

Data Management: More Than a Matter of CS

Andreas Grillenberger and Ralf Romeike

Friedrich–Alexander–Universität Erlangen–Nürnberg (FAU)
Department of Computer Science, Computing Education Research Group
{andreas.grillenberger,ralf.romeike}@fau.de

1 Big Data and Data Management

Data management is one of the most exciting and challenging topics in computer science. Nowadays, handling large amounts of varying data in short processing times, which is Big Data, is intensively discussed not only in CS; it has also a tremendous influence on daily life and society, as handling data and dealing with the new chances and threats increases in relevance for everyone. However, data management is only a marginal topic in current education. Hence, we will outline the importance of this topic for school by presenting three aspects mainly affected by the current developments in this field: the paradigm change in data management, the influences of this field on everyday life and using it as a tool.

2 A new Paradigm in Data Management

Data management concerns with storing, managing and analyzing data, often using database management systems (DBMS). As today the requirements on such systems are clearly changing, new types of DBMS, summarized as NoSQL databases (“not only SQL” [1]), evolved. Hence, common concepts of DBMS are challenged, e. g. the relevance of redundancy: In RDBMS, preventing redundancies is a main aim, while NoSQL databases use redundancy intentionally for accelerating access to distributed stored data. The decision whether to handle or to prevent redundancy has to be made from case to case. Therefore, various publications state an ongoing paradigm change in data management (cf. [4]): in the future a great variety of DBMS can be expected.

3 Data Management in Daily Life

Today, people not only use but also generate and manage large amounts of data every day. These data need to be handled in a proper way, depending on their value. While private data need to be protected from unauthorized access, for example by encryption, in other cases protecting the device on which the data is stored is sufficient. Other requirements are coming from the large amount of data-driven devices and applications everyone uses: For example, the task of synchronizing data between devices includes several important aspects of data management like handling redundancy and inconsistencies. Therefore, the user is often confronted with phenomena like duplicat-

ed contacts, accidentally reverted changes or other synchronization conflicts that have to be resolved by hand. Summarizing, data management comprises various decisions everybody needs to make: where to store data, how to protect them, how to deal with threats for data privacy, and so on. As these tasks are strongly related to everyday life, data management becomes increasingly relevant for everyone.

4 Using Data Management as a Tool

The amount of publicly available data (“Open Data”) increases continuously. By analyzing such data, various information can be derived that are relevant for understanding decisions, political topics, and so on. These methods for analyzing data are not only relevant in CS and in daily life, but also in other fields and sciences. Not only the data, but also more and more easy-to-use tools for analyzing and visualizing them are made available. In addition, there are indicators that people want to understand such analyses, e. g. not only weather forecasts are provided as smartphone apps but also the satellite images used for deriving these forecasts. Today, when moving to a foreign city, one cannot only read others views on the possible neighborhood, but can in detail analyze data, provided for example by the city administration: e. g. New York City offers data on the calls to the service number 311 at their open data portal. When aggregating these by borough and mapping them, it is easy to discover boroughs with for example bad street conditions or high noise. As in the digital life more and more decisions base on the results of data analysis, this is an important method and competency. This relevance is not only limited to CS: as data analysis becomes increasingly important, even a new profession is rising—the “data scientist”.

5 Conclusions

Within the last years, data management clearly changed: not only does handling data follow a new paradigm, also its influences on everyday life become visible. In addition, it is used as a tool in other fields. These three aspects entail numerous new requirements for CS education (cf. [2]). By emphasizing data management at school, a main aim of general education can be reached: fostering knowledge, skills and competencies strongly related to everyone’s daily life (cf. [3]).

References

1. Edlich, S., Friedland, A., Hampe, J., Brauer, B., Brückner, M.: NoSQL [in German]. Hanser, Carl GmbH + Co. (2011)
2. Grillenberger, A., Romeike, R.: Big Data - Challenges for Computer Science Education. In: Proceedings of ISSEP 2014 (2014)
3. Grillenberger, A., Romeike, R.: Teaching Data Management: Key Competencies and Opportunities [in print]. In: Proceedings of KEYCIT 2014 (2014)
4. Mayer-Schönberger, V., Cukier, K.: Big Data: A Revolution That Will Transform How We Live, Work, and Think. Houghton Mifflin Harcourt (2013)