Information Overload: Causes and Countermeasure Approaches

Final Report

(2492 Words)

Philipp Klein, 1326251

December 29, 2019

User Research Methods
Winter Term 2019/20

066 935 Media and Human-Centered Computing
Institute of Visual Computing and Human-Centered Technology
TU Vienna

Contents

1 Introduction

In our increasingly digitized world technology is present almost everywhere. Every interaction with a digital device produces a lot of data. This data is often used in user interfaces to convey knowledge and support decisions. Thus it is important to show the relevant and right amount of information to the user. This can be a challenging task since too much information can be overwhelming for people which leads to the Information Overload (IO) problem.

In Section 2 this report gives an introduction to the term ”information overload”. It describes and discusses different definition based on literature and gives examples on some of its main causes. IO is a major challenge in Human-Computer Interaction (HCI) therefore Section 3 shows the importance of IO in HCI and gives examples of countermeasure approaches for today's and future applications.

2 Information Overload

Today, especially interactions with devices that are connected to an online service, create a vast amount of information. This so called digital footprint is a increasingly popular term to address the information people generate with every digital interaction. Blue et al. argue that these digital footprints log the habits, interests, relationships, and communications of peoples physical lives [1]. These enormous amounts of data make it difficult for designers and developers to distinguish between valuable and unnecessary information. New information and communications technology allow to display data in many different forms. But one of the problems that occur from the interconnection of the digital and physical world is the amount of information that people are confronted with. Humans often cannot process everything they perceive and some valuable information might be lost. This problem is frequently called *information* or *data overload*. Although this issue has become more significant in recent years, scientists already addressed this topic before computers were such an important part of society. Milord and Perry, for example, mention the issue of information overload in 1977. They argue that information overload is caused by an individual's perception not by the input-output capacity of a system [8]. For them it is a psychological phenomenon because the type and density of data should not be determined by the used technology. It should consider the human as the most important part of an information visualization. Since then further research has been conducted that examined the information and data overload problem in our current society [12][4][6][9]. All of them argue that IO is a major challenge of our digitized and globalized world.

According to Rachfall et al. it is hard to filter the important information if people are confronted with a flood of data [9]. They argue that, when people talk about IO, they usually mean that they get too much information they cannot process. But Rachfall et al. describe that information overload is tightly connected to the scientific discipline. This makes it difficult to come up with an universal definition for the term ”information overload”. Rachfall et al. argue that different situations require different descriptions for IO. Sometimes it is ”an accumulation of information which can't be processed because of the limited possibilities of the human brain” [9, p. 28]. In other situations IO is defined as ”an excessive amount of information of which just small amounts are relevant” [9, p. 28]. Based on a work by Tushman and Nadler ([11]) they

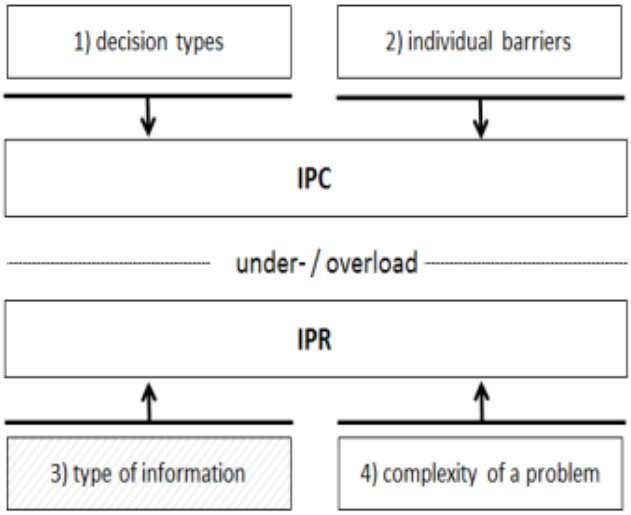

Figure 1: Information Under-/Overload Framework by Rachfall et al. (Source: [9])

came up with the formula: *information processing requirements (IPR) > information processing capacities (IPC)* to define IO [9]. According to Rachfall et al. IPC is based on the individual's cognitive capacity which is influenced by aspects like willingness, motivation, knowledge, etc. IPR on the other hand is characterized by non-individual aspects like presentation format, time pressure, etc. It describes the capacity that is needed to solve the problem or make a decision. Based on this assumptions Rachfall et al. came up with a framework to define the term "Information Overload". Figure 1 illustrates the framework. It suggests that IO occurs if the IPR is bigger than the IPC. Rachfall et al. argue that their framework captures definitions of various research on information overload

