

Chasing customers through the internet maze

When business started sniffing the potential of the worldwide web, few things excited it more than the discovery that it was so measurable.

Visitors left trails that showed where they had been, when and for how long; sometimes they even said who they were. For marketers, advertisers and publishers, this was a bone with an intoxicating smell.

But for many years, the bone remained buried. Only media publishers, needing data to keep advertisers happy, extracted much goodness.

E-commerce operators worried at it, but the relative crudeness of the tools made it hard to get much satisfaction. Others were defeated by the sheer quantity of data.

Suddenly however the bone seems worth the effort. "It's a hot topic now," says Ashley Friedlein of E-consultancy.com, who has written a report on website measurement and analytics. The main driver is the need to keep the finance director happy.

"The head of e-commerce is having to justify what he and his team are doing," he says. Ian Morgan, head of channel management for Barclays Retail Services, part of the UK banking group, agrees: "Everything we do has to be proven in terms of a robust business case."

Two other trends are fueling the industry. First, the increase in online marketing and advertising requires software that can track its effectiveness accurately. Second, multi-channel strategies mean that revenue measurement no longer reflects the true worth of a site.

E-commerce managers are using web measurement tools to prove the value they are adding to the core business," Mr Friedlein says.

The supply side is playing its part, with vendors providing greater sophistication and/or lower costs. The big shift in the last year or two has been towards web analytics - seeing how visitors behave within a site, rather than just counting them.

"People have historically treated websites like television channels - measuring how many people have tuned in," says Ian Thomas, marketing director of WebAbacus, a UK-based website performance analysis company.

"Now they tend to treat them like department stores - different individuals and groups come from different places through different doors, all with different issues and objectives."

The trick is to use data to work out who these groups are and how they behave. "There are still too many companies who just measure visitors and return on investment," says Dermot O'Mahony, a web development consultant in the online travel industry. "With just a little more analysis of key user behaviours on site, both can be dramatically

increased."

Although a few companies - notably NetIQ (owner of WebTrends), Nedstat and RedSheriff - are well-established, the industry is characterised by small companies.

There is a great deal of innovation, and a massive choice of commercial models. You can pay \$100,000 plus \$15,000 a year for a top of the range service from US-based SteelTorch; but you can also download a tool from Webalizer for nothing. The choice is impressive, but also bewildering: the trick is to identify your organisation's needs, and see which vendors serve them best.

E-businesses: Companies selling online need to see how well the site channels

visitors to the check-out, where they are dropping out, and how well promotions are performing.

Ideally, they need this information in real time. This requires a huge analysis effort, sorting the wheat from the chaff and making it digestible.

"You need to understand the 5 per cent who transacted and the 10 per cent who almost did," says John Woods, managing director of UK-based Site Intelligence. "You're not really interested in the other 85 per cent." (This is because e-commerce sites have found that ensuring online purchasers complete a transaction is generally a better way to increase ROI than attracting completely new customers.)

Top-end software is all-singing and all-dancing. SteelTorch works most effectively when integrated into application servers from the likes of BroadVision and ATG.

"We can spot that a visitor has done something and export the data to a personalisation agent so it takes action," says Andrew Stevens, UK country manager. It could, for example, be used to send an e-mail to everyone who has put more than \$50 worth of goods into a basket then abandoned the purchase, offering an incentive to complete.

Other vendors specialise in different areas, and in some cases offer complementary products. Clickstream and Speed-trap have

patented technology at the data collection level (See panel below), while others are strong on analysis and visualisation.

Site Intelligence generates screens that look like planetary systems: the bigger the planet, the more visits, and by lining up any series of planets (pages), unusual drop-offs can be spotted easily. WebAbacus offers three dimensional visualisation of paths, while Clicktracks uses the client's web page as a background for displaying information.

Boots, the UK retail chemist, started selling products online in March 2001, first as wellbeing.com, now as boots.com. For the first 15 months it had a simple package, but for the last year has used Site Intelligence.

"We could see before how many people were coming to the home page but we couldn't see where they were going next," says Carolyn Park, head of technology for boots.com. "We discovered using Site Intelligence that the left hand directory got far more clicks than we thought, while the images got fewer. As a result we are working on a redesign of the

home page.”

Ms Park has a data analyst working within her group. “We are always looking for those ‘Aha!’ moments,” she says. “But on its own the tool will give you only the obvious ones.” It can delve into the raw numbers and produce almost any report; but a human being is still needed to provide “true insight as opposed to just information.”

