# InfoCloud

## improving business intelligence usage by socialization

van der Wielen, F.D.

*Award date:*
2011

# InfoCloud
# Improving Business Intelligence Usage by Socialization

Florian van der Wielen

October 2011

# Abstract

Business Intelligence tools are used by professionals to obtain knowledge about the organization they work in which in turn is used to make good management decisions. One type of BI tools use visualizations to make the data in the organizations' databases more understandable and insightful. This report describes how the usage of such a visualization BI tool can be improved. After interviewing professionals and studying related research the addition of a social component is chosen. The resulting prototype in the visualization tool combines user comments with saved visualization states (called snapshots) into discussions and presentations. This prototype has been user tested, which showed that combining visualizations with discussions or presentations has its merit in an organization, that snapshots are a powerful and precise medium of expression, reference and argumentation in such communication. Adding a social component to a BI tool is deemed as a success, as it gives context, guidance and communication to the visualizations and motivates the users.

# Contents

CHAPTER 1

# Introduction

A picture says more than a thousand words. This is the core principle on which the visualization field is seated. When presented a table full of data, even the best analyst will take hours if not days to extract all the intricacies it encapsulates, whereas from a histogram or scatter plot of that same data, he knows it in much shorter time. It therefore has an immediate benefit and value to use such systems that can translate these big tables into meaningful images, and the user of such a system likely will not need to be convinced of the value of Information Visualization [**11**].

However, it is only within the last decades that this scientific field has found this value to be accessible by the public, as computer hardware is finally fast and widespread enough. Therefore people are trying to recontextualize the scientific field of infovis from being "by experts, for experts" towards a characterization as "for the people" [**8, 37**].

## 1.1. Business Intelligence

One of the ways to do this is through so called Business Intelligence (BI). Once introduced as "concepts and methods to improve business decision making by using fact-based support systems." (1989 Howard Dresner), Business Intelligence tools supply professionals with data of the company they work in. This gives them insight in the part of the company and work they are responsible for. This allows them to make better decisions, as the decisions are supported by actual facts and insights that are derived by using these systems. These decisions are often about management issues such as managing time, manpower or other resources.

One class of such Business Intelligence tools are visualization systems that visualize this company data. The professionals that need it can then use the benefits of visualizations to gain better insight. These benefits include using the capable human visual system, having both overview and details on demand and the interactivity of the visualization. This interactivity is beneficial as company data is typically huge, and the interactivity allows the user to focus the visualization on that part of the data he is interested in. This report focuses on BI tools based on visualizations, and we use these two terms interchangeably.

## 1.2. MagnaView WebLive

MagnaView BV makes such a BI tool. Started as a spin-off company from the department of computer science and mathematics of the Eindhoven University of Technology, MagnaView BV has developed a software solution that visualizes business data to make it more insightful. For end users, this software is called MagnaView WebLive, see Figure 1.1. As the name suggests, this is a web solution that shows visualizations to let employees get insight into the data of the organization. For an organization, this software is configured using a MagnaView project to extract the data from the company database and visualize this into multiple visualizations, called views. Users can browse between the views to look at the data in different ways. When looking at a view, they can interact with it in several ways. They can change filter settings to focus the visualization on a small part of the dataset according to attributes, like showing data only after a certain date. Data can also be hidden in a more freeform kind of way, by means of selecting the data in the view and clicking hide. Further interaction with the visualization is possible via hovering or clicking with the mouse on certain parts of the view, which respectively shows more data of the appropriate datarecord or browses to another view. All in all, MagnaView WebLive allows its users to see and interact with understandable and insightful images of complex company data.

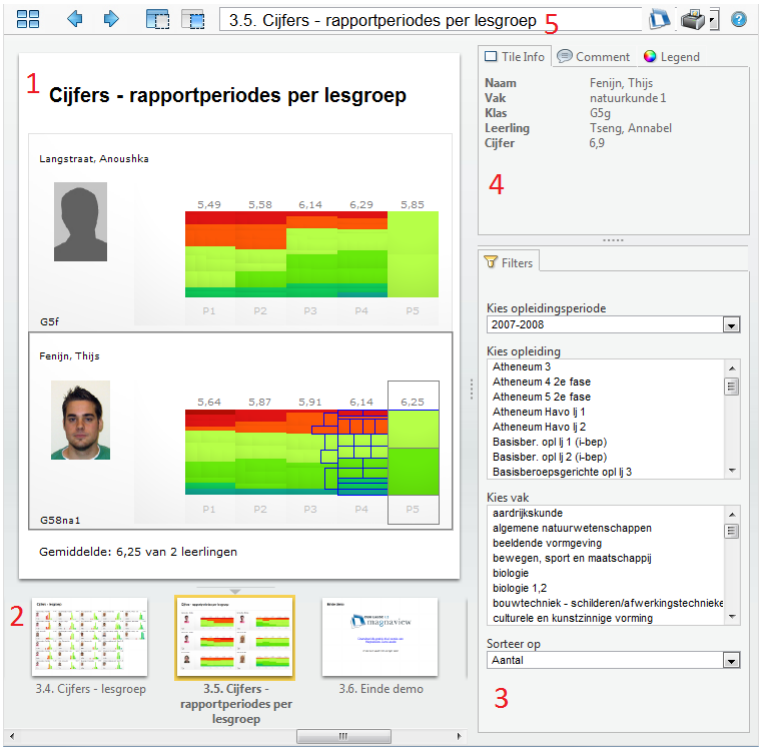



FIGURE 1.1. A screenshot of the MagnaView WebLive demo. The main part (1) of the screen is used for the visualization. In it, some data is selected, which give the selected data a blue outline. below it, three thumbnails are shown (2) that can be clicked to go to that view (other thumbnails can be scrolled to). On the right side there are two things. There are the filters (3), which show the values per attribute on that can be selected to only show records with that value. Above that (4), the data values are shown of the data the mouse is currently over. At the top there is a button bar(5), which has amongst other the button to hide the selected records, as well as unhide hidden records and options to browse through the project.

### 1.3. Problem Definition

With its close ties to university, MagnaView BV is very apt to innovate its products. Therefore they always look forward and try to find points where they can keep on improving their software. As the keystone of their BI tools is visualization, it is important to them to make these visualizations as valuable as possible for their clients.

The value of visualization has been reviewed by Van Wijk [**47**]. From his economical model of the value of visualizations it is derived that a great visualization method is used by many people, who use it routinely to obtain highly valuable knowledge, without having to spend time and money on hardware, software, or effort. This textualized equation has nice concrete keywords in which we can find this room for improvement. Let us therefore see where we can find room for improvement by looking for the valuable parts of usage. Value is directly mentioned in the knowledge obtained, which depends on the content of the visualization and the existing knowledge base context of the user. Although this content and context make it a difficult to grasp aspect of value, current users of MagnaView often say that the visualizations are very insightful, so apparently this is already taken care of. Also money factors are mentioned, however these are outside of our scope as it is up to the marketing and sales department to improve these as much as we can, it is not a computer science job to fix these. This leaves the number of users, their usage frequency on the positive side, and effort and time needed for the use on the negative side. The recurring word here is usage, so apparently that is where we can find our gain. This gives us the following research question:

*How can the usage of a BI tool be improved?*

To improve the usage of a tool is to make it more usable for the user. Also, as Van Wijk showed, we want more people to use the tool more often, which is strongly connected to motivating them. This

introduces the terms usability and motivation as aspects that should be improved. Users play a central role in these qualities, so we need to know better how this user works. As our tool is a BI tool, it visualizes management information, therefore we want to know how these users deal with this information.

### 1.4. The structure of the report

As the main research question is quite generic, Chapter 2 describes interviews conducted to get a clearer grip on how users use the information of a BI tool. Chapter 3 and 4 explore usage of a BI tool in the dimensions of motivation and usability respectively. Chapter 5 then explores possible design directions possible for the solution, and chooses one of them, adding a social platform. Chapters 6 and 7 further develop the communication in the social environment forms and how the visualizations connect to them respectively. Chapter 8 describes the resulting prototype, which is user tested in Chapter 9. The report is ended with the conclusions of chapter 10.

CHAPTER 2

# Interviews

To understand how professionals deal with management information, we performed some interviews. This allows us to gain more insight how they think about, make use of, and would like to have this information. This then indicates how the interface between them, the users, and the information should be.

The interviewees were all employees of Dutch secondary schools. This is convenient as every person in the schools' hierarchy has something to manage, from the principal who has the whole school to manage to the chemistry teacher who needs to manage his classes. This makes sure that we can more easily get the common ground between higher and lower levels of management without switching context, namely education. What also helped was that MagnaView has experience in this context, as they have a product for this sector. Therefore they had many contacts in the field who were enthusiastic in participating in these interviews.

The organizational chart of a general Dutch high school is shown in figure 2.1. A typical Dutch school has a matrix organization where one dimension is focused on the different courses (like biology, economy, mathematics), and the other dimension is focused on the different classes, the groups of students taking the courses. On the lowest level, each course (for a class) is given by one teacher and each class is under the supervision of one mentor. The second level aggregates these into course sections and coordination groups respectively. These two pillars exist more or less disjunct from each other, and are in the end managed by the schools' principal and management. Separate from all of this there exist separate staff roles that support the schools' functioning, such as IT, but of these separate roles only the Business Intelligence manager is relevant to our problem. Note that these functions are roles rather than jobs of single employees. A teacher usually performs multiple roles besides teaching, such as being a mentor, section leader, coordinator, management or perform one of these separate roles such as BI manager. A BI manager tries to get people to use BI by supplying and promoting it throughout the organization.

The six people who were interviewed were mostly teachers, with one full time BI manager as exception. All five others were teachers, but next to that two were positioned in school management, another was responsible for the schools software, another for graduation percentage control and another as a coordinator of several first year classes. This makes the peculiarities of any particular level of management not stick out as much, improving the interview results, as we try to help all levels of management. There

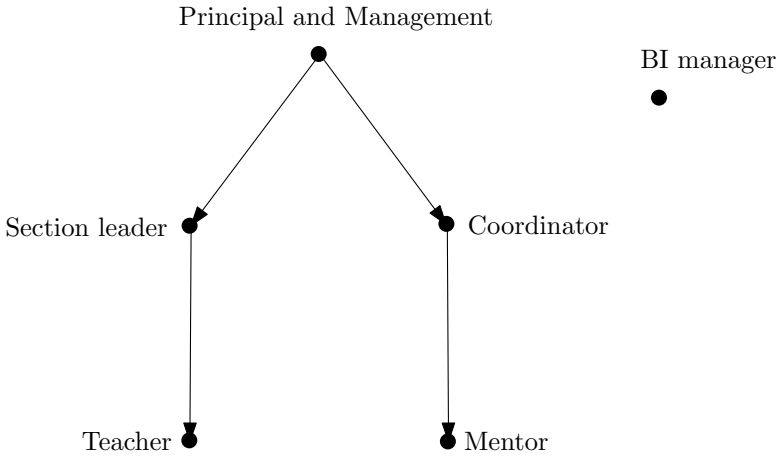



FIGURE 2.1. The hierarchy of a Dutch secondary school

were two people with computer affinity, so the difference of having this or not could also be noticed, but mainly the voice of lack of affinity would be heard, as that was more important.

The interviews were conducted in the school buildings where the interviewees work. This is a more safe setting, so the interviews were able to touch topics more related to personal experience and problems, which is what we are after. Also it is where the experience and problems occurred, so they were easier to retrieve from memory in this location. The questions were on the following general topics:

- Work situation and moments of usage;
- Importance of information;
- IT skills and motivation;
- Security and privacy; and
- Usability issues.

### 2.1. Work situation and moments of BI usage

Because work situation and moments of usage dictate how users can use the functionality, these are key aspects to how this usage is valuable to them. If the tool requires frequent usage whereas the user has limited time, the tool cannot be used and therefore cannot solve the problem. We therefore asked how their general working day looks like, regarding responsibility and work load as well. We also asked which goals in their work usually received their focus, both their short term daily tasks as well as on the long term. What came out was the following:

- **Busy during work** — They clearly said that they were scheduled too few hours for managing mentor responsibilities. One participant even said that he used at least the double of the amount of hours that he got scheduled. This shows that BI tool usage is done "in between" other duties. During the work day, BI tools are only consulted when something specific asks for it, but otherwise there is not really a moment to perform in depth analysis with it.
- **Analysis work in off time** — Managing, planning, and other related duties, such as administration, are usually postponed until for example the evening, when the teacher is not distracted. This is possible as they tend to be equipped with notebooks which can be taken home.

### 2.2. Importance of information

We are also interested in how important information is to the users. If the information is nice but there are more pressing tasks at hand, it is obvious that the BI tool will play second fiddle to the actualities. It is also interesting to see what information is important to the BI user, throughout the organization. Obviously, if the same data is of high importance for everybody in the organization, the BI tool based on that data must be very interesting and related to the core business of the organization. Therefore we asked all interviewees what information was important for which decisions, and if this information was accessed through a computer. We also asked how often decisions were taken that were based on data, and for what types of situations and questions they were counselled. It had the following results:

- **Demand for information** — the participants answered that "there is a demand for information", in that they feel a lack of information when they have to make decisions. For instance, they noted a lack of a student tracking system, in which they can easily observe the track record of a student.
- **Teamwork** — Another important fact is that the teaching job is really teamwork, all teachers involved try to get the student to perform best. This means that the teachers should help one another in keeping each other informed, as one interviewee put it.
- **Correctness** — Current information systems tended to be correct if and only if all relevant data was present. This means that if some part is missing, due to human or computer error, the information they get is not up to date, is too vague or even just blatantly incorrect. This makes it difficult to verify (or falsify) suspicions and intuitions.
- **Easy analysis** — participants noted that they want flexible overviews and points of reference to perform analyses more easily.
- **Information sharing** — it was disturbing for participants that they often heard information after the fact, for instance during a meeting from a colleague, not when the actual fact is happening. This is frustrating, as it keeps them running behind the facts. They noted that

a system that would deduce this kind of facts for them could be beneficial, as for example a student absent for 3 days is probably ill and this needs attention. All in all, a nice inbox where they would receive this type of information would be appreciated.

### 2.3. IT skills and motivation

Software is often complex to use. If users want to go anywhere further than the basics they need to look it up in the manual, often adding to their frustration and thereby demotivating. However, complexity grows with capability, as a program with many features needs a GUI that fits all the features inside. Such a complex GUI is by definition not easy to use by layman users. We are therefore limited to the IT skills of the user, as well as his motivation to use and develop such skills. The following remarks were made.

- **Web based apps** — They use emailing and web-based environments. Basically, just simple office software is used, like email, electronic calendars and word processors. If they use spreadsheet programs it is only on a basic level. Past that, the typical teacher does not really use a computer much so new applications should not be too complex, like web apps tend to be.
- **High tech is expert only** — More advanced software, such as the MagnaView visualization designer, was limited to enthusiasts such as the interviewed BI expert. This is no surprise, as also other people that were more involved in the schools' IT policies are also more prone to use IT solutions to other problems that they experience. However, this is only a select group of people,
- **Clumsy but effective** — As the employees needed to use more complex or unknown software, they experienced more difficulty in this usage. For instance, when trying to get the software to reach a certain state, if this state was reached by using a side effect of a different button and then cancelling its main purpose, the users were already happy.
- **Bureaucracy bores** — As everybody is, a teacher too is bored with bureaucracy. This results in slow data entry, especially when the culprit can hide in anonymity. One school therefore even let the janitor insert the data each Friday afternoon, to make sure that it at least happened, if only a week late.

