A Report on the GeoInformatics

Just under two years ago, in a short article published in the April 2001 issue of GeoInformatics, the World Wide Links were introduced as a publicly and freely available service to readers of this magazine. It has proven to be very popular - with the database containing the links being accessed by 550 individual users per week - and also quite useful - as attested by numerous e-mail messages and comments received since its introduction. Besides which, its content has been expanded greatly. Currently it contains 4,425 separate entries classified into 113 categories and 360 sub-categories covering the fields of Surveying, Geodesy, Imaging, Photogrammetry, Remote Sensing, Mapping, Cartography and GIS. Thus it now contains nearly double the number of links (2,300) that it contained two years ago. Furthermore the links are now held in a properly structured database with the provision of a fully categorized index and a search engine to help enquirers reach the specific entries or groups of entries that they are looking for, in as short a time as possible. This structured database has taken the place of the much simpler list of headings and links that was presented to users when the project started.

By Prof. Gordon Petrie

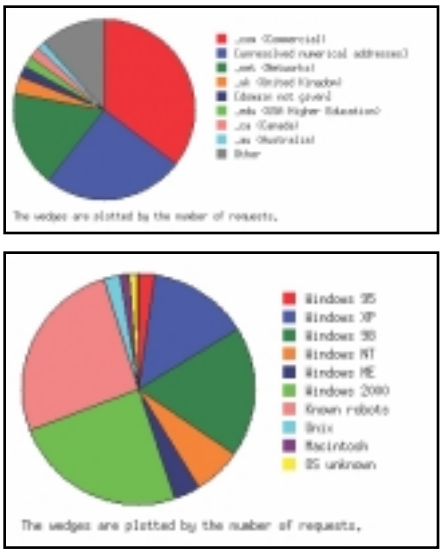
While I am responsible for the content of the database, the IT aspects of the project have been carried out by my youngest son, Colin, who is a graduate in electronics and computing science. Thus he has implemented the fully structured version of the database, using as its basis the Apache Web Server and the MySQL database engine, both of which run under the Linux OS and are available as open source software on the Web. He then wrote a series of scripts, first in Perl and most recently in PHP, linking these various components together. Afterwards he added a search engine to produce the database in the form that it stands at the moment. The high-speed connection of the Web Links database to the Internet is provided by NTL in Glasgow. However, the main entry points to the database remain the GeoInformatics Web site (http://www.geoinformatics.com/) and that of the University of Glasgow (http://www.geog.gla.ac.uk/~gpetrie/). The database is also indexed by the Google search engine on a regular basis. This ensures that the contents of the database are kept reasonably up-to-date within Google. Nowadays this source provides quite a large proportion of the enquiries received by the database.

Content of the Database

With regard to the content of the database, as reported previously in the April 2001 issue, the biggest single category remains University & College Departments & Institutions. Overall this category contains 732 entries, of which no less than 382 are links to American institutions. These figures compare with the overall total of 375 entries in this category, including those for 138 American institutions, mentioned in the original April 2001 article. Obviously education and research in the geoinformatics field is a thriving activity and so is the growth in the number of relevant Web sites and in our ability to discover and keep track of them. The second largest group remains the Surveying, Photogrammetric & Mapping Service Providers. Currently there are 284 of these in total (against 211 entries previously), of which no less than 140 entries are from the U.S.A. (as compared with 95 American entries previously). The next three categories have all been expanded considerably. Thus the Digital

Mapping & GIS Software Suppliers category still retains its third place, but the number of links now total 240 (as compared with 104 before), of which 123 are based in the

World Wide Links Project



These two diagrams are based on the statistics for all the queries made over the three-month period October to December 2002, showing (a) the domain addresses; and (b) the operating systems being used to access the GeoInformatics World Wide Links database.

