ISSN 2186-7437

NII Shonan Meeting Report

No. 173

TAT: Toughening the Foundation of Abstraction in Visualization Techniques

Organizers:
Hsiang-Yun Wu (St. Pölten University of Applied Sciences, Austria and TU Wien, Austria)
Tobias Isenberg (Inria, France)
Ivan Viola (KAUST, Saudi Arabia)
Xiaoru Yuan (Peking University, China)

September 11–14, 2023

National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo, Japan

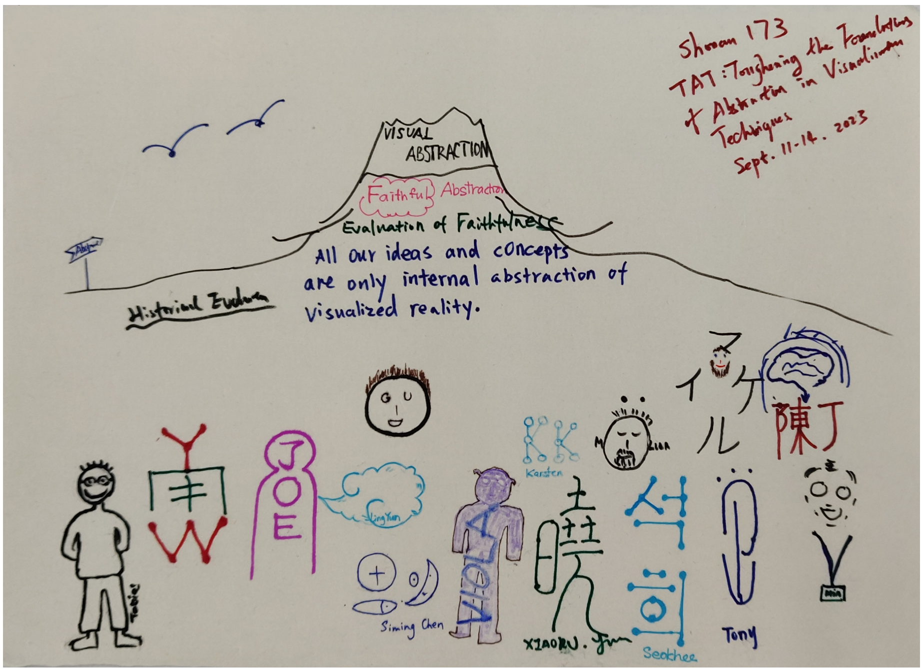
TAT: Toughening the Foundation of Abstraction in Visualization Techniques

Organizers:
Hsiang-Yun Wu (St. Pölten University of Applied Sciences, Austria and TU Wien, Austria)
Tobias Isenberg (Inria, France)
Ivan Viola (KAUST, Saudi Arabia)
Xiaoru Yuan (Peking University, China)

September 11–14, 2023

Abstract

In this meeting we discussed the concept of visual abstraction in visualization. Abstraction is essential to the field yet we have been using the term "abstraction" largely by intuition. Based on presentations from representatives from the field that cover a large variety of backgrounds and research focuses, in the meeting we worked toward a formalized understanding of the concept to better clarify its theory, its potential, and its uses.



Background and Introduction

Abstraction is considered as a conceptual process, whose outcome stands for the corresponding subordinate concepts as a whole. This representative notion allows users to facilitate comprehensive understanding and better memorability of the relevant complex knowledge in a more hierarchical or structured fashion. The concept of abstraction is especially significant when the information space is large and when it comes to real-world applications. The Tube Map of London Underground, for instance, is a classical map that removes unnecessary detail but retains sufficient information for better usability. The map successfully simplifies the geometry of the transportation structure, and facilitates users to effectively perform their tasks, which often includes (1) which route provides the optimal path in terms of distance or price, (2) how many stops still remain until the destination, (3) where to transfer to another line, and so on. Other than cartography, abstraction is powerful in many application domains, such as mathematics, biology, physics, etc., because experts in these fields can focus on vital elements of their data and think through the problem on a more abstract level.

Similarly, in data visualization, abstraction plays an important role because the outcome of the abstraction process needs to reflect what an expert has in mind and the transition process between a set of different visual representations should support the comprehension of the high-level concept in various aspects. To visualize essential aspects of the data, however, we need dedicated mechanisms that abstract the unnecessary detail to allow the viewer of a visualization to focus on the important elements, depending on the given task. Viola and Isenberg [26, 27] have done an initial investigation on this topic and confirmed that the crucial problem in this context is that it is impossible to know what is important and what is not in a general way—importance changes based on the research question, on the application domain, on the data size, on the user, on the specific situation or task, etc. Visualization techniques, therefore, are expected to support a dynamic change of data's visual abstraction to reflect these contextual changes.

Beyond 2D representations such as metro maps, Rautek et al. [20] and other researchers have established an understanding of visual abstraction for spatial 3D data to classify low-level and high-level visual abstraction techniques, while the support of reasoning and insight communication for abstraction techniques is still a fundamental challenge that remains today. Recently, Viola and Isenberg [26, 27] have intensively explored the concept of abstraction as it is used in visualization. They realized that researchers so far have used the concept of abstraction largely by intuition without a precise meaning, and thus initialized the pioneering discussion on theoretical foundations of abstraction in visualization. This lack of specificity left questions on the characteristics of abstraction, its variants, its control, or its ultimate potential for visualization and, in particular, illustrative visualization mostly unanswered. After this investigation, several open questions are still awaiting to be answered, while it requires not only visualization researchers, but also domain experts in order to formulate a research agenda for the practical usage of abstraction for the visual representation and exploration of data.

With the growth of data complexity, the need for abstraction techniques is increasing. Although there have been a few approaches studied along this line, the theory, technology, and applications have so far integrated less well in terms of the stability and usability of abstraction models for visual analytics. These challenges concern approach scalability, the applicability of the models in application areas, as well as the technology of the environment in which the abstraction is performed. In summary, the theory and application of abstraction in visualization poses a research challenge not only due to the complex nature of the data, but also its dynamics and semantics. In this pioneering Shonan meeting we thus gathered researchers from visualization, information theory, and applied science to address the topic.

We have invited every participant to suggest any topic they would like to focus on during the meeting. All topics suggested by the participants, detailed through abstracts in the next section, extend the visual abstraction formalism in the four distinct categories:

- **Theoretical Formalization of the Abstraction Concept as it relates to Visualization:** The term abstract is an antonym of concrete or tangible, resulting in an inherent difficulty to describe it. Viola and Isenberg have initiated the definition of abstraction as a transformation which preserves one or more key concepts and removes detail. Together with Chen [26], they have linked the initial definitions with abstraction in philosophy of science and have proposed to quantify it by means of the information theory. During our Shonan meeting we added a new essential component of the formalism by introducing the notion of *meaningfulness*.

- **Abstraction Techniques:** During the Shonan meeting, several presentations introduced abstraction in the context of specific aspects of the visualization pipeline. Some abstraction techniques were centered on specific type of data, such as the networks or graphs for example. Other presentations showcased visual abstractions from a historical perspective, or by using specific visual encoding strategy, such as visualization sequence or glyphs. We discussed multiple *axes* of visual abstraction, and the role of abstraction in user interaction.

- **Visual Abstraction Evaluation:** A key theme of the meeting was related to abstraction evaluation and quantification. We have built upon the concept of *faithfulness* [16] which measures how a visualization faithfully represents ground truth information of the abstract data.

  Inspired from this existing concept we define the abstraction *meaningfulness*. Furthermore, we have discussed how the information theory toolbox can quantify the degree of visual abstraction.

