FP7-SME-2012-315631

UGCS authoring evaluation report D4.2

Deliverable lead contractor: MMV

Main contributors:

Dimitris Skoutas   dskoutas@imis.athena-innovation.gr
Nikolaos Papadakis   npapadakis@imis.athena-innovation.gr
Dirk Morgenroth   dm@mmm24.de

Due data: 31.07.2014

Abstract

This document reports the results of the user survey conducted to evaluate the authoring environment. The participants of the survey were developers of the SME partners. Questions included in the survey cover all aspects and functionalities of the GeoStream demo developed so far, into which the authoring environment is integrated, thus the results can serve as an overall assessment and evaluation of the system to guide further development.

Research Center "ATHENA," Greece
FU BERLIN, Germany
Fraunhofer, Germany
Michael Mueller Verlag, Germany
TALENT, Greece
WIGEOGIS, Austria

Table of Contents

1 Introduction

The previous work in the project so far, i.e. conducted in WPs 1 and 2, focused on collecting user-contributed geospatial data from various Web sources, to be integrated and analyzed in the GeoStream database. In WP4, we explore the other aspect of this process, in particular how to support and facilitate the users in creating such content. For this purpose, we have developed an authoring environment, in the form of a Web application but also accessible via an API, which is available online (integrated in the GeoStream online demo[1]) and has been reported in Deliverable D4.1.

In parallel, we have designed and conducted a survey among the SME partners in order to allow them to assess and evaluate the features of the developed application and provide feedback to further improve the work. The target group for this survey was focused mainly on developers of the SME partners, or more generally users with a considerable level of technical knowledge and skills, who are expected to eventually reuse, adapt and potentially extend the developed system to incorporate and use it in their own final products. To familiarize the participants with the functionalities offered by the system and explain how to navigate the user interface and perform various actions, an accompanying manual was provided (also available online[2] on the project website).

The surveys were conducted by means of an online questionnaire[3] that the participants were asked to fill in after using the system. These sessions took place in the context of workshops organized at the SME's locations, as described in Deliverable D7.3

In this deliverable, we present the questionnaire that was used for the survey and we report the findings, describing our observations and proposing corresponding actions to guide and improve future work. In particular, Section 2 gives an overview of the questionnaire, while Section 3 provides information on the participant profiles. Section 4 presents and discusses the results, highlighting observations and proposed actions. Finally, Section 5 summarizes the conclusions.

---

1   http://dataminer.geocontentstream.eu
2   http://geocontentstream.eu/resources
3   http://survey.geocontentstream.eu

2 Overview of the questionnaire

The questionnaire was designed trying to achieve a good balance between being relatively short, so that it is not cumbersome and time consuming to fill in, while ensuring that it can capture all needed information and cover all main features of the system without leaving out important aspects of its functionality. Prior to the design of the Geostream questionnaire, a set of objectives and list of the information we would like to capture were set. This list of objectives and research goals guided our plan for the survey questionnaire. The questionnaire can be completed in about 15 minutes and comprises both structured or fixed response questions and non-structured or open questions.

It is composed of 21 questions and includes 7 different sections covering basic participants' demographics as well as more specific questions about the user experience, functionality and effectiveness of the system. In the following paragraphs these sections are presented together with their questions.

The questionnaire was deployed using a well-known open source application, LimeSurvey[4]. LimeSurvey allows users to quickly create intuitive, powerful, online question-and-answer surveys that can work for tens to thousands of participants without much effort. The survey software itself is self-guiding for the respondents who are participating. It includes a backoffice that allows the administrator to install, monitor, administer the installation, support survey creators and report generation users alike. Performing the survey using an online questionnaire had the benefit of making it easier and faster to collect the results electronically and process them. Moreover, it is easier to perform the survey remotely, as well as to repeat it in the future when needed to get more feedback.

The online survey is running at a Virtual Machine in ~okeanos[5], a cloud service designed and developed by the Greek Research and Technology Network (GRNET) for the Greek Research and Academic Community. The data is stored in a MySQL (version 5.5.38) database. The LimeSurvey Version is 2.05+ and the specific build is 140618.

The questionnaire starts with a short section that aims at selecting some basic information about the participant's profile. Optionally, the participant may also indicate his/her name. This option was provided for two reasons: (a) to give us the opportunity to contact the participant in case any clarifications or more detailed comments were needed, and (b) to be able to contact the participant again at a later point in the future, after revisions of the system have been performed, in order to get an updated feedback and verify whether the corrective actions were successful.

The second and main section of the questionnaire comprises a set of questions, mostly formulated as multiple choice questions, asking the participants to evaluate the various aspects and functionalities of the system. In some cases, the question is stated both in terms of how useful a particular feature is considered and how good its current implementation in the system is considered. The ratings are provided in a scale of 1 to 5, with 1 corresponding to "very poor" and 5 corresponding to "very good". The questions in this section address the following main aspects of the system:

- sources used for content retrieval

- data integration and mining functionalities

- search and browsing

- dashboards and visualizations

---

4   www.limesurvey.org
5   https://okeanos.grnet.gr

- authoring tools

- overall assessment

The complete questionnaire is presented in Appendix 6.1.

3 Participant information

In total, in this evaluation round conducted by the SME partners, 9 questionnaires were completed. The distribution among the SME partners was as follows:
- 6 responses by MMV
- 1 response by WIGEOGIS
- 2 responses by TALENT

In terms of roles, the distribution was as follows:
- 7 participants responded "Developer"
- 1 participant responded "Consulting and Sales"
- 1 participant had no reply.

