



HAL
open science

Enhancing User Interaction with Business Intelligence Dashboards

Micheline Elias

► **To cite this version:**

Micheline Elias. Enhancing User Interaction with Business Intelligence Dashboards. Other. Ecole Centrale Paris, 2012. English. NNT : 2012ECAP0039 . tel-00969170

HAL Id: tel-00969170

<https://theses.hal.science/tel-00969170>

Submitted on 2 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Doctoral Thesis

Enhancing User Interaction with Business Intelligence Dashboards

by

Micheline ELIAS

*A thesis submitted in fulfilment of the requirements
for the degree of*

Doctor of Philosophy

Ecole Centrale Paris, MAS Laboratory
SAP Research, Business Intelligence Practice

October 2012

Committee in charge:

Supervisors	:	Marie-Aude AUFAURE	Professor at Ecole Centrale Paris
		Anastasia BEZERIANOS	Assistant Professor at Paris-Sud 11 University
President	:	Jean-Daniel FEKETE	Research director at INRIA Saclay
Reviewers	:	Mountaz HASCOËT	Assistant Professor at Université Montpellier II
		Margit POHL	Associate Professor at Vienna University of Technology

Abstract

Enhancing User Interaction with Business Intelligence Dashboards

by Micheline ELIAS

Doctor of Philosophy

Ecole Centrale Paris

MAS Laboratory

The goal of this thesis is to improve visualization dashboards in order to help decision making and Business Intelligence (BI) analysis, and make them accessible to a larger user audience. We identify current trends and use practices of visualization dashboards in the BI domain, and following a user-centered design approach we provide enhancements and solutions. More specifically we address the following challenges: making dashboard construction and use accessible to visualization novices, enhancing dashboards with advanced annotation capabilities to help BI analysts in the analysis process, and add storytelling functionality to dashboards to help the communication between analysts and decision makers. Our user-centered approach consists of gathering user requirements (through background search, in-depth interviews or prototyping sessions with BI experts), iterative prototype development, and evaluation of our prototypes with representative users.

First, our work reports gathered insights regarding novice user practices on Exploration Views (EV), a system that allows novice visualization users to easily build and customize BI information dashboards, but also provides functionality needed by experts. We evaluated EV with both expert and novice visualization users, and found differences in how experts and novice users interact with the interface. Based on those observations we provide new guidelines that augment previous work on designing for visualization novices, in the context of interactive visualization systems in the form of dashboards.

BI analysts using dashboards (experts or novices) need support in record-keeping of their analysis, their data discoveries and gathered insights. This support is lacking in most dashboard visualization systems. We conducted in-depth interviews with BI experts that led to new annotation needs for multi-chart visualization systems (such as dashboards), on which we based the design of a dashboard prototype that supports data and context aware annotations, shared across visualizations and datasets. We focused particularly on novel annotation aspects, such as multi-target annotations, annotation transparency across charts, and annotations that can be shared among different data-sources and BI

applications. We identified and provided solutions to challenges in using context aware annotations, notably issues arising when the annotation's "context" changes (e.g. annotated data are deleted or changed). Our prototype was evaluated with BI experts that were able to easily perform several tasks that they deemed important in their work.

To increase the value of an analysis, recorded insights and knowledge need to be organized and communicated to others, and made available for future analysts to learn from. This communication step is very important, as often the analyst of data and the decision maker are two different people. To achieve this we need an easy insight organization and sharing mechanism to facilitate the transfer of knowledge, experiences, and stories, to decision makers and other analysts. We interviewed BI analysis experts and collected new requirements related to BI visual storytelling. Based on these requirements we designed and implemented a storytelling prototype that is integrated in a dashboard analysis tool, and allows easy transition from analysis to story creation and sharing. Our system was identified by experts as having great potential for training other analysts. Moreover, it can help BI analysis authors organize and communicate their findings faster to decision makers, and they can reach a broader audience since the results of a data analysis presented in the storytelling tool can be read with little to no training.

The main contributions of this work are:

- (1) Report how novice users interact with the EV system and the differences in how novice and expert users react to EV dashboard system that targets both.
- (2) Provide new design guidelines for practitioners building dashboard applications, on the needs of novice visualization users.
- (3) Define a new set of requirements for annotations, targeting dashboards and other coordinated view visualization systems.
- (4) The design and evaluation of a prototype dashboard that supports context-aware annotations.
- (5) Identify and provide solutions to challenges when using context aware annotations, such as issues arising when the "context" changes (e.g. annotated data are deleted).
- (6) Based on in-depth interviews with expert BI analysts, we derive and provide a better understanding of the actual practices in creating BI stories, and extract new requirements for enhancing BI visual analysis tools with storytelling capabilities.
- (7) Following these requirements and using a user centered design approach, we design and implement a prototype storytelling tool that is fully incorporated in an existing visual analysis dashboard. We report on feedback from BI experts creating a story on the usefulness of the prototype as a communication and teaching medium, and from BI novices reading a story on its effectiveness for story communication.

Keywords: Visualization Dashboard; Business Intelligence; Annotation; Storytelling; User Experience.

Résumé

Amélioration de l'interaction des utilisateurs avec des dashboards de Business Intelligence

par Micheline ELIAS

Docteur de Philosophie

[Ecole Centrale Paris](#)

Laboratoire MAS (Mathématiques Appliquées aux Systèmes)

L'objectif de cette thèse est d'améliorer les tableaux de bord (dashboard) de visualisation afin d'aider à la prise de décision et d'analyse BI, et les rendre accessibles à un public plus large. Nous identifions les tendances actuelles et pratiques d'utilisation des dashboards de visualisation dans le domaine BI, et après une approche de conception centrée, nous proposons des améliorations et des solutions. Plus précisément, nous relevons les défis suivants: faire la construction du dashboard et d'utiliser accessible aux novices de visualisation, l'amélioration des dashboards avec des capacités d'annotation pour aider les analystes BI dans le processus d'analyse, et ajouter des fonctionnalités du storytelling pour aider à la communication entre les analystes et les décideurs. Notre approche centrée-utilisateur consiste à recueillir les besoins des utilisateurs (grâce à la recherche de fond, des entrevues en profondeur ou des séances de prototypage avec des experts BI), le développement itératif des prototypes, et l'évaluation des prototypes issus avec des utilisateurs représentatifs.

Dans la première étape, notre travail rapporte idées recueillies au sujet des pratiques des utilisateurs novices sur Exploitation Views (EV), un système qui permet aux utilisateurs novices de visualisation à créer facilement des dashboards et les personnaliser, mais fournit également des fonctionnalités nécessaires aux experts. Nous avons évalué EV à la fois d'experts et des utilisateurs novices de visualisation, et constaté des différences dans la façon dont les experts et les utilisateurs novices interagissent avec l'interface. En se basant sur ces observations nous proposons de nouvelles lignes directrices, qui augmentent les travaux antérieurs, sur la conception pour les novices, dans le contexte des systèmes de visualisation interactifs sous forme de dashboard.

Les analystes BI utilisant des dashboards (experts ou novices) ont besoin d'aide dans la tenue des dossiers de leur analyse, de leurs découvertes et des idées recueillies des données. Ce soutien fait défaut dans la plupart des systèmes de visualisation de dashboard. Nous avons effectué des entretiens approfondis avec des experts en analyse BI, qui ont révélé de

nouveaux besoins d'annotation pour les systèmes de visualisation multi-graphiques (tels que des dashboard), sur lesquels nous avons basé la conception d'un prototype de tableau de bord qui gère les données et les annotations contextuels, qui peuvent être partagées à travers visualisations et des ensembles de données. Nous nous sommes concentrés en particulier sur les aspects d'annotation innovantes, tels que les annotations multi-cibles, la transparence des annotations à travers des graphiques, ainsi que les annotations qui peuvent être partagées entre les différentes sources de données et des applications de BI. Nous avons identifié et fourni des solutions aux problèmes en utilisant les annotations contextuels, notamment les questions qui se posent lorsque le contexte de l'annotation change (par exemple les données annotées sont supprimés ou modifiés). Notre prototype a été évaluée avec des experts, utilisateurs BI, qui ont réussi d'effectuer facilement plusieurs tâches qu'ils estiment importants dans leur travail.

Les analysis BI (experts ou novices) ont besoin d'enregistrer leurs idées et leurs connaissances les organiser et communiquer aux autres, et les mis à disposition pour les futurs analystes. Cette étape de communication est très important, comme souvent l'analyste de données et le décideur sont deux personnes différentes. Pour atteindre cet objectif nous avons besoin d'une mécanisme de partage facile afin de faciliter le transfert des connaissances, des expériences et des histoires avec les décideurs et d'autres analystes. Nous avons interviewé des experts en analyse BI et recueillies de nouvelles exigences liées à la narration visuelle BI. Sur la base de ces exigences, nous avons conçu et mis en œuvre un prototype de narration (storytelling) qui est intégré dans un dashboard, et permet une transition facile entre l'analyse et la création d'une histoire visuelle et son partage. Notre système a été identifié par les experts comme ayant un grand potentiel pour la formation d'autres analystes. En outre, il peut aider les auteurs à l'analyse BI d'organiser et de communiquer leurs conclusions rapidement aux décideurs, et les analyse peuvent atteindre un public plus large, puisque les résultats d'une analyse présentée dans l'outil de narration peut être lu avec peu ou pas de formation.

Les contributions principales de ce travail sont:

- (1) Rapport de la façon dont les utilisateurs novices interagissent avec le système EV, et les différences entre la façon dont les utilisateurs novices et les experts réagissent au système de dashboard EV.
- (2) Fournir des lignes directrices de conception des applications de dashboard, liées aux besoins des utilisateurs novices de visualisation.
- (3) Définir des exigences pour les annotations, sur les dashboard et autres systèmes de visualisation de vues coordonnées .
- (4) La conception et l'évaluation d'un dashboard qui supporte des annotations contextuels.

(5) Identifier et apporter des solutions aux problèmes de l'utilisation des annotations contextuels, comme les questions qui se posent lorsque le contexte change (par exemple données annotées sont supprimées ou modifiées).

(6) Basé sur des entretiens avec des experts analystes BI, nous dérivons une meilleure compréhension des pratiques réelles dans la création d'histoires BI, et nous extrairons de nouvelles exigences pour améliorer les outils BI d'analyse visuelle avec les capacités de narration visuelle.

(7) En se base sur ces exigences et en utilisant une approche de conception centrée utilisateur, nous concevons et mettons en œuvre un outil de narration qui est entièrement incorporé dans un dashboard, pour transmettre fluidement entre l'analyse et la création d'une narration. Nous rapportons les commentaires des experts BI qui créent une narration, sur l'utilité du prototype comme une moyenne de communication et d'enseignement. Nous rapportons aussi sur l'utilisation du prototype par de novices BI qui lisent une narration fait par notre prototype.

For my father

Dr.Emad ELIAS

Acknowledgements

The acknowledgements and the people to thank go here, don't forget to include your project advisor...

Publications from this thesis

Papers

1. Exploration Views: understanding dashboard creation & customization for visualization novices. Micheline Elias & Anastasia Bezerianos. In Proceedings of INTERACT 2011, the 13th IFIP TC13 Conference on Human-Computer Interaction, Volume Part IV, pages 274-291. Lisbon, Portugal, 5-9 September 2011.
2. Annotating BI Visualization Dashboards: Needs & Challenges. Micheline Elias & Anastasia Bezerianos. In Proceedings of CHI 2012, the ACM SIGCHI Conference on Human Factors in Computing Systems, pages 1641-1650. Austin, Texas, USA, 5-10 May 2012.
3. Storytelling in Visual Analytics tools for Business Intelligence. Micheline Elias, Marie-Aude Aufaure & Anastasia Bezerianos. *Submitted* to the 16th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW 2013) will be held February 23-27 in San Antonio, Texas, USA.

Patents

4. Data tracking visualization. Micheline Elias, Alexis Naibo, Christophe Favart & Anastasia Bezerianos. 2010.
5. Context aware annotations in BI documents. Micheline Elias & Anastasia Bezerianos. 2011.
6. Creating business storytelling from BI exploration/annotation systems. Micheline Elias 2012.

Contents

Abstract	ii
French Abstract	iv
Acknowledgements	viii
Publications from this thesis	ix
List of Figures	xiv
List of Tables	xvii
1 Introduction	1
1.1 Abstract	1
1.2 Research context	2
1.2.1 Business Intelligence Systems	3
1.2.2 Data Visualization and Visual Analytics	5
1.2.3 Dashboards	7
1.2.4 Methodology	8
1.3 Thesis Focus	9
1.4 Thesis Contributions	11
1.5 Thesis Outline	11
1.6 Summary	13
2 Related Work	14
2.1 Abstract	14
2.2 Introduction	15
2.3 Dashboards	15
2.3.1 Dashboards creation	15
2.3.2 Visualization creation	16
2.3.3 Visualization for novices	16
2.3.4 Novices in visualization	18
2.4 Annotations	20
2.4.1 Annotation of Textual data	20
2.4.2 Annotation of Image Data	21

2.4.3	Annotation of Visualizations	22
2.4.4	Context-Aware Annotations in Visualizations	22
2.5	Storytelling	24
2.5.1	Stories in Business	24
2.5.2	Stories in sense making	25
2.5.3	Stories in data visualization	25
2.6	Summary	28
3	Understanding Dashboard Creation and Customization for Visualization	
	Novices	29
3.1	Abstract	29
3.2	Introduction	30
3.3	Exploration Views (EV) Prototype	31
3.3.1	Easy Creation	32
3.3.2	Ease of Customization	35
3.3.3	Supporting Visual Analytics Process	36
3.4	Exploration Views Technical Platform	40
3.5	Study Design	41
3.5.1	Participants and Pilot Studies	41
3.5.2	Procedure and Apparatus	42
3.5.3	Tasks and Dataset	42
3.5.4	Data Analysis	44
3.6	Findings	45
3.6.1	Discoverability and Ease of Use of Interface	46
3.6.2	General Observations on Dashboards for Novices	47
3.6.2.1	Chart Creation and Customization	47
3.6.2.2	Interaction	49
3.6.2.3	Design Tradeoffs	52
3.7	Discussion and Design Guidelines	52
3.8	Summary	55
4	Annotating BI Visualization Dashboards	56
4.1	Abstract	56
4.2	Introduction	57
4.3	Approach	58
4.4	Interview 1: Annotation Requirements & Needs	59
4.4.1	BI Dashboards Annotation Practices	59
4.4.2	BI Dashboards Annotation Needs	59
4.4.2.1	Multiple Target Annotations	61
4.4.2.2	Chart Transparent Annotations	61
4.4.2.3	Visible Annotation Representation on the Dashboard	61
4.4.2.4	Granularity Transparent Annotations	61
4.4.2.5	Annotation Validity & Lifetime	62
4.4.2.6	Sharing BI Dashboard Annotations	62
4.4.2.7	Formal Annotation Representation on Dashboards	62
4.4.3	Summary	62
4.5	Context Aware Annotations Prototype	63

4.5.1	Architecture	64
4.5.2	Interface	65
4.6	How to Use Context	67
4.6.1	Transparency	67
4.6.2	Chart linking	69
4.6.3	Recommendations	69
4.6.4	External search	69
4.6.5	Automatic annotation	70
4.7	Interview 2: Design Challenges & Solutions	70
4.7.1	Dynamic context vs. dynamic annotations	70
4.7.2	Granularity	71
4.8	User Study	75
4.8.1	Participants	75
4.8.2	Procedure and Apparatus	75
4.8.3	Tasks and datasets	75
4.9	Results	76
4.9.1	Supporting Analysis	78
4.9.2	Visual Design	80
4.10	Summary	82
5	Storytelling in Visual Analytics tools for Business Intelligence	83
5.1	Abstract	83
5.2	Introduction	84
5.3	Interviews	85
5.3.1	BI Reports	85
5.3.2	What is a BI Story	86
5.3.3	Teaching BI storytelling	89
5.3.4	Creating a BI Story	89
	Interactivity	90
	Story templates	91
5.4	Participatory design session	92
5.5	Requirements for Creating BI stories	92
5.6	BI Narrative Prototype	96
5.6.1	Exploration/annotation dashboard (Figure 5.8)	96
5.6.2	Narrative board (Figure 5.9)	97
5.7	User Feedback	100
5.7.1	Story creation by experts	100
	5.7.1.1 General feedback and recommendations	100
	5.7.1.2 Constraining and opening up the story	101
5.7.2	Story reading by novices	102
	5.7.2.1 Reading the report	103
	5.7.2.2 Reading the story	103
5.8	Summary	106
6	Conclusions and Future Work	107
6.1	Review of Thesis Results and Contributions	107
6.2	Limitations, Discussions and Future Work	109

6.2.1	BI Dashboard Creation and Customization	109
6.2.2	Context-Aware Annotations	110
6.2.3	Storytelling	111

Bibliography	113
---------------------	------------

List of Figures

1.1	OLAP database schema for Sales database	4
1.2	OLAP data cube example	5
1.3	ISO UCD	8
2.1	Commercial BI dashboards	17
2.2	Improvise	18
2.3	Snap-Together	19
2.4	Visual Analytics (VA) systems	23
2.5	Types of Narrative Visualization proposed by Segel et al. [1]	26
2.6	Geotime Interface	27
3.1	Dashboard life-cycle	30
3.2	Exploration Views	32
3.3	Exploration Views creation dialog	33
3.4	Exploration Views Chart Suggestion	34
3.5	Exploration Views Data Category and measures selection.	35
3.6	Exploration Views Tool Bar	35
3.7	Drag and Drop visualization	37
3.8	Filters and Faceted navigation creation.	38
3.9	Filters and Faceted navigation creation, drag-and-drop.	38
3.10	Active Filters	39
3.11	Graphical explanations	43
3.12	eFashion	45
3.13	follow-up interview results	46
3.14	An example of EV Filtering functionality	49
3.15	Understanding Linking & coordinated views by novice and experts	50
4.1	Our dashboard prototype with context aware annotations. (a) The main dashboard contains icons with the number of annotations for each data-point (b). Below the list of all dashboard annotations in (c), is a set of recommended annotations (d) emphasizing their similarity.	58
4.2	Three examples of the actual practice of annotating visualizations provided by our participants. (1) A dashboard represented in the middle, and associated annotations on right and left. The left annotations represent the bad remarks, and the right side contains the good remarks. (2) and (3) are two slides from a report training powerpoint presentation, that contains annotations made on the presentation tool.	60

4.3	This figure shows the context-aware annotation structure followed in our prototype. An annotation can be attached to multiple targets, and each target is identified by its data context (dimensions and measures, names and values). The annotation saves with each target the dataset name, visualization type and a snapshot at the time of annotating. Moreover, an annotation has different properties such as: author, creation date, annotation ranking, categorization and lifetime. Targets	65
4.4	Detailed annotation view, including (a) all related data contexts & detailed data context in text form, (b) content, (c) metadata, (e) discussion thread, and (d) annotation properties (lifetime, privacy, etc).	66
4.5	Chart Transparent Annotations	68
4.6	This figure represents how our prototype distinguish invalid archived annotations (grayed out and written in red "Invalid"), from the valid annotations, in the annotation list.	72
4.7	Two variations of the dialog for creating a new annotation: (a) permanent annotation that would be always valid, (b) periodic annotation in which the user has to identify the life time of the annotation in time periods. The user's choice of "permanent" or "periodic" periods defines the view of the dialog.	73
4.8	Level Transparent Annotations	74
4.9	An annotation example that the participant explored during performing Task 3. (a) Radar chart presents the "50 Richest Americans in 2006". (b) Annotation icons that declares the existence of annotations attached to two data points on the chart. (c) In a new window we see the content of an annotation that says "Second Rich man!". (d) The original chart type (scatter plot) when the annotation was created. (e) Context of the data point, the annotations is attached to. (f) List of all existing annotations.	78
4.10	Three designs linking annotations in the annotation list with their data targets. Experts preferred the last design.	81
5.1	First page of a BI report given by an expert: a dashboard snapshot that presents the bug status in a specific development project. (Used with the experts' permission)	86
5.2	An example of a BI report given by an expert that was used in the evaluation of our system. This dashboard represents the "technical backlog" & "functional users stories" items status in different perspectives. (1) Green group, represents the status of the items per team. (2) Yellow group, shows the created items versus the completed "done" items, and the category (critical, major, etc) of each. (3) Pink visualization presents the remained items to be completed and its progression through time.	87
5.3	An example of explanations included in a power point presentation used as an introduction session to report readers. Example provided by one of our interviewed experts.	88
5.4	Genres of Narrative Visualization selected by BI experts from a group of templates defined by Segel et al.	91
5.5	Participatory design session	93
5.6	Story created by an expert during the design session.	94

5.7	The final story mockup created by an expert (fully annotated and sequenced story seen in Figure 5.6). It represents how a static story on a system should look like.	94
5.8	Dashboard with visualizations on the left and list of annotations on the right.	96
5.9	Narrative board (a), Sequence of entities (b), Relational (c) and Organization (d) entities, and (e) Playback control.	97
5.10	Storytelling playback modes.	99

List of Tables

3.1	Tasks users performed during each session of the study, and their purpose.	44
3.2	Average ratings of EV for different attributes, by Novice and Expert users (* sig: indicates significant difference between the 2 groups)	47
4.1	Tasks users performed during the study of context-aware annotations on dashboards.	77
4.2	Post-test (context-aware annotations tasks) questionnaire.	78
5.1	Users were asked this group of comprehension question related to the con- tent of the story and charts, after reading first the report then the story. .	103

Chapter 1

Introduction

1.1 Abstract

The explosive growth in the amount of data poses challenges in analyzing large data sets. Among the large number of available BI systems that help with large data set analysis, are BI dashboards. BI dashboards provide BI analysts with advanced visual analytics capabilities. But business users still spend too much time creating dashboards, analyzing their data, and preparing them for sharing findings with others. The motivation of our work is to improve the above processes. In this chapter we present the broad context of this work that flows between HCI, BI, and Data visualization, with a particular focus on dashboards, which are the main focus of this thesis. We present in this chapter the thesis motivation, approach, challenges that we address, as well as the content of the following chapters.

1.2 Research context

"There is nothing better than a picture for making you think of questions you had forgotten to ask (even mentally)."

— John W. Tukey (1985)

The amount of digital data in our world is growing at an incredibly quick rate according to research by the International Data Corporation (IDC)[2], and the increasing volume and detail of information captured by enterprises will fuel exponential growth in data for the foreseeable future. Therefore, analyzing large datasets is a key focus in business [3]. Extraction of hidden predictive information from large databases helps companies predict future trends and behaviors, and make decisions. This leads more people in every business to try to understand and exploit big data, not just managers that used to be the main target group of Business Intelligence systems. This emphasizes the need to reduce the cost of finding information as well as getting relevant answers to complex questions easily.

Business Intelligence (BI) deals with the collection of processes and software that supports organizations in understanding large business datasets, retrieving and analyzing information and making decisions. The value of visual presentation of data was identified early on in this field, and it usually takes the form of dashboards, interfaces resembling an automobile's dashboard that organize and present information in an easy to read manner [4]. BI visualization dashboards provide collections of multiple visual components, such as charts, on a single view so that information can be monitored at a glance [5].

The current life-cycle of a BI dashboard involves multiple actors [6], including end-users and business analysts. In user-centered BI [7] end-users intervene and provide feedback to the business analysts that create customized dashboards to meet the end-users needs. This feedback comes at different stages of the dashboard design and setup, and involves a large amount of communication between business analysts and end-users. Thus there is an intrinsic delay introduced in any end-user requirement change or customization request. When end-users needs are understood, the analysts conduct data analysis, and prepare their findings to present to their end-users or other analysts.

In this thesis we focus on enhancing and improving user interaction with BI dashboards in the different parts of a dashboard life cycle. Our work falls under the Human-Computer Interaction (HCI) field, that aims to improve the interactions between users and computers by making computers more usable and receptive to the user's needs. Following a popular HCI methodology, we adopted a user-centered design approach to understand the users'

requirements toward dashboard creation and analysis, and propose solutions for each step in the life of a dashboard.

Successfully analyzing big data sets offers considerable benefits to companies and organizations (e.g. a competitive edge against competitors), but also to individual users (e.g. a better investment plan). Thus we have seen an increase of visualization platforms targeting visualization novices. Researchers are now focusing on gathering insights regarding novice user practices. But dashboards are still only created by experts due to a complex and long creation process, that often cannot be conducted by end-users, and that involves a large amount of communication between business analysts and end users. This long creation process that requires special expertise makes dashboards less popular with novice users, and less accessible to a broader audience; in our work we focused on making dashboard creation accessible to visualization novices, developing and testing a dashboard that is easy to create and customize.

While attempting to aid users in their analysis, we identified the need of annotation and organization of findings as of particular interest, as in other visual analysis domains. Despite the growing use and research work on annotations for visual data analysis in general, related research on the domain of BI dashboards and other complex multi-chart visualization environments (coordinated views), has not equally progressed. Our work attempts to rectify that. Through user interviews we found out that users need more sophisticated solutions for annotating dashboards. Those requirements were implemented in a context-aware annotation dashboard prototype that was evaluated by BI experts.

In the last part of the dashboard's life, the experts' insights, knowledge and analysis stories need to be recorded in order to communicate them to others (e.g. end-users) more efficiently and make them available for future analysts to learn from. This communication step is very important, as often the analyst of data and the decision maker are two different people. Based on requirements from analysts we designed and implemented a storytelling prototype tool that is integrated in a dashboard analysis tool, and allows easy transition from analysis to story creation and sharing.

In the next sections we delve into each research field related to this thesis.

1.2.1 Business Intelligence Systems

"Business Intelligence is a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making." [8]

This thesis focuses on Business intelligence systems (BIS). BIS are interactive systems intended to help decision makers use communication technologies, data, documents, knowledge, and analytical models to identify and solve problems, thus a BI system can be called a decision support system (DSS)[9]. We focus on Online Analytical Processing (OLAP) tools, that enable users to interactively analyze multidimensional data from multiple perspectives. The core of the OLAP technology is the data cube Figure 1.2, which is a multidimensional database model. The model consists of dimensions and measures. The measures are numerical data such as revenue, cost, sales and budget. Those are categorized by dimensions, which are used to group the data similar to the "group by" operator in relational databases. Typical dimensions are time, location and product, and they are often organized in hierarchies. A hierarchy is a structure that defines levels of granularity of a dimension and the relationship between those levels. A time dimension can for example have hours as the finest granularity, and higher up the hierarchy can contain days, months, quarters and years.

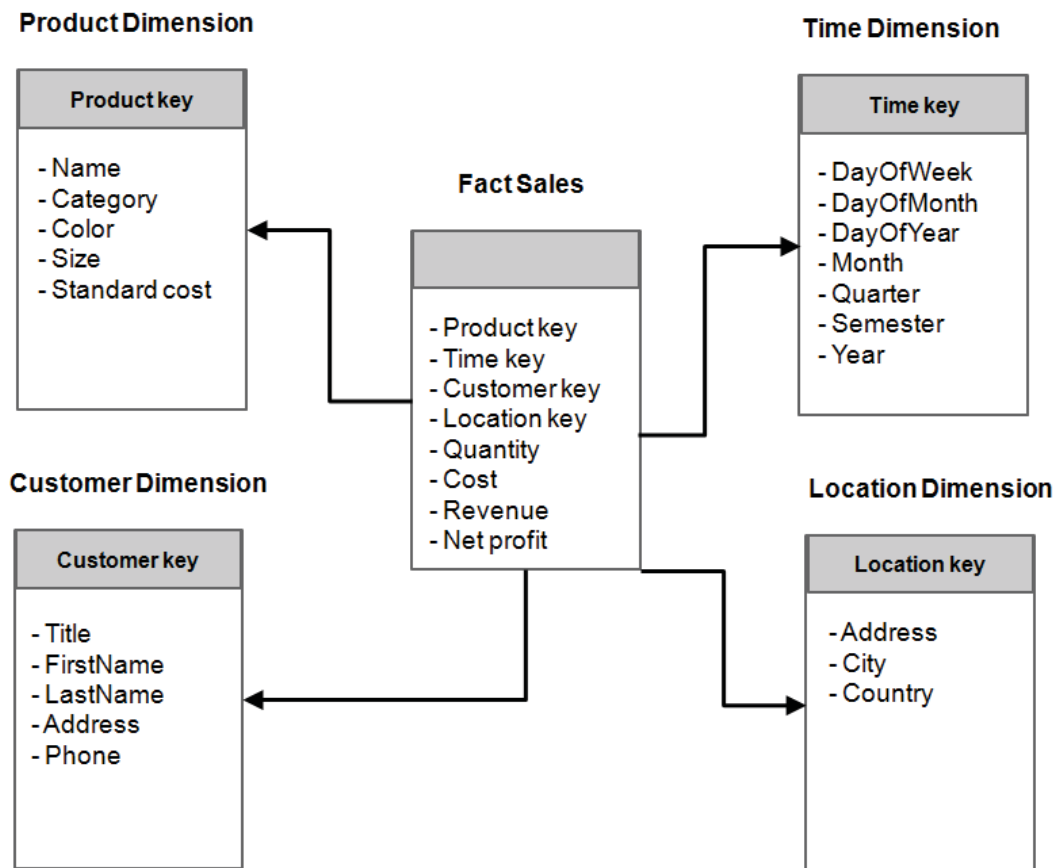


FIGURE 1.1: Example of an OLAP database schema for Sales database. The fact table usually represents measures. In this example the fact table represents Sales which are categorized by dimensional tables (Product, Customer, Time and Location). Each dimensional table needs to include a primary key that corresponds to a foreign key in the fact table that represents the sales measures.

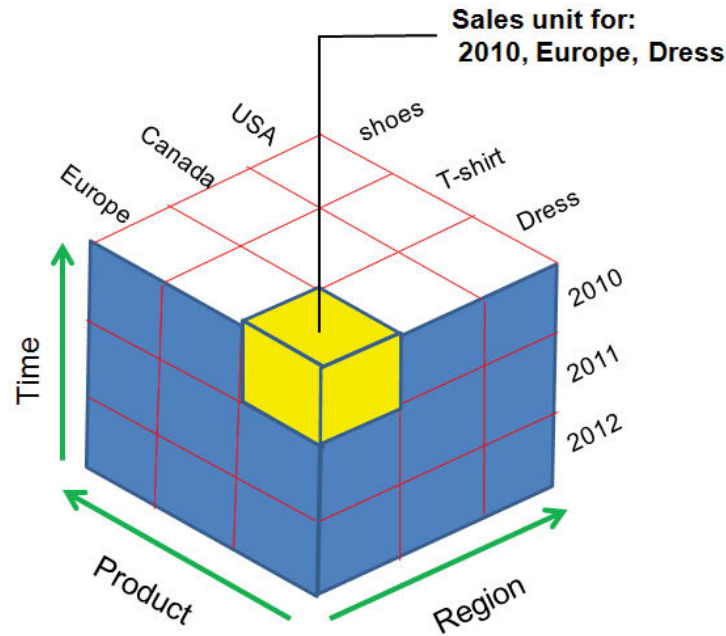


 FIGURE 1.2: OLAP data cube example.

In the past, business intelligence systems were used to answer OLAP queries in table based reports that answered multiple queries separately. The new generation of BIS has two salient features: integration of multiple query answers and visualization of these answers (e.g. using visualization dashboards).

This work has been heavily influenced by the industrial environment in BI practice of SAP Research. SAP is one of the market leaders in enterprise application software, providing important enterprise tools such as Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), and BI systems such as SAP Crystal Reports, SAP BusinessObjects Dashboards, SAP BusinessObjects Explorer.

1.2.2 Data Visualization and Visual Analytics

Friendly [10] has identified data visualization as "the study of the visual representation of data, meaning -information that has been abstracted in some schematic form, including attributes or variables for the units of information-."

The main goal of data visualization is its ability to visualize data, communicating information clearly and effectively. It doesn't mean that data visualization needs to look boring to be functional or extremely sophisticated to look beautiful. To convey ideas effectively, both aesthetic form and functionality need to go hand in hand, providing insights into a rather sparse and complex

data set by communicating its key-aspects in a more intuitive way.

— Vitaly Friedman[11]

Data visualization plays a major role in supporting data analysis and summarize the main characteristics of a data set in an easy-to-understand form, often with visual graphics or charts. It permits the analytical reasoning process to become faster and more focused [12]. Visualization research seeks to supplement human cognition by leveraging human visual capabilities to make sense of abstract information [13].

Visualizing multidimensional data is a challenging problem, because the human perceptual system is well equipped to process data presented in maximum three dimensions. Beyond this limit we cannot simply map the data into a graphical construction. Usually data is considered multidimensional if it has more than three dimensions. An important challenge in multidimensionality is the visualization of the relationships between data items across dimensions.

Nevertheless as explained by James et al. [14], specific types of visual representations and interaction techniques allow users to see, explore, and understand large amounts of information at once. In the domain of multidimensional visualization these include: scatterplot matrices [15], multidimensional glyphs (such as StarGlyphs [16] [17] [18] and Chernoff Faces), parallel coordinates [19] [20] [21] or starplots, multidimensional scaling [22], tabular visualization techniques [23] [24] and coordinated and multiple views [25]. This thesis focuses on coordinated and multiple views such as dashboards, described next, as they are the most popular within the BI domain.

Coordinated and Multiple Views represent a highly interactive visualization environment that enables the user to see diverse datasets in various forms, to manipulate the visual presentation in different ways, and also interact and coordinate the interaction between the different views [25]. Coordinated and Multiple Views help users to generate new information and understand their data better by interacting with the presented information and viewing it through different representations, that are all visible at the same time.

Often the discussed visualization techniques are part of a larger set of visual analytics tools (VA). Visual Analytics has been identified as "the science of analytical reasoning supported by highly interactive visual interfaces" [26]. VA combines automatic data analysis methods (e.g. data cleaning and data mining) and interactive visualizations. People use VA tools and techniques to derive insight from complex and large data and synthesize information into knowledge, that directly support assessment, planning, and decision making. There is a demand for VA support in different application areas where large information spaces have

to be processed and analysed, such as monitoring climate and weather, biology, medicine and business. Our work focus is Business Visual Analytics, that aids in analysing huge business data sets and permits taking right decisions. Nevertheless our focus is mainly on the interaction and visualization challenges, and less on the automatic analysis methods.

1.2.3 Dashboards

A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance.

— Stephen Few[27]

In the domain of business intelligence (BI) the main visualization tool is BI dashboards [28]. Dashboards can provide an effective solution to the overwhelming amount of data that business users experience every day [29]. Current market trends (seen in the Gartner Survey [3]), recognize dashboards and data analysis are one of the driving forces adding value to companies. This indicates a tendency for an expanding end-user population with diverse needs.

Dashboards can be classified according to their role. Few [30] has identified the following dashboard types: *strategic, analytical, operational, or informational*.

Most often, dashboards are used for *strategic* purposes [29], they provide a general and quick overview of the organization's "health". They provide quick overviews and comparisons to target values as well as simple performance evaluations (e.g. good or bad).

The *analytic* dashboard, is designed to help in data analysis. This can include making comparisons, reviewing extensive histories, evaluating performance, and supports interactions with the data such as drilling down into different data levels.

Operational dashboards are for monitoring data, that needs real-time, and minute-by-minute attention, and has no statistics or analysis functionalities. They provide alerts and clear presentation of the appropriate response in an emergency event. They are generally used by those on a departmental, rather than executive level [31].

Dashboards are tailored to fit specific requirements, and a dashboard does not need to fit a single category in order to be successful. Rather, it needs to provide clear and immediate information in order to improve decision making and understanding of data[27]. In this thesis we kept the dashboard purpose open to meet different requirements of BI domain analysts.

1.2.4 Methodology

“Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them.”

— Hewett, Baecker, Card, Carey, Gasen, Mantei, Perlman, Strong and Verplank [32]

The main goal of Human Computer Interaction HCI is to make computers more user-friendly by minimizing the barrier between the human’s goal of what they want to accomplish and the computer’s understanding of the user’s task.

One of the most significant HCI design methodologies is *User-centered design* Figure 1.3, in which users take center-stage in the design of any computer system, and is the design methodology we followed to understand our users’ needs and propose solutions in this work.

