Educating students in official statistics using embedded geovisual analytics storytelling methods

Mikael Jern[1]

[1]NCVA – National Center for Visual Analytics, ITN, *Linkoping University, Sweden*

Abstract

Official statistics such as demographics, environment, health, social-economy and education from national and sub-national sources are a rich and important source of information for many important aspects of life and should be considered to be more used and acknowledged in education. Educators and their students would be able to get informed and at the same time participate in increasing the knowledge on how life is lived and can be improved. A lot of this statistics information can be reached on the Internet. This is producing what is often called information overload and causing people to be increasingly faced with the problems of filtering and interpreting enormous quantities of information. We know that official statistics are used as a more or less important background for decisions especially in public planning and policy making. However, in education, official statistics are much less recognized and used than they ought to be and among the informed public they are even less used. Web-enabled GeoAnalytics is a technique that can help illustrating complex statistical data which for the eye are hard to uncover or even are not possible to perceive or interpret. In this paper, we introduce novel "storytelling" means for the author (educator) to 1) select any spatio-temporal and multidimensional national or sub-national statistical data, 2) explore and discern trends and patterns, 3) then orchestrate and describe metadata, 4) collaborate with colleagues to confirm and 5) finally publish essential gained insight and knowledge embedded as dynamic visualization "Vislet" in blogs or wikis with associate metadata. The author can guide the reader in the directions of both context and discovery while at the same time follow the analyst's way of logical reasoning. We are moving away from a clear distinction between authors and readers affecting the process through which knowledge is created and the traditional models which support editorial work. Value no longer relies solely on the content but also on the ability to access this information. Audiences are increasingly gathered around Web enabled technologies and this distribution channel is, more than ever, in control of the information value chain.

Categories and Subject Descriptors: I.3.2 Distributed/network graphics, I.3.6: Interaction Techniques, I.3.8: Applications

1. Introduction

We live in a data-rich global world where students have become familiar with notions like GDP and sustainable development. There are surveys on the socio-economy progress and economic performance. At the same time, educators want to lecture about statistics that capture the quality of their students own lives, taking into account a broader perspective beyond the economic one. How can official statistics be made available in a way to change the terms and structures for learning, as well as for the young generation?

The role that geovisual analytics (GeoAnalytics) plays in our research is not always paid attention to when it comes to new ideas of education. Our daily life is reflected by many decisions in government administration and planning that are based on official statistics. Indeed the real value of statistics depends on whether or not they influence our opinion in some way – especially in generating decisions leading to desired results rather than the opposite. However, in education, official statistics are much less recognized and used than they ought to be and among the informed public they are even less used. The potential is vastly under-utilised and there can be several explanations to this. Do educators know about existing statistics? Can they find them and then apply them in their daily lectures? – Not as easily as statisticians tend to think. And if they eventually obtain them, do they actually understand them in such a way that they can bring statistics knowledge to their students? These are issues that should be dealt with in official statistical organisations. The Internet provides a communication infrastructure which offers the opportunity to also participate in the production of content. A better understanding of how educators and their students can elicit better user understanding and participation by exploiting dynamic web-enabled visualization and its associated science of perception in learning is the focus of this paper in relation to the use of multidimensional spatio-temporal statistical data. Tools are introduced that help and engage educators to communicate progress initiatives measuring economic, social, educational, health and environmental developments to students.

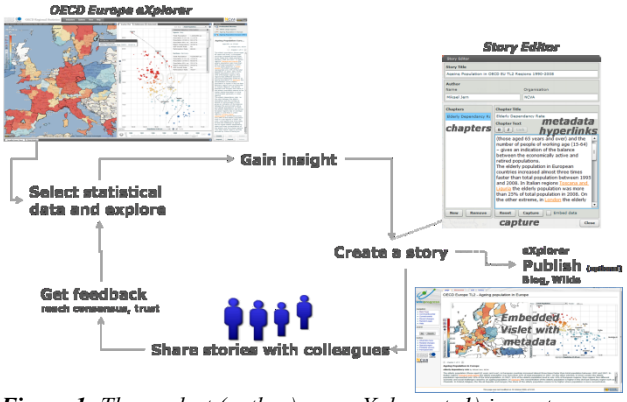


Figure 1: *The analyst (author) uses eXplorer to1) import any national or sub-national statistical data, 2) explore and make discoveries through trends and patterns and derive insight. Gained knowledge is the foundation for 3) creating a story that can be 4) shared with colleagues and reach consensus and trust. The visual discoveries are captured into snapshots together with descriptive metadata and hyperlinks in relation to the analytics reasoning. The author gets feedback from colleagues, adopts the story and 5) finally publishes "tell-a-story" to the community using a "Vislet" that is embedded in blogs or wikis.*

The term 'Web 2.0' has become undisputed linked with developments such as blogs, wikis, social networking and collaborative software development. While the benefits of GeoAnalytics tools are many, it has been a challenge [JTB08] to adapt these tools to the Internet and reach a broader user community. Research has so far focused more on tools that explore data [AA05] [GCM*06] while methods that efficiently publish gained knowledge have not achieved the same attention. Publication is the part of the analytical process that is visible to the consumers and the visual sparks it generates could take on new value in a social setting and become a catalyst for discussion. Emerging GeoAnalytics tools need to address challenges in support of both editorial and related authoring process with the goal to advance research critical to educational production and publishing. Seamless integration of exploration, collaboration and publication is required. The storytelling mechanism described in this paper could enable the transition of tedious statistics data into heterogeneous, open and communicative sense-making news entities with integrated contextual metadata that will emphasize on content creation aspects such as aesthetics analysis or "infosthetics" and where dynamic embedded visualization will engage the user.

