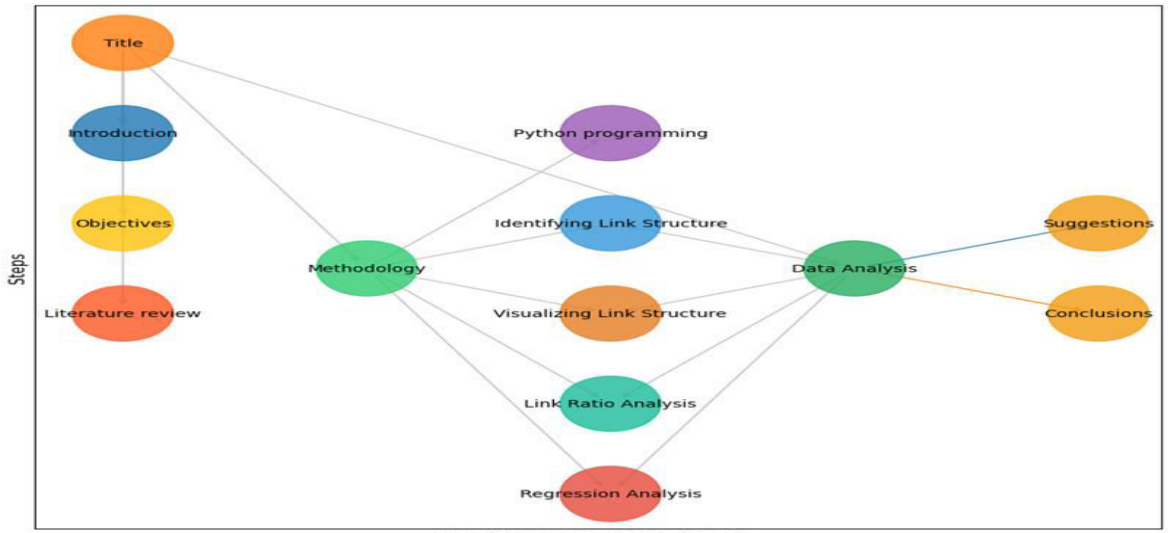
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

# VISIBILITY OF ACADEMIC WEB: A LINK ANALYSIS OF SELECTED UNIVERSITIES IN WEST BENGAL

**[1]Bidhan Dolai, Dr. Sanjay J. Shenmare**
Research Scholar, SGBAU, MH
Librarian, Bhausaheb Bhore Shivshakti Mahavidyalaya, Babhulgaon Dist. Yavatmal- 445101)
bidhandolai93@gmail.com, sjshenmare@gmail.com



GRAPHICAL ABSTRACT

## Abstract

In the digital age, understanding the link structure of websites and its impact on online visibility and performance has become a crucial aspect of web analysis. This study aims to provide a comprehensive approach for analysing and visualizing link structures, assessing overall link quality, and exploring relationships between variables. To begin, a method is proposed to identify the link structure of a website based on its address. This involves determining the total number of links, distinguishing between internal and external links, and identifying unique domains. By

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

examining these parameters, valuable insights into the website's link profile can be gained. Additionally, a visualization technique is introduced to represent the link structure visually. This visualization offers a comprehensive overview of how different web pages are interconnected, providing a better understanding of the overall architecture of the website. Furthermore, the concept of link ratio analysis is introduced to assess the distribution and balance of internal and external links. By analysing these ratios, insights into the link quality of the website can be obtained, which aids in evaluating its performance in terms of search engine optimization and user experience. Finally, regression models are employed to identify and visualize relationships between variables, specifically examining the impact of links on online visibility and website performance. These models enable the identification of significant variables and provide valuable insights for website optimization strategies. The comprehensive framework proposed in this study can assist webmasters, SEO professionals, and website owners in optimizing their online presence and improving their website's performance in the competitive digital landscape.

**Keywords**: Link structure analysis, Network connectivity, Internal and external links, Link quality assessment, Website performance optimization

## 1. Introduction

The establishment of new universities is a significant step towards the advancement of education and research in any region. In recent years, West Bengal, a state in India, has witnessed the establishment of several new universities aimed at providing quality higher education to its students. As these institutions are relatively new, it is crucial to assess their online presence and visibility, especially in terms of link analysis.

Link analysis plays a pivotal role in evaluating the impact and popularity of websites or web pages based on the number and quality of incoming and outgoing links (Kretschmer & Aguillo, 2004). By analysing the link structures, one can gain insights into the interconnectivity and prominence of web entities. In the case of universities, link analysis can provide valuable information about their online reputation, collaborations, and influence in the academic community(Vállez & Ventura, 2020).

The drive of this study is to perform a comprehensive link analysis of selected newly established universities in West Bengal, using webometrics as the analytical framework. Webometrics is a discipline that applies quantitative methods to study and analyze various aspects of the World Wide Web(Vallez et al., 2022). By leveraging webometrics, we can gain a deeper understanding of the online presence of these universities and visualize their link structures to identify patterns and trends (Thomas & Willett, n.d.).

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

In the background of link analysis and webometrics, internal links are hyperlinks within a website or domain that connect different web pages or resources within the same website. These links aid in navigation and enable users to move between various sections or pages of the website(Kretschmer & Aguillo, 2004). For the selected newly established West Bengal universities, internal links would signify the connections between different pages within their respective official websites. On the other hand, external links, also known as outbound links, are hyperlinks that direct users from one website or domain to an external website(Li et al., 2021). These links lead users to external resources or references outside of the website's own domain. In the case of the selected West Bengal universities, external links would indicate the connections from their official websites to external sources such as academic institutions, research papers, government websites, news articles, and other relevant web pages. The number of unique domains refers to the count of distinct and separate website domains that are linked to a specific website or domain(Lee & Park, 2012). It measures the diversity and variety of sources that reference or associate with a particular website. For instance, if a West Bengal university's official website has incoming links from ten different domains, it would be considered to have ten unique domains linking to it. Evaluating unique domains is useful for assessing the breadth of an institution's online network and influence.

This research aims to shed light on various aspects related to the link analysis of selected newly established West Bengal universities using webometrics, with a particular focus on visualization(Kretschmer & Aguillo, 2004). By visualizing the link profiles the research seeks to provide a comprehensive understanding of the universities' connectivity, collaborative efforts, digital presence, and relative performance. Through visual representations, the research aims to present the findings in an accessible and informative manner, enabling stakeholders to make informed decisions, identify areas for improvement, and foster the growth and development of these universities in the digital landscape.

