ARE YOU BEING SERVED? INITIAL FINDINGS OF THE IMAGE2000 WEBSITE USER SURVEY

CZY JESTEŚ OBSŁUGIWANY? WSTĘPNE WNIOSKI Z ANALIZY ANKIETY UŻYTKOWNIKA STRONY INTERNETOWEJ IMAGE2000

Joanna Nowak, Conrad Bielski

Spatial Data Infrastructures Unit,
Institute for Environment and Sustainability,
European Commission Joint Research Centre
Brussels, Belgium

Keywords: Image2000 website, public spatial information, user experience, geo-portal
Słowa kluczowe: strona internetowa Image2000, publiczna informacja przestrzenna,
doświadczenie użytkownika, geoportal

Introduction

Portal statistics and user registration information from the Image2000 (http://image2000.jrc.it/) and EU SOIL (http://eusoils.jrc.it/) websites were used as a proxy to understand usage patterns of user access to freely available geo-spatial information and imagery (Nowak and Bielski, 2007). It was clear after that experiment that such basic website statistics were unable to provide the necessary information to understand user experiences and usage patterns. In order to get the required information, a short survey was prepared and launched in May 2008. The IMAGE2000 user experience and satisfaction survey was sent to all registered users of the Image2000 image portal in order to identify and profile user communities, understand how the available data benefits the user communities and profile user requirements for future data holdings and services development (e.g. IMAGE2006, JRC SDI).

In this paper we present the Image2000 user experience and satisfaction survey and some preliminary results. The survey responses provided us important information that is impossible to derive from server logs. Specifically, we were able to get a better idea of how users, especially EU citizens, discovered geo-portals and the availability of free spatial information. Such information is paramount for socio-economic analysis and normally difficult to acquire especially for specific user communities such as those interested in tourism expansion at the local and regional levels or those trying to heighten regional awareness.

The Image2000 User Experience and Satisfaction Survey

As mentioned above, user experiences cannot be modelled based on server statistics and therefore we went ahead and compiled a survey. The primary intention of the survey was to try and acquire quantitative data on the experience of registered users of the Image2000 website (i.e. those users who decided to download data and not just view the imagery) and their level of satisfaction both with the data accessible through the website and with navigating the website. Depending on the type of user, the number of questions that needed to be answered varied from 25 to 32.

Through the survey, we attempted to distinguish four main user groups based on their reason for accessing the data: those that accessed the data for research purposes, for educational purposes, for work or for profit purposes (excluding research or education) and for leisure purposes. While a person filling out the survey was able to choose all the user groups that applied, each user group also had some specific questions directly related to the purpose for which the Image2000 data was accessed. Finally, some general information about the user (which was not mandatory to finish the survey) was requested in order to acquire information of the location of the user, the level of education and approximate age.

The survey itself was only made available for approximately 5 weeks, from May 6 to June 7, 2008 and was open only to registered Image2000 website users. It was made possible based on the Free and Open Source Software (FOSS) Limesurvey (www.limesurvey.org). The total number of surveys sent out was 3170, however due to the age of some of the addresses (over three years old) it is possible that the email invitation did not reach all of the recipients. Furthermore, since no test of the registered email address was made during registration, it is also possible that many addresses were incorrectly saved in the database. Even with these issues, the total number of responses was 582, which is 18% (a minimum of 10% response rate was considered to be acceptable). Of these 582 surveys received, 432 or 74% were fully completed, 111 surveys were partially complete and 39 respondents did not agree with the privacy policy.

Preliminary Results of the Image2000 Survey

The following section presents some preliminary results of the Image2000 user experience and satisfaction survey. Unfortunately, not all results could be presented in this short paper.

User Communities

The first question that survey respondents had to answer was the following: What was the main purpose for accessing and using the Image2000 data? For this particular question, it was possible to choose more than a single answer thus generating many possible combinations of interest.

Based on the results presented in Figure 1, the largest user community, and making up almost half of the respondents accessing the Image2000 data was the research community. While research, education and work are typical communities that can take advantage of remotely sensed imagery, those accessing the data purely for leisure activities are much more

difficult to define and have a very wide and ingenious scope of uses. We were glad to see that a significant number of users (6%) accessed the registered area for downloading data for leisure activities. 32% of them just wanted to see what their home neighbourhood or their region looks like from above. Of those accessing the data for leisure activities, the majority (39%) accessed it for tourism or for something related to a sporting activity, while 12% of them were just curious.