In this context, we build upon previous collaborative research [*Jer01*], [JRÅ08] and [JPR02] and our latest web-enabled application "*OECD eXplorer*" [JTB08] and [JB09]. The eXplorer platform is customized for interactively analyzing and collaborating gained insights and discoveries about statistics events using a snapshot mechanism. A first version of eXplorer was published in November 2008 at the OECD web-site and gained positive attention. An extended version was released in October 2009 that includes many new features based on user evaluation and suggestions. The *Open eXplorer* platform is now emerging as a de facto standard in the statistics community for exploring statistics data.

In this paper, we now extend previous research in response to requirements from eXplorer users to integrate seamless publication of stories. This new feature based on embedded dynamic visualization and conceptual metadata paradigms could elicit understanding and participation and also have a positive impact on the accessibility and notions of trust of statistics information to education and citizens. Students will also be able to discuss relevant issues thus raising awareness and increasing the common knowledge on a certain phenomenon.

In summary, this paper introduces, a novel storytelling mechanism for the author (educator) to first explore large, spatio-temporal and multidimensional statistical data, then orchestrate and describe metadata, collaborate with colleagues to reach consensus and finally publish essential gained insight and knowledge as a Vislet [*ageing population in Europe*] with metadata and two chapters embedded into blogs or wikis. This seamless integration (figure 1) of an authoring, storytelling and publishing solution is based on experience and prototypes developed during the last decade and could advance the state of the art in statistical geovisualization publishing facilitating:

- **Authoring**: data provider (spreadsheet and database), data manager, choropleth map, scatter plot, table lens, parallel coordinates plot (PCP), time graph, data grid, coordinated views, map layers, analytic tools (dynamic query, filter, regional categorization, profiles, highlight, motion charts), advanced dynamic colour scale and legend, create HTML code for Vislet.
- **Storytelling:** snapshot mechanism, metadata with hyperlinks, story and chapters, edit, capture, save, export story, embed story.
- **Vislet:** embeddable interactive motion visual representations based on statistical data including choropleth map, scatter plot, parallel coordinates (profile plot), table lens and metadata for publishing in blog, wikis etc. HTML code that characterizes the story-to-be-told is automatically created by the eXplorer authoring tool.

This paper is organized as follows: Section 2 describes the needs for such a tool as seen from the OECD followed by related work in Section 3. The architecture and main components of the storytelling are described in Section 4, an example for producing a Vislet in Section 5. Section 6 recalling the main issues emerged during the evaluation phase of OECD eXplorer and how they have been addressed. The last section concludes and sets out future project developments.

2. The user perspective

The interest for global and regional progress has increased in recent years [*regional statistics portal*]. The performance of regional economies and the effectiveness of regional policy help determine a nation's growth and shape the measure of well-being across countries. OECD [Odata09] and Eurostat study sub-national disparities and development patterns, while a community uses official statistics to learn, plan and communicate for the development of local regions.

Analysts want to create content and express themselves through "user-created knowledge" and have a more pro-active, collaborative role in content creation, distribution and shared use. More active users and user-centred innovation could have increasing social impact and importance. A possible target group is educators, who can make use of this tell-a-story tool to publish trusted official statistics with metadata and increasing the knowledge on how life is lived – and can be improved – from region to region.

