

Towards Interactive Visualization of Longitudinal Data to support Knowledge Discovery on Multi-Touch Tablet Computers

Andreas Holzinger¹, Michael Schwarz¹, Bernhard Ofner¹, Fleur Jeanquartier¹
Andre Calero-Valdez², Carsten Roecker^{1,2}, Martina Ziefle²

¹Research Unit Human-Computer Interaction,
Institute for Medical Informatics, Statistics and Documentation, Medical University Graz,
<a.holzinger,m.schwarz,b.ofner,f.jeanquartier}@hci4all.at

²Human-Computer Interaction Center, RWTH Aachen University, Germany
<calero-valdez,roecker,ziefle}@comm.rwth-aachen.de

Abstract. A major challenge in modern data-centric medicine is the increasing amount of time-dependent data, which requires efficient user-friendly solutions for dealing with such data. To create an effective and efficient knowledge discovery process, it is important to support common data manipulation tasks by creating quick, responsive and intuitive interaction methods. In this paper we describe some methods for interactive longitudinal data visualization with focus on the usage of mobile multi-touch devices as interaction medium, based on our design and development experiences. We argue that when it comes to longitudinal data this device category offers remarkable additional interaction benefits compared to standard point-and-click desktop computer devices. An important advantage of multi-touch devices arises when interacting with particularly large longitudinal data sets: Complex, coupled interactions such as zooming into a region and scrolling around almost simultaneously is more easily achieved with the possibilities of a multi-touch device than compared to a regular mouse-based interaction device.

Keywords: Data Visualization, Longitudinal Data, Time Series, Mobile Multi-Touch Device

1 Introduction and Motivation for Research

One of the grand challenges in modern data-centric medicine is dealing with large, complex, heterogeneous and weakly structured data sets and large amounts of unstructured information. This calls for new, efficient and user-friendly solutions for handling such data – with raising expectations of end-users. Traditional approaches for data handling often cannot keep pace with demand, also increasing the risk of delivering unsatisfactory results. Consequently, to cope with this rising flood of data, new user-centered approaches are vital [1-4].

Particularly, the advent of mobile devices and ubiquitous smart sensors has led to an ongoing trend to record all sort of personal biomedical data over time [5, 6]. These recordings lead to a growing amount of so-called longitudinal data, in the engineering domain maybe better known as time series data [7].

A major challenge is how to deal with such time-dependent data, and not only to deal with it, but to *discover knowledge* from it.

Generally, there are two different ways to accomplish such tasks:

- 1) Applying *mathematical models* for description and prediction purposes [8, 9]; or
- 2) Using *visual inspection* to generate general assertions in regard to the properties of the underlying data set.

For the second kind of knowledge discovery, interactive data visualization plays an important role. As stated by Tufte (1983) [10], graphical representations of data sets are instruments for reasoning about quantitative information. A recent example for such analysis and sensemaking of complex biomedical data can be found in [11]. Therefore, often the most effective way to summarize a large set of quantitative information is to look at “pictures” of these numbers. However, during this visual information seeking process it is often necessary to *interact* with the displayed data: It is important to include the human expert into the data exploration process, and to combine the flexibility, creativity, general knowledge and pattern recognition abilities (in low dimensions) of the human with the enormous capacity, analytical power, and pattern recognition abilities (in high dimensions) of computer solutions. A recent scientific approach is in combining the best of these two worlds [12] and a concrete topic is interactive visual data mining (VDM) [13, 14], which aims to integrate the human expert into the whole data exploration process and to effectively represent data visually, so to benefit from the human perceptual abilities and allowing the expert to get *insight* into the data by direct interaction with the data [15]. VDM can be particularly helpful when little is known about the data sets and/or the exploration goals are ill-defined or evolve over time [16]. The aspect of “time” in data visualization is generally most underrepresented in such approaches, yet, it is of vital importance, particularly in dealing with biomedical data [17].

Thereby, common tasks include: zooming into a portion of the overall data, changing the reference scale, comparisons with other data points and getting detailed, underlying information on some specific data points.

2 Theory and Background

2.1 Longitudinal Data

Longitudinal data (or time series data) arises when a certain random variable is recorded as a sequence over time, whereas the measurement of some characteristics at (roughly) one single point in time is called cross-section data. For a general introduction into time series data refer to [18-23].

Depending on the measurement, differentiated distinction can be drawn between *discrete* and *continuous* time series.

