

# Assessment of Digital Reference Tools/Websites Related to Tourism Information using Webometrics and Qualitative Analysis

Shiv Shakti Ghosh\* and Sunil Kumar Chatterjee

Department of Library and Information Science, Jadavpur University, Jadavpur – 700032, Kolkata, India;  
ghoshshiv@gmail.com

## Abstract

The study provides an assessment of digital reference tools/websites disseminating tourism information. It also notes the keywords and other drivers which trigger tourism information seeking from digital reference tools/websites.

**Keywords:** Digital Reference Tools, Impact Factor, Qualitative Analysis, Tourism Information, Webometrics

## 1. Introduction

The ALA Glossary of Library and Information Science defines a reference source as “A book designed by the arrangement and treatment of its subject matter to be consulted for definitive items of information rather than to be read consecutively”. This definition can be comfortably extended to the case of digital reference tools. Digital reference tools have additional features opening up new vistas for research, and sufficing the very requirements of reference service to survive in the digital environment. Further, they also help librarians to serve users better. Free Digital reference tools help libraries overcome the resource crunch faced during procurement of traditional reference tools. Digital reference tools are available in huge numbers today. Thus, it is crucial to identify the tools required to satisfy the needs of the user. In this study, the tools related to tourism domain have been chosen for evaluation.

The World Tourism Organization views tourism as a “key driver of the global economic recovery and a vital contributor to job creation, poverty alleviation, environmental protection and multicultural peace and understanding across the globe” (“Why tourism? World Tourism Organization UNWTO, n.d.). Recent years have seen the proliferation of travel related information

through personal blogs, commercial websites and social media. This calls for a commentary on the quality of the tools that disseminate tourism information.

## 2. Methodology/Approach

For a qualitative analysis of the digital reference tools/websites we have listed their dominant features. SocSciBot (2018, n.d.) has been used to crawl the websites and the number of outlinks from each website has been counted. Data about the inlinks and keywords that bring search traffic to each website have been gathered from Alexa (“Website Traffic”, 2018, n.d.).

## 3. Limitations

The free version of the crawler SocSciBot which has been used in this study cannot crawl links from secure websites, thus, some popular websites could not be included. Also, freely available information from the website Alexa has been used.

## 4. Analysis

The following features have been identified from the websites:

