









An Altmetric Alternative for Measuring the Impact of University Institutional Repositories' Grey Literature

Miguel Valles^(✉) , Richard Injante , Edwin Hernández , Juan Riascos ,
Marco Galvez , and Juan Velasco 

Universidad Nacional de San Martín, Tarapoto, Peru
mavalles@unsm.edu.pe

Abstract. To get the license, universities must meet a set of standards to guarantee the education of their students. This involves installation and maintenance of an institutional repository to publish their grey literature. This research aims to create an alternative altmetric tool to measure the impact of such literature. To get it done, we worked the institutional repository customization and optimization, we created procedures to retrieve records, transform and load altmetric data in a new database and we built a dashboard-type tool using the altmetric data of the institutional repository grey literature. Now, there is an alternative to measure the altmetric indices from different dimensions and analysis perspectives and it makes possible to take the necessary actions to apply strategies that aim to increase the repository visibility. The tool implementation in the study case allowed the monitoring and control of the institutional repository altmetric indices.

Keywords: Altmetric · Institutional repositories · Grey literature · Bibliometric · Dashboard

1 Introduction

Nowadays, Latin America and world universities are required a set of demands to fulfill the basic conditions of educational quality that has forced them to guarantee better conditions for a comprehensive education of professionals graduating from their schools. The trends is to look for a less heterogeneous, less unequal university scenario towards the future, but above all it is sought that they assume the role that allows building a society with better living conditions for its population, a democracy and economies strengthened as a result of increase science, education and culture [1]. Peru is no exception. In that sense, the new university law No. 30220 [2] has had a special impact on process of professionals education by demanding that universities comply with 55 minimum standards, which form a protection mechanism for students, their families and society as a whole [3, 4].

Part of the requirements implies research results publication in open access through institutional repositories to be used free of charge as a background for research or to

strengthen the bibliographic review of a topic to be investigated. Once the research results are available, the design of strategies that allow to increase its visibility and impact factor remains, incorporating as part of repository functions the sharing of resources in social networks as a marketing strategy to achieve a greater audience than the academic environment itself.

In that sense, Peruvian government has been making efforts along with universities to give licensing, demanding updated and accessible repositories [4], so that the revision of their content serves to formulate research projects leading to academic and/or professional degrees. This is done through the National Register of Research Works (RENATI from Spanish initials), the Free Access to Scientific Information for Innovation (ALICIA from Spanish initials) project, with Law No. 30035 endorsement that regulates the Open Access National Repository of Science, Technology and Innovation [5] and its corresponding regulations [6].

Therefore, Peruvian universities have been working to setup the infrastructure, specifications, as well as on the identification of requirements and processes training for the human resource responsible to guarantee that thesis and every material published in their repositories are available. Peruvian universities use DSPACE to support these processes. The Standard 37, Condition V of the Licensing Model, requires the institutional repository maintenance where research documents like thesis, research reports and scientific publications, among others should be registered and stored. Ferreras-Fernandez et al. [7] state that these open access repositories are valid channels for grey literature publication. Grey literature are publications that do not go through a strict quality validation process such as peer revision or have a limited circulation. This literature is usually recommended as bibliographic revision source in research documents, but recommendation is to avoid it, nevertheless, for A. Paez [8] grey literature can make great contributions to a systematic revision.

Measurement of repository performance indices and measurement of quality and quantity bibliometric indices of scientific production used to take actions to increase these indicators are complicated to monitor without a proper knowledge of techniques and tools available for this purpose. This also involves design of strategies that guarantee the use of the available material, visibility increase and impact factor of publications using social networks as a marketing and advertising strategy. Thelwall et al. [9] state that references in social networks have become a valuable marketing tool for publishers who try to advertise their articles. For S. Santana [10] current research processes are increasingly collaborative and communication between academics and researchers is developed through social interchange spaces, and this accelerates the diffusion of scientific knowledge, so the job of diffusion in social networks is necessary. In addition, according to the altmetric manifesto [11] since altmetrics are in themselves diverse, they are excellent at measuring the impact of the diverse academic ecosystem.

