Digital Visual Exploration Library

Nicholas Tan Jerome and Andreas Kopmann
Institute for Data Processing and Electronics, Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany

Keywords: Digital Library, Visual Exploration, Scientific Visualisation, Information Visualisation.

Abstract: With the advancement of instrument precision, research facilities are generating data at an unprecedented rate. These experimental results are stored in a digital library platform which the contents are later accessible from within the facility or the public. However, the sheer volume of collected data is overwhelming the capacity of researchers and impedes the process of browsing for the desired data. In this paper, we present a concept of Digital Visual Exploration Library (DVEL) based on the confluence of two major research domains—digital library and visualisation—that enables efficient browsing of the growing data within a digital library. We complement the current state-of-the-art textual metadata description by integrating visual exploration to address big complex data, i.e., data of large size, multimodal data and multivariate data. We describe our concept based on use cases from three unique domains: climate research with Doppler wind lidar, X-ray-imaging for entomology research, and medical imaging with ultrasound computer tomography.

1 INTRODUCTION

We are entering an era of data-intensive science where significant data are produced by scientific experiments worldwide. These data hold the key to verify scientific hypothesis and having more data can help drawing more reliable research conclusions. Hence, numerous works focus on advancing the instrument precision which allows researchers to produce data at an unprecedented rate. For instance, the Large Hadron Collider (LHC) particle accelerator can generate 60 terabytes of data per day (Brumfiel, 2011). With LHC serving as an extreme case in the data-intensive science, such large data are no stranger in most experiments. Often, the data are stored in a digital library which the contents are accessible from within the research facility or from the public. By storing all the information into the digital library, the sheer volume of data complicates the data exploration process. The main problem lies with users having little knowledge of the large amounts of data.

To date, many scientific experiments are starting to provide open data access to the public (Molloy, 2011) which at the same time enables citizen science movement—citizens (non-scientists) taking part in real-world experiments. There is a growing demand to access scientific data from the public and improving the digital library framework is essential. To facilitate the exploration and analysis of a digital library collection (Mathew et al., 2017), researchers emphasise metadata analysis, i.e., word-frequency analysis (Shubankar et al., 2011), co-occurrence word analysis (Isenberg et al., 2017), and probabilistic methods like Latent Dirichlet Allocation (Griffiths and Steyvers, 2004). While these methods offer a generic solution, but they fail to provide information related to a particular domain (Mesbah et al., 2017). Moreover, if the metadata is not complete or not informative, users have no other options but to download a copy for offline verification which is cumbersome.

We propose instead to integrate visualisation techniques into the framework of a digital library—Digitial Visual Exploration Library (DVEL). In particular, we want to utilise visualisation to improve the data browsing experience. With little knowledge of the data, the visual presentation in the DVEL enables users to explore the data faster and with higher confidence. Since the user is the main actor throughout the exploration process, the interactive interface of the DVEL allows users to formulate hypothesis while adjusting their exploration goals (Keim, 2001). Keim outlined the benefits of visualisation techniques for browsing large data sets but has hitherto received little attention in the digital library domain.

In this paper, we discuss design considerations to provide appropriate visualisation interface as part of the DVEL framework. We present three use cases that illustrate the large data size, multimodal data and multivariate data. As digital library often provide its data through a web portal, we are inherently limited by the

Jerome, N. and Kopmann, A.
Digital Visual Exploration Library.
DOI: 10.5220/0006650003410348
In *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018) - Volume 3: IVAPP*, pages 341-348
ISBN: 978-989-758-289-9

Table 1: The main components of the web portal for a list of digital libraries in the Physics domain (Last visited 21-07-2017).