2.1 Causes of Information Overload

According to Rachfall et al. different types of information have different effects on IO. Their developed framework (see Figure 1) was used in a case study to distinguish between bad-, irrelevant-, or relevant information. Its goal was to find a possible relationship between the types of information and IO. They conducted their study in four companies with management accountants. Rachfall et al. discovered that too much bad information leads to the information overload phenomenon. But too much relevant information has a positive effect on IO as well. On the other hand they found no correlation between the amount of irrelevant information and IO. Overall Rachfall et al. argue that there is a connection between information type and information overload.

A related concept to information overload is *Data Overload* (DO). Both terms are often used interchangeably although they are different. Like already mentioned, Rachfall et al. ([9]) define the term information overload as an individuals capability to process information in relation to the displayed information. They think that IO is heavily influenced by the user and peoples

ability to process information. Data overload on the other hand is more focused on the data that needs to be handled. Woods et al. think that DO is one of the major problems in our current world [12]. Especially in our digital world everybody creates a lot of data that needs to be processed. They argue that it is crucial to find the root causes of data overload first. Otherwise it is not possible to suggest appropriate solutions. In their work Woods et al. describe three basic characteristics of DO. They argue that DO is a difficult problem and the characteristics allow to better understand its causes. The first characterization is that often *too much data* is present. They describe it as a clutter problem because too much stuff is displayed. Although a reduction of the number of displayed data bits is a possible solution, Woods et al. think that this misinterprets the problem. They argue that this approach is often not fruitful because in many contexts, some of what is removed is valuable information. Woods et al. call the second characterization *workload as bottleneck*. They describe it as a time related problem because too much data needs to be analyzed in too few time. They argue that this problem can be partly solved by automation technologies. Such technologies should support the user during their activities. Although this is a potentially useful approach, they also argue that it is not sufficient all the time. The last characterization is the *finding of significance in data*. Woods et al. argue that the last characterization is the most important one, because it allows to create truly useful solutions. They think that machine learning allows to better organize data. The biggest problem is that the significance of data depends on the context. [12] argue that the value of the information in the data is not absolute and therefore the finding of significance is a challenging task.

Until now this report focused on the ”unintentional” causes of IO like browsing the web or using all kinds of digital services. Normally those can only be avoided by not using those products but not all data generation fall in that category. Gamzu et al. argue that some of the data overload people are confronted with, is self inflicted [4]. They examined the information overload created by emails, especially from junk mails. According to them machine-generated mail make up more than 90% of non-spam mail traffic, which means that users demanded those mails in the form of subscriptions. They argue that this is major contributor to information overload in email. But Gamzu et al. also found that there is a significant gap between users wish to reduce the number of those mails and their active behavior to do so. They think that this lack of willingness to reduce the IO in their mail traffic is due to inaccessible and complex mechanisms to unsubscribe. Although Gamzu et al. conducted their study in the context of subscriptions to email services, the overall message can be derived to IO in general. Gamzu et al. suggest that users need a simple and accessible way to decrease the IO in digital products, otherwise the complexity and effort to reduce IO exceeds the perceived benefits.

3 Information Overload and Human-Computer-Interaction

The low cost of data collection and storage led to a vast number of possible information to display. Designers often need to confront users with information that is considered crucial to the product but the problem of information overload can occur if too much information is shown to users. Thus IO is a very important topic in HCI nowadays and designers have to consider it in their products. To give the best user experience it is necessary to show the right amount of information in the most efficient way possible. Like mentioned IO is dependent on the individ-

uals perceptions and processing abilities. Therefore every person has a different IO threshold. This is a major problem when defining the right amount of information that can be included in a product without overwhelming the user.