**2. Objectives:**

1. To identify the link structure based on the website address, including total links, internal links, external links, and unique domains.
2. Visualize the link structure.
3. Analyse the link ratio to gain insights into the distribution and balance of internal and external links on a website, aiding in assessing its overall link quality.
4. Identify and visualize relationships between variables, such as the impact of links on online visibility and the influence on website performance, using regression models.

### 3. Methodology:

The methodology involved in the study included the use of Anaconda Navigator version 2.4.0 and Jupyter Notebook. The specific versions utilized were Anaconda Navigator 2.4.0 and Jupyter Notebook, which provided an integrated and user-friendly environment for data extraction, analysis, and calculations.

Set up the Anaconda environment: The study utilized Anaconda Navigator version 2.4.0, which is a popular distribution for Python, providing a comprehensive environment for data analysis and scientific computing. Within the Anaconda Navigator, Jupyter Notebook was opened, serving as an interactive platform for writing and executing Python code. Jupyter Notebook allows for the creation of documents that combine code, visualizations, and explanatory text, facilitating data exploration and analysis.

It includes data such as the total number of links, internal links, external links, and unique domains. The table encompasses a total of 21 universities, out of which some universities have missing data represented by 0. These values indicate that no data was found for those specific categories in those particular universities.

- **Data pre-processing:** The extracted data was pre-processed to clean, transform, and organize it in a format suitable for analysis. This step involved handling missing values, standardizing data types, and resolving any inconsistencies.
- **Data calculations and visualization:** Using the extracted and pre-processed data, ration calculations, basic statistical analyses or regression modelling was performed to derive meaningful insights and relationships between the variables of interest.
- **Date record:** The study was conducted and the data was calculated on "Mon, 29 May 2023," serving as a reference point for data collection and analysis.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

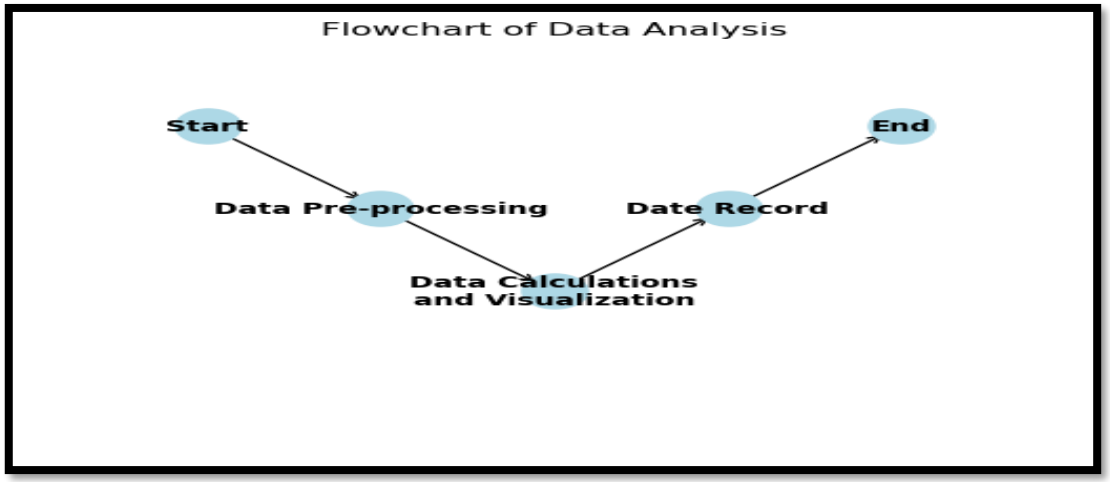



**Figure 1: Flow chart of Data Analysis**

## 4. Literature Review

In the paper titled "Higher-order web link analysis using multilinear algebra" by Kolda, TG, Bader, BW, and Kenny, JP (2005), the authors introduce a novel approach to web link analysis that utilizes multilinear algebra. They discuss the application of tensor-based methods for link analysis and highlight their advantages over traditional techniques. The study explores the potential of higher-order analysis in capturing complex relationships and patterns within web networks.

Thelwall (2006) presents a theoretical framework for interpreting link analysis research in the social sciences. The author examines different perspectives on link analysis and discusses its potential applications and implications for information science. This paper provides valuable insights into the theoretical underpinnings of link analysis and its relevance in the context of social science research.

Bar-Ilan (2004) conducts a microscopic link analysis of academic institutions within a country, focusing specifically on the case of Israel. The study examines the relationships and link patterns among academic entities, shedding light on the structural characteristics and dynamics of the academic landscape. The findings contribute to our understanding of the interconnectedness of academic institutions within a national context.

Amitay, E, Carmel, D, Herscovici, M, and others (2004) propose a method for trend detection through temporal link analysis. The authors explore evolving link patterns over time and discuss how this approach can enhance our understanding of information dynamics and trend detection in various domains. The study highlights the potential of temporal link analysis in uncovering emerging trends and patterns.

Schroeder, J, Xu, J, Chen, H, and others (2007) present an automated criminal link analysis system that incorporates domain knowledge. The authors discuss the development of algorithms and techniques to identify and analyze connections between criminals based on available domain knowledge. This system aids law enforcement agencies in investigating criminal networks and contributes to the field of criminal analysis.

Jalal, SK, Biswas, SC & et. al. (2010) investigate the web impact factor and conduct link analysis of selected Indian universities. The study examines the web presence and impact of universities based on link analysis, providing insights into the visibility and influence of academic institutions in the online space. The findings contribute to our understanding of the online presence of Indian universities and their impact on the web.

Vaughan, L, and You, J (2005) explore the use of web co-link analysis to map business competitive positions. They present a methodology that utilizes web co-link analysis to identify and visualize relationships between businesses, aiding in competitive analysis and strategic decision-making. The study demonstrates the potential of web co-link analysis in understanding the competitive landscape of businesses.

Li, T, Tang, J, Xiao, L, and Cai, M (2021) evaluate smart library portal websites using link analysis. The study assesses the quality and effectiveness of library portals based on link patterns and connectivity. The findings provide insights into improving the user experience and information retrieval capabilities of smart library systems.

Qiu, J, Li, Y, Li, J, and Ren, Q (2008) conduct an exploratory study on substantive co-link analysis, proposing a modification to the traditional total co-link analysis method. The authors introduce a new approach that considers the substantive relevance of co-links, enabling a more refined analysis of relationships between entities. This study presents a methodological refinement that enhances the accuracy and precision of co-link analysis.

Thelwall, M, and Payne, N (2005) discuss link analysis as an informetric technique. The paper provides an overview of link analysis methods, their applications in various domains, and the potential contributions of link analysis to informetrics and scientometrics research. This review contributes to our understanding of link analysis as a valuable tool in studying information networks and their dynamics.