- **Visual Abstraction Applications:** Visual abstraction can serve as a means to reduce various *costs*, such as costs of computing, costs of drawing, costs of interpreting, or costs of communicating. This implies opportunities where abstraction can be highly beneficial: in characterization, standardization and exchange formats of visualization design, in data and visualization provenance, in dealing with contextual and cultural biases, or missing data.

Overview of Talks

Thoughts on (Visual) Abstraction

Ivan Viola; KAUST, Saudi Arabia

In my talk I introduce basic definitions of abstraction and terminology that relates abstraction to visualization. First, I clarify its use in other disciplines, such as fine arts, geography, mathematics, programming paradigms, and I present how abstraction is rooted in the philosophy of science.

Axes of (Visual) Abstraction

Tobias Isenberg; Inria, France

I talk about the notion of abstraction axes that emerge in many visualization papers/contributions that look at different levels of scale of a given subject matter. I will base this discussion on our earlier formalization attempts of visual abstraction and will use examples from the visualization literature to illustrate my discussion.

Visual Abstraction from a Historical View

Xiaoru Yuan; Peking University, China

In this talk, I discuss the broad spectrum of visual abstraction from the perspective of historical visualizations. More specifically, I also share a few examples of earlier stage visualizations from China.

A Formal Definition of Abstraction

Min Chen, University of Oxford, United Kingdom

In my talk, I will examine a general definition of "abstraction" (including computational and cognitive abstraction) from an information-theoretic perspective. In particular, I will discuss the ability for information theory to help explain, in abstraction, various phenomena in visualization and visual analytics. Explanation is the first step in any theoretical development. Hopefully, this work will stimulate further theoretical developments in Measurement and Prediction.

Abstraction and Bias

Guido Reina, University of Stuttgart, Germany

Visualization and any subsequent abstraction need to perform some kind of reduction. I am interested in the fine line we walk on our way to insights: between preservation of features, guidance, story-telling, and the potential biasing of the user. I hypothesize we need to better define and capture the intent or task correspondence for visualizations and abstractions, given the loss of generality inherent to these aspects. This is compounded when we take into account focus and context techniques: the context moves on a continuum between informer and distractor, but the coupling to the focus and its perception and interpretation make these two very hard to disentangle and set up appropriately in a general way.

Exploring Abstraction in Immersive Environments: Navigating Multiscale Visualizations and User Interaction

Lingyun Yu, Xi'an Jiaotong-Liverpool University, China

Visual abstraction focuses on distinct features or essential concepts within complex structures. It aims to emphasize key information while simplifying details, thereby facilitating effective data analysis and communication. Immersive environments create captivating shared spaces, enabling multiple users to engage closely with data visualization. However, these environments also pose challenges when it comes to exploring multiscale data. This presentation will address fundamental questions concerning immersive multiscale visualizations and user interaction: 1) Navigation in Multiscale Environments: How do we allow people to navigate freely through a multiscale visualization space without confusing them? This is particularly relevant in the context of multiuser and multiscale visualizations within shared environments; 2) Adaptive Interaction for Multiscale Visualizations: What design strategies promote natural interaction across various levels of abstraction; 3) Designing Immersive Interaction Techniques for Multiscale Visualizations: Prioritizing User Needs and Enhancing Understanding.

Visualization Sequence as an Abstraction Technique

Puripant Ruchikachorn, Chulalongkorn Business School, Thailand

A visualization sequence, especially in the form of an animation, is generally considered to be a visual representation of temporal data. However, non-temporal data, such as multidimensional data, is often shown in a sequence, such as a slideshow, in which each visualization represents a subset of the data. The process of transforming data into a visualization can occur gradually, and the intermediate representations can be shown in sequence as well. Even in a dashboard with multiple views, a viewer typically observes one visualization at a time.

Based on prior work, a visualization sequence may need a compositor similar to one used for multidimensional data abstraction. It should also avoid the same pitfalls as other abstraction techniques, such as being misinterpreted as a temporal sequence or having inconsistent visualization types across a sequence. Still, most research on visual abstraction has been for a single display, so the potential of visualization sequence as a tool of visual abstraction is unexplored. These are some questions that need to be addressed. When data is split into several chunks, are their permutations cognitively different? Are there any references in other fields on particular visualization sequences? Is there any information-theoretic cost or benefit of certain sequencing strategies?

Visual Abstraction for Network Representation

Karsten Klein, University of Konstanz, Germany

Networks are used as a data model in many areas, in particular in the life sciences. A plethora of research has been conducted to advance network analysis and visualisation over several decades. However, the increasing scale, complexity, and heterogeneity of the underlying data as well as the developments in the hardware and software available for visual analysis, such as immersive environments, have rather increased the number of questions than to close the gap of data complexity and user capacity to perceive and understand it. Abstraction can support understanding but also introduce issues regarding confidence and faithfulness. I gave a short overview on potential contributions of abstraction research, some pointers to challenges and requirements, and discussed my viewpoint on several related aspects.

Faithful Abstraction for Big Complex Graphs

Seok-Hee Hong, University of Sydney, Australia

In this talk, I will discuss several recent approaches and examples for Data Abstraction, Task Abstraction and Visual Abstraction for big complex graphs.

More specifically, Data abstraction methods include filtering, sampling, dimension reduction and clustering. Task abstraction examples include the metro map visualisation and clustered graph visualisation. Visual abstraction examples include edge bundling methods and graph map representations.

I will conclude my talk with *faithful abstraction* including Information faithfulness, Task faithfulness and Change faithfulness.

Towards measuring meaning in / meaningful abstractions

Torsten Möller, University of Vienna, Austria

If we'd like to know whether abstractions are meaningful, we would need to ask 'For Who?' and then how to measure them. This is the connection to visual (data) literacy, which is slowly becoming of interest to the visualization community.

I want to demonstrate a recent representative survey of the Austrian population with a bar chart and a line chart as found in news articles. I will argue that we are far from properly assessing such a multi-dimensional concept as 'meaning.' Previous work has focused on task-based assessment. I argue to extend this assessment to several aspects of self-perception, including complexity and abstraction, graph familiarity, aesthetics, critical thinking, topic knowledge, and numeracy. There are lots of open questions, which hopefully stimulate a discussion.

Building Visual Metaphors for Visual Abstraction

Siming Chen, Fudan University, China

Visual metaphors and glyphs play an important role in visualization for understanding complex data. However, how to create suitable and novel metaphors is a challenging research question. With my previous experiences in designing map and bridge metaphors, I would like to discuss the following issues: 1) What is the design space for using visual metaphors and glyphs for abstraction? How to make use of the metaphor for readers to understand complex information in a simplified and vivid way? 2) How to design and build visual metaphors and glyphs for visual abstraction based on the design space? Moreover, which visual channel and data mining methods can better be integrated for designing the metaphors? 3) How to leverage recent AI models to automatically/semi-automatically create the metaphors for visual abstraction? I will discuss my opinions through the above three points.

Abstract Visualizations in Life Sciences: Bridging Insights Across Complex Data Landscapes and Scientific Areas

Michael Schwärzler, Takeda - Pharmaceutical Sciences, Japan

Abstraction in visualization techniques for life sciences offers a valuable approach for comprehending intricate biological systems and discovering candidates for new treatments. After a molecule has been selected as a candidate for a product, it enters the *pharmaceutical process development stage*, where stakeholders deal with large amounts of heterogeneous data sources generated over a decade or more until market entry. This data, spanning structured and unstructured formats, includes molecular structures, time series data from devices, sample results, image data, chromatography curves, experimental documentation, but also budgetary considerations, strategic objectives, and more.

