A Digital Library and Archive about the Marshall Islands
EXPERIENCES AND CHALLENGES

DIRK H. R. SPENNEMANN

The Johnstone Centre and School of Environmental and Information Sciences,
Charles Sturt University, P. O. Box 789, Albury 2640, Australia
e-mail: dpennemann@csu.edu.au

Abstract

The development of the World Wide Web has allowed for the establishment of on-line information warehouses, either in centralised or distributed form. A digital library and archive about the Marshall Islands, a small Pacific Island nation comprises of primary and secondary sources which are often scattered and, especially for remote communities, had to come by. This paper discusses the development, current use and future of the library, looking at users and subject areas of demand.

Introduction

The phenomenal growth of the World Wide Web (WWW) since its popularisation in 1993 (Cailliau 1995) has spawned a plethora of web sites serving an incredible number of users. It is estimated that over 745 million people have internet connections, with the largest number of users (28%) living in the USA (Global Reach 2003). Many sites are government or institutionally sponsored and deliver information to the end-user. However, unlike the content of physical libraries, the content of the websites is of varied quality, ranging from the authoritative to the mundane which has limited, on occasion little veracity as on-line publications can be a private matter of vanity publishing.

One of the major problems with WWW resources is the relative impermanence of many sites. All too often one encounters the "Error 404 File not found" message. While major countries, such as Australia for example, have been able to maintain archives of WWW site content through its Pandora project (NLA 2003), smaller countries have not. As a result, much of what is published on the WWW is ephemeral in nature. At the beginning of the internet 'boom' it was feared that smaller countries might be digitally colonised by larger ones due to their lesser technological and economic base (Spennemann *et al* 1996). While the growing number of internet connections in smaller countries does not bear out this feat (cf. Global Reach 2003), the electronic colonisation of the *content* has commenced. In a parallel development, the increasing commercialisation of the web has seen much of the amassed content being acquired, rephrased and locked up on commercial sites. This trend is bound to continue

Before we look a case study of a major 'free-to-web' digital library, however, let us consider some of the background of organising and presenting primary and secondary source information on the WWW.

Structuring and accessing sources

Search engines, such as Google, Altavista and the like provide access to the material based on search algorithms, but return much without quality control. While the WWW is the ultimate anarchic system of the delivery of information and disinformation, the power of the search engines means that, following a query, only the top ten or twenty sites listed are visited by the vast number of users. As any user of search engines has experienced, the sites thus returned are of, at best, varied relevance, in particular if the choice of search keyword was kept too general.

An early way of organizing the material on the web were the WWW Virtual Libraries (WWW VL 2003a), such as VL-Pacific Studies (WWW VL 2003b) and VL-Education (WWW VL 2003c). Here links to various thematically related sites are moderated, assessed and the currency of the links maintained by a selective group of individuals (WWW VL 2003a). Compared to that centralized model, the Open Directory Project (Netscape 2002) employs a large range of individuals with specific interests to compile a structured directory. Both Virtual Libraries and the Open Directory Project are human-based and structured, exercising some level of quality control. Despite quality control, many Virtual Libraries are in essence a collection of links to various resources distributed across the WWW, many of which in turn are collections of links.

This paper deals with, "The Marshall Islands. An Electronic Library & Archive of Primary Sources" and discusses it genesis, management and the experiences gained in developing and managing the electronic library as well servicing the needs of its users.

The Marshall Islands

The Republic of the Marshall Islands (RMI) extends over about 1,950,000 km^2 of the Pacific Ocean but has only 171 km^2 of land. It comprises 29 coral atolls and five islands that are arranged geographically in two chains trending north-west to north-east in an area between 4° - 19° North latitude and 160° - 175° East longitude. The environment is tropical, but, given the limited land mass, governed by oceanic climatic conditions, with limited vegetation coverage. While there is a large range of flora, the terrestrial fauna is limited.

European contact was essentially ship based until 1859 when a permanent mission station and then trading stations were established on Ebon, a southern atoll. The German trading interests gained ascendancy in the 1870s, leading to the formal annexation of the islands by Germany in 1885 and the establishment of a colonial administration from 1886 until 1914. Having taken the islands in the early stages of World War I, Japan was handed the Marshalls by the League of Nations, together with the rest of the German possessions, for administration as Mandated Territory. Japan used the atolls of the Marshalls as bases to launch attacks against US possessions in Micronesia during World War II. The US captured the islands in 1944 after a prolonged bombing campaign. Following World War II the Marshalls, like the rest of Micronesia, were administered by the USA as a Trust Territory on behalf of the United Nations. During the 1950s and 1960s the Marshalls, in particular Bikini Atoll, gained world-wide recognition as the location of nuclear and thermo-nuclear tests conducted by the USA. By 1989 the Marshalls had gained independence as the Republic of the Marshall Islands, closely allied to the USA through a Compact of Free Association.

Following the Compact of Free Association, which *inter alia* permits Marshallese citizens to reside in the US, large Marshallese expatriate communities were established in Hawaii, Costa Mesa (California), Oregon and more recently also in Arkansas.

Concept

The aim of the site "The Marshall Islands. An Electronic Library & Archive of Primary Sources" (http://marshall.csu.edu.au) is (i) to provide a digital library and archive of material on the Marshall Islands, (ii) make it available in an easy-to-use fashion, (iii) design it to be usable efficiently with slightly out-of-date technologies and/or at slower connection speeds; (iv) to do so both without fees, *ie* free-to-web publishing, and without advertising, and (v) to ensure permanency of the material hosted.

The site started in 1999 with the publication of electronic versions of three hard copy books *Marshallese Legends and Traditions* (Downing *et al* 1993), *Tattooing in the Marshall Islands* (Spennemann 1992), and *Essays on the Marshallese Past* (Spennemann 1993). Since then the site has been systematically expanded by the editor. The site, which is hosted by Charles Sturt University (Australia), is and will remain non-commercial and advertising free.

Table 1 Milestones of the Marshall Islands site

1 August 1999	Web site established on a Charles Sturt University server located in Albury.
1 November 1999	Three digital books published
17 February 2000	Invisible web counter set up for the portal page of the site.
18 June 2001	The site formally archived by the National Library of Australia as part of its PANDORA Archive of Australia's Networked Documentary Resources. Archiving continues in annual instalments, with a complete archive
6 September 2001	Counter of total site usage established. The counter commenced count at 44,948, which is the sum of all main pages with counters to date. The count for the site portal alone stands at 21, 272.
7 December 2001	Total site count exceeds 100,000
15 March 2002	Virtual domain marshall.csu.edu.au established
17 January 2003	Total site count exceeds one million
23 May 2003	The site has been moved to a new university server physically located in Wagga Wagga, as the old server ('life') in Thurgoona is being discontinued.

Site Structure

The site comprises of a total of 3452 text (html and pdf) and 7063 image files, structured in 230 directories. A further 300 text and over 1300 image files provide structural and navigational support as well as the 'look and feel' of the site. In total 1.63 gigabytes of data are uploaded on the server, with a further 0.53 gigabytes (in the form of over 900

text and over 2200 image files structured in 160 directories and sub-directories) in various stages of preparation.

Figure 1. Portal page of "The Marshall Islands. An Electronic Library & Archive of Primary Sources"

All files mounted on the site are either accessed directly through search engines, or accessed via the site portal (Figure 1) and then through the structured first level directory pages such as 'Culture & Society,' 'Politics & Economy,' 'European History,' 'World War II,' 'Environment,' Or 'Public Health,' 'Literature & Art' and through a number of first level directory pages that cut geographically, such as 'Map,' 'Individual Atolls,' and 'Country of Origin.'

