

Journal of Contemporary European Research

Volume 12, Issue 3 (2016)

Commentary

Pirating European Studies

Natalia Timuş, *Maastricht University*

Zakaria Babutsidze, *SKEMA Business School*

Citation

Timus, N. and Babutsidze, Z. (2016). 'Pirating European Studies', *Journal of Contemporary European Research*. 12 (3), 783 - 791.

First published at: www.jcer.net

Abstract

Open Science has gained a lot of attention not only within the academic community but also among policy-makers. Some international publishers have been active in moving towards open access publications and research data, but, overall, modest results have been achieved so far. In this context, the digital piracy engines emerge as vital actors in disseminating and determining the impact of research. This study examines the Sci-Hub downloads data in order to uncover patterns of piracy in European Studies research. We identify journals and the subjects of articles that have been pirated the most. We also study the geographical distribution of download requests. The analysis reveals that the readers are mostly interested in subjects reflecting the current major European challenges, specifically populism and the economic crisis. Both developing countries as well as the 'old' EU members are active in illegal downloads.

Keywords

Open science; Open access; European Studies; Downloads; Piracy

The days of keeping our research results to ourselves are over. There is far more to gain from sharing data and letting others access and analyse that data (European Commissioner Carlos Moedas, 4 December 2015).

Open Science, or 'Science 2.0', is dominating contemporary academic and societal debate. The speed and scale of digitalisation are transforming the way we produce, circulate and consume knowledge. We are witness to an ongoing evolution in 'the modus operandi of doing research and organising science' (European Commission 2015: 1), driven by digital technologies and Big Data, in the context of scientific globalisation and growing public demand to tackle societal challenges. The underlying idea behind the open science concept is that scientific knowledge 'should be openly shared as early as is practical in the discovery process' (Nielsen 2012). Thus, it entails multiple ongoing transitions of the research cycle and scientific practice.

Although open science is a relatively new concept, it has gained a lot of attention not only within the academic community, but also among policymakers. The European Commission and specifically its Directorate General for Research and Innovation has made open science a top priority on their agenda and has made a significant breakthrough in this field. The Commission concern with open science is natural given that the European Union is one of the world's leading producers of scientific knowledge.

One of the Commission initiatives is the Open Research Data (ORD) pilot of Horizon 2020, aimed at improving and maximising access to and re-use of the data of Horizon 2020 research projects (European Commission 2016a). The Commission has brought together key conceptual insights in its book about the three Os: *Open Innovation, Open Science, Open to the World* (European Commission 2016b). In May 2016, the Competitiveness Council of the EU set the year 2020 as the target for all publicly funded research published in Europe to be open access (Enserink 2016).

The debate around knowledge production has crystallised into two trends: open access and open source. The former refers to online, free access, peer-reviewed publications, with limited or no restrictions on copyright and licensing. Open source is about co-created free software without proprietary restriction (European Commission 2016b).

Discontent with paywall protected research has been growing for decades. The main reason for this is that the generation of the scientific content is largely funded by public funds. The creator, the author, provides this content for free to the publishers, who extract significant profit from users of this knowledge (other scientists, who are also funded by public funds and want to access copyrighted knowledge). Before the internet, publisher profit was relatively easy to justify, after all they were playing the crucial role in knowledge dissemination. However, the rise of the internet has changed this drastically. Today authors can share scientific knowledge with peers easily, directly through the internet. Thus, publishers have effectively lost their crucial role. This movement has disrupted the scientific publishing industry and resulted in the establishment of open access as a norm.

The rise of digitisation of intellectual property and of the usage of the internet as a distribution channel has also created a shift in consumer attitudes. Technology commentators have argued that digitisation and the internet would push businesses (of which publishing is one) to provide large parts of their content for free (Anderson 2009). As per the argument, this would be necessary to persuade consumers to pay for premium quality service, possibly because of the non-excludability feature of digital products. This is indeed what has happened in publishing with a number of projects like Google Books, as well as open access issues of scientific journals.

