Articles of Latin American companies on Wikipedia: Causal relationship between access, participation, and quality variables

Artículos de empresas latinoamericanas en Wikipedia: Relación causal entre las variables acceso, participación y calidad

Sergio Llano Aristizábal[1] | ORCID ID
sllano@uninorte.edu.co
Universidad del Norte, Colombia
Héctor López Ospinal[2] | ORCID ID
halopez@miuandes.cl
Universidad de los Andes, Chile

Recepción: 14/06/2023 Revisión: 04/12/2023 Aceptación: 20/12/2023 Publicación: 29/12/2023
http://dx.doi.org/10.5783/revrrpp.v13i26.817

Abstract
Wikipedia is a paradigmatic example of online collaboration. At an organizational level, the presence of content related to companies and institutions has made it an alternative scenario for social media management and interaction with stakeholders. To establish the eventual causal relationship between outstanding variables (access, participation, and quality), an inferential and correlational statistical analysis on data from articles from 40 Latin American companies was carried out. The results show conditional causality of the increase in participation over the quality of the articles, as well as a positive effect on visits. The result implies for organizations to face the dilemma of strengthening their presence on Wikipedia, while not violating the principle of editorial neutrality, due to conflict of interest.

Keywords: Wikipedia, social media, communication users, audience participation, quality control

Resumen
Wikipedia es un ejemplo paradigmático de colaboración en línea. A nivel organizacional, la presencia de contenidos relacionados con empresas e instituciones la ha convertido en un escenario alternativo para la gestión de redes sociales y la interacción con los grupos de interés. Para establecer la eventual relación causal entre variables destacadas (acceso, participación y calidad), se realizó un análisis estadístico inferencial y correlacional de datos de artículos de 40 empresas latinoamericanas. Los resultados muestran una causalidad condicional del aumento de la participación sobre la calidad de los artículos, así como un efecto positivo en las visitas. El resultado implica para las organizaciones enfrentar el dilema

[1] Sergio Llano Aristizábal es profesor en la Universidad Uninorte, Colombia. ORCID ID: https://orcid.org/0000-0003-1481-0621. Contacto principal para la correspondencia editorial.
[2] Héctor López Ospina es profesor de la Universidad de Los Andes, Chile. ORCID ID: https://orcid.org/0000-0001-8144-6394.

de fortalecer su presencia en Wikipedia sin violar el principio de neutralidad editorial por conflicto de intereses.

Palabras claves: Wikipedia, medios sociales, usuarios de la comunicación, participación de la audiencia, control de calidad

Summary

Sumario

1. INTRODUCTION

The Wikipedia encyclopedia is referenced as an exceptional online collaboration system (Aaltonen & Seiler, 2016; Ortega, 2012; Tang et al., 2014). Currently, it has 316 versions, in the same number of languages, and 55,875,354 articles. The English version, the largest, includes 6,251,922 articles, and the Spanish version, the ninth, 1,661,570 articles ("List of Wikipedias," 2021). It is estimated that the content related to organizations in Wikipedia represents 5.18% of the total (Lewoniewski et al., 2019).

At the access level, Wikipedia occupies, in the public rankings of user analytics, the privileged position among the most visited sites in the world; eighth position on Similarweb (n.d.). In the past year, the number of page views in the English version was 129,265,460,237, made from a monthly average of 199,375,282 unique devices. In the Spanish version, 16,478,519,216 page views were recorded from a monthly average of 199,375,282 unique devices (Wikistats, n.d.).

Wikipedia content is the result of the participation of a network of volunteers, who are not always identified. They do not receive any form of financial compensation or recognition for their work ("Comunidad de Wikipedia," 2021). Specifically, the encyclopedia is the product of 2,922,712,745 editions made by 94,403,941 users. In the English version, there are 1,002,766,451 editions made by 40,976,542 users, 148,305 of them active in the past month. The Spanish version has received 132,858,737 editions from 6,123,917 users, 17,805 of them active in the last month (Wikistats, n.d.).

Unlike access and participation, quality is a factor that has been heavily debated, because content creators and editors are not required to prove their training and experience, and there is no centralized editorial system, nor quality assurance reviewers ("Wikipedia Quality," 2021). Despite being a matter that has received wide attention by the encyclopedia itself, quality is understood as a mere matter of reliability ("Reliability of Wikipedia," 2021).

On the other hand, Wikipedia has been an active object of scientific study (Firer-Blaess, 2011). The research has mainly looked at the production and reliability of the content, in addition to related social aspects ("Academic Studies of Wikipedia," 2020).

Regarding research on content associated with companies and organizations, it is recognized as a relevant communicative resource for public evaluation (Etter & Nielsen, 2015), and a

sign of legitimacy for companies and institutions (Kaplan & Haenlein, 2014). In general, the available research emphasizes its value to corporate reputation.

In this regard, the categories of reputational measurement systems can be associated with the contents of articles about companies and organizations in the encyclopedia (González, 2020; Roessing & Enwiller, 2016). Without distinction, both positive and negative content coexist on Wikipedia (DiStaso & Messner, 2010; DiStaso & Messner, 2012). Indeed, Wikipedia is included within the set of media in which digital reputation (e-reputation) is built (Khelladi & Boutinot, 2017), among other reasons, thanks to the fact that user access to this type of articles is reactive to situations of public interest that involve the companies themselves (Llano et al., 2021).

These general features of Wikipedia: its popularity, the participatory model, and its quality, alongside the particular interest of its contribution for the purposes of corporate communication, converge in this article. It presents partial results of an investigation that addresses a variety of aspects regarding the value of Wikipedia for Latin American companies. The question of what causal relationships can be found between indicators related to the access, participation, and quality variables is answered.

2. THEORETICAL FRAME

The prominent position of Wikipedia among global sites is similar in the selected sample countries; it appears at the 11th position in Chile, 12th in Colombia, 13th in Mexico, and 11th in Peru (Similarweb, n.d.). In all cases, it is also the most consulted source in the category of reference materials.

Wikipedia traffic comes from searches in 86% of the cases, 10% from shortcuts, and the remaining from referral links. That is, there is no access from advertising campaigns or paid ads (Similarweb, n.d.). This implies, as is typical of natural traffic, that its popularity is an effect of search engine positioning, largely driven by factors associated with content and social preferences.

The first global study of Wikipedia users by Glott et al. (2010) showed that the average user is young and male: an average age of 25.22 years, half of them under 22 years, and a percentage of 68.99% of male readers, compared to 30.52% of women. Its distribution in educational levels was 11.75% with primary education, 36.11% with secondary education, 25.47% with undergraduate degrees, 17.68% with masters, and 2.95% with doctorates.

Despite the global nature of Wikipedia, access is not uniform throughout the planet, since it differs by time zones (Reinoso et al., 2009); there is a significant increase in hours of highest work activity in the Western Hemisphere (the Americas and Europe), and a significant reduction on weekends.

Although the general estimation is that there is no predominant motivation that drives access to the encyclopedia (Singer et al., 2017), the main extrinsic motivations have been measured for the English version: topics covered in the media (30%), issues that arise in conversations (22%), homework (16%), and current events (13%). Among the intrinsic motivations, (Kiesel et al., 2017): the desire to learn (25%), boredom (20%), and personal decision-making (10%).