Most of the first level directory pages have deeper-level overview pages, such as historic images or pages specific to major phases of Marshallese history. To provide a third approach, a list of all files hosted on the site can be accessed from the alphabetical list provided, as well as from a thematic file index, which is in preparation.

The portal page provides a 'link of the day' and an 'historic image of the day,' both of which are generated from a table of calendar days, cycling at monthly intervals. These cater for the curiosity of 'accidental tourists' arriving at the site. In addition a 'What's New?' page and a page of Editorial Comment has been provided.

As with any structuring of content, there are problems of classification, with some material belonging into two categories. Unlike physical libraries, where volumes can only be shelved at one location if confusion is to be avoided, digital libraries allow for double entries and thus alternative pathways of access. The site is structured along easily recognisable classes, which are all represented at the portal page level. There are two breaches of the overriding logic: the categories of 'World War II' and 'Stamps.' Both categories were places on the portal page after extended correspondence with other website managers, as cases could be made that these cater for high-volume special interest groups. The usage statistics (Table 7) certainly bear this out.

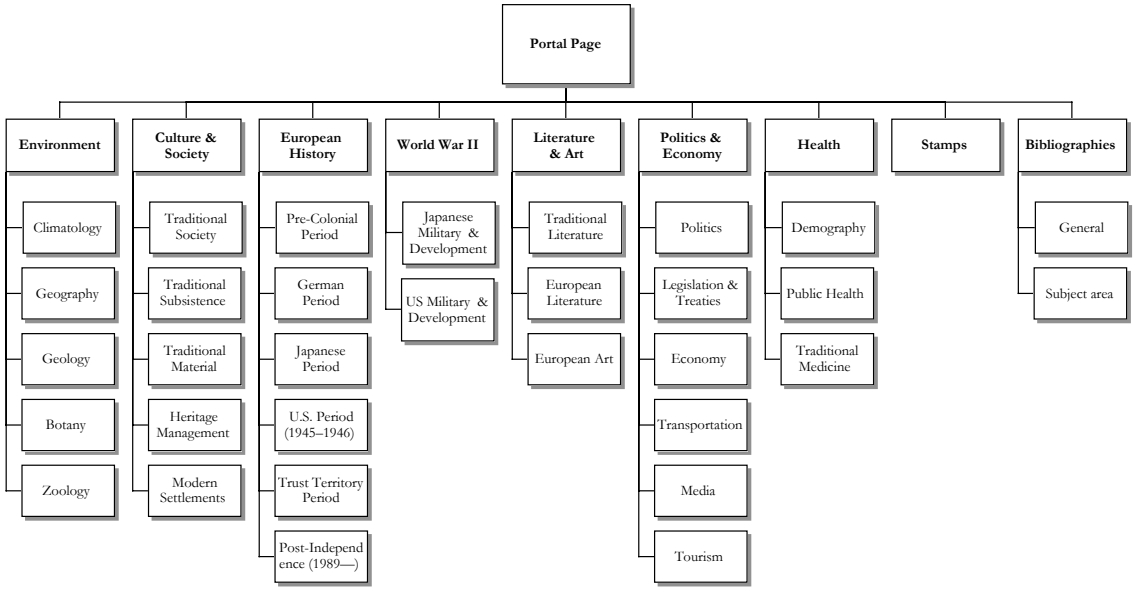

Figure 2. Structure of the first and second-level directories of "The Marshall Islands. An Electronic Library & Archive of Primary Sources"

Page Design

The site places emphasis not on form, but on the quality of the data presented and on the permanency of the data sources. Consequently, the main directory pages, as well as the pages for the subordinate level of directories, are designed to be image poor, so that they load faster. This was a conscious decision to enable such users to access the material who have older and thus slower machines, or connect via slower modes of access. This is in particular the case in developing nations. Document pages may contain a number of images. Where these abound, limited resolution copies are mounted on the page, with each image linking to a higher quality image presented on a separate page.

The overall lay-out is purposefully designed in an elegant, yet plain, fashion in order to avoid 'trendy' lay-out schemes, and animated images or graphic front pages, which swiftly show their age. The text is black on white background, using serif fonts (Times) for body text and sans serif fonts (Helvetica/Arial) for headings. All pages have a similarly designed masthead with a image that varies between the major categories. Likewise, all pages have a footer section that provides site navigation options as well as bibliographical information.

The citation of WWW pages has been giving headaches to librarians tasked with developing citation and bibliographic style manuals. An abundance of approaches exists.

To influence the outcome a preferred citation is provided at the bottom of each page as shown in the example below:

Spennemann, Dirk H.R. (2002). Postage Stamps used in the German Marshall Islands. German Postal Services on Jaluit. URL: http://marshall.csu.edu.au/html/Stamps/Stamps_History_Jaluit.html

Since early 2003 all pages have been designed with a fixed text width of 640 pixels, which prevents users from widening or narrowing the text on the screen. This has been done intentionally to ensure maximum readability by limiting the body text to about 60 characters per line which is deemed to be the ideal line length, and allows the adding in a white 'margin' at the left and right which has been shown to add readability (Youngman & Scharff 1998). Furthermore, the fixed with of the text block allows to 'format' documents in way that images remain at their intended placement. Progressively, older pages will be reformatted.

Language

The language used for the pages and documents is English. While some German archival documents are reproduced on the site, where possible English translations are provided side-by side. Two childrens' books exist in bilingual English and Marshallese. Clearly, the language ability of the user limits access to the library. Other research has shown that non-English speakers prefer to read websites in their language rather than English (IDC 2001 quoted after Global Reach 2003), with 34% of French users being prepared to read English-language pages, compared to only 18% of German users and as little as 8% of Japanese users. Statistics by Global Reach (2003b) suggest that at present some 35.6% of the world's on-line population speak English. Projections suggest however, that the number of non-English speaking on-line users is bound to increase dramatically. Realistically, it is not possible to make too many of the documents on the server available in languages other than English, with the possible exception of overview review narratives to be developed for major subject areas.

Usage Statistics

Rather than counting 'hits', *ie* system calls to the server which include calls to image files, the statistics presented here only account for html and pdf files. Some of the web browsers, such as Internet Explorer re-deliver a page image without down loading again if it is kept in the individual computer's cache. Thus the actual number of 'viewings' of a given page is higher than the number of requests sent to the server. This is particularly true for directory and overview pages. As a result the server log statistics only count 'pages delivered' or 'impressions.'

How many Impressions?

The usage statistics for the site have been collected by three different means, which causes some problems of comparison. The first counter was activated on 17 February 2000, but only measured the frequency with which the portal page was called up. Subsequent counters were placed on second-level pages (eg. culture, history, politics etc). As most search engines index and then target individual pages, and because many users may never access the site portal, these early statistics grossly undercounted the total site usage. The statistics until March 2002 were courtesy of Netscape's www.hitometer.com

On 6 September 2001 a new counter was established to measure the total usage of all pages on the site. That counter was set to commence counting at 44,948, which is the sum of the usage of all main pages with counters to that date. It is obvious that this figure grossly undercounts the total site usage, but was the only verifiable value available. From 6 September 2001 on, all sub pages had the same counter so that direct visits to lower-order pages contributed to a unified overall site visit count. A comparison between the total site count and the count of the visits to the portal page alone (period October 2001 to August 2002) suggests that less than 10% of the visitors use the portal page at any point of their visit. Based on this ratio, the total site count on 6 September 2001 would have been closer to 260,000 than the 44,948 used as the starting point of the current total count figure.

