RENDER
FP7-ICT-2009-5
Contract no.: 257790
www.render-project.eu

RENDER

Deliverable D3.1.3

Reusable Implementations of Diversity-aware Interfaces

Editor:	Patrick Wollner, Instrata
Author(s):	Patrick Wollner, Instrata; Dr Rachel Jones, Instrata; Tanya Goldhaber, Instrata
Deliverable Nature:	Report (R)
Dissemination Level: (Confidentiality)	Public (PU)
Contractual Delivery Date:	September 2012
Actual Delivery Date:	March 2013
Suggested Readers:	Developers creating tools or case study prototypes
Version:	1.0
Keywords:	visualisation techniques, visualisation guidelines, heuristic evaluation of tools

Full Project Title:	RENDER – Reflecting Knowledge Diversity
Short Project Title:	RENDER
Number and Title of Work package:	WP3 Diversity representation and processing
Document Title:	D3.1.3 - Reusable Implementations of Diversity-aware Interfaces
Editor (Name, Affiliation)	Patrick Wollner, Instrata
Work package Leader (Name, affiliation)	Andreas Thalhammer, UIBK

Executive Summary

We aim to provide an overview of visualisation techniques from a Human-Computer Interaction perspective relating to the diversity-aware tools delivered as part of the RENDER project. The overview comprises four distinct sections: a review of related literature, the development of guidelines, a discussion of selected case studies that illustrate examples of best practice, and the in-depth review of three tools that have been developed as part of RENDER, namely DiversiNews, Link ExtrActor (LEA) and the Drupal extension. The guidelines are focused on diverse datasets and support diversity in the capabilities and the intentions of tool users, as well as offering a diversity of information in terms of sentiment, opinion, and bias of searchable data. We introduce the concept of a visualisation pipeline that encompasses the data input, the visualisation schema and the consumers of the visualisation. The pipeline provides a novel approach to collating guidelines that span information visualisation, interacting with visualisations, and the interplay between the various components in the pipeline.

Table of Contents

Abbreviations

API	Application Programming Interface
HCI	Human-Computer Interaction
ICT	Information and Communications Technology
RDF	Resource Description Framework

Definitions

Schema	A representational framework in which information is gathered for a purpose, processed and presented in order to achieve a deeper understanding of it. Adapted from Russell et al. [10].

1 Introduction

The wealth of information emerging in our modern world necessitates new aggregation techniques and various types of visualisation to ensure that people can explore and make sense of the collected data. Specifically, RENDER aims to facilitate the improved search for and exploration of data that is diverse not only in terms of topic but also in terms of sentiment, bias, and opinion, and which should be navigable independently or in parallel to traditional result outputs. RENDER aims to provide a comprehensive conceptual framework and technological infrastructure for enabling, supporting, managing and exploiting information diversity in Web-based environments. Hence, the resulting tools may be broadly classified as visualisation tools for flexibly displaying diverse information. Our aim is to provide guidance towards achieving the optimal interaction with these tools in order to devise the most effective means by which different types of users can interact with the diversity of information.

This document introduces the concept of a visualisation pipeline, which relates input data, the visualisation process and users. We present key concepts from the literature to inform the development of guidelines, such as dependencies on levels of visualisation and the underlying cognitive processes. Many such concepts were developed at a time when interacting with a visualisation was mainly presentational, and as a result, we introduce interactivity as a concept and derive guidelines for designing for it. By further defining the visualisation pipeline and incorporating degrees of interactivity, we are able to determine further guidelines. We apply the overall set of guidelines to selected case studies in order to illustrate how the guidelines can be applied and to provide examples of best practice.

The final section of this document evaluates three Render tools using these guidelines. There are about ten main tools developed within the RENDER project, and several of these tools have multiple components. In addition, some of the tools share similar functionality, e.g. the Drupal extension and Enrycher. We have therefore decided to group the tools into three main categories, and provide an in-depth analysis of one tool from each of these categories. The guidelines that we proposed for the selected tools can be extended to other tools within the same category.

The three selected RENDER tools are: DiversiNews, Link ExtrActor (LEA) and the Drupal extension. For each of these tools we outline a set of recommendations for improving their interfaces.

2 Visualisation Pipeline

The increased breadth of data necessitates enhanced information visualisation techniques to ensure that diverse audiences are able to deal with the abundance of diverse content. While guidelines exist that reduce the barriers to accessing Information and Communications Technology (ICT), they tend to be generic and focussed on neither data visualisation nor diversity-awareness. Therefore, we require targeted, specific guidelines to ensure that information visualisation developments cater to a wide range of users while displaying equally diverse datasets. In this context, we introduce a pipeline, extended from Card et al. [2], as a way to consider the process of visualisation.

Because RENDER deals largely with search operations, it is important in the context of this deliverable to differentiate between exploratory and query-based search, since RENDER's tools support both kinds of search to varying degrees. Query-based search is commonly used to search for a specific piece of information or a well-defined information set. For example, Link ExtrActor (LEA) [1]is used for query-based search. The user searches for a specific article and receives a well-defined, predictable and consistent dataset in return.

In contrast, exploratory search deals with more complex datasets, such as those where the user does not have a predefined or specific question to answer (e.g. "I would like to know more about Henry VIII" instead of "Who was Henry VIII's third wife?") or searches that aim to compare or analyse information from different sources (e.g. "How has the media perception of global warming shifted over the past five years?"). This can present challenges in search optimisation. As explained by Wilson and Schraefel [3]:

> *"Exploratory search has been defined as more complex IR activities that may be required by users who do not have clearly defined goals, have changing complex needs or may be using a system that is poorly indexed. Marchionini suggested some of these possible activities, including: Aggregation, Comparison and Evaluation. In these conditions, simple keyword search may not support users effectively and convoluted user coping strategies have been recorded that involve iterative tentative guessing of keywords. Clearly, what these users require is an alternative method of search that may involve more browsing and exploring activities to achieve their goals."*

The idea of using interactive visualisations as an alternative method of search has already been explored in the literature (Dörk et al. [4], Lee et al. [5]), but left an incomplete picture of how such visualisations should be implemented. This report aims to provide guidelines that address this issue.

One of the additional challenges with exploratory search is that the dimensions of the results are not based on information that can be considered objective "facts" in a strict sense. In particular, a search will return information with a diversity of biases, opinions and sentiments, dimensions which should ideally be as easy to navigate as the topic itself. It is in this space that RENDER is operating and where diversity-enabled visualisations are most useful.

RENDER's purpose is to harness the "unprecedented success for facilitating the publication, use and exchange of information, at planetary scale, on virtually every topic, and representing an amazing diversity of opinions, viewpoints, mind sets and backgrounds" [6]. At the same time, "the Web is also confronted with fundamental challenges with respect to the purposeful access, processing and management of these sheer amounts of information, whilst remaining true to its principles, and leveraging the diversity inherently unfolding through world wide scale collaboration" [6]. In this context, RENDER aims to "leverage diversity as a crucial source of innovation and creativity, whilst providing enhanced support for feasibly managing data at very large scale, and for designing novel algorithms that reflect diversity in the ways information is selected, ranked, aggregated, presented and used" [6].

[1] http://toolserver.org/~RENDER/toolkit/LEA/

With this goal in mind, RENDER tools must combine the input data described above, process the data and visualise it to provide purposeful access to a diverse user population. We introduce a framework that captures these three stages as “Data Input”, “Visualisation Schema” and “Visualisation Consumption”, as outlined in Figure 1. Using this holistic model, we can establish guidelines that are generic enough to be applicable to all of RENDER’s and all likewise constructed diversity-aware tools, and specific enough to provide valuable input in their analysis and improvement.

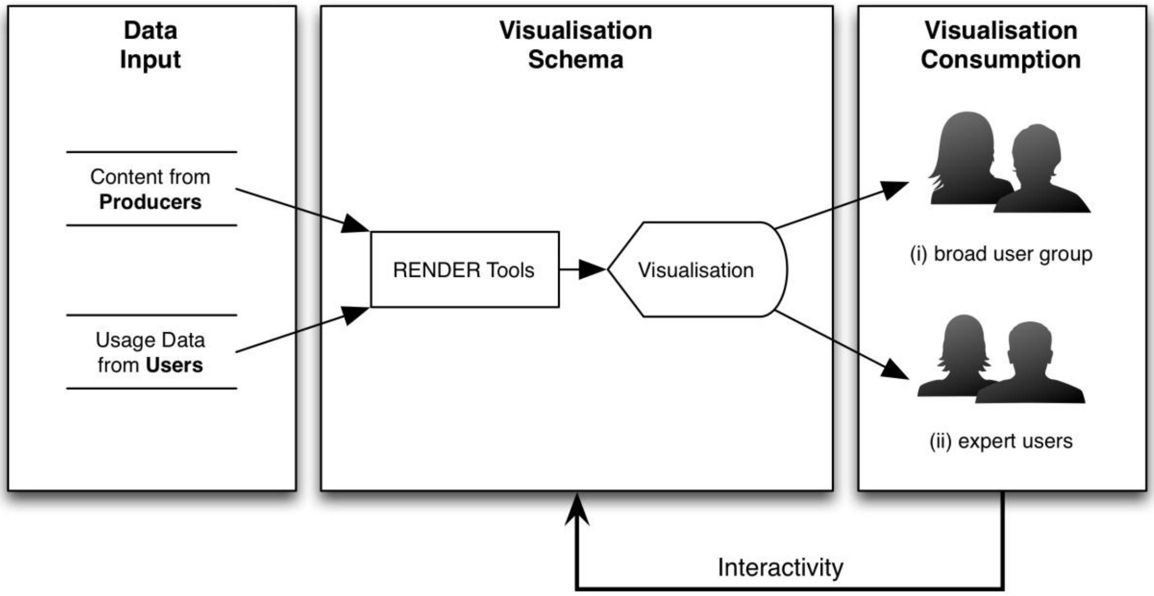


Figure 1: A representation of varied datasets being visualised for diverse audiences

In order to better understand the idea of a visualisation pipeline, let us take the example of a modern map tool – one where users can specify a location and be given information about that location, whether it is how to get there, the current weather, reviews of restaurants in the area, etc. From the initial user input (search query) to the final map, there are several stages that the map programme must go through to produce a useful visualisation. The first step is to use the search query to reduce down from the set of all available data (i.e. all location data for the entire world) to just the relevant data (i.e. location data for the requested area). The set of all available data is called the Data Input, and the reduced data set forms the Data Table for the visualisation.

The next step is to figure out what the visualisation will actually look like, as an actual table of data is unlikely to be useful in most circumstances. There are a variety of options for most applications. For example, on a map, a query about restaurants in a certain area could result in a list of restaurants or in an annotated map showing their physical location. Different applications will require different kinds of visualisations, and one of the most important design steps is deciding the kind of visualisation that will best display the data in a useful way. The kind of visualisation is called the Visualisation Schema.

Once the visualisation is created, it is then given to users, and their interaction with it is called Visualisation Consumption. This can be either passive or active; some visualisations do not change with user input, but some do, and changing with user input is called Interactivity, which will be covered in more detail in Section 5.

Below is a list of terms defined in this section:

Data Input. The data in RENDER may originate from *publishers or users producing content*, such as news articles and comments on those or encyclopaedia entries, or from *users producing usage data*, such as page views on news sites or encyclopaedia entries, as identified by Fortuna et al. [7].

Data Table. The subsets of data input relevant to the specific schema. This may be regarded as an implicit selection of the information that will be displayed in the respective schema.

Visualisation Schema. Based on the requirements of RENDER, we identify a number of visualisation schemas, which define how the data is best displayed and which provide the basis for establishing guidelines for visualisation tools, such as those developed as part of RENDER. The next section (3) focuses on the underlying mechanisms and requirements of visualisation schemas.

Visualisation Consumption. There are two different populations for which RENDER's tools cater. This necessitates two separate design categories: (i) enhanced design for a broad user group aiming to satisfy the cognitive demands and capabilities that make information visualisation a powerful tool to a wide audience and (ii) enhancements that cater specifically for an intended user group, such as an integration of the mental model and prior experience of expert users (e.g. for the RENDER Telefonica opinion mining tool).

Interactivity. Feedback from visualisation consumption to schema may be defined as interactivity. This outlines the degree to which properties of the visualisation process can be changed by the viewer or user.

