



# How Much is Much?: a Conceptual Study of Web Traffic

**Tord Høivik**

Oslo and Akershus University College,  
[tordhoivik@gmail.com](mailto:tordhoivik@gmail.com)

## Abstract

Valid and relevant statistics are required for library planning and advocacy. As libraries and their users turn to the web, library statistics must follow. In this paper we explore the use of three common traffic indicators to measure the impact of web resources from national libraries. We present and discuss the use of data on on page views, virtual visits and unique users, with examples from national, academic and public libraries. These indicators are in an early stage of development and need some conceptual and much empirical work to become good tools for strategic planning. But we note four findings: (1) the ratios between the three indicators are very unstable, so we must measure and interpret all three; (2) we find substantial differences between countries, with Denmark in a leading position; (3) in academic and national libraries the number of virtual visits is likely to overshadow the number of physical visits; (4) analysis of web traffic must be based on an understanding of J-shaped distributions ('power laws') rather than concepts drawn from ordinary well-behaved bell curves ('normal distributions').

**Key Words:** web analytics; national libraries; web traffic indicators; physical visits; virtual visits

## Introduction

In January 2008, the British Library (BL) presented an important report on user behaviour in the virtual environment (British Library, 2008). The Director, Dame Lynne Brindley, described her institution as follows: 'We are a trusted

and independent source, both in cyberspace and through our vast printed collections, with more than 67 million hits on our website in the past 12 months and 500,000 readers passing through our doors every year.' The emphasis on visits, both physical and virtual, rather than on collections or loans, was highly interesting. The numbers are large, since we speak of millions. But it is impossible to interpret them without a proper scale of measurement. The library used *hits* (= page views or impressions or downloads) as the counting unit. But what do half a million physical visits and sixty-seven million page views tell us about the British Library? What do the numbers *mean*? And how do these indicators compare with other measures of physical and virtual traffic?

In this paper we present, discuss and compare a number of different indicators that are or could be used to measure library traffic on the web. We hope national libraries, with their large economic and intellectual resources, will take a leading role in implementing this type of usage statistics. This will allow us to understand, predict and plan the interaction between users and libraries on a more solid basis. Web statistics in libraries are still at an early stage of development. Market actors are far ahead in their use of *web analytics* (Wikipedia, henceforth Wp) to monitor their customers and their results. But it is time to make a start. Rather than discussing concepts and proposals for indicators in the abstract, I believe it is time to start to measure, to interpret and to argue about the meaning of empirical data gathered from the field. National libraries operate in the same environment, and compete for the same customers, as other institutions devoted to education, research, media and culture. In this article traffic data from large academic and public libraries are compared. In the case of Norway, statistics are also compared with the most frequently visited web sites in the country.

Concepts and methods that work in the North will work equally well in the East, the West and the lovely Mediterranean South. In the future, libraries must deliver their services through three different channels: through a revitalized physical library, through branded sites on the web and, most radically, through popular web services and personalized web sites beyond the libraries proper. Norway is just one case among many. Our National Library cooperates closely with national libraries throughout Europe. We participate in the Europeana Project, which opened its virtual doors to 4.5 million digital objects at its beta launch in November 2008 (Europeana, Wp). The driving forces of change, we may name them digitalization, globalization

and mass education (or professionalization), define the new rules of the game for all knowledge institutions. Most people, I am afraid, regard statistics as a trivial and boring subject best left to 'dismal scientists' like economists, demographers and statisticians. This *may* have been true in the past, but it will certainly not be true in the future. Digital technology provides us with vast amounts of informative data and the tools to handle them rapidly and efficiently. Globalization increases competition and forces us to compare our own countries with the rest of the world. Professionals cannot just think or talk their way to the truth. They must *investigate* cases and contexts and develop knowledge-based practices. The demand for good statistical data and professional statistical interpretation is increasing in all countries.