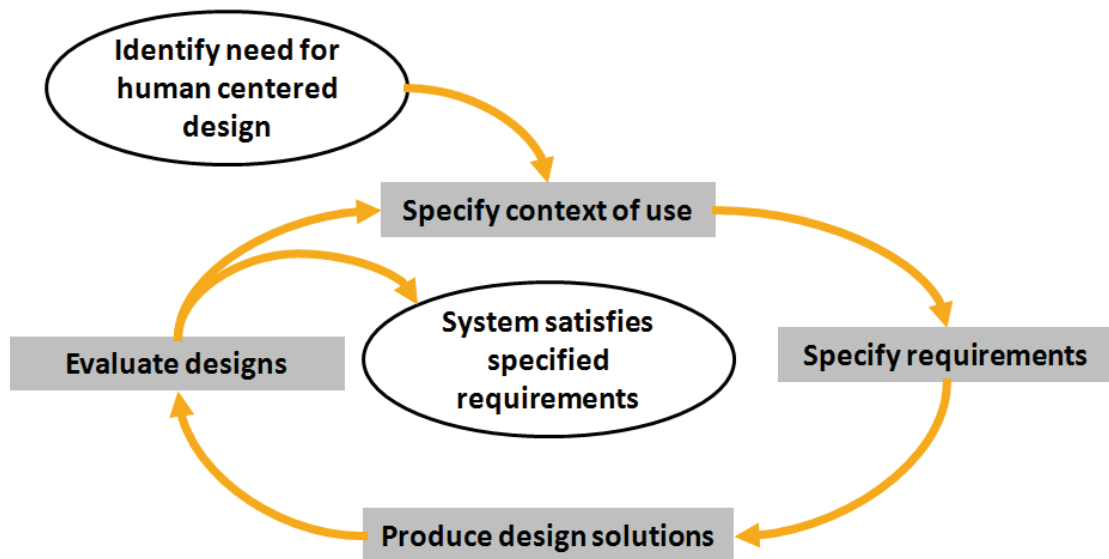


FIGURE 1.3: (ISO 13407: Human-centred design process) defines a general process for including human-centered activities throughout a development life-cycle. The four activities are: (1) **Specify the context of use:** Identify the people who will use the product, what they will use it for, and under what conditions they will use it. (2) **Specify requirements:** Identify any business requirements or user goals that must be met for the product to be successful. (3) **Create design solutions:** This part of the process may be done in stages, building from a rough concept to a complete design. (4) **Evaluate designs:** The most important part of this process is that evaluation - ideally through usability testing with actual users - is as integral as quality testing is to good software development.

A central concept in HCI is that of *usability*[33], ie "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use". [ISO definition]

Many methods exist for studying usability, but the most basic and useful is user testing, which has 3 components as identified by Nielsen [34]:

- Get hold of some representative users, in our case Business Intelligence experts or novices.
- Ask the users to perform representative tasks with the design.
- Observe what the users do, where they succeed, and where they have difficulties with the user interface.

This is the method we have followed to improve the usability of the different prototypes implemented during this thesis.

1.3 Thesis Focus

This work followed a user centred design approach to enhance user interaction and address the challenges when creating, customizing and using dashboards or coordinated view systems, in the BI domain for expert and novice users. We intervene in all aspects of life cycle of a dashboard: creation, analysis and communication.

Creation: The tendency for expanding dashboard end-user population with diverse needs, requires users, who are often novices in information visualization, to get quick access to customizable and easy to build Business Intelligence information dashboards. This is reflected in advances in easily customizable dashboard visualization systems (e.g. [35][36]), it coincides with advances in easily accessible information visualization environments for novices (such as ManyEyes [37]), and is also reflected in the increasing focus of the InfoVis research community on how to support novice users [38][39]. Nevertheless, there are still many questions about how novice users interact with actual information visualization systems and more so with dashboards. In the BI analysis domain, software vendors focus usually on adding more features into the next dashboard release, with no time to do the research needed to discover what actually works, or even to step back and observe how their products are really being used.

To find out the answers to such questions we gathered insights regarding novice user practices on Exploration Views (EV), a system that allows novice visualization users to easily build and customize Business Intelligence information dashboards. EV provides an intuitive environment for dynamically creating, rearranging, searching and exploring multiple visual data representations from diverse data-sources. These aspects aid users

to better retrieve, experiment and familiarize themselves with their data. We provide new guidelines that augment previous work on designing for visualization novices, in the context of dashboards. We expand on this aspects in Chapter 3.

Analysis: Just collections of visual representations do not speak for themselves; to become meaningful they require interpretation and explanation. But because of the limits of human memory [40] remembering all reasoning details and milestones in the analysis process becomes challenging. We report dashboard limitations in analysis record-keeping, data discovery and gathering insights.

We conducted in-depth interviews with dashboard experts and reported the need of making sense of their visualizations by annotating them with their findings. This led to new annotation needs for multi-chart visualization systems, on which we based the design of a dashboard prototype that supports data and context aware annotations that can be shared across visualizations and datasets. We focused particularly on novel annotation aspects, such as multi-target annotations, annotation transparency across charts, as well as attaching annotations to a data layer shared among different data-sources and BI applications, allowing cross application annotations.

This prototype was evaluated by BI experts by performing several important tasks in their work. We also identified and provided solutions to challenges in using context aware annotations, notably issues arising when the annotation's "context" changes (e.g. annotated data are deleted). We discuss this work in detail in Chapter 4.

Communication: Although the work practices of BI analysts' may differ, in the end they all construct stories that include: collections of visual components (dashboards, charts or tables) to support their stories, verbal explanations during presentation(s), or detailed textual annotations and descriptions in the form of written reports and external resources such as wikis. This process remains cumbersome and requires frequent switching between different software (dashboard analysis tools, screen captures, text editors, presentation or drawing tools, etc). BI analysts need an easier mechanism to share their analysis results and tell stories with their data.

Our work aims to aid BI analysts with their storytelling needs, in order to develop a high analysis value. Our interviews indicates that their recorded insights and knowledge, need to be organized and transferred to others, as well as remain available for future users. To achieve this we need a designing and sharing mechanism to facilitate the transfer of knowledge and experiences. We thus collected new requirements related to BI visual storytelling. Based on these requirements we designed and implemented a storytelling prototype tool that is integrated in a dashboard analysis tool and allows easy transition from analysis to story creation and sharing. Chapter 5 provides details on this work.

1.4 Thesis Contributions

The main contributions of this work are:

- (1) Report of how novice users interact with the EV system and the differences in how novice and expert users react to EV dashboard system that targets both.
- (2) New design guidelines for practitioners building dashboard applications, for novice visualization users.
- (3) New set of requirements for annotations, targeting dashboards and other coordinated view visualization systems.
- (4) Design and evaluation of a prototype dashboard that supports context-aware annotations. Using this system, we explain how annotations support advanced analysis functionality such as: "annotate once, see everywhere", annotations for multi-chart visualizations, annotation recommendations to promote knowledge discovery, and the use of the common data model layer present in BI systems to link annotations with external data sources and enable cross application annotations and information foraging.
- (5) Identification of challenges and possible solutions when using context aware annotations, such as issues arising when the "context" changes (e.g. annotated data are deleted or changed at the data source).
- (6) A better understanding of the actual practices in creating BI stories, based on in-depth interviews with expert BI analysts, and emphasize the need of BI storytelling support. We extracted new requirements for enhancing BI visual analysis tools with storytelling capabilities.
- (7) The design and implementation of a prototype storytelling tool that is fully incorporated in an existing visual analysis dashboard, to fluidly transition between analysis and story creation. We report on feedback from BI experts creating a story on the usefulness of the prototype as a communication and teaching medium, and from BI novices reading a story on its effectiveness for story communication.

1.5 Thesis Outline

CHAPTER 2 begins by covering related work. The remainder of the thesis is organized into three areas: (1) Understanding Dashboard Creation and Customization for Visualization Novices, (2) Annotating BI Visualization Dashboards, followed by (3) Storytelling in Visual Analytics tools for Business Intelligence.

In *CHAPTER 3* we introduce Exploration Views (EV) a dashboard prototype build specifically to empower end users, following guidelines from previous work on visualization novices, and evaluate its effectiveness. Finally, we derived a set of additional design guidelines for dashboards targeting novice users.

CHAPTER 4 investigates the use on annotations on data visualization, and present new annotation requirements for BI dashboards that are supported in our prototype of context-aware annotations. We discuss challenges in supporting context aware annotations in dashboards and other visualizations, and we report a user study and evaluation for the prototype, conducted with BI domain experts, and present faced challenges.

In *CHAPTER 5* we discuss the importance of adding storytelling mechanisms in dashboards. We report in-depth interviews with BI expert analysts and collect new requirements related to BI visual storytelling. We design and implement a storytelling prototype tool that is integrated in a dashboard analysis tool based on BI experts requirements. We present how his prototype that was evaluated by experts, helps BI analysis authors to organize and communicate their findings faster, and reach a broader audience (e.g. novices without any domain training).

1.6 Summary

The amount of data in our world has been exploding, and successfully analyzing large data gives an edge to both companies as well as individuals. BI helps better analyze large data sets, and one of the more important tools available to BI analysts are visualization dashboards. Nevertheless, these dashboards remain hard to create and customize by individuals with little training. Moreover, they still fall short in supporting detailed analysis record keeping, and communication of analysis results to others. This thesis focuses on enhancing user interaction with BI dashboards, through (1) easy creation & customization of dashboard for a broader user range, (2) adding novel annotation support for analysis, and (3) use of visual storytelling for communicating dashboard analysis results.

Chapter 2

Related Work

2.1 Abstract

In this thesis we focus on three aspects of interaction with BI dashboards, beginning with (1) dashboard creation and customization. Through the related work we show the limited focus of research in supporting novice users, and how current dashboard creation practices requires an expert to create a suitable dashboard for a consumer user, involving multiple customization iterations between the two, which limits dashboard use from reaching broader groups of users. (2) Second we show how annotations are considered as an important aid in record keeping and data discovery in different contexts such as text, image & data visualization, but how similar research for supporting BI visualization annotation, more specifically dashboard annotations, is lacking. (3) Third we focus on communicating business analysis results to others, and how storytelling that has been used with other data visualization domains can inspire similar work in dashboards.

2.2 Introduction

This chapter reviews the background work related to this thesis, organized in the three main areas studied by this thesis: dashboard construction and customization, visualization annotation, and narrative visualizations (storytelling).

2.3 Dashboards

There are dozens of software packages that mine data repositories and present details on how a business is performing. Sometimes called analytics or business intelligence software, they aggregate data from disparate internal and external sources and display it in the form of customized views. The fashionable term for these views is digital dashboards.

— Jeffrey Schwartz[41]

In this section we will present related research in the fields of dashboard creation and visualization construction, especially in the context of novice visualization users.

2.3.1 Dashboards creation

Few [30] has identified the characteristics of a well-designed dashboard, that can be used to *evaluate* dashboards. Information should be

1. Well organized. Dashboards should help people to immediately recognize what needs their attention.
2. Condensed, primarily in the form of summaries and exceptions. Summaries represent a set of numbers (often a large set) as a single number such as sums and averages, while exceptions represent something unusual happening that could be a problem or an opportunity. Thus, viewers don't wade through hundreds of values when only one or two require attention.
3. Specific to, and customized for, the dashboard's audience and objectives.
4. Displayed using concise and often small media that communicate the data and its message in the clearest and most direct way possible.

The majority of the commonly used commercial BI dashboards (e.g. Dundas [42], OracleBI 10g [43], SAP BusinessObjects dashboard [44], see (Figure 2.1)) assume that from the

end-user perspective the dashboard design will remain consistent, and focus more on easy report generation.

To create new or customize existing dashboards, end users need to either contact trained dashboard designers, or install separately provided dashboard design components and spend time to train themselves. There are some notable exceptions. Tableau [36] incorporates the design functionality in the dashboard system. Spotfire [35] also accommodates novice users with a flexible drag-and-drop interface for customizing visualizations on the fly. Finally, the prototype Choosel environment [45] allows for easy visualization creation and further exploration of the visualized data, but does not yet provide advanced functionality (e.g. hierarchical data representation as in OLAP data models [46], with drill-up/down functions), required in real life business data analysis. Thus with few exceptions, dashboards are hard to create and customize.

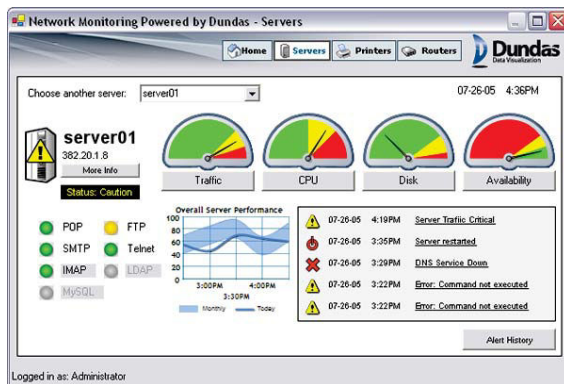
2.3.2 Visualization creation

A variety of visualization toolkits allow the creation of sophisticated visualization charts and combinations of them (e.g. Flare1, Silverlight2, JavaScript InfoVis tk3 [47], ivtk4 [48], D3 [49]). These do not target novice visualization users, and usually require additional programming to create the visualizations, to synchronize them in dashboard views, and to connect them to data-sources.

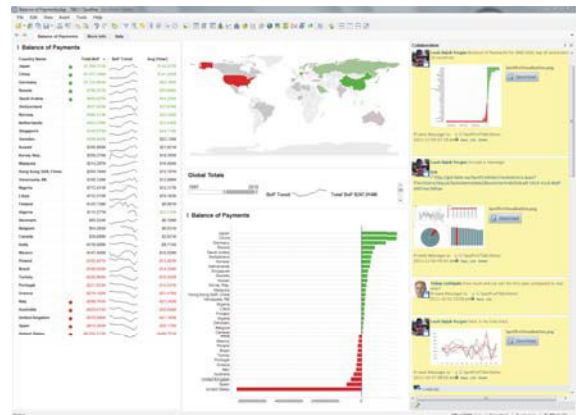
Several environments allow the construction of highly customized, powerful and fully linked visualizations. For example Improvise [50] (Figure 2.2) provides dynamic queries where users can interactively explore and change parameters of representations. Snap-Together [51] (Figure 2.3) allows the user to mix and match visualizations, load relations into visualizations then coordinate them based on the relational joins between them to construct custom exploration interfaces. (For a comprehensive and more detailed list of work on coordinated views see [25]). Nevertheless their setup cost and learning curve is usually high and they target visualization experts rather than novice users.

2.3.3 Visualization for novices

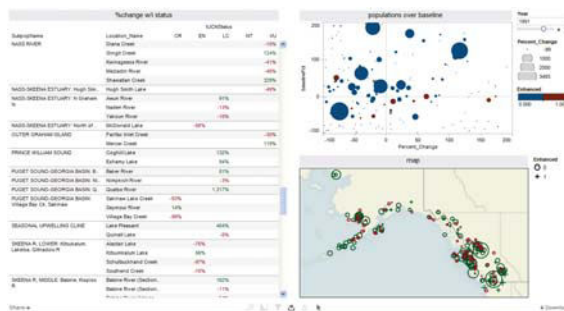
On the other hand, there is a set of visualization platforms that specifically target novice users. For instance, the Google analytics platform [28] provides a highly customizable dashboard with good default visualization charts for web-related data. Other web-based visualization platforms easily accessible to novice visualization users, like ManyEyes [37], Sense.us [52], or Polstar [53], usually restrict users to a viewing and sharing a single visible



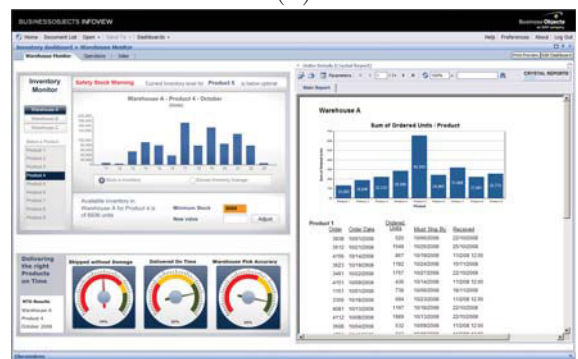
(a)



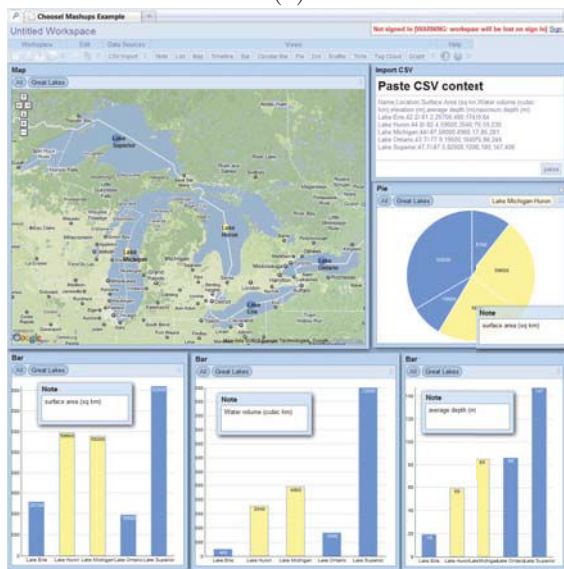
(b)



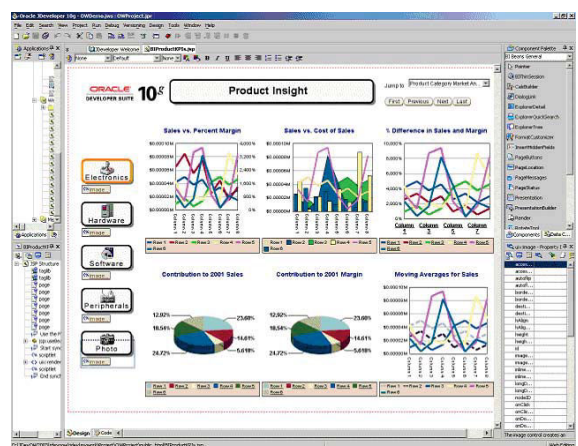
(c)



(d)



(e)



(f)

FIGURE 2.1: Commercial BI dashboards: (a)Dundas [42], (b)Spotfire [35], (c)Tableau [36], (d)SAP BusinessObjects dashboard [44], (e)Choosel [45], (f)OracleBI 10g [43].

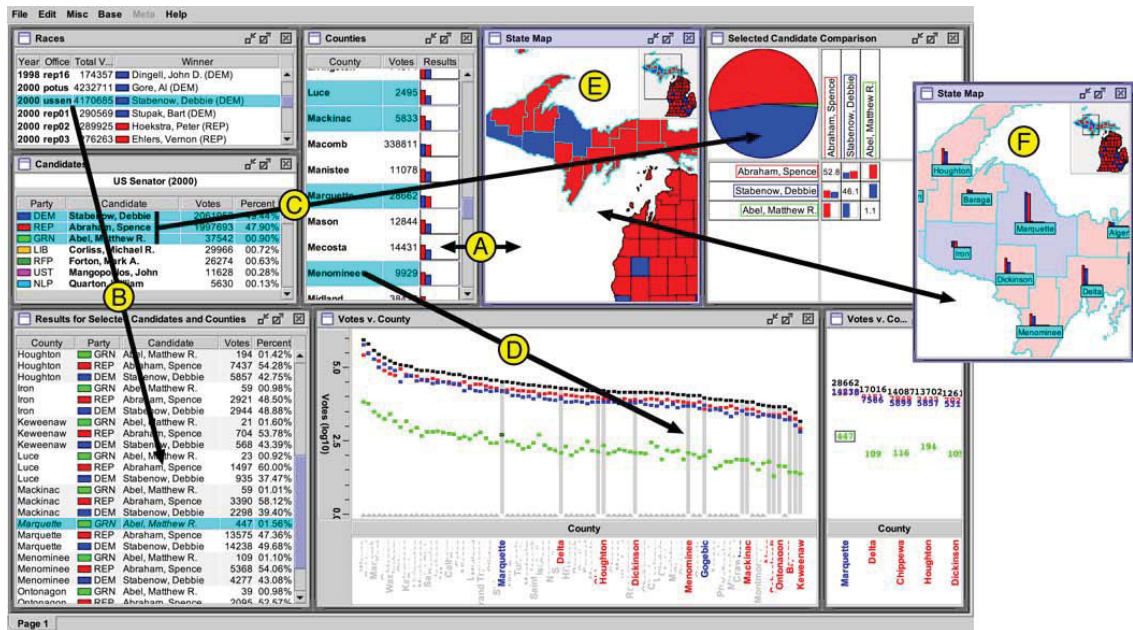


FIGURE 2.2: Improvise [50]. The visualization represents the election results in Michigan from 1998 to 2002. (A) Selection of counties visualized in table view and a map. (B) When selecting a race from races visualization, the election results for that race will be loaded (from a file) and shown in the "Results for selection candidates and counties" visualization. (C) When selecting from the "Candidates" table, the pie chart will be filtered to compare results for selected candidates only. (D) Selected counties in the scatterplot visualization are highlighted with gray bars. (E) A four-layer scatterplot colors counties by winning candidate party. (F) A semantic zoomed in "state map" visualization labels counties with nested bar plots.

chart at a time. Thus creating a dashboard is a laborious process performed outside the platform, and linking of multiple charts is no longer possible.

2.3.4 Novices in visualization

With the increase of visualization platforms targeting novices, researchers have started gathering insights into their practices and behaviors. Heer et al. [39] group different visualization systems by how skilled their users have to be, and note that tools supporting detailed data exploration target mostly experts, whereas it is mostly communication/reporting systems that accommodate novices.

They also suggest that visualization tools for novices should:

- (G1) Allow user-friendly data input for common data formats.
- (G2) Automatically provide visualizations or reasonable defaults based on data types.
- (G3) Use contextual information to clarify the displayed data and encoding.

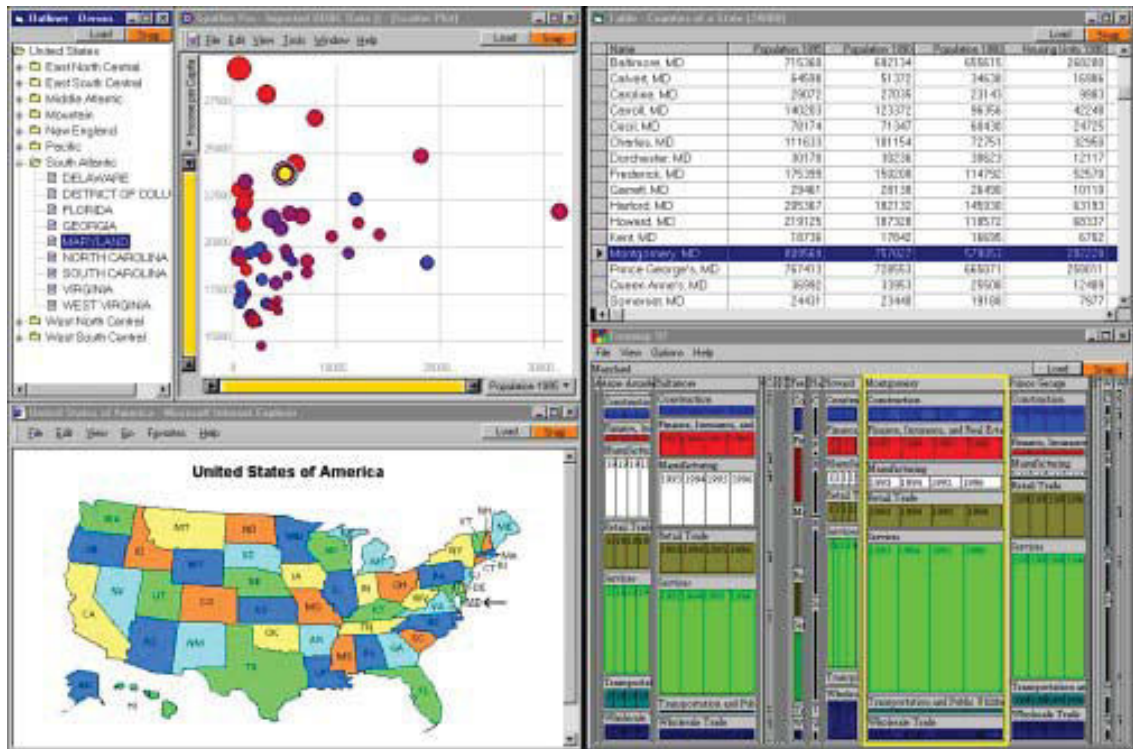


FIGURE 2.3: Snap-Together [51]. A coordinated visualization environment constructed using Snap-Together Visualization. This visualization represents the Census data of U.S. states and counties. When selecting a state from the left side visualizations, displays detailed county and industry information for that state in the table and Treemap on the right.

Grammel et al. [38] investigate how novices create single charts, using a human mediator to interact with the creation interface. They found that the main activities performed by novices (apart from communicating their specifications to the mediator) are data attribute selections, visual template selections, and viewing and refinements of the above.

They observed that subjects often had partial specifications for their desired visualization, and faced three major barriers: selecting correct data attributes, choosing appropriate visual mappings, and interpreting the visual results.

They also verified suggestions from [39], and provided additional guidelines, such as:

- (G4) Facilitate searching of attributes.
- (G5) Automatically create visualizations when possible.
- (G6) Provide explanations.
- (G7) Promote learning.

Our work follows the above seven guidelines from Grammel et al. [38] and Heer et al. [39], and extends those design guidelines for BI information dashboards, respecting Few's [30]

dashboard *creation* guidelines:

- Fit a great deal of information into a small amount of space, resulting in a display that is easily and immediately understandable.
- Base efforts on a firm understanding of how people perceive and think, building interfaces, visual displays, and methods of interaction that fit seamlessly with human ability.
- The ability to see everything you need at once in one view instead of multiple views, without fragmenting data.
- Try to present all within eye span. Something critical is lost when losing sight of some data by scrolling or switching to another screen to see other data.

We build a fully functional dashboard visualization system, and observed how novice users use the interface to create dashboards (without a human mediator [38]). In *CHAPTER 3* we report our observations of novices and experts reactions, and derive additional design guidelines for visualization dashboards that target both user groups.

2.4 Annotations

An annotation is a commentary on an object, in order to add information. This information could be represented by text, link, photo, audio or a video. It is a common practice during long term human activities to take notes and annotate documents in order to indicate or highlight interesting items that need further investigation, mark items that have been completed or remain to be done or add explanations and insights to facts. Current visualization and BI systems provide minimal (if at all) annotation capabilities in the form of simple text. As BI dashboards are highly visual in nature, there is a need to support annotations and comments on the visualizations.

In order to understand the challenges of annotating on dashboards, determine the best annotation suitable support, and ways to navigate through these annotations, we report on existing digital annotation scenarios in recent research.

2.4.1 Annotation of Textual data

An annotation is frequently used to support critical thinking and learning while reading, a process called "active reading" by Adler et al. [54]. Marshall [55] has examined the practice of annotation on textbooks to help designing annotation tools, specifically for readers working in a digital library tool. This study gave a clear identification of annotation usage, such as:

- *Annotations as procedural signals*: annotations are used in anticipation of future attention, to designate reading assignments, responsibility for "knowing", and desire to reread.
- *Annotations as place-markings and aids to memory*: a way of remembering or remembering to remember.
- *Annotations as in situ locations for problem-working*: markings were used to indicate action (or to help visualize action) to take at a later date.
- *Annotations as a record of interpretive activity*: as notes of interpretations of unfamiliar language, interpretation of structures, or adding insights.
- *Annotations as a visible trace of reader's attention*: when the material is difficult and in narrative form; attention is easier to maintain if the material is relatively accessible.
- *Annotations as incidental reflections of the material circumstances*: they reflect circumstances that are entirely outside of the realm of the text.

Moreover, text annotations can be used to support feedback exchange through discussion threads, that can benefit Information Retrieval systems [56].

These important practices of text annotation for reading support, are translated to digital libraries using "annotation systems" such as Annotea [57]. Digital annotations can persist throughout the lifetime of a document, and can easily be shared. But the variation of reading platforms poses a unique challenge in the adaptation of freeform pen-and-paper annotation to the digital domain: Each time the content of a digital document reflows to a new layout, any digital ink annotations must also reflow to keep up with it as done in Barger et al. work [58].

The annotated data in this context is text, so it will appear in the form of letters on lines only, while in chart visualization, data has different visual representations (visualization types) and different granularities.

2.4.2 Annotation of Image Data

Image annotation deals with annotating and commenting directly on an image. This means that the annotation is associated with a collection of pixels and is usually stored as a layer on top of images, such as in I2net [59]. When accessing annotations on large image corpora (in the form of images), one of the most important tasks is the need to find the images that have specific objects or to find keywords that best describe the image content [60]. Hence, this raises the possibility of several interesting applications such as image annotation with the semantics of the picture, to help in retrieving images in response to textual queries that correspond to its annotations semantics [61].

Libraries have traditionally used manual image annotation for indexing and then later retrieving their image collections. However, manual image annotation is an expensive and labor intensive procedure and hence there has been great interest in coming up with automatic ways to retrieve images based on content such as [62][63][64].

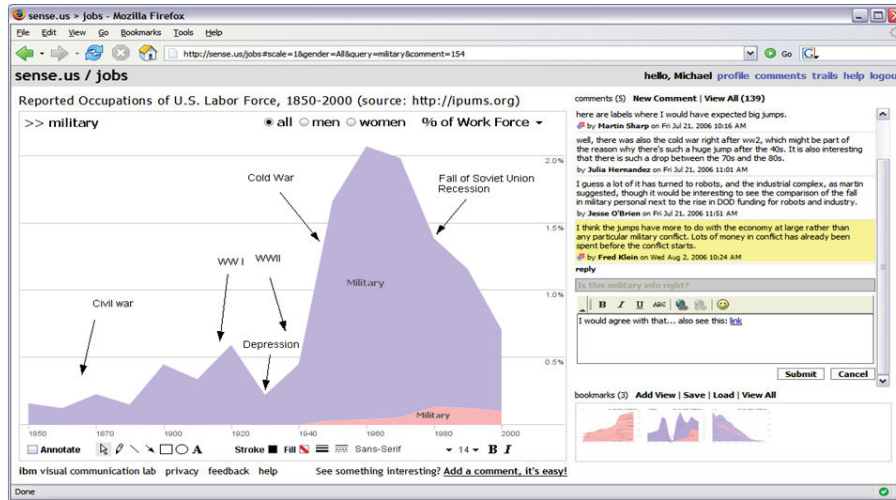
2.4.3 Annotation of Visualizations

To help analysts track their analysis process and findings, researchers have suggested the use of history tools (e.g. Aruvi [65], Tableau [66], CzSaw [67], Visage [68]) to review, re-visit and retrieve previous analysis steps [69]. Augmenting visual representations with annotations to record analysis details and findings has also been recommended by Heer et al.[52] and implemented in information visualization systems such as: Voyagers and voyeurs [52], HARVEST [70], and GeoTime[71], allowing analysts to re-visit their analysis and re-use previous findings [72].

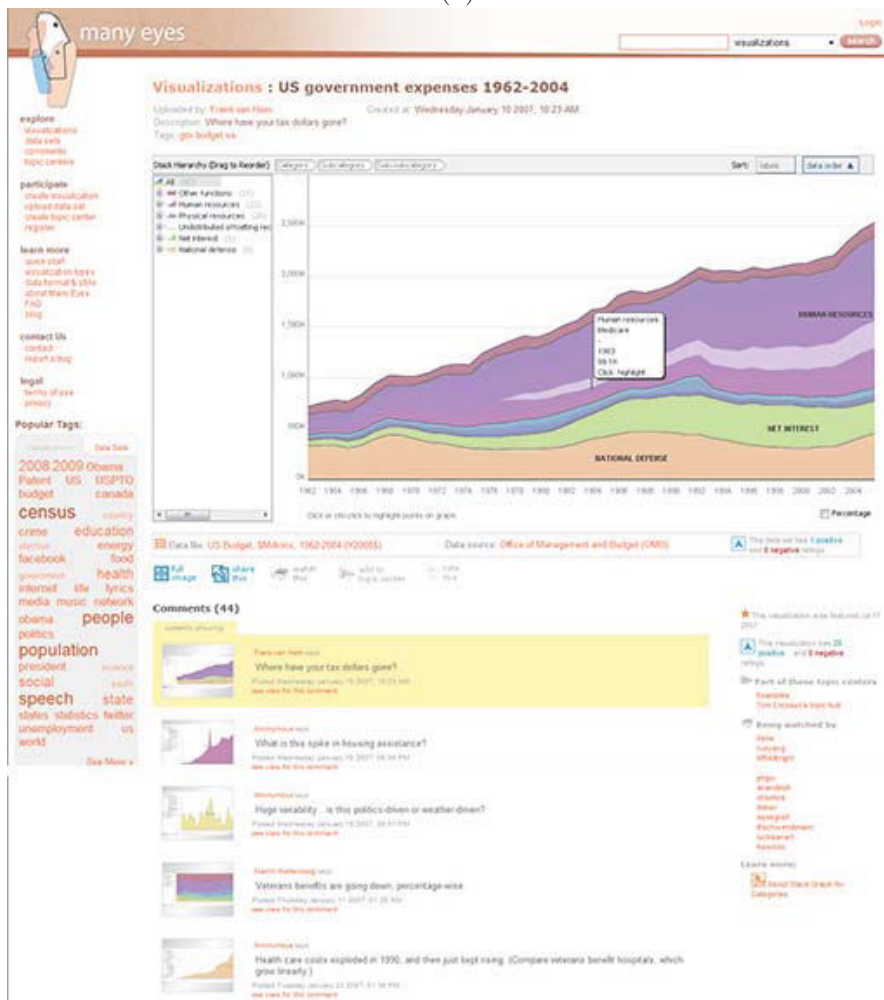
Marshall [72] identified three main processing activities during analysis: searching, interpreting, and presenting. Pirolli and Card [73] refer to the first two activities as the information foraging and sense making loops. Annotations are traditionally used to support interpretation (sense making). Visual Analytics (VA) systems like ManyEyes [37], sense.us [52], and CommentSpace [74] allow multiple users to annotate their insights on a single visualization, and search through them, encouraging peer learning and social sense making [75]. Others, like Sandbox [76] and VisualLinks [77], aid users to collect and group multiple discoveries and insights on a single visualization, and record them in the form of annotations. These support the sense making process of framing relevant information together, clarifying connections and information importance [78]. During analysis, fragmented information pieced together often reveals connections and further opportunities for investigation [79], contributing to the searching (foraging) aspect of sense making. Thus annotation support is clearly important in visual analysis.

2.4.4 Context-Aware Annotations in Visualizations

Recent VA systems have moved beyond this simple information grouping. Annotations are no longer simply attached to visualizations or chart graphics, but to the data, moving towards more context aware annotations. These annotations can then be used explicitly to search (forage) for related visualizations or annotations. The sense.us [52] system provides doubly linked discussions, allowing users to move from visualizations to annotations and vice versa. Aruvi [65] allows users to arrange their annotations in a "mind map", a diagram linking related visualization views and their annotations. Users can access past analysis



(a)



(b)

FIGURE 2.4: Visual Analytics (VA) systems that supports annotating a single visualization: (a) sense.us [52] (b) ManyEyes [37]

views using keyword searches and text similarity metrics. Shrinivasan et al. [70] recommend related annotations based on the context of the current analysis. Finally, Chen et al. [80] use annotation context as a way to semi-automatically annotate facts belonging to the same analysis categories across visualizations.

Most existing VA work supports context aware annotations on a single visualization view. Traditionally BI analysts conduct their analysis using complex visualization dashboards. Dashboards visualize multi-dimensional datasets through a combination of linked charts [5], and are very widespread in BI (e.g. [42][43][44][35][36]). Annotations are supported in most dashboards, and are usually attached to data points (not the chart graphic). Nevertheless, these systems do not adequately support several needs identified by BI experts (detailed later in *CHAPTER 4*), such as annotations spanning multiple charts, or annotation transparency across data dimensions and visualizations (annotate once, see everywhere, as is done in text [81]). These advanced functionalities, such as multi-chart annotation, can be hard to achieve, as annotating multiple points is an involved process: since only one visualization is shown at a time, an annotation can only be added to one query. To add more queries to an existing annotation, the user needs to continuously switch visualizations.