An analogous study, expanded to 14 versions of Wikipedia, established differences in motivations and behaviors, from which common traits were identified in countries with similar characteristics (Lemmerich et al., 2019). Motivations related to homework and boredom presented the greatest discrepancies among the different versions. It is inferred that English Wikipedia, the version most attended by research, cannot be taken as representative of the behavior in other versions.

Although the specificities of access to Wikipedia for organizational content are not widely referenced in the reviewed literature, there is empirical evidence regarding popularity indexes, access levels, positioning, and behavior of visits (DiStaso, 2013; DiStaso & Messner, 2010; González, 2020; Llano et al., 2021).

The users who participate in the content production process are called Wikipedians or editors. Anyone can become an editor by making a simple change to the text of the articles. But editing includes multiple additional tasks ("Wikipedians," 2021). Editors are not always identified, so they can contribute anonymously. These users are the ones that generate the greatest inconveniences, since this type of profile tends to abusee the editing guidelines ("Comunidad de Wikipedia," 2021).

The editing work varies by versions. English Wikipedia presents the highest value with 160.3 average edits per article. Spanish is seventh, with 79.9 on average. The number of user edits per article averages to 6.6 in English and 3.7 in Spanish ("List of Wikipedias by edits per article," 2021). It should be noted that automated edits are also performed by bots, especially useful for repetitive tasks ("Bots," 2021).

However, contributions per user show that collaboration in real terms is not as active as it may seem. (Swartz, 2006). This has been typified as a frequent behavior on social web platforms, even giving rise to the formulation of the so-called participatory inequality theory, or the 90-9-1 theory (Nielsen, 2006). According to this, 1% make 50% of the contributions; 9%, 45%; and 90%, only 5% of them. When testing the postulate on different platforms, Tancer (2008) identified only 3.5% of users acting as editors in Wikipedia.

In the profiles of Wikipedians, there are greater differences in gender participation, with 86.73% of editors being male, versus 12.64% being women (Glott et al., 2010). This gender gap has been confirmed in subsequent studies. Recently, only 11.6% of users performing editing tasks on Spanish Wikipedia were found to be women (Minguillón et al., 2021). This lower participation has been attributed to the lack of motivation to collaborate, driven by the absence of content of interest, and to the lack of experience in accumulated over time editing (Hinnosaar, 2019).

However, internet using skills have been identified as a factor that reduces the gender gap in Wikipedia contributions (Hargittai & Shaw, 2015). Education and gender, as argued by Shaw and Hargittai (2018), are therefore the main predictors of content editing work. Contribution to Wikipedia does not stand out on the Internet skills spectrum, the authors say, which, from the point of view of the democratization of knowledge, becomes a type of digital inequality that undermines the potential of the Internet in this collaborative setting. Similarly, a study of 40 versions of the encyclopedia (Miquel-Ribé & Laniado, 2018) revealed notable imbalances in the content affected by the impact of the cultural context of the editors.

One of the fundamental principles of Wikipedia is the neutral point of view, which seeks balance and proportionality as a mechanism to reduce any bias ("Reliability of Wikipedia," 2021). This issue is particularly sensitive in corporate-type information, since editing is discouraged for people linked to companies and organizations, due to the conflict of interest it entails ("Conflict of interest," 2021). However, it has been proven that it is a rule that communication professionals and public relations agencies do not comply with (DiStaso, 2012, 2013).

When it comes to the quality of Wikipedia, this is a subject of extensive academic discussion. This is a complex issue because the encyclopedia itself understands it as a reliability problem: the content is valid and verifiable, rather than truthful (Reliability of Wikipedia, 2021). Therefore, it declares that, although it may not be a reliable source, since it is not subject to an exhaustive review by expert peers but by chance ("General disclaimer," 2020), this does not omit the possibility that it provides valuable and correct information.

Wikipedia, however, has a content evaluation system that classifies articles according to their quality level. This is particularly developed in the English language version, in which more than 5 million articles are classified (Content assessment, 2021). This system, based on a manual evaluation process, has the disadvantage of the considerable volume of data (Wang & Li, 2020), which means that the vast majority of Wikipedia articles are not evaluated, reaching 99% in some languages ("Wikipedia Quality," 2021).

In the Spanish version, there is the denomination of featured articles (Artículos destacados, AD) ("Qué es un artículo destacado," 2021). They are articles that meet the characteristics of being based on reliable sources, being verifiable, neutral, well written; they are complete, extensive, and deep; meet the standard for style and structure and are stable. In Spanish, there are 1,189 articles with this characteristic, that is, only 0.07%.

Automatic evaluation systems are an alternative to the manual evaluation of the quality of the articles. This type of evaluation, as Lewoniewski (2019) notes, poses the challenge of integrating sufficient dimensions of quality. There are marked differences in the studies regarding the set of measures to be taken, as well as in the configuration of the automation algorithm. Wang and Li (2020), in an analysis of the current quality assessment models, conclude that they all have shortcomings by not providing satisfactory results and failing to adopt a comprehensive framework of characteristics.

Wikirank (n.d.) is a free access online tool for automatic evaluation that has gradually incorporated several of these characteristics. It has begun to be referenced in research projects that require quality standardized data ("Wikirank," 2020). In this system, the average quality of the English version, a value established based on all articles in the same language is 61.58. For Spanish Wikipedia, the value is 15.80. This means that the encyclopedia in Spanish reaches only 25.70% of the average quality it reaches in English (Wikirank, n.d.).

The quality variable has also been addressed in content studies of an organizational type in Wikipedia. Within them, it was verified the hypothesis that the quality of an article affects the digital reputation of the organization (Khelladi & Boutinot, 2017). Also, Llano and González (2021) refer to the prevalence of high-quality articles of a corporate type,

regarding the standard of Wikipedia in that language, in contrast to the fact that the individual values were medium and low compared to the measurement scale.

3. METHODOLOGY

In methodological terms, the component of the research on the value of Wikipedia articles for corporate communication that is presented in this article corresponds to a section in which the research was of the inferential type. The analysis and the models were applied to determine causality between the variables and/or indicators of access, participation, and quality.

The sample of articles was made up of 40 companies from Chile, Colombia, Mexico, and Peru that were selected from a sampling frame of the brands with the highest value in each country (BrandZ, 2018). These four countries are part of the Pacific Alliance (Alianza del Pacífico, n.d.), the most outstanding economic bloc in the region, representing 41% of the GDP of Latin America and the Caribbean.

Specifically, the articles analyzed were the following, according to the title of the article of the respective companies in Wikipedia:

- Chile (n=10): Falabella, Compañía de Petróleos de Chile, Banco de Chile, Líder (supermarkets), Latam Airlines Group, París (department store), Parque Arauco, Entel Chile, Compañía de Cervecerías Unidas, Ripley (store).

- Colombia (n=10): Cervecería Bavaria, Tigo Une, Bancolombia, Ecopetrol, Davivienda, Grupo Sura, Grupo Nutresa, Éxito (supermarkets), Banco de Bogotá, Avianca.

- México (n=10): Grupo Modelo, América Móvil, Bodega Aurrerá, Televisa, Grupo Bimbo, Grupo Financiero Banorte, Cemex, El Puerto de Liverpool, Cervecería Cuauhtémoc Moctezuma, Banco Azteca.