Users that considered themselves to be a part of the education user group made up 28% of the people who accessed the Image2000 data with those accessing it for work taking up the remaining 18% of the overall users. It should be noted that within the education user group, the majority of respondents (58%) were students. The subjects that were taught using Image2000 as a teaching aid covered a wide range of themes where the largest group was for GIS (52%). The category 'other' included such themes as archaeology, landscape architecture, optics, biology, history and Health Information Systems (Fig. 2).

Of course, distinguishing between user communities is not that simple because a single respondent could have used the data for a variety of reasons. As shown in figure 1 (centre), those that used the data for research mainly used it for research (50%) but also for education, work and leisure. The same pattern goes for the work and leisure groups. However, it is interesting to note that for those using the data for education, most used it for both education and research (64%) rather than only for teaching and studying (28%) (Fig. 1 right side).

Table 1. General themes for which 'work' and/or 'research' type respondents accessed Image2000

	% responses
Environmental	27
General mapping (land surveying, GIS, cartography...)	22
Agriculture	11
Forestry	11
Urban planning	7
Natural resources exploration	7
Web portals (i.e. geo-portals)	3
Cultural heritage / history	3
Transportation	3
Public health	1.5
National or global security	1
Real estate	0.5
Telecommunications	0.5
Law	0.5
Entertainment or media	0.3
Insurance	0.2
Other themes not mentioned above	1.5
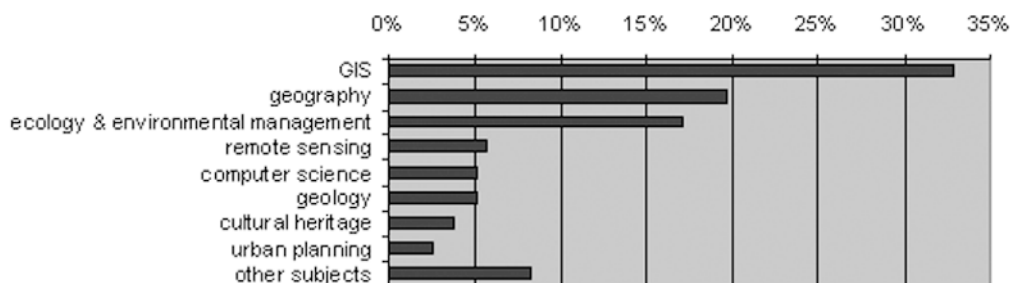
Figure 2. Subjects for which Image2000 data was used as a teaching aid
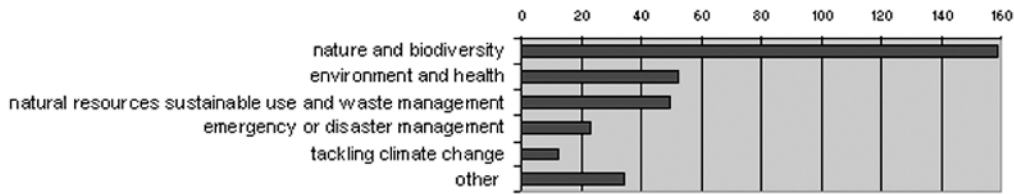
Figure 3. Respondents who chose «environmental application» could further specify their user group. The chart presents the number of respondents who chose a specific environmental theme

Those people that responded as using the data for work or research were able to choose from a provided list of general themes and choose all those that applied to their case. Even though users were able to choose multiple themes, the majority of respondents chose a maximum of two. Table 1 presents the results of this question and shows that most respondents used the data for environmental purposes, which is not very surprising. The respondents also mentioned such interesting themes as tele-medicine, animal health, ham radio, public awareness, archaeological research, geology, tectonics, tourism, resort development, socio-economic studies which were found under 'other'.

For respondents that chose environmental applications above, the choices were further subdivided (Fig. 3). A significant part of the 'environmental' users group (10%) did not find their themes in the pre-defined list and therefore their comments in the 'other' field include such interesting applications as social-environmental feedbacks, urban dynamics, landscape and land-use dynamics, morphotectonics, and hydrological modelling.

Perception of Availability of Free Spatial Data on the WWW

The perception of what people thought about their access to free spatial data on the Internet was also studied based on the Image2000 user survey. An amazing 48% of our respondents felt that a lot of spatial information is freely accessible, while 40% did not feel there was a lot of spatial information available for free. Those that did not know or had no opinion on the subject made up 12% of the respondents. What is very promising is the fact that 44% of the survey respondents also are of the opinion that the geospatial data that is available for free is useful to them.