A discrete time series is one in which the set T_0 of times at which observations are made is a discrete set.

On the other hand, if observations are recorded continuously over a time interval, e.g. $T_0 = [0,1]$ a continuous time series arises [24]. In contrast to analog recording, the process of digital recording is inevitably connected to a discrete sample frequency. Therefore, strictly speaking, although the lag between two measurements may be in the magnitude of milliseconds, we always have discrete time series data in the electronic data processing domain.

The same is true for the recorded value. To be exact, we always record discrete variable values. Nevertheless, the underlying natures of, for instance, height, blood pressure, or weight of a person, are examples for continuous variables, whereas the number of patient visits on one day is an example for discrete ones. For continuous values it is therefore important to choose a recording sampling frequency that draws an adequate picture of the underlying process.

Furthermore, for discrete time series, we can differentiate between *evenly* and *unevenly* spaced time series. Evenly spaced time series have constant time intervals between measurement points, whereas time intervals for unevenly spaced series can vary over time. Unevenly spaced time series are also called event-based time series [25], while evenly spaced series are called time-based records.

Most of the long-term clinical longitudinal data falls in the category of discrete, unevenly spaced longitudinal data, as time intervals between medical checkups may vary in most of the cases [26].

An additional distinction regarding the recorded values can be made into *qualitative* and *quantitative* data.

Quantitative data is always numerical. It arises when certain characteristics are measured or counted. The number of patients in an ambulance is quantitative data, since it involves a count of the number of patients. Equally, the blood pressure of a patient is quantitative data, since the answer involves measuring the blood pressure.

Qualitative data is information that ranks or labels items, but does not measure or count them. For instance, if information about the drug name that is used for medication in a certain therapy is collected from patients, that data would be qualitative. If patients are asked during a medical checkup, whether they feel “very well”, “well”, “average”, “bad” or “very bad”, their subjective health status is converted into a ranking. Therefore, qualitative data is generated also in this case.

Furthermore, depending on the recorded variable, measured values can be assigned to different type classes, namely *nominal*, *ordinal*, *interval* or *ratio* [27]. Thereby, the type defines the recorded data’s level of structure. In general, qualitative data is either nominal or ordinal, whereas quantitative data is either interval or ratio data.

Nominal data is the type with the least structure. Its values are simple labels that cannot be ordered or ranked in a meaningful way. The name of the drugs given to several patients would be such kind of data.

In contrast, ordinal data can be ordered or ranked, but does not measure or count any data characteristics. Questions about, e.g. the subjective health status or satisfaction level, generally involve a ranking.

Interval data does measure or count any characteristics, but ratios between two measured values have no intrinsic meaning. This applies for measurement scales, where the zero point does not describe a state of absence of a quantity (e.g. the absolute lowest value on a scale). Temperature measuring in degrees of Fahrenheit is an example of interval data.

Ratio data means that ratios between two measurement points have an intrinsic meaning. For instance, if one patient has a dosage of 400 mg and another patient a drug dosage of 200 mg, the former has a dosage that is twice as high as the second one.

Longitudinal data can also be classified by the number of independent quantities that are recorded for each observation. If a physician examines a patient and only records the heart rate, the data has just one independent quantity and is called univariate. On the other hand, data that involves more than one variable, is called multivariate. In special cases, when exactly two variables are measured, the data is called bivariate.

With regard to predictability of a time series we differentiate between *deterministic* and *stochastic* time series. If a time series can be predicted precisely, it is considered deterministic (e.g. if we look at the sinus wave). However, most of the time series fall in the category of stochastic time series [28]. Thereby, future events are only partly determined by past behavior and exact predictions are thus impossible and must be replaced by the idea that future values have a probability distribution that is conditioned by the knowledge of past values [21].

A time series is called *strictly stationary*, if the joint probability distribution does not change when shifted in time, i.e. $P_{t_1, \dots, t_n}(x_1, \dots, x_n) = P_{t_1+c, \dots, t_n+c}(x_{1+c}, \dots, x_{n+c})$.

A time series is called *weakly stationary* if $\mathbb{E}(x_t) = \mu$ and for the autocovariances $\mathbb{E}[(x_t - \mu)(x_{t-c} - \mu)] = \gamma_c$, $c = 0, \pm 1, \pm 2, \dots$

This means, that the parameter mean and variance do not change over time, or follow any trends.