According to Priem et al. [12], altmetrics is the study and use of academic impact measures based on activity in online tools and environments. The term describes the metrics themselves—one could propose in plural a “set of new altmetrics.” Altmetrics is in most cases a subset of both scientometrics and webometrics; it is a subset of the latter in that it focuses more narrowly on scholarly influence as measured in online tools

and environments, rather than on the Web more generally. The sources used for altmetrics are heterogeneous and include mentions and citations in blogs, Wikipedia, Twitter, Facebook, Mendeley or reader counts on social reference managers, bookmarking platforms and bibliographic references software. In that sense I. Aguillo [13] mentions that altmetrics is being explored as a potential tool for measuring research impact beyond the scientific communities, the so-called societal impact.

Why use altmetrics then? According to L. Palmer [14], because it provides new services and demonstrates content evaluation and assessment, integrates social network data, shows social importance of our content, adds context and meaning to resources downloads, plus the badges and API are free. L. Palmer [14] also mentions the concern about the future costs of aggregate exploitation services of the altmetric data and its complexity of use, justifying once more the need to create a dashboard based monitoring tool that makes use of the Altmetrics API functions.

The possibility for using social networks to divulge the scientific production results, specifically for repository's grey literature, has already been implemented in Peru (as shown in L. Elespuru & L. Huaroto [15] study). Thanks to configuring features for resource sharing on social networks and incorporating free altmetrics donuts provided by altmetrics.com [16] they have increased their visits, reading, quotations and impact factor according to E. Adie & W. Roe [17]. It has also allowed them complementing existing usage statistics to help plan repository development and the allocation of resources to improve marketing strategies for its content [18].

This implies the need to measure, control and improve the altmetric and bibliometric indices of institutional repository's grey literature. Many repository managers create in-house-generated spreadsheets and monthly statistics information that cannot be tracked easily or efficiently by repository software [19]. A way to achieve this is through subscription to services of existent alternative metrics, for example: Altmetrics, PlumX, Sciencecard, SciVerse Scopus API and others [9, 19, 20]. However, due to its high costs, complexity of initial setup work and data loading, it has become a distant possibility for universities. They want not just to integrate use of social networks as a marketing strategy to advertise their resources but they want use the information of relevant discussions of their content, analyze and compare the scope that dependencies have achieved within the organization or know the audience demographic data that only paid subscription services provide.

Another way, without paid subscription, is using an application-programming interface (API) which includes routines, data structures, objects classes, variables and response codes that subscription services provide. It is free, it simplifies the process and allow access to Altmetrics.com database for free [9] to those who wish to understand the patterns of academic communication in the web [21], but their use conditions force attribution to source from which data has been mined.

During literature review, we found proposals such as Zervas et All's [22] who converted the Cyprus University of Technology's repository into a Current Research Information System (CRIS). They mention that altmetrics is already implemented in the CRIS and their most important result is that now they have an institutional repository that can collect, manage, preserve and disseminate all information about their university's research and its performance. Also Konkiel & Scherer's [18] who contextualized

the importance of use Altmetrics for measurement of valuable indicators of interest to complement statistics and traditional use, however they also mentioned the services cost (as equal as L. Palmer [14]) to obtain aggregate indicators of Altmetric information that generates the use of its resources in social networks. And finally I. Aguillo's [13] who mentions that repositories play a key role in success of open access publishing. However, he indicates that making them available in this format is not enough to achieve visibility and impact, so it is important to use both public and academic social networks to achieve larger audiences and increase not only academic but also social impact.

However, we do not identified proposals to measure grey literature impact of university institutional repositories in social networks on the way that we can analyze and compare metrics graphically grouped by faculty, dependency, social network, demographics and any other indicator that has been achieved internally in the university without having to pay for the service. Therefore, this work proposes an alternative to integrate altmetric functions within repository and building a tool to monitor the altmetric indices using the API for developers provided by altmetric.com.

As a study case, we implemented the proposal with Universidad Nacional de San Martin repository, achieving improvements on the impact and visibility of published research.

