BUSTECH 2022

The Twelfth International Conference on Business Intelligence and Technology

ISBN: 978-1-61208-955-3

April 24 - 28, 2022

Barcelona, Spain

BUSTECH 2022 Editors

Lorena Parra, Universitat Politecnica de Valencia, Spain

BUSTECH 2022

Forward

The Twelfth International Conference on Business Intelligence and Technology (BUSTECH 2022), held on April 24 - 28, 2022, continued a series of events covering topics related to business process management and intelligence, integration and interoperability of different approaches, technology-oriented business solutions and specific features to be considered in business/technology development.

Similar to the previous edition, this event attracted excellent contributions and active participation from all over the world. We were very pleased to receive top quality contributions.

We take here the opportunity to warmly thank all the members of the BUSTECH 2022 technical program committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to BUSTECH 2022. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the BUSTECH 2022 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope BUSTECH 2022 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the area of business intelligence and technology. We also hope that Barcelona provided a pleasant environment during the conference and everyone saved some time to enjoy the historic charm of the city

BUSTECH 2022 Steering Committee

Malgorzata Pankowska, University of Economics in Katowice, Poland
Jürgen Sauer, Universität Oldenburg, Germany
Pierre Hadaya, ESG-UQAM, Canada
Hans-Friedrich Witschel, University of Applied Sciences and Arts, Northwestern Switzerland
Anbang Xu, IBM Research – Almaden, USA

BUSTECH 2022 Publicity Chairs
Javier Rocher, Universitat Politecnica de Valencia, Spain
Lorena Parra, Universitat Politecnica de Valencia, Spain

BUSTECH 2022

Committee

BUSTECH 2022 Steering Committee

Malgorzata Pankowska, University of Economics in Katowice, Poland
Jürgen Sauer, Universität Oldenburg, Germany
Pierre Hadaya, ESG-UQAM, Canada
Hans-Friedrich Witschel, University of Applied Sciences and Arts, Northwestern Switzerland
Anbang Xu, IBM Research – Almaden, USA

BUSTECH 2022 Publicity Chairs

Javier Rocher, Universitat Politecnica de Valencia, Spain
Lorena Parra, Universitat Politecnica de Valencia, Spain

BUSTECH 2022 Technical Program Committee

Emna Ben Abdallah, University of Sfax, Tunisia
Unal Aksu, Utrecht University, Netherlands
Reza Barkhi, Virginia Tech, USA
Khouloud Boukadi, Faculty of Economics and Management of Sfax, Tunisia
Simone Braun, Offenburg.University, Germany
Karl Cox, University of Brighton, UK
Beata Czarnacka-Chrobot, Warsaw School of Economics, Poland
Giuseppe A. Di Lucca, University of Sannio - RCOST (Research Center on Software Technology), Italy
António Dourado, University of Coimbra, Portugal
Helena Dudycz, Wroclaw University of Economics, Poland
Johannes Edler, University of Applied Sciences Upper Austria Campus Hagenberg, Austria
Bedilia Estrada-Torres, Universidad de Sevilla, Spain
Jiri Feuerlicht, University of Economics, Prague, Czech Republic
Lixin Fu, University of North Carolina at Greensboro, USA
Faiza Ghozzi, University of Sfax, Tunisia
Todorka Glushkova, Paisii Hilendarski University of Plovdiv, Bulgaria
Manuel Gomez-Olmedo, Universidad de Granada, Spain
Foteini Grivokostopoulou, University of Patras, Greece
Pierre Hadaya, ESG-UQAM, Canada
Rawad Hammad, University of East London, UK
Mariem Haoues, Mir@cl Laboratory - University of Sfax, Tunisia
Uswatun Hasanah, Universitas Amikom Purwokerto, Indonesia
Ioannis Hatzilygeroudis, University of Patras, Greece
Hmida Hmida, Higher Institute of Technological Studies of Bizerte, Tunisia
Wladyslaw Homenda, Warsaw University of Technology, Poland
Wei-Chiang Hong, School of Computer Science and Technology - Jiangsu Normal University, China