Experiment	**Navigation**	**Preview+Detail**	**Retrieve**
NASA Exoplanet Archive (Akeson et al., 2013)	List	Time Series	tbl
Sloan Digital Sky Survey (Eisenstein et al., 2011)	Guided search	2D Image	jpg
Crystallography Open Database (Gražulis et al., 2009)	List	Jmol (Hanson, 2010)	cif
CERN Open Data Portal (South et al., 2011)	List	—	aod

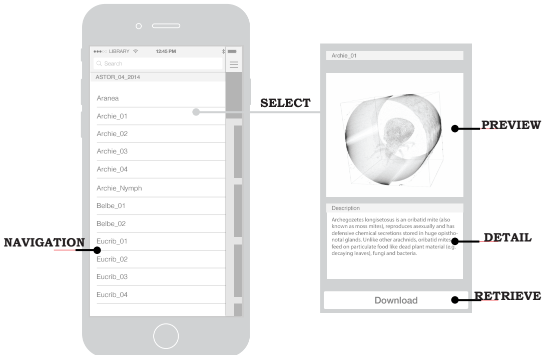

Figure 1: The main components in a web portal of the digital library. Upon selecting a data set, a 3D preview of the sample is shown together with its details.

bandwidth limitation. Hence, we perform data processing based on the data state reference model (Chi, 2000). In particular, we prepare a compact representation of the data which is small in size and yet best describe the data.

2 CONCEPT

The first use of the term *digital library* dated back in 1988 by (Kahn and Cerf, 1988), where they presented an open architecture of a digital library system that allows convenient access, either locally or remotely, to geographically distributed users. Back then, most information is not in digital form, and they described their architecture as an extension to the traditional library which users must travel physically to the library to utilise the resources (Kaur, 2015). However, such community facility is losing interest from the public in favour of digital technology advancement—a new paradigm dubbed the digital library. With information resources being digitalised, users are no longer required to travel physically to a library; but instead, they can access the information through any digitally connected medium, i.e. web browser.

Numerous definitions attempt to define the digital library (Cleveland, 1998; Lynch and Garcia-Molina, 1996; Seadle and Greifeneder, 2007; Trivedi, 2010), which can be broadly distinguished either by its physical space and collections, by its digital environment, or by its services. In this paper, we define library according to its primary purpose—a facility which provides public access to its materials. The term *digital* refers to the digital technology which includes the mobile phone, laptop, computer, network connection, and cloud service. Hence, a digital library consists of an extensive collection of digital information with public access. Similarly, research facilities store their experiment data in digital form which are accessible from within the institution or the public.

With open data providing a web portal for remote data access, users can search and download their desired data conveniently. We identified the main components in the web portal usage based on a set of collective actions to find the desired data set: to search (Navigation), to verify (Preview+Detail), and to download (Retrieve). Figure 1 shows the main components which are listed as below:

Navigation. A list and a search function to allow data browsing.

Preview+Detail. A page view presenting details—usually in tabular format—describing the selected data set. Optionally, an interactive data preview is present to emphasise essential features of the data set, i.e., an interactive 3D visualisation showing the spatial information.

Retrieve. The information of the data set can be acquired or downloaded in various formats.

We studied four experiments in the physics domain and discovered that the corresponding web portals mostly provide the Navigation and Retrieve components (Table 1). However, the Preview+Detail is not a prerequisite component and still mainly limited to textual descriptions. Instead, we highlight the unique features of each data set by integrating interactive visualisation techniques such as 3D visualisation (Levoy, 1988) or information visualisation (Buja et al., 1991). Although most information can be described in texts but conveying complex information solely with texts can be a daunting task. For instance, an interactive 3D visualisation would provide better spatial information. Taking the Crystallography Open Database as an example, the molecular structure is best described by an interactive 3D visualisation—Jmol (Hanson, 2010).

To support better data exploration, we propose to integrate the visual exploration technique as part of

Table 2: The data stages of each use cases following the data state reference model (Chi, 2000). The Value stage represents the raw data. The Analytic shows the post-processed data state. The Visual stage depicts the input formats required for the final visualisation (View).