One approach to find appropriate measures to detect individuals threshold of workload was done by Gevins and Smith. They argue that one of the key challenges regarding cognitive workload is the problem of how to accurately measure it [5]. An electroencephalogram (EEG) allows to monitor brain function, thus it is very sensitive to alertness and attention. Gevins and Smith argue that an EEG can measure cognitive workload because tasks that require different mental effort cause changes in brain activity. They reviewed a long-term study that aimed to monitor cognitive workload methods based on EEG measures. They evaluated different tasks in various computer based contexts to achieve more natural results. Gevins and Smith think that EEG measures can be effective to detect variations in cognitive workload in laboratory settings. But they also suggests that a real-world environment study would be beneficial for further applications. Overall Gevins and Smith argue that their work has shown "initial evidence for the feasibility of creating EEG-based technologies for monitoring cognitive workload during human–computer interaction" [5, p. 125].

There are many problematic fields of application where IO can cause security issues or might be harmful, for example new technologies, like smart cars. Hwangbo et al. evaluated the complexity of in-vehicle interface system [6]. They argue that drives have to process visual information in split seconds within a minimal field of view area. Hwangbo et al. conclude that the "development of in-vehicle technology has brought about several critical issues such as driver distraction" [6, p. 82]. Therefore their research shows the necessity for research on IO and display complexity in critical systems like smart cars.

3.1 Countermeasure Approaches

Information overload is a complex topic in HCI that needs to be considered carefully. The information handling capabilities of everyone is different and the amount of information that a person can process without being overwhelmed by the complexity or amount of information depends on various aspects. But the problem of IO has already been addressed by researchers and designers that came up with different countermeasure approaches. This section gives three possible techniques how current or future products can reduce IO.

3.1.1 Data Visualization

Big data is a major topic in computer science nowadays. Like already mentioned, everybody creates a lot of data constantly. Cota et al. argue that big data is used to generate knowledge and facilitate decision making [3]. But the major challenge is, how to visualize all the data without losing crucial information. Currently a lot of techniques exist to represent data, like tables, histograms, charts, treemaps, etc. Cota et al. think that today the main challenge for various forms of data representation are the scalability and dynamism of data. Current technologies represent data in 2D but Cota et al. argue that future visualizations will move towards three-dimensional graphs since it allows to display more information in an efficient way.

3.1.2 Simplicity

A widespread countermeasure is to display only the necessary and useful information. So called, minimalism or simplicity is a very popular term when designing user interfaces. Although Chang et al. focused on interaction design in general, their research can be applied to the representation of information as well [2]. They argue that attaining simplicity is a key challenge in interaction design because designers tend to have *featuritis* which they define as "the tendency for designers to emphasize the number or novelty of features over core usability" Chang et al.. According to Chang et al. the mental and physical clutter in user interfaces increases with device functionality. This leads to the assumptions that current and future applications tend to include more features that increase the IO of products. Other researchers and design best practices argue that simplicity is crucial to increase user experience and build better user interfaces, but it needs to be carefully implemented [10]. Otherwise too much information gets lost and users might get frustrated when using a product.

3.1.3 Recommendation Systems

Khusro et al. highlight the importance of recommendation systems to handle IO [7]. They define a recommender system as an "Information Retrieval technology that improves access and proactively recommends relevant items to users by considering the users' explicitly mentioned preferences and objective behaviors" [7, p. 1179]. Such preferences are based on a users profile which consists of personal information, items visited, rated, purchased, etc. Khusro et al. argue that today's recommender systems are not precise enough to fulfill all users' needs. Thus the major challenge for future applications are to create fine-tuned and high-quality recommender systems. According to them, future recommender systems should 1) use demographic filtering and clustering, 2) obtain personal information of newly registered users, 3) maintain two recommendation lists for users who frequently change their preferences, and 4) filter out obsolete and old items.