Wassim Jaziri, Taibah University, Saudi Arabia
Maria João Ferreira, Universidade Portucalense, Portugal
Jānis Kampars, Riga Technical University, Latvia
Mouna Torjmen Khemakhem, University of Sfax, Tunisia
Wiem Khlif, University of Sfax, Tunisia
Alexandr Kormiltsyn, Estonian Research Information System, Estonia
Adrianna Kozierkiewicz, Wroclaw University of Science and Technology, Poland
Kyriakos Kritikos, ICS-FORTH, Greece
Amir Laadhar, IRIT, France
Hongxiu Li, Tampere University, Finland
Audrone Lupeikiene, Vilnius University, Lithuania
Alex Norta, Tallinn University of Technology, Estonia
Hanen Ochi, Efrei Paris, France
Malgorzata Pankowska, University of Economics in Katowice, Poland
Ripon Patgiri, National Institute of Technology Silchar, Assam, India
Isidoros Perikos, University of Patras, Greece
Erwin Pesch, University in Siegen / HHL - Leipzig Graduate School of Management | CASiM - Center for Advanced Studies in Management, Germany
Stavros Ponis, National Technical University Athens, Greece
Patricio Ramírez-Correa, Universidad Católica del Norte, Chile
Franck Ravat, IRIT - Toulouse University, France
Ulrich Reimer, University of Applied Sciences St. Gallen, Switzerland
Michele Risi, University of Salerno, Italy
Gunter Saake, Otto-von-Guericke-University of Magdeburg, Germany
Luis Enrique Sánchez Crespo, University of Castilla-la Mancha, Spain
Jürgen Sauer, Universität Oldenburg, Germany
Isabel Seruca, Portucalense University, Porto, Portugal
Paweł Skrzypek, 7bulls.com - SME IT company, Poland
Stanimir Stoyanov, Plovdiv University "Paisii Hilendarski", Bulgaria
Zoltán Szabó, Corvinus University of Budapest, Hungary
Mohan Tanniru, University of Arizona / Henry Ford Health System, Detroit / Oakland University, Rochester, USA
Chibuzor Udokwu, University of Applied Sciences Upper Austria, Austria
Henrique Vicente, University of Évora, Portugal
Krzysztof Węcel, Poznań University of Economics and Business, Poland
Katarzyna Wegrzyn-Wolska, EFREI Paris | AlliansTIC & MINES ParisTech, France
Dirk Werth, August Wilhelm Scheer Institute for digitized Products and Processes, Saarbrücken, Germany
Janusz Wielki, Opole University of Technology, Poland
Hans-Friedrich Witschel, FHNW University of Applied Sciences and Arts, Northwestern Switzerland
Shaomin Wu, Kent Business School - University of Kent, Canterbury, UK
Anbang Xu, IBM Research - Almaden, USA
Sira Yongchareon, Auckland University of Technology, New Zealand
Zornitsa Yordanova, University of National and World Economy, Sofia, Bulgaria

Table of Contents

An Investigation of the Quality of Altmetric Data with Altmetric.com and PlumX as Examples

Tianhui Gong
Kent Business School,
University of Kent,
Canterbury, Kent CT2 7FS, UK
e-mail: thg5@kent.ac.uk

Wenbin Liu
Research Center for Mathematics,
Beijing Normal University,
Zhuhai, 519087, China
e-mail: wbliu@uic.edu.cn

Shaomin Wu
Kent Business School,
University of Kent,
Canterbury, Kent CT2 7FS, UK
e-mail: s.m.wu@kent.ac.uk

***Abstract*—Altmetrics, as an open-source tool to present the social impact of publications, has drawn great attention in the scientometrics and informatics fields. The quality of altmetric data is fundamental for applying and analyzing altmetrics in research evaluation, and altmetric data providers are crucial in ensuring data quality. This paper selects the two most commonly used altmetric data providers for an empirical study on analyzing data quality, coverage, consistency and heterogeneity. It proposes suggestions on the use of altmetrics data of the two platforms in the hope of providing a useful reference for other researchers.**

Keywords-altmetrics; data quality; Altmetric.com; PlumX;

I. INTRODUCTION

Modern scientific researchers are accustomed to the mainstream trend of online academic exchanges. The development of various user-oriented online social media platforms such as blogs, Twitter, Facebook, has not only attracted ever more researchers to utilize online search tools for literary works but also provided increasing opportunities to exhibit, share, comment, and discuss research outcomes online. After the term *altmetrics* was first coined in [1], it has gradually emerged as a prevalent research field for contemporary scientometrics and informatics [2] [3] [4]. Undoubtedly, altmetrics provide alternative data sources for academic evaluations. It is expected to improve the single evaluation method, which is currently widely used but merely relies on the citations in traditional measurement evaluation methods. The advent of altmetrics is also expected to compensate for the shortcomings of traditional metrological evaluation methods such as time lag.

Data quality remains the core of the analyses and applications of altmetrics. The altmetrics platforms are crucial in providing data and guaranteeing data quality. There has been a series of altmetrics platforms in use, such as Altmetric, Plum analytics (PlumX), and ImpactStory. With the gradual advancing of the research associated with altmetric data, several authors have expressed their concerns on data quality, data consistency and reliability [5] [6] [7]. Studies about altmetrics data and platforms appear relatively distributed, with the applicability of the indexes still being a matter of debate and the presence of divergent views about the data consistency and variation. Therefore, this study uses the data from Altmetric.com and Plumanalytics.com (PlumX) as examples to carry out an empirical research on their data coverage, consistency, and reliability.

The aim of the study is to provide a better understanding of the associated altmetric data and furnish valuable references for subsequent research.

