Deliverable 5.1

Preliminary requirements for Visual Analytics, Data exploration and Decision Support system

Grant Agreement number:	688082
Project acronym:	SETA
Project title:	An open, sustainable, ubiquitous data and service ecosystem for efficient, effective, safe, resilient mobility in metropolitan areas
Funding Scheme:	H2020-ICT-2015

Authors
Vitaveska Lanfranchi vita@k-now.co.uk
Suvodeep Mazumdar s.mazumdar@sheffield.ac.uk

Internal Reviewer(s)

Neil Ireson n.ireson@sheffield.ac.uk
Sam Chapman sam@thefloow.com

State: Final
Distribution: Public

Deliverable History

Date	Author	Changes
16.03.2016 30.04.2016	Vitaveska Lanfranchi Suvodeep Mazumdar	Initial draft Initial draft on Visual Analytics
15.06.2016	Vitaveska Lanfranchi	Contribution on decision support
28.06.2016	Vitaveska Lanfranchi	Requirements
12.07.2016	Vitaveska Lanfranchi	Minor changes and corrections
18.07.2016	Suvodeep Mazumdar	Major contribution to Visual Analytics state of the art
21.07.2016	Neil Ireson	General review, editing and corrections
21.07.2016	Suvodeep Mazumdar	Expanding State of the art
22.07.2016	Suvodeep Mazumdar	Expanding State of the art and requirements analysis
25.07.2016	Suvodeep Mazumdar	Contribution on UI mockup and requirements analysis
25.07.2016 28.07.2016	Vitaveska Lanfranchi Suvodeep Mazumdar	Changes to Decision Support and Introduction Changes to mockup section and minor changes in overall document
28.07.2016	Vitaveska Lanfranchi	Changes to mockup section and requirements section and minor changes in overall document
28.07.2016	Sam Chapman	General review, editing and corrections
29.07.2016	Vitaveska Lanfranchi	Addressing reviewers’ comments, editing and corrections
29.07.2016	Suvodeep Mazumdar	Addressing reviewers’ comments

Contents

1. Summary

Deliverable D5.1 firstly presents an introduction to the activities in the Work Package 5, Methodologies and Technologies for Visual Analytics, Data Exploration and Decision Support System. The document follows with a detailed discussion on the state of the art for Visual Analytics and Decision Support. The deliverable further discusses how the activities in each task would progress the state of the art and provides some initial ideas about the direction of development in the Work Package. The deliverable then combines the two activities of Visual Analytics and Decision Support to derive requirements based on state of the art, experience of project partners, study of the domain as well as requirements collection activities conducted in the Work Package 1 within the context of the case studies of the project. Following an analysis of these requirements, the deliverable presents a User Interface mockup that is based on free-hand sketches to be used for walkthrough sessions while going through the scenarios defined in Work Package 1. The document is concluded with a brief overview and plans for the next phases of activities within Work Package 5.

2. Glossary of Terms

DoW	*Description of Work*
D	*Deliverable*
DSS	*Decision Support System*
UI	*User Interface*
VA	*Visual Analytics*
WP	*Work Package*

3. Introduction

The SETA project is aimed at developing and deploying a technology and methodology to change the way mobility is organised, monitored and planned in large metropolitan areas. The goal is transforming the role of citizens into active players in the mobility process and provide services for citizens and decision-makers to better manage their own mobility and through this contribute to achieve better mobility in an area.

To this end, the project must develop solutions to visualize and use the data, models and predictions developed in the other work packages, giving access to all stakeholders, by allowing sharing and exchange of information and feedback. Citizens will be enabled to provide information both consciously (as human sensors) and unconsciously (as social networks participants), they will be able to provide feedback about authorities' decisions and will be provided with advanced services and apps based on sensors and data.
Decision-makers, on the other hand, will benefit from an integrated environment supporting their decision process and powered by a new kind of aggregated information.

This deliverable starts with the presentation of existing technologies and methodologies for visual analytics (Section 4) and decision support (Section 5) and how they can be adapted to the mobility domain. Input from WP1 is then taken into account to contextualise the visual analytics design and development into the case studies scenarios and requirements. This work allowed us to create a set of specific requirements, presented in Section 6, that WP5 will be in charge of implementing in the decision support technologies in the next project phase. Building on the requirements, this deliverable presents a mockup User Interface of the solution foreseen (Section 7), and describes the various features and components and how a user might use the solution. Finally, this deliverable concludes with plans of future activities within the Work Package.

4. Methodologies and technologies for Visual Analytics and Data Exploration

Visual Analytics, or the "science of analytical reasoning facilitated by interactive visual interfaces" [Thomas 2005] provides means for handling massive, heterogeneous and dynamic volumes of information, therefore it is a fundamental part of any Big Data decision support system. In the context of Big Data analysis, Visual Analytics has a significant role: essentially turning the massive information overload into an opportunity [Keim 2008]. Thomas and Cook presents the large number of research areas that contribute toward the scope of Visual Analytics, as illustrated in Figure 1, Left. As can be observed, Information Visualisation (Presentation, production and dissemination) is one of many aspects essential for Visual Analytics, which involves integration of a variety of other areas such as Interaction Design, Cognitive Studies, Data Management and so on.

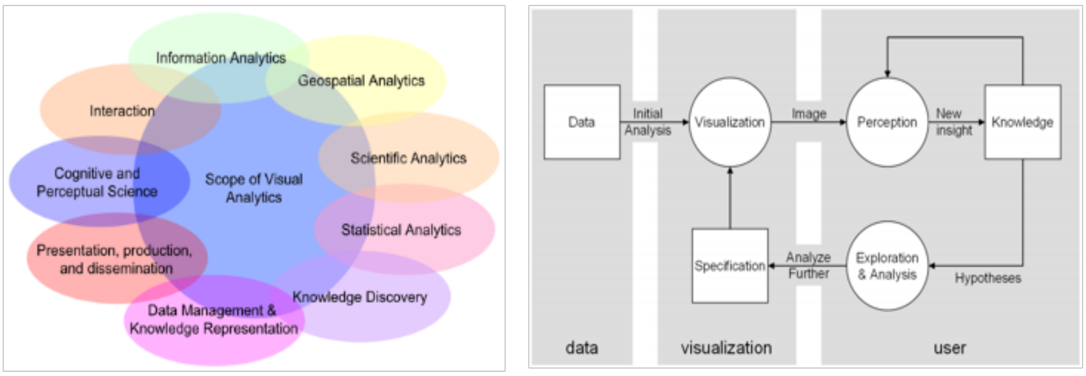

Figure 1- Left- Scope of Visual Analytics incorporating several areas of research, source: [Thomas 2005]; Right- Sensemaking loop [Wijk 2005] for the process of Visual Analytics involves a seamless integration of automated processing with human factors.

Keim et. Al, in [Keim 2008] presents this view by contrasting Information visualisation with Visual Analytics: just as the former has fundamentally changed the way we look at databases, the latter aims at making "our way of processing data and information transparent for an analytic discourse". This is achieved by the process of seamlessly combining human and electronic data processing as a coherent approach: Visual analytics is more than just visualization. In fact, the need for involving humans as a part of tackling the challenge associated with Big Data owing to the cognitive and analytic capabilities of humans to identify trends and patterns despite the advances in machine analysis and computation is well established - and Visual Analytics can serve as one of the ways human and machine processing can be harmoniously combined [Jagadish, 2014]. It can rather be seen as an integral approach to decision-making, combining visualization, human factors and data analysis (Figure 1, Right). Following on from the notion of Visual Analytics being a holistic approach toward facilitating exploring and understanding Big Data, our approach in the project does not only involve visualising large sets of data, but understanding how we can combine sensemaking, cognition, interaction design, visualisation and so on into a coherent and seamless visual analytics approach to facilitate decision making. This section discusses Visual Analytics in the context of Mobility and Traffic, however, often taking examples from other domains where the relevant data share the same characteristics of mobility data. An excellent example domain is networking and communications: datasets describing data packets shared between various nodes geographically distributed and timestamped can share the same spatio-temporal attributes of a vehicle moving across geographical locations.

The next section details the plans and requirements from the description of work as defined within the scope of the project. This, then follows with a more in-depth study of the relevant state of the art to identify gaps in the literature. Finally, we discuss the way identified gaps will be addressed within the Seta project, and discuss how the requirements of the various stakeholders shape the activities in this task.

4.1. Plans and Requirements from the Description of Work

Initial research work carried out when writing the DoW identified four key requirements for Visual Analytics in SETA:

1. Context specific visual representations, interactions and animations

Visualising data instances requires understanding the different contexts associated with a piece of information. For example, a ‘cycling’ activity reading for a user would require understanding the context behind the particular activity such as location, time and duration of the activity. The visualisations that present such data is hence, expected to preserve the context behind such data, so that by using interactions and animations, the end user can have access such information.

2. Multi-level, multi-dimensional hierarchical clustering and categorising

Multi-dimensional data can be presented to users using a multi-visualisation paradigm where for example temporal features of data can be presented as timelines, while spatial data can be presented in maps.

3. Aligning contexts of exploration and contexts of data capture

Decision makers can be provided with settings and preferences that can enable them to focus on the right dataset quickly. These preferences can be adapted as they explore the data and the domain, and hence provide means for understanding the context of exploration.

4. Extend existing visualisation, interaction and design paradigms to innovative visualisations

The domain of mobility and traffic is a complex one, generating massive volumes of heterogeneous multidimensional data. In order to effectively and intuitively explore such a dataset, it is important to develop new visualisations (requiring more learning), extend existing ones (requiring less learning) as well as combining multiple ways of visualisation to provide a number of exploratory and querying paradigms for understanding data. For example aggregate charts on timelines or maps can provide an interesting way of exploring spatial or temporal data by traversing different categories of data[1].

4.2. State of the Art in Visual Analytics

Visualising is essential for gaining an understanding of data and underlying phenomena, and hence a key aspect of Visual Analytic processes. Integrating automated analysis with interactive visualisations can provide excellent analytic capabilities [Keim, 2008, Keim, 2010]. Several excellent examples exist (as surveyed by [Sun, 2013]) that have been used for a wide range of applications such as astronomy [Ma, 2005], finance [Dang, 2013], studying election results [Wood, 2011], query analysis [Shi, 2012] to analyse spatial data [Slingsby, 2011], temporal data [Krstajic, 2011], spatio-temporal data [Tominski, 2012, Adrienko, 2011, Adrienko, 2013, Scheepens, 2011], multivariate data [Joia, 2011,Turkay, 2012], topic and

[1] https://github.com/FlowingMedia/TimeFlow/wiki/Calendar-View and http://kb.tableau.com/articles/knowledgebase/creating-filled-maps-pie-charts provides some interesting examples where category visualisations have been combined with timelines and maps.

feature data [Xu, 2013, Cao, 2010, Zhao, 2012] and graph based data [Von Landesberger, 2011].

While a large set of different types of data exist within the Mobility and Traffic domain, most relevant are the visual analytics approaches for exploring geospatial data, temporal data and spatio-temporal data. Several other types of data also exist within the field, such as categorical, qualitative, quantitative etc. and each type has a large number of ways for visualisation and subsequent analytics. Categorical data has been typically visualised as aggregate visualisations such as one-dimensional pie/donut charts, word/tag clouds, hierarchical pie charts, multi-dimensional bar charts, histograms and column charts. Hierarchical data as well as relations can be visualised as trees, force-directed graphs (Figure 2, Top Left), circular node-link tree (Figure 2, Top Center), treemaps (Figure 2, Top Center). Multi-dimensional categorical/quantitative data has also been visualised in different ways such as Parallel Plots (Figure 2, Bottom Left), or animated scatter plots (Figure 2, Bottom Right).