Use case	Value	Analytic	Visual	View
NOVA (Sec. 4.1)	Image Database	Aggregated Database	Slicemaps	Interactive 3D
KITcube (Sec. 4.2)	Relational Database	Event Summary Database	Encoded Images	2D Heatmap, Plan Position Indicator (PPI), Range Height Indicator (RHI)
USCT (Sec. 4.3)	Image Database	Aggregated Database	Encoded Slicemaps	3D Fusion

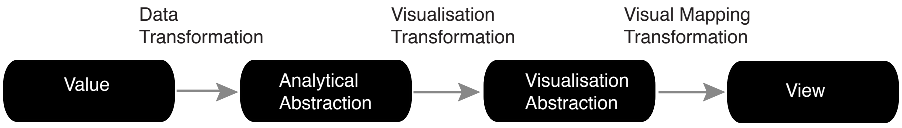

Figure 2: The data state reference model by (Chi, 2000). The black box denotes the state of the data, whereas the arrow shows the action performed from one state to another.

the Preview+Detail component within the DVEL framework. The DVEL can provide more information to complement any shortcomings from the textual metadata descriptions. Tominski described the visual exploration as an undirected search for relevant information within the data (Tominski, 2006). Tominski emphasised visual exploration as the process to provide an overview of the data, to browse through subsets of the data, and to interact with the data intensively. These requirements define what functions a useful visual exploration tool should have.

Within the DVEL framework, the visualisation component (Preview+Detail) acts as an add-on feature to enhance existing solutions. In particular, the viewing component strives to better inform users about the data without suffering from the scale of the massive data.

3 DESIGN STRATEGY

Our approach to design the DVEL revolves around two main considerations: technology limitations (data preprocessing) and the effectiveness of data exploration visualisation (three-step process). The technology limitations refer to the limited network bandwidth and the diverse client requirements.

The final visualisation should be served at an interactive rate. Hence, data transferred has to be small and yet still able to represent the original data. Moreover, the transferred data should match the diverse client requirements, i.e., mobile phone, laptop or desktop. To this end, we perform a data processing stage which follows the data state reference model (Chi, 2000) (Figure 2). We decided on the type of final visualisation and prepared the required input data accordingly. We reduce the data for a better network bandwidth (data reduction) and prepare a hierarchy of data with different data granularity to serve various client requirements. Table 2 shows the data stages of each use cases. The Value stage shows the raw data input which can be an image database or a relational database. Then, we perform data processing methods on the raw data such as data reduction and data fusion. Depending on the type of final visualisation, we transform the data into the required visual formats such as slicemaps (Tan Jerome et al., 2017b) or encoded images (Tan Jerome et al., 2017a).

To design the visual interface in the Preview+Detail component, we involve the domain experts in formulating primary features of the data through the human capabilities (Binnig et al., 2016). This approach allows the visual element to target the primary users' needs better. We follow the "information seeking mantra" (Shneiderman, 1996) which serves as our design guidelines—three-step process: (1) overview first, (2) zoom and filter, and (3) details on demand. In the overview step, we identify interesting patterns which best describe the data and enable users to drill down the details of the data.

4 APPLICATION SCENARIOS

In this section, we discuss the design strategies—data preprocessing and the three-step process—and appropriate visualisation techniques that fit into the Preview+Detail component in three different scenarios:

(1) visualising data of large size (gigabytes of data), (2) merging multimodal data, and (3) summarising multivariate data into a single representation.

4.1 Visualising Large Data

The NOVA[1] project (Schmelzle et al., 2017) serves as an ideal use case where hundreds of small animals such as insects and other arthropods are scanned using the synchrotron-based X-ray microtomography (SRμCT) imaging technique. These tiny animals are digitised and turned into large data sets with each size ranging from 30 GB to 120 GB. The data are stored in multiple collections of stack images where each stack of images represents volumetric data. The primary challenge lies in the size of the data, where transferring gigabytes of data through the network introduces high data latency.

Data Preprocessing. To overcome the network limitation, we reduce the data size but at the same time preserving the quality of the data. As we are limited by the network bandwidth, we adopted the multi-resolution slicemap approach to serving various clients (Tan Jerome et al., 2017b). The slicemap is a 3D data structure in the form of a mosaic-format image, which comprises a series of cross-section images. For better user experience, we first load a low-resolution slicemap (128^3 voxels) for instant visualisation and followed by a high-resolution slicemap (256^3 / 512^3 voxels) loading in the background (Figure 4: Tier I). However still, the multi-resolution slicemaps are reduced data which the quality is not comparable to the image rendered directly from the original image. To address this drawback, we utilise the resource of the High-Performance Computing (HPC) server to stream images which are rendered directly from the original data (Figure 4: Tier II).