The remainder of this paper is structured as follows. Section II introduces the data collection and processing methods used in this paper. Section III investigates and classifies the coverage difference of the data sources. Section IV analyses the consistency and heterogeneity of altmetric data provided by the two platforms and discusses their differences in detail. Section IV proposes suggestions on the method for selecting altmetric data for use.

II. DATA COLLECTION

Data were collected from the following platforms: Web of Science, Altmetric.com, and PlumX. The collection dates were in March 2020. This data sample contains 12,000 research papers derived from 76 random selected journals in the social science field, and all papers are published in year 2017. We first extracted the DOI of each paper from Web of Science, and then acquired associated altmetrics data via the APIs of Altmetric.com and Plumanalytics.com. To compare the data quality in a targeted manner, this study only focuses on the common data shared by both platforms. Table I summarizes the interpretation of data sources that were studied.

TABLE I. ALTMETRIC DATA AND SOURCES INVOLVED

		Interpretation of source	
No.	***Data***	***Altmetric.com***	***PlumX***
1	Blogs	Altmetric.com curates a list of blogs and collects links to relevant academic content through RSS feeds.	PlumX curates a list of blogs and collects the number of blogs mentioning the scholarly output.
2	CiteULike	The number of users of relevant scholarly outputs saved in CiteULike (not included in Altmetric Score, only in Altmetric Explorer).	The number of scholarly outputs added to the literature management tool-CiteULike.
3	Facebook	Only tracks the posts on Facebook's public pages and prioritizes popular pages.	Counts the number of times the links to scholarly outputs have been shared, commented, or liked on Facebook.

4	Mendeley	The number of users who save scholarly outputs to Mendeley Library is recorded as the number of readers. Attention Score no longer counted, only data provided.	The number of users who add scholarly outputs to Mendeley Library.
5	News	Altmetric.com curates a list of news sources and the third-party providers and RSS feeds provide data directly.	PlumX curates a list of news sources and collects the number of news mentioning the scholarly outputs.
6	Twitter	Count the public tweets, reposts or quoted tweets of related scholarly outputs on Twitter, and monitor suspicious activities.	The number of tweets and retweets mentioning scholarly outputs via Gnip.
7	Wikipedia	Collects the mentions of scholarly outputs in Wikipedia references; only for English Wikipedia.	The number of references cited in Wikipedia.

III. COVERAGE

The coverage of altmetric data determines the range of its application, accordingly we first compare the coverage of each altmetric data source of the two platforms. The results are shown in Table II.

The data source with the highest number of valid readings in Altmetric.com (hereinafter referred to as "Platform A") is Mendeley, and its coverage ratio is about 66.9% (8,031 valid readings/12,000 DOIs = 66.9%). The coverage ratios of Twitter, Blogs, CiteULike, Facebook, and News in Platform A are 64.1%, 13.6%, 4.8%, 26.1%, and 18.5%, respectively, while the coverage ratio of Wikipedia stands to be the lowest at about 2.7%. The data source with the highest number of valid readings in PlumX (hereinafter referred to as "Platform P") is Mendeley as well, with a coverage ratio of 97.6% (11,713/12,000 = 97.6%). The coverage ratios of Twitter, Blogs, CiteULike, Facebook, and News in Platform P are 54.7%, 8.3%, 4.2%, 13.2%, and 14.1%, respectively while the coverage ratio of Wikipedia stands to be the lowest at about 3.4%. Platform A appears to have higher coverage ratios in Blogs, Twitter, News, CiteUlike and Facebook, while Platform P has higher coverage ratios in Mendeley and Wikipedia.

TABLE II. ALTMETRICS DATA COVERAGE RATIO AND OVERLAPPING RATIO OF PLATFORM A AND P

	Altmetric.com		**PlumX**		**Altmetric.com - PlumX**		
	#	***%***	***#***	***%***	***Overlapping #***	***ratio in A***	***ratio in P***
Blogs	1630	13.6%	1000	8.3%	633	38.8%	63.3%
CiteULike	578	4.8%	500	4.2%	423	73.2%	84.6%
Facebook	3135	26.1%	1587	13.2%	903	28.8%	56.9%
Mendeley	8031	66.9%	11713	97.6%	7859	97.9%	67.1%
News	2214	18.5%	1693	14.1%	1524	68.8%	90.0%
Twitter	7691	64.1%	6564	54.7%	6471	84.1%	98.6%
Wikipedia	329	2.7%	405	3.4%	300	91.2%	74.1%

Furthermore, we counted the number of overlapping data (intersection) of each altmetric data source for the two platforms, and the proportion of the number of the overlapping data (intersection) in the respective data source of the two platforms. The results are also shown in Table II. Based on the proportion of overlapping data in each data source, the cross coverage of each data source in the two platforms can be derived. It is worth emphasizing, in this paper we only analyze the number of available data for each altmetric data source and the associated numerical value is not concerned. According to the results shown in Table II, we divide the status of intersection of the altmetric data sources among two platforms into three types:

(a) Dominant coverage: it means that one platform has an obvious coverage advantage than another. For example, the overlapping News data readings among two platforms is 1524, which accounts for 90% of News data coverage in Platform P's and 68.8% in Platform A's. This result implies that more News data has been collected in Platform A, and it covers 90% of Platform P's News data. Obviously, the News data source coverage of Platform A appears significantly higher than that of Platform P. For Twitter data, the overlapping Twitter data readings among two platforms is 6.471, which accounts for 84.1% of Platform A's reading and 98.6% of Platform P's. Hence, the coverage of Twitter in Platform A is greater than that of Platform P. Meanwhile, Platform P retains clear data coverage advantages in Mendeley and Wikipedia. The intersection data of Mendeley in the two platforms is 7,859, which accounts for 97.9% of Mendeley coverage in Platform A and 67.1% in Platform P. It means 97.9% of Mendeley data collected by Platform A, that has been covered by Platform P. Therefore, the Mendeley data coverage of Platform P appears to be much wider than Platform A. Similarly, the Wikipedia data coverage of Platform P is better than that of Platform A. The data sources with dominant coverage are drawn in Figure 1.

(b) Different coverage: it means the cross coverage of two platforms is limited and both platforms have their own unique data sources. For Blog and Facebook, the intersection of the data source of the two platforms does not account for a large proportion in any of the platform. The two platforms have 633 overlapping Blog data readings, accounting for 38.8% and 63.3% of data coverage in platform A and P, respectively. Although Platform A has more data than Platform P, Platform P possesses different data coverage from Platform A and its own unique data source. This means the data sources of the two platforms only represent a limited crossover range. A similar situation goes for Facebook: data source coverages of the two platforms seem dissimilar, and there exist certain overlaps and unique parts. The data sources with different coverage are drawn in Figure 2.

(c) Similar coverage: it means that the intersection of the same data source accounts for similar proportions in both of the platforms. In terms of CiteULike data, the two platforms show 423 intersection data readings, which account for about 84.6% of Platform P's CiteULike, and 73.2% of Platform A's CiteULike. Although Platform A covers more CiteULike data, the coverage of the CiteUlike data on the two platforms tends to be similar, as shown in Figure 3.

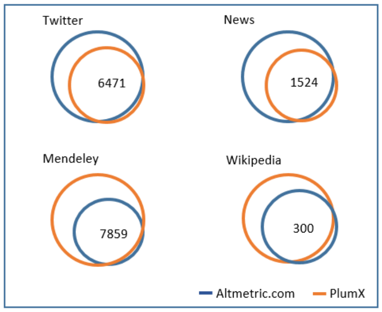

Figure 1. Schematic diagram for coverage advantages of data source in one platform

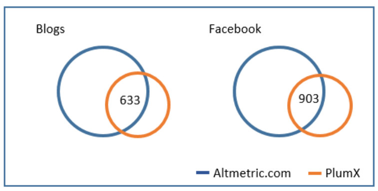

Figure 2. Schematic diagram for cross & different coverage of data source in two platforms

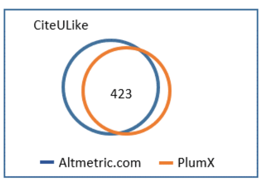

Figure 3. Schematic diagram for similar coverage of data source in two platforms

IV. CONSISTENCY AND HETEROGENEITY

In this section, a statistical method, the paired hypothesis test, is implemented to evaluate the data consistency and heterogeneity for the two platforms. Since there often exist unavoidable differences between individuals, such as experimental errors caused by the artificial factors or several uncontrollable aspects, it is essential to determine whether differences between the paired samples are merely random errors, or caused by different treatment effects such as collection methods. Prior to the hypothesis test, we tested whether the data follow a normal distribution and the result showed that all data source are not normally distributed. Therefore we employed the Wilcoxon signed rank test, a non-parametric test method for paring samples, to conduct the significance test of difference. The basic idea of this test is to assume that the two treatment effects of the pairing are the same, then the population distribution of the difference is symmetrical [11].

In the data pre-processing, we matched each altmetric data with the same article and only retained the overlapping data of the two platforms. For example, if an article has Blog data on Platform A but not on Platform P, the Blog data of this article is not included. Table III shows descriptive statistics of the overlapping data in the seven altmetric data sources studied. It can be seen that the minimum data value of CiteULike, Blogs, News, Twitter, Facebook and Wikipedia is 1 as both platforms do not present media data with a value of 0 (we assume the empty data is not available and could not be replaced by 0). While the value 0 in Mendeley comes from the Mendeley API so we retained it. The Wilcoxon Signed Rank Test was performed on the overlapping altmetric data sources of the two platforms, and the test results are shown in Table IV, in which Z stands for the value of the test statistic. In the following section, we will discuss the consistency and heterogeneity of each altmetric data source according to the statistical results, and analyze the reasons of the heterogeneity in detail.