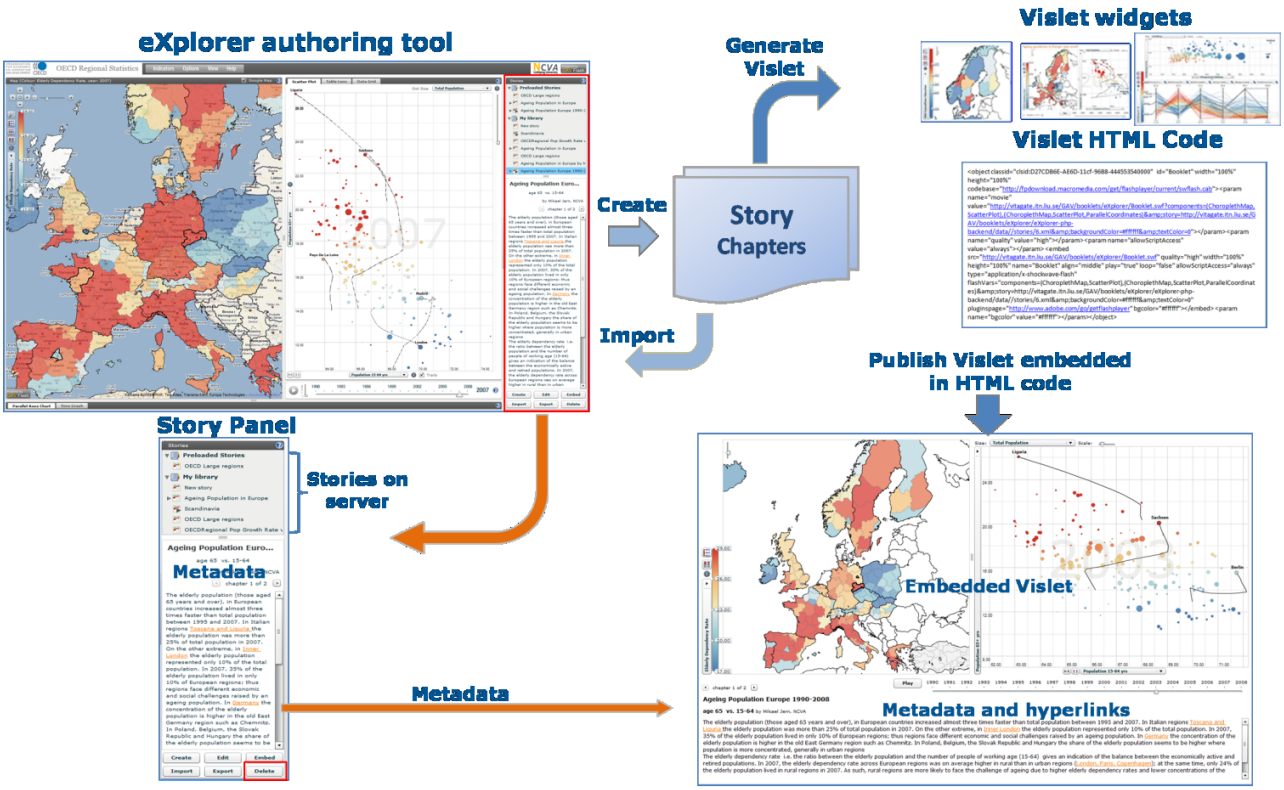


Figure 2: *From creating a story with the authoring tool "EU OECD eXplorer" to a published Vislet embedded in a Web page with associate metadata and hyperlinks. This three-step approach that seamless integrates tools for "story authoring" and "story telling" provides an education platform for both analysts and teachers to publish highly interactive statistical educational material.*

3. Related work

Volumes of official national and sub-national statistical data are today generated by statistics offices all over the world and stored in public databases such as the *OECD Regional database* but not used as effectively as one would wish for. Little focus has been given to make GeoAnalytics technologies useful and accessible to statisticians and advance visual presentation to the public. For example, information visualization research methods, such as the PCP used, for example, in CommonGIS [AA03] or the table lens are still unknown methods to the broad statistics user community and could be adopted to efficiently visualize multidimensional statistics patterns [Eds03],[AAF*06],[JJ07].

Despite the advances in Web graphics technologies, comparatively little research has been focused on applying more sophisticated GeoAnalytics applied to official statistics users. The most well-known tool is *GeoVISTA Studio*, an open source Java-based visual programming environment and is commonly used for developing geospatial applications within the research community. Rosling [Ros06] dramatically increased the awareness of geovisualization among the world of statisticians for education purposes by exploring the *Gapminder*, a time-animated bubble plot that was acquired by Google and now free available as a Google motion widget. But generally, the geovisualization research community has failed to bridge the gap between research and the official statistics analysis community.

The importance of a capacity to snapshot explorative sessions and then reuse them for presentation and evaluation within the same environment was early demonstrated by MacEachren [MB01] and Jern [J01] in geovisualization and incorporated features to capture and reuse interactions and integrate them into electronic documents. CCMaps [CWM05] presents a conditioned choropleth mapping tool that allows users to save snapshot events and reuse them for presentation purpose. Another effort was made by *Visual Inquiry Toolkit* [GCM*06] that allows users to place pertinent clusters into a "pattern-basket" to be reused in the visualization process. [Rob06] describes a method they call "Re-Visualization" and a related tool ReVise that captures and re-uses analysis sessions Keel [Kee06] describes a visual analytics system of computational agents that support the exchange of task-relevant information and incremental discoveries of relationships and knowledge among team members commonly referred to as sense-making. Wohlfart [WH07] describes a storytelling approach combined with interactive volume visualization and an annotated animation.

Many capture and reuse approaches are limited to be used within the same application environment that may well require a software license and are not always easily accessible to team members without installing external software [JJÅ08]. Increased computer security practice for statisticians could also limit this possibility.

Research has so far focused on tools that explore data [AA05] while methods that publish gained knowledge with clarity, precision, and efficiency has not achieved the same attention. The next generation of GeoAnalytics tools should also address challenges in support of an integrated editorial and related authoring process with the goal to advance research critical to educational production and publishing.

4. Conceptual implementation

The conceptual approach to our authoring and publishing concept is based around three complementary characteristics:

- **eXplorer**: Represents the expert and authoring process facilitating web-enabled GeoAnalytics technologies to explore national or regional statistics data, collaborate gained insight to colleagues, adopt notions of consensus and trust and finally deliver a story about an important statics event that could enlighten the community. eXplorer facilitates the visual analysis of spatiotemporal and multi-dimensional statistics patterns through highly interactive and time-linked visual representations (figure 3).
- **Storytelling**: Through the story mechanism facilitating descriptive metadata, textual annotations hyperlinked through the snapshot mechanism and integrated with interactive visualization, the author can guide the reader in the directions of both context and discovery while at the same time follow the analyst's way of logical reasoning. Support means of saving, packaging and sharing the gained knowledge of an explorative GeoAnalytics process.
- **Vislet**: An embeddable and stand-alone application (widget) implemented with Web 2 technologies Flash and HTML. Publish dynamic visualization with integrated metadata into blogs, wikis or enrich any editorial HTML content thus eliciting easier and better user understanding and participation of statistics reasoning and knowledge.

