A Web-Enabled Visualization Toolkit for Geovisual Analytics

Quan Ho*, Patrik Lundblad, Tobias Åström, Mikael Jern
National Center for Visual Analytics (NCVA), Dept. of Science and Technology, Linköping Univ./
SE-581 83 Linköping, Sweden

ABSTRACT

We introduce a framework and class library (GAV Flash) implemented in Adobe's ActionScript, designed with the intention to significantly shorten the time and effort needed to develop customized web-enabled applications for visual analytics or geovisual analytics tasks. Through an atomic layered component architecture, GAV Flash provides a collection of common geo- and information visualization representations extended with motion behavior including scatter matrix, extended parallel coordinates, table lens, choropleth map and treemap, integrated in a multiple, time-linked layout. Versatile interaction methods are drawn from many data visualization research areas and optimized for dynamic web visualization of spatio-temporal and multivariate data. Based on layered component thinking and the use of programming interface mechanism the GAV Flash architecture is open and facilitates the creation of new or improved versions of existing components so that ideas can be tried out or optimized rapidly in a fully functional environment. Following the Visual Analytics mantra, a mechanism "snapshot" for saving the explorative results of a reasoning process is developed that aids collaboration and publication of gained insight and knowledge embedded as dynamic visualizations in blogs or web pages with associative metadata or "storytelling".

Keywords: web-enabled visualization toolkit, geovisual analytics, storytelling, layered component architecture, multiple-linked views

1. INTRODUCTION

The term 'Web 2.0' has become undisputedly linked with developments such as blogs, wikis, social networking and collaborative software development. While the benefits of geovisual analytics tools are many, it has been a challenge to adapt these tools to the Internet and reach a broader user community. Research has so far focused more on tools that explore data [1, 2, 3] while tools that efficiently communicate and also publish gained knowledge have not achieved the same attention. In this context, we introduce an extended Flash version of the previous GAV toolkit [10], programmed in Adobe's ActionScript including a collection of common geo- and information visualization methods extended and optimized for this environment. An integrated storytelling mechanism will enable collaboration and transition of spatio-temporal and multivariate data into communicative sense-making news entities with contextual metadata. The analyst uses tools (figure 1) to: 1) input data, 2) explore and discern trends, 3) orchestrate with snapshots and metadata, 4) collaborate with colleagues to confirm and 5) finally publish essential gained insight and knowledge embedded as dynamic visualization "Vislet" in a blog or web page with associative metadata.

Data are analyzed through the use of time-linked views controlled by a time slider (figure 2). Trends are detected through several visual representations simultaneously, each of which is best suited to highlight different patterns and can help stimulate the analytical visual thinking process so characteristic for geovisual analytics reasoning. Interactive features include tooltips, brushing, highlight, visual inquiry, and conditioned statistics filter mechanisms that can discover outliers and simultaneously update all views. Of particular interest are our Flash implementation of scatter matrix, table lens and parallel coordinates plot (PCP) applied in three characteristic case studies that are included to demonstrate and evaluate the tools: Statistics data visualization (OECD), Self-organizing mobile networks (Ericsson) and Road emergency (SMHI) developed in close collaboration with our domain experts. In summary, this paper introduces a web-enabled visualization and dissemination toolkit supporting geovisual analytics tasks facilitating:

- Web enabled motion visual representations such as table lens, PCP, scatter matrix, treemap, scatter plot, choropleth map;
- Google map or Bing map as layered background maps to guide users in their analysis of data;

[*] ho.van.quan@itn.liu.se; phone 46 736 726 174; ncva.itn.liu.se

Visualization and Data Analysis 2011, edited by Pak Chung Wong, Jinah Park, Ming C. Hao, Chaomei Chen,
Katy Börner, David L. Kao, Jonathan C. Roberts, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7868,
78680R · © 2011 SPIE-IS&T · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.872250

- Coordinated multiple, time-linked views;
- A data cube model for fast access to spatial-temporal and multivariate attribute data required for time animation;
- Component-embedded interactions including brush, pick, highlight, filter, dynamic sliders, focus & context and other special interaction facilities;
- Framework for the creation of user components such as data transformers and also for making changes to existing low-level visualization components so that ideas can be tried out rapidly in a fully functional useful environment;
- Integrated snapshot mechanism for saving and packaging the results of a geovisual analytics reasoning process – the foundation for Storytelling;
- Publishing (HTML code) discoveries and knowledge in blogs or web pages as embedded dynamic visualizations;

The remainder of the paper is organized as follows. Section 2 gives an overview of related work. Section 3 presents our GAV Flash framework. Concepts of snapshots and storytelling are presented in section 4. Then three case studies are presented in section 5 with user feedback and the conclusion in section 6.

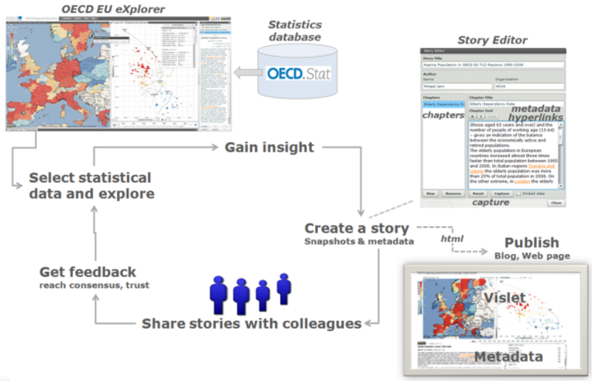


Figure 1. A scenario in OECD eXplorer [4], a statistics geovisual analytics application for exploring and publishing statistical data on the web and developed with the GAV Flash toolkit.

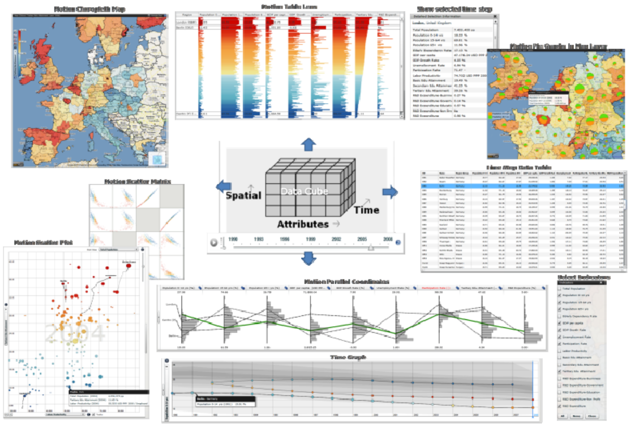


Figure 2. GAV Flash views are time-linked, important in the synthesis of animation within explorative data analysis. The user can stop the time animation and start interacting with the data at any time step.