2.2 Human Computer Interaction (HCI) & Interactive Visual Analytics

A good example of HCI research is visual analytics [29], which had its origins in the well-known *visual analytics mantra* (“overview first, zoom/ filter, details on demand” [30]) and Keim’s *extended mantra* (“Analyse First – Show the Important – Zoom, Filter and Analyse Further – Details on Demand” [29], [31]). Keim et al. [29] provided a formal description of the visual analytics process, following some notions of Van Wijk [32], which is highly relevant to solve problems involved with complex data: Let the input data sets be $S = S_1, \dots, S_m$ whereas each set $S_i, i \in (1, \dots, m)$, consists of attributes A_{i1}, \dots, A_{ik} . The output of this process will be called **Insight I** which is either obtained directly from a set of created **Visualizations V**, or through confirmation of stated **Hypotheses H** as the results of the analysis.

Formally this process can be described as a transformation $F: S \rightarrow I$, whereas F is a concatenation of functions $f \in \{D_W, V_X, H_Y, U_Z, \}$ defined as follows:

D_W describes the basic pre-processing functionality with $D_W: S \rightarrow S$ and $W \in \{T, C, SL, I\}$, including the data transformation functions D_T , data cleaning functions

D_C , data selection functions D_{SL} , and data integration functions D_I , which are required to make the analysis functions applicable to the data set.

$V_W, W \in \{S, H\}$ symbolizes the visualization functions, which are either functions visualizing the data $V_S: S \rightarrow V$ or functions visualizing hypotheses $V_H: H \rightarrow V$.

$H_Y, Y \in \{S, V\}$ represents the hypothesis generating process.

Important in this view is that Keim et al. distinguished between functions that generate hypotheses from data $H_S: S \rightarrow H$ and functions that generate hypotheses from visualizations $H_V: V \rightarrow H$.

Interactions $U_Z, Z \in \{V, H, CV, CH\}$ are thus an integral part in this formal model. Such interactions may either affect visualizations only (i.e. selecting, zooming, ...) $U_V: V \rightarrow V$, or the user interactions can affect hypotheses only by generating new hypotheses on the basis of the given ones $U_H: H \rightarrow H$.

Insight can be gained from visualizations $U_{CV}: V \rightarrow I$ or from the hypothesis itself $U_{CH}: H \rightarrow I$. This leads to the question: "What is interesting?" [33]. Close connected is the approach of attention routing [34] to overcome one critical problem in visual analytics: to help end users locate good starting points for analysis.

3 Longitudinal Data Visualization

Nowadays, the most common visualization techniques for longitudinal data include point charts, bar charts, line graphs, sequence graphs and circle graphs [35]. For a general overview on visualization techniques refer to [36] and an excellent work on the visualization of time-oriented data is [37]. In the following paragraphs, longitudinal data visualization techniques are briefly introduced.

Sequence graphs represent time-dependent data on one dimension by indicating each data point with a mark on the axis. The distance between each mark on the axis represents the time span passing between the events. With sequence charts, it is not possible to visualize a second dimension for a data point.

Figure 1 shows a sequence graph. In this example the graph visualizes treatment frequency information.

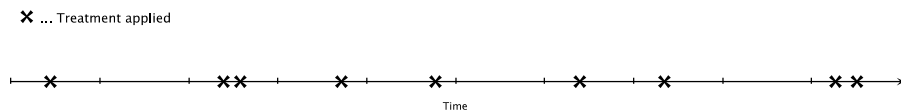


Figure 1 - Sequence graph

This mode of presentation allows an easy recognition of treatment pilings and longer time frames without treatment. Nevertheless, it is not possible to add additional information, like a dosage quantity related to the treatment, to this visualization.

Therefore, **point graphs** extend sequence charts by using a second axis to display a further information dimension. The distance from the main axis thereby represents the

second data dimension. Figure 2 shows a point chart. In this example the point chart visualizes treatment frequency and quantity information.