This body of research uses context aware annotations for different purposes. Our work extends it in two ways: based on expert users' comments we identify desired behavior and possible uses of context aware annotations on visualization dashboards; and we explore challenges in their design and use, for example due to the dynamic nature of "context", discussed in *CHAPTER 4*.

2.5 Storytelling

Stories are series of events and facts, with an order and connections between them (Oxford English Dictionary). In the domain of intelligence analysis, it is furthermore an abstraction used by analysts to conceptualize threats and understand patterns in the analytical process [82], and communicate their findings to others [83].

2.5.1 Stories in Business

In recent years, organizations and their leaders have identified the importance and value of narrative and anecdotal information conveyed in the form of stories [84], to present either complex ideas, the context and details of crucial information, or personal interpretation of that information [85].

Research conducted to date has demonstrated the value of storytelling to improve organizational structure [86] and collaboration quality [87, 88], socialization and adaptation of new employees [89–91], organizational and financial success [92, 93], innovation and new product development [94], and teaching and learning [95].

The majority of this work is a meta-analysis of the effect of storytelling within an organization, rather than the development of tools to enhance the storytelling process as is our case. Moreover, the stories themselves discussed in this work, relate to the transmission of information and knowledge *within an organization*, mostly in *textual or verbal form*, rather than in visual form. The widespread use of visualization dashboards in the domain is a more recent development [5], and so is the transmission of knowledge between organizations (dedicated BI analysis organizations and their clients or end-users). Thus storytelling needs in the domain have evolved and in our work we take a step at characterizing them more precisely.

2.5.2 Stories in sense making

Baber et al.[83] point out that contemporary theories of sense making rely on the idea of 'schema', of a structure to organize and represent factual information, as well as the knowledge, beliefs and expectations of the people who are engaged in sense making. They can thus be considered as a collection of narratives. They further discuss the formalism of stories in sense making, and how the most effective stories are organized around the actors in the stories, their actions and rationale, events, their context, and most importantly the relationships between these. As further argued by Bier et al.[96], for effective collaboration and communication we need to use less text, and organize knowledge around entities (people, places, things, times etc.) rather than free form text. Similarly, Pirolli & Russell [97] propose the mapping of intelligence facts and insights into frames, that can be expressed in a variety of forms including stories, maps, organizational diagrams or scripts. It is thus clear that conducting intelligence analysis, communicating findings, and organizing knowledge in stories, has a strong visual component that represents entities and their connections.

2.5.3 Stories in data visualization

Often in narratives text or audio transmit the main story, while visualizations or images support the story or provide details. Comics and flowcharts are special types of narratives relaying mostly on visual components rather than text. Recently, we have seen an increase in integrating complex visualizations into narratives in many news organizations [98], journalism reports (e.g. New York Times, Washington Post, the Guardian), and television

reports [99]. Segel et al.[1] explore different aspects of narratives from a variety of sources and identify distinct genres of narrative visualization, see Figure 2.5.

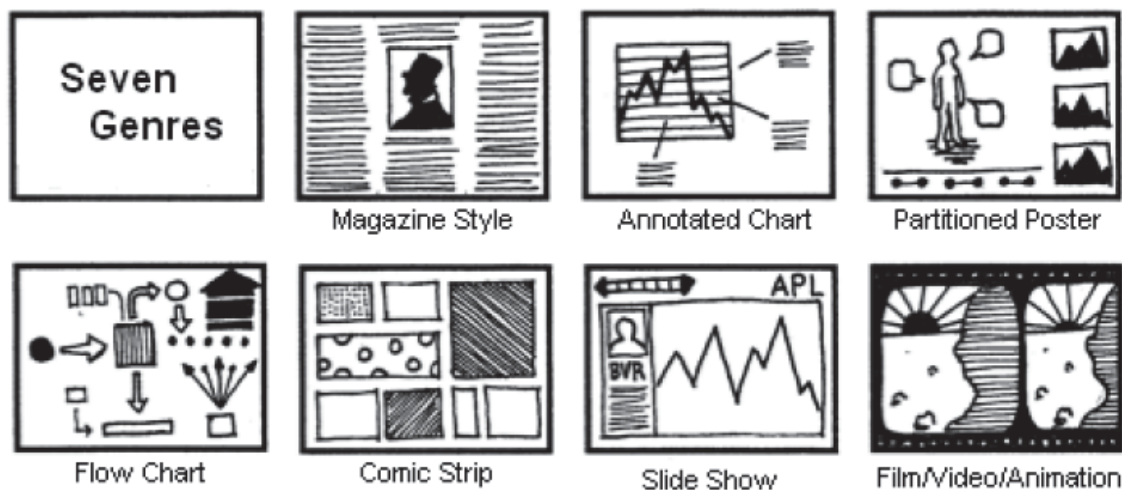


FIGURE 2.5: Types of Narrative Visualization proposed by Segel et al. [1]

In the domain of business intelligence the main visualization tool is BI dashboards [28], collections of multiple visual components (e.g. charts, tables) on a single view [100]. Nevertheless, as pointed out by Wojtkowski and Wojtkowski [101], dashboards and other visualization tools used to analyze complex content cannot simply be used to tell stories. They need to be “tailored” to accommodate storytelling so as to better highlight items of importance within very large data resources [102], in a way that is not cumbersome for the storyteller and clear for the audience.

Some visualization systems began to integrate storytelling. For example GeoTime [103], a geo-temporal event visualization system, integrates a story system that shows events in space and time, hypertext linked visualizations, and visual annotations to create an environment for both analytic exploration, and story creation and communication Figure 2.6.

Storytelling tools in the business domain are not yet as advanced. Systems like Sense.us [52] and Tableau [36] allow analysts to visualize their data, conduct their analysis, and store a history of the analysts exploration. This history can serve as a first step towards creating a story. Many Eyes [37], Tableau Public [104] and Sense.us [52] allow publishing of interactive visualizations online, and permit collaborative analysis through comments on a single visualization, creating an evolving analysis. Again this collaborative annotation can be seen as are a first step towards making a collaborative knowledge narrative, where an analysis story could be extracted from the visualizations and comments. Nevertheless, these tools do not provide explicit means to indicate story progression, and to highlight or explain relationships between multiple visualizations (seen in dashboards) which is key to intelligence analysis and communication [83].

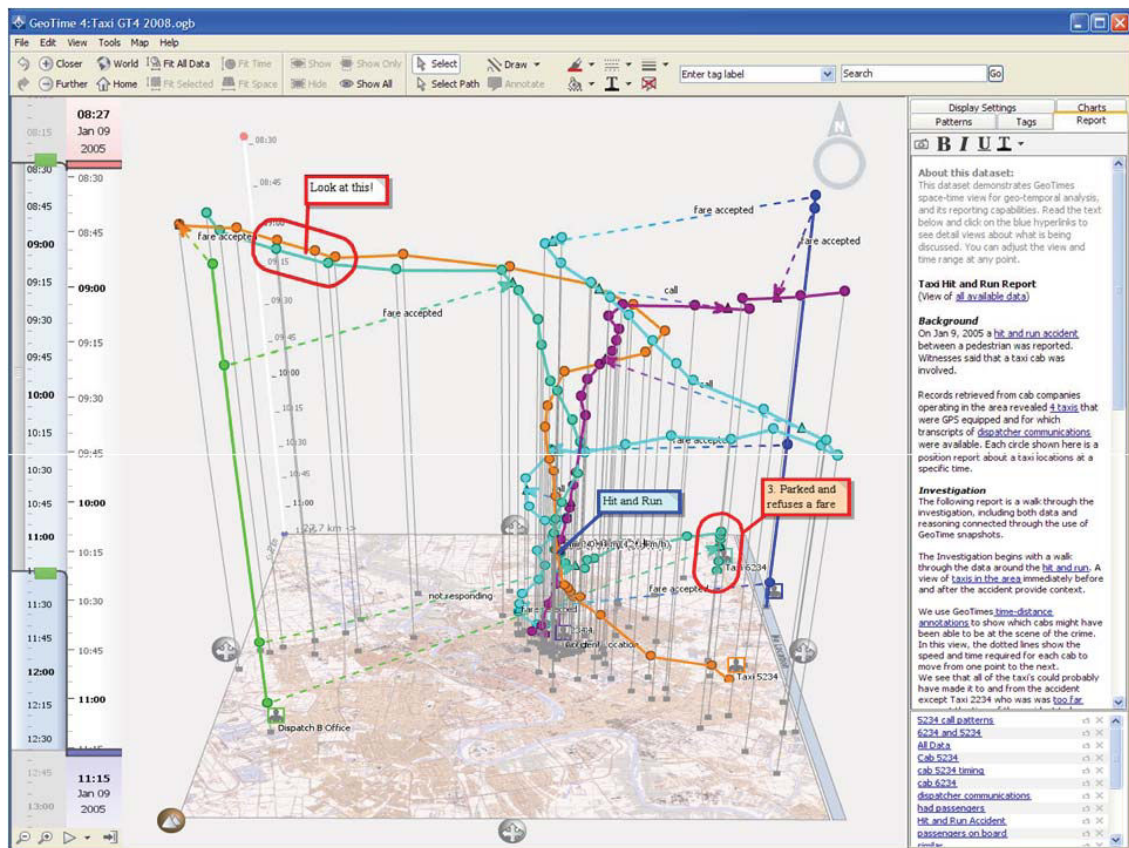


FIGURE 2.6: Geotime shows events in time and space, in a X, Y and T coordinating space. [103]

2.6 Summary

In this chapter, we have presented the related work to our three main research foci: (1) dashboard creation and customization for novices. We saw few systems that provide novice users with an easy to use interface to create dashboards. Moreover no work has studied novice user interaction with dashboards specially in the BI domain. (2) For enhancing dashboard analysis with annotations, we have presented the broad research done for annotating text, images and visualizations. The important functionalities in annotating text of "annotating once, seeing it everywhere", is not yet applied to visualization annotation. Rather, very often annotations were attached to images (visualization snapshots), and users loose their annotation if they change the data visualization type. Moreover, advanced functionalities such as multi-chart annotation is still missing or hard to achieve. (3) The third point we focused on, is how storytelling is important for presenting complex information in business, by organizing and representing factual information supported by strong visual component that represents entities and their connections. Then we presented some systems like GeoTime that supports storytelling, while we saw that storytelling tools in the business domain are not yet as advanced.

Chapter 3

Understanding Dashboard Creation and Customization for Visualization Novices

3.1 Abstract

With the increase of visualization platforms targeting novices, researchers are now focusing on gathering insights regarding novice user practices. This chapter focuses on novice and expert business users that require access to easily created and customized dashboards. We describe the design and evaluation of Exploration Views (EV), a system that allows novice visualization users to easily build and customize Business Intelligence information dashboards. EV provides an intuitive environment for dynamically creating, rearranging, searching and exploring multiple visual data representations from diverse data-sources. These aspects aid users to better retrieve, experiment and familiarize themselves with their data. We evaluated EV with both novice and expert dashboard designers and report here (i) how novice users interact with the system, (ii) differences in how novice and expert users react to a dashboard systems that targets both, and (iii) provide new design guidelines for practitioners building dashboard applications, on the needs of novice visualization users.

3.2 Introduction

Appropriate visual representations in dashboards, using colors, size and shape, are combined with interactive exploration [105] to amplify human cognition and enhance information understanding [13]. The current life-cycle of a BI dashboard involves multiple actors [6], including end-users and business analysts. In user-centered BI [7] end-users intervene and provide feedback to the business analysts that create customized dashboards to meet user needs. This feedback comes at different stages of the dashboard design and setup (Figure 3.1), and involves a large amount of communication between business analysts and end users, in order to define functional specifications and a positive user experience.

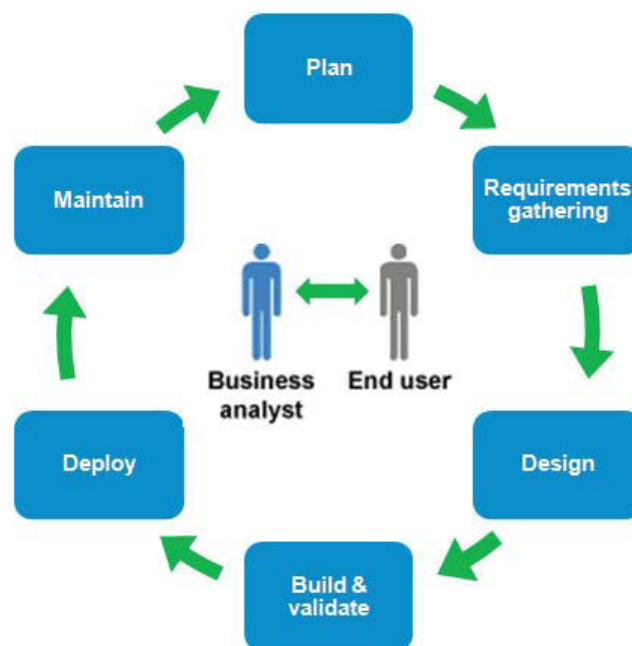


FIGURE 3.1: Current BI dashboard creation process involves multiple actors, end-users and business analysts, who require a large amount of communication in different stages of dashboard design.

Thus there is an intrinsic delay introduced in any end-user requirement change or customization request. Current market trends recognize dashboards and data analysis as one of the driving forces in business and are becoming more popular. This indicates an expanding end-user population with diverse needs, requiring quick access to customizable dashboard technology, ideally bypassing the cost of communication with dashboard designers. This is reflected in advances in easily customizable dashboard visualization systems (e.g. Spotfire [35], Tableau [36]), and can also be seen in related research in the InfoVis community [38],[39].

Nevertheless, there are still many questions about how novice users interact with actual information visualization systems. To better understand how novice users interact and build visualizations on their own (without dashboard designers or other human mediators) we created Exploration Views (EV) in collaboration with SAP engineers. EV is a dashboard prototype (Figure 3.2) build specifically to empower end users, following guidelines from previous work on visualization novices. To evaluate its effectiveness, we compared how dashboard experts and novices reacted towards EV when creating visualizations for realistic tasks.

3.3 Exploration Views (EV) Prototype

This section describes the design rationale behind Exploration Views (EV) and the functionality that supports novice users, based on guidelines from previous work Section 2.3.4. We will first explain the main goals of EV: easy creation, customization and interaction with multiple visual representations in a unified dashboard environment.

Although dashboard users may have clear questions regarding their data, they can be novices when it comes to visualization, and naive as to how dashboard software can support their needs. An easy creation process is essential for novices to experiment with dashboard designs fast. As seen in Grammel et al. work [38], visualization novices often have partial mental specifications for their visualization needs and tend to refine and change their designs. To ensure a user-friendly dashboard creation, the sequence of steps needs to be simple, with continuous visual feedback on the final outcome of user choices. Moreover, as novice users may have no previous knowledge of what visual templates and representations are possible for different data types, EV must provide chart suggestions and templates to choose from (also suggested by Grammel et al. [38], Heer et al. [39]). Finally, EV should support common data formats (recommendation by Heer et al. [39]), and users should not be restricted by data storage architecture, but be able to use data from multiple data-sources (e.g. personal data in excel and OLAP data from their enterprise).

As novice users create dashboards, they may need to explore and try out alternative visual templates and representations to learn what meets their needs. Requirements and tasks can also change over time, requiring different data representations and data filtering mechanisms, or new data sources. EV should support iterative visualization specifications

(suggested by Grammel et al. [38]) by being easily customizable and adaptable. It is important for novices to see the effect of their choices immediately, promoting exploration and experimentation with visualizations. Thus in EV customizations become active immediately, ensuring that customization is tightly coupled with the visual analytics process (recommended by Grammel et al. [38]). EV further supports the visual analytics process with functions such as saving and sharing, text search mechanisms, fully linked visualizations and visual queries, and other data exploration mechanisms. Finally, if the design of EV supports both low-cost experimentation with visual templates and visual analytics processes, it can help novice users to become more accustomed to visualization creation and analysis, promoting learning.



FIGURE 3.2: The Exploration Views consists of: (a) the main Dashboard and its components. (b) The Clipboard drawer widget that includes templates for charts (top) and data widget samples (bottom) that the user has created as a pallet. The drawer can be closed when not needed. (c) A visual representation of the global data filters that are active at any point. (d) And a set of other functions, such as search, save, share, etc.

3.3.1 Easy Creation

When creating a new dashboard the system provides miniature icons of saved dashboards and several template dashboards to choose from (Figure 3.3). The visual representation of possible dashboards helps users recall designs (as opposed to remembering dashboards

solely based on names [106]), and their titles are descriptive so novices can start from a reasonable existing dashboard design, minimizing new content creation. These dashboard templates can be also shared with others. If none of the existing dashboard templates meets the user needs, she can choose to create a dashboard from scratch.

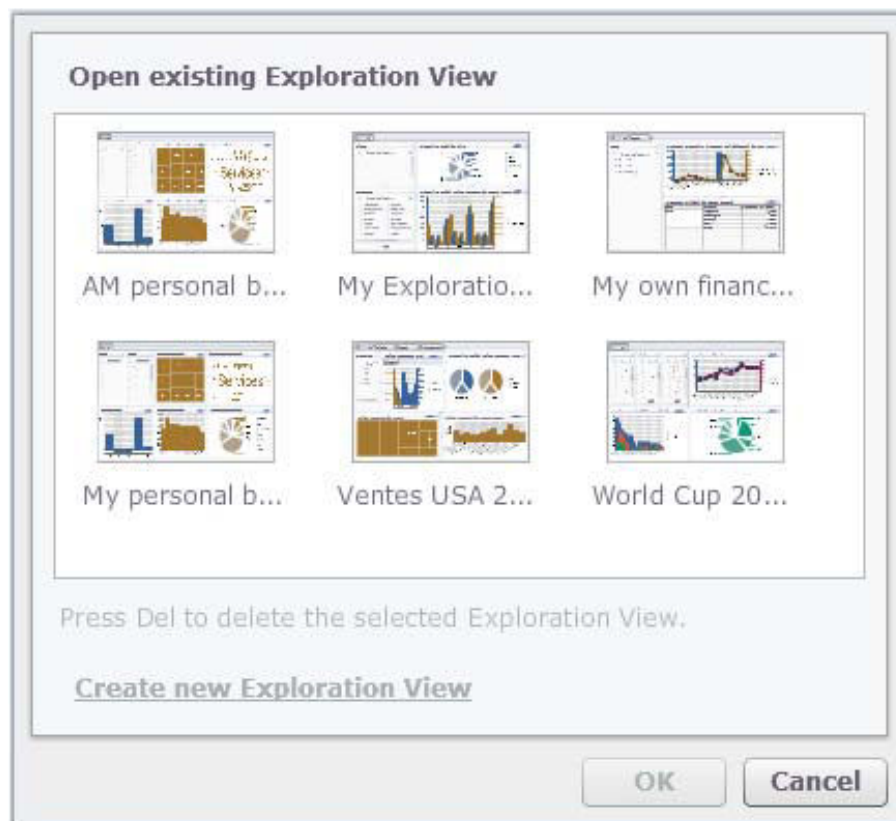


FIGURE 3.3: Creation dialogs with visual examples of templates or saved dashboards.

The user is guided by the system by first asking her to select data sources of various format types that are supported. When the data sources have been selected, the users can populate their dashboards with charts. The dashboard work area (Figure 3.2.a), acts as both the design space for the dashboard, as well as the active dashboard when content is added. Here users can add charts, tables, and data filters. Possible components to add can be found in the Clipboard drawer (Figure 3.2.b), which can be closed when creation (or customization) ends. Components from the Clipboard can be dragged-and-dropped in the work area, and EV automatically realigns existing components to accommodate new additions. When adding charts (i.e. visual templates), users need to select data attributes (e.g. sales, year of sale, etc), and match them to visual mapping (a visual property in the template). This is a challenging task for novice users [38] and EV aids them (as in

[105],[107]) by providing reasonable mapping and visual template recommendations (Figure 3.4).

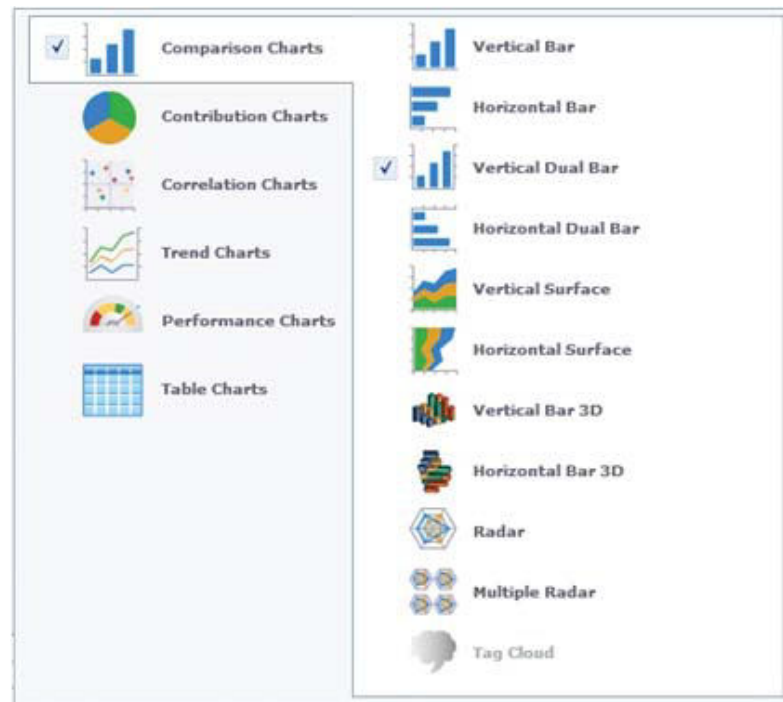


FIGURE 3.4: Chart creation dialog with recommendations of charts appropriate to data categories and analysis types, sorted from simpler to more complex.

After selecting one or more data attributes in the form of measures (data attribute, e.g. sales prices Figure 3.5) and categories (possible groupings of the attribute, e.g. per state, year), the system presents appropriate analysis types for this data. For example for a dataset of sales per state over the period of 2 years, possible analysis types include comparisons (sales by state or year), contributions (the percentage of sales per state over all sales), or trends (the evolution of sales over time). If the user selects one of these analysis types, the system automatically recommends visual templates (charts) suitable to the specific analysis and to the nature of the attributes (e.g. how many measures and categories are selected, if they are categorical, ordinal, etc), using a default visual mapping. These recommendations are ranked from simple and more commonly used charts to more complex ones, and aid novices to explore reasonable alternatives. Available visual templates include charts, gauges and tables.



FIGURE 3.5: Data Category and measures selection.

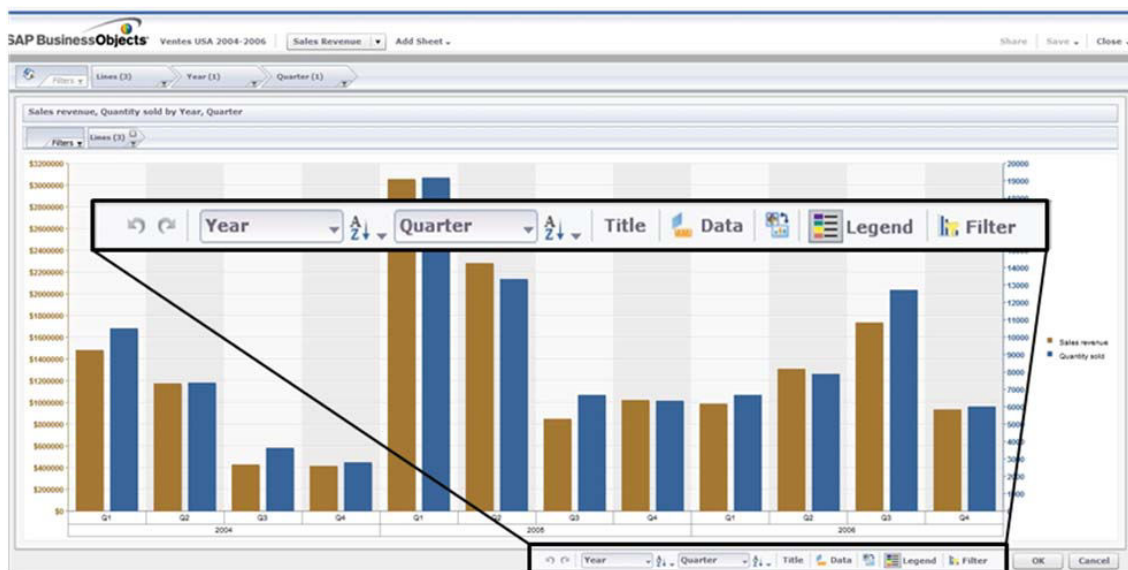


FIGURE 3.6: The "explore" mode is a detailed view of a chart, where users can undo/redo actions, choose data categories, sort items, change data sources, switch chart templates, and create local data filters. To avoid switching to the "explore" mode, part of this functionality is also available as a drawer toolbar for each chart in the main dashboard, with added options to lock the chart.

3.3.2 Ease of Customization

During the life of a dashboard, user needs may change; new hypothesis may need to be explored and new data sources to be added. Especially given novice users' tendency for iterative visualization specifications [38], EV has several features for user friendly customization, such as actions to add/remove components, change their properties, duplicate their content, and explore sub-parts of the data using data filters.

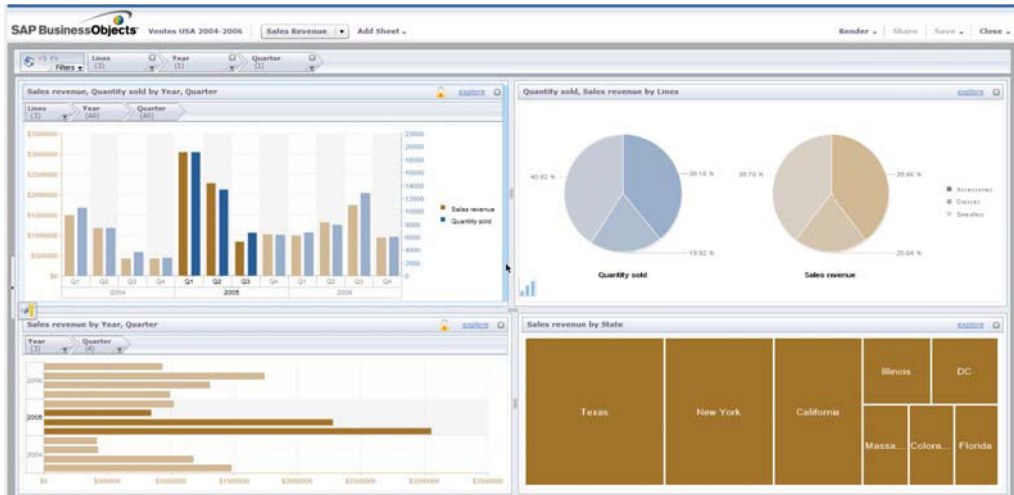
Users can add new components on the dashboard or drag-and-drop existing ones from the Clipboard. If the components are visualization templates, the user has to go through data, attribute and mapping selection as described before. If the components are saved in the Clipboard, we use the properties of the component as it was stored. Existing components on the dashboard can be refined and customized by clicking the "explore" icon of the component. This provides a detailed view of the component (Figure 3.6), together with choices for datasets, visual templates, attributes and mappings (as described in the Section 3.3.2). Changes are immediately reflected on the component in an automatic preview, so that users can experiment with different choices [38]. Users can also customize the dashboard layout itself through functionality for hiding/showing dashboard components, dragging to rearrange them, resizing them, etc.

EV intelligently resizes components to fit in the dashboard (e.g. Figure 3.7). At any point users can save content for later use (in this or other dashboards), by dragging them in the Clipboard Drawer. This allows exploration of alternatives without risking losing existing components. To further ensure that novice users are not afraid to experiment with their dashboards, all actions are reversible through undo history functionality applied on the entire dashboard and locally on specific charts.

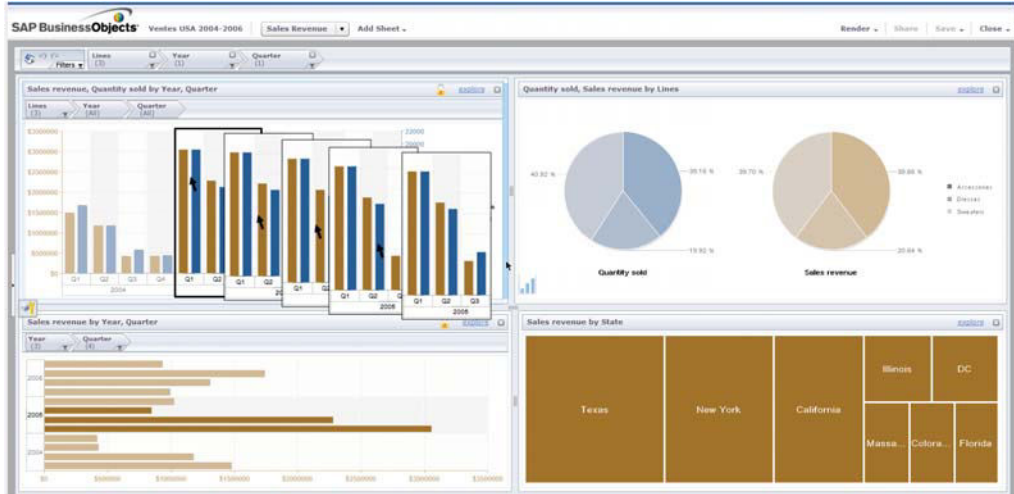
3.3.3 Supporting Visual Analytics Process

Apart from the creation and customization functionality, EV supports a series of functions that aid the visual analytics process. EV has text search support that follows the guidelines of dynamic queries [108]: as the user searches text the corresponding text components on the dashboard that match it (axis labels, categories, tags and any other textual information) get dynamically highlighted. All our visual representations are essentially linked coordinated views [25],[50] of the underlying datasets (unless users give them local filters, discussed later). We thus extend the dynamic query functionality to any visual selection or query in the dashboard: if the user selects in one chart bars representing the 3 quarters of a year, these data instances are also highlighted on all other charts (Figure 3.7.a).

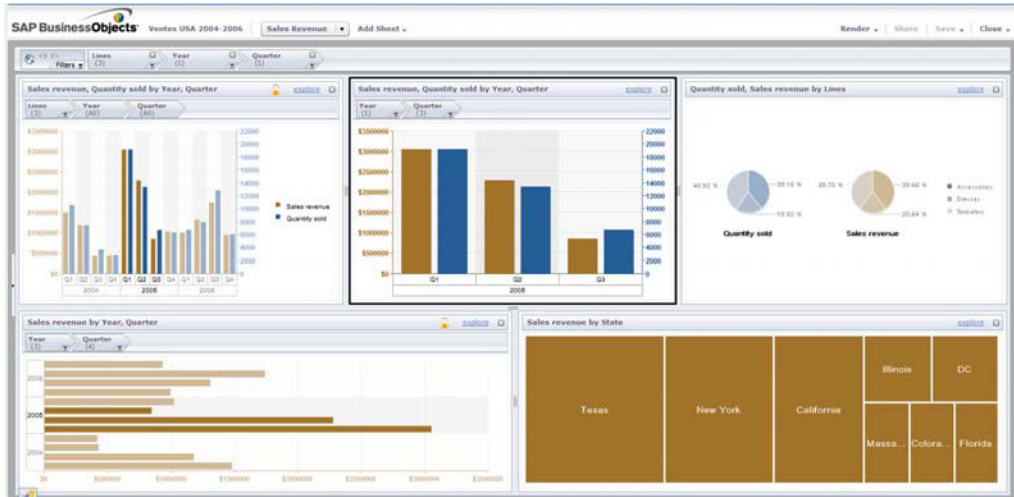
EV further encourages users to explore their data by providing global and local data filters, i.e. constraints on the data entries displayed. Filters can affect the entire dashboard (global), or specific components (local). Thus users can explore "what-if" questions without changing completely their dashboard. Our filters focus on data dimensions: enterprise data usually follow the multidimensional OLAP model [46], where data is organized in hierarchical categorical or ordinal dimensions (e.g. sales grouped per country and state,



(a)



(b)



(c)

FIGURE 3.7: (a) Users select subset of their data, that get highlighted in all charts. (b) They drag their selections outside the chart to create a cut-out clone visualization, of the same type (here bar chart) with the selected data only. A blue highlighted bar indicates the new chart will be placed on the right of the original, with existing charts shifting to accommodate it. (c) The new bar chart of the data subset created and the dashboard visualizations resized to accommodate the new visualization.

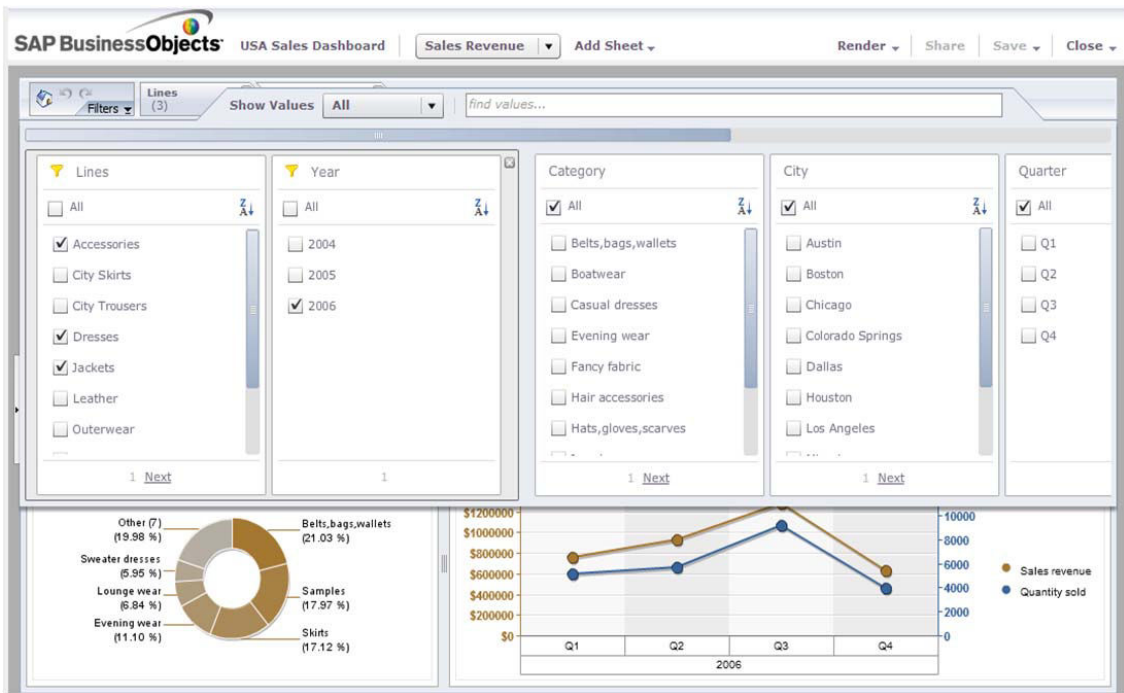


FIGURE 3.8: Filters and Faceted navigation creation. User can select global dashboard filters to apply directly on the whole dashboard charts.

