# Web-based Interactive and Visual Data Analysis for Ubiquitous Learning Analytics

Benjamin Weyers, RWTH Aachen University, Virtual Reality and Immersive Visualization Group, weyers@vr.rwth-aachen.de
Christian Nowke, RWTH Aachen University, Virtual Reality and Immersive Visualization Group, nowke@vr.rwth-aachen.de
Torsten W. Kuhlen, RWTH Aachen University, Virtual Reality and Immersive Visualization Group, kuhlen@vr.rwth-aachen.de
Mouri Kousuke, Kyushu University, Faculty of Arts and Sciences, mourikousuke@gmail.com
Hiroaki Ogata, Kyushu University, Faculty of Arts and Sciences, ogata@artsci.kyushu-u.ac.jp

**Abstract:** Interactive visual data analysis is a well-established class of methods to gather knowledge from raw and complex data. A broad variety of examples can be found in literature presenting its applicability in various ways and different scientific domains. However, fully fledged solutions for visual analysis addressing learning analytics are still rare. Therefore, this paper will discuss visual and interactive data analysis for learning analytics by presenting best practices followed by a discussion of a general architecture combining interactive visualization employing the *Information Seeking Mantra* in conjunction with the paradigm of *coordinated multiple views*. Finally, by presenting a use case for ubiquitous learning analytics its applicability will be demonstrated with the focus on temporal and spatial relation of learning data. The data is gathered from a ubiquitous learning scenario offering information for students to identify learning partners and provides information to teachers enabling the adaption of their learning material.

**Keywords**: interactive analysis; web-based visualization; learning analytics

## Introduction

Interactive visual data analysis is a well-established class of methods to gather knowledge from raw and complex data. Information visualization approaches and tools have shown high impact in various fields of research. For instance, the use of visual data analysis enables high performance computing experts to analyze code running on NUMA architectures (Weyers et al., 2014). In neuroscience, one major challenge is the analysis and interpretation of heterogeneous data resulting from simulations as well as biological experiments. Tools such as VisNEST provide coordinated multiple views which offer various perspectives on the data. Time-varying data is presented in such a way that domain scientists are able to navigate as well as analyze it (Nowke et al., 2013). From the data perspective, various types of visualization concepts can be found in literature, e.g., bar charts or pie charts (Spence, 2001) or representation concepts for relational data such as classic tables or graphs (Battista et al., 1999).

In learning analytics, only a few works can be found which address the benefits of visualization:

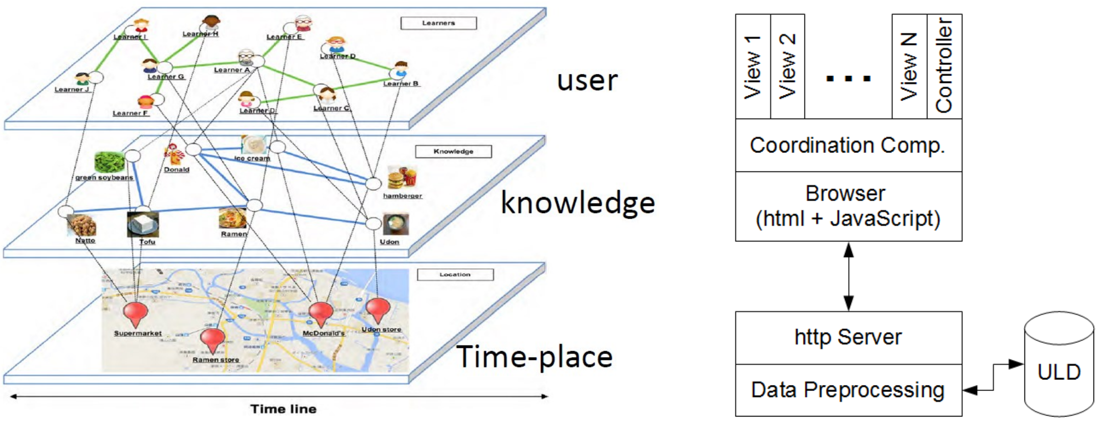


Figure 1: Left: three-layered representation of ubiquitous learning logs as generated in SCROLL. Right: Web-based visulaization architecture