National libraries used to be the playground of scholars. The general public was served by public libraries, while students and most researchers depended on university, college and special libraries for the services they needed. With the coming of the web, national libraries will be in a very different position. The National Library of Norway plans to digitalize the *totality* of its holdings within the next ten to twenty years. The numbers are impressive: 4.7 million newspapers (more than 60 million pages), 4 million manuscripts, 2 million periodicals, 1.9 million pamphlets ('småtrykk'), 1.3 million pictures (photos and postcards), one million hours of radio, 450 thousand books, 250 thousand hours of movies and TV, 200 thousand music scores, 200 thousand maps, 80 thousand hours of music and 60 thousand posters (Nasjonalbiblioteket, 2007).

With regard to the general public, national libraries are likely to play a much more active role on the web than in the physical world. Digital collections can be accessed by everybody. On the web, in fact, nobody needs to know that you *are* a national library. Norway's national librarian, Vigdis Moe Skarstein, points out that 'we can offer knowledge, but we cannot force the user to come to us. Knowledge is increasingly something people seek and find, rather than something given to them. ... We must respond to the challenge by offering access to our quality services within the user's environment' (Nasjonalbiblioteket, 2006). In Norway, large-scale digitization is progressing at great speed. But copyright issues crop up. The law will block access to all recent documents unless these copyright issues are resolved. Norway has a tradition for collective solutions in the copyright field, and the National Library has worked hard to find ways of compensating authors for republishing their work on the web.

A large-scale experiment was approved by the authors' organizations in 2009. The project Bokhylla.no (The Digital Book Shelf) will digitalize and publish all books issued in Norway during four decades: 1990–99, 1890–99, 1790–99 and 1690–99. At the moment, nearly fifty thousand titles from the 1990s have been released, as well as nine thousand titles from earlier periods (Bang, ND). To what extent will the general public, as opposed to historians, genealogists and humanistic researchers, be interested in using these resources? Will access to this vast cultural heritage influence cultural consumption patterns? Only statistics can answer that question.

Statisticians do more than collect numbers. We compare them. The ratio between two numbers is a basic tool in applied statistics. But the numbers must be well chosen. The number of public library staff (full-time equivalents, FTEs) per ten thousand inhabitants is highly meaningful. In Norway the number lies close to four. This tells us that the average (full-time) employee serves about two thousand five hundred persons. Since only half the population are library users, the actual number of people served is about twelve hundred and fifty. The number of staff members per ten thousand is less relevant, since it lumps part-time and full-time employees together.

The number of public libraries per ten thousand does not tell us anything about library resources, however. Since libraries are municipal entities, the number (which happens to be 0.9) just expresses the ratio between the number of municipalities (about 430) and the population (five million). The number of public library branches per ten thousand, which is used in some international statistics, is not very useful either. Whether we count the number of branch libraries or the number of library organizations, we combine very small with very large units. We might as well study the importance of agriculture by counting the number of animals per farm, lumping cows, horses, pigs and poultry together.

## **The British Library**

Most indicators in practical library use have the mathematical form of ratios: visits per inhabitant, accessions per hundred inhabitants, loans per FTE, and so forth. When we design and discuss such indicators, we must keep in mind what the components refer to. Words point to entities. What does the numerator (visits, accessions, loans) mean? How is it measured? What

entity does the denominator (population, FTE) describe? How is it measured? Finally, and most importantly: is there a meaningful conceptual relationship between the numerator, on the one hand, and the denominator, on the other? The number of staff per ten thousand inhabitants is meaningful. The number of library units per ten thousand is not.

Let me apply this thinking to the British Library and its visitors. BL is located in London, near St. Pancras station. The library serves the whole of the United Kingdom, with a population of more than sixty million. The metropolitan area of London has about thirteen million people, while the city itself has about 7.5 million inhabitants. Half a million readers a year corresponds to 6.5 visits per hundred London residents a year, four visits per hundred people living in the metropolitan area, and 0.8 visits per hundred UK inhabitants a year. These numbers are very low compared with public libraries, which typically attract a handful of visits *per inhabitant* per year. But such comparisons are nearly meaningless. The number of physical visitors to national libraries must be related to the purpose and intended audience of the library.