- Perú (n=10): Unión de Cervecerías Peruanas Backus y Johnston, Banco de Crédito del Perú, Intercorp, Corporación Lindley, El Pacífico Peruano Suiza Compañía de Seguros y Reaseguros, D'Onofrio, Tiendas Metro, Grupo Gloria, Mi Banco, Alicorp.

The analyzed information consisted of open data obtained with the help of visualization tools on the web. This type of data is those that can be freely used, reused, and redistributed (Open Knowledge Foundation, n.d.). Research with open data, an expression of the wide movement around the concept of *big data*, has become common as the online activities carried out by users are recorded automatically. By its nature, this trend has also permeated Wikipedia studies at different levels (Schroeder & Taylor, 2015). More specifically, the data collected and the tools with which they were obtained were the following:

For the access variable, the number of visits to each article and their daily, monthly, and annual averages were taken as indicators. For the participation variable, the total numbers of editions and editors, the editions in the observed period, the average of daily, monthly, and annual editions, total editors, and the number of editions per editor. In both cases, data was taken for the 2017-2020 period. The access data was extracted from the Pageview Analysis tool, a project hosted on the Toolforge (n.d.) cloud services platform for developers in the Wikimedia community. Participation data was obtained from the Page History

component, a resource that is part of the XTools toolkit of the same organization (Wikimedia Foundation, n.d.).

For the quality variable, the indicators provided by the Wikirank tool (Wikirank, n.d.), an automatic evaluation system of more than 40 million Wikipedia articles in 55 languages, were studied. Quality is assessed from 0 to 100, based on the normalization of the article length metrics, number of references, number of images, number of sections, and ratio of references per extension (Rt). The resulting added value is called the Quality Score. In this case, two data collections were made, one with a cut-off date of December 31, 2018, and the other for December 31, 2020. The selection of this tool is justified by the high volume of articles evaluated, the practicality of measurement compared to other systems, and its demonstrable use as a scientific research tool since its launch in 2015 (“Wikirank,” 2020).

The analysis consisted of the application of inferential statistical techniques, particularly correlational analysis, and simple and multiple linear regression models. The correlation matrices contain Pearson's Correlation Coefficient (rij) between the study variables. This statistical measure allows us to analyze the correspondence or linear relationship between two quantitative variables (i, j) (Weimer, 1996). Table 1 empirically defines the categorization of the correlation index.

Table 1. Correlation index categorization

R Value	Strength of relationship
$\|rij\| \geq 0,7$	Very strong
$0,5 \leq \|rij\| < 0,7$	Strong
$0,3 \leq \|rij\| < 0,5$	Moderate
$\|rij\| < 0,3$	Weak or none

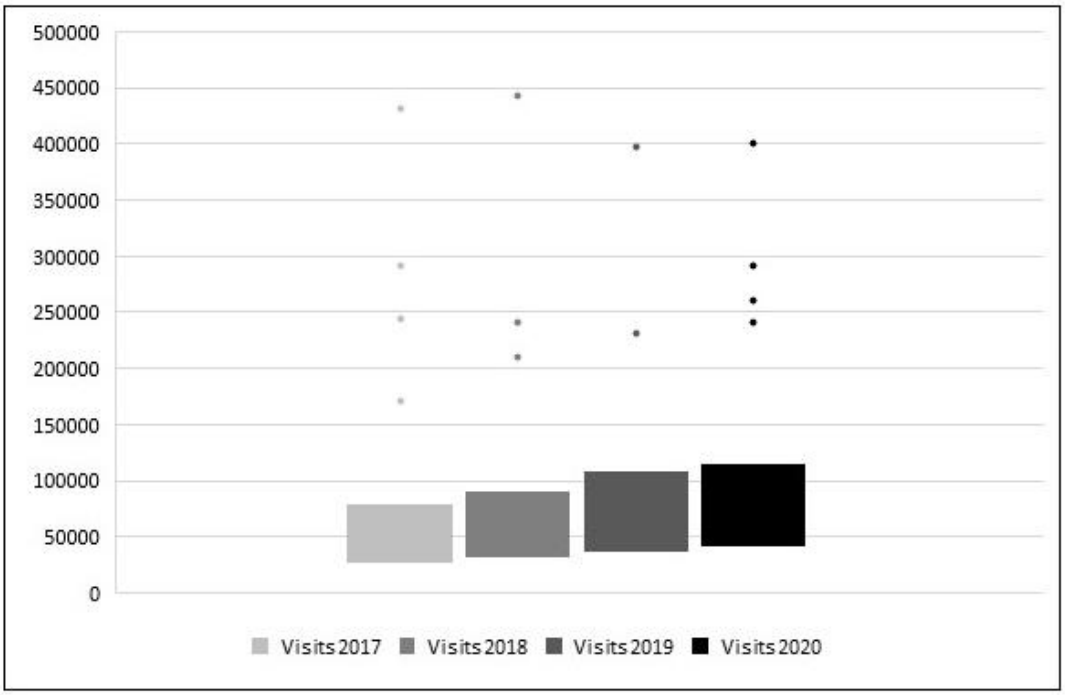
Source: own elaboration.

Subsequently, a linear regression model (Mazzei, et al., 2022; Kim et al., 2022) was applied to establish the causality between the participation and quality indicators with those of access. For this, the indicator with the greatest potential predictor of visits in each one was selected.

4. RESULTS

The results of the analyzed statistical indicators are presented below, in order to establish the existence of causal relationships between the study variables. First, Figure 1 shows the box and whisker plots of the grouped visits for the 40 Wikipedia articles belonging to the companies in the sample, for the 2017-2020 period. It can be seen that the position measures (quartile 1, median, and quartile 3) show slight growth throughout the analysis period. This growth can be seen in more detail in Table 2, which also shows the percentage increase that these values have. Both the median and the average had a greater percentage increase between the years 2019 and 2020. In all cases, atypical data was found. These extreme data correspond to the same companies in the observed time window: Avianca, Televisa, and Bimbo in all four periods, and Grupo Modelo in three periods.

Figure 1. Yearly visits 2017-2020 box and whiskers plots



Source: own elaboration.

Table 2. Yearly visits and percent increments 2017-2020

Statistical measure	Visits 2017	Visits 2018	Visits 2019	Visits 2020	Increment 2017 to 2018	Increment 2018 to 2019	Increment 2019 to 2020
1st quartile	26497	31405	36051	46014	19%	15%	28%
Median	44205	46688	52374	70065	6%	12%	34%
3rd quartile	77590	86378	101055	114110	11%	17%	13%
Average	72798	75615	83281	95041	4%	10%	14%

Source: own elaboration.

Table 3 shows the correlation matrix between the visits of the four years. It can be noted that there is a very high correlation between the view indicator, which implies a high linear relationship between them.

Table 3. Correlation matrix yearly visits 2017-2020

	Visits 2017	Visits 2018	Visits 2019	Visits 2020
Visits 2017	100,0%			
Visits 2018	97,4%	100,0%		
Visits 2019	95,2%	98,3%	100,0%	
Visits 2020	92,8%	93,7%	96,4%	100,0%

Source: own elaboration.