While the involved personas apply a multitude of advanced analytics and abstract visualization techniques to solve their tasks in their respective area of responsibility, the variety of perspectives and the differing understandings pose a major challenge - even though they all have a common goal and work on the same data. Additionally, varying visual abstraction techniques, cultural influences, and the loss of knowledge over time (due to personnel turnover or outdated technology) play roles. While "data silos" have become less of a problem, "knowledge silos" have become the most prominent challenge today, and missing visual abstraction standards and frameworks play a role in that. In fact, exchange and decision making happens mostly on a rather poor level of abstraction: using short summaries in PowerPoint presentations, giving up data provenance and the possibility to getting a deeper understanding of the underlying data.

At the data level, the use of knowledge graphs generated with taxonomies and ontologies aids in defining structures and relationships. However, finding ways to harmonize insight-generating abstract visualizations so that the barriers for knowledge exchange can be lowered as much as possible could have the potential to really achieve a coherent overview. Proposing a scientific definition of abstraction in visualization could establish harmonized standards or "visual knowledge interfaces". These would guide crafting of easier interpretable, meaningful abstract visualizations for stakeholders in highly heterogenous data contexts generated over a long time span for decision making on all levels. In my talk, I give an overview on the area, its data and visualization techniques, and trigger a discussion regarding observed aspects that might have a relevance for the theoretical foundations of abstraction for visualization techniques.

Abstraction by Human Visual Intelligence

Jian Chen, The Ohio State University, USA

I present how the human visual system simplifies information across different temporal scales to facilitate the processing of complex scenes. (1) At the scale of 200-500 milliseconds, global scene gist acts as a holistic feature representation. (2) Between 1 second and 3 seconds, memorable experiences often align with linguistic concepts. (3) Within a minute of viewing very large images, we can extract valuable signals to uncover meaningful patterns in the data. I provide empirical evidence demonstrating how we can harness these innate cognitive abilities to visualize complex simulation results, with the ultimate goal of understanding viewer behavior and developing a visual scene vocabulary.

Evaluation of Abstraction

Weidong Huang, University of Technology Sydney, Australia

As Viola and Isenberg mentioned for visualization, "the abstraction serves the goal of facilitating the understanding of the subject matter". Abstraction can be implemented in different forms in different contexts for different purposes for a given dataset. Once it is done, the question is whether or not an abstraction has achieved its goal. How do we know that? Does abstraction filter unwanted information and include only the information that we want? Does it introduce new or misleading information? To answer these questions, we need evaluation. What are the appropriate metrics, how do we measure them and how do we evaluate them? In this talk, I will present some challenges and share my views on these.

Quantifying Visual Abstraction: Visual Complexity or Data Insight Fidelity?

Yong Wang, Singapore Management University, Singapore

Visual abstraction is the process of transforming data into visual representations that can reveal data insights to viewers. For a given dataset, there can be multiple ways of visual abstraction, leading to different visual representations. But how abstract is a visual abstraction? Can we quantitatively compare the degree of abstraction of different visual abstractions? In my presentation, I will briefly discuss the two ways that seem able to quantify the abstraction degree of different visual abstractions: visual complexity and data insight fidelity. Further, I use concrete examples to illustrate that the measurement of abstraction degree of different visual abstractions essentially depends on the data insight(s) we intend to explore (or convey to viewers). Accordingly, I conjecture that visual complexity is probably NOT a good way to quantify the abstraction degree of visual abstractions, and data insight fidelity is a promising measurement to quantify visual abstraction. An overall framework for calculating the abstraction degree is also proposed by mathematically modeling the visual encoding process, the visualization decoding process and the data insight fidelity between insights existing in the original dataset and data insights kept in the decoded data. The visualization decoding process via human perception relies on 1) the visualization resulted from the visual abstraction process and 2) relevant factors of visualization viewers like their familiarity with data visualization, prior knowledge and/or culture background. Given that the relevant factors of human viewers are difficult to model at current stage, I would propose modeling human perception process by using image processing techniques (e.g., object detection and segmentation) to extract data from visualizations, which can lead to a simplified but practical framework for quantifying the abstraction degree of visual abstractions.

List of Participants

- Jian Chen, The Ohio State University, USA
- Min Chen, University of Oxford, United Kingdom
- Siming Chen, Fudan University, China
- Seok-Hee Hong, University of Sydney, Australia
- Weidong Huang, University of Technology Sydney, Australia
- Tobias Isenberg, Inria, France
- Karsten Klein, University of Konstanz, Germany
- Torsten Möller, University of Vienna, Austria
- Guido Reina, University of Stuttgart, Germany
- Puripant Ruchikachorn, Chulalongkorn Business School, Thailand
- Michael Schwärzler, Takeda - Pharmaceutical Sciences, Japan
- Ivan Viola, King Abdullah University of Science and Technology (KAUST), Saudi Arabia
- Yong Wang, Singapore Management University, Singapore
- Lingyun Yu, Xi'an Jiaotong-Liverpool University, China
- Xiaoru Yuan, Peking University, China

Meeting Schedule

Check-in day: September 10 (Sun)

- Welcome banquet

Day 1: September 11 (Mon)

- Talks and discussions
- Group photo shooting

Day 2: September 12 (Tue)

- Talks and discussions
- Break-out discussions

Day 3: September 13 (Wed)

- Break-out discussions
- Initial report writing
- Excursion and Main Banquet

Day 4: September 14 (Thu)

- Report writing
- Wrap-up

Summary of discussions

The following text reflects the state of our discussions at the meeting. The text is not meant to represent a final proposal of a new theory in the field of visualization, it is rather a draft that will be revised into a scientific article, where the nomenclature together with the relationships might be subjected to change.

1 Foundations

Before we can discuss particular abstraction techniques, we first need to establish the concept of abstraction. Figures 3 to 7 show several examples from related work that have already performed this process prior to an actual formalization of it.

Definition

The definition of *abstraction*, as given by Viola and Isenberg [26, 27], was that

> An abstraction is a process that transforms a source thing into a less concrete sign thing of the source thing. Abstraction uses a concept of point-of-view, which determines which aspects of the source thing should be preserved in its sign thing and which should be suppressed.

Three aspects related to this definition require clarity to avoid misunderstandings:

- **Transformations:** What is a transformation in the sense of the definition?
- **Reduction:** What is meant with *suppressing*?
- **Meaning:** The concept of a *point-of-view* is not clearly defined, but seems to be the essential ingredient for bringing in a cognitive aspect of the term *abstraction*.

Hence, we propose to properly define these three aspects and tie them into the existing discussion of the process of visualization from the visualization community.

1.1 Transformations T

Data Transformations are part of any data analysis process. Whether it is a change of basis function for metric spaces (such as Fourier- or Wavelet Transforms), a one-hot encoding for text data, or graph diffusion processes, these transformations can be expressed through a general (mathematical) mapping operation T:

$$T : \mathbb{A} \to \mathbb{A}' \tag{1}$$

Here, the spaces $\mathbb{A}$ and $\mathbb{A}'$ are seen in the most general way. Hence, this understanding is not constrained to *Data* transformations per se but includes *visual encoding* transformations just as well as transformations in the visual space. Hence, it properly ties into the wealth of work on the *Data Visualization Process / Pipeline* as already discussed by Card and Mackinlay [2] or Heer [4] up to such recent treatments by Kindlmann and Scheidegger [7] and van Wijk [25].

1.2 Reduction R

One crucial aspect of *abstraction* is that there will be *aspects* that are *suppressed*. Hence, some kind of reduction in data or information plays a central role. Therefore, we define a *reduction* operator R as a transformation, which is *surjective*. Hence, the space of all reduction operators is a proper sub-class of all transformation operators T, i.e., $R \subset T$.

Important examples of reduction operators include clustering operators, dimensionality reduction, as well as filtering.

As a part of reduction, new, transformed *reduced* data is generated from some *original* data. This new data set is strictly smaller than the original data set. However, while the reduction process compresses the data, if we consider original data together with the newly generated data, the overall data has actually increased. We highlight that the reduction only secures that the reduced data is a smaller set from the original data, while both of these data can be utilized in subsequent data processing. Nonetheless, as through abstraction, new data is created, the underlying amount of *information* $\mathbb{I}$ does not increase.

