Unveiling High-dimensional Backstage: A Survey for Reliable Visual Analytics with Dimensionality Reduction

Hyeon Jeon
Seoul National University
Seoul, Republic of Korea
hj@hcil.snu.ac.kr

Hyunwook Lee
UNIST
Ulsan, Republic of Korea
hwlee0916@unist.ac.kr

Yun-Hsin Kuo
University of California, Davis
Davis, CA, USA
yskuo@ucdavis.edu

Taehyun Yang
Seoul National University
Seoul, Republic of Korea
0705danny@snu.ac.kr

Daniel Archambault
Newcastle University
Newcastle upon Tyne, UK
Daniel.Archambault@newcastle.ac.uk

Sungahn Ko
UNIST
Ulsan, Republic of Korea
sako@unist.ac.kr

Takanori Fujiwara
Linköping University
Linköping, Sweden
takanori.fujiwara@liu.se

Kwan-Liu Ma
University of California, Davis
Davis, CA, USA
klma@ucdavis.edu

Jinwook Seo*
Seoul National University
Seoul, Republic of Korea
jseo@snu.ac.kr

Abstract

Dimensionality reduction (DR) techniques are essential for visually analyzing high-dimensional data. However, visual analytics using DR often face *unreliability*, stemming from factors such as inherent distortions in DR projections. This unreliability can lead to analytic insights that misrepresent the underlying data, potentially resulting in misguided decisions. To tackle these reliability challenges, we review 133 papers that address the unreliability of visual analytics using DR. Through this review, we contribute (1) a workflow model that describes the interaction between analysts and machines in visual analytics using DR, and (2) a taxonomy that identifies where and why reliability issues arise within the workflow, along with existing solutions for addressing them. Our review reveals ongoing challenges in the field, whose significance and urgency are validated by five expert researchers. This review also finds that the current research landscape is skewed toward developing new DR techniques rather than their interpretation or evaluation, where we discuss how the HCI community can contribute to broadening this focus.

CCS Concepts

• **Human-centered computing** → *Visual analytics*; • **Mathematics of computing** → **Dimensionality reduction**.

Keywords

Dimensionality reduction, Multidimensional projection, Reliability, High-dimensional data, Literature analysis, Survey

*Corresponding Author

1 Introduction

Dimensionality reduction (DR) serves as a backbone of data visualization and visual analytics [26, 111, 154, 212]. DR techniques such as t-SNE [199], UMAP [137], and RadViz [172] generate low-dimensional representations of high-dimensional data while preserving key structural characteristics of the original data, including local neighborhood structure and global cluster arrangement. DR projections are widely incorporated in many visual analytics systems and applications [37, 111, 151] across diverse domains such as bioinformatics [5, 18, 37, 143, 150], natural language processing [124, 194], and human-computer interaction [17, 126]. DR has also been used to explain machine learning models [166, 222], including cutting-edge foundation models like large language models [194].

Still, visual analytics leveraging DR can easily become unreliable: derived insights or knowledge from visual analytics may not accurately reflect the underlying data, potentially leading to flawed decisions. For example, because low-dimensional spaces have a fewer degree of freedom than high-dimensional spaces, distortions are inevitable in DR projections. These distortions can reduce the accuracy of the projections [13, 87, 152], misrepresenting the original data and potentially leading analysts to incorrect conclusions about its structure. Such erroneous insights or decisions can cascade downstream, negatively affecting subsequent analyses or research that relies on them. To address these challenges, the visual analytics community has thus conducted numerous studies to enhance the reliability of visual analytics using DR.

However, we need more comprehensive guidance to help analysts address the full scope of unreliability when using DR. Existing surveys [59, 154, 189] focus on guiding analysts in selecting DR techniques that produce accurate projections for specific data or tasks. For example, they recommend using t-SNE or UMAP to reliably support neighborhood investigation tasks. While these surveys

contribute to making visual analytics more reliable by increasing practitioners' awareness of the design and purpose of DR techniques, unreliability can still arise for various other reasons, such as suboptimal hyperparameter settings [106, 217] or the inherent instability of certain DR techniques [106, 137]. Other approaches beyond DR technique selection, including making DR more interactive [8, 71] or interpretable [21, 87, 122], are proposed to remedy these problems.

Our research aims to coordinate these diverse efforts, thereby contributing to achieving more reliable visual analytics with DR. We first carefully review the literature, identifying 133 relevant papers from diverse research fields, including data visualization, human-computer interaction, and machine learning. Based on this systematic review, we develop a **workflow model** that outlines the interaction between analysts and machines when conducting visual analytics with DR. Grounded in the identified papers and previous theories in visual analytics, this model enhances our understanding of how analysts use DR in practice. It offers a comprehensive explanation of DR usage, covering the entire analytic process—from preprocessing to visualizations and sensemaking. Finally, we organize the reviewed papers into a **taxonomy** structured around four dimensions: Stage, Problem, Aim, and Solution, each branch mapped to a component of the workflow model or defined by those components. A meta-analysis reveals six prevalent clusters of papers, offering deeper insights into the overall research landscape.

Our analysis identifies several **ongoing challenges**, including the limited discussion extending beyond 2D scatterplots and the lack of available libraries to support more advanced DR techniques. Eight expert researchers in the relevant fields validated the significance and urgency of these challenges. Furthermore, our survey reveals an imbalance in the research landscape, with many papers proposing new techniques but few focusing on their evaluation or interpretation. To fill this gap, we **call for the HCI community** to invest more efforts to enhance the interpretability of DR techniques and make it easier to use DR reliably. As a seminal step, we develop and release an actionable guide to help researchers navigate the literature included in our survey.

In summary, this research makes four key contributions:

- ☐ We present a **workflow model** that comprehensively details the visual analytics procedure using DR.
- ☐ We propose a **taxonomy** that categorizes 133 relevant studies in the field.
- ☐ We identify three significant and urgent **ongoing challenges** in DR-driven visual analytics.
- ☐ We reveal an imbalance in the relevant literature, where we **prompt the HCI community** to mitigate it.

To the best of our knowledge, this is the first survey specifically addressing the issue of unreliability in visual analytics for DR techniques. We thus hope this research will serve as a valuable reference not only for practical DR applications but also for future investigations into reliable visual analytics. We provide an interactive browser of this survey at https://dr-reliability.github.io/demo.

2 Related Work

We discuss three relevant areas of previous work: (1) existing surveys on DR techniques and evaluation metrics, (2) investigations on the practical usage of DR in visual analytics, and (3) theoretical models for visual analytics.

2.1 Surveys on Dimensionality Reduction

We review existing surveys on DR techniques and evaluation metrics. These surveys share a common goal with our research: enhancing the reliability of visual analytics using DR. Our research differentiates itself by addressing a broader spectrum of threats that can compromise reliability, spanning various stages of the analytical workflow beyond selecting DR techniques and metrics.

Surveys on DR techniques. These surveys aim to reveal the pros and cons of existing DR techniques, providing guidance for practitioners in selecting the most appropriate technique that matches the analysts' intentions, e.g., target tasks [44, 57, 65, 154, 189, 200, 224]. However, these literature reviews often lack quantitative validation of the performance of DR techniques, e.g., accuracy in preserving the structure of high-dimensional datasets, limiting the extent to which their guidance can be applied. For example, a technique may work appropriately for dense datasets but poorly for sparse datasets [59]. Researchers thus contribute benchmark studies that incorporate the measurements and analyses of projection accuracy or scalability [12, 28, 59, 200] to fill this gap. Recently, Espadoto et al. [59] conduct a large-scale study that shows how the performance of DR techniques is affected by the dataset traits (e.g., dimensionality, number of data points) or hyperparameter settings. Xia et al. [216] and Etemadpour et al. [60] also contribute large-scale benchmark studies, where they directly examine human task performance by conducting user studies.

Surveys on evaluation metrics. Although DR techniques are specialized to preserve certain aspects of the original data, they inherently cannot escape from distortions [43, 59, 87, 122, 154]. Thus, many quality metrics are proposed to assess the accuracy of DR projections in preserving structural characteristics of the original high-dimensional data [42, 87, 89, 202]. As with DR techniques, suitable use of quality metrics is essential in conducting reliable data analysis. For example, to identify DR projections that reliably support cluster analysis, metrics specialized to cluster structure (e.g., Steadiness & Cohesiveness [87]) are recommended to test the projections.

Reviews on DR quality metrics thus contribute as guidelines for selecting proper metrics. For example, Nonato and Aupetit [154] and Thurn et al. [196] taxonomize existing quality metrics based on their target structural characteristics (e.g., local neighborhood or global pairwise distance). Lee and Verleysen [116] provide an extensive review of the quality metrics that assess the preservation of neighborhood structure. A recent study from Jeon et al. [86] not only conducts the review but deploys the surveyed metrics as a unified Python library.

Our contribution. Existing literature reviews and benchmarks enhance the unreliability of visual analytics by addressing a lack of awareness about the disadvantages and advantages of DR techniques and quality metrics in supporting diverse analytic tasks. This is done by providing guidelines in selecting DR techniques and metrics that optimally aligns with the analysts' intention (e.g., task, prior domain knowledge) and data. In this work, we demonstrate that factors beyond optimality can compromise the reliability of

visual analytics (e.g., interpretability or stability of DR techniques). We also show that the efforts for addressing this unreliability go beyond the selection of proper techniques or quality metrics (Sect. 5). For example, our survey shows that visual analytics systems that explain DR techniques [46] or unified libraries [86] can even contribute to enhancing reliability. Doing so offers a more comprehensive taxonomy and actionable guidelines to help analysts conduct more reliable visual analytics using DR.

2.2 Investigation on the Practical Usage of Dimensionality Reduction

Given the wide usage of DR, several works investigate how DR is used for visual analytics [24, 58, 175]. Brehmer et al. [24] conduct interviews with data analysts, revealing a workflow of using DR for analyzing cluster structure or interpreting the meaning of dimensions. Based on task sequences, they derive guidelines for a more comprehensive evaluation of DR-based visualizations. Sacha et al. [175] conduct a literature review to identify how analysts interact with DR projections. They contribute a framework systematizing human-DR interaction and then use it to understand and compare visual analytics systems using DR. They also use the proposed framework to suggest research directions for enriching interaction with DR. Espadoto et al. [58] explore workflows for developing, evaluating, and deploying DR techniques for practical applications. They further identify the toolkits or libraries available for each workflow step. Furthermore, they discuss the challenges of using such toolkits in practice, e.g., the lack of benchmark datasets to evaluate DR techniques.

Our contribution. Previous research focuses on specific scenarios to offer insights into using DR in visual analytics. In contrast, we offer a more comprehensive explanation of DR usage, covering the entire workflow from data preprocessing to visualization and sensemaking. This approach gives a broader understanding of DR usage in practice. Grounded in this workflow, our study comprehensively captures reliability problems in DR-based visual analytics that previous studies have not fully explored.

2.3 Theoretical Models for Visual Analytics Workflow

Visual analytics and data mining communities have proposed diverse models that explain how analytic procedures derive knowledge from data. Van Wijk [201] propose a model of how each component in visual analytics (e.g., perception, interaction, visualization specification) affects knowledge. Van Wijk's model has been further complemented by Green et al. [79]. Fayyad et al. [63] propose the Knowledge Discovery in Databases (KDD) process, which describes the data mining stages that raw data should go through to generate knowledge. Chen and Jäenicke [36] explain the visual analytics process by aligning it with information theory. Keim et al. [101] review existing theories and models to comprehensively explain the visual analytics process. Based on the previous workflow models, Sacha et al. explain the role of visual analytics in generating knowledge [174].

Our contribution. This research introduces a theoretical workflow model describing how analysts and machines behave when

conducting visual analytics (Sect. 4). Our model uses these previous workflows as the theoretical backbone. For example, our model follows the global structure proposed by Van Wijk [201]. Nevertheless, our model examines each component of the structure in detail, providing a fine-grained explanation of how DR-based visual analytics takes place. Moreover, we not only contribute a workflow model but also discuss the unreliability that can arise within the workflow, providing practical guidance to help analysts make their visual analytics reliable.

3 Protocol

We discuss our search strategy for identifying relevant papers and provide an overview of the papers we select. We then explain in detail how we develop our workflow model and taxonomy.

3.1 Paper Selection

We outline our survey scope and the procedure for selecting papers. We also provide examples of papers included and excluded in our survey. Furthermore, we show how the papers are distributed across the years and fields, confirming the wide coverage of our survey.

Survey scope. We search papers that address unreliability from visual analytics using DR. The survey does not focus on unreliability primarily related to data-agnostic visualizations (e.g., overplotting in scatterplots) or other machine learning techniques (e.g., clustering or regression). The example papers for inclusion and exclusion are as follows:

Examples for inclusion:

- ☐ Studies that improve the accuracy of existing DR technique, e.g., by modifying loss functions or distance functions [88, 150, 195, 205]. These studies support analysts in conducting their tasks more accurately.
- ☐ Studies that propose new DR quality metrics [87, 147, 186] or improving the reliability of existing metrics [89, 117]. These studies help practitioners identify DR projections optimal for their tasks based on metric scores.
- ☐ Studies that propose visualizations or interaction techniques that explain distortions in projections [13, 43, 87, 122]. These studies contribute to making practitioners more aware of the existence of inaccuracy of projections due to distortions.
- ☐ Benchmark studies that provide an extensive comparison of existing DR techniques [59, 60, 200, 216]. These benchmarks help practitioners become more aware of the pros and cons of existing techniques and select the most appropriate techniques for given datasets and tasks.

Examples for exclusion:

- ☐ Studies that deal with reliability problems for general visualizations and thus are not always relevant to visual analytics using DR, e.g., visual perception of scatterplots [1, 90].
- ☐ Studies that mainly focus on improving the computational efficiency of existing DR techniques [69, 102].
- ☐ Studies that aim to remedy reliability problems related to other types of ML techniques, e.g., clustering [49, 123], or visualizations, e.g., Parallel Coordinates Plot, glyphs, that are used to analyze high-dimensional data [169, 221].
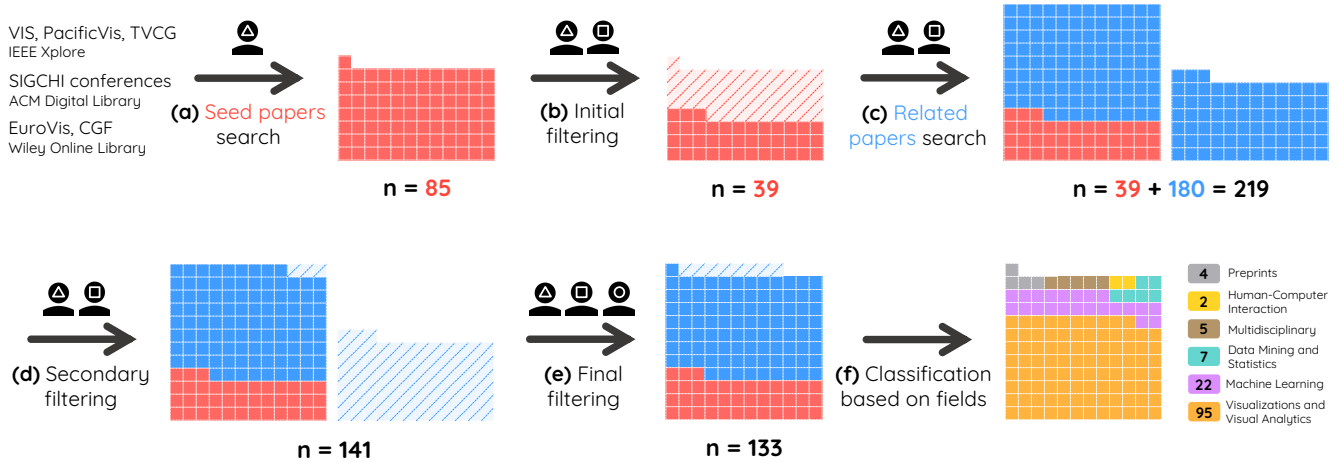
Figure 1: *Procedure of selecting papers (a-e) and their classification based on research fields (f).* (a) We first search seed papers published in major visualization and human-computer interaction venues (■), and (b) filter out papers that do not fall within our survey scope (▱). (c) We then extend our paper collection (■) by screening the related works and backgrounds of the seed papers. (d, e) Finally, by extensively reviewing and filtering out unrelated papers (▱), we finalize a total of 133 papers. (f) We classify the papers based on their fields. The icons above each arrow represent the authors (▲, ▣, and ◉) that are involved in the corresponding step.

Procedure. Our procedure of selecting papers is inspired by existing surveys in the data visualization field [165, 183, 207]. Three authors search the papers in May 2024.

First, we search seed papers and filter irrelevant papers. A primary author search papers published in the major visualization and HCI venues—VIS (InfoVis, VAST, and SciVis), TVCG, PacificVis, EuroVis, CGF, and SIGCHI affiliated conferences—using the corresponding digital libraries: IEEE Xplore, Wiley Online Library, and ACM Digital library (Fig. 1a). We focus on visualization and HCI venues as our primary focus is to investigate the use of DR in visual analytics.

The primary author query (*"Dimensionality Reduction"* OR *"Dimension Reduction"* OR *"Multidimensional Projection"* OR *"Multidimensional Scaling"*) AND (*"Reliability"* OR *"Trustworthy"* OR *"Distortion"* OR *"Uncertainty"* OR *"Quality"*) for paper title and abstract. We do not set specific time limits to make our list of papers more comprehensive. As a result, we identify 85 papers in total (Fig. 1a; ■). Then, two authors screen the title, abstract, and introduction of the papers and filter out papers that are out of our scope (Fig. 1b; ▱). We filter out papers only if two authors agree to reduce the risk of excluding relevant papers. The procedure ends up with 39 seed papers.

Second, we extend the search space to the related works of seed papers. The two authors conduct the second screening of the seed papers to curate a list of papers that likely are within our survey scope. We review the introduction section if a paper does not have such sections. Note that a paper is included in our collection if at least one of two authors thinks that the paper is relevant to our study to minimize the risk of overlooking relevant papers. As a result, we additionally find 180 papers (Fig. 1c; ■). However, as this extension is only done by reviewing the contents of seed papers, two authors again screen the title, abstract, and introduction of

these new papers and filter out irrelevant papers (Fig. 1d; ▱), where we conclude with 141 papers (including 39 seed papers) in total.

At last, we conduct the final screening and filtering of papers. We review the title, abstract, introduction, and method sections to ensure that our collection contains no false positives. An additional author is involved in this step to bring a fresh perspective to the screening. We keep papers only if all three authors agree to include the papers. We exclude eight papers, resulting in a final collection of **133** papers (Fig. 1e).

Metadata analysis. We want to identify the trends in relevant research over different periods and research fields. For this purpose, we classify the collected papers according to the research fields of the journals, conferences, and workshops in which they were published (Fig. 1f). We do so by reviewing the titles and descriptions provided by the venues. We then investigate how the total number of papers and the number of papers in each field have changed over the years.

Our classification divides the papers into six groups: *Visualization and visual analytics*, *Machine learning*, *Data mining and statistics*, *Multidisciplinary*, *Human-computer interaction*, and *Preprints* (Fig. 1 right bottom). The diversity of fields indicates that the reliability of data visualizations using DR is crucial in many research areas. This implies that our survey is likely to impact a variety of domains and research fields. The result is unsurprising since DR is widely employed to analyze high-dimensional data and present findings in various fields [17, 37, 124, 126, 194] (Sect. 1).