### 2.4. Security and privacy

As the data in the schools' information systems is very personal, and access to it is very restricted and regulated by Dutch law, the BI tool that uses this data needs to have sufficient security measures. However, security measures tend to block application functionality and speed, which therefore blocks the ease of use. Therefore this is relevant to us. There were two key aspects to this issue:

- **External privacy maintained** — However, during the interviews it came to show that the employees use preconfigured notebooks, which allows for setting up a secure connection easily. Therefore, only the sharing of the data within the organization is relevant.
- **Keep internal privacy sensible** — We therefore asked how they look at the usage of their own personal data, or less personal data such as mere averages. They said that security is indeed an important issue, but it is a question on which they did not have an immediate answer. However, as the government uses data for educational quality control, the participants said the school should do this as well. As a side note they also mentioned a flexible rights system, as they found rigid software security conflicting with the real world demands.

### 2.5. Usability issues

As we consider usability a big factor if a BI tool is used or not, we asked some questions regarding that. Every program has its own quirks and, despite being useful, nuisances. These nuisances tend to be focused on usability aspects, not features themselves. Therefore these quirks can inform us about important usability aspects of these kinds of programs. The interviewees noted the following:

- **Support core functionality** — The main target of the complaints was that the information supplied was not good. It could be unclear, not up to date or incorrect, even though the main

purpose of a BI tool is to supply information. Of course, an information system that fails in supplying information is not decent.

- **Support the standards** — Another source of annoyance was for instance the lack of copy paste functionality. Also other usability basics were noted as problems, such as a computer program feeling sluggish. Almost everybody mentioned some basic usability aspect of software that they experienced, and every time it was a great source of irritation.
- **Small deficits give big annoyance** — The interviewees did not only give examples where usability was lacking, but with it, we noticed quite some negative emotion coming along. They really stated the facts as examples of how annoying software sometimes is in usage. Therefore, it shows that usability is very important for software to connect with the user.

### 2.6. Solutions

The interviewees were also asked what they saw as viable solutions to their problems. This gives an interesting list of options which could then be further explored and indicate the play field. Several ideas were mentioned:

- **Show user trail** — One interviewee suggested to show user activity, for example by using messages. Even though it is a very rough idea, this will certainly give a different feel to the application as right now MagnaView does not collect data by itself, nor does it let its users create data in it, only viewing of data is possible.
- **Home screen with notifications** — Another suggestion was the creation of a home screen that is visible when people start to use the BI tool. This will give them quick access to things that need their attention, thereby giving a clear starting point in using it.
- **Experience points** — Another, more wild remark was that there would be some kind of "MagnaView experience points" as the interviewee dubbed them. Although very vague at best, it shows that users are open to radical changes in the usage model, even drawing inspiration from videogames.
- **Fact derivation** — As said before, one interviewee said that he would like some fact derivation from the system. He put it the following way "Can the system email me when it sees a student is absent for 3 days? This very obvious and simple fact clearly needs my attention, and it is too tedious a job to check it by hand when a computer can do this much better."
- **Create own overviews** — Some interviewees also said that they would like to be able to make their own overviews. In a BI tool, this means letting them make a visualization. The downside is that making visualizations is actually quite hard.

### 2.7. Summarizing

From all the user characteristics listed in the previous sections of this chapter we can distil our set of requirements. Some of these requirements are indeed a good lead in solving the problem at hand, but they fall out of the scope of the project due to their subject matter, so in the end they are dropped.

- **Correct Information** — Of course the information should be correct, otherwise it is worthless. However, that is not up to the tool in which the visualizations are displayed, but a requirement for the data shown. *dropped*
- **Automating human tasks** — Deriving new knowledge from existing knowledge is quite an interesting topic, which goes so deep as to be an entire sub-research field in computer science. However, as this graduation project is in the field of visualization, it falls out of scope. *dropped*
- **Privacy** — Like the interviewees mentioned themselves, security is important, but assuming vital information is not leaked out of the organization, it does not receive our main focus and is merely taken into account on a side note. *dropped*

- **Web app complexity** — Because a BI tool is meant for all professionals in the organization, it should not require much IT skills, as not everybody has those.
- **Teamwork and sharing** — Colleagues are valuable sources of information and teamwork is important, so supporting this in some way would help the user.
- **Pick up and use** — Due to busy working hours, it is best if using the BI tool in these hours requires a time investment as small as possible. Therefore, the application should be useful as fast as possible after starting it.

- **Fun to use** — Because deeper analysis is usually done in free time, and also to promote usage in free time, the application should not be boring and annoying to use. Instead the user should enjoy using it. This also promotes longer usage which allows for deeper analysis.

These aspects are all relatively clear. However, they do not show a general direction in which we can find a solution. Therefore, we review the topics of *motivation* and *usability* mentioned in the introduction, starting with the former.

CHAPTER 3

# Motivation

Recall from the economical model of valuable visualization that we want *more people* to use the BI tool *more frequently*. They must be motivated more to get this to happen, so we look into how to employ motivation for this purpose.

## 3.1. Volition

Volition or willpower is the cognitive process by which a person decides on and commits to a particular course of action. This is important as we want to have more people to commit themselves more often to BI tools. Volition is influenced by motivation, not reason, in contrary to common assumptions. Let us clarify this with two examples. An alcoholic can go to great extents when succumbed to his addiction, displaying a strong will in getting his fix. On other moments he can think that his drinking habits are ruining his life and go through the tough process of rehab, which requires great commitment as well. Although contradictory, they both make sense for the same person. However, logic would show that there is only one better choice. Therefore, it is not logic that is involved in this decision making. That is why we can focus on just inciting motivation, and there is no need for rationalizing this motivation.

## 3.2. Herzberg's two factor theory

As with many aspects of the human psyche, the true inner workings of motivation are still unknown. As such there is no one single theory explaining all its intricacies correctly. However, motivators (things that motivate) can be grouped into groups with similar traits. One such taxonomy is Herzberg's two factor theory [**22**], which is a common, clear and useful theory.

Herzberg's two factor theory on motivators says that there are two types of factors involved in motivation, namely

- **Motivational factors**, that when present really motivate the user. A typical motivational factors is satisfaction, which means the user feels the tool or action involved really helped him in in reaching his goal.
- **Hygiene factors**, that when present do not really motivate the user, but on the contrary demotivate when absent. Typical examples are stability and performance in software.

What is expected is that BI tools already in deal with hygiene factors properly, making fast programs that are not cumbersome. However, that is not truly motivating, which means that if true motivators are added, more people will use the BI tool more frequently.

EThe two-factor theory makes another distinction. Note that this distinction is disjunct from the previous, even though only the word motivator is mentioned here. The following types exist:

- **Intrinsic motivators** are motivators that come from the tool or usage of it itself. A typical intrinsic motivator is for example if the usage is fun.
- **Extrinsic motivators** are motivators that come from something else, but are involved in the usage. For example money or other rewards, or threat of punishment are typically extrinsic motivators.

The two distinctions, motivational/hygiene and intrinsic/extrinsic, span a two by two matrix, and for each combination of chosen attributes examples exist, see Figure 3.1.

In the context of a BI tool this distinction has two important parts. Firstly, only the intrinsic motivators can be influenced, as only they are part of the BI tool we can modify. However, the extrinsic motivators still play a role in the whole situation of usage, and should be taken into account.

| | *Motivator* | *Hygiene* |
|---|---|---|
| *Intrinsic* | Fun, Mastery<br>Satisfaction | Efficiency<br>Reliability |
| *Extrinsic* | Social interaction<br>Recognition | No threat of punishment<br>Salary |

TABLE 3.1. Examples of aspects that influence motivation. As these aspects can be either intrinsic or extrinsic, and motivating or hygienic (demotivating when absent), there are four types in all.

## 3.3. Intrinsic Motivators

Usability and user experience play an important role in intrinsic motivation, as intuitively a tool that does not feel usable and gives a poor user experience will not motivate usage. The definitions of usability and user experience confirm this connection. The ISO definition of user experience [**23**] is:

*a person's perceptions and responses that result from the use or anticipated use of a product, system or service.*

As the intrinsic motivation comes as a reaction or response on the usage and from this usage, it is part of the user experience. Additional notes mention users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after use as important factors. Therefore it is highly complex, and on top of that it depends also on the user and context, which makes it a difficult subject. Luckily, usability is more well-defined than user experience, and we review it in the next chapter. Therefore this section is limited to well-known motivators occurring from usage.

**3.3.1. Self-efficacy.** Self-efficacy is the belief of having the capability to reach a certain goal. Within the context of tool usage it means that the user thinks he can reach the goal by using the tool. In that sense it is closely related to mastery, as it is the expectation how much the tool is mastered to reach the goal. Typically, BI users are not tech savvy but need to use the software tool, which means that self-efficacy is low. This demotivates to use the tool, as they think they will fail using it.

**3.3.2. Fun.** If the tool is genuinely fun, the user is motivated to use it. However, it of course needs to be fun enough, and clearly not everything is always fun enough to really motivate. This "fun enough to motivate" is sometimes called engagement. The user experiences such engagement also when he is not using the tool. Also when he is not using he is reminded by the fun the usage gives, and therefore he longs for using it again. This "sticking to the mind" is a good way to have people recurringly use the tool. Very good examples of this are videogames, as their whole purpose for the player is to be fun. This makes videogames a good option for the general design direction of the solution.

## 3.4. Extrinsic Motivators

Hygiene factors are usually extrinsic. This is because hygiene factors tend to be about the situation, such as salary. They are more a requisite to the user for doing the activity. Although they are hygiene factors, they can motivate at least temporarily or superficially. In that sense they are a quick fix, as just adding that motivator, decreases the de-motivation or superficially improves it.

There is a flip side to adding these extrinsic motivators however, as it can lead to *overjustification*. Overjustification is the situation where the extrinsic motivators decrease the strength of the already present intrinsic motivators. For example, if toddlers are put to draw for themselves, the situation obviously revolves about them drawing which is genuinely fun for toddlers. This in contrast to toddlers being told to draw in reward of a nice ribbon, in which they would only draw for the ribbon, and when they have drawn enough for the ribbon they would already stop, no longer being motivated to draw. Therefore they spend less time drawing than in the situation without the extrinsically motivating ribbon. As visual analysis costs time this overjustification effect is something to be avoided.

In the following, a few extrinsic groups of motivators are discussed, showing their weaknesses and strong points.

**3.4.1. Incentives.** Incentives is the name the socio-economics field has given to motivators. Socio-economics looks at incentives to influence behaviour, for example a discounted price to improve product sales. As such, socio-economics do not focus on any specific behaviour, but the incentive or motivator. This also means that the extrinsic motivators are better represented, as they are not as dependant on the behaviour.

One of the main problems in the incentive theory is that incentives can motivate unintended behaviour. As people are creative, and always try to maximize the return of investment of their behaviour, they adapt their behaviour to get the biggest reward. This results in undesired side-effects, and the name to such an incentive is *perverse incentive*. An example is a $17th$ century English tax law that scaled along the number of windows of a house, then a measurement of luxury. This caused living standards of the poor to deteriorate, as landlords bricked up windows to reduce their tax liability, leaving the renters with insufficient light and fresh air.

**3.4.2. Social motivators.** Social motivators are motivators that show up in a social situation. In that sense they are extrinsic, as they do not come from the tool or usage itself, but from that usage in a social context. However, as almost everything done in life is done within some social context, it could be considered as intrinsic on some level. This social context especially exists for the work situation, in which our BI tool is used.

There are both negative and positive social motivators. Negatives are ones such as social pressure from peers or superiors. Positive motivators are for example social contact, status, influence or competition/winning. Encouragement to start or continue usage also motivates, and when it is mutual, brings about a positive atmosphere which many people like. Social motivators can prove to be important, as users base their attitude towards a system on their first experience. If this experience is bad, these users really need to be motivated to change their attitude. As they will not use the system to motivate themselves, the social factor is one of the few plausible ways to influence his attitude and motivation. In other words, the social actions occurring in an organization could improve the adoption of the BI tool throughout the organization. This adoption is an important aspect, as it means more people use the BI tool, one of the goals posed by the economic model of Van Wijk. All in all, the social situation introduces many motivators, which makes it a viable option for the general design solution.

**3.4.3. Rewards.** Rewards are one of the most well known incentives, and not only money works, but also for example more abstract rewards such as badges. This is more or less a version of the Skinner's box. B.F. Skinner was a behavioural psychologist who put a pigeon in a box with a button, with a food pellet delivered when the pigeon pressed the button. Like so, the pigeon was taught to press the button. This also works for human beings, however over-justification causes the motivation to be rather hollow, which therefore no longer works at some point.

### 3.5. Criteria and Options

As we are looking for a decent design direction to follow, this overview of motivation showed several relevant criteria that the design direction should perform well on, as well as two options for the design direction itself. The criteria are listed below, together with a small argument to signify their importance.

- **Intrinsic motivation** is important as it contains the longer lasting motivation to use the BI tool again and again.
- **Extrinsic motivation** is important as it allows for improvement of motivation when intrinsic motivation no longer works.
- **Adoption** is an important aspect, as it is a phase the BI tool undergoes when it is introduced, as well as the planned adaptation of it.

The options are:

- **Games** are software systems focused on offering such a nice experience that players are motivated to play it again and again, which sounds like the way to make many users use the BI tool more often.

- **Social** situations introduce new aspects on which users can be motivated to use the BI tool, as well as matching with the social situation of a company.

CHAPTER 4

# Usability

As mentioned before, we want to improve BI tool usage. This means that we want to make the BI tool more usable, which means in turn that we want to improve the usability. We also noticed in the previous chapter is that the usability of the tool influences the motivation of the users. We therefore look closer at usability. The ISO definition of usability is

*"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."*

Putting this definition in the context of the previous chapter, effectiveness and efficiency are hygiene factors, as they are required to make the usage 'nice', but they will not necessarily motivate the user.

Satisfaction on the other hand is a motivator. Also, satisfaction is a response on the usage, and therefore it is part of the user experience. However, user experience is something very complicated, which is why we stick to known concrete ways to influence it, especially those related or applicable to information visualization.

## 4.1. Effectiveness

Effectiveness in usability means, given a certain fixed amount of time and effort, how good a tool is in reaching a goal, like how accurately and complete the goal is reached. Our goal is to get as many people to use the BI tool as often as possible, gaining as much valuable information each time. A breakdown gives:

- **As many** — to reach as many people, we need accessibility. They must be able to access it, both physically as well as understand how to get to certain functionality. Therefore, this accessibility is a valuable criterion.
- **As often** — To use it often, the users must be motivated to do so, like described in the previous chapter. Therefore, we do not need to look further into it.
- **Valuable knowledge** — Information visualization has added value as user say it can do things for them that statistics simply cannot [**14**]. However, as Van Wijk [**47**] showed, it is difficult to assign actual definite value to certain information. We therefore leave it out of consideration.

## 4.2. Efficiency

Efficiency means that given our goal of getting many people use the BI tool often getting valuable information, that reaching this goal costs as few time and effort as possible. We again perform a breakdown:

- **Time** - To make our BI tool as time efficient as possible, the time of usage to obtain certain information should be as short as possible. Time of usage denotes here the period from the moment where the need for information arises up to the moment where a decision can be made or an insight was obtained. One aspect here is a good integration with the workflow of the user.
- **Effort** - To make the user input as few effort as possible, the ease of use should be good. This leads again to good accessibility.

Integration into the users' current work environment is an important factor for both the general successfulness [**27**] as well as the adoption of a visualization system by data analysts [**28**].

However, using two programs at the same time, such as two data analysis tools, disrupts the flow of usage [**15**], and statistics software is sometimes better than information visualization, usually in routine situations. Combining the two into one powerful application probably would not work however, as this would require a very smart user. Therefore it makes more sense for both types of applications to support only a little bit of the other ones functionality, so the switch between programs does not happen so often.

