

# Visual Storytelling with Snapshots applied to Official Statistics

Patrik Lundblad, Tobias Åström, Mikael Jern

<sup>1</sup>National Centre Visual Analytics, NCVA

Linköping University, Sweden

e-mail: patrik.lundblad@liu.se, tobias.astrom@liu.se, mikael.jern@liu.se

**Abstract**—The paper focuses on “visual storytelling” – exemplified through telling stories about official statistics development over time that could shape economic growth and well-being. Discoveries are made that draw the user into reflecting on how life is lived - and may be improved - from one region to another. In addition, the user can interactively participate in the web-based process, which is important to the education and dissemination of public statistics. A demonstrator World eXplorer with storytelling and an integrated snapshot mechanism is introduced, programmed in Adobe’s ActionScript and is based on the geovisual analytics paradigm. An interactive visual story mechanism assists the author to improve a reader’s visual knowledge through reflections on how life is lived by using a variety of demographics, such as healthcare, environment, and educational and economic indicators. Educators can develop interactive teaching material based on this storytelling mechanism and avoid boring statistics presentations. Integrated snapshots can be captured at any time during an explorative data analysis process and they become an important component of an analytical reasoning process. Students can access geovisual applications and explore statistical relations on their own guided by the stories prepared by the teachers. With the associated science of perception and cognition in relation to the use of multivariate spatio-temporal statistical data, this paper contributes to the growing interest in geovisual statistics analytics.

**Keywords**-Visual Storytelling, Geovisual analytics, web-based learning, statistics visualization, statistics database

## I. INTRODUCTION

The “participative web” is increasingly utilised by intelligent web services, which empower developers to customise web-enabled visualization applications that contribute to collaboration and communicate visual content (Figure 1). In this context, we introduce a collaborative geovisual analytics framework, based on the principles for Visual Analytics [1], for public statistics based on interactive statistics visualization [2] and its increasing role in sharing, collaborating and communicating gained knowledge. The collaboration and publication process facilitates storytelling aimed at producing statistical news content in support of an automatic authoring process. The author simply presses a button to publish the knowledge gained from a visual interactive discovery process. Storytelling, in our context, is about telling a story with associate snapshots about the statistics data and the related analytics reasoning processes on how knowledge was obtained. Storytelling within an

interactive web context could more engagingly draw the user into reflections and sometimes change a perspective altogether. The story is placed in the hands of those who need it, e.g., policy and decision makers, teachers and also informed citizens. Stories enable a leap in understanding by the user in order to grasp how statistical indicators may influence society. The conceptual approach and framework of the geovisual analytics storytelling implementation is based around three complementary characteristics:

- **Authoring (eXplorer):** data provider and manager, several motion visual representations, coordinated views, map layers, analytic tools (dynamic query, filter, regional categorization, profiles, highlight), and dynamic colour scale.
- **Tell-a-story:** snapshot mechanism that captures an interactive scenario (active views and indicators, attributes, time step, regions), and metadata with hyperlinks.
- **Publisher (Vislet):** import stories and create HTML code for embeddable interactive motion visual representations for publishing on a web site, interactive report or blog.

The rest of this paper is structured as follows. In section II related work is presented followed by our work on storytelling in section III. The paper is concluded with section IV presenting our conclusions.

## II. RELATED WORK

The importance to snapshot exploration sessions and then reuse them for presentation and evaluation within the same environment was demonstrated by MacEachren [3] and Jern [4] in geovisualization, and incorporated features to integrate them into electronic documents.

A variety of software is now available for creating snapshots. One of these is CCMaps [5], which is a conditioned choropleth mapping tool where the user can create snapshots of events and then reuse them for presentation. Another method is called “Re-Visualization” [6] and is used in the tool ReVise, which captures the analysis sessions and allows them to be reused. The Web-based Analysis and Visualization Environment (Weave) [7] is a Framework that uses session states and stores them on a server for later reuse. Another application that uses a similar approach where the user stores the data online is Many Eyes [8], which is a public website where novice users can upload their own data, create dynamic visualizations and participate in discussions. But, Many Eyes seems limited to showing

