Cronfa - Swansea University Open Access Repository

_____

This is an author produced version of a paper published in :
Computer Graphics Forum

Cronfa URL for this paper:

_____

Paper:

_____

Survey of Surveys (SoS) - Mapping The Landscape of Survey Papers in Information Visualization

Liam McNabb[1] and Robert S. Laramee[1]

[1]Department of Computer Science, Swansea University, Wales

Abstract

Information visualization as a field is growing rapidly in popularity since the first information visualization conference in 1995. However, as a consequence of its growth, it is increasingly difficult to follow the growing body of literature within the field. Survey papers and literature reviews are valuable tools for managing the great volume of previously published research papers, and the quantity of survey papers in visualization has reached a critical mass. To this end, this survey paper takes a quantum step forward by surveying and classifying literature survey papers in order to help researchers understand the current landscape of Information Visualization. It is, to our knowledge, the first survey of survey papers (SoS) in Information Visualization. This paper classifies survey papers into natural topic clusters which enables readers to find relevant literature and develops the first classification of classifications. The paper also enables researchers to identify both mature and less developed research directions as well as identify future directions. It is a valuable resource for both newcomers and experienced researchers in and outside the field of Information Visualization and Visual Analytics.

Categories and Subject Descriptors (according to ACM CCS): [Human-centered computing]: Visualization—Information visualization

1. Introduction and Motivation

"It used to be that SIGGRAPH was the only place that would publish computer graphics papers, and so all you had to do was read the SIGGRAPH Conference Proceedings and you knew you were up to date. But nowadays there's lots of other journals and it takes more and more effort to make sure that you know what's happening." This quote is taken from Jim Blinn's renowned keynote speech at SIGGRAPH 98 [Bli98]. Decades later this theme is still considered one of the most important challenges for research in any field.

Information Visualization is a rapidly evolving research field defined as "the communication of abstract data through the use of interactive visual interfaces" [KMSZ06]. Because of this, many researchers spend countless hours on research and development of Information Visualization techniques only to discover that research on a given topic has already been published. Survey papers and literature reviews are valuable and critical tools for managing the great number of previously published research papers. However even the number of survey papers themselves has reached a critical mass, thus inspiring a quantum step forward.

In this first undertaking of a survey of surveys (SoS), we aim to present the landscape of rapidly evolving research within Information Visualization. In order to emphasize open directions for future work, we present literature reviews of papers that survey research

topics, and extract essential information from them systematically as a guide for both newcomers and experts in the field and beyond. We then classify over 80 survey papers to examine trends and themes that have recently been published. Our contributions to the field include:

- A quantum step in literature review papers presenting the first meta-survey, i.e. a 'Survey of Surveys' (SoS).
- A novel classification of survey literature in the field of Information Visualization, which can be used as a guide for new researchers or a tool for field experts.
- The first classification of literature classification schemes.
- A structured overview of both mature and less developed future research directions that cover the domain of Information Visualization.

1.1. Literature-based Challenges in the Field

There are at least three major difficult challenges within this field:

1. *Understanding what has been already been done*: Many researchers end up losing time due to challenges associated with literature searches. As the Information Visualization landscape grows, so does the number of papers, conferences, and journals. This makes it increasingly difficult to find topic related papers.

Conference	Related Papers
The Annual EuroVis Conference	20
IEEE TVCG Journal	18
IEEE Pacific Visualization Symposium	2
IEEE VAST Conference	3
The Annual Eurographics Conference	3
Journal of Visual Languages & Computing	2
Information Visualization Journal	5
Computer Graphics Forum	3
ACM Computing Surveys	0
Other	30
Total:	86

Table 1: *A list of literature sources we search for survey papers, with the quantity of papers identified from each. For paper searching, we use IEE Xplore [iee16], ACM Digital Library [acm16], Google Scholar [sch16], and Vispubdata [IHK*17]*

2. *Understanding what areas in the domain have yet to be explored*: A researcher may not be truly sure whether a paper does not exist. This is a logical uncertainty. Until the point a paper is discovered, a researcher may still be unsure whether their development has been explored or not.
3. *Making sure your discoveries are not ignored*: Papers are published to present discoveries in their field. However, due to the volume of published material, papers can be missed, forgotten, or reinvented. This challenge is also highlighted by Jim Blinn [Bli98].

The SoS aims to address these challenges by surveying a collection of over 80 survey papers to enable a quick overview of the scope of research directions, what has already been done, and which directions are more open for research. We provide systematic summaries of these survey papers for those with interest in the field. This serves as a valuable starting point for young researchers and a practical reference guide for field experts. We also believe this survey of surveys will reach audiences beyond the field of information visualization and visual analytics, and entice more researches towards the area.

1.2. Literature Search Methodology

Our survey search methodology includes a combination of linear-search and relation-search. Our starting point includes reviewing previous EuroVis State-of-the-Art (STAR) papers. The linear search focuses on looking at each journal or conference and checking each paper that includes keywords such as 'Survey', 'Taxonomy', or 'State-of-the-Art'. The relation-search includes searching the references of each survey paper for related survey papers.

Sources that are searched within our SoS following this methodology are summarized in Table 1. Our survey paper search lasted over a year.
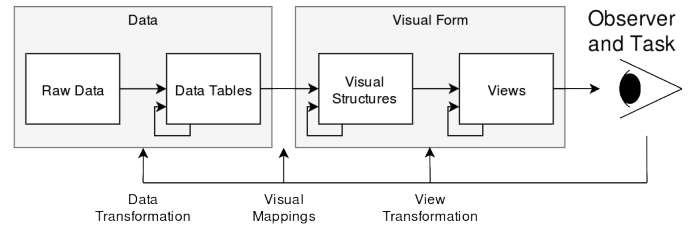
Figure 1: *The original Information Visualization Pipeline model created by Card et al. [CMS99] which we adapt to design our modified classification.*

1.3. Classification Overview

In order to classify each survey, we develop categorical dimensions. The dimensions are derived from previously published and well-known literature based on:

1. An adapted Information Visualization pipeline model originally presented by Card et al. [CMS99]. A visual aid to this pipeline can be found in Figure 1.
2. Subject-based clusters guided by SurVis [BKW16] and the survey paper topics themselves.

Both of these dimensions are explained in greater detail in Sections 1.3.1 and 1.3.2.

1.3.1. The Information Visualization Pipeline

The Information Visualization pipeline model we use to classify the surveys is based on that presented by Card et al. in their classic book *'Readings in Information Visualization'* [CMS99] (See Figure 1). The pipeline describes the transition of raw data into a visualization which is visible to a user. This consists of (1) raw data transformed into data tables via the use of data transformations. (2) Data tables transformed into visual structures via the use of visual mappings. (3) Visual structures transformed into views via the use of view transformations. The final step is (4) User input manipulating the data in order to feed back into the pipeline. For the purpose of the SoS, the pipeline is adapted in order to facilitate the categorization process. The following pipeline stages are used.

1. **Data Enhancement & Transformation** - Data Enhancement and Transformation is used to describe the raw data that is transformed or enhanced in order to derive a data structure(s) that can be used for visualization. The classification also includes how the data is captured and how the data is classified. Survey papers that are data-centric are placed in this category.
2. **Visual Mapping & Structure** - Visual Mapping and Structure defines the techniques to visualise data or data structures. This section also examines how a visualization is structured, such as how the mapping is used or the facets that are included. This category involves mapping the enhanced data to visual primitives, for example, color, opacity, textures, and geometry such as points, edges, as well as 2D and 3D shapes. Survey papers with an emphasis on visual mapping and structure are categorized here.
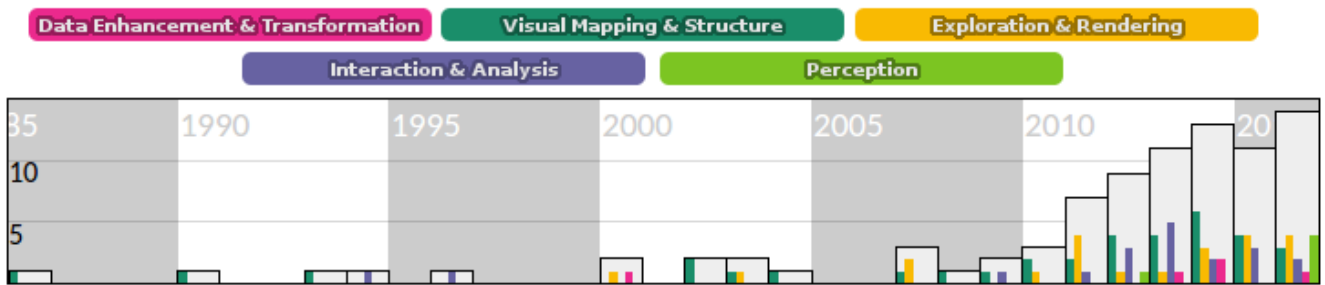
Figure 2: *A histogram providing the frequency of surveys found for each year mapped along the x-axis. each bar represents one year whilst the y-axis provides the amount of survey papers found for each year that meet our scope (Section 1.4). The colored bars represent a further breakdown of the survey papers based on their given classification dimension (discussed in Section 1.3.1). The visualization is taken from the SoS literature browser [McN17].*

3. **Exploration and Rendering** - The three common types of view transformation are location probes that use location to reveal additional information, viewpoint controls which are used to scale or translate a view, and distortions which modify the visual structure [CMS99]. Exploration and Rendering looks at these transformations along with the rendered representation views and projections. This is the presented state a visualization takes upon completion and aims to present the data to the user. Survey papers with a focus on rendering and exploration are classified here.

4. **Interactive Analysis** - Analysis refers to how the user provides feedback to a visualization. A user can connect with a visualization manually, by modifying or transforming a view state, or by reviewing the use, effectiveness, and their knowledge on the visualization. This also includes selection protocols and mapping techniques such as task taxonomy or other viariations of selection. Survey papers with a focus on interactive analysis or tasks are placed here.

5. **Perception** - Perception examines the cognitive interperatation of a visualization from the perspective of a user. Perception can be viewed through the design and creation of user studies or papers relating the visual system to Information Visualization. Survey papers emphasizing perception and user-studies are placed in this category.

These components form the basis of the first dimension of our classification. Figure 2 provides a breakdown of these classification by year.

1.3.2. Subject-based Clusters

Literature and subject-based clusters group similarly focused survey papers. SurVis [BKW16, McN17] is used as an aid in order to discover and diagnose suitable clusters, exploring keywords related to each paper. Survey papers that cite previous survey papers create natural topic clusters that are taken into account within our subject-based clusters. The subject clusters are as follows – Data-Centric contains literature that focus on types of data, or data itself. Multivariate & Hierarchical focus on structured data, or data that visualizes many dimensions. Graphs and Networks focus on liter-

● Data-Centric	◇ Data-Types
	◇ Text-Focus
● Multivariate & Hierarchical	◇ Hierarchical
	◇ High-Dimensional Overview
	◇ Parallel Coordinates
	◇ Glyphs
● Graphs & Networks	◇ Graphs
	◇ Networks
● Geospace + Time	◇ Temporal
	◇ Geospatial
● Coordinated Multiple Views	
● Real-World & Applications	◇ Finance
	◇ Healthcare
	◇ Security
	◇ Systems
	◇ SoftVis
	◇ Frameworks
● Overview	◇ Focus+Context
	◇ Provenance
	◇ General

Table 2: *Hierarchy of subject-based clusters.*

ature that discusses nodes and edges used in visualization, usually standard 2D views. Geospace+Time review surveys with that look at dimensional data. Coordinated Multiple Views (CMVs) centers around literature that examines the coordination or linkage between multiples. Real-World and Applications focuses on literature that reviews topics with an emphasis on practical data or with real-world users. Overview contains literature that provide a more broad survey of the information visualization landscape. A full breakdown of the topics can be found in Table 2.

1.4. SoS Scope

Restrictions are used in order to define, manage, and constrain the scope of the SoS.

Firstly, only survey papers with a focus on Information Visual-

ization and Visual Analytics found within the field are included. This means Scientific Visualization does not fall within the scope. There are several recent scientific visualization surveys not included. For purposes of this survey, we define scientific visualization (SciVis) as the following:

"Data that describes a physical phenomenon is defined as scientific data. Examples of this are fluid flow, living organisms, and data from the natural world." For the purpose of this survey, Euclidean space-time coordinate data is considered Scientific Visualization. Lipsa et al. survey visualization in physical sciences, this does not meet the criteria of our survey [LLC*12]. Edmunds et al. present a framework for flow visualization [ELC*12]. Flow visualization is considered SciVis. Blascheck et al. 's survey on Eye-Tracking data is a review of literature focused on data from a physical phenomenon and can therefore be classed as SciVis [BKR*14].

Papers that focus on Computer Vision are not considered within scope. Datondji et al. present a survey focused on vision based traffic monitoring of road intersections. This is considered a computer vision topic and therefore does not meet the scope of the paper [DDSV16].

Computer graphics and Graph theory literature surveys are not considered in the scope of the SoS. Ghosh and Goswami present a paper reviewing unsolved problems in visibility graphs [GG13]. Although the paper works on graphs, the review focuses on the mathematics of graph theory and not visualization. Biomedical and computational biology are beyond the scope of this SoS. They are covered in an affiliated SoS called the SoS-MDV [AAM*17].

Secondly, publication date is considered. This enables a clear emphasis on recent advancements, and what can be done at this time to increase our advancement in the field of Information Visualization. This is important as it allows us to provide a clearer message about future research fields, as the older survey papers may discuss research areas that are now mature. The emphasis for the SoS is between the years 2010 and 2016. As prior papers are still important, we include surveys that fall out of this time-frame onto our classification table, however, we do not include a detailed description of them. Older surveys that are included in the table only can shed historical light on where authors chose to publish visualization paper before the visualization community developed (prior to 1990).

Finally, the SoS emphasizes literature reviews, as opposed to comparison-oriented papers. Shneiderman et al. look at the innovation trajectories of treemaps, conemaps, and hyperbolic trees, but focus on the global state of each techniques' citations and papers, rather than a comparison of individual papers [SDSW12]. Sedlmair et al. look at scatter-plot and dimension reduction technique choices, as well as multiple reduction techniques for scatter-plots. These techniques are compared as abstractions [SMT13]. Papers such as these are not given a detailed summary, however, we include these in the classification table for completeness.

1.5. Related Work

To the best of our knowledge, there are no previous papers that attempt to review literature in this way. Other papers use alternative methods to address the literature explosion challenge. Laramee

et al. provide an in depth review of unsolved problems in human-centered visualization [LK07]. The review provides an in-depth understanding of challenges identified for each paper which differs from the solution our paper uses, that appropriates important research challenges by looking at the frequency of each challenge across papers. Henry et al. review 20 years of conference publications from CHI, UIST, AVI and InfoVis [HGEF07].

Isenberg et al. present a novel visualization of a database of papers across InfoVis, SciVis, VAST and Vis [IHK*17]. We provide a different view of the data by clustering papers together to quickly understand domains, with a focus on survey papers. Isenberg et al. present topic popularity using research paper keywords across four conferences and provide related papers [IIS*17]. Our paper differs by providing analysis of related papers with a focus on surveys to provide an overview and explore possible research areas in the field.

1.6. Organization of The SoS

Survey Literature is organized using a 2D matrix that incorporates the two classifications discussed in Section 1.3.2 and 1.3.1. Each survey paper is placed at the most relevant intersection of each classification criteria matched. Color is used to signify the depth of which the literature is reviewed in the paper. This is shown in Table 3.

The SoS is structured using the subject-based clusters as the primary organization. This groups related papers together. The modified pipeline classification is ignored in favor of chronological order within the paper's organization. This enables us to describe a natural progression within each section for papers that are intrinsically related.

2. A Classification of Classifications

Classifications are an integral and important part of a survey paper. Table 4 systematically indicates how each survey paper's classification of literature is represented. We provide three characteristics of classifications: **dimension**, **structure**, and **mapping schema**. For this discussion C denotes a classification topic.

The **dimensionality** organizes the space in which the classification is laid out. We sub-divide the dimensionality in three ways. One-dimensional (1D) classification presents the classification topics (C) in linear fashion. Two-dimensional classifications (2D) usually present more than one classification dimension, one on each axis (C), and are usually presented in the form of a table. The third category represents classification topics (C) with three or more dimensions. Common ways to represent additional attributes are through the use of color, shape, or symbols.

Structure represents the organization of the classification. This category is sub-divided into two columns, flat or hierarchical. Flat structures usually represent classification topics (C) with a discrete linear ranking or order. A hierarchy provides the classification topics (C) with a more complex structure by grouping similar items together.

