

Measuring New Media

Vasja Vehovar, Luka Kogovšek

University of Ljubljana, Faculty of Social Sciences, Kardeljeva pl.5, Ljubljana, Slovenia

Phone: +386 1 5805 297, Fax: +386 1 5805 101

info@ris.org

Keywords: Web, Survey, Media, Log analysis

Edited by: Matjaž Gams

Received: July 10, 2000

Revised: October 1, 2000

Accepted: November 15, 2000

After some years of uncertainty, confusion and high expectations the Web is slowly becoming a real alternative to the conventional media. Within this framework, the measurement of the Web activities is an extremely important, although difficult task. However, the overall technological trends offer a radical solution (e.g. Pc-meter) to the problem of the global standardized meter. Even more, it seems that the globalisation of the content also follows the path of the globalized measurements.

1 Introduction

The Web emerged rapidly as a new media, in only few years. However, despite its extreme public attention, the actual Web media consumption is still relatively small. Average daily-spent amount of time on the Web (calculated on the total population) is still measured in minutes, while classic media (i.e. TV, radio) can be measured in hours.

As for Fall 2000, in Slovenia the average daily Internet media consumption (including the non-users) is only 3 minutes (RIS, 2000). On the other hand, the TV is watched around two hours, and radio is listened over three hours daily. Of course, the numbers for Slovenian Web consumption are relatively low, particularly because in year 2000 only one third of the time Slovenians spent on the Web was located at Slovenian sites. However even in the US the average daily consumption on the Web stands around 16 minutes (9 minutes for home usage and 7 for business usage). Here, again, we calculate this approximation as the average across the active population (Nielsen Net-Ratings, 2000).

Accordingly, the advertising expenditures already allocate few percents of their budgets to the Web. Of course, this is the case only in the most developed countries, whereas in small markets (i.e. Slovenia) this share is about ten times smaller. Nevertheless, the advertising budgets closely follow the increased Web usage, so huge growth in Web advertising is expected in next years. This is the case also for the countries such as Czech and Slovenia, where the Web advertising is rapidly growing, not to mention the US where the projections estimate that in few years the percentage of advertising budgets allocated to the Web will surpass 10% of all advertising expenditures (IAB, 2000).

Simultaneously, the need for specific Web measurements also constantly increases. It is an extremely convenient fact that the Web itself offers very effective way of measuring users' activities. As Internet usage is basically one-person-to-one-device activity it is in principle possible to automatically track very detailed activities. Today, we can thus observe trends of globalisation and centralization in Web activities' measurements. With so-called PC-meter or Net-Meter we are basically ending up with only few players (i.e. MediaMatrix, Nielsen NetRatings, NetValue) that can maintain global panels of the Internet users who accepted the installation of the measurement software into their PC. In year 2000, the largest players already claim that with samples of standardized national PC-panels they already cover more than 80% of Web users in the world.

Of course, the global PC-meter-based measurement has also many weaknesses. The most problematic is the low resolution for smaller sites, which can hardly get enough traffic on top level (introduction pages), while it is practically impossible to estimate the number of users on subsidiary pages within any level of rational costs. Another disadvantage stems from difficulties with business usage, due to number of restrictions companies impose to such measurement, what results in low coverage. In addition, the response rates tend to be small in these types of survey panels.

Despite high costs for software solutions and expensive methodological developments, further expansion of these measurements is expected. In the near future there will be thus a possibility to observe, moment by moment, the representative picture of Web activities of the total world population active on the Web. Of course, we can also expect the extension and unification of these

measurements to other devices than PC (e.g. TV, mobile).

Despite these global trends, the Web measurement in small audiences, either geographical limited local communities or language specific populations, is still problematic. It seems that, here, the measurements will remain local also for next few years, as the revenues from such measurement cannot cover the costs of maintaining the appropriate panels. The main problem in this aspect is the fact that, as opposed to TV, radio or printed media, Web is much more diversified with thousands of Web locations. Accordingly, we need much larger samples to capture these activities and consequently the corresponding costs are relatively large.

2 Measuring the Web in Slovenia

As for now, the global companies are not yet interested in performing measurements for small audiences, because there is a clear lack of the critical mass to enable profitability of these business activities. There are, of course, also other reasons for a relatively low priority Slovenia has in the future plans of MediaMetrix and Nielsen NetRatings, because Ireland and Singapore are already included, despite their relatively smallness.