Three-step Process. In this application, biologists would like to identify the type of small animals without suffering from the high data latency. We present two approaches: thumbnail preview (additive blending) and the two-tier approach—surface rendering on demand. The first approach is integrated into the Navigation component, where we provide a list of thumbnail previews (Figure 3A). The thumbnail preview is a 3D image that enables biologists to accurately identify the specimen. Since the image placeholder in the data list is small, we use primitive techniques for fast render time by downscaling the data and performing additive blending. To inspect the data further, we provide an interactive 3D visualisation with quality comparable to the desktop application (Figure 3B).

[1]NOVA stands for Network for Online Visualization and synergistic Analysis of tomographic data

4.2 Visualising Multivariate Data

In a large-scale experiment, not only the increasing data size is a concern, but also the complexity of the data. These data usually are multivariate data consisting of different attributes such as temperature and humidity (Kehrer and Hauser, 2013). Despite the data being highly heterogeneous, researchers must, however, combine these data with various attributes and formats to formulate and verify their hypotheses. Using the Doppler wind lidar data from the KITcube mobile observation platform (Kalthoff et al., 2013), we demonstrate the DVEL which allows meteorologists to browse through multivariate daily data using the Doppler wind lidar interactive visualisation (Tan Jerome et al., 2017a).

Data Preprocessing. We separate the multivariate data based on three different measurement types into the Event Summary Database. Each type consists of further three attributes, in which we then normalise them with a constant interval, i.e., one second. We then reduce the size and merge the attributes into an encoded image (Figure 5). Although the encoded images are in the 2D format, the type of final visualisation determines how the 2D image is mapped, i.e., PPI plot maps the image into a circular scheme, and RHI plot maps the image into a hemisphere plot. The mapping is performed interactively using the GPU resource through the WebGL shader.

Three-step Process. The main objective is to provide meteorologists with an overview of the daily trend, in which we integrate three plots to study the wind velocity: Plan Position Indicator (PPI) scans, Range Height Indicator (RHI) scans, and Vertical Stare. In contrast to the previous list view, we design an aggregated view where we project daily information on the screen display (Figure 6). In the visual analytics, we use the daily vertical stare plot as the overview to detect interesting patterns. We also provide hourly view which shows a more detailed vertical stare, RHI and PPI plots. By choosing a particular date (Navigation component), meteorologists can analyse the daily overview interactively utilising the brushing technique (Spence, 2001) on multiple attributes.

4.3 Visualising Multimodal Data

When multiple sources are present, we describe the multivariate data as multimodal data. In the medical domain, multimodal data stem from different imaging techniques such as CT, MRI, or ultrasound

