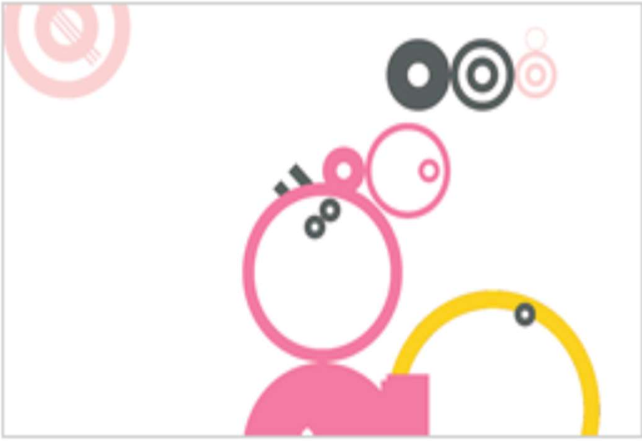

Figure 1 – Tableau Machine [18], an illustration of Casual InfoVis. This visualization displays residential activity in an aesthetically pleasant form.

Of all the contexts covered by self-impacted data, energy conservation provides the most extensive literature. Most governmental reports and studies reviewed came down to the same conclusion, albeit with few variations: so far, the benefits of smart meters are weak. Delmas et al. [8] wrote a meta-review and state that in several cases, eco-feedback could even backfire, and increase the energy use rather than decreasing it. Based on this meta-review, Buchanan et al. explain that the global energy savings offered by smart meters is of 2% [5]. The authors pointed out the lack of thorough evaluations proving the efficiency of In-Home Displays (IHDs). Rego Teixeira reported the results of several field experiments in Switzerland [20]. In one particular experiment, qualitative interviews revealed that users did not feel involved for several reasons: technical issues, lack of financial incentives (sparing energy did not save enough money) and lack of explanations regarding the system are among the most prominent barriers. During our literature reviews, we also came across experiments met with mixed results because of UE. "CommentSpace" [25] is an example of a collaborative visualization system that worked well in laboratory evaluation, but failed to attract users' attention in practise. During its one-month field deployment, "CommentSpace" gathered 180 accounts with only 32 of them leaving 123 comments. The authors observed that incentives and labour division were necessary for users to engage with a visual analytics system.

Many solutions were proposed to this challenging lack of UE. Among these:

**Storytelling & attractive visualizations:** One is to take down the barriers that repel end users. A Netherlands governmental report regarding smart meters [23] states that simple and attractive visualizations should be used as a "first step" for "computer-illiterates" to engage with the system. To this end, the field of "**Casual InfoVis**" - a subset of InfoVis aiming at non-expert users [18] – provides several techniques that may partially apply to our research.

**Monitoring:** Westkog et al. [24] report the results of several studies, which advocate the use of feedbacks in order to catch users' attention and make energy consumption "visible". Darby et al. [7] second this by stating that a monitoring approach is more effective than an "information-driven" approach.

**Nudges & gamification:** Behavioural theories suggest the use of Nudges and Gamification to drive user engagement. However, these techniques suffer from various shortcomings: Rathi and Chunekar [19] highlighted "paternalism" issues or unwanted behaviour changes when using Nudges. As for Gamification, despite many positive short-term results [10], its usefulness over time is yet to be proven. Bartle also theorized that gamification is likely to become less effective as it becomes ubiquitous [2]. However, several examples show that gamification can work over time such as "Stackoverflow", a website where developers ask for help. Points showing expertise on various topics reward those who are the most helpful.

The sustainability of a behaviour change is also problematic to evaluate. Klasnja et al. [16] proved how difficult it is to assess the efficiency of a technology in behaviour change. First, it would require "hundreds or even thousands of people matching control group" over several months or years. Then, technology and information alone are not enough to influence an individual "in the right direction": the success of a behaviour change is related to the socio-cultural environment of the people [17]. Existing evaluation methods of large group of users such as Randomized control trials, also fail to explain how and why users managed to change their behaviour or not. Klasnja et al. argue that early-stage technologies for behaviour change could be evaluated in a "tailored" way, depending on the intervention strategy of the technology – that is, if this technology fosters better behaviour by "self-monitoring", "reduction", "tunnelling", "tailoring", "suggesting at the right time", "conditioning" or "surveillance" (according to Fogg's taxonomy [9]). Each strategy uses different means to leverage user's behaviour and these means can be evaluated without requiring a large base of evaluators. Sprague and Tory [22] argue that new metrics are to be adopted to evaluate such systems, instead of mere efficiency.