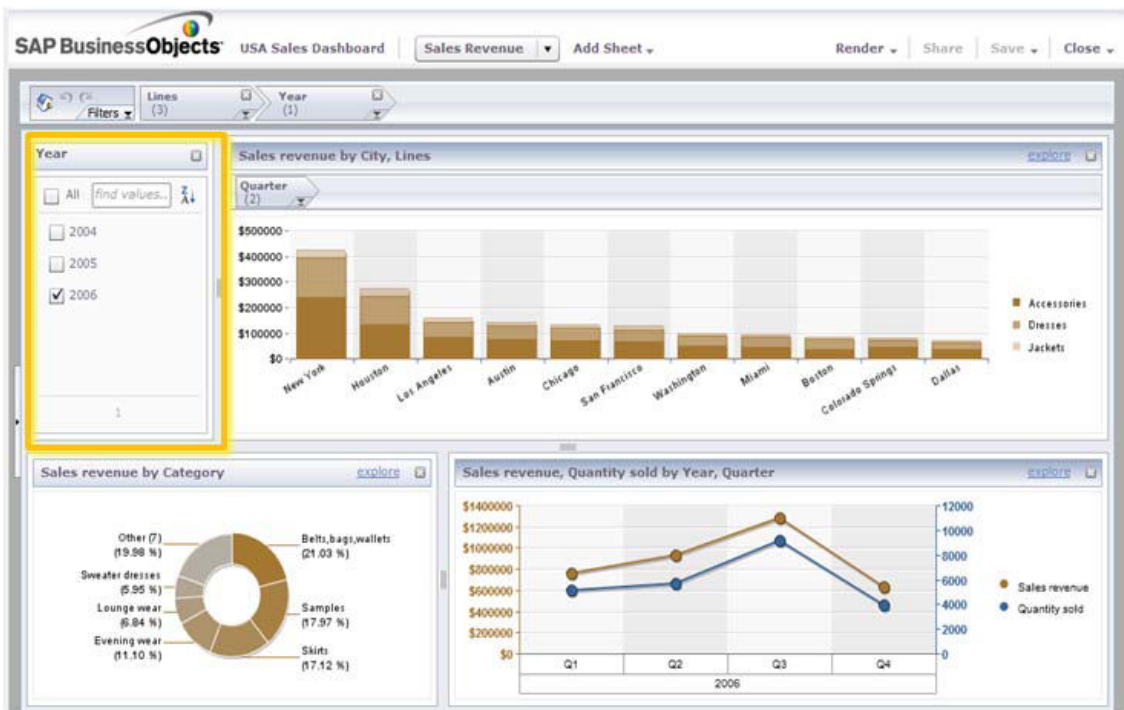


FIGURE 3.9: The user can drag-and-drop a new global filter and choose the data categories that will be visible on the dashboard. The filter is either directly applied, or can become a filter control component for faceted navigation.

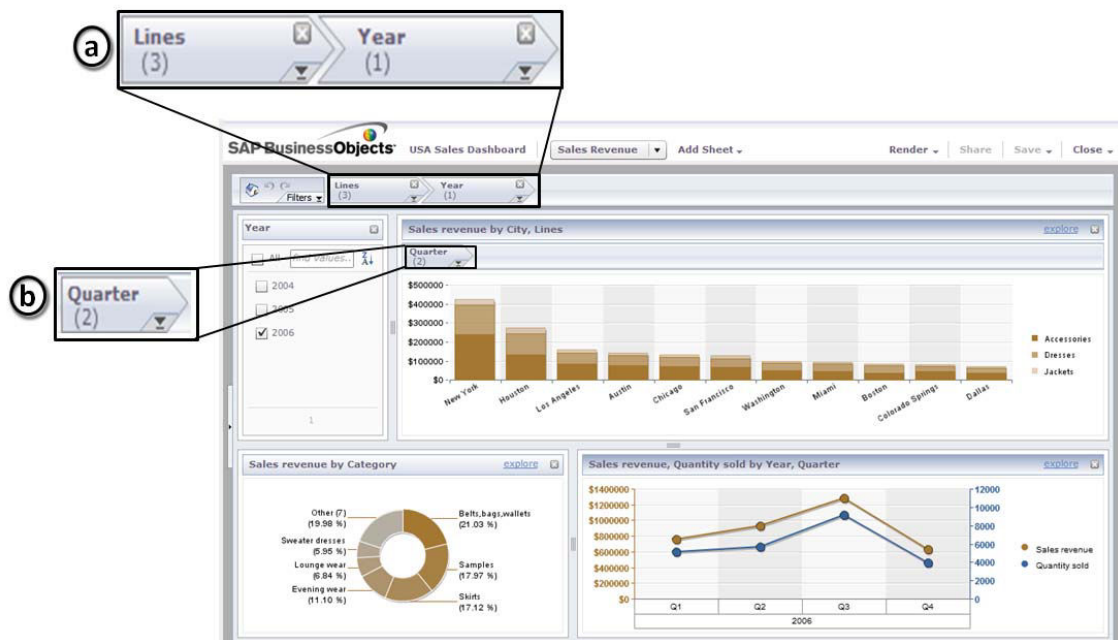


FIGURE 3.10: (a) All active global filters are visible at the top of the dashboard, (b) and local filters on the top of charts.

per year or month, etc). Each dimension has a set of values, which we call categories. Users can drag-and-drop Data-Filter components from the templates (Figures 3.8, 3.9), that can be used to restrict data categories.

To enable users to dynamically explore filtered data and compare categories, EV provides faceted navigation components [109][39]. Data-Filter-Controller components are visual components, which get attached to a data dimension, similarly to filters. But instead of constraining the dimension categories, they provide a visual list of the categories (Figure 3.10.b) that users can toggle on/off. Users can easily create and dismiss controllers to ensure that only the desired ones are visible at any given time, reducing clutter in the faceted navigation space (a challenge of faceted navigation [110]).

During data analysis and exploration, users may want to clone and explore in parallel interesting data. EV allows visualization cloning by using the Clipboard (full cloning) or by dragging and cloning parts of the data. For example in the bar chart in Figure 3.7, the user wants to explore in detail sales for the 3 quarters of 2005. By selecting and dragging only these bars, she creates a cut-out clone visualization of just the selected data, which keeps the initial component properties (dataset, visual template, categories, mappings). She can thus explore data subsets without creating components from scratch. This cloning process is essentially a fast local filter creation.

If the underlying data-sources follow the classic OLAP data organization, users have access to the traditional OLAP functionality [46], such as sorting, drill up/down, pivot, roll-up, etc. Users can also flip the chart axis to experiment with different chart layouts. This exploration functionality is necessary for expert users. It is present in the "explore" chart mode, but also in the main dashboard in the form of a widget drawer for each chart to avoid opening the "explore" view (similar to Tableau [36]). At any point users can undo/redo their actions on each chart.

Finally EV provides functions to support long term and collaborative visual analysis, such as save and share. To help users follow visual changes during data exploration, when using OLAP functions or filters, changes are smoothly animated to ensure visual momentum [111].

3.4 Exploration Views Technical Platform

EV is not just a lab prototype, but a fully operational BI dashboard that can handle large amounts of data from different sources. Although this chapter focuses on the interaction aspect of EV, we briefly describe its architecture. EV has 3 layers. The first is a web client based on the Adobe Flex that contains the dashboard layout, the exploration and faceted navigation components, and the chart manager.

The chart manager CVOM (Common Visualization Object Model) provides a set of services for chart creation and recommendation based on the analysis users want to perform and on data measures. CVOM provides three categories of visualizations: (1) traditional business graphics (line, bar, pie, 3D charts, surface, scatter/bubble, dual, radar, etc), (2) geographic visualizations, (3) tree/network relationship graphics such as star tree. Those visualizations support hierarchical data dimension labelling, and labels are laid-out according to dimension hierarchy and available space. CVOM also supports multiple visualization templates, animated transition and visual effects.

The second layer is a server containing a user session manager (as in Polstar [53]), a dashboard state management, and a component that synchronizes different data sources to enable the multi-chart visualizations. The third layer consists of an in-memory database for quick data storage and retrieval [112], and a semantic layer that helps us to connect visual queries from multiple charts and have synchronized views.

3.5 Study Design

In this section, we define the purpose of our study, its design and methodology. Our goal is to examine how novice visualization users create dashboards. We attempt to answer three questions: (i) is it easy for novices to discover visualization and dashboard functionality in an interface specifically designed using previous guidelines for novices; (ii) are the existing guidelines effective and what is missing from the literature, especially in the context of dashboard visualizations; and (iii) do novice and expert visualization users react similarly to such an interface.

Q1. Can novice users discover functionality and build dashboards? We created EV following guidelines by [38][39], and asked users to construct visualizations to answer questions, without the help of a human mediator (contrary to [38]). We wanted to see if novices can discover functionality, learn the interface and answer questions based on the created visualizations. Thus we provided no interface training and observed difficulties in interface discovery and visualization creation process.

Q2. Are the existing guidelines enough? Our users went through the entire process of dashboard design and customization, including data and visual template selection and attribute mapping, to answer multiple questions. We observed this process to verify existing guidelines and collect potentially new ones specifically for designing dashboard systems for novice users.

Q3. What are the differences between novice and expert users? Finally, we wanted to identify differences between the interaction needs and expectations of novices and BI visualization experts. Experts are knowledgeable in understanding which visualization templates and attributes mappings can answer specific questions, but not necessarily in how to interactively construct or customize a dashboard or visualization. In this respect all our users are novices in creation and customization of dashboards, but not in visualization choice and interpretation.

3.5.1 Participants and Pilot Studies

Overall 15 participants (4 women), ages between 24 and 43 years old (median of 29) took part in our study: 8 were visualization novice and dashboard design users, 7 were BI dashboard experts. Visualization and BI dashboard novices had no previous knowledge of visualization creation and were only familiar with basic charts (e.g. pie charts) seen in everyday life in newspapers or presentations. They reported being exposed to charts on a monthly basis. When asked, only 2 novice users reported that they knew what a dashboard was, and one of them had used the Google analytics dashboard.

All novice users were familiar with the concept of data filtering, but at least half did not understand concepts such as drill down-up. Novice users were recruited from our university community, with a wide professional expertise (IT, medicine, and pharmacology). Experts were product users of a BI dashboard provider (SAP) with different areas of expertise (finance, human resources, analysts, etc). Experts reported using charts and dashboards on a daily basis and were familiar with multiple types of charts, data analysis types (such as correlation, trends, comparisons, etc), and data exploration concepts (filtering, drill-down/up, pivot, roll-up, sorting, etc).

We conducted a pilot study with two novice participants and found that both had trouble understanding the difference between a measure and a category, as well as some data exploration concepts such as data filtering and drill-down/up. We thus decided to include a graphical explanation of these aspects (Figure 3.11) as training in the actual study. We also observed some interface problems: some icons were not clear enough, and were changed to more easily understandable icons and text descriptions.

3.5.2 Procedure and Apparatus

Participants performed the study in a usability lab using a 15.6" HD screen laptop. Sessions were audio/video recorded and screen captured. One observer was always present to give instructions, explain tasks and observe participants. Each session lasted between 1.5 to 2 hours. It started with a survey on the participant's background experience in computer usage, chart and dashboards, and specific data exploration functionalities such as filtering, drill down/up, etc. As one of our goals is to see how easy it is for novice users to discover functionality and create dashboards with EV, we provided graphical explanations and examples of concepts such as measures and categories, data filtering and drill-down/up, but no information about how to perform them in EV. Next, participants were introduced to a hypothetical situation and dataset (see Section 3.5.3).

We followed a think-aloud protocol, requesting users to vocalize their thoughts and actions while conducting their tasks. This phase lasted between 35-55 minutes. After finishing all the tasks in our scenario, users were asked to complete a Likert type questionnaire to elicit their opinion on specific aspects of EV, and were interviewed to clarify observations made by the experimenter, identify problems users faced, and suggest areas for improvement.

3.5.3 Tasks and Dataset

Our dataset is an e-fashion dataset (also used by Riche et al. [113]), containing 3860 records with 9 measures (e.g. sales revenue, quantity sold) and hierarchical categories



FIGURE 3.11: Graphical explanations used for training the novice participants in the user study: (a) Measures and dimensions on a chart axis (b) Filtering a chart (c) Drill down/up example.

(e.g. state/city /store, year/quarter/month, line/category). The dataset was new to both novice and experts so the impact of the dataset and domain on task performance is limited. The dataset is also complex enough to allow complicated analysis questions.

Our tasks were presented within a larger scenario. Participants were told they work in the financial department of a global chain of garment sales, monitoring financial aspects using a dashboard tool. We encouraged participants to think aloud, and if they failed to do so we prompted them to vocalize their insights and expectations. Apart from the graphical explanations (Figure 3.11), we also gave them a brief explanation of the dataset (what measures and categories exist), and told them to use the dashboard to create and rearrange visualizations.

Five tasks were performed, ordered from simple to complex (Table 3.1). Each task encouraged users to use specific operations and data exploration functionalities such as filtering.

#	Task script	Purpose of Task
T1	You are working on sales data for your company with a new dashboard tool (EV). Find and load the dashboard "Sales USA 2004-2006" and explore the new tool.	Create a new EV dashboard and explore its contents.
T2	Your manager calls to request an overview of Sales Revenue and Quantity Sold in the Year: 2006 in the States: California and Texas.	Select visual templates, measures and categories, and conduct data filtering.
T3	He later requests more information on the States of California and Texas in Year 2006: 1. Asks for a chart of the top 3 stores generating most Sales Revenue. 2. He requests to store a copy of this chart. 3. He asks for a comparison between Sales Revenue and Quantities Sold for Items accessories and dresses. 4. And finally he asks for an alternative visualization of the chart stored in T3.3 to choose from.	Customize charts, using options such as visual template changing, sorting, chart saving and reuse. Conduct different types of common data analysis (e.g. comparisons, extreme values).
T4	Your manager contacts you again to request the Year that your company sold most (Quantities sold) items from the product Line: "Leather".	Conduct different types of common data analysis (e.g. comparisons, extreme values.)
T5	Your manager needs a view showing the U.S. Sales for Years 2004-2006 based on a dashboard image he likes (given as an image). He asks you to create a new dashboard looking like this one (Figure 3.12) with the data from your dataset.	Design of a complete EV dashboard that incorporates global/local data filters, and requires layout management.

TABLE 3.1: Tasks users performed during each session of the study, and their purpose.

Tasks began by loading an existing dashboard and participants were given time to familiarize themselves with the dataset and dashboard. At the end of the session participants took part in an interview and filled out a questionnaire. The goal of the interview was to clarify observations made during the study, prompt users to identify difficulties/barriers they faced, and suggest improvements. The goal of questionnaire was to rate EV on a 7 point Likert scale on aspects such as its utility, ease of use, overall satisfaction of use, etc.

3.5.4 Data Analysis

We used four different data capturing and analysis methodologies: functionality check lists, observations, questionnaires and open-ended interviews. Observations made during

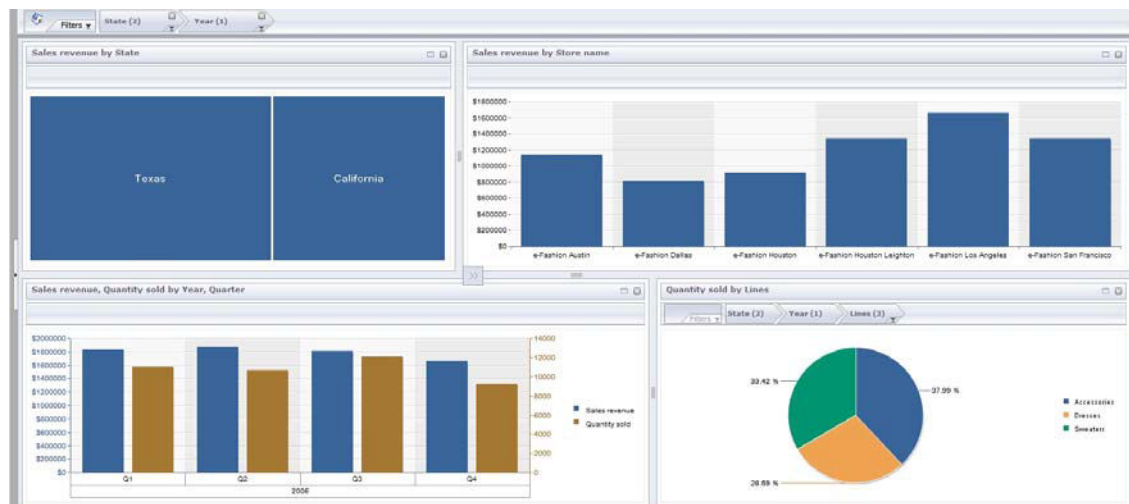


FIGURE 3.12: E-fashion dashboard that participants were asked to replicate.

the study were verified, and expanded upon, in a second pass through the recorded material (video, audio and screen captures).

To answer *Q1* on *discoverability* and *usability*, we kept a check list of functions in the interface and observed how easy it was for participants to discover and use them (results in Section 3.6.1). Questionnaire answers also contributed towards answering *Q1*.

To answer *Q2* on *guidelines for building dashboard systems* targeting novice users, we noted success events and barriers faced during dashboard creation and customization. These observations are augmented by user comments in the opened interview session. Findings are summarized in Section 3.6.2, and the resulting new guidelines are presented in Section 3.7.

To answer *Q3* on *reactions from novices and experts*, we report explicitly only the differences between the two.

3.6 Findings

We present our main findings on ease of use and discoverability of the EV interface (observed and user self-reporting, Section 3.6.1), as well as broader findings (based on observations), applicable to dashboard design for novice users in general (Section 3.6.2). We then provide design guidelines for creating dashboards for novices (Section 3.7).

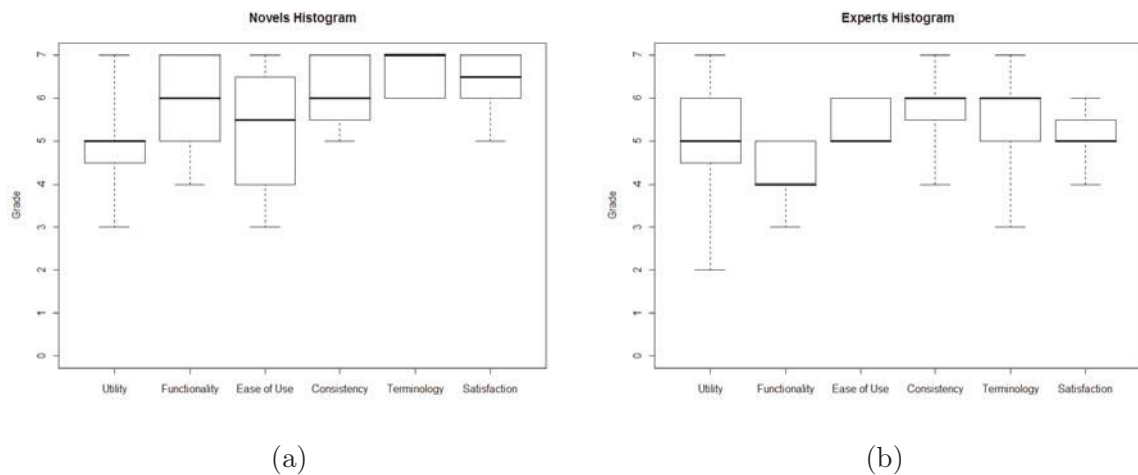


FIGURE 3.13: Results of 7-point Likert scale questionnaire: (a)Novel participants answers scale (b)Expert participant answers scale.

3.6.1 Discoverability and Ease of Use of Interface

All 8 novice participants and the 7 experts were able to successfully perform all tasks. To examine if the functionality of EV was easy to use and discover, an experimenter observed the study session and noted difficulties encountered by novice users (8 users). Almost all users easily discovered and performed actions to create new charts (7/8), take snapshots (7/8), manipulate measures, categories and visual templates (7/8), and exploration actions like sorting and drill-down/up (7/8). All novice users applied local filters easily (8/8), although 2 had initially added global filters, as they expected both global and local filters to be added in the same way.

Functionality that novice users had trouble discovering, was chart cloning (4/8), locking (8/8, see section on Locking below), and activating chart explore mode (2/8 see section on Terminology). Users who did not discover this functionality on their own, were able to complete actions after a small hint from the experimenter.

We examined if users reacted to the interface and functionality differently, based on their expertise by analyzing a 7-point Likert scale questionnaire. Analysis was conducted using a Wilcoxon rank sum test with continuity correction as our samples (Experts, Novices) had different sizes. Overall (Table 3.2), novices ranked EV higher, which can be attributed to the novelty of using a dashboard. Responses were statistically different for: functionality, consistency, and overall satisfaction (all $p < .05$). Experts would have been more satisfied with extra functionality (mentioned next), while both groups had some trouble with representation consistency of global vs. local filters (Section 3.6.2).

In the follow-up interview both groups of users reported EV was easy to use and customize. They appreciated that customization and data exploration actions became active immediately for quick feedback and verification of their choices. They stated that EV is intuitive and doesn't require training. Finally, mostly our expert users requested additional functionality, like extra customization (e.g. changing colors) and annotation features, and the ability to export dashboards in a printable version.

TABLE 3.2: Average ratings of EV for different attributes, by Novice and Expert users (* sig: indicates significant difference between the 2 groups)

User	Utility	Functionality	Ease of Use	Consistency	Satisfaction
Novice	5 (std. 1.12)	6* (std. 1.12)	5.5 (std. 1.58)	5.7* (std. 0.83)	6.5* (std. 0.74)
Expert	5 (std. 1.63)	4* (std. 0.75)	5 (std. 0.53)	6.1* (std. 0.95)	5* (std. 0.69)

3.6.2 General Observations on Dashboards for Novices

3.6.2.1 Chart Creation and Customization

Chart recommendations. As suggested in previous work on visualization novices, we provided a chart recommendation feature. All participants commented on the value of EV in helping them create appropriate charts fast. Although all users chose charts from the recommendation, only novices experimented with different alternatives, in order to "see what this new visual looks like", "look at a chart I've never seen before". We've had at least 2 users discovering and learning in this way a new visualization that they used often until the end of the study. Nevertheless, in most cases, novice and expert users ended up using familiar visualizations (bar and pie charts) for answering questions ("this new one looks complex, let's stick with what we have").

Base views and experimentation. We often found that users (novice and experts) created basic charts with measures they thought important (e.g. one for sales over time, one for number of items sold per state, etc). These charts which we refer to as base views, were usually 2-3 in number. They were used as a starting point for customizing and refining further in order to answer different questions.

When customizing a chart, we observed that novices experimented frequently with filters, drill-up/down operations and measures, to "play with things and understand the different options". On the other hand experts attempted to create the appropriate visualization without much exploration.

Chart per question. We observed that both user groups created a single visualization per task question. When asked if they see a benefit in using a dashboard, since one chart

was the focus of the analysis at a time, they all mentioned that it "helps keep the context of all the data" (benefit identified in BI dashboard literature [100]).

Use of categories. Novice participants often did not recognize the effect of adding hierarchical categories on a chart. In EV when a refined category is added in a chart, it affects aspects such as sorting, aggregation (averages), etc. For example a user had created a chart with sales sorted by year. She then added the category quarter, which is a drill-down category for year. At this point sorting was applied by quarter, but the user still expected sorting to be done by year. Thus automatically calculating metrics based on the finer categories of a hierarchy was not understood by novice users.

Data filtering. As we mentioned EV allows both global dashboard data filtering, as well as local filtering on a single chart. All users found this functionality useful, nevertheless it was often hard for novice users to make correct use of it.

There are 2 creation options for global filters: select part of the data on a chart and then apply the option "filter" (Figure 3.14) which will filter data based on the selected categories and values; or select the option "filter" and then go over the process of selecting on which categories and values to filter. We found that novices generally preferred the second option because they want to make sure they "selected all correct values". Contrary, experts created global filters with selections that they then refined further.

Another aspect where expert and novice users differed was the presentation of possible filter categories. Experts wanted categories to be presented grouped thematically, e.g. grouping of time (year/quarter), or geographical categories (state/city), while novices preferred an alphabetic presentation.

Users found the existence of both global and local filters useful and relevant when analyzing data as it allows advanced analysis, for example comparison of different products sold by stores for several periods. All participants commented that presenting global and local filters consistently is important, but that local filters should be further highlighted in charts. The EV interface gave such detailed filter views (Figure 3.8), and users commented on the bulkiness of the filter panels. There is thus a tradeoff between presenting filter details and space, novice participants want filter details to always be visible (both local and global), while experts found they take up too much space.

Finally, we found that sometimes users chose local filters that were in conflict with global ones. This caused problems for several novice users who could not understand why no data was displayed.



(a)

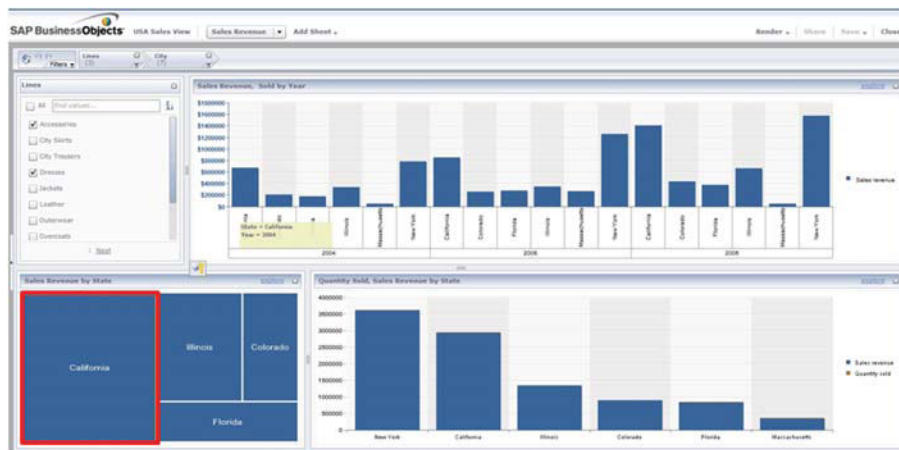


(b)

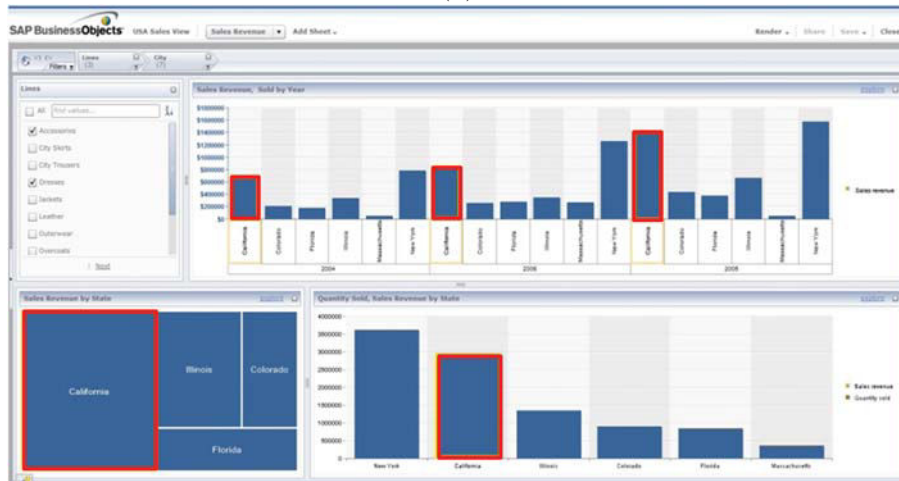
FIGURE 3.14: An example of EV Filtering functionality. In this example visualization of Sales Revenue (a) the user selects data for Q1 and Q2 in year 2005, then clicks on the "filter" button. (b) The filtering result presented in EV.

3.6.2.2 Interaction

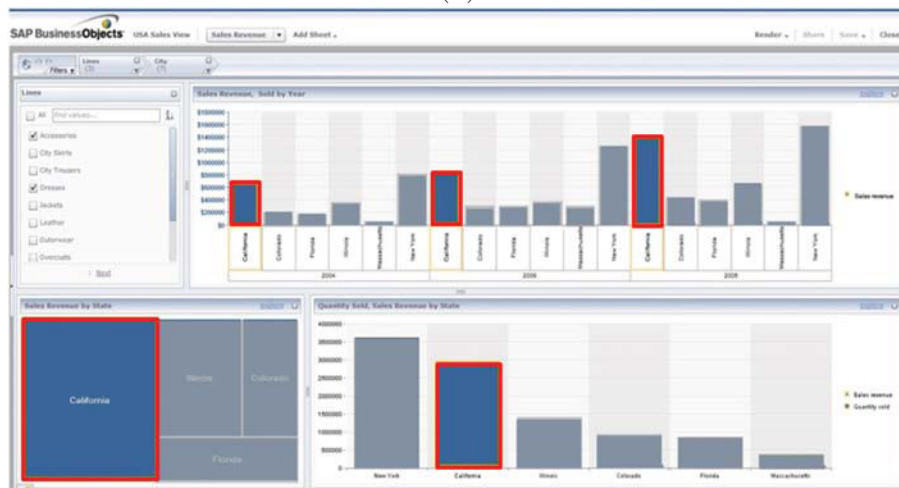
Linking & coordinated views. Most participants (experts and novices) expect a link between charts. Thus when selecting a portion of one chart to highlight Figure 3.15.a, they expect other charts with the selected data in EV to update accordingly. Nevertheless, the expectation of what "update" means differs across users. Most expert users expected selections on one chart to be reflected as selections in all the others Figure 3.15.b, in the classic brushing and linking approach in coordinated views (default behavior in EV, Figure 3.7). Nevertheless, most novice users expected selections to act as explicit filtering: when partial data was selected in a chart, they expected all charts to update such that only the selected data is reflected throughout the dashboard Figure 3.15.c. This interpretation would result in changes to measures such as averages, sorting, etc for different charts.



(a)



(b)



(c)

FIGURE 3.15: Understanding Linking & coordinated views by novice and experts when (a) performing a selection on a visualization: (b) Experts expected a reflection on the other dashboard charts as selection (brush & link) for the same context. (c) Novices expected other dashboard charts to update as if filtering was on, so only the selected data is reflected on the dashboard.

Undo as exploration strategy. All participants were familiar with undo/redo, but we found that it was novice participants that made extensive use of undo per chart. Undo was not used as a correction mechanism when mistakes were made, but rather as an exploration aid. Novice users reported that if they wanted to try something and see if their predictions were correct (e.g. drill-down), they could use undo to quickly dismiss their tests (instead of reversing their actions, e.g. drilling-up). They also used undo as a way of returning to base views (charts used as starting points for further exploration): from base charts they would drill-down/up or filter briefly to answer a question, and then use undo to return to the original chart.

Clipboard. This feature allows users to create a palette of active charts that they can later reuse. All participants found it useful for storing, but they reported it is most useful for sharing chart snapshots with others. This remark, reinforced by requests for ways to email or export charts to presentations, indicates that users also saw the clipboard as a storytelling creation platform for sharing their findings with others.

Participants from both groups also requested features such as annotation, putting titles, adding comments, and assigning the current date to a chart. Experts in particular requested stored charts to also clearly and visibly describe context (e.g. dataset, measures, filters, etc). These requests indicate that participants want to use the clipboard to bookmark their exploration process: keep track of interesting charts they explored, comment them, and store the date they conducted their exploration. Some novice users mentioned they were unsure exactly what is captured when selection that is active on a chart is placed on the clipboard (i.e. does the capture only include the selected data or not). This is an interpretation problem related the effect one action has to the remaining visualizations and data that are seen in coordinated views.

Visualization locking. In our prototype we provided functionality for visualization locking (not allowing editing or application of global filters). Expert participants expected this locking to apply to filters, but also visual properties like size and position in the dashboard. Some novices felt that locking a visualization "creates confusion because it takes out this particular chart from the context of the dashboard". They requested very clear visual representation to indicate locked charts, such as depressed charts.

Terminology. Overall users were satisfied with the terminology and labels used in EV. Some novice users suggested more precise terminology even for some features that are familiar from other UIs. For example our original icon for editing a chart was confusing and novice users were not sure if they were going to edit the title, the chart or the entire dashboard. Thus they often requested tool-tips detailing the function of icons and their context. Finally, while experts were accustomed to English terms (in terms of functionality,

dataset measures and categories), this was not true for novices with English as a second language who requested support in their mother tongue.

3.6.2.3 Design Tradeoffs

The design of EV balances several tradeoffs to support both expert and novice users. Functionality required by experts, such as chart locking, axis flipping, and extended chart recommendations, often confuses novices. In our work, we attempted to always provide the simpler and most used functionality in a more prominent way. For example options like sorting, filtering, and visualization changing are first in the chart widget drawer, while axis flipping or locking come last. In the chart recommendation list, most common charts always appear first. Moreover, when possible we use reasonable defaults, for example in assigning color and other visual variables to categories. Nevertheless, we often found that functionality required by expert users still confused novices occasionally (e.g. the locking option).

3.7 Discussion and Design Guidelines

All novice participants in our study were able to successfully perform the required tasks in EV, a system built using guidelines for novice user interaction. Based on our observations we have additional guidelines than the ones proposed by Grammel et al. [38] and Heer et al. [39], when designing visualization systems for novice users, especially in the context of dashboards.

Basic training on concepts such as categories: Our novice users were successful in creating visualizations and solving their tasks. But, based on our pilots, they were given small tutorials on the differences between measures, categories, and the nature of basic OLAP and data exploration functionality (drill-down/up, filtering, etc). Thus dashboard interfaces for novices should provide small tutorials of these concepts. This training is particularly important for hierarchical categories. Novice users need to know if metrics (averages, sorted lists, etc) are recalculated in charts when adding finer hierarchical categories. Based on user comments we recommend either not recalculating metrics unless users request it, or recalculating per category on demand.

Considerations in coordinated views: When creating a dashboard, novice users assume that charts are linked by default, and expect brushing and linking functionality. Its implementation for single data sources where charts share measures and categories is fairly easy. Nevertheless, if measures and categories need to be matched across data sources, then matching approaches need to be provided [51][50].

When providing coordinated views, designers should remember that novice users confuse linked selections of visual queries to applying global data filters based on the selections. To help disambiguation, they need clear visual differentiation between selection for linking and for data filtering. A linking selection should highlight selections without affecting the visual representation of unselected data; a filtering selection should also alter the visual representation of the unselected data (e.g. fade them out) to indicate that all data are being altered due to the selection and if the user requires it only show the filtered data.

Extensive exploration aids: Novice users created a small number of base view charts, most often of familiar visualization templates, that they later copied and refined to answer questions. When possible, such basic views should be provided.

Novices also made extensive use of chart undo/redo as an exploration aid, thus this functionality should be provided on a per chart basis.

Furthermore, clipboards were seen as good ways to share data for storytelling and communication, as well as bookmarks of the exploration process. For clipboard snapshots to be useful this way, they require functionality such as annotations and labeling, as well as ways to report their metadata, such as creation date and details of the context of the chart (i.e. dataset, active filters, etc).

Finally, although locking charts can be useful during long analysis (as indicated by experts), this locking should be very clearly visible to not confuse novices (e.g. fading the entire locked chart, or making it appear as depressed).

Data filtering considerations: Data filters, although very useful for data exploration, can be challenging for novice users. We found that novices preferred to create global filters from scratch, select categories, values, etc. While experts preferred to create them graphically (through visual selections) and then refine them. Both options should be available. When selecting categories to filter, we recommend presenting them in thematic groups (expert choice) that promotes dataset learning. But this organization should be supplemented by text search functionality and a possibility to order them alphabetically. Novice participants want filter details to always be visible (both local and global), while experts found they take up too much space. Designers should consider an auto-hide, or on/off visual representation of filters. Approaches such as interactive legends [113] could also reduce clutter of local filters, since most charts include labels of some sort.

For both groups, being able to distinguish between global and local chart filters is important, and it should always be clear which categories and values contribute to existing filters. Finally, since novice users have trouble predicting conflicts in global and local filters, it is important to clearly mark such conflicts.

Language & terminology: Novice users are often not aware of international terms for different charts, metrics or functions. Thus local language support is important.

Moreover, buttons, labels and icons (even universally understood ones) are not enough to provide novices with adequate information for their behavior in the context of dashboards. Detailed explanations need to be provided on demand (e.g. tool-tips).

3.8 Summary

We have presented Exploration Views, a prototype system that allows users to easily create, customize and interact with Business Intelligence visualization dashboards. EV was build following previous guidelines for novice visualization users. Through a user-study with novice and expert visualization users we verified that EV enables novices to create dashboards for answering complex questions. We found some differences in how experts and novice users interact with the interface, reported in our study. More importantly, based on our observations we provide new guidelines that augment previous work on designing for visualization novices, especially in the context of interactive visualization systems in the form of dashboards.

Chapter 4

Annotating BI Visualization Dashboards

4.1 Abstract

Annotations have been identified as an important aid in analysis record-keeping and recently data discovery. In this chapter we discuss the use of annotations on visualization dashboards, with a special focus on business intelligence (BI) analysis. In-depth interviews with experts lead to new annotation needs for multi-chart visualization systems, on which we based the design of a dashboard prototype that supports data and context aware annotations. We focus particularly on novel annotation aspects, such as multi-target annotations, annotation transparency across charts and data dimension levels, as well as annotation properties such as lifetime and validity. Moreover, our prototype is built on a data layer shared among different data-sources and BI applications, allowing cross application annotations. We discuss challenges in supporting context aware annotations in dashboards and other visualizations, such as dealing with changing annotated data, and provide design solutions. Finally we report reactions and recommendations from a different set of expert users.

4.2 Introduction

The main problem that plagues dashboards today is that they don't say enough and what they do say they don't say very well. In other words, they are not very informative.