This question of the survey also made available an area to provide the pros and cons of free geospatial information and it was used quite frequently. Generally, respondents contributed pros with respect to open access data policies. Interestingly, such data sources are the only available source of data for 23% of the respondents due to limited funds within an organisation such as Non-Governmental Organisations (NGO) or universities. Other important reasons for accessing freely available geospatial data were money savings (cost-effective) (12%), time savings and quick access (10%), and unfettered access (bypassing administrative bottlenecks) (7%). Of those that believe in free access to geospatial data, 19% consider it as a useful additional source of data (e.g. helps in planning and preparation), while 4% find the freedom of data choice is important to them.

From the teachers and students perspective, freely available spatial data enhances their curriculum (13%). Finally, 7% of respondents considered publicly available data as almost the only source of data for private use.

Only 15 cons were provided by respondents, and they can be summarised as: problems of low quality or accuracy, missing or insufficient metadata, problems with the consistency of the data, and interoperability problems. Unfortunately, no specific examples were provided for these cons and thus cannot be compared to other data sources.

Another interesting issue that was studied through the survey pertained to the manner in which people learned about the Image2000 website. Most survey respondents learned about the website through one of the European Commission websites (26%) with a colleague telling them about the website a close second (21%). What is very interesting is the fact that learning about the website from the media (3%) scored lowest from all possible choices except for the «other» choice (1%) which included three respondents saying they found out about the website from their teacher which could be considered also a friend or colleague.

So why did users of geospatial data seek out the Image2000 website? Figure 4 presents the number of respondents for each of the reasons provided on the survey. Based on the numbers, the fact that the data was free was the most important reason (369 responses). The fact that the data provided to users on the Image2000 website were appropriate for their needs and/or requirements came in second (215) with access to European high resolution satellite imagery a close third (205). The lowest number of respondents said that access to the Image2000 data provided them with the appropriate licensing agreement (52), which should be looked into further since it is an important issue when it comes to sharing data and information.
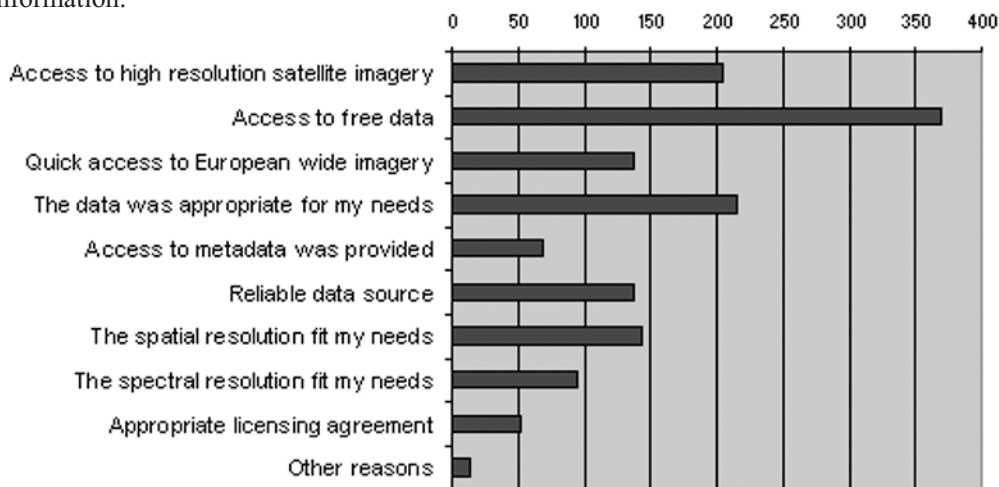

Figure 4. The reasons why Image2000 website users accessed the free imagery
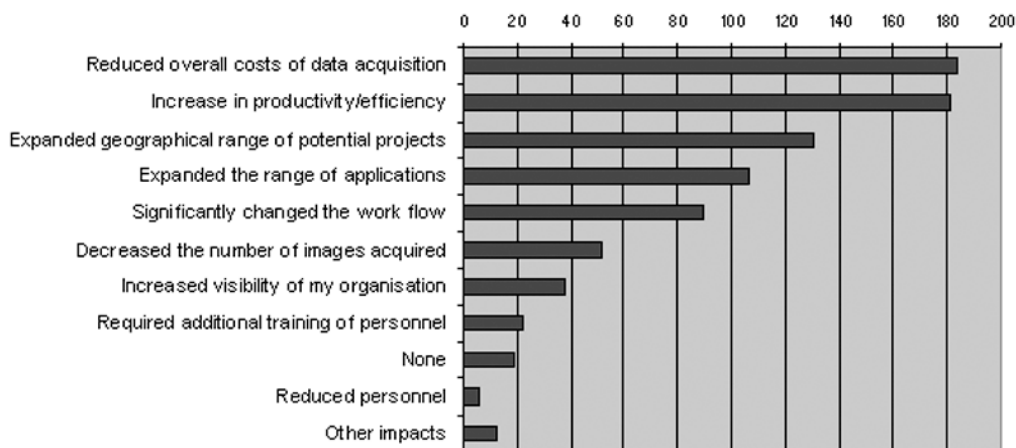
Impact of the Image2000 access to data