2. RELATED WORK

We divide generic tools for explorative data analysis into two categories: tools supporting for visualization and tools extended with built-in snapshot and storytelling capabilities.

2.1 Toolkits for spatial-temporal multivariate data

Visual exploration of spatial-temporal and multivariate data has been the subject of many research papers. Andrienko has illustrated several motivating approaches in earlier papers [5, 1]. Muller [2] and Guo [3] are other appealing examples. Many papers [6] emphasize the advantages of multiple linked views.

InfoVis Toolkit [7], CommonGIS [8], GeoVista [9], VIS-STAMP [3], GAV [R10] and CGV [11] are examples of exploratory data analysis (EDA) tools that all have evolved from research and can leverage visualization and computational methods to search for space-time and multivariate patterns. While the benefits of geovisual analytics tools are many, it has been a challenge to adapt these tools to the Internet and reach a broader user community. Tableau and its predecessor Polaris [12] are examples of a popular web-enabled tool applied to business analytics. Web-enabled tools are needed for applications explicitly designed with the purpose of exploring and communicating large spatial-temporal and multivariate data. Such tools should also employ data transformers and data providers, layout mechanisms, interaction, time animation and storytelling suited for a geovisual analytics' task.

Our applied research partners (see case studies) do not allow "unknown" software for security reasons to be installed in their computer administration, but Adobe Flash has apparently become an accepted web-enabled visualization platform in the commercial world. Another requirement, emphasized by our partners, is the possibility to disseminate gained discoveries and knowledge as dynamic visualizations combined with metadata.

2.2 Tools for storytelling

The importance of a capacity to snapshot EDA sessions [13] and then reuse them for presentation and evaluation was early demonstrated by MacEachren [14] and Jern [15] and incorporated features to capture and reuse interactions and integrate them into electronic documents. CCMaps [16] presents a mapping tool that allows users to save snapshot events and reuse them for presentation purposes. Another effort was made by Visual Inquiry Toolkit [3] that allows users to place pertinent clusters into a "pattern-basket" to be reused in the visualization process. Robinson [17] describes a method they call "Re-Visualization" and a related tool ReVise that captures and reuses analysis sessions. Keel [18] describes a visual analytics system of computational agents that supports the exchange of task-relevant information and incremental discoveries of relationships and knowledge among team members commonly referred to as sense-making. Many Eyes [28] is an interesting storytelling approach implemented for a public web site where novice users can upload their own data, create dynamic visualizations and participate in discussions.

We build upon previous research [20] and extend our toolkit and framework GAV for dynamic web visualization based on Adobe Flash. An integrated editorial and related authoring process with the goal to advance research critical to educational production and publishing is introduced based on a comprehensive snapshot mechanism supporting multiple-linked and motion geovisual analytics. Our tool facilitates support for both the expert and public user.

3. GAV FLASH FRAMEWORK

The GAV Flash framework is developed based on a recommendation from the visual analytics (VA) research program [21] "*support seamless integration of tools so that data requests, visual analysis, note-taking, presentation composition and dissemination all take place within a cohesive environment*" addressing the need for integrated exploratory, analytical reasoning and communicative tools. Common geovisualization and information visualization components are included that support interactive features such as tooltips, brushing, highlight, visual inquiry, and conditioned filter mechanisms that can discover outliers and methods supporting time-linked multiple views. Also tools that support data analysis algorithms, tools that connect the components to each other and data providers that can load data from various sources (figure 3). The GAV architecture allows new or existing components/classes to be incorporated with the already existing components, e.g. special methods used in our case studies (section 5). Means are also provided for a developer to extend and further customize the popular information visualization methods by breaking them into lower-level "atomic" components (figure 4 and 5).

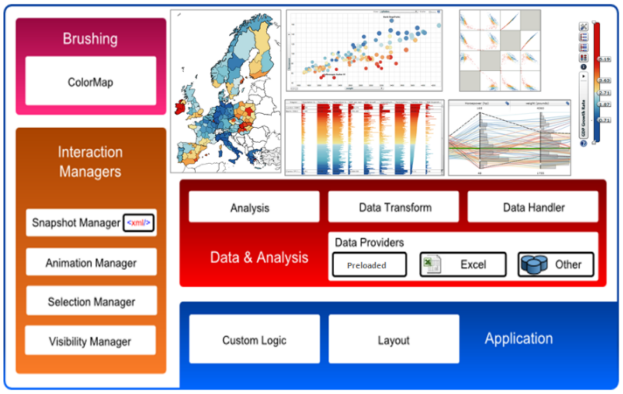


Figure 3. GAV Flash Framework

3.1 Framework design principle

The core philosophy of GAV Flash is modularity, we want application developers to be able to pick and choose from a wide range of visualizations, data providers and data transforms and combine them in various ways. This puts a high demand on each component of the framework to be generalized so it can receive and communicate data with others, but also be self contained so that the advanced functionality is always present, no matter which components are combined. The generalization is achieved through definition of interfaces, which detail only the necessary functions and properties in assets shared by components. An example of this is the data set, whose interface is limited to functions that supply data and metadata, all other functionality is encapsulated in the implementation. As the components are only aware of the interfaces, we can easily replace the data set with some other structure, for example a direct database connection, without re-implementing any visualizations or data processors. Apart from the data sets, GAV Flash applications are built using a combination of visualization components and linking modules that control selection, filtering, color and animation. Other components handle application level events such as menus, and a module for the snapshot mechanism, which is described further in section 4. The abstraction into interfaces also allows others to extend the framework with new functionality, be it new visualizations or data providers. They simply have to follow the framework definitions of how to access data and shared assets and then implement their own ideas.

3.2 Atomic and functional component architecture

The generalization of components coupled with advanced features can make it hard to encompass all data scenarios in a component. It could be faced with a large multivariate data set but also with a highly dense temporal set. These two types of data sets often require different solutions in terms of the data processing, the element drawing, and also the end user experience of the visualization. To facilitate this need for dynamic components we break them down into small blocks called atomic components. These atomics are used together to form a fully functional component but they are not dependent on each other, so they can be combined in any way. This concept can take many forms depending on the parent component, the clearest example being how the map uses different layers to display different levels of data. The map base class contains no visual parts at all, but it controls everything needed to display something within its context. It creates transformations to deal with projections and keeps track of the user's input as she zooms or pans. The visuals are instead created in layers, examples being polygon layer for colored regions and glyph layer for showing data related to points. By making the map independent of its layers, we can combine them in any way we choose, or create new ones as long as they adhere to the basic principles of the map. The same type of concept is used in other components as well, while not as obvious as the map example.

