The InterSecT project - mitigating barriers in the GIS usage in interdisciplinary archaeological research

Michał Rzeszewski1,3, Iwona Hildebrandt-Radke2
1,3 Faculty of Geographical and Geological Sciences, Adam Mickiewicz University
ul. Dziewielewski 27, 61-680 Poznan, Poland; e-mail: mrz@amu.edu.pl

Abstract

Tools of geographic information systems (GIS) are used in many aspects of archaeological research. Still, there are several barriers preventing GIS from being used on a larger scale and in more sophisticated ways than simple visualisations. A questionnaire survey was carried out among environmental archaeologists to identify those barriers and gather information about the general familiarity with GIS. The results suggest that the problem lies especially in the lack of funding, insufficient practical skills and inadequate data sources. A solution is proposed in the form of the Web-GIS framework that should mitigate some of the barriers. The software will also be designed to facilitate interdisciplinary research. This paper presents the results of the first stage of the InterSecT project.

Keywords: Web-GIS, interdisciplinary research, software, environmental archaeology

Introduction

Spatial information in environmental archaeology is used extensively from the very beginning of the research and on its every stage – from the choice of the study area, through the fieldwork and analysis, to the presentation of the results (Conolly & Lake 2006). Used within the geographic information systems, the location data can be used to merge the information from many fields of research, which would not be possible otherwise. However, interdisciplinarity often introduces heterogeneity in spatial data. This is hardly surprising, given the fact that the information comes from different sources not compatible in respect of the file formats, the level of detail, accuracy and precision as well as, most importantly, the perception of geographical space. The latter is particularly diverse – space is given unequal weight – e.g. many archaeologists place emphasis on the time rather than the location (Wagtendonk et al. 2009) and this is described in ontology specific to a given domain. This means that to use spatial information to its full potential it should be gathered, sorted and ordered in the most homogeneous and standard form possible.

Spatial data gathering, processing and presentation are the domain of geographic information systems (GIS) (Longley et al. 2006). They are used with great success in many aspects of everyday life and in the process of scientific knowledge production – either as a tool or the subject of research. Archaeology, environmental and classical alike, uses extensively GIS. In the next chapter the authors describe many of such uses and from a sheer number of the listed cases it can be concluded that GIS is hardly a novelty in archaeology.

One can therefore formulate an open question whether the potential of the GIS tools is fully utilised. It seems they are mainly used as a source of data or for the visualisation of the results, while much less often in the analysis, where they could be most beneficial. What is needed to change this situation is a better form of spatial communication where the spatial information becomes lingua franca of interdisciplinary research. This process is not without numerous barriers and it is worth pointing out that
many of them come from the nature of GIS itself; this issue will be discussed in more detail in one of the following chapters.

This paper presents the main assumptions, objectives and results of the first stage of the InterSecT project which aims at facilitating the use of geographical information systems in interdisciplinary research with the special case of environmental archaeology. It is worth mentioning that geographic information systems are perceived and described here as a set of tools and techniques used in the management of spatial data, rather than as a science on its own. This view, however limited (Wright et al. 1997), is much clearer in the given context and also popular among archaeologists (Conolly & Lake 2006).

GIS in environmental archaeology

GIS technology is extremely flexible and can be used in any spatial context. In archaeology GIS methods were first used at the turn of the 1970s and 1980s for statistical analyses, creation of databases and digital mapping. Due to the nature of GIS, the use of the environmental data, such as geographic coordinates as the criterion for organising spatial information, terrain models for analysis, etc., most of the applications of GIS in archaeology are related to the environment.

According to the area affected, the use of GIS in archaeology can be classified into two categories: inter-site and intra-site. The inter-site analyses are studies in the broader category of geographical settlement regions. Very often they result in creating spatial prognostic models (Gillings & Wheatley 2002). The second area of application are intra-site analyses of local archaeological sites, such as the analyses of the distribution of residual material at the site compared to its environmental parameters, such as slopes, humidity or fossil surface.