2 Materials and Methods

This work provides a dashboard based monitoring tool that institutional repository managers can use to monitor mentions in social networks of available resources. For this, we customized the institutional repository website optimizing it and incorporating it altmetric functions; then we mined the repository records that serve as income to perform requests to the altmetrics API, resulting in a dataset from those that have been shared. Finally, we built the dashboard interface.

We used data and website of Universidad Nacional de San Martin – Tarapoto institutional repository, where we incorporated function to share the available resources (mainly grey literature) in social networks and we deploy the dashboard based monitoring tool in the following URL: <http://altmetrics.unsm.edu.pe>. Up to September 2019, there were 2870 bibliographic resources available. Below, we described and outlined the stages that we have followed to achieve the objectives of the work.

The steps necessary to build, configure and use the monitoring tool are parameterizable and standard, thus ensuring that any institution that uses DSPACE as its repository manager can use it.

2.1 Customizing the Institutional Repository

To share the resources available at repository and to be able to track them by alternative metrics services, it is necessary to perform some setups and including new functions that we detailed:

1. Institutional repository interfaces redesign to improve its downloading performance and usability.

2. Creating and setting up unique identification service for handle resources [18].
3. Incorporating complements to share resources in social networks using the thisadd.com service.
4. Integrating alternative metrics of altmetrics.com in the institutional repository.
5. Incorporating Google Analytics to monitor visits traffic to the institutional repository website.
6. Implementing optimization techniques to improve performance and positioning of repository website in search engines using Google Webmasters.
7. Creating and setting up services to update resources in Google Scholar.

2.2 Automated Mining, Transformation and Data Upload Procedures

This work is based on J. Ortega [20], Robinson-Garcia et al [23] and E. Adie & W. Roe [17]’s data mining model, who mine an initial set of articles from public pages that then were searched in altmetric.com, using its API [24] as data provider interface for scientometric developers and researchers [21, 25]. Currently, the version available according to Altmetric.com [24] is version 1.

We designed a database model in PostgreSQL that guarantees the structured storage of data that where pivoted and integrated from the DSPACE data model with the organizational structure created, as well as the results of Altmetrics API (from where it was possible to obtain records of the resources that have been shared in social networks). We show it below (Fig. 1).

We designed a standard, automated and parameterizable flow processes as follows (Fig. 2):

A query is created to mine every register of the repository’s bibliographic resources which has a unique identification valid code; the university uses handle [26].