4 Conclusion

This report gave a brief introduction to information overload. Although information overload is a very complex topic and a definition depends on the context, the user is a key factor when determining the right amount of information integrated into a system.

The general definition is that information overload happens when the information processing requirements exceed the information processing capabilities of a user. Thus it is a important topic for HCI because IO can not be separated from the user. The causes of IO are rather complex but in general they can be unintentional or self inflicted. But the information processing abilities is different for every person. To determine the right amount of information various factors need to be considered. This report gave the example of an EEG to monitor the information workload threshold of people in a laboratory setting. Although this approach is not applicable to real-world applications it gives an example how to determine of the right amount of information in user interfaces.

Today many different countermeasure approaches exist and this report gave three examples. First it is possible to carefully think about the visualization technique that is used in today's interfaces. Cota et al. argue that big data led to a vast amount of information that can be displayed [3]. Future application will need to consider new techniques like 3D representations to handle all this information. Another very popular technique is to implement simplicity or minimalism in user interfaces although it can be challenging to display only the most important data. The last suggestions this report includes are recommender systems. These consider a users previous behaviour to find the most relevant information for them. Overall this report showed that information overload is a very challenging topic in HCI nowadays and trends like big data show that it will be important in future applications as well. Although various countermeasure approaches exits a combination of AI and human interaction techniques will be crucial to combat the issue of IO in the future.

References

[1] J. Blue, J. Condell, and T. Lunney. Digital Footprints: Your Unique Identity. In *BCS, The Chartered Institute for IT (ACM Proceedings) 32nd Human Computer Interaction Conference, BHCI 2018*, pages 1–4, 7 2018.

[2] A. Chang, J. Gouldstone, J. Zigelbaum, and H. Ishii. Simplicity in Interaction Design. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, TEI '07, pages 135–138, New York, NY, USA, 2007. ACM.

[3] M. Cota, M. Díaz Rodríguez, M. R. Gonzalez Castro, and R. Gonçalves. Analysis of Current Visualization Techniques and Main Challenges for the Future. *Journal of Information Systems Engineering & Management*, 2, 08 2017.

[4] I. Gamzu, L. Lewin-Eytan, and N. Silberstein. Unsubscription: A Simple Way to Ease Overload in Email. In *Proceedings of The 11th ACM International Conference on Web Search and Data Mining Los Angeles, California, USA*, pages 189–197, 02 2018.

[5] A. Gevins and M. Smith. Neurophysiological Measures of Cognitive Workload during Human-Computer Interaction. *Theoretical Issues in Ergonomics Science*, 4:113–131, 01 2003.

[6] H. Hwangbo, S. C. Lee, and Y. G. Ji. Complexity Overloaded in Smart Car: How to Measure Complexity of In-vehicle Displays and Controls? In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '16 Adjunct, pages 81–86, New York, NY, USA, 2016. ACM.

[7] S. Khusro, Z. Ali, and I. Ullah. Recommender Systems: Issues, Challenges, and Research Opportunities. In K. J. Kim and N. Joukov, editors, *Information Science and Applications (ICISA) 2016*, pages 1179–1189, Singapore, 2016. Springer Singapore.

[8] J. T. Milord and R. P. Perry. A Methodological Study of Overload. *The Journal of General Psychology*, 97(1):131–137, 1977.

[9] T. Rachfall, E. Williamson, and B. Temple. The Information Overload Phenomenon: The Influence of Bad - and (Ir) Relevant Information. *International Journal of Research in Engineering and Technology*, 03:27–32, 12 2014.

[10] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Pearson, 6th edition, 2016.

[11] M. L. Tushman and D. A. Nadler. Information Processing as an Integrating Concept in Organizational Design. *Academy of Management Review*, 3(3):613–624, 1978.

[12] D. D. Woods, E. S. Patterson, E. M. Roth, and K. Christoffersen. Can We Ever Escape from Data Overload? A Cognitive Systems Diagnosis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 43(3):174–178, 1999.