The environmental archaeology uses spatial analyses most widely. Initially, they were used mainly for the artistic visualisation of archaeological sites (Kvamme 1998, 1999; Van Leusen 2002). With time, GIS has become an analytical tool supported by statistical tests and software for the analyses of archaeological sites. Currently, GIS is very often applied for the spatial analyses of settlement distribution and the land use structure as well as the selected elements of the environment, at the regional level in terms of time (Llobera 2001; Winterbottom & Longham 2006). Another use of GIS is to study the relationship between the environmental parameters (variable environmental conditions) and settlement patterns. In most cases, the next step of such research is the development of predictive models in general, which in turn can be used in practice in search of archaeological sites (Kvamme 1989; Stančič & Veljanowski 2000; Fry et al. 2004; Howey 2007; Kay & Storm 2009). Such studies involve determining the correlation between the distribution of archaeological sites and the environmental variables in the region. The knowledge of environmental variables and the regression or discriminant functions helps determine the distribution of archaeological sites in the unexplored areas, their cartographic visualisation and testing the model on the basis of independent samples.

Other methods utilising GIS are the cost-surface and viewshed analyses. These methods are more often used in Europe than the US, although that is where the first publications originated. The literature indicates that the reason for this situation is a different approach to the study of the past in the Old and the New Continent. The American school of human-environment relations pays more attention to the characteristics of the physical environment, while the European studies often emphasise that human groups’ functioning is based on social determinants.
which are closely linked to the concept of territoriality (site catchment analysis) (Kvamme 1998, 1999).

In both methods the starting point of the analyses is a digital terrain model, which is used for identifying obstacles, i.e. "frictions", as part of the cost-surface. In most cases, these are physical parameters of the environment, such as watercourses, wetlands, steep slopes, etc. Using this method makes it possible to determine the old migration routes. The viewshed analysis determines the area seen from the place recognised as an indicator of territoriality, such as places of worship or burial mounds (Wheatley & Gillings 2000; Kay & Sly 2001; Lake & Woodman 2003; Ogburn 2006).

Today, an important aspect of the research is the four-dimensional GIS, which uses the fourth dimension - time, in addition to the coordinates x, y, z. Such analyses are not common as they require archaeological and environmental database for several time intervals. The most frequent problem is to obtain the paleogeographical data corresponding to the time of functioning of archaeological cultures. Most of these studies analyse the variability of settlement distributions and the land use changes over time (Spikins et al. 2002).

The countries with the largest-scale GIS usage include the US, UK, Australia, New Zealand, the Scandinavian countries (60% of the projects), and the Netherlands, followed by the Mediterranean countries, i.e. Italy, Spain and France. In the countries of Central and Eastern Europe the participation of GIS projects in the archaeological research is estimated at 15% (Djindjian 1998). Over time, however, the structure of participation in research projects has not changed much.

GIS is mainly used in archaeological heritage management (culture resource management - CRM). However, the projects analysing the relations between the settlement, landscape and culture within the broader framework of environmental archaeology, has only a bit smaller participation. Such projects are usually carried out in the US, Australia and the Scandinavian countries.

In Poland, despite the interest in the possibilities of GIS, especially among students and young archaeologists, there are few articles based on the GIS analyses. The method most often used is the digital terrain model which visualises archaeological sites (Dobrowolski et al. 2011; Zagórski 2009). More advanced studies include the analysis of prehistoric settlement preferences in relation to hydro-morphometric features of the site, which were obtained as a result of the transformation of the digital terrain model (Jasiewicz & Hildebrandt-Radke 2008 a, b, 2009). In addition to the settlement preferences, for the same set of the data settlement analyses were performed: density and trends in a number of time intervals (Jasiewicz & Hildebrandt-Radke 2011). However, the developments taking into account the viewshed analysis and the cost-surface analyses are at their initial phase (Jasiewicz & Makohonienko 2009).

This situation is mainly due to the fact that much of the documentation of archaeological research, coming both from the surface and the excavation studies, is still stored in the analogue form. This situation should gradually improve, as the construction and development of the geospatial database of monuments has been identified as one of the functions of the National Heritage Board of Poland. The new strategy of this institution implies moving away from the previous passive model in the administration of monuments towards the active management of the heritage. This is related to the implementation of the INSPIRE Directive in Poland, whose task will be primarily the collection and integration of the natural and geographical data (Kołodziej 2011; Banaszek 2010).
An example of activities in this regard is the computerisation of the data of the Polish Archaeological Record Project (Archeologiczne Zdjęcie Polski – AZP). The project is still in progress, criticised and recently converted into a spatial database. The previous actions led to the creation of a computer program AzpMax which was to be used to enter and analyse the AZP’s data (Prinke 1996, 1997, 1998, 2002). Practically, however, on a larger scale, it has not been realised. Today, it does not meet the criterion of a geospatial database. Therefore, there are local attempts to develop new programs for introducing and processing the AZP’s data as the GIS databases (Gawrysiak & Reder 2011).