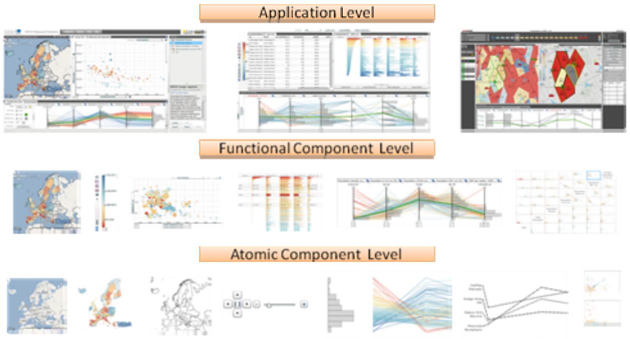


Figure 4. GAV Flash Atomic Component Architecture

The combination of a component base with one or several atomic parts forms a functional component. They are generally encapsulated together with the required GUI elements needed to control the visualization so that only the combined properties of all atomic parts are exposed to the surrounding application. The atomic parts allow the creation of custom functional components that can differ extensively depending on the end users' needs. A number of atomic parts can be reused in several components. For example, a circle glyph layer can be used in both the scatter plot and in the map or a range filter can be used in both the PCP and the color legend. The functional components can in turn be used by application developers and linked to each other through the use of linking components such as a selection manager, a visibility manager, and an animation controller to create quick prototypes and get a first look at their data. That first prototype can then determine which way the visualization needs to go, and if some kind of special atomic and/or functional components needs to be developed.

An example functional component: Parallel Coordinates

The building blocks for a standard PCP is axis and lines, commonly also the ability to filter along each axis and possibilities to see statistical analysis on the data. The GAV Flash PCP consists of a base plot model that keeps track of all common axis related information, such as order, visibility, positions and orientation. This model is available to all atomic layers added to the plot so that they can draw their contribution in the correct coordinate system. Each layer works independently of each other but they can all affect the common controls such as filter and selection lists by accessing the model. Figure 5 shows a number of typical layers of a PCP: axis layer, header layer, line layer, selected line layer, mean line layer, range filter layer, percentile filter layer, and histogram layer.

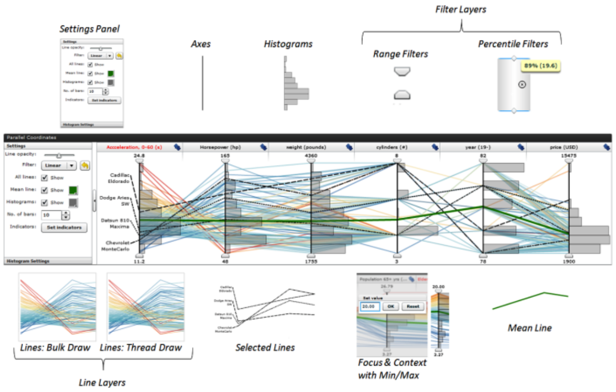


Figure 5. Functional PCP components and their atomic parts

3.3 Data model

GAV Flash uses a simple data set model as a base for storing and communicating data. It is designed to manage data in three dimensions, represented by attribute, space and time and can communicate the boundaries and content of these dimensions to the visualizations. To make it independent of actual storage structure, we have used a static and simple interface for data access. This allows different storage structures to be implemented to serve for different purposes. We have implemented a storage structure for GAV Flash that simply is an array optimized for fast access and can handle large spatio-temporal, multivariate datasets.

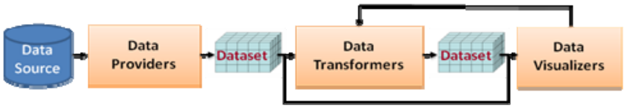


Figure 6. A typical pipeline architecture of VA applications

This data model also facilitates implementation of VA applications that typically have pipeline architecture of three modules, which is similar to the visualization pipeline of Card et al. [29], as illustrated in figure 6. In this architecture, data is first loaded from data sources into the data provider module and then passed to the data transformation module for analysis and/or processing before being passed to the visualization module for visualization. The transformation is optional, and an application can even combine visualizations that display the data both before and after transformation. Visualizations can be linked to control the transformation thus giving the user direct access to how the data is manipulated.

The open architecture, as presented below, allows developers to integrate their own components into our framework, for example, a new data transform component facilitating a needed data analysis algorithm that first process the data before visualization can be made. GAV Flash can then focus on visualization leaving data analysis to the experts.

3.4 Extendibility

To ensure that a specific VA task can be solved by the framework, the framework is designed so that components are extendable by developers. For example, developers can add new features or new layers into an existing component or rewrite a component to improve its performance or even replace an existing component by a new component. To make this possible, we have employed interface mechanism extensively in our design of components. Components are based only on interfaces that are necessary to make them work when connected to other components. Nevertheless, to make the implementation of components easier the interfaces are kept as simple as possible. This means that they only contain definitions necessary for communicating with the components. For example, in the data model above, datasets use a simple interface, called IDataset, which only includes definitions of a basis function getValue and some extra functions such as getMax, getMin, getMean to supply necessary data to other components as well as to avoid recalculation of the same operation in different components.

3.5 Performance and interactivity

For large datasets, we face two major challenges, performance and interactivity, when implementing advanced components such as PCP and choropleth map. For a PCP, the performance largely depends on the number of lines, the number of line segments of each line, line thickness, line opacity and the view size. For small datasets, of less than 2000 items over about 10 attributes, our approach is to represent a line by a Flash Shape object. This approach facilitates the implementation of line filtering and line color/opacity transforms through employing the abilities of Flash Player to show/hide objects and transform their color/opacity without redrawing. However for larger datasets the rendering time increases quickly. This is due to the pixel color blending process taking a longer time as the number of lines through a single pixel increases. To speed up this process we divide lines into groups, where each group includes a subset of lines, the number can be decided by users, but typically around 200. Each group is drawn into a Shape object and then cached as a bitmap. This approach will reduce the number of blending operations used, therefore reduce the rendering time. The two approaches have been tested on a dataset with 9000 items over 13 attributes for the operations: rendering and color/opacity transforms. It was found that the second approach with 200 lines in each group is 23 times faster than the first approach but takes 31 more MBs of memory due to the cached bitmaps.