4.1. eXplorer – a GeoAnalytics expert and authoring tool

Figure 3: *Explorer is developed from GAV Flash components.*

The eXplorer application platform is customized from our Web-enabled GAV Flash class library, programmed in Adobe's object-oriented language ActionScript and includes a collection of common geo- and information visualization representations such as choropleth map, scatter plot table lens, parallel coordinates plot (PCP), time graph, pie chart, table grid (figure 3) etc. Temporal statistical data are analysed through the use of multiple-linked views [Rob04], [JJP*05]. Complex patterns can be detected through a number of different visual representations simultaneously, each of which is best suited to highlight different statistics pattern and can help stimulate the analytical visual thinking process so characteristic for GeoAnalytics reasoning. All graphs are time-linked, an important feature in the synthesis of animation within explorative statistical data analysis targeted towards large multidimensional temporal data.

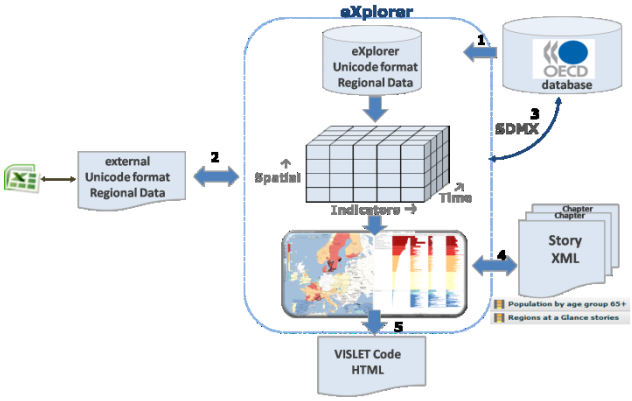


Figure 4: *External and internal data management in eXplorer*

Another key component is the *data providers* (figure 4) that open eXplorer to external statistical data sources and databases (*SDMX*) – an often under-estimated task. The Unicode format allows any type of geospatial statistics data to easily be loaded. The eXplorer data model is optimized for handling spatio-temporal and multivariate indicator data in a GeoAnalytics context [JJ06]. This conceptual data model can be seen as a data cube with three dimensions: space, time and indicators. The spatial dimension is represented by the regions and the indicators are various demographics measurements (GDP growth, elderly dependency rate, etc). Time is the data acquisition period. The general method for finding a value in the cube is by its position (space; time; indicator;). Space-time-indicator awareness means that the data cube can be analysed and visualized across all three dimensions simultaneously. eXplorer performs this task by integrating and time-linking all its motion graphs (figure 3): choropleth map, parallel coordinates, scatter plot, table lens etc.

Interactive features that support a spatial analytical reasoning process are exposed such as tooltips, brushing, highlight, visual inquiry, conditioned statistics filter mechanisms that can discover outliers and methods supporting dynamic linked multiple views.

Of particular interest is the introduction of the common information visualization methods table lens and PCP [Eds03], to a great extent unknown to the statistics community. In this context, our GAV Flash PCP component (figure 5) has been extended with special features that are important to statistics exploration of multidimensional indicators, for example, compare the profiles of selected regions, motion to see these profiles change over time, frequency histograms and filter operations based on percentile statistics [JJ07], switch min and max on an axis, assign user values to min and max etc. The Flash motion PCP and table lens (figure 11) have slowly demonstrated to be not only functional but also productive [JB09] analysing patterns for multidimensional statistical (4-10) indicators.

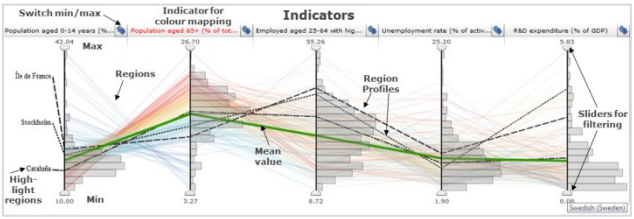


Figure 5: *PCP implemented in Flash and customized for the statistics visual analytics domain. Three European NUTS2 regions are highlighted and provide exceptional statistics means to compare them against several indicators a European median value (green). Patterns never before discovered are visually now obvious to the analyst.*

Collaboration [WH07], [JRÅ*08] is achieved through a mechanism in GAV Flash (figure 6) 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 explorative data analysis process and becomes an important task of the storytelling authoring process.