Mapping schema describes how the survey's reviewed literature

		Information Visualization Pipeline				
		Data Enhancement & Transformation	Visual Mapping & Structure	Exploration & Rendering	Interaction & Analysis	Perception
Data-Centric	Data-Types		[STMT12] [SMT13] [ZH15]	[dOL03] [IGJ*14a]		
	Text-Focus	[WSJ*14]	[SB10] [NVPMW14] [GZL*14] [JFCS15] [JFCS16]	[KK15a]	[AdOP12] [FHKM16a]	
Multivariate & Hierarchical	Hierarchical	[AMA*14b]	[EF10] [SHS11] [SDSW12]			
	High-Dimensional Overview			[BTK11] [LMW*15a]	[BCS96]	
	Parallel Coordinates		[HW13]		[DCK12]	[JF16]
	Glyphs		[War02] [BKC*13]			[FIBK16]
Graphs & Networks	Graphs	[Kob13]	[VLKS*11] [BBDW14b] [KKC14] [VBW15a]	[HMM00] [ZXYQ13]	[AP13] [KKC15] [KKCG15]	
	Networks		[ZMS14]	[BBR*16a]	[APS14]	
Geospace + Time	Temporal		[CLW12]	[BDA*14a]		
	Geospatial		[Tob04] [DWL08] [NK16b]	[CGW15]	[NK15]	
Coordinated Multiple Views (CMVs)			[JE12] [HSS15]	[Rob07] [GAW*11]		
Real-World & Applications	Finance	[KCA*16]		[FLVW16]		
	Healthcare		[KM02]	[RWA*11]		
	Security			[SSG12] [WFL*15]		
	Systems		[GBTS13]		[GZR*11] [ZSB*12]	
	SoftVis		[Mye86] [Mye90] [KM07]	[PBS93] [CZ11]		[MIK*16]
	Frameworks		[SNR14]	[HPK*16] [LH10]		
Overview	Focus+Context		[KHG03]	[TGK*14] [TGK*16]	[LA94]	
	Provenance				[GZ09]	
	General	[Chi00]	[DLR09] [WZM*16]	[ED07] [LCWL14]	[SWLL13] [BM13] [SNHS13] [IIC*13] [LCM*16]	[LBI*12] [BNW*16]

Table 3: *A 3-Dimensional hierarchical classification table depicting the categorization of all the survey papers found into the subject-based clusters and the modified Information Visualization pipeline model that have the strongest correlation.* Green Highlighting *represents survey's summarised within the SoS.*

Yellow Highlighting *represents surveys that were not summarised in detail due to prioritization of journals or size constraints.* Pink Highlighting *represents survey's not reviewed in detail within the paper due to year constraints discussed in the Scope (Section 1.4). This table corresponds to the classification example shown in Figure 3 (A).*

Literature:	Dimensions			Structure		Mapping Schema		
	1D	2D	3D+	Linear	Hierarchical	Indirect Mapping	Unique-Mapping	1-N Mapping
[APS14]		✓			✓	✓		
[AMA*14b]			✓		✓			✓
[BDA*14a]	✓				✓	✓		
[BBDW14b]		✓			✓		✓	
[BBR*16a]	✓			✓			✓	
[BTK11]		✓		✓				✓
[BKC*13]			✓		✓			✓
[CZ11]		✓			✓		✓	
[CGW15]		✓			✓		✓	
[CLW12]			✓	✓		✓		
[EF10]	✓				✓			✓
[FHKM16a]		✓		✓		✓		
[FIBK16]		✓			✓			✓
[GZR*11]	✓				✓		✓	
[HSS15]		✓		✓			✓	
[HPK*16]			✓	✓				✓
[HW13]	✓				✓	✓		
[IGJ*14a]		✓		✓				✓
[JFCS15]		✓			✓			✓
[JE12]	✓			✓			✓	
[JF16]			✓	✓				✓
[KKC14]			✓		✓		✓	
[KKC15]		✓			✓		✓	
[KKCG15]		✓		✓		✓		
[KCA*16]		✓		✓				✓
[KK15a]	✓			✓		✓		
[LCWL14]	✓				✓		✓	
[LMW*15a]		✓			✓		✓	
[NK15]		✓		✓			✓	
[NK16b]		✓		✓			✓	
[SHS11]		✓			✓			✓
[STMT12]		✓			✓	✓		
[SSG12]			✓		✓		✓	
[SNR14]		✓			✓	✓		
[TGK*14]		✓		✓				✓
[VBW15a]			✓		✓	✓		
[VLKS*11]	✓				✓	✓		
[WFL*15]		✓		✓				✓
[WSJ*14]		✓		✓				✓
[ZH15]			✓		✓			✓

Table 4: *A Categorization of classification tables found within each primary survey paper (highlighted* green *in Table 3). The table examines how many dimensions each survey table features, the structure of each survey classification, and the type of mapping schema it incorporates. This table uses the paper's visual representation of the classification. If there is more than one classification, the primary classification is shown. This table itself corresponds to the classification example shown in Figure 3 (B).*

(L) is mapped to classification topics (C). We introduce L to refer to a reviewed item (in most cases, the literature being reviewed). This is split into two categories, Unique-mapping and 1-N mapping. Unique-mapping schema map each reviewed item (L) once for every topic (C). This mapping schema is best for finding areas in the field with extensive or limited work, which may guide researchers to immature areas for new research possibilities. Figure 3 presents some examples of **unique mapping**.

Kerracher et al. use a unique mapping schema to plot the design space of temporal graphs [KKC14] by mapping classification topics to the x and y axis, and placing L at the intersection of the two criteria (Figure 14). Nusrat and Kobourov present a task taxonomy for cartogram visualization that conveys how different tasks can be classified. The tasks are uniquely mapped to 4 different classification topics (Figure 18) [NK15]. Wagner et al. present a malware visualization taxonomy and map reviewed literature directly

to the appropriate classification category (Figure 21) [WFL*15]. Our main taxonomy (Table 3) also uses a unique-attribute mapping schema to map our two classification topics, the modified InfoVis pipeline and the subject-based clusters to L. The table also displays how the literature is ordered in the SoS by mapping a unique color to each L.

1-N Mapping differs from the unique-mapping schema by allowing a reviewed item (L) to be mapped up to N times for each classification topic (C) where N is the number of available attributes. Examples of N-mapping can be found in Figure 4. Multiple-Attribute mapping matrices are most suited to comparing different elements, such as techniques or frameworks, against one another. These papers usually offer a checklist and present the criterion each paper fulfills or does not.

Borgo et al. compare different glyph-based visualization techniques using a multiple-attribute mapping matrix and identify

Figure 3: *Examples of classification schemes using **unique-mapping**. C refers to a classification topic and L refers to a reviewed item (in most cases, the literature reviewed). Examples (A) and (B) map L to each of C once. However, example (A) structures the table such that both classifcation topics are represented by an axis and map L to the appropriate intersection. Example (B) maps L to the Y-Axis and each classification topic C on the X-Axis. Example (C) links each of the reviewed items (L) to the appropriate classification topic in the form of a list. Examples (A) and (B) show the same information.*

Figure 4: *Examples of classification schemes using **1-N mapping**. C refers to a classification topic and L refers to a reviewed item. Examples A and B can map L to each of C multiple times. Example A structures the table such that both classifcation topics are represented by the X and Y axes and map the reviewed topics at their appropriate intersection. Example B plots reviewed items to the Y-Axis and each classification topic on the X-Axis. This examples gives a clear comparison of reviewed item's (L). Example C links each of the reviewed items (L) to the appropriate classification topics in the form of a list. Examples A and B show the same information.*

which papers exemplify the proposed design guidelines (Figure 11) [BKC*13]. Tominski et al. compare different magic lens approaches, and what data or tasks are applicable for each [TGK*14] using a N-mapping schema (Figure 24).

Some papers do not map L explicitly in their categorization and choose to display just their classification. We identify these survey papers as incorporating an **indirect mapping**. Some examples of this can be found with Sedlmair et al.'s taxonomy [STMT12] which classifies data characteristics between two different classification topics, Class-Factors and Influences (Figure 5). Another example of this is Heinrich and Weiskopf's state-of-the art report for

Parallel Coordinates [HW13], which presents a hierarchical view of the important topics within the field. This representation does not explicitly show how literature fills the specified topics (Figure 10).

The SoS aims to provide researchers with an understanding of open research areas. We use a 2D, Hierarchical, Unique-mapping table which follows the example found in Figure 3 (A). Our table is used to present a taxonomy which clearly conveys what areas are less developed in terms of survey papers (Table 3). We use a separate table to compare different types of classifications used (Table 4). This follows the same classification scheme but follows Figure 3 (B).

3. Survey Papers

This section provides a collection of summarised survey papers (see Table 3). Each paper is broken down to present the survey's concept, their classification schema, and open areas of research discovered.

3.1. Data-Centric Survey Papers

The Data-Centric section contains literature that focus on types of data or data itself. Of the survey papers reviewed, two categories were identified within as subtopics which include data-type papers that emphasize the type of data surveys and papers that focus on text.

3.1.1. Data-Type Focused Surveys

This subsection presents a diverse collection of survey topics including visual distance measures between images, hardware and software performance, and color maps.

Sedlmair et al. present a taxonomy of visual cluster separation factors in scatter-plots, as well as a qualitative evaluation of recently proposed separation measures [STMT12]. They provide a brief introduction to the area of dimension reduction, as well as their motivation behind the literature survey. The paper discusses chosen cluster separation measures: cluster identification and verification. This is followed by a section on related work before discussing their taxonomy which includes a qualitative data study. The paper presents its taxonomy on visual cluster separation factors, and evaluates them, before discussing the results of the study.

The proposed taxonomy includes four main categories: scale, point distance, shape, and position. These categories are examined during different states of observation. Within-class factors is the first state which includes variables such as density, curvature, and clumpiness. The second state is between-class factors which arise from the variance between two or more properties. For example, a variance of size between two time slices.

Sedlmair et al. suggest that the taxonomy could be extended with new dataset characteristics.

Isaacs et al. present a survey focused on reviewing hardware and software performance data as well as performance visualization techniques [IGJ*14a]. Performance data is defined as data generated to measure the effectiveness and behavior of a process. The
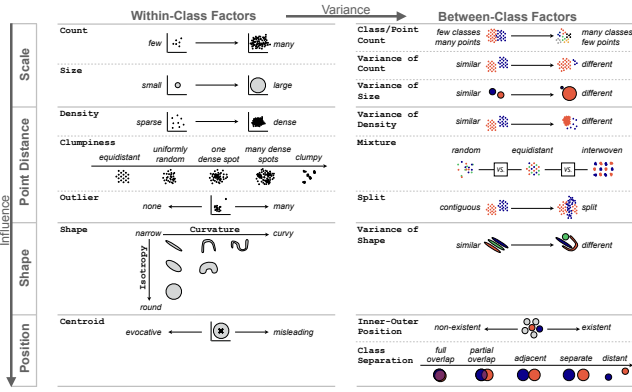
Figure 5: *An indirect mapping taxonomy of data characteristics with respect to class separation in scatterplots. Courtesy of Sedlmair et al. [STMT12]*

survey develops a taxonomy to aid selection of the appropriate techniques to display performance data for both analysts and developers. They introduce the concepts behind performance visualization, including how it is acquired, what form it can take, and the goals of using performance data. The paper discusses the taxonomy by discussing visualization types based on the source of performance data. Hardware visualization examines performance data of hardware and aims to visualize complex hardware topology. Software visualization describes performance data of software, such as software maintenance. Task performance investigates performance of tasks. The final category, application visualization, discusses context-specific performance data. Isaacs et al. also provide an interactive literature browser related to this topic. We provide a full list of these browsers in Table 5.

Isaacs et al. present their design space by looking at the context, scale and goal of each paper. How each paper fits in the presented taxonomy (hardware, software, task, application), scalability, what it visualizes, and what it presents are discussed (see Figure 6).

The paper presents future challenges that are found within the landscape of performance visualization. These include, scalability, the increase in system complexity and it's result on performance visualization, comparing performance differences and ensemble datasets, integration of performance visualization across each taxonomy point, and the creation of performance visualization that facilitates software development and debugging.

Zhou and Hansen present a comprehensive review of color-map generation techniques, and provide a reference for readers who are faced with color mapping decisions [ZH15]. The paper aims to provide a comprehensive overview of various color-map generation techniques. The paper also presents a hierarchical taxonomy of color-mapping literature to guide readers when choosing appropriate literature to review, and classifies representative visualization techniques that discuss usage of color-maps.

The hierarchical classification divides papers into various different color-map generation techniques based on the type of data used in the form of a flow chart. The taxonomy subdivides the data

Figure 6: *Isaacs et al. present a 1-N design space that classifies literature based on the context, scale and goal of each paper. Image courtesy of Isaacs et al. [IGJ* 14a]*

into either discrete or continuous. The discrete data type is split into either nominal or ordinal data types, another division of data comprehension. Both ordinal data and continuous data are linked to color-map transformation literature. The taxonomy also takes into consideration the type of paper and examples.

Zhou and Hansen recommend more research into multivariate or high-dimensional color-mapping as a future research direction. Further research into aesthetically pleasing color maps that still provide insight into data, and the use of meta-data to optimize color-mapping suggestions are also potential future research directions in the field.

3.1.2. Text-Focused Surveys

The text visualization literature is evolving rapidly. This section includes surveys focused on understanding and visualizing text. The section contains four survey papers. The first paper analyses how different text sources are used in visualization with event detection methods. The second focuses on advancement in the field of close and distant reading. The third presents a classification of text visualization techniques. The last survey focuses on the visualization of

Interactive Literature Browsers	URLs
Cartogram Visualization	[NK16a]
Dynamic Graph Visualization	[BBDW14a]
Financial Visualization	[DML14]
High-Dimensional Visualization	[LMW*15b]
Matrix Reordering	[BBR*16b]
Performance Visualization	[IGJ*14b]
Scientific Literature & Patents Vis	[FHKM16b]
Set Visualization	[AMA*14a]
Software Reuse Tasks	[SdVW14]
SoS Literature Browser	[McN17]
Space-Time Cube Visualization	[BDA*14b]
Text Visualization	[KK15b]
Time Visualization	[TA11]
Visualizing Group Structures in Graphs	[VBW15b]

Table 5: *The tables provides an overview of interactive literature browsers found during the literature search. Sorted in alphabetical order*

Figure 7: *This 1-N table identifies event detection techniques and displays the the use of visualization with each. The table reveals that clustering techniques are often represented using the river metaphor. Image courtesy of Wanner et al. [WSJ*14]*

'scientific literature and patent' text sources. Table 3 cites a number of other text surveys as well.

With the growth of social media and micro-blogging, Wanner et al. take the opportunity to explore the use of analytical processes on real-time data and divulge key event-detection approaches that can be used for textual data [WSJ*14]. They review the different available data sources when working the text-data streams including news, email, micro-blogging, research papers, and metadata. They discuss methods that are used to process text. The study examines the event detection methods that are grouped including: clustering techniques, classification-based, statistic based, and miscellaneous techniques such as the Kalman Filter and Fourier Analysis. They finish by discussing evaluation methods that can be applied to text-data.

Wanner et al. classify 51 papers across their survey. They classify these in various ways including: data source, text-processing methods, automatic event detection methods, visualization methods, and tasks supported. The main taxonomy displays the use of automatic event detection methods on the y-axis, and then characterizes each research paper via the visualization technique along the x-axis. By doing this, they can investigate the correlation between these two fields. (Shown in Figure 7).

The reviewed papers had little focus on discussion forums. Demand for more sophisticated techniques such as topic modeling is something that will become more apparent in the future, and event detection algorithms seem to exclude important information for deciding whether items are newsworthy or not. This could be considered new research in this area.

Jänicke et al. present recent advancements in the field of visualizations that support close and distant reading of textual data [JFCS15]. Nancy Boyles defines close reading as "reading to uncover layers of meaning that lead to deep comprehension" [Boy12],
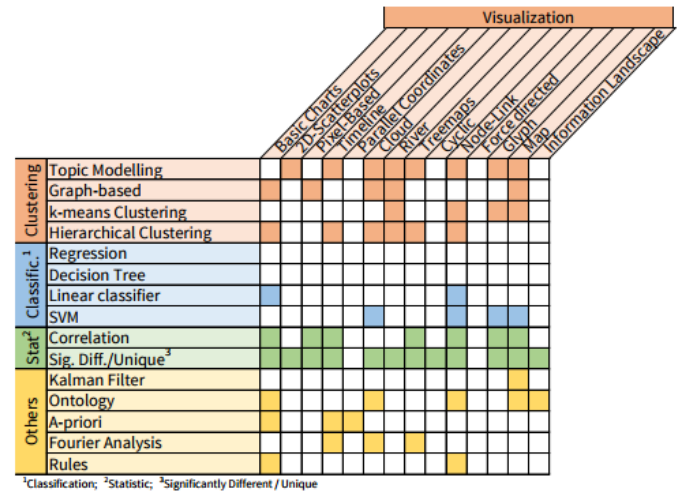
whilst Moretti describes distant reading with the statement, "a little pact with the devil: we know how to read texts, now let's learn how not to read them" [Mor13]. The paper focuses on quantitative literary text analysis using statistical analysis methods for visual analytics and visualization. Literature in the digital humanities is also covered. These are categorized using a taxonomy for applied methods [JFCS15]. The paper looks at different types of techniques, such as color mapping or heat-maps, for close-reading analysis, distant-reading analysis, and combinations of both.

Jänicke et al. classify papers based on the type of reading analysis provided, which is broken down for each paper. This is compared to the method of analysis, which includes single text analysis, parallel-text analysis, and corpus analysis.

They provide a large collection of areas for future research which includes novel techniques for close reading, visualizing transposition in parallel texts, geospatial and temporal uncertainty, usability studies, and the development of design guidelines for scholars. All these areas are ripe with unsolved problems in the field. Jänicke et al. published an extended version of this survey [JFCS16].