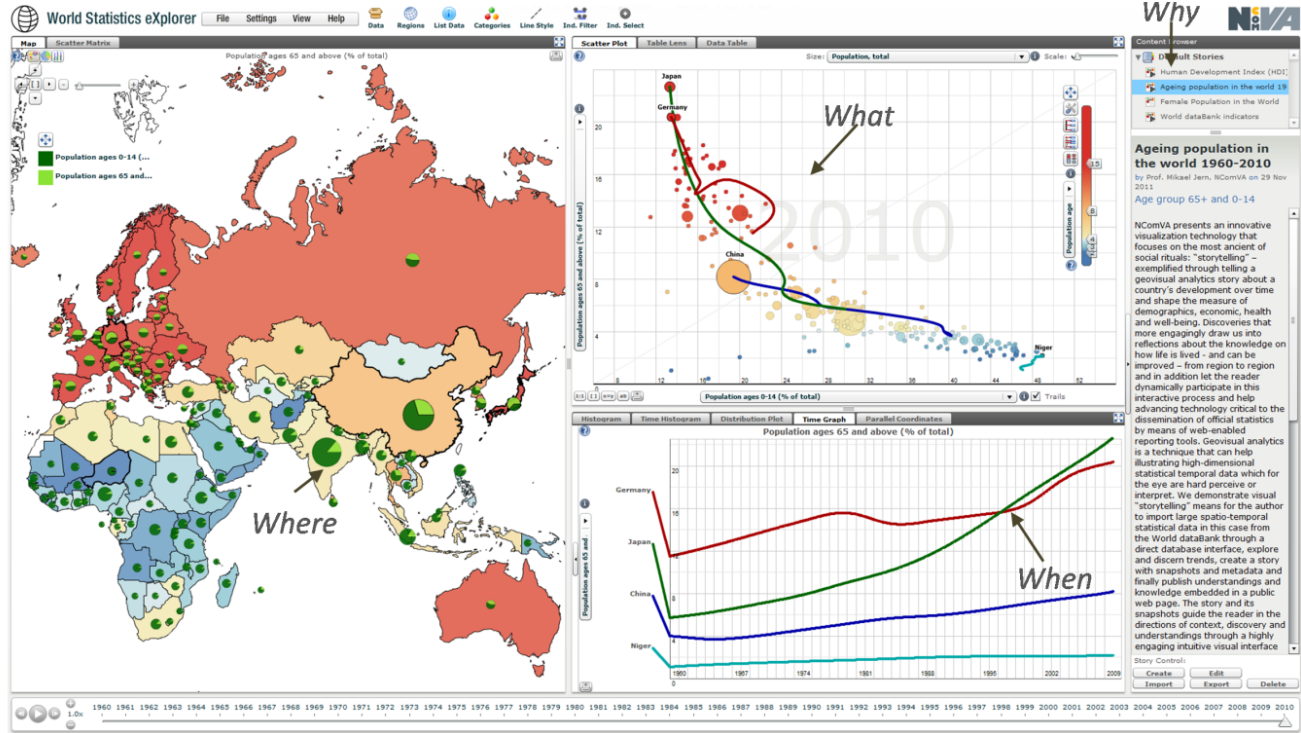


Figure 1. World eXplorer [9] based on the geovisual analytics concept “Where-What-When-Why” based on 3 time-linked views showing a worldwide ageing population during 1960-2010; map (age 65+), scatter plot (age 65+ vs. age 0-14) and time chart (65+); comparing 4 countries Nigeria, China, Japan and Germany. The story is published to the right side and includes linked snapshots. Users learn that Japan followed by Germany represents the countries with highest elderly population, while Niger has not changed. The users can interact and change indicators to discover reasons behind this trend.

only one visualization at a time and has no animation facilities for time sequences. For many of these systems, the snapshot has to be loaded into the same application environment as the one that created it, which puts a restriction on usage and sharing if the application requires a software licence. Such applications may not be easily accessible to team members without installing external software [10]. In this context, we introduce a web compliant layered component toolkit with a snapshot mechanism that captures, re-uses and shares active properties for individual functional components. We have demonstrated [11] that such an implementation could provide a more open and collaborative geovisual analytics framework for public use. Collaborative geovisual analytics tools have been proved to work well with statistics data [12]. Initial tests have also shown significant potential [13, 14] when making them accessible to educators and their students. With the availability of current web enabled geovisual analytics tools it is appropriate to explore the possibilities of using these tools within schools and allow educators to use these tools in different application areas [15]. There is significant future potential for educators to present and explore scientific data sets together with students.