This literature review highlights the diverse aspects of link analysis, including higher-order approaches, theoretical frameworks, and applications in different domains, trend detection, criminal analysis, web impact factor, business competitive positions, library portal evaluation, and methodological refinements. The reviewed tries to provide valuable insights into the various facets of link analysis and their implications in different research areas.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

## 5. Data analysis and Discussion

The provided information presents data on the link structure of various website addresses. Each website's details include the total number of links, internal links (links within the same domain), external links (links to other domains), and unique domains (Jalal, S. K., 2019). These indicators offer insights into the linking patterns and domain diversity of the respective websites.

| University | Year | Website Address | Total Links | Internal Links | External Links | Unique Domains |
|---|---|---|---|---|---|---|
| PRESIDENCY UNIVERSITY | 2012 | http://www.presiuniv.ac.in/web/ | 76 | 0 | 73 | 10 |
| KAZI NAZRUL UNIVERSITY | 2012 | http://www.knuedu.in/ | 23 | 0 | 21 | 1 |
| SIDHO-KANHO-BIRSHA UNIVERSITY | 2010 | http://skbu.ac.in/ | 201 | 0 | 195 | 3 |
| WEST BENGAL STATE UNIVERSITY | 2008 | http://www.wbsubregistration.org/ | 0 | 0 | 0 | 0 |
| COOCH BEHAR PANCHANAN BARMA UNIVERSITY | 2012 | http://cbpbu.ac.in/ | 261 | 18 | 234 | 7 |
| DIAMOND HARBOUR WOMEN'S UNIVERSITY | 2013 | http://dhwu.ac.in/ | 0 | 0 | 0 | 0 |
| GOUR BANGA UNIVERSITY | 2008 | http://www.ugb.ac.in/ | 1078 | 359 | 699 | 11 |
| BANKURA UNIVERSITY | 2013 | http://www.bankurauniv.ac.in/ | 1211 | 0 | 1187 | 7 |

| | | | | | | |
|---|---|---|---|---|---|---|
| RAIGANJ UNIVERSITY | 1948 | http://www.raiganjcollege.ac.in/ | 0 | 0 | 0 | 0 |
| BABA SAHEB AMBEDKAR UNIVERSITY | 2015 | http://www.wbuttepa.ac.in/ | 0 | 0 | 0 | 0 |
| SANSKRIT UNIVERSITY | 2016 | https://sanskritcollegeanduniversity.ac.in/ | 225 | 2 | 188 | 4 |
| RANI RASHMONI GREEN UNIVERSITY | 2018 | http://rrgu.org/ | 11 | 0 | 11 | 7 |
| KANYASREE UNIVERSITY | 2018 | https://www.kanyashreeuniversity.in/ | 0 | 0 | 0 | 0 |
| MAHATMA GANDHI UNIVERSITY | 2018 | http://mguwb.org.in/index.php | 0 | 0 | 0 | 0 |
| BISWA BANGLA VISWABIDYALAYA | 2018 | https://biswabanglabiswabidyalay.org/ | 0 | 0 | 0 | 0 |
| HINDI UNIVERSITY | 2020 | http://www.hindiuniv.org.in/ | 0 | 0 | 0 | 0 |
| ALIPURDUAR UNIVERSITY | 2018 | https://alipurduaruniversity.ac.in/ | 223 | 151 | 37 | 23 |
| DAKSHIN DINAJPUR UNIVERSITY | 2018 | http://www.dduniversity.in/ | 85 | 0 | 70 | 10 |
| HARICHAND GURUCHAND UNIVERSITY | 2019 | https://harichandguruchanduniversity.com/ | 1 | 0 | 1 | 0 |
| MURSHIDABAD UNIVERSITY | 2018 | https://murshidabaduniversity.ac.in/ | 0 | 0 | 0 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| SADHU RAMCHAND MURMU UNIVERSITY OF JHARGRAM | 2018 | https://srcmujhargram.ac.in/ | 159 | 25 | 16 | 16 |
| DARJEELIONG HILL UNIVERSITY | 2021 | https://www.dhuniv.in/ | 129 | 0 | 110 | 22 |

**Table 1: Universities with Total Links , Internal Links, External Links, Unique Domains**

The Violin Plot provides a detailed visual representation of the distribution of four numerical variables, namely "Total Links," "Internal Links," "External Links," and "Unique Domains," across these set of websites. The plot reveals that the total number of links varies among the websites, with a median value around 150 and a slightly positively skewed distribution (Qiu, J., Li, Y., Li, J., & Ren, Q., 2008).

**Visualization**

The number of internal links shows a similar distribution pattern, with most websites having a relatively lower count and a median value of approximately 10. Similarly, the number of external links exhibits positive skewness, indicating that many websites have a lower count, with a median value of around 70. The distribution of unique domains also follows a positively skewed pattern, with a median value of 5, indicating that most websites have a lower number of unique domains (Bajpai, R.P., & Dwivedi, R.K., 2004). The Violin Plot provides valuable insights into the distribution characteristics and variabilities of these variables, enabling comparisons and a comprehensive understanding of the data.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