Knowing the impact of the Image2000 data on users is necessary to understand which services were most important and what could be made available in the future. Most users downloaded the Image2000 imagery for viewing it off-line (31% of all respondents). The second reason was to further process the downloaded imagery (25%). The third one was 'superimposition with other data or images' (21%). In general, these three actions go hand-in-hand and are reflected in the responses of the user survey. The lowest priority was given to printing the downloaded data (6%).

We were also interested if Image2000 data is a reference or additional data for its users. In general, the Image2000 website was perceived as a source of additional information (49%). Respondents chose reference data as having the lowest priority (19%) for using Image2000 data while the third possible choice 'input data for further image processing' was selected by 30% of respondents. Finally, 2% of respondents did not find a corresponding answer and chose 'other'.

One of the advantages of the Image2000 dataset is the fact that it covers all European Union countries. In question 11, the extent that users typically work with was surveyed and the two highest responses in descending order of importance were that their work covered a specific region (31%) and that their work covered several countries within Europe (26%). The third biggest group were the people working within a single country (23%). Users that accessed the Image2000 data worked the least with extents that covered small local extents and those that worked with areas that covered large extents of the globe with 10% each.

Finally, for 67% of respondents, the Image2000 data and services did have an impact. This statistic is broken down into those for which it had a profound effect (8%) and those for which it made some difference (59%) in their work or life. However, 1/3 of respondents said that it did not have an impact.

Related to the question of how the data affected peoples' work or life is what type of impact it had (Fig. 5). Most responded that the Image2000 imagery reduced their overall costs of data acquisition (184) and/or increased their productivity (181). This result is somewhat in line with the fact that most people accessed the website in order to retrieve data for free as mentioned above. A second grouping of impacts is the fact that users were able to expand the geographic range of their work (130) and also expanded the range of applications for which images were used (106). A considerable number of respondents (20%) stated that the access to the Image2000 data significantly changed their workflow. Note that access to the data also impacted the reduction of personnel in a small number of cases (6). Table 2 provides a listing of the eleven 'other' impacts survey respondents provided in this context.

Figure 5. The impact of access to the Image2000 data on work or life according to the users surveyed (by responses nb)

Table 2. "Other" ways in which access to the Image2000 imagery affected or improved users work

Allowed us to create river basin land-cover information, otherwise unavailable	Enabled qualitative assessment of image processing chains	Taught lessons on what has to be improved to arrive at a consistent Pan-European image layer
Test conformity of applications	Increased knowledge of territory	EU-wide reference geometry
Improved output graphics	Better coverage where other products were not available	Facilitated European-wide production of CLC2000
Useful for a specific project	Increased quality of work	

Discussion

A user community here is considered to be a group of individuals who use data or information in a similar manner or for similar goals. In this paper we mostly used the expression 'users' to denote people who registered on the Image2000 website to get direct access to the data set. 'Viewers' are all those people who came to the Image2000 website on purpose or by accident. They are also sometimes referred to as 'browsers' (Wang et al., 2004) and did not try to access the Image2000 data.

While it is obvious that most Image2000 website users can be considered as 'viewers' based on the site statistics, the users that registered wanted more. We assume this because the registered users were willing to register themselves which is generally considered a deterrent for most internet browsers since they need to leave personal information behind and access is not instantaneous.

Identification of the Image2000 website user communities was one of the main goals of our analysis. Little attention has been given to the research of user communities on the internet compared to researching internet communities that are viewers of related web pages (Murata, 2004).