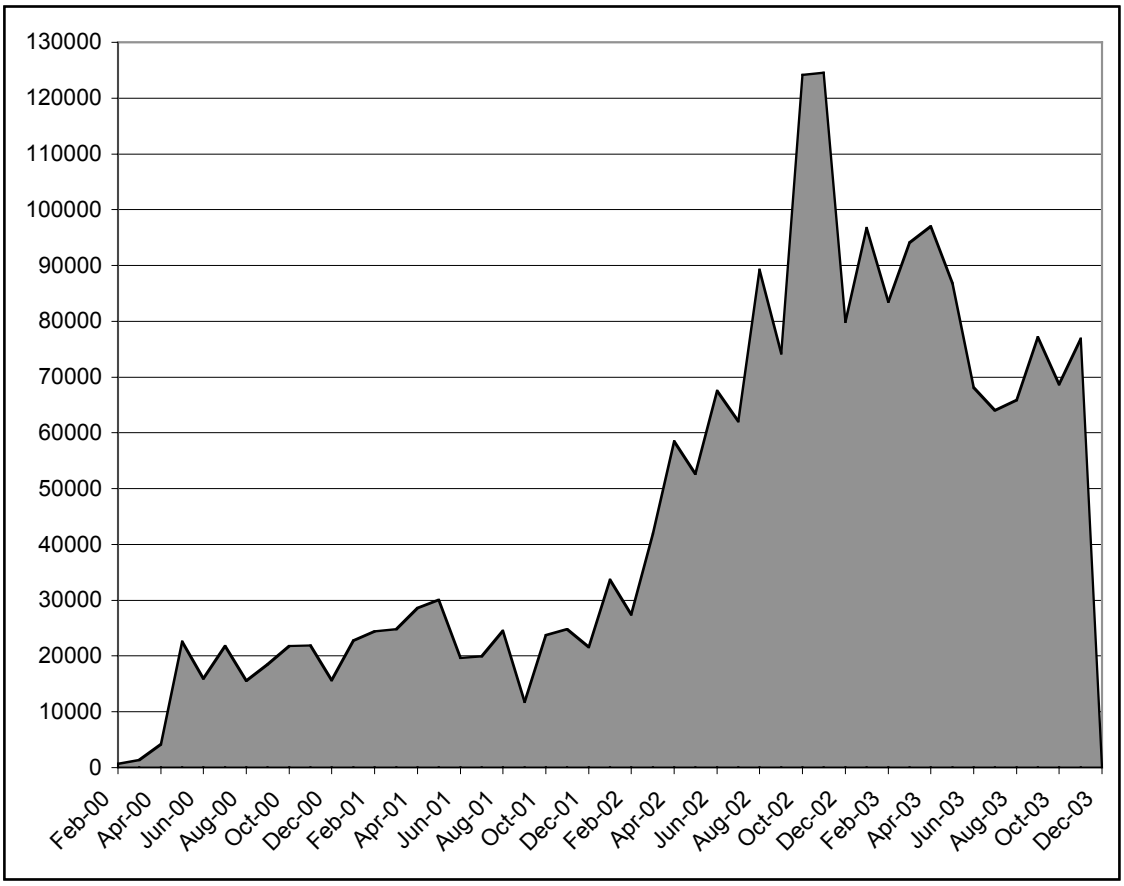

Figure 3. Pages delivered between February 2000 and November 2003

Commencing April 2002 accurate server-level statistics (through logs) became available. Combined with the fact that the free Netscape hitometer service was discontinued in August 2002, the counter system had to be revised. Some of the statistics provided by the hitometer service can be compared to the data from the server logs. A comparison of the count of pages delivered for the period 21 March to 30 June 2002 showed that the hitometer statistics undercounted the actual delivery of pages by an average of 16.2%. This is caused mainly by the fact that those users who do not allow 'cookies' to be placed, were not counted by Netscape's hitometer system.

In May 2003 the site was moved to a new server. After this move, the number of log entries dropped by 17%. Because the new server has a different page request log reporting system. Figure 3 graphs the development of the adjusted statistics, with the post-May

2003 data plotted uncorrected. During October and November 2002 the current maximum monthly usage rate occurred, with over 120,000 pages delivered. During the second part of 2002 and early 2003 the monthly average was around 85,000, and since then the monthly average has declined to just over 70,000.

How many Users?

Given that a single user can request a number of pages in succession, the total number of impressions provides a useful statistic for the relevance of the digital library, but does not provide an insight into the number of actual users. The server log files, however, allow the assessment of the number of unique hosts requesting page impressions. This is equivalent to the number of unique users. Except for situations where a user is dynamically allocated an IP address, as in the case of some dial-up connections, these unique hosts actually represent unique users, irrespective of whether they came and went on visits several days apart. The frequency of unique users as shown in Figure 4 demonstrates an average in excess of 11,000 unique users for the past twelve months.

To place this in context: unique users of the Marshall Islands site are equivalent of about 10% of all unique users accessing the *entire* public webserver of Charles Sturt University, and about 8-9% of all unique users accessing the *entire* public webserver of the University of Melbourne.

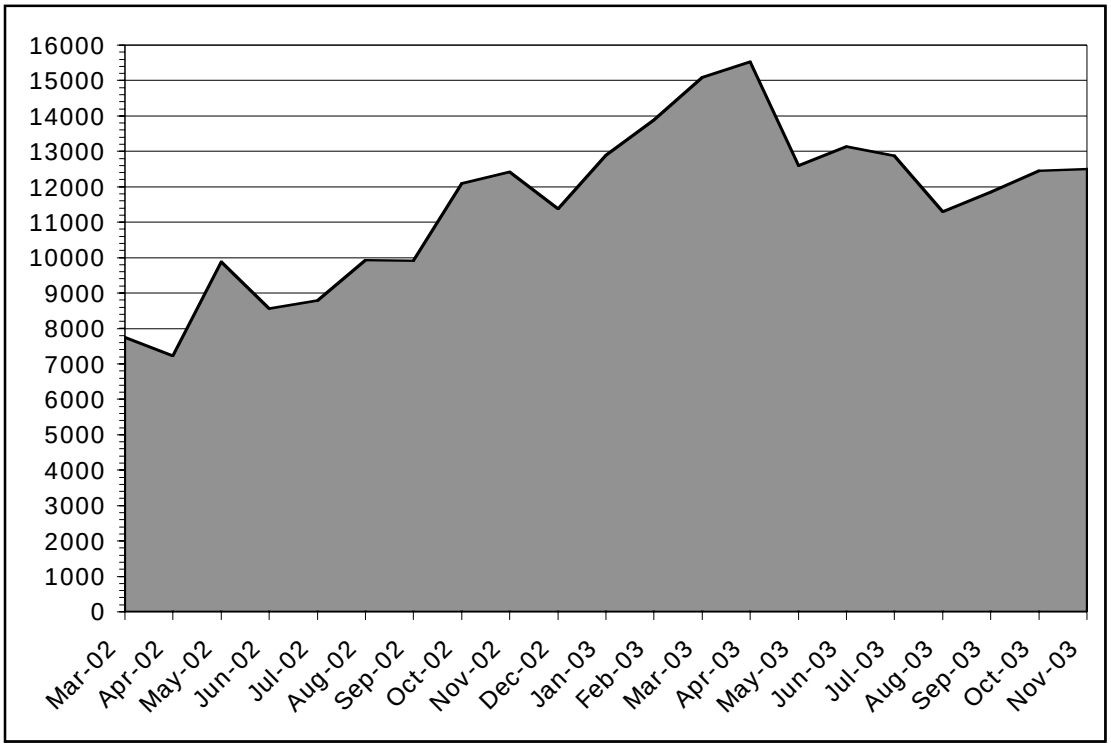

Figure 4. Number of unique users between March 2002 and November 2003

Geographic origin of users

In total, there are 247 extant, and historic, country domains (ISO 2003), 156 of which, or 63.1%, are represented in the logs for the Marshall Islands site since its inception (Table 2; see Table 9 for details for 2002 and 2003 records). These are plotted in Figure 5, showing that with the exception of a few countries in Asia and South America the major area

without requests remains to be Africa. While total number of users is very low there, all African countries have internet access at least in their capital cities (APC 2003).

