Advanced Data Visualization in Business Intelligence Software

Making data insights easier

Monday, February 15, 2010

ClicData SARL
Telmo Silva & Brock Rowlands

54/58 rue Nationale
F-59800 Lille
France
www.clicdata.com

TABLE OF CONTENTS

PREFACE

Data visualization is an exciting topic of discussion since it offers the promise of providing great insight into existing data within the enterprise. However, as with any new technology, the prospects and advantages of data visualization can be over-stated. Business intelligence was been put forward to offer many benefits including the improvement of data visualization; implemented as an extension of enterprise software solutions or as stand-alone product, but for the most part the promises of business intelligence vendors haven't been realized. Therefore it is important to examine the proper approach to and proposed benefits of data visualization carefully before selecting and implementing a Business Intelligence solution. This reference is going to review several components of data visualization systems, discuss implementation concerns, issues and practices and provide advice on how to ensure your data visualization solution is properly provisioned.

THE PURPOSE OF DATA VISUALIZATION

Data visualization is a way of making complex data simple and easy to understand and consume. As Edward Tufte stated "Graphical excellence consists of complex ideas communicated with clarity, precisions and efficiency." He adds, "Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space."

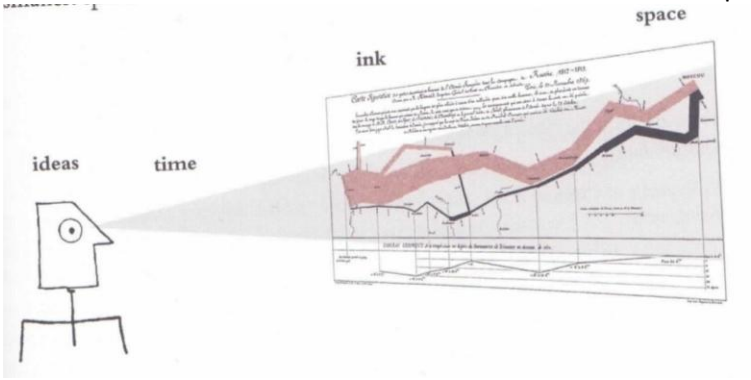

Figure 1 - Principles of Graphical Excellence

Good data visualization conveys more information to the user than numbers or text could, however if the data representation is incorrect, incomplete, or misleading data, it has obscured the information rather than revealed new insight and understanding.

There has been very little innovation in data visualization since its inception and adoption in business and scientific applications. An example is Microsoft Excel, one of the most widely used data analysis tools. Examining the chart wizard shows that in over 10 years there have been very few changes in how data visualization is implemented, other than updating the chart wizard's graphics.

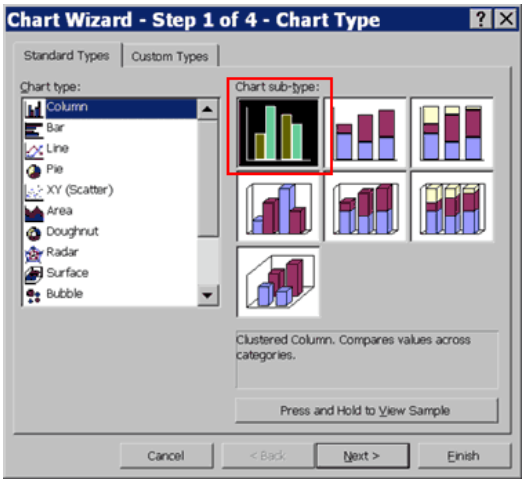

Figure 2 - Excel 98 Chart Wizard

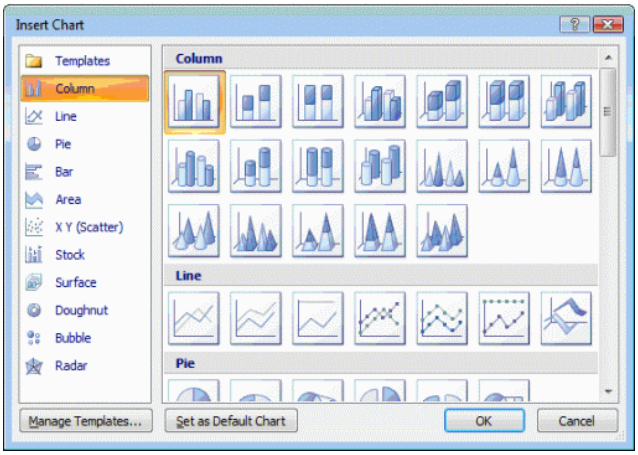

Figure 3 - Excel 2007 Chart Wizard

An argument as to the reason for the lack of change could be that charting must follow certain convention to be useful as an analytical tool, but innovation in this area would provide great benefits to both users and software providers.