One could argue that the abundance of free content provided by copyright holders would leave consumers wanting more free content and push them towards pirating the part of the content that is not free. This is highly likely for products like music where end-consumers personally pay for products. However, we would argue that this is unlikely to be the reason when it comes to scientific journal content. In this case the end-users, researchers, do not personally pay for access to this content. Rather, most scientific journal sales go through subscriptions that are paid for by libraries. Besides, despite the rise of open access journals in recent years, a great majority of the high quality research is still behind paywalls. Thus, the role of free access to part of the content in encouraging consumer expectations for more free content is not clear in the case of scientific publishing.

Although open access emerged as an international desirable practice a few decades ago, modest results have been achieved so far in providing the public with free access to scholarly research. This is, arguably, due to the fact that scientific journal publishers turn significant margins from peer-reviewed journal subscriptions and are reluctant to open up to free access.

However, some international publishers have been moving towards (partial or full) open access publications and research data. In the field of European Studies (ES), online journals like this one (the JCER) and the *European Integration Online Papers* have been pioneers in making top-level research accessible for free worldwide. Moreover, generally these types of journals also prove to have a higher research outreach than non-open access content (Antelman 2004).

Compared to the open access journals, the creation of digital piracy engines for sharing international scientific publications represents a more radical development of open access practice. These new players emerge as vital actors in disseminating and determining the potential of research outreach regarding various academic and societal stakeholders. Piracy services are seen as the portal that gives a chance to scholars from poorer countries to access paywall-protected scientific research.

Online piracy represents a copyright infringement whereby copyrighted material is reproduced or distributed without appropriate permission. For the purpose of this piece, the copyrighted material

represents scientific articles. In a classic scientific publishing model, the copyright holder is a publisher to which the author assigns the right. Therefore, any further use of the scientific content requires explicit agreement by the publisher. The situation when a scientific article is reproduced or distributed without such agreement constitutes a copyright infringement, which has come to be known as online (or internet) piracy.

However, it is appropriate to point out the difference between online piracy and the original meaning of the term. Piracy originally describes the theft of excludable goods (meaning that consumption by one person precludes the consumption by another), while online piracy concerns a non-excludable good — scientific research. In the latter case, the consumption of the good by one person does not preclude others from consuming the same good.

Online piracy involves (at least) two parties: end-users and facilitators. Facilitators are individuals or organisations that acquire the content (usually through legal means) and share this content with other individuals who do not have sufficient legal rights to access the content. The latter actors are end-users of the act of piracy. It needs to be stressed that in most legal systems the responsibility lies with the facilitators, not the end-users.

The most well-known piracy engine (facilitator) for scientific papers is Sci-Hub, which has been rapidly gaining in popularity since its creation in 2011. Up to very recently, however, not much has been known about the size of its operations. Thus the cash-strained-researcher enabler status of Sci-Hub was never verified. However, recently data on five months of downloads from Sci-Hub service has emerged (Elbakyan and Bohannon 2016). The analysis of the raw data delivers a fascinating picture (Bohannon 2016). The concern of Bohannon (2016) is with the geographical distribution of download requests. The conclusion is that not only researchers in less-developed countries use the engine, but the Sci-Hub seems to be popular in developed countries too.

The picture presented by Bohannon (2016) is an aggregate view that relies on all download requests for all fields of study. Given that real science publications are more numerous compared to their social science counterparts, these findings might be hiding interesting patterns within social sciences.

In this work, we examine the Sci-Hub downloads data in order to uncover patterns of piracy in ES. We identify those ES journals that have been pirated the most. After identifying the top downloaded articles, we examine the subjects that most interest readers. We also analyse the geographical distribution of download requests, with a particular focus on developing countries and advanced Western countries and regions (e.g. the EU).

DATA

We use the data comprising all download requests received by the Sci-Hub servers between October 2015 and February 2016 (Elbakyan and Bohannon 2016). This represents a total of 22,915,621 download requests. The data have been anonymised in order to protect the identity of the user. For this the IP addresses have been aggregated to the nearest city location. Thus the data contains the city and the country from where the download request was received. The data contains the Digital Object Identifier (DOI) of the article requested. There is no other information about the requested article.