\*Author for correspondence

- *Country Studies*: It contains detailed description on historical, economic, political, geographical aspects of all selected countries (“Country Studies”, n.d.).
- *UNEP Environmental Data Explorer*: This is an initiative by United Nations Environment Programme. Spatial and Non-spatial data on population, climate, forests, emissions etc. are displayed (“UNEP Environmental Data Explorer-The Environmental Database”, n.d.).
- *LinkedGeoData.org*: It is a RDF knowledge base of spatial data (“linkedgeodata.org: datasets”, n.d.) formed according to the Linked Data principles and interlinked with DBpedia.org and GeoNames.org.
- *NASA’s Socioeconomic Data and Applications Center (SEDAC)*: It provides free GIS data about interactions of human being with environment. Maps can also be viewed here (“Socioeconomic Data and Applications Center. SEDAC”, n.d.).
- *World by Map*: Entries begin with maps and name of countries which are arranged in descending order on the basis of different statistical parameters such as birth rate, growth rate, education rate, total area of land etc. (“World by Map: Statistics, Maps and Charts”, n.d.)
- *Atlapedia Online*: Atlapedia Online contains full color physical maps, political maps, key facts and statistics on countries, etc. (“Atlapedia Online”, n.d.).
- *Columbia Gazetteers*: With over 170,000 entries, the Columbia Gazetteer of the World is an authoritative A to Z encyclopedia of geographical places and features (“Columbia Gazetteers”, n.d.).
- *Diva GIS*: Provides access to GIS datasets that are related to boundaries of countries, transport root, climate data, crop collection data etc. (“Free Spatial Data. DIVA-GIS”, n.d.).
- *Fallingrain.com*: This gazetteer covers information about countries worldwide and their states, cities and places (“Global gazetteers”, n.d.).
- *Gazetteer.org.uk*: Provides an exhaustive Place Name Index to Great Britain (“Gazetteers of British place names”, n.d.).
- *Geography-site.co.uk*: Provides information on geography and other topics related to it. (“The Geography Site”, December, 11, 2012).
- *Geonames.org*: It contains geographical names, alternate place names, features related to a place etc. GeoNames is integrating geographical data such as names of places in various languages, elevation, population and others from various sources (“GeoNames”, n.d.).
- *India Tour*: It is a comprehensive travel and tour guide on India (“India Tour”, n.d.).
- *mapcruzin.com*: Provides free and open source city, state and national level maps, geospatial data and shape files (“mapcruzin”, n.d.).
- *One World-Nations Online*: An online guide to cultures and nations around the world. Contains information on world population, languages, cities of the world, international codes, flags, international currencies, airlines & airports, news, famous monuments, places of worship or power etc. (“One World-Nations Online”, n.d.).
- *Natural Earth Data*: Best suited for cartographers. Data has been classified according to scale, namely Large, Medium and Small which is aimed to serve different needs. It supports cultural data, base map data and physical data (“Natural Earth”, n.d.).
- *Open Topography*: Provides access to high spatial resolution topographic data and tools. It also provides the sources where we can find LiDAR (Light and Radar) data (“NSF Open Topography”, n.d.).
- *Travelinfoz*: It contains travel map, information on food and drinks, place to go, history of tourist place, flights, hotels, sights, vacation, night clubs, best time to travel, language and culture, airports, weather report, travel articles etc. (“Travelinfoz”, n.d.).
- *World Travel Guide*: With the description of tourist place, hotels, this guide-book provides road map, video of famous places, latest news and features, reviews, information on airline, local people and culture etc. (“World Travel Guide”, n.d.).

Table 1 shows the number of inlinks, keywords used by users which lead them to the websites included in our study and the rank of the websites based on the inlinks:

**Table 1.** Number of inlinks, keywords associated with each website (“Website Traffic”, 2018, n.d.)

Site	Keywords (Decreasing percentage of search traffic)	Inlinks	Rank
countrystudies.us	South Korea us relations How the economy works Germany GDP Green revolution	2,302	4
geodata.grid.unep.ch	Lake balkhash Environmental baseline study French southern Territories Environmental impact assessment GEO data explorer	2,140	5
linkedgeodata.org	Open GEO data Sparql endpoint Geodata.org Sparql endpoint Example	90	16
sedac.ciesin.columbia.edu	Columbia University Columbia Canvas Columbia	60,456	1
bymap.org	Literacy rate Age of countries Australia land area Area of USA Average age in Japan	149	14
atlapedia.com	Where is Singapore Laos capital Where is Nepal What country borders Monaco Latvia map	1,425	8
columbiagazetteer.org	Gazetteer Columbia gazetteer of The world World gazetteer Gazetteer online	121	15
diva-gis.org	Diva GIS Simple GIS software Free GIS data Spatial data download	627	10
fallingrain.com	Raipur Pakistan Fallingrain Weather zaaihoek dam	1,458	7
gazetteer.org.uk	Gazetteer British place names UK place names English place names	35	19

geography-site.co.uk	Rock flour Valley glacier Flow map Isoline map	254	12
geonames.org	Postal code California zip code New York postal code California zip codes	7,391	2
www.india-tour.com	Indian tour North India tour Packages	38	18
mapcruzin.com	World map high Resolution High resolution world Map Free map Israel	546	11
nationsonline.org	US map United States map USA map Map of Europe	2,817	3
naturalearthdata.com	Natural earth Natural earth data Natural earth populated Places	1,096	9
opentopography.org	Open topography Open topo Open lidar	250	13
travelinfoz.com	Best time to visit Singapore Best time of year to visit Miami	43	17
worldtravelguide.net	Singapore currency Tanzania visa Germany weather	1,539	6

Table 2 shows the number of pages and the number of outlinks associated with each website.