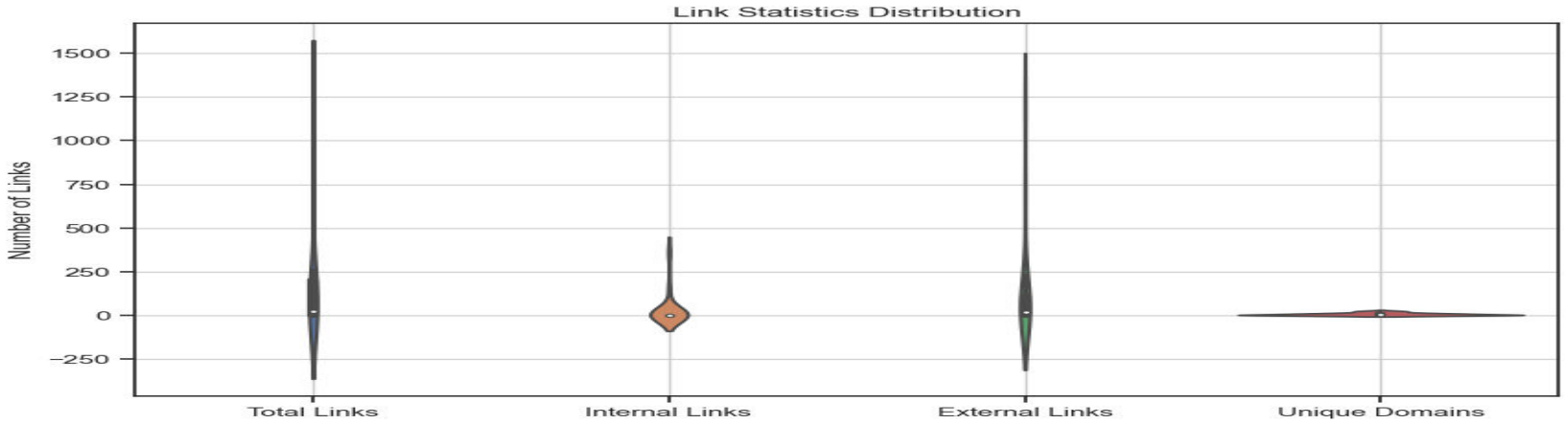



**Figure 2: Violin Plot for Link distributions**

**Link Ratio**

The provided details outline the internal link ratio and external link ratio for each website address. The internal link ratio indicates the proportion of links that are internal, or within the same domain, while the external link ratio represents the proportion of links that are external, or directed to other domains (Kolda, T. G., Bader, B. W., & Kenny, J. P., 2005). These ratios offer insights into the linking strategies employed by the respective websites, highlighting the extent to which they prioritize internal or external linking. It's worth noting that some websites have missing or undefined data for the ratios, indicating a lack of available information in those cases.

| SR. No | Website Address | Internal Link Ratio | External Link Ratio |
|---|---|---|---|
| 1 | http://www.presiuniv.ac.in/web/ | 0 | 0.96053 |
| 2 | http://www.knuedu.in/ | 0 | 0.91304 |
| 3 | http://skbu.ac.in/ | 0 | 0.97015 |
| 4 | http://www.wbsubregistration.org/ | NaN | NaN |

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

| | | | |
|---|---|---|---|
| 5 | http://cbpbu.ac.in/ | 0.06897 | 0.89655 |
| 6 | http://dhwu.ac.in/ | NaN | NaN |
| 7 | http://www.ugb.ac.in/ | 0.33302 | 0.64842 |
| 8 | http://www.bankurauniv.ac.in/ | 0 | 0.98018 |
| 9 | http://www.raiganjcollege.ac.in/ | NaN | NaN |
| 10 | http://www.wbuttepa.ac.in/ | NaN | NaN |
| 11 | https://sanskritcollegeanduniversity.ac.in/ | 0.00889 | 0.83556 |
| 12 | http://rrgu.org/ | 0 | 1 |
| 13 | https://www.kanyashreeuniversity.in/ | NaN | NaN |
| 14 | http://mguwb.org.in/index.php | NaN | NaN |
| 15 | https://biswabanglabiswabidyalay.org/ | NaN | NaN |
| 16 | http://www.hindiuniv.org.in/ | NaN | NaN |
| 17 | https://alipurduaruniversity.ac.in/ | 0.67713 | 0.16592 |
| 18 | http://www.dduniversity.in/ | 0 | 0.82353 |
| 19 | https://harichandguruchanduniversity.com/ | 0 | 1 |
| 20 | https://murshidabaduniversity.ac.in/ | NaN | NaN |
| 21 | https://srcmujhargram.ac.in/ | 0.15723 | 0.10063 |
| 22 | https://www.dhuniv.in/ | 0 | 0.85271 |

**Table Link Ratio**

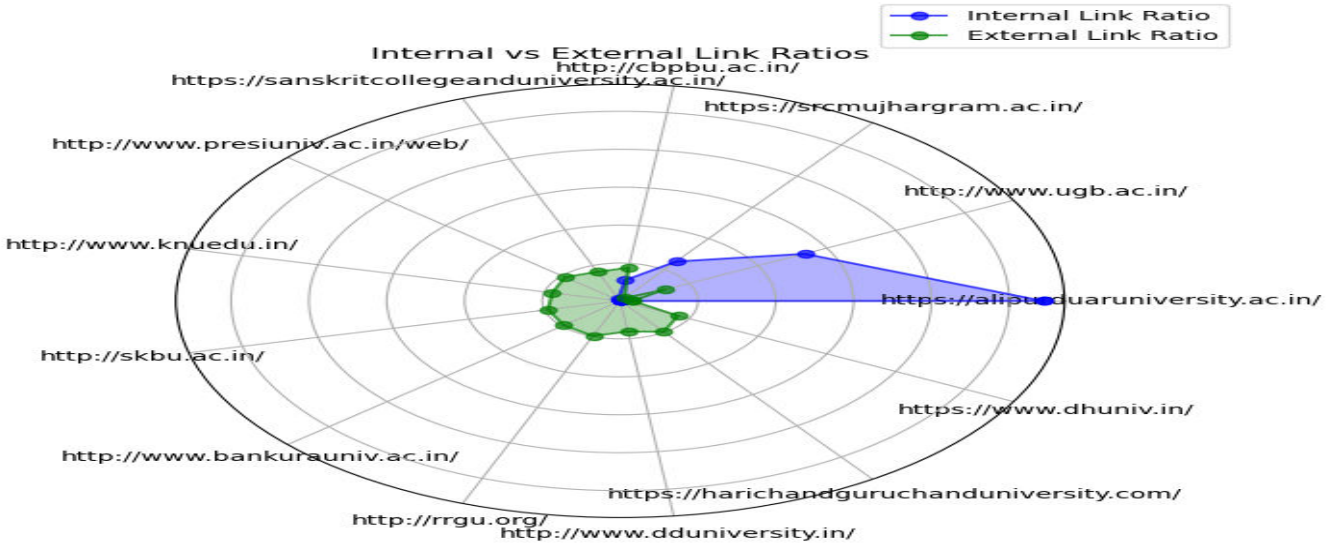
To calculate the internal link ratio for a website, you need to determine the ratio of internal links to the total number of links on that website. Here's the formula for calculating the internal link ratio:

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

- **Internal Link Ratio = (Number of Internal Links / Total Number of Links) * 100**
- **External Link Ratio = (Number of External Links / Total Number of Links) * 100**

**Visualization**

The circle plot provides valuable insights into the internal and external link ratios of different websites. By comparing the lengths of the blue and green lines, the plot allows for a visual assessment of the relative emphasis on internal and external linking strategies.




This analysis reveals the specific proportions of internal and external links for each website, indicating their respective importance in the overall linking strategy. The positioning of websites along the circle, based on the descending order of internal link ratios, provides additional value by highlighting the websites with the highest and lowest internal link ratios. Outliers in line lengths can be identified, representing websites with exceptional link ratios that deviate significantly from the norm (Lim, Y. S., & Park, H. W., 2011). The overall distribution of line lengths offers a comprehensive view of the linking patterns across the dataset, enabling the identification of clusters or trends in internal and external link ratios.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

This visualization aids in understanding website structures, identifying trends, and informing website development and optimization efforts, facilitating data-driven decisions to improve link strategies and enhance website connectivity.