Moreover, the plots below illustrate statistics of participants profiles regarding age, sex and level of education.
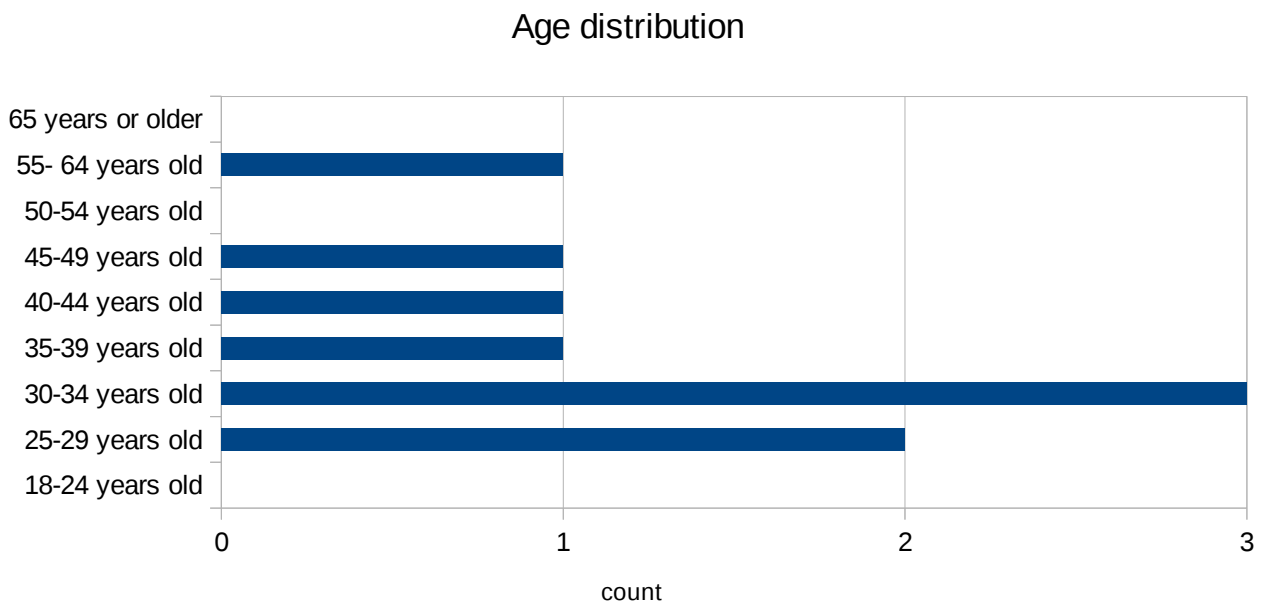
Age distribution

Figure 1: Age distribution of participants.
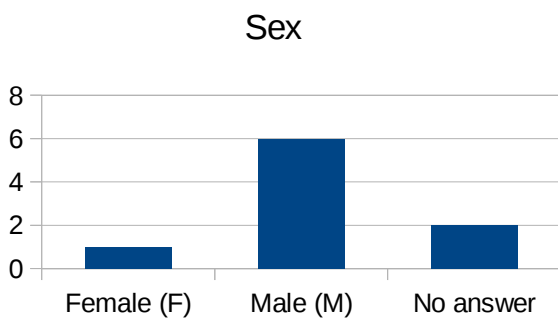
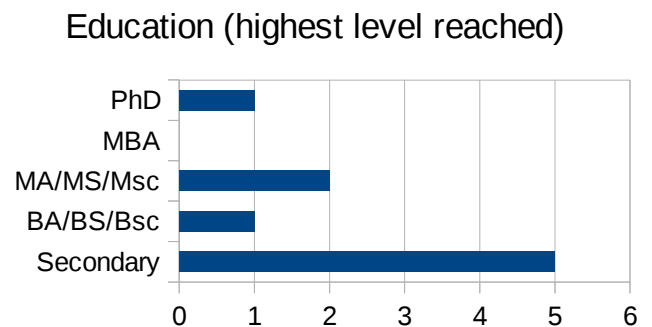*Figure 2: Sex distribution of participants.*

Figure 3: Education level distribution of participants.

4 Analysis of responses

In this section, we present the results of the evaluation and we report our observations and proposed actions.

4.1   Sources used for content retrieval

4.1.1   Results

The results are shown in Figure 4. In addition, the participants suggested to collect content from blogs, Facebook, Instagram and Picasa.