For choropleth maps, the performance depends on the number of regions, the number of polygons per region, the number of vertices of a polygon, the view size and the zoom level of the maps. We choose to represent a region by the aforementioned Flash Shape object, that, similar to the PCP case, facilitates implementation of filtering, selection and color/opacity transforms. This works well with normal maps of about 2000 regions, 5000 polygons and 250000 vertices, and even with larger maps that include 9000 regions, 24000 polygons and 1100000 vertices in total when we use our threading technique presented below. The map does not suffer from the color blending problem of the crossing PCP lines. However, for maps even larger, we are planning to use a tiled approach, common to web map services like Google maps, and employ the shader technology recently supported by Flash Player to enhance the performance further.

The second challenge comes from the fact that Flash Player does not support multithreaded ActionScript, which makes the application unresponsive during the execution of tasks that sometime can take ten seconds such as drawing a large number of items. Unresponsiveness for long time can distract the user and make exploration process of the user less efficient. To address this problem in a similar manner to [22], we have implemented our own library to simulate a multithreading environment. Our approach is to assign each large task, e.g. drawing 10000 lines, to a so-called micro thread or pseudo thread and to divide the large task into a number of smaller tasks, e.g. drawing 100 lines at a time, so that each small task can be executed in a reasonable time period. Then small tasks of threads are scheduled into frames so that the total execution time of each frame is less than what the given frame rate allows, typically 30ms per frame. After each frame the control is returned to the Flash Player and allows the user to interact with the application. Threads can be cancelled early due to user interaction and new threads can be scheduled and executed to reflect user updates.

Although used to enhance interactivity during execution of large tasks, it is important to realize that our threads are, due to the nature of Flash Player, still executed in a single ActionScript thread and thus do not employ multi core processors. A demonstration of aforementioned techniques can be seen in our demonstration application GridMap [23].

4. STORYTELLING

Collaboration [20, 24] is achieved through a mechanism in GAV Flash (figure 1) that supports the storage of interactive events in an analytical reasoning process through "memorized interactive visualization views" or "snapshots" that can be captured at any time during an EDA process and becomes an important task of the storytelling authoring analytical reasoning process.

4.1 Snapshot

When exploring and making sense of, for example, comprehensive statistics data, we need a coherent cognitive workspace to hang our discoveries on for organizing and navigating our thoughts. The GAV Flash toolkit includes such means by capturing saving and packaging the results of an exploration "gain insight" process in a series of "snapshots" that could help the analyst to highlight views of particular interest and subsequently guide other analysts to follow important discoveries. The snapshot tool creates a single or a continuous series (story) of visualization captures during the exploration process. In a typical scenario the analyst has selected relevant attributes, time step (temporal data), data items-of-interest, color class values, filter conditions for selected attributes and finally highlights the "discovery" from a certain angle (viewing properties).

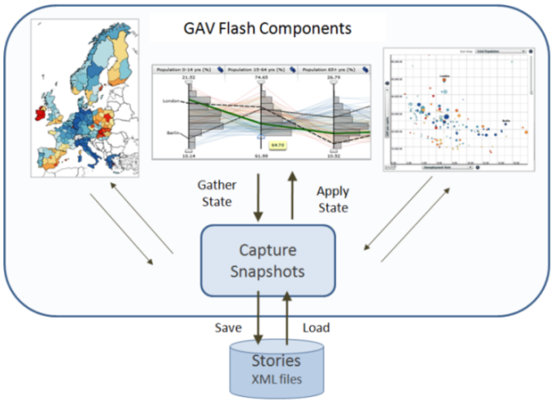
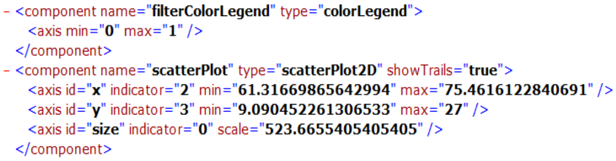


Figure 7. The snapshot system scans through all active views to gather and apply states.

The analyst requests a snapshot with the Capture function that results in a snapshot class operation scanning through all its connected GAV Flash components for properties to be captured. Each of these properties will then be parsed into XML and written to a file that also contains details on which data and variables were used and a unique name for each component. When a snapshot is activated, the saved state of the Snapshot class will be read from the XML file and parse its nodes back into component properties again. The previously marked properties will then be applied and set the state of the application.

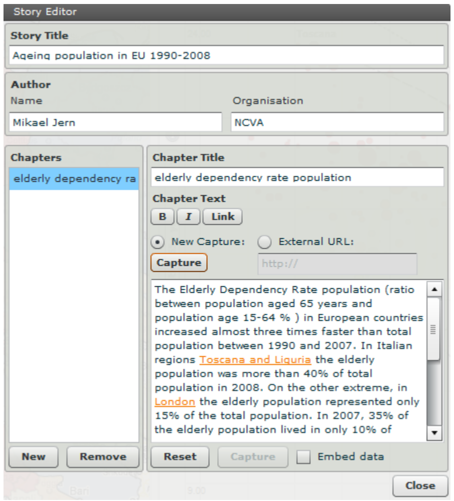


```
- <component name="filterColorLegend" type="colorLegend">
    <axis min="0" max="1" />
  </component>
- <component name="scatterPlot" type="scatterPlot2D" showTrails="true">
    <axis id="x" indicator="2" min="61.31669865642994" max="75.4616122840691" />
    <axis id="y" indicator="3" min="9.090452261306533" max="27" />
    <axis id="size" indicator="0" scale="523.6655405405405" />
  </component>
```

Figure 8. Example of a small set of XML snapshot code. Each component maintains its status.

4.2 Storytelling

Storytelling, in our context, is about telling a story on the subject of explored data and related analytics reasoning about how gained knowledge was achieved. Snapshots that instantiate a GAV Flash state are a central feature of our storytelling mechanism together with associated descriptive text that could guide the reader in the analyst's way of thinking. The author creates a single or a discrete series of captures during the explorative process by electing relevant indicators, regions-of-interest, color schema, and filter conditions focusing on the data-of-interest or a time step for temporal statistics (figure 9).



Story Editor

Story Title: Ageing population in EU 1990-2008

Author
Name: Mikael Jern
Organisation: NCVA

Chapters: elderly dependency ra

Chapter Title: elderly dependency rate population

Chapter Text
B I Link
● New Capture: ○ External URL:
Capture http://

The Elderly Dependency Rate population (ratio between population aged 65 years and population age 15-64 %) in European countries increased almost three times faster than total population between 1990 and 2007. In Italian regions Toscana and Liguria the elderly population was more than 40% of total population in 2008. On the other extreme, in London the elderly population represented only 15% of the total population. In 2007, 35% of the elderly population lived in only 10% of

New Remove Reset Capture □ Embed data
Close

Hypertext, meaning "more than just text", provides a richer functionality than simple metatext by allowing the reader to click on key words and learn about topics in the story. A story hyperlink is here a reference in the story metatext that links to an external URL web site or a captured snapshot. A hyperlink in the metatext can be related to a snapshot or an external URL.