Therefore, identifying the articles from the ES field represents a challenge. Clearly, all ES articles cannot be identified. Therefore, we followed Babutsidze (2016) and proceeded as follows. We relied on information provided by UACES, the University Association for Contemporary European Studies;

UACES provided the list of ES journals, which comprises 25 journals. The impact factor across all these journals varies. The general consensus in scientific research is that the top journals aggregate the most robust and cutting edge research. Therefore, articles published in high impact-factor journals are more likely to be sought after, and thus pirated. Thus, in order to determine the extent of pirated consumption of ES research it sufficed to focus on high impact factor journals. From the list of 25 ES journals provided by UACES,¹ we chose the journals that have the highest ISI (Institute for Scientific Information) impact factor above 1. This criterion was satisfied by six journals: *West European Politics* (WEP), the *Journal of European Public Policy* (JEPP), the *European Journal of Political Research* (EJPR), *European Union Politics* (EUP), the *Journal of Common Market Studies* (JCMS) and *European Political Science Review* (EPSR). We identified the articles published in these six journals that were downloaded through Sci-Hub service between October 2015 and February 2016.

To identify the articles from the selected set of journals, we analysed the DOI assignment procedures by publishers of each of the journals. We identified unique character strings contained by digital object identifiers for each of the six journals that allowed us to separate ES publications from the rest of the observations. This reduced the working dataset drastically to 2,390 observations. This represents only 0.01 per cent of the whole dataset.

Following Babutsidze (2016), before carrying out the analysis, we removed duplicate downloads from the raw data by Bohannon (2016). The original data contains all page load requests received by Sci-Hub servers. When the user refreshes the browser that is in the process of loading the article, the server registers an additional download request. If we had the original IP data, these kinds of downloads could have been perfectly screened out. However, given the anonymised data we had to work with download time - download location pair of variables - in order to identify duplicates. To screen out multiple records for one actual download, we identified groups of downloads for the same paper that occurred from the same city within five minutes of one another. For each of these identified groups we retained only one download in our final dataset. This eliminated 80 observations and left us with the final dataset of 2,310 downloads for 1,537 distinct papers.

ANALYSIS

Our analysis includes 2,310 downloads over the span of five months. This implies 462 downloads on average per month for all the content generated by the six journals in our sample. One can conclude that the ES piracy numbers are not so impressive. In our opinion, this phenomenon can be explained by the fact that ES researchers are mostly concentrated in European countries where scholars have good access to scientific journals and thus do not have the financial need to pirate. Another reason for such a low number could be that researchers do not know about the existence of Sci-Hub as an alternative open resource to access the scientific output.

Table 1: Top downloaded European Studies articles

Authors	Year	Title	Journal	No of downloads
M. Rooduijn & T. Pauwels	2011	Measuring Populism: Comparing Two Methods of Content Analysis	WEP	10
L. Curini & F. Zucchini	2012	Government Alternation and Legislative Party Unity: The Case of Italy, 1988–2008	WEP	10
A. Niemann & D. Ioannou	2015	European economic integration in times of crisis: a case of neofunctionalism?	JEPP	8
C. Lyrintzis	1987	The power of populism: the Greek case	EJPR	7
G. Majone	1994	The rise of the regulatory state in Europe	WEP	7
C. Mudde	1995	Right-wing extremism analyzed	EJPR	7
M. Bovens	2010	Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism	WEP	7
I. van Biezen, P. Mair & T. Poguntke	2012	Going, going, . . . gone? The decline of party membership in contemporary Europe	EJPR	7
S. Engler	2015	Corruption and Electoral Support for New Political Parties in Central and Eastern Europe	WEP	7
G. Cordero & P. Simón	2016	Economic Crisis and Support for Democracy in Europe	WEP	7

Table 1 shows the ranking of the most downloaded papers. It shows that the two most pirated ES articles have collected only 10 downloads each over the five month period. Another observation is that people pirate both old and new articles. This is different than in the case of other disciplines, such as Economics, where a similar study has found that it is the download of recent articles that prevails (Babutsidze 2016).

One of the most interesting discoveries is linked to the subjects of the top ten downloaded articles. It appears that the highest demand for illegally downloaded articles is in the field of party politics (six out of ten). Populism and extremism are among the top searched subjects. Other leading subjects are the economic crisis and the related topics of the regulatory state and democratic accountability. Similar to populism and extremism, the latter subjects reflect the current major societal challenges confronting Europe.