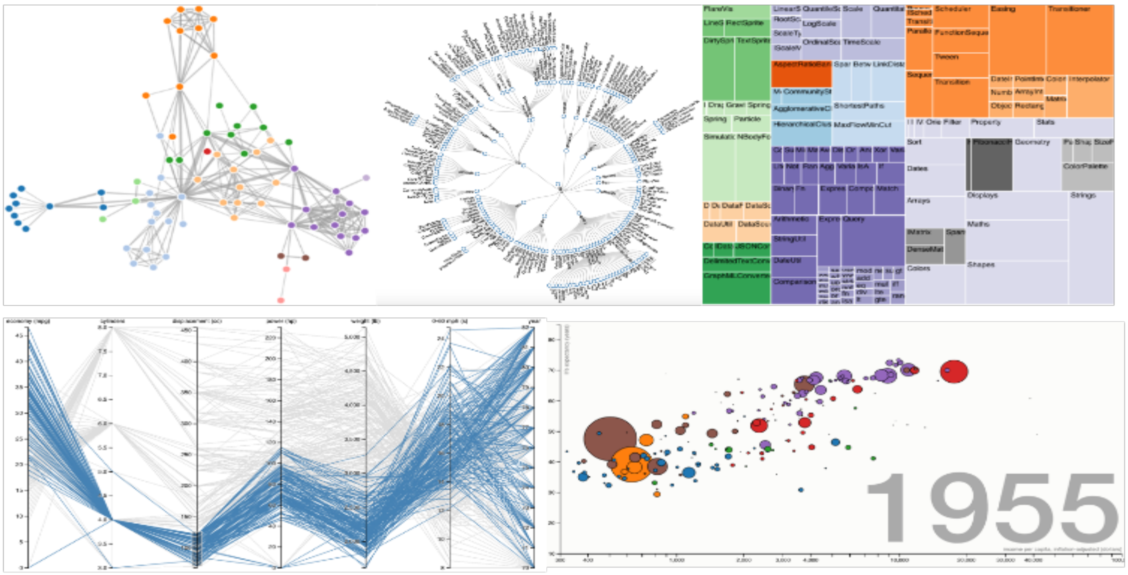

Figure 2 - Different visualisations for a variety of data types, source: all examples from https://bost.ocks.org/mike/, https://bl.ocks.org/jasondavies/ and https://github.com/d3/d3/wiki/Gallery

Qualitative data on the other hand provides subjective views of users and hence, is mostly content-based rather than objective scores and numerical quantities. As a result, few approaches have been developed to present such datasets, among which most common is to apply NLP techniques to identify keywords and concepts and visualise statistics on their occurrence. Tag/Word/Keyword clouds are one of the most common ways of visualising such information. Often, sentiment analysis on subjective feedback can provide a lot of interesting ways for visualisation, such as rose plots [Gregory 2006], treemaps [Gamon 2005], simple pie charts, visual encoding of text [Marcus 2011], or visually encode other visualisation components like timelines [Diakopoulos 2010], node-link graphs [Mazumdar 2012] and so on. Indeed, a lot of visualisation approaches have been developed for text analysis and more recently, for Social Media analysis [Shreck 2013].

Challenges in Spatial Visual Analytics

In typical visual analytic systems, interactive querying facilities are provided by interactive visualisations that use several techniques to visualise large datasets to perform analytic tasks for supporting knowledge construction, problem solving and decision making [Keim, 2010, MacEarchen, 2004, Boulos, 2011]. Visual Analytics has been extensively used for supporting decision making in various fields such as organisational management, spatial decision support, medical disease diagnosis, health analytics and epidemiologic health analytics [Al-Hajj, 2013]. This section will present the state of the art in Visual Analytics for Mobility and Traffic, particularly relevant within the scope of the Seta project.

Traditionally, visualising movement data has relied mostly on arrows or flowlines drawn on maps or images [Yu 2006], Minard's map of Napolean's 1812 Russian campaign being one of the first and most famous examples. Some of the best examples of visual analytic systems for large data collections exist in the scope of traffic and movement, possibly arising out of the unique characteristics of such datasets: movement data can be highly varied, comprising of a (single or multiple) geospatial, temporal, categorical, qualitative, quantitative components. While other types of data such as temporal, categorical etc. have a large number of visualisation techniques that can be employed for exploration and analysis (as discussed in the previous section), geospatial data is often the most challenging. This is mostly due to the complexities involved in mapping software, tools and spatial services.

Figure 3 - Left:Overlapping points introduce readability issues to maps, source: [Chua 2016]. Right: Too many trajectories in maps can be difficult to read, source: [Keim, 2003]

Particularly challenging is the variety of geospatial data that require mapping - geolocated points, trajectories (lines and polylines connecting multiple points to indicate movement), bounded areas (multiple points connected to form closed boundaries to indicate regions/buildings/areas). With large datasets however, each of these geospatial data types can introduce significant challenges [Keim, 2003]. Though dot maps can be elegant for point data in smaller scales, when dealing with large datasets can be extremely difficult to deal with - e.g. areas with greater instances can be overplotted and overpopulated with highly dense data, while areas with smaller instances are prone to be almost invisible. The same issues arise while visualising trajectories or bounded areas, where a large dataset can introduce a lot of clutter in locations where greater number of network connections (or areas) are visible, rendering maps extremely difficult to read. Figure 3 presents how overlapping data in dot maps and cluttered map networks can be difficult to read.

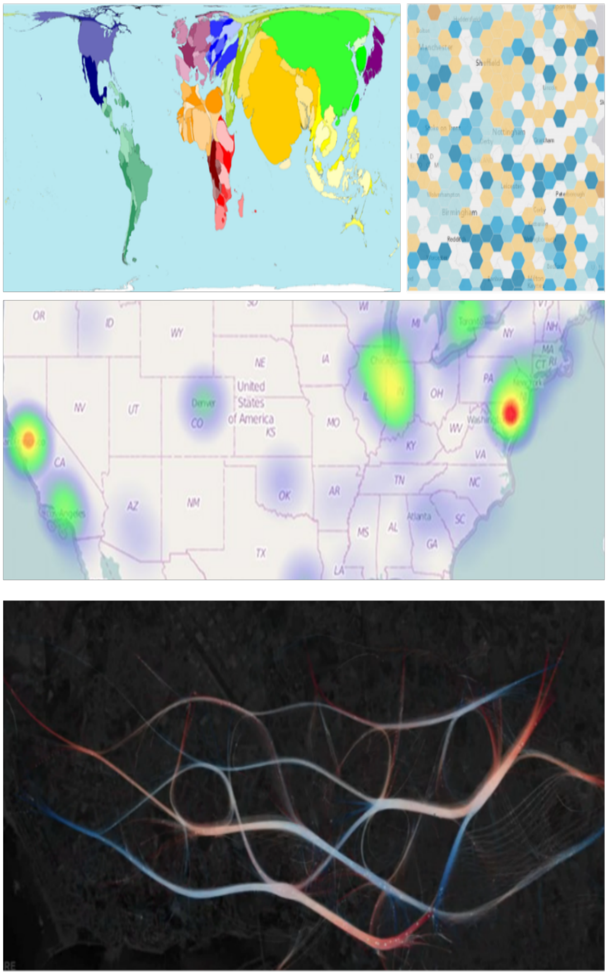

Figure 4 - Top-left: Cartograms can visually encode and distort map polygons to reduce large datasets, source: worldmapper.org; Top-right: Using hexagonal polygon heatmaps can help further reduce very large point datasets, source: https://blogs.esri.com/esri/arc

[Keim, 2003] discusses a few approaches to dealing with these challenges such as introducing a 2.5D visualisation for dot maps where pixels overlapping could be shifted to display denser data. Cluttered networks could be simplified by using highly interactive 2D and 3D techniques for users to focus on trends and patterns. Some other standard approaches that have also been employed are heatmaps and hexpolygon heatmaps[2]. Another approach could be to use similar techniques as hierarchical edge bundling [Holten, 2006] in mapping tools (though has been traditionally used in graph visualisations). Keim also discusses using Cartograms[3] [Hearnshaw 1994] to reduce complexities of large data sets, where geographical boundaries are to be considered. Figure 4 presents the various ways the problem of overlapping datasets, cluttered trajectories and dense area maps can be addressed.

While the visualisation focus in the field has typically been in visualising spatial data and trajectories, several other types of data are expected in the Seta project such as temporal, categorical, qualitative and quantitative. This involves understanding how such datasets can be visualised and subsequently explored. Understanding and classifying the large variety of visualisation approaches for different types of data has a significant history - several researchers have attempted a taxonomy based approach toward understanding how has the field typically visualised different types of data. [Chi 2000; Schneiderman 1996; Tory 2004] are some excellent founding examples of such attempts, investigating how different tools visualise data based on a variety of parameters such as data type, task type, interaction mechanisms and so on. The following discussion provides some examples of the variety of ways the different data types have been typically visualised so far.

Geospatial Data

Visualisation of geospatial data (though not only limited to spatial data) has seen a significant growth in the last few years, initially with a variety of (freely available open source or commercially available) GIS (Geographical Information Systems) tools and software which are Desktop based (e.g. ArcGIS Desktop[4], GRASS GIS[5], Quantum GIS[6], uDig[7]), Server based (e.g. ArcGIS Online[8], PyWPS[9], ZOO-Project's SAGA GIS[10]), Mobile (gvSIG[11],

---

2 https://blogs.esri.com/esri/arcgis/2012/06/08/using-a-binning-technique-for-point-based-multiscale-web-maps/
3 http://www.worldmapper.org/
4 http://www.esri.com/software/arcgis/arcgis-for-desktop
5 https://grass.osgeo.org/
6 http://www.qgis.org/en/site/
7 http://udig.refractions.net/
8 http://www.esri.com/software/arcgis/arcgisonline
9 http://pywps.org/
10 http://www.zoo-project.org/docs/kernel/sagagis.html
11 http://www.gvsig.com/en

Quantum GIS for Android - experimental[12]) and so on. Some GIS tools like GeoDA[13] [Anselin, 2005], PySAL[14], HiDE[15], however provide exploratory mechanisms for data analysis that can help identify patterns, hotspots, clusters and so on using different visualisation approaches such as scatter plots, box plots, parallel coordinates, thematic maps. [Steiniger 2012] presents a thorough survey of GIS tools and server technologies (e.g. Geospatial databases) that are essential for analysing geospatial data. With the recent movement toward Cloud-based services, web-based tools and increasing efforts in GI-Science by developer communities, several well-known tools and applications are now available to be used for GIS applications. The most well-known and obvious example would be Google Maps API[16], which served as one of the first (releasing in Mid-2005) online developer tools available for free to developer communities. Several other examples emerged later on such as Bing Maps[17], Open Layers[18], promising similar features as Google Maps, with high adoption by developers. Several new mapping startups such as Leaflet.js[19], MapBox[20] and MapMe[21] have also introduced their mapping solutions, heavily adopted by large organisations (who, often acquire Mapping startups[22]) and developer communities.

Temporal Data

A large number of ways have been proposed for visualising temporal data, though most of them have been variations of a timeline, which is a linear graphical visualisation of events over time. Timelines have taken a variety of forms over the years, and can communicate a large amount of information in a relatively short space - more so, with interactive features such as click and drag, scroll and zoom. Figure 5 presents some examples of timelines, presenting different types of information.