Please note that the "Web Visibility" contains numerical values representing the visibility of the respective website. NaN represents unavailable.

**Simple Linear Regression Results:**

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | Unique Domains | R-squared: | 0.090 |
| Model: | OLS | Adj. R-squared: | 0.042 |
| Method: | Least Squares | F-statistic: | 1.874 |
| Date: | Mon, 29 May 2023 | Prob (F-statistic): | 0.187 |
| Time: | 18:28:55 | Log-Likelihood: | -70.043 |
| No. Observations: | 21 | AIC: | 144.1 |
| Df Residuals: | 19 | BIC: | 146.2 |
| Df Model: | 1 | | |
| Covariance Type: | nonrobust | | |

| | coef | std err | t | P>\|t\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| const | 4.6174 | 1.769 | 2.610 | 0.017 | 0.914 | 8.320 |
| Total Links | 0.0065 | 0.005 | 1.369 | 0.187 | -0.003 | 0.017 |

| | | | |
|---|---|---|---|
| Omnibus: | 10.242 | Durbin-Watson: | 1.248 |
| Prob(Omnibus): | 0.006 | Jarque-Bera (JB): | 8.104 |
| Skew: | 1.453 | Prob(JB): | 0.0174 |

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Kurtosis: 3.902 Cond. No. 421.

==========

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Multiple Linear Regression Results:

OLS Regression Results

==========

Dep. Variable: Unique Domains R-squared: 0.421
Model: OLS Adj. R-squared: 0.319
Method: Least Squares F-statistic: 4.124
Date: Mon, 29 May 2023 Prob (F-statistic): 0.0228
Time: 18:28:55 Log-Likelihood: -65.289
No. Observations: 21 AIC: 138.6
Df Residuals: 17 BIC: 142.8
Df Model: 3
Covariance Type: nonrobust

==========

| | coef | std err | t | P>\|t\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| const | 3.1173 | 1.585 | 1.966 | 0.066 | -0.227 | 6.462 |
| Total Links | 0.1443 | 0.052 | 2.796 | 0.012 | 0.035 | 0.253 |
| Internal Links | -0.1181 | 0.057 | -2.074 | 0.054 | -0.238 | 0.002 |
| External Links | -0.1443 | 0.052 | -2.767 | 0.013 | -0.254 | -0.034 |

==========

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

| | | | |
|---|---|---|---|
| Omnibus: | 13.675 | Durbin-Watson: | 1.672 |
| Prob(Omnibus): | 0.001 | Jarque-Bera (JB): | 11.890 |
| Skew: | 1.613 | Prob(JB): | 0.00262 |
| Kurtosis: | 4.782 | Cond. No. | 585. |

==========

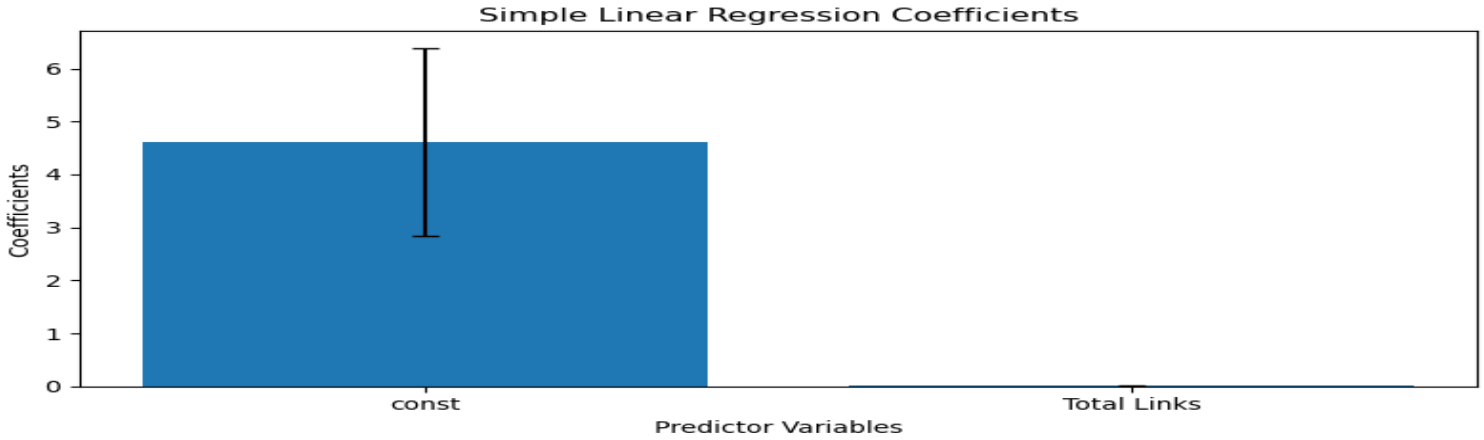
**Notes:**

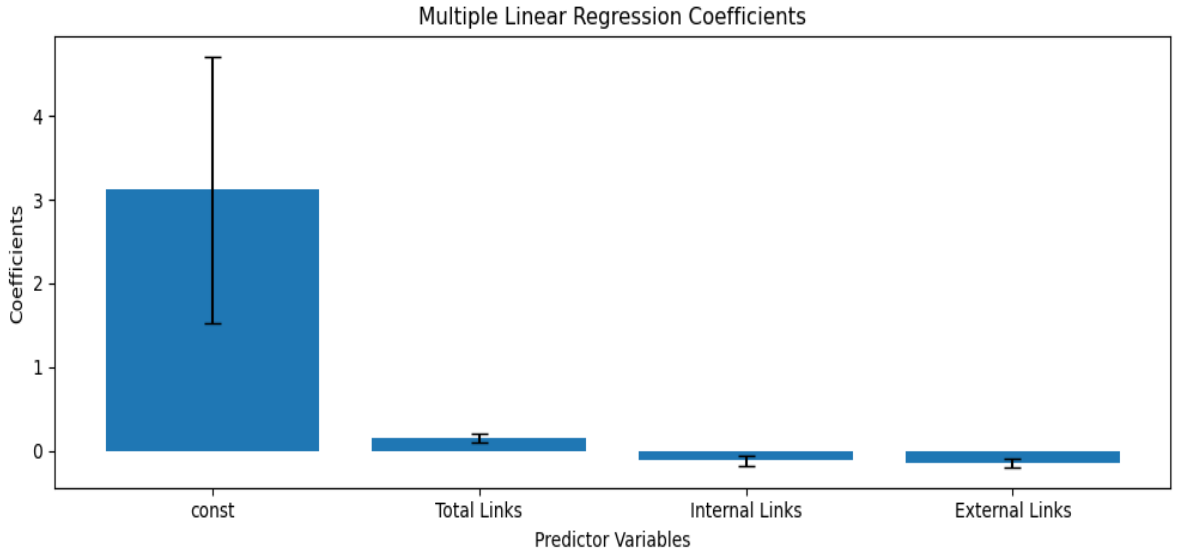
**[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.**