Text visualization is becoming a more mature field within information visualization and Kucher and Kerren aim to classify the literature using an interactive browser [KK15a]. See also Table 5. They examine the field of text visualization and closely related surveys. Kucher and Kerren then present the taxonomy they have created to look at different areas within text visualization and how they can be subdivided (the figure is provided in the supplementary material).

Kucher and Kerren categorize text visualization techniques into 5 main categories: Analytic Tasks, Visualization Tasks, Domain, Data, and Item Visualization.

			Close Reading					Distant Reading						
		References	Plain	Color	Font size	Glyphs	Connections	Structure	Heat maps	Tag clouds	Maps	Timelines	Graphs	Miscellaneous
Single Text Analysis	enhanced text views	[Pie10], [CGM*12], [Pie13], [GWFI14]				x								
		[PSA*06], [CTA*13], [Ben14], [BJ14]		x										
		[ARLC*13]				x	x							
		[WMN*14]		x	x									
	both	[VCPK09], [BGHJ*14], [KJW*14]		x				x		x				
		[WJ13b], [CMLM14], [KZ14]		x			x							
		[Cay05]	x					x						
		[CDP*07]		x					x					
		[WV08]		x									x	
		[MFM13]	x										x	
		[RSDCD*13]	x											x
	abstract text views	[KO07], [FS11], [CTA*13], [OKK13], [Ben14]							x					
		[Pie05]						x						
		[PBD14]												x
Parallel Text Analysis	section alignments	[WH11], [HKTK14]		x										
		[Cor13], [WJ13b]		x			x							
		[JRS*09]		x				x						
		[GCL*13]		x						x				
		[JGBS14b]		x			x		x					x
	sentence alignments	[BGHE10]		x			x							
		[JGBS14a]			x		x							
Corpus Analysis	statistics for textual entities	[Bea08], [Bea11], [Bea12], [BJ14]								x				x
		[WJ13a], [HCC14]												x
		[CWG11]		x	x					x				
		[Mur11]		x						x				
		[FKT14]	x							x	x			
	relationships between texts	[EX10], [Gal11], [WH11], [Joc12], [CEJ*14], [Ede14]											x	
		[RRRG05]								x				
		[OST*10]		x										x
		[Wol13]		x									x	
	relationships between textual entities	[RRRG05], [AGL*07], [vHWV09], [KKL*11], [MLSU13], [WJ13a], [Arm14]											x	
		[GZ12], [RFH14]		x									x	
		[MH13]		x							x		x	
		[AKV*14]		x							x			
	social networks	[Cob05], [CSV08], [BDF*10], [RD10], [BHW11], [Kle12], [Boo13], [KOTM13], [Töt13], [Pet14]											x	
		[KLB14]	x										x	
	space and time	[JHSS12], [JW13], [DNCM14], [GDMF*14], [ÓML14]									x	x		
		[Wea08]								x	x	x		
		[BPBI10]		x							x	x		
		[DWS*12]	x								x	x	x	
		[HACQ14]		x							x	x	x	
	space	[MBL*06], [DFM*08], [Tra09], [GH11b], [EJ14]									x			
	time	[KBK11], [ARR*12], [LWW*13]										x		
		[CLT*11], [CLWW14]								x		x		
		[HSC08]	x									x	x	
		[DWS*12]	x								x	x	x	
		[ESK14]										x		x
		[HPR14]		x								x		

Figure 8: *A 1-N Taxonomy by Jänicke et al. to map reading techniqes found within different analysis methods. Image courtesy of Janicke et al. [JFCS15]*

Federico et al. present a survey that addresses literature papers with an emphasis on visual approaches for scientific literature and patents. Literature is classified looking at the data-type and the fulfilled task of each [FHKM16a]. The paper identifies four types of data found within scientific literature papers and patents. These include text, citations, authors, and meta-data. The data types create the main structure of the paper, with each data-type divided to explore how tasks are analyzed for each category. The paper also provides a breakdown of literature that handles multiple data types, which focus on different tasks.

Their indirect classification uses two tables. The first table displays the total number of publications that match the criteria at the axis intersections, which examine the four data-types and different tasks such as lookups, relation seeking, and patterns. The second table is similar but records how multiple data-types are mapped to a new set of tasks. This includes aggregation, labeling, composition, tight integration, and multiple views. The table shows that most literature in the field attempts to analyze patterns within text.

Federico et al. break down future research directions for each data-type. Research on text data identifies contextual identification

in compact space. Research into author data suggests uncertainty analysis of ambiguities with synonyms and homonyms. Limited work is provided with citation data and meta-data, with citation data focusing on citation count and meta-data ignoring many pieces of gathered data, which narrows the fields significantly. Some other broader examples include quantitative and qualitative evaluations, scalability, user interaction, and research into user-tasks.

3.2. Multivariate & Hierarchical

This category discusses multivariate, high-dimensional, and hierarchical visualization. These are grouped together due to their association within large datasets. The reviewed content on this subject can be broken down to look at hierarchical visualization, an overview of high-dimensional visualizations, parallel coordinate plots, and glyph-based visualization.

3.2.1. Hierarchical Surveys

This section includes surveys that have an emphasis on hierarchical structures. The first survey focuses on the classification of hierarchical aggregation strategies for visualization. The second survey provides a design space of implicit hierarchy visualization to compare literature in the field. The final survey looks at set-typed data and how set-typed data visualizations relate to different tasks.

Elmqvist and Fekete review the use of hierarchical aggregation within Information Visualization. Hierarchical Aggregation is based on iteratively building a tree of aggregate items. Elmqvist and Fekete use this review to present a model that enables augmentation of existing techniques with multiscale functionality [EF10]. They first describe related reading before discussing hierarchical aggregation and presenting various related techniques (scatter-plots, parallel coordinates, etc). The paper then presents examples of hierarchical aggregation within visualization before presenting their classification and guidelines (the figure is provided in the supplementary material). Elmqvist and Fekete end by describing design guidelines for hierarchical aggregation.

Elmqvist and Fekete derive a classification of visual aggregation strategies. The table looks at the data-structure used, visualization mapping type, type of visualization, type of aggregation, the visual aggregate, and what is being visualized.

Elmqvist and Fekete discuss future work which includes model refinements, as well as investigating the trade-off between accuracy and usability. Finally, they look at the idea of reviewing different hierarchical structures such as Directed Acyclic Graphs (DAGs).

Schulz et al. construct a survey categorizing the design space of techniques for hierarchy visualization, with the aim of guiding researchers to unexplored research areas within the field [SHS11]. Examples of techniques for implicit hierarchy visualization include spatial dimensions and node representation. Schulz et al. present their aims, before presenting the design space for hierarchy visualization (see Figure 9). They follow this by discussing some of the limitations of the design space such as techniques with *mixed Treemaps*. Using the design space, they present novel techniques that are not explored with visual representations, using their own rapid visualization prototyping software.

Technique	Year Published	Figure in the Appendix
Treemap [10],[11]	1991	12(c)
Contrast Spiral Treemap [53]	2007	13(o)
2 1/2D Treemap [65]	1992	12(d)
Polar Treemap [15]	1993	12(e)
Jigsaw Map [59]	2005	13(g)
Generalized Treemap (Pie) [8]	2006	13(j) - left
Generalized Treemap (Pyramid) [8]	2006	13(j) - middle
Generalized Treemap (Pie+Pyramid) [8]	2006	13(j) - right
Tree Cube [30]	2003	13(c)
3D Treemap [15]	1993	12(h)
Cushion Treemap [51]	1999	12(m)
Voronoi Treemap [44]	2005	13(f)
Circular Partitions [45]	2008	13(q)
Quantum Treemap [66]	2001	12(q)
Data Jewelry Box [58]	2002	12(r)
Cascaded Treemap [19]	2008	13(p)
Ellimap [48]	2007	13(k)
Treemaps with Ovals [15]	1993	12(f)
Pebble Map [36]	2003	13(b)
CropCircles [37]	2006	13(h)
Lifted Treemap [68]	2009	13(r)
3D Nested Treemap [25]	1999	12(n)
Information Cube [29]	1993	12(i)
Nested Columns [15]	1993	12(g)
2D Icicle Plot [12]	1983	12(b) - left
Castles [17]	1981	12(a)
Cushioned Icicle Plot [52]	2007	13(l)
Triangular Aggregated Treemap [48]	1998	12(l)
Sunburst [40]	2000	12(p)
Interring [42]	2002	12(s)
PieTree [38]	2000	12(o)
Radial Edgeless Tree [46],[47]	2007	13(m)
Steptree [20]	2004	13(d)
3D Icicle Plot [12]	1983	12(b) - right
Nested Hemispheres [26]	2004	13(e)
3D Nested Cylinders and Spheres [21]	2006	13(i)
3D Sunburst [70]	2007	13(n)
3D Beamtree [23]	2002	13(a)
Information Pyramids™ [28]	1997	12(k)
Cheops™ [50]	1996	12(j)

(Columns in the figure: 2D, 3D, Adjacency, Overlap, Inclusion, Rectangles/Cuboids, Squares/Pyramids, Triangles/Trapezoids, Polygons, Circle segments/Cylinders, Circles/Ellipsoids, Ellipses/Ellipsoids, 3D Subdivision, 3D Subdivision, 1D Subdivision, 2D Packing, 3D Packing, Year Published, Figure in the Appendix)

Figure 9: *Design space for implicit hierarchy visualization created by Schulz et al. to compare techniques in the field. Image courtesy of Schulz et al. [SHS11]*

The survey paper investigates four main classification topics. Spatial dimensionality, how nodes are represented (such as their shape), how edges are represented (Do they overlap? Are they included?), and the layout (subdivision or packing).

Schulz et al. propose that the design space presented in the paper can be used to create more surveys within the field of hierarchy visualization. New layouts can also be created using the characterization of implicit hierarchy visualization.

Visualization of sets can be a demanding task due to the wide variety of possible relations between them. Sets are defined as items that are grouped into sets based on specific properties. Alsallakh et al. present an overview of state-of-the-art techniques for visualizing different forms of set data (defined as a collection of unique objects called set elements) which can be used to select appropriate techniques for different scenarios [AMA*14b]. They first define set-type data before looking at some common tasks. The tasks are either related to elements, element attributes, or relationships between sets. The paper then provides examples of different ways to visualize set-typed data such as using euler diagrams, bubble-sets, pivot-paths, and scatter views.

Alsallakh et al. classify literature by constructing an overview of different tasks and techniques that are supported, partially supported, or supported with an interaction requirement. These techniques include Euler-based techniques, overlays, node-link's, matrices, aggregation, and scatter techniques (the figure is provided in the supplementary material).

The paper presents an abundance of future research in this area. These include scalability within set-typed data, re-ordering sets to reveal clusters, a user study on the effectiveness of techniques, visualizing uncertainty, temporal set-typed data, generating euler diagrams with specific properties, visualizing set in context of other data types, and comparing multiple set families. There is also a discussion of improvements with coordinated multiple views and matrix-based representations of set-type data.

3.2.2. High-Dimensional Surveys

The high-dimensional section focuses on surveys that provide a broad overview of the field of high-dimensional visualization. "High-Dimensional" is defined as 'any data set with a dimensionality that is too high to easily extract meaningful relations across the whole set of dimensions' by Bertini et al. [BTK11]. The section covers two surveys papers. The first paper reviews quality metrics for high-dimensional visualization. The second survey discusses the recent advancements for visualization of high-dimensional data.

Bertini et al. review techniques that use quality metrics, defined as *"a metric calculated at any stage of the information visualization pipeline, that captures properties useful for the extraction of meaningful information about the data"*, to find meaningful results for the exploration of high-dimensional data. By analyzing literature related to the topic, Bertini et al. provide a systematization of approaches that use quality metrics in high-dimensional data. [BTK11]. They begin by describing the background and search methodology used within the survey. The paper presents the quality metrics pipeline, a modified version of Card et al.'s information visualization pipeline [CMS99], which we also use, before discussing their systematic analysis of the literature. Bertini et al. then give examples of quality metrics factors such as what is measured, where, and its purpose.

The survey classifies each work of literature by examining the visualization technique, the metrics measured which range from clusters to image quality, where the measurements take place, the purpose of the metrics, and any interaction techniques available.

Four main areas for future research are highlighted in the paper. Evaluation is poorly represented within the field with a small representation of the papers using a real-world setting and data. This could pave a path for future research. Perceptual Tuning is another area with weak focus, as well as scalability of data representation. Metrics systematization is the final area for directed future research. There are a wealth of different metrics which makes them very hard to compare and without guidance could end up as redundant metrics usage. Composite visualization of quality metrics is also discussed.

Liu et al. provide a comprehensive survey focused on the advances of high-dimensional visualization techniques between 2000

Figure 10: *An indirect hierarchical taxonomy of important topics withn Parallel Coordinate creation. Created by Heinrich and Weiskopf, the taxonomy reflects the different sections of their survey. Courtesy of Heinrich and Weiskopf [HW13] .*

and 2014 [LMW*15a]. The study aims to help practitioners understand the recent advances in this area with the hope of inspiring the creation of new visualizations and the comprehension of future opportunities with high-dimensional data. They first discuss their classification and methodology before discussing various sections or the information visualization pipeline. Similar to ours, this includes data transformation, visual mapping, and view transformation.

The taxonomy of the survey is based on the information visualization pipeline (the figure is provided in the supplementary material). The classification clearly represents when techniques are to be used within the pipeline, as well as some action-driven classification signifying the time of use for each technique. Some examples of this are histograms, jigsaw maps, and glyphs. Liu et al. have also created custom action-driven classifications to further improve their taxonomy.

Liu et al. take great care to expose multiple directions for future research. Subspace clustering is an important technique used when visualizing high-dimensional data, and exploring non-axis-aligned methods may lead to new view-selection techniques. Understanding uncertainty within data is also an important research area. As the amount of data increases, so does the understanding of the quality of data and the best way to visualize this high-dimensional data will become more apparent. Liu et al. also discuss model manipulation, topological data analysis, as well as problems occurring with multivariate volume visualization and machine learning with a link to high-dimensional data visualization.

3.2.3. Parallel Coordinates Surveys

Parallel Coordinates are a technique used to visualize multivariate and high-dimensional data. Parallel coordinates are constructed by placing axes in parallel, the choice of layout depends on the number of axes, and the range of data [HW13]. Included here are two survey papers focusing on Parallel Coordinates. One focuses on the design and structure of Parallel Coordinates and the second collects the findings of user-evaluation studies, comparing the use of a standard 2D parallel coordinate plot with other variations.

Heinrich and Weiskopf present the state of the art in visualization techniques for parallel coordinates, which is well-known for exploratory data analysis [HW13]. The paper describes the parallel coordinate plane, and reviews different variations of the methodology, with the hope of directing research in new areas related to the

topic. The paper aims to provide (1) a taxonomy of techniques related to parallel coordinates, (2) a review of challenges in the field, (3) a reference to important literature in the domain, and (4) a guide to the use of parallel coordinates for applications.

Heinrich and Weiskopf split their hierarchical taxonomy into four main topics: Geometry which focuses on the coordinate system creation; Image Generation maps data to the coordinate system; Image Analysis highlights visual perception of the mapped parallel coordinate; and Interaction emphasizes manipulation of the parallel coordinates (see Figure 10).

They state that an evaluation of existing tools would be required to identify issues in the implementation of parallel coordinate techniques. Additionally, more in-depth studies can test each category in order to find underrepresented research areas. They avoid evaluating techniques based on their applicability, correctness, usability, and performance. These are considered areas for future work within the field.

Johansson and Forsell provide a thorough literature review survey studying user-centered evaluations that investigate use and usability issues in the field of parallel coordinates. The paper aims to address issues within the domain and provide a set of guidelines for future research [JF16]. They first divide the 23 papers identified into 4 categories: evaluating axis layouts, comparing clutter reduction methods, showing practical applicability of different parallel coordinates, and comparing parallel coordinates with other analysis techniques. These categories are discussed in detail with reference to the evaluated papers. The categories include cluster analysis, correlation analysis, outlier search, value retrieval, pattern detection, and line tracing.

The paper provides a classification of how each paper compares with standard 2D Parallel Coordinates (2DPC). The table (provided in the supplementary material) maps the selection of aspects evaluated. The third dimension uses color to indicate how the advanced feature compares to a standard 2D Parallel Coordinates: Yellow identifies no significance variation in the results, Green signifies the extension is reviewed to be better than 2DPC, and red signifies the reviewed item compares unfavorably with 2DPC.

Johansson and Forsell provide a plethora of different research directions in the field. The paper proposes that existing axis configurations in parallel coordinate plots have not been thoroughly studied, with an emphasis on 3D parallel coordinates. They discuss the lack of conclusive clutter reduction comparisons, and propose

Figure 11: *A 1-N categorization of glyph-based approaches created by Borgo et al. In Desgin Guideline 2, -3 represents a small amount of complex glyphs with +3 displaying a large number of simple glyphs. Courtesy of Borgo et al. [BKC*13]*

Figure 12: *A 2-Dimensional table showing the classification of the literature in the glyph-based user-study survey. Courtesy of Fuchs et al. [FIBK16]*

research in evaluating clutter reduction for complex patterns in parallel coordinates. Some other areas for future research discussed include longitudinal studies, temporal views in 3D parallel coordinates, and the aesthetic design of parallel coordinates.