3 Visualisation Schemas

As outlined above, schemas constitute the basis for the presentation aspects of the tools that are being developed in RENDER – a process in which datasets are processed and visualised for a group of users. This section draws on the literature and particularly the area of crossover between the fields of information visualisation and Human-Computer Interaction (HCI), which are two distinct areas of research and development to establish a better understanding of visualisation tools in this context. Whilst there is much discussion in the literature around representational structures such as bar charts and graphs, we have selected two particular concepts, levels of visualisation and the underlying cognitive processes, in order to help us to determine a set of guidelines that can be applied across the RENDER toolset while being specific enough to be useful.

3.1 Levels of Visualisation

Different levels of visualisation exist and represent an important element in analysing visualisation schemas. These levels can be segmented by the intention of use and the potential mental models the user may already have as outlined by Popovic [8]. From an HCI perspective, intended use presents a key design challenge while also presenting an opportunity to improve the visualisation of the data. The intended use is often ambiguous, but, once determined, lays the foundation for designers for the development of the visualisation tool. The primary levels of visualisation are broken down by Card [9] into four distinct levels:

I. **Visualisation of information spheres** refers to information that exists in repositories beyond the user's workspace (such as in the cloud or stored in web services). It is task-independent and focused on the delivery of general information.

II. **Visualisation of an information workspace** outlines the active representations of information sources required to perform a specific task. This implies the adopted representation embodies the information required for the task and consequently reduces the time spent performing a specific task.

III. **Visual knowledge tools** rearrange the information in the workspace to reveal patterns or allow manipulation in the search for patterns. These tools are usually interactive representations (such as looking for hotels in a particular location).

IV. **Visually enhanced objects** represent data packed into virtual physical objects (such as a model of the earth or a representation of a biological system), in which data is stored and can be recalled by the user.

If the designer classifies the intended use of a specific tool into one or more of the above categories, the resulting visual representation can be optimised to satisfy the tool's intended use.

This concept is further enforced by the relevant work by Popovic [8], which examines the impact that knowledge, in particular domain-specific knowledge, has "in distinguishing […] user[s] and the way in which they use technologically interactive devices". The resulting models are differentiated by the degree to which the user is familiar with the technology from prior experience. In all models (and hence all degrees of prior experience) the intended use plays a key role in ensuring that the user can interact with a given system. In the context of this project, the level of visualisation may be seen as an implicit form of intended use and as such, this becomes our first guideline.

Guideline

1. Intended use should be easily identifiable by ensuring that the level of visualisation falls clearly into one of the four specified categories.

3.2 Underlying Cognitive Processes

A second essential concept that assists designers in schema selection (beyond comprehending the levels of visualisation) is an understanding of human cognition in order to predict how users will mentally process different kinds of visualisations. Once it is evident that visualised information simply acts as a cognitive aid, the ideal underlying cognitive process may be defined and – to the best of the designer's ability – be replicated in an HCI context. As explained by Card et al.: "Without external aids, memory, thought, and reasoning are all constrained. But human intelligence is highly flexible and adaptable, superb at inventing procedures and objects that overcome its own limits. The real powers come from devising external aids [such as visualisation tools]: it is things that make us smart" [2].

Card et al. [2] summarise cognition in the context of visualisation to be "the acquisition or use of knowledge". This maps onto three key goals of (i) discovery, (ii) decision making, and (ii) explanation in terms of the cognitive activities that may be aided through improvements in the visualisation technique or paradigm employed. By taking this intention as the key motivator in selecting a visualisation technique, it allows highly diverse content to be outlined in generally applicable guidelines.

More specifically, Card et al. [2] refine the task of visualisation to be one that is based on a schema that follows the path of what a user or viewer of visualised information may have attempted to accomplish with the same input data solely through their cognitive abilities. We have adapted the schema originally articulated by Card et al. [2], shown in Figure 2, by removing the complex dependencies between its components in order to simplify it and relate it to the visualisation pipeline, whilst still recognizing these interdependencies exist. A description of the components of Figure 2 is given below.

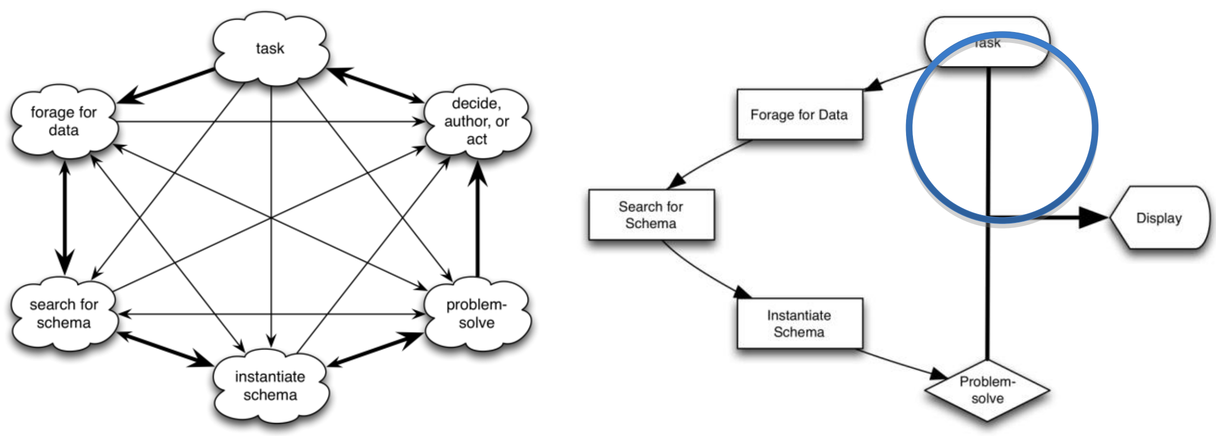


Figure 2a (left): The process of knowledge crystallisation, taken from Card et al. [2] and **Figure 2b (right):** The interpretation of this process applied to the interactive visualisation tools introduced as part of RENDER

Given a specific task, the user will search for relevant data. Once this data is located, the user attempts to construct a representational mental framework [10]. This schema has the purpose of accomplishing the initial task with the selected input data. The final action of the user is to reprocess the information with the schema before assuming to be able to solve the initial problem addressed by the task. In summary, such a decision-making flow relates to the selection of a specific visualisation schema.

Card et al. [2] argues that this cognitive process (that most humans undergo) should also be the underlying framework for information visualisation tasks. Therefore, information should be presented in such a way that it exploits pathways that already exist in human cognition. This implies that projects like RENDER should aim to support the cognitive process of its users. Furthermore, it is a form of applied reasoning similar to the process that humans complete on a daily level.

Card et al. [2] further introduces a process for knowledge crystallisation, which outlines a potential model of how the human mind acquires, processes and uses data with the intent of understanding the underlying trends, hidden information or simply in the attempt of finding solutions to problems by reprocessing the available data. This cognitive process is depicted in Figure 2a. The model in Figure 2b is our interpretation of Card et al.'s model applied to the visualisation tasks introduced in RENDER's tools. We use this as the basis of defining effective information visualisation methods of diverse data and hence as the basis of further guidelines in this process.

The process contains the following four key elements:

Forage for Data. In terms of finding information, visualisation tools may support the discovery of datasets previously not known to the user (e.g. serendipitous discovery, as outlined by Beale [11]) or give the user the right tools should he or she know the specific data that is required for completing the desired task.

Search for Schema. Once this information is selected, the visualisation tool may help in searching for a particular representation that best allows the reprocessing of the available data and hence aids the cognitive process as outlined above.

Instantiate Schema. The tool supports the user – and this is the key element of visualisation tasks – in representing the dataset in a format that is aligned with a selected schema.

Problem-solve. Depending on the nature of the visualisation at hand, the tool may also "problem solve to trade off features" [2], such as showing a particular subsection of a map or selecting a variable by which the contents of a list should be reordered.

The approach of using applied cognition as a basis for defining how to best visualise information in an HCI context is based upon a wealth of previous research. Larkin and Simon [12] conducted a study, which outlines the differences in users solving natural science exercises with and without visual aids. These aids, integrated into the framework outlined in Figure 2, allow for a metric of the effectiveness of the proposed methodology. Their experimental work focussed especially on establishing methods that captured search, recognition, and inference, and hence allowed an objective interpretation of their participants' performance with and without diagrammatic aids in solving the problems. Larkin and Simon's conclusions are highly applicable to today's challenges in visualisation paradigms: (i) search can be reduced by grouping sets of relevant information, (ii) complexity can be reduced by using the (geo)graphic location associated with a particular subset of information for grouping purposes, and (iii) certain principles are immediately recognisable to the participants that would – without a diagrammatic representation – require an explanation.

Based on the conclusions of this experiment, Card et al. extract "six major ways in which visualisations can amplify cognition":

1. By increasing the memory and processing resources available to the users
2. By reducing the necessity to search for information
3. By using visual representations to enhance the detection of patterns
4. By enabling perceptual inference operations (e.g. filling gaps by using non-verbal sensory information such as colours or sounds)
5. By using perceptual attention mechanisms for monitoring (using the subject's ability to perceive or ignore stimuli, both task-related and non task-related)
6. By encoding information in an interactive medium.

The first of these depends on the context of use and the task the user is performing, and cannot be transformed into a design guideline. However, we are able to transform the remaining five into a set of the guidelines we introduce for selecting and implementing a visualisation schema.

Guidelines

2. Reduce the necessity to search for information.
3. Use visual representations to enhance the detection of patterns.
4. Enable perceptual inference operations.
5. Use appropriate visual cues to draw attention to relevant data representations.
6. Encode information in a form in which it can be manipulated.

4 Interactivity

The previous section explores how to determine a schema in the context of the input data and the specific requirements of the viewers of the visualisation. In this section, we consider the degree of interactivity that the resulting visualisation tool supports. The concept of interactivity is commonly overlooked by the traditional information visualisation literature, probably because interaction paradigms are changing rapidly with the emergence of new technological capabilities, such as the recent support for touch and gesture interaction.

We define interactivity in this context as the degree to which one or more variables in the visualisation can be manipulated in order to modify the way in which the output is displayed. One extreme might be a static representation, which has the same viewer impact on and off screen. The other extreme allows full interaction and intervention of all or nearly all variables of the displayed dataset.

It is important to note that even in contemporary visualisation tasks, the range of interactivity that a particular visualisation paradigm may allow varies greatly. Some of RENDER's tools display a nearly static representation of a specific dataset that enhances the information flow to the user but does not allow for interaction or manipulation of any of the variables of how the dataset is displayed. Other visualisation techniques may allow full manipulation of the displayed data, including a selection of input datasets and their comparison. Similarly, changes in views and dynamic queries are common features of RENDER's tools. Manipulation of the data while being displayed and the linkage to related information are similar examples in RENDER's context.

Both the diversity of the data input and the consumers' ability impacts on the degree of interactivity that is feasible. The skills and mental models of a target user group directly impacts the degree and complexity of interactivity required to ensure the best possible visual representation of a particular dataset. Likewise, the nature or content of the data displayed influences the extent to which the representation can be interactive, i.e. how many variables can be modified.

In the design of an interactive visualisation, we define two novel guidelines that focus on a balance of interactivity. It is important to maintain a good understanding of the degree of interactivity that is:

1. Achievable by the data input and
2. Feasible for the viewer(s) of the visualisation.

Hence, the designer should maximise the potential of (1) without breaching the range of (2). This will ensure the most immersive experience without limiting the potential experience of the user.

Guideline

7. Maximise the degree of interactivity based on the limitations of the available data whilst ensuring that the assumed capabilities of the user support the chosen degree of interactivity.

5 Visualisation Pipeline Revisited

This section refines the definition of the Visualisation Pipeline introduced earlier by breaking the visualisation schema into three further elements and developing the concept of interactivity introduced in the previous section. As before, we segment the visualisation process into three main components: the data input, the visualisation schema and the visualisation consumers or consumption. The model introduced in Figure 4 (below) is an extension of the foundation outlined by Card et al. [2] in an attempt to create a unified approach to interpreting the pipeline of visualisation tasks.

The visualisation schema comprises three elements, as determined by Card et al.:

Data Tables contain the subsets of data input relevant to the specific schema. This may be regarded as an implicit selection of the information that will be displayed in the respective schema. This concept was also covered in Section 2.

Visual Structures relate to the levels of visualisation introduced earlier. Here, the level of visualisation is selected and implemented. It also defines whether the resulting representation will be in the form of a graph, table or annotated text, or some other representation.

Views are a range of different ways in which the specified visual structure may be displayed. In the simplest case, this could be different locations on a map-based representation or various orderings of a displayed table.