Publishers: Advertising-driven sites need accurate, auditable figures, and as much visitor profiling as possible. Content management systems are likely to include measurement tools, and registration is used to build profiles.

But the accuracy of Clickstream is attractive, not least because it spots activity that would otherwise be missed. Publishers may also want comparative numbers.

Hitwise uses data from internet service providers to give its 500 clients daily reports on their market share - in visitors - against their competitors. It is thus closer to being a market research than a usability tool. “As the internet population starts to mirror overall

population, people are using it to monitor the changing preferences of the public,” says Simon Chamberlain, UK general manager.

Information providers and “brochureware” sites: People responsible for sites that do not generate direct revenue will have a tough job justifying large expenditure on analytical software. Fortunately, they do not need to.

Barclays chose Sawmill, which starts at \$99 for a single licence and “is already providing high quality material on how we can improve,” says Mr Morgan of Barclays Retail Services. The tool is used to “monitor journeys, find dead ends, spot disproportionate drop-outs.”

A number of issues are common to all sectors. One is integration. Mr Friedlein says that organisations that already have management information systems need to consider whether they need additional software. On the one hand these systems offer one view regardless of the touchpoint. On the other the web has certain characteristics, such as path analysis, that require skill to analyse and display. “Specialist tools

do things the big offline tools can’t do,” he says.

Then there is the “will they be here next year?” problem. Smaller companies often have a technical edge, but they are more likely to be bought than buy. Mergers, takeovers and collapses are likely, but there will also be alliances between companies that appear to compete but are in fact complementary - data collection and data analysis specialists, for example.

Finally, privacy could rear its head. Although companies are more interested in trends than individuals, analytical software can easily be seen as “spyware”. “We are acutely aware of data protection issues,” says Ms Park at Boots.com. European Union legislation on data mining now being ratified originally sought to make all cookies - messages sent from web servers to individual users’ browsers to identify them - “opt-in”. That has been watered down, but the legal and technical trains are moving in only one direction. “We’re in the hands of legislation and Microsoft,” says Mr Thomas at WebAbacus. “It is a bit of a worry”.

TRACKING THE TRENDS

Many methods of measurement

A wide range of technology is available for data digging on websites. Here is a summary, from the least to the most sophisticated:

Logfiles: Every web server generates logfiles - a record of all the activity on a site over a particular period. It will include requests for each page, broken down by date and time of day, internet protocol (IP) addresses of visitors, sites from which visitors came, computer platforms they used, and more.

Pros: An easily-accessible source of vast quantities of data. Useful for tracking trends, and also for comparing traffic between sites

service providers to calculate “market share”).

Cons: Logs tend to be cumbersome, not least because they pick up the automated traffic of “bots” scouring the web. Inaccurate for three other reasons. Most ISPs circulate the same IP address around different users, so it is impossible to equate an IP address with an individual. Proxy servers allocate one IP address to a number of people. And logs do not spot pages drawn from the cache (for example by hitting the back button on the browser).

Tagging: The site’s owner puts a small piece of code on each page. When a visitor looks at the page, the code sends a call to another

server, which logs it. It may also include a plain English description of the tagged page.

Pros: Greatly simplifies the logfiles, both by reducing their size and listing pages by description rather than address.

Cons: Burdensome to tag every page. Does not in itself solve the IP address, proxy or caching problems.

Tagging with cookies: The tag is used to set a cookie on a visitor’s hard drive, so that a particular computer, rather than IP address, is logged. Cookies can be used to send information either to the website owner’s server (first party cookie), or to a server run by the analysis

company (third party).

Pros: Greatly increases accuracy, by solving the dynamic IP and proxy problems. Returning visitors are recognised and logged as the same person (or at least user of the same computer).

Cons: Will still be confused by people using caches (eg by hitting the Back button). There is also potential for data protection problems. Microsoft is already making it easier for users to block third party cookies on Internet Explorer 6.

Tagging with advanced cookies: Some vendors send small but sophisticated programs to users’ browsers. Speed-trap sends a Java applet that can follow a

mouse movement across the screen, and replay it. Clickstream’s cookies incorporate Javascript that records every keystroke made by the user, including backspacing and offline browsing. It stores this information and sends it back when it sees a clear path to the server.

Pros: Clickstream claims to be the only vendor to provide 100 per cent accurate results, while Speed-trap offers a remarkable form of remote usability testing.

Cons: Potentially vulnerable to tightening data protection rules. More functionality than most organisations need.

David Bowen