'The Library is open to everyone who has a genuine need to use its collections. ... Historically, only those wishing to use specialised material unavailable in other public or academic libraries would be given a Reader Pass. Recently, the Library has been criticised for admitting numbers of undergraduate students, who have access to their own university libraries, to the reading rooms. The Library replied that it has always admitted undergraduates as long as they have a legitimate personal, work-related or academic research purpose.' (Wikipedia article on the [British Library](#)).

Half a million visits a year corresponds to about forty thousand visits per month, or about ten thousand visits per week, or about fifteen hundred visits per day. The primary BL audience consists of specialized researchers and scholars in the humanities and social sciences who live in London, or who are visiting the area for the purpose of research and study. But we should note that the sheer number of library visits is not of great interest as such. It is the content and duration of the visit, we could almost say: *the meaning of the visit*, that is important.

Visits are easy to measure by the use of turnstiles or electronic counters. But the value and impact of the library depends on what people actually do once they have passed the gates. The number of visits published by the British

Library, nearly half a million in 2007, refers to the use of the reading rooms only (British Library, 2007, p. 27). The library does not, in other words, follow the IFLA guidelines, where a physical visit is defined as ‘the act of a person’s entering the library premises’ (Poll, 2007, p. 112). Had they behaved ‘properly’ and counted every entry visitor, numbers would have been substantially higher. National libraries often arrange exhibitions, lectures, guided tours and other open events. In London, the BL website explains, ‘the general public can see important works like the Magna Carta, Captain Cook’s journal, Charlotte Brontë’s *Jane Eyre*, ... for free in the Sir John Ritblat Gallery.’ Most national libraries would probably cast their net more widely and include all persons that enter the building, whether as scholars or as gawking tourists.

The actual design and lay-out of the building can also influence the counting. The coffee shop and the local bookstore may, for instance, be located inside or outside the library space proper. For planning purposes it is important to separate the primary users (scholars, researchers and students) from the secondary users (tourists, occasional visitors). I believe libraries ought to do rather more practical and conceptual work in this area. By a variety of simple observation methods we can, in fact, distinguish between types of use and gather detailed data on duration and content (Høivik, 2008). But such methods fall beyond the scope of this paper, which focuses on web traffic.

## Web Analytics

Three indicators are widely used to measure traffic on the web: page views (PV), visits or sessions (VS) and the number of unique visitors (UV) during a given period. These are the three variables that TNS Metrix, the main provider of web analytics in Norway, publishes on their open web site. The Danish firm KPI Index, which has been engaged to measure library traffic on the web, uses the the same indicators. Wikipedia sees PV and VS as rather primitive indicators: ‘Two units of measure were introduced in the mid 1990s to gauge more accurately the amount of human activity on web servers. These were page views and visits (or sessions). A page view was defined as a request made to the web server for a page, as opposed to a graphic, while a visit was defined as a sequence of requests from a uniquely identified client that expired after a certain amount of inactivity, usually 30 minutes. The page views and visits are still commonly displayed metrics, but are now considered

rather rudimentary. (Wikipedia article on Web analytics). But since libraries are beginners in this field, they need to start with the rudiments of learning.

The first two indicators are additive. The number of page views, or the number of virtual visits, in weeks A and B can be added together. The number of unique visitors must be handled with care, however. UV behaves in a very different manner. The number of unique visitors in week A cannot be added to the number of unique visitors in week B in order to find the number of unique visitors in the period A+B. People who came to visit on both weeks should only be counted once. When we want to know the number of *different visitors*, repeat visits should be excluded. This basic fact is very often forgotten (or has never been learned). TNS Metrix measures UV on a weekly basis. The British Library used a full year.