We also find that relevant papers have increased significantly since early 2000 (Fig. 2). This rise indicates that our topic is gaining more importance in these fields. Overall, our metadata analysis confirms the need for comprehensive knowledge of how the reliability of visual analytics using DR can be improved.
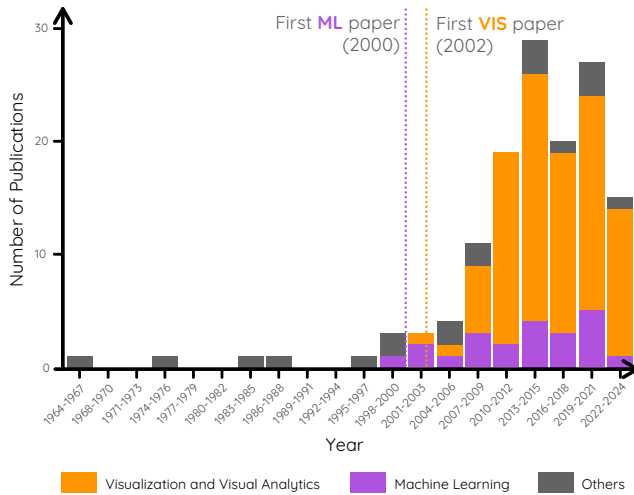
Figure 2: *The distribution of collected papers over the years.* **The papers we identify are dominantly incorporated in** machine learning (ML) **and** visualization (VIS) **fields. The number of published papers dramatically increases around the early 2000s.**

3.2 Workflow Model and Taxonomy Design

We establish our workflow model (Sect. 4) and taxonomy (Sect. 5) by conducting the iterative thematic analysis. Three authors involve the process as coders. The detailed process is as follows.

Initial review and discussion. Three coders individually read 133 collected papers and write a description describing when and why unreliability occurs in visual analytics and how the papers address such a problem. In this stage, the coders note their findings in free-form texts to add depth to the descriptions.

Workflow model design. Based on the descriptions, three coders design the workflow model through iterative discussions. The coders first establish three main considerations in the design process through iterative discussions.

- ☐ **(C1)** The model should clearly distinguish the role of analysts and machines (i.e., computers).
- ☐ **(C2)** The model should cover all analytic procedures (e.g., preprocessing, execution of DR, evaluation of projections).
- ☐ **(C3)** The model should correspond to or translate into existing models explaining visual analytics workflow.

We set C1 to clearly distinguish the role of humans (analysts) or computers (machines) in conducting visual analytics using DR. C2 is established to ensure that the workflow and the taxonomy (Sect. 5) can comprehensively cover the reliability problems stemming from all steps of the analytic workflow. Lastly, C3 is set to anchor our model in well-established and validated previous research, reducing its vulnerability to unexpected errors – a common approach in the visualization field for building new theoretical models [19, 78, 174, 208].

Three coders collaboratively revise the model after one coder drafts the model based on these considerations. The coders conduct four meetings to reach an agreement, where three more authors

occasionally participate in this discussion. Please refer to Sect. 4 for a detailed model description.

Taxonomy design and paper classification. Based on the initial description of papers and workflow model, three coders iteratively design the taxonomy and classify papers. At first, three coders draft the taxonomy together. The main consideration in this step is the alignment of the workflow model and the taxonomy; every branch within the taxonomy should be a component of the workflow or explained by the components.

Using our initial taxonomy, three coders independently classify the papers. The initial agreement measured by Cohen's κ is 0.63 on average. Three coders then iteratively revise the taxonomy and classification results until they reach an agreement. After two iterations, the final taxonomy and classification are established, which we detail in Sect. 5.

4 Workflow Model

Grounded by the papers we identified (Table 1) and previous theories on visual analytics (Sect. 2.3), we design a workflow model explaining how DR is used in visual analytics. The workflow model enhances our understanding of how analysts conduct visual analytics using DR. In this study, the workflow contributes to building a more comprehensive taxonomy of the papers we identify (Sect. 5).

Our workflow model (Fig. 3) consists of two main actors (Analysts and Machines) and six stages (Preprocessing, DR or Dimensionality Reduction, Quantitative Evaluation, Visualization, Sensemaking, and Interaction). The first four stages are carried out by Machines, and the last two stages are executed by Analysts. Inspired by van Wijk's visualization model [201], we describe visual analytics using DR as an iterative, looped process where analysts continually adjust configurations to update visualizations based on newly acquired knowledge. Our model also reflects a common workflow of using machine learning techniques through iterative monitoring and refinement of specifications [56, 98]. We further suggest that high-dimensional data cascades from previous stages to subsequent stages executed by Machines, with each stage enriching the data, for example, by generating DR projections (DR) or evaluating their accuracy (Quantitative Evaluation).

In the following sections, we first detail the labor of Analysts and Machines (Sect. 4.1). We then describe each stage of the analytics workflow (Sect. 4.2).

4.1 Analysts and Machines

Visual analytics cannot be established without both data Analysts and Machines. analysts initialize and update configurations that capture both their intentions (e.g., task, requirements, hypothesis) and knowledge, which is subsequently fed into Machines to guide their behavior.

Analysts. As demonstrated in the knowledge generation model proposed by Sacha et al. [174], analysts' main goal in our workflow is to perform Sensemaking , thereby generating knowledge that is useful and also reliably reflects the original data. Detailed descriptions of analysts' actions are as follows.
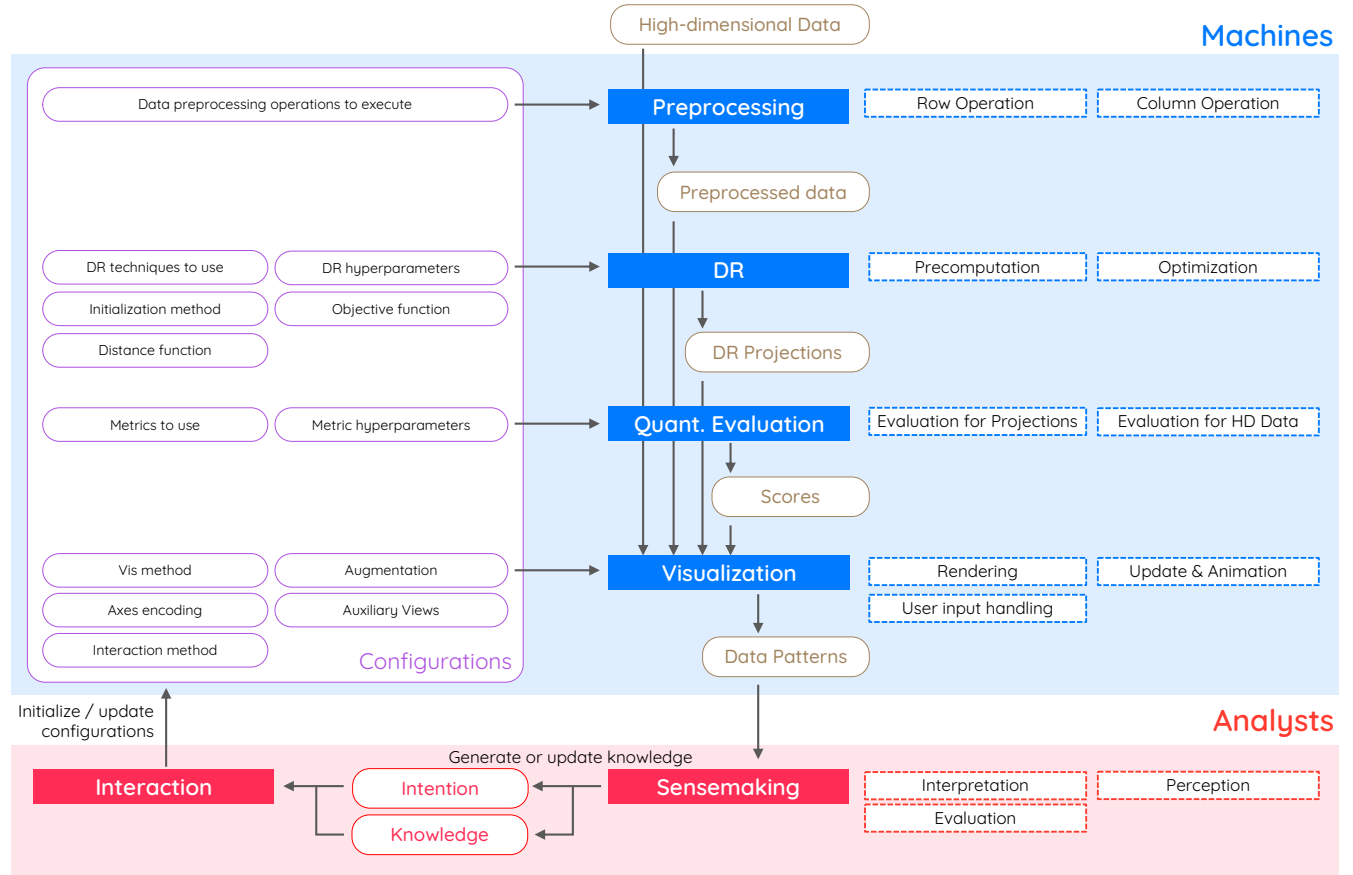
Figure 3: *The illustration of our workflow model.* The model explains how an Analyst and a Machine interact while conducting visual analytics using DR. Each stage of visual analytics executed by analysts and machines is represented by red and blue rectangles, respectively, and the input and output of each stage are designated by arrows.

At first, analysts initialize configurations. This is done by interacting with Machines, e.g., by manipulating visualizations or writing a program (Interaction stage). The setup is done based on analysts' knowledge (e.g., domain knowledge about the target data) and intention. For example, analysts may configure machines to execute DR techniques that focus on local structure (e.g., *t*-SNE or UMAP) to perform local neighborhood investigation. Similarly, when analysts' domain knowledge informs that certain attributes have low significance in the analysis, they can reduce the weight of those attributes in the distance function. Operational knowledge, encompassing the analysts' expertise in configurations, serves as a "mental toolbox" to support this process. Without prior knowledge of the configurations that fit their needs, analysts will have difficulty in establishing appropriate configurations. For example, analysts may hardly use *t*-SNE or UMAP for tasks related to the local structure if they do not know that such techniques fit the task well.

Analysts then generate or update knowledge based on output visualizations. After analysts deliver the configurations, the machines process the input high-dimensional data and visually convey data patterns to the analysts. Then, analysts perform Sensemaking,

e.g., perceiving and interpreting data patterns, thereby updating knowledge and intention. Please refer to Sacha et al. [174] or Kim et al. [104, 105] for a general explanation of the sensemaking procedure in visual analytics.

Analysts again interact with machines to update configurations based on the updated knowledge and intention. This happens when analysts think the current configurations and resulting data patterns can hardly validate new hypotheses. For example, analysts may newly want to explore the detailed attribute values of clusters they discovered in the current DR projection. In such a case, the analysts need to reconfigure the Visualization, e.g., by adding auxiliary views like parallel coordinate plots that can better depict the detailed attribute values of the clusters.

Machines. In our workflow model, machines process the input high-dimensional data based on the given configurations. It comprises four stages: Preprocessing, Dimensionality Reduction, Quantitative Evaluation, and Visualization. Each stage the machine executes enriches the data. The enriched data are summarized as data patterns and delivered by visualizations. We detail the labors of machines while we detail each stage in Sect. 4.2.

4.2 Workflow Stages

We outline six workflow stages of our model. The stages' composition draws inspiration from the KDD process model [63], which describes how raw data passes through preprocessing, transformation, and data mining to yield final visualizations. Here, we substitute the transformation and data mining steps to `DR` and `Quantitative Evaluation` stages to better match with visual analytics workflow using DR. We thus describe the stages in the sequence that are aligned with the KDD process. It is still important to note that the nature of interactive visual analytics causes the execution of these stages to be nonlinear: Analysts can interact with machines to update configurations, which can cycle the analysis back to the Machine-side stages (Sect. 4.1).

`Preprocessing`. The preprocessing stage involves a variety of data manipulations that are executed before applying the DR techniques. The step receives raw *high-dimensional data* as input and outputs the *preprocessed data* based on the configurations specifying preprocessing operations to execute. It can be largely divided into row operations and column operations. Data imputation, subsampling, or outlier removal are representative row operations. Regarding column operations, attribute selection that samples out or less weighs non-significant attributes is widely adopted. Attribute selection can also be conducted automatically, e.g., to reveal hidden patterns [70] or maximize specific patterns like class separation [186, 209]. Another typical column operation is to introduce new columns that represent the original structure's feature, e.g., by executing clustering techniques and adding the resulting clusters as labels to the datasets [212, 213].

`Dimensionality Reduction` or `DR`. This stage gets the *preprocessed data* as input and executes DR algorithms, producing *DR projections* as outputs. The stage is controlled by five configurations: DR techniques, technique-specific hyperparameters, initialization method, objective function (i.e., loss function), and distance function. The first two configurations should be necessarily set to execute the stage. The last three configurations are widely set for nonlinear DR techniques (e.g., UMAP, *t*-SNE, and Isomap [195]), which work by optimizing the 2D positions of data points to preserve the original distances between them. The initialization method defines how the points will be positioned before starting optimization, and the objective function defines how the preservation of distances will be quantified. These configurations substantially impact the resulting DR projections. For example, Kobak et al. [106] show that initializing points using PCA before the optimization leads to more accurate DR results than using random initialization. Lee and Verleysen [118] show that the design of the distance function affects the projection accuracy.

`Quantitative Evaluation`. This stage gets *high-dimensional data*, *preprocessed data*, and *DR projections* as inputs and produces evaluation results as numerical *scores*. These scores (1) explain how well the projections support the analytic tasks of investigating high-dimensional data or (2) the validation of hypotheses. For example, when analysts want to conduct cluster analysis, it is important to evaluate the projections in advance using the evaluation metrics assessing the preservation of cluster structure [87, 186]. This stage

is affected by two configurations: metrics and metric hyperparameters. For example, Trustworthiness & Continuity [202], which examine how well DR projections preserve the local neighborhood structure of original data, require a hyperparameter that designates the number of nearest neighbors to be considered. Note that these quality metrics differ from "perceptual metrics" (e.g., Scagnostics [48] or the class separation measure [16]), which focus on the perceived patterns of 2D projections without accounting for the original high-dimensional (HD) data. Please refer to Behrisch et al. [19] for detailed descriptions on these metrics.

`Visualization`. Machines get all data cascaded from the previous stages (raw *high-dimensional data*, *preprocessed data*, *projections*, and *scores*) as input, generating a visualization or a set of visualizations delivering *data patterns*. The main configuration that affects the patterns is the visualization methods for depicting DR projection. This incorporates the selection of not only the type and number of visual idioms (e.g., 2D or 3D scatterplots [178]) but also displays (e.g., 2D screen or VR headsets [84]) or the use of animations [11]. The way of encoding axes is another configuration that affects the knowledge generation process. For example, when the projection is made by selecting a few subspaces in `Preprocessing` step, such information can be encoded in axes. Configuring how monochrome scatterplots can be augmented also substantially affects the resulting data patterns. Color-encoding the classes, for example, could cause analysts to overlook the separation between intrinsic clusters within the data [179]. Also, how and how much projections suffer from distortions can be augmented by encoding the scores from the evaluation on the projections [13, 87, 122], making analysts be more aware of distortions. Finally, auxiliary visualizations can be configured to make the original high-dimensional data more interpretable [33, 112, 219].

`Sensemaking`. Analysts make sense of data by investigating *data patterns* delivered by the `Visualization`. The procedure starts by perceiving the patterns, e.g., cluster [1, 90] or local neighborhood [15, 121] patterns, aligned with the target task. The perception usually relies on the Gestalt principle of proximity and similarity [159], where analysts perceive data points that have high proximity or similar color to be closely located in the original high-dimensional space.

`Interaction`. Analysts can also interact with data, meaning they "signal" Machines to update configurations, thereby updating data patterns to align with their intentions. To do so, Analysts first evaluate whether the data patterns are suitable for validating hypotheses. If not suitable, they can interact with Machines to correct the visualizations. For instance, when analysts notice that the evaluation scores of DR projections indicate insufficient accuracy to properly support the target task, they can reconfigure the DR techniques or adjust hyperparameters to correct the projections. The Machines then again process the data based on the updated configuration, which leads to a new cycle of sensemaking.

Model applications. In Appendix A, we verify the completeness and applicability of our workflow model by instantiating it on two research works, contributing interactive visualizations leveraging DR: *AxiSketcher* [112] and *CommentVis* [219].

5 Taxonomy

We derive four dimensions—Stages, Problems, Aims, and Solutions—that explain existing research papers that address the unreliability of visual analytics using DR. Our taxonomy builds upon the workflow model (Sect. 4): every branch of the taxonomy is incorporated as an element in the workflow model or can be explained using the elements. Our taxonomy provides a solid foundation for organizing and classifying relevant papers (Table 1), thereby helping practitioners gain a clearer understanding of the field.

5.1 Stages

This dimension designates which workflow stage (Sect. 4.2) causes unreliability. This dimension encompasses four Machine-side stages: Preprocessing , DR , Quantitative Evaluation , and Visualization . Note that we do not include the Analyst-side stages (Sensemaking and Interaction) because unreliable visual analytics implies these stages to be problematic. In other words, all papers we identify face issues in the Analyst-side stages. For example, if DR projections become inaccurate due to distortions, this leads to inaccuracies in both the sensemaking of the original structure and the interaction with the projection (e.g., the clusters brushed by Analysts may not represent actual clusters in the original space). Please refer to Sect. 4.2 for detailed explanations of these stages.

Note that the stages where the unreliability is caused may not necessarily align with the stages where the Solutions are implemented. For example, accuracy issues in a dimensionality reduction (DR) technique may be addressed through the development of novel DR approaches [88, 94] or by employing auxiliary visual encodings to demonstrate where and how inaccuracies occur.

5.2 Problems

This dimension details why visual analytics using DR become unreliable. This dimension consists of eight classes.

Inaccurate. A stage is considered inaccurate when errors occur during its execution, even though Analysts have set configurations that align with their intentions. In terms of DR stage, this problem is widely represented by the term "distortion" [13, 87, 122, 154], which is frequently characterized as an unavoidable [89, 186] error in projection arising from the intrinsic complexity of high-dimensional data. For example, Lespinats and Aupetit [121] address distortions of neighborhood structure in DR projections, and Martins et al. [134] focus on the distortions that occur at the cluster level. Few papers state that Quantitative Evaluation can also be unreliable due to inappropriate design of DR evaluation metrics [89] or study design [14].

Suboptimal. Papers addressing suboptimality problems focus on situations where a particular configuration does not align with the target task. For example, Narayan et al. [150] and Jeon et al. [88] improve UMAP to better support the analytic tasks that examine global structure (e.g., cluster arrangement, pairwise distances between points). Similarly, Meng et al. [138] and Hajderanj et al. [81] claim that the original *t*-SNE does not well fit to classification task and thus improved it be supervised. Some papers also tackle the suboptimality in Visualization stage. For instance, Bian et al. [21]

improve the conventional scatterplot to better support the exploration of subspace by changing the encoding of dots representing data points.