1.3 Meaningfulness M

An essential concept of *abstraction*, as used within the visualization community, is the fact that it carries *meaning* for an individual within the context of a task or embedded within a mental model or cognitive construct. This is expressed as the *point-of-view* aspect in the original definition of *abstraction*. In order to conceptually cast it as a formal concept, we need to introduce the knowledge space $\mathbb{K}$ that would be required to assign *meaning* to some data. Hence, we define the meaningfulness M as a mapping from space $\mathbb{A}$ to a scalar value between zero and one with respect to some knowledge $\mathbb{K}$:

$$\mathbb{M} : \mathbb{A} \bigoplus \mathbb{K} \underset{\mathbb{T}}{\rightarrow} [0, 1] \tag{2}$$

Hence, assigning *meaningfulness* clearly is a form of reduction to a very special space (the unit interval). Hence, it is a proper subset of all reduction operators $M \subset R$. Abstraction A is a proper subset of all reductions R, i.e., $A \subset R : \mathbb{A} \bigoplus \mathbb{K} > 0$. Some reductions R might not have any meaning ($= 0$) for a particular task $\mathbb{T}$ and knowledge $\mathbb{K}$ in mind. Such reductions are not abstractions. Abstraction A is a reduction that is at least somewhat meaningful (> 0).

1.4 Coupling Reduction and Meaning

A complementary approach to the above would be the coupling of both concepts. We start from the assumption that some data $\mathbb{D}$ contains a certain amount of information $\mathbb{I}$, and with respect to a task, this information will carry some meaning $\mathbb{M}$ for the user such that $M(\mathbb{I},\mathbb{T}) > 0$. Note that, as already outlined in section 1.3, meaning is undefined unless there is a specific task.

Given that an abstraction reduces information, $A : \mathbb{I} \rightarrow \mathbb{I}' + \mathbb{R}$,with $\mathbb{I}' \subset \mathbb{I}$, we can exactly quantify the reduced/hidden/lost information as $\mathbb{R}$. In this light, the design of an appropriate abstraction has to optimize user effort (cognitive load) while retaining as much meaning as possible: $arg\,max\mathbb{M}(\mathbb{I}',\mathbb{T})$. Conversely, we can say that an accurate abstraction for a task does not lose what we now can define as *meaningful information*, or that the reduced information should be irrelevant for the task: $\mathbb{M}(\mathbb{R},\mathbb{T}) \rightarrow 0$.

1.5 Faithfulness F

Closely related to our *meaningfulness* quantity, is previously defined concept of *faithfulness*. The *faithfulness* metrics measure how a visual representation $\mathbb{V}$ accurately represents the ground truth information of an abstract data $\mathbb{D}$ [16]. These metrics complement existing human perception based quality metrics (i.e., how human perceive and understand the visualizations), for example, the *readability* metrics such as edge crossings in network visualization.

Roughly speaking, the outcome $\mathbb{V}$ of visualization V is *information faithful*, if the ground truth information of $\mathbb{D}$ can be reconstructed from $\mathbb{V}$. $\mathbb{V}$ is *task faithful*, if it displays sufficient information to perform a specific task $\mathbb{T}$ on $\mathbb{D}$ accurately. For example, the *cluster faithful* metrics measure how a visualization V accurately represents the ground truth cluster of an abstract data D as a geometric cluster in V [12].

For the purpose of abstraction, the faithfulness model can be extended to include abstraction faithfulness and perception faithfulness, see Figure 1.

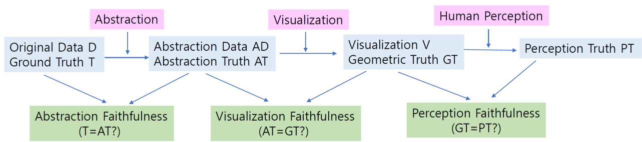


Figure 1: Faithfulness pipeline for abstraction.

1.6 The process of abstraction

Now that we have a stronger formal footing of abstraction, we embed it into the concept of the visualization pipeline. First, data about a particular phenomenon or system under investigation are collected, e.g. by measuring certain characteristics or by running a simulation code. When we abstract, we initially reduce our data $\mathbb{D}$ by first removing items or part of their attributes $A_1 : \mathbb{D} \underset{\mathbb{T}}{\rightarrow} \mathbb{D}_1$ based on the knowledge (or assumption) that a specific task $\mathbb{T}$ at hand does not require the information contained within them at the available level of detail[1].

[1]Please note that *data abstraction* as defined by Munzner [15], refers to various data types and data set types. In Munzner's nomenclature, these could have been termed data idioms, consistently with the visualization idioms. In our definition we refer to *data abstraction* as to a meaningful data-reducing transformation

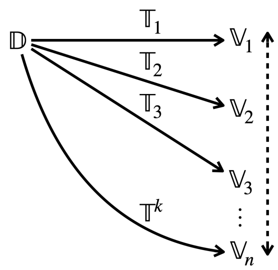
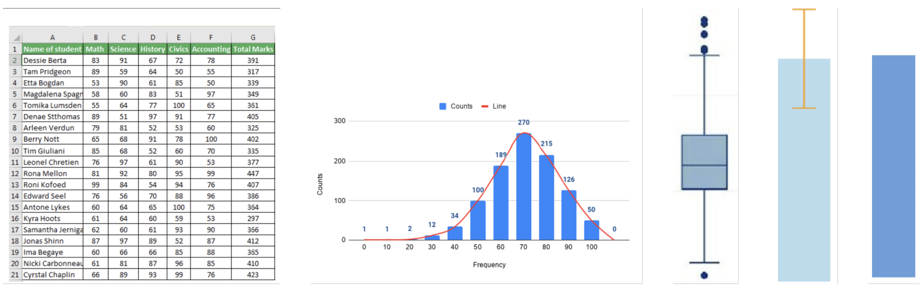


Figure 2: Data $\mathbb{D}$ can be abstracted to different visual representations $\mathbb{V}_i$ given certain tasks $\mathbb{T}_i$. As shown by the vertical dashed arrow, there can be connections between the visual representations in the visual-abstraction space as well—typically in the form of (human-constructed!) visual explanations that were called abstraction axes or abstraction space in the past [26, 27]. The mental operations in these abstraction space can be supported by superposition, animation, or morphing through the use of a metaphor.



	A	B	C	D	E	F	G
1	Name of student	Math	Science	History	Civics	Accounting	Total Marks
2	Dessie Berta	83	91	67	72	78	391
3	Tam Pridgeon	89	59	64	50	55	317
4	Etta Bogdan	53	90	61	85	50	339
5	Magdalena Spagn	58	60	83	51	97	349
6	Tomika Lumsden	55	64	77	100	65	361
7	Denae Stthomas	89	51	97	91	77	405
8	Arleen Verdun	79	81	52	53	60	325
9	Berry Nott	65	68	91	78	100	402
10	Tim Giuliani	85	68	52	60	70	335
11	Leonel Chretien	76	97	61	90	53	377
12	Rona Mellon	81	92	80	95	99	447
13	Roni Kofoed	99	84	54	94	76	407
14	Edward Seel	76	56	70	88	96	386
15	Antone Lykes	60	64	65	100	75	364
16	Kyra Hoots	61	64	60	59	53	297
17	Samantha Jerniga	62	60	61	93	90	366
18	Jonas Shinn	87	97	89	52	87	412
19	Ima Begaye	60	66	66	85	88	365
20	Nicki Carbonneau	61	81	87	96	85	410
21	Cyrstal Chaplin	66	89	93	99	76	423

Figure 3: Series of different chart visualization of information, each shows a different purpose in communicating statistical results.