[12] https://play.google.com/store/apps/details?id=org.qgis.qgis
[13] http://www.csiss.org/clearinghouse/GeoDa/
[14] http://pysal.readthedocs.io/en/v1.11.0/index.html
[15] https://editor.giscloud.com/
[16] https://developers.google.com/maps/
[17] https://www.microsoft.com/maps/choose-your-bing-maps-API.aspx
[18] http://openlayers.org/
[19] http://leafletjs.com/
[20] https://www.mapbox.com/
[21] https://mapme.com/
[22] http://www.recode.net/2015/9/16/11618606/apple-acquires-mapsense-a-mapping-visualization-startup

Figure 5 - (1) Non-interactive timeline displaying distinct events, source: [Karam 1994]; (2) Non-interactive timline displaying continuous events, source: [Karam 1994]; Interactive zoomable and scrollable styled timeline displaying (3) distinct events, source: Simile Timeline at http://www.simile-widgets.org/timeline and (4) continuous events, source: D3 example page at http://bl.ocks.org/bunkat/2338034

Figure 5 (1) shows a timeline presenting distinct events as an impulse on a horizontal X-Axis, where each line/bar indicates only the occurrence of an event [Karam 1994]. The timeline is non-interactive hence, resulting in a number of overlapping events if occurring at very short intervals, without the ability to focus on an individual section of the timeline to gather a better understanding. Figure 5 (2) shows a non-interactive timeline presenting continuous events, further indicating another parameter that can display the intensity of a feature (e.g. amplitude), and hence can display graphs [Karam 1994]. Figure 5 (3) shows a more stylised timeline, that is interactive - clicking and dragging scrolls the timeline to focus on a different section of the timeline. The timeline can also be zoomed-into by using mouse wheel gestures.

The Y-Axis positioning of the events is not significant, implying that there is no relation to where a particular event has been positioned on the Y-Axis and the only positioning that is relevant is the X-Axis temporal position. An example of such a timeline is the Simile timeline, developed in the MIT Simile project[23]. Figure 5 (4) displays continuous events on a timeline that offers interactivity (mouse wheels to zoom, click and drag to focus on different sections) as well as positions the events vertically to indicate which category the event belongs to. Furthermore, the timeline also provides visual encoding to display category information (colors indicate which type of events). The example shown in the figure was developed as a D3 example application, which builds on a similar approach as Simile[24].

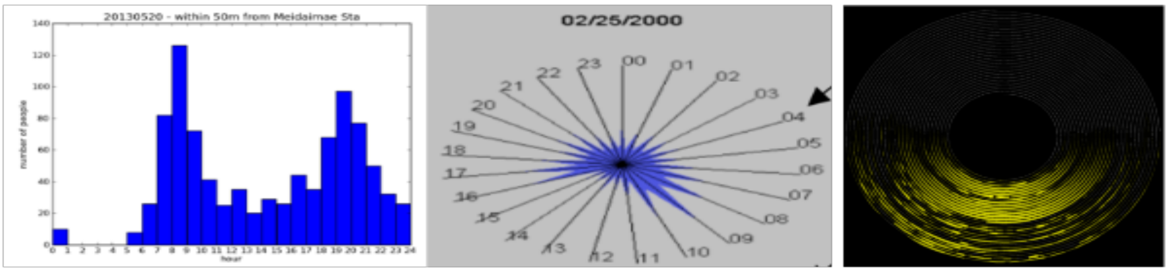

Figure 6 - Several examples exist where temporal data has been visualised differently as compared to traditional timelines; Left: an hourly view of temporal data to indicate number of instances occurring in the relevant hour, source: [Nishi, 2014]; Center: a radial star representation of the hour of the day, similar to a spider diagram, source: [Daasi, 2005]; Right: a spiral visualisation provides easy ways of exploring large temporal datasets, and enabling comparison of instances between days. source:[Weber, 2001].

Several other approaches have been developed for visualising temporal data. For example, further processing temporal data and classifying into regular sections (e.g. every hour of the day, every day of the week etc.) can provide a histogram or bar chart based view on temporal data (Figure 6, Left). The same approach can be extended to a star representation [Daasi 2005] (Figure 6, Center). [Weber 2001] presents a spiral visualisation technique that can display large nominal, quantitative or ordinal temporal datasets (Figure 6, Right). In fact, the number of ways of visualising temporal data is such varied [Aigner, 2011] that the Timeviz browser[25] provides an interactive catalogue of the different visualisation approaches that have been developed for temporal data visualisation.

---

[23] http://www.simile-widgets.org/timeline/

[24] http://bl.ocks.org/bunkat/2338034

[25] http://www.timeviz.net/

Spatio-Temporal Data

Research in Visual Analytics for smart cities, particularly mobility has focussed mostly on spatio-temporal analysis of movement data as abstractions (by visual representations of multivariate data) on geographical layouts [Adrienko, 2012] such as combining different overviews to present spatial and temporal evolution of the data [Kapler, 2005, Guo, 2011, Andrienko, 2011b]. With the advent of interactive and animation techniques, spatio-temporal data is typically visualized as animated maps and interactive cubes. [Adrienko 2000] proposed a dynamic map display controlled by a set of interactive time controls for supporting exploratory analysis of movement data of objects. Different colors and hues were used to differentiate routes, while animations with intervals display how objects dynamically move, displayed as "worms". Longer lengths of worms represent a faster movement, while shorter represent slower movements. Another approach toward visualising spatio-temporal data has been in employing 3D visualisations. Using 3D visualisations and cluster techniques to qualitatively analyse spatio-temporal patterns for movement and relationships between times of travel and land use, [Liu, 2009] presented means to spatially and temporally quantify and visualise urban mobility patterns (using smart cards and taxi GPS traces as data) for real time decision support. [Kapler 2004] developed GeoTime, a combined temporal and geospatial display, that shows interconnecting streams of events over a range of time in a single 3 dimensional image. X and Y axes of events position them in a geo-coordinated location, while the Z axis positions them temporally.

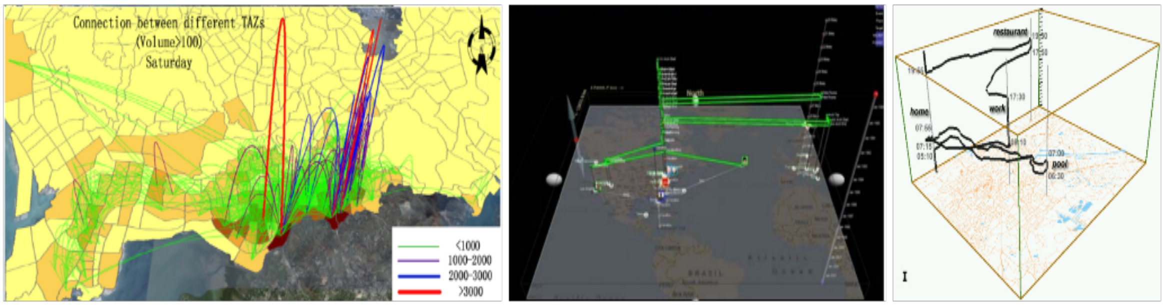

Figure 7 - 3D visualisations of spatio-temporal data. Left: [Liu 2009], Center: [Kapler 2004], Right [Kraak 2003]

[Kraak, 2003] employed a similar approach toward visualising spatio-temporal data in a space-time cube where the cube's base represents geography, while the height represents time. This provides a good way of visualising such data as when viewed from a three-dimensional perspective, the inclination of a line indicates the speed of movement of the object. This, integrated with several charts and diagrams provide further context and display various aspects of movement such as parallel plots for representing distances, and other multivariate characteristics of data [Dykes 2003], timelines and radial plots [Mountain 2005].

The very nature of big data (high volume, veracity, velocity, variety and value of multi-dimensional, multi-modal, heterogeneous data) presents significant challenges for visualisation, that has been identified as a major challenge for big data in the literature [Bizer, 2012, LaVelle, 2013, Bollier, 2010]. Most of the approaches toward handling large volumes of data has exploited data aggregation. Several aggregation techniques have been suggested such as temporal histograms, traffic density surface, accessibility surface (e.g. travel times from selected locations) [Adrienko 2007], or visualising clusters of similar mobility behaviour as trajectories and clusters [Adrienko 2009, Liu 2013]. However, summarizing movement data can result in missing the dynamics of movement itself [Adrienko 2007]. While several approaches stay true to their given spatial layouts by using location-specific maps to visualise movement data, some [Crnovrsanin, 2009, Laube, 2002, Weaver, 2008, 29] provide abstracted views. There are also a few commercial systems which provide customers with mobile and tablet apps adapted to present interactive dashboard for mobility data such as Tableau[26], Spotfire[27], QlikView[28] and JMP[29] [Zhang, 2012].

4.3. SETA beyond the state of the Art

SETA will exploit visual analytics as means to explore and deal with massive volumes of real-time and historical datasets. This requires the immediate involvement of user communities and in order to serve a variety of stakeholders (as identified earlier - citizens and public, businesses and authorities), there is a requirement to provide a variety of visual mechanisms for exploration of data. Furthermore, users needs vary across stakeholder communities, and as a result, may need more focus on different aspects of data. For example a citizen making use of a visual analytic system may only be interested in data that pertains to their immediate context (e.g. search criteria, locality or preferred mode of travel). Alternatively, decision makers may be interested in all data or a very specific subset of the data. For e.g., while monitoring, decision makers need to be aware of all situations on the ground and be alerted to any issue that may be of concern; in the case of investigating a

[26] http://www.tableausoftware.com
[27] http://spotfire.tibco.com/
[28] http://www.qlikview.com
[29] http://www.jmp.com/

particular event/scenario on the other hand, decision makers need a very specific subset of the data to focus on. Hence, one of the key aspects of the visual analytic solutions will need to be the ability to focus on different subsets of underlying data - pertaining to either geographic, thematic, modality or any other context. At the same time, this will help decrease information overload in dynamic and critical decision making tasks such as traffic management and emergency response.

In order to deal with Big Data, the visual analytic approach will need to address individual characteristics of such data, or otherwise, have mechanisms in place to address high Volume, Velocity, Veracity, Variety and Variance. Although different researchers mention different characteristics of such data, essentially, the need from Visual Analytics is to address how large volumes of rapid and real time data, with a variety of provenance, significance and value can be explored. We discuss our approach for the individual aspects in the next few sections, while proposing how Seta will attempt to address each aspect.

Volume and Velocity

High Volume and Velocity of data presents challenges that require massive scaling up in multiple respects: scaling up to ensure maximum utilisation of screen-space as well as scaling up to ensure maximum data instances are visualised as well as scaling up to ensure the visualisation is dynamic, flexible and fast. Volume and Velocity will be addressed by creating iterative visualisation techniques that compute intermediate features and parameters and create visualisations on the fly. New parameters for the selected features will be calculated on-the-fly when the user requests them, not continuously. SETA will employ a multi-dimensional, hierarchical, clustering and categorisation approach where visual spaces can provide a massive amount of information by employing clustering techniques in multiple levels, based on context of exploration and navigation of the decision maker. An example of such an approach is shown in the Figure 8, where a large number of points in a map can be clustered into heat maps or aggregated into representative regions such as postcodes, electoral wards, counties or even arbitrary sections as deemed necessary. The Visual Analytics approach would therefore need to index data in a variety of ways in order to facilitate real-time querying based on such varied parameters. In addition to querying datasets in the backend, frontend querying would also be used to enable quick filtering and 'drill-down' actions via interactive gestures. Typically this is done by making use of queryable caches that can index data once accessed from backend datastores.

Figure 8 - Mechanisms of reducing a very large spatio-temporal dataset (taxi trips for a week) - Left shows plots of taxi trips for a week. Center reduces the plot by presenting the same dataset as a heatmap. Right further reduces the dataset by summarising the trips and visually encoding neighbourhood boundaries. Source: [Ferreira 2013].

SETA will also use context-based visual representations and animations (e.g. an urgent alert on specific road junctions, triggered as a result of an incident) to communicate new information added into the system, however not overloading users with constant stream of data. These animations and visual representations for critical/non-critical information pieces will be designed from techniques developed in several areas such as usability and interaction design [White 2007] , preattentive processing [Healey 2012] and generic data visualisation [Heer 2012]. By employing such approaches, SETA will address challenges in the Big Data setting specifically, while ensuring a generic holistic approach is preserved. Moreover we will use parallelisation techniques, that process data in independent units and then put them together at interface level.

Veracity and Variety

High Veracity and Variety of data presents challenges in heterogeneity of content, accuracy and authenticity. Considering the significant limitations of screen space and computation power for real-time visual processing, this is a highly challenging task. Additionally, visually encoding millions or billions of data points with such information will introduce a massive cognitive burden on users.