Incomplete. If analysts cannot find a single configuration that matches their intention, the corresponding stage is incomplete. This problem mainly occurs during Quantitative Evaluation stage. Sips et al. [186] and Friedman et al. [67] account for the problem that no existing DR evaluation metric examines class separability. Johansson and Johansson [93] propose weighting DR evaluation metrics to make the evaluation complete against arbitrary user tasks. The problem also occurs in DR stage, where existing DR techniques cannot cover diverse analytic tasks [67, 113].

Unstable. A stage is unstable if the output (e.g., projections in DR stage) varies as configuration changes. For example, nonlinear DR techniques like *t*-SNE and UMAP can produce different projections when different hyperparameter settings or initialization methods are used [61, 106]. The randomness in the optimization procedure can also make DR techniques unstable [95, 211]. The papers thus make DR stage more robust to such changes [61] or recommend the best configuration [106]. The researchers also provide similar endeavors to remedy instability in DR evaluation metrics (Quantitative Evaluation stage) [7, 92, 117].

Uninterpretable. A stage is less interpretable if (1) the underlying mechanism of the stage or (2) the input data to the stage is not well explained to the Analysts. The former problem widely occurs in DR and Quantitative Evaluation stages due to the intrinsic complexity of DR techniques [33, 39] or evaluation metrics [116]. The latter problem mainly stems from Visualization stage. For example, scatterplots depicting DR projections can hardly inform the original attribute values to the analysts [53, 62, 112]. The scatterplots also mostly represent data items using dots, which have problems depicting the original data format, e.g., text [97].

Unscalable against dimensionality. The unscalability problem occurs when the complexity of final data patterns grows with increasing dimensionality, exceeding the available display pixels or human cognitive capabilities. For example, DR stage can be unscalable for subspace analysis as the number of subspaces to explore increases exponentially against dimensionality [67, 114, 149, 198]. Some papers claim that the unscalability originates from the Visualization stage due to the fundamental limitation of scatterplots: each scatterplot can only depict a single projection at once [96, 181].

Irreflective of domain knowledge. If a stage hardly reflects the domain knowledge of Analysts, it means that resulting data patterns (external representation of data) do not align with the internal cognitive map of the analysts [9, 10]. Kim et al. [103] and Kwon et al. [112] address the problem that Visualization stage becomes irreflective of domain knowledge as static scatterplots cannot receive user input. Xia et al. [214] and Brown et al. [25] claim that the problem originates from the DR stage as DR techniques are executed independently with domain knowledge.

Uninformed. Analysts are uniformed to a stage if they are unaware of (i.e., lack operational knowledge on) diverse techniques used in the stage. Every stage can thus be uninformed to analysts.

Analysts are often uninformed about the other reliability problems, e.g., suboptimal Quantitative Evaluation [14, 89]. The difference between **uninformed** and **uninterpretable** is that the former can lead analysts to establish inappropriate configurations, while the latter incurs unreliability during Sensemaking and Interaction.

5.3 Aims

The Aim dimension is the ultimate objective that each paper seeks to accomplish. While the first class (Enhance reliability) aims to directly "fix" the problem, the latter two classes (Enhance awareness, Enhance approachability) purpose to empower Analysts to help them address the problem on their own.

Enhance reliability. The papers in this category aim to enhance reliability by providing newly designed techniques regardless of a stage or a problem [4, 85]. Some papers do so by introducing the improved version of the existing techniques [88, 138, 150].

Enhance awareness. The papers inform Analysts to be more aware of the possible configurations, addressing the problem that the stages are **uninterpretable** [96, 112] or **uninformed** [154, 216]. This incorporates the efforts of making Analysts more aware of all kinds of reliability problems, e.g., **inaccuracy** [13] and **suboptimality** [45].

Enhance approachability. These papers aim to make Analysts easier to **interpret** techniques [31, 33], handle the **instability** of the techniques [141], and more easily inject their **domain knowledge** [103, 112].

5.4 Solutions

This dimension explains the contribution of the papers that materializes their Aim. The dimension is classified into eight classes. Note that the first three categories (Improvements in Preprocessing, Quantitative Evaluation, and Visualization) differ from the categories in the Stages; the former indicates where reliability is addressed, while the latter identifies where the reliability issues originate. There is no class entitled "Improvements in Preprocessing stage" as we find no papers that provide solutions that improve Preprocessing stage.

Improvements in DR stage. Papers in this category provide solutions for improving the DR stage, thus addressing the problems in the same stage. This includes the introduction of new DR techniques [61, 190], improvement of DR techniques [88, 138, 205], or the improvement of a certain subpart of DR techniques, e.g., initialization [106] or distance functions of nonlinear DR techniques [25]. Jeon et al. [88], for instance, addresses the **inaccuracy** and **suboptimality** of UMAP in representing global structure by altering the optimization process.

Improvements in Quantitative Evaluation stage. As with the previous category, these papers here provide solutions related to the Quantitative Evaluation stage, aiming to address problems in the Evaluation and previous stages. For instance, the papers introduce new DR evaluation metrics both to address **inaccuracy** of DR or Preprocessing stage [147] or to resolve **incompleteness** or **instability** of the Evaluation stage [87, 116]. Another case

is a library offering DR evaluation metrics, which improves their **approachability** [86].

Improvements in Visualization stage. These papers improve the Visualization stage, addressing problems in not only the Visualization but also previous stages. One notable types of paper in this category depict where and how much the DR projections are distorted by overlying such info over projections [87, 121] or dynamically relocating points to "correct" the distortions [82, 85, 191], dealing with **inaccuracy** in DR stage.

DR framework. These papers contribute a framework that automatically specifies configurations across multiple stages. For example, Choo et al. [38] propose a framework that establishes a configuration that mixes two DR techniques, aiming to address each technique's **suboptimality**. Fujiwara et al. [70] aim to find a set of configurations in Preprocessing and DR stage that maximizes the diversity of the patterns found in subspace analysis.

Visual analytics system. The papers in this category contribute a visual analytics system, These works often address **interpretability** issues of DR or Visualization stage to **enhance awareness** of DR techniques [33, 46, 91] and resulting projections [215]. We classify the papers to **improvements in Visualization stage** category if they are more focused on proposing novel visualization design and classify them to this category if their main contribution is a dashboard system consisting of linked visualizations.

Literature review. The researchers contribute literature review on DR techniques (DR stage) [57, 154] and evaluation metrics (Quantitative Evaluation stage) [86, 154], mainly aiming to **enhance awareness** of Analysts by informing their characteristics. These papers serve as "guidelines" for Analysts in setting up appropriate configurations. We can interpret the contribution of these papers as the authors' effort to share their operational knowledge.

Human-centered experiment. This type of research work is similar to **literature reviews**, but provides evidence about their guidelines by conducting user studies [60, 216]. For example, they determine the usability of DR techniques based on human task performance on their projections. Their guidelines are thus easy to apply in practice: configurations are directly mapped to suitable tasks. These papers can be interpreted as the authors performing Quantitative Evaluation and Sensemaking to gain and share operational knowledge.

Computational experiment. The papers in this category serve the same role as those in the **human-centered experiment** category, except that the evaluation of DR techniques is conducted computationally using DR evaluation metrics. Their guidelines are thus less applicable compared to the ones from human-centered experiment. For example, even if we use DR evaluation metrics that examine the preservation of cluster structure (e.g., Steadiness & Cohesiveness), we are not sure about whether metric scores directly align with human performance in conducting cluster analysis.

6 Meta-Analysis on the Taxonomy

We systematically investigate how the classified papers (Table 1) can be categorized at a higher level. Our analysis results help practitioners quickly overview the papers we identify, thereby reducing

Table 1: *The list of 133 papers we survey (rows) and their classification based on the proposed taxonomy (columns).* Papers within each group and type are arranged in descending order of their total citation as of September 2024, based on Google Scholar.

Group	#	Paper	Year	Preprocessing	Dimensionality Reduction	Quantitative Evaluation	Visualization	inaccurate	suboptimal	incomplete	unstable	uninterpretable	unscalable	irreflective	uninformed	Enhance reliability	Enhance awareness	Enhance Approchability	Improvements on DR	Improvements of Evaluation	Visualization Stage Solutions	DR framework	Visual Analytics System	Literature Review	Human-Centered Experiment	Computaional Experiment
Pioneer (type 1)	1	Roweis and Saul [171]	2000		✓				✓							✓			✓							
	2	Tenenbaum et al. [195]	2000		✓				✓							✓			✓							
	3	Friedman and Tukey [68]	1974		✓				✓							✓			✓							
	4	Friedman [67]	1987		✓					✓						✓			✓							
	5	De silva and Tenenbaum [185]	2002		✓				✓							✓			✓							
	6	Geng et al. [73]	2005		✓					✓						✓			✓							
	7	Venna et al. [203]	2010		✓	✓			✓							✓	✓		✓							
	8	Paulovich et al. [162]	2008		✓					✓						✓		✓	✓							
	9	Lespinats et al. [122]	2007		✓				✓					✓			✓			✓							
	10	Koren and Carmel [107]	2003		✓				✓	✓						✓			✓							
	11	Wang et al. [209]	2018		✓				✓							✓			✓	✓						
	12	Choo et al. [38]	2009		✓				✓							✓			✓				✓			
	13	Lee et al. [119]	2014		✓				✓							✓			✓				✓			
	14	Lehmann and Theisel [120]	2013		✓			✓								✓			✓							
	15	Novakova and Stepankova [155]	2009		✓			✓		✓						✓			✓							
	16	Narayan et al. [150]	2020		✓				✓							✓			✓							
	17	Lee and Verleysen [118]	2011		✓								✓			✓	✓		✓					✓		
	18	Fadel et al. [61]	2015		✓				✓		✓					✓			✓							
	19	Wang et al. [210]	2017		✓	✓				✓	✓					✓		✓	✓							
	20	Zhou et al. [227]	2016	✓	✓					✓						✓			✓							
	21	Hajderanj et al. [81]	2019		✓				✓							✓			✓							
	22	De bodt et al. [50]	2019		✓				✓							✓			✓							
	23	Colange et al. [42]	2020		✓			✓		✓						✓			✓							
	24	Pagliosa and Telea [158]	2019		✓								✓			✓			✓							
	25	Amorim et al. [6]	2014	✓	✓						✓				✓							✓			
	26	Castermans et al. [30]	2019		✓			✓						✓	✓			✓							
	27	Jeon et al. [88]	2022		✓			✓	✓							✓			✓							
	28	Hossain et al. [83]	2021		✓			✓	✓	✓	✓					✓			✓							
	29	Molchanov and Linsen [142]	2019		✓				✓							✓			✓							
	30	Vu et al. [205]	2021		✓				✓					✓		✓			✓							
	31	Meng et al. [138]	2024		✓				✓							✓			✓							
	32	Alcaide and Aerts [4]	2021		✓				✓							✓			✓							
	33	Matute and linsen. [136]	2020		✓					✓						✓			✓							
Pioneer (type 2)	34	Joia et al. [94]	2011		✓								✓			✓		✓	✓		✓					
	35	Endert et al. [55]	2011		✓								✓			✓		✓	✓		✓					
	36	Liu et al. [128]	2018		✓	✓		✓	✓							✓			✓					✓		
	37	Paulovich et al. [161]	2011		✓								✓			✓			✓		✓					
	38	Gleicher [76]	2013		✓					✓			✓			✓			✓		✓					
	39	Dos santos amorim et al. [52]	2012		✓	✓						✓				✓			✓	✓	✓					
	40	Liu et al. [130]	2015		✓	✓						✓	✓			✓			✓		✓					
	41	Zhou et al. [226]	2015		✓			✓				✓				✓			✓	✓	✓					
	42	Rubio-sanchez and Sanchez [173]	2014		✓							✓				✓			✓							
	43	Xia et al. [214]	2023		✓			✓					✓			✓			✓		✓					
	44	Jackle et al. [96]	2017		✓	✓							✓			✓			✓		✓					
	45	Cutura et al. [45]	2020		✓	✓			✓			✓				✓	✓	✓			✓					
	46	Bian et al. [21]	2021		✓	✓			✓	✓						✓					✓					
	47	Lai et al. [113]	2018		✓					✓			✓						✓	✓						
	48	Spathis et al. [190]	2018		✓								✓			✓			✓	✓						
Judge	49	Kruskal [110]	1964			✓		✓								✓				✓						
	50	Sips et al. [186]	2009		✓	✓		✓	✓							✓				✓						
	51	Johansson and Johansson [93]	2009			✓			✓			✓				✓		✓		✓				✓		
	52	Leban et al. [114]	2006		✓								✓			✓	✓			✓						
	53	Lee and Verleysen [117]	2010			✓			✓		✓					✓	✓			✓						
	54	Albuquerque et al. [2]	2010			✓	✓		✓							✓				✓						
	55	Rieck et al. [167]	2015			✓			✓							✓					✓					
	56	Albuquerque et al. [3]	2011			✓			✓							✓				✓						
	57	Motta et al. [147]	2015		✓	✓				✓						✓	✓			✓						
	58	Aupetit [14]	2014			✓		✓					✓			✓				✓						
	59	Rieck and Leitte [168]	2017			✓						✓				✓				✓				✓		
	60	Tian et al. [197]	2021			✓							✓			✓					✓					
	61	Jeon et al. [87]	2021		✓	✓		✓	✓							✓	✓			✓						
	62	Angelini et al. [7]	2022			✓		✓			✓					✓			✓	✓						
	63	Jeon et al. [89]	2024			✓		✓								✓	✓			✓						
	64	Jeon et al. [86]	2023													✓		✓		✓					✓	
	65	Zhang et al. [225]	2023			✓			✓							✓		✓					✓			

Category	#	Reference	Year	Preprocessing	Dimensionality Reduction	Quantitative Evaluation	Visualization	inaccurate	suboptimal	incomplete	unstable	uninterpretable	unscalable	irreflective	uninformed	Enhance reliability	Enhance awareness	Enhance Approchability	Improvements on DR	Improvements of Evaluation	Visualization Stage Solutions	DR framework	Visual Analytics System	Literature Review	Human-Centered Experiment	Computaional Experiment
Instructor	66	Van der maaten and Postma [200]	2009		✓										✓		✓									✓
	67	Cunningham and Ghahramani [44]	2015		✓										✓		✓		✓					✓		
	68	Sorzano et al. [189]	2014		✓										✓		✓							✓		
	69	Venna and Kaski [202]	2006		✓				✓						✓		✓		✓							✓
	70	Lee and verleysen [116]	2009			✓		✓	✓	✓					✓		✓									✓
	71	Espadoto et al. [59]	2021		✓										✓		✓							✓		
	72	Nonato and Aupetit [154]	2019		✓	✓	✓								✓		✓							✓		
	73	Kobak and Linderman [106]	2021		✓							✓			✓		✓		✓							
	74	Sedlmair et al. [178]	2013				✓								✓		✓								✓	
	75	Engel et al. [57]	2012		✓				✓						✓		✓							✓		
	76	Gisbrecht and Hammer [75]	2015		✓										✓		✓							✓		
	77	Lueks et al. [140]	2013			✓						✓	✓			✓	✓									✓
	78	Rubio-sánchez et al. [172]	2016		✓										✓		✓									✓
	79	Tatu et al. [193]	2010			✓						✓			✓		✓								✓	
	80	Etemadpour et al. [60]	2015		✓	✓			✓					✓	✓		✓								✓	
	81	Xie et al. [218]	2017	✓	✓			✓				✓			✓		✓						✓			
	82	Xia et al. [216]	2022		✓										✓		✓								✓	
	83	Cutura et al. [46]	2018		✓										✓		✓							✓		
	84	García fernández et al. [72]	2013	✓	✓				✓	✓					✓		✓									✓
	85	Ventocilla and Riveiro [204]	2020		✓				✓	✓					✓		✓							✓		
	86	Atzberger et al. [12]	2024	✓	✓										✓		✓									✓
	87	Johannemann and Tibshirani [92]	2019			✓					✓	✓			✓		✓									✓
Explorer	88	Asimov [11]	1985				✓					✓						✓			✓					
	89	Cook et al. [51]	1995				✓					✓						✓			✓					
	90	Yi et al. [223]	2005				✓				✓	✓						✓			✓					
	91	Nam and Mueller [149]	2013				✓					✓						✓					✓			
	92	Wang and Mueller [206]	2017	✓								✓						✓					✓			
	93	Liu et al. [127]	2016				✓					✓						✓				✓				
Explainer (type 1)	94	Aupetit [13]	2007		✓			✓									✓			✓	✓					
	95	Martins et al. [134]	2014		✓			✓			✓						✓				✓					
	96	Lespinats et al. [121]	2011		✓	✓		✓									✓				✓					
	97	Schreck et al. [177]	2010		✓	✓		✓								✓	✓				✓					
	98	Martins et al. [135]	2015		✓	✓		✓								✓	✓				✓					
	99	Seifert et al. [180]	2010		✓	✓						✓					✓				✓					
	100	Heulot et al. [82]	2013		✓		✓	✓								✓	✓				✓					
	101	Colange et al. [43]	2019			✓	✓	✓									✓				✓					
	102	Jeon et al. [85]	2022		✓		✓	✓								✓	✓				✓					
Explainer (type 2)	103	Kim et al. [103]	2016				✓		✓			✓		✓			✓				✓					
	104	Kandogan [97]	2012				✓					✓					✓				✓					
	105	Kwon et al. [112]	2017		✓		✓		✓				✓	✓			✓				✓					
	106	Kaski et al. [99]	2000		✓		✓										✓				✓					
	107	Da silva et al. [184]	2015		✓							✓					✓				✓					
	108	Coimbra et al. [41]	2016		✓		✓					✓					✓				✓					
	109	Faust et al. [62]	2019		✓		✓					✓				✓	✓				✓					
	110	Pagliosa et al. [157]	2016		✓		✓					✓					✓				✓					
	111	Sohns et al. [188]	2021		✓		✓					✓					✓				✓					
Architect	112	Jeong et al. [91]	2009		✓							✓					✓						✓			
	113	Brown et al. [25]	2012		✓									✓			✓						✓			
	114	Choo et al. [39]	2010		✓							✓					✓						✓			
	115	Stahnke et al. [191]	2016		✓		✓	✓				✓					✓						✓			
	116	Turkay et al. [198]	2011		✓						✓	✓	✓				✓						✓			
	117	Chatzimparmpas et al. [33]	2020		✓	✓			✓							✓	✓						✓			
	118	Buja and Swayne [27]	2002	✓	✓				✓			✓					✓						✓			
	119	Cavallo and Demiralp [31]	2018		✓		✓				✓					✓	✓						✓			
	120	Xia et al. [215]	2017	✓	✓					✓		✓				✓	✓						✓			
	121	Mamani et al. [132]	2013		✓		✓					✓				✓	✓						✓			
	122	Liu et al. [129]	2014			✓						✓					✓						✓			
	123	Dowling et al. [53]	2019		✓		✓					✓					✓						✓			
	124	Choo et al. [40]	2013	✓	✓							✓					✓						✓			
	125	Fernstad et al. [64]	2013		✓	✓			✓	✓		✓					✓						✓			
	126	Marcílio-jr and Eler [133]	2021		✓		✓					✓					✓						✓			
	127	Self et al. [181]	2018		✓		✓					✓				✓	✓								✓	
	128	Ma and Maciejewski [131]	2021		✓		✓									✓	✓						✓			
	129	Sun et al. [192]	2022	✓	✓												✓						✓			
	130	Li and Zhou [125]	2023		✓							✓	✓				✓						✓			
	131	Ghosh et al. [74]	2022			✓	✓					✓	✓	✓			✓						✓			
	132	Fujiwara et al. [70]	2023		✓						✓	✓					✓				✓					
	133	Molchanov et al. [141]	2020		✓						✓	✓					✓			✓						

Pioneer

Type 1

Jeon et al.
[VIS 2022]

t-SNE   UMAP   **UMATO**

Contributed UMATO to address **suboptimality** of UMAP in supporting
global structure investigation task

Type 2

Joia et al.
[TVCG 2011]

Contributed LAMP, a DR technique that dynamically updates the projection
algorithm based on user input to better **reflect domain knowledge**

Judge

Aupetit
[BELIV 2014]

Proposed to revise class labels to make a label-based evaluation of DR
projections more **accurate**. Jeon et al. [TVCG 2024] later introduced a DR
evaluation metric that shares the common goal.