This can be caused, for example, by the sheer volume of data, or the high frequencies contained within it. Then we perform some operation that transforms the reduced data (or part of it) into another form $A_2 : \mathbb{D}_1 \underset{\mathbb{T}}{\rightarrow} \mathbb{D}_2$. The union of all data resulting from *abstractions* can be expressed as $\mathbb{D}' = \mathbb{D}_2 \cup \mathbb{D}_1 \cup \mathbb{D}$ and all this digital representation of information $\mathbb{I}$ can be used in the next steps of the analytical or visualization pipeline. When we perform visualization V of this new form $\mathbb{D}'$, we end up with a visual representation $\mathbb{V}$ that has a reduced complexity in the sense that it is easier to interpret, faster to render or is more expressive for conveying information related to task $\mathbb{T}$:

$$V : \mathbb{D}' \underset{\mathbb{T}}{\rightarrow} \mathbb{V} \tag{3}$$

In case, this *visual transformation* does not show the full set of information contained in the original data $\mathbb{D}'$, and it *meaningfully* presents relevant information with respect to the task, we call this transformation a *visual abstraction*. This means, even if all information from $\mathbb{D}_i$ is visually conveyed, and no reduction happens during the transformation $\mathbb{D}_i \underset{\mathbb{T}}{\rightarrow} \mathbb{V}$, the composite meaningful reduction $\mathbb{D} \underset{\mathbb{T}}{\rightarrow} \mathbb{D}_i \cdots \underset{\mathbb{T}}{\rightarrow} \mathbb{V}$ is a visual abstraction. As a part of this process, each

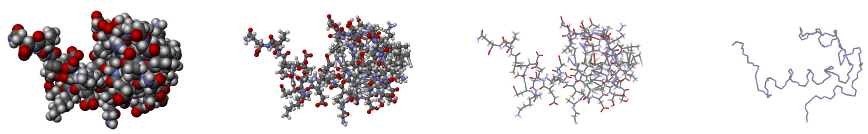


Figure 4: Series of different views of structural abstraction of a molecule, each generated with a differently parameterized visualization pipeline that individually focused on different aspects of the molecular data. Images from [24], images are in the public domain.

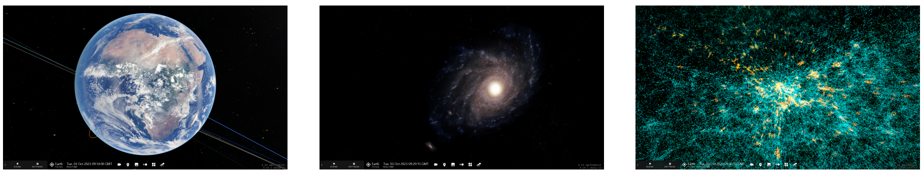


Figure 5: Series of views from OpenSpace [1], showing different levels of detail when zooming out from the earth over the milky way to the extents of the Sloan Digital Sky Survey.

time we visualize data, some form of abstraction A typically takes place. In each realistic scenario where visualization is utilized, abstraction, or several abstractions take place between observation of a given phenomenon up to the synthesis of the rendered representation. Also note that, strictly speaking, abstraction augments the data in that it creates additional abstracted representations $\mathbb{D}_i$ or $\mathbb{V}$ that are needed for the task, similarly as a Level of Detail (LOD) scheme or a MIP map image pyramid. We can keep the original data to facilitate switching between visualizations, or remove the unnecessary data $\mathbb{D} \setminus \{\mathbb{D}_i\}$.

1.7 Characterizing Meaningfulness

Depending on the task $\mathbb{T}$, a task-dependent ground truth $\mathbb{D}_{\mathbb{T}}$ can be derived from $\mathbb{D}$. There can be potentially many task-dependent ground truths $\mathbb{D}_{\mathbb{T}_i}$ depending on specific task $\mathbb{T}_i$. One can think of $\mathbb{D}_{\mathbb{T}}$ as the digital representation of the information $\mathbb{I}'$ that needs to be communicated to the user to accomplish the task $\mathbb{T}$, i.e. the subset of the information that carries most of the meaning relevant for the task (as outlined in section 1.4): $arg\, max M(\mathbb{I}', \mathbb{T})$. The task-dependent ground truth is normally the last part of the visualization pipeline, somewhat close to the *insight* or to the *take-home message*. Task-dependent ground truth $\mathbb{D}_{\mathbb{T}}$ will serve the purpose of understanding and quantifying the meaningfulness of an abstraction. It represents a theoretical construct that contains only data highly relevant to a particular task and nothing else. It represents the entire information a user needs to comprehend from visualization, nothing more, nothing less. While several closely related concepts such as expressiveness [10], appropriateness [17], and faithfulness [16] describe the relation of visually represented data and ground truth data, these concepts do not include human perception and interpretation of the visualization.

We say that a visualization V is meaningful if the visualization user gains at

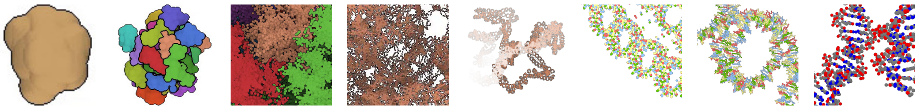


Figure 6: Series of different depictions of the DNA, from the cell nucleus to individual atoms. Each view shows a different level of detail and a different subset of the data. Images by Halladjian et al. [3],

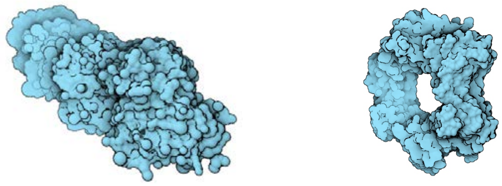


Figure 7: Examples of visual representations that transition between different visualization parameterizations depending on the distance from the viewer. These can be seen as results from individual visualization pipelines that are finally composited into a single view. Images from [18], used with permission.

least some information contained in the task-dependent ground truth $\mathbb{D}_{\mathbb{T}}$. This can be measured by any of the facets of understanding: explanation, interpretation, application, perspective, empathy, and self-knowledge [28]. For example, the *meaningfullness* can be measured by psychophysical studies that quantify the match between $\mathbb{D}_{\mathbb{T}}$ and human $\mathbb{H}$ reconstruction of $\mathbb{D}_{\mathbb{V},\mathbb{H}}$ after viewing $\mathbb{V}$. $\mathbb{D}_{\mathbb{V},\mathbb{H}}$ is the reconstructed task-relevant data that the user verbally reported on, has sketched, or has correctly identified in the context of a multiple-choice test. Such measurement inherently includes perception and cognition and forms the feedback loop shown in Figure 8.

1.8 Abstraction Characteristics

Abstractions along the visualization pipeline affect the domain $\mathbb{A}$ of the attributes, for example by compressing, binning, or lifting an attribute. Therefore, we also need to define $A : \mathbb{A} \underset{\mathbb{T}}{\rightarrow} \mathbb{A}'$. This effect can be observed in Figure 3, where the histogram is a binning of one of the columns in the table and the boxplot compresses all bins into a single one that represents the whole domain, using less screen space but a more complex mark (box, line, whiskers, outliers instead of one box per bin). As defined above, the reduction is a subjective transformation, which leads to the fact that the co-domain is smaller than the domain. Therefore, *domain shrinking* is a natural characteristic of an abstraction process.