The volume of page impressions requested can be compared the total number of internet users per country as estimated by Global Reach (2003). Users from US domains dominate the requests, but also have the highest number of internet users. The figures are roughly comparable until we include page requests from.com domains (Table 3) as well. Then the US dominate totally. Apart from the US the greatest discrepancy can be observed for Australia. While it makes up only 1.7% of WWW users, between 17.9 and 21.9% of pages requests originated here. Both the case for the total US as well as for Australia underline, of course, a salient fact: demand for information is not uniform, but driven by the needs of the clientele. The US has historic and current political ties with the Marshall Islands, while the Pacific is, current political rhetoric notwithstanding, essentially Australia's backyard. Lack of interest in the subject matter is understandable due to political and historic factors when we consider the relative under representation of China, the Russian Federation or India. Yet, geographic distance is not necessarily a deciding factor, as the examples of Germany (under represented by 55%) and the Netherlands (over represented by 270%) demonstrate. Despite the Marshalls once being a German colony, it seems to be more the factors of language ability (more Dutch than Germans are fluent in English) and overall intellectual outlook (Germany on the whole currently being more inward looking) that are influencing the use of a digital library.

Some requests are understandable in view of the geographic similarities of the requesting nations, like small island states and countries, such as Malta, the Faroe Islands and several small states in the Caribbean. But why, then, are there requests from some of the poorest countries in Africa without connections to the Marshall Islands, and without geographic similarities, such as Ivory Coast, Ghana, Benin and the like? The answer lies in easy gains to be made from fools: by now everyone will have received the incessant stream of the digital version of the infamous Nigeria mail scam and its variations. It is very likely that at least some these requests are part of an e-mail harvesting program.

Table 2. Geographic origin of users: top fifteen domains (based on 2003)

	People connected		Origin of site user			
	2003		2002		2003	
Site type & Country	N (mill)	%	N	%	n	%
United States (*)	208.4	27.9	45376	34.25	48224	31.71
Australia (au)	12.9	1.7	28980	21.88	27267	17.93
Canada (ca)	19.3	2.6	6092	4.60	8442	5.55
Netherlands (nl)	10.4	1.4	2705	2.04	7867	5.17
United Kingdom (uk)	34.8	4.7	3324	2.51	7616	5.01
Hungary (hu)	1.6	0.2	293	0.22	4653	3.06
Japan (jp)	69.4	9.3	5546	4.19	4388	2.88
Germany (de)	45.5	6.1	3169	2.39	4219	2.77
Switzerland (ch)	4.3	0.6	1284	0.97	3522	2.32
France (fr)	22.6	3.0	3053	2.30	3467	2.28
New Zealand (nz)	2	0.3	2963	2.24	3143	2.07
Italy (it)	22.7	3.0	1781	1.34	2892	1.90
Poland (pl)	6.4	0.9	10253	7.74	2683	1.76
Belgium (be)	4	0.5	2432	1.84	1784	1.17
Austria (at)	3.5	0.5	833	0.63	1640	1.08
TOTAL			132456		152098	

(*) combined total of edu, gov, mil and us domains.

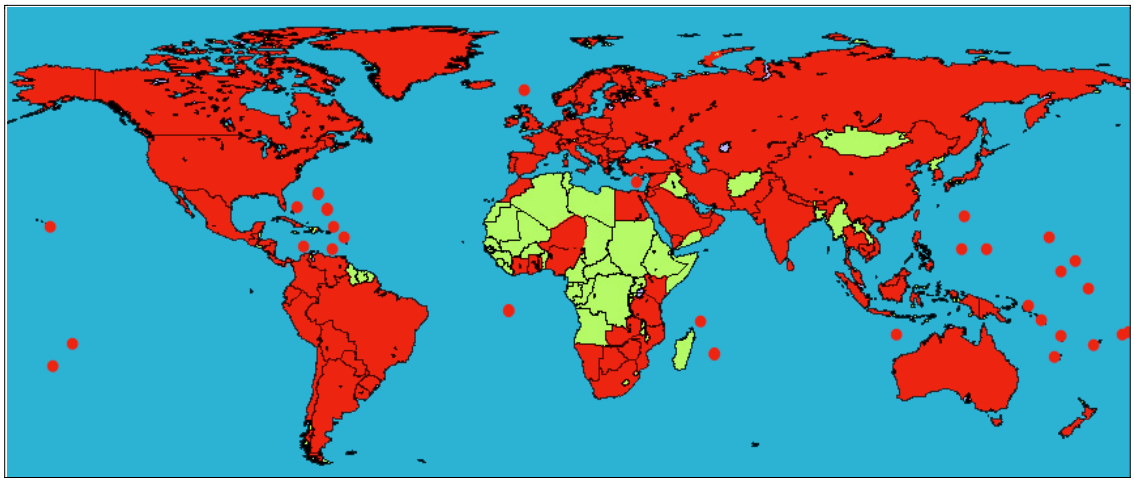

Figure 5. Geographic origin of users (dark areas and dots)

In addition there are users from the domains of Business (biz), Commercial (com), Network (net), Organization (org), International (int) and the like (Table 3). These are usually US addresses, but given the ubiquity of US companies in the global telecommunications market place, some may well be the sole providers for entire countries (cf palaunet.com). A substantial percentage of user origins cannot be resolved at all, as the servers are only identified by their numeric IP address. While it is possible to identify the location of each server using finger programs, it is not economical to do so.

Table 3. Geographic non-specific user domains

	2002	2003
Commercial (com)	395,636	359,665
Network (net)	108,421	154,630
Organization (org)	10,276	8,974
local	—	326
arpa	928	106
int	64	99
info	17	76
biz	20	20
Other (combined)	17	31
TOTAL	515,379	523,920

Some IP addresses of commercial domains (.com,.net etc) allow further breaking down, depending on how the service providers code their regional server hubs. For the major US and Australian providers this allows a breaking down a to t least the state level which is sufficient for the purposes of understanding the origin of users. Table 4 sets out the regional breakdown for the US for the periods July to December 2002 and for 2003 (excluding December), while Table 5 provides the regional breakdown for Australia. For the US the 2003 pattern reveals that populous states, such as New York and California, together with North Carolina and Hawaii are heading the list. The high percentage of Hawaiian users is understandable as the island state for various reasons has connections to the Marshalls and the Pacific in general. The high percentage from North Carolina cannot be explained at this point in time (Table 4). The Australian data (Table 5) reflect population densities across the continent as well as the fact that the eastern seaboard borders the Pacific.