**Table 2.** Number of pages, outlinks associated with each website (SocScibot, 2018 n.d.)

Site	Pages	Outlinks
countrystudies.us	12045	43599
geodata.grid.unep.ch	43	107
linkedgeodata.org	358	1572
sedac.ciesin.columbia.edu	5245	60127
world.bymap.org	81	159
atlapedia.com	13	68
columbiagazetteer.org/main/	60	1135
diva-gis.org	72	843
www.fallingrain.com/world	263331	3127489

www.gazetteer.org.uk	9	38
www.geography-site.co.uk	5280	62235
www.geonames.org	13710	136144
www.india-tour.com	429	16549
www.mapcruzin.com	19622	354419
www.nationsonline.org	2742	210563
www.naturalearthdata.com	2074	37170
opentopography.org	247	1964
www.travelinfoz.com	547	26291
www.worldtravelguide.net	6	15

The formulas which have been used to calculate three types of WIFs are as follows:

- Simple Web Impact Factor(SWIF): Total number of links / Total web pages;
- External Link Web Impact Factor(ELWIF): Total number of external links / Total web pages;
- Inlink Web Impact Factor or ILWIF : Total number of inlinks / Total web pages

After the calculation of WIFs, the correlation coefficient of these websites has been calculated using the following formula:

$$\text{Correlation}(r) = (N \sum pq - (\sum p)(\sum q)) / (\text{sqrt}((N \sum p^2 - (\sum p)^2) (N \sum q^2 - (\sum q)^2)))$$

where, p,q are variables and N is the no of values that p and q have.

Table 3 shows the necessary calculations for SWIF, ELWIF, SLWIF.

**Table 3.** Table showing the necessary calculations for SWIF, ELWIF, ILWIF, correlation coefficient

Digital Reference Tool/Website	countrystudies.us	geodata.grid.unep.ch	linkedgeodata.org	sedac.ciesin.columbia.edu	world.bymap.org	atlapedia.com	columbiagazetteer.org/main/	diva-gis.org	www.fallingrain.com/world
SWIF* ELWIF	13.79384	130.0319	20.38532	263.551	7.464106	600.7337	395.9889	1782579	141.1205
ILWIF* ELWIF	0.69178	123.8399	1.103898	132.135	3.610882	573.3728	38.14861	101.9601	0.065758
SWIF* ILWIF	0.728306	2600.638	1.167098	264.9931	6.994665	12588.91	42.21556	1325833	0.065789
(ELWIF)^2	13.10206	6.191996	19.28142	131.416	3.853224	27.36095	357.8403	137.0851	141.0547
(ILWIF)^2	0.036526	2476.798	0.0632	132.858	3.383783	12015.53	4.066944	75.83507	0.000031
(SWIF)^2	14.52214	2730.67	21.55242	528.544	14.45877	13189.64	438.2044	416.8402	141.1863
ELWIF	3.619676	2.488372	4.391061	11.46368	1.962963	5.230769	18.91667	11.70833	11.87665
ILWIF	0.191117	49.76744	0.251397	11.52641	1.839506	109.6154	2.016667	8.708333	0.005537
SWIF	3.810793	52.25581	4.642458	22.99009	3.802469	114.8462	20.93333	20.41667	11.88218
Total links	45,901	2,247	1,662	120,583	308	1,493	1,256	1,470	3,128,947

Table 3 Contd...