presenting raw and complex data in an easy to understand and interpretable manner. In context of learning dashboards, Verbert et al. (2013) present results how visualization can be used in learning analytics. Their work concentrates on a learning analytics process model in which information visualization is used to enrich learning dashboards to support students and teachers in their daily work. A similar approach has been presented by Leony et al. (2012). They present a web-based tool called GLASS that is based on a four-layered architecture. The architecture addresses the use of classic visualization components such as bar charts and the application of filters for pre-processing data. However, both works do not focus on interactive data analysis and omit the discussion of how to create suitable visualizations for learning analytics. The work of Liddo et al. (2011) presents a successful use of visualization techniques for learning analytics by using graph visualizations to represent the course of a discussion. Nonetheless, they constrain their work on a very specific use case. For location-based and mobile learning, some works can be found such as this presented by Melero et al. (2015). They present a set of visualizations for abstract information (including location) for a specific scenario, which do not offer interactive analysis features but show a relevant set of visualizations that can be used in location-based learning.

To our knowledge there is no work on the support of ubiquitous learning analytics by using interactive visualization techniques. Hence, the main contribution of this paper consists in the presentation of an architecture for ubiquitous learning analytics which is based on well-established concepts in information visualization. The applicability of this architecture will be demonstrated by means of a use case for the visualization of ubiquitous learning logs (ULLs). ULLs are gathered from learning scenarios where students use ubiquitous devices, e.g., mobile devices to track learning progress in the wild or in the class room. Logging generates spatiotemporal data that represents what the learner has acquired at which point in time (Ogata et al., 2011). ULLs can be interpreted as a four dimensional space where the first three dimensions represent the learner, the acquired knowledge, and the location. The fourth dimension represents time at which a student created the log. Beside the creation of these logs, further aspects can be included into the analysis, such as how often the ULL has been used to recall knowledge. In general, the first three dimensions can be interpreted and visualized as a graphical structure such that these dimensions can be interpreted as three layers which are related to each other as shown in Figure 1.

The paper is structured as follows. The next section presents an analysis of ULL and their specific requirements regarding interactive data analysis tools and the role of users. This discussion is followed by the presentation of an architecture for visual ubiquitous learning analytics which facilitates the information seeking mantra and the concept of coordinated multiple views. A use case is discussed which presents the feasibility of the architecture in real world scenarios. The paper concludes with a short summary and the discussion of future work.

## VISUAL DATA ANALYSIS FOR UBIQUITOUS LEARNING ANALYTICS

Ubiquitous learning data (ULD) in general and ULLs in specific present a special challenge for interactive visual data analysis approaches and tools. First, such data is heterogeneous comprising spatial, temporal as well as learning specific data (D1). For instance, ULLs contain a unique identifier for the student, data addressing the time and place where the student acquired knowledge. Second, ULD datasets can be large due to the number of users and stored data items (D2). Last, ULD has a tendency to be incomplete or to contain corrupted data by faulty entries (D3).

Beside the data requirements, a visualization architecture (VA) has to consider the analysis tasks of the user. These tasks are therefore user centric. This paper concentrates on two user roles: the teacher and the student. A user who is a teacher has certain requirements for visual analysis such as to include the obtained information contained in the ULD in the preparation of courses, extend and change course material or specifically include the ubiquitous learning infrastructure into the course. The latter could address to use of the system during the course or in between the course as well as for the definition of the final grade for a student. The student has a slightly different requirement on visual analysis system. A student can be interested in tracking achievements, in the learning progress or finding learning partners who have the same learning goals. In addition, she could be interested in where to learn best or where to find relevant learning material in her nearby surroundings. In summary, the following requirements can be identified:

- V1: The VA has to offer an overview of the data as well as specific details in a certain context depending on user roles
- V2: The VA has to offer various perspectives on the data, which reflect different requirements on the data depending on the user role
- V3: The VA has to offer presentations which consider different combinations of data dimensions of the underlying dataset to address the individual needs of a user

Most ubiquitous learning systems are implemented based on web technologies. To seamlessly integrate the analysis with the ubiquitous use of such learning technologies, the VA should be also based on web technologies leading to the last requirement:

V4: The VA should be implemented using web-based technologies

## ULD Visualization Architecture

To address the identified requirements, the realization of a VA has to follow two methodical paradigms which are already well-established in the visualization community: the *information seeking mantra* and the concept of *coordinated multiple views*. The information seeking mantra introduced by Shneiderman (1996) is based on interactive visualization which has been shown as successful for visual data analysis (Fuchs & Hauser, 2009). Interactive visual data analysis is understood as the analysis of data using visualizations which are customizable during runtime by the user. This customization can be defined by various types of manipulations such as the application of filters, the selection of data items in a visualization, details of this selection, or the change of the visualization technique used to display a dataset. The information seeking mantra proposes and specifies a general workflow for the visual analysis process: *overview first, zoom and filter, and details on demand*. Thus, as a first step, a visualization tool should present an overview of the dataset. Following this, the user should be able to explore the dataset's representation interactively by zooming into it, e.g., selecting a subset of data items, and apply filters, e.g., restricting the datasets dimensions. An architecture realizing the information seeking mantra addresses in particular D2 as well as V1 and V3.

Roberts (2007) presents an overview on coordinated multiple views. Coordinated multiple views "is a specific exploratory visualization technique that enables users to explore their data. In fact, the overall premise for the technique is that users understand their data better if they interact with the presented information and view it through different representations" (Roberts, 2007, p. 1). This is specifically true when the interpretation task benefits from different perspectives on the data by utilizing various visualization techniques and the data is multi-dimensional or heterogeneous as it is the case with ULD. The coordinated multiple view paradigm combines different types of visualizations of the same data with a coordination mechanism of views that react to user's interaction intents accordingly. For instance, this can be a selection of a data item in one view which is then propagated to all coupled views displaying this subset in their perspective on the data (i.e., a coordination mechanism termed brushing). Analogously, the zoom and filter step of the information seeking mantra can be coordinated between views by applying the same zooming and filtering operation to all other views. A visualization architecture implementing coordinated multiple views addresses the requirements D1, V2 as well as V3. V1 is implicitly addressed because an overview of the data can be obtained by various views showing different perspectives but in combination present the whole dataset at one glance. V4 is addressed by implementing the visualization architecture as web-based components as presented in Figure 1, right. In a web-based environment, the complete dataset must be accessible by the server. By a combination of a server for the communication with the client-side visualization and the data pre-processing, a reduction of the data size can be considered to make the communication more efficient. The server-side data preprocessing should be able to pre-compute data structures, such as graph-based representations from a table-based dataset. The client-side application should offer a coordination component that provides interfaces and communication logic for the propagation of interaction events between the views. A controller view should be provided to offer a graphical user interface for general control operations, e.g., triggering the loading of datasets and the application of client- or server-side filtering operations.

## Use Case – Learning Analytics for Ubiquitous Learning Logs

The applicability of the ULD visualization architecture will be shown by means of two perspectives on an analysis use case: (a) a student tries to improve her learning workflow by analyzing her own ULLs and (b) a teacher who would like to extend her learning material and activities. The gathered data originates from the SCROLL system (Ogata et al., 2011), a system for gathering ubiquitous learning logs which collects words a student observed in their everyday surrounding and which they learned for their individual vocabulary. The current implementation consists of three major visualization designs: two force-directed-layout-based interactive graph visualizations, a circular graph visualization using edge bundling, and a Google earth-based visualization of the position the ULLs have been captured, thus showing in which spatial context students learned words (see Figure 3, left). The presented visualizations are based on D3.js, a JavaScript library capable of visualizing data in web-based applications. We used different open source extensions for D3 to build these visualizations. Figure 2 shows two graph-based visualizations which represent the relation between students according to similar learned words. The left representation shows the words as light blue circles where students are represented in dark blue. The right representations only shows the relation of one specific student (here Sophie) to other