Table 4. Origin of page requests from US commercial providers.

	Jul-Dec 2002		Jan-Nov 2003			Jul-Dec 2002		Jan-Nov 2003	
Alaska	1	0.00	132	0.16	Montana	46	0.17	71	0.08
Alabama	72	0.27	560	0.67	North Carolina	4062	15.39	6290	7.53
Arkansas	27	0.10	502	0.60	North Dakota	37	0.14	269	0.32
Arizona	318	1.20	1025	1.23	Nebraska	85	0.32	153	0.18
California	1534	5.81	11005	13.17	New Hampshire	15	0.06	97	0.12
Colorado	270	1.02	1432	1.71	New Jersey	220	0.83	1279	1.53
Connecticut	52	0.20	303	0.36	New Mexico	168	0.64	417	0.50
Washington DC	36	0.14	435	0.52	Nevada	125	0.47	419	0.50
Delaware	—	—	251	0.30	New York	10717	40.59	13725	16.42
Florida	1126	4.26	4958	5.93	Ohio	480	1.82	2124	2.54
Georgia	310	1.17	1272	1.52	Oklahoma	75	0.28	731	0.87
Hawaii	1382	5.23	8422	10.08	Oregon	370	1.40	1648	1.97
Iowa	131	0.50	307	0.37	Pennsylvania	257	0.97	1759	2.10
Idaho	100	0.38	794	0.95	Rhode Island	41	0.16	212	0.25
Illinois	397	1.50	2271	2.72	South Carolina	112	0.42	605	0.72
Indiana	289	1.09	888	1.06	South Dakota	10	0.04	412	0.49
Kansas	107	0.41	512	0.61	Tennessee	45	0.17	773	0.92
Kentucky	14	0.05	78	0.09	Texas	720	2.73	4677	5.60
Louisiana	142	0.54	387	0.46	Utah	267	1.01	728	0.87
Massachusetts	221	0.84	901	1.08	Virginia	228	0.86	1308	1.57
Maryland	119	0.45	1076	1.29	Vermont	10	0.04	22	0.03
Maine	4	0.02	81	0.10	Washington	394	1.49	2535	3.03
Michigan	402	1.52	1849	2.21	Wisconsin	320	1.21	1390	1.66
Minnesota	380	1.44	1599	1.91	West Virginia	2	0.01	70	0.08
Missouri	121	0.46	709	0.85	Wyoming	6	0.02	51	0.06
Mississippi	33	0.12	59	0.07	TOTAL	26401		83573	

Table 5. Origin of page requests from Australian commercial providers (Jan to Nov 2003).

State		
Australian Capital Territory	72	0.87
New South Wales	3435	41.63
Northern Territory	22	0.27
Queensland	1878	22.76
South Australia	349	4.23
Victoria	2085	25.27
Western Australia	270	3.27
Tasmania	141	1.71
Grand Total	8252	

Another measure of the geographical distribution of the users across the globe is the time of demand on the server to deliver pages. At any given location, the demand is likely to be higher during the waking hours of the average user, peaking during office hours, while troughing out in the early hours of the morning. Thus the demand reflects a diurnal curve. The hourly demand curves for individual locations would cancel each other out were the users evenly distributed across the globe, resulting in a perfectly flat line. In the light of the vast expanse of the Pacific Ocean, this is unlikely to occur. However, as a rule, the flatter the curve, the more geographically even the demand. Figure 6 demonstrates that while in 2001 the demand reflects the diurnal pattern it has become increasingly global since then. Given the overall large percentage of US-based users (Table 2) it is not surprising that the curve maintains a weak diurnal pattern.

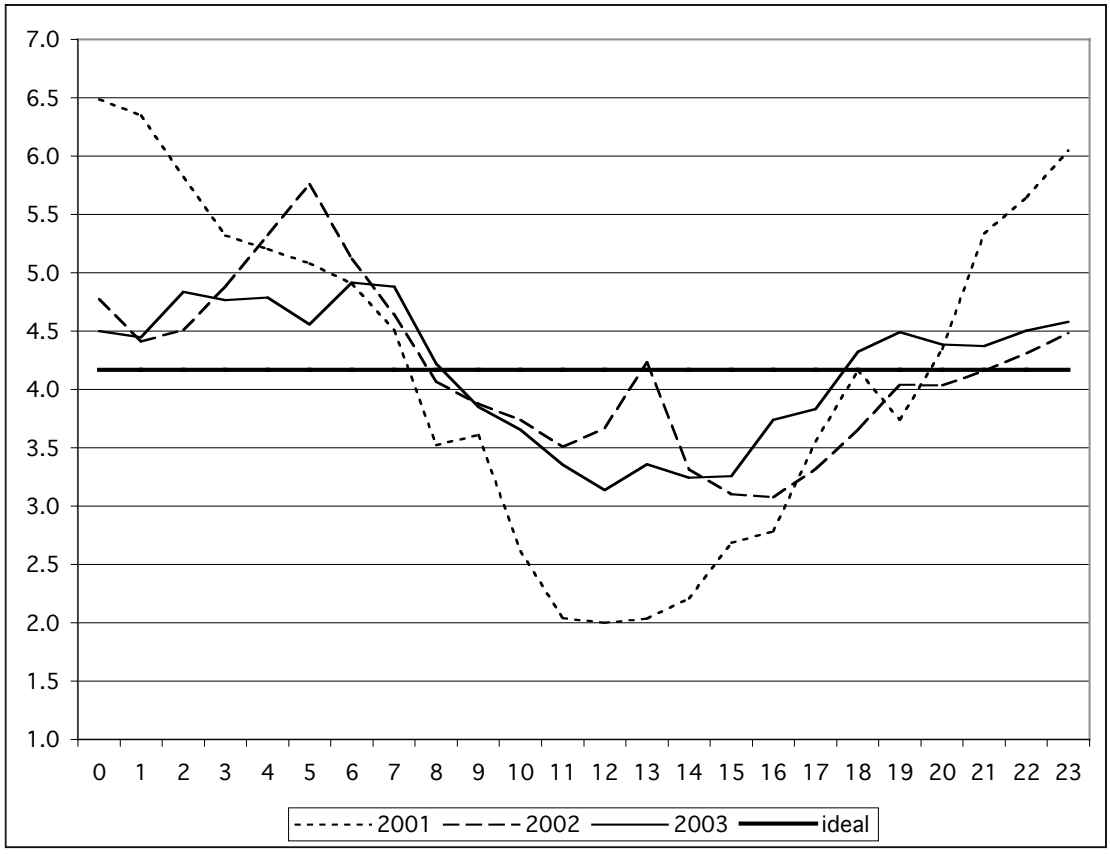

Figure 6. Average hourly demand for pages (in % of requests per day) for 2001 to 2003, compared to the ideal load (time zone is Greenwich Mean Time)

What users are looking for

Any library will provide for a wide range in clientele. The digital library on the Marshall Islands is no different in that regard. Table 6 lists the number of pages, both html and pdf, mounted on the site in November 2003 and compares this with the average monthly requests for pages. The pages have been grouped into major subject areas. A number of pages provide individual historical pictures and their sources. These have been assessed separately. The number of pages by subject area is not necessarily indicative of the relative significance of this area to the manager of the website, but may be a function of the way the information is presented. The nature of postal history and philately, for example, required that the information be broken up more than other subject matter, resulting in a greater number of pages. Table 7 allows for this and provides the average frequency with which each page in the subject area was called up per month.