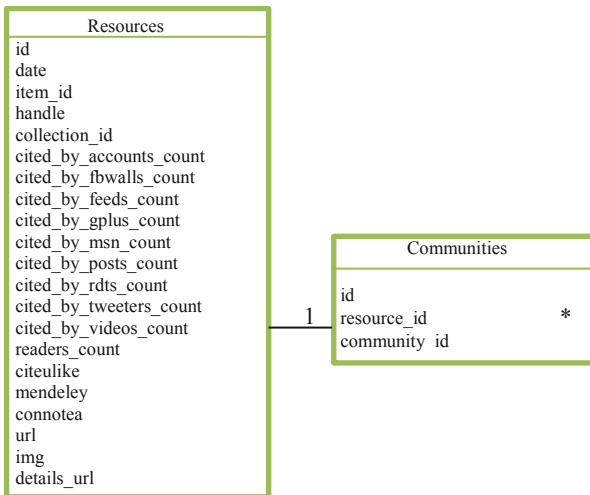


Fig. 1. Database of the tool

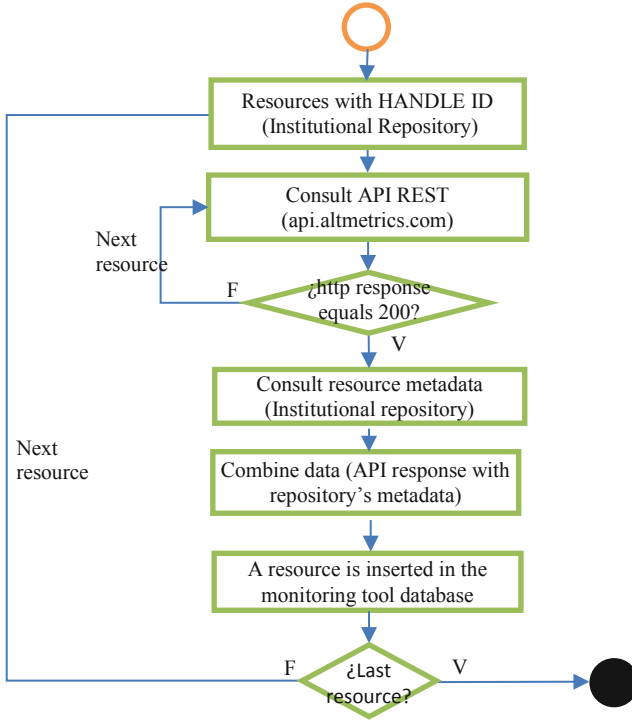


Fig. 2. Flow of standard and parameterizable processes

With this data an array is uploaded to perform timed requests to the Altmetrics API REST, then using the http methods available from the API REST a response is expected, which has one of the following codes [24] (Table 1):

Table 1. Requests response code.

Status code	Description
200	Success. The body of the response should contain the data you requested
403	You are not authorized for this call. Some calls and query types can only be made by holders of an API key and/or a license for the Full Access version of the API
404	Altmetric does not have any details for the article or set of articles you requested
429	You are being rate limited. If you have not already then apply for an API key
502	The Altmetric Details Page API version you are using is currently down for maintenance

If we found the resource in the altmetrics database, the API retrieves a metadata set in JSON format; then we combined it with the data obtained from a relational query

to the institutional repository database to obtain detailed information and the resource hierarchical structure.

This combined data is stored in the database of dashboard based monitoring tool that we created.

2.3 Building the Monitoring Tool and Mining for Altmetric Data

Components of Monitoring Tool

The tool has three components: Data access, business logic and presentation.

Data access component contains normalized data of repository database and the ones mined from the API REST. Business logic component of the tool is an algorithm that implements flow of standard and parameterizable processes mentioned in point 2.3. For presentation component, we used the code editor Visual Studio Code that is compatible with the technologies: PHP, CSS, HTML5, Bootstrap, AngularJS and Laravel applying the Model View Controller (MVC) software design pattern to integrate all these technologies. We designed the interface very similar to <https://www.altmetric.com/explore>.

3 Results and Discussion

3.1 Customization of the Institutional Repository

First, the project requires repository customization; it includes user interfaces redesign, analytical functions incorporation with Google Analytics and repository website optimization with Google Webmasters. Now, repositories managers have a set of tools to monitor analytical data where there are different segments, metrics and analysis dimensions available [27].

The following graph shows analytical indices evolution of institutional repository website:

After customization of institutional repository website, we can see in Fig. 3 that visits increased by an average of 65% compared to previous period. With a minimum increase of 16% and a maximum of 156%. This indicates that interface redesign, social networks inclusion and site optimization suggested by Google Webmaster improved the repository visibility.

To determine statistically that the interface redesign, social networks inclusion and site optimization suggested by Google Webmaster improve the visibility of the repository we have:

Null hypothesis. $H_0: \mu_x \leq 0$ (If initial mean difference of repository interface redesign, social networks inclusion and site optimization suggested by Google Webmaster is less than or equal to zero, then the repository visibility is not improved).

Alternate hypothesis $H_1: \mu_x > 0$ (If final mean difference of repository interface redesign, social networks inclusion and site optimization suggested by Google Webmaster is greater than zero, and then the repository visibility improves) (Table 2).