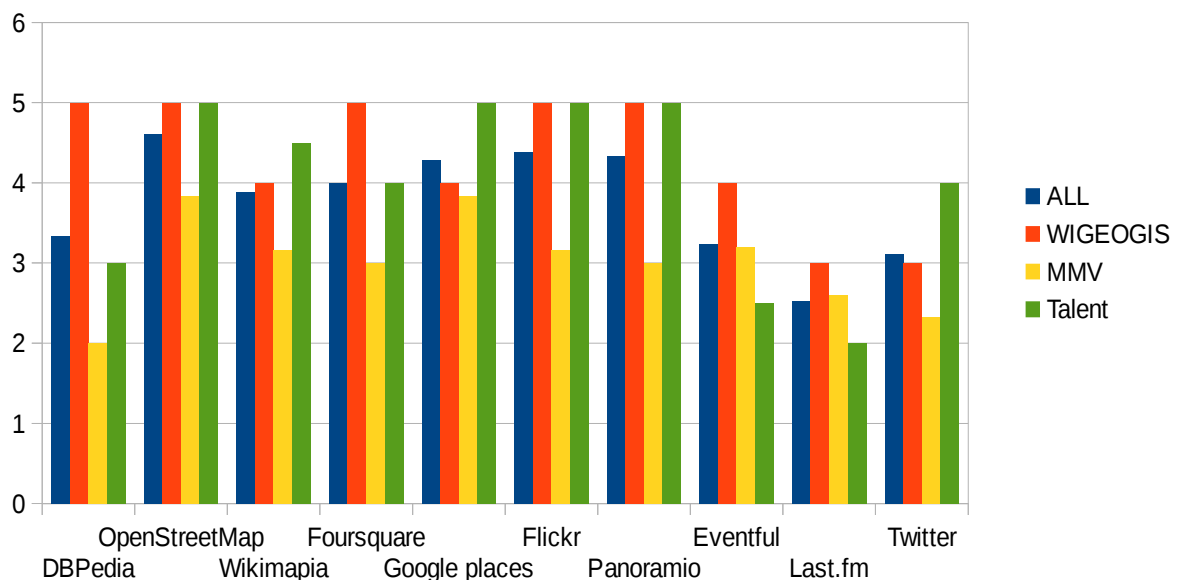How useful do you consider each of the following sources used for data collection in GeoStream?

Figure 4: Evaluation of sources.

4.1.2   Observations

A first rather surprising observation is that the use of DBpedia is rated low by MMV and TALENT. It is not clear what is the reason for this. A possible explanation is that such information, which is rather "objective" compared to the other sources of user-generated content, may already be available to the partners, and thus provides little new information. We will check this point with the partners to explain the result.

In contrast, OpenStreetMap, Wikimapia, Foursquare and Google places have received a high rating, as anticipated. The same applies for Flickr and Panoramio, which shows that photos are highly appreciated. The latter is even further supported by the fact that, as mentioned above, the participants suggested also the inclusion of Instagram and Picasa, which are also sources for photos.

However, sources for events (i.e. Eventful and Last.fm) have received a low rating. Surprisingly, this also applies for Twitter. We believe that this may not necessarily reflect that such content is not useful, but may rather have to do with the fact that this content is not sufficiently exploited in GeoStream at the moment. Hence, this is also a point for further investigation with the SMEs.

4.1.3   Proposed actions

Despite the fact that DBpedia, Eventful, Last.fm and Twitter have received a lower rating compared to the rest of the sources, we believe that such content is equally interesting and useful. Hence, we will try to better exploit and make prominent this content in GeoStream to demonstrate its use, while also getting more detailed feedback from the SMEs.

Regarding the additional sources that were suggested, we will check from a technical perspective if and how these sources can be included. In the case of Instagram and Picasa this may be easier, especially if they provide an API similar to that of Flickr and Panoramio that are already used. For Facebook, there are privacy restrictions that typically make its use difficult in applications as they limit significantly the amount of content that could be used. Finally, blogs are an interesting source that we plan to investigate, but this requires a different approach than the one currently used for all other sources, since it involves crawling several web sites and making use of NLP techniques to extract content. Hence, it requires a totally different content collection process than the one developed for the other sources.

4.2   Data integration and mining

4.2.1   Results

In this part, the questions are answered from two perspectives: how useful the proposed functionalities are considered, and how well they are currently implemented in the system. The results are plotted in Figures 5 and 6.

How useful do you consider each of the following functionalities provided by Geostream?
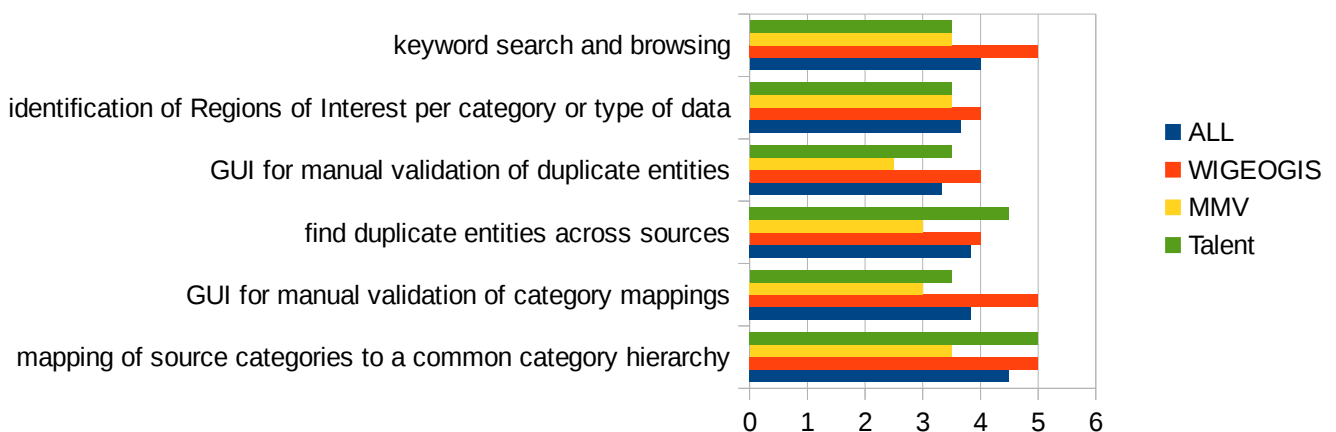

Figure 5: Evaluation of data integration and mining functionalities (I).