Commercial actors have moved beyond these three and apply much more sophisticated methods to study how customers navigate their web sites. Such data are generally secret, but we can get an idea of the detail involved from Ibsen.net, which publishes detailed statistics on their web traffic (Ibsen.net, Usage statistics). The same is true in the physical world. Commercial visitor studies are an important field within marketing research. The measures that were used in the past, such as the number of physical visits, are being replaced by measures that reveal the *activities* undertaken by the visitors as they move around in physical (or virtual) space. The study of *user behaviour* is more advanced in the commercial than in the public sphere. But the relevant methods are basically the same.

## **Page Views and Sessions**

Libraries tend to see themselves as public service providers. They live in protected environments, beyond the cut-throat competition of commercial markets. The web, however, is inherently competitive. Free or fee does not matter. All web sites compete for the scarcest commodity in the world, which is neither gold nor diamonds, but genuine human attention. Survey data from Nielsen Ratings illustrate the situation. In March 2008 the typical internet user spent thirty-three hours on the web and visited seventy different domains. He or she downloaded 1,550 web pages — and looked at each page for less than a minute. Each surfing period on the web lasted about an hour on average.

In July 2011, British Library virtual visitors spent about three minutes on the site, looking at four pages on average (Alexa, bl.uk). This is quite similar to the global Nielsen data for 2008: about forty-five seconds per page. I assume, therefore, that this was the case for BL in 2008 as well. In 2008, the web site contained about ten thousand individual pages. The *average* page was therefore downloaded 6,700 times a year or about twenty times a day. Sixty-seven million page views per year translate into fifty million minutes or about:

- three thousand hours of reading or attention per day
- twenty thousand hours per week
- eighty thousand hours per month
- one million hours per year.

Here I focus on scale rather than on numerical accuracy. To understand a new phenomenon, we need to know its relative size or scale, not the precise numbers. We may compare the time spent on the web with the time spent by physical visitors and by employees. Every day in 2008:

- virtual visitors spent about 400 working days at the BL web site
- physical visitors (using the seats) spent *at most* 1.500 working days at the London site
- staff spent 2,500 working days at the physical sites (London and Boston Spa).

So far we have considered page views. But the number of virtual visits, relative to the population served, is a more intuitive indicator. We start by defining the concept. According to ISO 2789:

- 'a virtual visit is a series of requests for data files from one and the same website visitor.
- a website visitor is either a unique and identified web browser program or an identified IP address that has accessed pages from the library's website.
- the interval between two consecutive requests must not be longer than a time-out period of 30 minutes if they are to be counted as part of the same virtual visit. A longer interval initiates a new visit.'



Poll (2007, p. 114) adds that a virtual visit must come from

- 'outside the library premises in order to use one of the services provided by the library.
- the population to be served is the number of persons to whom the library is commissioned to provide its services.'

In 2008 the British Library did not report the number of visits. But if we assume that the number of page views per visit remained constant, 67 million page views translates into

- fifty thousand visits per day
- three hundred and fifty thousand visits per week
- 1.4 million visits per month
- 17 million visits per year.

In 2011 the BL reported that nearly four hundred thousand persons visited the reading rooms. This corresponds to eight thousand visits per week or about 1,300 visits per week day. The library has 1,200 seats, which tend to be occupied most of the day. Internet traffic was probably higher in 2011 than in 2008. Clearly, the *number* of virtual visits is very much higher than the number of physical visits, by a factor of forty or more. But the *time* they spent at the library (web site) was substantially less. Fifty thousand virtual visits, lasting four minutes on the average, boils down to about four hundred working days. So physical *duration* dominates over virtual duration.

This point needs to be stressed: physical and virtual visits are very different entities. They should not be lumped together under the general heading of 'Visits'. When we work with statistics, we must compare like with like. The same principle applies to libraries and librarians: a person working 3.5 hours a week is not equivalent to a person working thirty-five hours a week. A big metropolitan library with a staff of one hundred is not equivalent to a one-person operation in a small rural community.