— Stephen Few[29]

Well-designed interactive visualizations help users gain insights into their data, identify patterns, and make decisions. During visual analysis, a single analyst or a group, access information from multiple sources to combine insights. These analysis sessions may last from a few hours to many weeks. Given the limits of human memory [40] remembering all reasoning details and milestones in the analysis process becomes challenging. Thus the importance of "externalizing" memory, knowledge, and insights [114] has been identified early on. Analysts achieve this by making sense of their visualizations, annotating them with their findings [115], organizing their findings, and presenting them to others [52][65]. Annotations are thus an integral part of complex visual analyses, aiding analysts in re-visiting and reusing past analysis sessions [70]. Recent work has started to explore the potential of utilizing the context surrounding annotations to better support users in data discovery and analysis (e.g. [80][70]).

Despite the growing use and research work on annotations for visual data analysis in general Section 2.4, related research on the domain of BI dashboards and other complex multi-chart visualization environments (coordinated views [25]), has not equally progressed. Our work attempts to rectify that.

Based on interviews with expert BI analysts, we derive new requirements for dashboard annotations that current dashboard tools fail to support. Following these requirements and using a user-centered design approach, we develop and evaluate an annotation-enabled dashboard, addressing several challenges.

The contributions presented in this chapter are: (1) A new set of requirements for annotations, targeting dashboards and other coordinated view visualization systems. (2) The design and evaluation of a prototype dashboard that supports context-aware annotations. Using this system, we explain how annotations support advanced analysis functionality such as: "annotate once, see everywhere" annotations for multi-chart visualizations (as was done for text [81]), annotation recommendations to promote knowledge discovery, and the use of the common data model layer present in BI systems (e.g. [116] [117]) to link annotations with external data sources and enable cross application annotations and information foraging. (3) We identify and provide solutions to challenges when using context aware

annotations, such as issues arising when the "context" changes (e.g. annotated data are deleted or changed at the data source).



FIGURE 4.1: Our dashboard prototype with context aware annotations. (a) The main dashboard contains icons with the number of annotations for each data-point (b). Below the list of all dashboard annotations in (c), is a set of recommended annotations (d) emphasizing their similarity.

4.3 Approach

In our work we followed a user-centered design approach. Based on in-depth interviews with expert BI analysts, we derive requirements and needs for annotations on visualization dashboards: this led to the choice of context aware annotations. Following these requirements and scenarios of use provided from our experts, we developed a dashboard prototype with annotation functionality that supports different analysis tasks.

While building our prototype, we identified challenges in using context aware annotations with dynamic data and evolving annotation context. Based on a second interview with experts, we refined our system to address these challenges, following user recommendations.

Finally, we evaluated our prototype with different experts and report results and reactions on the benefits and drawbacks of our designs.

4.4 Interview 1: Annotation Requirements & Needs

To better understand BI expert annotation needs, we conducted a set of in depth interviews with 8 BI experts. Our goal was to investigate their annotation activities when performing analysis using dashboards, as well as their hopes for future annotation enabled dashboards. Experts' experience in the BI domain ranged from 3-11 years. Three used dashboards daily while 5 several times a week.

Interviews lasted 50-90 minutes, and they were recorded and transcribed, and the emerging themes identified using a card sorting analysis approach. Bellow we summarize our major findings:

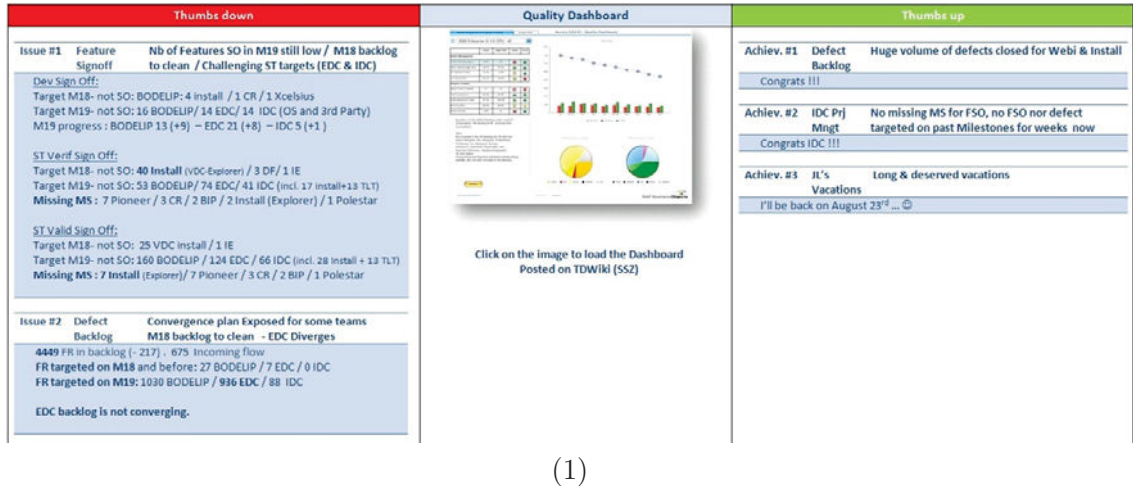
4.4.1 BI Dashboards Annotation Practices

All experts confirmed the importance of annotations in their analysis and the need of having annotations directly embedded in their BI dashboard tools. This way analysts can share their analysis with tactical and operational decision-makers simply by using their dashboards and relevant annotation indicators [118]. An expert explained: "The data has no meaning without our analysis and contribution we add to it, so charts and dashboards would be useless without explanations".

When asked about their actual analysis annotation practices, all experts mentioned taking static snapshots of dashboards (Figure 4.2) and including them in emails or power point presentations, with text comments accompanying them. Their annotated comments are either for dissemination purposes to a larger audience, or serve as personal reminders and analysis findings records. Experts explained that for each annotation they also copy a data query, as a reference to the exact data the annotation refers to.

4.4.2 BI Dashboards Annotation Needs

We further prompted experts to elaborate on their ideal vision of annotation-enabled dashboards. Experts brought to the interviews annotated examples made by them or other BI experts in their groups. These examples indeed consisted of either screenshots of dashboards on paper with hand-written comments, snapshots attached to emails with annotated explanations, or presentation slides with annotated text (Figure 4.2). By examining this material with them, we extracted annotation functionality they would like to have embedded in their dashboards systems. Seven major needs emerged from the interview and expert's examples:

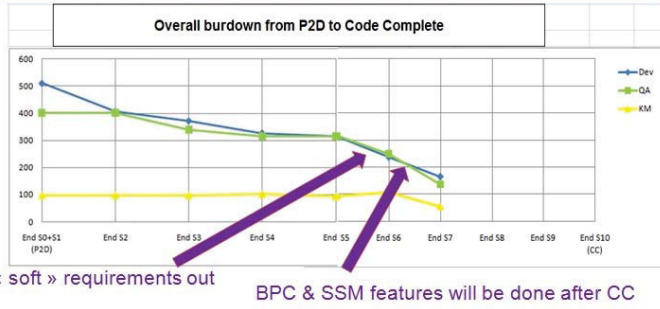


(1)

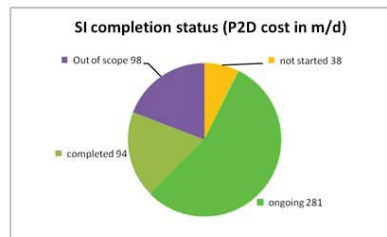
FIM Overall progress

•FIM team

– Overall burn-down:
= how many man/days remain

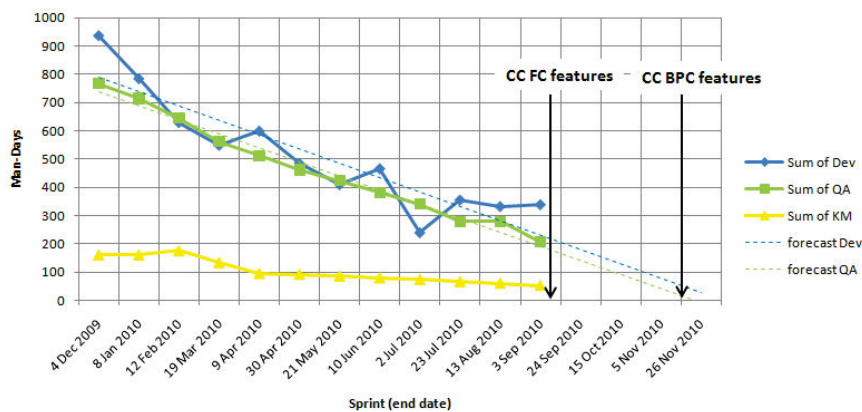


– SI progress status:
= part of the P2D scope being developed, in man/days



(2)

EPM Add-in burndown chart from P2D to D2P



(3)

FIGURE 4.2: Three examples of the actual practice of annotating visualizations provided by our participants. (1) A dashboard represented in the middle, and associated annotations on right and left. The left annotations represent the bad remarks, and the right side contains the good remarks. (2) and (3) are two slides from a report training powerpoint presentation, that contains annotations made on the presentation tool.

4.4.2.1 Multiple Target Annotations

Experts need to attach their annotations to multiple targets (multiple charts on the dashboard). Seven out of eight experts annotated multiple charts on a single dashboard. An expert commented, "The aim of a dashboard is to tell a story, that's why we need a way to link all the dashboard's charts that hold a connection and give the story we need". Another expert said "two different dimensions that cannot be gathered in one chart could answer a specific question, so we need to link them by one annotation to explain and show the result or solution".

4.4.2.2 Chart Transparent Annotations

Six experts mentioned they need to attach annotations to the actual data points, rather than static visualizations "because usually I change the chart view and then have to rewrite or re-comment for the new representation of data". Another explained that this is useful as a verification or exploration mechanism: "it is good not only for having the comment for both chart representations, but when seeing the comment on a different chart, it may show that the interpretation of this point was not right, or maybe a different chart would give a better interpretable view of your data". One expert mentioned that it "would be useful to find other visualizations or annotations for a specific event in that data", i.e. use the annotated data as a focus point from which to find related annotations and visualizations. Here experts highlighted the importance of being able to go from an annotation, or a list of annotations, to the referenced real data on the dashboard and back.

4.4.2.3 Visible Annotation Representation on the Dashboard

Five experts preferred to see visual indicators (icons, flags or special visual highlights) of annotations attached to the actual data "like small indicators on the dashboard and clicking them to see the content". They also wanted to have the option to expand and collapse all annotations' text: expand them "to see all dashboard related annotations at a glance and their story", and collapse them "because having them all on the dashboard would clutter it". Moreover they commented that a list of all dashboard annotations should be available, with visual links to their data targets.

4.4.2.4 Granularity Transparent Annotations

Four experts mentioned they would like to have annotations "transparent" through different data dimension granularities (e.g. *country/city*, *year/quarter/month*), and mentioned

there are business use cases where they want to mark a specific event of importance in all visualized levels (e.g. mark a market trend). They would like this feature as an option in special business cases, and deactivate it when not relevant.

4.4.2.5 Annotation Validity & Lifetime

The lifetime of dashboard annotations was an important issue raised, since there are cases where the data referred to by an annotation can change. All experts confirmed the need to archive annotations "we would like to have them archived to be used in future analysis and to access them for comparisons". They would also like to archive with the annotation a snapshot of the visualization on which it was made. Seven experts preferred to have the option to explicitly define the lifetime of annotations either in time periods (sessions, days, weeks, months, etc), or based on validity rules (e.g. valid while the value of *sales* ranges between 3000 and 5000).

4.4.2.6 Sharing BI Dashboard Annotations

Experts expressed the need to share annotations, and in particular being able to define user groups to share annotations with. They also requested to be alerted when a specific user or group of users edits or shares an annotation. Nevertheless, all experts expressed the need for private annotations to use as reminders or analysis record keeping.

4.4.2.7 Formal Annotation Representation on Dashboards

When prompted, all experts stated a preference for a more formal annotation design (e.g. use a toolbar with predefined annotation categories, shapes, arrows and text boxes), rather than hand-drawing shapes or notes [119]. They explained that they very often share annotated dashboards with peers and embed them in presentations, thus they see a "first priority for structured, clean annotations" that look more professional.

4.4.3 Summary

To summarize, we identified the following annotation needs expressed by BI experts, grouped in two categories:

a. Functional

F1. *Multiple Target Annotations*

Annotations should have **multiple data and chart targets**. Navigation to the different

targets should be possible from the annotation. With few exceptions (e.g. [70]), most visualization systems annotate single visualization charts.

F2. *Chart Transparent Annotations*

Annotations need to be **attached to data-points**, rather than charts. Few systems (e.g. [36]) actually allow data, rather than visualization annotations, and generally they do not show a data point's annotation in different charts.

F3. *Granularity Transparent Annotations*

Annotation transparency should be optionally preserved **across dimension granularities**. We are aware of no system that supports this.

F4. *Annotation Validity & Lifetime*

Annotations should be **archived** even if the context of the annotation (e.g. annotated data) changes, together with a **visual snapshot of the data** at the time of the annotation. Moreover, users should also be able to define the **annotation lifetime**, based on a time period or data related rules. We are aware of no system that supports both these aspects.

F5. *Sharing Annotations*

Users should have the option to **share** annotations with specific users groups, or kept them private for personal use. Systems supporting sharing (e.g. [37]), do so for single charts.

b. Design

D1. *Visible Annotation Representation on the Dashboard*

Annotations should be **attached to the annotated data** using small visual indicators that won't clutter the overall view, but with an option to expand them in place.

D2. *Formal Annotation Representation on Dashboards*

BI experts prefer **structured** and professional looking annotations to hand drawn ones, for presentation purposes.

4.5 Context Aware Annotations Prototype

Based on these findings we implemented a dashboard prototype which supports annotations that keep a record of their surrounding context (i.e. the multiple data points and charts they are attached to), as well as special annotation properties, such as their validity and lifetime. Using annotation context also allows us to provide annotation recommendations [70].

In this section, we provide an overview of our prototype. Apart from the described interview, experts were engaged throughout the design of our system.

4.5.1 Architecture

BI dashboards provide users with a multi-chart interface, connecting one or more data sources and data sets that are represented as charts. Based on our user interviews, we want seamless annotations that work across data sources and charts. To that end we allow annotations to refer to one or many "data targets" on a common data layer.

A common data layer is present in most BI systems (e.g. Oracle, SAP, Microsoft) and provides a layer of data abstraction that unifies different data sources. Thus data queries are performed directly on this layer and retrieve relevant results from multiple sources. Our annotations work on top of this layer: each annotation is stored in a database, together with a reference to its data targets on the common data layer. Although the common data layer is not present in all coordinated views system, there is usually a common querying layer (e.g. a database with SQL queries) that works across charts and controls chart coordination. In such systems our annotations can be connected to the querying layer.

Given this approach, our system permits context aware annotations that can be shared among different applications with a common data layer. Once an annotation is made on data points in one application, it is attached to this data irrespective of application, allowing cross-application annotation and annotation re-use.

The context of an annotation includes the following entities Figure 4.3, similar to the model proposed by [120]:

- Content: the text provided by the user.
- Authoring metadata: automatically captured information, such as the author and creation date of the annotation.
- User defined properties: these include annotation ranking and categorization of annotations (derived from previous research [121][122][72]). As well as new concepts such as annotation "validity rule" or "lifetime" (F4 discussed in great detail in the Annotation Challenges section). Defining these properties is optional.
- Data targets: these are one or more entities that the annotation refers to (or is attached to). Each target includes: an Overview data context that defines the larger data context of the annotation, for example the chart or data table on which is was made; and more importantly, a Detailed data context, the specific data points that the user is annotating. Both Overview and Detailed data context are stored as data

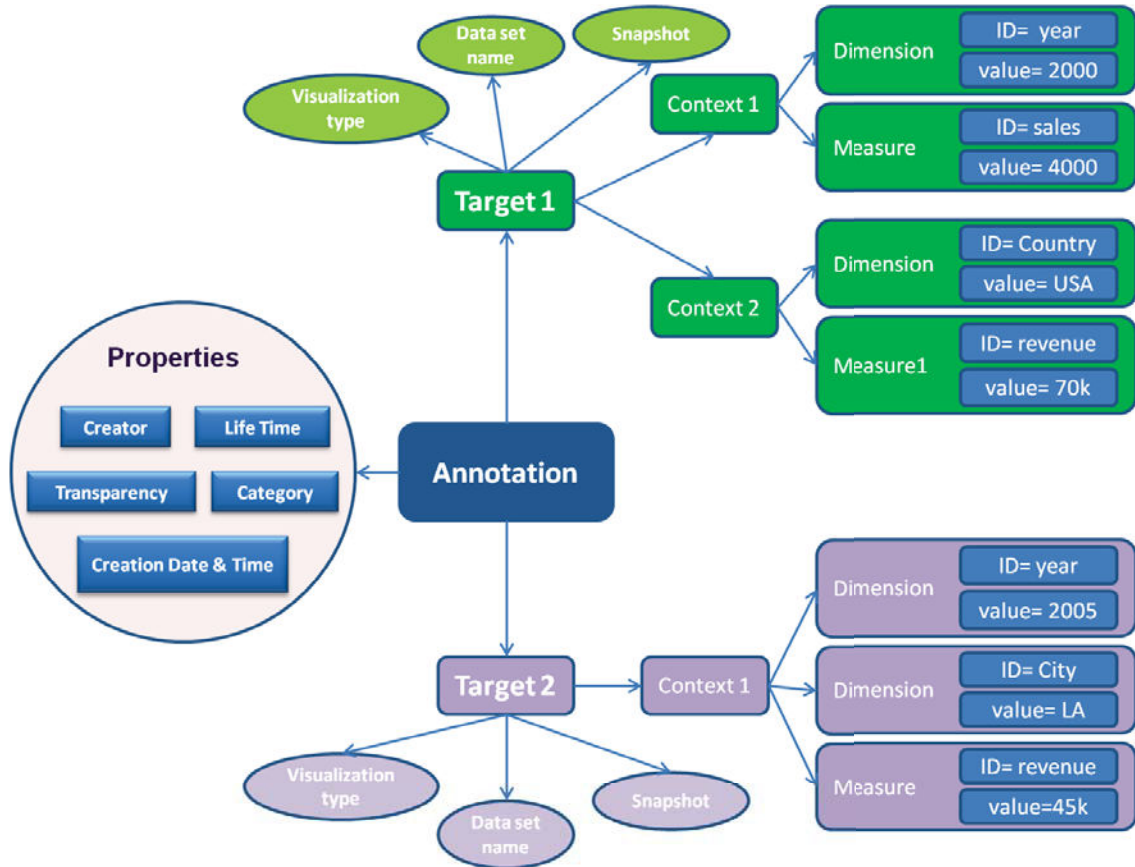


FIGURE 4.3: This figure shows the context-aware annotation structure followed in our prototype. An annotation can be attached to multiple targets, and each target is identified by its data context (dimensions and measures, names and values). The annotation saves with each target the dataset name, visualization type and a snapshot at the time of annotating. Moreover, an annotation has different properties such as: author, creation date, annotation ranking, categorization and lifetime. Targets

dimension IDs (e.g. *year*, *country*) and their values, and measure IDs (e.g. *sales prices*, *population*) and values.

- Threads: other annotations connected to this annotation (e.g. a discussion thread in collaborative annotating).

4.5.2 Interface

Figure 4.1 shows a screenshot of our prototype's main user interface. The left portion of the dashboard screen is the main canvas where the retrieved data is visualized (Figure 4.1.a). Users explore and analyze data by interacting with each visualization chart on the dashboard. We provide functionality such as data filtering, sorting, as well as advanced OLAP functionality [46] (e.g. drill-down/up). On the main canvas (Figure 4.1.b),

annotations are shown as icons (D1) containing the number of attached annotations to the specific data.

The right hand side of the screen (Figure 4.1.c) is the list of all annotations that users can navigate. For each annotation in this list we provide a snippet of its content and some authoring meta-data to help user recall. When users click on a specific annotation in the annotations list, they are provided with an overview of related and recommended annotations (Figure 4.1.d). We discuss this recommendation feature in the Section "How to Use Context".

Annotations in the list are searchable and can be double clicked to pop-up a detailed view of the annotation content (Figure 4.4). The detailed view includes content and metadata, as well as a list of all target data contexts the annotation is attached to (overview context charts and tables with the detailed data context highlighted, Figure 4.4.a). A textual description of the context (dimensions, measures, values) is also provided.

For example the annotation seen in Figure 4.4 is attached to 2 data contexts, part of a doughnut chart from one dataset, and part of a bar chart from another. When seen in a different application sharing the same data layer, this annotation will still show this exact snapshot, including the two original charts and the highlighted points, regardless of how the data is visually presented in the different applications.

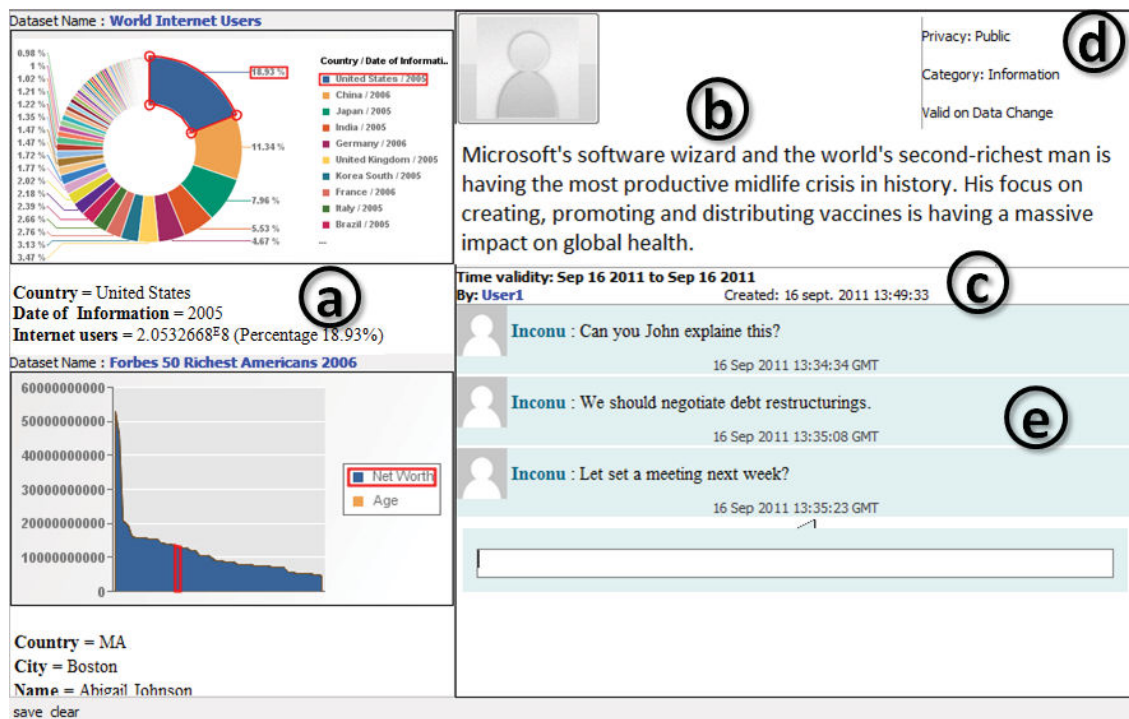


FIGURE 4.4: Detailed annotation view, including (a) all related data contexts & detailed data context in text form, (b) content, (c) metadata, (e) discussion thread, and (d) annotation properties (lifetime, privacy, etc).

Annotations are created in the following way: When a user selects one or multiple regions on different charts (representing different data contexts), she can right click to annotate them. A new detailed annotation window appears showing all charts in the dashboard containing the target data context (even if not selected when creating the annotation), with the specific context highlighted.

User can share their annotations (**F5**) with others through a private, public option, and can respond to other user's annotations in the form of an annotation thread (Figure 4.4.e), supporting social sense-making [75].

4.6 How to Use Context

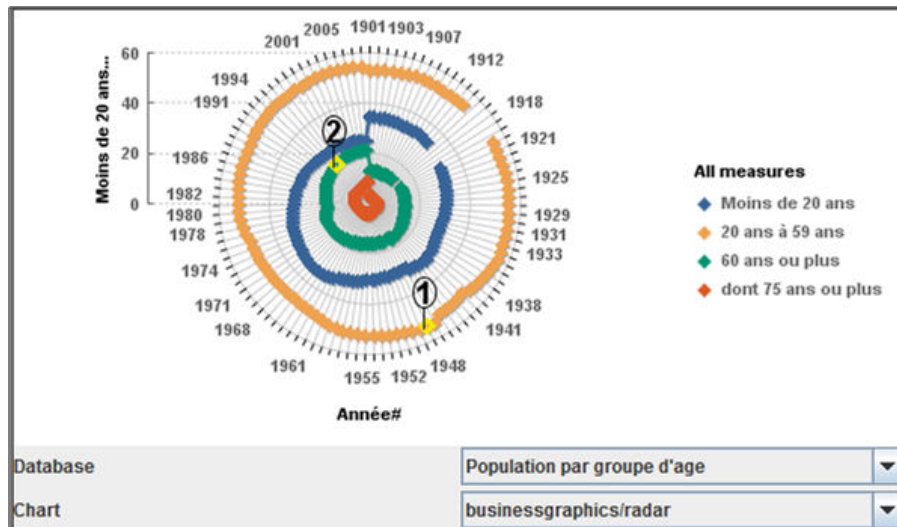
In this section we discuss in detail how our annotation prototype follows the experts' requirements and aids analysts in their tasks on dashboards. The tasks and scenarios described here come either from the interview or from intermediate feedback sessions with experts.

4.6.1 Transparency

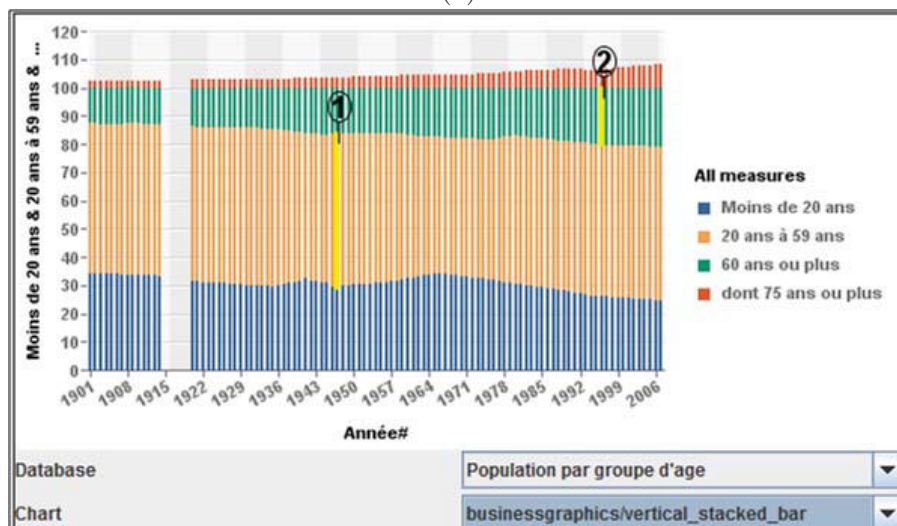
Our annotations are transparent through applications, dashboards, charts, and datasets following a common data layer. Moreover, they can be transparent through different hierarchical data granularities, discussed later on.

Our annotations store detailed data context, so whenever this data context appears, so does the annotation (annotate once, see everywhere as it is done for text in [81]). During the analysis phase this can help in 4 different ways:

- (i) Analysts can switch visualizations (chart types) without losing their annotations (**F2**), experimenting with alternative representations that may be better suited to their data.
- (ii) Our annotations provide a list of all other charts containing the same detailed annotated context. This allows analysts to find new charts that may present these data in different analysis scenarios, e.g. in combination with different measures or data categories. Thus analysts may gain new insights from their data by seeing them in different analysis situations.
- (iii) Through annotations, analysts can reach different material (other applications or data-sources) that incorporate the same data context, aiding in knowledge discovery.
- (iv) Finally, users have the option to see annotations from lower or higher data dimension granularity (**F3**). For example when viewing sales data for France they can make



(a)



(b)

FIGURE 4.5: Chart transparent annotations are attached to the actual data points, rather than static visualizations, that permits to show a data point's annotation in different charts. (a) A radar chart shows the attached annotations to two data points. (b) The visualization is switched to a vertical stacked bar chart, and still shows the attached annotations to the two data points.

an annotation transparent for the hierarchical dimension 'location' and see it appearing in sales in Paris. Noticing such annotations that are explicitly visible across dimension granularities can lead to new exploration paths and insights. We elaborate on this last point and discuss challenges with annotations across granularities in the next interview section.

4.6.2 Chart linking

Our system provides an easy way to add/remove data contexts to/from the annotation, either at the detailed annotation window, or by selecting interactively a data context (e.g. a bar in a bar-chart) and choose the option to link/unlink it to an existing annotation. Annotations can target multiple detailed data contexts from multiple charts, as well as multiple data sets (Figure 4.4.a). This was a specific requirement from our interviews (F1), made possible through the use of the common data layer across charts and data-sources. Analysts can thus use annotations as a means to chunk or group diverse material together for insight externalization and organization (an important task in sense-making [70], [78]).

Moreover, annotations can be seen as a way to navigate through the linked charts: by expanding the annotation of a data context in one chart, analysts can see all related contexts in other charts, select them and load them in the dashboard. This cross-linking and navigation has been identified as important in reading sense-making tasks (e.g. [123]) and visualizations representing one chart at a time (e.g. [52]).

4.6.3 Recommendations

We use annotation context similarly to [70] on text, to recommend annotations or charts of similar data contexts, promoting knowledge discovery and peer learning from annotations of others.

To identify similar contexts we compute a simple similarity measure. Similarity between two data contexts is the arithmetic average of all common dimension IDs across contexts, all common measures IDs and their values.

Recommended annotations are presented in a list by order of similarity (Figure 4.1.d). A set of two visual similarity bars was included initially in our prototype, the 1st to show the overall similarity of all dimensions and measures between two annotations, and the 2nd to show the similarity of dimensions only. The goal of this was to help analysts understand the ranking and, if they choose so, to focus the recommendations to dimensions only. Nevertheless experts indicated that the overall similarity was enough for their purposes and we adjusted our prototype (see "User Study").

4.6.4 External search

To aid knowledge foraging and discovery, our system allows users to use their annotation context as a search term outside applications with the same data model. For example we

use it to perform system wide searches for documents, as well as to construct a set of well formatted search terms to use in web search engines, like Google.

4.6.5 Automatic annotation

We do not perform automated or semi-automated annotation [80]. Although we do have all the required components, this previous work targets single chart visualizations to define annotation templates, and requires further work to extend to multiple charts. Nevertheless, in our system there is an implicit automated annotation taking place through the transparency of annotations across applications and data sources.

4.7 Interview 2: Design Challenges & Solutions

Early on in our interviews, we came across two major design challenges for annotations not addressed in previous work: (1) how should annotations behave when their underlying data has changed? (2) When are annotations applicable across granularities? We conducted a follow-up interview with the same 8 experts, where we discussed alternatives and collected their reactions and suggestions.

4.7.1 Dynamic context vs. dynamic annotations

Often annotations in BI target dynamic data points (e.g. the averages of sales records that are being updated every month). There are two design alternatives for how annotations could behave during such data changes.

The first is to make annotations dynamic: when the context of the annotation changes in a small way, then the annotation can be attached to this new context. The challenge is to clearly define what is "a small change".

Based on input from all our participants this "depends on the situation". Some mentioned that changes in data value only could be considered as a small change in some scenarios (while measure and dimension IDs remain the same). For example a participant attaches an annotation regarding sales peak on the data point with dimension ID= *month*, value = December, and measure ID = *sold-quantity*, value= 3000. They want to keep this annotation valid even if the value of the measure changes, e.g. becomes 3500. Others mentioned that sometimes if the dimensions or measures go through small changes, the annotation can remain attached to the data points. Clearly the nature of "small change" is unique to each analysis scenario and requires users to define "change" thresholds. On the

flip side, this requirement essentially turns annotations into data monitoring mechanisms: when the annotation is deleted a specific data threshold has been reached.

The second alternative is to acknowledge dynamic context, and assume that the annotation is no longer attached to the data. When the data context changes the system assumes the annotation is no longer valid, and keeps a snapshot of the context under which the annotation was made (dimensions, measures, values), as well as the visual chart it was made on. Nevertheless, all participants stressed that this should not mean the annotation gets lost, but archived.

This last approach is the default approach we chose in our system. The unattached annotation obviously has a context very similar to that of the changed data. Thus if the user revisits the changed data, the old annotation will appear as a highly ranked recommendation, easily accessible. The visual representation of such "archived" annotations should be clearly distinguishable from "active" ones: following the inactive metaphor of UI interfaces, archived annotations are grayed out (Figure 4.6). We chose this way of addressing dynamic data context, as it is simple to define, and provides a way to archive data changes (see Sec. User Study).

An orthogonal treatment comes as a direct suggestion from our experts. They stated that their analysis is often affected by events through time, or is valid for short time periods. This prompted the notion of a user-defined "validity period" of an annotation. If she so chooses, the user can explicitly define an expiration date or a validity period for annotations, after which annotations get archived or deleted from the system (Figure 4.7). At this point experts requested the option to be sent a notification inquiring if the annotation should be archived or deleted. This lifetime functionality is also implemented.

4.7.2 Granularity

In hierarchical datasets, as is the case in many BI data sources (e.g. OLAP databases), users can drill down or up in dimensions. For example for the dimension *time* we can have 3 hierarchical levels: *year*, *quarter* and *month*, where a *year* has 4 *quarters*, and a *quarter* 3 *months*. The question that arises is: if an annotation is made at a lower/higher hierarchy level, when should it appear on the dashboard?

In our system, users can reach annotations from different dimension levels through annotation recommendations (as their contexts are similar). Nevertheless, experts asked to have some annotations of other hierarchical levels visible as icons on the dashboard all the time. When asked to identify under which conditions this makes sense, they commented that this also depends on the nature of the analysis. They underlined the importance of

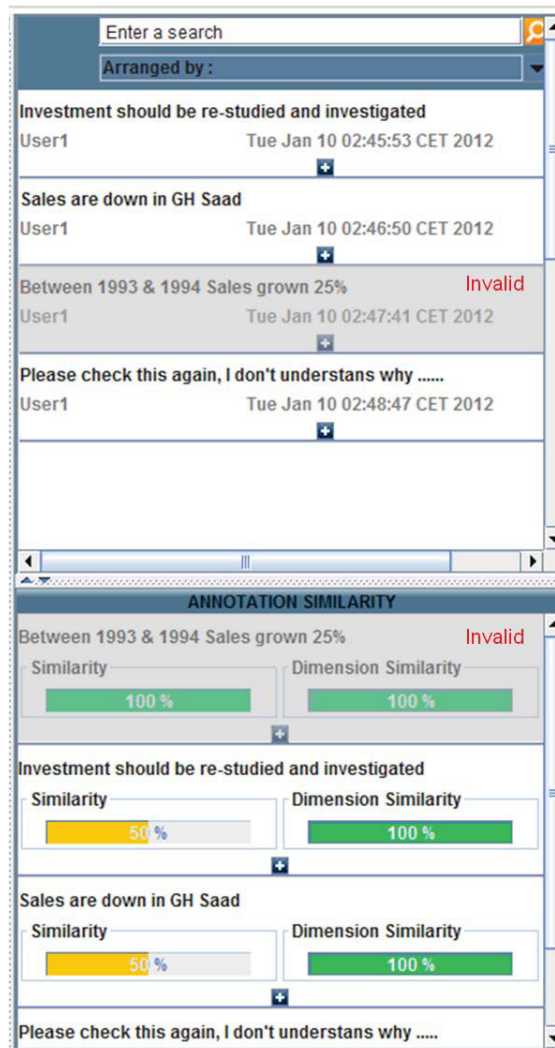


FIGURE 4.6: This figure represents how our prototype distinguishes invalid archived annotations (grayed out and written in red "Invalid"), from the valid annotations, in the annotation list.

viewing such important cross-level annotations for 2 reasons: First, to use annotations as a "bookmark" to an important point in the analysis, while the analyst explores new paths (e.g. annotate 2008 to mark the economic recession, while exploring stock market values per month). Second, to highlight important findings that could influence future analysis. The use of annotations to bookmark important analysis points is closely tied to the idea of "base views" discussed in Chapter 3.