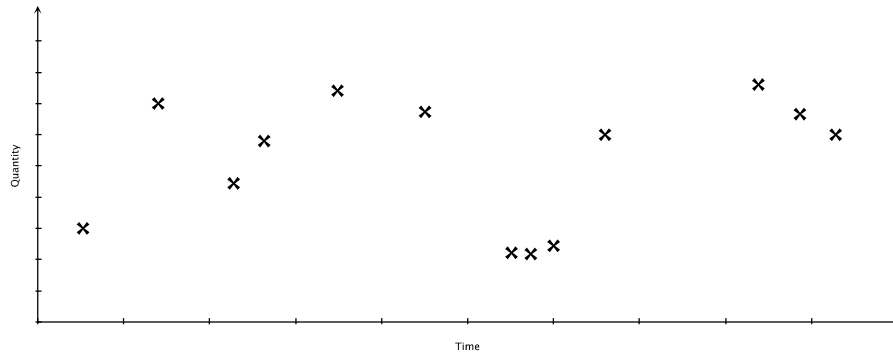


Figure 2 - Point graph

The vertical distance to the origin represents the quantity information, whereas the horizontal distance to the origin stands for the elapsed time. Point charts are useful to detect pilings within two-dimensional data.

Bar graphs replace the points with bars, which increases the comparability between the data points. Figure 3 shows the same information as represented in Figure 2, visualized as a bar chart. A comparison of these two visualizations shows that the ability to compare data point quantities is enhanced by using bar charts.

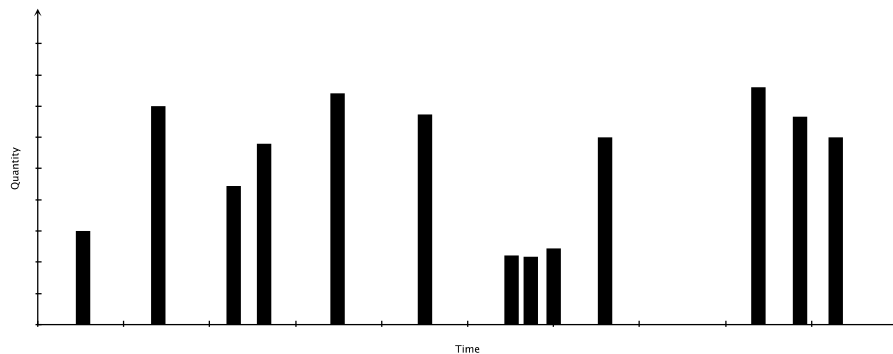


Figure 3 - Bar graph

Line graphs extend point charts by connecting the dots with lines to emphasize the temporal aspect of data. Line charts are very helpful for indicating trends over time.

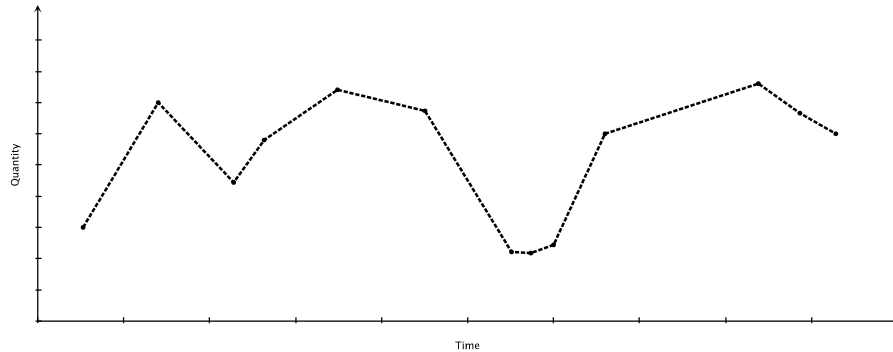


Figure 4 - Line graph

Circle graphs map the time series data into a spherical domain. They are commonly used to visualize periodic data with a known cycle length. Figure 5 shows an example of a circle graph, in which a fictive average internet usage time of a user is plotted as circle graph. As is shown in the picture the area within the data line is usually filled in circle graphs [38]. The shape of the area makes different periodic behavior comparable. Such approaches can be very helpful in the medical domain [39].

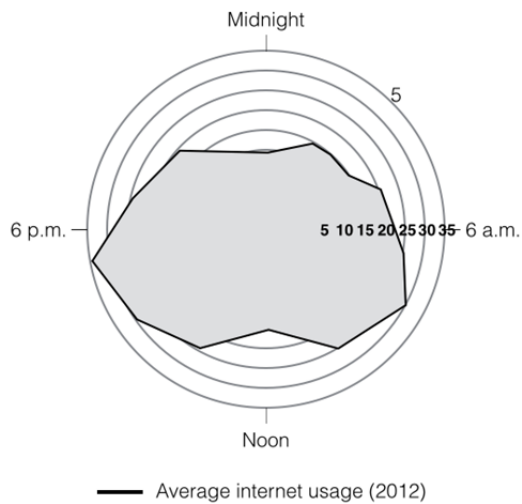


Figure 5 - Circle graph

According to Tufte [10], time series graphics are most suitable for big, complex data sets with real variability. Simple linear changes should better be summarized with one or two numbers. Tufte also introduced the following common guidelines that should be considered within every graphical display task:

- show the data
- induce the viewer to think about the substance, rather than about methodology, graphic design, the technology of graphic production, or something else
- avoid distorting what the data should express
- present many numbers in a small space
- make large data sets coherent
- encourage the eye to compare different pieces of data
- reveal the data at several levels of detail – from broad overview to fine structure
- serve a reasonably clear purpose: description, exploration, tabulation, or decoration.
- be closely integrated with the statistical and verbal descriptions of the data set.

To sum it up, the following table gives an overview on the different types of longitudinal data visualization and the corresponding use cases.

Table 1 - Overview of different methods of longitudinal data visualization

Type	Use case
Sequence-Graph	For one dimensional data (time); for detecting pilings within the datasets
Point-Graph	Two dimensional data (e.g. time/quantity); detecting pilings within two dimensional datasets
Bar-Graph	Two dimensional data; special focus on comparison of nearby data-points
Line-Graph	Two dimensional data; Emphasizing the temporal aspect of the data
Circle-Graph	Two dimensional data; displaying periodicity

Plotting Value-Ranges

At times, it can be necessary to visualize not only a single point, but a range of values for each observation. This can be the case, if the observed data points comprise some margin of error or if multiple observations are merged into one time point of the plot. The latter is often done with stock market data. As stock market prices vary over a whole trading day, valuable information would be lost if only the closing price was plotted on a daily basis. Therefore, not just closing, but day-open, day-close, day-high and day-low are visualized.

In the following paragraph some techniques for visualizing such range information will be introduced. Although many of the techniques are most frequently used with stock-market data, they can easily be applied in other areas of use.

Figure 6 shows three common visualizations for range information. The first way of introducing range into a plot is the incorporation of **error bars** [40]. The horizontal bar symbols indicate some lower and upper threshold for each value. Error bars provide no possibility to introduce more information, like day-start and -end values, into the visualization.



Figure 6 - Different techniques to incorporate range information, from left to right: error bars, envelopes and candlesticks. The examples are typical for use in financial market analysis.

A second way to include range information is by using **envelopes**. As shown in Figure 6, envelopes provide information about the range as well as day-start and -end values.

The same kind of information can be visualized with **candlesticks** [41]. Whereas envelopes encode the information about start and end values by adding a small line to the left and right of the range line, candlesticks use the coloring of a rectangle. Black coloring means that the top line of the rectangle is the start- and the bottom line is the end-value. White coloring inverts this convention.

Figure 7 shows a box plot visualization. **Box plots** [42] can provide information about the basic distribution properties of the data. The bottom and top of the box are first and third quartiles, and the band inside the box is the median. The ends of the lines, extending the box, are called whiskers. They can represent several possible alternative values. In this example, they stand for the minimum and maximum of the data set.

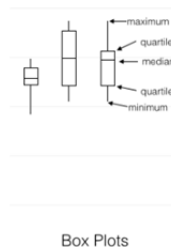


Figure 7 - Box plot

4 Longitudinal Data Interaction

As already emphasized in the introduction, besides choosing the right data-visualization it is also important to offer the user the possibility to interact with the displayed data in manifold ways. Especially within large-scale graphs, this interaction capability allows to see the overall picture but also to focus on interesting details.

According to Weber et al. (2001) [35], the following interaction methods can enhance the information perception process with the mentioned graph types.

- *Zooming* – initially, a high level overview of the course of the time series is given. By zooming, the user can obtain a more detailed view on a subarea of interest.
- *Scrolling* – if the area of display is not large enough to fit the whole chart (e.g. after zooming in), the user can scroll through areas.
- *Focusing and linking* – extends the idea of zooming by providing not only a zoomed-in version of the data, but also applying different, more effective visualization techniques to the zoomed-in dataset.
- *Brushing* – provides the idea of extended data visualization by automatically displaying pop-ups as a roll-over effect.
- *Filtering* – taking away (ignoring) irrelevant data objects.