How satisfying do you find each of these functionalities as currently implemented by Geostream?
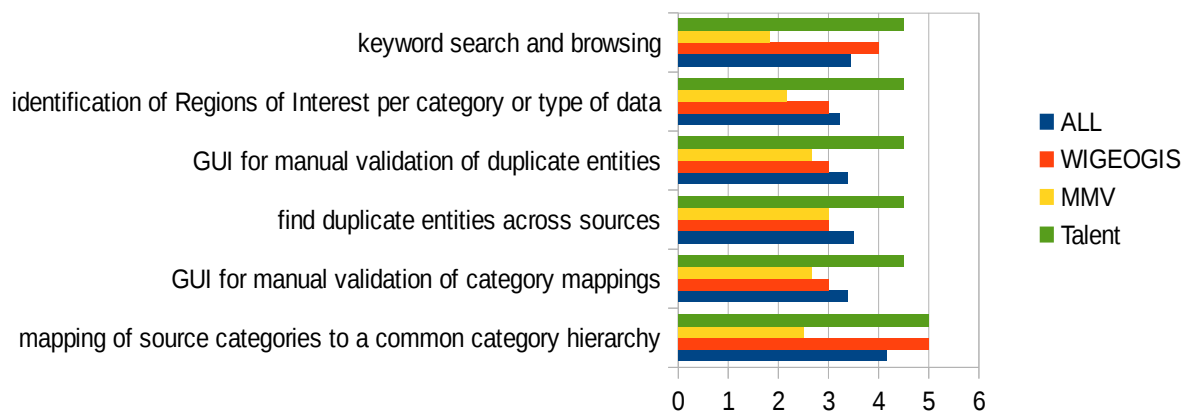

Figure 6: Evaluation of data integration and mining functionalities (II).

4.2.2　Observations

All features have received a positive score (i.e. above average), which shows that they are considered useful. The highest ones are keyword search browsing and the mapping of categories to a common schema. The evaluation in terms of current implementation of those features follows the same trend but with a bit lower scores, which means that there is space for improvement. A surprising observation is that the identification of regions of interest has received a rather low score, although we consider it to be one of the novel and more interesting features of the system.

4.2.3　Proposed actions

One important observation is that the keyword search and browsing, although considered quite useful, has received a relatively low score in term to current implementation. In fact, further evidence for this comes also from other questions that are included later on in the questionnaire. Hence, this is a feature that requires improvement, and we already have identified which specific points can be improved in upcoming versions of the system, as also discussed in following sections.

The relatively low scores for entity deduplication and for the GUI for validating category mappings are not very surprising since these tasks, by nature, are cumbersome and time consuming to perform. We have already tried in our design to make these GUIs easier for the user, e.g. by allowing to order or filter results by score or source, but still one needs to go through hundreds or thousands of categories and entities. In fact, we anticipate that when such functionalities are deployed and used in a system that has a large user base, these tasks could be crowdsourced rather than being performed by the system administrator.

4.3　Search and browsing

4.3.1　Results

The results of the evaluation for the search and browsing functionalities are presented in Figure 7.
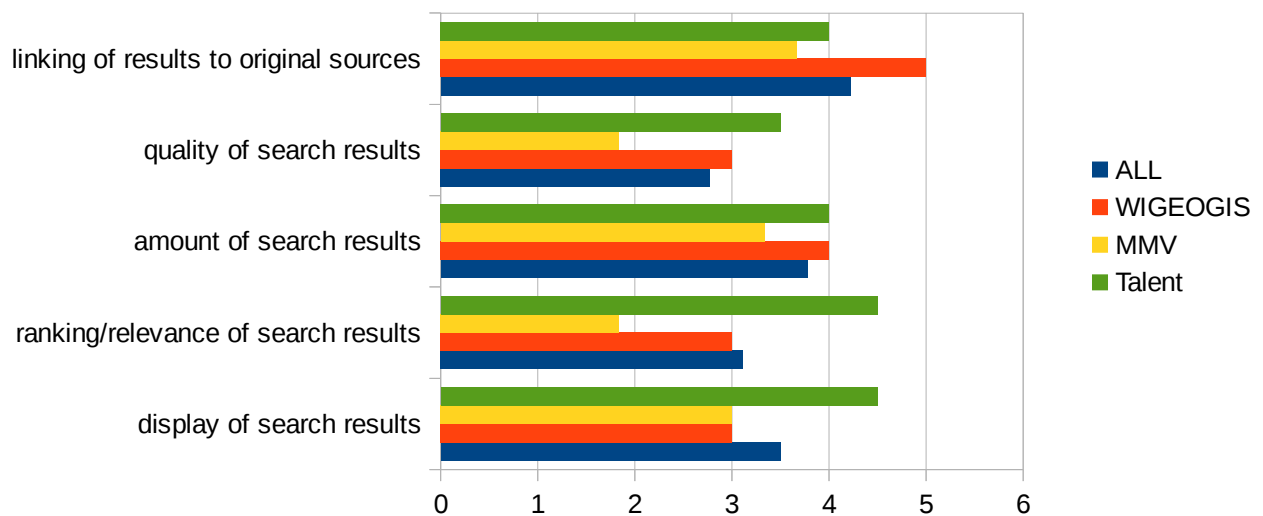
How would you rate the following?

Figure 7: Evaluation of search and browsing.

4.3.2   Observations

Search and browsing is apparently a key feature of the system, hence the results of this section are important. As shown, only the linking of results to original sources is rated very positively at the moment (which also shows that provenance of the information is considered important). Also, the amount of results is found satisfactory, taking also into consideration that data are currently collected only for specific areas and mostly for testing purposes.