It is important to do this right, however MagnaView is already an established tool with many smart design choices and a complex architecture, so adapting it to improve workflow integration is too difficult.

### 4.3. Satisfaction

Satisfaction does not have a definite underpinning theory splitting up user satisfaction into components. As mentioned before, satisfaction is a component from the user experience. User experience lacks theory as well, so none can be borrowed from here. However, some concepts do exist in the user experience that intuitively influence satisfaction, as they give the user a positive feeling.

**4.3.1. Flow.** Flow is the mental state of a person doing an activity who is fully immersed in a feeling of energized focus, full involvement, and success in the process of the activity [**29**]. There are several prerequisites to the situation, which are that the activity needs to be intrinsically rewarding and to have clear goals, direct immediate feedback and a balance of ability and control. The person also needs to have some sense of control over the situation.

Then, if the user would obtain flow, he would lose contact with his sense of time, his self-consciousness and his self-awareness of the moment. Instead he becomes very concentrated and completely focused on the activity. An important factor here is matching of the person's ability with the activity's challenge, in which the user would solve successive problems from the activity and that his thoughts are completely in sync with the problems of the situation. When these two synchronize with each other, they augment one another, and this is what we call flow [**29**].

To get this to happen, the flow of the program needs to get equal with the flow of the user. Too slow, and the user gets bored. Too fast, and it is too difficult for him to use the program, and the complexity frustrates him. Therefore it is about setting goals and reaching those goals in a rapid pace. When flow is attained, users become more focused and perform better. This makes the pace even higher. Therefore, the balance of flow is highly dynamic, so attaining a high flow is difficult [**34**].

**4.3.2. Interactivity.** Interactivity is the gradation in which a computer program is continuously open for input from the user. A batch program is non-interactive, whereas programs that have a GUI are typically interactive. Interactivity is a neutral metric. For example, power users, such as system administrators, typically do not want high interactivity, as the continuous prompts from the program for input makes their usage less efficient. These power users do not need the guidance that these prompts provide as they already know what to do. In the contrary they prefer efficiency as they need to do the complex tasks regularly. Layman users on the other hand do not perform complex tasks regularly and do need guidance, which means that interactivity is beneficial for them.

Interactivity is a very important aspect of information visualization software [**35, 26**]. Because of this interactivity, visualization users are able to browse through large datasets. This agrees with Shneiderman's mantra [**43**]: *Overview first, zoom and filter, then details-on-demand.* With this interaction the users can steer the focus of the visualization in a natural way, showing exactly that part of the huge dataset that they want to see. This means that the flow of thoughts is not disrupted, which improves the thinking process during using the visualization. It has also been shown that interaction gives joy and increases affect, emotion and perceived so-called aesthetic [**5**]. Regarding all these benefits, it is not surprising to see a move to systematize information visualization research and bring more than standard interactivity into data analysis and exploration [**17**]. An exponent of this move is the call for a new field within information visualization, which is called information aesthetics. This sub-field says that not only data, but also aesthetics and interaction play a central role in visualization design. Because of this focus shift towards aesthetics and interactivity, information aesthetics makes a new connection to information visualization with visualization art, which gives new light to old problems. Therefore, this information aesthetics could be a potential design solution to our problem and we add it to the list.

### 4.4. Criteria and Options

Again several useful criteria and another design option showed up in this review. The criteria are:

- **Accessibility** conforms with the earlier noticed property of a web app complexity level of the GUI.
- **Flow** is difficult to attain, but it highly augments the attained pleasure and focus, which combines nicely with the focus needed in analysis, also visual analysis.

- **Interactivity** is needed because layman users need to use the BI tool.

The new option for the general design direction is:

- **Aesthetics** has good bonds with interactivity, and also shows its strength in simple, elegant GUIs.

CHAPTER 5

# General Design Direction

As we noticed from the interviews, there are still several features lacking for some users of MagnaView, like easy data transfer functionality, performing some automated analysis and allowing some more sophisticated end-user analysis. This corresponds to observations done about the development and evolution of the visualization field and tools [**36**]. Even though obviously improving functionality, it just means implementing some nice features. These features do have enough depth to be a graduation topic when taken separately, for example fact learning as automated analysis. However this is not a visualization topic which we set out with. We therefore look at the three design directions found in the previous chapters. They are gaming, social and aesthetics. We start with the first.

## 5.1. Games

Video games (or games for short) are games played on computers. As the name implies, the video image on the screen plays a key role in the game. As such, video games have many similarities with visualizations. Video games are highly interactive pieces of software, and meant to be fun, so these aspects are well designed, even so much as to be known to incite flow [**34, 42**], as well as satisfaction [**38**]. One of the main reasons for games to be fun is that they are a safe learning environment, in which the players can improve their skill.

**5.1.1. Examples.** Gamification is a buzzword that means the idea of applying the good qualities of games to everyday situations such as work to improve them. Often what is brought under the buzzword banner of gamification is the behaviour-reward schemes of the Skinner box, with examples such as Foursquare that gives badges to people often visiting certain stores. However, videogames give the rewards only after behaviour that requires a clear skill, and it is this acknowledgement of skill that makes this reward meaningful, not the reward itself. Therefore, we do not want to simply apply such a reward scheme.

A real example of video games within the context of information visualization exists as well, luckily. Diakopoulos has tested several game mechanics based around a visualization [**9**], see figure 5.1. Results showed that the games were fun, however, the fun was a perverse incentive, as players were focused on the scoring and the game mechanic, not the actual data. However, as such they did serve as a playful introduction to using data visualizations.

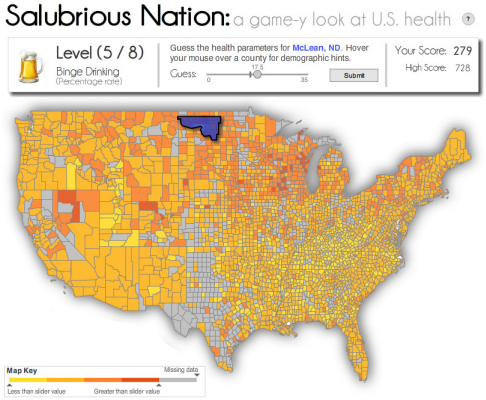



FIGURE 5.1. Salubrious Guess, one of the tested "Game-y Infographics", in which the player needs to accurately guess the magnitude of a given health parameter for a randomly selected target county [**9**].

**5.1.2. Strengths.** Games show that autonomy, competence, relatedness, fantasy, curiosity and challenge are important factors of enjoyment [**32, 38**]. Especially clear specific goals are enjoyable, which does conflict with the explorative analysis for which information visualization is lauded. Despite this minor setback, games show a major connection with fun, which could be applied to BI tools via these keywords.

**5.1.3. Playing is learning.** When we think of games, we think of fun. When we think of learning we think of work. Games show us this is wrong [**12**]. Actually, the theory of why games are fun is because we learn through playing them [**29**], as they keep giving us increasingly difficult situations to overcome, which we eventually do. In short, this makes the body generate endorphins which give pleasure. Games are also good in teaching factual knowledge on the side [**24**], and making this enjoyable [**12, 3**]. This playing is learning aspect combines nicely with the obtaining of information goal of a BI tool.

**5.1.4. Multiplayer.** Another important component in games is that quite often they are played with others. For multi-player games this is one of the bigger sources of enjoyment they tap into, the social aspect. The competition and collaboration this generates gives fun in itself and also gives meaning [**50**]. This meaning is important too as making something meaningful for a user makes it more important. This multiplayer aspect does give a whole new dynamic to playing games, with competitors and collaborators that have either focus on playing the game or playing with others [**1, 25**]. This needs to be taken into account into the design of such a game, and makes it more complex.

### 5.2. Social

Social information visualization is allowing some form of communication to be combined with an information visualization tool. Already quite some attention has been put into this sub-field of infovis. This is amongst others because it is one of the proven ways of opening up the possibilities of infovis for the main public [**51**], something the research field struggled with for quite some time. The strong points of social infovis are the collaborative analysis users can perform, by sharing visualizations in some way. This improves the analysis of the visualizations, as knowledge becomes known to more people faster and analysis is deeper [**18**]. Also, because people publicize the visualizations using this social infovis, they encourage each other to use it.

There is a transition in placing infovis tools not as individual applications but as part of the online platform of the organization, which therefore connects to the organization as a mini-society with social activities [**52**]. This concept has several interesting aspects. As Heer puts it, " "the social life of visualization" (to borrow a phrase from Wattenberg) has shown to be an important factor shaping the adoption, use, and efficacy of a visualization", and collective use gives deeper analyses [**18**].

**5.2.1. Examples.** The most famous example of social infovis is ManyEyes [**49**]. This is a web visualization platform on which users can upload datasets, visualize these datasets and discuss them. To organize this, topic centres exist to which visualizations, discussions or datasets belong. Another aspect is that a discussion belongs to one made visualization, and the settings of that visualization can be uploaded with a comment the user makes in a discussion. Discussions also exist separate from the visualizations. See Figure 5.3 for a picture of ManyEyes.

Another example is Aruvi [**46, 45**], a visual data analysis tool (see Figure 5.2). Aruvi has a data view, a knowledge view and a navigation view, used for visually examining the data, writing down findings and navigating through the findings. All visualizations states are saved, and a tree exists with connected states when the user went from one state to the other. Findings can be linked to visualization states, and these special states are also denoted. As such, users analyses are recorded, and these recordings can be later retrieved, also those of other users. To show how a retrieved analysis went, there is an interface allowing overview, searching and retrieving through such a (single) analysis. This allows for collaboration in the analysis.

**5.2.2. Strengths.** These two examples show the two main research directions of social infovis. The first track, signified by ManyEyes, focuses on the social part as a goal in itself, trying to tune infovis to be for the people, making it accessible, fun, publicly available and easy to adopt. The second track, signified by Aruvi, focuses on collaborative analysis, testing properties and processes involved in this analysis.

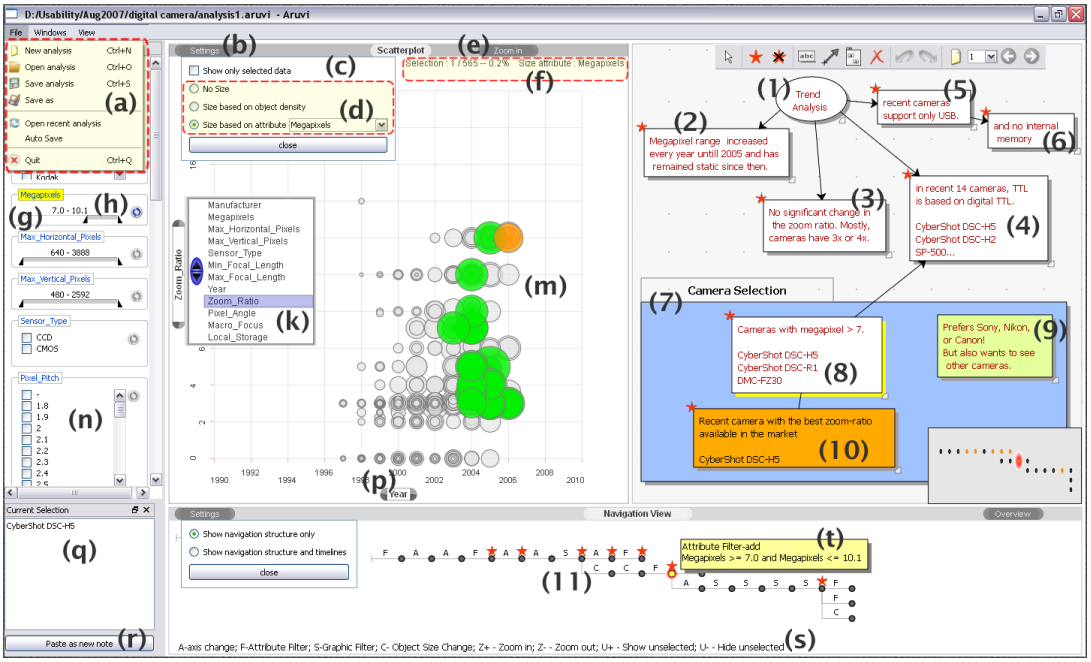



FIGURE 5.2. Aruvi, an information visualization tool that supports the analytical reasoning process. The letters and numbers denote different aspects of the GUI.

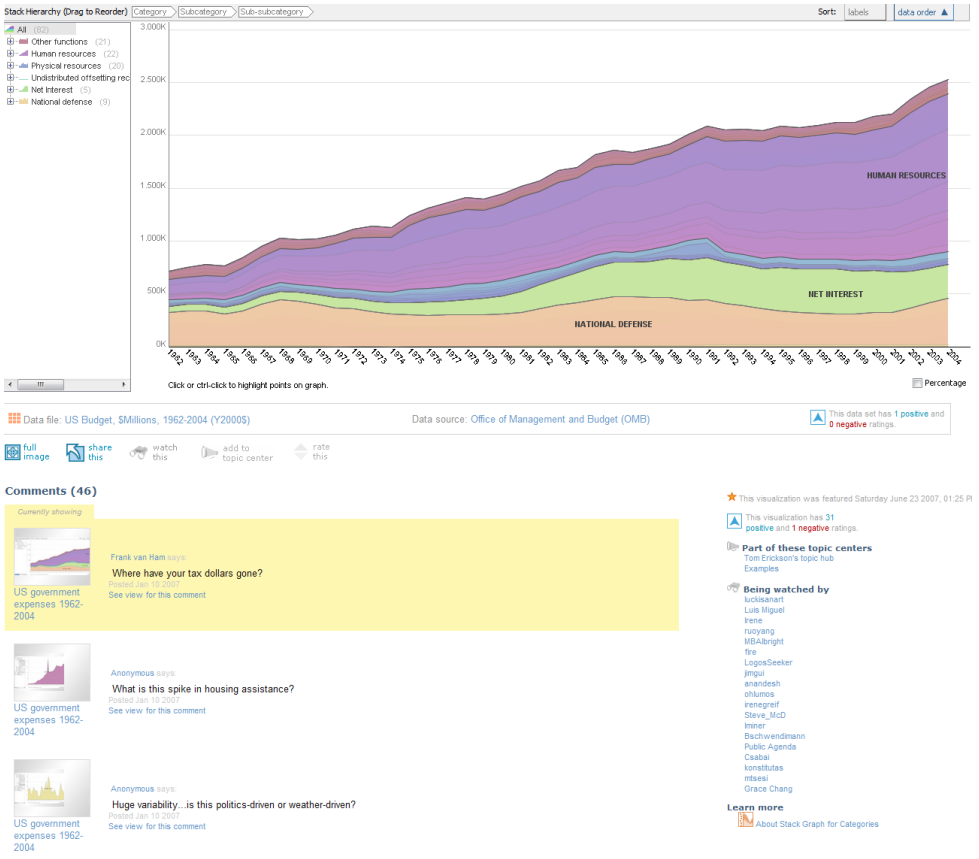



FIGURE 5.3. ManyEyes, an online social visualization tool, in which users can upload data, create visualizations and discuss them, `www-958.ibm.com/software/data/cognos/manyeyes/`,

**5.2.3. Adoption.** Socializing visualizations motivate their users automatically [**48**]. This is because social interaction and approval is a great source of motivation [**6**]. Also, this social aspect allows it to be used as a medium for encouragement and communication. This makes that adoption of the software in an organization should go easy, as people are easily motivated to use it, intrinsically as well as extrinsically by others.

**5.2.4. Collaborative analysis.** On a very general level, socializing information visualization has shown to be very beneficial. The use of visualizations in everyday has been suggested to be that of a communication medium [**7**], which very much relates to the visualizations' purpose of displaying information in an understandable way. Also, a social component makes interaction with visualizations easier [**31**] and making analysis easier. On top of that indications exist that combining conversation and visual data analysis can help people explore a data set both broadly and deeply [**19**].