Through our discussions with the experts, we could not conceive an automated way of detecting such important annotated points, and thus provided an option to explicitly make some annotations visible across hierarchical levels. We also provided users with an option to toggle on/off the visibility of all annotations from different levels.

The visual representation of such annotations was deemed very important by our experts: it should be different from annotations on the same level, but not too complex. After a

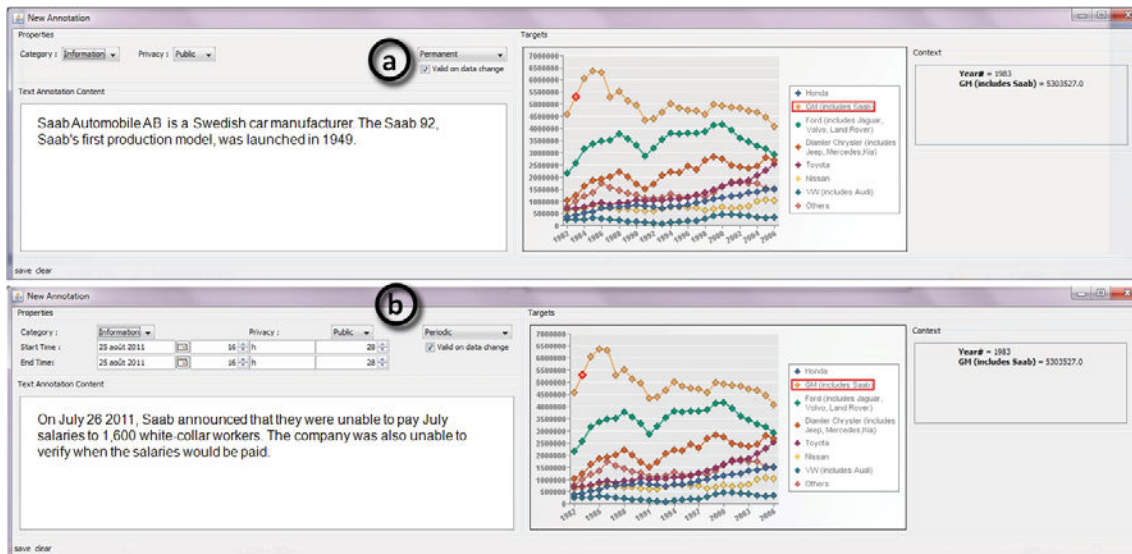
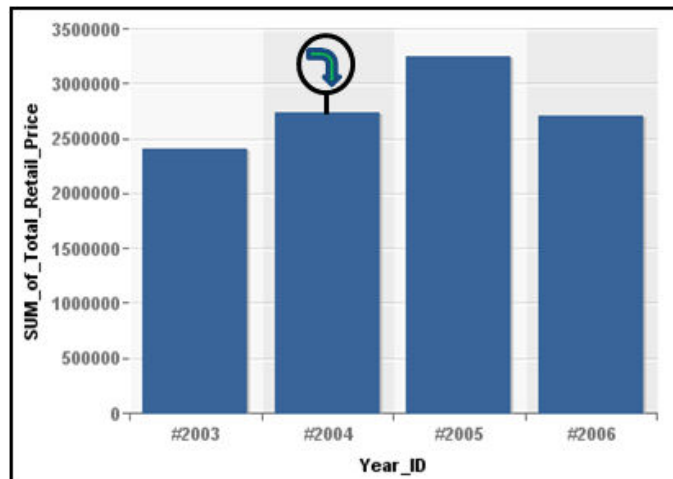
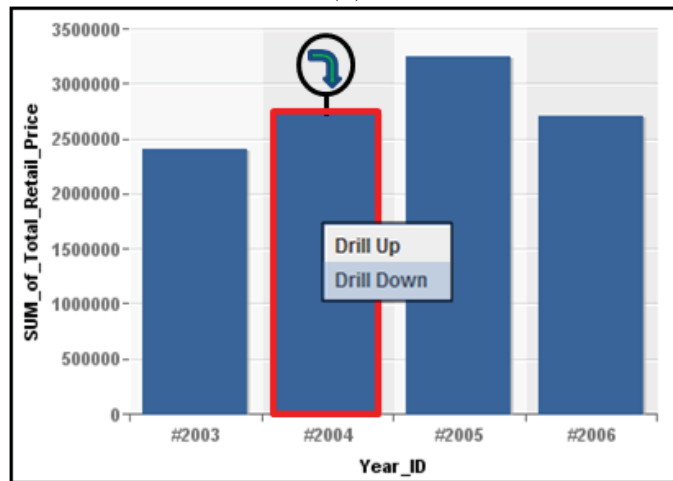


FIGURE 4.7: Two variations of the dialog for creating a new annotation: (a) permanent annotation that would be always valid, (b) periodic annotation in which the user has to identify the life time of the annotation in time periods. The user's choice of "permanent" or "periodic" periods defines the view of the dialog.

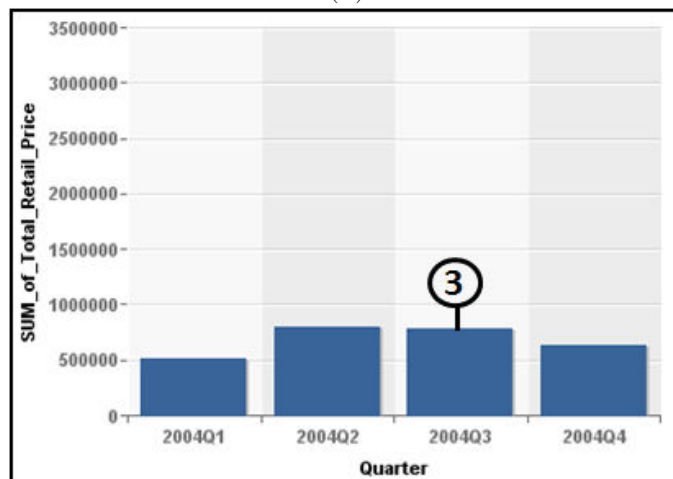
few design iterations, most agreed on using the same icon used for navigation across dimensions in the dashboard (up/down icon for annotations, drill-up/down in the hierarchy) Figure 4.8.



(a)



(b)



(c)

FIGURE 4.8: Annotation transparency through dimension levels. (a) A bar chart presents the total retail price from 2003 to 2006, a "drill down" icon indicates the existence of annotations in a lower dimension level for year 2004. (b) The user selects the 2004 bar, and right click, a list of existing actions will appear, drill up or drill down, the user selects drill down. (c) The drilled down chart to the Quarter dimension level for year 2004, shows the existence of 3 annotations attached to that level on Quarter 3.

4.8 User Study

After several iterations of the prototype with the original group of experts, we evaluated it with a different group of experts, to explore if they can use it and benefit from the different aspects of context aware annotations for their analysis. We requested that participants conduct a set of tasks similar to the ones identified in our interviews and user-centered design process. We report our observations and experts' comments and suggestions.

4.8.1 Participants

Seven participants took part in our study. They were BI experts, and only one was previously involved in the design of the system. They all had experience using dashboards. All were familiar with the concept of annotation, mostly in annotating text documents while conducting analysis tasks.

4.8.2 Procedure and Apparatus

Participants performed the study on a 15.6" HD screen laptop. One observer was present to give instructions, explain tasks and observe participants. Sessions lasted from 45 to 90 minutes, were recorded and later analyzed. They started with a survey of participants' annotation practices and ended with an exit interview. During the study we followed a think-aloud protocol, requesting users to vocalize their thoughts and actions while performing tasks.

4.8.3 Tasks and datasets

We used 4 different datasets, each presented on a different chart on the dashboard: (i) Forbes 50 Richest Americans in 2006 (50 records, 4 dimensions); (ii) Population per age group (110 records, 5 dimensions); (iii) World internet users (220 records, 3 dimensions); and (iv) US historical auto sales (25 records, 2 dimensions). Datasets were new to users, but their type was familiar (e.g. the measure *sales* and dimension *year*).

As indicated in our interviews, real analysis contains different data sources. Thus using 4 datasets allows us to examine if users could easily use annotations to link different contexts from different datasets.

Participants were told they are analysts exploring different datasets to find trends, and record their findings. We instructed them to think loud, and if they failed to do so, the observer prompted them to vocalize their insights and expectations. Participants

were given a brief explanation of the prototype functionalities (e.g. adding new and explore existing annotations). They performed tasks in 3 main groups, focusing on different functionalities of context aware annotations. (The detailed task scenarios are presented in (Table 4.1))

T1: *Usability and understandability.* Participants were requested to analyze part of a dataset and create a new annotation with their insights, attached to specific data. Then they explained their understanding of characteristics of that annotation, such as lifetime, privacy, category, and transparency across dimension granularity levels. Finally they were asked to experiment with other visualizations for the same data found through their annotations.

T2: *Retrieval of annotations of the same context.* Participants were asked to create an annotation on a specific data context and look for other annotations on the same exact context, or ones linking multiple contexts, including the current. When they detected these annotations, they were asked to determine if they were related to their analysis and explain why (based on the creator, context, datasets, characteristics, etc).

T3: *Knowledge discovery.* Participants were asked to search for possible similar annotations, and retrieve external search results related to the annotated data (through search in the databases sharing a data layer, or external searches on the hard disk and Google). Participants were asked to collect relevant material and add them to their analysis annotations.

Tasks began by loading a dashboard containing 4 charts, each representing a data set. Participants could change the chart visualizations and were given time to interact and familiarize themselves with the datasets and dashboard before starting the tasks. At the end of the tasks a short interview (Table 4.2) was conducted to understand how context aware annotations impacted participants' analysis, to clarify observations made during the study, to prompt users to identify difficulties/barriers they faced, and to suggest improvements.

4.9 Results

We present here feedback from participants on how the system supported their analysis, and on its visual design.

#	Task Scenario	Purpose of Task
Context	<p>You are an analyst that explores datasets, and analyzes them to find trends and contribute information. We provide you with an exploration system (dashboard), that permits you to:</p> <ul style="list-style-type: none"> - annotate data points - annotate multiple targets (charts) - explore annotations done by other users - retrieve previous annotations context - recommend annotations - give a similarity measure for recommended annotations - retrieve information from Google search related to a context on your dashboard <p>We would like to get your feedback of those functionalities on your information retrieval while analyzing different datasets.</p>	Permit participant to understand the context of the tasks and introduce the system use.
T1	<p>Read the first chart top left in the dashboard and describe your interpretation (Internet usage)</p> <ul style="list-style-type: none"> - What are the top 3 countries in internet usage in the world? - Annotate France with its order over the world. - Explore the creation interface that appears and explain your interpretation of the presented features. 	Usability and understandability
T2	<p>Explore the second chart on the first row (Population by group age).</p> <ul style="list-style-type: none"> - Find trends (min/max values for each age group) - Annotate the max population value for the age group 20-59 in the last 5 years - Check similar annotations <p>*Point with the mouse the area you're reviewing. *Express out loud what you expect as a type of action or result before you click on an item. *Explain your interpretation of similarity and difference between global similarity and dimension similarity.</p>	Retrieval of annotations of the same context
T3	<ul style="list-style-type: none"> - Explore the third chart (50 Richest Americans in 2006) (Figure 4.9.a) - Analyse the chart and express out loud what you interpret. - Show annotations on the dashboard and see if any related to this chart. - If any express out loud what you find. - Who is the second richest person in USA in 2006? Do you know him? Lets search on Google! 	Knowledge discovery

TABLE 4.1: Tasks users performed during the study of context-aware annotations on dashboards.

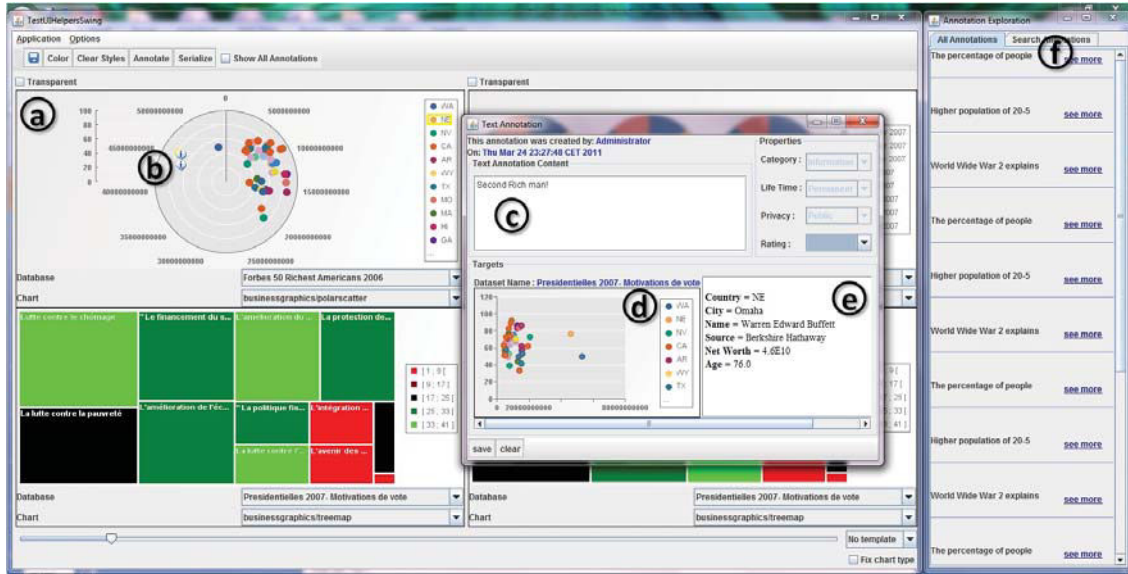


FIGURE 4.9: An annotation example that the participant explored during performing Task 3. (a) Radar chart presents the "50 Richest Americans in 2006". (b) Annotation icons that declares the existence of annotations attached to two data points on the chart. (c) In a new window we see the content of an annotation that says "Second Rich man!". (d) The original chart type (scatter plot) when the annotation was created. (e) Context of the data point, the annotations is attached to. (f) List of all existing annotations.

#	Question
Q1	Do you find adding annotations can add any value in the analysis process, why?
Q2	Do you think reading other users' annotations useful?
Q3	What do you think about attaching annotation to data points regardless the visualization type? (transparency)
Q4	How relevant were the recommended similar annotations to your selected context?
Q5	What are the system difficulties/limitations/missing functionalities?

TABLE 4.2: Post-test (context-aware annotations tasks) questionnaire.

4.9.1 Supporting Analysis

Participants completed all tasks successfully without any help from the experimenter. We present here observations made regarding the use of context aware annotations in our tasks.

T1: *Usability and understandability.*

- All participants understood the detailed data context(s) of the annotations, and explained that understanding the context was easy when "having the annotated data visually highlighted in a chart", and having a "detailed textual description of the data context" (dimensions' and measures' names and values) on the annotation body.
- The transparency of annotations between charts allows users to understand new chart types, using the annotated data as reference points across the two visualizations. A participant mentioned "the radar chart was unfamiliar for me, so I switched it to a scatter plot and I was able to still see the annotated data points. The annotation helped me to understand the radar by reviewing the scatter plot and see that I am looking at the same data point".
- Annotating data points vs. charts helps users see changes of important data points. Our system takes snapshots of annotated data points whose context changes. A participant commented that "annotating important data points acts as a way to see their changes through time".

T2: *Retrieval of annotations of the same context.*

- All participants found attaching annotations to data points (vs. entire charts) very important for verifying annotation relevance to their task, both for searched and recommended annotations (relevant to **T3**). As one participant mentioned, this "maximizes the analyst's focus on targeted data and not all the chart, and prevents losing time trying to verify the context for each searched or recommended annotation".
- All participants appreciated the use of annotations as reminders of past work, and as highlights of the important data in charts ("at a glance I can determine the important data points, no matter where I am").
- Finally they all appreciated the use of annotation transparency for all users across all visualizations, as a way to avoid work replication (for example prevent duplication of annotations among multiple analysts).

T3: *Knowledge discovery.*

- All participants were able to navigate the list of recommended annotations, and interpret their similarity. They mentioned they understood why each recommended annotation had the presented similarity, by comparing the context of the targeted

data points and that of the ranked recommendations. Five found that the global similarity measure that combines dimensions and measures is enough, while two would have also liked a text similarity metric.

- All noted the usefulness of recommending annotations for re-use or peer learning ("from notes of others"). They commented that it is especially useful because the reason for the recommendation is clear, due to clear data contexts.
- All participants found retrieving external data related to the annotation context very useful. They all commented that searching for supporting material during analysis will be easier if they can retrieve data from their company's portal using the data context to search other BI applications with the same data layer, or search for files containing the context keywords on their machine or on Google.

4.9.2 Visual Design

Participants provided us with design suggestions to improve the visual interpretation of annotated targets. Their feedback was used to reach the final design presented in this chapter. Here we present some of their comments that influenced the look and feel of our prototype. The first point raised was the use of icons on the main dashboard to indicate annotation on specific data points. Although they found this useful, they suggested that if a data point has many annotations we should use a different icon, that when clicked shows a drop down list of all annotations. This would introduce another annotation icon (apart from that of simple annotations and annotations at different granularities). So we chose an icon containing the number of annotations attached to the data point. The two annotation icons are seen in Figure 4.1 and Figure 4.8.

The other recommendation was to visually link annotations targeting multiple charts with lines across the dashboard. One participant had at first some trouble understanding that the charts in the detailed annotation window showed the multiple targets of an annotation. She explained that this could become more intuitive if annotations with many targets on the dashboard were connected with lines to the targeting charts. Thus at the end of the study we provided three design suggestions to our participants: the first (Figure 4.10.a) highlights targets in red, the second (Figure 4.10.b) draws arrows to the targets, and the third (Figure 4.10.c) fades out all related targets. All participants preferred fading out. As one explained 'this shows only points I should focus on while still seeing the whole context; moreover I can print this representation and share it because it has less clutter and more interpretable'.



FIGURE 4.10: Three designs linking annotations in the annotation list with their data targets. Experts preferred the last design.

4.10 Summary

We have collected a new set of requirements from BI experts for annotating dashboards. Context-aware annotations across multiple visualizations on a dashboard can meet these users needs. Based on the collected requirements and following a user-centered design approach we built our context-aware annotation prototype. We identified and provided solutions to challenges in using these types of annotations, notably issues arising when the annotation's "context" changes (e.g. annotated data are deleted or changed). Based on feedback from experts, such annotations should be achieved with a snapshot of their context, as long there is an easy way to retrieve them (e.g. through recommendations). Users should also be able to define the lifetime of annotations.

Our prototype was evaluated with a different set of BI experts that were able to easily perform several tasks that they deemed important in their work: use annotations to discover other visualizations with the same data context and look at data in a new perspective; use annotations to link and organize multiple charts related to their analysis; discover charts and annotations through annotation recommendation and understand the reasons behind it; and use annotation context to find related information outside the system.

Chapter 5

Storytelling in Visual Analytics tools for Business Intelligence

5.1 Abstract

Stories help us communicate knowledge, share and interpret experiences. In this chapter we discuss the use of storytelling in Business Intelligence (BI) analysis. We derive the actual practices in creating BI stories from in-depth interviews with expert BI analysts (both story "creators" and "readers"). These interviews revealed the need to extend current BI visual analysis applications to enable storytelling, as well as new requirements related to BI visual storytelling. Based on these requirements we designed and implemented a storytelling prototype tool that is integrated in an analysis tool used by our experts, and allows easy transition from analysis to story creation. We report experts' recommendations and reactions to the use of the prototype to create stories, as well as novices' reactions to reading these stories.

5.2 Introduction

Stories are one of humanity's communication structures and storytelling a means of passing on wisdom and culture. Individually and collectively, stories help us make sense of our past and reason about the future. Johnson [124] and MacIntyre [125] argue that story narrative goes beyond communication, it is also a process of extracting meaning from events, that is central to human experience and conduct. It is thus not surprising that intelligence analysts, people who make sense of data and identify links between disparate pieces of intelligence, are interested in storytelling. Their analysis process is supported by the construction of stories and narratives. Bier et al. [82] point out that "a story is a powerful abstraction used by intelligence analysts to conceptualize threats and understand patterns as part of the analytical process".

As analysts continue to work with increasingly large data sets, data visualization has become an incredibly important asset during analysis, and when communicating findings to other analysts or to a broader public [52, 65]. Given the importance of storytelling in the analysis process it is clear there is a need to enhance visual analysis tools with storytelling support. Nevertheless this process is not simple [101, 126], as analysts need to highlight and explain items or events of importance and their connections, within very large data resources. Despite the growing use and research work on storytelling and narrative visualization in the visualization domain [1, 52, 103], related research on the domain of BI has not equally progressed (see section 2.5). Our work attempts to rectify this.

As previously discussed, in Business Intelligence (BI) analysis the most popular visualization tools are dashboards. Nevertheless, just collections of visual representations do not speak for themselves; to become meaningful they require interpretation and explanation, often presented in a story narrative.

Our findings indicate that although the work practices of BI analysts' may differ, in the end they all construct stories that include: collections of visual components (dashboards, charts or tables) to support their stories, verbal explanations during presentation(s), or detailed textual annotations and descriptions in the form of written reports and external resources such as wikis. This process remains cumbersome and requires frequent switching between different software (dashboard analysis tools, screen captures, text editors, presentation or drawing tools, etc).

Our work aims to aid BI analysts with their storytelling needs, and makes the following contributions: (1) Based on interviews with expert BI analysts (story "creators" and "readers"), we derive and provide a better understanding of the actual practices in creating BI stories. (2) We emphasize the need of BI storytelling support and extract new requirements for enhancing BI visual analysis tools with storytelling capabilities. (3) Following

these requirements and using a user-centered design approach, we design and implement a prototype storytelling tool that is fully incorporated in an existing visual analysis dashboard, to fluidly transition between analysis and story creation. (4) We report on feedback from BI experts creating a story on the usefulness of the prototype as a communication and teaching medium, and from BI novices reading a story on its effectiveness for story communication.

5.3 Interviews

To support visual storytelling in BI, we attempted to understand current practices in transmitting BI knowledge and analysis results, and identify the challenges experts face when creating and sharing their stories or reading and interpreting the stories of others. We interviewed 5 BI experts occupying positions in SAP which is a leading business intelligence development company, with experience in the BI domain ranged from 6 months to 12 years. Participants job descriptions included development project manager, project manager, development director, data warehouse engineer and delivery manager. Three experts used dashboards and communicated their analysis or read analysis from others daily, while 2 several times a week. Interviews were held either in person or by phone and lasted on average 1 hour. We report next our findings.

5.3.1 BI Reports

All experts communicate their analysis or read analysis from others in the form of *BI reports*. These contain an entire dashboard, often accompanied by several single charts and tables. Their purpose is to help monitor aspects of business performance, by highlighting with charts key performance indicators (KPIs) that indicate success of project management and progression of development teams. Our experts were report creators, but occasionally acted as readers of others' reports.

A BI report can be a single page, with only a title and a dashboard as in Figure 5.1 and Figure 5.2. This one page report summarizes the most important data, and can serve as a starting point for longer reports, up to 50 pages in length, that give more details. Details can be additional visualizations, tables, annotations, links to the data used in the visualizations, and finally block text, although all participants preferred limiting text to 5% of the space in a report. Thus BI reports have a very strong visual component, with little text added for explanations.

The one page report is preferred by clients such as company managers that want an "at-a-glance view" of a project status. They usually see many such reports during a

day, thus require answer to their question(s) in a timely and targeted manner. Longer detailed reports are used more to communicate findings to other analysts, project teams and product owners.

Experts mentioned that BI reports are used to: spacing

label Answer specific questions (e.g What is the progress of code development in project X the last 3 weeks?);

label Investigate specific data (e.g Investigate the increase of bugs in section Y of project X);

label Manage conflicts and highlight problems (e.g Team X completed less use cases than planned, because test team Y did not complete testing);

label Interpret past data and predict future trends (e.g Given last year’s sales last, what is the sales projection for this year).

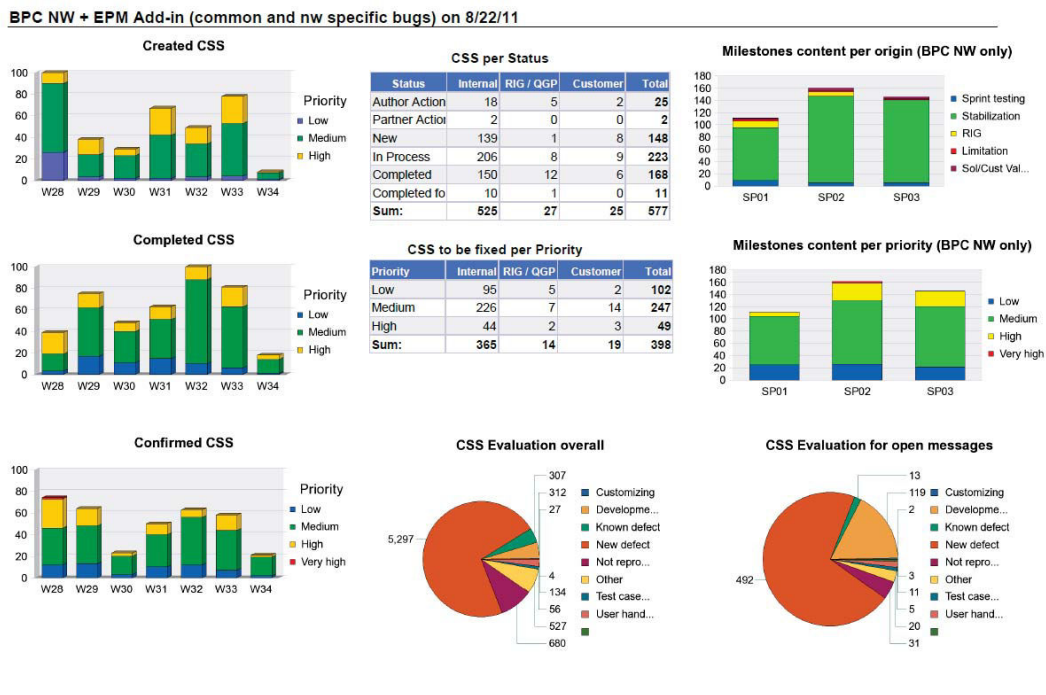


FIGURE 5.1: First page of a BI report given by an expert: a dashboard snapshot that presents the bug status in a specific development project. (Used with the experts’ permission)

5.3.2 What is a BI Story

Experts explained that reports are difficult to understand without detailed explanations from their creator. Thus they don’t represent a complete BI story on their own. Before

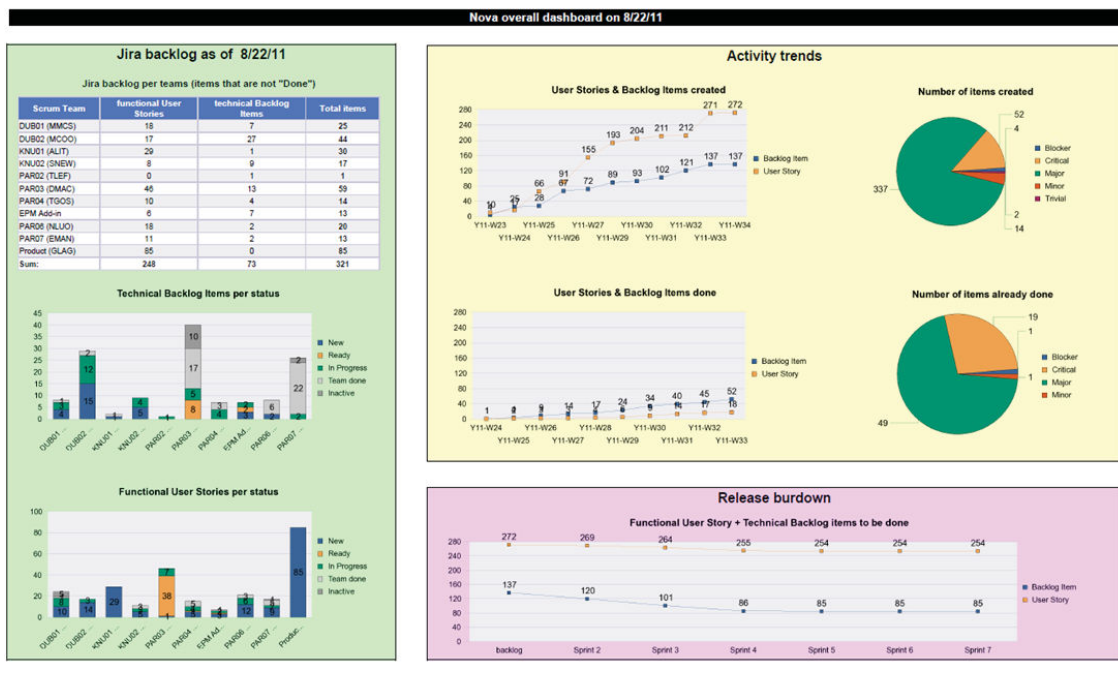


FIGURE 5.2: An example of a BI report given by an expert that was used in the evaluation of our system. This dashboard represents the "technical backlog" & "functional users stories" items status in different perspectives. (1) Green group, represents the status of the items per team. (2) Yellow group, shows the created items versus the completed "done" items, and the category (critical, major, etc) of each. (3) Pink visualization presents the remained items to be completed and its progression through time.

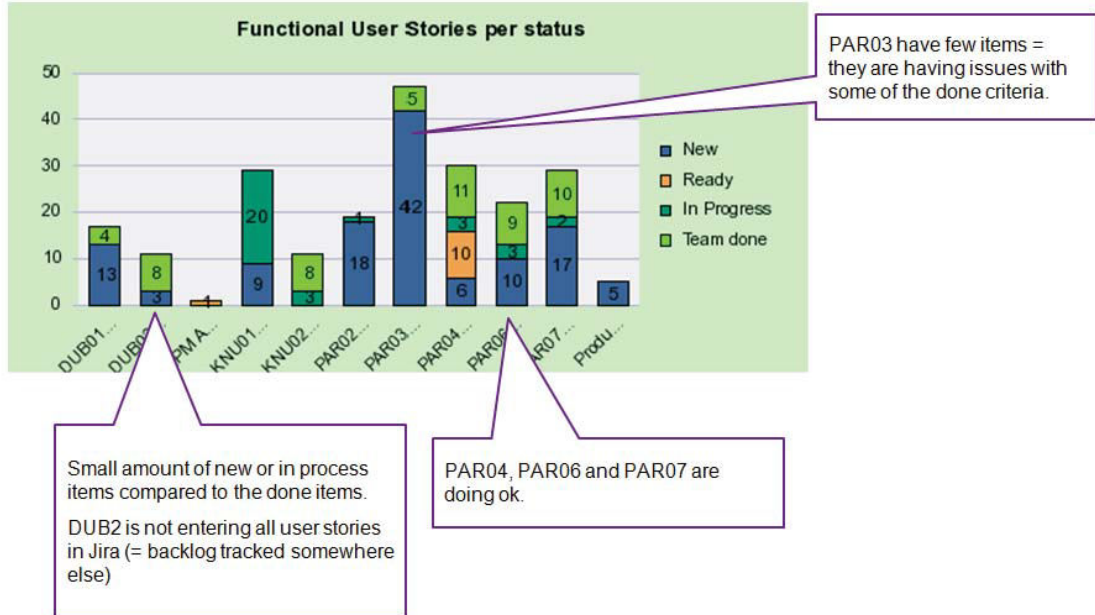
publishing a new report, the creator provides an introduction session to all report readers. During this session (that lasts usually an hour), she explains with a presentation the entire story and goal of the report, the meaning of each chart, the relations between different KPIs in different charts, as well as in what sequence to read it. Thus their verbal explanations and presentation slides are ways for creators to explain their analysis path. These presentations often show relational diagrams, text explanations and interpretations, highlight specific visualizations or parts of visualizations using colored highlights, arrows and other symbols see Figure 5.3. The report creator often adds by hand a drawing on a chart identifying the ideal data values, to help compare the current situation to the target goals.

The audience can ask questions during the training session or afterwards, through emails or arranged meetings.

Content very similar to the presentation, with more textual explanations is often put on a wiki page to help the audience and answer follow-up questions. The session recording, wiki page, and presentation slides, are made available to the audience as reminders and reference material. Experts explained that this material is not a complete story either, as it does not include the visualization and data details shown on the BI report.

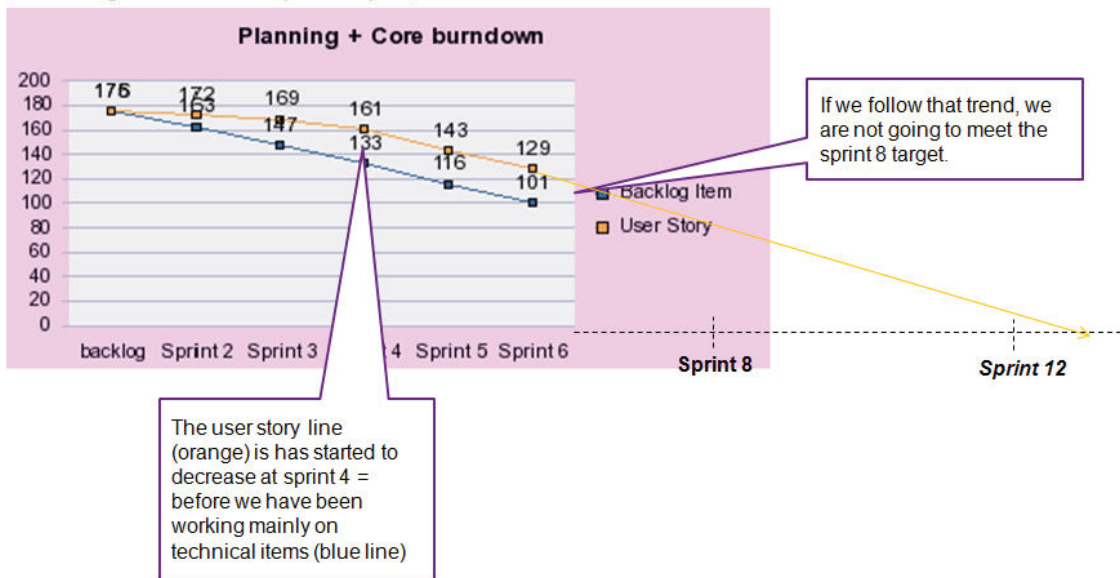
What can teach the dashboard?

This chart gives the number of User stories in each team's backlog per status.



(a)

This chart gives the number of user stories and backlog items that are remaining after each completed sprint



(b)

FIGURE 5.3: An example of explanations included in a power point presentation used as an introduction session to report readers. Example provided by one of our interviewed experts.

Thus a complete BI story is a BI report with visual representations of the most important data, accompanied by instructions on how to read the report, in the form of presentations, and verbal or textual instructions on a wiki.

5.3.3 Teaching BI storytelling

According to our experts, the material making up a BI story is also a tool for training new analysts. As one recently trained report creator explained, this material taught her the key aspects of a BI report and how to interpret them. When analysts first start out, they read in detail older reports and all their supporting material to understand how to read and analyze visualizations, see their relations, and how to identify the important points and their link to KPIs.

Experts mentioned that designing a report from scratch is hard and needs a lot of experience, thus they often use a template that they modify according to their needs. Three stressed the influence of a senior report creator (the 5th interviewed expert), and their reliance on his BI report templates to create their own. They still occasionally contact the senior creator when they face difficulties or are unsure of the clarity and completeness of their messages in a report.

The supporting material is also often reused. All experts explained that data (and thus reports) change, but often the structure of BI stories repeats itself. Moreover, they sometimes create a sequence of reports on dynamically evolving data (e.g. sales over several years). These stories change very little. Thus material on how to read reports can be reused by adapting it to new stories. Nevertheless, as experts pointed out, this adaptation can be a tedious and repetitive process.

5.3.4 Creating a BI Story

Experts explained that report creation follows several iterations of communication with a client (usually a decision maker, like a manager or executive), who is one of many possible report readers (a process identified in Chapter 3). Thus the focus and structure of the story changes and evolves through the communication between the story teller and the story reader over time, a process that does not appear in most other visual analysis domains and visualization storytelling systems, where the needs of the reader are not directly reflected into the creation of the story. This interactive report evolution process can be heavyweight (through phone calls, emails, etc), and can take several iterations to iron out, while the communication details are often lost between different versions of the report.