Instructor

Sedlmair et al.
[TVCG 2013]

Compared the utility of 2D scatterplots, 3D scatterplots, and SPLOMs in class
separability task, making analysts **more aware of** appropriate visual encodings

Explorer

Nam and Mueller
[TVCG 2017]

Introduced TripAdvisor^{N-D}, a visual analytics system that supports the
scalable exploration of data **subspaces**

Explainer

Type 1

Lespinats & Aupetit
[CGF 2011]

Contributed CheckViz that resolves **inaccuracies** in DR projections by
visualizing distortions

Type 2

Faust et al.
[TVCG 2019]

Contributed DimReader that **interprets** how high-dimensional
attribute values change over the nonlinear DR projections

Architect

Chatzimparmpas et al. [TVCG 2020]

Proposed visual analytics system called t-viSNE, which aims to better **interpret**
underlying data and help analysts to be **more aware of** t-SNE technique

Figure 4: *Example papers in each cluster we identify by conducting meta-analysis* **(Sect. 6). The problems (Sect. 5.2) that each
system resolves are highlighted in bold. Reference papers:** Pioneer **[88, 94],** Judge **[14],** Instructor **[178],** Explorer **[149],**
Explainer **[62, 121], and** Architect **[33].**

their cognitive load to understand the research landscape and make their visual analytics more reliable. The analysis findings also serve as a foundation for identifying ongoing research challenges (Sect. 7) and imbalances in the research landscape (Sect. 6.3, 8.4).

6.1 Objectives and Design

We group papers that are similarly classified according to our taxonomy (Sect. 5). We follow the data-driven approach of clustering papers proposed by the survey conducted by Shin et al. [183].

Cluster identification. We first convert the categorization of papers into a data table. The rows of the data table correspond to the papers, and the columns correspond to each category in our taxonomy (Table 1). Each paper is thus represented as a binary vector, where each element is 1 if the paper belongs to the corresponding category and 0 if it does not.

We then apply three different clustering algorithms (hierarchical clustering [148], DBSCAN [108], and HDBSCAN [29]) while using Jaccard distance to measure the dissimilarity between data points. To find the optimal clustering results, we test various hyperparameter settings with Bayesian optimization [187] while using Silhouette coefficient [170] as a target variable. The total number of iterations for Bayesian optimization is 100. We select the clustering results with the best Silhouette coefficient score. We use `scikit-learn` implementation (`AgglomerativeClustering`) [163] for clustering algorithms and Nogueira's implementation [153] for the Bayesian optimization.

Manual validation. As clustering algorithms do not always output optimal clustering results [20], two authors manually screen the papers to validate the individual compactness and mutual separability of the clusters [20]. The authors revise the grouping by merging clusters that have papers with similar characteristics (See Sect. 6.2 for results).

6.2 Identified Clusters

We find that hierarchical clustering yields the best result, consisting of seven clusters with a Silhouette score of 0.292. After manual validation, the authors merge two clusters with papers with similar purposes, resulting in six clusters: Pioneer , Judge , Instructor , Explorer , Explainer and Architect . The following are detailed explanations of these clusters, characterized by their role in visual analytics using DR. Please refer to Table 1 to the list of papers within each cluster (linked by color). Also refer to Fig. 4 for the example figures of the papers within each cluster we discuss.

Pioneer . Papers in this cluster "pioneer" reliable visual analytics by contributing methodologies to produce better DR projections. One branch of papers within this cluster proposes new DR techniques [94, 136] or improves the existing DR techniques [73, 88, 138] to address **incompleteness**, **inaccuracy**, and **suboptimality** in DR stage (**type 1**; row 1–33 in Table 1). For instance, Jeon et al. [88] address the limitations of UMAP in supporting global structure investigation tasks. These papers contribute static DR techniques, which means that they are not updated or affected by Interaction . In contrast, the remaining papers propose DR techniques that interactively update the inner logic based on user input (**type 2**; row

34–48). For example, Joia et al. [94] propose to dynamically update the projection algorithm by reflecting the user interaction that updates the 2D positions of points.

Judge . The papers in this cluster address problems regarding Quantitative Evaluation stage by proposing new DR evaluation metrics or strategies, supporting Analysts to reliably judge the quality of DR projections. For example, Aupetit [14] reveals **inaccuracy** in evaluating cluster structure of projections by using class labels as ground truth clusters. The paper not only makes aware of the issue but also remedies it by proposing to revise the class labels to better reflect cluster structure. Jeon et al. [89] later provide a more tangible solution to this issue by introducing new DR evaluation metrics.

Instructor . The papers in this cluster teach Analysts (i.e., make them **aware** of) how to establish effective configurations, mostly for DR stage. These papers mostly contribute **literature reviews**, **human-centered experiment**, and **computational experiment**. One exception is Cutura et al. [46], where the authors contribute to a **visual analytics** system that explains and compares different DR techniques. A few papers in this cluster also guide setting effective configurations in Visualization stage. For example, Sedlmair et al. [178] conduct an experiment comparing different visualization types (2D scatterplot, 3D scatterplot, and SPLOM) in supporting the class separability tasks using DR projections.

Explorer . These papers aim to make the exploration of subspaces more **approachable**. These papers address the **scalability** problem, which arises from the increasing number of subspaces that need to be investigated as dimensionality increases. For example, Asimov [11] propose to animate scatterplots to allow analysts to "tour" the possible set of subspaces, and Nam and Mueller [149] elaborate this approach based on the metaphor of tourism in our real life.

Explainer . The papers here mostly address **inaccuracy** or **uninterpretability** of DR or Visualization stages by augmenting scatterplots. In terms of accuracy, this is typically done by overlaying how and where projections are distorted [13, 121, 135] or resolving distortions by moving them [82, 85] (**type 1**; rows 94–102). Similarly, interpretability is addressed by overlying high-dimensional attribute values to the scatterplots [62] (**type 2**; rows 103–111).

Architect . These papers architect **visual analytics systems** that help people better understand DR techniques or underlying data. For example, Jeong et al. [91] and Chatzimparmpas et al. [33] contribute visual analytics systems that enhance Analysts's awareness of the inner logic of PCA and t-SNE, respectively. On the other hand, Xia et al. [215] and Liu et al. [129] contribute visual analytics systems that help users conduct more reliable analysis despite DR techniques' inherent **incompleteness** and **instability**.

6.3 Implication: Imbalance in the Research Landscape

Our meta-analysis reveals the imbalance of research topics in the field. While we identify 48 Pioneer -type papers (36.1%), we find a relatively small number of papers for other types of papers. For example, there exists only 17 (12.8%), and 22 (16.5%) of Judge -type papers and Instructor -type papers. The situation is worse for

Explorer , as it only incorporates six papers. The finding suggests that the research community invests substantial effort in developing new DR techniques but comparatively less effort in evaluating them and making them interpretable. From the perspective of our taxonomy, this result suggests that while we have thoroughly studied the improvement of low-level performance in DR techniques (e.g., **inaccuracy**, **suboptimality**, **incompleteness**), the challenges associated with their high-level usage remain underexplored (e.g., **uninterpretability** and the **lack of reflection to domain knowledge**).

Such imbalance can potentially harm the reliability of visual analytics. For example, the absence of Instructor -type papers can result in Analysts **being unaware of** newly proposed DR techniques, making it challenging to determine optimal configurations. We discuss tangible solutions to remedy such imbalance in Sect. 7 and Sect. 8.4.

7 Research Challenges

After finishing the literature review and analysis, we again review the papers we identify and conduct a post-review discussion to reveal open challenges in the field. Six authors iteratively participate in the discussions. As a result, we find three crucial challenges in the field.

In this section, we first outline the challenges and suggest which direction the visualization research community should pursue to address them (Sect. 7.1). We then detail expert interviews that verify the significance, urgency, and completeness of our challenges (Sect. 7.2).

7.1 Challenges

Challenge 1: We lack human-centered evaluations. Human-centered evaluations provide applicable guidelines to Analysts (Sect. 5.4). However, we have identified a lack of human-centered evaluations in the field. Among 22 Instructor -type papers, which is already small (Sect. 6.3), only five papers contribute to human-centered experiments.

We moreover identify that we lack recent human-centered evaluations. For DR stage, while we have three papers contributing human-centered experiments to examine the utility of DR techniques in supporting analytics tasks, they either do not cover certain modern techniques (e.g., *t*-SNE or UMAP) [60] or limit their focus to specific analysis tasks (cluster analysis) [216]. The situation is similar for other stages (Preprocessing , Quantitative Evaluation , and Visualization); we need new experiments covering recently proposed techniques that can keep our guidelines up-to-date. For Preprocessing stage, there is no human-centered evaluation at all. Addressing this problem is crucial as preprocessing, like normalizing data attribute values, significantly affects the resulting data patterns [70]. This finding highlights a need for a comprehensive human-centered experiment on modern popular techniques to offer more valuable insights to Analysts.

Call for the research community. We suggest visual analytics researchers to conduct more human-centered experiments across all stages of our workflow model. We first highlight the importance of benchmarking recent DR techniques as they affect data patterns the most. We also want to prioritize the importance of conducting

human-centered experiments on DR evaluation metrics, examining whether their scores align well with human task performance [193]. By knowing such alignment, we can approximate the potential human performance on different DR projections based on **computational experiments**, which is much more cost-efficient. This will thus enable practitioners to identify the most appropriate configurations for DR stage that match the target task just by evaluating DR projections using the metrics. The endeavor will also contribute to mitigating the scarcity of Instructor - and Judge -type papers (Sect. 6.3) by prompting practitioners to pay more attention to the importance of evaluations.

We also recommend the community to incentivize replication studies. As mentioned in Quadri et al. [164], visualization and HCI communities lack replication studies, which prohibits knowledge hardly maintained up-to-date or re-validated. Encouraging researchers to conduct more replication studies will strengthen the robustness of knowledge not only related to DR but also the ones relevant to other areas in visual analytics and HCI. We thus also resonate with existing efforts on incentivizing replication studies[1].

Finally, we suggest researchers build computational models that simulate human perception or cognition [1, 90, 160]. These models cannot fully replace but still can approximate human-centered evaluations in a scalable and cost-effective manner. Moreover, such models can be integrated into the loss function of DR techniques, thus contributing to not only finding effective DR techniques but optimizing their projections [90].

Challenge 2: We require considerations of visual representations beyond 2D scatterplots. While reviewing the list of identified papers, we notice that the visual representations used to depict DR projections are excessively bound to 2D scatterplots. 2D scatterplots are certainly effective at showing spatial proximity between data points and also have high **approachability**, i.e., do not require high visualization literacy to understand. However, we also notify that many reliability problems originate from the use of 2D scatterplots in the Visualization stage. As described in Sect. 5.2, the visualization can be unreliable as 2D scatterplots are **suboptimal** or **less interpretable** [21, 45]. We also find that the 2D scatterplots are **unscalable against dimensionality** [11, 149].

Moreover, our survey implies that scatterplots may not be the best visual representations to make visual analytics more explainable. In Explainer cluster, **accuracy** and **interpretability** problems in DR techniques are addressed by augmenting 2D scatterplots (See Fig. 4). However, interpreting such augmented plots requires additional efforts. This barrier makes it difficult for these advanced visualization techniques to be applied in everyday data analysis.
Call for the research community. We thus suggest investigating the utility of alternative visualization representations. One possibility is to investigate the utility of 3D scatterplots. As 3D space has a higher degree of freedom compared to 2D space, 3D scatterplots can reduce unreliability due to the distortion of DR. Previously, Sedlmair et al. [178] verify that 3D scatterplots have no clear benefit compared to 2D scatterplots in representing cluster separability of DR projections in 2D displays, but other studies have shown different results with immersive analytics technologies: although not yet focused on DR projections and visual analytics, Yang et al.

---

[1]https://ieeevis.org/year/2024/blog/vis-2024-OPC-blog-replication

[220] show that visual cues in immersive environments, especially motion cues, positively affect cluster perception of 3D scatterplots. Pursuing these efforts is also crucial as the effectiveness of immersive environments in data analysis for experts has just started to flourish [84].

We also recommend researchers investigate non-Euclidean layout space [139]. Chen et al. [34, 35] show that mapping graph (i.e., network) data in a torus or sphere instead of a 2D plane significantly improves the accuracy in identifying clusters. Since graph layout can be interpreted as "projecting" data into a low-dimensional space [109], these layout spaces can also be explored for DR. The machine learning community has made contributions in this area [32], but the utility of non-Euclidean spaces for DR has not yet been validated through user studies.

Finally, we encourage researchers to investigate the utility of visual idioms beyond scatterplots, such as parallel coordinates. To the best of our knowledge, this area has been largely unexplored and may hold significant potential for further study.

Challenge 3: We lack libraries. In our survey, we find only one paper that introduces a library [86], which implements a collection of DR evaluation metrics (Quantitative Evaluation stage). This is a surprisingly small number, particularly regarding the visualization community's acknowledgment of the success of libraries (e.g., D3 [23], Vega-Lite [176], and ggdist [100]) and their substantial contribution to the progress of research. The finding verifies that we lack libraries that serve diverse methodologies in a unified interface.

Outside of our survey scope, we have many libraries from the machine learning field for Preprocessing and DR stages. For example, scikit-learn [163] provides diverse DR techniques[2] and preprocessing functions[3] using the Python interface. From a visualization perspective, Druid$_{JS}$ [47] supports four DR techniques in JavaScript, aligning with the trend of web-based development in the field. However, these libraries only support widely used DR techniques (e.g., t-SNE, UMAP, Isomap), lacking support for diverse techniques contributed by the Pioneer -type papers (e.g., LAMP [94], LoCH [61], UMATO [88], NeRV [203]). Still, to the best of our knowledge, no library exists to support the Visualization stage.

More libraries will also help address the shortage of Instructor -type papers (Sect. 6.3), particularly those contributing to **human-centered** and **computational experiments**. For instance, a library that comprehensively encompasses DR techniques would greatly simplify the execution of controlled experiments comparing these techniques.

Therefore, although numerous methodologies to make visual analytics using DR reliable exist, these contributions can hardly be leveraged for practice. Few papers open-source their research artifacts [59, 70, 87, 94, 138], but it is tedious to install and test these codes individually. Analysts thereby may find it challenging to apply these techniques in practice. They might identify a suitable configuration that aligns with their intentions, perhaps by reading this paper, but will struggle to access and execute the corresponding techniques. To address this problem, we need more libraries that serve various techniques with a unified interface.

---

[2]https://scikit-learn.org/stable/modules/manifold.html
[3]https://scikit-learn.org/stable/modules/preprocessing.html

Call for the research community. We advocate visual analytics researchers to build and deploy more libraries that materialize the methodologies contributed by the papers we survey. We especially recommend to build libraries that support Visualization stages. This will help reduce the shortage of Explorer , Explainer , and Architect -type papers by allowing researchers to easily test their new contributions with existing research artifacts.

However, similar to replication studies (see Challenge 1), implementing existing techniques is unlikely to be seen as novel by the community at this time. We thus incentivize the development of libraries of "existing techniques" as valid research contributions. This will significantly save the time of researchers and analysts, allowing them to dedicate more time and effort to do something truly "novel." Here, we want to also emphasize the importance of avoiding overreliance on libraries, as it can lead to unreliable research caused by coding errors. To avoid such a problem, not only deploying libraries but maintaining them should also be incentivized.

We additionally suggest making these libraries more unified. As seen in our workflow (Sect. 4), Analysts should frequently move back and forth between the stages. However, existing libraries support each stage individually, requiring analysts to integrate multiple libraries manually. Providing libraries that allow analysts to establish and manage configurations of every stage simultaneously will substantially enhance the efficiency of data analysis.

7.2 Expert Evaluation

We conduct an expert evaluation to verify the significance, urgency, and completeness of the challenges we identify. Inspired by the Deimos study [115], we recruit eight experts in DR, data visualization, and machine learning and conduct semi-structured interviews that ask the experts to evaluate the challenges.

7.2.1 Study Design. We detail how we recruit and interview participants (i.e., experts).

Experts and recruitment. We recruit eight experts who have published more than three papers related to or leverage DR in major data visualization (e.g., TVCG, VIS, PacificVis, CGF, EuroVis), HCI (e.g., CHI, IUI, TiiS), and machine learning (e.g., NeuIPS, Neurocomputing) venues. We recruit experts by sending them e-mails. The demographics of the experts are depicted in Table 2. As of November 2024, the participants' average number of citations and h-index is 2165.9 (±3015.4) and 15.5 (±4.9), respectively, indicating their sufficient expertise in evaluating our challenges.

Interview protocol. We conduct an interview for each expert, where a single author manages all interview sessions. Once participants give their consent, the author explains the purpose of the research and the interview. The author explains that the challenges are derived from the literature review and details our workflow model and taxonomy. The main interview is then conducted. This part is divided into three sessions, each corresponding to the three identified challenges (Sect. 7.1). During each session, the author outlines the corresponding challenge, including relevant examples and our call for the community. This is done by making a presentation with prepared slides. The experimenter then asks experts to gauge the significance and urgency of the challenge and provide

Table 2: *Demographic information of the expert interview participants.* These experts confirm the significance, urgency, and completeness of the research challenges we identify. "Exp.", "Pub." and "Venue" stands for the years of research experience, the number of archived publications related to or leverage dimensionality reduction published before Nov 2024 (according to Google Scholar), and the venues in which these papers are published, respectively. The main research domains refer to the representative subareas of these DR-related papers. Also note that (A) and (I) stand for academia and industry, respectively.

	Occupation	Gender	Age	Exp.	Pub.	Main research domains	Venue
P1	Assistant Professor (Prof.) (A)	Male	32	10	7	DR acceleration and evaluation	TVCG, VIS, EuroVis
P2	Senior Research Scientist (A)	Male	49	20	20	DR evaluation and interaction	TVCG, CGF, NeurIPS, Neurocomp., etc.
P3	Assistant Prof. (A)	Female	31	7	3	Bioinformatics, XAI	TVCG, CHI
P4	Assistant Prof. (A)	Male	29	7	8	DR Interpretation, XAI	TVCG, CGF, EuroVis, IVIS, PacificVis
P5	Research Scientist, Adjunct Prof. (A)	Male	36	14	6	DR interaction, Clustering	TVCG, IUI, VDS, TiiS
P6	Research Engineer (I)	Male	30	5	11	DR algorithm and interpretation	TVCG, C&G, IVIS, JoV, etc.
P7	Software Engineer (I)	Female	30	7	4	Machine learning, XAI	NeurIPS, Distill
P8	Assistant Prof. (A)	Male	40	15	17	DR evaluation and interpretation	TVCG, CGF, EuroVis, EuroVA, etc.