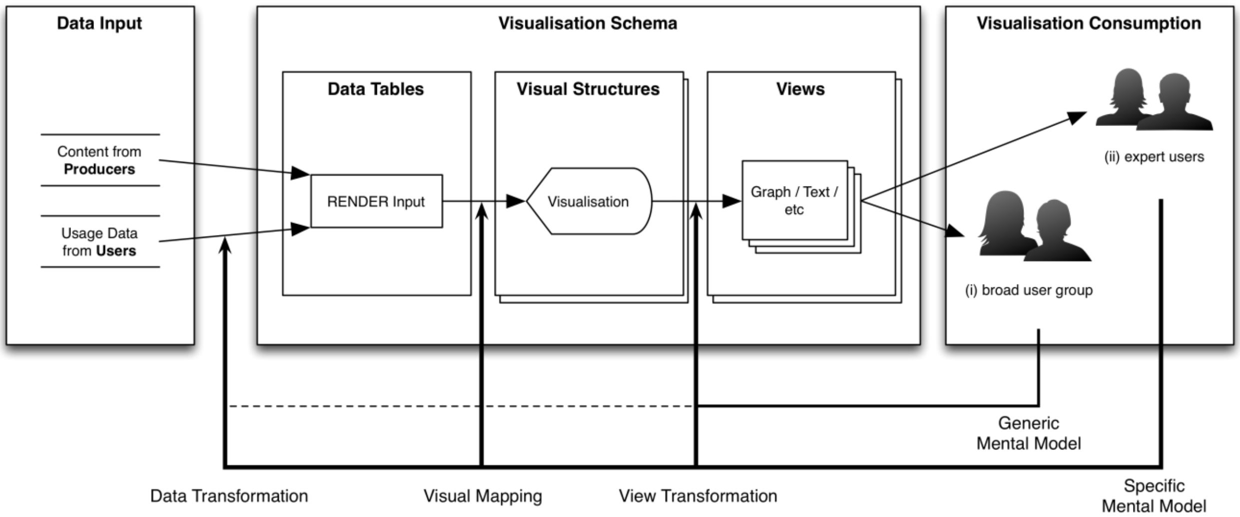


Figure 4: Reference model for visualisation, adapted from Card et al. and applied to the requirements set by RENDER.

Let us briefly come back to the map example from Section 2. Previously, we explained that there were a variety of steps in creating a visualisation: reducing down the available data to create a data table, choosing the kind of visualisation, and then using the data table to make the visualisation. However, there are varying levels of user input at each stage of the process. Some applications do not require any user input at all, for example general infographics on a webpage. Other applications require user input at the point of search; in other words, a user inputs into the process of going from input data to data table, usually by entering some sort of search query. In the map example, this would be the user specifying a desired location or activity. Once the data table has been formed, it is also possible in some applications for users to determine the kind of visualisation they want. For a map, this might be giving users the option to display information as an annotated map or in a list. Finally, there is an optional level of interactivity with the visualisation itself. Maps often allow users to interact with them directly, for instance via zooming in or out of the map or moving it around the screen, among other possibilities.

Visualisations are often dynamic in that user input at the various levels can alter how the desired data is displayed, or even which data is desired. The feedback process is a critical part of the user interaction. For example, a user having seen map data for restaurants in a certain area may decide to search for restaurants in a new area (requesting a new **Data Table**), select for restaurants to be listed instead of being presented on an annotated map (requesting a different **Visual Structure**), or zoom in on the map to see only restaurants within one mile of their current location (changing the **View**). A good design will not only provide useful data, visual structures, and views, but also will be designed for the optimal level of user feedback. For instance, allowing users to change the Visual Structure might not always be a good idea, even if designers want to let users manipulate the View.

In Figure 4, we have enhanced the feedback loop, as indicated by the black lines, showing how visualisations can allow consumers to interact with the respective boundaries between elements of the visualisation schema, and thus we are able to determine four degrees of interactivity. In the case of an expert user with a specific mental model, the viewer may have full control over the desired *data transformation*, which takes place in the selection of the subset of data to be displayed, the *visual mapping*, which selects the type of visualisation (i.e. a map or list view of the same data) and *view transformations* which allow the manipulation of the specific visual structure chosen. Broad user groups with a generic mental model (such as non-expert users) will have lower levels of feedback flexibility commonly, as indicated by the relative strength of the line [8].

Excellence in graphical output of information visualisation may be summarised by Tufte's simple rule of "clarity, precision, and efficiency" [13]. From a guideline perspective, this means that user interfaces that aim to communicate diverse information to a diverse population should "induce the viewer to think about the substance, present many numbers in a small space, make large datasets coherent, encourage the eye to compare different pieces of data, reveal the data at several levels of detail, from a broad overview to the fine structure" [2].

It is this compromise that determines our third set of guidelines, introduced by the interplay of the elements outlined in the above:

- Data tables should be selected to accommodate the diversity of the data input without creating unnecessary complexity in the visualisation schema. This process is called **Data Transformation**.
- Visual structures should be selected to best display the relevant data tables without having to filter data but enabling flexibility in the number of related views. This process is called **Visual Mapping**.
- The number of views should be limited to what is feasible by the interaction paradigm (e.g. map vs. table) and in full awareness of the resultant complexity created for the user. This is known as **View Transformation**.
- Feedback to data transformation, visual mapping and view transformations should be balanced to establish the maximum possible functionality without creating complexity in the user experience based on the consumers' ability.

Guidelines

8. Create data tables that accommodate the diversity of the data input without creating unnecessary complexity in the visualisation scheme.
9. Select visual structures that best display the relevant data tables without having to filter data.
10. Limit the number of views to what is feasible in the selected interaction paradigm (e.g. map vs. table) whilst being aware of the resultant complexity created for the user.

6 Application to RENDER

In analysing the various tools that have been developed as part of the RENDER project, we must find a balance between generic guidelines, which apply to a wide range of information visualisation applications, and specific guidelines, which can be used in the development or analysis of a particular type of visualisation.

RENDER's tools fall into one or more of the following visual structures as determined by the representation of the data that is output: (i) a (geo)graphic representation of a dataset, (ii) a text or query-based representation of a dataset, or (iii) the original representation of a dataset with annotations. This segmentation allows us to take a portfolio of tools that illustrate best practice in relation to each of the visualisation types.

We could also use a more finely granulated segmentation, such as one that categorises either the type of content or the type of usage. For example, Fortuna et al. [7] suggest that data that is generated by analysis of content may be categorised into aspects such as topic, opinion, or cross-lingual aspects. Likewise, data that is generated by the analysis of usage (or via usage statistics) may be categorised into aspects such as demographic, access method, or time. Fortuna et al.'s analysis clearly outlines the breadth of diversity-aware visualisations and we could employ such a detailed categorisation for our purposes. However, we suggest that guidelines based on such a detailed segmentation would be too specific and duplicative.

Table 1 indicates the RENDER tools segmented by the representation of the data that is output. Due to the functionality of some of the tools, they fall into more than one of the visual structures. We propose that this segmentation not only captures all of the tools but also allows for the best mapping to existing online tools, which provide valuable examples of best practice for the particular visualisation paradigms employed by RENDER's tools.

Table 1: Overview of visual structures employed by RENDER's established tools

(Geo)graphic Representation	**Text or Query-based Representation**	**Annotation-based Representation**
Wikipedia Map Corpex LEA – Link ExtrActor Change Detector WIKIGINI Opinion Mining Tool Newsfeed TwiDiViz DiversiNews (incl. TopicSum Summarizer)	Task List Generator (TLG) Enrycher Interactive Modeling Tool DiversiNews (incl. TopicSum Summarizer) TwiDiViz Drupal extension	WIKIGINI Article Statistics and Quality Monitor (ASQM) Enrycher

7 Conceptual Examples of existing diversity-aware visualizations

For each of the categories outlined in Table 1, we have identified a selection of third-party tools in order to show how the guidelines can be applied and to illustrate best practice.

7.1 (Geo)graphic Representation

We compare a selection of applications that display information on a map and consequently allow the manipulation of information based on location. The three case studies range from a generic package with no specific application to a specific application that only geographically represents one type of information on a map. This progression of specificity makes the selection of case studies especially relevant to RENDER's diversity in visualisation.

CartoDB[2] is a platform for presenting map-based data on desktops and mobile applications. As an open platform with a large number of APIs, this tool adopts an approach that does not constrain it to specific tasks. The degrees of interactivity are especially high, enabling a range of APIs to be utilised that allow a multitude of visual styles to be matched with a view landscape that is particularly open-ended.

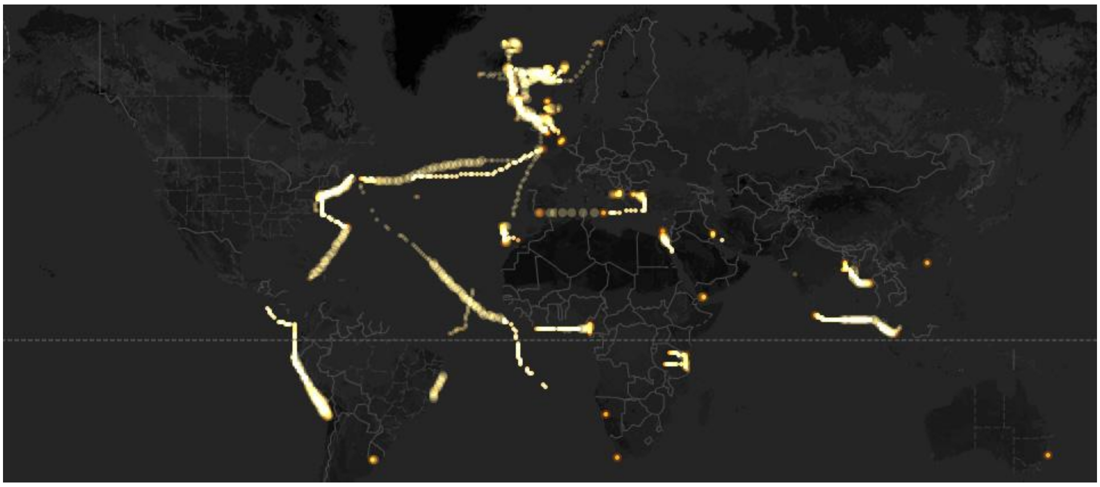


Figure 1. A screenshot of CartoDB map-based representation.

GeoCommons[3] aims, unlike CartoDB, to generate both a platform for presenting map-based data as well as aggregating the content to be displayed onto these maps. Similarly to CartoDB, all output is focused on map-based data from heterogeneous repositories. GeoCommons focuses on both the aggregation and presentation of data and hence allows the user a lower degree of data transformation, but maintains a high degree of flexibility in visual mapping and view transformation.

[2] http://cartodb.com/
[3] http://geocommons.com/

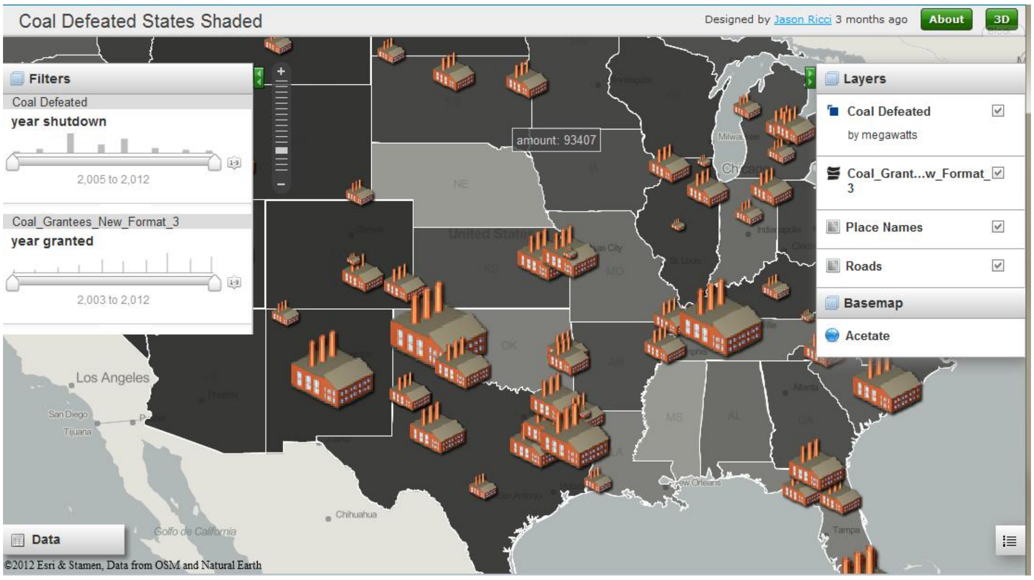


Figure 2. A screenshot of GeoCommons map-based representation.