At each communication iteration, creators explore the desired data using interactive visualization dashboard tools, create initial visualizations, provide additional information and details about the analytical techniques used, and finalize reading paths (i.e. how to read a report).

All interviewed creators use an in-house reporting tool, that gives access to different data sources and can extract interactive dashboards, individual charts and tables from their dashboard analysis system. It also provides the capability of adding text. Three experts also used Excel for creating less complex reports. Our experts identified several issues with this process. Firstly, when extracting interactive visualizations from their analysis tools, metadata and annotations from their analysis are not extracted and have to be recreated. Moreover, the annotation capabilities of report creating software is very limited. There is no ability to point to specific data points, and sequences and connections cannot be displayed graphically but have to be explained in one of the supporting material. Finally the report creation tool can be in-house software, thus to share their reports with others, creators often extract static snapshots of the tool output and save it in PDF format to share with their readers.

Some report creation tools give access to interactive visualizations and dashboards through hyperlinks. Our experts like this option and add these links to their non-interactive reports (PDF, Excel) and presentations. Readers within the organization may have access to the in-house reporting tool, and can then interact with individual visualizations or dashboards (e.g. query and filter data, or drill down/up). This interactivity is lost when reports are communicated outside the company.

To overcome the shortcomings of reporting tools, creators are forced to provide the supplementary material (presentations, wiki pages) and when possible links to interactive visualizations for in-company clients. They described this process as limiting and requiring a lot of work duplication.

We requested clarification on two points regarding the possible integration of analysis and storytelling tools: (i) Given that analysis dashboards are exploratory environments that allow users to interact fully with data, should such interactions be allowed in the storytelling? (ii) If creators had the opportunity to create their entire story with all support material in one place, how would they prefer the story to be visualized?

Interactivity Due to the constraints of current reporting tools, shared reports are often non-interactive when accessed outside the organization. When asked if it were possible for the visualizations in their reports to be completely interactive, would they encourage their readers to interact with them (e.g. using drill down/up, filter, link & brush), four of our

experts prefer to have interactive visualizations that permit linking and brushing (i.e. data selection). But they would limit the more advanced interactions such as drill down/up or filtering. They felt that all needed data to tell the story should be displayed clearly in the report without the need to explore the data further. The fifth participant would not be opposed to fully interactive visualizations. Thus authors feel business stories and related visualizations should be mostly author-driven and constrained, a practice that is known to work best when the goal of a visualization is storytelling or efficient communication [1].

Story templates Going from the current practices of storytelling (BI report, presentations, wiki) to a dedicated storytelling tool is not straightforward. We thus showed our experts a group of story templates identified by Segel et al. [1] to see if they met their needs. All chose the "Annotated chart" as the preferred template (Figure 5.4), with the modification that it should have multiple charts on the same page (dashboard) that they can annotate. Four experts identified "Partitioned poster" as a potential template, where the side list of charts display details that support the main chart/charts in the central region. Three mentioned that the "Slide Show" template is useful both as a means to focus attention on each chart and zoom to details, and as a step-by-step presentation that clarifies the analysis path and the ideal reading sequence. One expert mentioned the "Flow Chart" would be useful to show some business scenarios, like following bugs during development (discovered, tested, fixed etc). And another mentioned that the "Comic strip" could be a useful template, but with added annotations.

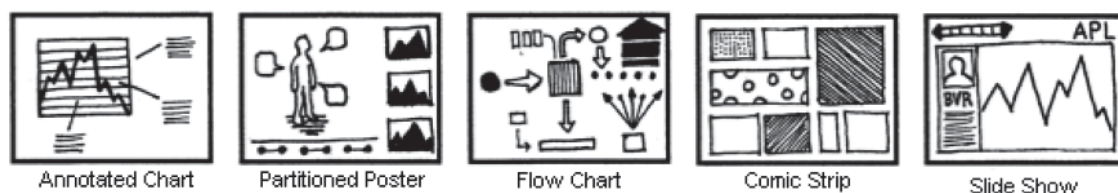


FIGURE 5.4: Genres of Narrative Visualization selected by BI experts from a group of templates defined by Segel et al.

So besides the templates identified by Segel et al.[1], for BI stories we need a new template that consists of an annotated dashboard, that combines multiple templates from that group. Our users attempt to do this with their current reporting tools, but they are limited (cannot annotate detailed points and relations), or they create it manually in their presentations.

5.4 Participatory design session

We brought back the senior BI expert, that was involved in training of other analysts, to identify the best design for a BI storytelling tool. She provided us beforehand with one of their old reports and supplementary material. During this session we gave the expert a copy of one of her material, and several cut-outs of the entities in the report (charts, titles or tables). She was also given "narrative aids", such as arrows, lines, numbered items, grouping containers of different colors, and annotation bubbles (Figure 5.5). She was asked to construct a stand-alone story of the report on an A3 paper, with accompanying audio if necessary, in a way that it can be interpreted by a general audience with no training. The session lasted 2 hours and was recorded.

Figure 5.6 shows the final version of the manually created story by the expert, where we can see the intense usage of annotations and explanations, connected by arrows and lines. Other structures used by the participant were chart grouping rectangles (enclosures for placing charts that need to be read together), sequence markers (what needs to be read first), and highlight markers (to draw attention to part of the story).

At the end of the session, we asked our expert how she would prefer that the readers see the story, as a static image or as an animated presentation. Our expert explained that both are needed, the static representation shows the "entire story and gives context", whereas the animation "focuses the audience where I want". She then played out for us how she wanted the animation to be presented by explaining when and how to zoom to specific areas of the story. This play-back was recorded using static shots.

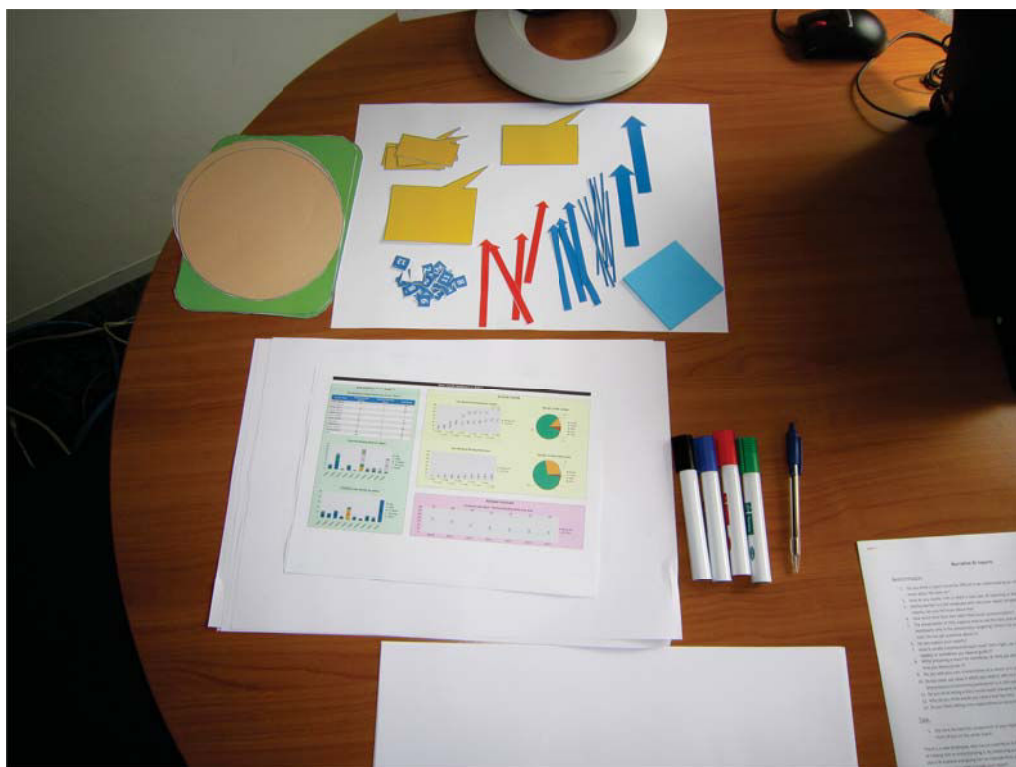
5.5 Requirements for Creating BI stories

BI reports, the main communication medium of BI analysis, consist of snapshots of visualizations and textual descriptions, made in a different environment than the one used for the analysis. These reports are fact-based and can be interpreted only by an expert audience. To increase their audience, report creators use supplementary material, such as wiki pages, presentations, and when possible links to interactive visualizations. Based on our experts' comments and the participatory design session. To create a more effective BI story, we identified a set of requirements for such as storytelling tool:

R1. Fluid Transition. Report creation tools are different than the analysis tools used to explore data and create visualizations. This costs time and effort to export visualizations and limits the possibility of embedding meta data or annotations created during the analysis. To ensure that creators do not recreate this missing information, they require a



(a)



(b)

FIGURE 5.5: Participatory design session settings, that took place at SAP UX room.

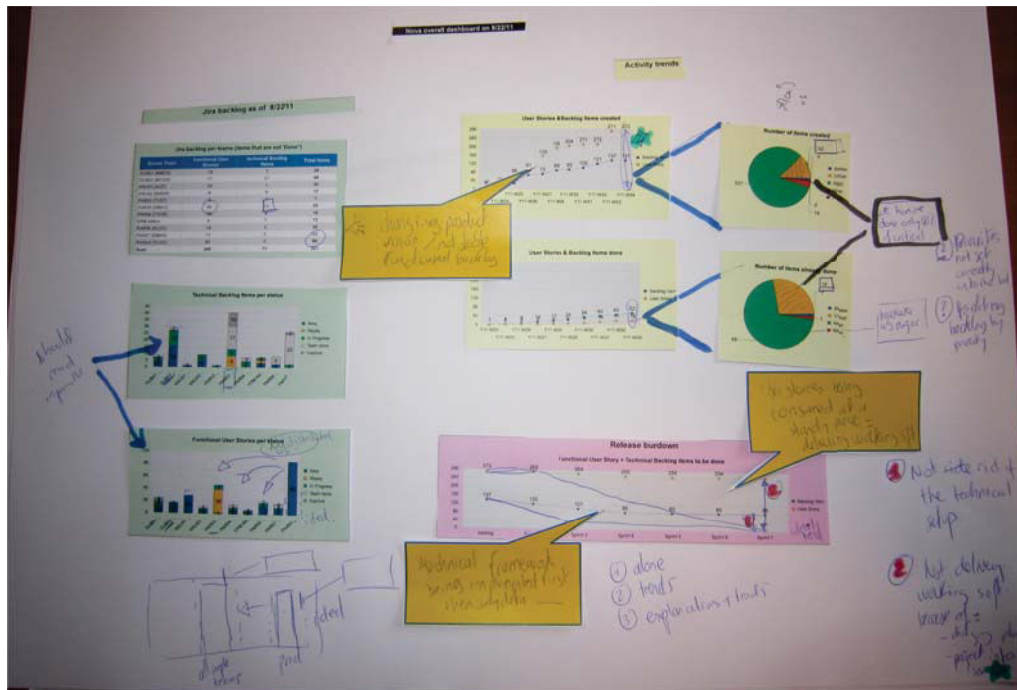


FIGURE 5.6: Story created by an expert during the design session.

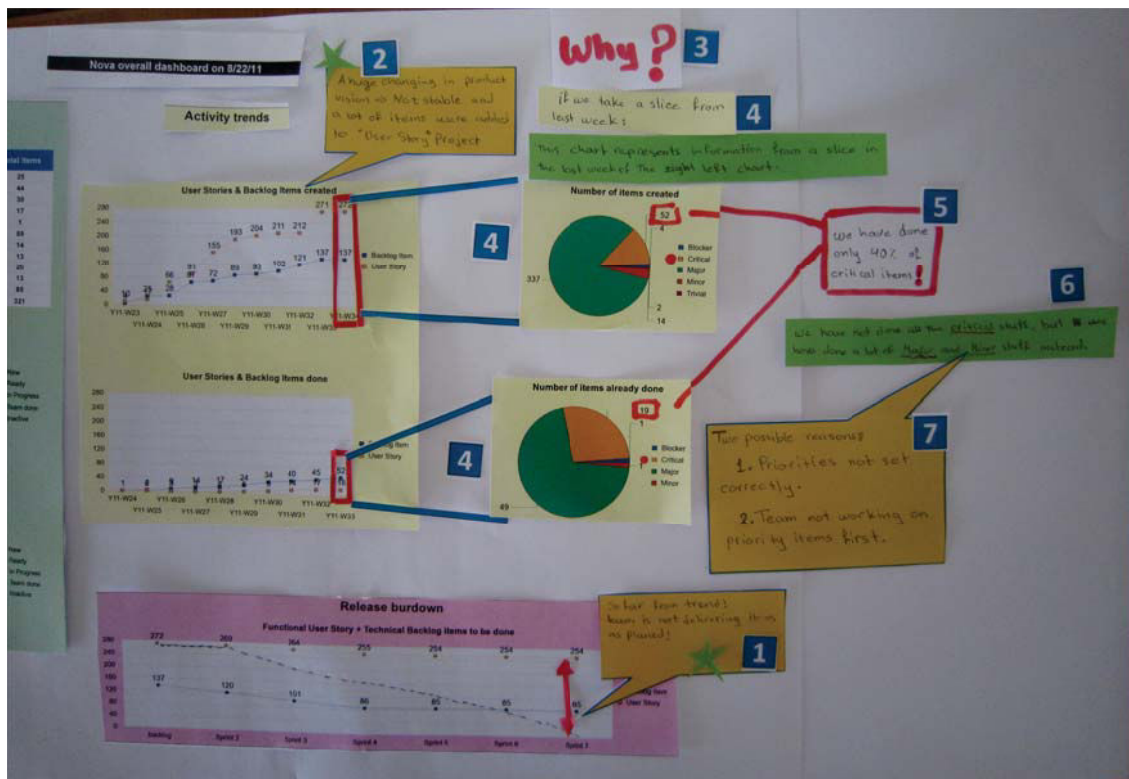


FIGURE 5.7: The final story mockup created by an expert (fully annotated and sequenced story seen in Figure 5.6). It represents how a static story on a system should look like.

fluid and integrated way to transition from their analysis and meta data associated with it, to report creation.

R2. Integration. To tell their stories, BI creators need tools that combine all materials used currently in their story creation: BI reports, interactive visualizations, ways to indicate story structure, highlighting capabilities, presentation of the story in sequence, and textual or audio explanations.

R3. Narrative visual aids. Report creators need to add focus expressions to draw attention to specific visualization data, such as highlighting, coloring, annotating and zooming. They also require ways to indicate reading sequence (e.g. vectorial references, such as arrows [127][128]). These are not available in reporting tools or BI analysis tools.

R4. Interactive visualizations. Visualizations on shared reports are often non-interactive when read outside the organization. A storytelling tool should have completely interactive visualizations, although the way that readers interact with the data should be limited (by default to brushing and linking) and be controlled by the creator. This balance has been identified as a challenging aspect of storytelling [126].

R5. Appropriate BI Story templates. BI stories have specific structure not necessarily shared by other story narratives identified by Segel et al.[1]. Our experts identified templates of interest (Figure 5.4) and highlighted the need for a new template that consists of an annotated dashboard.

R6. Reuse. Although BI reports and data changes from analysis to analysis, often the underlying structure of BI stories remains the same. It is thus important to be able to easily reuse the structure of stories created within the tool, both for stories of evolving data and similar future stories.

R7. Playback. Readers should be able to see a static representation of the whole story to get a general overview, but also be provided with a guided playback of the story to help them follow analysis paths that are not easy to discover, and ensure they understand the important points according to the creator.

Such storytelling support would alleviate some of the burden associated with creating and learning how to read BI stories. We are not aware of any BI dashboard or other multi-chart visualization system that currently provides all the above functionalities.

5.6 BI Narrative Prototype

Following these requirements, we extended the dashboard exploration tool described in Chapter 4, to support the creation of BI stories.

5.6.1 Exploration/annotation dashboard (Figure 5.8)

The user starts from a traditional analysis dashboard, a collection of coordinated (synchronized) charts connecting one or more data sets. Our dashboard provides advanced exploration capabilities such as data selection and filtering, and OLAP functionality (e.g. drill-up/down) [46]. We build our narrative tool on top of this visualization dashboard system [129] that supports annotations on "data targets".

After or while conducting her analysis on the dashboard, the analyst can choose to create a story. The menu option "Create story", opens up the *narrative board* window. All visualizations and their annotations (annotation text + data targets highlighted) are immediately placed in the narrative board. Thus the analyst can transition fluidly (**R1**) from analysis to story creation. Because dashboards can present evolving data, but a story can be an instance in time, by default each visualization is placed on the narrative board with a time-stamp to indicate when the visualization was taken from the data.

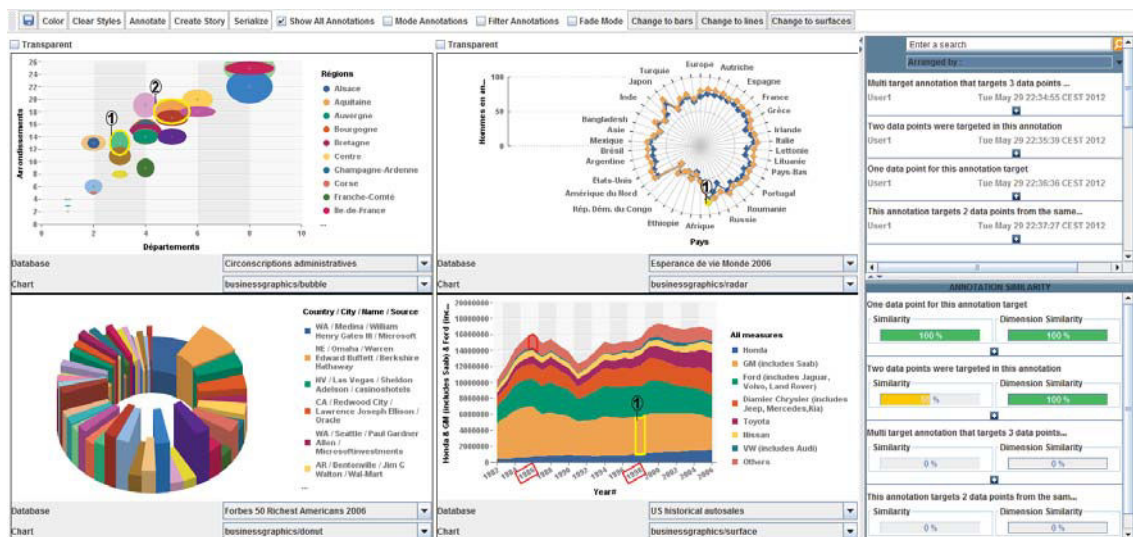


FIGURE 5.8: Dashboard with visualizations on the left and list of annotations on the right.

5.6.2 Narrative board (Figure 5.9)

Here users can create a BI story. Visualizations and annotations from the dashboard appear on the main part of the window. By default they appear as the dashboard (to support the dashboard template identified in our interviews **R5**). Users can then resize and re-locate all story entities freely, or choose to organize them using other appropriate narrative templates (**R5**). They can edit visualization titles, add/remove and edit annotations of entities, and add text entities. Users can add visual entities that indicate relations, grouping and emphasis (**R3**) to help create stories, available at a tool palette on the right.

The story can be seen in: a static representation, where all entities remain one screen; and a playback representation, where entities are highlighted sequentially according to an author defined sequence (**R7**).



FIGURE 5.9: Narrative board (a), Sequence of entities (b), Relational (c) and Organization (d) entities, and (e) Playback control.

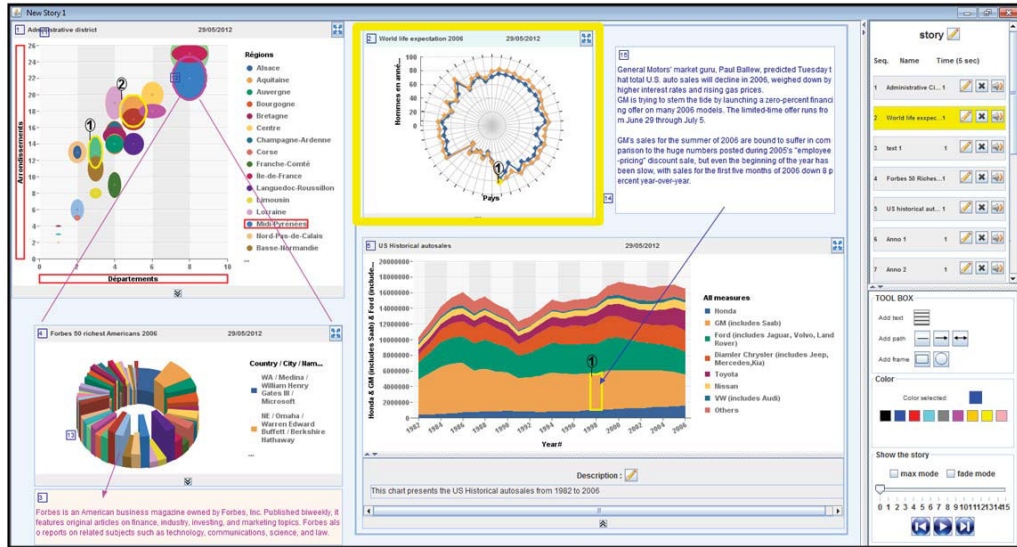
We categorize the entities available on the narrative board as:

1. *Information entities: visualizations, text, annotations.* These are mainly the imported visualizations and annotations. For annotated visualizations, the annotation text, the visualization, as well as the annotated data target (the specific part of the visualization that is annotated), are all defined as story entities. In addition to

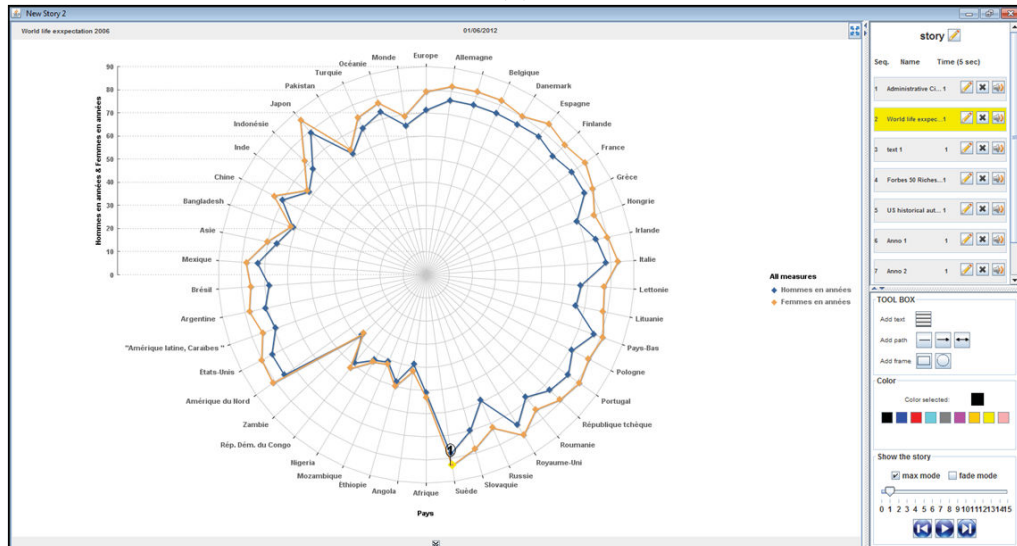
the entities created during analysis, users can place text anywhere on the narrative board to add further explanations to their story. Moreover, new data points on visualizations can be selected and highlighted with different colors to focus attention on and add them to the sequence of entities in the story.

2. *Relational entities: arrows, lines, html links.* A story can include relationships between entities, such as causality. We provide visual arrows, lines, html links and other vectorial references to define relations between story visualizations and text entities (Figure 5.9.c).
3. *Organization entities: visual grouping and sequences.* Previous work [97] and our interviews emphasized the need of visual grouping of entities that are to be seen together in a story. We support this with entity grouping borders. Moreover, our experts indicated the need to define a reading sequence for each entity to show readers how to move through the story. We achieve this by allowing users to define the order of appearance of all entities through a sequence list (Figure 5.9.d). Users can change the entity's sequence in the list, delete, or rename any entity. They can also define a playback time for each entity, a time for the entity to be in focus in the playback presentation mode. Audio can also be recorded for any entity, to be played when the reader watches the story in playback mode.
4. *Emphasis entities: highlighting and zooming.* To focus reader's attention toward a specific story entity, users can add color highlights to any entity (e.g. a visualization, an arrow, or text entity), or select and highlight data targets within visualizations (e.g. a bar in a bar chart visualization). This color highlighting can be present in the static story presentation (always visible), or in the playback representation (the color highlight appears at a specific point in the story timeline). Besides color highlighting, users can add zooming highlighting to any story entity, that will take place in a specific point in time during playback.

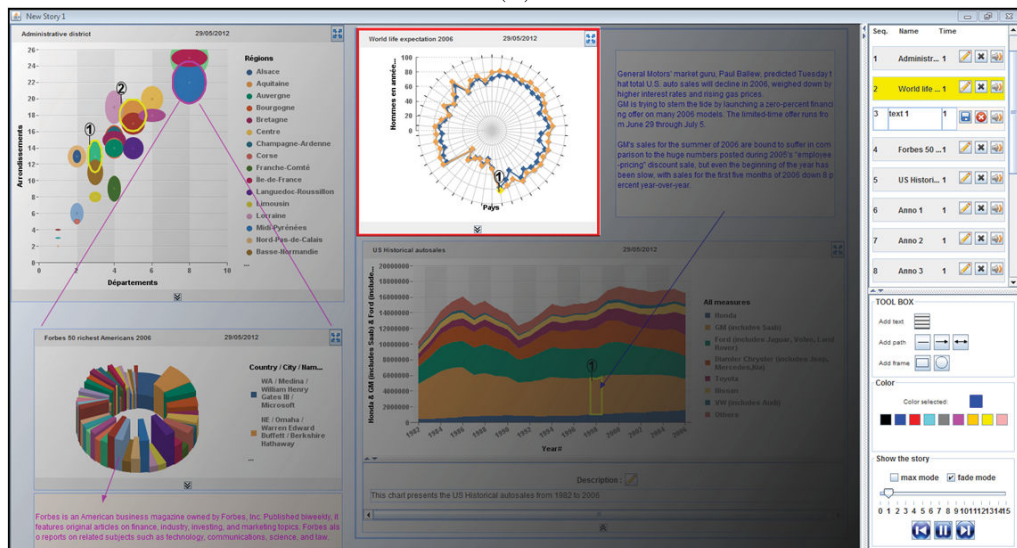
Playback: As discussed in **R7** our experts desire 2 ways of showing their stories to their readers, a *static overview* version of the story that provides context, and an *animated playback* of the story. Our narrative board acts as the static version. We provide three options of the animated playback where entities are given focus in sequence for a given duration. In the color highlight (Figure 5.10.a), entities that are in focus change color to grab attention (default playback). In the "max playback", entities that are in focus are zoomed-in to the maximum possible size, taking up the entire narrative board (Figure 5.10.b). Finally, in the "fade mode", during playback, all entities except the entities in focus fade out (Figure 5.10.c). As mentioned, authors can record audio to accompany the playback of entities.



(a)



(b)



(c)

FIGURE 5.10: Three storytelling playback modes. While playing the sequence of elements (the element sequence list on right) in a story, each element (highlighted in the element sequence list on right) will be: (a) Highlighted, (b) Maximized, (c) or Faded out, on the storyboard in the play back mode.

Interactive visualizations: Our visualizations are imported with time-stamps referring to a "snapshot" of the dashboard data at the time of the story creation. Nevertheless, they are still connected to that version of the data and are still live. Thus users can still interact with them to perform brushing and linking actions (**R4**). By default we have deactivated more advanced interactions such as drill-down/up, but the authors can choose to reactivate them through a properties window for the visualizations.

Moreover, the creators may decide to take a new snapshot of the data for the visualizations (i.e. change the timestamp for the visualizations and thus the values of the data). This way they can reuse the story structure for a new version of the data (**R6**). This is of particular importance for evolving dynamic data. Finally, users can select any visualization in the story and replace it by another visualization from another dashboard that they can load in the system. This way they can reuse the story structure not only for the same dynamic data, but also for completely different datasets.

Thus our narrative board integrates all the material used currently in BI story creation (**R2**): BI reports (visualizations and annotations), interactive visualizations, ways to indicate story structure, highlighting capabilities, presentation of the story in sequence, and textual or audio explanations. The entire narrative board can then be shared with others (e.g. readers) that have access to the prototype.

5.7 User Feedback

We conducted two user feedback sessions to assess the usability and benefits of the system, both from the creator's and the reader's perspective.

5.7.1 Story creation by experts

We first invited back two of the interviewed experts, and showed them the system in individual sessions (lasting 90 min). One of the experts has been an analyst for almost 12 years and has trained several other analysts. The second expert has been an analyst for a year and was trained by the first expert. They were able to explore the prototype, see story examples, and create stories themselves.

5.7.1.1 General feedback and recommendations

Both experts were very enthusiastic with the prospect of having access to such a system for their work. They found the prototype easy to use, and very helpful, specially as it

is integrated in the exploration/annotation dashboard tool. As one mentioned it "saves me from recreating any charts or annotations to present to others". The prototype saves a lot of time and effort compared to their actual practices, not just in terms of creation, but also in communicating the reports to others. As one expert explained, by creating a story with the prototype, there is no need for an explanation presentation to clients, for creating a wiki page or interactive charts.

Moreover the story can be archived and used for any new arriving employee that needs to learn how to create reports (comment from our trainer analyst). This expert stressed how important such an archiving is for knowledge passing between analysts. The recently trained expert mentioned how this tool could help him further improve his reports by looking at the story structure of others.

Both participants appreciated the ability to use the annotated dashboard template, saying that "this is how we want to present our analysis story, same as we present our dashboard". Moreover they both commented that annotations attached to data points are very important in pointing out important data values to focus on "I add a lot of manual arrows to point to annotations that refer to specific data points on a chart, and their ideal value".

When it came to how to show the story to readers, both creators reaffirmed that story reading must be strictly guided by the story creator, else the goal of the story may get lost. In their current practices they enforce this by the order of explanations used in their BI reports. They commented that the prototype supports this well by adding a sequence for the entities. Participants found the playback mode (with fade out and zoom) useful to guide readers through the story and keep their focus on a single entity each time. Both participants mentioned that they liked the fact they could create a story on one page (even if scrolling is required), as opposed to the current multi-page reports, because as one mentioned "people tend to read only the first page, and explore less the following pages". Finally, they appreciated the static representation of stories, because "the reader can always understand the position and link between parts and the overall story".

5.7.1.2 Constraining and opening up the story

Our experts suggested that our system should support three types of BI narrative stories:

- **Static stories.** Stories that present snapshots of data sets at specific points in time, yet are interactive (as is the case in our prototype), to be shared as tutorials, explanations, or reports with any audience. This is the case in our system.
- **Online or Live stories.** Stories that present live data. This is needed for dynamic data that often have the same analytic scenario, regardless the values of data. For

example they may have the same description for each chart (e.g. data measures and dimensions), have the same relation between KPIs, and the same reading sequence. In this case the visualizations in our stories are no longer snapshots, but completely live and updated with data changes. We have implemented this extension.

- **Collaborative stories.** Based on our interview findings, in our current version the analyst designs and creates the story, and controls all aspects of its presentation. One expert commented on how this storytelling prototype can also be used as a means to evolve stories. Multiple analysts can integrate their own comments and knowledge in the story, encouraging peer learning, but also creating more complete and detailed stories.

One expert also mentioned she could envision using the system to iterate the definition of the story directly with their clients (a process conducted now by email or phone-calls). Using our prototype they envisioned cases where their clients add further explanations on specific entities and their relations between entities (to explain patterns), possible summaries about the decisions they took based on the report or what was the information missing, etc. This goes beyond collaborative analysis and storytelling, as it empowers the reader and becomes a way for her to communicate what they want from the story.

5.7.2 Story reading by novices

We then ran a user feedback session to evaluate the prototype from a reader's perspective, and thus close the story communication cycle. We conducted private sessions that lasted between 40 and 50 minutes, with 5 BI novices. All were IT professionals, knew what a dashboard is but had never worked with one. Two had heard of BI reports but never used one.

Participants were asked to (1) Read a BI report created by one of our experts, (2) Read a BI story (created by an expert report creator in the previous session, and (3) Explore our prototype.

The story presented the progress of a system development project, from different perspectives: a General (how are the projects evolving in terms of finished code in a sprints time line, how many code components are added to a waiting list, and how many are done each week with a categorization of components as critical, major or minor), and a Detailed perspective (the progress of each development team in coding and testing each code component). In the story the development is not progressing according to plan because many bugs fixed are not critical, whereas new bug reports coming in add critical bugs to the

#	Question
Q1	Why is the trend line of "user story" in the pink line chart stable in the first spring?
Q2	Why is there no progress in the project "user story" in the pink chart?
Q3	What is the difference between the two yellow line charts?
Q4	What is the number of critical items done last week?
Q5	Why are the critical non developed items increasing?
Q6	Is the "User Story" project meeting the expectations? Why?

TABLE 5.1: Users were asked this group of comprehension question related to the content of the story and charts, after reading first the report then the story.

waiting list. On chart in the bottom right highlighted the problem and the rest provide details.

5.7.2.1 Reading the report

When given the report, all BI novice readers read the first page dashboard from left to right and up to down (which in fact is not the "author" suggested order). They characterized the report as a presentation of facts, that they understood the goal (progress of a project) and what the charts displayed (e.g. bugs in waiting list) with the aid of titles and legends on each chart. But they all struggled to finding what was the problem illustrated in the report. Only one participant figured out that the project development is not progressing over time which is not normal, nevertheless she could not understand why.

This clearly demonstrated our experts' comment that reports cannot be read without the supporting material.

5.7.2.2 Reading the story

Participants found that reading a story was easier "the story showed the facts in an understandable manner". The 5 readers found the system easy to use, understood the story, and were able to answer correctly several comprehension question related to the content of the story and charts Table 5.1.

We report here our main observations and readers' comments:

- Persistence: All users tended to remember even detailed aspects of the story. They only went back to the system when answering detailed quantitative questions (e.g. to retrieve numbers from charts), while all general comprehension questions were answered without looking back to the story visualizations and comments. They also tended to answer using the same terms used by the expert analyst in their comments. When asked similar questions after reading just the report, they went back to the report each time to search for answers.
- Confidence: Readers were not confident about their understanding of the report since they had to draw their own conclusions. They expressed worry that they may have misunderstand or not noticed some important points. While all readers felt confident when reading the story, specially since their interpretations are confirmed by another person (analyst comments).
- Guidance: all readers appreciated the guidance in reading the story both static and playback. Participants commented on the importance of both modes: "the static mode permits to understand the whole story", while playback "is easier to follow" the story sequence. Particularly for the playback, they commented on how it helped them focus and prevented them from spending time on unimportant data. Four participants preferred the fade mode over the default highlight mode, and 3 preferred to add the max mode when focusing on an entity.
- Understanding: all users understand the entire story using our tool, and found it easy to read and interpret. They also commented how it was easier to find answers both to qualitative and quantitative questions. The story structure showed clearly what is the problem, how to analyze it and how to find the cause. On the other hand they described the report as ambiguous, as they couldn't find the relation between charts or KPIs. When reading the simple report, each reader can choose his own way in structuring and analyzing the data and spend time in exploring in detail each chart to find relations. In the end most could just deduce that all charts represent the same data but from different perspectives.
- Transmitting knowledge: All readers found the annotations very helpful in explaining the relationships between charts and KPIs, and in teaching them the analysis logic they should follow. Four mentioned that the system can "aid in transmitting different knowledge in the company between different users", and 2 participants said they would like to be able to use the system to communicate with their team their own data (even if they are not analysts).
- Engagement: Three readers out of 5 got very engaged with the story and began searching in other detailed charts to find how this problem can be solved. As one mentioned "stories are more encouraging than static reports".

The comparison of BI reports without supplementary material over our story prototype is clearly unfair and our goal is not to prove its superiority in terms of communication value. It mainly serves to illustrate how our stand-alone prototype could be an effective means of communicating BI stories without additional material. The comparison of our tool against the collective BI story material used today in terms of comprehension is future work, but our experts already identified the gains it provides in terms of saving time on report creation, and allowing the sharing of interactive stories.