### III. STORYTELLING

Storytelling within a participative web context, could more engagingly draw the user into exciting reflections and sometimes change a perspective altogether. The story is placed in the hands of the users as an interactive guided

learning experience to achieve a more complete understanding of the knowledge using descriptive metadata integrated with “memorized interactive visualization views” or “snapshots” and external web links to relevant information. Visual storytelling is in this scenario an approach of telling more vital and engaging stories through interactive web-enabled visualizations. The snapshot mechanism helps the author of a story to highlight data views of particular interest and subsequently share or guide others to significant visual discoveries. These interactive events in an analytical reasoning process can be captured at any time during an explorative data analysis process and represent an important part of an analytical storytelling process (Figure 2).

#### A. Snapshots

When exploring and making sense of comprehensive statistics data there needs to be a way of storing discoveries in a coherent and cognitive workspace, which can be organized, navigated and maintained within the application so that discoveries may later be loaded. The GAV Flash toolkit [16, 17] includes such means by giving the analyst the option of capturing, saving and packaging the results of an exploration “gain insight” process. The insights are captured 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. Snapshots are created as a series of visualization captures during the exploration process and form the foundation for a story.

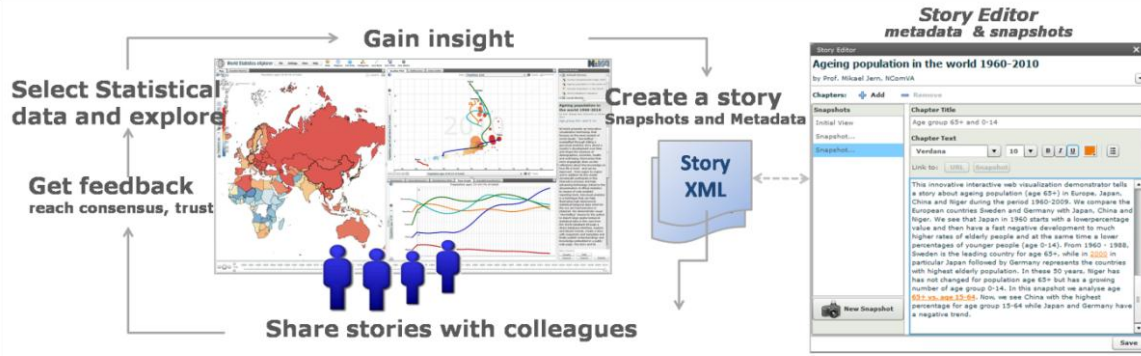


Figure 2. The analyst is the author in this storytelling loop

In a typical scenario the author selects relevant attributes for a snapshot, e.g., time step, highlighted regions for comparisons, class values for colour legend, filters inquiry conditions for a reduced dataset and finally highlights the “discoveries” in the world map view from a certain angle. When the analyst presses the capture button all the components within the application stores their properties, thus creating a state, which can later on be recreated (Figure 3). The snapshot is then inserted in the story by creating a hyperlink in the text that is highlighted. When the reader follows the story the reader can click the highlighted link and the state of the visualization when captured is recreated so the reader has the same view as the author. The reader may then also make changes to the application such as changing filter values or highlighting other areas of interest. This new discovery may then be stored as a new snapshot, which can be inserted into the story or replace the old snapshot.

In our story editor (Figure 6) the author has access to all stored snapshots in a list on the left side and can easily go to one of them and change the parameters. The author can also format the text and insert url links to other resources on the web. A story may also contain multiple chapters so that multiple discovery processes can be stored in the same story.

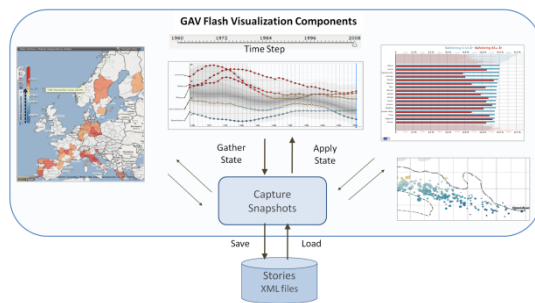


Figure 3. The snapshot tool creates series of captures during an exploration process that form the story.

**B. Publish Vislets**

Publisher is the tool that takes a story created by World eXplorer and generates the HTML code that can be used to publish a Vislet on the web (Figure 4). A Vislet is a standalone Flash application (widget) that has been

assembled from low-level GAV Flash components in a class library (Figure 5).