Table 3: *The list of questions we asked experts to evaluate our challenges.* Significance and urgency are asked for each question, while completeness is assessed across all three challenges collectively.

Criteria	No.	Questions
Significance	Q1	*The challenge addresses an important problem within its field.*
	Q2	*Addressing this challenge has the potential to significantly advance current knowledge of our research field.*
	Q3	*Addressing this challenge could lead to significant improvements in practical applications.*
Urgency	Q4	*The challenge is urgent in the context of current trends in research.*
	Q5	*Addressing this challenge will prevent negative consequences that can potentially happen in practice.*
Completeness	Q6	*The list is complete; they cover important challenges that hinder making visual analytics more reliable.*

their reasoning (questionnaires in Table 3; Q1–Q5). After the sessions are finished, the experimenter asks the experts to review the completeness of the set of challenges (Q6), also with reasonings. To fully elicit the experts' insights, we do not limit the interview duration. Still, all interviews end within 50 minutes. An equivalent of 20 USD is paid for the compensation.

7.2.2 Results and Discussions. We discuss how experts evaluated our challenges. Overall, experts agree on the significance and urgency of the challenges while having mixed perspectives on the completeness.

Significance and urgency of the challenge 1 (Lack of human-centered experiments). Experts generally agree with the significance and urgency of Challenge 1. As seen in Table 4, the average scores for all questions are greater than 4.00. Experts' reasoning also backs up the significance and urgency of this challenge. For example, P2 note, *"The way we perceive clusters or patterns is not the way the machine computes such clusters and patterns."*.

Significance and urgency of the challenge 2 (Lack of investigation on visual representations beyond 2D scatterplots). As with challenge 1, the evaluation scores indicate that experts overall agree with the significance of the second challenge (Table 4). Meanwhile, participants do not strongly agree on the urgency. Three participants (P1, P2, and P8) mention that the challenge will be difficult to address as they also hardly imagine visual representations other than scatterplots, which implies that more research endeavors should be invested to resolve this challenge.

Significance and urgency of the challenge 3 (Lack of libraries). The experts overall agree or strongly agree that this challenge is timely and significant (Table 4). Six experts strongly agree with Q3, which asks about the practical impact of addressing the challenge. The experts also agree on the importance of building libraries from a research perspective. P4 especially suggests that *"It will be good if we have a library in which everyone can build on top of it, just like D3."*, reaffirming the need for the unified library that covers the entire visual analytics workflow using DR.

Completeness of the challenges. The diverging scores highlight the eight experts' differing opinions on the completeness of the challenges (Table 4). For example, P5 generally agree that the main challenges were well-covered, stating, *"You have hit on three really interesting areas to advance this field."*. On the other hand, P2 and P3, and P8 raise concerns about the completeness of the study. For example, they note that although scalability and computational efficiency do not directly improve reliability, they can indirectly do so by enabling the testing of more configurations. The findings, overall, suggest that unknown challenges remain in the field, and we should invest more effort in systematically identifying and addressing them.

8 Discussions

We extend our discussion by emphasizing the practical value of our contributions and proposing future work to strengthen such aspects.

Table 4: *The experts' evaluation on significance (Sig.), urgency (Urg.), and completeness (Co.) of our challenges*. The experts evaluate the challenges by answering the questionnaires (Table 3) in Litert Scale (1: strongly disagree, 2: disagree, 3: Neutral, 4: agree, 5: strongly agree). We highlight the cell in red if the score indicates agreement, using an opacity scale where 0 represents neutral (score of 3) and 100 represents strong agreement (score of 5), with a linear gradient in between. Overall, the experts agree on the significance and urgency of our challenges while presenting mixed perspectives for completeness.

		P1			P2			P3			P4			P5			P6			P7			P8			Avg.		
		C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
Sig.	Q1	4	4	4	5	5	5	4	4	5	4	4	4	4	4	4	5	4	5	5	5	5	5	3	4	4.50	4.13	4.50
	Q2	2	4	2	5	5	5	4	4	4	5	4	3	4	5	4	5	4	5	4	3	5	5	4	4	4.25	4.13	4.00
	Q3	3	5	5	5	5	5	4	4	5	4	4	5	5	5	4	5	5	5	5	5	5	5	3	5	4.50	4.50	4.88
Urg.	Q4	4	4	5	5	4	5	4	4	4	5	5	5	5	3	4	5	3	4	4	4	5	5	3	4	4.63	3.75	4.50
	Q5	3	4	2	5	4	5	4	3	4	4	4	5	5	4	4	3	4	4	5	4	5	4	4	4	4.13	3.88	4.13
Co.	Q6		4			2			3			4			4			4			4			2			3.38	

8.1 Call for the Analysts

This research mainly motivates visualization researchers to make visual analytics using DR more reliable. We also detail research challenges (Sect. 7) that researchers can put effort into. We believe, however, that Analysts can also contribute to advancing visual analytics, as they are the end users of the products contributed by researchers. We provide two recommendations for analysts that can enhance the reliability of visual analytics using DR.

Share the analysis if possible. We suggest Analysts deploy their analysis if there are no ethical problems, e.g., data privacy. For example, analysts can open-source their code and dataset or publicize analytic dashboards made by business intelligence platforms, e.g., Tableau or Spotfire. This will help researchers build a knowledge base for understanding visual analytics practices [156]. For instance, the deployed codes and dashboards will support researchers in empirically validating and revising our workflow model (Sect. 4).

Moreover, researchers can leverage these deployments to address the challenges we identify (Sect. 7). For example, by investigating which DR techniques or evaluation metrics are widely used in visual analytics, researchers can determine which techniques or metrics should be prioritized when developing the libraries (Challenge 3). Deploying datasets can also benefit researchers by enabling them to include more diverse datasets in human-centered experiments, resulting in more generalizable findings (Challenge 1).

Consider diverse configurations. We also recommend Analysts consider diverse configurations. For example, we suggest analysts test and leverage diverse DR techniques beyond *t*-SNE, UMAP, and PCA. Although not yet validated, we observe that these three techniques dominate the selection of DR methods in visual analytics, despite being unsuitable for certain tasks like cluster density estimation or global pairwise distance investigation [86, 144, 150, 216]. Using alternative techniques that better align with these tasks will enhance the reliability of their analysis. Additionally, if analysts more actively consider diverse configurations, it will motivate researchers to contribute new techniques that address gaps in the literature, creating a positive feedback loop for making visual analytics more reliable.

Here, a barrier to considering diverse configurations will be a **lack of operational knowledge**. We thus suggest Analysts refer to the Instructor -type papers, especially the ones contributed **human-centered** and **computational experiments** to be aware of the pros and cons of diverse configurations.

8.2 Reliability Problems Derived from Analysts

In our survey, we find many papers that aim to make analysts more **informed**. They provide analysts with operational knowledge of different configurations, thereby reducing the chances of analysts making suboptimal configurations. However, though analysts are aware of reliability problems and various configurations, visual analytics can be unreliable due to human bias or errors. We discuss two cases in which awareness does not guarantee reliability.

Suboptimal configuration (despite awareness). Analysts can still make suboptimal configurations even though they are informed of corresponding operational knowledge. As discussed in Sect. 7.1, one notable factor is the availability of actionable codes or libraries. Another factor potentially leading analysts to suboptimal configurations is their preference for the projections. For example, Morariu et al. [145] and Bibal and Frénay [22] quantitatively show that separability between class labels significantly affect human preferences for the projections. Such preferences can bias analysts in setting configurations, e.g., by selecting DR techniques that emphasize cluster structure like UMAP [89]. Identifying factors that can potentially bias configurations and mitigate their effect will be an important future avenue to explore. For example, we can explore how visualizing distortions (Explainer) can mitigate the biased preferences towards well-separated clusters.

Inappropriate Sensemaking. Even if analysts establish appropriate configurations, human Sensemaking can be unreliable. At first, analysts are **unstable** in perceiving data patterns [89], which means that their sensemaking varies by individuals and also by time. Analysts may also place excessive trust in their configurations, even for tasks that the configurations were not originally designed to support (Sect. 8.1).

These examples indicate that our survey does not yet fully cover all potential reliability issues in visual analytics using DR, highlighting the need for future efforts to uncover problems that remain undetected.

8.3 Limitations and Future Work

Our three main contributions—workflow model, taxonomy, and challenges—are theoretically grounded by literature. As future work, we successively aim to verify their validity with real Analysts. For example, we could crawl source codes from GitHub that involve high-dimensional data analysis using DR and evaluate how well they fit with our workflow and taxonomy. Additionally, we can conduct an interview study with industry analysts to reveal their daily challenges in practical data analysis and see how these align with our taxonomy.

We also plan to make our findings to be more actionable. Our survey informs that using proper evaluation metrics is crucial to obtain **accurate** and **optimal** DR projections. However, as a paper, it does not recommend specific metrics to use reflecting the current intentions of analysts. Moving forward, we intend to develop a database that houses specific knowledge extracted from the papers we gather. This will facilitate the creation of an actionable framework that enables analysts to query their intention and obtain responses about the appropriate configurations, similar to the approach taken by Draco [146] for general visualization design.

8.4 Call for the HCI field: Mitigating Imbalance in Research Topics

Our meta-analysis (Sect. 6) reveals the imbalance of the relevant research landscape: while numerous DR techniques are proposed (Pioneer -type papers), only a few works that evaluate or interpret these techniques emerged (Judge , Instructor , Explorer , Explainer , and Architect -type papers) (Sect. 6.3). This imbalance aligns with the dominance of papers from the visualization and machine learning fields in our list, with only two papers originating from the HCI field. The true extent of the imbalance may be larger, as our survey started by searching for papers primarily in the visualization and HCI fields. Such findings indicate that the HCI field should pay more attention to this issue. We call on the community to invest more research efforts in two key perspectives: *System* and *Guidance*.

Efforts in a system perspective. The HCI community has already verified its capability to build new systems that enhance the explainability and usability of machine learning techniques. For example, the community substantially contributes to the explainable AI (XAI) systems [54, 66, 77] and also designs numerous methodologies to make AI techniques more user-friendly [80, 182]. We strongly suggest the HCI community utilize such capability to make DR-based visual analytics more reliable and reduce this imbalance. For instance, designing interactive systems that help novice analysts easily be **aware** of **inaccurcy** of **suboptimality** of DR techniques would be an interesting avenue to explore.

Efforts in a guidance perspective. The imbalance reveals that while visualization and machine learning fields dedicate substantial effort to advancing state-of-the-art DR techniques, their endeavors are less human- or user-centered. These fields seldom invest in making these techniques informed and approachable for everyday use by analysts. As a result, expert researchers and analysts are often familiar with these techniques, whereas novice analysts frequently remain unaware of their existence or applicability. To address this gap, we propose a shift in focus within the HCI community toward creating practical, user-friendly guidance. For example, interview studies can be conducted to investigate why practitioners are **uninformed** of DR techniques or evaluation metrics. Based on the insights from the studies, user-friendly guidance tailored to varying levels of expertise can be established to improve the adoption and usability of advanced methods.

As a seminal step in this direction, we provide a practical guide to accompany our survey (Table 1) to help analysts select the most relevant papers to read based on their expertise level and interests, thereby helping analysts improve the reliability of their visual analytics using DR. This guide addresses a key challenge: While our survey offers a comprehensive overview of the research landscape, providing the same references to all readers is not effective because analysts have varying levels of expertise in DR. In addition, many analysts may not accurately assess their own expertise, making it difficult for them to determine the appropriate next steps in their learning journey. To address these challenges, our guide includes a six-item checklist designed to help analysts evaluate their expertise level. Based on this self-assessment checklist, we provide tailored recommendations for reading papers. Our guide can thus be interpreted as a suggestion for an optimal order for exploring the literature, enabling a more comprehensive and systematic understanding of the relevant field. We release the guide as an interactive web page to enhance readability and approachability, available at https://dr-reliability.github.io/guide.

9 Conclusion

This paper presents several contributions based on a literature review of 133 papers. This includes (1) a workflow model that explains how analysts employ DR in visual analytics, (2) a taxonomy to organize the relevant literature, (3) a meta-analysis that groups the papers at a high level, (4) a discussion of ongoing challenges validated by experts, and (5) an identification of the imbalance in research landscape and the call for the action. Together, these contributions substantially benefit visual analytics using DR by making it more reliable. We expect our research to be a key reference that initiates new endeavors to enhance the reliability of visual analytics.

Acknowledgments

References

[1] Mostafa M. Abbas, Michaël Aupetit, Michael Sedlmair, and Halima Bensmail. 2019. ClustMe: A Visual Quality Measure for Ranking Monochrome Scatterplots based on Cluster Patterns. *Computer Graphics Forum* 38, 3 (2019), 225–236. https://doi.org/10.1111/cgf.13684

[2] Georgia Albuquerque, Martin Eisemann, Dirk J. Lehmann, Holger Theisel, and Marcus Magnor. 2010. Improving the visual analysis of high-dimensional datasets using quality measures. In *2010 IEEE Symposium on Visual Analytics Science and Technology*. 19–26. https://doi.org/10.1109/VAST.2010.5652433

[3] Georgia Albuquerque, Martin Eisemann, and Marcus Magnor. 2011. Perception-based visual quality measures. In *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 13–20. https://doi.org/10.1109/VAST.2011.6102437

[4] Daniel Alcaide and Jan Aerts. 2021. Spanning Trees as Approximation of Data Structures. *IEEE Transactions on Visualization and Computer Graphics* 27, 10 (2021), 3994–4008. https://doi.org/10.1109/TVCG.2020.2995465

[5] El-ad David Amir, Kara L Davis, Michelle D Tadmor, Erin F Simonds, Jacob H Levine, et al. 2013. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nature Biotechnolgy* 31, 6 (2013), 545–552. https://doi.org/10.1038/nbt.2594

[6] Elisa Amorim, Emilio V. Brazil, Luis G. Nonato, Faramarz Samavati, and Mario C. Sousa. 2014. Multidimensional Projection with Radial Basis Function and Control Points Selection. In *2014 IEEE Pacific Visualization Symposium*. 209–216. https://doi.org/10.1109/PacificVis.2014.59

[7] Marco Angelini, Graziano Blasilli, Simone Lenti, Alessia Palleschi, and Giuseppe Santucci. 2022. Effectiveness Error: Measuring and Improving RadViz Visual Effectiveness. *IEEE Transactions on Visualization and Computer Graphics* 28, 12 (2022), 4770–4786. https://doi.org/10.1109/TVCG.2021.3104879

[8] G. Appleby, M. Espadoto, R. Chen, S. Goree, A. C. Telea, E. W. Anderson, and R. Chang. 2022. HyperNP: Interactive Visual Exploration of Multidimensional Projection Hyperparameters. *Computer Graphics Forum* 41, 3 (2022), 169–181. https://doi.org/10.1111/cgf.14531

[9] Daniel Archambault and Helen C. Purchase. 2013. Mental Map Preservation Helps User Orientation in Dynamic Graphs. In *Graph Drawing*. Springer Berlin Heidelberg, 475–486.

[10] Daniel Archambault and Helen C. Purchase. 2013. The "Map" in the mental map: Experimental results in dynamic graph drawing. *International Journal of Human-Computer Studies* 71, 11 (2013), 1044–1055. https://doi.org/10.1016/j.ijhcs.2013.08.004

[11] David Asimov. 1985. The Grand Tour: A Tool for Viewing Multidimensional Data. *SIAM J. Sci. Statist. Comput.* 6, 1 (1985), 128–143. https://doi.org/10.1137/0906011

[12] Daniel Atzberger, Tim Cech, Matthias Trapp, Rico Richter, Willy Scheibel, Jürgen Döllner, and Tobias Schreck. 2024. Large-Scale Evaluation of Topic Models and Dimensionality Reduction Methods for 2D Text Spatialization. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 902–912. https://doi.org/10.1109/TVCG.2023.3326569

[13] Michaël Aupetit. 2007. Visualizing distortions and recovering topology in continuous projection techniques. *Neurocomputing* 70, 7 (2007), 1304–1330. https://doi.org/10.1016/j.neucom.2006.11.018

[14] Michaël Aupetit. 2014. Sanity Check for Class-Coloring-Based Evaluation of Dimension Reduction Techniques. In *Proc. of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*. 134–141. https://doi.org/10.1145/2669557.2669578

[15] Michaël Aupetit and Thibaud Catz. 2005. High-dimensional labeled data analysis with topology representing graphs. *Neurocomputing* 63 (2005), 139–169. https://doi.org/10.1016/j.neucom.2004.04.009

[16] Michael Aupetit and Michael Sedlmair. 2016. SepMe: 2002 New visual separation measures. In *2016 IEEE Pacific Visualization Symposium (PacificVis)*. 1–8. https://doi.org/10.1109/PACIFICVIS.2016.7465244

[17] Chongyang Bai, Xiaoxue Zang, Ying Xu, Srinivas Sunkara, Abhinav Rastogi, Jindong Chen, and Blaise Aguera y Arcas. 2021. UIBert: Learning Generic Multimodal Representations for UI Understanding. arXiv:2107.13731 [cs.CV]

[18] Etienne Becht, Leland McInnes, John Healy, Charles-Antoine Dutertre, Immanuel WH Kwok, Lai Guan Ng, Florent Ginhoux, and Evan W Newell. 2019. Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology* 37, 1 (2019), 38–44. https://doi.org/10.1038/nbt.4314

[19] M. Behrisch, M. Blumenschein, N. W. Kim, L. Shao, M. El-Assady, J. Fuchs, D. Seebacher, A. Diehl, U. Brandes, H. Pfister, T. Schreck, D. Weiskopf, and D. A. Keim. 2018. Quality Metrics for Information Visualization. *Computer Graphics Forum* 37, 3 (2018), 625–662. https://doi.org/10.1111/cgf.13446

[20] Shai Ben-David and Margareta Ackerman. 2008. Measures of Clustering Quality: A Working Set of Axioms for Clustering. In *Advances in Neural Information Processing Systems*, Vol. 21. Curran Associates, Inc. https://doi.org/10.5555/2981780.2981796

[21] Rongzheng Bian, Yumeng Xue, Liang Zhou, Jian Zhang, Baoquan Chen, Daniel Weiskopf, and Yunhai Wang. 2021. Implicit Multidimensional Projection of Local Subspaces. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1558–1568. https://doi.org/10.1109/TVCG.2020.3030368

[22] Adrien Bibal and Benoit Frénay. 2016. Learning Interpretability for Visualizations using Adapted Cox Models through a User Experiment. arXiv:1611.06175 [stat.ML]

[23] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D^3: Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2301–2309. https://doi.org/10.1109/TVCG.2011.185

[24] Matthew Brehmer, Michael Sedlmair, Stephen Ingram, and Tamara Munzner. 2014. Visualizing Dimensionally-Reduced Data: Interviews with Analysts and a Characterization of Task Sequences. In *Proc. of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization* (Paris, France) *(BELIV '14)*. ACM, New York, NY, USA, 1–8. https://doi.org/10.1145/2669557.2669559

[25] Eli T. Brown, Jingjing Liu, Carla E. Brodley, and Remco Chang. 2012. Dis-function: Learning distance functions interactively. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 83–92. https://doi.org/10.1109/VAST.2012.6400486

[26] P. Bruneau, P. Pinheiro, B. Broeksema, and B. Otjacques. 2015. Cluster Sculptor, an interactive visual clustering system. *Neurocomputing* 150 (2015), 627–644. https://doi.org/10.1016/j.neucom.2014.09.062

[27] Andreas Buja and Deborah F Swayne. 2002. Visualization methodology for multidimensional scaling. *Journal of Classification* 19, 1 (2002), 7–43.