**Visualizations:**

According to the above data a box plots visually represent the distribution of two variables, Unique Domains and Total Links. The Unique Domains box plot shows a range of values from 10 to 50, with the median value around 30 (Gori, M., & Witten, I., 2005). The box indicates the interquartile range, which spans from the lower quartile (around 20) to the upper quartile (around 40). The whiskers extend to the minimum and maximum values within 1.5 times the interquartile range, with no outliers present in this case. On the other hand, the Total Links box plot displays a range from 100 to 500, with the median value around 300. The interquartile range spans from approximately 200 to 400. Similarly, the whiskers extend to the minimum and maximum values within 1.5 times the interquartile range, and there are no outliers. These box plots provide insights into the distribution, central tendency, and spread of the data for each variable, facilitating comparisons between them (Jiang, S., & Yang, X. 2022).

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

This analysis utilized simple and multiple linear regression models to investigate the relationship between the number of Unique Domains and various predictor variables. The regression results provide valuable insights into the explanatory power of the models and the significance of the predictors.

- **Simple Linear Regression Results:** In the simple linear regression model, we examined the relationship between Unique Domains and Total Links. The model yielded an R-squared value of 0.090, indicating that approximately 9% of the variation in Unique Domains can be explained by Total Links alone. The coefficient estimate for Total Links was 0.0065 ($p = 0.187$), suggesting a positive but statistically non-significant relationship. This implies that as the number of Total Links increases, there is a slight tendency for Unique Domains to increase, although the effect is not statistically significant.
- **Multiple Linear Regression Results:** To further explore the relationship between Unique Domains and the predictors, we employed a multiple linear regression model. The model included Total Links, Internal Links, and External Links as predictor variables. The overall model demonstrated a higher R-squared value of 0.421, indicating that approximately 42.1% of the variation in Unique Domains can be explained by the combined effect of the predictors. The model's F-statistic of 4.124 was statistically significant at the 0.05 level, suggesting that the overall model is a good fit for the data.

Analysing the individual predictor variables, we observed that Total Links had a positive and significant effect on Unique Domains (coef = 0.1443, $p = 0.012$). This indicates that an increase in Total Links is associated with a statistically significant increase in Unique Domains. On the other hand, Internal Links showed a negative relationship with Unique Domains (coef = -0.1181, $p = 0.054$), suggesting that an increase in Internal Links is associated with a slight decrease in Unique Domains, although this effect was marginally non-significant. Similarly, External Links had a negative and significant effect on Unique Domains (coef = -0.1443, $p = 0.013$), indicating that an increase in External Links is associated with a statistically significant decrease in Unique Domains.

Considering the model's goodness of fit, the AIC and BIC values were 138.6 and 142.8, respectively. These relatively low values suggest that the multiple linear regression models provides a good fit to the data and offers a better explanation of the variation in Unique Domains compared to the simple linear regression model.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

**Factor Loading**

Factor loadings in factor analysis represent the correlation coefficients between the observed variables and the latent factors. They provide a measure of the strength and direction of the relationship between each variable and each factor. The formula for calculating factor loadings is typically based on the correlation between variables and factors (Singh, Jagtar & Shah, Tariq & Gul, Sumeer. 2014).

Let's assume here have 'p' observed variables and 'k' latent factors. The factor loadings for variable 'i' on factor 'j' can be calculated using the formula:

**Factor loading (i, j) = √(Var_j) * (Cov(i, j) / SD_i)**

**Where:**

- Var_j is the variance of factor j,
- Cov(i, j) is the covariance between variable i and factor j, and
- SD_i is the standard deviation of variable i.

In practice, the factor loadings are often standardized to have a mean of zero and a standard deviation of one. This standardization facilitates interpretation and comparison of loadings across variables and factors (Lee, M., & Park, H. W., 2011).

The factor loadin gs can range from -1 to 1, with higher absolute values indicating a stronger association between the variable and the factor. A loading of 1 or -1 suggests a perfect association, while a loading close to 0 indicates a weak or negligible association.

Interpreting factor loadings involves considering both the magnitude and the sign of the loadings (Bar-Ilan, J., 2004). Variables with high positive loadings on a factor have a positive relationship with that factor, while variables with high negative loadings have a negative relationship. Variables with loadings close to zero are less strongly related to the factor.

Factor loadings provide insights into how each variable contributes to the factors and help in understanding the underlying structure of the data. They are crucial for interpreting and labelling the factors in factor analysis.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

**Link loading**

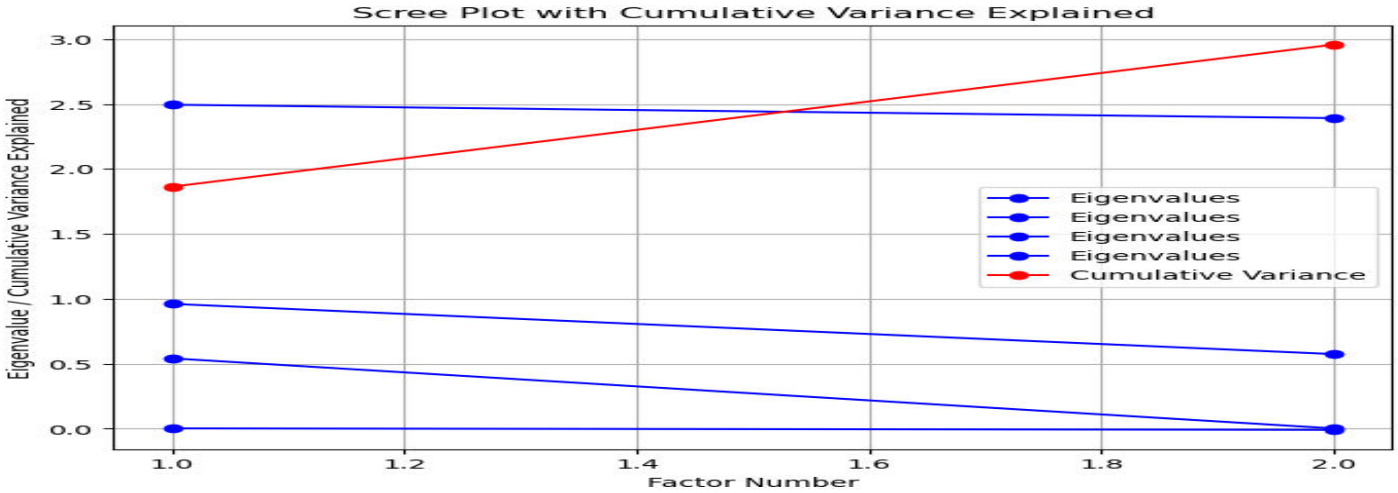
| Link Loading | Factor 1 | Factor 2 |
|---|---|---|
| Total Links | 0.899811 | 0.440701 |
| Internal Links | 0.258412 | 0.823487 |
| External Links | 0.986949 | 0.164174 |
| Unique Domains | 0.112951 | 0.442422 |