Before the actual capture is done, the user navigates, for example, the map view to a particular country, select indicators for the scatter plot, select time step. A new view such as PCP can be added to the story etc. A “Capture” is made and all preferred states are saved. When the story later is read, hyperlinks can be initiated and the application will display the state-of-the-snapshots.

Hyperlinks that instantiate snapshots and is associative descriptive text represent a central feature in our storytelling mechanism. These could guide the reader in the analyst’s way of thinking. While it’s true that a picture is often worth a thousand words, sometimes a few words and a snapshot provide the difference between a pretty picture and understanding. This focus on publishing through assisted content creation with emphasis on visualization and metadata represents a novel advantage of our storytelling.

4.3 Publish Geovisual Analytics results

A Vislet is a standalone Flash application (widget) assembled from low-level GAV Flash components in a class library and Adobe Flex GUI tools and is represented by, for example, a single map view or a composite time-linked map and scatter plot view (figure 10). A Vislet facilitates the transition of selected tedious statistics data into heterogeneous and communicative sense-making news entities with integrated metadata and dynamic embedded animated visualization that could engage the user.

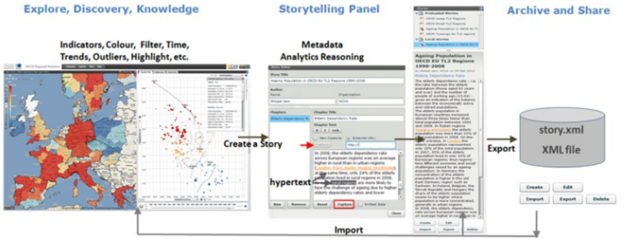
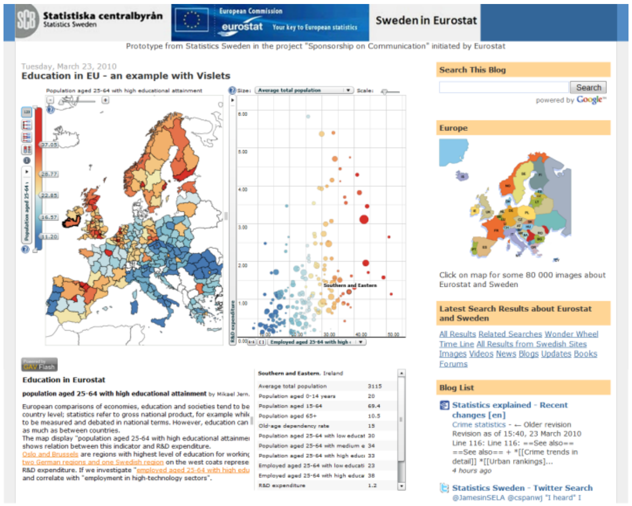


Figure 9. The GAV Flash storytelling mechanism

Publisher is the server tool that imports a story (figure 1) and generates the HTML code that represents the Vislet and metadata. First, the user selects appropriate visual representation for the Vislet e.g. map, scatter plot, parallel axes, table lens or time graph. Then the size of the Vislet window with metadata is set and finally Publisher generates the HTML code. This code snippet can then be copied and pasted into any web page system, such as a blog. The Vislet can now be opened in the reader’s Web browser and dynamically communicate the story. A Publisher server maintains the Vislet flash (swf) files together with a story repository, statistical data and regional shape maps. The Vislets run locally in the client’s Flash Player and can thus achieve dynamic interactive performance.



Statistiska centralbyrån Statistics Sweden | European Commission eurostat Your key to European statistics | Sweden in Eurostat

Prototype from Statistics Sweden in the project "Sponsorship on Communication" initiated by Eurostat

Tuesday, March 23, 2010

Education in EU - an example with Vislets

Education in Eurostat

population aged 25-64 with high educational attainment by Mikael Jern.

European comparisons of economies, education and societies tend to be country level; statistics refer to gross national product, for example whil to be measured and debated in national terms. However, education can as much as between countries.
The map display "population aged 25-64 with high educational attainme shows relation between this indicator and R&D expenditure.
Oslo and Brussels are regions with highest level of education for workin two German regions and one Swedish region on the west coats represe R&D expenditure. If we investigate "employed aged 25-64 with high edu and correlate with "employment in high-technology sectors".

Southern and Eastern, Ireland

Average total population	3115
Population aged 0-14 years	20
Population aged 15-64	69.4
Population aged 65+	10.5
Old-age dependency rate	15
Population aged 25-64 with low educat	30
Population aged 25-64 with medium e	34
Population aged 25-64 with high educ	33
Employed aged 25-64 with low educati	23
Employed aged 25-64 with high educa	38
R&D expenditure	1.2

Search This Blog
Search
powered by Google

Europe
Click on map for some 80 000 images about Eurostat and Sweden

Latest Search Results about Eurostat and Sweden
All Results Related Searches Wonder Wheel Time Line All Results from Swedish Sites Images Videos News Blogs Updates Books Forums

Blog List
Statistics explained - Recent changes [en]
Crime statistics - ← Older revision Revision as of 15:40, 23 March 2010 Line 116: Line 116: ==See also== ==See also== + *[[Crime trends in detail]] *[[Urban rankings]...
4 hours ago
Statistics Sweden - Twitter Search
@JamesinSELA @cspanwj "I heard" I

Figure 10. Example of a Vislet embedded in a web page (blog)

Interactive features in a Vislet are exposed to all visualizations including tooltips, brushing, highlight, filter that can discover outliers and dynamic multiple-linked views. Several specialist color legend tasks are supported e.g. show outliers based on 5th and 95th percentiles in certain colors or dynamic sliders that control class values etc.

5. CASE STUDIES

In this section we present three of our case studies to demonstrate how GAV Flash is used in development of visual analytics applications. These case studies have been conducted in collaboration with our partners within NCVA to produce applications used by our partners within statistics data visualization, self-organizing cellular radio network analysis and emergency scenarios.

5.1 OECD eXplorer

In Nov 2008, "OECD eXplorer" was introduced (http://stats.oecd.org/OECDregionalstatistics/), developed with GAV Flash in collaboration with OECD. It introduced common information visualization motion tools to the statistics community, such as PCP, table lens and time graph to illustrate complex statistical data. OECD explorer's user interface is divided into three main dynamic linked views (figure 11): a choropleth map, a scatter plot and a PCP. These views are separated by interactive splitters, allowing the user to scale the sizes of individual views. The scatter plot share views with a data grid and a table lens, while the PCP shares with a time graph for time animation.