This suggests the hypothesis that it is possible to predict visits for a given year with information from previous years. To test this hypothesis, a simple linear regression model

was performed between visits from 2017 to 2018, and another between 2018 to 2019, and its predictive power was analyzed for the next year. Table 4 shows the estimators of the linear regression models, as well as the value of R^2 that corroborate its validity. The R^2 value of 95% found by model 1 means that 95% of the variation in the visits of 2018 is due to the variation in the data of the visits of 2017. A very similar result was obtained in model 2.

Table 4. Linear model regressions for 2017-2018 and 2018-2019

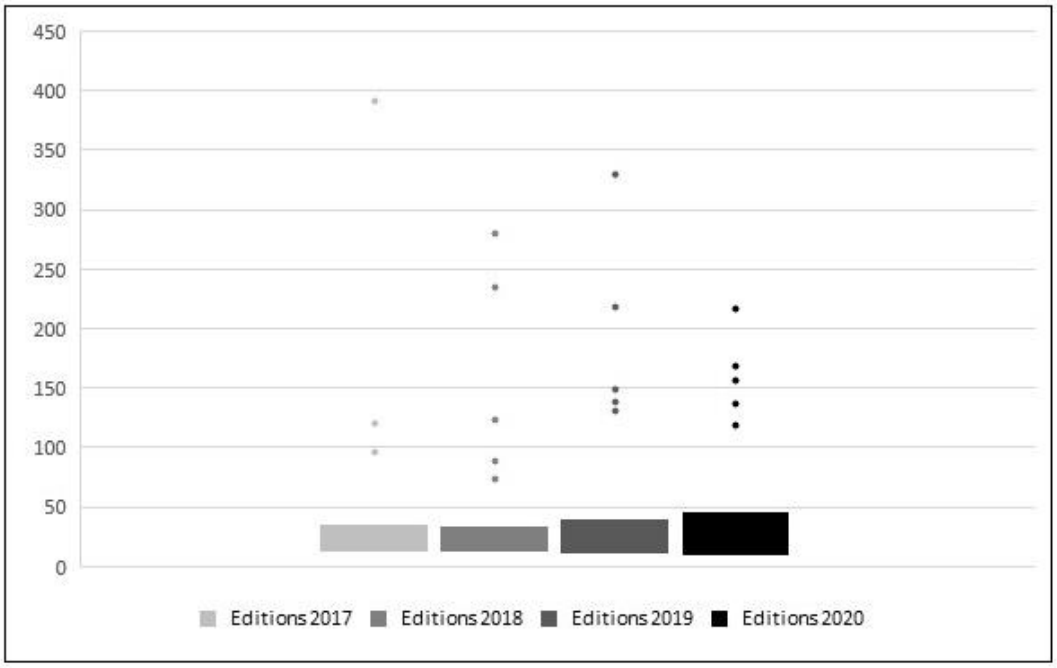
	Equation	R^2
Model 1	Visits 2018 = 8702,86 + 0,919146*Visits 2017	94,78 %
Model 2	Visits 2019 = 11669,1 + 0,947064*Visits 2018	96,57 %

Source: own elaboration.

The application of model 1 showed that, of the 40 companies, 20 (50% of the sample) had a prediction error of less than 10% (2017-2018). In model 2 (2018-2019), the number was only 10 companies (25% of the sample), which dismisses, despite the year-to-year correlation, that reliable predictions can be made with only the indicator of visits.

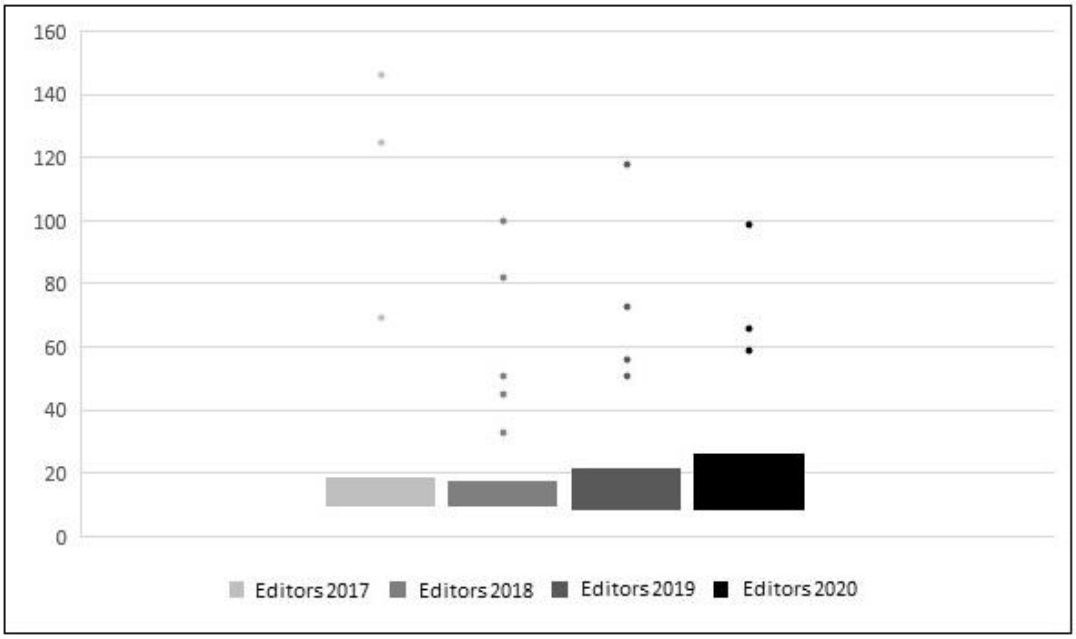
Figures 2 and 3 show the box and whisker plots of the editions and editor participation indicators for the 2017-2020 period in the analyzed set of 40 Wikipedia articles. The outliers were presented for the two indicators, almost identically, in the same organizations: Avianca, Televisa, and Bimbo in the 4 periods, and Falabella in 3 periods.

Figure 2. Yearly editions 2017-2020 box and whisker plots



Source: own elaboration.

Figure 3. Yearly editors 2017-2020 box and whiskers plots



Source: own elaboration.

Furthermore, Tables 5 and 6 present the measurements of quartile 1, median, quartile 3, and the annual average of editions and editors. The results show that there is no pattern in these indicators of the participation variable. That is, there was no increase or decrease over the years. For example, for editions, the highest median was in 2020 (22.5) and the lowest value was in 2019 (17.5). However, the highest average was in 2017 (44.45) and the lowest was in 2018 (39.47). As can be seen, these averages are significantly affected by the outliers that are illustrated in the respective box and whiskers plot (Figure 2).

Table 5. Central tendency averages in yearly editions 2017-2020

	Editions 2017	Editions 2018	Editions 2019	Editions 2020
1st Quartile	11,75	12	11	9,75
Median	19,5	18,5	17,5	22,5
3rd Quartile	36	33,5	38,25	44,25
Average	44,45	39,475	42,125	41,55

Source: own elaboration.

Table 6. Central tendency averages in yearly editors 2017-2020

	Editors 2017	Editors 2018	Editors 2019	Editors 2020
1st Quartile	9,75	9	8	8,75
Median	13	13	12	13,5
3rd Quartile	19	18	21,25	25,5
Average	21,25	18,825	19,625	20,35

Source: own elaboration.

