Annotating BI Visualization Dashboards: Needs & Challenges

Micheline Elias
Ecole Centrale Paris (MAS), France
SAP Research, France
micheline.elias@ecp.fr

Anastasia Bezerianos
Univ Paris-Sud & CNRS (LRI) - INRIA
F-91405 Orsay, France
anastasia.bezerianos@lri.fr

ABSTRACT
Annotations have been identified as an important aid in analysis record-keeping and recently data discovery. In this paper 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.

Author Keywords
Annotation, visualization dashboards and Business Intelligence.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

General Terms
Design.

INTRODUCTION
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 [15] remembering all reasoning details and milestones in the analysis process becomes challenging. Thus the importance of "externalizing" memory, knowledge, and insights [24] has been identified early on. Analysts achieve this by making sense of their visualizations, annotating them with their findings [21], organizing their findings, and presenting them to others [23][35]. Annotations are thus an integral part of complex visual analyses, aiding analysts in revisiting and reusing past analysis sessions [34]. 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. [11][34]).

In the domain of Business Intelligence (BI) analysis , the most popular visualization tools are dashboards [16], collections of multiple visual components (such as charts) on a single view [19]. 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 [32]), 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.

Our contributions 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 [14]), annotation recommendations to promote knowledge discovery, and the use of the common data model layer present in BI systems (e.g. [2] [5]) 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).

RELATED WORK
To help analysts track their analysis process and findings, researchers have suggested the use of history tools (e.g. [35] [22] [25] [18]) to review, re-visit and retrieve previous analysis steps [33]. Augmenting visual representations with annotations to record analysis details and findings has also been recommended [23] and implemented in information visual-

CHI 2012, May 5–10, 2012, Austin, TX, USA.

Figure 1. Our dashboard prototype with context aware annotations. (a) The main dashboard contains icons with the number of annotations for each data-point (b). Bellow the list of all dashboard annotations in (c), is a set of recommended annotations (d) emphasizing their similarity.

ization systems (e.g.[23][34][26]), allowing analysts to re-visit their analysis and re-use previous findings [29].

Marshall [29] identified three main processing activities during analysis: searching, interpreting, and presenting. Pirolli and Card [31] 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], and sense.us [23] allow multiple users to annotate their insights on a single visualization, and search through them, encouraging peer learning and social sense making [30]. Others, like Sandbox [38] and VisualLinks [8], 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 [27].

During analysis, fragmented information pieced together often reveals connections and further opportunities for investigation [10], contributing to the searching (foraging) aspect of sense making. 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 [23] system provides doubly linked discussions, allowing users to move from visualizations to annotations and vice versa. Aruvi [35] allows users to arrange their annotations in a "mind map", a diagram linking related visualization views and their annotations. Users can access past analysis views using keyword searches and text similarity metrics. Shrinivasan et al. [34] recommend related annotations based on the context of the current analysis. Finally, Chen et al. [11] 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 [16], and are very widespread in BI (e.g. [1][3][4][6][7]). Annotations are supported in all these 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), such as annotations spanning multiple charts, or annotation transparency across data dimensions and visualizations (annotate once, see everywhere, as is done in text [14]).

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".

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 lead 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.

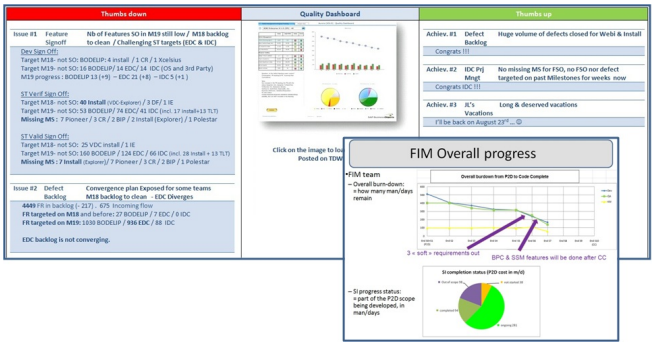

Figure 2. Two annotation examples from participants.

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 were recorded and transcribed, and the emerging themes identified using a card sorting analysis approach. Bellow we summarize our major findings:

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 [20]. 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 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.

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 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:

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 bold 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".

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 "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.

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 be not clear". Moreover they commented that a list of all dashboard annotations should be available, with visual links to their data targets.

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.

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).

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.

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 [17]. 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.

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. [34]), 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. [7]) 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.

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 [34].