Figure 4. The Exploration and Publishing process

When the user has a finished story that they want to publish they login in to the publisher portal and upload their story to the server. In publisher the user first chooses the layout of the Vislet to be used for the visualizations. This may be everything from a single view layout up to a divided area with multiple visualizations linked to each other. The user then chooses for each view the appropriate visualization to be used. After that the user has some extra options for functionality within the Vislet. When the user is ready the publisher then deploys the story as HTML code, which has all required links to geographic and data sources. The code can then be pasted into the user’s favourite web site. A Vislet is created and can now be opened inside the reader’s Web browser and communicates the story.

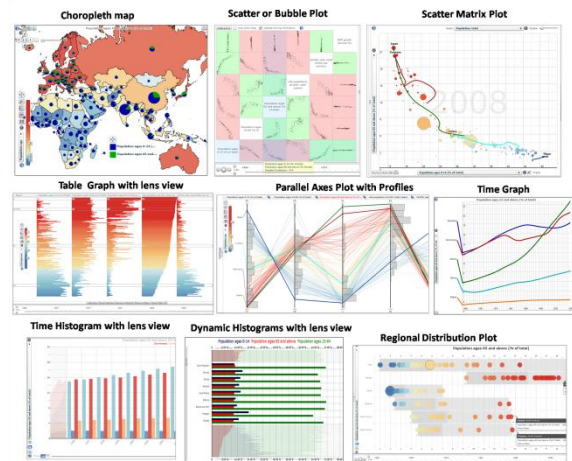


Figure 5. Example of standalone Vislets for publication

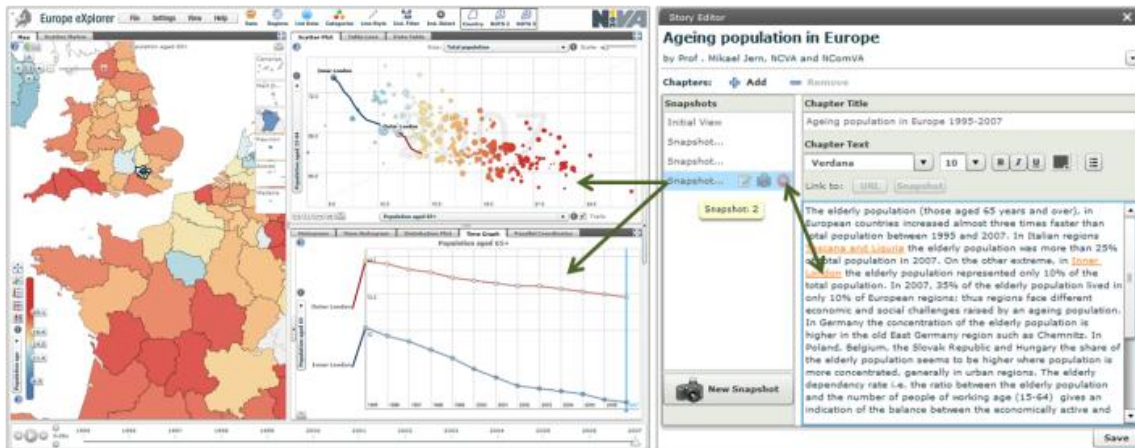


Figure 6. Story Editor with the associate Snapshot pointing to “Inner London” in all three views in this regional European visual analytics. The author has access to all stored snapshots in a list on the left side and can easily go to one of them and change the parameters. The Snapshot is linked to the metatext “On the other extreme, in Inner London the elderly population represented only 10% of the total population”. Clicking on this Snapshot in a published Vislet immediately initiates all eXplorer views to the state given by the Snapshot.

C. Interactive document

An interactive document is based on a wider storytelling concept, where Vislets play the role of images and figures. This adds another depth to the report or publication by making diagrams interactive, which allows the reader to reach a deeper understanding and further explore the subject. Readers can run animations, change indicators, and view more details on specific figures.

One such publication where an interactive document has supplemented the normal paper version is Organisation for Economic Co-operation and Development (OECD) Regions at a Glance (RaG) [18]. This bi-annual report now has an interactive version where selected chapters have been transferred from the publication to an interactive state and published on the web (Figure 7).