Global Twitter Heartbeat[4] aims to extract information from the density of specific twitter feeds (through searches and hash tags) based on geographic location. Similar to the previous two examples, the output is then represented graphically on a map. This tool offers the lowest degree of interactivity; it does not offer access to visual mappings or view transformations. It does support data transformation, i.e. how the data appears on the map.

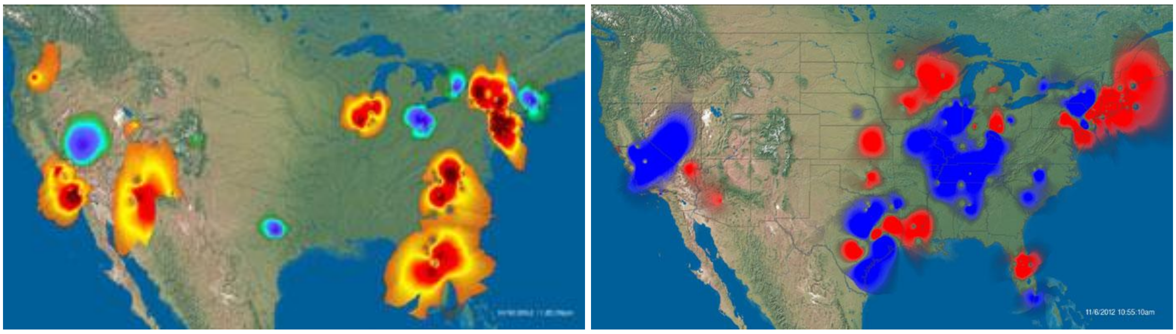


Figure 3. A screenshot of Global Twitter Heartbeat map-based representation.

These examples of third party geographic representations outline the progression of available degrees of interaction in this domain and so allow us to compare the effects of interactivity on map-based visual styles. For example, by limiting the feedback to view transformations and visual mapping, the interface becomes especially simple and can still offer a large array of data transformations.

7.2 Text or Query-based Representation

Visualisation-relevant case studies are more difficult to select for text-based information due to the diversity of the data that RENDER's tools manipulate. We introduce two examples to highlight best practice in existing implementations.

[4] http://www.sgi.com/go/twitter/

Social Mention[5] presents information from a range of online social media sources in a query-based form. By aggregating user-generated content from across the web into a single stream of information, it provides a measurement of what people are saying online about specific topics in real-time.

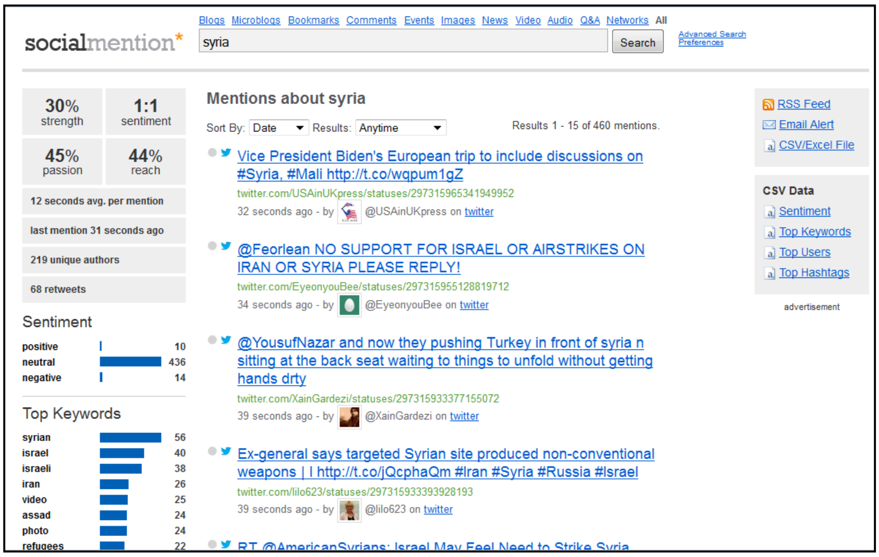


Figure 4. Screenshot of the Social Mention tool.

Google Activity Report[6], which was launched recently, allows users to see a primarily text-based representation of usage and access data of their Google account (shown in Figure 6). The diversity of the presented information makes this tool especially relevant to a comparison with RENDER's output.

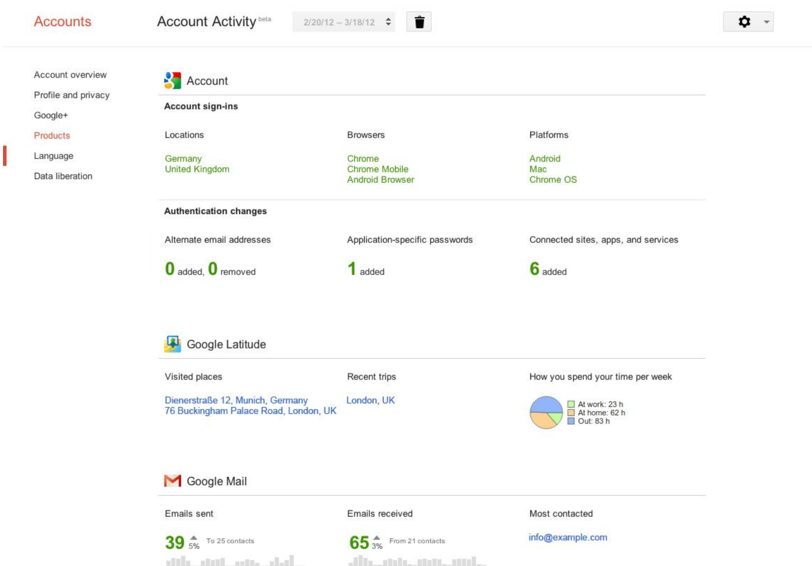


Figure 5. Screenshot of the Google Activity Report.

[5] http://www.socialmention.com/
[6] https://www.google.com/settings/activity/

7.3 Annotation Representation

We have selected two tool examples that represent annotation. The source representation of data (e.g. a document or Wikipedia article) is displayed in its original format and augmented with annotations that provide additional information to the user.

Turnitin[7] allows submitted work to be analysed for potential sources of plagiarism. A large set of information (websites, books, papers) is compared to sections of the uploaded document; a copy of the document with annotations indicating potential examples of plagiarism is displayed to the expert user. Due to the wide range of source data (electronic and analogue), this case study underlines the relevance of diversity in analysing information.

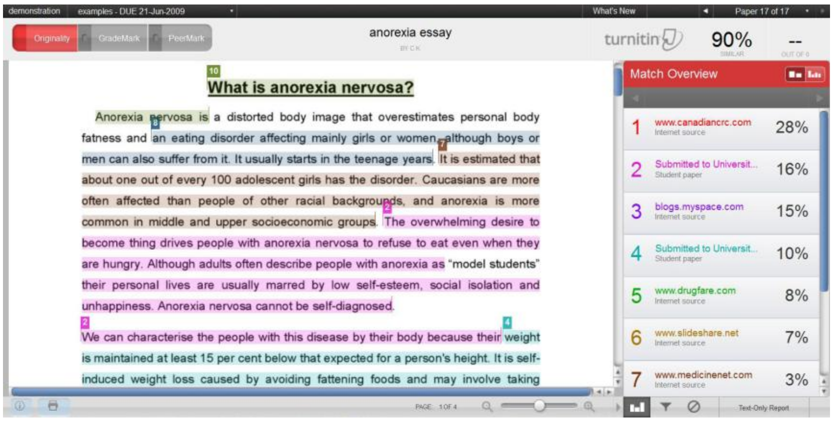


Figure 6. Screenshot of the Turnitin tool.

Linguee.de[8] is a translation tool that relates websites of similar content (i.e. websites that are assumed to have the same content translated from one language to another) and highlights the search query of specific words when translating from one language to another. This allows the user to base translations on entire sentences rather than individual words. The output of each query is depicted in the form of annotated representations of both language variants of the requested word or string of words. Due to the multiple languages utilised both in the underlying data and visual representation of this tool, it is especially relevant from a diversity perspective.

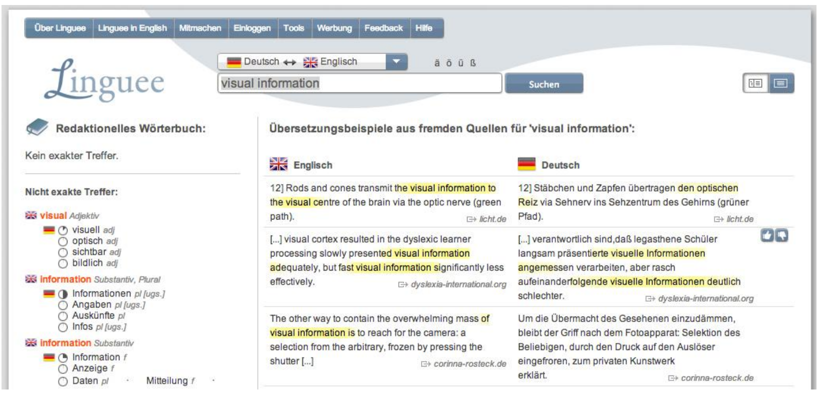


Figure 7. Screenshot of the Linguee tool.

[7] http://www.turnitin.com/
[8] www.linguee.de

7.4 Application of the Guidelines

We use the some of the above examples to exemplify each of the guidelines and to indicate best practice.

Guideline 1. Intended use should be easily identifiable by ensuring that the level of visualisation falls clearly into one of the four specified categories.

By specifying the Level of Visualisation of Turnitin as a Visually Enhanced Object, the designers of the tool implicitly optimised it for its intended use. Likewise, by focussing GeoCommons to act as a Visual Knowledge tool, the designers implicitly optimised its intended use for finding patterns rather than using a map as a Visually Enhanced Object.

Guideline 2. Reduce the necessity to search for information.

All visualisations outlined in the above examples aim to reduce the necessity to search for information. Map-based output by Global Twitter Heartbeat underlines the absence of formulating queries for identifying particular subsets of information; instead, a graphical representation utilises a model known to the user (a map) as the basis for reducing the search for information.

Guideline 3. Use visual representations to enhance the detection of patterns.

The annotation of pairs of related words in two languages by Linguee.de is representative of a visual representation. Simply using two strings of words next to each other diminishes the cognitive effort involved in detecting patterns.

Guideline 4. Enable perceptual inference operations.

Using the annotation feature of Turnitin, which makes use of colours and highlights potentially plagiarised text in documents, allows the user to quickly infer the degree to which a document could be breaching copyright regulations without the user having to read a report. The colour and placement of the annotations alone are sufficient to enable perceptual inference.

Guideline 5. Use appropriate visual cues to draw attention to relevant data representations.

By including additional graphical information to a particular query at the side of the results box, Social Mention utilises perceptual attention mechanisms to augment the visualisation experience.

Guideline 6. Encode information in a form in which it can be manipulated.

GeoCommons makes use of the standardised manipulation (or interaction) techniques that the user experiences in other popular tools, such as Google Maps, but also encodes information repositories in this representation. Similarly, Turnitin encodes information in a page-based document structure that can be manipulated.

Guideline 7. Maximise the degree of interactivity based on the limitations of the available data whilst ensuring that the assumed capabilities of the user support the chosen degree of interactivity.

CartoDB allows the areas displayed on a map to be limited to the regions for which data is available. This maximises the region that can be explored, whilst restricting it to the limitations of the dataset. The limited interactivity of Global Twitter Heartbeat ensures that the selected views are approachable by a wide user group. Yet, at the same time, it limits the functionality of the service.

Guideline 8. Create data tables that accommodate the diversity of the data input without creating unnecessary complexity in the visualisation scheme.

By creating a predefined sub-selection of relevant data tables, such as in the case of Global Twitter Heartbeat, the user is presented with predefined static images and the interactivity is limited.

Guideline 9. Select visual structures that best display the relevant data tables without having to filter data.

Given a selection of a dataset, such as the document to be analysed by Turnitin, the available visual structure (such as an annotated variant of the original document) best represents the basis for displaying *the intended visualisation. Colours and the annotated representation limit the need for filtering data (such as the type of plagiarism annotated).*

Guideline 10. Limit the number of views to what is feasible in the selected interaction paradigm (e.g. map vs. table) whilst being aware of the resultant complexity created for the user.