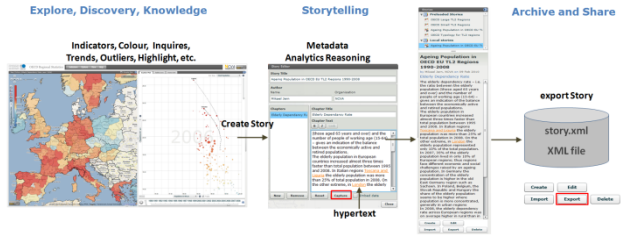


Figure 6: *The storytelling construction mechanism.*

4.2. Storytelling

Statisticians with diverse background and expertise participate in this creative discovery processes (figure 1) that transforms statistical data into knowledge. Storytelling tools integrate the GeoAnalytics process with collaborative means that streamline a knowledge exchange process of developing a shared understanding with other statisticians and after consensus has been reached is published. The snapshot mechanism helps the author to highlight data views of particular interest and subsequently guide others to important discoveries. The storytelling tool creates a single or a discrete series of captures during the explorative process. The author elects relevant indicators, regions-of-interest, colour schema, visual inquiries, filter conditions focusing on the data-of-interest and finally highlights the "discovery" from the most appropriate visual representations and include reasoning text. A sharable story (figure 2 and 6) is created and stored locally on the client or a server in a public repository. Readers can through hyperlinks in the metadata that instantiates a snapshot in the visual representation, follow the analyst's way of reasoning. This focus on publishing through assisted content creation with emphasis on visualization and aesthetics that takes into account heterogeneous nature of information represents the key advantage of our storytelling.

An eXplorer story can have several chapters where the contents relating to indicators and visual layout can change. For example, in *ageing population for Europe during 1990-2008* (figure 10), chapter 1 includes a map linked to a scatter plot while chapter 2 has the same map but here linked to both scatter plot and a PCP to simultaneously analyse three selected regions from different visual perspectives. A highlighted text item in the descriptive metadata could represent a hyperlink to a snapshot (discovery) or any external Web site. When mouse clicked, the story mechanism automatically initiates the snapshot, for example, a zoom to a certain region, change of indicators or links to referenced external Internet content. Hyperlinks to an eXplorer state are an important feature in our storytelling mechanism together with associated descriptive text and guide the reader in the analyst's way of way of thinking.

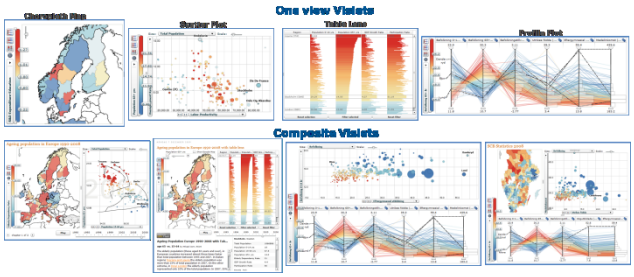


Figure 7: *Examples of one-view Vislets and composite Vislets*

4.3. Vislets

A Vislet is a standalone Flash application (widget) assembled from the GAV Flash class library and Flex GUI tools (figure 7). Integrates selected statistical indicators supported by highly interactive visualization with descriptive metadata that are embedded into blogs, wikis or any HTML document.

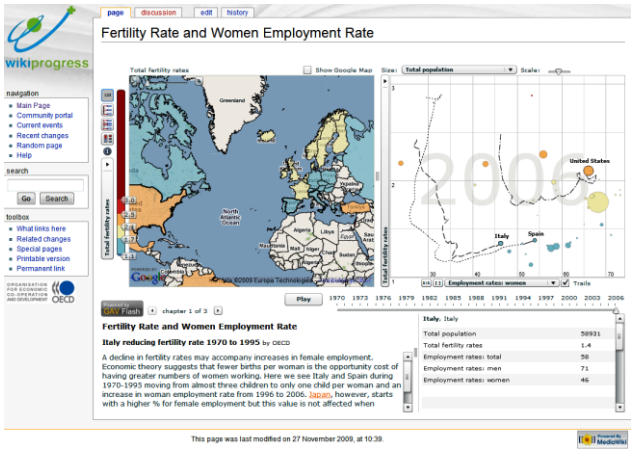


Figure 8: *A Vislet publishing a story about "Fertility and women employment rate 1970-2006" exploring the relations among indicators identifying common paths across countries using two linked motion graphs (map and scatter plot) and associate metadata embedded in MediaWiki.*

The Vislet is created from the contextual description of a geovisual analytics reasoning discovery process in an eXplorer generated Story. The Vislet, represented by, for example, a single eXplorer map view or a composite time-linked map and scatter plot view (figure 7), 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. A Vislet can be assigned any of the visual representation components (figure 7) e.g. map, scatter plot, PCP, table lens, time graph or data grid. Interactive features are exposed to all visualizations including tooltips, brushing, highlight, filter that can discover outliers and dynamic multiple-linked views. Several advanced colour legend tasks are supported e.g. show outliers based on 5th and 95th percentiles in certain colours or dynamic sliders that control class values etc.

The eXplorer server (figure 9) maintains eXplorer and Vislet flash (swf) files together with a story repository, statistical data and regional shape maps. eXplorer runs locally in Flash Player on the client machine and generates HTML code representing a story that is manually (copy/paste) embedded into a Web page. This resulting Vislet is then opened in the reader's Web browser and requested data, GIS, visualization widgets, eXplorer code and stories (metadata with snapshots) are accessed from the eXplorer server.