Only getting through this phase of the research, which involves the digitisation of both the data and the results of archaeological research, will make it possible to take a full advantage of the geospatial environmental data. This is especially important as such data is constantly growing and gaining wider access via the Internet platforms.

Having archaeological and geospatial data allows for putting forward and solving scientific problems, using a variety of queries and, most importantly, allows for using the GIS methods and geostatistical analysis, which will result in a completely new quality of research results.

InterSecT project

The aim of the InterSecT (the acronym for the Interdisciplinary Scientific Tool) is to create a tool in the form of a software framework that will facilitate the utilisation of GIS tool in interdisciplinary research. The main goal is to allow scientists from various disciplines to use spatial information easily, regardless of their backgrounds and lack of geoinformation knowledge, without the need for extensive training and with the existing hardware and software infrastructure. The latter is particularly important in the initial stages of scientific collaboration, when funds are often limited. Therefore, it can be said that the InterSecT is trying to overcome two barriers in the GIS use: the software interoperability and the disciplinary interoperability.

This aims may be achieved by using the web-based geographic information systems (Web-GIS). They were developed in recent years and have become a viable alternative to more traditional desktop solutions (Dragićević 2004; Rzeszewski & Jasiewicz 2009). Most of the everyday tasks given to the traditional GIS packages can be achieved in the web environment (Anderson & Moreno-Sanchez 2003) alongside the more complicated geospatial analyses possible thanks to the Web Processing Service (Diaz et al. 2008; Meng et al. 2010). Moreover, some feature exclusive to Web-GIS are capable of overcoming software interoperability:

- Modularity: geographical information systems working in the web environment are not complete and closed solutions. They are rather described as conglomerates of multiple, often independent modules. This allows for customisation of the features provided by a given framework according to the users’ needs and the existing hardware infrastructure. But this kind of the software architecture is also entirely dependent on smooth cooperation between different software packages, which can be a non-trivial task.

- Open Source Software: Web-GIS is one of the branches of geographic information systems rooted in the Free and Open Source Software for Geoinformation (FOSS4G) movement and the available software packages belong to the most advanced ones.
By utilising the FOSS4G packages one can drastically decrease the initial costs required for the development of GIS system (Anderson & Moreno-Sanchez 2003). Thus, the spatial information can be distributed more easily and to a greater number of users. There is also a higher probability that a non-geographer will know at least one Open Source software item, the one accompanying any professional package. However, there are risks involved: many Open Source solutions do not include proper support or documentation, and the available training is limited. Few vendors offer commercial support, Ubuntu by Canonical being a notable exception.

• Browser based interface: Web-GIS software can be based around a thin-client concept, where the main software and data is located on a remote server and the users are accessing the required features by a web site. It is fairly safe to assume that the browser graphical interface is well known to an average computer user and thus the training period can be short. One disadvantage of this solution is that adding non-standard features often requires extensive programming effort. Also, developers must always pay additional attention to security issues.

The second barrier between disciplines is much harder to overcome, because it is the result of the ontological differences in the perception of space. Differently perceived phenomena are also differently described, which can easily lead to a misunderstanding. To avoid this, spatial information should be as syntactically homogeneous as possible. The transformation and conversion of the data streams are therefore necessary and should be carried out invisibly to most of the users to eliminate the human error factor. Only in this way spatial communication can mitigate disciplinary barriers.

The framework will be first tested and introduced in environmental archaeology because of a wide range of sciences involved in this field and its strong connection to geographical space. The project will be conducted in five phases:

1. Composing and conveying a questionnaire to be used to gain knowledge about the needs and expectations of prospective users. The form will be electronically distributed amidst members of the Polish Association of Environmental Archaeology (Stowarzyszenie Archeologii Środowiskowej - SAS).

2. Constructing a spatial database that will encompass most of the potential data.

3. Choosing the software and composing the framework; adding missing functionalities and interoperability routines.

4. Testing the project in a real-life application.

5. Documenting and distributing.

The first stage of the project has been completed. The results and conclusions are included beneath.

**Questionnaire survey results**

The first stage of the project was carried out from January to May 2012. It consisted of a questionnaire survey that was sent to the Polish environmental archaeologists. There were 39 responds, out of which 33 were used for further analysis. The rejected ones were either inadequately filled in or were from the people who had previous knowledge about the purpose of the research and it
was believed that this might have introduced bias. The overall number of the respondents was relatively small rather than a single person. The full content of the questionnaire is available at http://intersect-gis.pl/ankieta.