Linguee.de does not allow re-representation or broadening of the selected query in the view (the annotated text). Instead, a new query needs to be placed. This puts the interactive emphasis on the data transformation stage instead of the view transformation stage and ensures that no additional complexity is generated for the user.

GeoCommons provides an elegant approach to the balance of feedback cycles in their interactivity model. The data transformation sets the stage of the displayed visualisation and allows the user to express the intention of the map he or she wants to view. The visual mapping is not accessible to the viewer (i.e. the user can only select a map-based visual structure). This ensures that the user interacts with the most effective representation of the dataset displayed. Once opened, layers and filters allow interactive view transformations that enable great flexibility.

7.5 Hybrid Approaches

It is important to note that some third-party tools, like some RENDER tools, display more than one visual structure at a time. This adds to the complexity and impacts the user experience. The resulting trade-off between visual complexity and functionality is addressed by Tovée's Gestalt principles [14] and can be considered as further guidelines.

The principles are a simplified summary of how certain graphic design decisions may influence the perception of the presented information. Given the intended use of the analysed tool, these principles can be used to advantage. If the principles are ignored, the resulting design may distract or confuse the user.

- **Pragnanz**. Every stimulus pattern is seen in such a way that the resulting structure is as simple as possible.
- **Proximity.** The tendency of objects near one another to be grouped together into a perceptual unit.
- **Similarity.** If several stimuli are presented together, there is a tendency to see the form in such a way that the similar items are grouped together.
- **Closure.** The tendency to unite contours that are very close to each other.
- **Good continuation.** Neighbouring elements are grouped together when they are potentially connected by straight or smoothly curving lines.
- **Common fate.** Elements that are moving in the same direction seem to be grouped together.
- **Familiarity.** Elements are more likely to form groups if the groups appear familiar or meaningful.

These are general guidelines that are particularly relevant to visualisation tools when more than one visual structure is displayed.

8 Summary

The purpose of this document up to now has been to determine a set of guidelines for diversity-aware visualisations that meet an appropriate level of granularity that can be usefully applied to evaluate RENDER tools. We build on work in the literature that crosses over between the fields of Human-Computer Interaction and Information Visualisation.

Firstly, we introduce the concept of a visualisation pipeline (see Figure 5) as the flow of information from data repositories (i.e. diverse data input), through a visualisation schema, to the display of information, and onto a diverse audience, ranging from expert users to generic users.

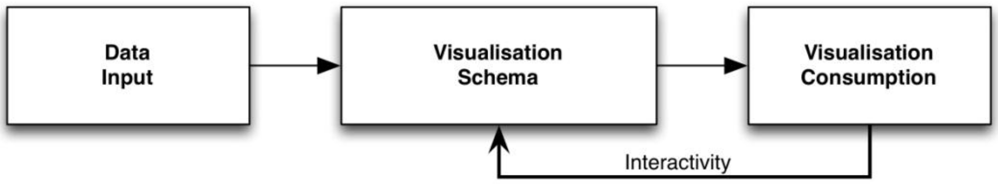


Figure 5: Simplified representation of the visualisation pipeline

Second, we present two concepts from the literature to inform the development of guidelines, levels of visualisation and the identification of the cognitive processes. Further refinement of the visualisation schema component in the pipeline identifies three elements: data tables, visual structures and views. We introduce the concept of degrees of interactivity and apply this to the pipeline to develop guidelines that create balance between user interaction and functionality.

Finally, we consider the RENDER tools and identify three categories of visual structure: (geo)graphic representations, text-based representations and annotation-based representations. Some of the tools use multiple visual structures. We select seven third-party tools that fit these categories and can act as examples of best practice for each guideline.

RENDER aims to develop solutions that increase diversity in content and consumption by drawing on improvements in functionality and usability of information visualisations. The guidelines introduced in this report aim to support this endeavour.

8.1 Diversity-aware Visualisation Guidelines

The following guidelines are the complete set we have determined based on recommendations for selecting and implementing visual schemas, defining the degree of interactivity, and balancing the overall implementation of the tool.

Guidelines

1. Intended use should be easily identifiable by ensuring that the level of visualisation falls clearly into one of the four specified categories.
2. Reduce the necessity to search for information.
3. Use visual representations to enhance the detection of patterns.
4. Enable perceptual inference operations.
5. Use appropriate visual cues to draw attention to relevant data representations.

6. Encode information in a form in which it can be manipulated.
7. Maximise the degree of interactivity based on the limitations of the available data whilst ensuring that the assumed capabilities of the user support the chosen degree of interactivity.
8. Create data tables that accommodate the diversity of the data input without creating unnecessary complexity in the visualisation scheme.
9. Select visual structures that best display the relevant data tables without having to filter data.
10. Limit the number of views to what is feasible in the selected interaction paradigm (e.g. map vs. table) whilst being aware of the resultant complexity created for the user.

9 Review of RENDER's Tools

We selected three RENDER tools based on visualisation structure, their development, and their availability and analysed them employing the guidelines introduced in this report. Further details of the review of the tools can be found in Appendices A, B and C. It is important to note that the tools were analysed together when considering the data tables but broken down into subcomponents when analysing visual structures and views.
Two of the authors carried out the analysis. The method of analysis involved both authors analysing a tool independently, and then the authors discussing their critiques and reaching a consensus on their evaluation. This method of analysis is considered to be more robust than one expert's evaluation.

9.1 DiversiNews

DiversiNews is an innovative and creative news visualisation service. In general, while the ideas behind the data visualisations are solid and in compliance with the design guidelines, the implementation of many of the features, in particular the topic map, need to be improved to be fully compliant. In order to bring the site into full compliance with the guidelines, it is not necessary to do a complete redesign, but rather to fix a few of the individual visualisations to make them either easier to understand, easy to manipulate in a meaningful way, or both. However, the map visualisation is a good example of a feature that is already in compliance with the guidelines. There are several key modifications that would improve DiversiNews:

(1) Move data visualisations to the left-hand side of the page to increase user awareness of the cause and effect mechanism between visualisation manipulation and summary/article updates.

(2) Clarify the purpose of the topic visualisation. Users should be able to draw meaning from how the topic map is laid out and how moving the red dot will affect search results.

(3) Add additional information to the map, such as making location markers bigger to indicate more news output from those locations.

(4) Change the slider so that its default location indicates the general feel of the news around a certain topic to the user (e.g. mostly positive).

(5) Introduce a "back" and/or "reset" button such that users can recreate past searches based on their modifications.

9.1.1 DiversiNews Evaluation

This appendix aims to evaluate the extent to which DiversiNews conforms to the guidelines defined in RENDER Deliverable D3.1.3. DiversiNews is a news aggregation service that allows users to dynamically visually adjust a news search across the dimensions of topic, location, and sentiment. For example, users could search for news on an upcoming smartphone release that was specifically about the operating system, generally positive, and published in the USA, all without having to type any new search terms. Figure 6 shows a screenshot of a DiversiNews search.

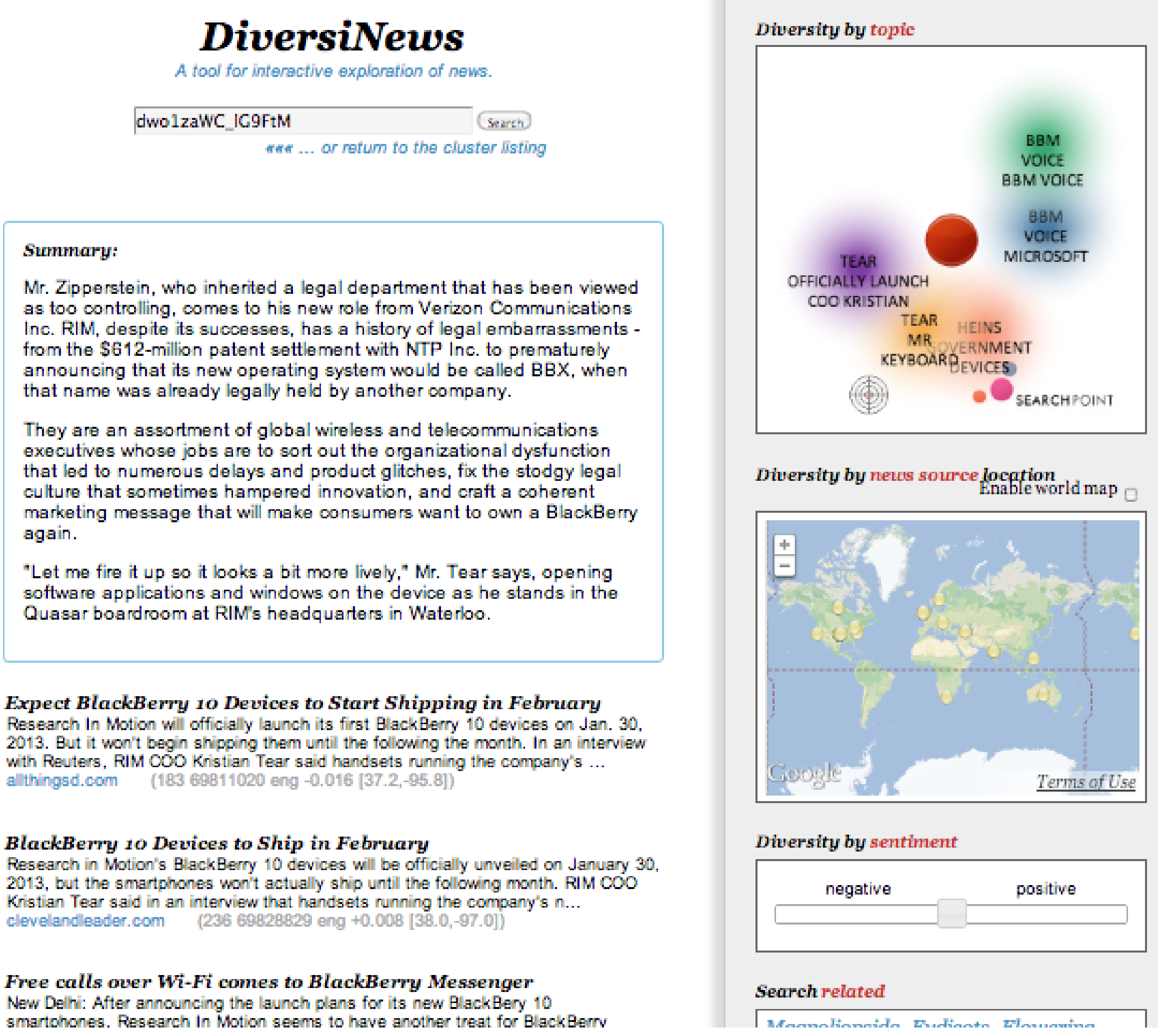


Figure 6: A DiversiNews search, showing topic, location, and sentiment on the right-hand side.

To better illustrate Guideline 7 and 8, we map the components introduced in Section 5 (Data Tables, Visualisation Structures and Views) onto the subcomponents of DiversiNews.

Data Tables. Data tables constitute the news articles, their summary, their semantic content and various metadata. This is uniform across all displayed visual structures in DiversiNews.

Visual Structures. DiversiNews' visual structures include a tabular view (text-based) of news articles on the left of the screen, a (graphic) visual representation of topics on the top right of the screen, and a geographic representation underneath. We do not consider the sentiment slider to be a visualisation – rather, it may be considered as a control for filtering the other visualisations.

Views. The types of views available for each of the three representations is similar to the ones outlined in the report. The overall news table to the left of the screen is composed of views that depend on the filtering that can be manipulated through the interactive representations on the right. The topic explorer provides feedback to when the target beacon is moved but only changes views when a new query is entered into the search field. The map provides the same types of views as the examples in the report due: manipulations by zooming, panning, etc.

This also allows us to analyse the details and balance of the three related feedback loops introduced in Section 5 (Data Transformation, Visual Mapping and View Transformation).

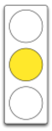
Data Transformation. Data transformation in the context of DiversiNews is achieved by entering a new search query and/or selecting one of the recommended search queries underneath the search box when first opening the tool.

Visual Mapping. The visual mapping process is not controllable by the user since the individual views are defined by the overall layout of the tool. There is no way of selecting which visualisation structure should be in focus – all are displayed at the same time. It is important to note that the directionality of the input (search, manipulate graphs, see articles vs. search, see articles, see graphs) implicitly defines the visual mapping.