## Denmark

Let us look at some other libraries. In late 2007, the Danish Agency for Libraries and Media hired KPI Index to set up a standardized system

for measuring web traffic. The three basic indicators (PV, VS, UV) are published every week, and are also available on a monthly and an annual basis. This is the most advanced web indicator system I am aware of in the library world. This tool provides quite a fine-grained mapping of the traffic from a national point of view. Libraries that join the system can generate a wide variety of more detailed traffic reports. To what extent such studies are actually carried out I do not know. So far the impact of these data on published library research seems modest. But KPI is definitely one of the finest data resources available for studies on web traffic.

Here I give a brief analysis of the 2010 data (Høivik, 2011). In 2010 the KPI index covered eighty-four public library systems (excluding Sydslesvig, which belongs to Germany). Their size went from half a million (central Copenhagen) to twelve thousand inhabitants. The typical size, defined as the libraries lying between the upper and the lower quartile, ranged from sixty to thirty thousand people. Half the libraries registered between 2.6 and 5.2 web visits per inhabitant in 2010. The median number was 3.6. This is somewhat below the number of physical visits, which was 5.5 in 2009. The typical number of pages downloaded per session spanned a quite narrow range, from 9.2 to 7.1 pages for the middle half. The number of visits per unique visitor also had a quite compressed distribution. The interquartile range went from 3.6 to 2.9. Bigger libraries seem to have more users than small libraries. In 2010, libraries serving more than fifty thousand had a median of seventy-nine visits per hundred inhabitants. Libraries in the 30–50 thousand range had sixty-seven. Libraries with less than thirty had fifty-nine. The differences are not major. Two likely explanations are (1) that big libraries serve more urbanized areas with more web oriented users, and (2) that big libraries offer better web sites which attract more users.

We can summarize the results as follows: the average Dane visits her public library on the web three to four times a year and consults an average of eight pages each time. Note that the number of unique users depends both on the interval (2010) and on the library system doing the counting. A roaming user will be counted as a unique visitor in several systems during the same interval. The sum of all unique users is in fact greater than the total Danish population — and belongs to the many numbers that may be computed *without referring to a meaningful social entity*.

In Danish public libraries web visits and physical visits have the same order of magnitude. The ratio between the two is therefore much lower than what we found in the British Library. To get an understanding of web traffic in general we explore this further. The Royal Danish Library serves both as a national and as a university library. It has reported on both types of visits since 2005. The virtual-to-physical ratio varied quite a bit: 6.5 in 2005, and then 8.5 - 18 - 26, and finally 11 in 2009 (Kongelige Bibliotek, 2010; Kongelige Bibliotek, 2011). Since the numbers are rough, I only use two significant digits. Some of the variation is due to changes in the actual procedures. Most libraries are still in a calibration phase. We are testing out the tools and trying to find out what to do with the numbers. So far we have found three different patterns:

- British Library: about forty virtual visits per physical visit
- Royal Danish Library: about ten virtual visits per physical visit
- Danish public libraries: less than one virtual visit per physical visit.

The double function of the Royal Danish Library probably explains the high number of physical visits. The British Library allows students, but is meant for research rather than for ordinary student use. As national libraries increase their resources and services on the web, the number of physical visits is likely to be swamped by the number of virtual visits. Virtual visits are not limited by physical space and have a much greater potential for growth. But they also take place in a much more competitive environment. People who visit physically are committing their bodies as well as their minds to the library. People who arrive through the web are fickle and insubstantial as ghosts. This is the secret of Google's search engine. More than half the web visitors to the British Library come directly from Google, and more than half go back to Google when they are finished (Alexa, bl.uk). Google does not tie people down. It just helps the ghosts to find the hosts (which they want to haunt, one might say).

The biggest public library in Norway, Deichmanske bibliotek in Oslo, is one of the few Norwegian libraries for which some web statistics are available. In recent years (2007–10) it has reported about twelve million page hits a year (Høivik, 2011). Since Oslo approaches six hundred thousand inhabitants, this corresponds to twenty downloads per inhabitant. In Denmark, the median number of downloads (2010) ranges from thirty-four for large to twenty-five

pages per inhabitant for smaller libraries. So public library web traffic is clearly higher in Denmark than in Norway. It is also very much better documented. A reasonable and testable hypothesis would be: web sites with user friendly design (information architecture) and rich resources will be used more intensively than sites with less content and a less attractive structure. This implies longer dwell times and more page views per visit. On top of that, I would expect central urban libraries to receive more web traffic than more provincial and rural libraries. The digital transformation starts in the more cosmopolitan urban environments.