Instead, quality, ranking and display of results have received low scores. These are important aspects and definitely need to be improved in future versions of the system.

4.3.3   Proposed actions

Currently, Apache Solr[6] is used for search and browsing. The lower scores for the evaluation above can be explained by the fact that Solr was used with its default configuration for search and ranking of results. In the next versions of the system, we plan to better configure and tune the ranking of Solr according to criteria needed for our type of content (e.g. higher weight for "title" than "description" or "tag", or preferring results of certain sources than others), which we believe can significantly improve the search results. Moreover, we will try to improve the display of search results, possibly having different ways to display different types of results, i.e. for POIs, photos, event or tweets.

4.4   Dashboards and visualizations

4.4.1   Results

Again here there are two perspectives evaluated: the usefulness of the feature and its current implementation. The results are illustrated in Figures 8 and 9.

---

6   http://lucene.apache.org/solr/

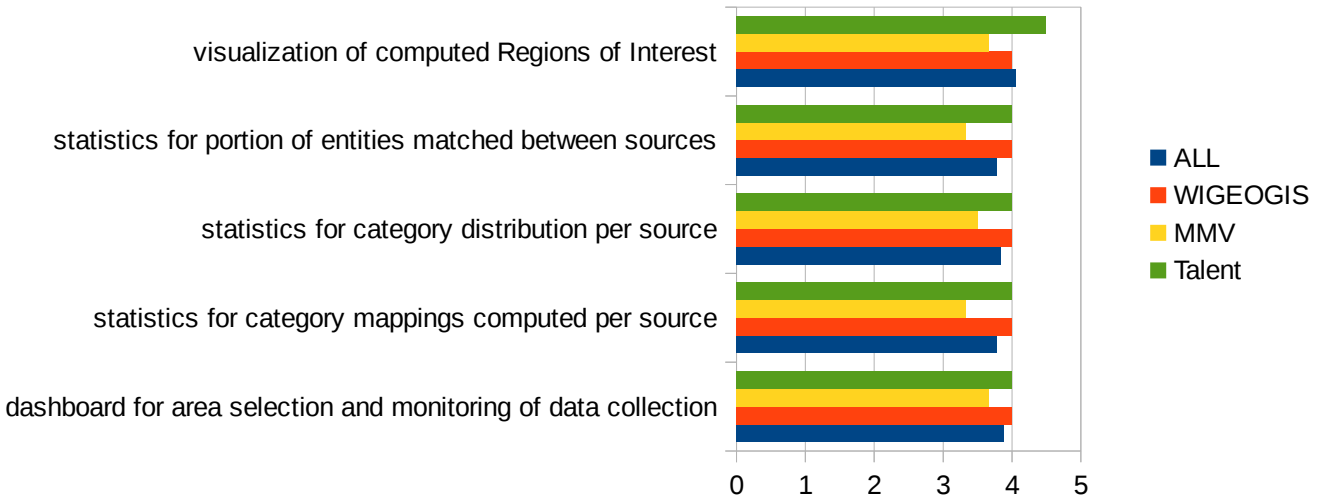## How useful do you consider the following visualizations provided by Geostream?

Figure 8: Evaluation for dashboards and visualizations (I).

How satisfying do you find the following visualizations as currently implemented by Geostream?
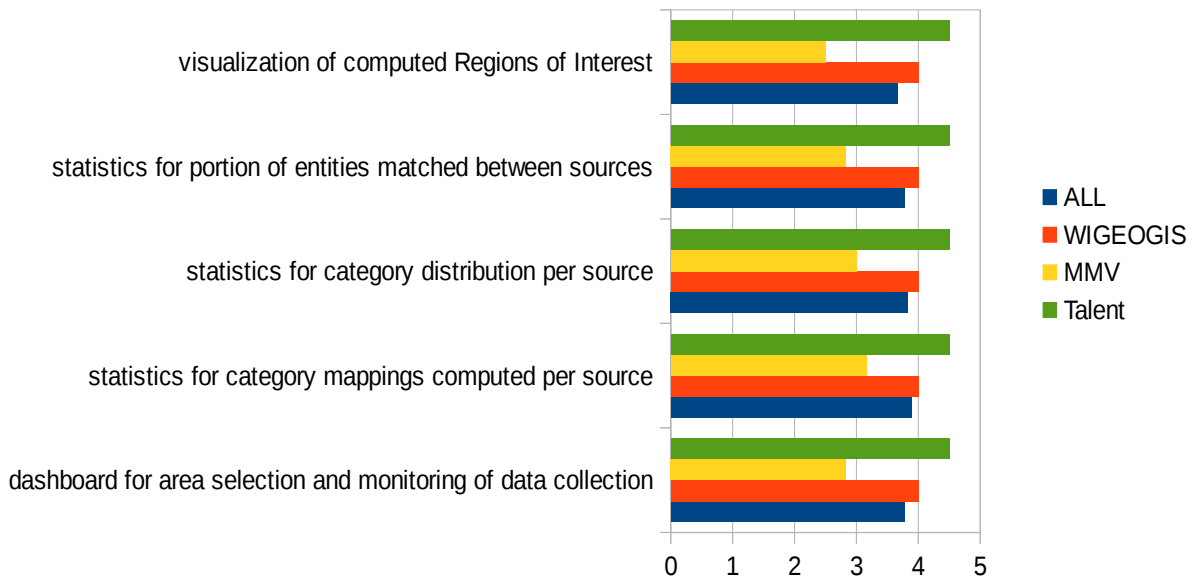

Figure 9: Evaluation for dashboards and visualizations (II).

