STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

23rd International Conference on Science and Technology Indicators (STI 2018)
"Science, Technology and Innovation indicators in transition"
12 - 14 September 2018 | Leiden, The Netherlands
#STI18LDN

Production and uptake of Open Access publications involving the private sector: the case of big pharma

Alfredo Yegros-Yegros[*], Thed van Leeuwen [**]

[*]*a.yegros@cwts.leidenuniv.nl*
CWTS, Leiden University, Nederlands

[**] *leeuwen@ cwts.leidenuniv.nl*
CWTS, Leiden University, Nederlands

Introduction

Over the last years Open Access (hereafter abbreviated as OA) has been ranked very high on science policy agenda's both internationally as well as nationally. This resulted in many national mandates and international guidelines on OA publishing of scientific results. One of the reasons OA has been pushed so strongly by science policy is found in the argument that what is financed publicly, should be publicly available. This argument, also known as the 'tax payers argument' is used to support and legitimize the push for open accessibility, not only of scientific publications, but also of the underlying research data, in order to guarantee the non-academic sector, with lower degrees of accessibility to otherwise 'behind-the-paywall' information, access to outcomes of scientific research in the public sector.

In this study we will focus on the developments in the OA publishing in one particular institutional sector, the private sector. Business enterprises represent the main sector in terms of R&D investments. According to Eurostat, in the year 2016 this sector represented 65% of the total R&D expenditures within the EU28. While objectives and incentives in the private sector might not always been aligned with the disclosure of research results in the open scientific literature, there is no doubt that this is the main actor when it comes to R&D performance.

Within the business sector, we will focus our study in the pharmaceutical sector, by selecting a number of large pharmaceutical companies (hereafter abbreviated as Big pharma). Pharmaceutical companies represent an interesting case of study, given that is it one of the most R&D intensive industries, with the highest investments in R&D (19.4% of all industries) according to the 2017 EU industrial R&D investment scoreboard[1]. Another reason for selecting Big pharma is due to the known shift in R&D orientation, from an in-house focus in the development of R&D towards a model much more open and collaborative, with more interactions with academic partners and other companies (e.g. Mignani et al, 2016; Schuhmacher, 2016). It has been also argued that OA should be mandatory for all research results related to drugs in order to make sure that these products are fit for use, in return for the government approvals for the commercialisation of these products (Doctorow, 2012).

[1] The 2017 EU Industrial R&D Investment Scoreboard (http://iri.jrc.ec.europa.eu/scoreboard17.html)

Despite the importance of industrial R&D, until now it remains relatively understudied the way in which private sector institutions active in R&D have embraced the OA movement. A study conducted back in 2008 (Stjernberg et al, 2008) by the university of Gothenburg reported the general lower awareness of OA in the business sector compares to universities at that time. More recent studies have started to consider companies in relation to the OA movement (Siler, 2018), but it remains relatively unknown how the private sector adapts to and can benefit from the new paradigm of open scholarship.

Our objective is to shed more light on the extent to which big pharma both has been publishing in OA and also has been benefiting from OA publications to build their own research.

The two main research questions that guide our research are :

- To what extent did Big pharma turn to the OA way of publishing scientific results themselves?
- To what extent has Big pharma benefit from OA publications?

Data and method

Selected pharmaceutical companies and scientific publications

We have considered 23 large pharmaceutical companies in this study[2]. This means that rather than including an exhaustive list of companies integrating the pharma sector, we focus on a small group representing the largest companies. Most of these companies are among the top 100 companies worldwide in terms of R&D investment. According to the EU Industrial R&D Investment Scoreboard produced by the JRC Economics of Industrial Research & Innovation, these 23 companies together invested more than 97 billion euros in R&D in 2016 and achieved sales for more than 600 billion in the same year.

In terms of the geographic distribution of headquarters, 11 of the selected companies are located in the US, 9 in Europe, 2 in Japan and one company in Israel. For the collection of their scientific publications, we considered all their subsidiaries as well as acquisitions. To define and delineate these companies we relied on Moody's Orbis, which is the largest commercial database providing corporate information, supplemented with other written sources (mainly Wikipedia and Rafols et al (2014)). Past publications produced by any acquisition or subsidiary have been assigned to the current parent company and all the analyses have been conducted at the level of the 23 current owners. There are two exceptions, the recent acquisition of Monsanto by Bayer AG as well as Eli Lilly's Elanco have not been considered in the analyses as their main focus is other than pharmaceuticals.

We collected from the CWTS in-house version of Clarivate Analytics' Web of Science Core Collection (WoS) all articles, reviews and letters published in the period 2009-2016, including at least one author affiliate address referring to one of the 23 pharmaceutical companies (or any of their subsidiaries or acquisitions).

2

Identification of Open Access publications

The methodological approach that we propose mainly focuses at adding different OA labels to the complete in-house version of the Web of Science database for the period 2009-2016, using various data sources to establish this OA status. Basic principles for this OA label are sustainability and legality. With sustainability we mean that it should, in principle, be possible to reproduce the OA labelling from the various sources used, again and again, in an open fashion, with a relatively limited risk of the source disappearing behind a pay-wall. The second aspect relates to the usage of data sources that represent legal OA evidence for publications, excluding rogue or illegal OA publications.