## Unique Visitors

On the web, the location of the library is no longer important. The virtual arm of the British Library serves not only the United Kingdom, but all speakers of English as long as they have access to the web. The number of people who can read the English language may be above one thousand million (Wikipedia). About five hundred million have access to the web (Internet world stats). These are all potential users of the BL web site. But people living outside UK will, in general, have national libraries of their own. The *primary* audience of the British Library web services must be the sixty million inhabitants of the UK. We know, however, that a few percent of the incoming traffic to BL comes via the Indian version of Google. The share was 4% on July 11, 2011, Alexa says. In addition to page views, BL also provided some data on unique visitors (or hosts). The numbers lie close to five million:

- 2005/06: 4.2 million unique visitors
- 2006/07: 4.9 million

The number of unique hosts is the best approximation available of the number of individual users of the website. During 2006/07 less than ten percent of the UK population consulted the BL web site. We do not know the distribution of traffic between the United Kingdom, on the one hand, and the Anglophone world beyond the Isles, on the other. In Germany, approximately forty percent of the NL web traffic came from outside the country in 2006 (top level domain = .de) (Deutsche Nationalbibliothek). If we assume an equal division between national and international traffic in Great Britain, we get a participation rate of 2.45 (million UV) / 60, or about

four percent of the population. The BL presumably has access to detailed traffic data that can give us better numbers, and much more information about the web users than the published statistics allow. Since we know the number of visitors, we can also calculate the average number of pages consulted:

- 2005/06: 11.7 pages per unique visitor
- 2006/07: 12.4.

With a viewing time of about one minute per page, we could imagine five million users spending twelve minutes each on the website during the year. This mental image should not be trusted too far, however. When we look at web traffic, we are dealing with data that do not follow the customary bell curves ('normal distributions'). Individual usage of library sites is characterized by J-shaped (or long-tailed) rather than by bell-shaped distributions. In the BL case, fifty-five million UK inhabitants do not consult the web site at all. Most of the active users are likely to be brief and occasional visitors.

In social life economists and sociologists often find distributions that follow the Pareto principle: that twenty percent of the individuals represent eighty percent of the activity (Wikipedia). We encounter many such distributions in studies of other library activities. I am willing to guess that we will find the same pattern on the web. If I am right, eighty percent of the sixty-seven million pages — or about fifty million pages, will have been downloaded by about one million 'heavy users'. This corresponds to fifty pages (or minutes) per person per year. The remaining thirteen million pages are shared among four million 'light users'. Each of them drops in for a brief visit or two, consulting, on average, three pages a year. Once we get more detailed data on unique users and other aspects of web traffic, we will be able to test whether this 'Pareto hypothesis' is correct.

## **Comparing National Libraries**

Until 1998, the Norwegian National Library combined the functions of a national and a university library. Today it has the same specialized profile as the British Library. Its traditional building in the centre of Oslo was nearly totally rebuilt a few years ago. This is where researchers come to study and the general public to attend lectures and to visit exhibitions. But the library

also has a vast storage, development and digitalization unit located at Mo i Rana, a former factory town in Northern Norway. Here visitors are few and far between. The library registered 170 thousand physical visits in 2008 and slightly fewer in 2009. It does not provide regular reports on its web traffic, but has released monthly data on page hits from January 2006 to April 2008 (Solbakk, 2008). At that time the annual number of page hits was around eight million per year. This implies that the Norwegian level of traffic, as measured by page hits, was about fifty percent above the British level:

- United Kingdom: 67 million page views / 60 million inhabitants = 1.1 pages per inhabitant
- Norway: 8 million pages / 4.6 million inhabitants = 1.7 pages per inhabitant.