[28] Kerstin Bunte, Michael Biehl, and Barbara Hammer. 2012. A General Framework for Dimensionality-Reducing Data Visualization Mapping. *Neural Computation* 24, 3 (2012), 771–804. https://doi.org/10.1162/NECO_a_00250

[29] Ricardo J. G. B. Campello, Davoud Moulavi, and Joerg Sander. 2013. Density-Based Clustering Based on Hierarchical Density Estimates. In *Advances in Knowledge Discovery and Data Mining*, Jian Pei, Vincent S. Tseng, Longbing Cao, Hiroshi Motoda, and Guandong Xu (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 160–172. https://doi.org/10.1007/978-3-642-37456-2_14

[30] Thom Castermans, Kevin Verbeek, Bettina Speckmann, Michel A. Westenberg, Rob Koopman, Shenghui Wang, Hein van den Berg, and Arianna Betti. 2019. SolarView: Low Distortion Radial Embedding with a Focus. *IEEE Transactions on Visualization and Computer Graphics* 25, 10 (2019), 2969–2982. https://doi.org/10.1109/TVCG.2018.2865361

[31] Marco Cavallo and Çağatay Demiralp. 2018. A Visual Interaction Framework for Dimensionality Reduction Based Data Exploration. In *Proceedings of the 2018 SIGCHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174209

[32] Ines Chami, Albert Gu, Dat P Nguyen, and Christopher Re. 2021. HoroPCA: Hyperbolic Dimensionality Reduction via Horospherical Projections. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 1419–1429. https://proceedings.mlr.press/v139/chami21a.html

[33] Angelos Chatzimparmpas, Rafael M. Martins, and Andreas Kerren. 2020. t-viSNE: Interactive Assessment and Interpretation of t-SNE Projections. *IEEE Transactions on Visualization and Computer Graphics* 26, 8 (2020), 2696–2714. https://doi.org/10.1109/TVCG.2020.2986996

[34] Kun-Ting Chen, Tim Dwyer, Benjamin Bach, and Kim Marriott. 2021. It's a Wrap: Toroidal Wrapping of Network Visualisations Supports Cluster Understanding Tasks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 465, 12 pages. https://doi.org/10.1145/3411764.3445439

[35] Kun-Ting Chen, Tim Dwyer, Yalong Yang, Benjamin Bach, and Kim Marriott. 2022. GAN'SDA Wrap: Geographic And Network Structured DAta on surfaces that Wrap around. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 135, 16 pages. https://doi.org/10.1145/3491102.3501928

[36] Min Chen and Heike Jäenicke. 2010. An Information-theoretic Framework for Visualization. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (2010), 1206–1215. https://doi.org/10.1109/TVCG.2010.132

[37] Furui Cheng, Mark S Keller, Huamin Qu, Nils Gehlenborg, and Qianwen Wang. 2023. Polyphony: an Interactive Transfer Learning Framework for Single-Cell Data Analysis. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2023), 591–601. https://doi.org/10.1109/TVCG.2022.3209408

[38] Jaegul Choo, Shawn Bohn, and Haesun Park. 2009. Two-stage framework for visualization of clustered high dimensional data. In *2009 IEEE Symposium on Visual Analytics Science and Technology*. 67–74. https://doi.org/10.1109/VAST.2009.5332629

[39] Jaegul Choo, Hanseung Lee, Jaeyeon Kihm, and Haesun Park. 2010. iVisClassifier: An interactive visual analytics system for classification based on supervised dimension reduction. In *2010 IEEE Symposium on Visual Analytics Science and Technology*. 27–34. https://doi.org/10.1109/VAST.2010.5652443

[40] Jaegul Choo, Hanseung Lee, Zhicheng Liu, John Stasko, and Haesun Park. 2013. An interactive visual testbed system for dimension reduction and clustering of large-scale high-dimensional data. In *Visualization and Data Analysis 2013*, Pak Chung Wong, David L. Kao, Ming C. Hao, and Chaomei Chen (Eds.), Vol. 8654. International Society for Optics and Photonics, SPIE, 865402. https://doi.org/10.1117/12.2007316

[41] Danilo B. Coimbra, Rafael M. Martins, Tácito T.A.T. Neves, Alexandru C. Telea, and Fernando V. Paulovich. 2016. Explaining three-dimensional dimensionality reduction plots. *Information Visualization* 15, 2 (2016), 154–172. https://doi.org/10.1177/1473871615600010

[42] Benoît Colange, Jaakko Peltonen, Michael Aupetit, Denys Dutykh, and Sylvain Lespinats. 2020. Steering Distortions to Preserve Classes and Neighbors in Supervised Dimensionality Reduction. In *Advances in Neural Information Processing Systems*, Vol. 33. Curran Associates, Inc., 13214–13225. https://doi.org/10.5555/3495724.3496832

[43] Benoît Colange, Laurent Vuillon, Sylvain Lespinats, and Denys Dutykh. 2019. Interpreting Distortions in Dimensionality Reduction by Superimposing Neighbourhood Graphs. In *IEEE Visualization Conference*. 211–215. https://doi.org/10.1109/VISUAL.2019.8933568

[44] John P Cunningham and Zoubin Ghahramani. 2015. Linear dimensionality reduction: Survey, insights, and generalizations. *The Journal of Machine Learning Research* 16, 1 (2015), 2859–2900. https://jmlr.org/papers/v16/cunningham15a.html

[45] Rene Cutura, Michaël Aupetit, Jean-Daniel Fekete, and Michael Sedlmair. 2020. Comparing and Exploring High-Dimensional Data with Dimensionality Reduction Algorithms and Matrix Visualizations. In *Proceedings of the 2020 International Conference on Advanced Visual Interfaces* (Salerno, Italy) *(AVI '20)*. Association for Computing Machinery, New York, NY, USA, Article 10, 9 pages. https://doi.org/10.1145/3399715.3399875

[46] René Cutura, Stefan Holzer, Michaël Aupetit, and Michael Sedlmair. 2018. VisCoDeR: A tool for visually comparing dimensionality reduction algorithms. In *The European Symposium on Artificial Neural Networks*. https://api.semanticscholar.org/CorpusID:53241481

[47] Rene Cutura, Christoph Kralj, and Michael Sedlmair. 2020. DRUIDJS — A JavaScript Library for Dimensionality Reduction. In *2020 IEEE Visualization Conference (VIS)*. 111–115. https://doi.org/10.1109/VIS47514.2020.00029

[48] Tuan Nhon Dang and Leland Wilkinson. 2014. ScagExplorer: Exploring Scatterplots by Their Scagnostics. In *2014 IEEE Pacific Visualization Symposium*. 73–80. https://doi.org/10.1109/PacificVis.2014.42

[49] George S. Davidson, Brian N. Wylie, and Kevin W. Boyack. 2001. Cluster stability and the use of noise in interpretation of clustering. In *IEEE Symposium on Information Visualization, 2001. INFOVIS 2001*. 23–30. https://doi.org/10.1109/INFVIS.2001.963275

[50] Cyril de Bodt, Dounia Mulders, Daniel López Sánchez, Michel Verleysen, and John A Lee. 2019. Class-aware t-SNE: cat-SNE.. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. 409–414. https://dial.uclouvain.be/pr/boreal/object/boreal:217888

[51] Javier Cabrera Dianne Cook, Andreas Buja and Catherine Hurley. 1995. Grand Tour and Projection Pursuit. *Journal of Computational and Graphical Statistics* 4, 3 (1995), 155–172. https://doi.org/10.1080/10618600.1995.10474674

[52] Elisa Portes dos Santos Amorim, Emilio Vital Brazil, Joel Daniels, Paulo Joia, Luis Gustavo Nonato, and Mario Costa Sousa. 2012. iLAMP: Exploring high-dimensional spacing through backward multidimensional projection. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 53–62. https://doi.org/10.1109/VAST.2012.6400489

[53] Michelle Dowling, John Wenskovitch, J.T. Fry, Scotland Leman, Leanna House, and Chris North. 2019. SIRIUS: Dual, Symmetric, Interactive Dimension Reductions. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 172–182. https://doi.org/10.1109/TVCG.2018.2865047

[54] Upol Ehsan, Philipp Wintersberger, Q. Vera Liao, Elizabeth Anne Watkins, Carina Manger, Hal Daumé III, Andreas Riener, and Mark O Riedl. 2022. Human-Centered Explainable AI (HCXAI): Beyond Opening the Black-Box of AI. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 109, 7 pages. https://doi.org/10.1145/3491101.3503727

[55] Alex Endert, Chao Han, Dipayan Maiti, Leanna House, Scotland Leman, and Chris North. 2011. Observation-level interaction with statistical models for visual analytics. In *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 121–130. https://doi.org/10.1109/VAST.2011.6102449

[56] Alex Endert, William Ribarsky, Cagatay Turkay, B.L. William Wong, Ian Nabney, Ignacio Díaz Blanco, and Fabrice Rossi. 2017. The State of the Art in Integrating Machine Learning into Visual Analytics. *Computer Graphics Forum* 36, 8 (2017), 458–486. https://doi.org/10.1111/cgf.13092

[57] Daniel Engel, Lars Hüttenberger, and Bernd Hamann. 2012. A Survey of Dimension Reduction Methods for High-dimensional Data Analysis and Visualization. In *Visualization of Large and Unstructured Data Sets: Applications in Geospatial Planning, Modeling and Engineering - Proceedings of IRTG 1131 Workshop 2011 (Open Access Series in Informatics (OASIcs), Vol. 27)*, Christoph Garth, Ariane Middel, and Hans Hagen (Eds.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 135–149. https://doi.org/10.4230/OASIcs.VLUDS.2011.135

[58] Mateus Espadoto, Eduardo Faccin Vernier, and Alexandru Telea. 2020. Selecting and Sharing Multidimensional Projection Algorithms: A Practical View. In *VisGap - The Gap between Visualization Research and Visualization Software*, Christina Gillmann, Michael Krone, Guido Reina, and Thomas Wischgoll (Eds.). The Eurographics Association. https://doi.org/10.2312/VISGAP.20201105

[59] Mateus Espadoto, Rafael M. Martins, Andreas Kerren, Nina S. T. Hirata, and Alexandru C. Telea. 2021. Toward a Quantitative Survey of Dimension Reduction Techniques. *IEEE Transactions on Visualization and Computer Graphics* 27, 3 (2021), 2153–2173. https://doi.org/10.1109/TVCG.2019.2944182

[60] Ronak Etemadpour, Robson Motta, Jose Gustavo de Souza Paiva, Rosane Minghim, Maria Cristina Ferreira de Oliveira, and Lars Linsen. 2015. Perception-Based Evaluation of Projection Methods for Multidimensional Data Visualization. *IEEE Transactions on Visualization and Computer Graphics* 21, 1 (2015), 81–94. https://doi.org/10.1109/TVCG.2014.2330617

[61] Samuel G. Fadel, Francisco M. Fatore, Felipe S.L.G. Duarte, and Fernando V. Paulovich. 2015. LoCH: A neighborhood-based multidimensional projection technique for high-dimensional sparse spaces. *Neurocomputing* 150 (2015), 546–556. https://doi.org/10.1016/j.neucom.2014.07.071

[62] Rebecca Faust, David Glickenstein, and Carlos Scheidegger. 2019. DimReader: Axis lines that explain non-linear projections. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 481–490. https://doi.org/10.1109/TVCG.2018.2865194

[63] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. 1996. From Data Mining to Knowledge Discovery in Databases. *AI Magazine* 17, 3 (Mar. 1996), 37. https://doi.org/10.1609/aimag.v17i3.1230

[64] Sara Johansson Fernstad, Jane Shaw, and Jimmy Johansson. 2013. Quality-based guidance for exploratory dimensionality reduction. *Information Visualization* 12, 1 (2013), 44–64. https://doi.org/10.1177/1473871612460526

[65] Imola K. Fodor. 2002. *A survey of dimension reduction techniques*. Technical Report. Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States). https://computing.llnl.gov/sites/default/files/148494.pdf

[66] Jules Françoise, Baptiste Caramiaux, and Téo Sanchez. 2021. Marcelle: Composing Interactive Machine Learning Workflows and Interfaces. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 39–53. https://doi.org/10.1145/3472749.3474734

[67] Jerome H. Friedman. 1987. Exploratory Projection Pursuit. *J. Amer. Statist. Assoc.* 82, 397 (1987), 249–266. https://doi.org/10.1080/01621459.1987.10478427

[68] Jerome H. Friedman and John W. Tukey. 1974. A Projection Pursuit Algorithm for Exploratory Data Analysis. *IEEE Trans. Comput.* C-23, 9 (1974), 881–890. https://doi.org/10.1109/T-C.1974.224051

[69] Cong Fu, Yonghui Zhang, Deng Cai, and Xiang Ren. 2019. AtSNE: Efficient and Robust Visualization on GPU through Hierarchical Optimization. In *Proc. of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 176–186. https://doi.org/10.1145/3292500.3330834

[70] Takanori Fujiwara, Yun-Hsin Kuo, Anders Ynnerman, and Kwan-Liu Ma. 2023. Feature Learning for Nonlinear Dimensionality Reduction toward Maximal Extraction of Hidden Patterns. In *2023 IEEE 16th Pacific Visualization Symposium (PacificVis)*. 122–131. https://doi.org/10.1109/PacificVis56936.2023.00021

[71] Takanori Fujiwara, Xinhai Wei, Jian Zhao, and Kwan-Liu Ma. 2022. Interactive Dimensionality Reduction for Comparative Analysis. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 758–768. https://doi.org/10.1109/TVCG.2021.3114807

[72] Francisco J. García-Fernández, Michel Verleysen, John A. Lee, and Ignacio Díaz. 2013. Stability Comparison of Dimensionality Reduction Techniques Attending to Data and Parameter Variations. In *EuroVis Workshop on Visual Analytics using Multidimensional Projections*, M. Aupetit and L. van der Maaten (Eds.). The Eurographics Association. https://doi.org//10.2312/PE.VAMP.VAMP2013.005-009

[73] Xin Geng, De-Chuan Zhan, and Zhi-Hua Zhou. 2005. Supervised nonlinear dimensionality reduction for visualization and classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 35, 6 (2005), 1098–1107. https://doi.org/10.1109/TSMCB.2005.850151

[74] Aindrila Ghosh, Mona Nashaat, James Miller, and Shaikh Quader. 2022. VisExPreS: A Visual Interactive Toolkit for User-Driven Evaluations of Embeddings. *IEEE Transactions on Visualization and Computer Graphics* 28, 7 (2022), 2791–2807. https://doi.org/10.1109/TVCG.2020.3039106

[75] Andrej Gisbrecht and Barbara Hammer. 2015. Data visualization by nonlinear dimensionality reduction. *WIREs Data Mining and Knowledge Discovery* 5, 2 (2015), 51–73. https://doi.org/10.1002/widm.1147

[76] Michael Gleicher. 2013. Explainers: Expert Explorations with Crafted Projections. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2042–2051. https://doi.org/10.1109/TVCG.2013.157

[77] Jochen Görtler, Fred Hohman, Dominik Moritz, Kanit Wongsuphasawat, Donghao Ren, Rahul Nair, Marc Kirchner, and Kayur Patel. 2022. Neo: Generalizing Confusion Matrix Visualization to Hierarchical and Multi-Output Labels. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 408, 13 pages. https://doi.org/10.1145/3491102.3501823

[78] Tera Marie Green, William Ribarsky, and Brian Fisher. 2008. Visual analytics for complex concepts using a human cognition model. In *2008 IEEE Symposium on Visual Analytics Science and Technology*. 91–98. https://doi.org/10.1109/VAST.2008.4677361

[79] Tera Marie Green, William Ribarsky, and Brian Fisher. 2009. Building and Applying a Human Cognition Model for Visual Analytics. *Information Visualization* 8, 1 (2009), 1–13. https://doi.org/10.1057/ivs.2008.28

[80] Juhye Ha, Hyeon Jeon, Daeun Han, Jinwook Seo, and Changhoon Oh. 2024. CloChat: Understanding How People Customize, Interact, and Experience Personas in Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 305, 24 pages. https://doi.org/10.1145/3613904.3642472

[81] Laureta Hajderanj, Isakh Weheliye, and Daqing Chen. 2019. A New Supervised t-SNE with Dissimilarity Measure for Effective Data Visualization and Classification. In *Proceedings of the 8th International Conference on Software and Information Engineering* (Cairo, Egypt) *(ICSIE '19)*. Association for Computing Machinery, New York, NY, USA, 232–236. https://doi.org/10.1145/3328833.3328853

[82] Nicolas Heulot, Michael Aupetit, and Jean Daniel Fekete. 2013. ProxiLens: Interactive Exploration of High-Dimensional Data using Projections. In *EuroVis Workshop on Visual Analytics using Multidimensional Projections*. The Eurographics Association. https://doi.org/10.2312/PE.VAMP.VAMP2013.011-015

[83] Md Iqbal Hossain, Vahan Huroyan, Stephen Kobourov, and Raymundo Navarrete. 2021. Multi-Perspective, Simultaneous Embedding. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1569–1579. https://doi.org/10.1109/TVCG.2020.3030373

[84] Sungwon In, Eric Krokos, Kirsten Whitley, Chris North, and Yalong Yang. 2024. Evaluating Navigation and Comparison Performance of Computational Notebooks on Desktop and in Virtual Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 606, 15 pages. https://doi.org/10.1145/3613904.3642932

[85] Hyeon Jeon, Michael Aupetit, Soohyun Lee, Hyung-Kwon Ko, Youngtaek Kim, and Jinwook Seo. 2022. Distortion-Aware Brushing for Interactive Cluster Analysis in Multidimensional Projections. arXiv:2201.06379 [cs.HC]

[86] Hyeon Jeon, Aeri Cho, Jinhwa Jang, Soohyun Lee, Jake Hyun, Hyung-Kwon Ko, Jaemin Jo, and Jinwook Seo. 2023. ZADU: A Python Library for Evaluating the Reliability of Dimensionality Reduction Embeddings. In *2023 IEEE Visualization and Visual Analytics (VIS)*. 196–200. https://doi.org/10.1109/VIS54172.2023.00048

[87] Hyeon Jeon, Hyung-Kwon Ko, Jaemin Jo, Youngtaek Kim, and Jinwook Seo. 2021. Measuring and Explaining the Inter-Cluster Reliability of Multidimensional Projections. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 551–561. https://doi.org/10.1109/TVCG.2021.3114833

[88] Hyeon Jeon, Hyung-Kwon Ko, Soohyun Lee, Jaemin Jo, and Jinwook Seo. 2022. Uniform Manifold Approximation with Two-phase Optimization. In *2022 IEEE Visualization and Visual Analytics (VIS)*. 80–84. https://doi.org/10.1109/VIS54862.2022.00025