When it comes to implementing these interaction methods, touch input has advantages compared to traditional point-and-click interfaces. This results from the fact, that compared to mouse-based devices, multi-touch devices provide a greater richness of interaction possibilities. Richness of interaction refers in this case to the degrees of freedom in interaction supported by the technology [43], [44]. This allows the UI designer to encapsulate the various graph interaction methods into easy to learn and quickly executable gestural commandos. Moreover, multiple tests with end users have yielded the confirmed result that complex mouse tasks, such as rotating an object and scaling, are faster done with multi-touch than with mouse-based devices [45]. Furthermore, well-designed gestural interfaces can shorten the learning curve by replacing a maze of menus and controls with simple actions, gestures, affordances and feedback [46]. Precondition to capture the benefits of the greater degree of freedom in interaction is a set of established gestural interaction guidelines. Nowadays, there are already some de facto gesture standards established for common tasks (see table 2).

Table 2 - Overview of gestural commandos for graph interaction

Gesture	Use case
Pinch gesture	Continuous zooming-out
Spread gesture	Continuous zooming-in
Double-tap gesture	Discrete zooming-in and discrete zooming-out (fast overview and detail)
One-finger swipe	Scrolling continuous back and forth
Cut-Gesture	Zooming into a discrete portion of the graph (market by the two fingers executing the cut)
Long-Press	Focusing on a data-point, showing detailed information

5 Open Problems

Some open problems which we discovered during our work include:

- How to select the appropriate axis parameters? This is maybe the grand major challenge when dealing with interactive visualizations of longitudinal data sets.
- How to deal with missing data sets? Uncertainty is another crucial issue, particular in the medical domain.
- With the increasing amount of data we also have to find solutions for scalability and platform issues, especially when dealing with big data analysis on mobile devices.
- Another issue to name is dealing with high-dimensional longitudinal data.
- Combining the visualization of changes over time as well as showing trends, predictions and correlations is an open problem, too.
- A further open problem is to find methods for choosing optimal data resolutions (sampling frequencies) for the recording process as well as for the visualization step. The resolution should be chosen in dependence of the anticipated data variability and the expected cycle length of patterns within the set.
- Furthermore, it must be discussed how to handle changing visualization resolutions when continuously zooming in or out. Especially, methods for merging multiple data points into one.
- Bridging the gap between mathematical analyzing and visual inspection of longitudinal datasets is another issue, more concise: Create a user interface, which supports end users to utilize visual information for adjusting model parameters.

6 Conclusion and Future Work

Dealing with large longitudinal data sets is a hot and promising topic and the application of advanced multi-touch interaction, e.g. graph-based interactions [47] are a starting point for a number of unsolved problems, particularly in the visualization of complex, multidimensional and multivariate data sets, e.g. in summarizing and showing statistics and correlation of specific time sections; to get more insight into the applicability of such approaches in the real-world, some work on information perception would also be needed, e.g. to answer the question on what is a meaningful maximum number of simultaneously displayed plots for comparison purposes.

Acknowledgements

We are grateful for fruitful discussions within our meetings of the hci4all.at group.