Table 6. Average number of pages delivered per month (by subject area) between April 2002 and November 2003 in comparison with actual pages mounted on the server.

	raw numbers				percentages			
	number of pages mounted	requests 2002 Apr-Dec	2003 Jan-May	2003 Jun-Nov	number of pages mounted	requests 2002 Apr-Dec	2003 Jan-May	2003 Jun-Nov
Culture, Traditional	376	3,509	12,804	11,376	13.2	11.1	20.0	19.4
History, European	371	5,297	7,133	8,648	13.0	16.8	11.1	14.7
Stamps	713		15,982	8,064	25.1		25.0	13.7
Geography	224	6,060	7,099	7,315	7.9	19.2	11.1	12.5
History, WWII	77	1,180	5,748	4,830	2.7	3.7	9.0	8.2
Bibliographies	39	778	1,859	2,629	1.4	2.5	2.9	4.5
Environment	91	8,552	3,234	1,921	3.2	27.1	5.1	3.3
Health	55	569	1,508	1,457	1.9	1.8	2.4	2.5
Literature	66	395	598	1,399	2.3	1.3	0.9	2.4
Politics	56	564	1,393	1,186	2.0	1.8	2.2	2.0
Economy	68	712	953	1,135	2.4	2.3	1.5	1.9
Site Administration	135	814	1,180	1,070	4.7	2.6	1.8	1.8
Picture Pages								
Geography	10	551	337	151	0.4	1.7	0.5	0.3
History (*)	550	1,454	2,403	5,773	19.3	4.6	3.8	9.8
History, WWII	12	1,065	1,811	1,705	0.4	3.4	2.8	2.9
Average Total	2843	31,501	64,041	58,658	100.0	100.0	100.0	100.0

(*) Historic photographs of both traditional cultural subjects and European Colonial History, excluding World War II.

Table 7. Average frequency with which each page in a category was requested per month

	2002 Apr-Dec	2003 Jan-May	2003 Jun-Nov
Culture, Traditional	9.3	34.1	30.3
History, European	14.3	19.2	23.3
Stamps		22.4	11.3
Geography	27.1	31.7	32.7
History, WWII	15.3	74.6	62.7
Bibliographies	19.9	47.7	67.4
Environment	94.0	35.5	21.1
Health	10.3	27.4	26.5
Literature & Art	6.0	9.1	21.2
Politics	10.1	24.9	21.2
Eonomy	10.5	14.0	16.7
Site Administration	6.0	8.7	7.9
Picture Pages			
Geography	55.1	33.7	15.1
History (*)	2.6	4.4	10.5
History, WWII	88.8	150.9	142.1
Average Total	11.1	22.5	20.6

The adjusted monthly averages show that pages dealing with Word War II as subject matter had the highest request rate with each image page being requested on average 142 times and each text page 63 times. Given anecdotal evidence of WWII buffs on the WWW this was not surprising. What was surprising, however, is the high impression rate for bibliographies (67.4).

At the other, low end of the scale, as expected, are administrative pages (editorial, what's new, file lists, link of the day, etc), but surprisingly also historic images. The statistics show changes over time, with demand on some subject areas, such as the environment, in steady decline and other areas, such as bibliographies and literature and art, on a steady increase. These data allow the editor to respond to demand and update the relevant pages more frequently, but also give pause to rethink the approaches in less frequently requested subject areas.

Table 8. Ranking of the site "The Marshall Islands" in various popular search engines

Site	Search Engine					
	Google	Lycos	Altavista	Yahoo	Webcrawler	Excite
Republic of the Marshall Islands	1	1	1	1	1	1
CIA - The World Factbook -- Marshall Islands	2	9	3	2	2	2
Marshall Islands NIC	3	8	2	3	5	7
The Marshall Islands. An Electronic Library & Archive	**4**	**2**	**4**	**4**	**4**	**5**
Republic of the Marshall Is-Consular Information Sheet	5	42	>50	6	3	4
Lonely Planet World Guide | Destination Marshall Islands	6	13	29	5	11	10
Marshall Islands Stamp Center	7	5	25	7	20	16
Governments on the WWW: Marshall Islands	8	44	18	8	11	9
US Embassy Majuro	9		10	9	14	13
The Marshall Islands, General Inform.- Resources Available	10	49	>50	10	22	17
Total (approx)	3,610,000	2,900,020	112,661	2,470,000	n/a	n/a

Status and Visibility

Before we look at the challenges faced by maintaining such a library, some comments are in order to assess the relative status and visibility of the site. This can be achieved by assessing how search engines, the main access tool to information by the average user, rank the site.

A search for the generic term "Marshall Islands" was run on 17 December 2003 using the search engine Google, which is the most popular engine at the present time. This search found approximately 3,610,000 pages with the term 'Marshall Islands' somewhere on the page. The site discussed here ranked fourth after the official page of the Republic of the Marshall Islands' Embassy in Washington (which because of its official status will always rank first, regardless of breadth or depth of the site itself); the CIA World Factbook which

ranks highly for all smaller countries, again because of its perceived official status; and the official site running the Marshall Islands internet name register which contains no information on the islands at all. Table 8 provides a comparison of the top ten Google rankings with those of six other popular search engines. In all cases the RMI Embassy site was ranked number 1, while the sites under discussion in this paper ranked in the top five.

Challenges

One of the challenges to maintaining such a library and archive is to deal with the number of information requests that flow in. In order to allow users of the web pages the capacity to ask questions, an e-mail address was from the start on provided at the bottom of each page. An unwanted side effect of a WWW presence at the beginning of the twentyfirst century is that websites are being harvested for e-mail addresses by web crawlers and web robots and that these addresses are then sold on to marketers sending out mass e-mails to everyone on the list. As a result of the inclusion of the e-mail on the bottom of pages, the amount of such SPAM mail has increased disproportionately (approximately 30-fold) to that of fellow academics working in the same school. On balance, and with the aid of SPAM filters, it was decided that this is a 'price' worth paying.

Some of the legitimate questions are those faced by any librarian: requests for help expressed at various levels, from professionals requesting detailed information after other avenues have been exhausted, to high school and undergraduate students in essence requesting that someone writes their essay for them. These requests are responded to differentially, given that this library is run by a single academic and not by a public funded library system. Implicitly, demand on the editor's time due to standard academic teaching and research, as well as scholarly community obligations, detract from the ability to develop, expand and maintain such a library.

One of the unexpected side effects of the Marshall Islands digital library site is that it acts as a magnet for other material being 'donated.' While it has long been known that sites that have good resources to offer attract visitors and thus also direct them to other material, the 'honey-pot effect' (Green 1995). The Marshall Islands digital library has demonstrated a variation to the theme. As a result of the exposure of the Marshall Islands site, private archival material has been offered for inclusion in the site, which not only adds to the site, but also enhances the research capacity of the site's editor.

Conclusion

There are a number of ways of accessing information on the WWW with much or the information distributed. While an initial fear of electronic colonisation of smaller countries may be unfounded, the locking up of content on the WWW through increasing commercialisation may lock out many economically weak participants. There is a need for 'free-to-web' published libraries of primary and secondary source materials rather than mere collections links.

On-line technology in the form of server logs allows to study the nature of the library users and also the subject matter requests made by them. If a digital library is to be relevant to the audience it tries to address, such in-depth analyses are a necessity. Such an analysis can result in a strategy how the library 'collections' can be both added to, to cater

for the users, and restructured, to make existing content with low demand more accessible to potential users.