Table 2: Top downloaded European Studies journals

Journal	No of downloads	No of articles downloaded	No of downloads / journal's total output	No of downloaded articles / journal's total output
West European Politics	795	546	0,462	0,317
Journal of European Public Policy	748	482	0,506	0,326
European Journal of Political Research	517	353	0,195	0,133
European Union Politics	180	108	0,436	0,262
Journal of Common Market Studies	68	46	0,030	0,020
European Political Science Review	2	2	0,012	0,012

Table 2 presents analysis at the journal level. We have to acknowledge that journals have generated article stock of different sizes. Obviously, more articles imply more potential downloads. Therefore, we needed to control for the size of research at a journal level. We gathered the data from ISI Web of Science about the total number of articles published by each journal. Web of Science (WoS) records are not perfect; however, they span over an extended number of years. Using the WoS data, we could estimate the number of articles published by each of the journals yearly and construct an estimate of the journal's total output. The last two columns in Table 2 normalise download data using these estimates. When examining the most pirated journals, it is worth mentioning that there are only three journals mentioned in Table 1 out of the six journals that we studied: WEP, the EJPR, and the JEPP. The fact that WEP leads with the top ten downloaded articles (six out of ten) and the total number of downloaded articles (in absolute terms) is not a surprise. The journal has established itself as one of the most authoritative publications on Western Europe and provides comprehensive coverage of all the major political and social developments in the region, including the European Union. However, when the number of published articles is taken into account, the JEPP comes top.

Looking at the profile of the top three downloaded journals, we can also conclude that Sci-Hub users are mostly interested in EU studies. This finding highlights that ES research (in terms of both supply and demand) is largely dominated by the study of the European Union.

Table 3 presents the countries from where the content has been most frequently downloaded. Compared to the rankings reported by Bohannon (2016), our data reveals a different pattern of geographical downloading. The explanation seems quite straightforward as the European Studies field has a strong geographical determinant. However, it is noteworthy that Brazil, China and the United States still generate more than 10 monthly downloads each.

Table 3: Top downloading countries

Country	No of downloads	No of yearly downloads / UACES member
Spain	308	22
Italy	289	18
Germany	176	5
Ukraine	144	173
Poland	103	9
Brazil	96	-
Turkey	90	9
France	76	8
Russia	65	26
Belgium	59	2
Hungary	57	10
China	54	26
United States	51	6

A more accurate picture on the popularity of piracy in ES has to take into account the size of the research bodies in each of the countries. For this we require the number of researchers engaged in ES in each of the countries. Such data is not readily available. However, we can proxy the number of ES researchers by the data on UACES membership. One of the 13 countries in Table 3, Brazil, has no UACES members. The data on the remaining 12 countries is taken into account in the third column of Table 3.

Financial background and the likelihood of the illegally downloaded scholarly articles are expected to be highly correlated. The well-established Northern European universities, for example, have sufficient financial resources to provide legal access to electronic resources to its staff and students. Moreover, these institutions have a larger number of ES courses and ES programmes (Timus and Cebotari 2014; Timuş, Cebotari and Hosein 2016). However, the opposite is expected in developing countries, both on the European continent (e.g. post-Soviet EU neighbours) and worldwide. Paradoxically, our data reveals that the highest number of total downloads comes from 'old' EU members, such as Germany, Italy, Spain and not developing countries. We believe this may be explained by the higher demand for ES research in the West European countries. As Germany and Italy were among the founding members of the EU, they have a stronger tradition of ES research. But it is also likely that the poorer financial situation of Italian and Spanish universities may be the driving factor behind the highest ES piracy rates in absolute numbers.

Table 3 also reveals that Ukraine is on the top of illegal downloads per UACES member. First, we must acknowledge that there were only two officially registered UACES members in Ukraine in 2014. However, Ukraine also holds fourth place among the countries with the highest downloads in terms of absolute numbers. The Ukrainian context, in our opinion, combines a poor financial situation for higher education institutions and an increasing demand for ES research. This demand may also be linked to the pro-European orientation of Ukrainian political elites since the Orange Revolution in 2004 and particularly after the Maidan protests in 2014. One can speculate that the high number of country downloads in Turkey may also be explained by the ongoing EU accession process and the active political and societal debates linked to the European integration process.