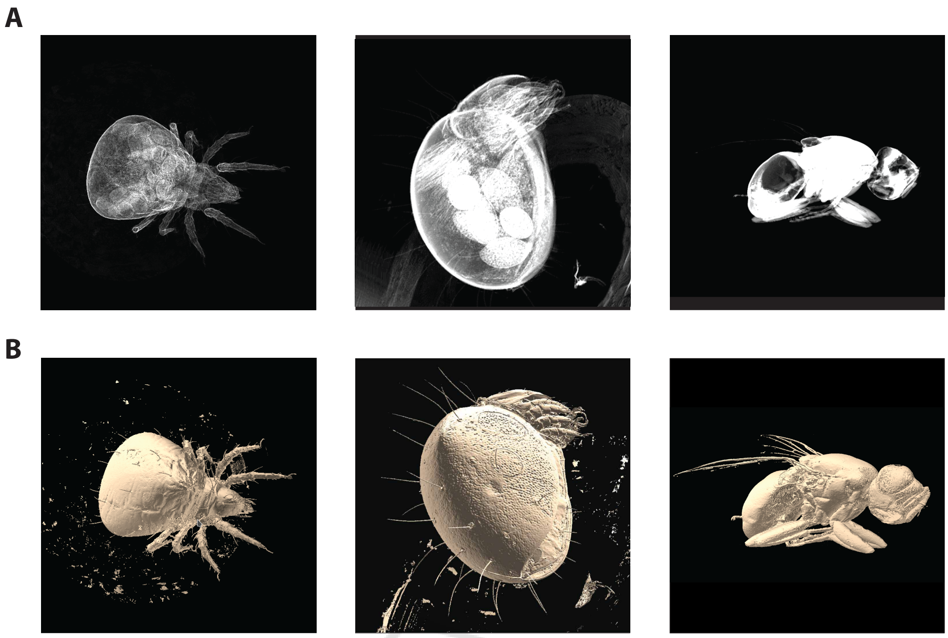

Figure 3: Three examples of the arthropods in the NOVA project rendered using (A) the additive blending and (B) the surface rendering approaches: the oribatid mite *Archegozetes longisetosus* (left column), the box mite *Euphthiracarus reticulatus* (middle column) and an unidentified Diptera species *fly* (right column).

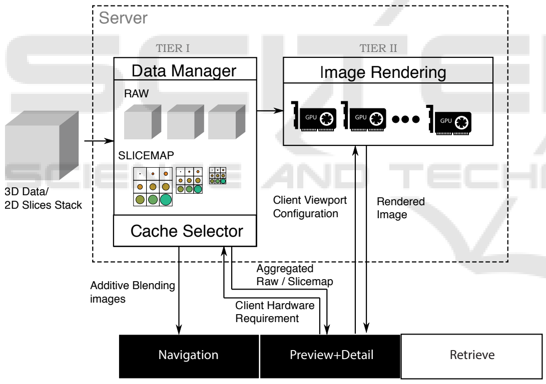

Figure 4: The data preparation of the original data for final visualisation at the client side. A set of preview images rendered using the additive blending approach is provided to the Navigation component. For the Preview+Detail, a set of high quality rendered images are streamed to the client.

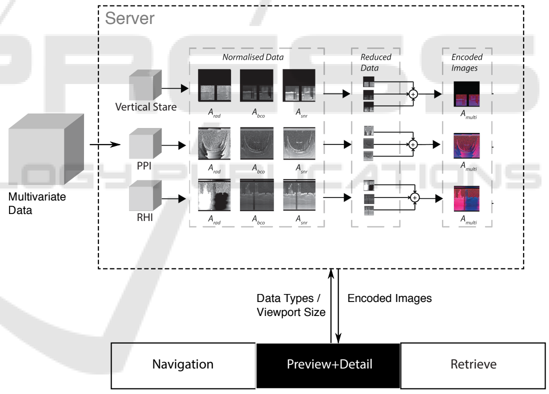

Figure 5: The data preparation of the multivariate data for final visualisation at the client side (Tan Jerome et al., 2017a). The multivariate data are processed at the server before serving to the clients.

data (Kehrer and Hauser, 2013).

Using the USCT [2] project as our use case (Gemmeke et al., 2017), we have a digital library filled with multimodal ultrasound data. The USCT project aims to develop new image methodology for early breast cancer detection. The USCT detector produces three different types of images, which are the reflection image, the sound speed image and the attenuation image. The sound speed and attenuation images allow doctors or scientists to classify lesions precisely, whereas the reflection image contains the information of the breast

[2] USCT stands for 3D Ultrasound Computer Tomography.

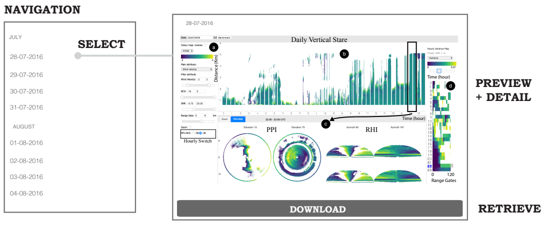

Figure 6: An aggregated view of the Doppler wind lidar data where each item on the Navigation component shows a daily overview of the wind velocity which consists of vertical stare, RHI and PPI plots.