3.2.4. Glyph Surveys

A glyph is defined as *'a small independent visual object that depicts attributes of a data record'* [BKC*13]. Glyphs use visual primitives to represent different attributes in multivariate data. This section contains two literature surveys. The first presents design guidelines for the creation of glyphs. The second paper presents a selection of glyph comparison papers to provide an evaluative perspective on glyph usage.

Borgo et al. examine the fundamental concepts and design guidelines of glyph-based visualization, and how current implementation techniques adhere to them [BKC*13]. They first discuss the concepts and history of glyph usage. The paper covers the design and usage guidelines for glyphs, data mapping, shape design, glyph appearance, glyph placement, rendering, and glyph interaction. Borgo et al. then discuss the application of glyph-based visualization in different visualization scenarios.

Borgo et al. categorize the literature by examining each technique and determining whether they fulfill the design guideline criteria. There are 13 tasks created that the papers are categorized according to, which range from complexity and prioritization to design and balance. The papers are also categorized according to the focus on different visual channels including color, shape, size, orientation, texture, and opacity (see Figure 11).

Borgo et al. suggest the need for a common framework to direct
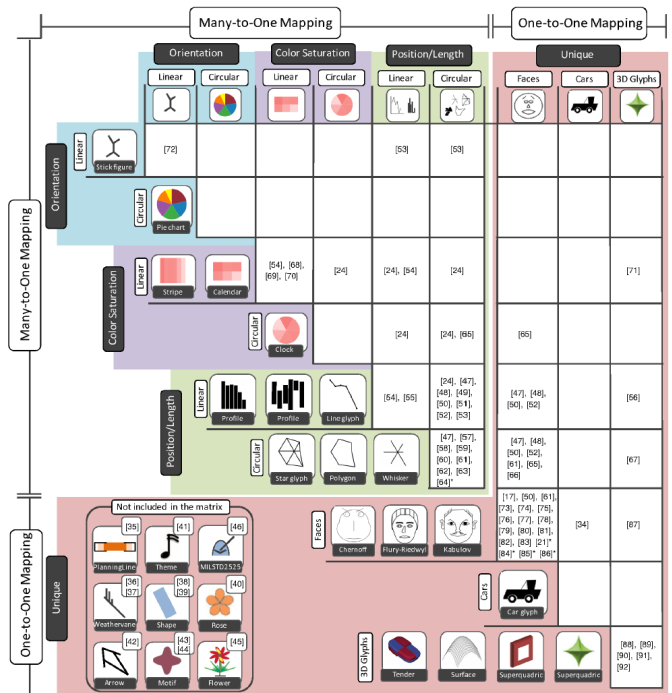
more energy to the field. They also suggest evaluating extended data dimensions using glyph representations.

Fuchs et al. aim to assist researchers in gaining a stronger understanding of user-studies within the glyph design space [FIBK16]. Data glyphs have a wide variety of designs and uses, with many studies carried out across different glyph-types, however an overview of these studies has never been recorded. They present how the expanse of user-studies in the field compare. There are many types of data glyphs analyzed within the paper including: orientation-based, which include angular representations to represent dimensions; color saturation, which uses various saturation's to represent dimension values; positional or length glyphs, which plot dimensional data to bars or lines; and Faces, which map dimensional data to different attributes of a face. They focus on studies with measurable tasks using data glyphs. Tasks are discussed within the classification. The three main goals identified are 1) a comparison of glyph designs based on observed performance with the aim of ranking each design. 2) A comparison of variation within glyph types to detect important features within the data type. 3) The comparison of glyphs as opposed to raw data tables in order to stimulate the use of visual representation rather than textual representations [FIBK16]. Fuchs et al. present the results of their survey by reviewing the influence of background information or the layout, the quantity of data, data dimensionality, the influence of the task, and the effect of metaphoric glyph design.

			Graph structural encoding				
			Space filling	Node-link	Matrix	Compound	Other
Temporal encoding	Multiple timeslices	Sequential view	TS08 (5)	BdM06 (47)	FQ08 (1)	RPD09 (4)	
		Juxtaposition	TJ92 (1)	RM13 (20)	PS12 (3)	BD08 (4)	*Graph statistics* GGK*11 (4) *Alluvial diagrams* RB10 (3)
		Additional spatial dimension		ADM*04 (12)	BPF14 (1)		GHW09 (1)
		Superimposition		FAM*11 (5)	MGK11 (1)		
	Embedded	Merged		GdBG11 (5)	BPF14 (1)		
		Nested	HDKS05 (2)	SLN05 (2)	YEL10 (3)		
		Time as a node		TDKB07 (2)	Hoe11 (1)		

Figure 14: *Design Space for Temporal Graph Visualization. Shading represents the number of papers found per combination (number in brackets). Courtesy of Kerracher et al. [KKC14]*

The literature is classified according to data-to-visual primitive mapping. This classification allows Fuchs et al. to understand some comparisons with limited discussion and discuss the benefits of further work within the area (see Figure 12).

Fuchs et al. provide many open research areas within their survey. There is a large contrast in the observations of data glyphs for quantitative data compared to observations with qualitative data. There is also little research on how data glyphs differ between synthetic and real data. Fuchs et al. find a high percentage of studies fail to look at the use of data-glyphs in exploring data and extracting information, but focus on presentation and simple output in it's stead. They propose that data glyphs can be explored in more complex layouts in order to understand further use and glyph influence in applications.

3.3. Graphs & Networks

There is a large body of work published on graphs and networks with over 10 surveys about the topic. Graph papers tend to have a concise range of citation years, which can be seen in Figure 13. This figure shows the quantity and range of citations found for each survey paper.

3.3.1. Graph Surveys

This section includes surveys with a focus on graph papers. A graph is defined as *'a diagram showing the relation between variable quantities'* [SP16]. The section includes six survey papers: an analysis of large graphs, a classification of dynamic graphs, the visualization of grouped structures in graphs, and three papers providing an understanding of temporal graph visualization.

Von Landesberger et al. examine the visual analysis of graphs designed for large data sets. The survey reviews current techniques, with regards to the types of graph supported, and aims to present open research challenges in the field [VLKS*11]. They give some key definitions for graphs and preprocessing techniques that are related to the survey, including different types of graphs (directed, dynamic, compound, etc.). They introduce some visual representations of these graphs, as well as how interaction in graphs is incorporated. Von Landesberger et al. finish the survey with graph analysis.

Von Landesberger et al. classify graphs according to both time-dependency and structure. Time dependency can be split into either static graphs, or time-dependent graphs, whilst the structure can be classified as a tree structure, generic graph structure (directed, undirected) or compound graphs (the figure is provided in the supplementary material).

They provide extensive insight into the future challenges within the field. These include scalability, uncertainty, perceptual evaluation, interaction, task evaluation, data-type analysis, generic frameworks, user studies, and graphing benchmark systems.

Using the SurVis system [BKW16], Beck et al. create a paper deriving a hierarchical taxonomy of dynamic graph visualization which is achieved by categorizing and tagging publications within the area, as well as evaluating and comparing the use of animated diagrams against time-line diagrams (see Table 5). The paper first classifies a static graph as $G := (V, E)$, where V represents the vertices and E defines the edges. Understanding this allows a representation of a dynamic graph as $G_i := (V_i, E_i)$, where i defines the number of time steps for the graph [BBDW14b]. The paper could also be considered a survey within our temporal categorization. The three main section discussed are animation, time-lines, and a hybrid of the two (the figure is provided in the supplementary material).

The hierarchical classification of dynamic graph visualization provides the density of reviewed papers per category, as well as some example techniques for each subsection. Beck et al. also provide an interactive system using SurVis [BKW16] to provide extra categorization for user-exploration of the dynamic graph visualization field.

Beck et al. point out a number of open research areas in the field including the evaluation of dynamic graph visualization techniques, scalability of dynamic graphs, a larger effort in hybrid visualization, extended dimensionality of data, new interaction approaches, and new applications for dynamic graph visualization itself.

Kerracher et al. present a paper which identifies the design space of temporal graph visualization [KKC14]. They aim to identify this design space, expanding the classification to create an open research area, and raising awareness of the temporal graph visualization field. The design space of temporal graph visualizations is divided into two categories, a graph structural dimension and a temporal dimension, before looking at a combined view. After showing a topology of their survey results, they discuss their findings.

The design space presented specifies a correlation between spe-
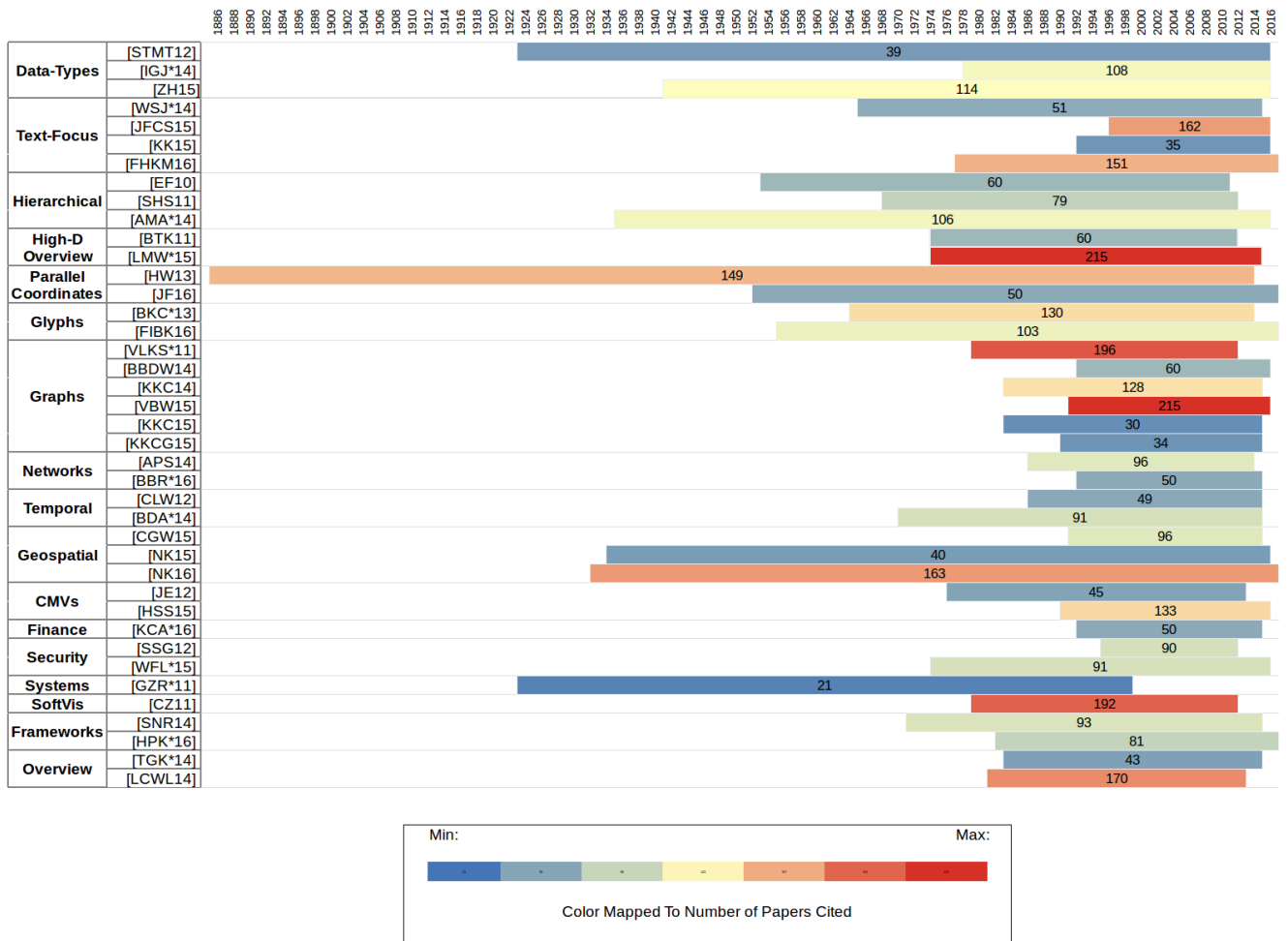
Figure 13: *Visualization of the number of citations within each survey paper discussed, where the years spanned is mapped to the length of each bar along the x-axis. The color represents the number of papers cited within each survey.*

cific temporal encodings and specific graph structural encodings. Kerracher et al. create a 7x5 matrix yielding a total of 35 different combination of temporal graph visualizations (Figure 14). Of these 35, 14 are not filled. The design space also color shades each cell based on the density of papers having this combination. This design space informs the reader that node-links within a sequential view are the most common type of temporal graph visualization.

Kerracher et al. suggest that there could be more research on temporal visualizations using space filling techniques, as well as matrices for dense networks. There are also many unexplored areas found within their design space. An example of this would be a space-filling graph with superimposed time-slices (refer to Figure 14).

'Graphs or networks are used to model relationships between objects of any kind' [VBW15a]. Vehlow et al. survey the use of these group structures within graph visualization in order to gain a meaningful insight into the underlying data. They provide a clas-

sification of techniques used to visualize these group structures. Vehlow et al. present an overview of group structures in graphs and their encodings before presenting their classification table. The survey discusses node attributes, juxtaposed visualization, superimposed visualization, and nested visualizations.

Vehlow et al. classify the surveyed papers based on group structure and group visualization. The group structure can be categorized as disjoint flat, overlapping flat, disjoint hierarchy and overlapping hierarchy (the figure is provided in the supplementary material). Of these four, disjoint hierarchy is the most densely populated. Color nodes, glyph nodes, juxtaposition, superimposition, and nested visualizations are identified visualization techniques (see Figure 15).

Vehlow et al. interview 7 domain experts to identify 5 research challenges in the domain. These are time-varying groups and comparison, data complexity, scalability, and the use of interaction techniques such as providing alternative group structures based on

Figure 15: *Taxonomy table created by Vehlow et al. correlating group visualizations and group structures. Courtesy of Vehlow et al. [VBW15a]*

users' instructions. The final research challenge identified focus on tasks and evaluation. This involves evaluating group structures based on given tasks in order to analyze the most appropriate group structure to visualize.

Kerracher et al. look at the quickly-developing research area of visual representations of temporal graphs. This taxonomy reviews the design space in a way that enables the reader to review overlapping task categories for the purpose of formalizing graph decisions based on these criteria. [KKC15, KKCG15]. Kerracher et al. incorporate the Andrienko Task Framework (ATF) [AA06] as the foundation of their classification (the figure is provided in the supplementary material). They analyze the limitations of the framework and propose an extension to consider structural tasks. They then present a summary of tasks primarily for the use of temporal graphs.

Kerracher et al. use the Andrienko Task Framework (ATF) [AA06] to create an indirect classification for their survey. This task categorization includes lookup tasks, comparison tasks, and relationship seeking tasks. This is combined with a classification of data items (the figure is provided in the supplementary material)

discussed in their previous work, *The Design Space of Temporal Graph Visualization* [KKC14].

Kerracher et al. examine the quickly-developing research area that is visual representation of temporal graphs, with a focus on techniques to support exploratory analysis tasks [KKCG15]. Kerracher et al. discuss a task classification which is used to create groups of exploratory tasks based on data attributes. The paper then maps visualization techniques to the quadrant classification as well as the tasks categories and provides examples of the related systems (the figure is provided in the supplementary material).

Kerracher et al. point out a few areas for future research in the field. Using the task taxonomy, a new classification can be made for more than graph areas, such as static graphs, multivariate graphs, and graph comparison. The taxonomy also opens up more research into temporal graphs, with a focus on classification of visualization techniques and mapping tasks to real world scenarios. These could lead to further research on edges cases or unsupported tasks.

They note that more research needs to be aimed at incorporating techniques from a wider range of research areas than the ones typically associated with temporal graph visualization. In particular, techniques used to support the comparisons of data items in temporal graph visualization.

3.3.2. Network Surveys

This section provides an understanding of the survey landscape for network visualization. Bertin defines a network as the following: '*when the correspondences on a plane can be established among all the elements of the same component, the graphic is a network*' [Ber83]. This SoS discusses two surveys on this topic which include a task taxonomy with focus on network evolution analysis and a classification of matrix reordering methods for network visualization.

Ahn et al. provide a literature review with a focus on visualization tasks for network evolution analysis. The paper surveys 53 existing systems and creates a taxonomy with the aim of providing suggestions for designing future visualization tools in the domain [APS14]. This paper could also be categorized in the temporal space of the SoS but is presented here since networks are the primary focus. The paper identifies three aspects of the systems: entities, properties, and temporal features. These aspects are broken down to create the design space for network evolution analysis. The paper uses the design space to analyze task frequency within network evolution analysis and surveys domain experts on their views of the design space. The results show 67% of domain experts rated the design space as very positive (see Figure 16).

The three aspects of the system identified (entities, properties, and temporal features) are used to create the design space for network evolution analysis. The entities are broken into three subcategories including node/link, groups, and networks. The entity signifies what is being analyzed. Temporal features are displayed on the Y-axis, looking at individual events, shape of changes, and the rate of changes. The properties of the task are displayed inside the table and provide guidance on when a task can be applied, and what information is needed.