In the exemplary visualizations depicted in Figures 3–7 we can see that abstractions can have multiple levels, and we can either have distinct processes that perform visualization directly from different abstractions A in an independent fashion, but can also have abstractions that operate on the result of previous ones $A_i : \mathbb{D}_{i-1} \underset{\mathbb{T}}{\rightarrow} \mathbb{D}_i$ that could result in different useful visual representations $\mathbb{V}$ on different levels, if applied partially. The latter gives us a sequence

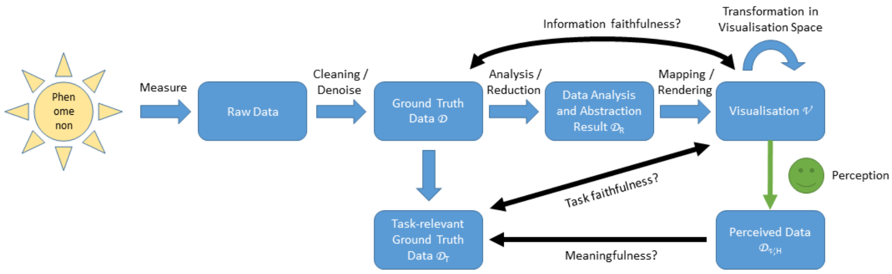


Figure 8: Data Processing and Visualisation Pipeline: Data from a phenomenon under investigation is collected and after pre-processing a ground truth data set $\mathbb{D}$ is derived. A non-empty part of this ground truth, $\mathbb{D}_{\mathbb{T}}$, will be task-relevant. The result of data analysis and abstraction, $\mathbb{D}'$, is mapped into the visualisation space where it is perceived by the human, resulting in the perceived data $\mathbb{D}_{\mathbb{V},\mathbb{H}}$ for which we can evaluate meaningfulness with respect to $\mathbb{D}_{\mathbb{T}}$.

of gradually increasing visual abstractions of the underlying data $\mathbb{D}$. But even without an increasing abstraction or increasing reduction of data, the different representations may be placed (by a human) in a meaningful order such that we can transition between them, in a way to explain a subject matter to others or to record insights gained from some data exploration. This is indicated by the vertical dashed arrow in Figure 2—and which were called axes of abstraction in the past [26, 27].

One very interesting challenge that relates to this sequence of meaningful abstracted representations is the free exploration of abstraction space, meaning the direct transition from one abstracted visualization to another (e.g., [13, 14, 24]. This could be done along one sequence of abstractions as described above, or one can also imagine a transition *across* different visual abstraction pipelines [5, 6]. While in principle the transition from one chain of transformations to another can be formalized, the actual implementation in most cases is probably rather challenging. This exploration seems to have similar properties to the faceting operation as defined by Satyanarayan et al. [22].

We propose that given a task $\mathbb{T}$ and a target user $\mathbb{H}$, we can design a tool or system that communicates some insight by offering an appropriate set of visual representations $\mathbb{V}_i$ and a means to arrange them. Depending on their respective domain, this can be juxtaposition, superposition, animated interpolation, ... [15]. Which of these variants are possible depends on the different domains $\mathbb{A}$ as described above: while juxtaposition is generally feasible, it can, however, be very expressive if there is an overlap in the domain, like in scatterplots combined with marginal histograms. Superposition, on the other hand, strictly requires a common domain.

We observe that in theory it is possible to move arbitrarily in abstraction space, meaning that all the visual representations in the images above could be applied to a single dataset, once the correct transformations have been applied. Some combinations might seem dubious at first glance, like a node-link diagram of data without connectivity, but this does not mean there is no possible transformation that can generate meaningful connectivity from the raw data. For example, we can generate hydrogen bonds for an atomistic dataset.

Finally, we need to consider the ground truth: generally, it would represent a data set as a whole, including all possible insights. In its raw form, this ground truth might be generally so large and complex to be unintelligible. The goal of domain research often is to condense this into a (sufficiently faithful) model of the data, such that it can be used interchangeably and allows for prediction of the real world, for example. In practice, given sufficiently complex data, we usually have to settle for a ground truth with respect to some specific task $\mathbb{D}_{\mathbb{T}}$. We can provocatively call this a *partial truth*, and it can itself be seen as abstraction of the ground truth. As long as we keep in mind the task $\mathbb{T}$, this serves its purpose, but there are risks: A visualization designed to communicate $\mathbb{D}_{\mathbb{T}}$ biases the user towards this partial truth. This is problematic since we want to allow for explorative analysis, but that will—by design—be constrained by the currently applied abstractions. Given that a reduction R has been performed, we lose generality with respect to the ground truth, since some insights (other tasks) are potentially coupled to the information contained in what has been removed: $M(\mathbb{R}, \mathbb{T})$. It is thus vital that we formally define and document the provenance of a visualization in order to avoid fostering erroneous deductions: one can mistake absence of a feature or a data item as information while it just represents the performed abstraction. This ties back into the observation that an abstraction is defined and meaningful exclusively with respect to a task.

2 Categorization of Abstraction Techniques

In this section we present categories of abstraction techniques, which, broadly speaking, encompass abstractions in the data domain, abstractions in visual representation, and abstractions in interaction.

First, regarding abstractions in the data domain, within the visualization process, these include data filtering (subset), abstractions in data scale, and data sampling. Through abstractions in the data domain, we can reduce the desired information to a meaningful subset, enabling further exploration and insight communication. In terms of visual exploration, a sequence of abstractions can traverse freely the abstraction space from one representation to another, or keep the representation fixed and explore inside a single representation.

Second, methods for abstracting from data to visual representations or between visual representations include aggregation, composition, and layout change. In aggregation methods, discrete numerical distributions can be converted into aggregated numerical distributions, such as transforming data from a table visualization into a bar chart or aggregating individual points on a map into a heat map. In composition methods, we can combine different visual forms, for instance, assembling multiple data dimensions into a form resembling parallel coordinates. We can also create visual metaphors for conveying vivid data patterns. In layout change, we can utilize superposition, juxtaposition, and implicit layout forms to arrange certain relevant visualizations.

Third, abstractions in interaction primarily support higher-level user interactions to select, highlight, and explore corresponding data subsets, making them visible within the visualization. The user interaction is typically designed such that with less of time spent on interaction more of a task can be achieved.[2]

[2]One example could be the "ortho" rule in Google Sketchup software, where modeling in

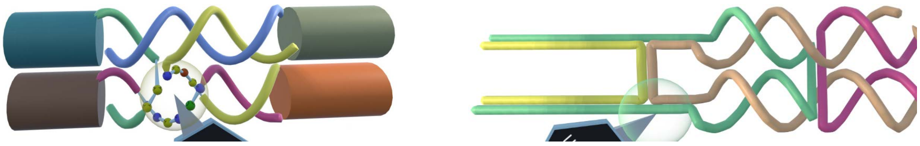


Figure 9: The Magic Scale Lens: within the Magic Scale Lens, the single strand scale (Focus) is embedded into the double strand scale (Context). Images from [8], 🅭🅯 CC-BY 4.0.

We can also employ a focus-and-context approach to direct users' attention towards the focal area, providing varying levels of detail, whether it is less or more detailed information. For example, the Magic Scale Lens and DNA Untwister in Figure 9 are employed to embed different representations into one single view using a lens, enabling users to inspect local representations as needed. These two interaction concepts tackle challenges related to occlusion and multiscale exploration. The first concept uses a lens to transform the representation of a specific focusing region. The focus region conveys information on a particular level of detail, while a more abstract representation serves as the context. This approach facilitates the seamless combination of multiple scales within a single view, allowing users to choose the level of detail they wish to display on demand. The second interaction concept is applied when an abstract representation suffices to convey the data or message effectively. For example, an untwisted DNA depiction simplifies the overall visualization without sacrificing the essential two-strand structure when intricate details are unnecessary. Within the lens, the DNA helix is untwisted into parallel strands for clarity and simplicity.