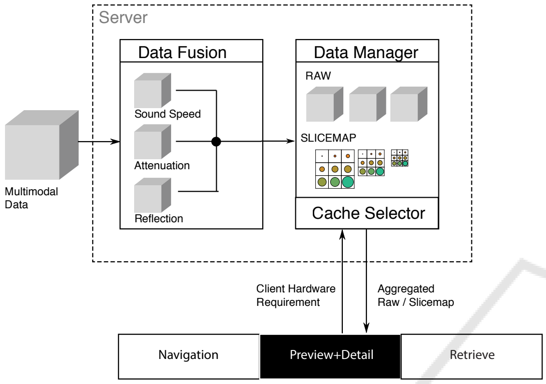

Figure 7: The main components in a web portal of the digital library. Upon selecting a data set, a 3D preview of the sample is shown together with its details.

structure (Hopp et al., 2015).

Data Preprocessing. Similar to the data processing in Section 4.1, we perform data reduction and data transformation on the original data to produce a series of slicemaps. Also, we encode the three imaging modalities into the slicemap which results in an encoded slicemaps (data fusion).

Three-step Process. The goal of this application is to enable doctors to identify possible tumour regions. Rather than studying each modality separately, we fuse the multimodal data and render the final image using the normal fusion approach (Stokking et al., 2001). The multimodal data are fused according to the Ranger's fusion scheme (Ranger et al., 2012). In Figure 8 shows the multimodal view of a data set, where:

- Cancer (orange): Sound image thresholded at $(1.52 \pm 0.03) \times 10^{-3}\,\mathrm{km\,s^{-1}}$ and attenuation image thresholded at $(0.16 \pm 0.04) \times 10^{-3}\,\mathrm{dB\,cm^{-1}}$.
- Fibroadenoma (green): Sound image thresholded at $(1.52 \pm 0.03) \times 10^{-3}\,\mathrm{km\,s^{-1}}$ and attenuation image not in the thresholded region of $(0.16 \pm 0.04) \times 10^{-3}\,\mathrm{dB\,cm^{-1}}$.
- Fibrous bands or architecture (grey): The whole reflection image range.

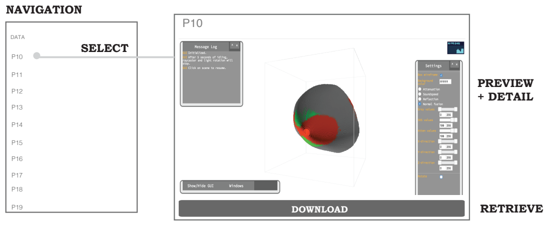

Figure 8: The DVEL for the multimodal USCT breast dataset. The orange region shows the area with high sound speed values suggesting possible tumour tissues. The green region is the auxiliary attribute (attenuation) to support the deduction of the tumour possibility.

We transform the fused dataset into 3D formats and perform surface rendering to obtain an illustration depicting the possible tumour region.

5 DISCUSSION AND CONCLUSIONS

The continuously growing data in a digital library will pose a challenge in data exploration. Thus, supports to enable users finding their data from the exhaustive list are increasingly important. We proposed Data Visual Exploration Library (DVEL) that integrates visual exploration techniques as part of the digital library framework, where the visual display can represent the prominent feature of the data better (Keim, 2001). We presented three different use cases and showed the relevance of our concept with the framework of the digital library.

Due to the bandwidth limitation, we applied a personalised processing which reduces the memory footprint and yet retains the identity of the data. Also, the smaller data enable a more responsive visual interaction on the client. Furthermore, we utilise the GPU resource by offloading graphics and data rendering at the client. Although image rendered from the original large data shows much more detail, the purpose of our concept is to allow fast data exploration and ultimately enable users to narrow down their search.

From our experience, the DVEL concept introduces another type of data to be stored—a compact visual data. Rather than storing only the original data, we need to define a compact visual data which can be either made of a single attribute or a collection of attributes. The compact visual data has to be designed by domain experts for better visual communication. The preparation of the visual data requires close collaboration among researchers of different domains which might be cumbersome.

Although our use cases differ in its visualisation purpose, we can set up a design strategy base on the data state reference model (Chi, 2000) and the three-step process (Shneiderman, 1996). However, the current state of the design strategy relies heavily on empirical approaches which are not scalable. To this end, designing a software tool that encompasses the design strategy proposed can be attractive.