On the other hand, for the editors indicator, the highest median was in 2020, with a value of 13.5, and the lowest was in 2019, with a value of 12. Furthermore, as in editions, the highest average was in 2017 (21.25) and the lowest in 2018 (19.63). It can also be seen that for this same indicator the statistical measures show an almost constant behavior.

After reviewing the central tendency measures, we proceeded to establish the correlation between the participation indicators of editions (2017 to 2020) and editors (2017 to 2020), with the indicator of visits for 2020. The last year was taken as a reference, since the visits showed high correlation in all periods observed, as shown in Table 7.

Table 7. Correlation matrix of 2017-2020 yearly editions and editors with 2020 visits

	Editions 2017	Editors 2017	Editions 2018	Editors 2018	Editions 2019	Editors 2019	Editions 2020	Editors 2020	Visits 2020
Editions 2017	100%								
Editors 2017	98%	100%							
Editions 2018	93%	93%	100%						
Editors 2018	93%	94%	98%	100%					
Editions 2019	74%	74%	79%	81%	100%				
Editors 2019	81%	83%	89%	92%	88%	100%			
Editions 2020	73%	75%	77%	81%	82%	84%	100%		
Editors 2020	82%	84%	89%	90%	85%	93%	88%	100%	
Visits 2020	65%	75%	69%	73%	69%	77%	70%	71%	100%

Source: own elaboration.

The values denote a high correspondence or linear relationship between almost all the indicators in the matrix. In particular, the 2019 editions are highly correlated with both editions and editors from the other years. In addition, it is the value that is most related to the 2020 visits.

As one of the purposes of the present research was to design an adequate linear regression model that associates the indicators of the quality variable with the visits of the year 2020, in order to establish the causality of the three study variables, 2019 editors was the only indicator selected as a potential independent variable of the statistical model.

Before referring to the model, a statistical analysis of the quality indicators was carried out with two statistical time cuts (2018 and 2020). In the first place, the correlation between the quality indicators of the year 2020 and those of the year 2018 is shown (Table 8). All the indicators of the year 2018 are highly correlated with their respective indicators of the year 2020. For this reason, it is not necessary to use the indicators of the two cut-off years in the multiple linear regression model, since said correlation allows for selecting one of the two to explain 2020 visits. As the most current set of data, the 2020 indicators were selected.

Table 8. Correlation matrix quality indicators 2018 and 2020

	Article length 2018	Number of references 2018	Number of images 2018	Number of sections 2018	Rt references per length 2018	Quality Score 2018	Quality percentile 2018
Article length 2020	85%	79%	54%	77%	10%	76%	45%
Number of references 2020	61%	77%	31%	61%	29%	65%	36%
Number of images 2020	56%	42%	87%	46%	12%	63%	63%
Number of sections 2020	67%	66%	46%	89%	20%	74%	58%
Rt references per length 2020	18%	49%	11%	27%	78%	50%	57%
Quality Score 2020	72%	78%	66%	76%	42%	87%	74%
Quality percentile 2020	34%	40%	55%	46%	48%	61%	84%

Source: own elaboration.

Table 9 shows the percentage of increase (positive) or decrease (negative) in the evaluations of the quality indicators in 2018 and 2020 for quartile 1, the median, quartile 3, and the average. It can be noted that the quality variable, expressed in the Quality Score, and that integrates all other indicators, has a relevant increase in all measures. For example, on the average, the increase was almost 16%, and on the median, it was 26.4%. Even so, some values showed a decrease in some of these statistical measures. For example, the valuation of the number of sections had an average decrease of almost 5%. The ratio of referrals by extension also had a decrease in the median of 8.3%, although its variation in the average level was almost nil. And it is important to note that the quality percentile also has an increase in its assessment because this variable is associated in part with the increase in the assessment of the Quality Score. This is particularly notable in quartile 1, which includes the articles with the lowest quality in the sample.

Table 9. Quality indicators variation 2018-2020

	Article length	Number of references	Number of images	Number of sections	Rt references per length	Quality Score	Quality percentile
1st quartile	16,4%	14,9%	-4,8%	-4,6%	9,2%	20,5%	24,6%
Median	-3,5%	6,6%	0,5%	-4,5%	-8,3%	26,4%	5,5%
3rd quartile	0,0%	-9,4%	-7,7%	-4,5%	7,2%	13,7%	1,0%
Average	3,5%	12,7%	3,2%	-4,8%	-0,1%	15,9%	7,6%

Source: own elaboration.

After the previous analysis, the aim was to study the impact of the described indicators on 2020 visits. Table 10 presents the correlation matrix of the quality indicators of the year 2020, and 2019 with them. Note that, individually, the length of the article, the number of references, Quality Score, and 2019 editors have a very strong relationship or linear correspondence with visits. In addition, the assessment of the number of sections for the year 2020 has a moderately strong relationship with a correlation of 59%.

Table 10. Correlation matrix of 2020 quality, 2019 editors, and 2020 visits

	Article length 2020	Number of references 2020	Number of images 2020	Number of sections 2020	Rt references per length 2020	Quality Score 2020	Quality percentile 2020	Editors 2019	Visits 2020
Article length 2020	100%								
Number of references 2020	83%	100%							
Number of images 2020	49%	23%	100%						
Number of sections 2020	80%	62%	46%	100%					
Rt references per length 2020	26%	59%	7%	26%	100%				
Quality Score 2020	85%	82%	59%	80%	60%	100%			
Quality percentile 2020	43%	39%	54%	54%	56%	74%	100%		
Editors 2019	82%	58%	41%	62%	15%	66%	26%	100%	
Visits 2020	80%	68%	48%	59%	33%	76%	39%	77%	100%

Source: own elaboration.

Based on these results, various multiple linear regression models were tested with the potential variables of the previous matrix. In particular, the goal was to establish the models that would reveal the eventual causal relationship between quality variables and access to Wikipedia.

After testing different models, the ones with the best fit were the one that includes 2019 editors and 2020 Quality Score with an R^ 2 of 87.3%, and the one that article length 2020, the ratio of references per length 2020 and editors 2019 with an R^2 of 87.5%. To establish the level of causality of the quality indicators independently, the second model did not consider the Quality Score, as it is an aggregate measure.

Table 11 shows the result of the estimation of the first model, in which the visits are related to the independent variables 2019 editors and 2020 Quality Score.

Table 11. Regression model 2019 editors, 2020 quality

Parameter	Estimation	Standard error	T statistic	P-value
Editors 2019	1991,94	427,676	4,6576	0,0000
Quality Score 2020	1396,79	266,694	5,23745	0,0000

Source: own elaboration.

The results of some statistical measures describe some of the benefits of the models:

- R-squared = 87,3088%
- Durbin-Watson = 1,89601

The linear regression model to describe the relation between 2020 visits and the independent variables is:

*Visits 2020 = 1991,94 * Editors 2019 + 1396,79 * Quality Score 2020*

The model shows that an increase in an editor translates into an increase in annual visits by an average value of 1991. Similarly, an increase of one point in the Quality Score translates into an increase of 1,396 average annual visits. Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables with a confidence level of 95.0%.

The R-Squared statistic indicates that the model explains 87.3088% of the variability in the 2020 visits. The adjusted R-Squared statistic, which is more appropriate to compare models with different number of independent variables, is 86.9748%.

