

**a short report on digitised
newspapers including web
analytics and other marvelous
stuff**

Frederick Zarndt, Chair
IFLA Newspapers Section
sponsored by
CCS, Digital Divide Data, DL Consulting












Photo by DAVID ILIFF. License: CC-BY-SA 3.0

reading rooms by the numbers

		Monthly average			
			Visitors		
		Population	Reading Room	Microform	Print
	Australia	22,876,000	5,130	345	240
	France	65,350,000	3,000	2,000	1,000
	Netherlands	16,847,000	NA	NA	NA
	New Zealand	4,414,000	NA	NA	NA
	Norway	4,985,000	600	400	NA
	Singapore	5,184,000	NA	300	NA
	UK	62,262,000	2,000	6,900	4,816
	USA	313,292,000	NA	NA	NA













digitised newspapers by the numbers

	Monthly average				
	Digitised Historical Newspapers				
	Population	Unique Visitors	Genealogist	Other	User Age
	22,876,000	150,000	50%	50%	>55
	37,692,000	12,800	65%	35%	>50
	5,405,000	NA	NA	NA	?
	65,350,000	22,000	NA	NA	?
	16,847,000	50,000	NA	NA	?
	4,414,000	83,333	50%	NA	>50
	4,985,000	1,500	NA	NA	?
	5,184,000	12,400	NA	NA	?
	62,262,000	NA	NA	NA	?
	313,292,000	NA	NA	NA	?



physical versus digital



		Monthly average	
		Requests for Newspapers Paper + Microform	Digitised Historical Newspapers Unique Visitors
	Population 22,876,000	585	150,000
	37,692,000	NA	12,800
	5,405,000	NA	NA
	65,350,000	3,000	22,000
	16,847,000	NA	50,000
	4,414,000	NA	83,333
	4,985,000	400	1,500
	5,184,000	300	12,400
	62,262,000	11,716	NA
	313,292,000	NA	NA



more numbers!

				Monthly average					
	Collection			Digitised Historical Newspapers					
	Population	Name	~Size [pages]	Unique Visitors	Genealogist	Other	Lines Corrected	User Age	
	22,876,000	Trove	5,000,000	150,000	50%	50%	220,000	>55	
	37,692,000	CDNC	495,000	12,800	65%	35%	31,000	>50	
	5,405,000	Historical Newspaper Library	2,000,000	NA	NA	NA	NA	?	
	65,350,000	Gallica	2,200,000	22,000	NA	NA	NA	?	
	16,847,000	Historische Kranten	5,000,000	50,000	NA	NA	NA	?	
	4,414,000	Papers Past	2,213,000	83,333	50%	NA	NA	>50	
	4,985,000	NBDigital Aviser	8,100,000	1,500	NA	NA	NA	?	
	5,184,000	Newspaper SG	2,400,000	12,400	NA	NA	NA	?	
	62,262,000	British Newspaper Archive	4,880,000	NA	NA	NA	NA	?	
	313,292,000	Chronicling America	4,100,000	NA	NA	NA	NA	?	

what is Alexa?











- Alexa collects and analyzes Internet data for purposes of web analytics. Web analytics is the measurement, collection, analysis and reporting of Internet data for the purposes of understanding and optimizing web usage. Alexa is now a subsidiary of Amazon.
- Alexa was founded in 1996 by Brewster Kahle (Internet Archive) and Bruce Gilliat.
- Alexa operations includes archiving of webpages as they are crawled. This database served as the basis for the creation of the Internet Archive accessible through the Wayback Machine.
- Alexa continually crawls all publicly-available websites to create a series of snapshots of the web.
- Alexa gathers information from a variety of sources to provide key statistics about each site on the web, for example, Traffic **Rank**, the number of **PageViews**, and site **Speed**, **Bounce Rate**, etc. This information is derived from Alexa toolbar users (~6,000,000 worldwide).

definitions











- A **PageView** is a request for a file whose type is defined as a page.
- A **Unique Visitor** is a uniquely identified client generating requests on the web server or viewing pages within a defined time period (i.e. day, week or month). A Unique Visitor counts once within the timescale.
- A **Visit** is a series of page requests from the same uniquely identified client with a time of no more than 30 minutes between each page request.
- **Bounce Rate** is the percentage of visits where the visitor enters and exits at the same page without visiting any other pages on the site in between.
- World | Country **Rank** is a function of the average daily unique visits and the number of unique pages requested.