4.4.2   Observations

These features have been received positively and the results are fairly consistent among the various features and among the partners, with some exceptions for MMV, where the scores are lower for some cases.

4.4.3   Proposed actions

We will communicate with MMV to get more detailed feedback on which specific parts could be improved and how. Especially for the case of regions of interest, we already plan to investigate ways to label or annotate such regions with summarized information (e.g. a tag cloud or top photos), which we anticipate will improve the way they are visualized and conceived. This suggestion was already discussed as a future feature in the last project meeting, so it is already included in the planned actions for next versions.

4.5   Authoring tools

4.5.1   Results

This section presents the results for the evaluation of the authoring tools[7]. The results are illustrated in Figures 10 and 11. In addition, for the question "What would be your preferred way to create a trip?", the following responses were received:

- 4 participants answered "Both"

- 4 participants answered "via search and browsing, and then adding results to the trip"

- 1 participant responded "None".

For the questions included in this section, please indicate your response
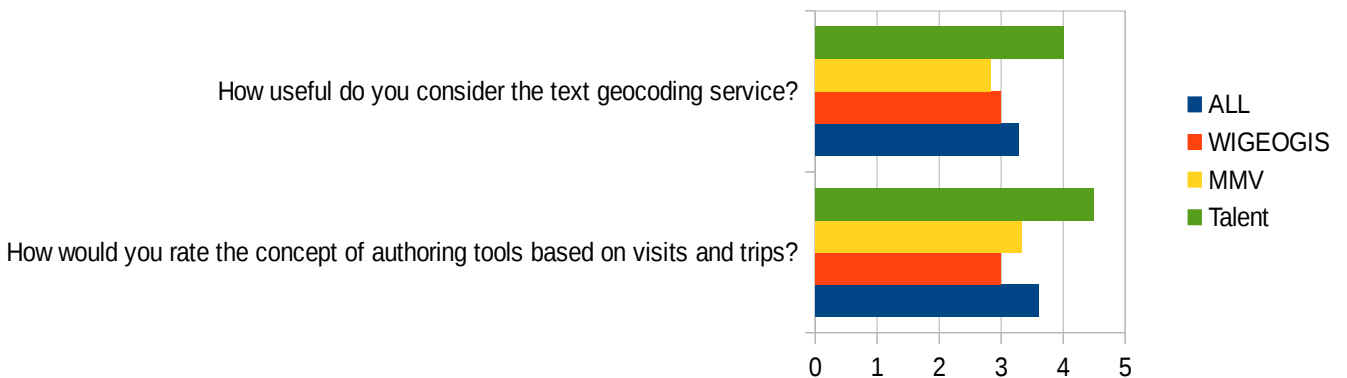
Figure 10: Evaluation of authoring tools (I).

---

7   Note that although the term "authoring tools" is used to refer to this specific part of the system, as opposed to the parts performing, e.g., data collection and aggregation, the search and authoring of content involves all parts and steps of the system and workflow. Hence the authoring environment is considered as a whole, and an overall evaluation of all features has been included in this survey, rather than restricting to a partial evaluation of a specific part.

How useful do you consider the following attributes for describing visits and trips:
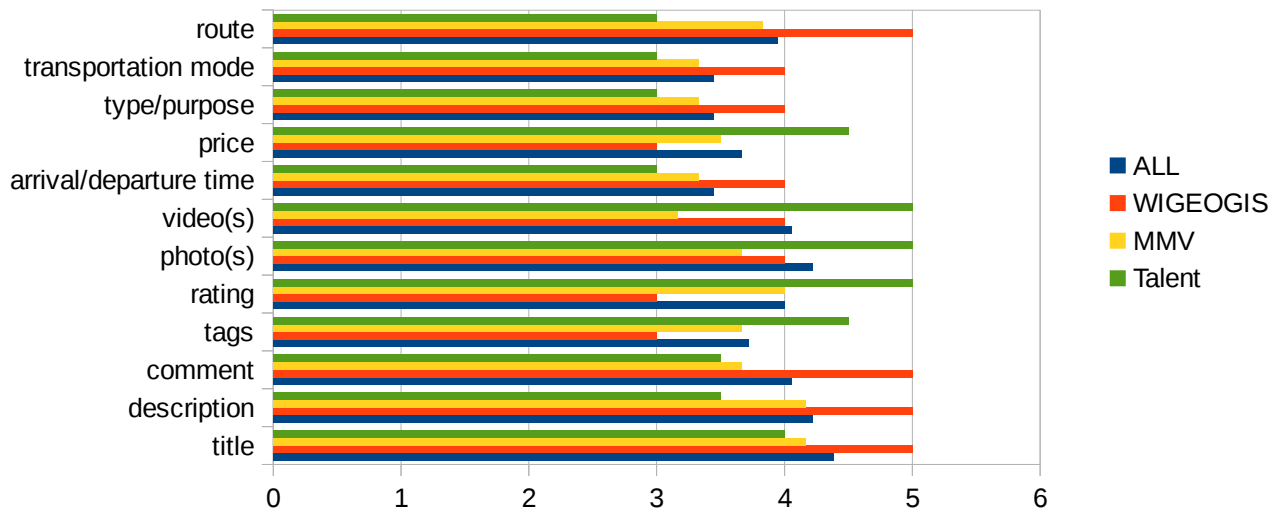
Figure 11: Evaluation of authoring tools (II).