To determine if the model can be simplified, note that the highest P-value of the independent variables is 0.0000, which corresponds to Editors 2019. Since the P-value is less than 0.05, this term is statistically significant with a confidence level of 95.0%. Consequently, no variables should be eliminated from the model. Table 12 presents the confidence intervals for the estimates of the coefficients of the linear regression model.

Table 12. 95% confidence intervals for coefficient estimations

Parameter	Estimation	Standard error	Lower limit	Upper limit
Editors 2019	1991,94	427,676	1126,16	2857,73
Quality Score 2020	1396,79	266,694	856,9	1936,69

Source: own elaboration.

Table 13 shows the result of the estimation of the second statistical model that relates visits to the independent variables article length 2020, ratio of references per length 2020, and editors 2019. The statistical measures that describe the benefits of the model are presented below.

Table 13. Regression model 2020 article length, 2020 ratio of references per length, 2019 editors

Parameter	Estimation	Standard error	T statistic	P-value
Article length 2020	1710,83	617,318	2,77139	0,0087
Rt references per length 2020	437,373	170,243	2,56912	0,0144
Editors 2019	1490,51	596,262	2,49976	0,0170

Source: own elaboration.

- R-squared = 87,4645%
- Durbin-Watson = 2,01504

The linear regression model to describe the relation between 2020 visits and the independent variables is:

*Visits 2020 = 1710,83 * Article length 2020 + 437,373 * Rt references per length 2020 + 1490,51 * Editors2019*

The model shows that a percentage point in article length translates into an increase in annual visits by an average of 1710. Similarly, the increase in one percentage point in the Rt of references per length translates into an average increase of 437 visits. This shows the causal relationship between quality indicators and access.

The R-Squared statistic indicates that the model explains 87.4645% of the variability in the 2020 visits. To determine if the model can be simplified, note that the highest P-value of the independent variables is 0.0170, which corresponds to 2019 editors. Since the P-value is less than 0.05, that term is statistically significant with a confidence level of 95.0%.

5. DISCUSSION AND CONCLUSIONS

This research exercise, in which the causal relationships between the variables of access, participation, and quality were analyzed, based on their statistical description, shows a high linear relationship between them, although with some particularities that are explained below.

The atypical data of yearly visits showed stability, given the little significant variation. This type of results tended to be presented in the same companies. This suggests that, as has been previously mentioned, the characteristic features of companies are likely to be associated with the behavior of visits (González, 2020; Llano et al., 2021). Wikipedia, therefore, deserves to be seen as a space for communication with variable impact according to the types of companies and specific cases.

In the research, there was homogeneity in the growth of the grouped visits, that is, there was a constant increase in the periods observed. The increase in visits, which means greater visibility for organizations, is explained in the linear regression including yearly visits, which showed a positive marginal impact equal to 1. Since a new visit in a particular year becomes an additional visit the following year, the model proved causality in access. It should be noted, however, that the model does not offer a high predictive level.

The access also shows that the potential impact of Wikipedia as a source of information for the external audiences of companies tends to increase. This has been reported in similar studies, in which behavior was reviewed over time is corroborated, such as DiStaso (2010), in which there was an increase in the number of visits.

If increased access is translated into greater visibility for corporate information on Wikipedia, the direct effect on reputational issues, a commonly accepted assumption in the literature (DiStaso & Messner, 2010, 2012; Khelladi & Boutinot 2017; Roessing & Einwiller, 2016) is reaffirmed in this research. Although there are no studies that account for the motivations for specific access to corporate content, it is possible to associate the interest motivated by business events covered in the media (Lemmerich et al., 2019), with extreme changes in visitor behavior as reported by Llano et al. (2021).

Regarding participation, the stability in the edition indicators contrasts with the increase in visits since the aggregated data did not show variations, although they did occur in some cases. Again, the values indicate that each company deserves an individual review to detail its own specificities.

The fact that the aggregate number of users who performed editing tasks did not increase confirms the type of participation that predominates in 2.0 platforms (Nielsen, 2006; Swartz, 2006; Tancer, 2008) where the greatest editing activity is concentrated in a small percentage of editors.

Indeed, participation shows dichotomous values compared to the averages of the encyclopedia. On the one hand, the average number of editions per article in the reviewed sample: 231 in Chile, 820 in Colombia, 957 in Mexico, and 159 in Peru ("List of Wikipedias by edits per article," 2021), surpasses the average of 106.3 of English Wikipedia, and the 79.9 average of Spanish Wikipedia (Toolforge, n.d.). On the contrary, the average number of user edits per article, which is 6.6 in English and 3.7 in Spanish, exceeds the averages of 2.04 in the sampled articles for Chile, 2.26 for Colombia, 2.14 for Mexico, and 2.01 for Peru. This means that, although participation is sufficiently active, the number of participants is small. Furthermore, if one takes into consideration that the main motivations for getting involved in editing are the desire to share knowledge and the desire to contribute (Glott et al., 2010), it is to be assumed that corporate information arouses less interest the encyclopedia users in Spanish.

On quality, the added value in the two observed periods showed a notable increase. Between 2018 and 2020, the quartile of companies with the lowest quality (quartile 1) showed the highest level of improvement, with 20.5% in the Quality Score as a whole, and 24.6% in the quality percentile, although in the higher quality quartile (quartile 3), there was no significant variation. In the grouping by countries, the increase was notable, since it went

from 29.8 to 44.1 in the group of Chilean companies, from 39.8 to 45.1 for the Colombians, 45.5 to 53.5 for the Mexicans, and 26.5 to 28.0 for the Peruvians.

If the total number of articles that make up Spanish Wikipedia (Wikirank, n.d.) is taken as a reference, where the Quality Score went from 14.86 in 2018 to 15.80 in 2020, the quality of the articles in the sample was always higher. In addition, all the articles tended to be in the highest quality percentiles for the 2020 measurement (98 for Mexico, 96 for Chile and Colombia, and 86 for Peru) when accounting for all the articles in the Spanish language encyclopedia. In contrast, the quality of the sampled articles, in all cases (with the exception of articles from Mexico in 2020), is below the average value (50) of the highest possible quality, which speaks of a very large room for improvement.

In summary, the set of articles in the sample is part of the highest quality articles that currently make up Spanish Wikipedia, but most do not reach the average quality of the measurement standard. In this order of ideas, although corporate information occupies an intermediate place among the topics with the highest volume (Lewoniewski et al., 2019), the fact that the articles of large companies are among the highest quality in the Spanish version, does not necessarily make them of lower quality compared to those of the most popular subject categories.

On the other hand, the correlation obtained between the indicators of participation and quality reflects that as the former increases, the length of the articles becomes greater. However, if the fact that the correlation between participation and quality is medium is taken into account, it follows that an increase in the first variable will not necessarily have an effect on the second. In other words, it is a conditional causality, since to have a real effect, it must be a qualified participation.

Proof of this is that in the quartiles in which the quality increased the most between 2018 and 2020 (quartile 1 and median), those of the companies with the lowest quality values, there was a more stable number of editors. On the contrary, in the companies with the least increase (quartile 3), those with the highest quality values, there was an increase in participation. Participation in quartile 1 and a half was, more than abundant, sufficiently qualified.

The regression model also demonstrated effects of both participation and quality on visits. That is, if participation in an article is increased, an increase in visits is to be expected. Similarly, an improvement in quality, particularly in the length of the article and the number of references, also translates into an increase in visits.