In the years 2002–07 the National Library of Finland registered between 1.0 and 1.5 page views per inhabitant, placing it between Norway and the UK. In 2010 the library had 170 thousand physical and 1.9 million virtual visits. This places it close to the Danish National library:

- Finland: 11 virtual visits per physical visit (10)
- Denmark: 11 virtual visits per physical visit (09).

## Web Traffic: the Case of Norway

TNS Gallup dominates the Norwegian market for analysis of web traffic. They publish a weekly summary from more than one hundred organizations, covering the three indicators we have mentioned. The information released by the National Library only gives the number of pages. I will therefore concentrate on the page indicator — with data from Week 20, 2008 (= May 12–18) (TNS Gallup, 2008):

- two sites registered more than one hundred million page views per week: Norway's biggest newspaper (VG) and Finn, which is the biggest site for classified advertisements
- nine sites notched up between ten million and a hundred million downloads: the Norwegian MSN site, two national newspapers, one digital news service, the main national broadcasting web site, two search portals, one phone directory and one financial magazine.

- thirty-nine sites delivered more than one million, but less than ten million, page views.
- thirty-seven sites had more than one hundred thousand, and less than one million, page views.

We find the National Library web site, with a weekly average of 160 thousand hits, towards the lower end of this last category. To place the traffic in context, we may look at the site's immediate neighbours. These range between one hundred and two hundred and fifty thousand page impressions per week. Just above the National Library we find:

- one small regional newspaper (*Østlandets Blad*)
- one small national newspaper for farmers (*Nationen*)
- an important soccer support club (*Rosenborg*)
- a regional news site (*iBergen.no*)
- a very small phone directory (*Folk.no*)
- a magazine on technology (*Teknisk Ukeblad*).

Just below we encounter

- a marketing magazine (*Kampanje*)
- a youth-oriented magazine (*Det Nye*)
- a web site on buildings and construction (*Bygg.no*)
- consumer magazine on sound and image (*Lyd & Bilde*)
- a web site devoted to cell phones (*Mobilen*)
- a consumer technology web site (*Teknofil*).

These are, we may say, small-scale neighbours. The most popular web sites in Norway attract, in other words, a thousand times as much traffic as the National Library. Other major web services receive several hundred times more attention. Many regional and specialized web sites also attract much more traffic. With the ambitious and costly goal of digitizing *all* its collections within two decades, the Norwegian National Library has set its sight on the new, digital knowledge-based society. Involving the general public in this project is still a major task. In a country known for its many tunnels underground, roads still need a fair amount of traffic to justify their existence. Libraries that provide access without attracting users have not achieved their goals. I believe this is true in the rest of Europe as well.

## Conclusion

Libraries of the future will clearly deliver their services in three different environments: physical buildings, local web sites and distributed user environments. The physical libraries are caught up in a process of profound change. When their customers change, they must adapt or fade. It is the new ways of research, teaching and learning that give the process its momentum. Public libraries are becoming more physical and more digital at the same time. Academic librarians are taking a new look at their place within the intellectual division of labour. Student-oriented libraries redefine themselves as learning centres. Research-oriented libraries take on much more active roles in the research process. The term *cyberinfrastructure* is clumsy, but covers what I mean. National libraries face similar challenges. On the web they must shift from collection-based thinking to competitive customer services. This involves a deep change in organizational culture.

The speed and uptake of web-based practices will of course differ from country to country. But I see the end result as given. The digital revolution will impose its conditions on the world like the industrial revolution did two hundred years ago. We are all moving in the same direction. When we discuss the future of national libraries, we can therefore draw on trends and experiences from many neighbouring fields. Library statistics tend to focus on information that is easy to collect but hard to interpret. The number of physical visits does not indicate the time spent at, or the benefit gained from, using the library. The number of loans does not measure reading or understanding, learning or wisdom. If we look at web-based versus physical activities, we see that virtual visits are different in kind from physical visits.