One of the main goals of the questionnaire was to assess the general knowledge about geographic information systems among the environmental archaeologists. Firstly, the respondents were asked to self-evaluate their GIS expertise using a three-grade scale: professional, good or superficial. Some of the further questions were designed to verify this evaluation by asking about more detailed aspects of GIS. Figure 1 cross references the self-evaluation and the number of known GIS software packages. The latter question was open and participants were allowed to list any number of names. As it can be seen in the figure, both questions correlate positively and strongly. The respondents declaring superficial GIS expertise listed no more than 1-2 packages, while those perceiving themselves as professionals were able to list at least 3-4 and mostly 5-6 different software names.
There were some cases where people who declared superficial knowledge were able to enumerate more than 3 names of packages. This may result from either wrong estimation of the possessed knowledge or confusion between knowledge and proficiency in software. The latter seems more probable because there were no opposite cases. Not only is the number of packages interesting but also their names. Three most popular ones were Quantum GIS (14), ArcGIS (13) and Saga GIS (7). Significantly less popular were Mapinfo (3), gvSIG(3) and AutoCAD (2), and only single respondents mentioned Geomedia, GeoMap, Global Mapper, GRASS, ILWIS, Microstation, PostGIS, Spatialite and Surfer. A visible popularity of the Open Source Software is not surprising, as it reflects the increasing use of this kind of software by research community at large (Steiniger & Bocher 2009). This also fares well with the assumptions of the InterSecT project, which will be entirely based on FOSS4G.

2). The correlation is similar – the greater the expertise the more frequent spatial data usage, but not nearly as strong. There are groups of people using GIS data very frequently and yet having only a superficial
knowledge about geographic information systems. Again, this may be the result of wrong self-estimation.

But there are other possible explanations. For example, this could be a signal that there are scientists who use spatial data in research or even just perceive the potential of using it and, at the same time, unable to fully utilise GIS tools due to existing barriers, lack of knowledge or skill. This would confirm the hypothesis that there is a need for GIS tools in disciplines other than geography or geoinformation, and there is a niche for projects like the one presented here.

In the next part of the questionnaire the respondents indicated most frequently used data types (Fig. 3). Our hypothesis was that the “simpler” types of data, like scanned maps and GPS points, would be used more often. This is in fact the case – topographic maps and measurement data are among the most frequently used, ranked at the first and third place accordingly. But the differences between the data types are almost negligible and statistically insignificant. Only six less respondents listed text data, ranked at the last place, than topographic maps. This results show that there is a wide variety of data types and formats that must be taken into consideration in the framework database design. The authors hoped for the distribution to be more skewed towards some particular type. The almost uniform shape means that technically it will be more difficult to encompass all the possible data inputs. It may be even necessary to narrow down the choices available to future users. Lack of any prevalent standard whatsoever signifies that it is important to provide some methods of translation.

The diverse data types are usually the result of equally diverse tasks performed using geographic information systems. The respondents were asked to rank 13 common functions provided by the GIS packages in three-grade scale: necessary, useful or unnecessary (Fig. 4). The choice of wording was deliberate, with the intention of obtaining as much information as possible. The two extreme choices required some previous experience, while the middle one was not so restrictive and therefore was probably more often selected in case of unfamiliar functions. Also, to mark the functions as unnecessary, one must have a complete knowledge of their GIS needs. The results show that only three functions were designated in this way by more than one person. Surprisingly, one of them was editing vector objects, which at the same time was described as necessary by the greatest number of respondents. The other two were digital terrain models and web services. Equally puzzling is that only 10 persons gave GPS data the highest grade. It had been expected this would be one of the more useful tools for archaeologists. For an explanation it would be helpful to look into the workflows used to gather the location data. Generally, it is clear that GIS in the environmental archaeology is used in a rather basic way: vector data editing, georeferencing, map production, classifications. Therefore, one of the goals of the proposed framework should be to introduce more advanced tools and better ways of utilising those already being used, for example, scanned and georeferenced maps could be replaced by remote data sources from web services like the WMS or WCS. This must be done with care however, to ensure that the simpler functions are still available.

The respondents also had an opportunity to include their own desired functionalities and more than half of the respondents (22) took it. From the detailed descriptions that were given it was concluded that most of the propositions can be classified in the domains of spatial analysis or better graphical interfaces that would allow for simpler workflows. The last part of the questionnaire dealt with the barriers...
perceived in using GIS. This was an open question so the answers were classified into four categories: funding and availability, lack of skill, quality and availability of data and lack of knowledge (Fig. 5). It was hardly surprising that almost half of the answers (15) suggested lack of sufficient funding. The respondents complained that it was equally difficult to gain access to professional GIS software and hardware infrastructure. This partly explains the aforementioned popularity of the Open Source Software.