When contrasting the hypothesis at 5% significance level, we have 138.5 as difference result, with 72.931 as standard deviation and with 95% confidence interval that the

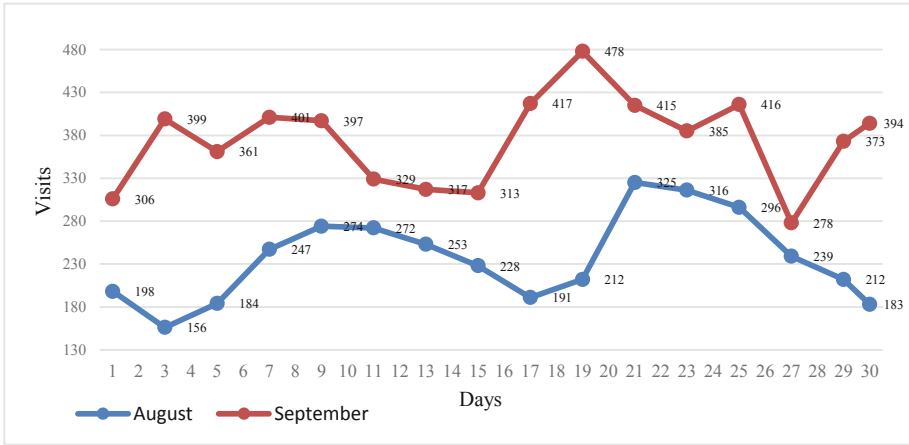


Fig. 3. Evolution of institutional repository website visits. August – September 2019 period.

Table 2. Paired samples test (SPSS.v.24)

	Paired differences					t	Df	Sig.
	Mean	Standard deviation	Mean standard error	95% TI				
				Lower	Higher			
Con_Interfaz-Convencional	138,5	72,931	18,233	99,638	177,362	7,596	15	<i>p</i> = 0,000

average difference is between [99.638; 177.632]. We conclude then that difference is significant with $p = 0.000$; that is to say the interface redesign, social networks inclusion and optimization of the site suggested by Google Webmaster significantly improves repository visibility.

3.2 Building of Automated Mining, Transformation and Uploading of Data Procedures

To use dashboard based monitoring tool, the system automatically perform recuperation process weekly on Sundays at 00:00. However, there is also the possibility to performing the procedure on demand. We can see the retrieved data in the following Table 3:

Table 3. Summary of retrieved data from altmetrics.

Date	Records in Dspace	Shared records	Times shared	Facebook	Twitter
08/12/2019	2,750	45	57	28	29
08/31/2019	2,767	51	64	34	30
09/17/2019	2,790	51	65	34	31
09/23/2019	2,801	51	65	34	31
10/05/2019	2,840	52	66	35	31
10/07/2019	2,857	57	71	40	31

3.3 Construction of the Altmetric Data Monitoring and Exploitation Tool

With this data, we obtained different kinds of queries in a graphic way. The most important ones are the following (Tables 4, 5, 6 and 7):

Table 4. Mentions in social networks summarized by university

Description	Mentions
At university level	71

Table 5. Mentions in social networks summarized by type of research

Description	Mentions	Percentage
Research and development	7	9.9%
Literary production	1	1.4%
Thesis	63	88.7%
Total	71	100.0%

Table 6. Mentions in social networks summarized by type of thesis

Description	Mentions	Percentage
Doctorate	2	3.2%
Master	2	3.2%
Undergraduate	59	93.7%
Total	63	63.0%

Finally, the monitoring tool interface, when used from a mobile device (Fig. 4).

Table 7. Mentions in social networks of undergraduate theses summarized by faculty

Description	Mentions	Percentage
Faculty of Agricultural Sciences	14	23,7%
Faculty of Health Sciences	2	3,4%
Faculty of Economic Sciences	6	10,2%
Faculty of Law and Political Sciences	0	0,0%
Faculty of Ecology	2	3,4%
Faculty of Education and Humanities	3	5,1%
Faculty of Agroindustrial Engineering	2	3,4%
Faculty of Civil Engineering and Architecture	6	10,2%
Faculty of Systems Engineering and IT	23	39,0%
Faculty of Human Medicine	1	1,7%
Total	59	100,0%



Fig. 4. Interface of monitoring tool “Observatory of the social networks”

4 Conclusions

We built a dashboard based monitoring tool where you may observe different indices of alternative metrics graphically, thanks to the mentions that repository resources get in social networks because marketing strategies, advertising and new academic uses given by the researchers.