References

1. Holzinger, A., Dehmer, M., Jurisica, I.: Knowledge Discovery and interactive Data Mining in Bioinformatics - State-of-the-Art, future challenges and research directions. *BMC Bioinformatics* 15(Suppl 6), I1 (2014)
2. Holzinger, A., Stocker, C., Ofner, B., Prohaska, G., Brabenetz, A., Hofmann-Wellenhof, R.: Combining HCI, Natural Language Processing, and Knowledge Discovery - Potential of IBM Content Analytics as an assistive technology in the biomedical domain. *Springer Lecture Notes in Computer Science LNCS 7947*, pp. 13-24. Springer, Heidelberg, Berlin, New York (2013)
3. Holzinger, A.: On Knowledge Discovery and Interactive Intelligent Visualization of Biomedical Data - Challenges in Human-Computer Interaction & Biomedical Informatics. In: *DATA 2012*, pp. 9-20. INSTICC, (2012)
4. Holzinger, A.: Weakly Structured Data in Health-Informatics: The Challenge for Human-Computer Interaction. In: *Proceedings of INTERACT 2011 Workshop: Promoting and supporting healthy living by design*, pp. 5-7. IFIP, (2011)
5. Culler, D.E., Mulder, H.: Smart sensors to network the world. *Scientific American* 290(6), 84-91 (2004)
6. Ghrist, R., de Silva, V.: Homological sensor networks. *Notic. Amer. Math. Soc* 54(1), 10-17 (2007)
7. Esling, P., Agon, C.: Time-series data mining. *ACM Computing Surveys (CSUR)* 45(1), 12 (2012)
8. Enright, C.G., Madden, M.G., Madden, N., Laffey, J.G.: Clinical time series data analysis using mathematical models and DBNs. *Artificial Intelligence in Medicine*, pp. 159-168. Springer (2011)
9. Sriyudthsak, K., Iwata, M., Hirai, M.Y., Shiraiishi, F.: PENDISC: A Simple Method for Constructing a Mathematical Model from Time-Series Data of Metabolite Concentrations. *Bulletin of mathematical biology* 1-19 (2014)
10. Tufte, E.R.: *The Visual Display of Quantitative Information*. Graphics Press, Chesire (CT) (1983)
11. Mueller, H., Reihs, R., Zatloukal, K., Holzinger, A.: Analysis of biomedical data with multilevel glyphs. *BMC Bioinformatics* 15(Suppl 6), S5 (2014)
12. Holzinger, A.: Human-Computer Interaction & Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together? In: Cuzzocrea, A., Kittl, C., Simos, D.E., Weippl, E., Xu, L. (eds.) *Multidisciplinary Research and Practice for Information Systems*, Springer Lecture Notes in Computer Science LNCS 8127, pp. 319-328. Springer, Heidelberg, Berlin, New York (2013)
13. Grinstein, G., Ankerst, M., Keim, D.: "Visual Data Mining: Background, Applications, and Drug Discovery Applications", Tutorial at ACM SIGKDD2002, Edmonton, Canada (2002)
14. Keim, D.A.: Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics* 8(1), 1-8 (2002)
15. Beale, R.: Supporting serendipity: Using ambient intelligence to augment user exploration for data mining and Web browsing. *International Journal of Human-Computer Studies* 65(5), 421-433 (2007)
16. Otasek, D., Pastrello, C., Holzinger, A., Jurisica, I.: Visual Data Mining: Effective Exploration of the Biological Universe. In: Holzinger, A., Jurisica, I. (eds.) *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges*. Lecture Notes in Computer Science LNCS 8401, pp. 19-34. Springer, Heidelberg, Berlin (2014)