Figure 9: Four Different levels of detail, that can be explored by using interactive gestures, providing information regarding different levels. Left - UK regions, Center - UK Parliamentary Constituencies, Right Top - Birmingham Parliamentary Wards, Right Bottom - Birmingham Lower Layer Super Output Areas

SETA will employ a broad range of visual encoding and interaction techniques to complement large scale visualisations with such information from multiple levels of granularity. The visual encoding, representations and interactions pertaining to veracity and variety of data will provide as enrichment to large scale visualisations, and will be designed as a multi-level process. For example, the visual encoding will consider the impact of the veracity and variety of larger cluster of instances from a higher level. If user exploration moves to a higher granular level, the clusters will update based on their overall impact to the visualisation as a whole. This, for spatial data can be explained by imagining a map with a large number of levels as boundaries - each level of detail can have an associated bounding level. On the highest level, counties could be visualised, enriched with underlying statistics. Zoomed to a further level, ward level or postcode levels could be visualised. Figure 9 presents such an approach. Within the context of mobility and traffic, understanding the importance of trust, reliability and accuracy of data is extremely important. Critical decisions will be taken based on such information and hence each piece of actionable information needs to be analysed and verified. A lot of the sensors developed in Work Package 2 is aimed at drastically reducing costs as well as increase coverage to a great extent. Hence, it is expected that the data being generated may have different levels of accuracy. Furthermore, sensors can develop faults over a period of time, be inaccurate and may not always be calibrated. Traditionally, unreliable information is blacklisted until manually fixed, hence potentially leaving large areas without any means to sense for longer durations. SETA will develop large scale filtering and visual capabilities to identify, visualise and filter information based on their reliability and accuracy.

Variance

High variance in value of data presents a different challenge to visualisations, as it involves highly complicated means of assessing the intrinsic value of information pieces: while the value of a piece of information is necessary to be communicated to users, low value information for one type of analysis could be of utmost importance for another. SETA will address this challenge by modelling, representing and interpreting context in two levels: (i) decision makers who use the visual analytic system; (ii) citizens, communities or sensors that provide the information. This stems from the belief that visual exploration, navigation and searching occurs in multiple contexts (e.g. in major road incidents involving large number of commuters, a decision maker's visual exploration of maps for different contexts would be different: searching for routes to plan rescue operations, identifying safe zones or plotting stranded citizens) - SETA will aim to identify the context of exploration of decision makers and align this with the contextually relevant value of information pieces (e.g. in a rescue scenario during a fire, the value an induction loop in a closed down road would not be as relevant as that of a citizen's image of the fire reported to decision makers).

5. Methodologies and technologies for Decision Support

5.1. Plans and Requirements from the Description of Work

Initial research work carried out when writing the DoW identified two key requirements for Decision Support in SETA:

1. Support multiple levels of decision-making

All stakeholders need systems for supporting their decision-making process, accordingly to their roles and needs. SETA will address this by providing tools for combining quantitative data and qualitative knowledge/perceptions [González 2013].

2. Support analysis of large, dynamic datasets

Existing decision support systems for mobility employ visual analytic techniques to navigate, explore and analyse large datasets [Liu, 2009; Guo, 2013] but they normally focus on macro areas [Cho 2002] or use static, historic data [Nijkamp 2007]. SETA will address this challenge by focusing on visualization of dynamic, fused datasets that focus on micro-level data and real-time analysis.

5.2. State of the Art in Decision Support

An assessment of existing decision support for mobility solutions has been done in order to understand what the existing solutions could offer in terms of functionalities and features, with the goal of, if appropriate, matching the provided functionalities with user requirements and needs, and if any of the existing solutions to be integrated within SETA.

In order to evaluate the decision support for mobility solutions, a list of "interesting" features has been prepared, which defines functionalities and characteristics of the platforms.

After the evaluation of the solutions, a comparison matrix has been produced, which matches solutions against features[30].

This matrix can then be used to verify how well existing solutions match known user requirements prior to exploring its use within individual use cases.

Reviewed Platforms

EmerT - a web based decision support tool for traffic management

EmerT [1], Emergency mobility of rescue forces and regular Traffic is a web decision support application for real time traffic situation, prognosis and simulation. The main purpose of EmerT is to handle major incidents and large events which affect traffic demand by offering valuable information to both users and operators.

EmerT supports data from induction loops, floating car data and real time data of occupancy rate of parking spaces with the possibility of including new approaches like bluetooth detection, input from low resolution web cameras and more.

30 *Section 6.3 provide a simplified matrix, where each feature is simply marked as available or not available. More information about the availability of the feature (e.g. when a feature is partially available and what is provided) can be found in the complete matrix in Appendix 2.*

PROs	CONs
offers traffic simulation for different scenarios	the system is heavily dependent on location (because the data sources are set in that location)
uses real time data from multiple sources	needs a UI that allows for the creation of complex scenarios for training purposes
offers traffic routing in emergency scenarios	the system does not yet integrate the availability and usage of public transport
by offering traffic simulations, users can observe the potential effects of different management measures	

Mobility Analyser subsystem of IRMA

This framework is validated by Mobility Analyser which is a subsystem of Integrated Real-time Mobility Assistant (IRMA) and it is comprehensive, scalable and cost-effective [2]. It can process tweets for bus simulation and crowd area detection. Social data is collected and analysed and the presented in different components using GIS (Geographical Information System).

The system needs to be configured for the area that is to be monitored and once it has, it will monitor and gather tweets in that area. This framework offers a complete support for the social data analysis lifecycle: data collection, data preprocessor, data analyser, data presenter and data storage.

PROs	CONs
the system offers complete support for social data analysis lifecycle	has no other data source except social data
uses social data to offer real time information on a day to day basis	only provides relevant information for bus simulation

offers social data clustering for crowd area detection	depends on the availability of the public twitter API
is independent of location as long as public transport information is available.	

T-DSS - Transportation Decision Support System

A T-DSS [3] designed to help users make decisions about trips based on large amounts of geographical data and information. It makes use of a GIS for visualisation and it is based on a geospatial object-oriented database structure.

The entire platform is built on top of open-source technologies and it is able to generate route alternatives based on traces of a local transport system using previous user behaviour. It offers both a web interface and a mobile interface.

PROs	CONs
the system analyses previous user behaviour to offer better route alternatives	only designed to support users in taking trip planning decisions
offers web app and mobile app	only focused on public transport
the routing algorithm is responsive and precise	no traffic monitoring
promotes the use of public transport by improving user information	no other data sources except public transport information

Decision support system during evacuation using VISSIM

This specific system was used to model simulations for various evacuation strategies and scenarios for human caused or natural disasters. VISSIM was used to generate a simulated street network in order to evaluate pre and post disaster performance [4].

In order for this simulation to be possible, Google Maps and manual observations of the road network were used to configure VISSIM with the correct lane configuration, traffic signals. In addition, the network was coded to have peak hour volumes in order simulate the worst case scenario in terms of traffic.

PROs	CONs
specific model simulation for programmed section of road network	only for simulation purposes
capable of handling multiple scenarios	road network needs to be pre configured in VISSIM
offering routing alternatives	uses no real data source
	does not account for public transport

TSC's sentiment mapping project

TSC's Sentiment Mapping project[31] (carried out by DfT, Zipabout, Transport Focus, Nottingham University and Keolis) has created an analysis of "customer experience indicators" by overlapping historic operational data with sentiment analysis to understand and potentially influence customers' behavior.

PROs	CONs
uses social data to offer real time information on a day to day basis	has no other data source except social data
Explores user experience	Limited number of relevant tweets
Very friendly user interface	depends on the availability of the public twitter API

[31] https://www.youtube.com/watch?v=4UAyAQO1S0g

	Only works on historic data

DeSSIA - a decision support tool for evaluating mobility projects

DeSSIA - Decision Support System for Impact Assessment is a software primarily used to evaluate mobility projects at different levels: traffic, environmental, social and economic. The system is composed of several modules for graphics generation, map generation, etc.

The system allows a project manager to configure multiple maps with specific criteria (sensor positions, road sections, pollution); background algorithms then determine the efficiency of the implemented measures in the mobility project that is being studied.

PROs	CONs
offers a specific simulation for programmed section of road network	only for simulation purposes
capable of following the evolution of spatial and temporal data	road network needs to be pre configured in VISSIM
can be used to study the efficiency of any major decision	uses no real data source
	does not account for public transport

TSS – Aimsun traffic modelling software[32]

Aimsun is traffic modelling software developed by TSS to model traffic at a microscopic and mesoscopic scale, integrating traffic and infrastructure data with raw geographical and socio-economic data. It combines different models (Dynamic user equilibrium (DUE) techniques, stochastic/discrete route choice models, mesoscopic or microscopic modelling) in a user

[32] https://www.aimsun.com/aimsun/top-features/

friendly in interface. The system allows a mobility manager to simulate traffic management and consequences of infrastructure changes.

PROs	CONs
offers a specific simulation for programmed section of road network	only for simulation purposes
capable of following the evolution of spatial and temporal data	road network needs to be pre configured
can be used to study the efficiency of any major decision	uses no real time data sources

GPU Pedestrian Simulation

This is a Multi-agent simulation system [Karmakharm et al, 2010] that utilizes multiple GPU-enabled computers connected over a standard TCP/IP network to simulate pedestrian flow at faster than real-time speeds.

Whilst this allows running generating pedestrians and running simulations in response to changes in the environment, it uses no real-time or crowdsourced data for analysis.

PROs	CONs
offers a specific simulation for pedestrian	only for simulation purposes
capable of presenting results of changes	Limited to pedestrian traffic
	uses no real time data sources

✗✓		EmerT	Mobility Analyser IRMA subsystem	T-DSS	VISSIM based simulator	DeSSIA	TSC	AIMSUN	GPU
Data gathering	Public Transport Information	✗	✓	✓	✗	✗	✓	✓	✗
	Social data	✗	✓	✗	✗	✗	✓	✗	✗
	Other Real Time information(bluetooth, webcams, induction loops)	✓	✗	✗	✗	✗	✗	✗	✗
System features	Map generation	✓	✓	✓	✓	✓	✓	✓	✓
	Traffic simulation/ Emergency Simulation	✓	✗	✗	✓	✓	✗	✓	✓
	Emergency Routing	✓	✗	✓	✓	✗	✗	✗	✗
	Navigation	✗	✗	✓	✗	✗	✗	✗	✗
	Bus/ Public Transport	✗	✓	✓	✗	✗	✗	✓	✗

5.3. SETA beyond the state of the Art

Observing the matrix presented before, it is clear how a new system for decision support in mobility should focus on fusing real-time information collected from various sensors with social data and historic data. SETA will provide a step change in Decision Support by developing a holistic approach combining decision support and visual analytics in an integrated interface that provides real-time decision support features to increase the Situation Awareness of decision-makers.

Whilst supporting real-time analysis is fundamental we also aim to support planning in a novel manner, by integrating real-time, historic and crowdsourced data in a unique solution that allows for understanding influences between multiple, possibly previously unknown, factors using methodologies like direct comparison and simulation.

SETA decision support system must also be able to capture, formalise, store, and retrieve decisions (historical, present and future), their rationale and supporting evidences by developing lightweight logging mechanisms that will seamlessly combine interpretations from user interactions, user input, semantic content, and all available metadata to build a 'decision-graph'.

6. Decision Support and Visual Analytics Requirements

The aim of task 5.1 is to understand requirements to develop solutions for large scale visual analytics and decision support solutions for three categories of stakeholders:

- Group #1 – Citizens and communities with mobility needs in the three metropolitan areas where the SETA solution will be deployed. They will be data and information providers as well as users of data and services.
- Group #2 – Decision-makers in metropolitan city councils. They will be both providers and users of data and services, especially predictive and real-time analytics services designed to support daily and long term decision- making. This group is impacted mostly by data quality, data analytics and reporting requirements.
- Group #3 – Businesses that are present or plan to start/extend their presence in metropolitan city councils. They will be both providers and users of data and services, especially services designed to support business expansion activities or day-to-day management of mobility for logistic or workforce management purposes.