Figure 3. Comparison of self-evaluated GIS expertise with the most frequently used data types
The second significant barrier is the lack of skill in using GIS software. The reasons for this are manifold. Obviously, environmental archaeologists came from many different backgrounds, where geographic information systems are not routinely taught and every skill must be learned by experience rather than formal training. But even among geographers, where the situation should be much better, this is still a relatively new subject. It was not until the late 1990s that GIS started playing a more than marginal role in geography (Kistowski 2001). This is also related to the first of the barriers - lack of funding severely limits the access to the GIS software and therefore impairs the GIS skills in the scientific community.

Two other barriers were more rarely mentioned but they are significant nonetheless. The first one is Quality and availability of data. Once more, it is strongly dependent upon finances. Many datasets essential to archaeology, like the detailed digital terrain models, are expensive and beyond the reach of many scholars in Poland. Thankfully, the situation is gradually improving due to the INSPIRE project and the evolving National Spatial Data Infrastructure. The initiatives, like the Polish Geoportal (http://geoportal.gov.pl), make the spatial datasets available for viewing and even downloading by the means of various web services (WMS, WFS). But from the comments received from the respondents and the importance they placed on the web services in the questionnaire, it can be concluded this source of
data is still not widely recognised. Therefore, one of the goals of the InterSecT should be to propagate the knowledge about the remote data sources.

The lack of knowledge about GIS is the last category of the perceived barriers that was identified. It is interrelated with the lack of skill but we chose to separate it to underline some additional issues. The respondents drew the attention to the problems with the general spatial literacy, e.g. lack of basic cartography skills needed for simple tasks like the transformation between the coordinate systems or creating a readable diagram or a map. This group also includes people who have never had contact with any GIS software.

Conclusions - improving GIS usage with Web-GIS

It is clear that Web-GIS can be used to mitigate barriers existing between the environmental archaeology and a wider use of GIS tools. The proposed framework will be entirely based on the FOSS4G software and therefore the initial costs will be minimal. Without a large data load all the components of Web-GIS can work relatively smoothly even on low-tech hardware. Therefore, it is possible for interested users to conduct preliminary test-runs on the existing computer architecture. This will reduce funding problems indicated as important for the majority of the respondents. Still, the knowledge and skill barriers remain. Therefore, the project goals were modified to combat this problem.

Basing on the results of the questionnaire it seems that the proposed framework should fulfil both an educational and practical role as it may be equally important to introduce certain tools as it is to provide easy access to them. This issue will be addressed twofold. Firstly, the installation and data integration will be streamlined to the bare minimum. The finished framework will be published as a downloadable, preconfigured Linux distribution that can be either installed on a separate server machine or as a virtual machine. In this way the users will not be required to install any software components on their own. Secondly, the detailed documentation with tutorials for basic tasks will be provided. The comments from the survey suggest that even a simple introduction to the data models and spatial analysis would be beneficial. This is of course not without costs. Simpler, clearer and better documented workflows require time that could be otherwise allocated for the development of more advanced geoprocessing tasks. This means that the final project will not deliver all the planned functionalities but hopefully it will be better suited for the needs of the potential end-users.

The only important barrier that will be partially addressed is the accessibility and quality of the data. The framework will include various types and formats of the exemplary data and this should introduce users to the many possible data sources, not only in the form of locally stored files but also to the remote services, like the WMS or WFS. But the quality of the currently available datasets and their accessibility are still not sufficient for many tasks. The situation is improving thanks to various institutional efforts, like http://geoportal.gov.pl or http://geoserwis.gdos.gov.pl, but the majority of the data is still in the analogue form. However, the authors hope that this project will make the use of Web-GIS for research more popular and, since it is relatively easy to publish data gathered in this way, it could lead to more resources being available in public domain.

The status of the project and all the data and software gathered so far are available at the project website http://intersect-gis.pl.

Acknowledgements

The authors would like to thank the questionnaire respondents for their time and
willingness to help and share their thoughts and comments. The present research has been financed by the Nationa Science Centre (Narodowe Centrum Nauki) as the project N N305 071540.

References


Longley, P.A. et al., 2006. GIS - Teoria i praktyka, Warszawa: PWN.