The OECD analysts created stories and snapshots from the RaG data, wrote new or edited existing text and collected interesting links. The stories were used to create chapters in the interactive document, where visualizations reside together with the analysts text, relevant links, the corresponding chapter in the publication as a pdf and access to the source data

The interactive document platform used in this case is built to be simple, flexible and portable. Chapters are stored as eXtensible Markup Language (XML) files with lists of their internal sections, the text to display, and the settings for the visualization. These chapters are transformed to html using eXtensible Stylesheet Language (XSLT). This setup might sound simplistic, as using a data base would allow for more advanced and easier updates, but the choice was made with portability in mind. The aim was to allow for very simple installation, no creation of data bases and very limited configuration. In fact, the only setup needed is writing base urls and switching between using Hypertext Preprocessor (PHP) and Active Server Pages (ASP) for running the XSLT.

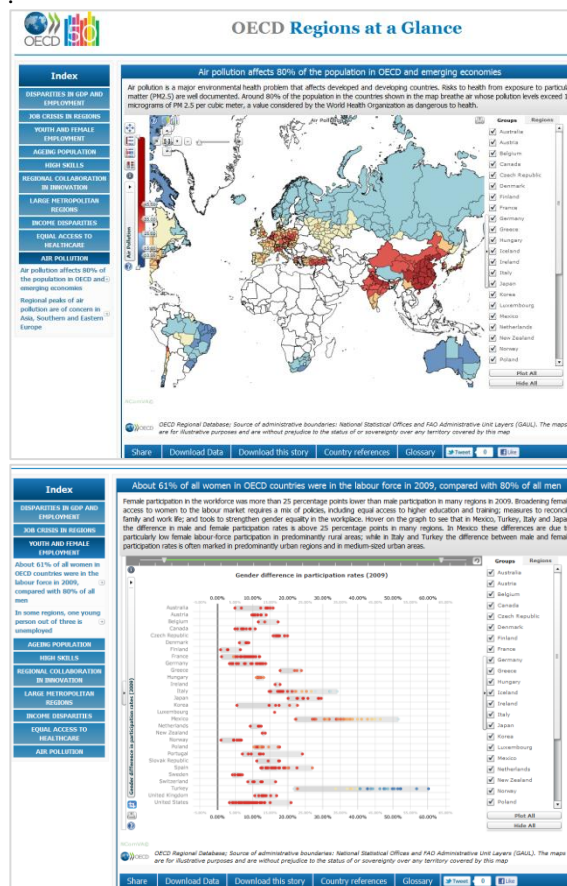


Figure 7. OECD interactive visualization report

Although the choice was made to not use an elaborate system for this publication, different approaches are available. The Vislet technology can be fully integrated into any existing content management system to achieve similar results.

The main benefit of making the report interactive is that it enables the user to visualize the data that is of interest to them, and also to show the variation over time using animation. The analysts at OECD have chosen snapshots that are of interest, but the user can interact with the visualization and change the indicators and select what they find interesting, and thus enhance the user experience.

IV. CONCLUSION

The primary objective of the introduction of visualization is to avoid boring statistics. We want official statistics to be exciting so that they invade people’s minds and create knowledge, thus allowing users to apply new insights as a basis for decisions. In order to do this the visualization should highlight structures and patterns in the data and allow the users to play around and test their own hypotheses and ideas. It is well known that interactive web based maps speak to the minds of many people. In addition to maps showing the data, we want to highlight interesting correlations through our snapshot mechanism between several indicators across the geographic areas, as well as dynamics over time. If users are to devote interest to statistics, gain knowledge, and use the information for decisions, the statistics must clearly convey a message. The visual analytics storytelling technique applied to statistics visualization therefore is essential.

Focus has historically been on text and picture comprehension but given the explosion of representations made available since the introduction of graphical interfaces, the field now considers all forms of representation including but not limited to, text, pictures, graphs, diagrams, concept maps, animations, equations, virtual reality, information-and scientific visualization, haptics, multimedia, hypermedia, and simulations. Research on learning, when using these aids, is essential. There is research of learning with multimedia environments in different experimental studies but there is hardly any research done of this in real school contexts i.e., in a socio cultural perspective. NCVA is also engaged in a project about social science education in Swedish schools [19].

Within an international perspective our research builds upon collaborating work with many world-leading statistical organizations including the Eurostat, European Commission, OECD [20], [21], Statistics Denmark, Finland, Sweden [22] providing innovative geovisual analytics technology to these organizations. We have also been involved in the development of visualization for the PISA2009 profiles [23]. The national Italian bureau (ISTAT) provides another very interesting and sophisticated learning material [24] about the development and progress in Italian regions in (Figure 8).