Figure 16: *Design Space of network temporal evolution tasks courtesy of Ahn et al. [APS14]*

Future research directions for this field include reviewing the importance of domain properties, additional research on temporal features such as *rate of changes*. Granularity or the scale for analysis had few related papers and is an option for future research. Finally, more research into compound tasks is considered a critical research direction by experts in the field.

Behrisch et al. provide an overview of algorithms used to reorder visual matrices of tabular data. A visual matrix is defined as a visual representation of tabular data used to depict graphs and networks [BBR*16a]. Their survey provides a guide to reordering algorithms in a unified manner to enable a wide audience to understand their differences and subtleties, and provides an overview of how, and when these algorithms are used. Behrisch et al. start by providing an introduction to the visual matrix. They discuss the different pattern types which can be used with matrices before deriving their taxonomy. They review some examples of each, before comparing the performance, and how they were tested. The paper finishes by describing directions on algorithm selection.

The matrix reordering algorithms are classified into seven families. These families are grouped by the type of algorithm which includes Robinsonian, Spectral, Dimension Reduction, Heuristic Approaches, Graph Theoretic, Bi-clustering, Interactive User-Controlled.

Some of the open research directions include hybrid solutions to 'global vs local' algorithms, and research into similarity, or distance calculation. Frameworks to assess the quality of patterns within matrices and craft objective functions to optimize algorithms' performance, and human assisted reordering are important.

3.4. Geospace + Time

The section presents at two different visualization types, geospatial visualization and temporal visualization. We place time and geospace together since they are both traditionally dimensional types of data.

3.4.1. Time Oriented Surveys

Here we cover surveys with a primary emphasis on time-series data or visualization across multiple time-slices. The SoS summarizes two surveys. The first paper provides a classification for visualization of dynamic data. The second paper provides an understanding of different ways to review slices of data within a Space-Time Cube.

Cottam et al. review the impact of dynamic data on Information Visualization, and how this data change can influence a visualization's discernability. This is done via the creation of a taxonomy that categorizes dynamic visualization techniques. The paper defines dynamic visualizations as "visualizations that change over time" [CLW12].

Cottam et al. present the dimensions of their classification before presenting their technique matrix. Each cell of the technique matrix is reviewed, which gives an understanding of how the axes interact with each other. These are then grouped into higher-level identity groups, which represent how each classification cell would update to a new state. The paper ends by matching techniques to task scenarios.

The classification has three dimensions. The first dimension envelopes retinal (visual-based) categories, which include: (1) unchanged (immutable) scales, (2) a known scale, (3) extreme bins which categorize catch-all bins such as "100+", and (4) mutable scales which are dynamic scales. The second dimension identifies spatial categories such as fixed spatial dimension, mutable spatial dimension, new spatial elements (create), and create and deleted spatial elements. The third dimension depicts the higher-level identity groups: identity preserving changes, transitional changes, and immediate changes (the figure is provided in the supplementary material).

Cottam et al. consider that the taxonomy could be extended by looking at additional spatial categories, such as "delete but not create." The categories are not distributed evenly, so a study into why is also suggested.

Bach et al. survey a variety of temporal data visualization techniques and discuss how their operations can be used with space-time cubes in order to create a simple visualization from the 2D+time model. [BDA*14a] The paper discusses common static space-time cube operations which includes time-cutting, time flattening, time juxtaposition, space cutting, space flattening, sampling, and 3D rendering. Bach et al. then present the taxonomy of
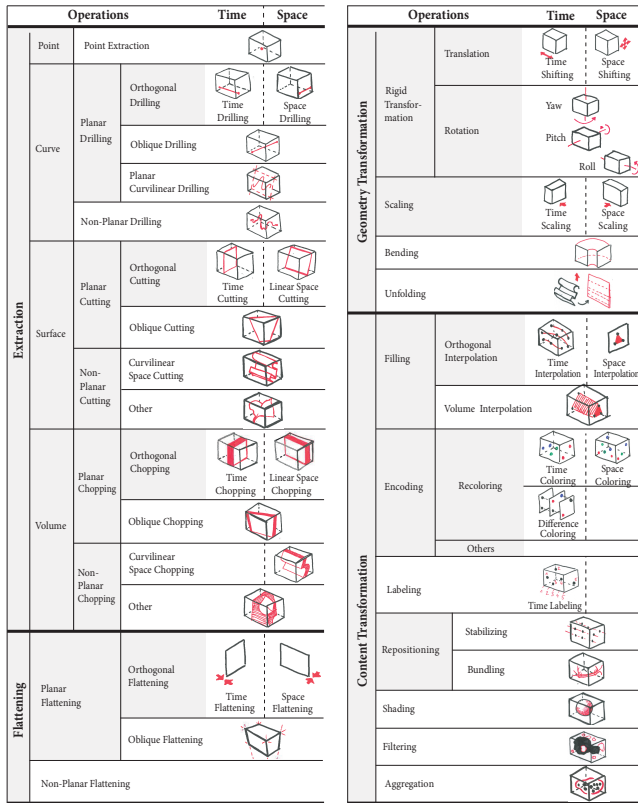
Figure 17: *Taxonomy of Space-Time cube operations created by Bach et al. [BDA*14a] Each operation gives a representation of how the operation may work. Bold font indicates complete operations. Gray shading indicates non-leaf nodes. Image courtesy of Bach et al. [BDA*14a]*

space-time cube operations that they have designed before giving the reader a selected sample of multi-operation systems.

The taxonomy (see Figure 17) presents a classification of elementary space-time cube operations such as drilling, cutting and chopping. These are broken down into sub-sections with schematic illustrations in order to enable the user to easily understand what effect the operation has. For example, the flattening section is broken down into planar flattening and non-planar flattening. Planar flattening is broken down into orthogonal flattening and oblique flattening.

There are many open research areas that are discussed within the paper. Some of these include interaction techniques such as focus+context to use with different operations, research into operations for extended data dimensions, and understanding which operation is most appropriate for a given task.

3.4.2. Geospatial Focused Surveys

This section focuses on surveys that examine geo-spatial visualization. The section provides an understanding, and classification, of tasks for cartograms, a view of geospatial traffic data, and another

review of the use of cartographic visualization in information visualization.

Chen et al. analyze various ways traffic data can be recorded as well as some different approaches that are brought forward to depict a combination of spatial, temporal, dimensional, and categorical visualization [CGW15]. They systematically review how traffic data is captured. This is divided into three unique categories: Location-Based, which records data as it appears at a fixed point, or sensor range; Activity-Based, which may record data when a specific event is started or finished; Device-Based, recording from a device which records information periodically such as a GPS. Each of these has their own unique benefits and uses. Each data type can be broken down into the four types mentioned previously: spatial; temporal; dimensional; and categorical. Chen et al. begin by looking at how the traffic data can be captured and discuss the different ways this data can be processed. The main focus of the survey is how the data is visualized. The paper provides examples of usual design for time, spatial properties, spatio-temporal data, and multi-property data.

To further improve the taxonomy of traffic data, Chen et al. structure the visual design types into the three main goals: Situation-Aware Exploration and Prediction; Pattern Discovery and Clustering; and Visual Monitoring of Traffic Situations. These goals are clearly explained and then presented with relevant examples [CGW15].

Chen et al. present analysis of situation-aware and immersive environments, as well as the design of huge spatio-temporal analysis of online or streamed data, as open challenges in the field. Visual Analysis of heterogeneous data, social transportation as an example, is another area that requires more focused research.

A cartogram is a type of visualization that aims to combine statistical and geographical information where areas are scaled dependent on statistical proportions. Nusrat and Kobourov study the effectiveness of cartograms as a visualization tool, as well as compare the effectiveness of different cartogram methods. The paper presents a set of cartographic visualization tasks and their application to information visualization [NK15]. Nusrat and Kobourov begin by providing an overview of cartograms, with related literature for those who want to expand their knowledge of the subject. They then present their design space for cartographic tasks and visual goals for cartogram usage. Both of these are coupled with clear examples of when they could be selected.

Nusrat and Kobourov classify tasks by reviewing two hierarchical dimensions, analytic tasks such as identification, location, sorts and clustering (see Figure 18). The second axis maps visualization goals which includes the goal of the visualization, how a task is carried out, features of the task, and the cardinality.

Nusrat et al. follow-up their previous survey with an extended version [NK16b]. They start by presenting a history of cartographic visualization, beginning with their origins in 1870. The paper examines the literature surveys related to cartograms, discussing what is presented in each, before discussing the different design-types of cartographic layouts as well as a task taxonomy extension of '*Task Taxonomy for Cartograms*' by Nusrat and Kobourov [NK15], their applications, and their effectiveness.

	Goals			Means				Characteristics		Cardinality		
	Query	Search	Extract	Map Relation	Data Relation	Navigation	Derive	Low Level	High-Level	Single	Multiple	All
Recognize	✓	✗	✗	✓	✗	✗	✗	✓	✗	✓	✗	✗
Detect Change	✓	✗	✗	✓	✗	✗	✗	✓	✗	✓	✗	✗
Compare	✓	✗	✗	✗	✓	✗	✗	✗	✓	✗	✓	✗
Find top-k	✗	✓	✗	✗	✓	✗	✗	✗	✓	✗	✗	✓
Filter	✗	✓	✗	✗	✓	✗	✗	✗	✓	✗	✗	✓
Cluster	✗	✓	✗	✗	✓	✗	✗	✗	✓	✗	✗	✓
Locate	✗	✓	✗	✗	✗	✓	✗	✓	✗	✗	✗	✓
Find Adjacency	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗	✓	✗
Summarize	✗	✗	✓	✗	✗	✗	✓	✗	✓	✗	✗	✓
Identify	✗	✗	✓	✗	✗	✗	✓	✓	✗	✓	✗	✗

Figure 18: *Task taxonomy for cartogram visualization courtesy of Nusrat and Kobourov [NK15].*

Figure 19: *Examples of multiple facet representations within visualization. Courtesy of Hadlak et al. [HSS15]*

They introduce the three major design dimensions of cartographic visualization: Statistical accuracy, geographical accuracy, and topological accuracy. In addition, cartograms are sub-divided into four different types which include contiguous, non-contiguous, Dorling, and rectangular cartograms.

Nusrat and Kobourov pose a multitude of areas for future cartogram research. Firstly, some of the design dimensions are under-utilized, which could facilitate a paper comparing usage with less used cartographic layouts. Cartograms normally only excel in one of the three design dimension, and a study to mitigate errors in the other two dimensions may allow for a better understanding of cartogram usage. Some other areas of interest are the mapping of multivariate data, memorability and recall within cartograms, uncertainty within cartograms, and 3D cartographic visualization.

3.5. Coordinated Multiple View (CMV) Surveys

Coordinated Multiple Views (CMVs) surveys focus on literature that examines the coordination or linkage between multiple views. This section summarizes two papers related to the subject. The first reviews the use of composite visualization and the second provides an understanding of multi-faceted graph visualization.

Javed and Elmqvist examine different Composite Visualization Views (CVVs), which are defined as *"a visual composition of two or more visual structures in the same view,"* and present their CVV design patterns created via a literature survey [JE12]. They start by discussing different CVVs, including their design patterns and existing formalisms. This is followed with an in-depth look at the different types of views (juxtaposed, integrated, superimposed, etc), with examples of each. They then present their design space and guidelines, as well as their 1-N classification table.

The authors look at different techniques and how the visualization's relations are classified. The grouping looks at the result of merging two visual designs, the composite relation (juxtaposed, integrated, superimposed, overloaded, nested), as well as what data-relation is created (item-item, item-group, etc).

Javed and Elmqvist suggest their design patterns are limited to literature reviewed and therefore work can be invested into extending their framework. The design pattern is also limited to only
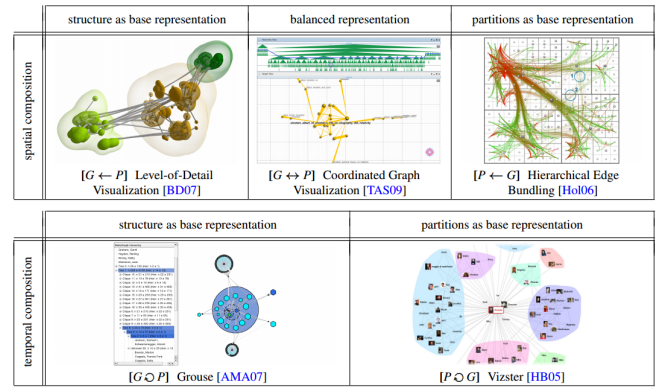
spatial relation and does not look at other composite visualization views, such as interaction or animation.

Many surveys focus on only a single additional facet in order to classify graphing techniques. Hadlek et al. aim to build on existing surveys in order to create a more in-depth observation of four common facets: partitions, attributes, time, and space. Each of these characteristics are discussed based on their relationship as well as examples of how these graphs can be represented depending on the hierarchy. Hadlek et al. focus on an output oriented perspective, and optimize facet selection by focusing on their composition. These compositions are given a representative visualization (seen in Figure 19) to discuss in detail in the content of the survey. Hadlek et al. analyse visual design of the graph structure with a single additional facet and graph structure with multiple additional facets. These are sub-divided for each common facet. This is followed by analyzing multiple instances of graph facets.

Each graph structure is split into five combinations. Whilst looking at the spatial composition, an example is given for structure as the base representation, partitions as the base representation, or a balanced representation. A temporal composition has an example for either structure as the base representation or partitions as the base representation.

The paper notes the exploration of geo-spatial graph visualization, by reviewing output or task taxonomy, has yet to be published. Hadlak et al. also point out that some of the facets discussed had very sparse usage such as temporal compositions. Finally, only four facets were examined but there are many other extensions such as provenance, uncertainty, heterogeneity, or text/annotations that have little-or-no exposure which could be a new thread of research to investigate.

3.6. Real-World and Applications

Many literature surveys focus on real-world scenarios and applications. This area covers a wide range of surveys including finance, health-care, security, systems, software visualization, and visualization frameworks.

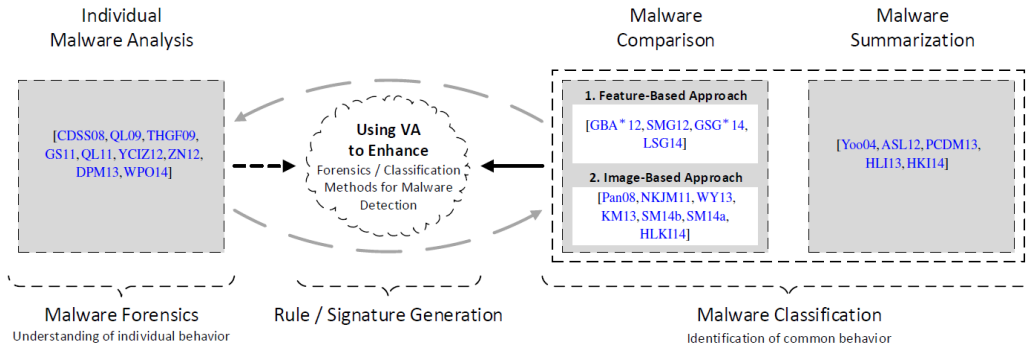**Figure 20:** *Ko et al.'s Categorization of surveyed papers via Keim's visualization techniques taxonomy [KCA*16, Kei02] .*

	Sum	From Industry
Std (2D)	16	3
Std (3D)	11	4
GT	12	6
Iconic	6	2
Dense Pixel	10	1
Stacked	7	3

3.6.1. Finance Focused Surveys

This section has a focus on survey papers related to finance visualization. The survey summarizes one paper focused on different sources of financial data and how they are visualized.

Ko et al. perform and present a study of visualization and visual analytics of financial data. Economy is an important field for any business which has led to financial data being a popular topic for visualization in industries. They aim to utilize existing papers in order to help researchers design better systems and understand new research fields within the area [KCA*16]. After discussing the survey scope, Ko et al. explore the types of data analyzed and some data sources for each. Some examples of this include stock data, transaction data, and fund data. The paper derives a classification of techniques and provides examples of these uses before examining the interaction methods and evaluation methods. This is followed by an evaluation of the papers.

Ko et al. examine multiple ways to classify the gathered information. The first classification tests the type of data used by each paper, which indicates a heavy focus on stock data. The second classification discusses papers based on automated techniques, such as K-means. The third looks at visualization techniques, based on Keim's technique taxonomy [Kei02] (see Figure 20). The fourth categorization considers interaction methods and the final organization examines evaluation methods.

They propose that there are nine more business domains which could have their visual analytics reviewed such as economic analysis, financial risk management, and portfolio management. They also discuss the lack of research into company performance with financial data and suggest this as an open field. An important area for research is with automated visualization techniques, which industry experts believe are important to facilitate a richer depth of information. The final point they discuss is the use of heterogeneous data to enable improved prediction models.

3.6.2. Security-based Literature Surveys

This section includes survey papers that present security systems. One features a focus on visualization systems for network security and another on malware analysis.

Shiravi et al. provide a comprehensive overview of network security visualization and present data sources for each. The paper also provides a taxonomy that includes literature across five use-cases [SSG12]. They present a table that provides potential data sources for security visualization before giving an in-depth view of five use-cases for reviewing network security and their related papers. The use-cases are host/server monitoring, internal/external monitoring, port activity, attack patterns, and routing behavior.