17. Gschwandtner, T., Gärtner, J., Aigner, W., Miksch, S.: A taxonomy of dirty time-oriented data. *Multidisciplinary Research and Practice for Information Systems*, Springer LNCS 7465, pp. 58-72. Springer, Heidelberg, Berlin (2012)
18. Harvey, A.C., Harvey, A.: *Time series models*. Harvester Wheatsheaf, New York (1993)
19. Hamilton, J.D.: *Time series analysis*. Princeton university press, Princeton (1994)
20. Box, G.E., Jenkins, G.M., Reinsel, G.C.: *Time series analysis: forecasting and control*. Fourth Edition. John Wiley & Sons, Hoboken (NJ) (2008)
21. Chatfield, C.: *The analysis of time series: an introduction*. Sixth Edition. CRC press, Boca Ration (FL) (2009)
22. Brockwell, P.J., Davis, R.A.: *Time series: theory and methods*. Springer (2009)
23. Shumway, R.H., Stoffer, D.S.: *Time series analysis and its applications: with R examples*. Third Edition. Springer, New York, Dordrecht, Heidelberg, London (2011)
24. Brockwell, P.: *Time Series*. In: Lovric, M. (ed.) *International Encyclopedia of Statistical Science*, pp. 1601-1605. Springer Berlin Heidelberg (2011)
25. Warner, R.M.: *Spectral analysis of time-series data*. Guilford Press (1998)
26. Simonic, K.M., Holzinger, A., Bloice, M., Hermann, J.: *Optimizing Long-Term Treatment of Rheumatoid Arthritis with Systematic Documentation*. In: *Proceedings of Pervasive Health - 5th International Conference on Pervasive Computing Technologies for Healthcare*, pp. 550-554. IEEE, (2011)
27. Stevens, S.S.: *On the theory of scales of measurement*. *Science* 103(677-680) (1946)
28. Gradišek, J., Siegert, S., Friedrich, R., Grabec, I.: *Analysis of time series from stochastic processes*. *Physical Review E* 62(3), 3146-3155 (2000)
29. Keim, D., Mansmann, F., Schneidewind, J., Thomas, J., Ziegler, H.: *Visual Analytics: Scope and Challenges*. In: Simoff, S., Böhlen, M., Mazeika, A. (eds.) *Visual Data Mining, Lecture Notes in Computer Science 4404*, pp. 76-90. Springer Berlin Heidelberg (2008)
30. Shneiderman, B.: *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations*. In: *Proceedings of the 1996 IEEE Symposium on Visual Languages*, pp. 336-343. (1996)
31. Keim, D., Kohlhammer, J., Ellis, G., Mansmann, F. (eds.): *Mastering the Information Age: Solving Problems with Visual Analytics*. Eurographics, Goslar (2010)
32. Van Wijk, J.J.: *The value of visualization*. In: *Visualization, 2005. VIS 05. IEEE*, pp. 79-86. IEEE, (2005)
33. Dervin, B.: *Sense-making theory and practice: an overview of user interests in knowledge seeking and use*. *J. Knowl. Manag.* 2(2), 36-46 (1998)
34. Chau, D.H., Myers, B., Faulring, A.: *What to do when search fails: finding information by association*. In: *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pp. 999-1008. ACM, (2008)
35. Weber, M., Alexa, M., Müller, W.: *Visualizing time-series on spirals*. In: *Information Visualization, IEEE Symposium on*, pp. 7-7. IEEE Computer Society, (2001)
36. Turkay, C., Jeanquartier, F., Holzinger, A., Hauser, H.: *On Computationally-enhanced Visual Analysis of Heterogeneous Data and its Application in Biomedical Informatics*. In: Holzinger, A., Jurisica, I. (eds.) *Interactive Knowledge Discovery and Data Mining: State-of-the-Art and Future Challenges in Biomedical Informatics. Lecture Notes in Computer Science LNCS 8401*, pp. 117-140. Springer, Berlin, Heidelberg (2014)
37. Aigner, W., Miksch, S., Schumann, H., Tominski, C.: *Visualization of Time-Oriented Data. Human-Computer Interaction Series*. Springer, London (2011)
38. Harris, R.L.: *Information graphics: A comprehensive illustrated reference*. Oxford University Press (1999)
39. Holzinger, A., Hoeller, M., Bloice, M., Urlesberger, B.: *Typical Problems with developing mobile applications for health care: Some lessons learned from developing*

- user-centered mobile applications in a hospital environment. In: International Conference on E-Business (ICE-B 2008), pp. 235-240. INSTICC, (2008)
40. Cumming, G., Fidler, F., Vaux, D.L.: Error bars in experimental biology. *The Journal of Cell Biology* 177(1), 7-11 (2007)
 41. Lee, K., Jo, G.: Expert system for predicting stock market timing using a candlestick chart. *Expert Systems with Applications* 16(4), 357-364 (1999)
 42. Williamson, D.F., Parker, R.A., Kendrick, J.S.: The box plot: a simple visual method to interpret data. *Ann. Intern. Med.* 110(11), 916-921 (1989)
 43. Hodges, S., Izadi, S., Butler, A., Rustemi, A., Buxton, B.: ThinSight: Versatile Multi-touch Sensing for Thin Form-factor Displays. In: *UIST 2007: Proceedings of the 20th Annual Acm Symposium on User Interface Software and Technology*, pp. 259-268. ACM, (2007)
 44. Buxton, B.: A Touching Story: A Personal Perspective on the History of Touch Interfaces Past and Future. In: *SID Symposium*, pp. 444-448. Wiley, (2010)
 45. Benko, H., Wilson, A.D., Baudisch, P.: Precise selection techniques for multi-touch screens. In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 1263-1272. ACM, (2006)
 46. Wigdor, D., Wixon, D.: *Brave NUI world: designing natural user interfaces for touch and gesture*. Morgan Kaufman, Burlington (MA) (2011)
 47. Holzinger, A., Ofner, B., Dehmer, M.: Multi-touch Graph-Based Interaction for Knowledge Discovery on Mobile Devices: State-of-the-Art and Future Challenges. In: Holzinger, A., Jurisica, I. (eds.) *Interactive Knowledge Discovery and Data Mining: State-of-the-Art and Future Challenges in Biomedical Informatics*, Springer Lecture Notes in Computer Science LNCS 8401, pp. 241–254. Springer, Berlin, Heidelberg (2014)