Knowing that self-impacted data appeal end users who are not expert in InfoVis, all these considerations are of prime importance to any visualization system in this context. Designers could stick to Fogg's taxonomy and define which persuasive strategies they plan to implement, evaluating them in a "tailored" way as suggested by Klasnja et al.

## 5 Effects of Collaboration

In the section above, we highlighted a serious issue with user engagement in the context of self-impacted visual analytics. While literature proposed several solutions to this problem, we believe that collaboration can be an additional way to solve this issue, as well as being an efficient way to provide end users with better findings.

Collaborative visualization is defined by Isenberg et al. as "the shared use of computer-supported, interactive, visual representations of data by more than one person with the common goal of contribution to joint information processing activities" [15]. The idea that visualization is a powerful tool to support decision-making and sense making is globally accepted in the HCI community. For instance, Bresciani and Eppler [4] conclude that the use of any kind of visualization (either optimal or suboptimal) for a managerial group discussing strategy implementation offers better results than no visualization at all.

However, and contrary to the popular view, the same cannot be said of collaboration. Shepperd [21], among others, observed that collaborating does not always result in higher productivity. One of the reasons for this productivity loss is "collaboration cost" – the time people spend to discuss and coordinate themselves. Lack of perceived gains is also part of the equation. Balakrishnan et al. [1] investigated this phenomenon further and suggested that collaboration works best when each member of the group knows only a portion of the data. They are thus appealed to discuss and share their knowledge, which prevents confirmation bias that can occur in groups where each participant knows all the information.

Literature in CSCW uses the "time-space matrix", a taxonomy based on two spatiotemporal dimensions: distance (co-located or remote collaboration) and time (synchronous or asynchronous collaboration). Of the four situations derived from this taxonomy, two were mostly studied: synchronous co-located collaboration [14] and asynchronous remote collaboration. While it is possible to imagine a social environment that promotes synchronous and co- located collaboration over self-impacted data, this study mainly focuses on remote and asynchronous platforms, similar to what most commercial solutions currently propose.

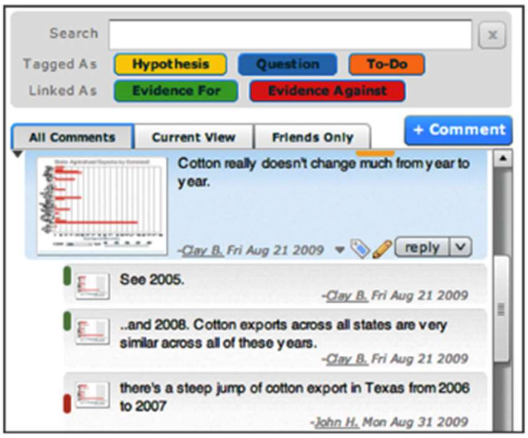


Figure 2 - CommentSpace search interface [12]. Users can create hypotheses, questions or "to-do"s. Other users can provide evidence either for or against this.

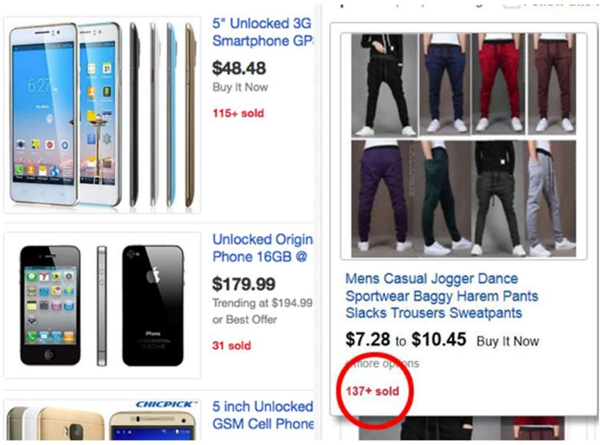


Figure 3 – Social navigation example. While searching on eBay, users are provided with several information such as how many times other users bought a given product, in order to show which are the most popular items.

Heer et al. [12] state that a fundamental aspect of collaborative visualizations is the possibility to share the current state of the visualization and discuss it. This led them to design "Sense.us", a collaborative visualization system that put in practice their theory. With "CommentSpace" [12], Willet et al. further explored this idea by structuring discussions with predefined tags (to categorize contributions) and links (to specify whether the contribution is an evidence for or against the initial assumption). Heer and Agrawala also proposed a precious list of 24 design considerations that take into account various dimensions such as "common ground", "incentives" or "division and allocation of work" [11].