However, besides global on-line PC-meters applied to the sample panels of Web users, the Web measurement can be performed also in other directions. As for now, in Slovenia we collect the data from the three sources presented in next subsections.

Recall survey measurements: user-centric approach

With user-centric approach we survey the Internet users as opposed to other approaches (log analysis) where units of measurement are basically the requests on servers. In the absence of the panel with the PC-installed programs (which also belongs to user-centric type of measurements), we can roughly estimate the monthly visits with recall techniques in telephone surveys. Such a data arise from a monthly RIS (Research on Internet in Slovenia) telephone survey where certain pages are asked to be evaluated by respondents. Of course, there do exist methodological problems with this methodology, however, the information is generally stable and consistent.

We should stress that the monthly reach of AltaVista and Yahoo still dominates among Slovenian users, with around 150,000 persons visiting them on a monthly basis. However the gap to Slovenian sites is rapidly narrowing each year.

Table 3 presents the list of top Web sites from the beginning of the year 2000. We should recall that there were around 280,000 users at that time.

Site	Reach
Matkurja.com	90,000
Tis.telekom.si	70,000
Arnes.si	65,000
Mobitel.si	50,000
Siol.net	50,000

Table 1: Monthly audience for top five domestic Web sites in Slovenia 2000/I

By the end of the year 2000 the above figures generally increased by 30-40%.

Log analysis - server analysis: Web-centric approach

Different results were obtained from the log analysis. The log file is a tracking that is recorded with every visit at certain Web site. The log files have been analysed within RIS project for 40 most visited sites in Slovenia at the beginning of the 2000. There does exist a complex methodological environment for understanding these results, as the standard results of Webtrends analysis gave different figures for a monthly number of user sessions.

We should keep in mind that each user could, of course, produce several sessions within a day or a month. A session is defined as a visit to a Web site by a single visitor, which is performed without a certain break (usually 30 minutes). Of course, this type of measurement is extremely vulnerable towards faking, specific technological settings at the server, and also to the patterns of visitation that is exhibited by the visitors. Certain sites have a lot of daily visitors, while other have majority of monthly visitors.

Site	Sessions
Siol	398,285
Slowwwenia	278,545
Mobitel	197,170
Arnes	167,623

Table 2: Top five sites according to monthly numbers of users session in Slovenia 2000/I

It is thus not surprising that another statistics, number of hosts, gave rather different results. The host is a device, usually the PC, with a specific IP number. The IP number thus tells from which computer the access was performed. The picture of top sites in this aspect is rather different:

Site	Distinct IP
CVI-sigov	28,794
Arnes	27,344
Tis.telekom.si	21,532
RCUM	21,098
Siol	13,243

Table 3: Top sites according to monthly number of distinct IPs in Slovenia 2000/1

Obviously, there exist certain ambiguities about these numbers too, however, it is very useful to expose them in a comparative manner.

The differences in measuring IP numbers as opposed to sessions well illustrates the problems of log analysis and the ambiguity when these data are used as a measure. On the other hand the gap compared to the telephone results reveal additional discrepancies, although the correlation was found to be high ($R=0.8$) between the IP measures and telephone results.

The difference between users centric and server centric statistics are common also with PC-meter technology approach. A well-known reconciliation study (MacEvoy, Kalyanam, 2000) reveals that the two types of data coincide only in the ranking of the first three sites, while with others considerable differences appear. Numerous methodological problems contribute to this gap and, of course, we will not address them here, but only repeat the general caution towards the usage of any type of this data.

We can thus conclude that, as for now, serious methodological problems exist in the field of measuring Web activities. However, the intensive work may soon remove the basic obstacles in this methodology, since it is only few years old and is thus lacking both, the experience and the research.

Ad-server analysis: ad-server centric approach

The measurement that is located at the server where the advertisements are launched is technologically much more precise than the previous demonstrated standardized log analysis. Here, the organization that buy certain advertising space at the publisher Web sites pushes the advertisement to the users, so they have a centralized opportunity to measure the visits in an interactive and standardised way. However, this statistics still suffers from all deficiencies of log analysis. In addition, this approach is limited only to the pages that are included in a certain syndicated Web advertising pool.