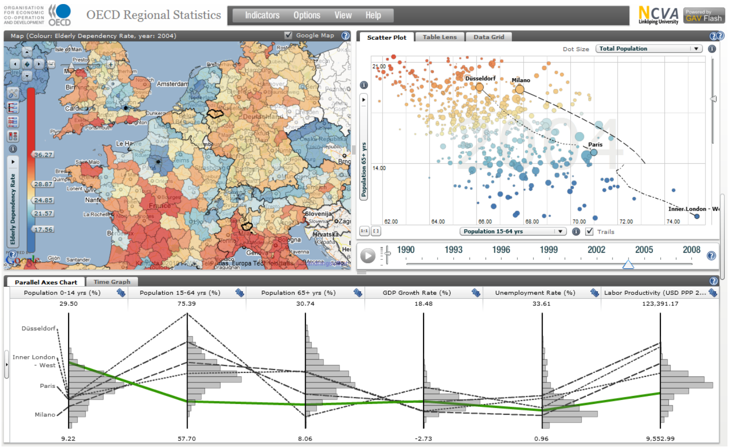


Figure 11. Example of the OECD eXplorer layout

The PCP was introduced to the statisticians supporting a number of tasks, for example, to analyze the relationships between indicators and to see a profile for selected regions. Each region is represented by a string passing through the parallel axes where each axis represents a single indicator. Differences between highlighted regions can be found by visually comparing the profiles representing them, a highly appreciated method to the statistics community.

The PCP has been extended with special features that are important to statistical exploration, such as histograms and filter operations based on percentile statistics. Histograms attached to each axis are used to visualize the distribution of indicator data, splitting the axes into a user defined number of equally high rectangular areas (bins). The width of a rectangle indicates the frequency of regions intersecting that bin. Statistical filter methods based both on ranges chosen by the user and on percentile calculations are embedded in the PCP attached to an indicator. Figure 12 shows only regions which fulfill the condition below 10th percentile and above 90th percentile, e.g. only regions representing outliers are displayed. Two regions Liguria and Paris which comply with these conditions are highlighted.

Our case study also demonstrates the storytelling process from creating a story with snapshots and metatext, saving the story and finally using Publisher to load the story and generate the HTML code that can be placed in a web site. A video showing the whole process can be found at [30] and an example of Vislet can be evaluated at [31].

Reviews from eXplorer partners who have evaluated the platform and available tools, highlight the following eXplorer characteristics as most effective:

- It enables the statisticians to (1) simultaneously analyze relations among different indicators, (2) explore trends over time and for different geographical boundaries (3) use different map layers for better locating places. It provides functions for analyzing data and benchmarking regions, presenting stories about the statistics and combining metadata and maps status.
- It captures the complexity of multi-dimensional regional data through dynamic multiple and time-link views.
- The structure of eXplorer encourages collaboration between statistics analysts and users of statistics;
- It encompasses data visualization and of possibility to capture, save and open discoveries (snapshots) with attached analytics reasoning metadata e.g. Storytelling to better support more educational use of official statistics;
- The publishing approach (through Vislets) is regarded as very attractive to a general public, since it does not require IT expertise. This publishing technique may become a strategic tool for news media to publish statistics news on the web.

- It is today a worldwide recognized tool for better understanding the socio-economic structure of OECD regions and their performance over time.

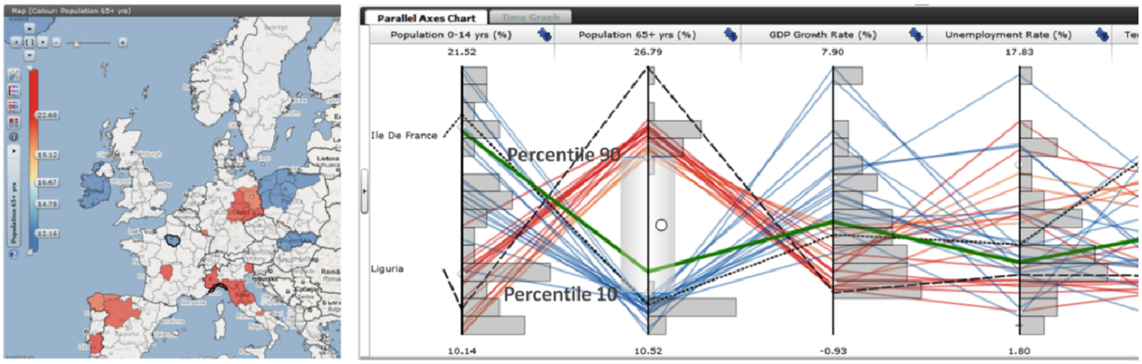


Figure 12. Regions below 10th percentile and above 90th percentile (bottom)

5.2 ANROSS

ANROSS, a new version of VoSON [25], is a visual analytics prototype developed in close collaboration with domain experts from the Swedish telecom company Ericsson to show how ANR [26] works. ANR is an algorithm invented by Ericsson to automatically detect and resolve physical cell identity (PCI) conflicts; therefore it can self-configure cellular radio networks that are typically mobile phone networks. To show how ANR work it is important to visualize and analyze the network data which are output of the process in which ANR configures a network.

The network in this context mainly includes two kinds of objects: cells and cell relations. A cell is a device that covers a geographical area and serves mobile phones in its area as well as handovers phone calls to other cells when necessary. Each cell has two identifiers: a globally unique cell identifier (CGI) and a PCI. CGIs are unique and constant over time but are more difficult and time consuming to detect for mobile terminals. In contrast, PCIs are easier to detect but they are not unique since there are at most 504 different PCIs in LTE [27]. The PCI of a cell can vary over time since it needs to be unique in a region and may need to change. A cell relation is a neighbor relation between two neighbor cells. Cell relations are used to handover phone calls from one cell (source cell) to a neighbor cell (target cell).

The network data in this context mainly includes three kinds: cell data, relation data and event data. All of them are time-varying and recorded for each time period which normally is 15 minutes or 1 hour. Cell data include cell position, cell coverage area, cell PCI, cell CGI, number of call drops, and number of out-going/in-coming handover successes/failures. Relation data include number of handover successes/failures. Event data include cell added/removed, relation added/ removed, PCI conflict detected/resolved.