4.5.2   Observations

The overall perception of the geocoding service and the authoring tool centered on creation of visits and trips was positive (above average) but with low score. This is however understandable, since this component is the latest one that has been developed and included in the system, so this was the first version that was presented to the partners and evaluated. It is recognized that further improvement can be done.

In terms of attributes used to describe visits and trips, all those included are considered useful; lower scores were assigned to some, such as transportation mode, price, arrival/departure time. Perhaps this has to do with the fact that those attributes are more specific to the notions of visits and trips and hence less applicable to other cases.

4.5.3   Proposed actions

There is already planned work to improve the performance of the geocoding engine, using data provided from the partners (or other sources) for training. Moreover, we will try to make the authoring of visits and trips more generic and improve the user interface. In fact, some of the functionalities involved here need to converge and be combined with the mobile application designed and developed by Fraunhofer, so some parts are subject to redesign and change.

4.6   Overall assessment

4.6.1   Results

The results for the overall assessment are shown in Figure 12. In addition, a comment that appeared in the question "List any other functionalities that you would like to be implemented in the context of the project", was the following: "could not find out how to combine search terms, e.g. for Vienna "Mozart Exibition""

Overall Assessment
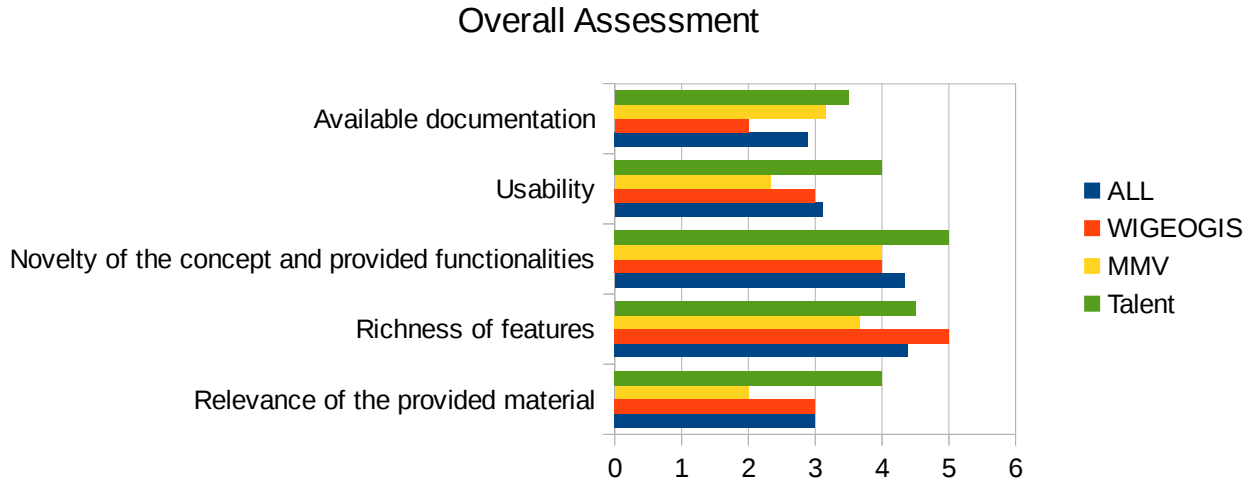
Figure 12: Overall assessment.

4.6.2   Observations

Overall, the system received a positive evaluation. Most importantly, it received high scores in terms of novelty and richness of features. This shows that the features are appreciated and considered useful for being included and used in future products, and that no important features are completely missing from the system at this stage. However, it received lower scores in terms of available documentation, material currently provided, and usability, which are points to be improved in the future.

4.6.3   Proposed actions

For the remaining of the project, we will try to improve the performance and usability of its features. In terms of available documentation, currently there exists a manual (in the form of a pdf file) that was available to the participants, explaining in a step-by-step approach how the system can be used. It seems that this documentation is not sufficient or convenient. Instead, we will embed instructions in the Web application itself, e.g. via a "Help" entry in the navigation menu, as well as use of tooltips on buttons that perform certain actions, which we anticipate will be an easier and more convenient way to guide the user rather than having to consult another document.

5 Conclusions

In this report, we have presented the survey that was conducted to evaluate the authoring environment, and more generally, the GeoStream demo that has been developed so far. The participants of the survey were selected by the SME partners in order to get feedback on how they perceive the provided features and functionalities from the perspective of potential future exploitation in their own products. We have presented the questionnaire that was used for the survey, providing an analysis of the results and pointing out observations and corresponding actions.

Based on the presented results, the overall assessment and feedback for the system was positive, identifying also points for improvement. What we consider important is that the participants have evaluated with a high score the novelty of the concept and its functionalities, and the richness of provided features. This shows that the work is progressing in the right direction and there are no major functionalities missing.

Also important is that the survey has helped us to identify points that need improvement, so as to focus on them for the remainder of the project. In particular, based on the results, we have identified the following main aspects to be improved in future versions:

- ranking and relevance of search results, as well as the way they are displayed

- better presentation and use of the regions of interested, as well as for events and tweets

- increase usability, both through the user interface and the API

- more and better documentation.