If data visualizations are poorly constructed the result can obscure the information the data contains. Whether data visualizations are executed well or poorly, experts in the area of data visualization agree that there are limits to the amount of useful information that can be placed on a screen. While adding mark-up information (3D, colours, additional mark-up such as arrows and highlights, etc.) to data visualization can add contextual information it should only be used when it can add clarity to the data analysis.

Data visualization can turn data into compelling information when properly executed. An example is using Jacques Bertin's "Semiology of Graphics". As seen below, both visualizations contain the same data, but representing the numbers graphically, much more information is conveyed by the dot density compared to the actual numbers. This means that areas of high and low density can be easily recognized and areas of further investigation quickly determined.

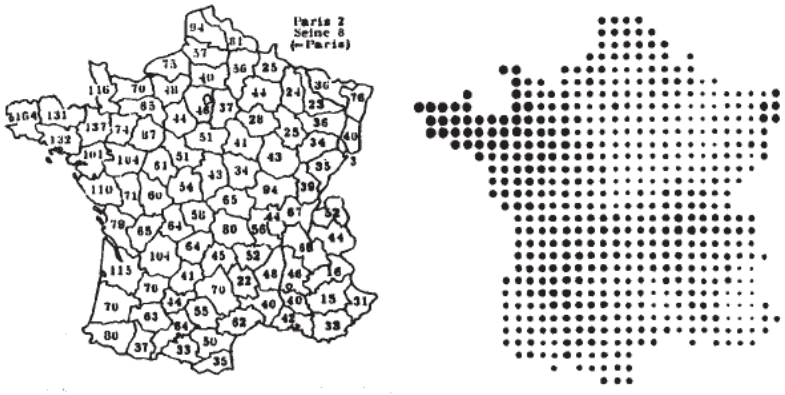

Figure 4 - Bertin's Semiology of Graphics

THE IMPACT OF TECHNOLOGY IN DATA VISUALIZATION

Using computers it is easy to provide more data and visualizations of the data, but with data visualization it is important to ask **what information is important**, and how to illustrate that data clearly, rather than continuing to add more data which could be confusing to the user.

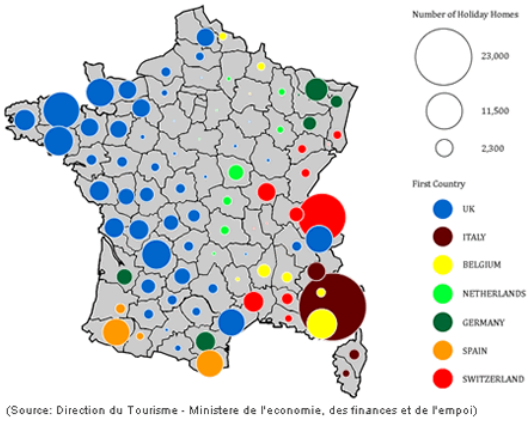

Figure 5 – A visualization showing many different data types

Benjamin Fry has said, “Because of the accuracy and speed with which the human visual system works, graphic representations make it possible for large amounts of information to be displayed in a small space.” However, with new technology the power and novelty of techniques can lead to using technology for technology’s sake instead of providing real benefits to users and providing more information through visualization.

Implementing a successful data visualization solution will require understanding of:

- infrastructure/hardware
- end user interaction
- user adaptation
- user adoption

All of these components contribute to proper data visualization and are a part of successful data visualization solutions.