The first experiment of our research on Image2000 usage patterns used portal statistics and user registration information as a data proxy for understanding user access to free geospatial information and imagery (Nowak and Bielski, 2007). Registered users are constantly being added and currently there are over 3200 records in the Image2000 registered user database which includes the following information: name, email address, organisation represented, and the purpose for which the data is to be downloaded. The problem with the current registration process however, is that it is based on free text answers. Consequently, the person registering can use any language, sometimes making it difficult to know the purpose of their registration, there is no verification of the submitted email address, and a person could decide to provide no information at all.

The main conclusion from that experiment was that such simple website statistics were unable to provide the necessary information to fully understand usage patterns, and in particular user experiences. Thus, the Image2000 user experience and satisfaction survey was prepared and sent to all registered users of the Image2000 image portal. The response was excellent with 582 of those invited answered the survey, which is a response rate of 18%.

An important question that needs to be better understood is whether the survey statistics can be extrapolated to better understand the Image2000 website and registration statistics? We believe that many survey respondents did not submit their opinions due to a language

barrier. The Image2000 user survey was only provided in English whereas approximately 50% of Image2000 registered website users provide their information in a language other than English. Another obstacle of the current user survey is the vocabulary used. While we attempted to avoid using «geospatial jargon», many expressions could not be avoided. The registered Image2000 users have a variety of experience levels and/or education and therefore could have encountered some difficulties, misunderstood some of the questions or even refrained from filling out the survey.

Our ultimate goal is to develop an automatic (or at least almost automatic) service to monitor and log usage patterns of users accessing publicly available geo-spatial data. The current data sets available to us, 'wide' user registration data and 'deep' user survey data, are valuable and needed in order to reach this goal. From a registration perspective, users cannot comment on data and services until after they have gotten access to the data and used it. Therefore, some feedback mechanism must also be put into place in order to record the entire user satisfaction story. Finally, the tracking of 'viewers' is also important but certainly more difficult because user identifiable information cannot be acquired.

With the user information gathered to this point, it may be possible to extrapolate «viewer» statistics based on 'user survey statistics'. If such an extrapolation would be possible, then it would also help in understanding viewer usage. Following this line of thinking, one would expect that the majority of viewers were also researchers. However, our gut feeling is that this is not the case because spikes in the viewers of the Image2000 website were usually found after the site was advertised somewhere.

Considering the millions of viewers of such geo-portals as Google Earth, Virtual Earth, etc., the Image2000 website pales in comparison. However, the purpose of the Image2000 website is significantly different and provides users with added value. Can the value of such a public asset be increased through an advertising campaign?

Preliminary Conclusions and Future Directions

The Image2000 website user survey was based on the findings from our initial experiment (Nowak and Bielski, 2007). One conclusion of that experiment was that a user satisfaction survey was needed to quantify user experience if one was to analyse the socio-economic impact of providing free access to spatial information. While we are aware of the fact that user satisfaction surveys are subjective, this survey did provide important information about who and for what reason Image2000 data was accessed and not just viewed.

In order to keep the survey simple and accurate, most questions provided a predefined set of answers (closed-ended question) because such data is more objective and easier to interpret. However, we are aware that it was impossible to categorise all responses. Therefore, certain free text comments were allowed which allowed respondents to better express their thoughts and provided a higher level of understanding of the users requirements.

Overall, the Image2000 user experience and satisfaction survey was a success and provided a plethora of new and interesting information about geo-spatial data users who accessed the Image2000 imagery directly. The goal was to have a minimum of 10% response rate and a rate of 18% was attained providing a statistically significant sample of the population of registered users.

The most important conclusion of this exercise is that it was clear that most users were satisfied with the Image2000 data they downloaded and benefited from this access is some manner. However, it is also clear that data and services can be improved.

Further analysis of the data is ongoing and there are several open issues that we want to deal with in the near future:

1. Profile user requirements for future data holdings and access to on-line services;
2. Extend user profile analysis in cooperation with sociologists to investigate socio-economic aspects of users interested in tourism expansion;
3. Support impact analysis of evolving SDI's around the world;
4. Fix quality issues of the data and services being provided.

At the end of this exercise, it is expected that it will be known why the registered Image2000 website users decided to access the data (and viewing was not sufficient) and whether the services satisfied their needs. If one understands the needs of the customer, then it is possible to provide them with better data and services.

References

Murata T., 2004: Discovery of User Communities from Web Audience Measurement Data. Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence table of contents, ISBN:0-7695-2100-2, pp. 673 – 676.