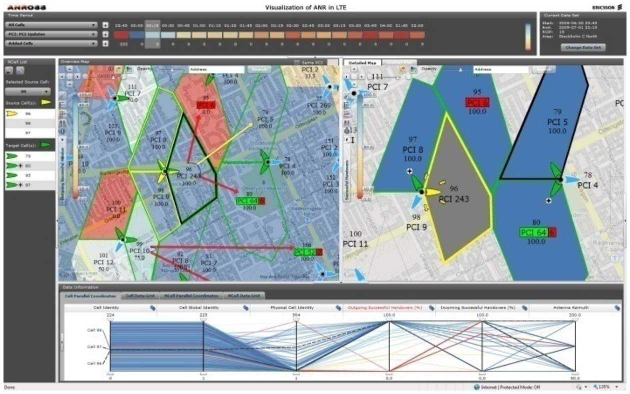


Figure 13. ANROSS layout including multiple linked and interactive views

To visualize all three kinds of data ANROSS includes multiple-linked and interactive views (figure 13): (1) an advanced time slider (top view) which displays aggregation data of each time period such as number of cells added, number of cells removed, number of relations added, number of relation removed, number of PCI conflicts detected, number of PCI conflicts resolved; through using a color map designed carefully it highlights time periods in which some indicator has extreme value; (2) a cell map (middle left view) presenting cell data and cell event data in a time period such as number of call drops, cell PCIs, cell PCI conflicts, cell PCI resolutions; (3) a relation map (middle right view) presenting relation data and relation event data in a time period such as number of handover failures of each relations, relations added, relations selected; (4) two PCPs (bottom view), one visualizing cell data and one visualizing relation data; (5) two data tables (or data grids) (bottom view), one displaying cell data and one displaying relation data; (6) a data selection panel (left) highlighting cells and relations being selected in cell map and relation map; (7) a search panel (right, hidden) which allows finding low/high performance cells or relations which will be then highlighted in the cell map or relation map respectively. Through these interactive views ANROSS allows users to see how the network evolves from different initial configurations under the control of ANR and answer various questions such as

- Why a cell changes its PCI, how a PCI conflict is detected, and what is the difference in performance of a cell (e.g. number of call drops) before and after a PCI change;
- Why a relation is added or removed, what is the performance (e.g. number of handover failures) of a relation;

In addition, through using filtering ability of PCPs and searching tool ANROSS allows users to find and supervise "problem" cells and relations such as cells/relations having a large number of call drops/handover failures, or cells having a potential PCI conflict in future.

ANROSS has been evaluated by network experts from Ericsson and has also been shown to Telia, one of the biggest network operators in Sweden. The overall feedback from both the network experts and the operator is very positive. They like the time slider, the two maps and the ability to see how the network changes over time. The ability to find 'problem' cells and relations were also considered as highly positive "-it looks useful and very good to me". As a result, they would like to collaborate to develop this prototype into product. Two videos showing the layout and functionality of ANROSS as well as how it is used to present ANR can be found at [30].

5.3 RoadVis

RoadVis is another application developed using the GAV Flash framework that is used to analyze and make decisions, often in time-critical situations, on the large and ever-increasing amounts of time-varying and geospatial digital weather information related to emergency scenarios. It has been developed and customized in close collaboration with domain experts from the Swedish Meteorological and Hydrological Institute (SMHI) and the Swedish road administration and is today used on a daily basis for analyzing and communicating information about road weather conditions, particularly during the Swedish winter months.

RoadVis consists of a multiple-linked visual user interface consisting of a collection of dynamic maps, weather and information visualization methods that enable the users to simultaneously analyze relations among several different attributes with the aim to augment an analyst and decision-maker capabilities to assimilate complex situations and reach important knowledge. The data visualized by RoadVis consists of weather observations from 770 automated observation stations around Sweden collected every half hour as well as a long 24 hour and a short six hour forecasts for each station that is made every hour. Using threading and dynamic queries the application is constantly updated with the latest data in the background and the analyst can still use the application during their vital work to keep the roads safe.

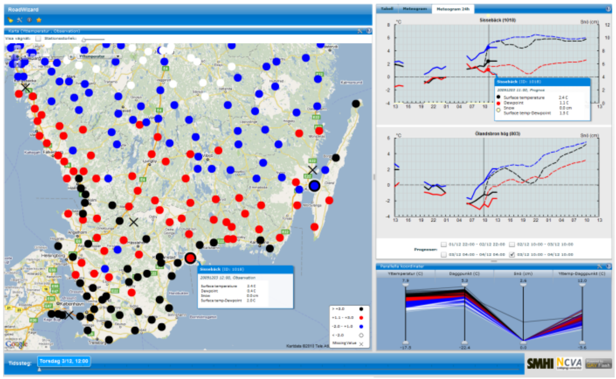


Figure 14. RoadVis multi-linked visual user interface, where a station, represented by a circle, is here colored according to temperature. Meteorological diagrams are shown for two selected stations in the map. A parallel coordinates plot makes dynamic inquires and filter the multivariate weather data. A time slider controls time steps for the time-linked views

The visual representation, as shown in figure 14, for RoadVis is based around a map of Sweden where each individual station is marked and colored depending on requested attribute value. The user can select relevant road condition such as ice slipperiness, frost, condensation and temperature etc. RoadVis then calculates both present and forecasted scenarios for the coming six hours as well as the next 24 hours using the short and the long forecast. Parallel Coordinates Plot is used to view, analyze and make visual inquiries for the multivariate weather data. Relations and trends are observed and the analyst can also filter data and discover areas with a higher risk for accidents unless preventive methods are taken immediately. Meteorological diagrams visualize both the observed values and the forecasted weather for the selected stations. A demonstration video of RoadVis can be found at [30].

6. CONCLUSION AND USER FEEDBACK

The authors expect that the three case studies introduced and now in full operation, will enhance the use and understanding of geovisual analytics applied to spatial-temporal and multivariate data. It will enable the analyst and operator to take a more active role in the discovery process of exploring. The tool will increase the interest in and knowledge of structures and development patterns among specialist as well as non-specialist users. Feed-back from domain experts from Ericsson, SMHI and OECD who have started using the tool shows that a sense of analytical reasoning and speed-of-thought interaction is achieved. Here are some features highlighted by our partners:

- A component-based geovisual analytics toolkit programmed in ActionScript for Web access;
- Ability to have dynamic time-link views and see the multi-dimensionality of complex data;
- IT expertise is not required to publish interactive visualization embedded in blogs and wikis;
- Flash implementation of practical information visualization methods PCP, Table lens, time graph etc. PCP with embedded fundamental statistics based on dynamic percentile inquiry and filtering and attached histograms;
- An architecture facilitating the analyst to explore data and simultaneously save important snapshots of discoveries or create a continuous story of snapshots (storytelling) to be communicated and shared with team or public users;

At the same time, GAV Flash will encourage the practical use of collaborative geovisual analytics through an integrated exploration, collaboration and publication process addressing editorial storytelling aimed at producing geovisual analytics-related news content in support of an automatic authoring process. The author simply presses a button to publish gained knowledge that clearly visualizes the result of an explorative data analysis process. A storytelling technology with the goal to advance research critical to, for example, official statistical production based on Vislets - embedded dynamic visualization with the analytics sense-making metadata joined together and publishable in HTML web pages such as blogs and wikis.