As the implementation of the DVEL in the three use cases—NOVA (Schmelzle et al., 2017), KIT-cube (Kalthoff et al., 2013), USCT (Gemmeke et al., 2017)—are still at its early stage, user feedbacks from the DVEL remains as an interesting component to shape the future research challenges. In this paper, we described mainly on the concept and design of the DVEL but not the relationship between the meta data and the visual components.

ACKNOWLEDGEMENT

We would like to thank the German Federal Ministry of Education and Research for funding of ASTOR and NOVA (grants 05K13VK5 and 05K16VKB, respectively). We also express our gratitude to Sebastian Schmelzle, Michael Heethof, Thomas van de Kamp, Torsten Hopp, and Michael Zapf for their helpful insights.

REFERENCES

Akeson, R., Chen, X., Ciardi, D., Crane, M., Good, J., Harbut, M., Jackson, E., Kane, S., Laity, A., Leifer, S., et al. (2013). The nasa exoplanet archive: data and tools for exoplanet research. *Publications of the Astronomical Society of the Pacific*, 125(930):989.

Binnig, C., Fekete, A., and Nandi, A., editors (2016). *HILDA '16: Proceedings of the Workshop on Human-In-the-Loop Data Analytics*, New York, NY, USA. ACM.

Brumfiel, G. (2011). High-energy physics: Down the petabyte highway. *Nature News*, 469(7330):282–283.

Buja, A., McDonald, J. A., Michalak, J., and Stuetzle, W. (1991). Interactive data visualization using focusing and linking. In *Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on*, pages 156–163. IEEE.

Chi, E. H.-h. (2000). A taxonomy of visualization techniques using the data state reference model. In *Information Visualization, 2000. InfoVis 2000. IEEE Symposium on*, pages 69–75. IEEE.

Cleveland, G. (1998). *Digital libraries: definitions, issues and challenges*. IFLA, Universal dataflow and telecommunications core programme.

Eisenstein, D. J., Weinberg, D. H., Agol, E., Aihara, H., Prieto, C. A., Anderson, S. F., Arns, J. A., Aubourg, É., Bailey, S., Balbinot, E., et al. (2011). Sdss-iii: Massive spectroscopic surveys of the distant universe, the milky way, and extra-solar planetary systems. *The Astronomical Journal*, 142(3):72.

Gemmeke, H., Hopp, T., Zapf, M., Kaiser, C., and Ruiter, N. V. (2017). 3d ultrasound computer tomography: Hardware setup, reconstruction methods and first clinical results. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*.

Gražulis, S., Chateigner, D., Downs, R. T., Yokochi, A. F. T., Quirós, M., Lutterotti, L., Manakova, E., Butkus, J., Moeck, P., and Le Bail, A. (2009). Crystallography Open Database – an open-access collection of crystal structures. *Journal of Applied Crystallography*, 42(4):726–729.

Griffiths, T. L. and Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl 1):5228–5235.

Hanson, R. M. (2010). Jmol–a paradigm shift in crystallographic visualization. *Journal of Applied Crystallography*, 43(5):1250–1260.

Hopp, T., Duric, N., and Ruiter, N. V. (2015). Image fusion of ultrasound computer tomography volumes with x-ray mammograms using a biomechanical model based 2d/3d registration. *Computerized Medical Imaging and Graphics*, 40:170–181.

Isenberg, P., Isenberg, T., Sedlmair, M., Chen, J., and Möller, T. (2017). Visualization as seen through its research paper keywords. *IEEE transactions on visualization and computer graphics*, 23(1):771–780.

Kahn, R. E. and Cerf, V. G. (1988). *An open architecture for a digital library system and a plan for its development*. Corporation for National Research Initiatives.

Kalthoff, N., Adler, B., Wieser, A., Kohler, M., Träumner, K., Handwerker, J., Corsmeier, U., Khodayar, S., Lambert, D., Kopmann, A., et al. (2013). Kitcube–a mobile observation platform for convection studies deployed during hymex. *Meteorologische Zeitschrift*, 22(6):633–647.