TABLE III. STATISTICS OF OVERLAPPING DATA OF EACH DATA SOURCE

	N	*Mean*	*Std. Deviation*	*Minimum*	*Maximum*
Citeulike – Platform A	423	1.29	0.758	1	6
Citeulike – Platform P	423	1.28	0.719	1	6
Blogs - Platform A	633	3.17	3.687	1	31
Blogs - Platform P	633	2.55	3.153	1	33
News - Platform A	1524	15.93	28.443	1	276
News - Platform P	1524	8.11	21.142	1	311
Twitter - Platform A	6471	34.39	172.117	1	8128
Twitter - Platform P	6471	29.01	163.116	1	9819
Facebook - Platform A	903	7.02	17.245	1	324
Facebook - Platform P	903	328.99	2684.636	1	39422
Wiki - Platform A	300	1.46	1.359	1	14
Wiki - Platform P	300	1.68	1.682	1	15
Mendeley - Platform A	7859	70.71	110.001	0	2372
Mendeley - Platform P	7859	70.60	109.527	0	2382

TABLE IV. STATISTICS OF WILCOXON SIGNED RANKS TEST

	Citeulike	*Blog*	*News*	*Twitter*	*Facebook*	*Wikipedia*	*Mendeley*
Z	-0.039[b]	-7.158[b]	-24.801[b]	-5.803[b]	-21.550[b]	-5.575[b]	-11.313[b]
Asymp. Sig. (2-tailed)	.969	.000	.000	.000	.000	.000	.000

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.

A. *CiteULike*

As shown in Table III, the overlapping CiteULike data for the two platforms is 423 (n = 423). The average data value of Platform A is 1.29 and that of Platform P is 1.28, showing very minute difference. The standard deviations of data of the two platforms are 0.758 and 0.719, respectively, which shows that their dispersions are similar. The test result for CiteUlike in Table IV is Z = - 0.039 and the two-sided value of asymptotic significance is P = 0.969. This suggests that we fail to reject the null hypothesis and conclude that, there is no significant difference between the CiteULike data of Platform A and Platform P.

The CiteULike data of the two platforms being not significantly different indicates that the data of the two platforms comes from the same population. Although minor differences could still be noticed in the values of CiteULike on the two platforms, it could be an error caused by artificial reasons and can be considered as a random error. Based on the current test results, the CiteULike data on the two platforms appears to be consistent. We may conclude that when using CiteULike data from either of the platforms

 ISBN: 978-1-61208-955-3

under consideration, Platform A or Platform P, it does not essentially affect the statistical results. Therefore, when using CiteULike data for research and analysis, the CiteULike data of the two platforms may be used at will.

B. Blogs

There are 633 overlapping Blog data for the two platforms (see Table III), and the average data value of Platform A is 3.17 which is bigger than that of Platform P. The standard deviations of data of the two platforms are 3.687 and 3.153, respectively, and the degree of dispersion shows a minor difference. According to the results of the Wilcoxon signed rank test, $Z = -7.158$, and the two-sided value of asymptotic significance is $P = 0.000$. With the significance level $\alpha=0.05$, $P<0.05$, which implies that the null hypothesis should be rejected, and the conclusion remains that the two matching variables appear significantly different. Therefore, the differences are not random errors such as manual input errors but induced by factors such as different data sources or different ways of data collection among the two platforms.

When analyzing the reasons of difference between two platforms, firstly it is noticed from Table I that the Blog data tracked by the two platforms belongs to the blog list curated by the platforms themselves. Since neither of the platforms discloses their Blog list information, the Blog lists may transpire to be different. According to Altmetric.com, the counting method of Blogs of Platform A is to search article's citations and references (especially links) across each Blog automatically based on their curated Blog list. Moreover, Platform A only searches for links related to the research outputs through the information in the RSS feed, which indicates that it can only search for the content in the blog text, excluding the mentions in the sidebar of the blog [12]. Platform P also cites on its official website that counts the mentions of Blogs by tracking all the links related to scholarly outputs in its curated Blog list. But unfortunately, neither of the platforms has published their Blog lists. Therefore, the data difference between the two platforms may be due to different data sources.

There are not many studies discussing Blog data so far. Ortega [13] compared the Blog data of Altmetric.com, PlumX, and Crossref Event Data, and found that the overlap rate between these platforms was very low. Besides, Shema [14] examined the peer-reviewed Blog posts discussing academic papers on ResearchBlogging.org, compared the WOS citations of related papers, and found that "Blog citations" are significantly related to journal citations. The sources of blog lists on two platforms still remain unclear at the moment, with low transparency. Consequently, we believe that the current Blog data quality is questionable and should only be used with discretion. If it is unavoidable to use Blog data, it is preferable to adopt the data of both platforms A and P at the same time for complementation.