However, there are unexplored dimensions concerning how users can finely control the level of detail within the focusing region. This concern is linked to our earlier discussion on transitioning between different visual abstractions and the methods by which users control this transition. Another interesting aspect relates to user interactions within these multiscale embedded visualizations. To effectively navigate such visualizations, which encompass diverse visual abstractions within a single view, it becomes crucial to consider adaptive interaction techniques. Given that users perceive varying levels of detail, especially when working with distinct representations, they may develop diverse mental models for data manipulation and exploration, as well as expectations regarding the outcomes of specific interactions. Thus, the effect of these interaction techniques should depend on the abstraction to which they are applied, and the visual abstractions perceived by users.

Other related techniques such as faceting and overview-and-details also similarly employ visual abstractions but in a different composition or arrangement. A data dashboard may be composed of more than two visual abstractions. For example, the histogram and line chart in Figure 3 are superimposed and can be shown to the left of a box plot to give a complete picture of the data. A dashboard user can go back and forth between visual abstractions or interact with one of them to gain data insights. Similarly, a visualization of quanti-

3D assumes that lines connect under 90 degree angle, when interpreting positioning of line endpoint in 3D space from a view projection image ties through gauges can represents visual abstractions that aggregate the gauge visualization on multiple levels of detail [11].

In scientific visualizations, blending and masking may be used to create a continuous impression between visual abstractions as in examples in Figure 7. These visual abstractions often happen by cleverly arranging (super-positioning in 3D space, potentially also combined with animations between different representations) different visual representations $\mathbb{V}_i$ such that the viewer mentally connects the different representations, as depicted as the vertical bar in Figure 2.

Given ever-increasing and more and more complex networks that analysts have to face in many application areas, it is natural that network abstraction is of increasing interest to support the network analysis process. One example of the successful application of abstraction is the use of powergraph analysis for protein network analysis. Royer et al. [21] show that on average over 50% of network edges are redundant due to the typical patterns found in these networks. In addition to this compression effect, powergraph analysis facilitates visual representations that greatly improve the identification of functional subunits by visual separation. A typical example for abstracting from data to visual representation of networks are Graph Thumbnails [29], small icon-like representation of networks that support quick browsing of large network sets and high-level comparison of networks. They abstract a network to the nesting structure of connected components and k-cores, which is represented in a nested circle packing.

An abstraction technique is often associated with a strategy for presenting data, typically in a linear sequence. The resulting visualizations can be arranged either temporally, as in animations, or spatially, as in dashboards, with the primary objective of conveying a singular message. When each visualization represents a distinct subset of a dataset, this exemplifies abstraction in the data domain (filtering). If each visualization shows a distinct representation of the same dataset, this corresponds to the second form of abstraction. And, when each visualization drills down on the data, it is not unlike abstraction in interaction. Previous work [9, 19, 23] on storytelling has explored into how to shorten or lengthen this sequence of visualizations or to control the abstraction parameter. For these purposes, animation serves as an important means of facilitating intuitive comprehension.

The design of glyph is a highly abstract visual design that must strike a balance in the selection of the most important data features for mapping, and organically combine these features into a whole. The abstraction here entails three layers of meaning. Firstly, Glyph is a form of abstract representation of complex data, with its elements being concise and understandable. Secondly, the design of the glyph requires the abstraction of data, selection of important data dimensions, aggregation of data, and acquisition of more abstract data features mapped within the glyph. Thirdly, the design of the glyph necessitates the combination of important data features, which is the process of obtaining abstract data representation.

In the generation of glyph, we should consider and practice "abstraction" in the following ways: Firstly, based on the attributes of the data (whether ordinal, sequential, or quantitative), we select dimensions for visualization mapping, find suitable combinations, and make choices based on the characteristics of the graphic elements. This process requires selecting Glyph designs that are both representative and not overly complex. Secondly, we need to sort or explore the correlations of data dimensions, select representative dimension combinations, and minimize meaningless or redundant dimensions. Thirdly, we need to organically combine the selected data dimensions, which presents a significant design space that requires further research and exploration.

3 Opportunities and Applications

In this section, we discuss potential application and industry opportunities derived from our findings regarding abstraction in visualization.

3.1 Main Impact in Industrial Scenarios

A formalization of abstraction in visualization could lead to benefits in industrial settings and applications:

- Using abstract visualization layers could act as "knowledge interfaces" between diverse groups of users, who are not the primary target audience of a developed visualization. Higher levels of abstraction could be used for presentations, decision making, and knowledge sharing in the industry, replacing the current standard of using slides, which is completely decoupled in terms of data provenance and data exploration, error-prone, introduces the chances for bias and manipulation, and is moreover creating additional preparation work.
- Formalized definitions and guidelines for abstraction in visualization bring opportunities for automation and reuse of visualization—on a software development and design level, but also through easier interpretation of users from other domains. We could even see standardized abstraction layers that different visualization tools agree and rely on to exchange knowledge through visualizations between user groups that work on the same data with a completely different focus.
- Temporal consistency and re-usability are currently not considered enough. Both abstract visualizations as well as the underlying data have to be accessible AND understandable decades later. Standardization (not only on a data level, but particularly for abstraction in visualization) helps to ensure that visualizations could be considered reliable and future-proof, reducing the hesitation to apply visualizations in such settings.

3.2 Role and value

Through a formalization of abstraction in visualization techniques we see a chance to increase the value and acceptance of the fundamental work of visualization researchers by redefining their perceived roles or tasks:

- Formal definition, guidelines, standards, etc. lower the barriers for adapting visualization and reduce the chance of failures.
- This could lead to the creation of expert roles for the coordination of data-visualization workflows, considering all aspects from data requirements, visualization software and data architectures, abstraction layers, design guidelines, performance, data and software life cycle management, etc.

- Formalization also allows to continuously re-evaluate and adapt existing guidelines and their applicability to the problem to solve, particularly when it comes to changing cultural background, etc.

- All this could help to define and reveal the true value of visualizations [25]—not just in the research area, but also to quantify and justify visualizations, their application, their risks, etc.

During our discussion, we came up with multiple roles names (reaching from "Visualization Scientist" over "Visualization Data Scientist" to "Data Intelligence Workflow Design and Optimization expert" (which is not considering the the term *visualization* at all, and was therefore perceived with ambivalent feelings by the participants).

The visualization community is at the interface between (business) users and data scientists, so aiming for the establishment of a term that reflects the corresponding role's responsibilities could help to make these more graspable and to increase the perceived value of visualizations in the industry. By understanding visual abstraction and abstraction in general (i.e., including statistical abstraction and algorithmic abstraction), a new generation of visualization researchers can play a more significant role in designing and optimizing data intelligence workflows.

The main aspect is to apply visual abstraction principles to guide and model the processes of establishing visualizations—but there is the need to clarify that this task is not only about design. It involves the high-level coordination to establish overarching visualization environments that provide the benefits described in Section3.1. This could lead to a higher perceived value for this "profession".

3.3 Chances in the Industry

Establishing visualization techniques in the industry is often a difficult endeavor. The reasons are manifold, and include

- The (business) value of visualizations are hard to define. Formalizing and standardizing abstraction in visualization could make this task easier.

- Visualizations envisioned by visualization experts are hard to grasp for users before they exist—particularly abstractions. Relating to existing standards might have a positive impact on this process.

- Visualization tools introduce a significantly higher complexity in terms of required resources for maintenance, computation, training, etc. than non-visual analytics. In a business setting, justifying these investments and being able to understand the risks better is crucial. Standardization increases comparability of solutions and allows better estimations.

- Related to this, standardization also increases re-usability of visualization techniques, having a benefit on all the aforementioned aspects.

- A standardization of abstraction in visualization techniques could potentially lead to the design of corresponding Domain Specific Languages (DSL)

One of the biggest advantages of well-defined abstraction layers in visualizations in the industry is workload reduction. By integrating for example standardized presentation layers, i.e., the highest level of abstraction, into standard presentation workflows, removes the task of creating additional, completely de-coupled presentation material without any provenance.