Digital Reference Tool/Website	www.gazetteer.org.uk	www.geography-site.co.uk	www.geonames.org	www.india-tour.com	www.mapcruz.in.com	www.nationsonline.org	www.natureearthdata.com	opentopography.org	www.travelinfoz.com	www.worldtravelguide.net
SWIF* ELWIF	34.24691	139.4988	103.9636	1491.506	326.7503	5975.866	330.6649	71.27303	2313.925	647.5
ILWIF* ELWIF	16.41975	0.567023	5.353364	3.416967	0.502601	78.89219	9.470777	8.047993	3.778339	641.25
SWIF* ILWIF	31.54321	0.569337	5.643988	3.424813	0.503375	79.94764	9.750034	9.072432	3.784519	66433.5
(ELWIF)^2	17.82716	138.9318	98.61026	1488.089	326.2477	5896.974	321.1941	63.22503	2310.147	6.25
(ILWIF)^2	15.12346	0.002314	0.290624	0.007846	0.000774	1.055453	0.279257	1.024439	0.00618	65792.25
(SWIF)^2	65.79012	140.0681	109.6076	1494.931	327.2537	6055.814	340.415	80.34546	2317.71	67081
ELWIF	4.222222	11.78693	9.93027	38.57576	18.06233	76.79176	17.92189	7.951417	48.06399	2.5
ILWIF	3.888889	0.048106	0.539096	0.088578	0.027826	1.027352	0.528447	1.012146	0.078611	256.5
SWIF	8.111111	11.83504	10.46937	38.66434	18.09015	77.81911	18.45034	8.963563	48.1426	259
Total links	73	62,489	143,535	16,587	354,965	213,380	38,266	2,214	26,334	1,554

The correlation between the three types of web impact factor, calculated from Table 3, gives the following values:

Impact factors involved	Correlation coefficient
<i>swif</i> and <i>ilwif</i>	0.331602
<i>ilwif</i> and <i>elwif</i>	-0.25739
<i>elwif</i> and <i>swif</i>	0.017303

The above data tells that the correlation between *swif* and *ilwif* is the highest. Since, *swif* is calculated from total links (sum of inlinks and outlinks) to a website and *ilwif* is calculated from the number of inlinks to a website, the number of inlinks to a website sufficiently influences the value of *swif*, in this case. On the other hand, *ilwif* and *elwif* are negatively correlated. The correlation between *elwif* and *swif* is positive but not as high as that between *swif* and *ilwif*. Also, none of the three types of impact factors have very high correlation existing between them. Taking these factors into account, the websites have been ranked only according to the number of inlinks received. This has been displayed in Table 1.

*sedac.ciesin.columbia.edu*, *geonames.org*, *nationsonline.org* are the top 3 websites according to the rank list formed using the number of inlinks to the websites. The number of outlinks and inlinks to these websites is quite higher and displays that they are well connected. The website *linkedgeodata.org* has less inlinks and outlinks, and thus has low values for *swif*, *slwif* and *elwif*. The traffic on this site is due to unique search features (Table 2). This is probably because they publish machine readable data. Thus, a quantitative analysis alone would have led us to a wrong conclusion, but the quality of information and the technology being used by the websites to disseminate information must also be taken into account for a proper assessment. Some intangible and non-quantifiable factors like human perception, psychology, etc. might also play an important role behind the impact that a website has on the society.

## 5. Conclusion and Future Work

Digital reference tools can help a library expand and extend their services. The tourism domain is growing

rapidly and hence use of digital reference tools will help cope with information needs related to this domain. The assessment strategy and results of this study will work as a guide in selecting the most suitable digital reference tool(s)/ website(s) according to purpose. Future work includes charting out assessment methodologies for evaluating personal blogs, social media, etc. through which tourism information is also disseminated.

The way tourism websites disseminate information is typical of the Web of Documents. The main lacuna of the Web of documents is that background knowledge is not captured here. Knowledge is multidimensional. Thus, it is better to publish data in a machine readable way, make their semantics explicit (Bizer, Heath & Lee, 2009). It will help users can the data according to their purpose. This is a way towards the Web of Data (Bizer, Cyganiak and Heath, 2014). It will enhance publishing and connecting data with related data (Bizer & Heath 2011). It will also enable different applications to interoperate and share their data. The basic assumption behind this is that there is increase in value and usefulness of datasets when interlinked with other data (Bizer, Heath & Lee, 2009). There is need for designing suitable architecture for publication of tourism information in a machine-readable format. The assessment methodology discussed here, can also be extended to evaluate digital reference tools/websites related to other domains.