The case study of a digital library on the Marshall Islands is a microcosm, demonstrating world-wide demand for such content. Given that some of the material hosted there are reports of limited distribution in paper form, or transcriptions of archival resources, this library provides an essential service in the dissemination of information.

Acknowledgements

The site started off in 1999 with the publication of three electronic books (*Marshallese Legends and Traditions, Tattooing in the Marshall Islands*, and *Essays on the Marshallese Past*), the raw coding for which was set as class exercises for the subject ITC 130 ' On-Line Publishing'. Fundamental to its success were Prof. David Green and Rob Stocker. Darren Stuart provided some of the further site development and recoding of the students' work into the present form of the three books.

Since then the site has been systematically enlarged by the editor. I am particularly indebted to all authors who graciously made their material available, as well as to the web administrators who shared lists and data collected by them.

References

APC (2003) The Association for Progressive Communications. Information & Communication Technologies (ICTs) Telecommunications, Internet and Computer Infrastructure in Africa. http://www3.sn.apc.org/

Cailliau, Robert (1995). A Little History of the World Wide Web from 1945 to 1995. World Wide Web Consortium URL: http://www.w3.org/History.html

Downing, Jane, Dirk H.R.Spennemann & Margaret Bennett (eds.) (1992) Bwebwenatoon Etto. A collection of Marshallese Legends and Traditions. With illustrations by Amram Enos and Paul Kingsbury. Marshall Islands Culture And History. Series E: Legends, Chants and Proverbs. Vol Nº 1. Majuro, Republic of the Marshall Islands: Historic Preservation Office.

Global Reach (2003a) Details of country/language analysis. Last revised on 30 Sept., 2003. URL: http://global-reach.biz/globstats/details.html

Global Reach (2003b) Global Internet Statistics (by Language). Last revised on 30 Sept., 2003. URL: http://global-reach.biz/globstats/index.php3

Global Reach (2003c) International Online Marketing to Increase Your Web Traffic -- and Sales -- from Abroad http://www.global-reach.biz/eng/index.php3

Green, D.G. (1995). From honeypots to a web of SIN - building the world-wide information system. In: Tsang, P., Weckert, J., Harris, J. and Tse, S. (eds.): *Proceedings of AUUG'95 and Asia-Pacific World Wide Web '95* Conference Charles Sturt University, Wagga Wagga, pp. 11-18.

ISO (2003) International Standards Organisation. VISO 3166 Code lists. English country names and code elements. URL http://www.iso.org/iso/en/prods-services/iso3166ma/02iso-3166-code-lists/list-en1.html

Netscape (2002). About the Open Directory Project. URL: http://dmoz.org/about.html

NLA (2003) National Library of Australia. Pandora Archive - Preserving and Accessing Networked Documentary Resources of Australia. http://pandora.nla.gov.au/index.html

Spennemann, Dirk H.R. (1992) Marshallese Tattoos. Marshall Islands Culture And History. Series C: Traditional Skills and Practices. Volume Nº 1. Majuro Atoll, Republic of the Marshall Islands: Historic Preservation Office. 233pp.

Spennemann, Dirk H.R. (1993) Ennaanin Etto A Collection of essays on the Marshallese Past. Marshall Islands Culture And History. Series F: Technical Studies and Miscellaneous. Volume Nº 1. Majuro Atoll, Republic of the Marshall Islands: Historic Preservation Office. 314.

Spennemann, Dirk H.R., Birckhead, Jim, Green, David G. & Atkinson, John S. (1996) The electronic colonisation of the Pacific. *Computer Mediated Communications Magazine* Vol. 3(2). URL: http://www.december.com/cmc/mag/1996/feb/spen.html

Williams, Stayce and Scharff, Lauren (1998) The Effects of Font Size and Line Spacing on Readability of Computer Displays. URL: http://hubel.sfasu.edu/research/SWExp.html

WWW VL (2003a) The WWW Virtual Library. URL: http://vlib.org/

WWW VL (2003b) Pacific Studies WWW Virtual Library http://coombs.anu.edu.au/WWWVL-PacificStudies.html.

WWW VL (2003c) Education Virtual Library The WWW Virtual Library. URL http://www.csu.edu.au/education/library.html

Youngman, Melissa and Scharff, Lauren (1998) Text Width and Margin Width Influences on Readability of GUIs (Presented at SWPA 1998) URL: http://hubel.sfasu.edu/research/textmargin.html

Table 9. Geographic origin of users of the Marshall Islands site compared to the estimated population of internet users in that country

	People connected		Origin of site user			
	2003		2002		2003	
Site type & Country	N (mill)	%	N	%	n	%
Andorra (ad)			2	>0.01	1	>0.01
Antigua and Barbuda (ag)			2	>0.01	2	>0.01
Albania (al)			1	>0.01	—	—
Algeria (dz)	0.6	0.1	—	—	—	—
American Samoa (as)			20	0.02	5	>0.01
Argentina (ar)	3.9	0.5	468	0.35	441	0.29
Aruba (aw)			—	—	12	0.01
Ascension Island (ac)			—	—	1	>0.01
Australia (au)	12.9	1.7	28980	21.88	27267	17.93
Austria (at)	3.5	0.5	833	0.63	1640	1.08
Azerbaijan (az)			1	>0.01	1	>0.01
Bahamas (bs)			—	—	2	>0.01
Bahrain (bh)	1.7	0.2				
Belarus (by)			9	0.01	7	>0.01
Belgium (be)	4	0.5	2432	1.84	1784	1.17
Belize (bz)			13	0.01	38	0.02
Benin (bj)			—	—	3	>0.01
Bermuda (bm)			3	>0.01	8	0.01
Bhutan (bt)			—	—	1	>0.01
Bolivia (bo)	0.1	0.0				
Bosnia and Herzegovina (ba)			1	>0.01	3	>0.01
Botswana (bw)			4	>0.01	2	>0.01
Brazil (br)	14.3	1.9	689	0.52	1153	0.76
Brunei Darussalam (bn)			16	0.01	17	0.01
Bulgaria (bg)	0.6	0.1	32	0.02	38	0.02
Cambodia (kh)			2	>0.01	—	—
Canada (ca)	19.3	2.6	6092	4.6	8442	5.55
Cayman Islands (ky)			41	0.03	3	>0.01
Chile (cl)	3.1	0.4	84	0.06	119	0.08
China (cn)	68	9.1	20	0.02	11	0.01
Cocos (Keeling) Islands (cc)			6	>0.01	30	0.02
Colombia (co)	1.6	0.2	26	0.02	87	0.06
Cook Islands (ck)			22	0.02	19	0.01
Costa Rica (cr)	0.4	0.1	38	0.03	23	0.02
Cote d'Ivoire (ci)			2	>0.01	—	—
Croatia/Hrvatska (hr)	1	0.1	253	0.19	178	0.12
Cuba (cu)	0.1	0.0	17	0.01	3	>0.01
Cyprus (cy)	0.2	0.0	90	0.07	91	0.06
Czech Republic (cz)	3.1	0.4	956	0.72	1394	0.92
Denmark (dk)	3.4	0.5	659	0.5	1170	0.77
Dominican Republic (do)			47	0.04	49	0.03
Dominica (dm)			1	>0.01	—	—
Ecuador (ec)			4	>0.01	35	0.02
Egypt (eg)	0.6	0.1	20	0.02	58	0.04
Estonia (ee)	0.4	0.1	312	0.24	701	0.46
Faroe Islands (fo)			6	>0.01	39	0.03
Fiji (fj)			241	0.18	820	0.54
Finland (fi)	3.2	0.4	510	0.39	1631	1.07
France (fr)	22.6	3.0	3053	2.3	3467	2.28
French Polynesia (pf)			12	0.01	22	0.01
Georgia (ge)			1	>0.01	—	—
Germany (de)	45.5	6.1	3169	2.39	4219	2.77
Ghana (gh)			—	—	2	>0.01
Gibraltar (gi)			11	0.01	7	>0.01