## 6 "SHARE FINDINGS, NOT DATA"

Techniques described in the above section took into account collaboration that occurs on a shared dataset. Self-impacted data being inherently personal, they cannot be shared so easily. One of the question that emerged from these readings was whether it is possible to collaborate on a common structure from several datasets, rather than on a single dataset: concretely, how can users share experience and findings without revealing their personal data.

**Figure 4** shows a preliminary design for a 3-steps system to share findings and observations without revealing data per se. This system is "context-agnostic", as we believe that all self-impacted data could be crossed and analysed: there might be relations between unexpected dimensions. The steps are as follows: (1) produce & gather data, (2) monitor & analyse data, (3) share ideas & adopt new behaviour.

There exist several ways to gather data (1) depending on their nature and availability, we intend to find a standardized way to import them and compare them. Being continuous, they will be characterized by their timestamp and their dimension. Once imported, the data will be displayed (2) with various visualizations as to allow end users to explore them freely. Ambient visualizations – such as power aware cords or hue changing lamps – will trigger higher user engagement, following the idea that a monitoring approach is the best way to drive people's attention. The collaborative layer will be omnipresent: a user's exploration of data will be structured by others' previous findings and hypotheses using techniques described in literature such as social navigation – the fact of displaying former users' trail over virtual objects. Then, in order to propose anonymous suggestions based on others' experience (3), the system will gather relevant data from their own annotations. This act of sharing will be encouraged by

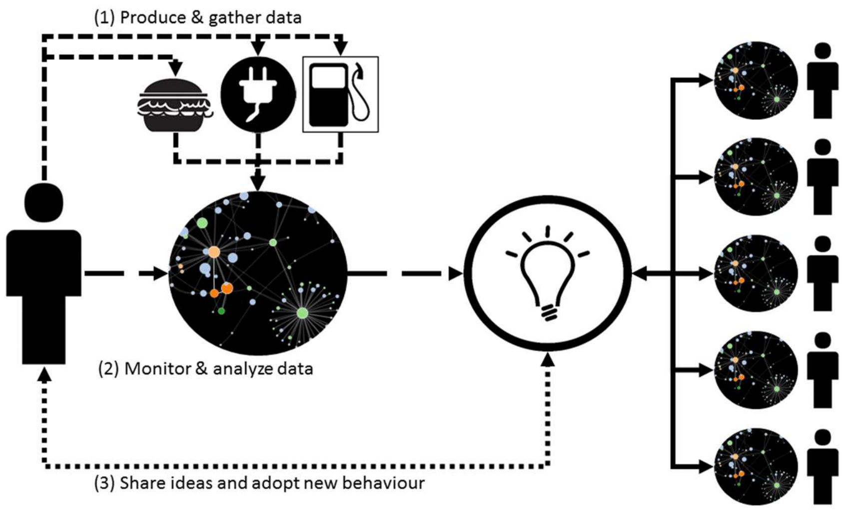


Figure 4 - Design of a collaborative visualization system for behaviour change

nudges and “gamified” mechanisms similarly to those of websites like “Stackoverflow”.

Using Fogg’s taxonomy, such a system would thus implement the “self-monitoring” and “suggestion at the right time” strategies. Surveillance – from other users – could also be implemented: at this stage of the design, we are yet to determine to what extend profiles should be private.

## 7 Conclusion

This paper reviewed the state-of-the-art literature according to two initial hypotheses.

Collaboration can bring better results when team members agree to share their information bias. The challenge lies in of how we are going to use this knowledge with collaborative visualization systems of self-impacted data: is it possible to generalize this observation to domains where data are personal and where there exists no clear “good” answer? How can we adapt techniques developed by Heer et al, and Isenberg et al to self-impacted data? We also assumed that it is possible to generalize techniques for all self-impacted data: this is largely arguable and experiments are required to validate this hypothesis.

This literature review also highlighted the many interesting trails to follow in order to solve user engagement problem. So far, none has been properly evaluated. This opens the door to various experiments.

Subsequent to this paper, we plan to design and develop the collaborative visualization system presented above, in order to push our hypotheses further. The focus of our research will be the sharing of findings and their application: how can an individual benefit from others’ findings and increase her understanding of her self-impacted data? The way such a system will recognize whether a finding made by another person can be applied to another, has yet to be defined.