## 6. References

- “World Tourism Organization UNWTO”. Accessed 1 August, 2018. <http://www2.unwto.org/>
- “SocSciBot: Link crawler for the social sciences”. Accessed 1 August, 2018. <http://socscibot.wlv.ac.uk/>
- “Website Traffic, Statistics and Analytics – Alexa”. Accessed 1 August, 2018. <https://www.alexa.com/siteinfo>
- “Country Studies”, Accessed 1 August, 2018. <http://countrystudies.us/>
- “UNEP Environmental Data Explorer-The Environmental Database”. Accessed 1 August, 2018. <http://geodata.grid.unep.ch/>
- “linkedgeodata.org: datasets”. Accessed 1 August, 2018. <http://linkedgeodata.org/Datasets>
- “Socioeconomic Data and Applications Center |SEDAC”. Accessed 1 August, 2018. <http://sedac.ciesin.columbia.edu/>
- “World by Map: Statistics, Maps and Charts”. Accessed 1 August, 2018 <http://world.bymap.org/>
- “Atlademia Online”. Accessed 1 August, 2018. <http://www.atlademia.com/>
- Almind, T.C. and Ingwersen P. (1997). Informetric analyses on the world wide web: Methodological approaches to ‘webometrics’. *Journal of Documentation*, 53(4):404–26. Data accessed 1 August, 2018. [https://www.ischool.utexas.edu/~i385df04/readings/Almind\\_Ingwersen-1997-Webometrics.pdf](https://www.ischool.utexas.edu/~i385df04/readings/Almind_Ingwersen-1997-Webometrics.pdf), <https://doi.org/10.1108/EUM000000007205>.
- Alastair, S. (1999). A tale of two web spaces: Comparing sites using web impact factors. *Journal of Documentation*, 55(5):577–92.
- Bjornborn, L. and Ingwersen, P. (2004). Toward a basic framework for webometrics. *J. of the Amer. Soc. for Inf. Sci. and Tech.*, 55(14):1216–27.
- Bizer, C., Heath, T. and Lee, T.B. (2009) Linked data - The story so far. *International Journal on Semantic Web and Information Systems*, 5(3):1–22, DOI: 10.4018/jswis.2009081901. <https://doi.org/10.4018/jswis.2009081901>.
- Bizer, C. and Heath, T. (2011). *Linked Data: Evolving the Web into a Global Data Space*, San Rafael: Morgan and Claypool.
- Bizer, C., Cyganiak, R. and Heath, T. How to publish linked data on the web. Data accessed June 2, 2014, <http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/>.
- “Columbia Gazetteers”. Accessed 1 August, 2018. <http://www.columbiagazetteer.org/>
- “Free Spatial Data. DIVA-GIS”. Accessed 1 August, 2018. <http://www.diva-gis.org/Data>
- “Global gazetteers”. Accessed 1 August, 2018. <http://www.fallingrain.com/world/index.html>
- “Gazetteers of British place names”. Accessed 1 August, 2018. <https://www.gazetteer.org.uk/>
- “The Geography Site”. Accessed 1 August, 2018. <http://www.geography-site.co.uk/>
- “GeoNames”. Accessed 1 August, 2018. <https://www.geonames.org/>
- “India Tour”. Accessed 1 August, 2018. <https://www.india-tour.com>
- “mapcruzin”. Accessed 1 August, 2018. <https://www.mapcruzin.com>
- “One World-Nations Online”. Accessed 1 August, 2018. <http://www.nationsonline.org>
- “Natural Earth”. Accessed 1 August, 2018. <http://www.naturalearthdata.com/>
- “NSF Open Topography”. Accessed 1 August, 2018. <http://www.opentopography.org/>
- Ingwersen P. (1998). The calculation of web impact factors Google Search. *Journal of Documentation*, 54(2):236–43. <https://doi.org/10.1108/EUM000000007167>.
- Maltese, V. and Farazi, F. (2011). *Towards the Integration of Knowledge Organization Systems with the Linked Data Cloud*. UDC Seminar; p.75.
- Ranganathan, S. R. (1967). *Prolegomena to Library Classification*. Asia Publishing House.
- Thelwall, Mike. (2009). *Introduction to Webometrics: Quantitative web research for the social sciences*. Morgan and Claypool, San Rafael. PMCid: PMC3188415.