Collaborative analysis exists of phases of independent and collaborative work. Ha et al.[**16**] noted that in the beginning the analysis activity and tools are more individual, but slowly become more collaborative. This could give a larger number of independent perspectives on the data, which improves comprehension, decision making and thereby action [**16**].

### 5.3. Aesthetics

Information aesthetics forms a cross-disciplinary link between information visualization and visualization art. Whereas visualization is traditionally more focused on development of new algorithms that try to actually visualize a difficult data set usefully, information aesthetics not only looks at how to present that data, but also to make this presentation aesthetically and interactively appealing. As this is a rather new field coming together from two extremes, the visualizations from the aggregated fields tend to be either technical or artistic [**8**]. This results in visualizations that aim at mapping the data in an understandable way on the technical side, and visualizations aimed at giving a message to the viewer by means of mapping data in an interpretive way on the artistic side, see Figure 5.4 [**30**]. Information aesthetics basically says that information visualization can learn and improve from visualization art, which will result in more aesthetical and interactive appeal for its users.

**5.3.1. Examples.** Because of it being a new field, not much research has been done on visualizations with a true focus on both aesthetics and functionality. However, some user tests have been done, exposing earlier work as examples of aesthetic information visualizations. SunBurst is one of those visualizations, see Figure 5.5. It is a circular hierarchical visualization, starting with the parent node of the tree in the center and subsequent levels around it. As aesthetic can come from many things, its real decomposition is not really known, but it did show that people liked it more in general too, interacting more and longer, and trying harder when using it for their analysis [**5**].

**5.3.2. Strengths.** As mentioned before, the look and feel, officially called aesthetics, highly determines how good it is in generating engagement [**30**]. This feel is more or less the interaction component of it, which means that Aesthetics has a strong connection with that usability aspect.

Aesthetics give delight to the user [**39**], which makes it an intrinsic source of motivation. Therefore a focus on having an appealing interface with satisfying sounds, attractive graphics, compelling content and alluring metaphors [**44**] makes sense. The delight factor also works in a corporate setting [**39**], which makes it very valuable to the situation at hand.

**5.3.3. Analysis with emotions.** Emotions play an important part in the human psyche, and they obviously influence our thoughts. The emotional system is even so intertwined with our cognitive abilities, that it is stated that "only through our emotions do we unravel problems" [**33**]. Therefore in good design, beauty and brains, pleasure and usability go hand-in-hand. It is therefore argued that aesthetic should no longer be seen as a cost to utility (utility not functionality), as it really has added value [**30**].

### 5.4. Comparison

With these three options to increase satisfaction, we need to select the most suitable one. To this end, we quantify how the three solutions score on the various criteria discussed in the previous chapters.

The scores are a rough estimate, based on the the literature we studied and observations we made. The scoring is done on the criteria found in previous chapters. Another criterion is added, namely

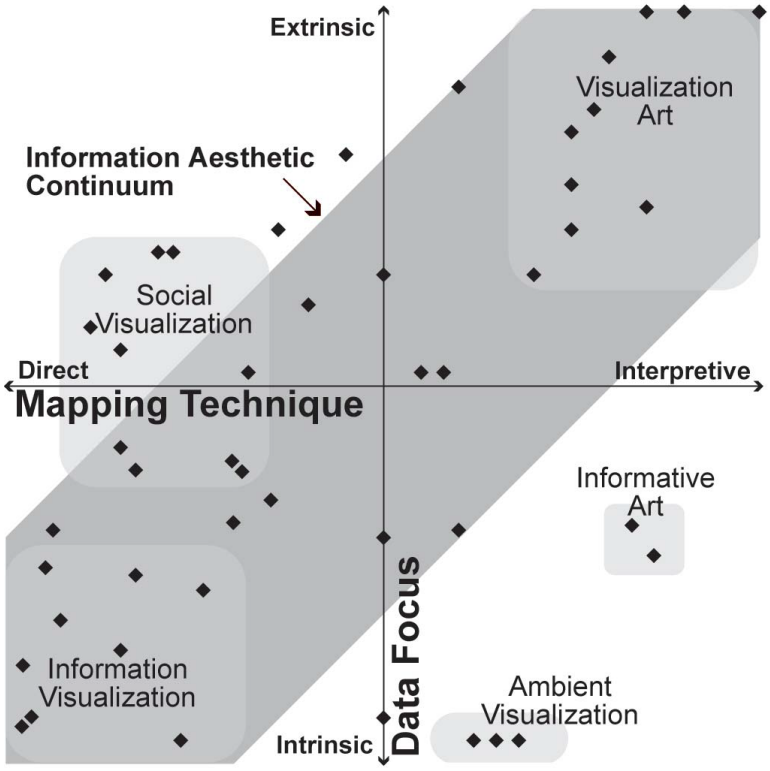



FIGURE 5.4. The dimensions of aesthetical information visualization, data focus and mapping technique. Note that in the aforementioned social visualization field the data has a more extrinsic focus, as there, the goal is also to socially interact [**30**].

| Solution | *Intrinsic m.* | *Extrinsic m.* | *Interactivity* | *Flow* | *Adoption* | *Accessibility* | *Maturity* | # |
|---|---|---|---|---|---|---|---|---|
| Social | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ | ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | 5 |
| Gaming | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ | ⋆ ⋆ ⋆ | ⋆ ⋆ | 3 |
| Aesthetics | ⋆ ⋆ ⋆ | ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ | ⋆ ⋆ | ⋆ ⋆ ⋆ | ⋆ | 1 |

TABLE 5.1. Comparison of social, gaming and aesthetic design solutions. The final column counts the number of wins on the attributes before

*maturity*. With maturity we try to quantify to what extent the solution direction has already proven itself, both in theory and out in the field. This is a very practical argument, as it increases the probability of success. Also, it allows some validation, as a solution that uses more established techniques allows for comparison with those techniques. It also allows us to use gained insight in the field to determine other choices we need to make later on.

See Table 5.1 for a comparison of the three design solutions. Each attribute is scored on a scale from zero to five, with zero worst and five the best score.

Social scores high on adoption and accessibility due to its integral support for communication to colleagues, which helps spread the system through the organization. It is also very mature as there already exist academic and commercial social visualization tools. Aesthetic scores generally low because it is such a new field, with not that many scientific reports on it yet. This makes it difficult to judge. Aesthetic does have at least one known strength, the integral interactivity, which is connected to flow. Also, because of its focus on attractive and appealing GUIs, it scores some points on accessibility. Gaming scores high on intrinsic motivation, interactivity and flow, which is precisely what actual video games do well, so this is no surprise. It also scores some on extrinsic motivation due to games often being

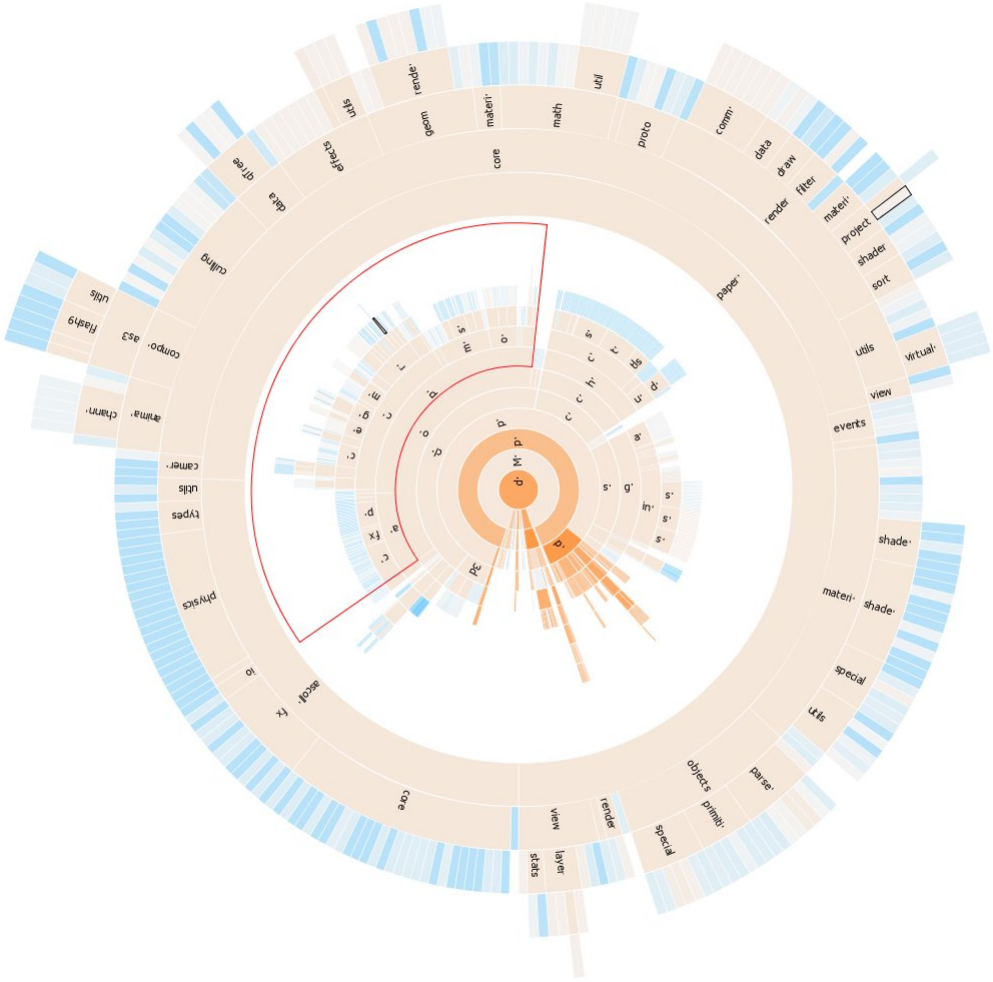



FIGURE 5.5. SunBurst, a circular hierarchical visualization. Amongst ten visualizations regarded as most aesthetically appealing, according to user tests [**5**].

,

played with or against others, and on accessibility, as each game need to be learned from scratch, and generally games succeed in getting gamers to do this.

Evaluating this table, it is immediately visible that the Aesthetics solution is not yet that developed, which results in most attributes lagging behind the other two. Social and Gaming are neck-and-neck, with social connecting more to motivation in general, and gaming connecting to the intrinsic side of it. The major downside of gaming is that it is not that focused on making the BI tool suited for an organization, and focuses on fun instead of accessibility and adoption. Also, there is not a concrete merger of visualization goals and game goals, as was visible in the gamey-visualizations of Diakopoulos [**9**]. This makes a concrete solution direction more difficult to conceive. Therefore we choose socialization as our design direction, which narrows down our initial problem of the introduction. The new problem is formulated in the next section.

## 5.5. Conclusion

With the new context of socialization added, the new problem is stated as follows.

*How can the usage of a BI tool be improved by introducing a social aspect?*

From the interviews and other previous chapters we have distilled some requirements. These are the following:

- **Simple GUI** — Because of the required accessibility and webapp simplicity, the GUI needs to be easy to use.
- **Support short bursts usage** — Deep analysis is valuable, but it is not possible for everyone to do this, due to lack of time or expertise. However, the other way of usage, in short bursts, is something everyone is capable of. Therefore we also look more into the direction of ManyEyes than Aruvi.
- **Incorporate adoption** — If we enable users to reach out easily to non-users, the organisation adopts the BI tool more easy. Also the community becomes bigger that way, and therefore more valuable according to Metcalfe's law [**41**] that says the value of any network is the square of its size.
- **Motivate intrinsically** — The motivation should come from the BI tool itself, as intrinsic motivation works better and stays longer, therefore the community has a better chance to mature itself. Luckily, implementing a social component internalises the extrinsic social motivators, making them intrinsic.
- **Online** — Because of the needed simplicity, type of usage and unrestricted accessibility there is almost no other choice than to make the BI tool available through a webpage.

For this social environment we need to develop some kind of social structure to which On top of that the visualization needs to be bound to the social activities, which is another important aspect of designing a social BI tool. These two steps are taken in the following chapters, starting with the first two.

CHAPTER 6

# Social Structure

The online BI tool should take the society of the organization into account. To that end theories of online societies concerning group decision and group sensemaking making are examined. Also some new aspects of social visualization are noted, namely the hierarchy of the organization and the lifecycle the business information goes through in the process of going from problem to solution, for which our information is typically used. Based on these insights a design decision is made, namely what the structure will be of the communication that the users will have.

## 6.1. Online societies

Because the BI tool is going to be an online social application, information is needed on those online societies. Bartle has made a taxonomy of personality types existing in an online environment [**1**]. In this taxonomy two axes are defined, the acting-interacting axis and the players-world axis. These two axes go respectively from *acting with* to *acting on*, and from *an emphasis on players* to *an emphasis on the environment*. In each of the four quadrants a personality type exists (see Figure 6.1). These are the following:

- **The explorer** is enjoying the world he is offered, playing along with it, discovering what it has to offer. Therefore he *interacts* with the *world*.
- **The achiever** is trying to improve his status in the society, trying to overcome all challenges. Therefore he *acts* on the *world*.
- **The socializer** tries to have fun with his fellow players, talking to them and doing things together. Therefore he *interacts* with the *players*.
- **The killer** tries to defeat other players, besting them regardless of how his victims feel. Therefore he *acts* on the *players*.

A user assumes one of these personality types when he is in this environment. Note that these are roles, and can be changed between or even within sessions of use [**25**]. All these types of behaviour should be taken into account. The killer is the most important to deal with as he is destructive. Therefore, he should have as few power as possible. All the others seem to have a positive effect on the situation, so for the rest it does not matter much which role is assumed.

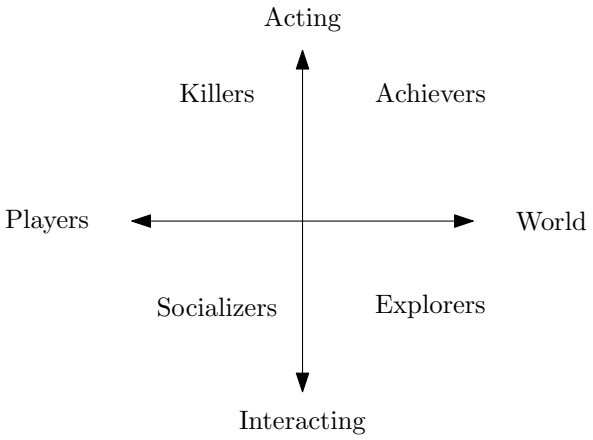



FIGURE 6.1. Bartle's four personality types in online MUDs or Multi-User Dungeons. As dungeon is meant in a very general sense, it can be seen as just any other environment.

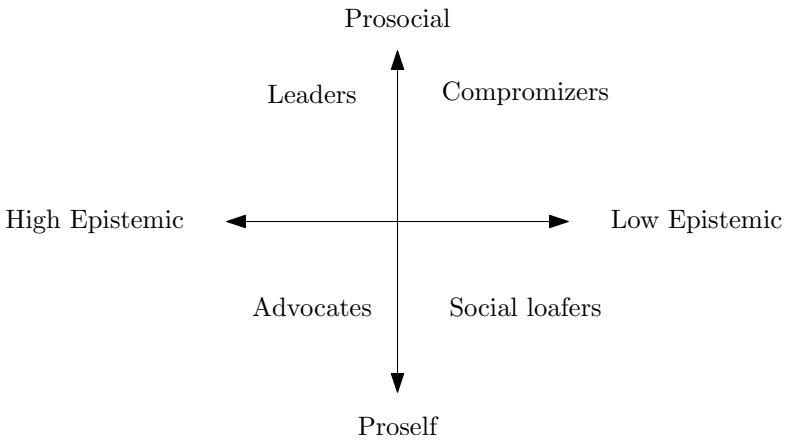



FIGURE 6.2. De Dreu's four personality types of persons involved in group decision making.