HARDWARE

Increasing amounts of data generated by your company place increasing demands on your hardware infrastructure, and with data visualization there can be a need for increased processing power to deal with the data processing requirements. Super computers exist such as Silicon Graphic’s Stokes, one of the top 200 fastest machines existing today, which is powered by 640 Intel® Xeon® processors (totaling 2,560 cores) and features 5.1 Terabytes (TB) of memory and 84TB Panasas ActiveStor parallel file storage solution. Considering the power that the average consumer computer has, super computers are not that impressive. If fact many large companies have deployed equivalent processing power, or better, to its users in the form of desktop and laptop computers. Setting up a

complete data infrastructure; could mean researching, purchasing, setting up and maintaining a large hardware installation, including the space, electricity, cooling equipment and personel requirements, or now there are cloud computing options available so that your infrastructure can be outsourced. This means that instead of large capital expenditures your infrastructure can be managed through cloud computing services. Setting up this service means additional data security, network setup and bandwidth concerns but it can provide a viable alternative to managing your own data centre.

So what does this mean? It means that now more than ever there are many computing options. For focused scientific research the supercomputer is still very important, but distributed computing which harnesses the power of your existing computers can be a powerful, accessible and affordable way of adding processing power for data visualization and other computationally intensive tasks. Cloud computing can offer a method of creating a scalable architecture by partnering with a service provider.

Previously limitations of computing involved processing power and memory constraints. While computers have grown more powerful with faster processors and more memory, the innovation in applications has not followed pace. Data visualization solutions need to provide innovation and methods of better leveraging your infrastructure decisions.

USER INTERACTION

Computer human interaction has not radically changed since computers were first introduced. The keyboard, first put in use in the post World War II era, has remained, for close to 70 years, as the preferred method of computer input device. The mouse has likewise reigned for close to 30 years without dispute and will most likely continue for some decades to come. There are many methods of interacting with computing systems, but none have come to challenge the dominance of the keyboard and mouse.

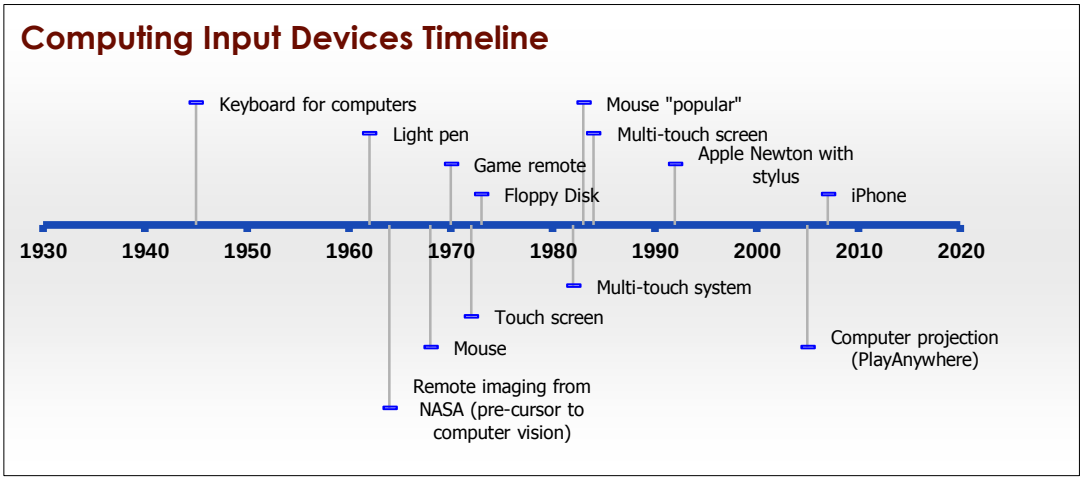

Figure 6 - Use of computing IO devices for the last 70 years

Computer-human interaction devices are expected to evolve as alternative ways of interacting with computing systems become more wide spread, including navigating 2D and 3D data models. Alternative methods have started to show up in personal computing and the electronic game industry. Change to standards is always slow, and touch screens, gestures and voice commands have existed for a long time but have yet to be adopted outside niche markets.

Voice recognition is an input method which has been around since the early 1990's and yet has not reached wide acceptance. Very few people actually use voice recognition to interact with their computer, or mobile phone. Very few people know that since Windows 98, voice recognition and voice output has been a part of the Windows operating system's standard set of functionality. Voice recognition is still severely underused for a variety of reasons. These include the effort and practice required to use the voice recognition systems, and the hardware limitations. Microphones require close placement to the user's mouth, or they need to be sensitive enough to pick up the commands of the user without capturing all the ambient noise or other conversations. Additionally current voice recognition systems do not properly contextualize user's input; there is difficulty in differentiating between commands, simply talking out loud, or conversations with another person in the room.