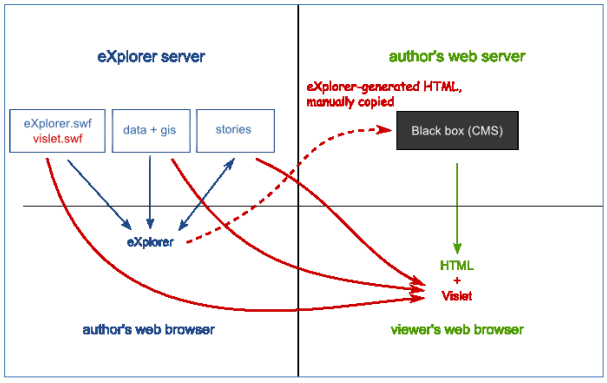
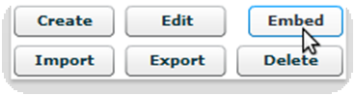


***Figure 9**: Client-server management of the required eXplorer code, indicator data, GIS, stories and Vislets.*

5. Procedure to embed dynamic Vislet in a HTML document

I. Create new or select existing Story to create a Vislet

Use the authoring tool eXplorer to create a new story and associate metadata, or choose one of the already existing stories.



II. Customize your Vislet

Set Vislet properties in the *Embed Story* panel e.g. height, width, graphics layout for visualization and metadata, background and text color.

III. Choose visual representations and copy HTML code

Select visual representations (map, scatter plot, table lens, parallel coordinates or time graph) in drop-down menus. The HTML-code is now automatically generated by eXplorer and copied when you press "Copy". This HTML code, when pasted into a Web page, will load your Vislet with associate metadata and hyperlinks using the settings you have chosen. See example below of generated HTML code for the Vislet.

```
<object classid="clsid:D27CDB6E-AE6D-11cf-96B8-444553540000" id="Booklet" width="100%"
height="100%"
codebase="http://fpdownload.macromedia.com/get/flashplayer/current/swflash.cab"><param
name="movie"
value="http://vitagate.itn.liu.se/GAV/booklets/eXplorer/Booklet.swf?components=(ChoroplethMap,
ScatterPlot),(ChoroplethMap,ScatterPlot,ParallelCoordinates)&amp;story=http://vitagate.itn.liu.se/G
AV/booklets/eXplorer/eXplorer-php-
backend/data//stories/6.xml&amp;backgroundColor=#ffffff&amp;textColor=0"></param><param
name="quality" value="high"></param><param name="allowScriptAccess"
value="always"></param><embed
src="http://vitagate.itn.liu.se/GAV/booklets/eXplorer/Booklet.swf" quality="high" width="100%"
height="100%" name="Booklet" align="middle" play="true" loop="false" allowScriptAccess="always"
type="application/x-shockwave-flash"
flashVars="components=(ChoroplethMap,ScatterPlot),(ChoroplethMap,ScatterPlot,ParallelCoordinat
es)&amp;story=http://vitagate.itn.liu.se/GAV/booklets/eXplorer/eXplorer-php-
backend/data//stories/6.xml&amp;backgroundColor=#ffffff&amp;textColor=0"
pluginspage="http://www.adobe.com/go/getflashplayer" bgcolor="#ffffff"></embed><param
name="bgcolor" value="#ffffff"></param></object>
```

IV. Paste the HTML-code into your Web page

Paste it either in a new HTML document or directly into your CMS or blog. *See the result in a blog*:

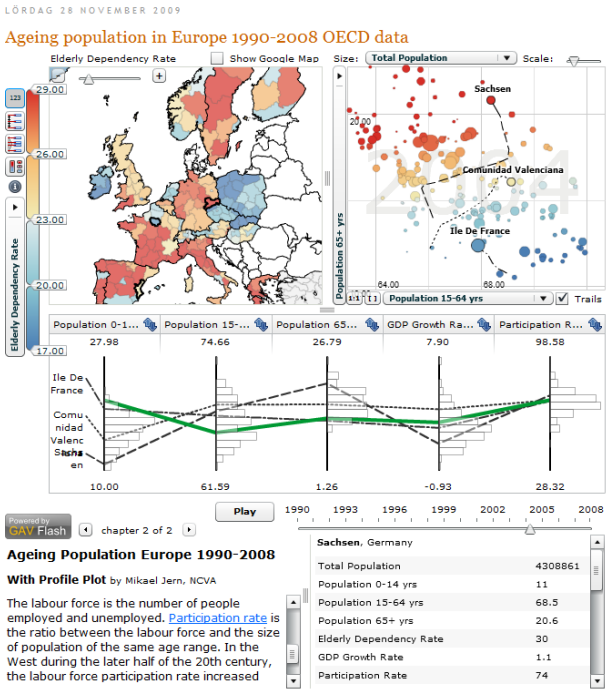


Figure 10: *Time animation in a Vislet showing ageing population for Europe during 1990-2008. Four Vislet views map, scatter plot, PCP and data are dynamically time-linked and animated.*

6. Regional statistical used in education

Official statistics such as demographics, environment, health, social-economy and education from national and sub-national sources are a rich and important source of information for many important aspects of life and should be considered to be more used and acknowledged in education. Students would be able to get informed about a nation's growth and shape the measure of well-being across countries or within local communities and at the same time participate in increasing the knowledge on how life is lived – and can be improved – from region to region. For example, learn about demographics such as population growth rate for sub-national (NUTS3) regions within Europe (figure 11).