### 6.2. Group decision making

Another important aspect is that the BI tool is used to get information for decision making. Placing this in a social environment makes group decision making important. For this too, a taxonomy is made with two axes (see Figure 6.2) [**10**]. Important for group decision making are epistemic motivation and group motivation. Epistemic motivation is motivation to obtain knowledge, group motivation is whether the person is motivated to aid the group or just himself.

- proself, low epistemic people try to stay out of the decision, trying to influence the decision with inflexibility and vetoing. We call these people the social loafers.
- proself, high epistemic people use advocacy, deception and withholding of knowledge to influence the decision in their favour. We call this group the advocates.
- prosocial, low epistemic people are not that bothered about the decision, and try to keep the group together. We call this group the compromisers.
- prosocial, high epistemic people are trying to collaborate with all, trying to solve the problem together. We call this group the leaders.

This last type is the most favourable one for the society of the BI tool, as they try to involve more people and information in the decision making. The proself types are not that bothered about other people's opinions and options, their solutions are more unique, leading to bigger diversity among the used information and therefore better decisions for which the information is used. Therefore they are not as bad as they may seem, and we are OK with their presence.

### 6.3. Group Sensemaking

The group sensemaking process consists of several phases [**16**], with earlier stages more individual and later stages more group-focused. Related to this is that the knowledge involved in this process evolves. This is actually the case for any sensemaking or analysis process in an organization. This related concept of evolution of knowledge in an organization is further explored and explained in section 6.5. Also an important aspect of working together is knowledge transfer. Storytelling can help in this situation [**13**], which means that users can "write stories" with visualizations to explain to other users.

### 6.4. Hierarchy and Organization Communication

Every organization has a hierarchy. Therefore, we already have a social structure. It highly influences communication, as you will not talk the same way you talk to your superior as to your peers or subordinates. It is good to match the structure of the online social environment to this existing offline structure, as it is the true social structure of the organization. Hierarchy is an important aspect of this social structure. There are three different situations for the communication:

- between peers
- superior to subordinate
- subordinate to superior

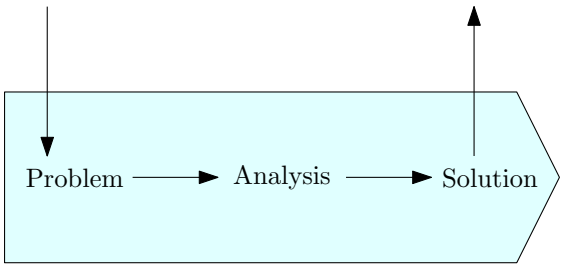



FIGURE 6.3. The life cycle of knowledge. The process starts with a lack of information, then through analysis the problem is understood and a solution is formed, after which it is complete and all necessary information is found and understood.

The first is for example seen during meetings, where people discuss among each other what is the best course of action. The second is like the superior giving an order to the subordinate. The third can be seen as the subordinate reporting about his work. These last two, which are along the lines of hierarchy, are actually kind of similar. In both cases, the communication only does not have much response, there is one sender and one receiver of the message. We therefore find two types of communication, without and with response. This distinction is intuitively not unique for hierarchy situations. With communication between peers or just unrelated different departments in the organization response can also be desirable or not. Therefore, this response seems an important property of communication in an organization in general.

This aforementioned difference between single peers and whole departments communicating gives rise to another important attribute in organization communication. This is audience. When talking with a colleague, an employee usually knows this colleague and both have a certain shared knowledge base that they are aware of. With groups communicating to groups, typically not everybody knows everybody, which makes this shared knowledge base much more difficult to know. Therefore, things need to be explained more proper concerning both correctness and clarity. Therefore, another important aspect is audience size.

### 6.5. Knowledge lifecycle

A BI tool is generally consulted for information whenever there is a decision to make that raises a question about which option to choose. From this question the usage starts, and in the end a decision is made when all necessary information is obtained and understood. Between these points the knowledge of the people involved changes and evolves, and this lifecycle of knowledge is captured in figure 6.3. This shows the development of the knowledge on a topic. The knowledge starts actually from the realization of the lack of information. Someone asks for information to help him make a decision. This lack of information is then the starting point from which people try to find an answer. How this process is actually done depends on the question and person(s) involved, but the main thing is that the answer does not come straightaway, as otherwise there would not really be a evolutionary process, so that is a different situation. This state of finding an answer is characterized by analysis. After this analysis, an answer is reached and the lifecycle is complete.

### 6.6. Existing communication forms

As mentioned before, in corporate communication both existence of response as well as audience size are important. As we try to find the correct communication media for our social platform, let us look how these properties look in the light of the communication itself. With response it is meant whether the medium supports recipients sending feedback along the medium. With audience the size of the group of recipients is meant. The smallest this group can be is one, otherwise there would not be any communication. This group can also be bigger, any arbitrary size, which is typically denoted with $N$. The following communication forms exist:

- **Message** — The message is the most simple form of communication. It consists of a sender sending one message to the receiver. It is basic, and all other forms of communication consists

| | | No response | Response |
|---|---|---|---|
| Audience | 1 | Message: $1 \to 1$ | Question: $1 \rightleftarrows N$ |
| | $N$ | Presentation: $1 \to N$ | Discussion: $N \rightleftarrows N$ |

TABLE 6.1. The four forms of communication. Note that with the discussion form both then $N$'s are actually the same population, it is only due to typesetting they are different.

of different applications of this building block. As it is most basic, a message is sent to an audience of 1, which cannot respond in that message.

- **Presentation** — A presentation is a small story about a certain subject created by a person, about a certain topic. As such, it is a string of several messages. This presentation is then presented or shown to a bigger audience, which might give rise to some questions, but not many as only a few minutes of the whole presentation time are reserved for this. Therefore a presentation is sent to an audience of $N$, which cannot respond.
- **Discussion** — In a discussion multiple people communicate with each other regarding a certain topic. They send messages to everyone involved telling what they think about the topic. Therefore a discussion has an audience of $N$, that all respond to each other.
- **Question** — A question is something asked by a person to one or more others, which should be eventually answered to the asker. As such, the question has an audience of one. Also, the questioned people do deliver feedback in the form of the answer, and can discuss among themselves about the answer, which means there is the possibility for response.

See Table 6.1 for the communication forms ordered according to the two axes. These axes are quite neutral. To get response is nice, but the communication forms that have these are not complete without this responds, which limits their value as people need to input the effort of giving response before the communication form becomes valuable. Also if the audience is bigger, the information in the communication form is accessible by more, however the communicator needs to take its audience into account, and a bigger audience is more difficult to cater to.

### 6.7. Comparison

As done with the general design decision, the four options are again scored on certain properties, which is done by rough estimation on a scale from zero to five stars. These attributes are:

- **Long term value** — Web applications are known for their search functionality, which is easy to use, useful and it allows the system to retain value. Therefore the searchables should be able to retain their value so this search functionality can be tapped into.
- **Group focus** — We are trying to make something social, so a structure that connects better to the whole society is better.
- **Expressivity** — As users try to express themselves in the system, expressivity is important.
- **Simplicity** — As repeated already often, the application needs to be simple for the users. Also simple structures are better to predict, which makes designing a program for them easier.
- **Collaboration** — group sensemaking is good, so structures that support this are better.

See Table 6.2 for the comparison table. We now give a motivation on the placement of the stars in this table. A question is dynamic, it starts being asked, then slowly an answer is formed after which it becomes answered. Because of this dynamics a question is not simple. Long term value for messages and questions is low, as they are aimed at an audience of one, and therefore it is doubtful if in the long run, after they have been actively created, they are still useful for people outside of this small audience of one. On the other side, people are able to collaborate with active questions, namely when forming the answers, which means there is a group focus and good possibility for group sensemaking. There are also some practical remarks. For instance, a message is very basic, and very similar to email communication. As people already use email it is better to not reinvent the wheel and focus on wholly new tools more powerful and suited to the situation. On the other side, it is rather difficult to predict the success of the complex question. No comparison can be made with already existing social (visualization) platforms and its state is dynamic (going from raising the question towards answer formed). It also requires good

| Solution | Long term value | Group focus | Expressivity | Simplicity | Collaboration | # |
|---|---|---|---|---|---|---|
| Message | ⋆ ⋆ | ⋆ | ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ ⋆ | ⋆ | 1 |
| Discussion | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | 4 |
| Presentation | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ ⋆ | ⋆ | 3 |
| Question | ⋆ ⋆ | ⋆ ⋆ ⋆ ⋆ | ⋆ ⋆ | ⋆ | ⋆ ⋆ ⋆ | 1 |

TABLE 6.2. Comparison of Message, Discussion, Presentation and question communication forms. The final column counts the number of wins on the attributes before, to give an estimate of suitability.

moderation tools including support to designate groups, to prevent killers, advocates and social loafers from doing things they should not do.

The evaluation of the four shows that the discussion and presentation are a better choice than the message and question. We opt to choose both of them, as they allow for two interesting forms of communication, more or less complementing each other in the different phases of the community sensemaking and lifecycle of knowledge processes. It also fits with the distinction often made between presentational and exploratory information visualization [**8**].

CHAPTER 7

# Binding the Visualization

As we have chosen discussions and presentations as communication forms, we need to decide how to link or bind visualizations to these. Binding the visualization has two sides, as the bound must be determined at the visualization side as well as the communication side. On the visualization side we need to determine what aspects are actually "bound", which properties will be saved in the system as belonging to the particular bind; on the discussion side, we need to determine to which level of the discussion (the comment level or discussion level) such a bind binds.

## 7.1. Snapshots

The common technique to bind a visualization to another object, such as a discussion, is by means of a so called snapshot. This snapshot is a saved version of the visualization state on a certain moment, which means that upon loading the snapshot, the saved version of the visualization is returned to the user. The information stored can range from a few user settings to practically the entire visualization, including sometimes even the dataset it visualizes as it is part of the state. Logically, this places the "create snapshot" action after the creation of the visualization [**31**], so this general technique also works in the case of MagnaView WebLive users.

These saved snapshots are often used in journalistic articles or other communication, as they are good in telling a story [**40**]. A certain state of the visualization is grabbed that shows something remarkable that the user can see immediately, so without tweaking the visualization towards this desired state. Also, the story in which the snapshots are used contextualizes the visualization, giving its usage a purpose or goal. This helps and focuses the otherwise overwhelmed user [**8**].

**7.1.1. Example.** To illustrate the choices to be made in this chapter it helps to look at ManyEyes, as its designers needed to make the same choices. There are actually two binds in ManyEyes. The first one appears when a user makes a visualization of his dataset. The whole visualization is saved (including which dataset it visualizes), which starts a new discussion attached to the visualization. The second bind occurs when a user is posting a comment in such a visualization discussion. He can opt to save the current state along with his comment, which saves settings such as filters, highlighted visualization parts and sorting. Therefore the visualization interface state is bound to the comment.

## 7.2. The visualization side

For the visualization side of the bind a decision must be made what level of information about the visualization is stored in the snapshot. The more in the snapshot, the more interactive and true to the state when saved it is. However, it also depends on what is technically possible or feasible, for example, if the whole dataset must be saved, this has some repercussions. In MagnaView, datasets can be the size of one Gigabyte uncompressed. Compressed it becomes five to twenty Megabytes, but it is still too big as the aim is to use the social component, and therefore snapshots, intensively.

**7.2.1. Options.** Within MagnaView there are various stages at which the visualization can be saved. These are listed below:

(1) **Image** - Saving the image of the visualization as the snapshot maker sees it makes sure that the snapshot viewer will see exactly that. This has the big benefit of the viewer seeing the aspect of the visualization that the snapshot maker wanted to. The big drawback however is that all interactivity disappears.
(2) **Filter state** - In MagnaView, the most visualization tweaking happens by using filters. These are GUI components such as drop-down boxes, checkbox lists and the like, with which the user can select which datarecords to use in the visualization according to their properties.

(3) **Hidden tiles state** - This GUI action is hiding selected data. A user can freeform select visualized datarecords in the visualization, and this selection can then be hidden.

**7.2.2. Choice.** Ideally, we would save all of the internal state, so both the filter state and hidden tiles state, as when we have those, we can completely construct the visualization as it was. However, saving the hidden tiles state has a big problem, namely that it is freeform, which has as result that its state is defined by which individual datarecords are hidden. This can be saved by storing the identifiers of those datarecords, something which intuitively most tables have.

Sadly, this intuition is not true. Because MagnaView works on complex data from big databases or data warehouses, the result from the data extraction query is not that simple. Therefore, it becomes unclear which column or columns from the result form a primary key. Therefore a primary key must be defined for the dataset of the visualization, making the creation of a visualization more difficult.

When these primary keys are taken along with the snapshot, the snapshot also becomes larger. As the data is stored internally via keeping track of the unique values in the columns, a significant amount of the 5 MB mentioned contains the big list of identifiers/primary keys, as they have the most unique values. This would make the snapshot quite big.

The biggest problem is changes in the data, however. If between the moments of saving the snapshot and loading the snapshot a lot of data was added or removed to the dataset, the visualization still looks very different between those moments. Therefore, the use of taking these identifiers along with the snapshot does not have that much benefit. Adding that with the drawbacks of a big size and extra difficulty when making the visualization, we decide to not put the hidden tiles state in the snapshot.

This leaves the filter state and the image as possibilities. The filter state has as drawback that it does not contain the selection, therefore the looks of the visualization upon loading the snapshot could still be quite different than when the snapshot was made. In that case the viewer does not see what the snapshot maker intended, and the use of the snapshot is lost. Taking only the picture makes the snapshot no longer interactive. However, as remarked earlier, the end users need this interactivity. Also, it removes much of the attractiveness of the snapshots, which would otherwise motivate the users. Therefore we store both the image and the filter settings in snapshots.

### 7.3. The discussion side

As the ManyEyes example shows, on the discussion side it matters on which entity in the social universe the snapshot binds. There are three options, which are all present in ManyEyes, with two of them already mentioned.

- **No binding** - This is what is seen on standard internet forums, there is no binding with other objects, just text. ManyEyes shows this is also an option, as it has discussions completely separate from the visualizations. The visualizations then exist separate from or next to the discussions. With the screens of discussion and visualization next to each other, it could be a desirable option.
- **Binding to a discussion** - This is like the first ManyEyes bind, with a complete discussion binding to the visualization. As noticeable in ManyEyes, this results in discussions with the visualization as its main topic. A similar model can be seen with YouTube, with a discussion in the comments below the movie, which is the main attraction of the site.
- **Binding to a comment** - As mentioned before, In ManyEyes this is the settings snapshot each user can add to his comment. An internet equivalent would be for example FaceBook, where links to other media such as images or movies can be inserted in all types of comments. The same happens on online imageboards, where users can add an image to each of their comments.

**7.3.1. Choice.** Due to benefits such as the attractiveness of the multimedial snapshots, attempting to make a homogenous GUI experience and preference of actual social infovis, choosing no binding does not compete with the other two options. This leaves only the other two.