To ensure that users would come forward and contribute to a good implementation of geovisual analytics methods applied to statistics, NCVA established a small group

consisting of users with different profiles including experts from international and national statistics organizations (OECD, ISTAT, Sweden Statistics), regional researchers, municipal planners and in 2011 was extended to also include school teachers and students. The group met first time already in 2010 and was invited to test and evaluate solutions based on Statistics eXplorer. The feedback has been very positive, as most user categories found that services along these lines would greatly help them understand and use the proposed visualization and storytelling methods. They also gave useful input regarding aspects of the web-enabled presentation, some asking for additional and more complex methods and interactive features, others for simplifications. This confirms the expectation that we must have different levels of service for different segments of users. This type of customization is possible to manage with our fundamental component based tools to assemble the eXplorer applications.

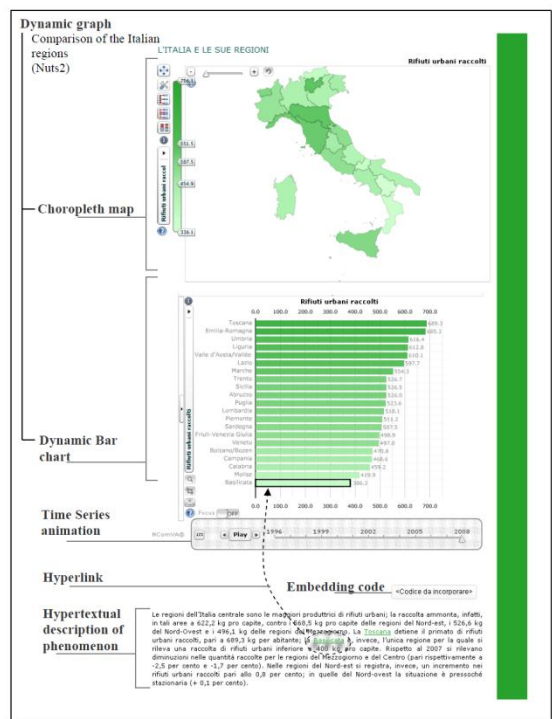


Figure 8. ISTAT – the Italian national statistical bureau has developed this innovative Vislet web site providing an interesting example of a sophisticated learning material.

This test and comprehensive evaluation practice has then continued over several new improved versions. Here are some key features that are highlighted by our partners:

- Mechanism for Storytelling and Snapshots for creating dynamic Web documents;
- Identify what’s meaningful in the data in ways that your eyes can see and brain can understand and keep focused on what’s important;
- Support large spatio-temporal and multi-dimensional regional data using a data cube;

- Established information visualization technology for multivariate data is adapted and customized to the statistics domain e.g., histogram with fish eye technique, table lens, parallel coordinates with profiles, scatter matrix linked to the scatter plot, choropleth map linked to a treemap;
- Data is simultaneously explored through multiple-linked and coordinated views;
- Map layer architecture - overlay several types of maps such as coloured statistical regions, country boundaries, background maps e.g., Google, pie chart – control transparency level for each layer;
- Dynamic state-of-the-art colour legend with integrated statistical methods such as percentiles;
- Visual inquiry and conditioned statistics filter mechanisms highlighting important discoveries or detecting outliers;
- Interactive time animation simultaneously visualized in all active views;
- Screen space usage is optimized for visualization–no unnecessary visible GUI panels;
- Support for categorical data and visual classification of, for example, urban, rural and intermediate regions;

#### ACKNOWLEDGMENT

This research and case studies were carried out by NCVA in close collaboration with OECD and Statistics Sweden who supplied data and comprehensive evaluation of the statistical storytelling system. The research is in part supported by funding from the Swedish Agency for Innovation Systems (VINNOVA), the “Visualization Program” coordinated by the Swedish Knowledge Foundation and the Swedish foundation “Marcus och Amalia Wallenbergs Minnesfond” for the educational project VISE “Visual Storytelling in Education”. The authors thank the entire research team at NCVA, Linköping University for valuable contributions and Linnéa Stenliden VISE.