For the past years the OECD has been studying regional disparities and regional economic growth in order to evaluate innovative strategies for development and spread successful

policies. The *OECD Regional database* is a set of comparable statistics and indicators on about 2000 regions in 30 countries. It currently encompasses yearly time-series for around 40 indicators of demography, economic accounts, labour market, social and innovation themes. Regions are classified on the basis of two territorial levels: the macro level consists of 335 large regions (figure 10) in the OECD countries, while the micro level comprises 1,681 small regions (figure 11). NCVA has in close collaboration with OECD helped to make regional data more easily accessible in an interactive and user-participative way. In particular, more extensive use of dynamic web-enabled information visualization techniques has effectively conveyed the three dimensions included in the regional database: spatial, multidimensional statistical indicator data and time steps.

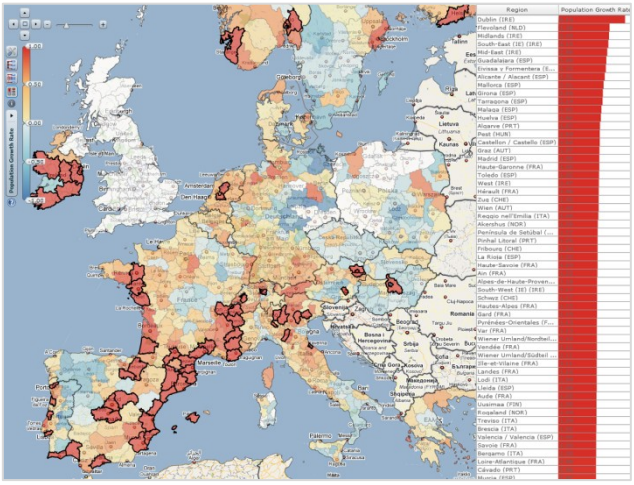


Figure 11: *Population growth rate is an important indicator to better understand a region's development. The 30 European regions with highest growth rate are discovered in the table lens and highlighted simultaneously in the map.*

The story about "fertility rate versus women employment rates" in figure 8 relates to statistics at country level and was produced using *Factbook eXplorer*. Many statistics on education outcome and input to compare different countries educational system are available here.

Statistics have, however, an unfortunate image of being boring, but they can in fact be fascinating and exciting if presented visually in an interactive way. But it is a historical fact that the representations of statistics in different media do not necessarily lend themselves easily to being absorbed by the citizens and this, in turn, makes them less interested in looking for it. A primary target group for our storytelling is the educators and students but also citizens, who can make use of this tool in their learning process.

Researchers and students can now analyse and create content expressing them through "user-created knowledge" with a more pro-active, collaborative role in shared analysis. Active users and user-centred innovation are increasingly important and could have a beneficial social and educational impact. Students can get informed and at the same time participate in increasing their knowledge on how life is lived – and can be improved – from region to region.

It is obvious that if the student is to attach interest to the statistics, gain knowledge and use the information for learning, the statistics must convey a message. In organisations specialising in analysing data, finding the most important messages is a core activity and expertise. Teachers (of all kinds) need to find such important stories and use them but it is not so easy. Therefore they may have to engage statistical experts and analysts in order to discover the messages in the first place. Ideally, our proposed storytelling tools should be able to help users here.

6. 1. Evaluation

The *OECD eXplorer* development followed a user-centric design approach [AA03]. A combination of domain experts from OECD, statistical managers and selected users of OECD statistics outside the organisation have been involved in the various stages of the prototype design and implementation, providing user feedback about usability and utility evaluation. The design process involved public access to beta versions of the authoring tool eXplorer and published Vislets, allowing focus group discussions. The overall involvement and reactions have been very positive. Many useful suggestions for improving the functionality were made and have been incorporated in successive implementation iterations.

A number of characteristics of the current version of OECD eXplorer were derived from comments received during the evaluation phase. In this context, for example, the PCP was considered not to be self-evident to traditional users of statistics, as this is a technique that has not previously been employed in the statistics community and is not described in the methodological literature on statistics, and therefore it was decided to keep it hidden in the start-up phase; at the same time it was regarded as a valuable addition to the statistical toolbox, especially the possibility of dynamically filtering to discover outliers and use profiles to make comparisons between highlighted regions. Finally, the dynamic links between views (context and focus map, scatter plot, PCP and table grid) were evaluated as very important. The spatial cognition for specially the time-linked views was evaluated by OECD and Statistics Swedish and found to be both intuitive and innovative. Statisticians discovered interesting trends in all views respectively.

7. Conclusions and future development

Traditional educational practices for GeoAnalytics methods applied to official statistics need tools for more user participation and increased expectations in terms of user experience. This paper presents a seamless integrated statistics exploration, collaboration and publication process that will address editorial storytelling aimed at producing statistical news content in support of an automatic authoring process. The author (educator) should be able to simply press a button to publish gained knowledge that efficiently and clearly visualize statistical data. A storytelling technology with the goal to advance research critical to official statistical production and publishing and to deliver this research into a web-enabled toolkit for the generation, management and publication of embedded dynamic visualization with the analytics sense-making metadata joined together and publishable in any HTML web pages such as blogs, wikis etc. Publishing official statistics through assisted content creation with emphasis on visualization and aesthetics represents another key advantage of our storytelling and could in many

ways change the terms and structures for learning.