Multi-level abstract visualizations should ultimately not just be a tool for a certain group (e.g., for expert users of a scientific area), but to support whole processes for knowledge discovery, knowledge generation, knowledge sharing, and decision making between multiple groups. Full traceability over years as well as relying on meaningful representations are extremely valuable aspects for any industry.

It seems important to point out that the current need is not to define tools and templates that provide an overarching solution, but to find applicable abstraction definitions, design rules, etc. that people can agree on and that can be implemented in various ways.

Another big chance for the industry lies in using an additional side effect of the establishment of visual abstraction standards: As soon as the value is understood and accepted, it could be a driving force for increasing the data quality and the underlying data capturing processes to allow the best possible facilitation. The intrinsic motivation of people to change their processes in industry can significantly rise with a clear purpose and value.

References

[1] A. Bock, E. Axelsson, J. Costa, G. Payne, M. Acinapura, V. Trakinski, C. Emmart, C. Silva, C. Hansen, and A. Ynnerman. OpenSpace: A system for astrographics. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):633–642, 2020. doi: 10.1109/TVCG.2019.2934259

[2] S. K. Card and J. Mackinlay. The structure of the information visualization design space. In *Proc. InfoVis*, pp. 92–99. IEEE Computer Society, Los Alamitos, 1997. doi: 10.1109/INFVIS.1997.636792

[3] S. Halladjian, H. Miao, D. Kouřil, M. E. Gröller, I. Viola, and T. Isenberg. ScaleTrotter: Illustrative visual travels across negative scales. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):654–664, Jan. 2020. doi: 10.1109/TVCG.2019.2934334

[4] J. Heer, S. K. Card, and J. A. Landay. Prefuse: A toolkit for interactive information visualization. In *Proc. CHI*, pp. 421–430. ACM, New York, 2005. doi: 10.1145/1054972.1055031

[5] J. Heer and G. Robertson. Animated transitions in statistical data graphics. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1240–1247, Nov. 2007. doi: 10.1109/TVCG.2007.70539

[6] Y. Kim and J. Heer. Gemini^2: Generating keyframe-oriented animated transitions between statistical graphics. In *Proc. Visualization*, pp. 201–205. IEEE Computer Society, Los Alamitos, 2021. doi: 10.1109/VIS49827.2021.9623291

[7] G. Kindlmann and C. Scheidegger. An algebraic process for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2181–2190, 2014. doi: 10.1109/TVCG.2014.2346325

[8] D. Kuťák, M. N. Selzer, J. Byška, M. L. Ganuza, I. Barišić, B. Kozlíková, and H. Miao. Vivern—A virtual environment for multiscale visualization and modeling of DNA nanostructures. *IEEE Transactions on Visualization and Computer Graphics*, 28(12):4825–4838, 2022. doi: 10.1109/TVCG.2021.3106328

[9] A. Lu and H.-W. Shen. Interactive storyboard for overall time-varying data visualization. In *Proc. PacificVis*, pp. 143–150. IEEE Computer Society, Los Alamitos, 2008. doi: 10.1109/PACIFICVIS.2008.4475470

[10] J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141, apr 1986. doi: 10.1145/22949.22950

[11] K. Matkovic, H. Hauser, R. Sainitzer, and M. E. Gröller. Process visualization with levels of detail. In *Proc. InfoVis*, pp. 67–70. IEEE Computer Society, Los Alamitos, 2002. doi: 10.1109/INFVIS.2002.1173149

[12] A. Meidiana, S. Hong, P. Eades, and D. A. Keim. A quality metric for visualization of clusters in graphs. arXiv preprint, 2019. doi: 10.48550/arXiv.1908.07792

[13] H. Miao, E. De Llano, T. Isenberg, M. E. Gröller, I. Barišić, and I. Viola. DimSUM: Dimension and scale unifying maps for visual abstraction of DNA origami structures. *Computer Graphics Forum*, 37(3):403–413, June 2018. doi: 10.1111/cgf.13429

[14] H. Mohammed, A. K. Al-Awami, J. Beyer, C. Cali, P. Magistretti, H. Pfister, and M. Hadwiger. Abstractocyte: A visual tool for exploring nanoscale astroglial cells. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):853–861, Jan. 2018. doi: 10.1109/TVCG.2017.2744278

[15] T. Munzner. *Visualization Analysis and Design.* CRC Press, Boca Raton, 2014. doi: 10.1201/b17511

[16] Q. Nguyen, P. Eades, and S.-H. Hong. On the faithfulness of graph visualizations. In *Proc. PacificVis*, pp. 209–216. IEEE Computer Society, Los Alamitos, 2013. doi: 10.1109/PacificVis.2013.6596147

[17] D. A. Norman. *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine.* Addison-Wesley Longman Publishing Co., Inc., USA, 1993.

[18] J. Parulek, D. Jönsson, T. Ropinski, S. Bruckner, A. Ynnerman, and I. Viola. Continuous levels-of-detail and visual abstraction for seamless molecular visualization. *Computer Graphics Forum*, 33(6):276–287, Sept. 2014. doi: 10.1111/cgf.12349

[19] O. Radiano, Y. Graber, M. Mahler, L. Sigal, and A. Shamir. Story albums: Creating fictional stories from personal photograph sets. *Computer Graphics Forum*, 37:19–31, 4 2018. doi: 10.1111/cgf.13099

[20] P. Rautek, S. Bruckner, E. Gröller, and I. Viola. Illustrative visualization: New technology or useless tautology? *ACM SIGGRAPH Computer Graphics*, 42(3), Aug. 2008. doi: 10.1145/1408626.1408633

[21] L. Royer, M. Reimann, B. Andreopoulos, and M. Schroeder. Unraveling protein networks with power graph analysis. *PLoS Comput Biol.*, 4(7), 2008. doi: doi: 10.1371/journal.pcbi.1000108

[22] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer. Vega-Lite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350, 2017. doi: 10.1109/TVCG.2016.2599030

[23] D. Shi, X. Xu, F. Sun, Y. Shi, and N. Cao. Calliope: Automatic visual data story generation from a spreadsheet. *IEEE Transactions on Visualization and Computer Graphics*, 27:453–463, 2021. doi: 10.1109/TVCG.2020.3030403

[24] M. van der Zwan, W. Lueks, H. Bekker, and T. Isenberg. Illustrative molecular visualization with continuous abstraction. *Computer Graphics Forum*, 30(3):683–690, May 2011. doi: 10.1111/j.1467-8659.2011.01917.x

[25] J. J. van Wijk. The value of visualization. In *Proc. VIS*, pp. 79–86. IEEE Computer Society, Los Alamitos, 2005. doi: 10.1109/VISUAL.2005.1532781

[26] I. Viola, M. Chen, and T. Isenberg. Visual abstraction. In M. Chen, H. Hauser, P. Rheingans, and G. Scheuermann, eds., *Foundations of Data Visualization*, chap. 2, pp. 15–37. Springer, Berlin/Heidelberg, 2020. doi: 10.1007/978-3-030-34444-3_2

[27] I. Viola and T. Isenberg. Pondering the concept of abstraction in (illustrative) visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(9):2573–2588, Sept. 2018. doi: 10.1109/TVCG.2017.2747545

[28] G. Wiggins and J. McTighe. *Understanding by Design*. Association for Supervision & Curriculum Development, 2nd expanded ed., 2005.

[29] V. Yoghourdjian, T. Dwyer, K. Klein, K. Marriott, and M. Wybrow. Graph thumbnails: Identifying and comparing multiple graphs at a glance. *IEEE Transactions on Visualization & Computer Graphics*, 24(12):3081–3095, dec 2018. doi: 10.1109/TVCG.2018.2790961