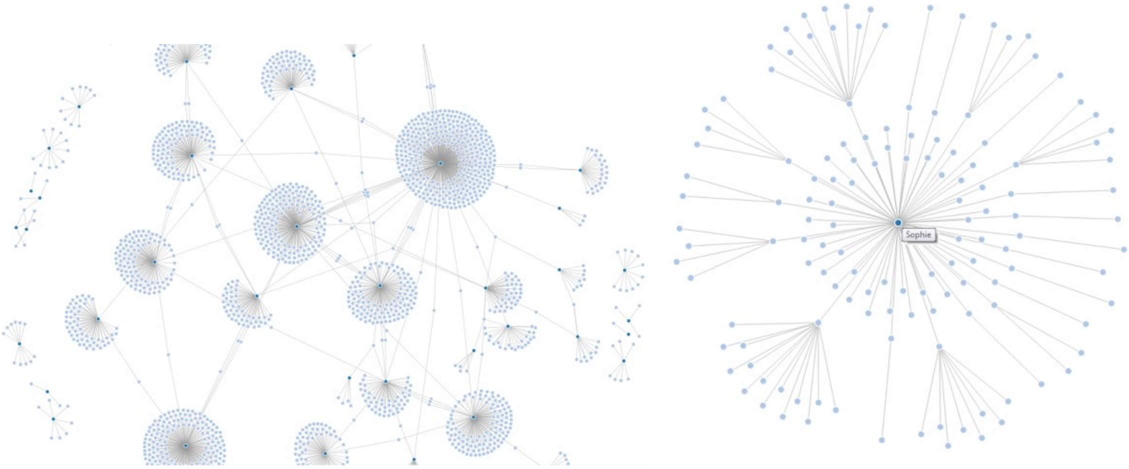


Figure 2: Left: Overview of the dataset showing the relation of students (dark blue) to learned words (light blue). Right: Graph visualization of the nearest neighbors of student “Sophie”.

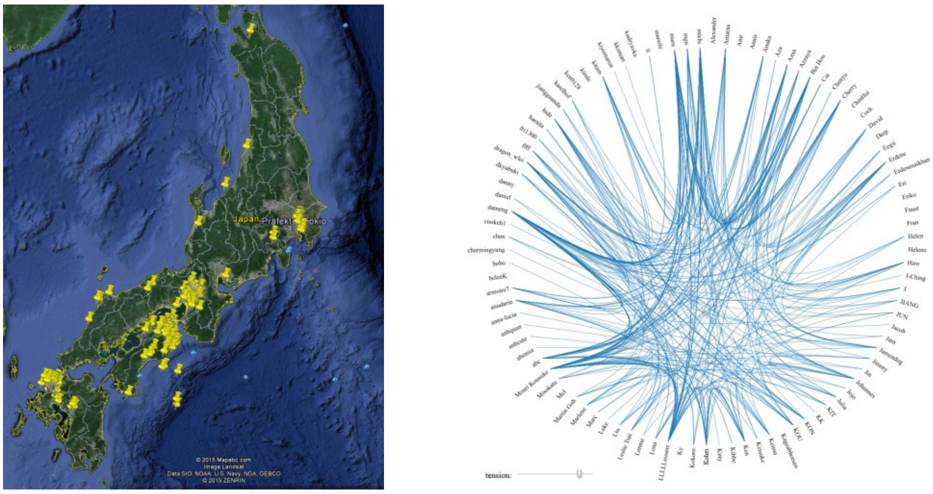

Figure 3: Left: Google Earth based visualizatoin of the geographic position of ULLs. Right: A circular graph view based on an edge bundling approach showing the relation between studens by means of learned words.

students and omits information on words they have in common. The latter information is also presented in Figure 3 on the right. Here, all students are visible and can be interactively selected such that the relations of students are highlighted. In the following, an informal description of use cases regarding the two user roles will be discussed and the potential benefit of the ULD visualization architecture identified.

## Use Case - Student’s Perspective

A student (Sophie) would like to find collaborators to learn new words in order to extend her vocabulary. Her first step is to find other students which have some overlap in their already learned words. Therefore, she uses the graph view presented in Figure 2, right. In a second step, she wants to find a student who knows as many words as possible that Sophie does not know yet. As both visualizations can be assumed to be coordinated linked, the selection of a student in the right view of Figure 2 is propagated to the left graph view. Sophie is now able to see which person fits to the previously defined criteria by inspecting the number of words connected to the selected student. Following this, Sophie inspects the Google earth representation which highlights the selected student’s location she is interested in to learn with. This will inform her whether the matching student is in a reachable distance. Finally, Sophie can identify the hot spots which students are visiting to learn vocabularies. This process follows the Information Seeking Mantra as Sophie first gets an overview and then selecting certain entities she is interested in. This can be interpreted as “zooming in” into the data set. Finally, by checking the location information, details are included into Sophie’s analysis.