Kaur, H. (2015). Role of Digital Libraries in the Present Era: Challenges and Issues. In Thanuskodi, S., editor, *Handbook of Research on Inventive Digital Tools for Collection Management and Development in Modern Libraries*, pages 86–102. IGI Global, Hershey, PA, USA.

Kehrer, J. and Hauser, H. (2013). Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE transactions on visualization and computer graphics*, 19(3):495–513.

Keim, D. A. (2001). Visual exploration of large data sets. *Communications of the ACM*, 44(8):38–44.

Levoy, M. (1988). Display of surfaces from volume data. *IEEE Computer graphics and Applications*, 8(3):29–37.

Lynch, C. and Garcia-Molina, H. (1996). Interoperability, scaling, and the digital libraries research agenda. *Microcomputers for Information Management*, 13(2):85–132.

Mathew, G., Agrawal, A., and Menzies, T. (2017). Trends in topics at se conferences (1993–2013). In *Proceedings* *of the 39th International Conference on Software Engineering Companion*, pages 397–398. IEEE Press.

Mesbah, S., Fragkeskos, K., Lofi, C., Bozzon, A., and Houben, G.-J. (2017). Facet embeddings for explorative analytics in digital libraries. In *International Conference on Theory and Practice of Digital Libraries*, pages 86–99. Springer.

Molloy, J. C. (2011). The open knowledge foundation: open data means better science. *PLoS biology*, 9(12):e1001195.

Ranger, B., Littrup, P. J., Duric, N., Chandiwala-Mody, P., Li, C., Schmidt, S., and Lupinacci, J. (2012). Breast ultrasound tomography versus mri for clinical display of anatomy and tumor rendering: Preliminary results. *American Journal of Roentgenology*, 198(1):233–239.

Schmelzle, S., Heethoff, M., Heuveline, V., Lösel, P., Becker, J., Beckmann, F., Schluenzen, F., U. Hammel, J., Kopmann, A., Mexner, W., Vogelgesang, M., Tan Jerome, N., Betz, O., Beutel, R., Wipfler, B., Blanke, A., Harzsch, S., Hörnig, M., Baumbach, T., and van de Kamp, T. (2017). The nova project: maximizing beam time efficiency through synergistic analyses of srμct data. In *Developments in X-Ray Tomography XI*, volume 10391, pages 10391 – 10391 – 17. International Society for Optics and Photonics.

Seadle, M. and Greifeneder, E. (2007). Defining a digital library. *Library Hi Tech*, 25(2):169–173.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pages 336–343. IEEE.

Shubankar, K., Singh, A., and Pudi, V. (2011). A frequent keyword-set based algorithm for topic modeling and clustering of research papers. In *Data Mining and Optimization (DMO), 2011 3rd Conference on*, pages 96–102. IEEE.

South, D. M., Group, I. D. S., et al. (2011). Data preservation in high energy physics. In *Journal of Physics: Conference Series*, volume 331, page 012005. IOP Publishing.

Spence, R. (2001). *Information visualization*, volume 1. Springer.

Stokking, R., Zuiderveld, K. J., and Viergever, M. A. (2001). Integrated volume visualization of functional image data and anatomical surfaces using normal fusion. *Human Brain Mapping*, 12(4):203–218.

Tan Jerome, N., Chilingaryan, S., Kopmann, A., and Wieser, A. (2017a). An interactive web-based doppler wind lidar visualisation system. In *Workshop on Visualisation in Environmental Sciences (EnvirVis)*. The Eurographics Association.

Tan Jerome, N., Chilingaryan, S., Shkarin, A., Kopmann, A., Zapf, M., Lizin, A., and Bergmann, T. (2017b). WAVE: A 3d online previewing framework for big data archives. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 3: IVAPP, (VISIGRAPP 2017)*, pages 152–163. INSTICC, ScitePress.

Tominski, C. (2006). *Event based visualization for user centered visual analysis.* PhD thesis, University of Rostock.

Trivedi, M. (2010). Digital libraries: functionality, usability, and accessibility. *Library Philosophy and Practice (e-journal)*, page 381.