## References


[1] Balakrishnan, A., Kiesler, S., Kittur, A., & Fussel, S. (2010). Pitfalls of Information Access with Visualizations in Remote Collaborative Analysis. *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 411-420). ACM.

[2] Bartle, R. A. (2011). *Gamification: Too much of a good thing.* Retrieved July 14, 2015, from Presentation given at 'Digital Shoreditch': http://www.scribd.com/doc/81716927/Gamification-Too-Much-Of-a-Good-Thing

[3] Bouton, M. (2014). Why behavior change is difficult to sustain. *Preventive medicine*, 29-36.

[4] Bresciani, S., & Eppler, M. (2009). The Benefits of Synchronous Collaborative Information Visualization: Evidence from an Experimental Evaluation. *Visualization and Computer Graphics, IEEE Transactions on, 15*(6), 1073-1080.

[5] Buchanan, K., Russo, R., & Anderson, B. (2015). The question of energy reduction: The problem(s) with feedback. *Energy Policy*, 89-96.

[6] Costanza, E., Ramchurn, S. D., & Jennings, N. R. (2012). Understanding Domestic Energy Consumption through Interactive Visualisation: a Field Study. *UbiComp ’12*.

[7] Darby, S., Liddell, C., Hills, D., & Drabble, D. (2015). *Smart Metering Early Learning Project: Synthesis report.*

[8] Delmas, M., Fischlein, M., & Asensio, O. (2013). Information strategies and energy conservation behavior:. *Energy Policy, 61*, 729-739.

[9] Fogg, B. J. (2002). Persuasive technology: using computers to change what we think and do.

[10] Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? — A Literature Review of Empricial Studies on Gamification. *Proceedings of the 47th Hawaii International Conference* . Hawaii.

[11] Heer, J., & Agrawala, M. (2008). Design considerations for collaborative visual analytics. *Information Visualization, 7*, 49-62.

[12] Heer, J., Viégas, F. B., & Wattenberg, M. (2009). Voyagers and Voyeurs: Supporting Asynchronous Collaborative Visualization. *Communications of the ACM, 52*(1), 87-97.

[13] Huang, D., Tory, M., Aseniero, B., Bartram, L., Bateman, S., Carpendale, S., . . . Woodbury, R. (2014, September 22). Personal Visualization and Personal Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 420-433.

[14] Isenberg, P. (2009). Collaborative Information Visualization in Co-located Environments. Calgary.

[15] Isenberg, P., Elmqvist, N., Scholtz, J., Cernea, D., Ma, K.-L., & Hagen, H. (2011). Collaborative Visualization: Definition, Challenges and Research Agenda. *Information Visualization, 10*(4), 310-326.

[16] Klasnja, P., Consolvo, S., & Pratt, W. (2011). How to Evaluate Technologies for Health Behavior Change in HCI Research. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3063-3072). ACM.

[17] Owens, S., & Driffill, L. (2008). How to change attitudes and behaviours in the context of energy. *Energy Policy*, 4412-4418.

[18] Pousman, Z., Stasko, J., & Mateas, M. (2007). Casual Information Visualization: Depictions of Data in Everyday Life. *Visualization and Computer Graphics, IEEE Transactions on, 13*(6), 1145-1152.

[19] Rathi, S., & Chunekar, A. (2014). Not to buy or can be ‘nudged’ to buy? Exploring behavioral interventions for energy policy in India. *Energy Research & Social Science* , 78-83.

[20] Rego Teixeira, A. (2014). Le « Smart-metering » en Suisse Romande : conditions sociales d’acceptation et d’appropriation. Genève.

[21] Shepperd, J. (1993). Productivity Loss in Performance Groups: A Motivation Analysis. *Psychological Bulletin, 113*(1), 67-81.

[22] Sprague, D., & Tory, M. (2009). Motivation and Procrastination: Methods for Evaluating Pragmatic Casual Information Visualizations. *IEEE Computer Graphics and Applications - CGA, 29*(4), 86-91.

[23] van Elburg, H. (2014). Dutch Energy Savings Monitor for the Smart Meter.

[24] Westskog, H., Winther, T., & Sæle, H. (2015, May 5). The Effects of In-Home Displays—Revisiting the Context. *Sustainability, 7*, 5431-5451.

[25] Willett, W., Heer, J., Hellerstein, J., & Agrawala, M. (2011, May). CommentSpace: Structured Support for Collaborative Visual Analysis. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 3131-3140). ACM.