We customized and optimized institutional repository website incorporating altmetric and analytical functions, allowing higher visibility increasing visits by 58% obtaining higher mentions in social networks, mainly Facebook and Twitter.

The tool implementation of study case allowed us monitoring and control institutional repository altmetric indices.

An important aspect of this alternative is the low cost for implementation and customization compared to subscription to existent alternative metrics services that has high costs and some complexity degree for customization.

Thanks to features, functional and non-functional requirements implemented, as well as the standardized flow of proposed processes, this tool may be implemented at any university institutional repository and thus obtain altmetric measurements of its grey literature.

Acknowledgements. To Universidad Nacional de San Martín – Tarapoto – Peru, for financing the research project called “Observatory of social networks and alternative metrics and their relationship with different bibliometric indices of institutional repository of UNSM-T” from which the results of this article arise.

References

1. Gazzola, A., Didriksson, A. (eds.): Trends in Higher Education in Latin America and the Caribbean. UNESCO. IESALC-UNE, Caracas (2008). 16–17 p.
2. MINEDU. Ley Universitaria 30220, p. 35. MINEDU, Lima (2015). http://www.minedu.gob.pe/reforma-universitaria/pdf/ley_universitaria.pdf. Accessed 27 Mar 2019
3. Sunedu. Condiciones básicas de calidad (2019). <https://www.sunedu.gob.pe/8-condiciones-basicas-de-calidad/>. Accessed 4 Oct 2019
4. Sunedu. El Modelo de Licenciamiento y su Implementación en el Sistema Universitario Peruano. Primera. Minedu, editor. MINISTERIO DE EDUCACION. SUNEDU, Lima, p. 37 (2015). <http://repositorio.minedu.gob.pe/handle/123456789/4565>
5. Ley N° 30035. Ley que regula el repositorio nacional digital de ciencia, tecnología e innovación de acceso abierto, p. 2. Concytec, Lima (2013). <https://portal.concytec.gob.pe/images/stories/images2013/portal/areas-institucion/dsic/ley-30035.pdf>. Accessed 30 Mar 2019
6. Decreto Supremo N° 006-2015-PCM. Reglamento de la Ley N° 30035, Ley que regula el Repositorio Nacional de Ciencia, Tecnología e Innovación de Acceso Abierto, p. 4. PCM, Lima (2015). http://portal.concytec.gob.pe/images/stories/images2013/portal/areas-institucion/dsic/reglamento_repositorio_nacional_alicia.pdf. Accessed 13 Apr 2019
7. Ferreras-Fernández, T., García-Peñalvo, F.J., Merlo-Vega, A.: Open access repositories as channel of publication scientific grey literature. In: ACM International Conference Proceeding Series, pp. 419–26 (2015)
8. Paez, A.: Gray literature: an important resource in systematic reviews. *J. Evid.-Based Med.* **10**(3), 233–40 (2017) <http://doi.wiley.com/10.1111/jebm.12266>. Accessed 4 Aug 2019