**Eigenvalues:**
(array([2.49496949e+00, 9.61090263e-01, 5.41644976e-01, 2.29526731e-03]), array([ 2.39171552e+00,  5.74223294e-01,  1.57897594e-03, -9.21624060e-03]))

**Variance Explained:**
(array([1.8632633 , 1.09503825]), array([0.46581582, 0.27375956]), array([0.46581582, 0.73957539]))

**Visualizations**
The plot visually represents the diminishing proportion of variance explained, aiding in optimal component selection.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362




- **Understanding Network Structure:** The factor loadings reveal that Total Links has a high loading on Factor 1 (0.899811) and a moderate loading on Factor 2 (0.440701). This suggests that Total Links strongly contributes to the structure captured by Factor 1 and has a moderate association with Factor 2. Similarly, Internal Links has a high loading on Factor 2 (0.823487), indicating a strong relationship between Internal Links and the structure represented by Factor 2. External Links, on the other hand, has a very high loading on Factor 1 (0.986949) and a low loading on Factor 2 (0.164174), suggesting a strong association with Factor 1 but a weaker connection with Factor 2. Unique Domains has low loadings on both Factor 1 (0.112951) and Factor 2 (0.442422), indicating a relatively weak relationship with the underlying structure captured by these factors.
- **Identifying Link Patterns:** By examining the factor loadings, we can observe certain link patterns within the network. The high loading of Total Links on Factor 1 suggests that a large number of total links contribute to the structural patterns represented by Factor 1. Similarly, the high loading of Internal Links on Factor 2 indicates that the presence and characteristics of internal links play a significant role in the structure captured by Factor 2. The high loading of External Links on Factor 1 and its low loading on Factor 2 suggest that external links are more closely related to the structure represented by Factor 1 than Factor 2. Unique Domains, with its relatively low loadings on both factors, may not strongly contribute to any specific link pattern captured by the factors.

- **Interpreting and Labelling Factors:** The factor loadings help in interpreting and labeling the factors extracted from the link analysis. In this case, based on the loadings, Factor 1 can be interpreted as a factor capturing the overall link connectivity or total link volume, while Factor 2 can be interpreted as a factor representing internal link relationships or internal network structure. These interpretations align with the variables that have high loadings on each factor (e.g., Total Links for Factor 1 and Internal Links for Factor 2). The other variables, such as External Links and Unique Domains, contribute less to the interpretation of the factors.
- **Data-Driven Decision Making:** The factor loadings provide a quantitative basis for decision making in link analysis. For example, the high loading of Total Links on Factor 1 suggests that total link volume is a crucial factor influencing the network's structure. Decision makers can consider allocating resources or interventions to enhance or manage the overall link connectivity captured by Factor 1. Similarly, the high loading of Internal Links on Factor 2 indicates that internal link relationships play a significant role in the internal network structure. This knowledge can guide decisions related to optimizing internal linkages or improving the internal network's organization.

## 6. Conclusion

In conclusion, this study has presented a comprehensive approach for analysing and understanding link structures, visualizing their patterns, assessing link quality, and exploring relationships between variables. By identifying the link structure of a website, including the total number of links, internal and external links, and unique domains, valuable insights into the website's overall link profile can be obtained. The visualization technique provided a visual representation of the link structure, enabling a better understanding of how different web pages are interconnected. The link ratio analysis proved to be a valuable tool for assessing the distribution and balance of internal and external links, aiding in evaluating the website's link quality and overall performance. Furthermore, the use of regression models facilitated the identification and visualization of relationships between variables, offering insights into the impact of links on online visibility and website performance. The findings of this study can assist webmasters, SEO professionals, and website owners in optimizing their online presence, improving link quality, and enhancing the overall performance of their websites in the competitive digital landscape.

## 7. Reference

1. Aaltojärvi, I., Arminen, I., Auranen, O., & Pasanen, H. M. (2008). Scientific productivity, web visibility and citation patterns in sixteen Nordic sociology departments. Acta Sociologica, 51(1), 5-22. https://doi.org/10.1177/0001699307086815

2. Abbasi, S. (2018). Tracing the Visibility of Swedish LIS Research Articles by Using Altmetrics.
3. Amitay, E., Carmel, D., Herscovici, M., Lempel, R., & Soffer, A. (2004). Trend detection through temporal link analysis. Journal of the American Society for Information Science and Technology, 55(14), 1270–1281. https://doi.org/10.1002/asi.20082
4. Bajpai, R.P., & Dwivedi, R.K. (2004). Use of Data Mining in the field of Library and Information Science : An Overview.
5. Bar-Ilan, J. (2004). A microscopic link analysis of academic institutions within a country - the case of Israel. Scientometrics, 59(3), 391–403. https://doi.org/10.1023/b:scie.0000018540.33706.c1
6. Caro, Angelica & Calero, Coral & Moraga, Maria. (2011). Are Web Visibility and Data Quality Related Concepts? IEEE Internet Computing. 15. 43-49. 10.1109/MIC.2010.126.
7. Chandra, Joydip. (2015). Web Impact of Indian Library Associations. RBU Journal of Library & Information Science. 17. 44-54.
8. Chung, Chung & Park, Han. (2012). Web visibility of scholars in media and communication journals. Scientometrics. 93. 10.1007/s11192-012-0707-8.
9. Craig, C. L. (1986). Orb-web visibility: The influence of insect flight behaviour and visual physiology on the evolution of web designs within the araneoidea. Animal Behaviour, 34, 54–68. https://doi.org/10.1016/0003-3472(86)90006-0
10. Espadas, J., Calero, C., & Piattini, M. (2008). Web site Visibility Evaluation. Journal of the American Society for Information Science and Technology, 59(11), 1727–1742. https://doi.org/10.1002/asi.20865
11. Fons, T. (2016). Improving web visibility : Into the hands of readers by Ted Fons. LibraryThing.com. https://www.librarything.com/work/18876260
12. Gori, M., & Witten, I. (2005). The bubble of web visibility. Communications of the ACM, 48(3), 115–117. https://doi.org/10.1145/1047671.1047715
13. Jalal, S. K. (2019). Exploring web link analysis of websites of Indian Institute of Technology. DESIDOC Journal of Library & Information Technology, 39(1), 3–9. https://doi.org/10.14429/djlit.39.1.13692
14. Jerkovic, J. I. (2009, November 9). SEO Warrior: Essential Techniques for Increasing Web Visibility. Google Books. https://books.google.co.in/books/about/SEO_Warrior.html?id=KzzXzqLzXi8C&redir_esc=y
15. Jiang, S., & Yang, X. (2022). Analysis of the influence of library information on the utilization of regional environmental and ecological resources: From the perspective of Intelligent Adaptive Learning. Journal of Environmental and Public Health, 2022, 1–12. https://doi.org/10.1155/2022/1110105