The YouTube model places the focus on a single visualization instead of the discussion. This reduces the beneficial effect the context of the discussion gives, as it removes focus and attention from this social context. However, enabling people to add other visualizations to their reactions allows for more flexibility and therefore more sides of a discussion to be better argumented, improving nuance, analysis and the

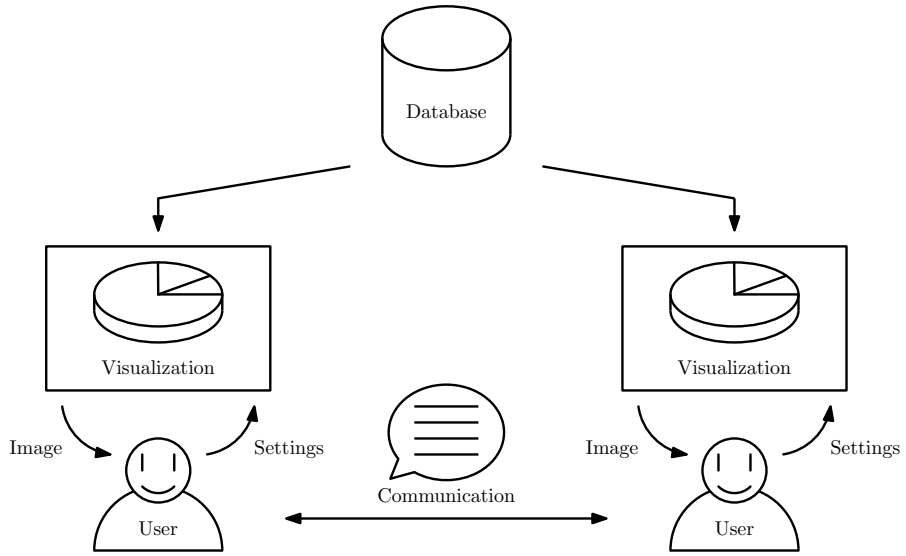



FIGURE 7.1. BI tool usage in an organisation. On the arrows objects that are communicated are noted, like the image and settings of a visualization.

decision making [**10**]. Therefore we add this to the snapshot, which only shows its value on the comment level, which thus determines our choice for the comment bind.

### 7.4. Information Chunk

The design decisions taken in this chapter lead us to define an elementary information chunk as the following:

- A comment, as this is the level on which the bind works.
- An image, which is exactly how the visualization looks like upon making the snapshot.
- The visualization state and current visualization, to allow further interactive analysis.

When looking at the situation in which the social extension works (see Figure 7.1), these three elements correspond with the communication arrows a user has, with communication generalized as the sending of messages. Apparently, the elementary information chunk that was created seems to cover all the interaction the user has with the rest of the system, which now has become digital and therefore more easily shareable with his colleagues. Also, with this information chunk as a building block, the user can create discussions with others or make presentations for them, allowing communication in all kinds of ways. Users do not always want to communicate with images. Therefore the snapshot part of the information chunk is optional.

This design seems promising, which makes the next step the testing of this design by means of a prototype.

CHAPTER 8

# Prototype

Vanilla MagnaView (so without the designed extension) as accessible from the web is called MagnaView Weblive, or WebLive in short. To make discussion easier, the social extension is called InfoCloud. The etymology is as follows. Technically, it does not have a cloud design, as it only uses one server. However, for the user it feels that discussions and presentations can be picked out of a big cloud containing these, ready to use. On top of that the cloud is a current IT buzzword gaining a lot of attention, so everything needs to be cloud anyways.

First WebLive is described in short to give a context in which the extension works. Next the InfoCloud extension itself is explained with all important features illustrated. A global description of the architecture and implementation is given in appendix A.

## 8.1. WebLive

See Figure 8.1 for an image of vanilla WebLive when started by a user. The left of the top bar is dedicated to navigate through a MagnaView project, consisting of multiple visualizations. There are also buttons to hide or unhide selected nodes in the visible visualization. The top right of the bar is dedicated to auxiliary functionality, such as downloading a related Excel file or image of the current view. The main part of the screen is dedicated to the visualization, with a sidebar containing related information, such as a description, legend for colors or icons or the information of the datarecord the mouse cursor is currently hovering on. It also has space for the filters, which allows the user to easily select which data he wants to be visible or hidden according to the values of certain fields. These simple actions together allow powerful manipulation of the data, such that the user can easily look at his big dataset.

## 8.2. InfoCloud

With InfoCloud present, a little bit of space on the left side of the WebLive screen is given to InfoCloud, see Figure 8.2 for a complete picture. This could make the screen slightly complex for the user, as two screens are more complex than one. However this has the great advantage that the visualization is always visible while reading a discussion or presentation. This is not possible in ManyEyes, which has the visualization on the top of the page and the discussion unfolding below it. This design of InfoCloud allows the comment to be read while the live visualization of the snapshot belonging to it is also visible, which is much easier than scrolling up and down the web page to get the idea across as needed with the ManyEyes design.

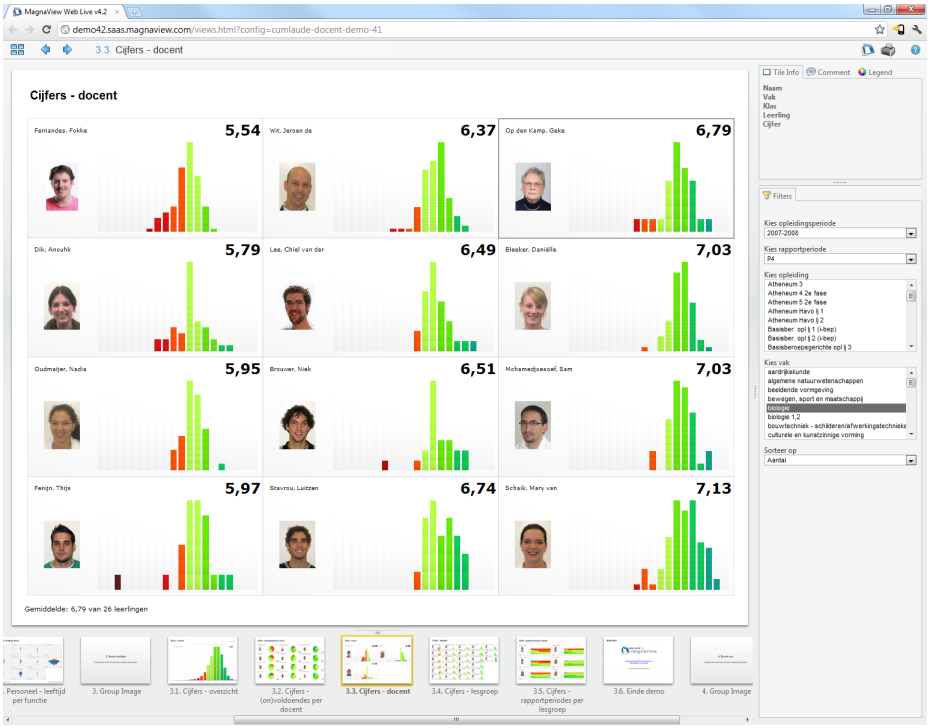



FIGURE 8.1. WebLive without the InfoCloud extension. This particular image is taken from the freely available demo, `http://demo42.saas.magnaview.com/views.html?config=cumlaude-docent-demo-41`

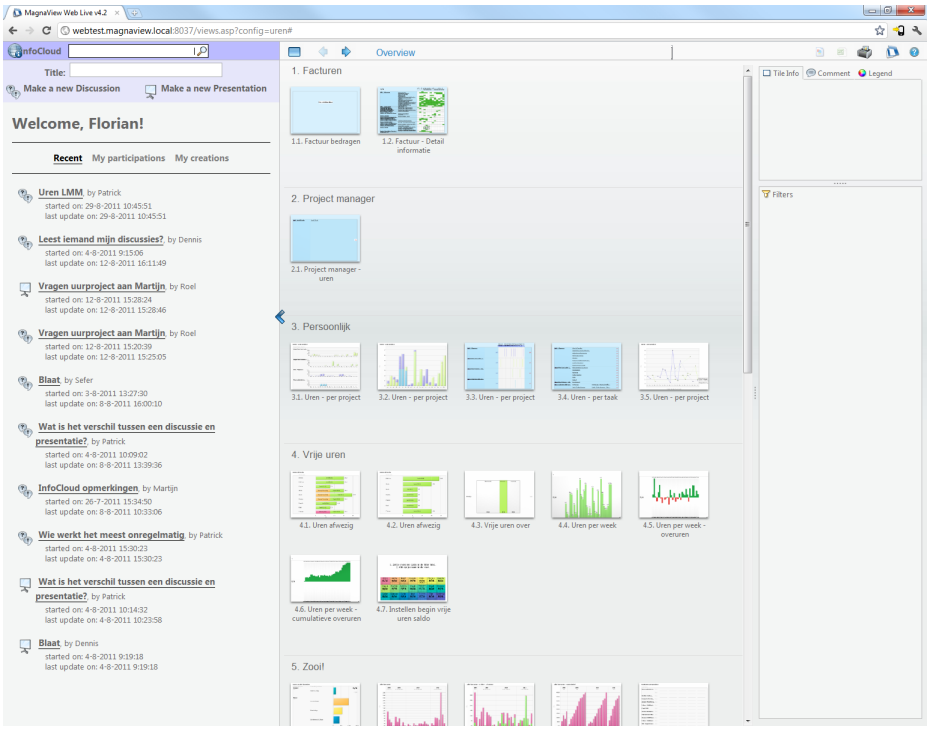



FIGURE 8.2. InfoCloud running next to WebLive. Although the subdivided screen makes them look as two separate entities, InfoCloud is an extension within WebLive

Figure 8.2 shows the screen of InfoCloud just after logging in, the homescreen. The user can immediately start to create a new discussion or presentation by filling in a title for it and clicking on the respective buttons. There are also three lists:

- A top 10 of the most recently changed discussions or presentations, to make a user's actions more visible and to give an overview of what is going on, which motivates [**18**].
- All the discussions and presentations the user has participated in.
- All the discussions and presentations the user has started.

Initially, the first option is shown, but a click on the name of another list shows that one in its place. The last two options give easy overviews of discussions and presentations important to the user.

## 8.3. Discussion and Presentation

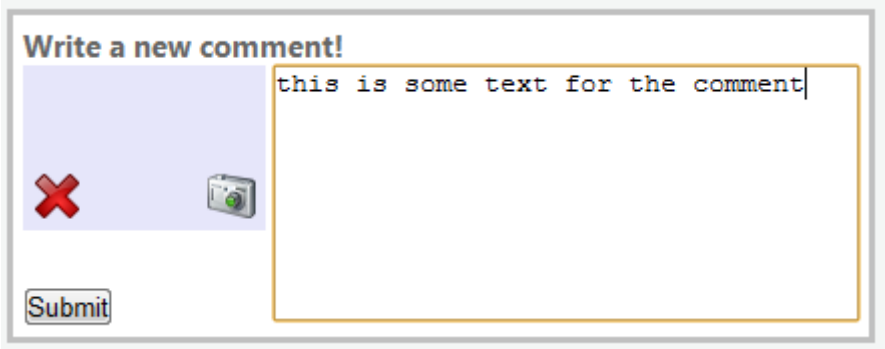


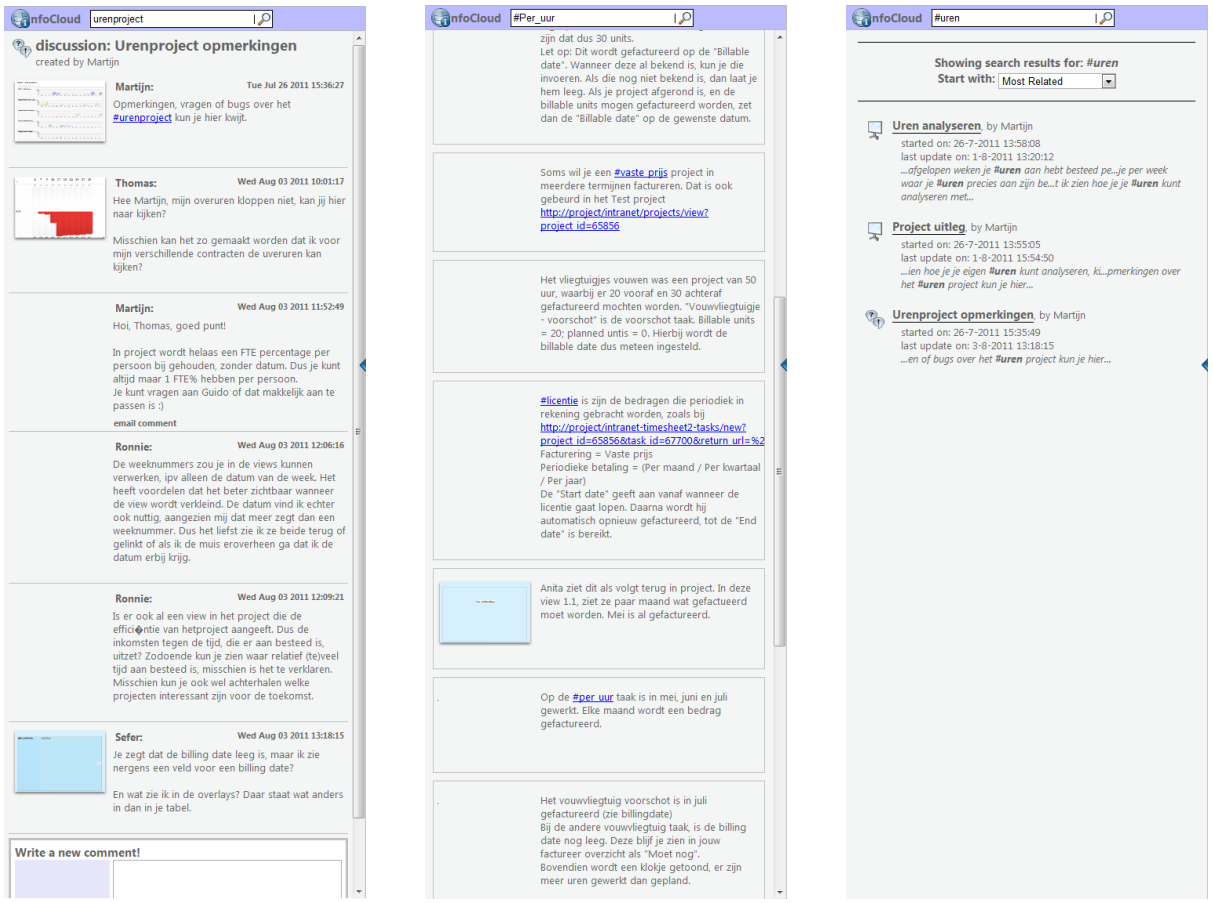

FIGURE 8.3. The new comment panel, with the icons to add or remove a snapshot visible. They are only visible when hovering over the panel to prevent a cluttered GUI and a confused user.

When clicking on a discussion or presentation in the home screen, or when starting a new one, users see that respective discussion or presentation. They have overlap, as both are a list of information chunks as described by the previous chapter. The difference is in their appearance, and allowed actions, see Figure 8.4. The panel to write a new comment, if possible, is always at the bottom of the discussion or presentation, as that is where the new comment would be. It is styled as a draft comment, with also a space for the thumbnail for the snapshot if the user wants to add one. Initially, this space is grey, with icons to add (a camera) or remove (a red cross) the snapshot appearing when hovering over the draft comment. When clicking the camera, InfoCloud takes a snapshot and shows its thumbnail in the gray space. When users have second thoughts about adding a snapshot, they can click the red cross, which removes the snapshot again, showing only the gray substitute for the thumbnail.

**8.3.1. Discussion.** See Figure 8.4a for a picture of a discussion in InfoCloud. Every user is allowed to post a comment in any discussion. The comments are displayed according to their posting date, which is also shown with the comment itself.

Users can still edit their comments, in case they want to improve their comment. This is because a typing mistake is made easily and this is standard functionality for online internet forums, which are very similar to these discussions. However, that this post was edited is again visible, to make sure that discussion vandalism does not get an upper hand. A vandal (a killer from Bartle's taxonomy [**1**], see section 6.1) could for example post an offensive comment, inciting spiteful replies from other users, and then edit out his offence. Readers can see in that case that it is edited, with which they know that there is something suspicious with that edited comment.