C. News

For News data, the average data value of Platform A is 15.93 and that of Platform P is 8.11. The data standard deviations of the two platforms are 28.443 and 21.142, respectively. The News data readings of the two platforms are evidently different. The Wilcoxon signed rank test result for News data suggests that the null hypothesis is rejected, and the conclusion is the two matching variables appear significantly different. Therefore, the difference between the two platforms is not a random error, that is, the two platforms may possess different data sources or different methods of data collection.

For tracking reasons behind the heterogeneity, it is noticeable that the two platforms use their own edited News lists (Table I), which may lead to varied third party data sources. We found Platform A has exposed its main third-party news sources and tracking methods on their websites [12], including identifying links to academic papers in news reports and tracking information related to academic articles in the news content, which is conducive for enhancing the transparency and reliability of the News data. While Platform P has not published a specific list of its news sources on its website yet, but it has disclosed that Newsflo is its news data provider, covering more than 55,000 diverse news sources [15]. Moreover, a Blog post on Platform A announced that the platform has been cooperating with the news service provider, Moreover Technologies, since May 2015, and has expanded its track list from approximately 1,300 news media to more than 80,000 news media. However, Ortega [16] pointed out that Moreover Technologies was later acquired by Lexis-Nexis and the cooperation with Platform A ended, leaving 19% of the links invalid, so that Platform A only manages 2,900 news media. Platform A's official website updated the information about News data sources on April 7, 2020, and stated that it still tracked more than 5,000 global mainstream news media portals [12]. As the two platforms possess diverse data sources with evident differences, it is not recommended for direct comparison. However, in this comparison study, the News data of Platform A appears to be higher than Platform P in terms of total coverage, coverage of disciplines, and value of each data. Platform A's news sources are also more appreciably transparent than that of Platform P as well. Therefore, when using the News data of the two platforms, the News data of Platform A is deemed to be more reliable.

D. Twitter

The total overlapping Twitter data of the two platforms is 6,471. The average data value of Platform A is 34.39, and that of Platform P is 29.01, seemingly less than the former. The result of the Wilcoxon signed rank test shows that the null hypothesis is rejected, and there is a significant difference in the Twitter data of the two platforms. Therefore, the difference of Twitter among two platforms cannot be considered as a random error but may have different methods of collecting or processing data.

The first reason for the difference in Twitter data could be attributed to the different measurement standards for the data by the two platforms. As the statistical methods of the two platforms in Table I, Twitter count of Platform A = the number of tweets + the number of retweets + the number of quoted tweets; while the Twitter count of Platform P = the number of tweets + the number of retweets. Therefore, statistically, Platform A additionally includes the quotation tweets. As shown in Table III, the average Twitter data value of Platform A is indeed higher than that of Platform P. Secondly, when comparing the collection of Twitter data resources by the two platforms, it was found that the data on the two platforms is not directly obtained from Twitter. Platform A's Twitter data comes from a third party that is not disclosed by the official website. Platform P's data comes from data provider GNIP, and the tweet data provided by GNIP does not contain quoted tweets [17]. Evidently, the Twitter data source of Platform A is different from that of GNIP. Therefore, we suggest that the Twitter data of both the platforms come from different third parties. The different counting methods of the two platforms may also be caused by the different measurement standards of the third-party data sources. Hence, we assume that the difference in the Twitter data of the two platforms is essentially due to the two different third-party data sources, and the measurement standards of the data sources for Twitter counting may be different.

E. Facebook

For Facebook data, the overlapping data in the sample is 903, and the average data value of Platform P is 328.99, much larger than Platform A's 7.02. The data standard deviation on the two platforms also appears quite different. According to the test result in Table IV, $Z = -21.550$, $P = 0.000 < 0.05$, the null hypothesis is rejected, so that the Facebook data from the two platforms is significantly different. Therefore, Facebook data from the two platforms may have different data sources or utilize different methods of data collection.

The first reason for the difference in Facebook data could be the fact that the two platforms possess different definitions of Facebook data measurement. Platform A only counts Facebook's public posts mentioning links to research outputs and prioritizes statistics of popular pages, excluding personal posts on the timeline and the number of likes [12]. Platform P counts the total number of shares, likes, and comments on the research link in Facebook. Furthermore, Platform P's early counting for Facebook data was to incorporate the shares and likes of links related to the scholarly outputs in public pages into the category of social media data, and links in comments content into the category of mention data. But since August 2016, Platform P added the number of comments on the personal page and integrated all of them into one Facebook data count [18]. This integrated data only records the number of comments and does not disclose the content of the comments, so it enhance the measurement of attention score of an academic link to the greatest extent and also enhance the visibility of impact measurement as well as complying with the privacy regulations. In our data sample, we can also see that the average Facebook data value of Platform P is much larger than that of Platform A. Combining the differences in the data source coverage of the two platforms in previous Section, the two platforms do demonstrate different data source coverage.