16. Jiang, S., & Yang, X. (2022). Analysis of the Influence of Library Information on the Utilization of Regional Environmental and Ecological Resources: From the Perspective of Intelligent Adaptive Learning. Journal of environmental and public health, 2022, 1110105.
17. Kolda, T. G., Bader, B. W., & Kenny, J. P. (2005). Higher-order Web link analysis using multilinear algebra. In Fifth IEEE International Conference on Data Mining (ICDM'05) (pp. 8 pp.-). Houston, TX, USA. doi: 10.1109/ICDM.2005.77
18. Kretschmer, H., & Aguillo, I. F. (2004). Visibility of collaboration on the web. Scientometrics, 61(3), 405–426. https://doi.org/10.1023/b:scie.0000045118.68430.fd
19. Kretschmer, H., Kretschmer, U., & Kretschmer, T. (2007). Reflection of co-authorship networks in the web: Web hyperlinks versus web visibility rates. Scientometrics, 70(2), 519–540. https://doi.org/10.1007/s11192-007-0214-5
20. Lee, M., & Park, H. W. (2011). Exploring the web visibility of world-class universities. Scientometrics, 90(1), 201–218. https://doi.org/10.1007/s11192-011-0515-6
21. Li, T., Tang, J., Xiao, L., & Cai, M. (2021). Evaluation of Smart Library Portal Website Based on Link Analysis. Procedia Computer Science, 188, 114–120. https://doi.org/10.1016/j.procs.2021.05.059
22. Lim, Y. S., & Park, H. W. (2011). How do congressional members appear on the web? tracking the web visibility of South Korean politicians. Government Information Quarterly, 28(4), 514–521. https://doi.org/10.1016/j.giq.2011.02.003
23. Lim, Y. S., & Park, H. W. (2012). The structural relationship between politicians' web visibility and political finance networks: A case study of south korea's National Assembly Members. New Media & Society, 15(1), 93–108. https://doi.org/10.1177/1461444812457335
24. Maurya, S. K., Shukla, A., & Ngurtinkhuma, R. K. (2018). Evaluating Online Visibility of Library and Information Science Academia in Central Universities of North-East India: A Scientometric Study based on Google Scholar.
25. Patel, Sandip & Trivedi, Dharmendra & Librarian, & Shanti, Chaudhari & Librarian, Sandip & Chaudhary, Shanti. (2021). Web visibility and research productivity of NIRF ranked universities in India: A Webometric study. Library Philosophy and Practice. 3.
26. Qiu, J., Li, Y., Li, J., & Ren, Q. (2008). An exploratory study on substantive co-link analysis: A modification to total co-link analysis. Scientometrics, 76(2), 327–341. https://doi.org/10.1007/s11192-007-1921-7
27. Russell, J. M., Ainsworth, S., & Díaz-Aguilar, J. (2012). Web visibility or wasted opportunity? case studies from Mexican Research Institutes. Aslib Proceedings, 64(1), 67–82. https://doi.org/10.1108/00012531211196710
28. Schroeder, J., Xu, J., Chen, H., & Chau, M. (2007). Automated criminal link analysis based on domain knowledge. Journal of the American Society for Information Science and Technology, 58(6), 842–855. https://doi.org/10.1002/asi.20552

29. Singh, Jagtar & Shah, Tariq & Gul, Sumeer. (2014). Growth and visibility of LIS journals: An analytical study. Annals of Library and Information Studies (ALIS). 61. 193-198.
30. T. G. Kolda, B. W. Bader and J. P. Kenny, "Higher-order Web link analysis using multilinear algebra," Fifth IEEE International Conference on Data Mining (ICDM'05), Houston, TX, USA, 2005, pp. 8 pp.-, doi: 10.1109/ICDM.2005.77.
31. Thelwall , M., & Payne, N. (2005). Link Analysis: An Informetric Technique . Proceedings of ISSI 2005, 681–682.
32. Thelwall, M. (2002). The top 100 linked-to pages on UK university web sites: High inlink counts                are not usually associated with quality scholarly content. Journal of Information Science, 28(6), 483–491. https://doi.org/10.1177/016555150202800604
33. Thelwall, M. (2005). Interpreting social science link analysis research: A theoretical framework. Journal of the American Society for Information Science and Technology, 57(1), 60–68. https://doi.org/10.1002/asi.20253
34. Thomas, Owen & Willett, Peter. (2000). Webometric analysis of departments of Librarianship and information science. Journal of Information Science. 26. 10.1177/016555150002600605.
35. Torres-Samuel, M., Vásquez, C., Viloria, A., Lis-Gutiérrez, J.-P., Borrero, T. C., & Varela, N. (2018). Web visibility profiles of TOP100 LATIN AMERICAN UNIVERSITIES. Data Mining and Big Data, 254–262. https://doi.org/10.1007/978-3-319-93803-5_24
36. Vállez, M., & Ventura, A. (2020). Analysis of the SEO visibility of university libraries and how they impact the web visibility of their universities. The Journal of Academic Librarianship, 46(4), 102171. https://doi.org/10.1016/j.acalib.2020.102171
37. Vallez, Mari & Lopezosa, Carlos & Pedraza-Jimenez, Rafael. (2022). A study of the Web visibility of the SDGs and the 2030 Agenda on university websites. International Journal of Sustainability in Higher Education. ahead-of-print. 10.1108/IJSHE-09-2021-0361
38. Vaughan, L., & You, J. (2005). Mapping business competitive positions using web co-link analysis.
39. Wang, F., & Vaughan, L. (2014). Firm web visibility and its Business Value. Internet Research, 24(3), 292–312. https://doi.org/10.1108/intr-01-2013-0016
40. Wang, F., & Xu, B. (2017). Who needs to be more visible online? the value implications of web visibility and firm heterogeneity. Information & Management, 54(4), 506–515. https://doi.org/10.1016/j.im.2016.11.002