Nowak J., Bielski C., 2007: A First Look at Usage Patterns of Publicly Available Spatial Information [In:] Geographic Information Systems Conference and Exhibition 'GIS Odyssey 2007', 3th–7th September 2007, Šibenik, – Richness and Diversity of GIS. (Kerekovic,D., ed.), Zagreb, ISBN 83-918826-3-2 , ISBN 953-6129-31-0, pp. 37-42.

Wang Q., Makaroff D.J, Edwards H.K., 2004: Characterizing Customer Groups for an E-commerce Website [In:] Proc., ACM Conference on Electronic Commerce (ACMEC 2004), New York, NY, pp. 218-227.

Abstract

Recently, the Image2000 user experience and satisfaction survey was made in order to identify and profile user communities, understand how the available data benefits the user communities and profile user requirements for future data holdings and services development. In this paper we present initial results and specifically we researched user communities, the perception of the availability of free spatial data, and the impact of the Image2000 access to data. The survey responses provided us with important information that could not be derived from server logs. It was possible to get a better idea of how users, especially European citizens became informed about the presence of geo-portals and the availability of public spatial information. A discussion is also made on the differences between «users» and «viewers» especially pertaining to the development of an automatic system to monitor usage patterns of people accessing publicly available geospatial data. An important conclusion of this survey is that most registered users were satisfied with the Image2000 data they downloaded and benefited from this access in some manner.

Streszczenie

Niedawno stworzona została ankieta badająca doświadczenie i zadowolenie użytkowników strony internetowej Image2000. Jej celem jest wyznaczenie i określenie profilu społeczności użytkowników, poznanie sposobu, w jaki dostępne dane wpływają na społeczności użytkowników oraz określenie profilu potrzeb użytkownika dla przyszłego rozwoju zasobów danych i usług. W niniejszym artykule prezentujemy wstępne wyniki. W szczególności zbadaliśmy: społeczności użytkowników, postrzeganie

dostępności bezpłatnych danych przestrzennych oraz wpływ dostępu do danych Image2000. Wyniki ankiety dostarczyły nam ważnych informacji, których nie można było uzyskać z rejestrów zdarzeń serwera. Możliwe było nakreślenie lepszego obrazu tego, jak użytkownicy, a w szczególności obywatele Europy, zostali poinformowani o istnieniu geoportali i dostępności publicznej informacji przestrzennej. Przeprowadzona jest również dyskusja na temat różnic pomiędzy «użytkownikami» a «osobami przeglądającymi», szczególnie w odniesieniu do rozwoju automatycznego systemu monitorowania wzorów użytkowania przez osoby korzystające z dostępu do publicznie dostępnych danych geoprzestrzennych. Ważnym wnioskiem tej ankiety jest to, iż większość zarejestrowanych użytkowników była zadowolona z tych danych Image2000, które pobrali, i z których w jakiś sposób skorzystali.

dr inż. Joanna Nowak
joanna.nowak@jrc.it
http://sdi.jrc.it/
phone: +39 33 278 58 54
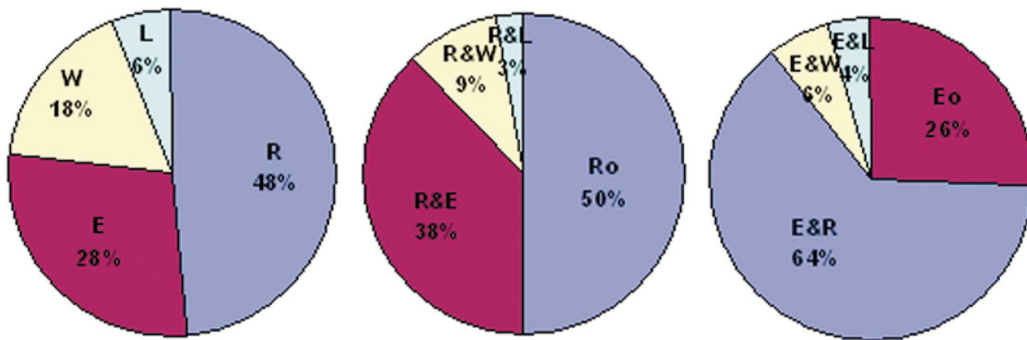
Figure 1. The purpose for accessing and downloading Image2000 data. The left pie chart presents the total percentage of respondents who chose the user communities possible: Research (R), Education (E), Work (W), or Leisure (L).

The chart in the centre and on the right show the mixed categories for the research and education user communities respectively (Ro – Research only, R&E or E&R – research&education, R&W – research &work, R&L – research&leisure, Eo – Education only, E&W – education &work, E&L – education &leisure)