[89] Hyeon Jeon, Yun-Hsin Kuo, Michaël Aupetit, Kwan-Liu Ma, and Jinwook Seo. 2024. Classes are Not Clusters: Improving Label-Based Evaluation of Dimensionality Reduction. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 781–791. https://doi.org/10.1109/TVCG.2023.3327187

[90] Hyeon Jeon, Ghulam Jilani Quadri, Hyunwook Lee, Paul Rosen, Danielle Albers Szafir, and Jinwook Seo. 2024. CLAMS: A Cluster Ambiguity Measure for Estimating Perceptual Variability in Visual Clustering. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 770–780. https://doi.org/10.1109/TVCG.2023.3327201

[91] Dong Hyun Jeong, Caroline Ziemkiewicz, Brian Fisher, William Ribarsky, and Remco Chang. 2009. iPCA: An Interactive System for PCA-based Visual Analytics. *Computer Graphics Forum* 28, 3 (2009), 767–774. https://doi.org/10.1111/j.1467-8659.2009.01475.x

[92] Jonathan Johannemann and Robert Tibshirani. 2019. Spectral Overlap and a Comparison of Parameter-Free, Dimensionality Reduction Quality Metrics. arXiv:1907.01974 [stat.ML]

[93] Sara Johansson and Jimmy Johansson. 2009. Interactive Dimensionality Reduction Through User-defined Combinations of Quality Metrics. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 993–1000. https://doi.org/10.1109/TVCG.2009.153

[94] Paulo Joia, Danilo Coimbra, Jose A. Cuminato, Fernando V. Paulovich, and Luis G. Nonato. 2011. Local Affine Multidimensional Projection. *IEEE Transactions on Visualization and Computer Graphics*. 17, 12 (2011), 2563–2571. https://doi.org/10.1109/TVCG.2011.220

[95] Myeongwon Jung, Jiwon Choi, and Jaemin Jo. 2023. Projection Ensemble: Visualizing the Robust Structures of Multidimensional Projections. In *2023 IEEE Visualization and Visual Analytics (VIS)*. 46–50. https://doi.org/10.1109/VIS54172.2023.00018

[96] Dominik Jäckle, Michael Hund, Michael Behrisch, Daniel A. Keim, and Tobias Schreck. 2017. Pattern Trails: Visual Analysis of Pattern Transitions in Subspaces. In *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 1–12. https://doi.org/10.1109/VAST.2017.8585613

[97] Eser Kandogan. 2012. Just-in-time annotation of clusters, outliers, and trends in point-based data visualizations. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 73–82. https://doi.org/10.1109/VAST.2012.6400487

[98] Ashish Kapoor, Bongshin Lee, Desney Tan, and Eric Horvitz. 2010. Interactive optimization for steering machine classification. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1343–1352. https://doi.org/10.1145/1753326.1753529

[99] Samuel Kaski, Jarkko Venna, and Teuvo Kohonen. 2000. Coloring that reveals cluster structures in multivariate data. *Australian Journal of Intelligent Information Processing Systems* 6, 2 (2000), 82–88. https://research.aalto.fi/en/publications/coloring-that-reveals-cluster-structures-in-multivariate-data

[100] Matthew Kay. 2024. ggdist: Visualizations of Distributions and Uncertainty in the Grammar of Graphics. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 414–424. https://doi.org/10.1109/TVCG.2023.3327195

[101] Daniel Keim, Gennady Andrienko, Jean-Daniel Fekete, Carsten Görg, Jörn Kohlhammer, and Guy Melançon. 2008. *Visual Analytics: Definition, Process, and Challenges*. Springer Berlin Heidelberg, Berlin, Heidelberg, 154–175. https://doi.org/10.1007/978-3-540-70956-5_7

[102] Hannah Kim, Jaegul Choo, Changhyun Lee, Hanseung Lee, Chandan Reddy, and Haesun Park. 2017. PIVE: Per-Iteration Visualization Environment for Real-Time Interactions with Dimension Reduction and Clustering. *Proceedings of the AAAI Conference on Artificial Intelligence* 31, 1 (Feb. 2017). https://doi.org/10.1609/aaai.v31i1.10628

[103] Hannah Kim, Jaegul Choo, Haesun Park, and Alex Endert. 2016. InterAxis: Steering Scatterplot Axes via Observation-Level Interaction. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 131–140. https://doi.org/10.1109/TVCG.2015.2467615

[104] Yea-Seul Kim, Paula Kayongo, Madeleine Grunde-McLaughlin, and Jessica Hullman. 2021. Bayesian-Assisted Inference from Visualized Data. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 989–999. https://doi.org/10.1109/TVCG.2020.3028984

[105] Yea-Seul Kim, Logan A. Walls, Peter Krafft, and Jessica Hullman. 2019. A Bayesian Cognition Approach to Improve Data Visualization. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300912

[106] Dmitry Kobak and George C. Linderman. 2021. Initialization is critical for preserving global data structure in both t-SNE and UMAP. *Nature Biotechnology* 39, 2 (01 Feb 2021), 156–157. https://doi.org/10.1038/s41587-020-00809-z

[107] Yehuda Koren and Liran Carmel. 2003. Visualization of labeled data using linear transformations. In *IEEE Symposium on Information Visualization 2003*. 121–128. https://doi.org/10.1109/INFVIS.2003.1249017

[108] Hans-Peter Kriegel, Peer Kröger, Jörg Sander, and Arthur Zimek. 2011. Density-based clustering. *WIREs Data Mining and Knowledge Discovery* 1, 3 (2011), 231–240. https://doi.org/10.1002/widm.30

[109] Johannes F. Kruiger, Paulo E. Rauber, Rafael Messias Martins, Andreas Kerren, Stephen Kobourov, and Alexandru C. Telea. 2017. Graph Layouts by t-SNE. *Computer Graphics Forum* 36, 3 (2017), 283–294. https://doi.org/10.1111/cgf.13187

[110] Joseph Bernard Kruskal. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29 (1964), 1–27. https://doi.org/10.1007/BF02289565

[111] Bum Chul Kwon, Ben Eysenbach, Janu Verma, Kenney Ng, Christopher De Filippi, Walter F. Stewart, and Adam Perer. 2018. Clustervision: Visual Supervision of Unsupervised Clustering. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 142–151. https://doi.org/10.1109/TVCG.2017.2745085

[112] Bum Chul Kwon, Hannah Kim, Emily Wall, Jaegul Choo, Haesun Park, and Alex Endert. 2017. AxiSketcher: Interactive Nonlinear Axis Mapping of Visualizations through User Drawings. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 221–230. https://doi.org/10.1109/TVCG.2016.2598446

[113] Chufan Lai, Ying Zhao, and Xiaoru Yuan. 2018. Exploring high-dimensional data through locally enhanced projections. *Journal of Visual Languages & Computing* 48 (2018), 144–156. https://doi.org/10.1016/j.jvlc.2018.08.006

[114] Gregor Leban, Blaž Zupan, Gaj Vidmar, and Ivan Bratko. 2006. Vizrank: Data visualization guided by machine learning. *Data Mining and Knowledge Discovery* 13 (2006), 119–136. https://doi.org/10.1007/s10618-005-0031-5

[115] Benjamin Lee, Arvind Satyanarayan, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. 2023. Deimos: A Grammar of Dynamic Embodied Immersive Visualisation Morphs and Transitions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 810, 18 pages. https://doi.org/10.1145/3544548.3580754

[116] John A. Lee and Michel Verleysen. 2009. Quality assessment of dimensionality reduction: Rank-based criteria. *Neurocomputing* 72, 7 (2009), 1431–1443. https://doi.org/10.1016/j.neucom.2008.12.017 Advances in Machine Learning and Computational Intelligence.

[117] John A. Lee and Michel Verleysen. 2010. Scale-independent quality criteria for dimensionality reduction. *Pattern Recognition Letters* 31, 14 (2010), 2248–2257. https://doi.org/10.1016/j.patrec.2010.04.013

[118] John A. Lee and Michel Verleysen. 2011. Shift-invariant similarities circumvent distance concentration in stochastic neighbor embedding and variants. *Procedia Computer Science* 4 (2011), 538–547. https://doi.org/10.1016/j.procs.2011.04.056 Proceedings of the International Conference on Computational Science, ICCS 2011.

[119] Jenny Hyunjung Lee, Kevin T. McDonnell, Alla Zelenyuk, Dan Imre, and Klaus Mueller. 2014. A Structure-Based Distance Metric for High-Dimensional Space Exploration with Multidimensional Scaling. *IEEE Transactions on Visualization and Computer Graphics* 20, 3 (2014), 351–364. https://doi.org/10.1109/TVCG.2013.101

[120] Dirk J. Lehmann and Holger Theisel. 2013. Orthographic Star Coordinates. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2615–2624. https://doi.org/10.1109/TVCG.2013.182

[121] Sylvain Lespinats and Michaël Aupetit. 2011. CheckViz: Sanity Check and Topological Clues for Linear and Non-Linear Mappings. *Computer Graphics Forum* 30, 1 (2011), 113–125. https://doi.org/10.1111/j.1467-8659.2010.01835.x

[122] Sylvain Lespinats, Michel Verleysen, Alain Giron, and Bernard Fertil. 2007. DD-HDS: A Method for Visualization and Exploration of High-Dimensional Data. *IEEE Transactions on Neural Networks* 18, 5 (2007), 1265–1279. https://doi.org/10.1109/TNN.2007.891682

[123] Joshua Lewis, Margareta Ackerman, and Virginia de Sa. 2012. Human cluster evaluation and formal quality measures: A comparative study. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 34. https://escholarship.org/uc/item/8rr3t8dx

[124] Jiwei Li, Xinlei Chen, Eduard Hovy, and Dan Jurafsky. 2016. Visualizing and Understanding Neural Models in NLP. arXiv:1506.01066 [cs.CL]

[125] Jie Li and Chun-qi Zhou. 2023. Incorporation of Human Knowledge into Data Embeddings to Improve Pattern Significance and Interpretability. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2023), 723–733. https://doi.org/10.1109/TVCG.2022.3209382

[126] Chungman Lim and Gunhyuk Park. 2023. Can a Computer Tell Differences between Vibrations?: Physiology-Based Computational Model for Perceptual Dissimilarity Prediction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 774, 15 pages. https://doi.org/10.1145/3544548.3580686

[127] S. Liu, P.-T Bremer, J. J. Jayaraman, B. Wang, B. Summa, and V. Pascucci. [n. d.]. The Grassmannian Atlas: A General Framework for Exploring Linear Projections of High-Dimensional Data. *Computer Graphics Forum* 35, 3 ([n. d.]), 1–10. https://doi.org/10.1111/cgf.12876

[128] Shusen Liu, Peer-Timo Bremer, Jayaraman J. Thiagarajan, Vivek Srikumar, Bei Wang, Yarden Livnat, and Valerio Pascucci. 2018. Visual Exploration of Semantic Relationships in Neural Word Embeddings. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 553–562. https://doi.org/10.1109/TVCG.2017.2745141

[129] S. Liu, B. Wang, P.-T. Bremer, and V. Pascucci. 2014. Distortion-Guided Structure-Driven Interactive Exploration of High-Dimensional Data. *Computer Graphics Forum* 33, 3 (2014), 101–110. https://doi.org/10.1111/cgf.12366

[130] S. Liu, B. Wang, J. J. Thiagarajan, P.-T. Bremer, and V. Pascucci. 2015. Visual Exploration of High-Dimensional Data through Subspace Analysis and Dynamic Projections. *Computer Graphics Forum* 34, 3 (2015), 271–280. https://doi.org/10.1111/cgf.12639

[131] Yuxin Ma and Ross Maciejewski. 2021. Visual Analysis of Class Separations With Locally Linear Segments. *IEEE Transactions on Visualization and Computer Graphics* 27, 1 (2021), 241–253. https://doi.org/10.1109/TVCG.2020.3011155

[132] Gladys M. H. Mamani, Francisco M. Fatore, Luis Gustavo Nonato, and Fernando V. Paulovich. 2013. User-driven Feature Space Transformation. *Computer Graphics Forum* 32, 3pt3 (2013), 291–299. https://doi.org/10.1111/cgf.12116

[133] Wilson E. Marcílio-Jr and Danilo M. Eler. 2021. Explaining dimensionality reduction results using Shapley values. *Expert Systems with Applications* 178 (2021), 115020. https://doi.org/10.1016/j.eswa.2021.115020

[134] Rafael Messias Martins, Danilo Barbosa Coimbra, Rosane Minghim, and A.C. Telea. 2014. Visual analysis of dimensionality reduction quality for parameterized projections. *Computers & Graphics* 41 (2014), 26–42. https://doi.org/10.1016/j.cag.2014.01.006

[135] Rafael Messias Martins, Rosane Minghim, and AC Telea. 2015. Explaining Neighborhood Preservation for Multidimensional Projections. In *Computer Graphics and Visual Computing*. Eurographics-European Association for Computer Graphics, 7–14. https://doi.org/10.2312/cgvc.20151234

[136] Jose Matute and Lars Linsen. 2020. Hinted Star Coordinates for Mixed Data. *Computer Graphics Forum* 39, 1 (2020), 117–133. https://doi.org/10.1111/cgf.13666

[137] Leland McInnes, John Healy, and James Melville. 2020. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426 [stat.ML]

[138] Linhao Meng, Stef van den Elzen, Nicola Pezzotti, and Anna Vilanova. 2024. Class-Constrained t-SNE: Combining Data Features and Class Probabilities. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 164–174. https://doi.org/10.1109/TVCG.2023.3326600

[139] Jacob Miller, Dhruv Bhatia, and Stephen Kobourov. 2024. State of the Art of Graph Visualization in non-Euclidean Spaces. *Computer Graphics Forum* 43, 3 (2024), e15113. https://doi.org/10.1111/cgf.15113

[140] Bassam Mokbel, Wouter Lueks, Andrej Gisbrecht, and Barbara Hammer. 2013. Visualizing the quality of dimensionality reduction. *Neurocomputing* 112 (2013), 109–123. https://doi.org/10.1016/j.neucom.2012.11.046

[141] Vladimir Molchanov, Sagad Hamid, and Lars Linsen. 2020. Efficient Morphing of Shape-preserving Star Coordinates. In *2020 IEEE Pacific Visualization Symposium (PacificVis)*. 136–145. https://doi.org/10.1109/PacificVis48177.2020.8199

[142] Vladimir Molchanov and Lars Linsen. 2019. Shape-preserving Star Coordinates. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 449–458. https://doi.org/10.1109/TVCG.2018.2865118

[143] Kevin R Moon, David van Dijk, Zheng Wang, Scott Gigante, Daniel B Burkhardt, et al. 2019. Visualizing structure and transitions in high-dimensional biological data. *Nature Biotechnolgy* 37, 12 (2019), 1482–1492. https://doi.org/10.1038/s41587-019-0336-3

[144] Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt. 2020. Topological Autoencoders. In *Proceedings of the 37th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 119)*, Hal Daumé III and Aarti Singh (Eds.). PMLR, 7045–7054. https://proceedings.mlr.press/v119/moor20a.html

[145] Cristina Morariu, Adrien Bibal, Rene Cutura, Benoît Frénay, and Michael Sedlmair. 2023. Predicting User Preferences of Dimensionality Reduction Embedding Quality. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2023), 745–755. https://doi.org/10.1109/TVCG.2022.3209449

[146] Dominik Moritz, Chenglong Wang, Greg L. Nelson, Halden Lin, Adam M. Smith, Bill Howe, and Jeffrey Heer. 2019. Formalizing Visualization Design Knowledge as Constraints: Actionable and Extensible Models in Draco. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 438–448. https://doi.org/10.1109/TVCG.2018.2865240

[147] Robson Motta, Rosane Minghim, Alneu de Andrade Lopes, and Maria Cristina F. Oliveira. 2015. Graph-based measures to assist user assessment of multidimensional projections. *Neurocomputing* 150 (2015), 583–598. https://doi.org/10.1016/j.neucom.2014.09.063

[148] Fionn Murtagh and Pedro Contreras. 2012. Algorithms for hierarchical clustering: an overview. *WIREs Data Mining and Knowledge Discovery* 2, 1 (2012), 86–97. https://doi.org/10.1002/widm.53

[149] Julia EunJu Nam and Klaus Mueller. 2013. TripAdvisor^{N-D}: A Tourism-Inspired High-Dimensional Space Exploration Framework with Overview and Detail. *IEEE Transactions on Visualization and Computer Graphics* 19, 2 (2013), 291–305. https://doi.org/10.1109/TVCG.2012.65

[150] Ashwin Narayan, Bonnie Berger, and Hyunghoon Cho. 2021. Assessing single-cell transcriptomic variability through density-preserving data visualization. *Nature Biotechnology* 39, 6 (2021), 765–774. https://doi.org/10.1038/s41587-020-00801-7

[151] Arpit Narechania, Alireza Karduni, Ryan Wesslen, and Emily Wall. 2022. VITALITY: Promoting Serendipitous Discovery of Academic Literature with Transformers & Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 486–496. https://doi.org/10.1109/TVCG.2021.3114820

[152] Quan Nguyen, Peter Eades, and Seok-Hee Hong. 2013. On the faithfulness of graph visualizations. In *2013 IEEE Pacific Visualization Symposium (PacificVis)*. 209–216. https://doi.org/10.1109/PacificVis.2013.6596147

[153] Fernando Nogueira. 2014–. Bayesian Optimization: Open source constrained global optimization tool for Python. https://github.com/bayesian-optimization/BayesianOptimization

[154] Luis Gustavo Nonato and Michaël Aupetit. 2019. Multidimensional Projection for Visual Analytics: Linking Techniques with Distortions, Tasks, and Layout Enrichment. *IEEE Transactions on Visualization and Computer Graphics* 25, 8 (2019), 2650–2673. https://doi.org/10.1109/TVCG.2018.2846735

[155] Lenka Nováková and Olga Štepanková. 2009. RadViz and Identification of Clusters in Multidimensional Data. In *2009 13th International Conference Information Visualisation*. 104–109. https://doi.org/10.1109/IV.2009.103

[156] Michael Oppermann, Robert Kincaid, and Tamara Munzner. 2021. VizCommender: Computing Text-Based Similarity in Visualization Repositories for Content-Based Recommendations. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 495–505. https://doi.org/10.1109/TVCG.2020.3030387

[157] Lucas Pagliosa, Paulo Pagliosa, and Luis Gustavo Nonato. 2016. Understanding Attribute Variability in Multidimensional Projections. In *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. 297–304. https://doi.org/10.1109/SIBGRAPI.2016.048

[158] Lucas de Carvalho Pagliosa and Alexandru C. Telea. 2019. RadViz++: Improvements on Radial-Based Visualizations. *Informatics* 6, 2 (2019). https://doi.org/10.3390/informatics6020016

[159] Stephen E. Palmer. 1999. *Vision science : photons to phenomenology*. MIT Press, Cambridge, Mass. 810 pages. https://doi.org/10.1086/420636

[160] Anshul Vikram Pandey, Josua Krause, Cristian Felix, Jeremy Boy, and Enrico Bertini. 2016. Towards Understanding Human Similarity Perception in the Analysis of Large Sets of Scatter Plots *(CHI '16)*. ACM, New York, NY, USA, 3659–3669. https://doi.org/10.1145/2858036.2858155