In short, the linear regression model of visits, quality, and participation showed that a greater number of editors and an improvement in the quality of the articles have a positive effect on visits, thus proving the existence of a causal relationship between the participation and quality variables on the access variable.

Given that, in practical terms, a considerable and growing volume of visits means greater effects on the communicative and participatory processes in social media, the results of this research serve to draw the attention of organizations with a presence on Wikipedia to the challenges posed by this space of non-conventional interaction.

By revalidating the assumption that the visibility of products, brands, and organizations have turned Wikipedia into an alternative space for interaction with its different audiences on the social web (Etter & Nielsen, 2015; Kaplan & Haenlein, 2014), companies face the challenge to find mechanisms that give impetus to the editing tasks by volunteers, without directly assuming the task or violating the policies of the encyclopedia.

But the low levels of user participation, as the study makes clear, add a restrictive condition that is difficult to solve. The need to correct, update, and improve articles, especially with issues that require immediate attention, is faced with the low availability of editors. This work, as it falls exclusively on volunteers ("Wikipedians," 2021), creates a dilemma for the organizations involved when faced with the recommendation not to edit when there is a "Conflict of Interest" (2021). The lack of volunteers exerts a clear pressure on the companies and agencies that provide them with services that may be forced to undertake editing by themselves, as indeed, happens (DiStaso, 2012, 2013).

This can be particularly sensitive in matters where reputation comes into play, especially those that are negative (DiStaso & Messner, 2010; 2012). Consider the prevailing emphasis in the literature on corporate content on Wikipedia that attributes the greatest effects to reputational issues (DiStaso & Messner, 2010, 2012; Khelladi & Boutinot, 2017; González, 2020; Llano et al., 2021; Roessing & Enwiller, 2016), the same ones that companies naturally tend to protect.

While this research makes it clear that qualified participation compensates for the small and stable number of editors, it also means that editing is always focused on the same participants. This suggests that the articles are not constructed in a process as democratic as the Wikipedia ideal is.

Ultimately, if organizations choose to respect the policies that Wikipedia proposes to companies but want to turn it into a platform that is integrated into their social media strategies, designing alternative mechanisms to expand the number of volunteer editors is recommended. For example, motivating ambassadors, fans, and admirers of their brands to join this work as a way of awakening the respective interest (Hinnosaar, 2019), rather than leaving it to the occurrence of incidental events (Zhang et al., 2019). It would be that they contribute to give an indirect impulse to the edition in such a way that the improvements of the articles are organically translated into a much greater editing activity (Aaltonen & Seiler, 2016). The cultural particularities of editors (Miquel-Ribé & Laniado, 2018) should also be considered by the companies for the construction of better strategies to motivate edition.

In any case, it is understood that the need for companies to see their complete articles, without errors, without biased or false information, without direct participating, confronts their obligation to defend their corporate value with the conditions of the Wikipedia guidelines. Therefore, it is not surprising that the pressures and restrictions described lead them to find ways to evade the policies of the encyclopedia. Characterization of this type of practice remains to be identified.

6. REFERENCES

Aaltonen, A. & Seiler, S. (2016). Cumulative growth in user-generated content production: evidence from Wikipedia. *Management Science, 62*(7), 2054-2069. https://doi.org/10.1287/mnsc.2015.2253

Academic Studies of Wikipedia. (2020, October 22) In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Academic_studies_of_Wikipedia&oldid=984931689

Alianza del Pacífico (n.d.). ¿Qué es la Alianza del Pacífico? Retrieved April 22, 2019, from https://alianzapacifico.net/que-es-la-alianza

Bots. (2021, April 11). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Bots&oldid=1017176057

BrandZ. (2018). *BrandZ top 50 most valuable Latin American brands 2018.* WPP. https://shorturl.at/fiv16

Comunidad de Wikipedia. (2021, April 3). In *Wikipedia*. https://es.wikipedia.org/w/index.php?title=Comunidad_de_Wikipedia&oldid=134496074

Conflict of interest. (2021, January 28). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Conflict_of_interest&oldid=1003365958

Content assessment. (2021, March 8) In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Content_assessment&oldid=1011022543

Distaso, M. W. (2012). Measuring public relations Wikipedia engagement: How bright is the rule. *Public Relations Journal,* 6(2), 1-22.

DiStaso, M. W. (2013). Perceptions of Wikipedia by public relations professionals: A comparison of 2012 and 2013 surveys. *Public Relations Journal, 7*(3), 1-23.

DiStaso, M. W., & Messner, M. (2010). Forced transparency: Corporate image on Wikipedia and what it means for public relations. *Public Relations Journal, 4*(2), 1-23.

DiStaso, M.W., & Messner, M. (2012). Wikipedia's role in reputation management: An analysis of the best and worst companies in the USA. *Digithum* (14), 1-9.

Etter, M. & Nielsen, F. (2015). Collective remembering of organizations: Co-construction of organizational pasts in Wikipedia. *Corporate Communications, 20*(4), 431-447. https://doi.org/10.1108/CCIJ-09-2014-0059

Firer-Blaess, S. (2011). *Wikipedia: State of the art of the research.* Academia. https://www.academia.edu/1539270/Wikipedia_State_of_the_art_of_the_research_April_2011_

General disclaimer. (2020, January 10) In *Wikipedia.* https://en.wikipedia.org/w/index.php?title=Wikipedia:General_disclaimer&oldid=935141538

Glott, R., Schmidt, P., & Ghosh, R. (2010). *Wikipedia survey–overview of results.* Wikipedia Study. http://www.wikipediastudy.org/docs/Wikipedia_Overview_15March2010-FINAL.pdf

González, L. E. (2020). *Visibilidad, calidad y pertinencia para la comunicación corporativa de artículos de Wikipedia de empresas colombianas que participan en el mercado bursátil* [Master's Thesis, Universidad del Norte]. Repositorio Digital de la Universidad del Norte. https://manglar.uninorte.edu.co/bitstream/handle/10584/8827/140073.pdf

Hargittai, E., & Shaw, A. (2015). Mind the skills gap: the role of Internet know-how and gender in differentiated contributions to Wikipedia. *Information, communication & society, 18*(4), 424-442. https://doi.org/10.1080/1369118X.2014.957711

Hinnosaar, M. (2019). Gender inequality in new media: Evidence from Wikipedia. *Journal of Economic Behavior & Organization, 163*, 262-276. https://doi.org/10.1016/j.jebo.2019.04.020

Kaplan, A., & Haenlein, M. (2014). Collaborative projects (social media application): About Wikipedia, the free encyclopedia. *Business horizons, 57*(5), 617-626. https://doi.org/10.1016/j.bushor.2014.05.004

Khelladi, I., & Boutinot, A. (2017). The role of Wikipedia on corporate e-reputation: Evidence from French companies. *International Studies of Management & Organization, 47*(1), 23-41. https://doi.org/10.1080/00208825.2017.1241087

Kiesel, J., Potthast, M., Hagen, M., & Stein, B. (2017). Spatio-temporal analysis of reverted Wikipedia edits. In *Proceedings of the International AAAI Conference on Web and Social Media, Canada* (pp.122-131). AAAI.