**8.3.2. Presentation.** See Figure 8.4b. A presentation is made and controlled by one user. As such, the presentation is more open to editing, which is no longer restricted to just the text of the comments. For example, the order of the comments can be changed by dragging and dropping these in the right place. The creator can also remove a comment if it is no longer necessary. As such, users can shape their presentations to their liking, allowing editing in almost all aspects.




(A) A discussion (B) A presentation (C) A searchscreen

FIGURE 8.4. Three screens of InfoCloud. Note that with each of the three screenshots the WebLive part has been cut out, to save space.

## 8.4. Comment actions

Both authors and viewers can perform actions with the comments. The two most important actions are *view image* and *load live*. View image shows the saved image of the snapshot, superimposed over WebLive, along with the text of the comment, see Figure 8.6. Load live loads the visualization state in the snapshot into WebLive itself, resulting in the live visualization (hence the name), which looks as much as the saved snapshot as possible, see Section 7.2.


FIGURE 8.5. A detail of Figure 8.6, the comment with thumbnail of the snapshot. Also the actions on the comment are visible, which are to view the image of the snapshot, load it live, or start an email with URL to this comment.


FIGURE 8.6. After clicking on view image action on the comment, see Figure 8.5, the image of the snapshot becomes visible, superimposed over WebLive, with the text of the comment below the image

Next to those actions, it is also possible to email a link to that comment, where browsing to that link results in a view as if already clicked on view image. Clicking on that button starts a draft email with that URL inside, already properly filled with text, ready to be finished with receivers and then sent. This allows the prosocial high epistemic motivated people mentioned in Section 6.2 to contact their colleagues easier, tapping into their motivation to spread the information and the tool usage across the organization. To promote the tool even more in the organization, InfoCloud also sends a similar email fully automatically to all participants of discussions (who posted comments in it) when a new comment is posted in a discussion. It also makes the system more usable, as the participants do not need to constantly poll the discussion for new comments.

## 8.5. Navigation

Even though they are both visible at the same time, InfoCloud and WebLive can be browsable independently from each other. This allows users to analyse other aspects of the dataset when viewing certain discussions with snapshots, improving group sensemaking.

**8.5.1. URLs.** As shown by the URLs that can be generated for each comment with the email button, the navigation state of InfoCloud is stored in the URL of the browser. This allows users to navigate using the browser's back button, and giving the users a history of what they have browsed through so far. More tech savvy users can use this by bookmarking certain comments or their containing discussions with the browser's bookmark system. This means that InfoCloud is well integrated with the browser's known capabilities, which is convenient and easy for layman users, and even allows browser plugins that use these aspects, such as bookmark sharing across computers, to integrate with InfoCloud, even further extending capabilities.

**8.5.2. Search.** As the number of discussions or presentations can become large, navigation though this big collection is a practical problem at hand, which should be solved correctly. An important aspect is that this navigation should be simple, as the users are not tech savvy. Therefore a Google-like keyword search is featured, which is always visible at the top of InfoCloud, see Figure 8.4c. The results (presentations or discussions containing those keywords) are then returned sorted on relevancy (number

of hits with the keywords), but can also be sorted on other properties such as start date or last date changed.

**8.5.3. Hashtags.** Another search technique that is supported is the use of hashtags. In a comment a user can type a hashtag (a string of characters without spaces, started with a hash, # ) to signify a topic. When this comment is posted, this hashtag becomes a clickable link, which directs the user to the search on that hashtag (see Figure 8.4c for an example), thereby showing all discussions or presentations also containing that hashtag. This allows users to easily mark and group discussions and presentation according to topics, making navigation easier for themselves and others.

A faceted search has also been considered to improve navigation, which is more capable offiltering messages by various attributes, thereby composing more complex queries. However, a query that cannot be answered with keywords search on its content but still on other properties is pretty vague, like for example " all the discussions in which I replied a month ago to John, but did not get a reply yet and I forgot what it was about". A GUI capable of describing such a query would be quite complex, and therefore too difficult to use for the typical user. Also, it is unlikely for a discussion to be important, but not important enough that a user forgot its content. On top of that, important discussions are likely discussions that the user created or participated in, which can already be found on the home screen. Therefore a faceted search is expected to be used only on rare occasions, as usually discussions are important by their content, which is how they are remembered by the user, and can be searched on by the search box.

CHAPTER 9

# User tests

To evaluate the prototype, we have carried out a small user test.

## 9.1. Test conditions

The user test is performed internally at MagnaView BV. Of course a test with the teachers of the interview would have been better, but placing a prototype at customers is not possible, therefore this was the only option. The test is performed at the office of MagnaView in Eindhoven, in between normal work. Because the prototype is a web app no software needed to be installed, which allowed testing to be literally in the normal situation, the normal working desk of the participant.

There were seven participants. All of them were employees from MagnaView, male and had finished a technical education. Six of the seven were standard employees which mainly need to manage their own work and one of them was the manager. Also, they were not quite used to structurally using a BI tool for their own work.

The project loaded into WebLive is the so called *"uren-project"*, which is a MagnaView project on the hours administration of MagnaView, hence the name. Three participants said that they were not that familiar with the dataset or usage of WebLive, somewhat mitigating the influence of the technical background. The data visualized is also actually the live data obtained from the database, which means that the users were really testing on their own data, they had a relation to it and it had value for them.

## 9.2. Test description

The test was started by sending the participants an email with four short tasks, the first and most important two consisted of reading specific InfoCloud presentations that would explain InfoCloud and the *uren project* respectively. After this they were simply told to use the system by giving opinions on InfoCloud and the *uren project*, and finally they were suggested to use the system further.

During this activity the test examiner took notes of the usage. The main focus was amongst other looking if people liked the new social visualization concept of presentation, how easy they picked up the tool without help (as the instruction email was rather short, there was no real help), how they responded to the social context the MagnaView visualizations were newly placed in, and how much they enjoyed using the tool.

## 9.3. Results

As the prototype has been under limited development, the test results showed multiple bugs. There were also several interesting results. These are listed below.

- **Discussions** were a concept easily understood, showing their low entry level as expected.
- **Presentations** gave some difficulty to understand, although the users were following the instruction presentations easily. The main issue was that users desired to reply on presentations.
- The **show image** action of a snapshot was well appreciated. They responded audibly content when they were able to exactly see what the comment was talking about, right next to the comment. They also liked the slideshow feel of it.
- Also using **load live** was appreciated. This could have several reasons, but the enjoyment was visible.
- The interface of **two independent windows** works. The first few minutes showed a little getting used to the set up, as the users indeed needed to get acquainted with two windows at once. Afterwards the users were used to it and using either part of the screen separately went without effort, as well as performing the load live action of the snapshot.
- The **hashtags** were understandable for the users, the search results that were shown after clicking on them were not unexpected by the users.

- The **enjoyment** of the users was noticeable with the social and creative aspects of GUI, like making comments, and the viewing images and loading live parts of the snapshots.
- One user mentioned to expect some life cycle of some sort for the discussion or presentation objects. He would like an object that more or less followed the start of the problem until it was solved, which matches the **life cycle of knowledge**.
- Another user made a positive comparison to an earlier combination of visualizations and social interaction that he once used, and he liked this version definitely more.
- The resulting discussions and presentations after the user tests showed that the users also used the system for basic banter to each other, showing that they intrinsically liked this social aspect [**20**].
- A final nice result from the prototype is that it already has proven its value. One user helped another user with finding two days worth of holiday hours, of which he was previously unaware!

**9.3.1. Prolonged usage results.** The prototype was also accessible to the participants after the actual test for a prolonged period of time, about a month. This was mainly done to see if the users would have a tendency to return to the prototype. Therefore, to keep it natural and not influence the results, no effort was done to stimulate return to the prototype.

A few people indeed did return to the prototype. Although it were not many, there were only seven participants to begin with, of which two to three returned. Given the situation and the timespan this occurred in, it is seen as a not too bad number.

In all, during the user test, all of the seven participated made about 40 snapshots and about 100 comments in 15 discussions and/or presentations. almost all of these were during the initial round of user tests, with a few comments and snapshots made outside of this. The number of logins and emails sent with the system was not stored in the database, which makes it more difficult to estimate. Of this the emails is most interesting, which were roughly 10 and almost all were sent by the same person, therefore we expect this person to be a leader in the classification of De Dreu [**10**], see Section 6.2.

CHAPTER 10

# Conclusion

This report has as its outset the following question:

*How can the usage of a BI tool be improved?*

After finding a direction for the design of the solution, this design has been further developed, a prototype has been build and user tests were conducted, with promising results. From this it can be deduced that the usage of a BI tool in an organization can be improved by adding a social aspect, as this gives fun, context, guidance and communication to the user, which improve the usability of the tool and the motivation users get. Another benefit is that the social aspect allows for better group sensemaking of the data. A minor drawback is that this causes an attention shift, as the social aspect seems already very attractive in itself, which could reduce the attention to the visualization and data itself.

After this first question, the report was set out to answer how the usage of a BI tool can be improved by introducing a social aspect. In this context, design decisions have been made that introduced the concepts of discussions, presentation and the bind between visualization and social platform on a comment level. Besides in the general use context as seen in ManyEyes, discussions also show merit in an organization. This is because they allow asynchronous non co-located communication about the visualizations, which allows more people from different locations and time zones to collaborate. This also allows users to communicate in their own time, location and pace, not needing to adapt to the group.

Presentations also showed merit, as their expressive and controlling capabilities exceed that of a discussion. They therefore allow for better storytelling about the visualizations and data. However, they do allow less social interaction compared to discussions.

Placing the bind of visualization and social platform at the comment level proved its worth, allowing each comment to use snapshots. Snapshots showed to be powerful and precise tools of expression, reference and argumentation when discussing the visualizations and data they contained.

## 10.1. Evaluation of requirements

In chapter two and five we have developed some requirements that we wanted our adaptation to fulfil. It is interesting to see how much we have succeeded in doing so.

**10.1.1. Interview requirements.** The interview requirements came as a result from performing the interviews, described in Chapter two.

- **Web app complexity** — as MagnaView WebLive was the starting point of our adaptation, the result is still an actual web app. However, because of the two subscreens in the system, one for InfoCloud and one for WebLive, the tool was not that simple anymore, which showed during the user tests.
- **Teamwork and sharing** — Because the general design solution is social the system is very suitable for teamwork and sharing. The snapshots even allow sharing with actual visualizations.
- **Pick up and use** — Because the navigation is stored in the URLs, the system is easily opened navigated to a certain part, ready to use.
- **Fun to use** — Users said explicitly that they liked to use the system, with both the social side as well as the interactivity from the snapshot noted as fun.

**10.1.2. Theoretical requirements.** The theoretical requirements come from the theory overview from Chapters three and four. They were used to compare social, games and aesthetics as design direction in Chapter five, specifically in Table 5.1

- **Intrinsic motivation** — As noted in the interview requirements, the participants really experienced the prototype as fun.

- **Extrinsic motivation** — As the social aspect of the system allows others to easily coerce the user towards using the system, this seems as covered. The others can also easily email URLs that link to and within the system, making this push even stronger.
- **Interactivity** — The snapshots provided nice extra interactivity, but for the rest no new special interactive components were introduced by InfoCloud.
- **Flow** — The prototype was not engineered towards supplying flow, and the user tests did not show that a state of flow was reached by the participants.
- **Adoption** — The system sends emails to participants of a discussion on a new comment, and users can email from within the system. Therefore it seems as if the system has apt tools to promote itself through an organization.
- **Accessibility** — The two subscreens made the prototype slightly less accessible, but the participants of the user test showed to be quite capable to use the system,

### 10.2. Prototype improvements

As the prototype was developed under time and effort limitations, the user tests of course showed several bugs that could be improved. Besides that the user tests also showed more meaningful aspects of improvement.

As the prototype is an example of a web 2.0 application, it could deliver an even richer experience by adding hotkeys, loading rings and the like. This gives more feedback to the user, making the application more usable.

The users liked to use the slideshow aspect of viewing the image of a snapshot. Therefore, this aspect should be augmented, by making clicks or keypresses go to the next slide of the presentation. As this would also accentuate the general presentation aspect of a InfoCloud presentation, this makes it easier to distinguish presentations from discussions. On the other side this feature could also be added to discussions, as it allows easy viewing of both presentations and discussions according to the user tests. If a reply feature would then be added to presentations (for example by linking a feedback discussion to it), the urge to reply on presentations would also be satisfied.

The addition of hashtags was very convenient for users. This concept could be extended by adding easy ways to link to anything within InfoCloud. Going even further, allowing easy ways to add certain markup to comments allows for even more flexibility and expression, for example when displaying lists, adding in images or even uploading whole images.

Right now in MagnaView, and therefore cascading to WebLive and InfoCloud, one project only has one dataset. However, an organization often has multiple important datasets, like a school having absence and grading data. The user can use the same InfoCloud application for multiple projects, and view the images of the snapshots without problem, no matter which project they are from. However, this is not the case for load live, as in that case the other project should be loaded that conatins the visualization. This breaks user flow. Therefore load live could support multiple projects at the same time, making WebLive switch to the correct MagnaView project when necessary.

Also, the social component can benefit from support for groups and a moderation and revision system. This would allow multiple people to work on the same presentation (or other object) more collaboratively, as potential bad changes can be reversed with the moderation and revision system (like for example present in Wikipedia). Such systems also make the more difficult questions from Chapter 6 a more viable option, as the unknown factors that could be negative can be kept under better control, because all actions can be moderated.

### 10.3. Future research

As one of the participants to the user tests noted, the use of a BI tool and the related knowledge has a certain lifecycle: there is some rise of a problem, then a solution is found, which then solves the issue. This seems kind of similar to the dynamics of the question communication object. Therefore, instead of providing several social objects that could be used to follow this lifecycle, one type of object that can evolve along with this knowledge and problem solving could be easier for the users.

This prototype created a social environment within the BI tool. However this environment may as well exist outside of the tool [**51, 7**]. Many organizations use online platforms such as Sharepoint or WebSphere to manage their information, documents and content, often allowing communication about it. If the BI tool could be plugged into this type of platforms, the organization would have similar online access. However, as the content is then served on a platform out of control of the initial BI tool, which makes security, screen use, and saving snapshots become more complex.

This report explored on taking the aspects of an organization into account when developing a BI tool. This business side could be taken into account even more. For example, you do not say the same things the same way to your colleagues as to your boss and also, not all discussions are relevant for all people Therefore, the inherent groups and hierarchy of a company could be taken more into consideration, allowing for better community support [**49, 2**].

Snapshot are a nice construct to point others at certain aspects of the dataset and MagnaView project. This could be taken even further by supporting annotation of the snapshots. This annotation, such as drawing arrows or rectangles over the snapshot, or placing text on it, allows for even more expression. Snapshots also show which parts of the dataset and MagnaView project are so important to users that they want to share it. This is valuable information for both MagnaView and the company that uses it, which makes it a sensible source for datamining. Also, this could be visualized in some way, to show highly discussed parts of the data and visualizations, and parts that are unexplored.

As discussed before, a social situation is a great source of motivation. One great social motivator is social approval. The system does support approval implicitly via comments. This social approval can be made more explicit and visible, for example with like buttons found in FaceBook, improving motivation [**6, 21**].

People are also motivated by goals and challenges that are more specific [**21, 2**]. Even though this contradicts with the typically orientational analysis of visualization, these specific goals and challenges would allow focusing of user attention on certain known problem areas. Therefore, a solution that aims to motivate in this way would be very interesting. Note that this does not need to happen in a social environment per se.