CONCLUSION

Despite the limited time period of data on illegal downloads, this study has revealed several factors explaining the illegal download of ES scholarly publications. First, the limited number of downloads in the ES field, compared to other disciplines, may be due to the geographical concentration of the ES research on the European continent. The high degree of access to scientific journals among European universities, particularly within the EU, could be a major explanation for the lack of financial need to pirate. However, we do not rule out the possibility of a low degree of knowledge about Sci-Hub as an alternative open source of downloading ES research.

Second, the most pirated journals are the well-established top ES journals like *West European Politics*, the *European Journal for Political Research* and the *Journal of European Public Policy*. The readers are mostly interested in subjects reflecting the current major European societal challenges, specifically populism and the economic crisis.

Last, but not least, the analysis presents mixed findings regarding the worldwide downloading pattern. On the one hand, the financial situation of universities in certain countries appears to be accountable for a higher number of downloads within Europe. On the other hand, the higher demand for ES research represents another key determinant of the number of downloads worldwide. Our analysis reveals that an older EU membership explains a stronger tradition of ES research and, consequently, higher illegal downloads of ES publications. Moreover, a pro-European political discourse (Ukraine) or an ongoing EU accession process (Turkey) can also account for a higher demand for pirated ES research.

Ultimately ES might be getting something good from the Sci-Hub: publishers are not losing much revenue (due to the small number of downloads), while researchers in under-developed countries and cash-strapped universities are getting access to important content. This is what open science should be about!

ACKNOWLEDGEMENTS

We thank Emily Linnemann for providing us with the UACES membership counts for the countries of interest as of 2016.

CORRESPONDENCE ADDRESS

Natalia Timuş, Maastricht Graduate School of Governance and UNU-MERIT, Boschstraat 24, 6211 AX Maastricht, The Netherlands [n.timus@maastrichtuniversity.nl].

¹ List of European Studies Journals, available online <http://www.uaces.org/resources/list-of-european-studies-journals> [last accessed 06 June 2016].

REFERENCES

- Anderson, C. (2009) *Free: How Today's Smartest Profit by Giving Something for Nothing*. London: Random House.
- Antelman, K. (2004) 'Do Open-access Articles Have a Greater Research Impact?' *College and Research Libraries* 65(5): 372-382.
- Babutsidze, Z. (2016) 'Pirated Economics,' *MPRA Paper No: 71703*. University Library of Munich, Germany, <https://mpra.ub.uni-muenchen.de/id/eprint/71703> [last accessed 20 June 2016].
- Bohannon, J. (2016) 'Who's Downloading Pirated Papers? Everyone,' *Science* 352(6285): 508-512. DOI: 10.1126/science.352.6285.508.
- JCER (2016) 'Continued Journal Ranking Success,' *Journal of Contemporary European Research*, available online: <http://jcer.net/index.php/jcer/announcement/view/35> [last accessed 20 May 2016].
- Elbakyan, A. and Bohannon, J. (2016) 'Data From: 'Who's Downloading Pirated Papers? Everyone'', *Dryad Digital Repository*. DOI: 10.5061/dryad.q447c.
- Enserink, M. (2016) 'In Dramatic Statement, European Leaders Call for 'Immediate' Open Access to All Scientific Papers by 2020', *Science*, 27 May, DOI: 10.1126/science.aag0577 [last accessed 28 May 2016].
- European Commission (2015) 'Public consultation 'Science 2.0': Science in Transition,' *Background Paper*, available online: <https://ec.europa.eu/research/consultations/science-2.0/background.pdf> [last accessed 20 May 2016].
- European Commission (2016a) *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020*, available online: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf [last accessed 9 May 2016].
- European Commission (2016b) *Open Innovation, Open Science, Open to the World*. Brussels: Directorate-General for Research & Innovation (RTD).
- Moedas, Carlos (2014) Speech at European Research and Innovation for Global Challenges, Lund, 4 December 2015.
- Nielsen, M. (2012) *Reinventing Discovery: The New Era of Networked Science*. Princeton: Princeton University Press.
- Timuş, Natalia and Cebotari, V. (2014) *INOTLES Survey on Teaching European Studies (2014)*, available online: <http://www.inotles.eu/survey-teaching-es> [last accessed 9 May 2016].
- Timus, N., Cebotari, V. and Hosein, A. (2016) 'Innovating Teaching and Learning of European Studies: Mapping Existing Provisions and Pathways', *Journal of Contemporary European Research* 12 (2): 653 - 668.