The taxonomy takes the form of a 2D table, sub-divided into five sections representing the five use-cases. The taxonomy reviews the type of visualization techniques and data source of each research paper. It also includes the number of citations a paper has to emphasize systems that have more references, as well as signifying whether the system is available online (the figure is provided in the supplementary material).

Shiravi et al. present future research topics including situation awareness in presenting information, user experience evaluation, scalablity, occlusion in network security, privacy preservation and novel ways to provide 3D imagery of the data.

Wagner et al. present a systematic overview and classification of malware visualization systems used for visual analysis. The field is gaining more and more interest due to the increasing threat of malware attacks on user systems. Malware is defined as *"any software that does something that causes harm to a user, computer or network"* [WFL*15]. They focus on malware systems for visual analysis. They first provide tools, discussed as *data providers*, and provide a comparison of their usage. They present their malware visualization taxonomy and categorize each data provider in a number of different classification groups.

Wagner et al. create three categories for malware visualization systems: (1) Individual malware analysis which enables the system to look at a single malware sample and learn about its individual behavior. (2) Malware comparison which facilitates comparison of a range of malware for viewing. (3) Malware summarisation that outlines the behavior of different malware samples (see Figure 21).

They find that malware visualization is cleanly partitioned between each classification category and believe more work needs to be done in creating a connection between categories. Using different systems could cause unnecessary overlap for users which could

Figure 21: *The 1D malware visualization taxonomy courtesy of Wagner et al. [WFL*15]*

be minimized with a system that could move through this categorization. Some other challenges in the field include the integration of more data sources, a stronger understanding of the requirements for malware visualization, enabling expert analysis and externalization, and an increased focus on the relation between analysis and visualization for malware.

3.6.3. Systems-based Surveys

This section focuses on Surveys that have an emphasis on classifying systems. We summarize one survey paper in this section which looks at performance visualization of large-scale systems.

Gao et al. present a review on papers designated for researching performance visualization on large-scale systems. Gao et al. define performance visualization as *'the use of graphical display techniques for the visual analysis of performance data'* [GZR*11]. The paper aims to shed light on open research areas, and increased discussion on the design of visual tools for these systems. The paper provides a brief background of performance visualization and how it functions.

Gao et al. classify performance visualization techniques using four main categories. Simple visual structures which include statistical charts with one or two variables, composed visual structures that include a combination of simple chart views, interactive visual structures featuring structures that provide a variety of user interactions, and focus+context which refers to visualizations with mapping that is automatically modified without the need of user interaction (the figure is provided in the supplementary material).

Some future work areas presented by Gao et al. include scalability, user studies, and the synthesis of high-level context with low-level detail.

3.6.4. Software Visualization Surveys

Software Visualization papers focus on visualizing aspects of software creation. Diehl defines the topic as follows: *'Software visualization encompasses the development and evaluation of methods for graphically representing different aspects of software, including its structure, its execution, and its evolution'* [Die07]. The SoS summarizes one recent survey paper in this section that focuses on the static aspects of software visualization.

Figure 22: *Caserta and Zendra present a table that classifies methods that visualise the static aspects of software and the associated literature [CZ11] .*

Caserta and Zendra categorize visualization techniques that represent the static aspects of software and it's evolution. The paper defines visualization of the static aspects of software as *'visualizing software as it is coded, and dealing with information that is valid for all possible executions of the software'* [CZ11]. Evolution of software adds a temporal dimension to the visualization of the static aspects of software. The paper provides a guide to papers that feature code-line-centered visualization, class-centered visualization, architecture visualization, and the visualization of software evolution. They provide examples of different types of visual design for each and how they may be applied.

The hierarchical, 1-N, classification presents the representation and visualization techniques used for each paper, and shows how they fit into their taxonomy (see Figure 22).

Caserta and Zendra propose that it would be beneficial to invest research into usability evaluation to find the most effective ways to visualize the static aspects of software. There is also limited research in navigation and interaction for 3D visualization within the field.

3.6.5. Surveys of Frameworks

This section examines the review of frameworks, which is defined as *'a basic structure underlying a concept'* [fra17]. The section covers two surveys: the first presents a design framework survey for bi-cluster visualizations and the second presents a framework for emphasis in visualizations.

Sun et al. provide a survey focused on bi-cluster visualization, design considerations, and applications. Bi-clusters *"provide a rich high-level abstraction that represents coordinated relationships between groups of entities of different types"* [SNR14]. The advantages and disadvantages are compared, and a five-level relationship is presented to assist in design options that support user-tasks. The paper describes the concept of bi-clustering and the five relationship levels of the bi-cluster visualization design framework. These include: entity level (single entity relationships), group level (entity group relationships), bi-cluster level (coordinated relationships), chain level (chained coordinated relationships), and the schema level (schema level relationships). The paper also examines four levels of interaction design: readability, navigation, parameter, and object level.

Sun et al. provide a summary of the five-level design framework for bi-cluster visualization. This incorporates the interaction design level, major tasks, design choice, and trade-offs.

They identify three challenges in the field. The first challenge is the creation of an example that implements design options across all five levels of the framework. The second challenge discusses traversal between the different levels. The final challenge suggests optimal layout of bi-cluster chains in visualization.

Hall et al. present a mathematical Framework for Information Visualization Emphasis (FIVE) by reviewing existing emphasis literature and frameworks [HPK*16]. Some examples of emphasis provided include highlighting regions of interest, animating data points, and altering the size of data points. They first present a language for emphasizing sub-sets of data, and present a table displaying how their framework compares to previous solutions (see Figure 23). The paper then discusses different types of emphasis effects and how they can be used such as position, color, motion and transparency. Finally, the paper discusses the opportunities to use FIVE and some future directions for research.

The frameworks are split into three categories. (1) *Magnification* - papers that describe magnification emphasis effects. (2) *Beyond magnification* - papers describing non-magnification emphasis effects. (3) *Data suppression* - papers that focus on the creation of emphasis effects through data suppression.

They provide four areas for future work: creating emphasis effects using under-explored visual variables and time variation, exploring alternative ways to vary data point prominence to create emphasis effects, providing a richer space of how to define and implement emphasis effects, and conducting empirical studies of emphasis effects.

Figure 23: *Hall et al. present a table used to classify previous emphasis frameworks to diagnose what types of emphasis are discussed, and compare them to FIVE [HPK*16] .*

3.7. Overview Surveys

Surveys that attempt to cover information visualization as a whole are presented here. This section summarizes two papers that review recent advancements in information visualization.

We summarize one survey that reviews the use of interactive lenses in visualization. Tominski et al. aim to analyze the use of interactive lenses in the context of visualization by reviewing the different techniques used to create lenses, whilst also helping researchers identify when to use interactive lenses. They discuss applicable data types such as geo-spatial data, why incorporating lenses is beneficial to the user experience, and some important techniques to be aware of if you are interested in this type of visualization. An interactive lens is defined as *'to provide an on demand alternative visual representation of the data underlying a local area of the screen'* [TGK*14]. After introducing the model of the interactive lens as well as the discussion of lens usage, they segment the

Techniques	Data							Tasks						
	Temporal	Geospatial	Flow	Volume	Multivariate	Graph	Document	Select	Explore	Reconfigure	Encode	Abstract & Elaborate	Filter	Connect
[SB92, SB94] Fisheye Views		●			●							●		
[RM93] Document Lens							●					●		
[CMS94] MagicSphere				●				●			●	●	●	
[RC94] Table Lens					●							●		
[VCWP96] 3D Magic Lenses				●								●	●	
* [FG98] Lenses for Flow Visualization			●								●	●		
[FP99] Excentric Labeling					●						●	●		●
[SHER99] Interactive 3D Lens			●	●	●	●					●	●	●	
[LHJ01] Volume Magnification Lens				●								●		
[SFR01] Time Lens	●								●					
[BCPS02] Fuzzy Lens, Base-Pair Lens, Ring Lens						●				●	●	●		
[MTHG03] 3D Flow Lens			●								●	●		
* [WCG03] EdgeLens						●					●			
* [vHvW04] Graph Abstraction Lens						●						●		
[RHS05] Magic Lenses in Geo-Environments		●										●	●	
* [EBD05, ED06b, ED06a] Sampling Lens					●								●	
[RLE05] Temporal Magic Lens	●								●					
* [WZMK05] The Magic Volume Lens				●								●		
[TAvHS06] Local Edge Lens						●							●	
[KSW06] ClearView				●								●	●	
[TGBD08] 3D Generalization Lenses		●										●		
* [TAS09] Layout Lens						●					●			●
[BRL09] Enhanced Excentric Labeling					●						●	●		●
* [MCH*09] Bring &Go						●					●			●
[ACP10] High-Precision Magnification Lenses	●	●	●	●	●	●	●	●				●		
[JDK10] Network Lens						●					●			
* [KCJ*10] Detail Lenses for Routes		●										●		●
[Kin10] SignalLens	●											●		
* [SNDC10] PushLens						●				●				
* [STSD10] Tangible Views		●			●	●		●	●	●	●	●	●	●
* [EDF11] Color Lens	●	●	●	●	●						●			
[GNBP11] FlowLens			●						●		●		●	
[HLTE11] SemLens					●						●		●	
[HTE11] MoleView		●		●		●					●			
[LWG11] Facet Lens							●					●	●	
[PBKE11] EdgeAnalyser					●	●					●	●		
[ZCB11] MagicAnalytics Lens	●													●
* [ZCPB11] ChronoLenses	●												●	●
[TSAA12] Time Lens		●									●	●		
[PPCP12] JellyLens		●										●		
* [KTW*13] TrajectoryLenses		●											●	
[PPCP13] Gimlenses				●								●	●	
[UvK13] Magic Lenses for Hypergraphs						●						●		
* [CC13] Lens for Querying Documents							●						●	●
[AACP14] RouteLens		●										●		
[BHR14] PhysicLenses		●										●		
* [MW14] Bubble Lens					●			●				●		
[DMC14] VectorLens	●					●		●					●	
[KRD14] Multi-touch graph Lenses						●					●		●	●
[DSA15] 3DArcLens		●				●					●		●	

Figure 24: *Lens Techniques categorised according to data types and task. Courtesy of Tominski et al. [TGK*16]*

research into interaction types. This includes examination of mouse and keyboard interaction, touch and multi-touch interaction, tangible interaction, tangible views and spatial interaction, gaze-based interaction and head tracking. [TGK*14]

The taxonomy of Tominski et al's survey examines both the data types that each visualization technique demonstrates (temporal, geospatial, flow, etc.), as well as the task that is achieved using it (across a total of 43 papers). This taxonomy enables discussion of possible future work such as the use of lenses for multi-user work (see Figure 24).

After reviewing some survey notes, they suggest that the need for more dynamic lenses with flexibility is an important design note as well as useful for a more varied use of functionality. Although mentioned in the survey, Tominski et al. would like to further investigate the development of multi-user or shared lenses, due to the growth in high-resolution and interactive displays. Finally, the idea of lens tool kits is an important focus area. Lenses are globally recognized visualization types but are in low use due to how interwoven they are with visualizations. With big data becoming a strong focus, this is something that needs further development. Tominski et al. also provide an extended version of the classification (see Figure 24) [TGK*16].

Liu et al. create a comprehensive study on the domain of information visualization. The aim of the paper is to derive an organization of the field, describing features, goals and state-of-the-art approaches for each category [LCWL14]. The paper opens by examining the visualization pipeline and classification schemes. They proceed to present their 1-N taxonomy of the literature landscape within recent years, and present each topic whilst giving examples of papers in the related area. This continues on to communicating some technical challenges.

Liu et al. break their taxonomy down into four main categories: empirical methodologies, interactions, frameworks, and applications. These categories are sub-divided into sub-categories (the figure is provided in the supplementary material). The classification has a lot of overlap with our organization.

Liu et al. describe an abundance of open research areas including usability, scalability, heterogeneous data, real-time visualization, and uncertainty.

4. Future Work

At the end of most survey papers, it is common for the authors to discuss future areas of work that are discovered over the course of the survey. These challenges and research directions have been compiled using SurVis, an online literature browser [BKW16, McN17]. This enables us to find a number of future research areas within Information Visualization.

The research topics found to have a high frequency (over 15%) are listed in this section. Each topic provides a percentage of recently *summarized* papers (40) that address a challenge found within the surveys.

1. **Evaluation (50%):** The most frequent topic discussed for open-research directions is visualization evaluation. This includes user studies, qualitative studies, quantitative studies, and longitudinal studies. There is a strong focus on perceptual surveys, that would clearly show where studies are limited within each topic.
2. **Missing Scenarios (40%):** Many topics point out vacant research scenarios within their classification tables. All of these topics are subject specific, and can be viewed using our literature browser [McN17].
3. **Scalability (35%):** Scalability is still a very important trend in visualization design at the moment. This includes large datasets, the ability to move between views clearly, reducing clutter, and improving visual understanding of complex views.
4. **Interaction Techniques (28%):** The use of interaction techniques continues to be an essential part of visualization design in the recent decade. Research that discusses ways to filter or manipulate visualization seem to be an important topic in the upcoming future.
5. **Extended Data Dimensions (25%):** Many surveys suggest that new data dimensions can be explored in the field. This differs from Item 2 by discussing the extension of a taxonomy.
6. **Tools (20%):** A number of papers describe the need for new tools for their domain which includes a need for both generic tools or frameworks to enable users to quickly use techniques

LITERATURE:	EVALUATION	MISSING SCENARIOS	SCALABILITY	INTERACTION TECHNIQUES	EXTENDED DATA DIMENSIONS	TASKS	TOOLS	FUZZYNESS & UNCERTAINTY	TIME-VARYING DATA	DESIGN GUIDELINES	HETEROGENEITY
[APS14]		X	X						X		
[AMA*14b]	X	X		X			X	X	X		
[BDA*14a]	X			X		X					
[BBDW14b]	X	X	X	X							
[BBR*16a]	X						X				
[BTK11]	X		X								
[BKC*13]							X				
[CZ11]	X			X							
[CGW15]							X				X
[CLW12]	X					X					
[EF10]					X						
[FHKM16a]	X	X	X	X							
[FIBK16]	X	X									
[GZR*11]	X		X								
[HSS15]		X						X			X
[HPK*16]	X										
[HW13]						X					
[IGJ*14a]	X		X								
[JFCS15]	X	X						X		X	
[JE12]				X					X		
[JF16]	X										
[KKC14]		X									
[KKC15]							X				
[KKCG15]		X									
[KCA*16]	X	X				X					
[KK15a]	X										
[LCWL14]	X	X				X					
[LMW*15a]				X				X			
[NK15]											
[NK16b]		X	X		X			X			
[SHS11]										X	
[STMT12]											
[SSG12]		X		X							
[SNR14]							X				
[TGK*14]				X							
[VBW15a]	X		X	X	X				X		
[VLKS*11]			X	X				X			
[WFL*15]			X	X	X		X				
[WSJ*14]	X	X	X								
[ZH15]	X				X					X	

Table 6: *The tables shows a breakdown of research directions discussed for each primary survey paper (highlighted green in Table 3). The directions displayed represent research areas that are discussed in more than one survey. This table corresponds to the 1-N classification example shown in Figure 4 (B).*

over multiple pieces of software, or context sensitive tools to allow specific test or an understanding within the field.

7. **Tasks (20%):** 20% of papers suggest that a stronger focus needs to be placed on exploring different tasks within each topic. This would enable researchers more understanding when a technique is appropriate, what approaches would produce the best output, and what can be gathered from visualization design choices.

8. **Fuzziness and Uncertainty (18%):** Fuzziness and Uncertainty are a growing topic within Information Visualization and the results of our survey shows that this is a positive step. Visualization aims to represent clear findings and it is therefore essential that and uncertainty is represented. This open research topic was mainly suggested for text-focused surveys and multivariate surveys. Although there are uncertainty surveys published, only one of these fulfills one of our topics [DCK12], while the other two are SciVis papers [PWL97, BAOL12], so this research area is still open for research.

Our paper presents some interesting findings. Graph surveys and text surveys have a large quantity of survey work in recent years. This enables quite a large overview of the current landscape of the topics but will also make it difficult to justify the creation of new surveys in the field. There is a large quantity of user studies across many fields yet there is little evidence of their use. Perceptual surveys are a great tool to analyze and document user-studies within a field, with the two papers summarized in this category giving a greater understanding of their benefit and contribution to the field [JF16, FIBK16] (see Table 6 and figure 25). We also look how these papers are used in the field. We found that over 50% of survey topics show a positive trend between surveys before and after 2010 (see Figure 26).

5. Limitations

There are some important considerations when looking at the implementation of the Survey of Surveys. The SoS uses natural topic clusters to classify literature in the field of information visualization. This means that topics reviewed are naturally biased towards surveys that have been published. A second limitation is that open-research is only based on what is discussed within each survey and this does not necessarily fully represent the current landscape of the domain, as there is a possibility that papers have been presented that fulfill open research directions between the publication of the survey and the publication of the SoS. This means that the older the paper, the more uncertain we are that the open research areas have matured.