While the Visual Analytic approach and Decision Support System discussed how Seta will address the different challenges of Big Data within the context of Mobility, it is important to understand how different requirements of the solution will be addressed. To this extent the requirements from stakeholders and the scenarios identified in WP1 have been analysed to

extract specific visual analytics and decision support needs and challenges, that need to be addressed by this workpackage. To achieve this, we followed the following process:
Analysis of the requirements and scenarios presented in D1.1 for (1) stakeholders and decision makers and (2) public/citizens. Deliverable D1.1 presents a set of scenarios and storyboarding, which identified how the different technical work packages within the project would be integrated and fit with each other:

- All the tasks relating to WP5 were assimilated.
- Each of the roles/tasks was colour coded to highlight different features/datasets and assigned to distinct sections of the final system.
- For each of the roles/tasks we identify the corresponding DSS (Decison SUppoort System) or VA (Visual Analytics) requirement.

The requirements are discussed in the following two sections.

It is important to note that the requirements do not necessary refer to a self-contained piece of software developed within SETA, but to all the technologies and tools we are going to use to implement the environment. For instance, some functionalities could be provided by developing a SETA portal, but others will be supplied by mobile apps, existing web portals or by using specific functionalities of social networking sites. This will reduce the time to delivery during the first phase and will ensure that users can provide relevant feedback during the first evaluation to support the second phase of the project.

6.1. Decision Makers Requirements

Figure 10 shows the analysis of decision maker and stakeholder storyboards, identifying a set of features and aspects of the system envisaged.

maps and overlay - She can see the traffic heat map layer understanding which is the real time traffic situation. **event calendar**- Switching to event calendar section to check particular events in Turin. **planning interface** - For this reason she plans some extra pickup in the stadium area with the Garbage Truck planner. - She wants to know if that accident could be a problem for the trucks planned routes. So she selects the map and choose to show the truck routes layer. She recognizes that it could be a problem for truck #134. **app push notification** - Barbara sends a notification to the workforce mobile application. **Dashboard Interface** - On his work calendar he can see he has blocked today as a day to work on deciding where best to place new bicycles routes across the city and understanding the impact that a new roundabout will have on pollution around the city. Mark opens the SETA Decision Maker Planning Platform and opens the historic data dashboard. **Map Interface and Layers** - Mark looks at the map of historic data and decides to have a look at the movements of bicycles in the city over the past year. Clicking on the right hand side of the screen he brings up a contextual menu that allows to choose which data needs to be shown on the map. Mark clicks on bicycle. The map updates showing only bicycles tracked over last year (Mark decided to not change the temporal horizon in the date slider). - **Map interface and layers** - Looking at the map, Mark can clearly see areas with higher bicycle traffic. Bringing up the layer menu again, Mark chooses to visualise the existing bicycle lanes. - **Zooming capabilities** + - **Time sliders** - Mark uses this visualisation to understand where in the city there are areas of high bicycle traffic but no bicycle lanes. He zooms in on certain areas to better understand the distribution and uses the time slider to see how the traffic changes over the day/week/month/year. - **Annotations** - Satisfied with his theory, Mark adds an annotation to the map for the areas where he would like to add bicycle lanes. Mark takes screenshots of these areas results to place them in his report. - **Map**

visualisation and layers - He now decides to compare the areas with high bicycle traffic and no bicycle lanes with the other traffic. He brings up the menu and enables the public transport and private cars layers **- Modelling dashboard** - Looking at the comparison, Mark decides to focus on two main areas and checks the pollution levels of each. Seeing the pollution level, Mark estimates putting cycle lanes in those areas will improve both traffic and pollution in the city. To check his estimate, Mark opens the Modelling dashboard. **- Zoomable maps and contextual menu** - Moving on to the task of analysing the impact of a new roundabout in the city centre, leading to the shopping mall, Mark goes back to the platform home page. He clicks on the map of historic data and zooms in to the shopping centre. Using the contextual right menu he chooses to visualise public transport and private traffic converging to the area. Immediately he can see the location of the proposed roundabout being a hotspot for traffic and pollution. Looking at the origin destination trajectories he can see the hotspot is exactly where roads that connect Birmingham to nearby cities are merging and where there is lots of constant traffic between the cities. **- graph dashboard + - historic traffic data** - He then decided to investigate the traffic using the graph dashboard, to compare traffic levels over time of the day. Looking at monthly, weekly and daily trends he can see there are higher values in rush hours (8-9 in the morning and 5-6 in the afternoon) but traffic is consistently high throughout the day. Mark saves the graphs to insert in his report. **- Zoomable maps, contextual menu and layers**. - Going back to the map, Mark uses the contextual right menu to bring up the pollution layer. In this case he wants to see where the pollution sensors are placed and he notices they are all at least 10m away from the proposed roundabout. He decides, if the roundabout is built, to place a new pollution sensor on it to get better readings. **- alerts** - After zooming in he chooses a new component from the menu and he adds a roundabout, making sure he connects it to all the incoming roads. He launch the simulation and he gets a warning that the simulation will take over an hour to perform therefore he chooses to receive an alert when it's done.

Figure 10: Systematic text analysis of requirements of stakeholders studying scenarios in D1.1, colour-coding different aspects of the requirements

The solution envisaged is proposed to be an online/standalone dashboard interface, which comprises of several distinct information spaces - namely an interactive map, calendar, modelling section and graph section. The final implementation, however will depend on a variety of other variables that will results from the participatory design and development process.

For example, the outcome of the first brainstorming has identified the need of a map as the primary interaction object, with interactive pop-up interface elements or layers. This has led to the definition of the first set of low-fidelity mock-ups (Section 7 as an example) that will be subsequently refined with our stakeholders using participatory design techniques.

Hence, while the final implemented solution is open, this section attempts to analyse the functional elements of the envisaged solution. One of the most important sections identified was the need to display spatial data on geographical maps. Temporal data will also need to be displayed in a variety of ways - e.g. timeline, navigable facets and calendar views.

Functionality	**Requirements**	**VA/DSS**
Login	Logging of personnel - load personal schedules / plans/ context	DSS
Dashboard interface	Give decision makers an immediate picture of issues/priorities through multiple visualization widgets	

Search	Allow people to search information in different ways, exploiting the richness of the information available (models, data, predictions): ● hierarchy of categories browsing ● faceted browsing ● full text search ● search based on tags that are aimed to classify information according to different dimensions	
Geographic Map	Traffic and highlight hotspots Heatmap of mobility data (e.g. static users/locations of high running activities) Historic data of static points / trajectories / paths Trajectories for Origin-Destination data of different modes Overlay with data other than mobility (e.g. pollution) Static points such as sensors	VA
Calendar View	Planned events Schedules (personal/group)	VA + DSS
Modelling View	Impact of events on routes Impact of routes on pollution	DSS
Graphs	Compare traffic levels of different modalities over daily/weekly/monthly/yearly trends Compare traffic levels over hourly (rush hours) intervals	VA
Contextual Menu	To enable different layers of data presentation - Truck routes - Cycle routes - Pollution - Public transport - Private vehicles	VA
Zooming and Filters	Zoom-in to different regions/zones of interest Filters for visualising different temporal resolutions	VA
Annotations	Virtually add infrastructure components (e.g. lanes / routes / roundabouts)	VA
Exploitation of information	Social networks monitoring (Facebook, Twitter, Flickr) Social Network predictions	DSS+VA

implicitely provided by citizens	Visualization through charts/timelines	
Capture additional information on-demand	Geospatial localization map about cooperating citizens registered in the system Support for sending requests to citizens	DSS+VA
Content validation	Verification of information provided by cititziens, their veracity and reliability Visual display of veracity and reliability	DSS+VA
Planning and Interaction with predictive tools	New schedules (e.g. garbage pickups) New routes and roundabouts	VA + DSS
Support strategic decision-making	Visualization of: ● Mobility predictions ● Probability of traffic/congestion in some areas ● Consequences of decisions	VA+DSS
Notifications	Provide push updates to teams via mobile apps Provide updates when simulations/processes complete execution	DSS + VA
Reporting	Capture Screenshots and Graphs to create reports	DSS + VA

6.2. Public/Citizens Requirements

A similar process was performed for understanding the requirements of the decision support and visual analytics solutions for public/citizens.

Login interface - Joe logs in to the Birmingham Real-Time Mobility Portal to check his journey before setting off. **Zoomable Dashboards for different data sources** - Zooming-in on the city-centre Joe can see that there is an 85% chance of rain between 3 and 4 pm, the time he would be cycling home, so he decides to walk instead. **- Historic journeys** - He selects one of his saved journeys, a walk between home and the University. **-Heatmaps and overlays** - Looking at the colour-coded pollution map for the route, he can see that pollution levels are quite high, the traffic overlay shows slow-moving traffic. **- Link to live-streams feeds** + **- Citizens data collection and visualisation interface:** He clicks on a webcam icon and opens a live stream; he can see that a bus has broken down. Warning icons start appearing on the map, he clicks one which says 'number 52 bus has broken down on Church Street, nothing moving!' - **Map-based routing** + **Real-time alerts and notifications** + **Citizens data collection:** Joe would rather avoid the pollution so he selects an alternative route via the park; there is one alert on this route dated 4 pm yesterday: 'watch-out for deep puddles on the path here' - **Citizens observatory - commenting and feedback:** A member of the council parks department has marked the message 'resolved' and commented 47 minutes ago 'thank you for reporting it,

we have cleared the drains and the puddles have drained'. **- Smartwatch app:** Remembering his umbrella, Joe leaves the house at 8:25, giving himself plenty of time to detour via the park. On the way, his smart-watch monitors his walking and congratulates him on achieving his 30 minute daily exercise target. **- live data streams display** Joe checks the webcam on the bus, the images of individuals are deliberately blurred so he can't recognise anyone, but he can see that there are several spare seats upstairs. **- Citizens observatory app for reporting:** Joe and his girlfriend share his umbrella and walk to the bus-stop. Unfortunately someone has vandalised the bus-stop and there is broken glass on the ground. Joe reports this via the portal, taking a photo as evidence. Avoiding the glass they wait for only a few minutes until the bus arrives on time. **Journey Assistant Interface:** George, after looking at his calendar, starts thinking about the best possible way to be in Milton Keynes University at 11 tomorrow for a meeting. His meeting is 11- 1 and then he needs to come back to Birmingham by 3pm to pick up his child from school **Journey Assistant Interface - OD form:** George opens the SETA Personal Journey Assistant app, chooses the Journey Assistant functionality and inputs his home postcode and his destination postcode and time of departure and arrival. **Journey Assistant Interface - user preferences Setting** - Before pressing "Search" George checks his personal preferences. George has set up his preference when he first used the app, choosing cycling and public transport as favourite modalities of transport. Happy that the system has remembered his preferences, George presses "Search".**WP4/5 Generation of best multimodal path:** George can see 3 different options for his journey:car travel, highlighted as fastest and therefore most efficient; bus travel, highlighted as lowest cost; bicycle + train travel, highlighted as the one best matching his preferences**. Visualisation of route on a map:** George clicks on the car travel and views the directions on a map, including the potential traffic blocks, miles to destination, car emissions (NOx, greenhouse gasses) and cost of travelling by car (given by maintenance, insurance, parking etc.) George can choose to press a button to view different parking options and availability but chooses instead to go back to the results screen. George now has a quick look at the bus journey, that includes connections from the bus stop at home to the interchange and then the bus to Milton Keynes and a bus from the Milton Keynes interchange to his final destination. Although George can see this travel is very cheap and easy to book (the app offers a link for ticketless booking) the time to reach his destination is over 3 hours and George would be late at his meeting. George goes back to the Results Page and presses the bicycle + train option. The Journey Assistant presents him the total cost of travelling, the amount of calories/activity he will be doing and the dynamic pollution levels of the route chosen and the weather forecast. George chooses this option.**Connection to booking API:** It's 8am in the morning and George is having breakfast and checking his plan for the day. Having decided yesterday to go to Milton Keynes by bicycle and bus George wants to check if it's sensible for him to cycle to the train station, take the bicycle on the train and then cycle in Milton Keynes or if it's better to leave the bicycle in the parking at the train station. The other alternative he considers is public transport to the station. **Visualisation of possible cycling routes on a map** - Given he has plenty of space, before setting off George wants to ensure he will cycle through a non-polluted route.George asks the Journey Assistant to check for alternatives routes based on pollution levels. **replanning**: Given the delay, the app offers George the possibility of re-planning his journey. George presses "Re-Plan" and the app comes back with suggestions on how to reach Milton Keynes University as fast as possible. **user history** - The Journey Assistant knows that George has left his bicycle in Birmingham therefore it recommends a few options to reach its final destination: an Uber Taxi, as this is the fastest option; a city-bike to rent, as this is George's usual preferred means of transport **alerts** - At 10.40 the SETA Journey Assistant alerts George that it's time to book the taxi.George clicks on the Uber link but there are no taxis available to be booked at that moment in time. George goes back to the SETA Journey Planner and chooses the alternative option. **feedback** - The SETA Journey Planner asks for feedback on the reason of the different choice, George clicks "No taxi available" between the options.**voice commands/smartwatch:** Before getting off the train, George clicks on the possible routes for reaching Milton Keynes University, choosing in this case the shortest route. As he does not know Milton Keynes but he needs to cycle, George chooses to use voice directions.