In the context of information contribution to a community, two motivators that have been left relatively unexplored are unique contributions and the greater good for society [**2, 4**]. This is because they are rather difficult to wield, as they require some good to contribute to, as well as a certain uniqueness metric to be derived from contributions and a clear definition of the greater good that the organization has as a goal. As organizations are typically companies, that greater "good" is company profit which is not a greater good for society.

This report focused on improving a BI tool with socialization, however the other general design decisions considered in Chapter 5 can still be explored. Gaming is good in grabbing the attention of users during usage. Even though this removes the focus from the dataset, it could be used to introduce the WebLive environment as well as the basics of the dataset to new users. As games are a fun way to support this, this lowers the mental threshold of usage [**3**]. In a more general sense, as games include many motivators in a strong way, such as competition, fun (even augmented by flow), creativity, social interaction, and mastery, game concepts can be used to increase the motivation of usage [**21, 50**].

Also aesthetics can be employed, as its increase in interactivity in analysis is fun, gives satisfaction and feels creative. Although not available for end users, the MagnaView Designer, used to design the visualization projects, allows for interaction in many interactive and aesthetically appealing ways, such as using dragging and dropping to change color or size attributes, or even completely changing the tree layout of the generalized treemap. All this interaction allows for better manipulation and therefore deeper analysis, so it would help the end users to give them similar interaction possibilities.

# Bibliography


[1] R. Bartle. Hearts, Clubs, Diamonds, Spades: Players who suit MUDs. 1996.

[2] G. Beenen, K. Ling, X. Wang, K. Chang, D. Frankowski, P. Resnick, and R. E. Kraut. Using social psychology to motivate contributions to online communities. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, CSCW '04, pages 212–221, New York, NY, USA, 2004. ACM.

[3] F. Bellotti, R. Berta, A. D. Gloria, and L. Primavera. Enhancing the educational value of video games. *Comput. Entertain.*, 7:23:1–23:18, June 2009.

[4] Y. Benkler. Coase's penguin, or, linux and "the nature of the firm". *The Yale Law Journal*, 112(3):369, Dec. 2002.

[5] N. Cawthon and A. Moere. The effect of aesthetic on the usability of data visualization. In *Information Visualization, 2007. IV '07. 11th International Conference*, pages 637 –648, 2007.

[6] C. Cheshire. Selective incentives and generalized information exchange. *Social Psychology Quarterly*, 70(1):82–100, Mar. 2007.

[7] C. M. Danis, F. B. Viegas, M. Wattenberg, and J. Kriss. Your place or mine?: visualization as a community component. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 275–284, New York, NY, USA, 2008. ACM.

[8] M. Danziger. *Information visualization for the people.* PhD thesis, Massachusetts Institute of Technology. Dept. of Comparative Media Studies., 2008.

[9] N. Diakopoulos, F. Kivran-Swaine, and M. Naaman. Playable data: characterizing the design space of game-y infographics. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pages 1717–1726, New York, NY, USA, 2011. ACM.

[10] C. K. W. D. Dreu, B. A. Nijstad, and D. van Knippenberg. Motivated information processing in group judgment and decision making. *Personality and Social Psychology Review*, 12(1):22–49, Feb 2008.

[11] J.-D. Fekete, J. J. Wijk, J. T. Stasko, and C. North. The value of information visualization. Springer-Verlag, 2008.

[12] J. P. Gee. Learning by design: Games as learning machines. *Interactive Educational Multimedia*, 8(8):15–23, Apr. 2004.

[13] N. Gershon and W. Page. What storytelling can do for information visualization. *Commun. ACM*, 44:31–37, August 2001.

[14] V. González and A. Kobsa. Benefits of information visualization systems for administrative data analysts. In *Information Visualization, 2003. IV 2003. Proceedings. Seventh International Conference on*, pages 331 – 336, 2003.

[15] V. González and A. Kobsa. A workplace study of the adoption of information visualization systems. In *I-KNOW'03: 3rd International Conference on Knowledge Management*, pages 92–102, 2003.

[16] D. Ha, M. Kim, A. Wade, W. Chao, K. Ho, L. Kaastra, B. Fisher, and J. Dill. From tasks to tools: A field study in collaborative visual analytics. In *IEEE Symposium on Visual Analytics Science and Technology, 2007.*, pages 223 –224, nov 2007.

[17] J. Heer. Prefuse: a toolkit for interactive information visualization. In *CHI 05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 421–430. ACM Press, 2005.

[18] J. Heer. Socializing visualization. In *CHI 2006 Social Visualization Workshop*, 2006.

[19] J. Heer. Voyagers and voyeurs: Supporting asynchronous collaborative information visualization. In *Proceedings of the CHI07*, pages 1029–1038, 2007.

[20] J. Heer and D. Boyd. Vizster: visualizing online social networks. *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.*, October:32–39, 2005.

[21] J. Heer, F. van Ham, S. Carpendale, C. Weaver, and P. Isenberg. Creation and collaboration: Engaging new audiences for information visualization. In A. Kerren, J. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization*, volume 4950 of *Lecture Notes in Computer Science*, pages 92–133. Springer Berlin / Heidelberg, 2008.

[22] F. Herzberg, B. Mausner, and B. Snyderman. *Motivation to Work.* John Wiley, New York, 1959.

[23] ISO. Iso fdis 9241-210:2009. Ergonomics of human system interaction - part 210: Human-centered design for interactive systems (formerly known as 13407)., 2009.

[24] E. Ju and C. Wagner. Personal computer adventure games: their structure, principles, and applicability for training. *SIGMIS Database*, 28(2):78–92, Apr. 1997.

[25] F. Karlsen. Media complexity and diversity of use: Thoughts on a taxonomy of users of multiuser online games. In *Other Players conference proceedings*, number December, 2004.

[26] D. A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8:1–8, January 2002.

[27] A. Kobsa. An empirical comparison of three commercial information visualization systems. In *Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)*, pages 123–130, Washington, DC, USA, 2001. IEEE Computer Society.

[28] A. Kobsa. User experiments with tree visualization systems. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 9–16, Washington, DC, USA, 2004. IEEE Computer Society.

[29] R. Koster and W. Wright. *A Theory of Fun for Game Design*. Paraglyph Press, 2004.
[30] A. Lau and A. Moere. Towards a model of information aesthetics in information visualization. In *Information Visualization, 2007. IV '07. 11th International Conference*, pages 87 –92, 2007.
[31] H. Macdonald, J. Yuille, R. Stanton, and S. Viller. The social life of visualization. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, OZCHI '09, pages 193–200, New York, NY, USA, 2009. ACM.
[32] T. W. Malone. *Heuristics for designing enjoyable user interfaces: lessons from computer games*, pages 1–12. Ablex Publishing Corp., Norwood, NJ, USA, 1984.
[33] D. A. Norman. *Emotional design*, volume 2004. ACM, New York, NY, USA, January 2004.
[34] R. Pausch, R. Gold, T. Skelly, and D. Thiel. What HCI designers can learn from video game designers. In *Conference companion on Human factors in computing systems*, CHI '94, pages 177–178, New York, NY, USA, 1994. ACM.
[35] D. Pfitzner, V. Hobbs, and D. Powers. A unified taxonomic framework for information visualization. In *Proceedings of the Asia-Pacific symposium on Information visualisation*, volume 24 of *APVis '03*, pages 57–66, Darlinghurst, Australia, Australia, 2003. Australian Computer Society, Inc.
[36] C. Plaisant. The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces*, AVI '04, pages 109–116, New York, NY, USA, 2004. ACM.
[37] T.-M. Rhyne. Computer games and scientific visualization. *Commun. ACM*, 45:40–44, July 2002.
[38] R. Ryan, C. Rigby, and A. Przybylski. The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, 30:344–360, 2006.
[39] M. Sedlmair, P. Isenberg, D. Baur, and A. Butz. Information visualization evaluation in large companies: Challenges, experiences and recommendations. 10(3):248–266, 2011.
[40] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1139 –1148, nov.-dec. 2010.
[41] C. Shapiro and H. Varian. *Information rules*. Harvard Business School Press, Boston, Massachusetts, 1999.
[42] I. D. H. Shepherd and L. D. Bleasdale-shepherd. Towards effective interaction in 3D data visualizations: What can we learn from videogames technology? pages 1–29, 2000.
[43] B. 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 336–343, Washington, DC, USA, 1996. IEEE Computer Society.
[44] B. Shneiderman. Designing for fun: How can we design user interfaces to be more fun? *Computing Systems*, September:48–50, 2004.
[45] Y. B. Shrinivasan and J. van Wijk. Supporting exploration awareness in information visualization. *IEEE Comput. Graph. Appl.*, 29:34–43, September 2009.
[46] Y. B. Shrinivasan and J. J. van Wijk. Supporting the analytical reasoning process in information visualization. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1237–1246, New York, NY, USA, 2008. ACM.
[47] J. van Wijk. The value of visualization. In *Visualization, 2005. VIS 05. IEEE*, pages 79 – 86, oct. 2005.
[48] F. B. Viegas, M. Wattenberg, M. McKeon, F. v. Ham, and J. Kriss. Harry Potter and the meat-filled freezer: A case study of spontaneous usage of visualization tools. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, HICSS '08, page 159, Washington, DC, USA, 2008. IEEE Computer Society.
[49] 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(6):1121–8, 2007.
[50] P. Vorderer, T. Hartmann, and C. Klimmt. Explaining the enjoyment of playing video games: the role of competition. In *Proceedings of the second international conference on Entertainment computing*, ICEC '03, pages 1–9, Pittsburgh, PA, USA, 2003. Carnegie Mellon University.
[51] M. Wattenberg. Baby names, visualization, and social data analysis. In *Proceedings of the 2005 IEEE Symposium on Information Visualization*, pages 1–7, Washington, DC, USA, 2005. IEEE Computer Society.
[52] M. Wattenberg and J. Kriss. Designing for social data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):549–57, 2006.

APPENDIX A

# Implementation

In this appendix we give an overview of the prototype implementation and describe the architecture, the database component and the communication between the browser and the server. Because both discussions and presentations are lists of comments, both are meant when only one of them is mentioned.

## A.1. Architecture

See Figure A.1 for the call graph of the prototype from the GUI in the browser to the database storing the discussions. InfoCloud is built inside the MagnaView WebLive Software solution. As it is part of WebLive, it can be reached over the internet, and uses Microsoft Internet Information Services (IIS, version 6.1) to expose itself as a website. It uses html and javascript for the layout and behaviour the user see on the browser side, and ASP as an entrance point for the communication on the server side. This ASP script sends the communication to the WebLive COM object, which is started for the user when he starts to use WebLive and InfoCloud. This COM object is the Delphi code which is the heart of WebLive.

To save the snapshots, discussions and comments, InfoCloud communicates with a database. The communication from the Delphi COM object towards this database does not go directly, because InfoCloud was built by using older code. This older code used this construction, so it was easier to adopt to this design than to start from scratch. This design has the added benefit that the WebLive/InfoCloud part that runs for each user does not need to be on the same machine a the database part.

## A.2. Database interface

As mentioned, the database component is stateless. This means that all actions that it supports keep the database in a correct state. The database used is a Postgres 9.0 database(`www.postgresql.org/`), because the license of this software allows usage for commercial purpose without fee or agreement from its creators. It is also a stable and fast database with many features, which makes it competitive with commercial alternatives.

The database model can be seen in figure A.2. The structure is fairly simple, as both discussions and presentations are just lists of comments, only different in how the user interacts with them. Comments

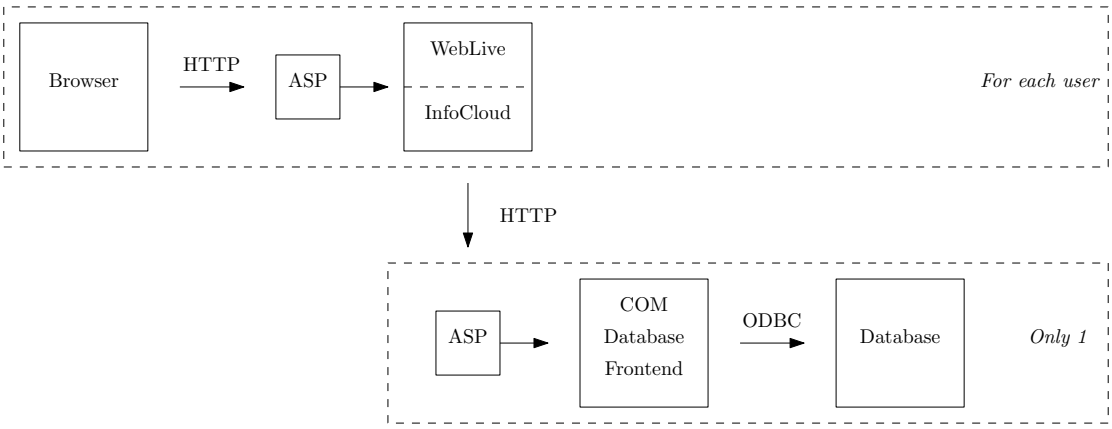



FIGURE A.1. The architecture of MagnaView WebLive including InfoCloud. The arrows denote the direction of the calls, and a reply is always sent, either as the retrieved object from the database or that the called procedure finished successfully.

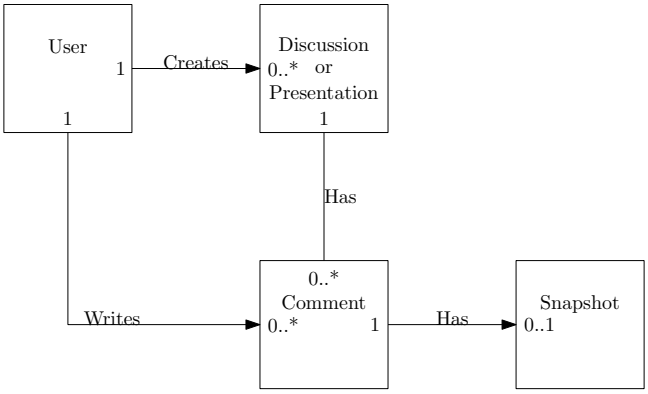



FIGURE A.2. The database diagram of InfoCloud. The four squares are tables and the arrows denote relations. The cardinality is written as follows: a user can create zero to many discussions, therefore it says "0..*" on the discussion side of this relation.

are the central object in the database, as they are inside discussions or presentations, made by users and contain a snapshot. The design is intentionally kept this simple, as it makes extensions of the system easier. As during development it was not clear what features were going to be implemented, this flexibility allowed easier planning.

More complex queries can be obtained from this simple database with SQL queries, which is fully supported by Postgres. Queries are created in the COM object according to the calls it gets from WebLive, which are then sent to the Postgres database with an ODBC connection. One of the more difficult queries sent is for the keyword search, which uses regular expressions, subqueries and aggregates to obtain results with relevancy to the query.

### A.3. Browser - server communication

As mentioned, the standard HTML+CSS+JavaScript combination is used for the website GUI of InfoCloud. The Javascript library JQuery is also used, as it has many useful functions. Communication to and from the WebLive server is done as much as possible in JSON (JavaScript Object Notation), as JSON is small (usually smaller than XML), easy to read and easy to use in Javascript. As said, the standard communication line of WebLive is used, which sends all the data towards a certain function in the Delphi WebLive code (the name of the function is `Get`). In this `Get` function, code was added that routes the calls for InfoCloud to InfoCloud code. within the Delphi InfoCloud some data was cached, as communication with a database is slow.

Also, the data in the database is confidential, not public, therefore all actions in the WebLive COM object are protected by a login check, so if the user is not logged in, he cannot do anything. Also on the browser side, calls for this confidential data are only done when the browser knows the user is logged in, directing user flow towards the login section when the user is not logged in. This results in InfoCloud only showing the login bar when the user is not logged in, and after logging in the navigation information is extracted from the URL, which is used to determine what the user wants to see in InfoCloud.