ACKNOWLEDGEMENTS

The applied research case studies were carried out by NCVA (http://ncva.itn.liu.se) in close collaboration with OECD, SMHI and Ericsson Research who supplied data and comprehensive evaluation of the application. The research is in part supported by funding from the Swedish Knowledge Foundation and the Swedish Agency for Innovation Systems (VINNOVA). The authors thank colleague Markus Johnsson.

REFERENCES

[1] Andrienko, G. and Andrienko, N., “Visual exploration of spatial distribution of temporal behaviors,” Proc. the 9th International Conference on Information Visualisation (IV05), 799-806 (2005).

[2] Muller, W. and Schumann, H., “Visualization methods for time-dependent data – an overview,” Proc. the 2003 Winter Simulation Conference, 737-745 (2003).

[3] Guo, D., Chen, J., MacEachren, A. M. and K. Liao, “A visualization system for space-time and multivariate patterns (VIS-STAMP),” IEEE Transactions on Visualization and Computer Graphics, 12(6), 1461-1474 (2006).

[4] OECD eXplorer. [Online]. Available: http://www.oecd.org/gov/regional/statisticsindicators/explorer.

[5] Andrienko, N. and Andrienko, G., “Interactive visual tools to explore spatio-temporal variation,” Proc. the Working Conference on Advanced Visual Interfaces (AVI 2004), 417-420 (2004).

[6] Dykes, J., MacEachren, A. M. and Kraak, M.J., [Exploring Geovisualization], Pergamon Press, Oxford, 159-180 (2005).

[7] Fekete, J. and Plaisant, C., “InfoVis Toolkit,” Proc. the 10th IEEE symposium in Information Visualization (InfoVis ’04), 167-174 (2004).

[8] Voss, H., Andrienko, N. and Andrienko, G., “Commongis - common access to geographically referenced data,” ERCIM News, (41), 44–46 (2000).

[9] GeoVista Studio. [Online]. Available: http://www.geovistastudio.psu.edu.

[10] Jern, M., Åström, T. and Johansson, S., “GeoAnalytics tools applied to large geospatial datasets,” Proc. the 12th International Conference Information Visualisation (IV08), 362-372 (2008).

[11] Tominski, C., Abello, J. and Schumann, H., "CGV-An interactive graph visualization system," Computers and Graphics, 33(6), 660-678 (2009).

[12] Stolte, C., Tang, D. and P. Hanrahan, "Polaris: A system for query, analysis, and visualization of multidimensional relational databases," IEEE Transactions on Visualization and Computer Graphics, 8(1), 52–65 (2002).

[13] Shrinivasan, Y. B. and Wijk, J. J., "Supporting the analytical reasoning process in information visualization," Proc. the 26th Annual SIGCHI Conference on Human Factors in Computing Systems, 1237-1246 (2008).

[14] MacEachren, A. M., Brewer, I. and Steiner, E., "Geovisualization to mediate collaborative work: Tools to support different-place knowledge construction and decision-making," Proc. the 20th International Cartographic Conference, 6-10 (2001).

[15] Vince, J. A. and Earnshaw, R. A., [Digital Content Creation], Springer-Verlag, New York, 140-162 (2001).

[16] Carr, D. B., White, D. and MacEachren, A. M., "Conditioned choropleth maps and hypothesis generation," Annals of the Association of American Geographers, 95(1), 32-53 (2005).

[17] Robinson, A., "Re-visualization: Interactive visualization of the progress of visual analysis," Proc. VASDS, (2006).

[18] Keel, P., "Collaborative visual analytics: Inferring from the spatial organisation and collaborative use of information," Proc. the IEEE Symposium on Visual Analytics Science and Technology (VAST), 137-144 (2006).

[19] Withall, M., Phillips, I. and Parish, D., "Network visualisation: a review," IET Communications, 1(3), 365-372 (2007).

[20] Jern, M., Rogstadius, J., Åström, T. and Ynnerman, A., "Visual analytics presentation tools applied in HTML documents," Proc. the 12th International Conference Information Visualisation (IV08), 200-207 (2008).

[21] Thomas, J. and Cook, K., [Illuminating the Path: The Research and Development Agenda for Visual Analytics] IEEE Computer Society Press, Los Alamitos, (2005).

[22] Piringer, H., Tominski, C., Muigg, P. and Berger, W., "A multi-threading architecture to support interactive visual exploration," IEEE Transactions on Visualization and Computer Graphics, 15(6), 1113-1120 (2009).

[23] GridMap. [Online]. Available: http://vitagate.itn.liu.se/GAV/gridmap/swedenzip/.

[24] Wohlfart, M. and Hauser, M. H., "Story telling for presentation in volume visualization," Proc. EuroVis 2007, 91-98 (2007).

[25] Ho, Q., Åström, T. and Jern, M., "Geovisual analytics for self-organizing network data," Proc. the IEEE Symposium on Visual Analytics Science and Technology (VAST), 43-50 (2009).

[26] Amirijoo, M., Frenger, P., Gunnarsson, F., Kallin, H., Moe, J. and Zetterberg, K., "Neighbor cell relation list and physical cell identity self-organization in LTE," Proc. the 2008 IEEE International Conference on Communications Workshops, 37-41 (2008).

[27] Dahlman, E., Parkvall, S., Sköld, J., and Beming, P., [3G Evolution: HSPA and LTE for Mobile Broadband], Academic Press, California, (2007).

[28] Viégas, F. B., Wattenberg, M., Ham, F. v., Kriss, J. and McKeon, M., "Many Eyes: a Site for Visualization at Internet Scale," IEEE Transactions on Visualization and Computer Graphics, 13(6), 1121-1128 (2007).

[29] Card, S. K., Mackinlay, J. D. and Shneiderman, B., [Readings in Information Visualization: Using vision to think]. Morgan Kaufman Publishers, San Fransisco, (1999).

[30] Demonstration videos. [Online]. Available: http://vitagate.itn.liu.se/GAV/videos/VDA.

[31] An example of Vislet. [Online]. Available: http://vitagate.itn.liu.se/GAV/vislets/html/test1.html.