6. Conclusion

The SoS contributes a quantum step forward in literature surveys. We present a novel classification of survey papers that enables the reader to find recently published literature among a wide variety of topics. The classification also enables users to easily spot areas of open-research for survey publication, as well as an understanding of broad open research topics in the field of information visualization. The paper provides a basic systematization of classification tables among the existing survey literature. It provides a valuable starting point for both newcomers and experienced researchers in
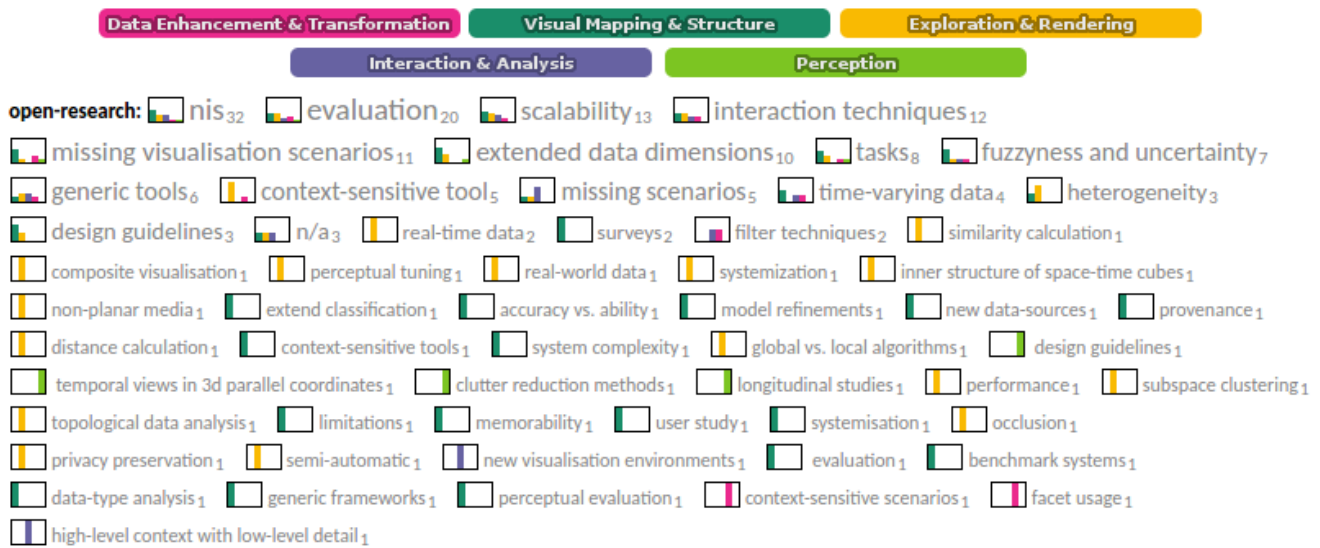
Figure 25: *Open research keywords collected across all recent, reviewed survey papers (2010-2017). 'nis' refers to papers not summarised. 'n/a' refers to summarised papers with no explicit open-research. Collected Keywords are reviewed using the SurVis Literature Browser [BKW16, McN17].*

visualization. We also believe it provides a valuable resource to readers outside of the information visualization and visual analytics communities.

7. Acknowledgments

References

[AA06] ANDRIENKO N., ANDRIENKO G.: *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer Science & Business Media, 2006. 16

[AAM*17] ALHARBI N., ALHARBI M., MARTINEZ X., KRONE M., ROSE A., BAADEN M., LARAMEE R. S., CHAVENT M.: Molecular visualization of computational biology data: A survey of surveys. In *EuroVis 2017 - Short Papers* (2017), The Eurographics Association. forthcoming. 4

[acm16] ACM digital library. http://dl.acm.org/, 2016. Accessed: 2016-9-26. 2

[AdOP12] ALENCAR A. B., DE OLIVEIRA M. C. F., PAULOVICH F. V.: Seeing beyond reading: a survey on visual text analytics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2*, 6 (2012), 476–492. 5

[AMA*14a] ALSALLAKH B., MICALLEF L., AIGNER W., HAUSER H., MIKSCH S., RODGERS P.: Visualizing sets and set-typed data. http://www.cvast.tuwien.ac.at/SetViz, 2014. [Online: accessed 05-January-2016]. 9

[AMA*14b] ALSALLAKH B., MICALLEF L., AIGNER W., HAUSER H., MIKSCH S., RODGERS P.: Visualizing sets and set-typed data: State-of-the-art and future challenges. In *Eurographics conference on Visualization (EuroVis)–State of The Art Reports* (2014), The Eurographics Association, pp. 1–21. 5, 6, 11, 24

[AP13] ARCHAMBAULT D., PURCHASE H. C.: The 'map' in the mental map: Experimental results in dynamic graph drawing. *International Journal of Human-Computer Studies 71*, 11 (2013), 1044–1055. 5

[APS14] AHN J.-W., PLAISANT C., SHNEIDERMAN B.: A task taxonomy for network evolution analysis. *IEEE transactions on visualization and computer graphics 20*, 3 (2014), 365–376. 5, 6, 16, 17, 24

[BAOL12] BRODLIE K., ALLENDES OSORIO R., LOPES A.: *A Review of Uncertainty in Data Visualization*. Springer London, London, 2012, pp. 81–109. 24

[BBDW14a] BECK F., BURCH M., DIEHL S., WEISKOPF D.: Dynamic graph visualisation browser. http://dynamicgraphs.fbeck.com/, 2014. [Online: accessed 05-January-2016]. 9

[BBDW14b] BECK F., BURCH M., DIEHL S., WEISKOPF D.: The state of the art in visualizing dynamic graphs. In *EuroVis - State of the Art Reports* (2014), The Eurographics Association. 5, 6, 14, 24

[BBR*16a] BEHRISCH M., BACH B., RICHE N. H., SCHRECK T., FEKETE J.-D.: Matrix Reordering Methods for Table and Network Visualization. *Computer Graphics Forum 35*, 3 (2016), 693–716. 5, 6, 17, 24

[BBR*16b] BEHRISCH M., BACH B., RICHE N. H., SCHRECK T., FEKETE J.-D.: Matrix reordering survey. http://matrixreordering.dbvis.de/, 2016. [Online: accessed 05-January-2016]. 9
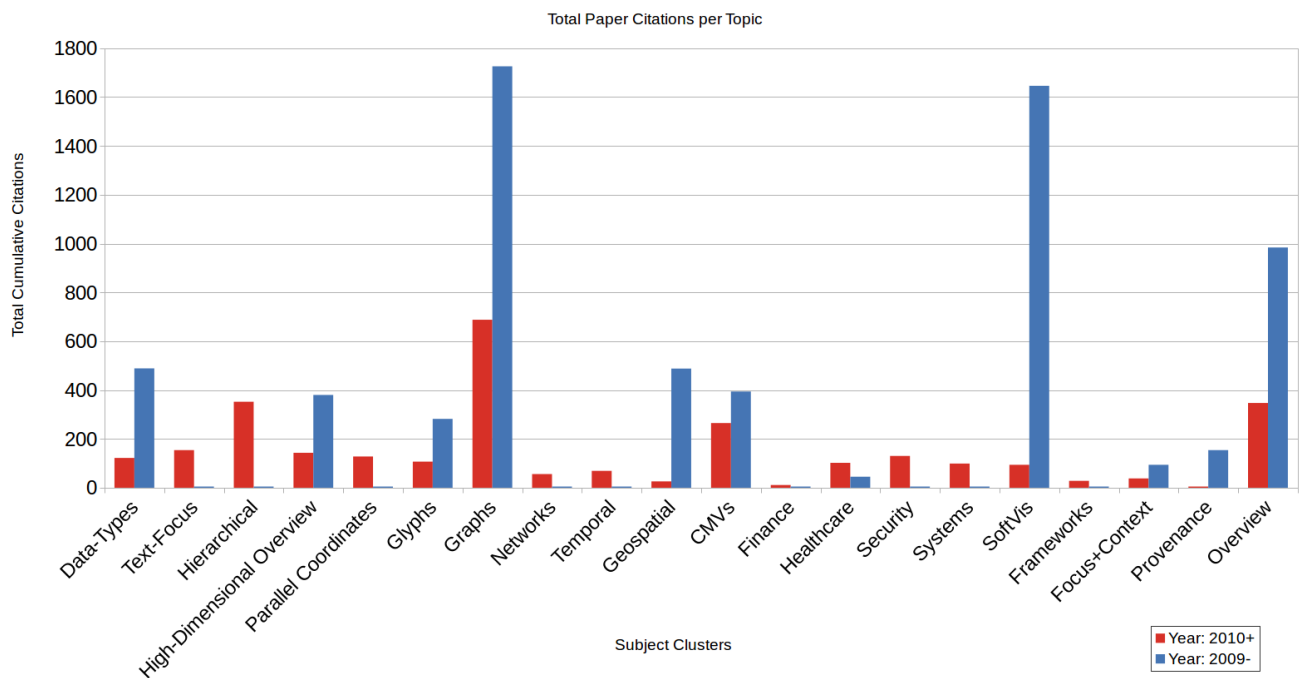
Figure 26: *The graph provides a visualization of paper citations, ordered by topic, highlighting the difference between surveys published before and after 2010. The graph shows that 11 of the 20 show a positive trend, 10 of which are considered new subject clusters.*

[BCS96] BUJA A., COOK D., SWAYNE D. F.: Interactive high-dimensional data visualization. *Journal of computational and graphical statistics 5*, 1 (1996), 78–99. 5

[BDA*14a] BACH B., DRAGICEVIC P., ARCHAMBAULT D., HURTER C., CARPENDALE S.: A review of temporal data visualizations based on space-time cube operations. In *EuroVis 2014 - State of the Art Reports* (2014), The Eurographics Association, pp. 1–21. 5, 6, 17, 18, 24

[BDA*14b] BACH B., DRAGICEVIC P., ARCHAMBAULT D., HURTER C., CARPENDALE S.: A review of temporal visualizations based on generalized space-time cube operations. http://spacetimecubevis.com/, 2014. [Online: accessed 05-January-2016]. 9

[Ber83] BERTIN J.: *Semiology of graphics: diagrams, networks, maps*. University of Wisconsin press, 1983. 16

[BKC*13] BORGO R., KEHRER J., CHUNG D. H., MAGUIRE E., LARAMEE R. S., HAUSER H., WARD M., CHEN" M.: "glyph-based visualization: Foundations, design guidelines, techniques and applications". *"Eurographics State of the Art Reports"* (may "2013"), "39–63". 5, 6, 7, 13, 24

[BKR*14] BLASCHECK T., KURZHALS K., RASCHKE M., BURCH M., WEISKOPF D., ERTL T.: State-of-the-art of visualization for eye tracking data. In *EuroVis - State of the Art Reports* (2014), The Eurographics Association. 4

[BKW16] BECK F., KOCH S., WEISKOPF D.: Visual analysis and dissemination of scientific literature collections with SurVis. vol. 22 of *TVCG*, IEEE, pp. 180–189. 2, 3, 14, 23, 25

[Bli98] BLINN J. F.: Ten more unsolved problems in computer graphics. *IEEE Computer Graphics and Applications 18*, 5 (Sep 1998), 86–89. 1, 2

[BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2376–2385. 5

[BNW*16] BLUMENSTEIN K., NIEDERER C., WAGNER M., SCHMIEDL G., RIND A., AIGNER W.: Evaluating information visualization on mobile devices: Gaps and challenges in the empirical evaluation design space. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (New York, NY, USA, 2016), BELIV '16, ACM, pp. 125–132. 5

[Boy12] BOYLES N.: Closing in on close reading. In *On Developing Readers: Readings from Educational Leadership*. 2012, ch. 9, pp. 89–99. 9

[BTK11] BERTINI E., TATU A., KEIM D.: Quality metrics in high-dimensional data visualization: An overview and systematization. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2203–2212. 5, 6, 11, 24

[CGW15] CHEN W., GUO F., WANG F.-Y.: A survey of traffic data visualization. *IEEE Trans. Intelligent Transportation Systems 16*, 6 (2015), 2970–2984. 5, 6, 18, 24

[Chi00] CHI E. H.-H.: A taxonomy of visualization techniques using the data state reference model. In *Information Visualization, 2000. InfoVis 2000. IEEE Symposium on* (2000), IEEE, pp. 69–75. 5

[CLW12] COTTAM J. A., LUMSDAINE A., WEAVER C.: Watch this: A taxonomy for dynamic data visualization. In *Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on* (2012), IEEE, pp. 193–202. 5, 6, 17, 24

[CMS99] CARD S. K., MACKINLAY J. D., SHNEIDERMAN B.: *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999. 2, 3, 11

[CZ11] CASERTA P., ZENDRA O.: Visualization of the static aspects of software: A survey. *IEEE transactions on visualization and computer graphics 17*, 7 (2011), 913–933. 5, 6, 21, 24

[DCK12] DASGUPTA A., CHEN M., KOSARA R.: Conceptualizing visual uncertainty in parallel coordinates. In *Computer Graphics Forum* (2012), vol. 31, Wiley Online Library, pp. 1015–1024. 5, 24

[DDSV16] DATONDJI S. R. E., DUPUIS Y., SUBIRATS P., VASSEUR P.: A survey of vision-based traffic monitoring of road intersections. *IEEE transactions on intelligent transportation systems 17*, 10 (2016), 2681–2698. 4

[Die07] DIEHL S.: *Software visualization: visualizing the structure, behaviour, and evolution of software*. Springer Science & Business Media, 2007. 21

[DLR09] DRAPER G. M., LIVNAT Y., RIESENFELD R. F.: A survey of radial methods for information visualization. *IEEE transactions on visualization and computer graphics 15*, 5 (2009), 759–776. 5

[DML14] DUMAS M., MCGUFFIN M. J., LEMIEUX V. L.: Finance-vis.net: A visual survey of financial data visualizations. http://financevis.net/, 2014. [Online: accessed 05-January-2016]. 9

[dOL03] DE OLIVEIRA M. F., LEVKOWITZ H.: From visual data exploration to visual data mining: a survey. *IEEE Transactions on Visualization and Computer Graphics 9*, 3 (2003), 378–394. 5

[DWL08] DODGE S., WEIBEL R., LAUTENSCHÜTZ A.-K.: Towards a taxonomy of movement patterns. *Information visualization 7*, 3-4 (2008), 240–252. 5

[ED07] ELLIS G., DIX A.: A taxonomy of clutter reduction for information visualisation. *IEEE transactions on visualization and computer graphics 13*, 6 (2007), 1216–1223. 5

[EF10] ELMQVIST N., FEKETE J. D.: Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines. *IEEE Transactions on Visualization and Computer Graphics 16*, 3 (May 2010), 439–454. 5, 6, 10, 24

[ELC*12] EDMUNDS M., LARAMEE R. S., CHEN G., MAX N., ZHANG E., WARE C.: Surface-based flow visualization. *Computers & Graphics 36*, 8 (2012), 974–990. 4

[FHKM16a] FEDERICO P., HEIMERL F., KOCH S., MIKSCH S.: A survey on visual approaches for analyzing scientific literature and patents. *IEEE Transactions on Visualization and Computer Graphics* (2016). 5, 6, 10, 24

[FHKM16b] FEDERICO P., HEIMERL F., KOCH S., MIKSCH S.: A survey on visual approaches for analyzing scientific literature and patents. http://ieg.ifs.tuwien.ac.at/~federico/LiPatVis/, 2016. [Online: accessed 05-January-2016]. 9

[FIBK16] FUCHS J., ISENBERG P., BEZERIANOS A., KEIM D.: A systematic review of experimental studies on data glyphs. *IEEE Transactions on Visualization and Computer Graphics* (2016). forthcoming. 5, 6, 13, 24

[FLVW16] FLOOD M. D., LEMIEUX V. L., VARGA M., WONG B. W.: The application of visual analytics to financial stability monitoring. *Journal of Financial Stability 27* (2016), 180–197. 5

[fra17] *English Oxford Living Dictionaries*. Oxford University Press, Jan 2017. Defined: 'Framework', https://en.oxforddictionaries.com/definition/framework. 22

[GAW*11] GLEICHER M., ALBERS D., WALKER R., JUSUFI I., HANSEN C. D., ROBERTS J. C.: Visual comparison for information visualization. *Information Visualization 10*, 4 (2011), 289–309. 5

[GBTS13] GRAMMEL L., BENNETT C., TORY M., STOREY M.-A.: A survey of visualization construction user interfaces. *EuroVis-Short Papers* (2013), 19–23. 5

[GG13] GHOSH S. K., GOSWAMI P. P.: Unsolved problems in visibility graphs of points, segments, and polygons. *ACM Computing Surveys (CSUR) 46*, 2 (2013), 22. 4

[GZ09] GOTZ D., ZHOU M. X.: Characterizing users' visual analytic activity for insight provenance. *Information Visualization 8*, 1 (2009), 42–55. 5

[GZL*14] GAN Q., ZHU M., LI M., LIANG T., CAO Y., ZHOU B.: Document visualization: an overview of current research. *Wiley Interdisciplinary Reviews: Computational Statistics 6*, 1 (2014), 19–36. 5

[GZR*11] GAO Q., ZHANG X., RAU P.-L. P., MACIEJEWSKI A. A., SIEGEL H. J.: Performance visualization for large-scale computing systems: A literature review. In *Proceedings of the 14th International Conference on Human-computer Interaction: Design and Development Approaches* (2011), HCII'11, Springer-Verlag, pp. 450–460. 5, 6, 21, 24

[HGEF07] HENRY N., GOODELL H., ELMQVIST N., FEKETE J.-D.: 20 years of four hci conferences: A visual exploration. *International Journal of Human-Computer Interaction 23*, 3 (2007), 239–285. 4