U.S.A. However the current fourth and fifth placed categories are quite different to those occupying these positions in the original introductory article. In fourth place is the GIS Consultants & Solution Providers category which has 217 entries (of which 113 are from the U.S.A.). Occupying fifth place is the Geographic, Cartographic & Spatial Data Providers category, which now contains 207 links (including 122 American suppliers). Much expanded, but still quite some way down the table are the entries for Airborne Laser Scanning. Currently these comprise 46 service providers; 13 U.S. government agencies; 9 university research groups; 8 system suppliers; and 3 general entries. Readers may also be intrigued by the number of entries - 4 sub-categories and a total of 31 links - now populating the category Balloon, Blimp & Kite Imaging & Photogrammetry. If readers know of missing links in any of the 113 categories contained in the database, then please do let me know about them.

Currency of the Database

As mentioned in the previous article from April 2001, the maintenance and updating of the database is a major issue, requiring a considerable amount of time for its implementation. New entries are found mainly through the detailed inspection of other specialist databases and the reading of articles, news items and adverts in trade magazines and newsletters. Besides which, some people have sent me e-mail messages asking if their Web sites could be added to the database. With regard to the existing entries in the database, with nearly four and a half thousand links needing to be checked, it would be a full-time job if this task was to be done manually. However, once again, Colin has eased the task by writing a program that checks through all of the links automatically once per week. Thus, every Sunday morning, I am presented with a list of all those links that were not responding at the time when the check was carried out overnight on Saturday.

Maintenance of the Database

However, even with this assistance, an average of one and a half days has to be spent every week on the maintenance of the database. Quite a lot of time has to be spent in trying to find out what has happened to each inactive link that has been flagged by the checking program. Has the company concerned been taken over or gone out of business and its Web site closed? With small companies, has their Web site been moved to a new Internet service provider (ISP) as happens quite frequently? Or was it simply the case that the Web site was not responding at the time of the check? It really is a very dynamic situation and one that needs constant attention. Most frustrating of all is the situation that occurs with some of the largest companies and government organizations which apparently have specialists who are dedicated to looking after their often extensive Web sites. These people help to justify their employment by continually tinkering with their site. Frequent tiny changes are made to the URL or to its suffix (as well as to its content), apparently without any realization or understanding of the impact that this has on Web users' personal bookmark lists or on reference databases or directories such as ours! Nor is there any attempt by these people to provide links between the old and new versions of the URL. The failure to do so is most frustrating and time consuming. In these situations, Google is an absolute godsend in trying to resolve the matter! But it still takes up a lot of time to sort out these matters.

Users and Usage of the Database

Besides the list of inactive or unresolved Web links, once per week (also on Sunday morning), I am presented with a full log of the usage during the previous week. Again this information is provided by yet another of my son's programs. This list is always very interesting and useful. Users of the database come from every part of the globe and not just Europe where most of GeoInformatics readers are based. However, as one might expect, there are very few users from Africa. More surprisingly perhaps, there are comparatively few users from countries in the Middle East where facilities are better and there are not too many users either from Latin America. By contrast, Asian users are much more numerous, coming mainly from India, the south-eastern Asian countries and East Asia (especially Japan, Korea and Taiwan). The vast majority of enquirers will only inspect the one or two categories that interest them and then activate the specific links that will give them the information that they are looking for. Most users will activate fewer than 5 or 10 categories or links during a single session. But, every week, a few users will interrogate the database intensively. For example, this week, a single user from a university in Thailand activated 220 categories or links contained in the database. However this was completely overshadowed by another individual user from Israel who activated 821 individual links while logged on to the database - which is quite unbelievable! Other interesting points are that there are apparently many enquiries from home users, especially from North America. Also some parts of the database are copied regularly by one or two other universities [- even the misspellings have been copied! -] and these then appear on their Web sites, presumably for the benefit of their own students.

Looking forward to receiving still more entries for the database from our readers! Happy New Year and good browsing!

Professor G. Petrie (g.petrie@geog.gla.ac.uk), Department of Geography & Topographic Science, University of Glasgow,Glasgow, G12 8QQ, Scotland, U.K
Web Pages: www.geog.gla.ac.uk/~gpetrie/