View Transformation. View transformations occur on two levels in the context of DiversiNews – once when using the graphs to the right of the results box to refine the search criteria, and once, on a lower level within the map view when zooming, panning, etc the search view.

Below, each guideline is listed followed by a brief report on how well DiversiNews conforms to that guideline.

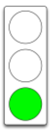
Guideline 1: Intended use should be easily identifiable by ensuring that the level of visualisation falls clearly into one of the four specified categories.



DiversiNews does not fall squarely into any of the visualisation categories, rather sitting somewhere between levels II and III: Visualisation of an Information Workspace and Visual Knowledge Tools. Visualising an Information Workspace involves representing the workspace itself so that users can get a sense of what data is available and how it is related and connected. Visual knowledge tools, in contrast, allow users to rearrange information to actively look for patterns. In the context of DiversiNews, while it is clear that the use pathway begins with a search query, it is unclear if the user is meant to look at the results first and use the visualisation on the right-hand side to better understand the search results or if the user is meant to look at the right-hand side first to manipulate the data that is displayed.

The user experience could be improved by deciding which of the two paradigms is the best fit for the target user group or target user activity and then tailoring the use pathway to better fit just one of the visualisation categories. This will allow users to better understand how they are meant to use the tool. For this reason, the tool is **not in full compliance** with this guideline.

Guideline 2: Reduce the necessity to search for information.

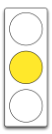


Overall, DiversiNews complies well with this guideline. Through introducing a search query and three variables on the right side, the search for information given a specific topic is reduced. Furthermore, the suggested topics when first opening the page is also an implicit form of reducing the search for information compared to looking through a comparable body of news articles. The summary at the beginning of the search results is a nice addition making an additional search technically unnecessary so long as the summary gives you the required information.

The search can be dynamically adjusted, and adjusting the red dot on the topic map updates the summary as well as the search results, which is a beneficial addition. However, it is unclear how the map is operated and what the marker that can be activated on the map implies. In addition, because this is a news site, there is not a lot of background information (everything is recent developments, which means users who are not familiar with the story threat might be confused). It might be good to include some background information in the summary so that users do not have to perform additional searches.

Overall, DiversiNews **mostly conforms** to this guideline.

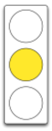
Guideline 3: Use visual representations to enhance the detection of patterns.



The main tool for detecting patterns would seem to be the map and topic visualisations. The map tool conforms to this guideline but could be improved by giving more information, such as the size of dots on the map changing with the volume of news output from those locations. However, the topic diagram is very hard to understand. The spatial relationship of topic clusters to each other does not have any obvious relationship to the news being presented, nor is it easy to immediately pinpoint what each colour-coded topic cluster represents. Although it allows for easy dynamic search modification, it is hard to figure out how one should move the red dot in the visualisation for a particular search goal.

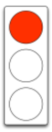
Therefore, the topic visualisation **does not comply** with this guideline, whereas the map visualisation **does.**

Guideline 4: Enable perceptual inference operations.



By placing the news-related topics into a specific visual representation (top right of page), the user is supported in making physical, geometric representations in the context of how specific news stories are related. However, the disconnectedness of the four visualisations and lack of clarity of how they interact (e.g. does the selection of one impact the output of another?) are not in support of visual inference. In addition, because adjusting the visualisations on the right-hand side affects the search results, it would make more sense in terms of perceptual inference to actually put the visualisations on the left-hand side as, at least in the Western world, that orientation makes it easier to perceive cause and effect. On a **micro level** (per visual structure)**, this guideline is complied with, but on a macro level** (in terms of the overall visualisation), **it isn't**.

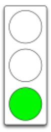
Guideline 5: Use appropriate visual cues to draw attention to relevant data representations.



The tool displays multiple variables, but none of the visualisations draw the attention of the user in a directed way. In other words, the nature of the visualisation does not immediately instruct the user in how to best use the tools at hand for their search goal, with the possible exception of the sentiment slider. The first step in fixing this problem might be to increase compliancy with Guideline 1, followed by a better implementation of the visualisation tools, particularly the topic map.

At present, the tool **does not conform** to this guideline.

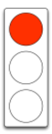
Guideline 6: Encode information in a form in which it can be manipulated.



Manipulation is easily achieved given the controls introduced, but there is a lack of the directionality of manipulation. In other words, the user can modify the search using the visualisation tools, but the same user might have trouble recreating a past set of results if he or she wanted to go back to a previous set of results. While this is a problem, the style of manipulation does facilitate serendipitous discovery of new information. Depending on the tool's intended use (whether is intended to be used to discover new information or intended to better understand the information sphere of a specific topic), which is not clear, the aspect of manipulation could be improved.

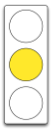
Although it is not easy at present to predict the effect that manipulation will have on search results, the tool nonetheless **complies with this guideline**, albeit not in an ideal fashion.

Guideline 7: Maximise the degree of interactivity based on the limitations of the available data whilst ensuring that the assumed capabilities of the user support the chosen degree of interactivity.



Even though the degree of interactivity is maximised, the tool **does not comply** with the second part of this guideline. The capabilities of the user are unlikely to match the multidimensionality of the controls. This means that most users will find it challenging to optimise their search results based on the tools available. Furthermore, given the degree of feedback, the chosen interaction paradigm exceeds the likely ability level of the user.

Guideline 8: Create data tables that accommodate the diversity of the data input without creating unnecessary complexity in the visualisation scheme.

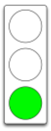


In DiversiNews, the data tables generated upon formulating a search query accommodate the diversity of the data; this is supported through the presence not only of metadata in the view of individual news stories in the overview field on the left, but also through the diversity of the available manipulation tools to the right of the main view. By including topics, location-based data and sentiment data, the diversity of the displayed news articles can be accommodated.

At the same time, the multitude of the available controls creates a complex interface with a bidirectional way of interacting (where the tools on the right may be used as further analysis of the data displayed in the middle as well as a way of manipulating the filters of what is displayed in the main search window on the left.

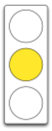
Overall, this implies that the **first half of the guideline is satisfied**, whereas the **second half is not fully satisfied**.

Guideline 9: Select visual structures that best display the relevant data tables without having to filter data.



Different parts of DiveriNews comply with this guideline to varying extents. The map feature complies well as a map is an appropriate way of displaying location information. However, the topic diagram, as specified previously, is a confusing way to display the data. In addition, the sentiment slider does not indicate the overall sentiment bias in the news, which is a critical piece of information that is left out that could be easily visualised. Although the **guideline is met overall, the implementation needs improvement**.

Guideline 10: Limit the number of views to what is feasible in the selected interaction paradigm (e.g. map vs. table) whilst being aware of the resultant complexity created for the user.



There are six views that the user deals with: the summary, topic visualisation, sentiment slider, map, news article list, and search box. Items like the summary help the user deal with complexity, but, as previously, items like the topic map and slider are not optimised to facilitate a greater understanding of the data space. The concept of the views do not necessarily need to be changed, but their implementation could be improved. In terms of the number of views, it would be visually cumbersome to have more, so the number of views is at the extreme of what is feasible within the guideline. The **first part of this guideline seems to be satisfied, whilst the second half is not**.

Conclusions

DiversiNews is an innovative and creative news visualisation service. In general, while the ideas behind the data visualisations are solid and in compliance with design guidelines, the implementation of many of the features, in particular the topic map, need to be improved to be fully in compliance. In order to bring the site into full compliance with guidelines, it is not necessary to do a full redesign, but rather to fix a few of the individual visualisations to make them either easier to understand, easy to manipulate in a meaningful way, or both. However, the map visualisation is a good example of a feature that is already in compliance with the guidelines. There are several key modifications that could improve DiversiNews:

(1) Move data visualisations to the left-hand side of the page to increase user awareness of the cause and effect mechanism between visualisation manipulation and summary/article updates.

(2) Clarify the purpose of the topic visualisation. Users should be able to draw meaning from how the topic map is laid out and how moving the red dot will affect search results.

(3) Add additional information to the map, such as making location markers bigger to indicate more news output from those locations.

(4) Change the slider so that its default location indicates the general feel of the news around a certain topic to the user (e.g. mostly positive).

(5) Introduce a "back" and/or "reset" button such that users can recreate past searches based on their modifications.

9.2 Link ExtrActor (LEA)

LEA has the potential to be a very useful tool for users of Wikipedia, particularly those who frequently write or edit pages. However, its current layout and interaction method make it very difficult to use. The way that data is presented does not necessarily help the user understand the data, and it is unclear what the visual cues are supposed to represent such as the colour-coding and pie chart. The following aspects would help to improve LEA:

(1) After a search, the page should look noticeably different, such that the results of the search are immediately apparent.

(2) The presentation of the data should instruct users where to look on the page. The visualisations should have a clear "flow" between them.

(3) The nature of the visualisations should make it clear to the users how they are meant to be interpreted. At present, a novice user would struggle to figure out how the colour-coding or pie chart related to the actual content of a Wikipedia article. The visualisation itself should make this clear. In this capacity, a simple table or pie chart is unlikely to suffice.

(4) At the moment, the only method of interactivity is through conducting a search. While this may be sufficient, the developers of LEA should consider whether increasing the degree of interactivity would be useful to users.

(5) The large tabular representation of data should allow the user to access more views through filtering and/or sorting, thereby enabling a less complex interpretation of particular subsets of data.

(6) By making the same tabular representation include excerpts of the relevant text in which the Wikipedia articles are cited, the user will have better access to the relevant information (the context in which these links occur in the original text) and hence find it easier to use the tool as part of its intended use.

9.2.1 LEA Evaluation

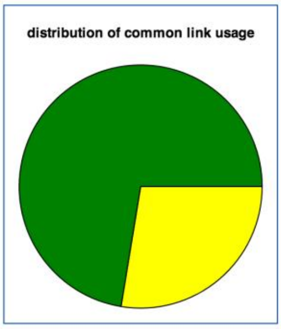
This appendix aims to evaluate the extent to which Link ExtrActor (LEA) conforms to the guidelines put forth in Deliverable D3.1.3. LEA is a tool that allows users to find missing links in a Wikipedia page by comparing it against the same page written in other languages. It also extracts concepts that may not be linked at all and informs the user if the concepts have existing Wikipedia articles that the target article does not link to. Figure 7 shows a screenshot of an LEA search.

The article Cambridge exists in 81 further language versions.

Requested language version (en)	570 links
The following 3 language versions contain the most links to other Wikipedia articles:	
Cambridge (it)	171 links
Cambridge (fr)	172 links
Cambridge (cs)	208 links
Intersection	58 links

no article	0 links
not linked	16 links
linked	42 links

Link to this analysis



English	Italiano	Français	Čeština
1025	1025	1025	1025
1284	1284	1284	1284
1446	1446	1446	1446
1515	1515	1515	1515
1534	1534	1534	1534
1610	1610	1610	1610
1845	1845	1845	1845

Figure 7: Sample query with LEA - Link ExtrActor

To better illustrate Guideline 7 and 8, we map the components introduced in Section 5 (Data Tables, Visualisation Structures and Views) onto the subcomponents of LEA.

Data Tables. The data table generated after a query to LEA is composed of sets of terms that exist as titles of Wikipedia entries within multiple language variants of a specific Wikipedia article. For each of the languages, the table indicates whether the link exists.

Visual Structures. The visual structures present in LEA break down into three separate elements: (1) a textual representation that indicates the overall statistics of the query previously entered at the centre top; (2) a pie chart, which acts as a graphic representation indicating the number of links in a specific Wikipedia article in a specific language correspond with the number (and association) of links that the same article in different languages include; and (3) a table that acts as a Visually Enhanced Object (Level IV Visualisation), indicating the presence of a specific link in multiple variants of the same article (in different languages).

Views. The views of each of the three structures present in the visualisation are limited to the static unfiltered view the user is presented upon first instantiation with the exception of the pie chart, which has three distinct views, dependent on whether the legend is shown or not. Here, the views are switched by hovering the mouse above the relevant area of the pie chart.

This also allows us to analyse the details and balance of the three related feedback loops introduced in Section 5 (Data Transformation, Visual Mapping and View Transformation).

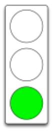
Data Transformation. The data transformation stage of LEA is dictated by the search query formulation when first opening the tool. Here, the user has two choices to make: the article he/she chooses to analyse and the language in which he or she wants to analyse it.