definitions adapted from Wikipedia http://en.wikipedia.org/wiki/Web_analytics



Alexa ranking world view

Alexa 3 month trailing averages 2-Apr-2012			
	Population	Website	World rank [Lo is good]
	313,292,000	http://www.loc.gov/index.html/	3,122
	22,876,000	http://trove.nla.gov.au/	16,700
	65,350,000	http://www.bnf.fr/	17,096
	62,262,000	http://www.bl.uk/	27,079
	4,414,000	http://www.natlib.govt.nz/	123,976
	62,262,000	http://www.britishnewspaperarchive.co.uk/	155,259
	16,847,000	http://www.kb.nl/	155,363
	5,184,000	http://www.nl.sg/	156,610
	4,985,000	http://www.nb.no/	189,940
	5,405,000	http://www.nationallibrary.fi/	3,212,803

Alexa ranking country view











Alexa 3 month trailing averages 2-Apr-2012				
	Population	Website	World rank [Lo is good]	Country rank [Lo is good]
	5,405,000	http://www.nationallibrary.fi/	3,212,803	199
	22,876,000	http://www.nla.gov.au/	16,700	375
	4,414,000	http://www.natlib.govt.nz/	123,976	515
	65,350,000	http://www.bnf.fr/	17,096	727
	4,985,000	http://www.nb.no/	189,940	891
	313,292,000	http://www.loc.gov/index.html/	3,122	1,011
	5,184,000	http://www.nl.sg/	156,610	1,208
	62,262,000	http://www.bl.uk/	27,079	2,245
	16,847,000	http://www.kb.nl/	155,363	3,450
	62,262,000	http://www.britishnewspaperarchive.co.uk/	155,259	15,692

where visitors go

Alexa 3 month trailing averages 2-Apr-2012					
	Population	World rank [Lo is good]	Country rank [Lo is good]	Where visitors go [sub-domain]	
	5,405,000	3,212,803	199	NA	NA
	22,876,000	16,700	375	http://trove.nla.gov.au/	57.2%
	4,414,000	123,976	515	http://paperspast.natlib.govt.nz/	50.9%
	65,350,000	17,096	727	http://gallica.bnf.fr/	52.0%
	4,985,000	189,940	891	NA	NA
	313,292,000	3,122	1,011	http://chroniclingamerica.loc.gov/	4.8%
	5,184,000	156,610	1,208	http://newspapers.nl.sg/	28.0%
	62,262,000	27,079	2,245	http://newspapers11.bl.uk/blcs/	2.5%
	16,847,000	155,363	3,450	http://kranten.kb.nl/	22.4%
	62,262,000	155,259	15,692	NA	NA












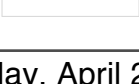
lots of numbers

(sorted by time on site)

		Alexa 3 month trailing averages 2-Apr-2012				
Website		Speed [Hi is good]	Bounce rate [Lo is good]	Reputation [Hi is good]	Page views per visitor [Hi is good]	Time on site [Hi is good]
	http://www.britishnewspaperarchive.co.uk/	51%	28%	485	13.0	11m 40s
	http://www.bnf.fr/	71%	35%	13,744	14.9	8m 30s
	http://www.natlib.govt.nz/	96%	44%	2,480	5.3	6m 49s
	http://trove.nla.gov.au/	42%	55%	9,514	5.4	4m 52s
	http://www.loc.gov/index.html/	67%	51%	91,331	5.3	3m 55s
	http://www.kb.nl/	89%	54%	3,295	5.0	3m 42s
	http://www.bl.uk/	54%	52%	16,191	3.8	3m 2s
	http://www.nb.no/	59%	47%	1,579	3.0	2m 57s
	http://www.nationallibrary.fi/	NA	54%	199	3.1	2m 6s
	http://www.nl.sg/	72%	65%	802	2.0	2m 4s