5.8 Summary

BI analysts often need to organize visually and communicate their analysis story and findings to others. Nevertheless their tools do not allow them to easily transitioning from visual analysis to storytelling. They are often required to use multiple tools, replicate work, and train their audience to understand their analysis. Based on interviews and a paper prototyping sessions with expert BI analysts, we derived requirements on how to extend BI visual analysis dashboards to support storytelling, including the need for reuse of story structure, and the constraints to impose interactive visualizations.

Using these requirements we implemented a prototype tool that is integrated in BI analysis dashboard and allows fluid transition from analysis to storytelling. We then evaluated in two sessions. First we gain feedback from story creators on the use and benefits of the tool. Our findings indicate that the requirements and resulting prototype meet the authors needs and prompted further suggestions for collaborative storytelling and story archiving. We then presented a story created by an expert to novice readers, as a first step of testing the understandability of stories seen in our system, and its potential value as a stand-alone means of communicating BI analysis stories. Our readers were all able to understand the created stories without any training and to answer complex comprehension questions.

Chapter 6

Conclusions and Future Work

The incredibly quick rate of growth of data poses challenges in analyzing large data sets. BI as a field aims to better analyze large data sets related to business and one of the more important tools available to BI analysts are visualization dashboards. Dashboards are still only created by BI experts due to a complex and long creation procedure, that often cannot be conducted by end-users, involving a large amount of communication between business analysts and end users. Moreover, they still fall short of supporting detailed analysis record keeping, and extracting results to communicate to others is a tedious process. To address these challenges, this thesis focuses on enhancing user interaction with BI dashboards, in three aspects of the life of a dashboard (1) creation: easy creation & customization of dashboards for a broader user audience, (2) analysis: adding novel annotation support and (3) communication: the use of visual storytelling in communicating dashboard analysis results.

6.1 Review of Thesis Results and Contributions

The central problem addressed by this thesis is how to enhance user interaction with BI dashboards in order to access a broader user audience and enable them to analyze large data sets and communicate their findings with others.

Creation: To that end we have presented Exploration Views, a prototype system that allows users to easily create, customize and interact with Business Intelligence visualization dashboards. EV was build following previous guidelines for novice visualization users.

Through a user-study with novice and expert visualization users we verified that EV enables novices to create dashboards for answering complex questions. We found some differences in how experts and novice users interact with the interface, such as different

functionality needs for each group of users, for example visualization novice users need more interface exploration and navigation capabilities (e.g. base views, undo) and more explanations and tutorials, while experts require advanced data exploration and customization functionality (e.g. OLAP functions and color/label customization). Moreover each group of users has different understanding of concepts, for example when novices select a part of a visualization in the dashboard, they expect a filter action on the dashboard visualizations, while the same action for experts translates to traditional crashing and linking.

More importantly, based on our observations, we provide new guidelines that augment previous work on designing for visualization novices, in the context of interactive visualization dashboards, such as the need for extensive exploration aids (e.g. basic views and self reviling functionality).

Analysis: After an easy creation of a BI dashboard, we need to support advanced VA in order to answer complex questions. Recent years have seen an increase of use of context aware annotations to support VA tasks. Based on in-depth interviews with expert Business Intelligence analysts, we extend this previous work by defining a new set of requirements for annotations, targeting visualization dashboards and other coordinated view systems. With these requirements and a user-centered design approach, we developed a dashboard prototype that supports context aware annotations.

We discuss how such annotations support new functionality, like "annotate once, see everywhere" for visualizations (not just text [81]), multi-chart annotations, and annotations that are transparent across hierarchical data dimensions. They also support knowledge discovery through annotation recommendations, as well as the use of the common data model layer present in most BI systems (e.g. [116][117]) to perform cross-application annotation, and information foraging outside the visualization system.

While building our prototype, we identified and provided solutions to challenges in using context aware annotations, notably issues arising when the annotation's "context" changes (e.g. annotated data are deleted). Based on feedback from experts, such annotations should be archived with a snapshot of their context, as long there is an easy way to retrieve them (e.g. through recommendations). Users should also be able to define the lifetime of annotations.

Our prototype was evaluated with a different set of BI experts that were able to easily perform several tasks that they deemed important in their work such as: use annotations to discover other visualizations with the same data context and look at data in a new perspective; use annotations to link and organize multiple charts related to their analysis;

discover charts and annotations through annotation recommendation and understand the reasons behind it; and use annotation context to find information outside the system.

Communication: Finally, BI analysts often need to organize visually and communicate their analysis story and findings to others. Nevertheless their tools do not allow them to easily transition from visual analysis to storytelling. They are often required to use multiple tools, replicate work, and train their audience to understand their analysis. Based on interviews and a paper prototyping sessions with expert BI analysts, we derived requirements on how to extend BI visual analysis dashboards to support storytelling, including the need for reuse of story structure, and the constraints to impose on interactive stories.

Using these requirements we implemented a prototype tool that is integrated in a BI analysis dashboard and allows fluid transition from analysis to storytelling. We then evaluated it in two sessions. First we gained feedback from story creators on the use and benefits of the tool. Our findings indicate that the requirements and resulting prototype meet the authors needs and prompted further suggestions for collaborative storytelling and story archiving. Our system was also identified by experts as having great potential for training other analysts. We then presented a story created by an expert to novice readers, as a first step of testing the understandability of stories seen in our system, and its potential value as a stand-alone means of communicating BI analysis stories. Our readers were all able to understand the created stories without any training and to answer complex comprehension questions. Our preliminary results indicate that our prototype can help authors reach a broader audience, and little or no training is required to read the results of an analysis.

6.2 Limitations, Discussions and Future Work

By conducting research into enhancing user interaction with BI dashboards, this thesis opens the door to new lines of inquiry; we hope it serves as a prelude to a continuing stream of research. There is a great deal of future work that can be done to further enhance dashboard usability. Here we elaborate on some of the limitations of this thesis and corresponding opportunities for future research.

6.2.1 BI Dashboard Creation and Customization

As described in Chapter 3, Exploration Views was evaluated in a short term study. Further work is needed to gain more insights in how novices use dashboards, and it would be interesting to observe their exploration behavior in a long-term study. Such a study

could provide us with a clearer understanding of the evolving needs of a novice user that gains more and more experience. Indeed an important question to explore is if in fact systems designed for novices with limited functionality, actually help novices to evolve and transition to expert users.

Moreover, as study participants were not data experts, and they performed tasks on unfamiliar data sets, it could be interesting to see if our findings extend to data experts, and find out how would they use EV to explore and analyze their data sets. We also tested EV for business data, but it would be interesting to discover its usability in other different domains' data, such as science, mathematics and weather forecast, etc.

Finally, our work focused more on the user interface design and use of a visualization environment for novices, and less on the insights gained from the data to help improve decision-making. Clearly this is an avenue of future work. For example it would be interesting to observe the evolution of exploration strategies (e.g. use of base views and undo), how the different visualization combinations effect understanding and decision making, and whether mechanisms such as the clipboard and history are enough for archiving analysis findings and resuming older analysis sessions.

6.2.2 Context-Aware Annotations

Our work in Chapter 4 highlights the importance of data-centered annotations, and how analysts need to attach annotations to data points (rather than images), so that annotations remain attached to different data visual presentations used by analysts to see data from different perspective, in order to gain better insights. In our work, we have presented a new set of requirements for context-aware annotations for the BI domain, and it would be valuable to test if those requirements are applicable to other visualization domains. For example level transparent annotations, could be useful in layered visualizations (e.g. systems supporting semantic zooming). Annotation validity and lifetime would be of interest in dynamic data visualization systems (e.g. stock market visual analysis tools). The set of requirements and design decisions in this work are highly influenced by the domain of focus (BI). Thus some may not apply to other domains (e.g. the need for formal annotation presentation). Nevertheless, some of the more general requirements on annotation transparency across charts and dimensions, as well as ways of addressing changing data context can be applied to other visualization domains.

For example: annotation transparency and re-use saves user effort in any visualization system with many possible data views (e.g. multidimensional visualization systems); granularity transparent annotations that highlight important data across data levels can benefit any system with a hierarchical or layered data representation (e.g. systems supporting

semantic zooming); whereas issues related to annotation validity and lifetime affect any dynamic data visualization system (e.g. stock market visual analysis tools). Nevertheless, further work is needed in this regard to clarify the importance of the identified needs and to validate the proposed designs with domain expert users of these visualization systems.

More generally, our work highlights a deeper need for data-aware and transparent note taking: for example in today's spreadsheet applications users can comment specific cells, but when cells are turned into charts the comments are not visible on them. Our system takes advantage of a common data model to annotate specific data contexts and have them available in all data representations and views (even across applications). But the idea of annotating directly data or queries for annotation re-use and recommendation, can be modified to work with any data archiving system that has a querying mechanism for the annotation to be attached to, and a notion of "distance" of the archived objects for recommendations. Thus it can be applicable to interactive applications ranging from spreadsheets to medical imaging archives or large document collections.

Although initial reactions to our system are very positive, a long-term study is needed to determine how context-aware annotations can impact analysis behavior in the long run. Our prototype design is now integrated in different analysis systems that would permit a deeper study of its usability, as well as how central annotations will be used in different analysis systems by using the common data layer.

Another topic of interest would be collaborative annotations. The capability to extend an annotation to an information exchange medium, where each user can add new comments to an existing annotation, could effect decision making and raises questions regarding the quality of new contributions and trust of its author. Further more, tracking new annotations on a specific data context could be frustrating, as a big amount of annotations could cause clutter, and furthermore if an opinion conflict exists between annotations on the same context it can become confusing for other analysts.

6.2.3 Storytelling

Although our work focuses on the BI domain, the broader need for storytelling is true for all visual analysis domains where analysts need to communicate their findings. Nevertheless such support exists in only few visual analysis systems. We hope our system motivates more work on visual storytelling, we feel that some of our storytelling requirements, for example the need to both playback the story as an animation but also have the option for a static overview, can be directly applied to other domains. Other requirements, such as the need to often reuse the same story structure and constraint user interaction, maybe domain specific and need to be re-examined for other domains.

Our system was identified by experts as having great potential for training other analysts. Although it was designed as a stand alone tool, this does not mean that it aims to replace the expert trainer analysts. We think that in training scenarios, as mentioned by one expert, it will act to aid novice analysts understand how to read the analysis results so they can then focus their questions to their trainer on more complex aspects, such as analysis methodology, the analysis goals and content details. It can help authors reach a broader audience if little or no training is required to read the results of an analysis.

BI analysis is highly driven by the analysts' client and BI stories evolve based on client needs and feedback, a characteristic not present in most other analysis domains. Nevertheless the details of this process (done through emails, and phone calls) are often lost. Experts suggested using our prototype as a collaborative storytelling system shared between analysts and clients, that can help reduce the turnaround time for the final story, and that captures the evolution of the story structure and focus. Thus the storytelling tool can be act as means for the audience to influence the story outcomes themselves. We currently exploring this possibility.

Moreover, it was suggested by our experts that our prototype can help to improve BI story quality and creators expertise, through communicating with readers and learning from their comments and feedback directly on the story (i.e. collaborative annotation). Some experts even expressed the desire for a collaborative storytelling tool, where the story (not just the annotations) is open to being edited by both readers and authors.

In both cases of collaborative story editing, the prototype can move from simple communication to becoming a collaboration medium, where the authoring of a story opens-up and evolves with the contributions of many users, and acts as an archive of knowledge and different points of view. This opening up of the story is somewhat contradictory to the need to control the story, expressed by our experts in their requirements. In such a system our experts explained that readers need to be able to clearly differentiate between the contributions of individual users (e.g. filtering the author of added explanations and annotations). To support this process we need to also provide a powerful history and comparison mechanism, that displays different story versions and shows to the reader clearly what changed through time or from a storyteller to another, while providing advanced temporal navigation and filtering functionalities. We are currently exploring this extension.

Our prototype is currently being finalized to be given to BI analysts in order to study its long term use and how its impact on analysis collaboration and communication practices.

Bibliography

- [1] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)*, 2010. URL <http://vis.stanford.edu/papers/narrative>.
- [2] <http://www.idc.com/>.
- [3] Lopez J. Raskino, M. Early findings from the 2010 gartner ceo and business executive survey, December 2009.
- [4] M. Sorenson. Dictionary definition of dashboards, searchcio.com, February 2002.
- [5] J. Cutroni. *Google Analytics*. O'Reilly Media, Inc, 2007.
- [6] A. Chiang. Creating dashboards: The players and collaboration you need for a successful project. *Business Intelligence Journal*.
- [7] P. Foody. User-centered business intelligencet. *Business Intelligence Journal*.
- [8] B. Evelson. Topic overview: Business intelligence. Technical report, Forrester Research, June 2006.
- [9] D.J. Power. A brief history of decision support systems. DSSResources.COM, World Wide Web, <http://DSSResources.COM/history/dsshistory.html>, March 2007. version 4.0.
- [10] M. Friendly. Milestones in the history of thematic cartography, statistical graphics, and data visualization. In *13TH INTERNATIONAL CONFERENCE ON DATABASE AND EXPERT SYSTEMS APPLICATIONS (DEXA 2002)*, AIX EN PROVENCE, pages 59–66. Press, 1995.
- [11] V. Friedman. Data visualization and infographics. Technical report, in: Graphics, Monday Inspiration, January 2008.
- [12] E. R. Tufte. *The visual display of quantitative information*. Graphics press LLC, 2006.

- [13] Mackinlay J.D. Shneiderman B. Card, S.K. *Readings in information visualization: Using vision to think*, volume Learning with animation. Morgan Kaufmann Publishers, San Francisco, 1999.
- [14] James J. T. and K. A. Cook. *Illuminating the Path: The R&D Agenda for Visual Analytics*. National Visualization and Analytics Center., 2005.
- [15] J.A. Hartigan. Printer graphics for clustering. *Journal of Statistical Computation and Simulation*, 4(3):187–213, 1975. doi: 10.1080/00949657508810123. URL <http://www.tandfonline.com/doi/abs/10.1080/00949657508810123>.
- [16] William Ribarsky, Eric Ayers, John Eble, and Sougata Mukherjea. Glyphmaker: Creating customized visualizations of complex data. *Computer*, 27:57–64, 1994. ISSN 0018-9162. doi: <http://doi.ieeecomputersociety.org/10.1109/2.299412>.
- [17] Christopher D. Shaw, James A. Hall, Christine Blahut, David S. Ebert, and D. Aaron Roberts. Using shape to visualize multivariate data. In *Proceedings of the 1999 workshop on new paradigms in information visualization and manipulation in conjunction with the eighth ACM international conference on Information and knowledge management*, NPIVM '99, pages 17–20, New York, NY, USA, 1999. ACM. ISBN 1-58113-254-9. doi: 10.1145/331770.331777. URL <http://doi.acm.org/10.1145/331770.331777>.
- [18] Matthew O. Ward. Xmdvtool: integrating multiple methods for visualizing multivariate data. In *Proceedings of the conference on Visualization '94*, VIS '94, pages 326–333, Los Alamitos, CA, USA, 1994. IEEE Computer Society Press. ISBN 0-7803-2521-4. URL <http://dl.acm.org/citation.cfm?id=951087.951146>.
- [19] M. d'Azavedo. *Coordonnées parallèles et axiales: méthode de transformation géométrique et procédé nouveau de calcul graphique déduits de la considération des coordonnées parallèles*. Gauthier-Villars, 1885.
- [20] A. Inselberg. *The plane with parallel coordinates*. Visual Computer, 1985.
- [21] E. J. Wegman. *Hyperdimensional data analysis using parallel coordinates*, 1990.
- [22] I. Borg and P Groenen. *Modern Multidimensional Scaling: Theory And Applications*. Springer Series in Statistics, 2005.
- [23] E. F. Codd. A relational model of data for large shared data banks. *Commun. ACM*, 26(1):64–69, January 1983. ISSN 0001-0782. doi: 10.1145/357980.358007. URL <http://doi.acm.org/10.1145/357980.358007>.
- [24] D. Bricklin. Visicalc, <http://www.bricklin.com/visicalc.htm>, 2006.

- [25] Jonathan C. Roberts. State of the art: Coordinated & multiple views in exploratory visualization. In *Proceedings of the Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization*, pages 61–71, Washington, DC, USA, 2007. IEEE Computer Society. ISBN 0-7695-2903-8. doi: 10.1109/CMV.2007.20. URL <http://dl.acm.org/citation.cfm?id=1270380.1270575>.
- [26] Ieee visual analytics science and technology (vast) symposium, <http://conferences.computer.org/vast/vast2007/>.
- [27] Stephen F. *Dashboard Confusion*. Perceptual Edge, 2004.
- [28] J. Cutroni. Google analytics. *O’Reilly Media, Inc*, 2007.
- [29] Stephen F. *Information Dashboard Design: The Effective Visual Communication of Data*. O’Reilly Media, 2006.
- [30] Stephen F. *Dashboard Design for Real-Time Situation Awareness*. Inova Solutions, 2007.
- [31] N. RASMUSSEN, C. Y. CHEN, and M. BANSAL. *Business Dashboards, A Visual Catalogue for Design and Deployment*. John Wiley & Sons, Inc, 2009.
- [32] T. T Hewett, R. Baecker, S. Card, T. Carey, J. Gasen, M. Mantei, G. Perlman, G. Strong, and W. Verplank. *ACM SIGCHI Curricula for Human-Computer Interaction*. ACM SIGCHI, 1992.
- [33] John M. Carroll. *Human Computer Interaction (HCI)*. The Interaction-Design.org Foundation, Aarhus, Denmark, 2009. URL http://www.interaction-design.org/encyclopedia/human_computer_interaction_hci.html.
- [34] Nielsen J. Usability 101: Introduction to usability. *Jakob Nielsen’s Alertbox*.
- [35] Spotfire, <http://spotfire.tibco.com/>.
- [36] Tableau, <http://www.tableausoftware.com/products/desktop>, .
- [37] F. B. Viegas, M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. Manyeyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13:1121–1128, November 2007. ISSN 1077-2626. doi: <http://dx.doi.org/10.1109/TVCG.2007.70577>. URL <http://dx.doi.org/10.1109/TVCG.2007.70577>.
- [38] Tory M. Storey M.-A. Grammel, L. How information visualization novices construct visualizations. *IEEE InfoVis - Transactions on Visualization and Computer Graphics*, 17, 2010. ISSN 943â&AŒ952.

- [39] van Ham F. Carpendale-S. Weaver C. Isenberg P. Heer, J. Creation and collaboration: Engaging new audiences for information visualization. *In: Kerren, A., Stasko, J.T., Fekete, J.-D., North, C. (eds.) Information Visualization .LNCS*, 4950:92–133, 2008.
- [40] Nelson Cowan. *Attention and Memory: An Integrated Framework*, volume 26 of *Oxford Psychology Series*. Oxford University Press, 1998.
- [41] J. Schwartz. *New Digital Dashboards Help Drive Decisions*. BtoB, 2003.
- [42] Dundas dashboard, <http://www.dundas.com/microsite/dashboard>.
- [43] Oracle bi 10g dashboard, <http://www.oracle.com/technology/products/bi/enterpriseedition-platform-components.html>.
- [44] SAP business objects xcelsius, <http://www.sap.com/solutions/sapbusinessobjects/large/businessintelligence/dashboard-visualization/xcelsiusenterprise/index.epx>, .
- [45] Choosel: <http://code.google.com/p/choosel/>, 2009.
- [46] OLAP and OLAP server definitions. The OLAP Council, 1995.
- [47] <http://thejit.org/>.
- [48] <http://ivtk.sourceforge.net>.
- [49] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. D3: Data-driven documents. *IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)*, 2011. URL <http://vis.stanford.edu/papers/d3>.
- [50] C. Weaver. Building highly-coordinated visualizations in improvise. *IEEE Symposium on Information Visualization, InfoVis*, pages 159–166, 2004.
- [51] C. North and B. Shneiderman. Snap-together visualization: A user interface for coordinating visualizations via relational schemata. *Proc. AVI*, 2000.
- [52] J. Heer, F. B. Viégas, and M. Wattenberg. Voyagers and voyeurs: Supporting asynchronous collaborative visualization. *Commun. ACM*, 52:87–97, January 2009. ISSN 0001-0782. doi: <http://doi.acm.org/10.1145/1435417.1435439>. URL <http://doi.acm.org/10.1145/1435417.1435439>.
- [53] D. Cubrarnic. Polstar: Assisted navigation for exploring multi-dimensional information spaces. *Huma-Computer Information Retrieval, HCIR*, 2008.
- [54] M.J. Adler and C.L. Van Doren. *How To Read A Book*. Simon and Schuster, 1972.

- [55] C. C. Marshall. Annotation: from paper books to the digital library. Technical report, In Proceedings of the Second ACM international Conference on Digital Libraries, Philadelphia, Pennsylvania, United States, 1997.
- [56] Guillaume Cabanac, Max Chevalier, Claude Chrisment, and Christine Julien. Collective annotation: perspectives for information retrieval improvement. In *Large Scale Semantic Access to Content (Text, Image, Video, and Sound)*, RIAO '07, pages 529–548, Paris, France, France, 2007. LE CENTRE DE HAUTES ETUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE. URL <http://dl.acm.org/citation.cfm?id=1931390.1931440>.
- [57] José Kahan and Marja-Ritta Koivunen. Annotea: an open rdf infrastructure for shared web annotations. In *Proceedings of the 10th international conference on World Wide Web*, WWW '01, pages 623–632, New York, NY, USA, 2001. ACM. ISBN 1-58113-348-0. doi: 10.1145/371920.372166. URL <http://doi.acm.org/10.1145/371920.372166>.
- [58] David Bargeron and Tomer Moscovich. Reflowing digital ink annotations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '03, pages 385–393, New York, NY, USA, 2003. ACM. ISBN 1-58113-630-7. doi: 10.1145/642611.642678. URL <http://doi.acm.org/10.1145/642611.642678>.
- [59] C. E. Chronaki, X. Zabulis, and S. C. Orphanoudakis. I2net medical image annotation service. *Informatics for Health and Social Care*, 22(4):337–347, 1997. doi: 10.3109/14639239709010905. URL <http://informahealthcare.com/doi/abs/10.3109/14639239709010905>.
- [60] P. Duygulu, Kobus Barnard, J. F. G. de Freitas, and David A. Forsyth. Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary. In *Proceedings of the 7th European Conference on Computer Vision-Part IV*, ECCV '02, pages 97–112, London, UK, UK, 2002. Springer-Verlag. ISBN 3-540-43748-7. URL <http://dl.acm.org/citation.cfm?id=645318.649254>.
- [61] S. L. Feng, R. Manmatha, and V. Lavrenko. Multiple bernoulli relevance models for image and video annotation. In *Proceedings of the 2004 IEEE computer society conference on Computer vision and pattern recognition*, CVPR'04, pages 1002–1009, Washington, DC, USA, 2004. IEEE Computer Society. URL <http://dl.acm.org/citation.cfm?id=1896300.1896446>.
- [62] J. Jeon, V. Lavrenko, and R. Manmatha. Automatic image annotation and retrieval using cross-media relevance models. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, SIGIR

- '03, pages 119–126, New York, NY, USA, 2003. ACM. ISBN 1-58113-646-3. doi: 10.1145/860435.860459. URL <http://doi.acm.org/10.1145/860435.860459>.
- [63] Ilaria Bartolini and Paolo Ciaccia. Multi-dimensional keyword-based image annotation and search. In *Proceedings of the 2nd International Workshop on Keyword Search on Structured Data*, KEYS '10, pages 5:1–5:6, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0187-9. doi: 10.1145/1868366.1868371. URL <http://doi.acm.org/10.1145/1868366.1868371>.
- [64] Jia Li and James Z. Wang. Real-time computerized annotation of pictures. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(6):985–1002, June 2008. ISSN 0162-8828. doi: 10.1109/TPAMI.2007.70847. URL <http://dx.doi.org/10.1109/TPAMI.2007.70847>.
- [65] Y. B. Shrinivasan and J.J. van Wijk. Supporting the analytical reasoning process in information visualization. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1237–1246, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-011-1. doi: <http://doi.acm.org/10.1145/1357054.1357247>. URL <http://doi.acm.org/10.1145/1357054.1357247>.
- [66] Jeffrey Heer, Jock Mackinlay, Chris Stolte, and Maneesh Agrawala. Graphical histories for visualization: Supporting analysis, communication, and evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 14:1189–1196, November 2008. ISSN 1077-2626. doi: 10.1109/TVCG.2008.137. URL <http://dl.acm.org/citation.cfm?id=1477066.1477414>.
- [67] Nazanin Kadivar, Victor Y. Chen, Dustin Dunsmuir, Eric Lee, Cheryl Z. Qian, John Dill, Christopher D. Shaw, and Robert F. Woodbury. Capturing and supporting the analysis process. In *IEEE VAST*, pages 131–138, 2009.
- [68] Mark Derthick and Steven F. Roth. Data exploration across temporal contexts. In *Proceedings of the 5th international conference on Intelligent user interfaces*, IUI '00, pages 60–67, New York, NY, USA, 2000. ACM. ISBN 1-58113-134-8. doi: <http://doi.acm.org/10.1145/325737.325782>. URL <http://doi.acm.org/10.1145/325737.325782>.
- [69] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, pages 336–, Washington, DC, USA, 1996. IEEE Computer Society. ISBN 0-8186-7508-X. URL <http://dl.acm.org/citation.cfm?id=832277.834354>.
- [70] Yedendra Babu Shrinivasan, David Gotz, and Jie Lu. Connecting the dots in visual analysis. In *IEEE VAST*, pages 123–130, 2009.

- [71] Thomas Kapler and William Wright. Geo time information visualization. *Information Visualization*, 4:136–146, July 2005. ISSN 1473-8716. doi: 10.1145/1103520.1103526. URL <http://dl.acm.org/citation.cfm?id=1103520.1103526>.
- [72] Catherine C. Marshall. Work practice study: Analysts and notetaking. Technical report, Xerox Palo Alto Research Center, 1990.
- [73] Peter Pirolli and Stuart Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of International Conference on Intelligence Analysis*, 2005.
- [74] W. Willett, J. Heer, J. Hellerstein, and M. Agrawala. Commentspace: Structured support for collaborative visual analysis. In *ACM Human Factors in Computing Systems (CHI)*, 2011. URL <http://vis.stanford.edu/papers/commentspace>.
- [75] Les Nelson, Christoph Held, Peter Pirolli, Lichan Hong, Diane Schiano, and Ed H. Chi. With a little help from my friends: examining the impact of social annotations in sensemaking tasks. In *ACM CHI*, pages 1795–1798, 2009.
- [76] William Wright, David Schroh, Pascale Proulx, Alex Skaburskis, and Brian Cort. The sandbox for analysis: concepts and methods. In *Proceedings of the SIGCHI conference on Human Factors in computing systems, CHI '06*, pages 801–810, New York, NY, USA, 2006. ACM. ISBN 1-59593-372-7. doi: <http://doi.acm.org/10.1145/1124772.1124890>. URL <http://doi.acm.org/10.1145/1124772.1124890>.
- [77] Visual analytics inc., <http://www.visualanalytics.com>.
- [78] Gary Klein, Jennifer K. Phillips, Erica L. Rall, and Deborah A. Peluso. A data-frame theory of sensemaking. In *Expertise Out of Context Proceedings of the Sixth International Conference on Naturalistic Decision Making*, pages 113–154, 2007.
- [79] Simon Attfield, Sukhvinder Hara, and B. L. William Wong. Sense-making in visual analytics: processes and challenges. In *EuroVAST 2010: The 1st European Symposium on Visual Analytics Science and Technology*, 2010.
- [80] Yang Chen, Scott Barlowe, and Jing Yang. Click2annotate: Automated insight externalization with rich semantics. In *IEEE VAST*, pages 155–162, 2010.
- [81] Wen-Huang Cheng and David Gotz. Context-based page unit recommendation for web-based sensemaking tasks. In *Proceedings of the 14th international conference on Intelligent user interfaces, IUI '09*, pages 107–116, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-168-2. doi: <http://doi.acm.org/10.1145/1502650.1502668>. URL <http://doi.acm.org/10.1145/1502650.1502668>.

- [82] E. A. Bier, S. K. Card, and J. W. Bodnar. Entitybased collaboration tools for intelligence analysis. *IEEE Symposium on Visual Analytics Science and Technology*, pages 99–106, 2008.
- [83] C. Baber, D. Andrews, T. Duffy, and R. McMaster. Sensemaking as narrative: Visualization for collaboration. *VAW2011, University London College*, 2011.
- [84] D. Sole and D. G. Wilson. Storytelling in organizations: The power and traps of using stories to share knowledge in organizations. *Harvard Graduate School of Education*, 2002.
- [85] D. Snowden. The art and science of story or 'are you sitting uncomfortably?'. *Business Information Review*, pages 215–226, 2000.
- [86] I. Mitroff and R. H Kilmann. Stories managers tell: A new tool for organizational problem-solving. *Management Review*, 1975.
- [87] W. McWhinney and J. Battista. How remythologizing can revitalize organizations. *Organizational Dynamics*, pages 46–58, 1988.
- [88] M. E. Boyce. Collective centring and collective sense-making in the stories and storytelling of one organization. *Organization Studies*, pages 107–137, 1995.
- [89] M. H. Brown. That reminds me of a story: Speech action on organizational socialization. *University of Texas at Austin*, 1982.
- [90] M. R. Louis. Surprise and sense making: What newcomers experience in entering unfamiliar organizational setting. *Administrative Science Quarterly*, 1980.
- [91] M. R. Louis. Organizations as culture-bearing milieu. *Organizational symbolism: Monographs in organizational and industrial relations*, pages 39–54, 1983.
- [92] D. M Boje. The storytelling organization: A study of story performance in an office-supply firm. *Administrative Science Quarterly*, pages 106–126, 1991.
- [93] D. M Boje. Stories of the storytelling organization: A postmodern analysis of disney as "tamara-land". *Academy of Management Journal*, pages 997–1035, 1995.
- [94] W. McWhinney. A health confluence. *Unpublished paper*, 1995.
- [95] J. E. Helmer. Between horses: An ethnographic study of communication and organizational culture at a harness track. *University of Illinois at Urbana-Champaign*, 1989.
- [96] E. A. Bier, S. K. Card, and J. W. Bodnar. Entity-based collaboration tools for intelligence analysis. *IEEE Symposium on Visual Analytics Science and Technology*, 2011.

- [97] P. Pirolli and D.M. Russell. Introduction to this special issue on sensemaking. *Human-Computer Interaction*, 2011.
- [98] P. Farhi. Cnn hits the wall for the election, <http://bit.ly/cnn-wall>, 2008.
- [99] Gapminder, <http://www.gapminder.org>, 2010.
- [100] S. Few. Show me the numbers: Designing tables and graphs to enlighten. *Analytics Press*, 2004.
- [101] W. Wojtkowski and W. G. Wojtkowski. Storytelling: its role in information visualization. *European Systems Science Congress*, 2002.
- [102] N. Gershon and W. Page. What storytelling can do for information visualization. *Commun. ACM*, 44(8):31–37, August 2001.
- [103] R. Eccles, T. Kapler, R. Harper, and W. Wright. Stories in geotime. *IEEE VAST*, pages 19–26, 2007.
- [104] Tableau public, <http://tableausoftware.com/public>, .
- [105] Stephen Few. *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. Analytics Press, 2004.
- [106] J. Nielsen. *Usability Engineering*. Academic Press, 1994.
- [107] G. Zelazny. The say it with charts complete toolkit, December 2006.
- [108] Shneiderman B. Williamson, C. The dynamic homefinder: Evaluating dynamic queries in a real-estate information exploration system. *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 339–346, 1992.
- [109] M. Hearst. User interfaces and visualization. In: Baeza-Yates, R., Ribeiro-Neto, B. (eds.) *Modern Inf. Retrieval*, ch.10, 1999.
- [110] M. Hearst. Uis for faceted navigation, recent advances and remaining open problems. In: *HCIR*, 2008.
- [111] D. Woods. Visual momentum: a concept to improve the cognitive coupling of person and computer. *Int. J. Man-Mach. Stud.*, 1984.
- [112] Popovici N. Boshmaf Y.-Plattner H. Zeier A. Schaffner J. Willhalm, T. Simd-scan: Ultra fast in-memory table scan using on-chip vector processing units. *VLDB*, 2009.
- [113] N. H. Riche, B. Lee, and C. Plaisant. Understanding interactive legends: a comparative study with standard widgets. *Proc. Eurovis*, 29(3):1193–1202, 2010.

- [114] Edwin Hutchins. *Cognition in the Wild*. MIT Press, Cambridge, MA, USA, 1994.
- [115] Jeffrey Heer and Maneesh Agrawala. Design considerations for collaborative visual analytics. In *Proceedings of the 2007 IEEE Symposium on Visual Analytics Science and Technology*, pages 171–178, Washington, DC, USA, 2007. IEEE Computer Society. ISBN 978-1-4244-1659-2. doi: 10.1109/VAST.2007.4389011. URL <http://dl.acm.org/citation.cfm?id=1549822.1549974>.
- [116] Microsoft UDM. [http://technet.microsoft.com/enus/library/ms345143\(sql.90\).aspx](http://technet.microsoft.com/enus/library/ms345143(sql.90).aspx).
- [117] SAP businessobjects semantic layer: <http://www.sap.com/solutions/sapbusinessobjects/large/businessintelligence/dashboard-visualization/xcelsiusenterprise/index.ep>, .
- [118] Matteo Golfarelli, Stefano Rizzi, and Iuris Cella. Beyond data warehousing: what’s next in business intelligence? In *Proceedings of the 7th ACM international workshop on Data warehousing and OLAP, DOLAP '04*, pages 1–6, New York, NY, USA, 2004. ACM. ISBN 1-58113-977-2. doi: <http://doi.acm.org/10.1145/1031763.1031765>. URL <http://doi.acm.org/10.1145/1031763.1031765>.
- [119] Ivan Denisovich. Software support for annotation of visualized data using hand-drawn marks. In *Proceedings of the Ninth International Conference on Information Visualisation*, pages 807–813, Washington, DC, USA, 2005. IEEE Computer Society. ISBN 0-7695-2397-8. doi: 10.1109/IV.2005.118. URL <http://dl.acm.org/citation.cfm?id=1084015.1084480>.
- [120] Yang Chen, Jing Yang, and William Ribarsky. Toward effective insight management in visual analytics systems. In *PacificVis*, pages 49–56, 2009.
- [121] Narges Mahyar, Ali Sarvghad, and Melanie Tory. A closer look at note taking in the co-located collaborative visual analytics process. In *IEEE VAST*, pages 171–178, 2010.
- [122] Yang Chen, Jing Yang, and William Ribarsky. Toward effective insight management in visual analytics systems. In *Proceedings of the 2009 IEEE Pacific Visualization Symposium, PACIFICVIS '09*, pages 49–56, Washington, DC, USA, 2009. IEEE Computer Society. ISBN 978-1-4244-4404-5. doi: <http://dx.doi.org/10.1109/PACIFICVIS.2009.4906837>. URL <http://dx.doi.org/10.1109/PACIFICVIS.2009.4906837>.
- [123] Craig S. Tashman and W. Keith Edwards. Liquidtext: a flexible, multitouch environment to support active reading. In *ACM CHI*, pages 3285–3294, 2011.
- [124] M. Johnson. The moral imagination. *Chicago: University of Chicago Press*, 1993.
- [125] A. MacIntyre. After virtue. *Notre Dame, IN: University of Notre Dame Press*, 1981.

-
- [126] K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H.-N. Kostis. Scientific storytelling using visualization. *Computer Graphics and Applications, IEEE*, 32(1):12–19, Jan.–Feb. 2012.
- [127] C. Forceville. Educating the eye? kress and van leeuwen’s reading images: The grammar of visual design (1996). *Language and Literature*, 8(2):163–178, 1999.
- [128] B. Tversky, J. Heiser, S. Lozano, R. MacKenzie, and J. Morrison. *Enriching Animations*, volume Learning with animation. Cambridge University Press, 2007.
- [129] M. Elias and A. Bezerianos. Annotating bi visualization dashboards: Needs & challenges. In *ACM SIGCHI Conference on Human Factors in Computing Systems, CHI’12*. Springer-Verlag, 2012.