Kim, S., Austin, L., Liu, B. F., & Jin, Y. (2022). Exploring differences in crisis literacy and efficacy on behavioral responses during infectious disease outbreaks. *Public Relations Review, 48*(3), 102204. https://doi.org/10.1016/j.pubrev.2022.102204

Lemmerich, F., Sáez-Trumper, D., West, R., & Zia, L. (2019). Why the world reads Wikipedia: Beyond English speakers. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (pp. 618-626). ACM. https://doi.org/10.1145/3289600.3291021

Lewoniewski, W. (2019). Measures for Quality Assessment of Articles and Infoboxes in Multilingual Wikipedia. In Abramowicz, W., Paschke, A. (Eds), *Lecture Notes in Business Information Processing* (Vol 339, pp. 619-633). Springer. https://doi.org/10.1007/978-3-030-04849-5_53

Lewoniewski, W., Węcel, K., & Abramowicz, W. (2019). Multilingual ranking of Wikipedia articles with quality and popularity assessment in different topics. *Computers*, *8*(60), 1-31. https://doi.org/10.3390/computers8030060

Llano A., S. & González, L. E. (2021). Comunicación corporativa y Wikipedia: artículos de empresas colombianas cotizadas en bolsa. *Área abierta, 21*(2), 153-170. https://doi.org/10.5209/arab.72710

Llano A., S., López-Ospina, H., & González, L. E. (2021). Influencia potencial de artículos de empresas latinoamericanas en Wikipedia. *Signo y Pensamiento, 40*(78). https://doi.org/10.11144/Javeriana.syp40-78.ipae

List of Wikipedias. (2021, January 17). In *Meta-Wiki.* https://meta.wikimedia.org/w/index.php?title=List_of_Wikipedias&oldid=20978317

List of Wikipedias by edits per article. (2021, January 5). In *Meta-Wiki*. https://meta.wikimedia.org/w/index.php?title=List_of_Wikipedias_by_edits_per_article&oldid=20930559

Mazzei, A., Ravazzani, S., Fisichella, C., Butera, A., & Quaratino, L. (2022). Internal crisis communication strategies: Contingency factors determining an accommodative approach. *Public Relations Review, 48*(4), 102212. https://doi.org/10.1016/j.pubrev.2022.102212

Minguillón, J., Meneses, J., Aibar, E., Ferran-Ferrer, N., & Fàbregues, S. (2021). Exploring the gender gap in the Spanish Wikipedia: Differences in engagement and editing practices. *Plos one, 16*(2), 1-21. https://doi.org/10.1371/journal.pone.0246702

Miquel-Ribé, M., & Laniado, D. (2018). Wikipedia culture gap: Quantifying content imbalances across 40 language editions. *Frontiers in Physics,* 6(54). 1-14. https://doi.org/10.3389/fphy.2018.00054

Nielsen, J. (2006, October 8). *The 90-9-1 rule for participation inequality in social media and online communities.* Nielsen Norman Group. https://www.nngroup.com/articles/participation-inequality/

Open Knowledge Foundation. (n.d.). What is open? Retrieved May 10, 2021 from https://okfn.org/opendata/

Ortega S., J. F. (2012). *Wikipedia: A quantitative análisis* [Doctoral dissertation, Universidad Rey Juan Carlos]. BURJC digital. http://hdl.handle.net/10115/11239

Qué es un artículo destacado. (2021, February 10). In *Wikipedia.* https://es.wikipedia.org/w/index.php?title=Wikipedia:Qu%C3%A9_es_un_art%C3%ADculo_destacado&oldid=133112989

Reinoso, A. J., Gonzalez-Barahona, J. M., Robles, G. & Ortega, F. (2009). A quantitative approach to the use of the Wikipedia. In *2009 IEEE Symposium on Computers and Communications* (pp.56-61). IEEE. https://doi.org/10.1109/ISCC.2009.5202401

Reliability of Wikipedia. (2021, January 23) In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Reliability_of_Wikipedia&oldid=1002133079

Roessing, T., & Einwiller, S. (2016). Portrayals of large corporations in the English and German version of Wikipedia - Exploring similarities and differences. *Corporate Reputation Review, 19*(2), 108-126. https://doi.org/10.1057/crr.2016.3

Schroeder, R., & Taylor, L. (2015). Big data and Wikipedia research: social science knowledge across disciplinary divides. *Information, Communication & Society*, *18*(9), 1039-1056. https://doi.org/10.1080/1369118X.2015.1008538

Shaw, A., & Hargittai, E. (2018). The pipeline of online participation inequalities: The case of Wikipedia editing. *Journal of Communication, 68*(1), 143-168. https://doi.org/10.1093/joc/jqx003

Similarweb. (n.d.). Clasificación de los principales sitios web. Retrieved May 22, 2021, from https://www.similarweb.com/es/top-websites/

Singer, P., Lemmerich, F., West, R., Zia, L., Wulczyn, E., Strohmaier, M., & Leskovec, J. (2017). Why we read Wikipedia. In *Proceedings of the 26th international conference on world wide web* (pp.1591-1600). ACM. https://doi.org/10.1145/3038912.3052716

Swartz, A. (2006, September 4). *Who writes Wikipedia?* Raw Thought. http://www.aaronsw.com/weblog/whowriteswikipedia

Tancer, B. (2008). *Click: what millions of people are doing online and why it matters.* Hyperion Books.

Tang, X., Wang, H., Ouyang, Z. & Yu, W. (2014). Seek the Consent, Respect the Dissent: An Analysis of User Behaviors in Online Collaborative Community. In J. S. Pan, V. Snasel, E. Corchado, A. Abraham & S. L. Wang (Eds), *Intelligent Data analysis and its Applications* (Vol. 1, pp. 223-234). Springer. https://doi.org/10.1007/978-3-319-07776-5_24

Toolforge. (n.d.). *Pageview analysis. Comparison of pageviews across multiple pages*. Retrieved February 3, 2021, from https://pageviews.toolforge.org/

Wang, P., & Li, X. (2020). Assessing the quality of information on Wikipedia: A deep-learning approach. *Journal of the Association for Information Science and Technology, 71*(1), 16-28. https://doi.org/10.1002/asi.24210

Weimer, R. C. (1996). *Estadística.* Cecsa.

Wikimedia Foundation Labs. (n.d.). *XTools. Feedeing your data hunger.* Retrieved February 4, 2021, from https://xtools.wmflabs.org/

Wikipedia Quality. (2021, January 23). In *Wikipedia Quality*. https://wikipediaquality.com/index.php?title=Wikipedia_Quality&oldid=26916

Wikipedians. (2021, March 2) In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Wikipedians&oldid=1009822494

Wikirank. (2020, September 24). In *Wikipedia Quality*. https://wikipediaquality.com/index.php?title=WikiRank&oldid=25408

Wikirank. (n.d.). *Quality and popularity assessment of Wikipedia.* Retrieved February 5, 2021. Available at: https://wikirank.net/.

Wikistats. (n.d.). *Wikimedia statistics.* Retrieved May 20, 2021 from https://stats.wikimedia.org/

Zhang, A. F., Wang, R., Blohm, E., Budak, C., Robert, L. P., & Romero, D. M. (2019). Participation of new editors after times of shock on Wikipedia. *Proceedings of the International AAAI Conference on Web and Social Media*, *13*(1). 560-571.