Greece (gr)	2	0.3	307	0.23	455	0.3
Greenland (gl)			1	>0.01	1	>0.01
Guatemala (gt)			23	0.02	25	0.02
Honduras (hn)			—	—	2	>0.01
HongKong (hk)	4.6	0.6	158	0.12	194	0.13
Hungary (hu)	1.6	0.2	293	0.22	4653	3.06
Iceland (is)	0.2	0.0	45	0.03	159	0.10
India (in)	16.6	2.2	118	0.09	71	0.05
Indonesia (id)	4	0.5	153	0.12	209	0.14
Iran (ir)	2	0.3	6	>0.01	1	>0.01
Ireland (ie)	1.1	0.1	50	0.04	103	0.07
Isle of Man (im)			—	—	1	>0.01
Israel (il)	3.7	0.5	227	0.17	347	0.23
Italy (it)	22.7	3.0	1781	1.34	2892	1.9
Jamaica (jm)			11	0.01	4	>0.01
Jersey (je)			6	>0.01	—	—
Japan (jp)	69.4	9.3	5546	4.19	4388	2.88
Jordan (jo)	0.2	0.0	1	>0.01	1	>0.01
Kenya (ke)			3	>0.01	11	0.01
Kiribati (ki)			4	>0.01	—	—
Korea, Republic of (kr)	28.6	3.8	34	0.03	22	0.01
Kyrgyzstan (kg)			—	—	3	>0.01
Kuwait (kw)	0.2	0.0				
Latvia (lv)	0.2	0.0	48	0.04	46	0.03
Lebanon (lb)	0.4	0.1	7	0.01	15	0.01
Liechtenstein (li)			—	—	1	>0.01
Lithuania (lt)	0.3	0.0	380	0.29	244	0.16
Luxembourg (lu)	0.2	0.0	44	0.03	60	0.04
Macau (mo)			—	—	1	>0.01
Macedonia (mk)			4	>0.01	3	>0.01
Malaysia (my)	6.7	0.9	196	0.15	207	0.14
Malta (mt)			111	0.08	101	0.07
Marshall Islands (mh)			2	>0.01	—	—
Martinique (mq)			—	—	2	>0.01
Mauritius (mu)			78	0.06	36	0.02
Mexico (mx)	6.7	0.9	671	0.51	622	0.41
Micronesia, Federated States (fm)			276	0.21	317	0.21
Moldova, Republic of (md)			—	—	4	>0.01
Morocco (ma)	0.5	0.1	—	—	1	>0.01
Mozambique (mz)			1	>0.01	3	>0.01
Namibia (na)			10	0.01	10	0.01
Nauru (nr)			26	0.02	2	>0.01
Nepal (np)			10	0.01	9	0.01
Netherlands (nl)	10.4	1.4	2705	2.04	7867	5.17
New Caledonia (nc)			4	>0.01	28	0.02
New Zealand (nz)	2	0.3	2963	2.24	3143	2.07
Nicaragua (ni)			14	0.01	2	>0.01
Niger (ne)			3	>0.01	—	—
Nigeria (ng)			3	>0.01	—	—
Niue (nu)			27	0.02	10	0.01
Northern Mariana Islands (mp)			—	—	4	>0.01
Norway (no)	2.45	0.3	371	0.28	608	0.40
Oman (om)	0.1	0.0	2	>0.01	—	—
Panama (pa)			3	>0.01	—	—
Pakistan (pk)	1.2	0.2	37	0.03	34	0.02
Papua New Guinea (pg)			60	0.05	78	0.05
Paraguay (py)			—	—	3	>0.01
Peru (pe)	3	0.4	16	0.01	17	0.01
Philippines (ph)	2	0.3	248	0.19	317	0.21
Poland (pl)	6.4	0.9	10253	7.74	2683	1.76

Portugal (pt)	4.4	0.6	329	0.25	583	0.38
Puerto Rico (pr)	0.6	0.1				
Qatar (qa)	0.1	0.0	—	—	3	>0.01
Romania (ro)	2	0.3	274	0.21	341	0.22
Russian Federation (ru)	18	2.4	495	0.37	823	0.54
Rwanda (rw)			—	—	2	>0.01
Saint Kitts and Nevis (kn)			—	—	1	>0.01
Saint Lucia (lc)			1	>0.01	—	—
Saudi Arabia (sa)	1.6	0.2	321	0.24	316	0.21
Serbia (cs)	0.3	0.0				
Seychelles (sc)			—	—	3	>0.01
Singapore (sg)	2.8	0.4	950	0.72	976	0.64
Slovak Republic (sk)	0.9	0.1	98	0.07	166	0.11
Slovenia (si)	0.6	0.1	60	0.05	39	0.03
Solomon Islands (sb)			9	0.01	11	0.01
South Africa (za)	3.1	0.4	412	0.31	302	0.20
Spain (es)	20	2.7	1342	1.01	1117	0.73
Sri Lanka (lk)	6.7	0.9	—	—	3	>0.01
Sweden (se)			863	0.65	1178	0.77
Switzerland (ch)	4.3	0.6	1284	0.97	3522	2.32
Syrian Arab Republic (sy)	0.1	0.0	—	—	5	>0.01
Taiwan (tw)	11.6	1.6	225	0.17	782	0.51
Tanzania (tz)			152	0.11	2	>0.01
Thailand (th)	2.3	0.3	—	—	314	0.21
Tonga (to)			1	>0.01	2	>0.01
Trinidad and Tobago (tt)			43	0.03	40	0.03
Tunisia (tn)	0.5	0.1				
Turkey (tr)	4.9	0.7	159	0.12	187	0.12
Turks and Caicos Islands (tc)			—	—	5	>0.01
Tuvalu (tv)			7	0.01	6	>0.01
Ukraine (ua)	0.8	0.1	44	0.03	52	0.03
United Arab Emirates (ae)	1.2	0.2	14	0.01	154	0.1
United Kingdom (uk)	34.8	4.7	3324	2.51	7616	5.01
United States (us)	208.4	27.9	8616	6.5	9277	6.1
Uruguay (uy)	0.4	0.1	52	0.04	45	0.03
US Educational (edu)			16964	12.81	18367	12.08
US Government (gov)			3712	2.8	4383	2.88
US Military (mil)			16084	12.14	16197	10.65
Uzbekistan (uz)			—	—	6	>0.01
Vanuatu (vu)			2	>0.01	17	0.01
Venezuela (ve)	1.3	0.2	10	0.01	29	0.02
Vietnam (vn)	1.5	0.2	3	>0.01	8	0.01
Virgin Islands (USA) (vi)			11	0.01	6	>0.01
Yemen (ye)	0	0.0	1	>0.01	—	—
Yugoslavia (yu)			44	0.03	65	0.04
Western Samoa (ws)			3	>0.01	77	0.05
Zambia (zm)			3	>0.01	1	>0.01
Zimbabwe (zw)			2	>0.01	3	>0.01
TOTAL	746.75	100.0	132456		152098	