3 Conclusion

At the beginning of the millennium we can observe the Web activities' measurements being in a sharp transition towards a global measurement. Of course, it will take few more years until the standards will be widely accepted. In addition to this global statistics, there will perhaps remain a certain space also for areas, where the global measurements cannot be performed. This is somehow similar to the TV audience measurement or radio measurement for small areas with few hundreds of thousands consumers, where the national-oriented surveys cannot provide representative data. Therefore, alternative and local surveys will be still preformed to provide the necessary data.

Today, the global (PC-meter based) measurements of Web activities already cover a large majority of World Web users and we can expect that the coverage will further increase. It is thus already possible almost instantly to measure certain significant global reaction to specific news, events or new sites that appeared on the Web. Such a high coverage of centralized global measurement is unique to the Web and cannot be compared to any other media.

However, we should add that it is not only the measurement that is becoming much more global with the Web than with any other media. It is also the content that is much more global with the Web than with other media too. Even a brief look at the top ranked sites (Nielsen Netratings, 2000) suffices to confirm that almost in all countries we find Yahoo, AltaVista, Excite, Microsoft, and MSN among the top five or top ten sites. Similarly, the Amazon.com is appearing among top shopping sites all over the world. It is not surprising that it is also the top-shopping site for Slovenian users. Such a globalisation is not at all the case with any other media (i.e. TV, radio, newspapers, magazines).

The Web content is thus becoming also extremely centralized. Of course, on the other hand, a massive stream of specific sites with narrow audience appears, too. However, the software, investment and promotion needed to establish and maintain a highly-visited Web site demand extreme resources, so we can expect that only the largest players will be able to keep the race of offering best information, technology, design and promotion. In a certain aspect we could even say that the global measurement itself also promote the globalisation of the content, as only the sites that appear in global top rankings can receive larger stream of advertisements what enables their further development and promotion.

In this aspect things are changing very rapidly. Only two years ago, when preparing a chapter for the monograph *Cyber-imperialism* (Ebo, 2000) we projected that the global English-language based sites will keep the lead also in small countries. At that time the visitation of

Slovenian Web sites was way behind the English-language well-known Web sites (e.g. Yahoo, Altavista). It thus seemed that Slovenian companies would have to locate and pay the advertisement at these foreign Web sites to target the Slovenian consumers. Eventually, that did not happen. There were no Slovenian advertisements on Yahoo or AltaVista yet. This was partially due to the fact that these sites were losing their advantage as more and more Slovenian users without English knowledge appeared, but also due to the increased Slovenian content. The cheaper home-based advertising and the lack of competent comparison also strongly contributed to this fact.

Instead, however, the beginning of another much more radical trend can be observed. It seems that in near future the global companies will – directly or indirectly – perform not only the advertising but all Web-related operations: measuring the Web site visits, distributing the global and local (e.g. Slovenian) Web advertisement, and at the final stage they will also purchase all the worthy top Slovenian Web sites.

Thus, the trends in technology and the very nature of the Web strongly stimulate the globalisation of the industries linked to the Web, being it the advertising, the measurement or the content. It seems, however, that the measurement is the first activity that will become completely globalized. Of course, with globalisation, we mean both, the covering of the whole world Web

population as well as a uniform methodology that provides an instant global insight. However, the globalisation also means that only few private but global companies will own this whole of this specific industry.

4 References

- [1] Ebo, B (2000): *Cyber-imperialism : Global Relations in the New Electronic Frontier*, Northwest Publishing Group.
- [2] Internet Advertising Bureau (2000). URL: www.iab.net.
- [3] Kogovšek L: Analiza logov – jim lahko verjamemo, *Marketing Magazin*, 2000/junij.
- [4] Nielsen Net Ratings, 2000. URL: www.nielsen-netratings.com
- [5] MacEvoy, B, Kalyanam, K. (2000): Data reconciliation, reducing discrepancies in audience estimates from web survey and online panels. ESOMAR Online Media Measurement Conference and Exhibition, Paris, 2000.
- [6] Research on Internet in Slovenia (2000). URL: www.ris.org.
- [7] Vehovar V. (2000): Je slovenski trg za Internet premajhen? 3. marketinška konferenca. URL: <http://www.ris.org/si/ris2000/novice/20000619.htm>