Figure 11: Systematic text analysis of requirements of public by studying scenarios in D1.1, colour-coding different aspects of the requirements

The following table summarises the requirements that emerged.

Functionality	**Requirements**	**VA/DSS**
Logging to user account	Preserve historical route interests or preferences (e.g. modality/price)	DSS
Geographical Map	Weather prediction Heatmap for point data Pollution map Traffic overlay Warnings and alerts on the map (e.g. road incidents) Multiple routes and directions Historic route data	VA + DSS

	Route map (cycling/walking routes)	
Generic Interface visualisation	Data from external source (e.g. webcam live streams) Citizen Observatory data (e.g. citizen generated data) Location-based services	DSS + VA
Journey Assistant / Planning	Connection with Calendar Origin-destination Select preferences (e.g. modes) Present multiple routes/modes Display characteristics of routes (e.g. costs, miles to destination, car emissions, connections, pollution exposure, weather on the way)	DSS + VA
Multimodal data capture and access (e.g. smartwatch or mobile app)	Monitoring activities/steps Present / query possible alternative routes Push information	DSS
Reporting	Report issues/concerns via Citizen Observatory Report reasons for choosing alternative routes Submit reports/data seamlessly across multiple devices Geolocalization of reports Support for uploading text, images, and video/audio Content tagging/categorization. Implicit and explicit sharing (e.g. reporting activity through senssors)	DSS
Alerts and notifications	Reminders and display incidents Registration to news/alert/report categories Support for choosing the location to receive alerts about, and the size of the area Support for mail, SMS, RSS notification	DSS + VA
Make the access to the platform and the data as open as possible	Data API Dedicated apps and interfaces	DSS

6.3. Administration/Management Functionalities

In addition to the specific requirements identified for Decision-Makers and Public/Citizens we have identified a set of requirements that address admin and management needs.

Need	Functionality
Build up the SETA community	User registration
Get the wider citizens involvement	• Anonymous access to some resources/functionalities • Language localization (to offer the platform in the native language)
Support user interaction through different devices	Interface customization according to the set of allowed devices, at least for web and mobile.
Provide different functionalities to different users (e.g. authorities vs. citizens)	Support different roles, with associated sets of access privileges to resources/functionalities

7. Prototyping a Solution

An analysis of the domain and the subsequent analysis of the requirements presented an opportunity to develop low-fidelity mockups as initial attempts to describe a solution to end users. The mockups will then be refined using an iterative process: we will create mock-ups for every stage of the scenarios described in D1.1 and validate them during a series of walk-through sessions with end users.
As a result, the SETA-UC4 scenario (see D1.1) has been used as a starting point for the following mock-ups.

Data Visualisation and Analysis

1	It's 9am. Mark sits down at his desk to start the working day. On his work calendar he can see he has blocked today as a day to work on deciding where best to place new bicycles routes across the city and understanding the impact that a new roundabout will have on pollution around the city. Mark opens the SETA Decision Maker Planning Platform and opens the historic data dashboard.
2	Mark looks at the map of historic data and decides to have a look at the movements of bicycles in the city over the past year. Clicking on the right hand side of the screen he brings up a contextual menu that allows to choose which data needs to be shown on the map. Mark clicks on bicycle. The map updates showing only bicycles tracked over last year (Mark decided to not change the temporal horizon in the date slider).

3	Looking at the map, Mark can clearly see areas with higher bicycle traffic. Bringing up the layer menu again, Mark chooses to visualise the existing bicycle lanes.
4	Mark uses this visualisation to understand where in the city there are areas of high bicycle traffic but no bicycle lanes. He zooms in on certain areas to better understand the distribution and uses the time slider to see how the traffic changes over the day/week/month/year.

Figure 12 presents an example walkthrough for the Decision Maker Planning Platform interface: there are three main columns - distinctly dividing the workspace into three sections. The first (left column, Section 5) is essentially a filtering section that presents different facets of the data, each facet providing means of accessing a specific aspect of the underlying data. The central column (Sections 1-4) presents different views of the data, while the right column (Section 6, 7) presents tools specific to decision makers.

Looking at the tasks performed by Mark in the scenario, this interface allows for:

- **Temporal filter (e.g. to choose only data from the past year):** This filter firstly provides an overview of the volume of the data being presented as well as provides a way for users to select a time period for which the analyst can view the data. Filters are real-time and coordinated - hence, any selection of filters will update other widgets/elements in the interface.
- **Modality of transport filter (e.g. to visualize only bicycle trace data):** Every modality of transport is a different layer that can be turned on and off using a filter.
- **Location filter (e.g. to focus only on the city centre)** provides a way to directly query underlying data regarding specific postcodes, or type of location (below, location type).

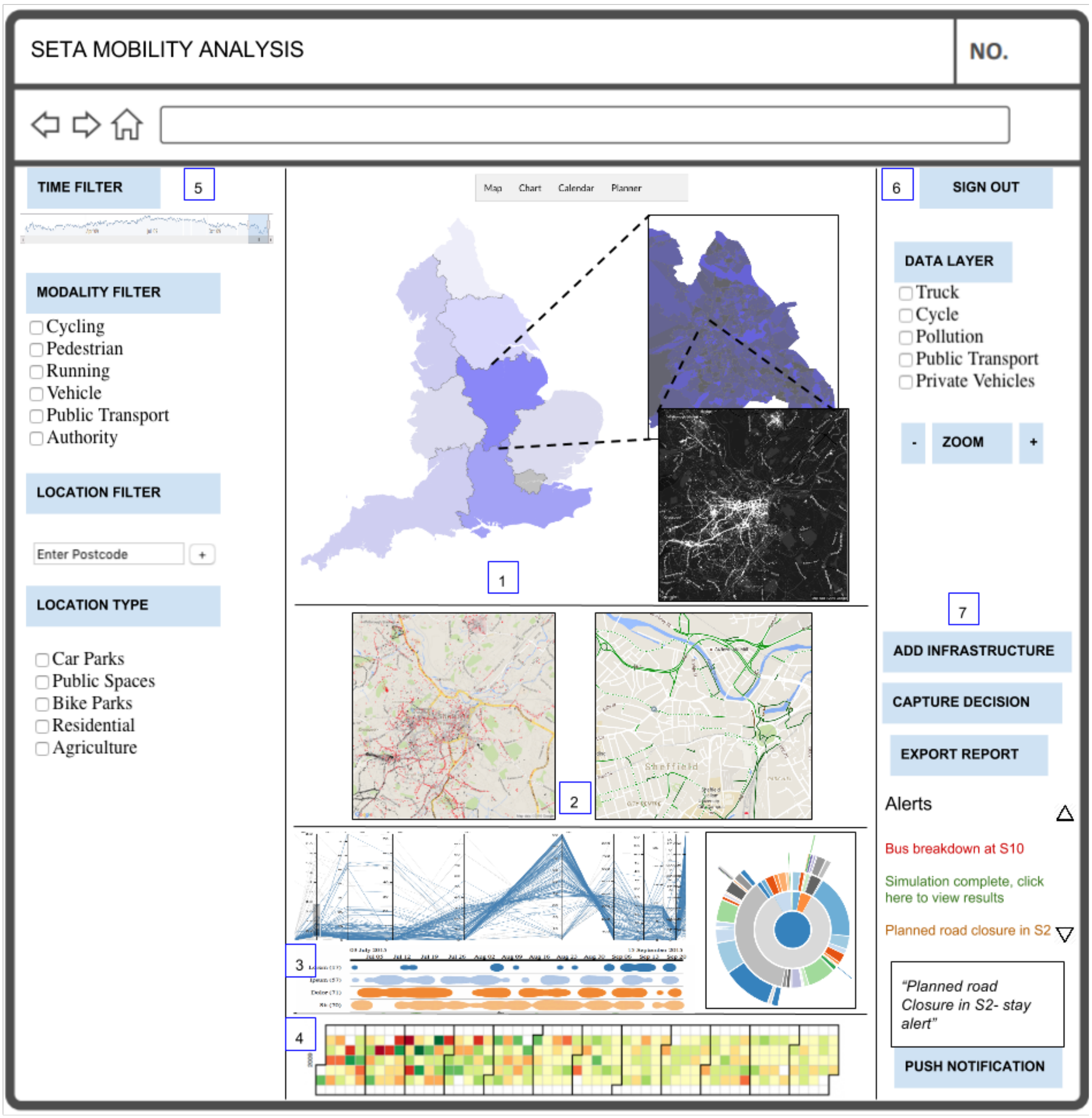

Figure 12: Developing a prototype solution. Three core sections of the interface divide the different aspects of the proposed solution; Left - a faceted filtering interface; Center - presentation of the data, Right - Decision-support tools. Some of the figures are generated by Seta modules currently in development. The remaining images are created by D3.js, and are available as examples from http://d3js.org

Selecting the right data, Mark can then focus on the different views available to explore the underlying data - map, chart, calendar, planner.
The map section is visualised by selecting the Map tab on top of the central data view column. The Map view (Section 1, 2) presents geographical views of the location-specific data. Two main methods of exploration are offered:

- **Hierarchical exploration:** This visualization technique can help reduce massive volumes of geographical data by clustering underlying data into zones/boundaries. For example, millions of map points can be summarised into a higher level group (e.g. postcode or ward level). In Figure 12, Section 1 shows an exploration carried out using a map of England, visualising different zones (South, Midlands, North etc.), color-coded on the basis of a certain parameter (for example, population). Clicking on an individual section zooms the map into a higher level of detail, here - a city wide ward level. This presents color coded ward boundaries that clusters the respective data. Finally, clicking on the ward level provides access to the raw data instances on a map.
- **Thematic exploration of maps** can also be performed by selecting different themes of data - here, Mark can click on any one of the data layers to explore the context of the data. Figure 13 illustrates a scenario where the user can thematically explore different features of the underlying data, incrementally adding layers such as road average speed (1), building occupancy (2), public park occupancy (3). All the layers are built upon underlying activity, speed and location data collected by WP2 modules.