## Use Case - Teacher’s Perspective

The teacher’s use case is in various points similar to this of the student’s but follows slightly different objectives. A teacher could be interested in the interpretation and analysis of ULLs to extend or to adapt the content and the pedagogical approach she is following in class. For instance, the teacher could define learning groups for the classes along the same analysis process as described in the student use case above. On top, the teacher can identify new vocabulary to be taught in class based on prior knowledge of students and identify words unknown to most students. By zooming in and highlighting details on demand she is able to identify well-known words and topics. Finally, she can use the circular graph visualization in Figure 3 (right) to identify which students have comparable vocabulary skills.

## Conclusion and Future Work

This paper explored the application of well-established visual analysis methods into the field of Ubiquitous Learning Analytics. We proposed a web-based architecture for interactive and visual ubiquitous learning analytics following two main concepts: the information seeking mantra and coordinated multiple views. By means of a simple use case, the applicability of the architecture has been shown.

For future work, the implementation and evaluation of such a framework is planned. It is planned to realize the proposed architecture in the context of e-book based instructions in various scenarios, such as MOOCs. Furthermore, we plan to analyze further visualization approaches and techniques to offer various types of analysis workflows in the future.

## Acknowledgement

The research have been supported by “Research and Development on Fundamental and Utilization Technologies for Social Big Data” (178A03), the Commissioned Research of the National Institute of Information and Communications Technology (NICT), Japan.

## References

Battista, D., Eades, P., Tollis, I. G., & Tamassia, R. (1999). *Graph drawing: algorithms for the visualization of graphs*. Upper Saddle River: Prentiece Hall.

De Liddo, A., Shum, S. B., Quinto, I., Bachler, M., & Cannavacciuolo, L. (2011). Discourse-centric learning analytics. In Proceedings of the 1st International Conference on Learning Analytics and Knowledge (pp. 23-33). ACM.

Fuchs, R., & Hauser, H. (2009). Visualization of multivariate scientific data. *Computer Graphics Forum, 28*, 1670-1690.

Hänel, C., Weyers, B., Hentschel, B., & Kuhlen, T.W. (2016). Visual Quality Adjustment for Volume Rendering in a Head-Tracked Virtual Environment. *Transactions on Visualization and Computer Graphics*, in press.

Melero, J., Hernández-Leo, D., Sun, J., Santos, P., & Blat, J. (2015). How was the activity? A visualization support for a case of location-based learning design. *British Journal of Educational Technology*, *46*(2), 317-329.

Nowke, C., Schmidt, M., van Albada, S. J., Eppler, J. M., Bakker, R., Diesmann, M., Hentschel, B., & Kuhlen, T. (2013). VisNEST—Interactive analysis of neural activity data. In *Proc. of IEEE Symposium on Biological Data Visualization* (pp. 65-72). IEEE.

Ogata, H., Li, M., Hou, B., Uosaki, N., El-Bishouty, M.M., & Yano, Y. (2011). SCROLL: Supporting to share and reuse ubiquitous learning log in the context of language learning. *Research and Practice in Technology Enhanced Learning, 6*(2), 69-82.

Roberts, J. C. (2007). State of the art: Coordinated and multiple views in exploratory visualization. In *Coordinated and Multiple Views in Exploratory Visualization*, (pp. 61-71). New York, IEEE.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In Visual Languages, 1996. In *Proceedings of IEEE Symposium on Visual Languages* (pp. 336-343). New York, IEEE press.

Spence, R. (2001). Information visualization (Vol. 1). New York: Addison-Wesley.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. American Behavioral Scientist,.

Weyers, B., Terboven, C., Schmidl, D., Herber, J., Kuhlen, T. W., Müller, M. S., & Hentschel, B. (2014). Visualization of memory access behavior on hierarchical NUMA architectures. In *Proceedings of the First Workshop on Visual Performance Analysis* (pp. 42–49). New York: IEEE.