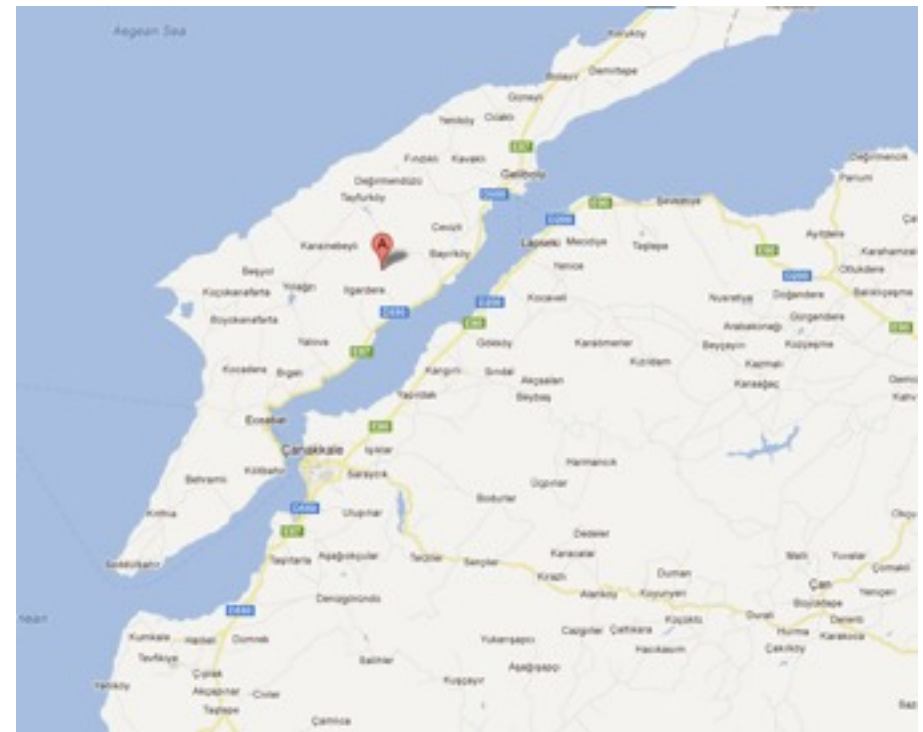
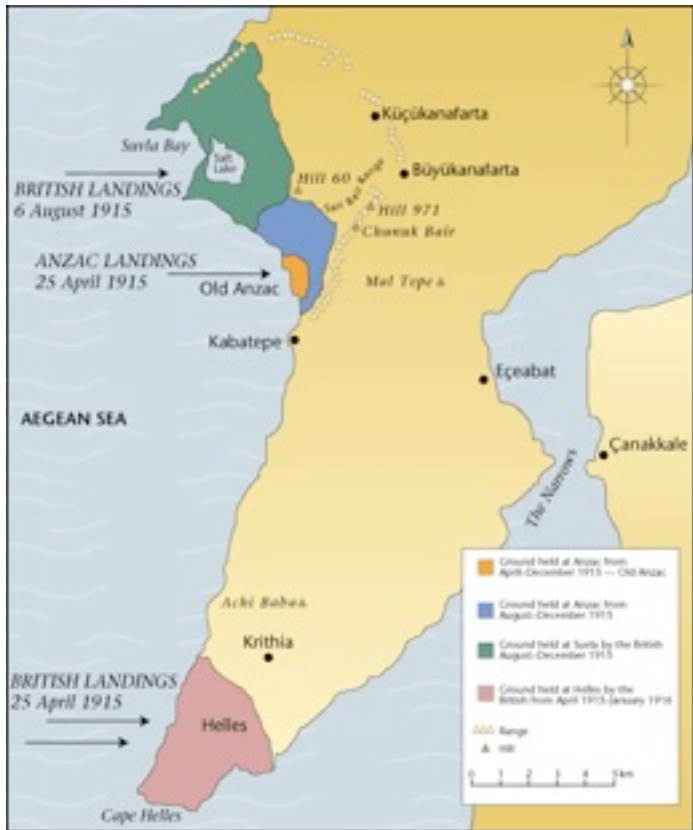
even more numbers

(sorted by time on site)