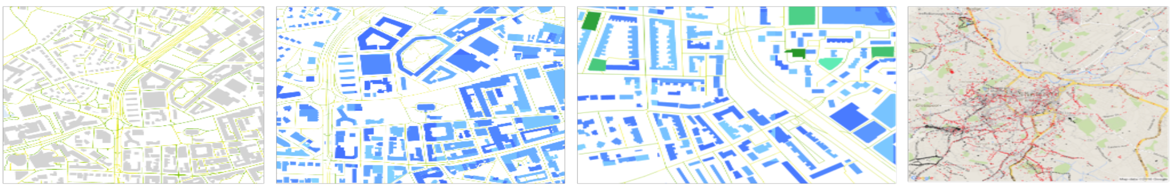

Figure 13: Incrementally adding layers of data can help analysts understand large volumes of data - (1) a map presenting colour coded roads reflecting average speed; (2) road average speed is provided more context by adding a color coded building occupancy layer; (3) further context is added by presenting public parks occupancy level layer; (4) raw underlying large scale data that is aggregated to generate the maps (1-3) - this data is collected by modules developed in WP2. All figures are generated from Seta modules currently in development.

For what regards the possibility of zooming in to see different levels of granularity in the data, the Data layer (Figure 12, top right) enables Mark to add/remove layers of information for analysis - these are reflected in the maps / calendar, parallel plots etc. The Zoom button indicates a semantic zoom feature, where a hierarchy of contexts encapsulates different data elements - this is particularly relevant to zooming in the map.

10	Moving on to the task of analysing the impact of a new roundabout in the city centre, leading to the shopping mall, Mark goes back to the platform home page. He clicks on the map of historic data and zooms in to the shopping centre. Using the contextual right menu he chooses to visualise public transport and private traffic converging to the area. Immediately he can see the location of the proposed roundabout being a hotspot for traffic and pollution. Looking at the origin destination trajectories he can see the hotspot is exactly where roads that connect Birmingham to nearby cities are merging and where there is lots of constant traffic between the cities.
11	He then decided to investigate the traffic using the graph dashboard, to compare traffic levels over time of the day. Looking at monthly, weekly and daily trends he can see there are higher values in rush hours (8-9 in the morning and 5-6 in the afternoon) but traffic is consistently high throughout the day. Mark saves the graphs to insert in his report.

The SETA solution will include a graph dashboard for creating custom data visualization that can facilitate comparison and analysis.

Looking at the tasks performed by Mark in the scenario, this interface allows for presenting different summaries of data from a variety of perspectives. For example, a parallel plot (Section 3, top) is used to present a large number of data instances that contain categorical data. Another visualisation is a parallel timeline, presenting a large number of data instances containing temporal details. For example, such activities could be: walking, cycling, running, moving in vehicle etc. Presenting a summary in such a way can provide useful insight into when specific activities occur. The Calendar view presents temporal data as a color-coded matrix that can indicate when critical activities are planned.

Data manipulation, annotation and reuse

12	Going back to the map, Mark uses the contextual right menu to bring up the pollution layer. In this case he wants to see where the pollution sensors are placed and he notices they are all at least 10m away from the proposed roundabout. He decides, if the roundabout is built, to place a new pollution sensor on it to get better readings.

13	Again Mark needs to check his hypothesis using the modelling and simulation platform. He opens the modelling dashboard and he zooms in on the desired area.
14	After zooming in he chooses a new component from the menu and he adds a roundabout, making sure he connects it to all the incoming roads. He launch the simulation and he gets a warning that the simulation will take over an hour to perform therefore he chooses to receive an alert when it's done.
15	After 45 mins Mark receives an alert from the SETA planning platform, both via email and text message. He logs back into the platform and clicks on the received message. This opens the modelling dashboard with the results of the simulation. Mark checks the results and takes a screenshot of the prediction.

Figure 12 shows a specific section of the interface dedicated to ways for decision makers to act upon the data they view. Mark is presented with a login interface, to customise and save his work in a dedicated area.

Looking at the tasks performed by Mark in the scenario, this interface allows for:

- **Adding custom annotations to a map**, for example clicking on 'Add Infrastructure' button presents a dialog for the prompt of the type of infrastructure and specific location (e.g. a roundabout, bus stop). This is then captured and fed to the modelling modules.
- **Capturing decisions taken:** Clicking on the 'Capture decision' button prompts users for the details about the decision, as well as 'freezes' the current context of the decision maker (e.g. a screenshot of their view, underlying data, timestamp as well as filter settings) so that the same conditions can be recreated for the analyst for post event analyses.
- **Receiving alerts:** Alerts are presented as color-coded messages in a scrollable panel at the bottom of the interface. The color coding is done on the basis of priority, as defined by either the decision maker or the Seta data analysis system.
- **Sending alerts:** A text input provides Mark with means to type short messages which can then be broadcast to all recipients (all mobile app users, on-the-ground volunteers etc.).
- **Exporting the work done:** All activities can be exported as a PDF or HTML report, which is generated once Mark clicks on the 'Export Report'.

The mockup interface is a preliminary attempt at summarising the requirements as a demonstrable solution, incorporating some of the modules already currently in development within the Seta project. The next steps involve developing basic versions of different components in the interface, which will then be (either individually or combined) presented to stakeholders and decision makers. This will be conducted as a set of structured/unstructured participatory design sessions, over the next phase of the project.

8. Conclusion

This deliverable summarized the activities in the Work Package 5 during the requirements gathering phase. The plans and state of the art for Visual Analytics and Decision Support activities are presented independently of each other in Sections 4 and 5: they discuss how the activities will further the state of the art in mobility and present some ideas of how the SETA project will be developing some technologies. For example, visual analytics will explore the idea of multi-dimensional multi-hierarchical visualisations for reducing the size of data in a Big Data setting. Decision Support on the other hand will develop tools to capture the contexts of decision makers to better align their information and task needs with the right set of data. This also can help in addressing the significant demands for Big Data. Decision Support activities will also explore mechanisms for real-time communication between decision makers as well as personnel on the ground. These ideas have evolved from studying the domain of mobility for large scale areas as well as the state of the art. Additionally, requirements gathering sessions in parallel seeded new ideas for the two areas of activity within this Work Package.

A significant effort was involved in requirements collection activities (e.g. stakeholder focus groups, online questionnaires etc.) within Work Package 1 and hence, a lot of requirements and ideas were collected which reflected on the tasks of the project as a whole. This provided an excellent opportunity for the Work Package to explore how the two activities (VA and DSS) could fit and integrate together. This helped draw out specific requirements for the Work Package, as presented in Section 6. Finally, studying the requirements, a User Interface Mockup was created, which was presented as a walkthrough based on a scenario developed in Work Package 1 (Section 7).

Following the creation of the mockups, the first step would be to re-iterate the process of cognitive walkthroughs. This will be an informal conversation between the Work Package participants to ensure the mockups capture all the requirements within a variety of contexts and then by validation with end users. This will be then followed by independent implementations of the technology, ensuring that the implementations can be integrated in a coherent system finally. This will then be followed by incremental user evaluations, following the process of an iterative user centered design.

9. References

1. Adrienko N, Adrienko G. Spatial generalization and aggregation of massive movement data (2011). IEEE Transactions on Visualization and Computer Graphics, 17(2): 205-219.

2. Aigner, W., Miksch, S., Schumann, H., & Tominski, C. (2011). Visualization of time-oriented data. Springer Science & Business Media.

3. Al-Hajj, S., Pike, I., & Fisher, B. (2013). Interactive Dashboards: Using Visual Analytics for knowledge Transfer and Decision Support. Public Health's Wicked Problems: Can InfoVis Save Lives? IEEE Visual Analytics Science and Technology, Information Visualization, Scientific Visualization, 1-8.

4. Andrienko G, Andrienko N, Hurter C, Rinzivillo S, Wrobel S. (2013). Scalable analysis of movement data for extracting and exploring significant places. IEEE Transactions on Visualization and Computer Graphics, 19(7): 1078-1094.

5. Andrienko, G., Andrienko, N., & Wrobel, S. (2007). Visual analytics tools for analysis of movement data. ACM SIGKDD Explorations Newsletter, 9(2), 38-46.

6. Andrienko, G., Andrienko, N., Hurter, C., Rinzivillo, S., & Wrobel, S. (2011). From movement tracks through events to places: Extracting and characterizing significant places from mobility data. In Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on (pp. 161-170). IEEE.

7. Andrienko, G., Andrienko, N., Rinzivillo, S., Nanni, M., & Pedreschi, D. (2009). A visual analytics toolkit for cluster-based classification of mobility data. InAdvances in Spatial and Temporal Databases (pp. 432-435). Springer Berlin Heidelberg.

8. Andrienko, N., & Andrienko, G. (2012). Visual analytics of movement: An overview of methods, tools and procedures. Information Visualization, 1473871612457601.

9. Andrienko, N., Andrienko, G., & Gatalsky, P. (2000, May). Supporting visual exploration of object movement. In Proceedings of the working conference on Advanced visual interfaces (pp. 217-220). ACM.

10. Anselin, L. (2004). Exploring spatial data with GeoDaTM: a workbook. Urbana, 51, 61801.

11. Bieker, L., Behrisch, M. and Ruppe, S. (2012). EmerT – a web based decision support tool for Traffic Management. [online] ResearchGate. Available at: https://www.researchgate.net/publication/259902407_EmerT_-_a_web_based_decision_support_tool_for_Traffic_Management [Accessed 3 Mar. 2016].

12. Bizer, C., Boncz, P., Brodie, M. L., & Erling, O. (2012). The meaningful use of big data: four perspectives--four challenges. ACM SIGMOD Record, 40(4), 56-60.

13. Bollier, D., & Firestone, C. M. (2010). The promise and peril of big data (p. 56). Washington, DC, USA: Aspen Institute, Communications and Society Program.

14. Bonczek, R. H., Holsapple, C. W., & Whinston, A. B. (2014). Foundations of decision support systems. Academic Press.

15. Boulos, M. K., Viangteeravat, T., Anyanwu, M. N., Nagisetty, V. R., & Kuscu, E. (2011). Web GIS in practice IX: a demonstration of geospatial visual analytics using Microsoft Live Labs Pivot technology and WHO mortality data.International journal of health geographics, 10(1), 19.

16. Byrnes, J., & Dollery, B. (2002). Local Government Failure in Australia? An Empirical Analysis of New South Wales. Australian Journal of Public Administration, 61, 54- 64

17. C. Weaver. Cross-dimensional visual queries for interactive+animated analysis of movement (2008). In Workshop in Geospatial Visual Analytics.

18. Cao N, Sun J, Lin Y R et al (2010). Facetatlas: Multifaceted visualization for rich text corpora. IEEE Transactions on Visualization and Computer Graphics, 16(6): 1172-1181.

19. Carsjens, G. J., & Ligtenberg, A. (2007). A GIS-based support tool for sustainable spatial planning in metropolitan areas. Landscape and urban planning, 80(1), 72-83.

20. Chi, E. H. H. (2000). A taxonomy of visualization techniques using the data state reference model. In Information Visualization, 2000. InfoVis 2000. IEEE Symposium on (pp. 69-75). IEEE.

21. Cho, C.-J. (2002). The Korean growth-management programs: issues, problems and possible reforms. Land Use Policy, 19, 13-27.

22. Chua, A., & Vande Moere, A. (2016). BinSq: visualizing geographic dot density patterns with gridded maps. Cartography and Geographic Information Science, 1-20.

23. Crnovrsanin, T., Muelder, C., Correa, C., & Ma, K. L. (2009). Proximity-based visualization of movement trace data. In Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on (pp. 11-18). IEEE.

24. Daassi, C., Nigay, L., & Fauvet, M. C. (2005). A taxonomy of temporal data visualization techniques. Information-Interaction-Intelligence, 5(2), 41-63.

25. Dang T N, Anand A, Wilkinson L. (2013). TimeSeer: Scagnostics for high-dimensional time series. IEEE Transactions on Visualization and Computer Graphics, 19(3): 470-483.