Touch interfaces have been made popular by Apple and their iPod/iPhone hardware. While touch interface computing has been tried before by many different hardware vendors, including Apple, it wasn't until the iPod and iPhone that touch interfaces became widely used and accepted. The Apple Newton had a stylus and handwriting recognition, and the Palm PDA was a market leader with user stylus interaction and graffiti system which recognized handwriting, but these devices were never as widely adopted as touch interfaces have now become. Apple has announced the iPad, a large scale tablet device which could replace smaller laptops but while there is much discussion surrounding the iPad, it remains to be seen what impact it will have on the market and if it will be readily adopted by users.

Science fiction has inspired many of the technology advances that we enjoy today and looking at the movie "Minority Report" is a perfect example of a seemingly amazing user interface that combines personalization, gestures and large scale displays to convey large amounts of data, in a flexible and easy to use manner. An example of science fiction coming to life is the AtracSys' touch screen system which gaining some popularity with marketing departments to showcase their products.

Figure 7 - AtracSys touch display used for marketing purposes

The interface interacts extremely well with single touch based manipulation and has an advantage of the transparent projection surfaces.

Data visualization will be of limited value if it is only accessible through traditional input/output methods of the keyboard and mouse. Being able to use touch interfaces or interact with the data during presentations will add to a data visualization solution's applicability and usefulness. And while

touch interfaces are now common for mobile applications, they have yet to be widely used on larger form factors, although touch and multi-touch screens are becoming more readily available.

USER ADAPTATION

Business and scientific applications are not typically designed with robust and easy to use interfaces. In fact, most programmers are not user interface or user experience experts. Enterprise software solutions are well known for complex and difficult interfaces, with a focus on engineering and features rather than usability. Applications should be designed so that users can easily navigate the interface, and perform basic tasks in the software without extensive training. An example of intricate yet easy-to-use interface can be found in video games.

Take for example any of the most popular games available in the market today. Their interfaces are such that the user is encouraged to complete their primary tasks, which can be exploring a world, saving the galaxy or fighting off an alien invasion. The games allow players to navigate in a three dimensional (3D) world easily using simple input devices such as a mouse and keyboard or a controller. They do that by constantly relaying information through dashboard types of display, but more importantly by allowing the player to orient themselves via the actual virtual landscape. Seen below, Gears of War 2 shows a scene where the user navigates based on their position to other objects and actors. The engagement of the user is maintained, and they don't realize the breadth and depth of data that they are receiving and reacting to through this visual interface.

Figure 8 – Gears of War - the interface is the user

New generations of scientific and business analysts will have higher expectations in navigation and reaction times, based on their experience with consumer technology, and games. Many application interfaces could benefit from redesigns with many ideas and methods taken from consumer technology like the iPhone and games.

Google has been very successful at dominating the search market by providing a simple and easy to use interface while providing fast and relevant search results. For many people who now use computers, it is more important to know how to find information that it is to know all the information.

Data visualization solutions should make it easy for you and your users to find the information that they need. It is important to be able to ask questions, using real language, and not having to use programming languages to form complex queries to get answers to your business questions. By incorporating natural language searching your data visualization solution will be more accessible and usable to a much larger audience than a solution that relies on only a complex query language.

USER ADOPTION

Technology in its many forms can be complex and confusing, even for information technology professionals. As computers grow more powerful and capabilities are ever expanding, applications for technology exponentially increase as does their complexity. Technology imposes changes in process and behaviour and changes that can be difficult for people to accept. However, as Peter M. Senge states “People don’t resist change. They resist being changed.” The adoption of a data visualization solution depends on the benefits to the users, whether they are tangible, intangible or even emotional. Users typically have to adapt to a new technology and the steepness of the learning curve involved in adapting to a new technology can spell the difference between success and failure. User centric design is very important for technology adaptation and how the technology fits the user’s methods. Data visualization is becoming a mainstream tool, as can be seen in television broadcasts, newspapers and websites.