		Alexa 3 month trailing averages 2-Apr-2012				
Website		Speed [Hi is good]	Bounce rate [Lo is good]	Reputation [Hi is good]	Page views per visitor [Hi is good]	Time on site [Hi is good]
	http://www.ancestry.com/	32%	24%	20,055	29.9	23m 54s
	http://www.familysearch.org/	50%	18%	9,832	15.8	16m 19s
	http://www.britishnewspaperarchive.co.uk/	51%	28%	485	13.0	11m 40s
	http://www.bnf.fr/	71%	35%	13,744	14.9	8m 30s
	http://www.natlib.govt.nz/	96%	44%	2,480	5.3	6m 49s
	http://trove.nla.gov.au/	42%	55%	9,514	5.4	4m 52s
	http://www.loc.gov/index.html/	67%	51%	91,331	5.3	3m 55s
	http://www.kb.nl/	89%	54%	3,295	5.0	3m 42s
	http://www.bl.uk/	54%	52%	16,191	3.8	3m 2s
	http://www.nb.no/	59%	47%	1,579	3.0	2m 57s
	http://www.nationallibrary.fi/	NA	54%	199	3.1	2m 6s
	http://www.nl.sg/	72%	65%	802	2.0	2m 4s

ok, so what! not sure i
believe that Alexa stuff. is
that the best you can do?

Gallipoli Campaign April 1915 to January 1916 aka Battle of Gallipoli Dardanelles Campaign Battle of Çanakkale





search








(battle OR campaign)

AND

(Gallipoli OR Dardenelles OR Çanakkale)

date range 1-Jan-1915 to 31-Dec-1916

search results

	URL	Collection Name	~Size [pages]	Number of results
	http://trove.nla.gov.au	Trove	5,000,000	16,131 articles
	http://cdnc.ucr.edu	CDNC	495,000	6 articles
	http://www.nationallibrary.fi/	Historical Newspaper Library	2,000,000	333 results
	http://gallica.bnf.fr	Gallica	2,200,000	222 results
	http://kranten.kb.nl	Historische Kranten	5,000,000	34,399 articles
	http://paperspast.natlib.govt.nz	Papers Past	2,213,000	5,910 articles
	http://www.nb.no/aviser/	NBDigital Aviser	8,100,000	539 articles
	http://newspapers.nl.sg	Newspaper SG	2,400,000	689 articles
	http://britishnewspaperarchive.com	British Newspaper Archive	4,880,000	97 articles
	http://chroniclingamerica.loc.gov	Chronicling America	4,100,000	82,265 hits

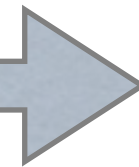
search words

(battle OR campaign)

AND

(Gallipoli OR Dardenelles OR Çanakkale)

date range 1-Jan-1915 to 31-Dec-1916



	http://www.google.com/
	http://www.google.co.uk/
	http://www.google.com.au/
	http://www.google.co.nz/
	http://www.google.com.sg/



Google Search

I'm Feeling Lucky

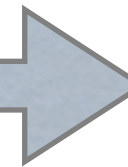
search results

(battle OR campaign)

AND

(Gallipoli OR Dardanelles OR Çanakkale)

date range 1-Jan-1915 to 31-Dec-1916



	http://www.google.com/
	http://www.google.co.uk/
	http://www.google.com.au/
	http://www.google.co.nz/
	http://www.google.com.sg/

IN 1st 100 GOOGLE RESULTS, NOT A SINGLE RESULT FROM A LIBRARY!



search words

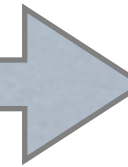
News

(battle OR campaign)

AND

(Gallipoli OR Dardenelles OR Çanakkale)

date range 1-Jan-1915 to 31-Dec-1916











	http://news.google.com/
	http://news.google.co.uk/
	http://news.google.com.au/
	http://news.google.co.nz/
	http://news.google.com.sg/
	http://news.google.no/
	http://news.google.nl/
	http://news.google.fr/

search results

(battle OR campaign)
AND
(Gallipoli OR Dardenelles OR Çanakkale)

date range 1-Jan-1915 to 31-Dec-1916

			Results mostly from
	http://news.google.com/	8,210 results	Papers Past, NY Times, Google News
	http://news.google.co.uk/	3,340 results	Papers Past, NY Times, Google News
	http://news.google.com.au/	3,340 results	Papers Past, NY Times, Google News
	http://news.google.co.nz/	3,340 results	Papers Past, NY Times, Google News
	http://news.google.com.sg/	3,340 results	Papers Past, NY Times, Google News
	http://news.google.fr/	0 results	
	http://news.google.no/	0 results	
	http://news.google.nl/	0 results	

in conclusion

know your customers

make your stuff visible

Merci

Frederick Zarndt

frederick@frederickzarndt.com

sponsored by

CCS, Digital Divide Data, DL Consulting