During its first decade (1991–2000) the web was mainly used to strengthen existing structures and services. Web sites were typically designed as mirror images of the organization chart. Beaujolais Nouveau was poured into old bottles. During the second decade, from 2001 to 2010, the web started to challenge the structures themselves. This third decade (2011–2020) will be a period of major readjustment and deep structural change. All the libraries I have visited in this paper are confronting the challenges of the web. Their strategic plans are clear enough: a new era is upon us. What I argue for is better documentation. We need statistics that show us what is happening. Some of the numbers will be encouraging. Others will not. Professionals should gather, publish and discuss both kinds, beyond hope or fear.



In Norway, massive digitalization of books started after this paper was first drafted. So far we have only scattered data on the corresponding traffic. But the volume is clearly increasing. In the Summer of 2010 the national librarian reported that traffic to the book site now 'represents roughly one book page per second, day and night.' (Høivik, 2010). This implies thirty million pages a year, moving the library from the low to the high end of its category (= 0.1–1 million page hits per week). However, to properly interpret these changes we need to look at all three indicators. The book readers on the web form a relatively small group. They make a big impact because they remain on-site for long periods and download lots of pages: 65 pages during the average visit. To understand web traffic, we need to separate high-intensity from low-intensity users. Since we lack data about user profiles from other national and academic libraries, this is a task for the future.

## References

Alexa, bl.uk.

Bang, S. (ND): Om bokhylla (About the Digital Book Shelf).

British Library (2007): The British Library Annual Report and Accounts 2006/07, URL= <http://www.bl.uk/about/annual/2006to2007/performstats.html>

British Library (2008): Information Behaviour of the Researcher of the Future. URL= <http://www.bl.uk/news/pdf/googlegen.pdf>

Deutsche Nationalbibliothek (2006): Report on usage in 2006. Page no longer available.

Høivik, T.: KPI Index Denmark

Høivik, T. (2008): Count The Traffic. Paper for the 2008 IFLA conference in Quebec.

Høivik, T. (2010): Få, men trofaste brukere. Samstat blog, July 2010.

Høivik, T. (2011): ST 54/11: Data fra Deichman 2010.

Ibsen.net, usage statistics. <http://www.ibsen.net/stats/> [accessed 11 December 2011]

Internet World Stats. Internet world users by language.

Kongelige Bibliotek, Det (2010): Årsberetning 2009 (Annual report 2009). København: Det Kongelige Bibliotek.

Kongelige Bibliotek, Det (2011): Årsrapport 2010 (Annual report 2010). København: Det Kongelige Bibliotek.

KPI index. Homepage (Danish).

Nasjonalbiblioteket (2006): Nasjonalbibliotekarens kommentar til bibliotekutredningen (The National Librarian's response to the Library White Paper). Oslo: Nasjonalbiblioteket.

Nasjonalbiblioteket (2007): Digitalisering av bøker i Nasjonalbiblioteket - metodikk og erfaringer (Digitalization of books in the National Library - methods and experiences). Oslo: Nasjonalbiblioteket. PDF.

Poll, R. and Peter te Boekhorst (2007): *Measuring quality. Performance measurement in libraries*. München: Saur, 2nd edition.

Solbakk, S.A. (Director of digital services, Nasjonalbiblioteket) (2008): Personal communication, May 23-24.

TNS Gallup. Topplisten uke 20/2008.

## Wikipedia Articles

British Library, [http://en.wikipedia.org/wiki/British\\_Library](http://en.wikipedia.org/wiki/British_Library)

English language, [http://en.wikipedia.org/wiki/English\\_language](http://en.wikipedia.org/wiki/English_language)

Europeana, <http://en.wikipedia.org/wiki/Europeana>

Pareto principle, [http://en.wikipedia.org/wiki/Pareto\\_principle](http://en.wikipedia.org/wiki/Pareto_principle)

Web analytics, [http://en.wikipedia.org/wiki/Web\\_analytics](http://en.wikipedia.org/wiki/Web_analytics)