Data visualization solutions need to provide clear benefits to the users, not just expected corporate ROI and improvements to features, process, or reporting. Information technology projects have a history of failure when a new technology is imposed on a user base. Rapid technology adoption can happen, and a recent example is the iPod and iPhone from Apple. While the devices were innovative in their approach to providing technology for users, it was successful not because it was a leading edge technology (it wasn’t) but how it made the user’s life easier and it was fairly simple to adopt. For successful technology implementations, change management has a role to play. Project failures are a result of many factors, but the most important part of any technology project is to create user buy-in and engagement in the solution.

Historically technology has been designed for technical users, even consumer products typically require a significant learning curve to properly use the software even the basic functions. User centric design is not just about simplifying the user interface, but making the software accessible to users by understanding the environment and how they use it. For software to be easy to use, it is important that its features are discoverable. It is common knowledge among software development professionals that software manuals and help files are the last resort of users; if they are used at all. Software needs to be designed to be simple, (but not simplistic) which usually takes a significant amount of thought and engineering.

Data visualization can be a very powerful business tool since the human brain is the most powerful pattern recognition system. Visualizations, when properly created and presented will lead to new business insight, ability to understand and address issues and to make better informed business decisions. However, this is only the case if the data visualization software solution can be easily adopted by users. No matter how advanced the functionality, if there is not a simple and easy way to get started using the data visualization solution, it will not be used.

SUMMARY

In looking at data visualization solutions and business intelligence software we found that there wasn’t a product on the market that adequately addressed these issues for users. Recognizing the market need was what brought us to develop Infinis.

Infinis can be provisioned as a solution running in an existing data centre, as software-as-a-service or as part of a cloud computing solution. The flexibility in managing Infinis and the hardware requirements was thought to be important because of the variety of computing solutions available

today. There are pros and cons to each hardware implementation method, however since Infinis can be deployed in any manner, the decision to be made is, which one is right for your situation?

Infinis provides many different user interaction methods. This is because Infinis was designed to be a collaborative analysis and presentation tool. Users can interact with Infinis and their data using the traditional keyboard and mouse, but also using touch. Additionally, when using projectors with Infinis, IR remotes can be used to select, move and manipulate data during a presentation. This multiplicity of user interface methods means that the right one can be chosen for a given situation and Infinis is able to be used in a variety of scenarios.

Infinis was designed to provide a visually compelling interface that promoted a visual experience and provided many methods of user interaction. Infinis uses virtual workspaces that can have text, graphics, video, included with data to provide additional information and inputs to your data analysis. By providing additional context for your data analysis, new conversations and discussions can be held.

With any new technology or software there is a learning curve for users, and there can be resistance to adopting new methods of accomplishing tasks. Infinis seeks to overcome the typical user resistance by creating a simple and easy to user environment, where a user can be almost immediately productive in data analysis tasks. Natural language search is one method of making tasks that are traditionally difficult, like obtaining data from business intelligence software, simpler for users.

Infinis is a new approach to business intelligence software and incorporates many new technology and user-centric changes to business intelligence software. We have looked at many of the challenges facing businesses today; with the continuing rapid change in technology, the globalization of competition and business, the continuing growth of data generation and data consumption by businesses, increased competition as well as the ever increasing demands of customers. Infinis has been designed to provide you with faster and better access to your data, so that you can meet these challenges, and expand your business.

Visit us at **InfinisWorld.com** to see and understand how Infinis is changing the approach to business intelligence applications and how Infinis can help you achieve better results for your business.

REFERENCES

1. Edware Tufte, *The Visual Display of Quantitative Information,* p.51, 52, Graphics Press LLC, 2007 2nd edition
2. Jacques Bertin, *Sémiologie graphique*; Editions de l'Ecole des Hautes Etudes en Sciences, 1999. [In French]
3. Government of France, Economic Ministry, Tourism, website: http://www.industrie.gouv.fr, Direction du Tourisme – Ministère de l'econonomie, des finances et de l'emploi, accessed June 6, 2009
4. Benjamin Jotham Fry, Organic *Information Design*. p.14,. Carnegie Mellon University, 1997
5. Atracsys website, http://www.atracsys.com/_products/beMerlin.php, accessed February 10, 2010
6. Screen capture from the game, Gears of War 2, for the Xbox 360 console