Visual Mapping. Due to the emphasis on a specific intended use of this tool, it does not enable the user to interact with the selection of visual structures; instead he/she is presented with the three structures outlined above.

View Transformation. Finally, there is only one available view transformation: the display of annotations on the pie chart to the top left of the output.

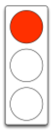
Below, each guideline is listed followed by a brief report on how well LEA conforms to that guideline.

Guideline 1: Intended use should be easily identifiable by ensuring that the level of visualisation falls clearly into one of the four specified categories.



LEA fits well within the first visualisation category: visualisation of information spheres. The purpose of the tool is to find missing links in Wikipedia articles and therefore to help users visualised concepts that are and are not linked. The data all exists outside of the user's workspace and does not require manipulation past the initial search. Therefore, LEA **complies well with this guideline**.

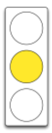
Guideline 2: Reduce the necessity to search for information.



After entering a search query, users are directed to a page that is the same at the top as the initial search page, meaning that on a small screen they may not even realise that the search results are lower in the page unless they scroll down. Once users scroll down, the information provided is somewhat confusing. The language of the articles in comparison languages is given in that language, meaning that users not familiar with the language abbreviations may not know what languages their topic is being compared against. LEA informs users how many different languages the article exists in, but does not give an option to see what those languages are.

There is a pie chart for distribution of common link usage, but users have to hover over the pie chart in order to see the numbers, which is not obvious or intuitive. Finally, the colour-coding of the link chart does not make it clear which language the links are missing in (if a link is green in the English version, does that mean that it is also linked in the other language versions?). LEA does very little to minimise the need to search for information and therefore **does not comply** with Guideline 2.

Guideline 3: Use visual representations to enhance the detection of patterns.

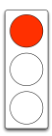


It is not clear if LEA is actually meant to enable the detection of patterns, since users only look at one article at a time. It is unlikely that pattern detection or recognition would increase the usability or usefulness of LEA, as there may not be a pattern to which links are missing (if any).

However, the visual representations given are confusing, as it is not clear what the colour-coding actually corresponds to (does it represent which links are present in other articles but not in the requested-language version? Does it represent which links are present in the requested-language version but not the English version?).

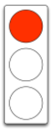
Furthermore, the purpose of the pie chart for distribution of common link usage is not clear, nor is it clear what information it is meant to present. Although pattern-detection may not be necessary for LEA users, LEA nonetheless only **partially complies** with this guideline.

Guideline 4: Enable perceptual inference operations.



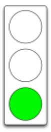
The two perceptual inference operations that are facilitated by LEA are seeing the distribution of common link usage (pie chart) and the colour-coding of which concepts are and are not linked. The colour-coding would work well if it were made clear what each colour represented in the context of the table (it is not clear what the comparisons with other language pages actually represented). This is more of a fault of the instructions than of the interface, although there is probably potential to redesign the table in a way that clarifies how the user is meant to interpret the information. The pie chart is similarly confusing and does not give the user enough information to figure out what inferences they are meant to make. For these reasons, LEA **does not comply** with this guideline.

Guideline 5: Use appropriate visual cues to draw attention to relevant data representations.



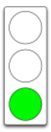
As with Guideline 1, because the page looks the same before and after search unless the user scrolls down, users of devices with a small screen may never even see the data. It would be advisable to have the screen look noticeably different after a search so the user knows where the search results will be. Second, there is no clear visual flow to the search results once they are presented; it is not clear where users are meant to look or how they are meant to interpret the information given, and therefore the data tables and charts do not give a feel for the order in which users should look for information or process the data on the page because is attention actively drawn to any single component. The page should be laid out so that there is a clear visual flow. For these reasons, LEA **does not comply** with this guideline.

Guideline 6: Encode information in a form in which it can be manipulated.



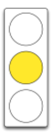
In LEA, information manipulation requires a new search. Whilst this may be sufficient to satisfy the guideline, we recommend that the data presented in the table (the main results section, indicating compliancy in multiple languages) should be sortable or filterable. This would help the user in their interaction with the data and allow expert users (such as users that use the output from LEA for moderating Wikipedia entries) to better navigate through the information that is presented and visualised with this tool. Hence, we conclude that while the tool **complies with** this guideline, the extent of manipulation could be enhanced to meet the demands of the tool's intended use.

Guideline 7: Maximise the degree of interactivity based on the limitations of the available data whilst ensuring that the assumed capabilities of the user support the chosen degree of interactivity.



The only interactivity allowed by LEA is that the distribution of the common link usage pie chart shows actual numbers when the mouse hovers over it. Other than that, the only interaction is with conducting a new search or clicking on links. Clicking on links does not open a new window, meaning that users will lose their search in some browsers, so the interface should be changed to open a new window when a user clicks on a link. It would also be useful to change the way that pages are linked such that when users are taken to a linked page, the information they have searched for is highlighted or shown first. Nonetheless, for the purpose of LEA, the current degree of interactivity is appropriate, so it **complies with** this guideline.

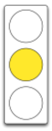
Guideline 8: Create data tables that accommodate the diversity of the data input without creating unnecessary complexity in the visualisation scheme.



The data tables that are generated are sufficient for the first two visualisations (the top left and top right) which both contain summary information regarding the article that was the basis of the search query. The bottom visualisation structure, which goes into greater depth of the analysed depth, however, is limited by the limited data available in the data tables. Here, if excerpts of the context of each of the words were included, the visualisation structure could provide far better information. Once such recommendation is including short annotated examples of each of the links in each of the languages. This would help users make better use of the data displayed as the output of LEA. An example of how this could be achieved was outlined as one of the case studies presented in this report – linguee.de, Section 7.3.

Hence, we conclude that the **first half of this guideline is met,** while the second half of the guideline is **not satisfied.**

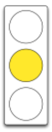
Guideline 9: Select visual structures that best display the relevant data tables without having to filter data.



Because of the issues with clarity of the data, it is unlikely that the data tables are currently laid out in the best possible form. The summary visual structures (the two structures displayed at the top) provide relevant information to the user, whilst the representation of data in the central results section (below) does not provide the best possible information to the user, given the tool's intended use.

Instead, we recommend that the central table fully engages with the concept of a Visually Enhanced Object (Visualisation Level IV), hence enabling the user to explore a corpus of data – the compliancy of links in Wikipedia articles across a range of languages – through presenting the user with a table that provides an interactive basis for exploring relevant information to each of the cells displayed within it. The top two visual structures of LEA **satisfy this guideline**, while the **central visualisation does not**.

Guideline 10: Limit the number of views to what is feasible in the selected interaction paradigm (e.g. map vs. table) whilst being aware of the resultant complexity created for the user.



As there are only a very limited number of views of one of the visual structures (the pie chart) and only one view associated with the other two visualisation structures, this guideline is met. Nonetheless, the lack of clarity and resultant complexity of the central visualisation imply that the limitation of views (and inability to filter the representation) actually is detrimental to the user being able to process the resulting visual output. We strongly recommend that, given the complexity created for the user, the number of views of the central data table is increased, such as by enabling filters, sorting the rows by colour or name, etc.

Hence, we conclude that even though the **first two visualisation structures meet this guideline, the third does not.**

Conclusions

LEA has the potential to be very useful for users of Wikipedia, particularly those who frequently write or edit pages. However, its current layout and interaction method makes it very difficult to use. The way that data is presented does not necessarily help the user understand the data, and it is unclear what things like the colour-coding and pie chart are supposed to represent. The following things could help improve LEA:

(1) After a search, the page should look noticeably different, such that the results of the search are immediately apparent.

(2) The presentation of the data should instruct users where to look on the page. The visualisations should have a clear "flow" between them.

(3) The nature of the visualisations should make it clear to users how they are meant to be interpreted. At present, a novice user would struggle to figure out how the colour-coding or pie chart related to the actual content of a Wikipedia article. The visualisation itself should make this clear. In this capacity, a simple table or pie chart is unlikely to suffice.

(4) At the moment, the only method of interactivity is through conducting a search. While this may be sufficient, the developers of LEA should consider if increasing the degree of interactivity would be useful to users.

(5) The large tabular representation of data should allow the user to access more views through filtering and/or sorting, thereby enabling a less complex interpretation of particular subsets of data.

(6) By making the same tabular representation include excerpts of the relevant text in which the Wikipedia articles are cited, the user will have better access to the relevant information (the context in which these links occur in the original text) and hence find it easier to use the tool as part of its intended use.

9.3 Drupal Extension

We conclude that RENDER's Drupal Extension, which extends the sentiment-based indexing and analysis of Drupal articles, is of benefit to the discovery and navigation through Drupal-based sites. The simplicity of its interactions and the scope for diverse and complex datasets results in high compliancy with the guidelines set out in this report. Nonetheless, due to its unique positioning in the context of RENDER's other tools, the analysis is not directly comparable to the analysis of the previous two tools. We recommend the following changes to improve the Drupal Extension tool:

(1) Separate the RDF views from the navigational views – this will allow for a clearer separation between the two user groups the tools are catered for and allow a wider scope for implementation across Drupal sites.

(2) Add more views (with filters and reordering functionalities) to the two RDF representations, facilitating the navigation without limiting the user with complex controls.

(3) Add a visual segmentation between the sentiment categories, allowing for inference operations in the process of understanding the implications and trends of the sentiment of all available articles relating to a specific topic or keyword.

9.3.1 Drupal Extension Evaluation

This report aims to evaluate the extent to which RENDER's Drupal Extension conforms to the guidelines put forth in Deliverable D3.1.3. The Drupal Extension allows Drupal articles to be explored in their diversity, providing an analytical framework that is front-ended with controls that allow users to explore further (related) articles given extracted metadata from the displayed article. In many ways, this tool is unique in the context of RENDER, as it does not provide a dedicated visualisation or representation of a particular data table, as all other RENDER tools do. Rather, it acts as a method of enriching the navigation through a Drupal installation by providing direct access to similar articles with a negative, neutral or positive sentiment. This is achieved by analysing the article's topics, which is a further navigational tool presented as part of this extension.

Please note that this review is incomplete due to the lack of documentation and access to the tool.

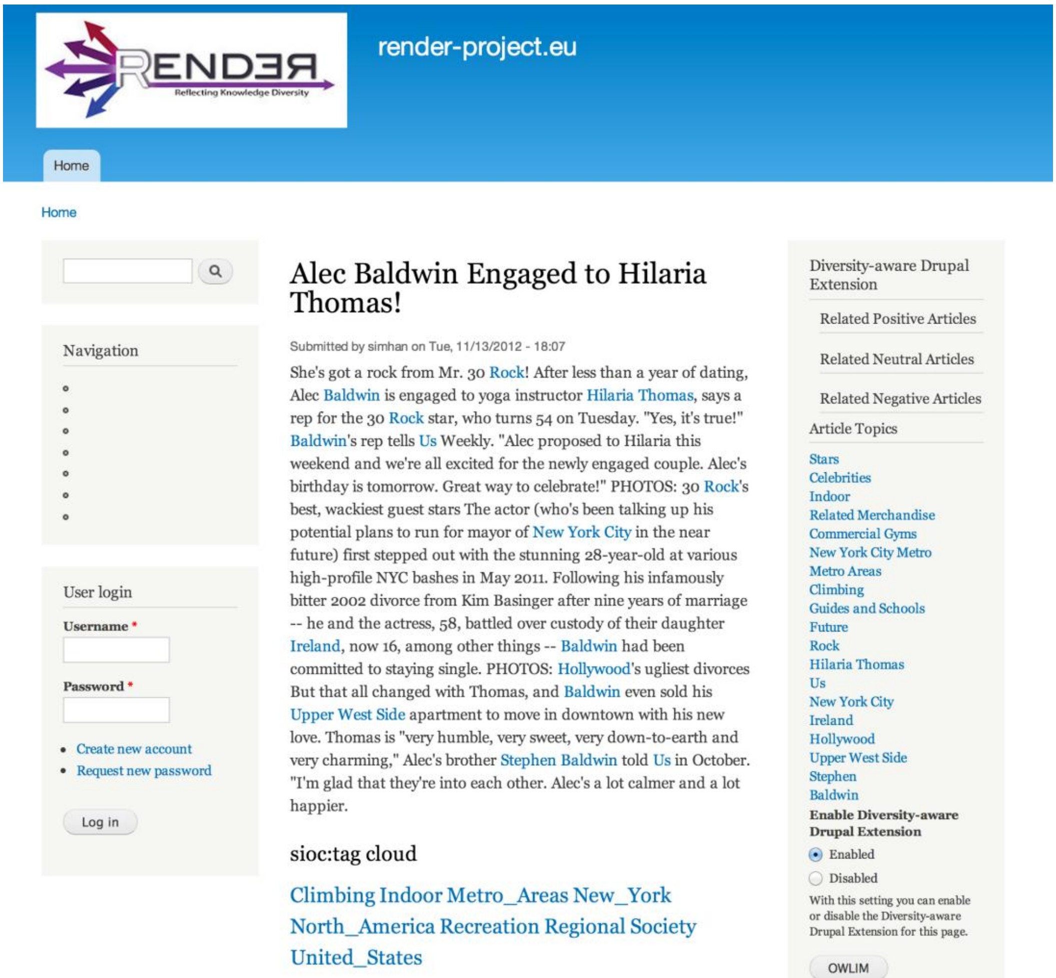


Figure 8: Sample query with the Drupal Extension introduced in RENDER

To better illustrate Guideline 7 and 8, we map the components introduced in Section 5 (Data Tables, Visualisation Structures and Views) onto the subcomponents of the RENDER Drupal Extension.