26. Thompson, T., Levine, J., Bennett, J., Bremer, P., Gyulassy, A., Pascucci, V., Pébay, P. (2011) Analysis of large-scale scalar data using hixels, in: 2011 IEEE Symposium on Large Data Analysis and Visualization (LDAV), pp. 23–30.

27. Diakopoulos, N., Naaman, M., & Kivran-Swaine, F. (2010). Diamonds in the rough: Social media visual analytics for journalistic inquiry. In Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on (pp. 115-122). IEEE.

28. Dykes, J. A., & Mountain, D. M. (2003). Seeking structure in records of spatio-temporal behaviour: visualization issues, efforts and applications.Computational Statistics & Data Analysis, 43(4), 581-603.

29. Endsley, M. R. (1995) Toward a theory of situation awareness in dynamic systems: Situation awareness. Human factors 37, 1.

30. Ferreira, N., Poco, J., Vo, H. T., Freire, J., & Silva, C. T. (2013). Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. IEEE Transactions on Visualization and Computer Graphics, 19(12), 2149-2158.

31. Gamon, M., Aue, A., Corston-Oliver, S., & Ringger, E. (2005). Pulse: Mining customer opinions from free text. In international symposium on intelligent data analysis (pp. 121-132). Springer Berlin Heidelberg.

32. González, A., Donnelly, A., Jones, M., Chrysoulakis, N., & Lopes, M. (2013). A decision-support system for sustainable urban metabolism in Europe.Environmental Impact Assessment Review, 38, 109-119.

33. Gregory, M. L., Chinchor, N., Whitney, P., Carter, R., Hetzler, E., & Turner, A. (2006). User-directed sentiment analysis: Visualizing the affective content of documents. In Proceedings of the Workshop on Sentiment and Subjectivity in Text (pp. 23-30). Association for Computational Linguistics.

34. Guo H, Wang Z, Yu B, et al (2011). TripVista: triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In: Proceedings of the pacific visualization symposium PacificVis, IEEE,

35. Guo, B., Zhang, D., & Wang, Z. (2011). Living with internet of things: The emergence of embedded intelligence. In Internet of Things (iThings/CPSCom), 2011 International Conference on and 4th International Conference on Cyber, Physical and Social Computing (pp. 297-304). IEEE.

36. Guo, B., Zhang, D., Yu, Z., Liang, Y., Wang, Z., & Zhou, X. (2013). From the internet of things to embedded intelligence. World Wide Web, 16(4), 399-420.

37. Holten, D. (2006). Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. IEEE Transactions on visualization and computer graphics, 12(5), 741-748.

38. Thomas J. and Cook, K. (2005) Illuminating the Path: Research and Development Agenda for Visual Analytics. IEEE-Press.

39. Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big data and its technical challenges. Communications of the ACM, 57(7), 86-94.

40. Joia P, Paulovich F V, Coimbra D et al (2011). Local affine multidimensional projection. IEEE Transactions on Visualization and Computer Graphics, 17(12): 2563-2571.

41. Ma, K., Lum, E., Yu, H., Akiba, H., Huang, M.-Y, Wang, Y. and Schussman, G. (2005) “Scientific discovery through advanced visualization,” in Proceedings of DOE SciDAC 2005 Conference, San Francisco.

42. Kapler, T., & Wright, W. (2005). GeoTime information visualization. Information Visualization, 4(2), 136-146.

43. Kapler, T., & Wright, W. (2005). GeoTime information visualization. Information Visualization,4(2), 136-146.

44. Karam, G. M. (1994). Visualization using timelines. In Proceedings of the 1994 ACM SIGSOFT international symposium on Software testing and analysis (pp. 125-137). ACM.

45. Keim, D. A., Kohlhammer, J., Ellis, G., & Mansmann, F. (Eds.). (2010).Mastering the information age-solving problems with visual analytics. Florian Mansmann.

46. Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008).Visual analytics: Scope and challenges (pp. 76-90). Springer Berlin Heidelberg.

47. Keim, D. A., Panse, C., & Sips, M. (2003). Visual data mining of large spatial data sets. In International Workshop on Databases in Networked Information Systems (pp. 201-215). Springer Berlin Heidelberg.

48. Kek, A. G., Cheu, R. L., Meng, Q., & Fung, C. H. (2009). A decision support system for vehicle relocation operations in carsharing systems. Transportation Research Part E: Logistics and Transportation Review, 45(1), 149-158.

49. Klein, G. A. (1989) Recognition-primed decisions. Advances in man-machine systems research (Ed: W. B. Rouse).

50. Kr´stajic M, Bertini E, Keim D A. (2011) CloudLines: Compact display of event episodes in multiple time-series. IEEE Transactions on Visualization and Computer Graphics, 17(12): 2432-2439

51. Kraak, M. J. (2003). The space-time cube revisited from a geovisualization perspective. InProc. 21st International Cartographic Conference (pp. 1988-1996).

52. LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2013). Big data, analytics and the path from insights to value. MIT Sloan Management Review, 21.

53. Liu, L., Biderman, A., & Ratti, C. (2009). Urban mobility landscape: Real time monitoring of urban mobility patterns. In Proceedings of the 11th International Conference on Computers in Urban Planning and Urban Management (pp. 1-16).

54. Liu, S., Pu, J., Luo, Q., Qu, H., Ni, L. M., & Krishnan, R. (2013). Vait: A visual analytics system for metropolitan transportation. Intelligent Transportation Systems, IEEE Transactions on, 14(4), 1586-1596.

55. MacEachren, A. M., Gahegan, M., Pike, W., Brewer, I., Cai, G., Lengerich, E., & Hardistry, F. (2004). Geovisualization for knowledge construction and decision support. Computer Graphics and Applications, IEEE, 24(1), 13–17.

56. Marcus, A., Bernstein, M. S., Badar, O., Karger, D. R., Madden, S., & Miller, R. C. (2011). Twitinfo: aggregating and visualizing microblogs for event exploration. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 227-236). ACM.

57. Mazumdar, S., Ciravegna, F., Gentile, A. L., & Lanfranchi, V. (2012). Visualising context and hierarchy in social media. In International Workshop on Intelligent Exploration of Semantic Data (IESD'2012) at EKAW (Vol. 2012).

58. Mountain, D. (2005). Visualizing, querying and summarizing individual spatio-temporal behaviour.Exploring geovisualization, 181-200.

59. Mountain, D. (2005). Visualizing, querying and summarizing individual spatio-temporal behaviour.Exploring geovisualization, 181-200.

60. Nijkamp, N., Borzacchiello, M.T., Ciuffo, B., Torrieri, F. (2007). Sustainable Urban Land Use and Transportation Planning: A Cognitive Decision Support System for the Naples Metropolitan Area. International Journal of Sustainable Transportation Vol. 1, Iss. 2.

61. Nishi, K., Tsubouchi, K., & Shimosaka, M. (2014). Hourly pedestrian population trends estimation using location data from smartphones dealing with temporal and spatial sparsity. In Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 281-290). ACM.

62. Omrani, H., Gerber, P., Ion, L. and Trigano, P. (2008). A decision support tool for evaluating mobility projects. [online] ResearchGate. Available at:

https://www.researchgate.net/publication/242563821_A_decision_support_tool_for_evaluating_mobility_projects [Accessed 4 Mar. 2016].

63. Laube P. and Imfeld, S. (2002). Analyzing relative motion within groups of trackable moving point objects. In GIScience ’02: Proceedings of the Second International Conference on Geographic Information Science, pages 132–144.

64. Scheepens R, Willems N, van de Wetering H et al. (2011). Composite density maps for multivariate trajectories. IEEE Transactions on Visualization and Computer Graphics, 17(12): 2518-2527.

65. Shi C, Cui W, Liu S, Xu P, Chen W, Qu H. (2012). RankExplorer: Visualization of ranking changes in large time series data. IEEE Transactions on Visualization and Computer Graphics, 18(12): 2669-2678.

66. Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In Visual Languages, 1996. Proceedings., IEEE Symposium on (pp. 336-343). IEEE.

67. Slingsby, A., Dykes, J. & Wood, J. (2011). Exploring Uncertainty in Geodemographics with Interactive Graphics. IEEE Transactions on Visualization and Computer Graphics, 17(12), pp. 2545-2554. doi: 10.1109/TVCG.2011.197

68. Steiniger, S., & Hunter, A. J. (2013). The 2012 free and open source GIS software map–A guide to facilitate research, development, and adoption. Computers, Environment and Urban Systems, 39, 136-150.

69. Sun GD, Wu YC, Liang RH et al. (2013) A survey of visual analytics techniques and applications: State-of-the-art research and future challenges. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 28(5): 852–867. DOI 10.1007/s11390-013-1383-8

70. Karmakharm, T., Richmond, P., Romano, DM (2010). Agent-based Large Scale Simulation of Pedestrians With Adaptive Realistic Navigation Vector Fields. TPCG.

71. Tominski C, Schumann H, Andrienko G, Andrienko N. (2012) Stacking-based visualization of trajectory attribute data. IEEE Transactions on Visualization and Computer Graphics, 18(12): 2565-2574.

72. Tory, M., & Moller, T. (2004). Rethinking visualization: A high-level taxonomy. In Information Visualization. INFOVIS 2004. IEEE Symposium on (pp. 151-158). IEEE.

73. Turkay C, Lundervold A, Lundervold A J, Hauser H. (2012) Representative factor generation for the interactive visual analysis of high-dimensional data. IEEE Transactions on Visualization and Computer Graphics, 18(12): 2621-2630.

74. van Wijk, J.J. (2005). The value of visualization. In: IEEE Visualization, p. 11

75. Vitale, A., Festa, D., Guido, G. and Rogano, D. (2014). A Decision Support System based on Smartphone Probes as a Tool to Promote Public Transport. Procedia - Social and Behavioral Sciences, 111, pp.224-231.

76. Von Landesberger T, Kuijper A, Schreck T, Kohlhammer J, van Wijk J J, Fekete J D, Fellner D W. (2011). Visual analysis of large graphs: State-of-the-art and future research challenges. Computer Graphics Forum, 30(6): 1719-1749.

77. Weber, M., Alexa, M., & Müller, W. (2001). Visualizing Time-Series on Spirals. In Infovis (Vol. 1, pp. 7-14).

78. Wood J, Badawood D, Dykes J, Slingsby A. (2011). BallotMaps: Detecting name bias in alphabetically ordered ballot papers. IEEE Transactions on Visualization and Computer Graphics, 17(12): 2384-2391.

79. Xu P, Wu Y, Wei E, Peng T Q, Liu S, Zhu J H, Qu H. (2013). Visual analysis of topic competition on social media. IEEE Transactions on Visualization and Computer Graphics, 19(12)

80. You, L., Motta, G., Sacco, D. and Ma, T. (2014). Social data analysis framework in cloud and Mobility Analyzer for Smarter Cities. Proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics.

81. Yu, H. (2006). Spatial-temporal GIS design for exploring interactions of human activities, Cartography and Geographic Information Science, 33(1), pp. 3-19

82. Yu, J., Pande, A., Nezamuddin, N., Dixit, V. and Edwards, F. (2014). Routing Strategies for Emergency Management Decision Support Systems During Evacuation. Journal of Transportation Safety & Security, 6(3), pp.257-273.

83. Zhang, L., Stoffel, A., Behrisch, M., Mittelstadt, S., Schreck, T., Pompl, R., ... & Keim, D. (2012). Visual analytics for the big data era—A comparative review of state-of-the-art commercial systems. In Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on (pp. 173-182). IEEE.

84. Zhao J, Chevalier F, Collins C, Balakrishnan R. (2012). Facilitating discourse analysis with interactive visualization. IEEE Transactions on Visualization and Computer Graphics, 18(12): 2639-2648

85. Zhong Liang, ChiTian He, Zhang Xin, (2010). Feature based visualization algorithm for large-scale flow data, in: Second International Conference on Computer Modeling and Simulation, 2010, ICCMS ’10, vol. 1, pp. 194–197