F. Wikipedia

As shown in Table III, there are 300 overlapping Wikipedia data readings for the same article on the two platforms. The average Wikipedia data value of Platform A is 1.46, and that of Platform P is 1.68, appearing only minutely different. The data standard deviations of the two platforms are 1.359 and 1.682, respectively. The result of the Wilcoxon signed rank test is $Z = -5.575$, $P = 0.000 < 0.05$, so the null hypothesis is rejected and the conclusion remains that the two matching variables are significantly different.

Since the Wikipedia data from the two platforms is significantly different, we check the measurement method of two platforms first. As shown in Table I, Platform A only counts the references to a certain scholarly output in the reference section of English Wikipedia, while Platform P counts all scholarly outputs cited as references in Wikipedia, that is, the latter possesses a wider measurement range. Ortega [7] considers the differences of Wikipedia data between different platforms to be systemic errors caused by coverage issues. A survey by [10] revealed that Platform A obtains its data via Wikipedia API, while Platform P obtains its data by combining different search engine results, including Wikipedia full-text search and literature citation searches. Therefore, the Wikipedia data value of Platform P is more than that of Platform A. The findings in [10] can be further elucidated the differences between the data sources of the two platforms. From the data available in this study, the average Wikipedia data value of Platform P indeed appears slightly higher than that of Platform A.

G. Mendeley Reader Count

The overlapping Mendeley data is 7,859 for the same article among the two platforms. The average Mendeley data value of Platform A is 70.71 and that of Platform P is 70.60. The data standard deviations of the two platforms are 110.001 and 109.527, respectively. From the view of the descriptive statistics, there is not much of a difference between the two platforms. The result of the Wilcoxon signed rank test is $Z = -11.313$, $P<0.05$, so the null hypothesis is rejected, and the conclusion remains that the two matching variables are significantly different. There is no evident difference in the description of the measurement method between the two platforms in Table I, and both platforms obtain the data directly from the Mendeley API. But the question remains as to why there is a difference. Ortega [7] assumed that when Platform P counts the number of Mendeley readers, it may duplicate records of documents of similar titles, years, and authors; while Platform A counts the number of readers based on a unique identifier and eliminates duplicate entries. Zahedi [10] compared the data of journal PLOS One on different altmetrics platforms and found that 97.9% of the Mendeley counts recorded by Platform A were the same as Mendeley.com, while Platform P had only 30% of the same count. In other words, the two platforms may perform differently when processing Mendeley data.

V. Conclusion and Recommendation

This paper compared and analyzed the data quality of two altmetrics platforms: Altmetric.com and PlumX, and undertook detailed statistical analysis in terms of the data coverage, data consistency, and reliability of the data sources. Firstly, in terms of the coverage ratio of altmetrics data in the two platforms, the coverage ratio of Almetric.com in social media data including Twitter, Facebook, Blogs, and News appears higher than that of PlumX. The data coverage ratio of CiteULike in Almetric.com also appears slightly higher than that of in PlumX, while PlumX shows a higher Mendeley and Wikipedia data coverage ratio than that of Altmetric.com. By evaluating the proportions of the overlapping data in the two platforms, we found: (a) *Dominant coverage:* the data coverage of one platform appears evidently larger than that of the other platform, such as Twitter, News, Mendeley, and Wikipedia in Figure 1; (b) *Different coverage:* the data coverage of the two platforms limitedly overlaps but each platform has its own unique data source, such as Blogs and Facebook in Figure 2; and (c) *Similar coverage*: the data coverage of the two platforms may remain similar, such as CiteULike in Figure 3.

To discuss whether the overlapping data of the two platforms appears to be consistent, we employed the non-parametric Wilcoxon Signed Rank test to compare if any significant differences exist between the data of the two platforms. The results signify no significant difference between the CiteULike data of the two platforms, while all the other altmetrics data show significant differences between the two platforms. It indicates that the CiteULike data of the two platforms belongs to the same population, and the difference of a small amount of data value is caused by random errors. The differences between other altmetric data source could possibly be due to the factors including different ways of data counting (definitions), data processing, and different coverage of data source. Section IV showed more detailed analysis of reasons on heterogeneity between two platforms. Based on the above analysis results, we propose the following suggestions for the use of Altmetrics data of the two platforms.