Data Tables. Defining the data tables in the context of the Drupal Extension is not as straight-forward as with some of the other tools introduced in the project. This is mainly due to the fact the extension, as explained above, is not a visualisation. Rather, it employs some of the concepts that are otherwise used as the basis of a visualisation for ordering Drupal articles. The data tables hence are a catalogue of all articles, with a predefined query that matches the keywords and content of the displayed article and various fields that allow an analysis by keyword, relatedness and sentiment.

Visual Structures. Apart from Lists and a tabular view of the related articles, related by columns that relate to the sentiment of the articles, there are two tabular structures with a colour coding – the RDF (Resource Description Framework) View and the RDF Structural View, both aimed at expert users. The tabular representation by keywords is shown as a popup and depicted in Figure 9.

Views. The views of the text-based representation is defined by the state of state of the tree structures in which the keywords are displayed. By default all trees (to the right of the article) are collapsed. Due to the possible combinations of tree expansions, there are multiple different views that may be seen as types of filters.

Figure 9: Detailed tabular view of articles relating to sub-topic and sentiment given a particular keyword present in the originally displayed article

This also allows us to analyse the details and balance of the three related feedback loops introduced in Section 5 (Data Transformation, Visual Mapping and View Transformation).

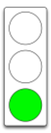
Data Transformation. Unlike most of the other tools introduced by RENDER, the Drupal Extension does not have an explicit search tool. It relies on serendipitous discovery of related articles. Hence, the data transformation happens implicitly and selects the relevant articles, given the currently displayed article, for all available and analysed articles.

Visual Mapping. Due to the singularity of the visual structures employed in this visualisation, the mapping process into a list view with a tree structure is simple. The additional visual structures can be requested by clicking the related buttons and are displayed in a popup within the webpage.

View Transformation. The view transformations in the context of this tool are extensions of the list view, generally operated by expanding or collapsing the trees on the side of the article or in the dedicated visualisation. The RDF views are static and are hence presented in only one view.

Below, each guideline is listed followed by a brief report on how well RENDER's Drupal Extension conforms to that guideline.

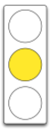
Guideline 1: Intended use should be easily identifiable by ensuring that the level of visualisation falls clearly into one of the four specified categories.



The Drupal Extension clearly satisfies this guideline – its intended use is clear and falls into Level II: Visualisation of an Information Workspace. Within this context, the tool helps related articles to be identified and selected by the additionally available metadata outlining the sentiment of the related article. The RDF views extend into Level III, allowing developers and expert users to see and analyse patterns in a technical syntax.

Therefore, even though the levels of visualisation in this tool spread across Level II and III, this **guideline is conformed** to. The reason the spread is acceptable is because the visualisation structures that fall into Level III are a subset that allow analysis by a different (expert) user group.

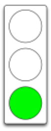
Guideline 2: Reduce the necessity to search for information.



By supporting the concept of serendipitous discovery, this tool helps reduce the necessity to search for information given the information sphere of a Drupal distribution with rich, diverse content. Its sole purpose is to shorten the path to finding relevant and related information to the article the user is currently viewing, segmented by sentiment. Hence, within the user journey, it allows the user to explore and search for information in a subtle way.

Nonetheless, due to the depth of the information that is displayed in the form of keywords, the tool may not in practice reduce the necessity to search for information if the article keyword density or length is especially high. Hence, we conclude that this guideline is **partially conformed** to.

Guideline 3: Use visual representations to enhance the detection of patterns.

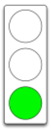


In order to respond to the conformity of this guideline, we must consider the tool in its separate intended use cases. In the context of using the Drupal as an extension to navigating through articles, the tool does not use any visual representations to enhance the detection of patterns. This, however, is in compliance with the limited functionality of presenting related articles, segmented by sentiment. The only possible extension would be by using a colour scheme to indicate the sentiment of the links.

The second intended use (for expert users) is the RDF representation, which is shown in two different variants – in the text (xml) form, which is difficult to navigate, and in a “Structured” variant, which uses a tabular and colour-coded scheme. Hence, this uses visual representation to enhance the detection of patterns in the coding.

Given the simplicity of the frontend of the tool, this guideline **is conformed to** by RENDER’s Drupal extension.

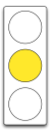
Guideline 4: Enable perceptual inference operations.



Given the lack of graphic visualisations in this tool, the number and scope of perceptual inference operations is very limited. The notable exception is the column-based view of the sentiment analysis related keywords, as displayed in Figure 9. We recommend that a visual structure be presented to the user in the side view that allows him or her to infer the sentiment of all related articles. This would add value to the tool (given its intended use) and support the previous guidelines.

Given the type of data displayed, this **guideline is adhered to**.

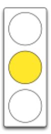
Guideline 5: Use appropriate visual cues to draw attention to relevant data representations.



Apart from the structured RDF visualisation, in which colour coded table cells are employed to relate to content contained in the RDF representation of the data, no visual cues are employed throughout this tool. We recommend that colours, were appropriate, are utilised as the basis drawing attention to relevant representations; this might be to emphasise the most relevant articles or repeated keywords or might be through indicating which articles had previously been linked to or visited by the user.

In the absence of visual cues and a pure text representation, we conclude that this guideline **is partially followed**.

Guideline 6: Encode information in a form in which it can be manipulated.



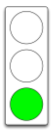
The data presented as the output of this tool is not interactive, apart from the ability to expand and contract the resultant list views. Given the complexity and diversity of the displayed keywords, we recommend that the interactivity of some of the features be increased. As previously outlined, this may happen on two levels – one for general users, based on the navigation functionality of the related articles and keywords, and one on the basis of the interactive potential that the structured RDF view caries. We suggest that a search and order functionality is presented in the structured RDF view and that the basic RDF view allows individual nested statements to be expanded and collapsed, allowing the user to quickly process large amounts of data. This could also be accompanied by a table of contents on the left of the xml view, allowing statements to be presented in an indexed fashion. Furthermore, we suggest that the generation of data tables (search, or suggested terms) is implemented in the side panel, allowing a wider array of data to be processed by the user.

Apart from these modifications, given the simplicity of the data available in the context of this tool, the degree to which manipulation is possible is limited and hence we conclude the tool displays partial conformity with Guideline 6.

Guideline 7: Maximise the degree of interactivity based on the limitations of the available data whilst ensuring that the assumed capabilities of the user support the chosen degree of interactivity.

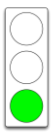
Given the low level of interactivity, we conclude that this guideline is of limited relevance to this tool.

Guideline 8: Create data tables that accommodate the diversity of the data input without creating unnecessary complexity in the visualisation scheme.



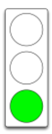
The generated data tables nicely accommodate the diversity of the data (spread across a spectrum of sentiments and keywords) without creating complexities in the resulting visualisation. The only complexity generated is the nature in which the date is displayed in a mixed format between tree representations, actions that allow articles to be loaded, the RDF views and an action that applies any changes to the displayed data. We recommend that the controls targeted at different data tables (expert vs. general users) be separated. Overall, this **guideline is satisfied** by the tool.

Guideline 9: Select visual structures that best display the relevant data tables without having to filter data.



The separation of the RDF presentation (and two-staged approach to RDF interpretation) is an elegant way of ensuring that the entirety of the data tables are displayed without having to create complex filtering schemes and/or complex interfaces. The tree view is equally a good choice, as long as the displayed data is limited in the number of keywords. Once the diversity of keywords grows, a different visual structure (as outlined in Guideline 2) may be a better way of displaying the relevant data tables. We recommend that a topic visualisation, as employed by DiversiNews (see Appendix A) be employed in the context of the displayed data of the RENDER Drupal Extension. We conclude that **Guideline 9 is followed** by this tool.

Guideline 10: Limit the number of views to what is feasible in the selected interaction paradigm (e.g. map vs. table) whilst being aware of the resultant complexity created for the user.



The limitation of the number of views in the context of this tool may be considered as too limited for the host of data available in its context. Especially in terms of the RDF visual structures, multiple views would aid users to explore the diversity in the datasets without a compromise for a more complex interface. While this **guideline is satisfied** in principle, adding additional views would not compromise the complexity and be of benefit to the user.

Conclusions

We conclude that RENDER's Drupal Extension, which extends the sentiment-based indexing and analysis of Drupal articles, is of benefit to the discovery and navigation through Drupal-based sites. The simplicity of its interactions and scope for diverse and complex datasets results in a generally high compliancy with the guidelines set in this report. Nonetheless, due to its unique positioning in the context of RENDER's other tools, the analysis is not directly comparable to the other two. We recommend the following changes to improve the Drupal Extension:

(1) Separate the RDF views from the navigational views – this will allow for a clearer separation between the two user groups the tools are catered for and allow a wider scope for implementation across Drupal sites.

(2) Add more views (with filters and reordering functionalities) to the two RDF representations, facilitating the navigation without limiting the user with complex controls.

(3) Add a visual segmentation between the sentiment categories, allowing for inference operations in the process of understanding the implications and trends of the sentiment of all available articles relating to a specific topic or keyword.

10 References

[1] http://render-project.eu/about-us/approach/

[2] Card, S. K., Mackinlay, J. D., and Schneiderman, B., 1999, “Information visualization: Using vision to think”, San Francisco: Morgan-Kaufmann.

[3] Wilson, M. L., 2007, "Bridging the gap: Using IR models for evaluating exploratory search interfaces.”.

[4] Dörk, M., Williamson, C., and Carpendale, S., 2012, Navigating tomorrow's web: From searching and browsing to visual exploration. ACM Trans. Web 6, 3, Article 13 (October 2012), 28 pages.

[5] Lee, B., Smith, G., Robertson, G. G., Czerwinski, M., and Tan, D. S., 2009, FacetLens: exposing trends and relationships to support sensemaking within faceted datasets. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, NY, USA, 1293-1302.

[6] RENDER Annex I - "Description of Work"

[7]Fortuna, B., Rusu, D., Trampus, M., Dali, L., Stajner, T., and Grobelnik. M., 2011, “Prototype of the Fact Mining Toolkit”, RENDER Project Deliverable D2.2.1.

[8] Popovic, V., 2003, “Expert and Novice Users Model and their Application to the Design Process”, Journal of the Asian Design International Conference.

[9] Card, S., 1996, “Information visualization and information foraging”, In Catarci, Tiziana, Costabile, Maria Francesca, Levialdi, Stefano and Santucci, Giuseppe (eds.) AVI 1996 Proceedings of the workshop on Advanced visual interfaces May 27-29, 1996, Gubbio, Italy. pp. 12.

[10] Russell, D. M., Stefik, M. J., Pirolli, P., and Card, S. K., 1993, “The cost structure of sensemaking”, Paper presented at the INTERCHI '93 Conference on Human Factors in Computing Systems, Amsterdam.

[11] Beale, R., 2007, "Supporting serendipity: Using ambient intelligence to augment user exploration for data mining and web browsing", International Journal of Human Computer Studies 65, pp. 421—433.

[12] Larkin, J. and Simon, H., 1987, “Why a diagram is (sometimes) worth ten thousand words”, Cognitive Science, 11:65-99.

[13] Tufte, E. R., and Graves-Morris, P. R., 1983, *The visual display of quantitative information*. Vol. 31. Cheshire, CT: Graphics press.

[14] Tovée, M. J., 2008, “An Introduction to the Visual System”, 2nd ed. Cambridge University.