At the same time, it will encourage the practical use of more advanced, collaborative GeoAnalytics science technologies because of its easy accessibility to import any regional statistical data from national statistics to statistics from a neighbourhood. It could also enable the readers to take a more active role in the discovery process of exploring regional indicators, for example, to identify those regional areas that outperform other regions of their country or mean values. The tool will increase the interest in and knowledge of regional structures and development patterns among specialist as well as non-specialist users. Comments from our NCVA partners who started to evaluate the tool highlights the following features:

- Open eXplorer is free available;
- Easy-to-use external statistical data access;
- Ability to have dynamic time-link views and see the multi-dimensionality of regional development;
- Possibility to capture, save and open discoveries (snapshots) with attached analytics reasoning metadata;
- IT expertise is not required to publish interactive visualization embedded in blogs and wikis;
- Important tool to publish statistics news on the Web;
- Increased expectations in terms of user experience;
- Will encourage more educational use of official statistics;

NCVA is a partner in the Global Project on "Measuring the Progress of Societies" and will now implement our storytelling mechanism for evaluation in "*Wikiprogress*". This project represents a catalyst of initiatives existing around the world on the measurement of progress, informing through statistical indicators describing economic, social and environmental trends based on solid evidence.

8. Acknowledgements

This research and case studies were carried out by NCVA in close collaboration with OECD and Sweden Statistics who supplied data and comprehensive evaluation of the statistical storytelling system. The research is in part supported by funding from the "Visualization Program" coordinated by the Swedish Knowledge Foundation. The author thanks in particular to research engineer Markus Johnsson but also the entire research team at NCVA, Linkoping University for valuable contributions. The author also wishes to thank Monica Brezzi at OECD Paris for their dedicated contribution that made this project possible.

9. References

[AA05] Andrienko G, and Andrienko N,: Visual Exploration of Spatial Distribution of Temporal Behaviors, In *Proceedings of IEEE IV2005*.

[AAF*06] Andrienko, G, Andrienko, N., Fischer.R., Mues.V, Schuck, A. : The parallel coordinate plot in action: design and use for geographic visualization. International Journal of Geographical Information Science, 20(10), 2006.

[And06] Andrienko G. et al.: Reactions to geovisualization: an experience from a European project, Int. Journal of Geographical Information Science, 2006.

[BG03] Brodbeck D, and Girardin L.: Design study: using multiple coordinated views to analyze geo-referenced high-dimensional datasets. In Proceedings of The Coordinated and Multiple Views in Exploratory Visualization, pages 104–111. IEEE Computer Society, July 2003.

[CWM05] Carr D., White D, and MacEachren A.M.,: Conditioned choropleth maps and hypothesis generation. In *Annuals of the Assoc of Am. Geographers,* vol 95, no. 1, 2005.

[Eds03] Edsall R.: Design and Usability of an Enhanced Geographic Information System for Exploration of Multivariate Health Statistics. Professional Geographer, 55, no. 2, 2003.

[GCM*06] Guo D., Chen J., MacEachren A.M., Liao K.: A visualization system for space-time and multivariate patterns, IEEE Visualization and Computer Graphics, Vol 12, No 6, 2006.

[Jer01] 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.

[JPR02] Jern M., Palmberg S., Ranlöf M: "Visual Data Navigators Collaboratories", Invited Paper to *Computer Graphics International 2002* (CGI2002), International Conference, July 2002, Bradford, UK, Published in "Advances in Modelling, Animation, and Rendering", Vince J. A. and R. A. Earnshaw (Eds) Springer Verlag, June 2002.

[JJP*05] Jern M., Johansson S., Pettersson M., Feldt A.**:** "Tailor-made Exploratory Visualization for Statistics Sweden", *CMV 2005*, London, July 2005, published by IEEE Computer Society.

[JJ06] Jern M., Franzén J.: "*GeoAnalytics – Exploring spatio-temporal and multivariate data*", Reviewed proceedings, IV 2006, London, published by IEEE Computer Society.

[JJ07] Johansson S., Jern M.: GeoAnalytics Visual Inquiry and Filtering Tools in Parallel Coordinates Plot, Seattle, ACM GIS 2007.

[JRÅ*08] 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.

[JTB08] Jern M., Thygesen L., Brezzi M: "A web-enabled Geovisual Analytics tool applied to OECD Regional Data", Reviewed Proceedings in Eurographics 2009, Munchen, March 2009

[JB09] Jern M., Brezzi M: "OECD eXplorer: Making regional statistics come alive through a geovisual analytics web tool", International conference, Statistics: investment in the future 2009, Prague.

[Kee06] Keel P.: Collaborative Visual Analytics: Inferring from the Spatial Organisation and Collaborative use of information, VAST 2006, pp.137-144, IEEE.

[MB01] MacEachren A.M., Brewer I., et al.: Geovisualization to mediate collaborative work: Tools to support different-place knowledge construction and decision-making. In *20th International cartographic conference*, Beijing, China. 2001

[Rob04] Roberts, J. C.: Exploratory Visualization with Multiple Linked Views, *Exploring Geovisualization*, J. Dykes, A.M. MacEachren, M.-J. Kraak (Editors) 2004.

[Rob06] Robinson A.: Re-Visualization: Interactive Visualization of the Progress of Visual Analysis, workshop proceedings, VASDS. 2006.

[Ros06] Roslin, L.Gapminder.2006.

[WH07] Wohlfart, M., Hauser, H.: Story Telling for Presentation in Volume Visualization, EuroVis2007.