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

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, similar to the model proposed by [12]:

- 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 [28][13][29]). 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

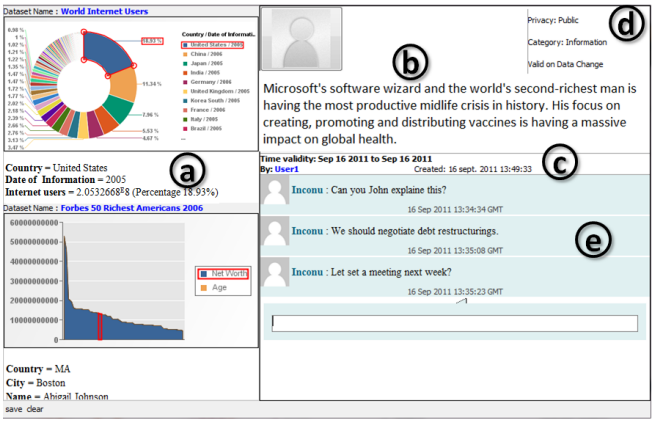

Figure 3. Detailed annotation view, including (a) all related data contexts & detailed data context in text form, (b) content, (c) metadata, (e) discussion thread, and (d) annotation properties (lifetime, privacy, etc).

data context, the specific data points that the user is annotating. Both Overview and Detailed data context are stored as data 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).

Interface

Figure 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 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 [9] (e.g. drill-down/up). On the main canvas (Figure 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 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 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 3). 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 3.a). A textual description of the context (dimensions, measures, values) is also provided.

For example the annotation seen in Figure 3 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.

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 3.e), supporting social sense-making [30].

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.

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 [14]). During the analysis phase this can help in 4 different ways:

(i) Analysts can switch visualizations (chart types) without loosing 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 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.

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 3.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 [34], [27]).

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. [36]) and visualizations representing one chart at a time (e.g. [23]).

Recommendations

We use annotation context [34] 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 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").

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.

Automatic annotation

We do not perform automated or semi-automated annotation [11]. 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.

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.

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 all our input from 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. 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. At this point experts requested the option to be sent a notifica-

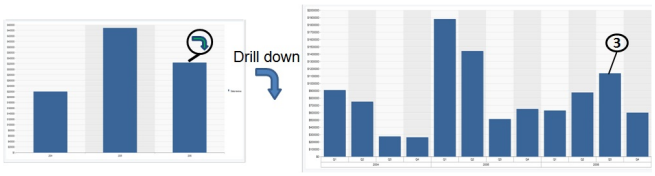

Figure 4. Annotation transparency through dimension levels.

tion inquiring if the annotation should be archived or deleted. This lifetime functionality is also implemented.

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 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.

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 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.

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.

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.

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.

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:

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 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.

RESULTS

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

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 loosing 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.

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 paper. 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 1 and Figure 3.

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 5.a) highlights targets in red, the second (Figure 5.b) draws arrows to the targets, and the third (Figure 5.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 5. Three designs linking annotations in the annotation list with their data targets. Experts preferred design (c).

DISCUSSION

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.

CONCLUSION AND FUTURE WORK

Recent years have seen an increase of use of context aware annotations to support VA tasks. We extend this work by defining a new set of requirements for annotations, targeting visualization dashboards and other coordinated view systems, based on in-depth interviews with expert Business Intelligence analysts. 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 [14]), 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. [2][5]) 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 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 information outside the system.

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.

ACKNOWLEDGMENTS

This work was supported by Business Intelligence Practice in SAP Research France.