- Twitter and News on Altmetric.com are significantly better than that on PlumX in terms of data coverage ratio, coverage range, and information transparency of data sources, respectively. Therefore, it is recommended to use Twitter and News data from the Altmetric.com.
- As for the Wikipedia data, the data coverage ratio and data source coverage of Plum X are far better than those of Altmetric.com. Hence, it is advised to use PlumX's Wikipedia data.
- In terms of Mendeley data of the two platforms, PlumX is appreciably better than Altmetric.com, and the Mendeley data of PlumX is recommended. However, the Mendeley data of the two platforms appears statistically significantly different from the number of readers of Platform Mendeley. Therefore, whenever conditions are permitted, we recommend employing the Mendeley data from its own application.
- There is no substantial difference being noted between the CiteULike data of the two platforms in this statistical test and the correlation seems strong, so either platform's CiteULike data can be used.
- The Blog data sources of the two platforms have a limited crossover range. Both platforms possess their own unique data sources. Therefore, the Blog data of the two platforms can be used in combined to complement each other.
- The Facebook data sources of the two platforms also maintain a limited crossover range and each one has its own characteristics; but the two platforms demonstrate very different definitions of Facebook counts. Altmetric.com only counts the links mentioning scholarly outputs in public posts, while PlumX counts the total shares + likes + comments of links mentioning scholarly outputs. Therefore, it is not convenient and accurate to merge the data of the two platforms, and they can only be selected with discretion.

Finally, we should address that the current results and analysis are limited to the specific situation of this paper's data sample so further research on different samples may be needed in the future.

References

[1] J. Priem, D. Taraborelli, P. Groth, and C. Neylon, "Altmetrics: A manifesto," [Online]. Available from: http://altmetrics.org/manifesto/ [Retrieved: Feb, 2022]

[2] C. L. González-Valiente, J. Pacheco-Mendoza, and R. Arencibia-Jorge, "A review of altmetrics as an emerging discipline for research evaluation," Learned Publishing, vol. 29, no. 4, pp. 229–238, 2016.

[3] P. Chellappandi and C. S. Vijayakumar, "Bibliometrics, Scientometrics, Webometrics/Cybermetrics, Informetrics and Altmetrics-An Emerging Field in Library and Information Science Research," Shanlax International Journal of Education, vol. 7, no. 1, pp. 5-8, 2018.

[4] X. Zhang, X. Wang, H. Zhao, P. Ordóñez de Pablos, Y. Sun, and H. Xiong, "An effectiveness analysis of altmetrics indices for different levels of artificial intelligence publications," Scientometrics, vol. 119, no. 3, pp. 1311–1344, 2019.

[5] A. Jobmann, C. P. Hoffmann, S. Künne, I. Peters, J. Schmitz, and G. Wollnik-Korn, "Altmetrics for large, multidisciplinary research groups: Comparison of current tools," Bibliometrie-praxis und forschung, vol. 3, 2014.

[6] Z. Zahedi, R. Costas, and P. Wouters, "How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications," Scientometrics, vol. 101, no. 2, pp. 1491-1513, 2014.

[7] J. L. Ortega, "Reliability and accuracy of altmetric providers: a comparison among Altmetric. com, PlumX and Crossref Event Data," Scientometrics, vol. 116, no. 3, pp. 2123-2138, 2018.

[8] H. Alhoori and R. Furuta, "Do altmetrics follow the crowd or does the crowd follow altmetrics?," IEEE/ACM Joint Conference on Digital Libraries, IEEE, 2014.

[9] C. Meschede and T. Siebenlist, "Cross-metric compatability and inconsistencies of altmetrics", Scientometrics, vol. 115, no. 1, pp. 283-297, 2018.

[10] Z. Zahedi and R. Costas, "General discussion of data quality challenges in social media metrics: Extensive comparison of

four major altmetric data aggregators,” PLoS One, vol. 13, no. 5, 2018.

[11] R. C. Blair and J. J. Higgins, “Comparison of the power of the paired samples t test to that of Wilcoxon's signed-ranks test under various population shapes,” Psychological Bulletin, vol. 97, no.1, pp. 119–128, 1985.

[12] Altmetric.com. *Solutions.* [Online]. Available from: https://help.altmetric.com/support/solutions/ [Retrieved: Feb, 2022]

[13] J. L. Ortega, “The coverage of blogs and news in the three major altmetric data providers,” The 17th International conference of the international society for scientometrics and informetrics (ISSI 2019), Sep. 2019, pp. 75-86.

[14] H. Shema, J. Bar-Ilan, and M. Thelwall, "Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics," Journal of the Association for information science and technology, vol. 65, no. 5, pp. 1018–1027, 2014.

[15] P. Allen, “World-wide news coverage in PlumX,” [Online]. Available from: https://plumanalytics.com/world-wide-news-coverage-plumx/ [Retrieved: Feb, 2022]

[16] J. L. Ortega, “Blogs and news sources coverage in altmetrics data providers: a comparative analysis by country, language, and subject,” Scientometrics, vol. 122, no. 1, pp. 555-572, 2020.

[17] A. Michalek, “NISO Altmetrics Working Group on Data Quality,” [Online]. Available from: https://support.gnip.com/sources/twitter/overview.html [Retrieved: Feb, 2022]

[18] P. Allen, “PlumX’s Facebook Altmetrics – Measure Up,” [Online]. Available from: https://plumanalytics.com/plumx-facebook-altmetrics-measure-up/ [Retrieved: Feb, 2022]