6 Appendix

6.1   Questionnaire used for the survey

Participant profile

First name

Please write your answer here:

Please insert your Given name

Last name

Please write your answer here:

Please insert here your Family Name

Please select your Age range *

Please choose **only one** of the following:

- ○ 18-24 years old
- ○ 25-29 years old
- ○ 30-34 years old
- ○ 35-39 years old
- ○ 40-44 years old
- ○ 45-49 years old
- ○ 50-54 years old
- ○ 55- 64 years old
- ○ 65 years or older

Sex

Please choose **only one** of the following:

- ○ Female
- ○ Male

Organization *

Please choose **only one** of the following:

- ○ WIGeoGIS
- ○ MMV
- ○ Talent

Role

Please write your answer here:

job title, such as software developer, manager, UI designer, etc

Education (highest level reached) *

Please choose **only one** of the following:

- ○ Secondary
- ○ BA/BS/BSc
- ○ MA/MS/MSc
- ○ MBA
- ○ PhD

Selection of sources for crowdsourced geospatial content

How useful do you consider each of the following sources used for data collection in GeoStream? *

Please choose the appropriate response for each item:

	don't know	very poor	poor	fair	good	very good
DBPedia	○	○	○	○	○	○
OpenStreet Map	○	○	○	○	○	○
Wikimapia	○	○	○	○	○	○
Foursquare	○	○	○	○	○	○
Google places	○	○	○	○	○	○
Flickr	○	○	○	○	○	○
Panoramio	○	○	○	○	○	○
Eventful	○	○	○	○	○	○
Last.fm	○	○	○	○	○	○
Twitter	○	○	○	○	○	○

List any other sources that you would be interested in for the purposes of the project

Please write your answer here:

Data integration and mining

How useful do you consider each of the following functionalities provided by Geostream? *

Please choose the appropriate response for each item:

	very poor	poor	fair	good	very good
mapping of source categories to a common category hierarchy	○	○	○	○	○

	very poor	poor	fair	good	very good
GUI for manual validation of category mappings	○	○	○	○	○
find duplicate entities across sources	○	○	○	○	○
GUI for manual validation of duplicate entities	○	○	○	○	○
identification of Regions of Interest per category or type of data	○	○	○	○	○
keyword search and browsing	○	○	○	○	○

How satisfying do you find each of these functionalities as currently implemented by Geostream? *

Please choose the appropriate response for each item:

	very poor	poor	fair	good	very good
mapping of source categories to a common category hierarchy	○	○	○	○	○
GUI for manual validation of category mappings	○	○	○	○	○
find duplicate entities across sources	○	○	○	○	○
GUI for manual validation of duplicate entities	○	○	○	○	○
identification of Regions of Interest per category or type of data	○	○	○	○	○
keyword search and browsing	○	○	○	○	○

List any other functionalities that you would like to be implemented in the context of the project

Please write your answer here:

Search and browsing

How would you rate the following? *

Please choose the appropriate response for each item:

	very poor	poor	fair	good	very good
display of search results	○	○	○	○	○
ranking/relevance of search results	○	○	○	○	○
amount of search results	○	○	○	○	○
quality of search results	○	○	○	○	○
linking of results to original sources	○	○	○	○	○

Dashboards and visualizations

How useful do you consider the following visualizations provided by Geostream? *

Please choose the appropriate response for each item:

	very poor	poor	fair	good	very good
dashboard for area selection and monitoring of data collection	○	○	○	○	○
statistics for category mappings computed per source	○	○	○	○	○
statistics for category distribution per source	○	○	○	○	○
statistics for portion of entities matched between sources	○	○	○	○	○
visualization of computed Regions of Interest	○	○	○	○	○

How satisfying do you find the following visualizations as currently implemented by Geostream? *

Please choose the appropriate response for each item:

	very poor	poor	fair	good	very good
dashboard for area selection and monitoring of data collection	○	○	○	○	○
statistics for category mappings computed per source	○	○	○	○	○
statistics for category distribution per source	○	○	○	○	○
statistics for portion of entities matched between sources	○	○	○	○	○
visualization of computed Regions of Interest	○	○	○	○	○

Additional comments

Please write your answer here:

Authoring tools

For the questions included in this section, please indicate your response *

Please choose the appropriate response for each item:

	very poor	poor	fair	good	very good
How would you rate the concept of authoring tools based on visits and trips?	○	○	○	○	○
How useful do you consider the text geocoding service?	○	○	○	○	○

How useful do you consider the following attributes for describing visits and trips: *

Please choose the appropriate response for each item:

	very poor	poor	fair	good	very good
title	○	○	○	○	○
description	○	○	○	○	○
comment	○	○	○	○	○
tags	○	○	○	○	○
rating	○	○	○	○	○
photo(s)	○	○	○	○	○
video(s)	○	○	○	○	○
arrival/departure time	○	○	○	○	○
price	○	○	○	○	○
type/purpose	○	○	○	○	○
transportation mode	○	○	○	○	○
route	○	○	○	○	○

What would be your preferred way to create a trip? *

Please choose **only one** of the following:

- ○ by geocoding a text document
- ○ via search and browsing, and then adding results to the trip
- ○ both
- ○ Other

Additional comments (e.g. additional/other attributes to describe visits and trips)

Please write your answer here:

Overall Assessment

Overall Assessment *

Please choose the appropriate response for each item:

	very poor	poor	fair	good	very good
Relevance of the provided material	○	○	○	○	○
Richness of features	○	○	○	○	○

	very poor	poor	fair	good	very good
Novelty of the concept and provided functionalities	○		○	○	○	○
Usability	○		○	○	○	○
Available documentation	○		○	○	○	○