9. Thelwall, M., Haustein, S., Larivière, V., Sugimoto, C.R.: Do altmetrics work? Twitter and ten other social web services. *PLoS One* **8**(5), e64841 (2013). <https://dx.plos.org/10.1371/journal.pone.0064841>. Accessed 5 Oct 2019. Bornmann, L. (ed.)
10. Santana, S.: Las métricas alternativas y sus potencialidades para el profesional de la salud. *Rev Médica Clínica Las Condes* **29**(4), 484–490 (2018). <https://linkinghub.elsevier.com/retrieve/pii/S0716864018300804>. Accessed 15 May 2019
11. Priem, J., Taraborelli, D., Groth, P., Neylon, C.: Altmetrics: a manifesto (2010). <http://altmetrics.org/manifesto/>
12. Priem, J., Groth, P., Taraborelli, D.: The altmetrics collection. *PLoS One* **7**(11), e48753 (2012). <https://dx.plos.org/10.1371/journal.pone.0048753>. Accessed 13 Jan 2020. Ouzounis, C.A. (ed.)
13. Aguillo, I.F.: Altmetrics of the open access institutional repositories: a webometrics approach. In: 23rd International Conference on Science and Technology Indicators (STI 2018). Centre for Science and Technology Studies (CWTS) (2018)
14. Palmer, L.: Altmetrics and Institutional Repositories: A Health Sciences Library Experiment. Library Publications and Presentations (2013). https://escholarship.umassmed.edu/lib_articles/142. Accessed 13 Jan 2020
15. Eléspuru, L., Huaroto, L.: Los repositorios institucionales como herramientas para medir los indicadores Altmetrics: experiencia de la Universidad Peruana de Ciencias Aplicadas (UPC). In: V Congreso Internacional de Bibliotecas Universitarias 2016 (CIBU), Santiago de Chile, p. 20 (2016). http://repositorio.pucp.edu.pe/index/bitstream/handle/123456789/52633/repositorios_institucionales_altmetric_elsespuru_huaroto.pdf?sequence=1&isAlloWed=y. Accessed 13 Apr 2019
16. Altmetrics.com. Altmetric badges – Altmetric (2019). <https://www.altmetric.com/products/altmetric-badges/>. Accessed 14 Jan 2020
17. Adie, E., Roe, W.: Altmetric: enriching scholarly content with article-level discussion and metrics. *Learn Publ.* **26**(1), 11–17 (2013)
18. Konkiel, S., Scherer, D.: New opportunities for repositories in the age of altmetrics. *Bull. Am. Soc. Inf. Sci. Technol.* **39**(4), 22–26 (2013). <http://doi.wiley.com/10.1002/bult.2013.1720390408>. Accessed 12 May 2019
19. Callicott, B.B., Scherer, D., Wesolek, A.: Making Institutional Repositories Work, 1st edn, pp. 12–26. Purdue University Press, West Lafayette (2016)
20. Ortega, J.L.: Reliability and accuracy of altmetric providers: a comparison among Altmetric.com, PlumX and Crossref Event Data. *Scientometrics* **116**(3), 2123–2138 (2018). <https://doi.org/10.1007/s11192-018-2838-z>. Accessed 16 May 2019
21. Altmetrics. Researcher Data Access Program, p. 1 (2019). <https://www.altmetric.com/research-access/>
22. Zervas, M., Kounoudes, A., Artemi, P., Giannoulakis, S.: Next generation institutional repositories: the case of the CUT institutional repository KTISIS. *Procedia Comput. Sci.* **146**, 84–93 (2019). <https://linkinghub.elsevier.com/retrieve/pii/S1877050919300882>. Accessed 13 Jan 2020
23. Robinson-García, N., Torres-Salinas, D., Zahedi, Z., Costas, R.: New data, new possibilities: exploring the insides of altmetric.com. *El Prof. la Inf.* **23**(4), 359–366. <https://recyt.fecyt.es/index.php/EPI/article/view/epi.2014.jul.03/16959>. Accessed 5 Oct 2019
24. Altmetric.com. Altmetric API documentation (2019). <http://api.altmetric.com/>
25. Mingers, J., Leydesdorff, L.: A review of theory and practice in scientometrics. *Eur. J. Oper. Res.* **246**(1), 1–19 (2015). <https://doi.org/10.1016/j.ejor.2015.04.002>

26. Manhique, I., Rodrigues, F., Sant'Ana, R., Casarin, H.: Indicadores altmétricos em periódicos brasileiros da Ciência da Informação. *Rev Ibero-Americana Ciência da Informação* **12**(2), 515–532 (2019). <http://periodicos.unb.br/index.php/RICI/article/view/9156>. Accessed 23 July 2019
27. Gadd, E., Rowlands, I.: How can bibliometric and altmetric suppliers improve? Messages from the end-user community. *Insights UKSG J.*, 31 (2018). <http://insights.uksg.org/articles/10.1629/uksg.437/>. Accessed 23 July 2019