REFERENCES

1. Dundas dashboard, www.dundas.com.
2. Microsoft UDM. technet.microsoft.com/en-us/library/ms174783(SQL.90).aspx.
3. Oracle bi 10g dashboard. www.oracle.com/technetwork/middleware/bi-enterprise-edition/overview/index.html.
4. SAP business objects xcelsius. http://www.sdn.sap.com/irj/boc/crystal-dashboard.
5. SAP businessobjects semantic layer. www.sdn.sap.com/irj/boc/index?rid=/webcontent/uuid/c0a477d8-c676-2b10-db80-8d84d1f8e317.
6. Spotfire, spotfire.tibco.com.
7. Tableau. www.tableausoftware.com/products/desktop.
8. Visual analytics inc. www.visualanalytics.com.
9. OLAP and OLAP server definitions. The OLAP Council, 1995.
10. Attfield, S., Hara, S., and Wong, B. L. W. Sense-making in visual analytics: processes and challenges. In *EuroVAST* (2010).
11. Chen, Y., Barlowe, S., and Yang, J. Click2annotate: Automated insight externalization with rich semantics. In *IEEE VAST* (2010), 155–162.
12. Chen, Y., Yang, J., and Ribarsky, W. Toward effective insight management in visual analytics systems. In *PacificVis* (2009), 49–56.
13. Chen, Y., Yang, J., and Ribarsky, W. Toward effective insight management in visual analytics systems. In *PacificVis* (2009), 49–56.
14. Cheng, W.-H., and Gotz, D. Context-based page unit recommendation for web-based sensemaking tasks. In *ACM IUI* (2009), 107–116.
15. Cowan, N. *Attention and Memory: An Integrated Framework*, vol. 26 of *Oxford Psychology Series*. Oxford University Press, 1998.
16. Cutroni, J. *Google Analytics*. O'Reilly Media, Inc, 2007.
17. Denisovich, I. Software support for annotation of visualized data using hand-drawn marks. In *International Conference on Information Visualisation* (2005), 807–813.
18. Derthick, M., and Roth, S. F. Data exploration across temporal contexts. In *ACM IUI* (2000), 60–67.
19. Few, S. *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. Analytics Press, 2004.
20. Golfarelli, M., Rizzi, S., and Cella, I. Beyond data warehousing: what's next in business intelligence? In *ACM international workshop on Data warehousing and OLAP*, DOLAP '04 (2004), 1–6.
21. Heer, J., and Agrawala, M. Design considerations for collaborative visual analytics. In *IEEE VAST* (2007), 171–178.
22. Heer, J., Mackinlay, J., Stolte, C., and Agrawala, M. Graphical histories for visualization: Supporting analysis, communication, and evaluation. *IEEE TVCG 14* (2008), 1189–1196.
23. Heer, J., Viégas, F. B., and Wattenberg, M. Voyagers and voyeurs: Supporting asynchronous collaborative visualization. *Commun. ACM 52* (2009), 87–97.
24. Hutchins, E. *Cognition in the Wild*. MIT Press, 1994.
25. Kadivar, N., Chen, V. Y., Dunsmuir, D., Lee, E., Qian, C. Z., Dill, J., Shaw, C. D., and Woodbury, R. F. Capturing and supporting the analysis process. In *IEEE VAST* (2009), 131–138.
26. Kapler, T., and Wright, W. Geo time information visualization. *Information Visualization 4* (2005), 136–146.
27. Klein, G., Phillips, J. K., Rall, E. L., and Peluso, D. A. A data-frame theory of sensemaking. In *Expertise Out of Context Proc. of the International Conference on Naturalistic Decision Making* (2007), 113–154.
28. Mahyar, N., Sarvghad, A., and Tory, M. A closer look at note taking in the co-located collaborative visual analytics process. In *IEEE VAST* (2010), 171–178.
29. Marshall, C. C. Work practice study: Analysts and notetaking. Tech. rep., Xerox Palo Alto Research Center, 1990.
30. Nelson, L., Held, C., Pirolli, P., Hong, L., Schiano, D., and Chi, E. H. With a little help from my friends: examining the impact of social annotations in sensemaking tasks. In *ACM CHI* (2009), 1795–1798.
31. Pirolli, P., and Card, S. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *International Conference on Intelligence Analysis* (2005).
32. Roberts, J. C. State of the art: Coordinated & multiple views in exploratory visualization. In *Proc. of the Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization* (2007), 61–71.
33. Shneiderman, B. The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages* (1996), 336–.
34. Shrinivasan, Y. B., Gotz, D., and Lu, J. Connecting the dots in visual analysis. In *IEEE VAST* (2009), 123–130.
35. Shrinivasan, Y. B., and van Wijk, J. J. Supporting the analytical reasoning process in information visualization. In *ACM CHI* (2008), 1237–1246.
36. Tashman, C. S., and Edwards, W. K. Liquidtext: a flexible, multitouch environment to support active reading. In *ACM CHI* (2011), 3285–3294.
37. Viegas, F. B., Wattenberg, M., van Ham, F., Kriss, J., and McKeon, M. Manyeyes: a site for visualization at internet scale. *IEEE TVCG 13* (2007), 1121–1128.
38. Wright, W., Schroh, D., Proulx, P., Skaburskis, A., and Cort, B. The sandbox for analysis: concepts and methods. In *ACM CHI* (2006), 801–810.