#### REFERENCES

- [1] Andrienko G., Andrienko N., Demsar U., Dransch D., Dykes J., Fabrikant S. I., Jern M., et al. (2010). “Space, Time and Visual Analytics”. *International Journal of Geographical Information Science*, 24(10), 1577-1600. Taylor & Francis.
- [2] Jern M., Thygesen L., Brezzi M. “Storytelling – “How to visualize statistics”, Reviewed paper, NITS 2011 (New Techniques and Technologies for Statistics), international scientific conference on the impact of new technologies on statistical collection, production and dissemination systems, Brussels, February 2011.
- [3] Maceachren A. M., Brewer I., Steiner E. (2001). “Geovisualization to mediate collaborative work”. *Tools To Support Different-Place Knowledge Construction and Decision-Making*. *Environment*, (3), 2533-2539. ICA.
- [4] Jern M., “Smart Documents for Web-Enabled Collaboration”, Published in “Digital Content Creation”, Vince J. A. and R. A. Earnshaw (Eds) Springer Verlag, June 2001
- [5] Carr D., White D., and MacEachren A.M., “Conditioned choropleth maps and hypothesis generation”. In *Annals of the Assoc of Am. Geographers*, vol 95, no. 1, 2005.
- [6] Robinson A., “Re-Visualization”: Interactive Visualization of the Progress of Visual Analysis, workshop proceedings, VASDS. 2006.
- [7] Baumann A., The design and implementation of Weave: A session state driven, web-based visualization framework, University of Massachusetts Lowell, 2011
- [8] Vie’gas FB., Wattenberg M., Ham FV., Kriss J. and McKeon M. “Many eyes: A site for visualization at Internet scale”. *IEEETrans Vis Comput Graph* 2007; 13(6): 1121–1128
- [9] World Statistics eXplorer with ageing population in the world: <http://www.ncomva.se/v4/world/alt.html#story=1>
- [10] Jern M., Rogstadius J., Åström T., and Ynnerman A: “Visual Analytics presentation tools applied in HTML Documents”, Reviewed proceedings, IV08, London, July 2008, published by IEEE Computer Society.
- [11] OECD web site for visualization of regional development: <http://www.oecd.org/GOV/regionaldevelopment>
- [12] M. Jern, “Collaborative Educational GeoAnalytics applied to large statistics temporal data”, Reviewed proceedings, CSEEDU 2010, Conference on computer supported education, Valencia, April 2010
- [13] Jern M, “Educating students in official statistics using embedded geovisual analytics storytelling methods”, Reviewed Proceedings in Eurographics 2010, Norrköping, May 2010.
- [14] Stenliden L., Jern M., “Educating official statistics using geovisual analytics storytelling methods”, Reviewed proceedings, International Technology, Education and Development Conference INTED, Valencia, 2010.
- [15] Kinzel M., Wright D., “Using Geovisualizations in the Curriculum: Do Multimedia Tools Enhance Geography Education?” Paper Number 1290, Environmental Systems Research Institute Education User’s Conference, 2008.
- [16] Quan H., Lundblad P., Åström T., and Jern M., “A web-enabled visualization toolkit for geovisual analytics”. *Journal Information Visualization*. Published online before print November 9, 2011, doi: 10.1177/1473871611425870.
- [17] Quan H., Lundblad P., Åström T., and Jern M., “A Web-Enabled Visualization Toolkit for Geovisual Analytics Visualization and Data Analysis”, Awarded best paper, SPIE: Electronic Imaging Science and Technology, Visualization and Data Analysis, Proceedings of SPIE, San Francisco Jan 2011
- [18] OECD interactive report of regional development <http://rag.oecd.org/>
- [19] Stenliden L., Jern M., “Visual Storytelling – Understanding and Knowledge in Education”, Awarded best paper, International Symposium on Academic Informing Science & Engineering: isAISE 2011 in the context of The 2nd International Multi-conference on Complexity, Informatics and Cybernetics: IMCIC 2011, Orlando, Florida, USA. ISBN-13: 978-1-936338-21-4.
- [20] Jern, M., Thygesen L., Brezzi M., “A web-enabled Geovisual Analytics tool applied to OECD Regional Data”, Reviewed Proceedings in Eurographics 2009, Munich, March 2009
- [21] OECD web site for visualization of regional statistics: <http://stats.oecd.org/OECDregionalstatistics/>
- [22] The Statistic Atlas from Statistics Sweden [http://www.scb.se/Kartor/Statistikatlas\\_KN/index.html](http://www.scb.se/Kartor/Statistikatlas_KN/index.html)
- [23] OECD PISA 2009 profiles by country/economy: <http://stats.oecd.org/PISA2009Profiles/>
- [24] Our italy, by the national Italian bureau ISTAT: <http://noi-italia.istat.it/>