[161] Fernando V. Paulovich, Danilo M. Eler, Jorge Poco, Charl Botha, Rosane Minghim, and Luis Gustavo Nonato. 2011. Piece wise Laplacian-based Projection for Interactive Data Exploration and Organization. *Computer Graphics Forum* 30, 3 (2011), 1091–1100. https://doi.org/10.1111/j.1467-8659.2011.01958.x

[162] Fernando V. Paulovich, Luis G. Nonato, Rosane Minghim, and Haim Levkowitz. 2008. Least Square Projection: A Fast High-Precision Multidimensional Projection Technique and Its Application to Document Mapping. *IEEE Transactions on Visualization and Computer Graphics* 14, 3 (2008), 564–575. https://doi.org/10.1109/TVCG.2007.70443

[163] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* 12 (2011), 2825–2830. https://www.jmlr.org/papers/v12/pedregosa11a.html

[164] Ghulam Jilani Quadri and Paul Rosen. 2019. You Can't Publish Replication Studies (and How to Anyways). arXiv:1908.08893 [cs.HC]

[165] Ghulam Jilani Quadri and Paul Rosen. 2022. A Survey of Perception-Based Visualization Studies by Task. *IEEE Transactions on Visualization and Computer Graphics* 28, 12 (2022), 5026–5048. https://doi.org/10.1109/TVCG.2021.3098240

[166] Paulo E. Rauber, Samuel G. Fadel, Alexandre X. Falcão, and Alexandru C. Telea. 2017. Visualizing the Hidden Activity of Artificial Neural Networks. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 101–110. https://doi.org/10.1109/TVCG.2016.2598838

[167] Bastian Rieck and Heike Leitte. 2015. Persistent Homology for the Evaluation of Dimensionality Reduction Schemes. *Computer Graphics Forum* 34, 3 (2015), 431–440. https://doi.org/10.1111/cgf.12655

[168] Bastian Rieck and Heike Leitte. 2017. Agreement Analysis of Quality Measures for Dimensionality Reduction. In *Topological Methods in Data Analysis and Visualization IV*. Springer International Publishing, Cham, 103–117. https://doi.org/10.1007/978-3-319-44684-4_6

[169] Greg Ross, Alistair Morrison, and Matthew Chalmers. 2004. Coordinating views for data visualisation and algorithmic profiling. In *Proceedings. Second International Conference on Coordinated and Multiple Views in Exploratory Visualization, 2004*. 3–14. https://doi.org/10.1109/CMV.2004.1319522

[170] Peter J. Rousseeuw. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20 (1987), 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

[171] Sam T. Roweis and Lawrence K. Saul. 2000. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science* 290, 5500 (2000), 2323–2326. https://doi.org/10.1126/science.290.5500.2323

[172] Manuel Rubio-Sánchez, Laura Raya, Francisco Díaz, and Alberto Sanchez. 2016. A comparative study between RadViz and Star Coordinates. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 619–628. https://doi.org/10.1109/TVCG.2015.2467324

[173] Manuel Rubio-Sánchez and Alberto Sanchez. 2014. Axis Calibration for Improving Data Attribute Estimation in Star Coordinates Plots. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2013–2022. https://doi.org/10.1109/TVCG.2014.2346258

[174] Dominik Sacha, Andreas Stoffel, Florian Stoffel, Bum Chul Kwon, Geoffrey Ellis, and Daniel A. Keim. 2014. Knowledge Generation Model for Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 1604–1613. https://doi.org/10.1109/TVCG.2014.2346481

[175] Dominik Sacha, Leishi Zhang, Michael Sedlmair, John A. Lee, Jaakko Peltonen, Daniel Weiskopf, Stephen C. North, and Daniel A. Keim. 2017. Visual Interaction with Dimensionality Reduction: A Structured Literature Analysis. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 241–250. https://doi.org/10.1109/TVCG.2016.2598495

[176] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. 2017. Vega-Lite: A Grammar of Interactive Graphics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 341–350. https://doi.org/10.1109/TVCG.2016.2599030

[177] Tobias Schreck, Tatiana von Landesberger, and Sebastian Bremm. 2010. Techniques for Precision-Based Visual Analysis of Projected Data. *Information Visualization* 9, 3 (2010), 181–193. https://doi.org/10.1057/ivs.2010.2

[178] Michael Sedlmair, Tamara Munzner, and Melanie Tory. 2013. Empirical Guidance on Scatterplot and Dimension Reduction Technique Choices. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2634–2643. https://doi.org/10.1109/TVCG.2013.153

[179] M. Sedlmair, A. Tatu, T. Munzner, and M. Tory. 2012. A Taxonomy of Visual Cluster Separation Factors. *Computer Graphics Forum* 31, 3pt4 (2012), 1335–1344. https://doi.org/10.1111/j.1467-8659.2012.03125.x

[180] Christin Seifert, Vedran Sabol, and Wolfgang Kienreich. 2010. Stress Maps: Analysing Local Phenomena in Dimensionality Reduction Based Visualisations.. In *EuroVAST @ EuroVis*. The Eurographics Association. https://doi.org/10.2312/pe/eurovast/eurovast10/013-018

[181] Jessica Zeitz Self, Michelle Dowling, John E. Wenskovitch, Ian C Crandell, Ming Wang, Leanna House, Scotland Leman, and Chris North. 2018. Observation-Level and Parametric Interaction for High-Dimensional Data Analysis. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 8 (2018), 1 – 36. https://doi.org/10.1145/3158230

[182] Woosuk Seo, Chanmo Yang, and Young-Ho Kim. 2024. ChaCha: Leveraging Large Language Models to Prompt Children to Share Their Emotions about Personal Events. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 903, 20 pages. https://doi.org/10.1145/3613904.3642152

[183] Sungbok Shin, Andrea Batch, Peter W. S. Butcher, Panagiotis D. Ritsos, and Niklas Elmqvist. 2023. The Reality of the Situation: A Survey of Situated Analytics. *IEEE Transactions on Visualization and Computer Graphics* (2023), 1–19. https://doi.org/10.1109/TVCG.2023.3285546

[184] Renato R. O. da Silva, Paulo E. Rauber, Rafael M. Martins, Rosane Minghim, and Alexandru C. Telea. 2015. Attribute-based Visual Explanation of Multidimensional Projections. In *EuroVis Workshop on Visual Analytics (EuroVA)*, E. Bertini and J. C. Roberts (Eds.). The Eurographics Association. https://doi.org/10.2312/eurova.20151100

[185] Vin Silva and Joshua Tenenbaum. 2002. Global Versus Local Methods in Nonlinear Dimensionality Reduction. In *Advances in Neural Information Processing Systems*, Vol. 15. MIT Press. https://proceedings.neurips.cc/paper_files/paper/2002/file/5d6646aad9bcc0be55b2c82f69750387-Paper.pdf

[186] Mike Sips, Boris Neubert, John P. Lewis, and Pat Hanrahan. 2009. Selecting good views of high-dimensional data using class consistency. *Computer Graphics Forum* 28, 3 (2009), 831–838. https://doi.org/10.1111/j.1467-8659.2009.01467.x

[187] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. 2012. Practical Bayesian Optimization of Machine Learning Algorithms. In *Advances in Neural Information Processing Systems*, F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger (Eds.), Vol. 25. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf

[188] Jan-Tobias Sohns, Michaela Schmitt, Fabian Jirasek, Hans Hasse, and Heike Leitte. 2022. Attribute-based Explanation of Non-Linear Embeddings of High-Dimensional Data. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 540–550. https://doi.org/10.1109/TVCG.2021.3114870

[189] Carlos Oscar S. Sorzano, Javier Vargas, and A. Pascual Montano. 2014. A survey of dimensionality reduction techniques. arXiv:1403.2877 [stat.ML]

[190] Dimitris Spathis, Nikolaos Passalis, and Anastasios Tefas. 2018. Fast, Visual and Interactive Semi-supervised Dimensionality Reduction. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. https://doi.org/10.1007/978-3-030-11018-5_44

[191] Julian Stahnke, Marian Dörk, Boris Müller, and Andreas Thom. 2016. Probing Projections: Interaction Techniques for Interpreting Arrangements and Errors of Dimensionality Reductions. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 629–638. https://doi.org/10.1109/TVCG.2015.2467717

[192] Guodao Sun, Sujia Zhu, Qi Jiang, Wang Xia, and Ronghua Liang. 2022. EvoSets: Tracking the Sensitivity of Dimensionality Reduction Results Across Subspaces. *IEEE Transactions on Big Data* 8, 6 (2022), 1566–1579. https://doi.org/10.1109/TBDATA.2021.3079200

[193] Andrada Tatu, Peter Bak, Enrico Bertini, Daniel Keim, and Joern Schneidewind. 2010. Visual quality metrics and human perception: an initial study on 2D projections of large multidimensional data. In *Proceedings of the International Conference on Advanced Visual Interfaces* (Roma, Italy) *(AVI '10)*. Association for Computing Machinery, New York, NY, USA, 49–56. https://doi.org/10.1145/1842993.1843002

[194] Adly Templeton, Tom Conerly, Jonathan Marcus, Jack Lindsey, Trenton Bricken, Brian Chen, Adam Pearce, Craig Citro, Emmanuel Ameisen, Andy Jones, Hoagy Cunningham, Nicholas L Turner, Callum McDougall, Monte MacDiarmid, C. Daniel Freeman, Theodore R. Sumers, Edward Rees, Joshua Batson, Adam

Jermyn, Shan Carter, Chris Olah, and Tom Henighan. 2024. Scaling Monose-manticity: Extracting Interpretable Features from Claude 3 Sonnet. *Transformer Circuits Thread* (2024). https://transformer-circuits.pub/2024/scaling-monosemanticity/index.html

[195] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. 2000. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* 290, 5500 (2000), 2319–2323. https://doi.org/10.1126/science.290.5500.2319

[196] Michael C Thrun, Julian Märte, and Quirin Stier. 2023. Analyzing Quality Measurements for Dimensionality Reduction. *Machine Learning and Knowledge Extraction* 5, 3 (2023), 1076–1118. https://doi.org/10.3390/make5030056

[197] Zonglin Tian, Xiaorui Zhai, Daan van Driel, Gijs van Steenpaal, Mateus Espadoto, and Alexandru Telea. 2021. Using multiple attribute-based explanations of multidimensional projections to explore high-dimensional data. *Computers & Graphics* 98 (2021), 93–104. https://doi.org/10.1016/j.cag.2021.04.034

[198] Cagatay Turkay, Peter Filzmoser, and Helwig Hauser. 2011. Brushing Dimensions - A Dual Visual Analysis Model for High-Dimensional Data. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2591–2599. https://doi.org/10.1109/TVCG.2011.178

[199] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9, 86 (2008), 2579–2605. https://jmlr.org/papers/v9/vandermaaten08a.html

[200] Laurens Van Der Maaten, Eric O Postma, H Jaap van den Herik, et al. 2009. Dimensionality reduction: A comparative review. *Journal of Machine Learning Research* 10, 66-71 (2009), 13. https://lvdmaaten.github.io/publications/papers/TR_Dimensionality_Reduction_Review_2009.pdf

[201] Jarke J. van Wijk. 2006. Views on Visualization. *IEEE Transactions on Visualization and Computer Graphics* 12, 4 (2006), 421–432. https://doi.org/10.1109/TVCG.2006.80

[202] Jarkko Venna and Samuel Kaski. 2006. Local multidimensional scaling. *Neural Networks* 19, 6 (2006), 889–899. https://doi.org/10.1016/j.neunet.2006.05.014

[203] Jarkko Venna, Jaakko Peltonen, Kristian Nybo, Helena Aidos, and Samuel Kaski. 2010. Information Retrieval Perspective to Nonlinear Dimensionality Reduction for Data Visualization. *Journal of Machine Learning Research* 11, 13 (2010), 451–490. https://doi.org/10.5555/1756006.1756019

[204] Elio Ventocilla and Maria Riveiro. 2020. A comparative user study of visualization techniques for cluster analysis of multidimensional data sets. *Information Visualization* 19, 4 (2020), 318–338. https://doi.org/10.1177/1473871620922166

[205] Viet Minh Vu, Adrien Bibal, and Benoît Frénay. 2021. HCt-SNE: Hierarchical Constraints with t-SNE. In *2021 International Joint Conference on Neural Networks (IJCNN)*. 1–8. https://doi.org/10.1109/IJCNN52387.2021.9534470

[206] Bing Wang and Klaus Mueller. 2018. The Subspace Voyager: Exploring High-Dimensional Data along a Continuum of Salient 3D Subspaces. *IEEE Transactions on Visualization and Computer Graphics* 24, 2 (2018), 1204–1222. https://doi.org/10.1109/TVCG.2017.2672987

[207] Junpeng Wang, Shixia Liu, and Wei Zhang. 2024. Visual Analytics for Machine Learning: A Data Perspective Survey. *IEEE Transactions on Visualization and Computer Graphics* (2024), 1–20. https://doi.org/10.1109/TVCG.2024.3357065

[208] Qianwen Wang, Kexin Huang, Payal Chandak, Marinka Zitnik, and Nils Gehlenborg. 2023. Extending the Nested Model for User-Centric XAI: A Design Study on GNN-based Drug Repurposing. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2023), 1266–1276. https://doi.org/10.1109/TVCG.2022.3209435

[209] Yunhai Wang, Kang Feng, Xiaowei Chu, Jian Zhang, Chi-Wing Fu, Michael Sedlmair, Xiaohui Yu, and Baoquan Chen. 2018. A Perception-Driven Approach to Supervised Dimensionality Reduction for Visualization. *IEEE Transactions on Visualization and Computer Graphics* 24, 5 (2018), 1828–1840. https://doi.org/10.1109/TVCG.2017.2701829

[210] Yunhai Wang, Jingting Li, Feiping Nie, Holger Theisel, Minglun Gong, and Dirk J. Lehmann. 2017. Linear Discriminative Star Coordinates for Exploring Class and Cluster Separation of High Dimensional Data. *Computer Graphics Forum* 36, 3 (2017), 401–410. https://doi.org/10.1111/cgf.13197

[211] Martin Wattenberg, Fernanda Viégas, and Ian Johnson. 2016. How to Use t-SNE Effectively. *Distill* (2016). https://doi.org/10.23915/distill.00002

[212] John Wenskovitch, Ian Crandell, Naren Ramakrishnan, Leanna House, Scotland Leman, and Chris North. 2018. Towards a Systematic Combination of Dimension Reduction and Clustering in Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 131–141. https://doi.org/10.1109/TVCG.2017.2745258

[213] John Wenskovitch, Michelle Dowling, and Chris North. 2020. With Respect to What? Simultaneous Interaction with Dimension Reduction and Clustering Projections. In *Proc. of the 25th International Conference on Intelligent User Interfaces*. 177–188. https://doi.org/10.1145/3377325.3377516

[214] Jiazhi Xia, Linquan Huang, Weixing Lin, Xin Zhao, Jing Wu, Yang Chen, Ying Zhao, and Wei Chen. 2022. Interactive Visual Cluster Analysis by Contrastive Dimensionality Reduction. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–11. https://doi.org/10.1109/TVCG.2022.3209423

[215] Jiazhi Xia, Fenjin Ye, Wei Chen, Yusi Wang, Weifeng Chen, Yuxin Ma, and Anthony K.H. Tung. 2018. LDSScanner: Exploratory Analysis of Low-Dimensional Structures in High-Dimensional Datasets. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 236–245. https://doi.org/10.1109/TVCG.2017.2744098

[216] Jiazhi Xia, Yuchen Zhang, Jie Song, Yang Chen, Yunhai Wang, and Shixia Liu. 2021. Revisiting Dimensionality Reduction Techniques for Visual Cluster Analysis: An Empirical Study. *IEEE Transactions on Visualization and Computer Graphics* (2021), 1–1. https://doi.org/10.1109/TVCG.2021.3114694

[217] Ruizhi Xiang, Wencan Wang, Lei Yang, Shiyuan Wang, Chaohan Xu, and Xiaowen Chen. 2021. A Comparison for Dimensionality Reduction Methods of Single-Cell RNA-seq Data. *Frontiers in Genetics* 12 (2021). https://doi.org/10.3389/fgene.2021.646936

[218] Haozhe Xie, Jie Li, and Hanqing Xue. 2018. A survey of dimensionality reduction techniques based on random projection. arXiv:1706.04371 [cs.LG]

[219] Guangjing Yan, Jinhwa Jang, and Jinwook Seo. 2024. CommentVis: Unveiling Comment Insights Through Interactive Visualization Tool. In *2024 IEEE 17th Pacific Visualization Conference (PacificVis)*. 353–358. https://doi.org/10.1109/PacificVis60374.2024.00050

[220] Fumeng Yang, James Tompkin, Lane Harrison, and David H. Laidlaw. 2023. Visual Cue Effects on a Classification Accuracy Estimation Task in Immersive Scatterplots. *IEEE Transactions on Visualization and Computer Graphics* 29, 12 (2023), 4858–4873. https://doi.org/10.1109/TVCG.2022.3192364

[221] Jing Yang, Matthew O. Ward, and Elke A Rundensteiner. 2003. Visual hierarchical dimension reduction for exploration of high dimensional datasets. In *Proceedings of the Symposium on Data Visualisation 2003 (VISSYM '03)*. Eurographics Association, Goslar, DEU, 19–28. https://dl.acm.org/doi/abs/10.5555/769922.769924

[222] Catherine Yeh, Yida Chen, Aoyu Wu, Cynthia Chen, Fernanda Viégas, and Martin Wattenberg. 2024. AttentionViz: A Global View of Transformer Attention. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 262–272. https://doi.org/10.1109/TVCG.2023.3327163

[223] Ji Soo Yi, Rachel Melton, John Stasko, and Julie A. Jacko. 2005. Dust & Magnet: Multivariate Information Visualization Using a Magnet Metaphor. *Information Visualization* 4, 4 (Oct. 2005), 239–256. https://doi.org/10.1057/palgrave.ivs.9500099

[224] Hujun Yin. 2007. Nonlinear dimensionality reduction and data visualization: a review. *International Journal of Automation and Computing* 4 (2007), 294–303. https://doi.org/10.1007/s11633-007-0294-y

[225] Yang Zhang, Jisheng Liu, Chufan Lai, Yuan Zhou, and Siming Chen. 2024. Interpreting High-Dimensional Projections With Capacity. *IEEE Transactions on Visualization and Computer Graphics* 30, 9 (2024), 6038–6055. https://doi.org/10.1109/TVCG.2023.3324851

[226] Fangfang Zhou, Wei Huang, Juncai Li, Yezi Huang, Yang Shi, and Ying Zhao. 2015. Extending Dimensions in Radviz based on mean shift. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*. 111–115. https://doi.org/10.1109/PACIFICVIS.2015.7156365

[227] Fangfang Zhou, Juncai Li, Wei Huang, Ying Zhao, Xiaoru Yuan, Xing Liang, and Yang Shi. 2016. Dimension reconstruction for visual exploration of subspace clusters in high-dimensional data. In *2016 IEEE Pacific Visualization Symposium (PacificVis)*. 128–135. https://doi.org/10.1109/PACIFICVIS.2016.7465260