[HMM00] HERMAN I., MELANCON G., MARSHALL M. S.: Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics 6*, 1 (2000), 24–43. 5

[HPK*16] HALL K. W., PERIN C., KUSALIK P., GUTWIN C., CARPENDALE S.: Formalizing Emphasis in Information Visualization. *Computer Graphics Forum 35*, 3 (2016), 717–737. 5, 6, 22, 24

[HSS15] HADLAK S., SCHUMANN H., SCHULZ H.-J.: A survey of multi-faceted graph visualization. In *Eurographics Conference on Visualization (EuroVis)* (2015), vol. 33 of *VGTC*, The Eurographics Association, pp. 1–20. 5, 6, 19, 24

[HW13] HEINRICH J., WEISKOPF D.: State of the art of parallel coordinates. *STAR Proceedings of Eurographics 2013* (2013), 95–116. 5, 6, 7, 12, 24

[iee16] IEEE Xplore. http://ieeexplore.ieee.org/Xplore/home.jsp, 2016. Accessed: 2016-9-26. 2

[IGJ*14a] ISAACS K. E., GIMÉNEZ A., JUSUFI I., GAMBLIN T., BHATELE A., SCHULZ M., HAMANN B., BREMER P.-T.: State of the art of performance visualization. *EuroVis - State of the Art Reports* (2014). 5, 6, 7, 8, 24

[IGJ*14b] ISAACS K. E., GIMÃL'NEZ A., JUSUFI I., GAMBLIN T., BHATELE A., SCHULZ M., HAMANN B., BREMER P.-T.: State of the art of performance visualization. http://cgi.cs.arizona.edu/~kisaacs/STAR/, 2014. [Online: accessed 05-January-2016]. 9

[IHK*17] ISENBERG P., HEIMERL F., KOCH S., ISENBERG T., XU P., STOLPER C. D., SEDLMAIR M. M., CHEN J., MÖLLER T., STASKO J.: vispubdata.org: A Metadata Collection about IEEE Visualization (VIS) Publications. *IEEE Transactions on Visualization and Computer Graphics 23* (2017). To appear. 2, 4

[IIC*13] ISENBERG T., ISENBERG P., CHEN J., SEDLMAIR M., MOLLER T.: A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2818–2827. 5

[IIS*17] ISENBERG P., ISENBERG T., SEDLMAIR M., CHEN J., MÖLLER T.: Visualization as seen through its research paper keywords. vol. 23, pp. 771–780. 4

[JE12] JAVED W., ELMQVIST N.: Exploring the design space of composite visualization. In *Visualization Symposium (PacificVis), 2012 IEEE Pacific* (2012), IEEE, pp. 1–8. 5, 6, 19, 24

[JF16] JOHANSSON J., FORSELL C.: Evaluation of parallel coordinates: Overview, categorization and guidelines for future research. *IEEE transactions on visualization and computer graphics 22*, 1 (2016), 579–588. 5, 6, 12, 24

[JFCS15] JÄNICKE S., FRANZINI G., CHEEMA M. F., SCHEUERMANN G.: On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In *Eurographics Conference on Visualization (EuroVis) - STARs* (2015), pp. 83–103. 5, 6, 9, 10, 24

[JFCS16] JÄNICKE S., FRANZINI G., CHEEMA M., SCHEUERMANN G.: Visual text analysis in digital humanities. In *Computer Graphics Forum* (2016), Wiley Online Library. 5, 9

[KCA*16] KO S., CHO I., AFZAL S., YAU C., CHAE J., MALIK A., BECK K., JANG Y., RIBARSKY W., EBERT D. S.: A survey on visual analysis approaches for financial data. *Computer Graphics Forum 30*, 3 (2016), 599 – 617. 5, 6, 20, 24

[Kei02] KEIM D. A.: Information visualization and visual data mining. *IEEE transactions on Visualization and Computer Graphics 8*, 1 (2002), pp. 1–8. 20

[KHG03] KOSARA R., HAUSER H., GRESH D. L.: An interaction view on information visualization. *State-of-the-Art Report. Proceedings of EUROGRAPHICS* (2003), 123–137. 5

[KK15a] KUCHER K., KERREN A.: Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific Visualization Symposium (PacificVis)* (2015), IEEE, pp. 117–121. 5, 6, 9, 24

[KK15b] KUCHER K., KERREN A.: Text visualization techniques: Taxonomy, visual survey, and community insights. http://textvis. lnu.se/, 2015. [Online: accessed 05-January-2016]. 9

[KKC14] KERRACHER N., KENNEDY J., CHALMERS K.: The design space of temporal graph visualisation. In *EuroVis- Short Papers* (2014), no. 3, The Eurographics Association, pp. 7–11. 5, 6, 14, 16, 24

[KKC15] KERRACHER N., KENNEDY J., CHALMERS K.: A task taxonomy for temporal graph visualisation. *Visualization and Computer Graphics, IEEE Transactions on 21*, 10 (2015), 1160–1172. 5, 6, 16, 24

[KKCG15] KERRACHER N., KENNEDY J., CHALMERS K., GRAHAM M.: Visual techniques to support exploratory analysis of temporal graph data. In *Eurographics Conference on Visualization (EuroVis) - Short Papers* (2015), The Eurographics Association, pp. 1–21. 5, 6, 16, 24

[KM02] KOSARA R., MIKSCH S.: Visualization methods for data analysis and planning in medical applications. *International journal of medical informatics 68*, 1 (2002), 141–153. 5

[KM07] KIENLE H. M., MULLER H. A.: Requirements of software visualization tools: A literature survey. In *2007 4th IEEE International Workshop on Visualizing Software for Understanding and Analysis* (2007), IEEE, pp. 2–9. 5

[KMSZ06] KEIM D. A., MANSMANN F., SCHNEIDEWIND J., ZIEGLER H.: Challenges in visual data analysis. *Tenth International Conference on Information Visualisation* (2006), 9–16. 1

[Kob13] KOBOUROV S. G.: Force-directed drawing algorithms. In *Handbook of Graph Drawing and Visualisation*. 2013, ch. 12, pp. 383–408. 5

[LA94] LEUNG Y. K., APPERLEY M. D.: A review and taxonomy of distortion-oriented presentation techniques. *ACM Transactions on Computer-Human Interaction (TOCHI) 1*, 2 (1994), 126–160. 5

[LBI*12] LAM H., BERTINI E., ISENBERG P., PLAISANT C., CARPENDALE S.: Empirical studies in information visualization: Seven scenarios. *IEEE transactions on visualization and computer graphics 18*, 9 (2012), 1520–1536. 5

[LCM*16] LU J., CHEN W., MA Y., KE J., LI Z., ZHANG F., MACIEJEWSKI R.: Recent progress and trends in predictive visual analytics. *Frontiers of Computer Science* (2016), 1–16. 5

[LCWL14] LIU S., CUI W., WU Y., LIU M.: A survey on information visualization: Recent advances and challenges. *The Visual Computer 30*, 12 (2014), 1373–1393. 5, 6, 23, 24

[LH10] LIANG J., HUANG M. L.: Highlighting in information visualization: A survey. In *2010 14th International Conference Information Visualisation* (2010), IEEE, pp. 79–85. 5

[LK07] LARAMEE R. S., KOSARA R.: Challenges and unsolved problems (in human-centered visualization). In *Human-centered visualization environments*. Springer, 2007, pp. 231–254. 4

[LLC*12] LIPŞA D. R., LARAMEE R. S., COX S. J., ROBERTS J. C., WALKER R., BORKIN M. A., PFISTER H.: Visualization for the physical sciences. In *Computer graphics forum* (2012), vol. 31, Wiley Online Library, pp. 2317–2347. 4

[LMW*15a] LIU S., MALJOVEC D., WANG B., BREMER P.-T., PASCUCCI V.: Visualizing high-dimensional data: Advances in the past decade. In *Eurographics Conference on Visualization (EuroVis) - STARs* (2015), VGTC, The Eurographics Association, pp. 21–40. 5, 6, 12, 24

[LMW*15b] LIU S., MALJOVEC D., WANG B., BREMER P.-T., PASCUCCI V.: Visualizing high-dimensional data: Advances in the past decade. http://www.sci.utah.edu/~shusenl/ highDimSurvey/website/, 2015. [Online: accessed 05-January-2016]. 9

[McN17] MCNABB L.: Sos literature browser. http://sos. swansea.ac.uk/, 2017. [Online: accessed 06-April-2017]. 3, 9, 23, 25

[MIK*16] MATTILA A.-L., IHANTOLA P., KILAMO T., LUOTO A., NURMINEN M., VÄÄTÄJÄ H.: Software visualization today: Systematic literature review. In *Proceedings of the 20th International Academic Mindtrek Conference* (New York, NY, USA, 2016), AcademicMindtrek '16, ACM, pp. 262–271. 5

[Mor13] MORETTI F.: *Distant reading*. Verso Books, 2013. 9

[Mye86] MYERS B. A.: Visual programming, programming by example, and program visualization: a taxonomy. In *ACM SIGCHI Bulletin* (1986), vol. 17, ACM, pp. 59–66. 5

[Mye90] MYERS B. A.: Taxonomies of visual programming and program visualization. *Journal of Visual Languages & Computing 1*, 1 (1990), 97–123. 5

[NK15] NUSRAT S., KOBOUROV S.: Task taxonomy for cartograms. *17th IEEE Eurographics Conference on Visualization (EUROVIS - short papers)* (2015). 5, 6, 18, 19, 24

[NK16a] NUSRAT S., KOBOUROV S.: Cartogram interactive browser. http://cartogram.cs.arizona.edu/ survis-cartogram/, 2016. [Online: accessed 05-January-2016]. 9

[NK16b] NUSRAT S., KOBOUROV S.: The state of the art in cartograms. In *Eurographics conference on Visualization (EuroVis)–State of The Art Reports* (2016), vol. 35, The Eurographics Association, pp. 619–642. 5, 6, 18, 24

[NVPMW14] NUALART-VILAPLANA J., PEREZ-MONTORO M., WHITELAW M.: How we draw texts: a review of approaches to text visualization and exploration. *El profesional de la informacion 23*, 3 (2014), 221–235. 5

[PBS93] PRICE B. A., BAECKER R. M., SMALL I. S.: A principled taxonomy of software visualization. *Journal of Visual Languages & Computing 4*, 3 (1993), 211–266. 5

[PWL97] PANG A. T., WITTENBRINK C. M., LODHA S. K.: Approaches to uncertainty visualization. *The Visual Computer 13*, 8 (1997), 370–390. 24

[Rob07] ROBERTS J. C.: State of the art: Coordinated & multiple views in exploratory visualization. In *Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV'07. Fifth International Conference on* (2007), IEEE, pp. 61–71. 5

[RWA*11] RIND A., WANG T. D., AIGNER W., MIKSCH S., WONGSUPHASAWAT K., PLAISANT C., SHNEIDERMAN B.: Interactive information visualization to explore and query electronic health records. *Foundations and Trends in Human-Computer Interaction 5*, 3 (2011), 207–298. 5

[SB10] SILIC A., BASIC B. D.: Visualization of text streams: A survey. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems* (2010), Springer, pp. 31–43. 5

[sch16] Google scholar. https://scholar.google.co.uk/, 2016. Accessed: 2016-1-20. 2

[SDSW12] SHNEIDERMAN B., DUNNE C., SHARMA P., WANG P.: Innovation trajectories for information visualizations: Comparing treemaps, cone trees, and hyperbolic trees. *Information Visualization 11*, 2 (2012), 87–105. 4, 5

[SdVW14] SCHOTS M., DE VASCONCELOS R. R., WERNER C.: why do visualization approaches support software reuse tasks? http:// www.cos.ufrj.br/~schots/survis_reuse/, 2014. [Online: accessed 05-January-2016]. 9

[SHS11] SCHULZ H.-J., HADLAK S., SCHUMANN H.: The design space of implicit hierarchy visualization: A survey. *IEEE transactions on visualization and computer graphics 17*, 4 (2011), 393–411. 5, 6, 10, 11, 24

[SMT13] SEDLMAIR M., MUNZNER T., TORY M.: Empirical guidance on scatterplot and dimension reduction technique choices. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2634–2643. 4, 5

[SNHS13] SCHULZ H.-J., NOCKE T., HEITZLER M., SCHUMANN H.: A design space of visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2366–2375. 5

[SNR14] SUN M., NORTH C., RAMAKRISHNAN N.: A five-level design framework for bicluster visualizations. *IEEE transactions on visualization and computer graphics 20*, 12 (2014), 1713–1722. 5, 6, 22, 24

[SP16] SEDIG K., PARSONS P.: Design of visualizations for human-information interaction: A pattern-based framework. *Synthesis Lectures on Visualization 4*, 1 (2016), 1–185. 14

[SSG12] SHIRAVI H., SHIRAVI A., GHORBANI A. A.: A survey of visualization systems for network security. *IEEE Transactions on visualization and computer graphics 18*, 8 (2012), 1313–1329. 5, 6, 20, 24

[STMT12] SEDLMAIR M., TATU A., MUNZNER T., TORY M.: A taxonomy of visual cluster separation factors. In *Computer Graphics Forum* (2012), vol. 31, Wiley Online Library, pp. 1335–1344. 5, 6, 7, 8, 24

[SWLL13] SUN G.-D., WU Y.-C., LIANG R.-H., LIU S.-X.: A survey of visual analytics techniques and applications: State-of-the-art research and future challenges. *Journal of Computer Science and Technology 28*, 5 (2013), 852–867. 5

[TA11] TOMINSKI C., AIGNER W.: A visual survey of visualization techniques for time-oriented data. http://www.informatik.uni-rostock.de/~ct/timeviz/timeviz.html, 2011. [Online: accessed 05-January-2016]. 9

[TGK*14] TOMINSKI C., GLADISCH S., KISTER U., DACHSELT R., SCHUMANN H.: A survey on interactive lenses in visualization. vol. 3, Citeseer. 5, 6, 7, 22, 23, 24

[TGK*16] TOMINSKI C., GLADISCH S., KISTER U., DACHSELT R., SCHUMANN H.: Interactive lenses for visualization: An extended survey. In *Computer Graphics Forum* (2016), Wiley Online Library. 5, 23

[Tob04] TOBLER W.: Thirty five years of computer cartograms. *ANNALS of the Association of American Geographers 94*, 1 (2004), 58–73. 5

[VBW15a] VEHLOW C., BECK F., WEISKOPF D.: The state of the art in visualizing group structures in graphs. In *Eurographics Conference on Visualization (EuroVis)-STARs* (2015), VGTC, The Eurographics Association, pp. 21–40. 5, 6, 15, 16, 24

[VBW15b] VEHLOW C., BECK F., WEISKOPF D.: Visualizing group structures in graphs - bibliography. http://groups-in-graphs.corinna-vehlow.com/, 2015. [Online: accessed 05-January-2016]. 9

[VLKS*11] VON LANDESBERGER T., KUIJPER A., SCHRECK T., KOHLHAMMER J., VAN WIJK J. J., FEKETE J.-D., FELLNER D. W.: Visual analysis of large graphs: state-of-the-art and future research challenges. In *Computer graphics forum* (2011), vol. 30, Wiley Online Library, pp. 1719–1749. 5, 6, 14, 24

[War02] WARD M. O.: A taxonomy of glyph placement strategies for multidimensional data visualization. *Information Visualization 1*, 3-4 (2002), 194–210. 5

[WFL*15] WAGNER M., FISCHER F., LUH R., HABERSON A., RIND A., KEIM D. A., AIGNER W.: A survey of visualization systems for malware analysis. In *Eurographics Conference on Visualization (EuroVis) - STARs* (2015), The Eurographics Association, pp. 105–125. 5, 6, 20, 21, 24

[WSJ*14] WANNER F., STOFFEL A., JÄCKLE D., KWON B. C., WEILER A., KEIM D. A., ISAACS K. E., GIMÉNEZ A., JUSUFI I.,

GAMBLIN T.: State-of-the-art report of visual analysis for event detection in text data streams. In *Computer Graphics Forum* (2014), vol. 33 of *EG*, Citeseer, The Eurographics Association. 5, 6, 9, 24

[WZM*16] WANG X.-M., ZHANG T.-Y., MA Y.-X., XIA J., CHEN W.: A survey of visual analytic pipelines. *Journal of Computer Science and Technology 31*, 4 (2016), 787–804. 5

[ZH15] ZHOU L., HANSEN C.: A survey of colormaps in visualization. *IEEE Transactions on Visualization and Computer Graphics 22*, 8 (2015). 5, 6, 8, 24

[ZMS14] ZAIDI F., MUELDER C., SALLABERRY A.: *Analysis and Visualization of Dynamic Networks*. Springer New York, 2014, pp. 37–48. 5

[ZSB*12] ZHANG L., STOFFEL A., BEHRISCH M., MITTELSTADT S., SCHRECK T., POMPL R., WEBER S., LAST H., KEIM D.: Visual analytics for the big data era - a comparative review of state-of-the-art commercial systems. In *Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on* (2012), IEEE, pp. 173–182. 5

[ZXYQ13] ZHOU H., XU P., YUAN X., QU H.: Edge bundling in information visualization. *Tsinghua Science and Technology 18*, 2 (2013), 145–156. 5