As main data sources we used:

- the DOAJ list (Directory of Open Access Journals) [https://doaj.org/],
- the ROAD list (Directory of Open Access scholarly Resources) [http://road.issn.org/],
- PMC (PubMed Central) [https://www.ncbi.nlm.nih.gov/pmc/],
- CrossRef [https://www.crossref.org/], and
- OpenAIRE [https://www.openaire.eu/]

We further distinguish between gold and green OA. A publication is considered gold OA if was published in a journal included either in the DOAJ or ROAD lists. If the publication is found in any of the other three sources (i.e. PMC, CrossRef or OpenAIRE), then it is consider green OA. In those cases in which the same publication is classified both as gold and green OA, then it is considered as gold OA only.

All these sources fulfil the above mentioned requirements while other popular 'apparent' OA sources such as ResearchGate and SciHub fail to meet these two principle requirements. Thus, it is important to highlight here that our approach has a more policy and strategy perspective than a utilitarian one. In other words, our approach aims to inform the number and share of sustainable and legal OA publications (i.e. publications that have been published in OA journals or archived in official and legal repositories), instead of the mere identification of publications whose full text can be retrieved online (regardless the source or the legal status of the access to the publication).

All these sources mentioned above were fully downloaded (as provided by the original sources) using their public Application Programming Interface (API). The obtained metadata has been parsed and linked to the CWTS in-house version of WoS.

More details on the approach used to identify OA publications and its validation can be found in van Leeuwen et al (2017).

Production/uptake of Open Access publications

In this study we analyse the interplay of large pharmaceutical companies with OA publications by looking at the publications produced and '*used*' by these companies.

The way we operationalise the production of OA publications is by looking at the publications to which the big pharma companies have contributed. In order to analyse the extent to which pharma companies '*use*' and benefit from OA publications is by looking at the reference list of publications produced by big pharma, and determining whether the cited documents were published as OA.

The uptake analysis is also limited to scientific publications cited by large pharmaceutical companies which in turn are covered by the WoS, more specifically to cited articles, letters and reviews. This means that publications in journals not covered by WoS as well as citations to any other types of documents not captured by WoS are not considered in this analysis.

An additional restriction in the uptake analysis is imposed by our current temporal coverage in the classification of OA publications, as currently we have complete information for the period 2009-2016. Therefore, any cited document published earlier than 2009 will be excluded of the analysis.

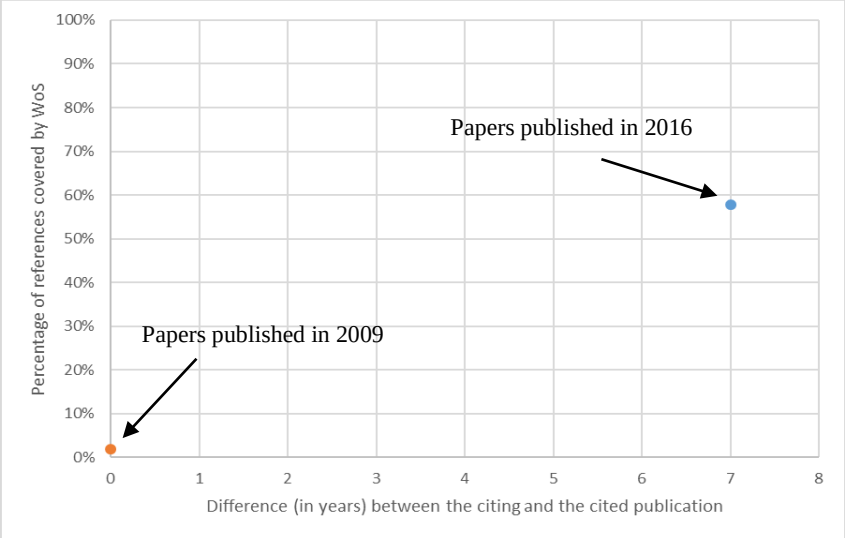
Table 1 shows the total number of publications cited by big pharma during the period 2009-2016 as well as the percentage of these cited publications that are covered by the WoS and the percentage of these corresponding to articles, letters and reviews. Therefore, from the perspective of the cited documents covered by WoS, we are able to include in our analysis most of the cited publications (~84%).

Table 1. Number and percentage of references covered by WoS

Publication year	All references	% references in WoS	% references (articles, letters and reviews) in WoS
2009	424,928	86.2%	84.2%
2010	431,958	86.4%	84.4%
2011	444,678	86.6%	84.6%
2012	433,047	86.8%	84.9%
2013	433,452	86.7%	84.7%
2014	432,733	86.9%	84.9%
2015	429,034	86.5%	84.5%
2016	490,754	86.8%	84.8%

Perhaps the most important limitation is our current coverage in the classification of OA publications, as it does not allow us to go back beyond 2009. The main reason for this limitation is the role played by DOIs in linking OA tags to publications in WoS. Most of the linking of external sources to WoS is done via DOIs, and we have seen that many older publications in our in-house version of WoS do not carry DOIs as extensively as from 2009 onwards (van Leeuwen et al, 2018). The implications for this part of our analysis can be observed with the example shown in Figure 1. For documents published by big pharma in 2009 we can only analyse references that were also published in 2009, which represent a very small amount of references (1.8%), while for publications by big pharma published in 2016 we can perform our analysis going back in time 7 years, which allow us to cover around 57.7% of their references.

Figure 1. Percentage of references covered by WoS by citing year



As the same publication might be cited by more than one paper published in a given year, for each citing year (from 2009 to 2016) we de-duplicated the cited publications in order to conduct this part of the analysis, and the same was done to analyse the whole period (2009-2016).

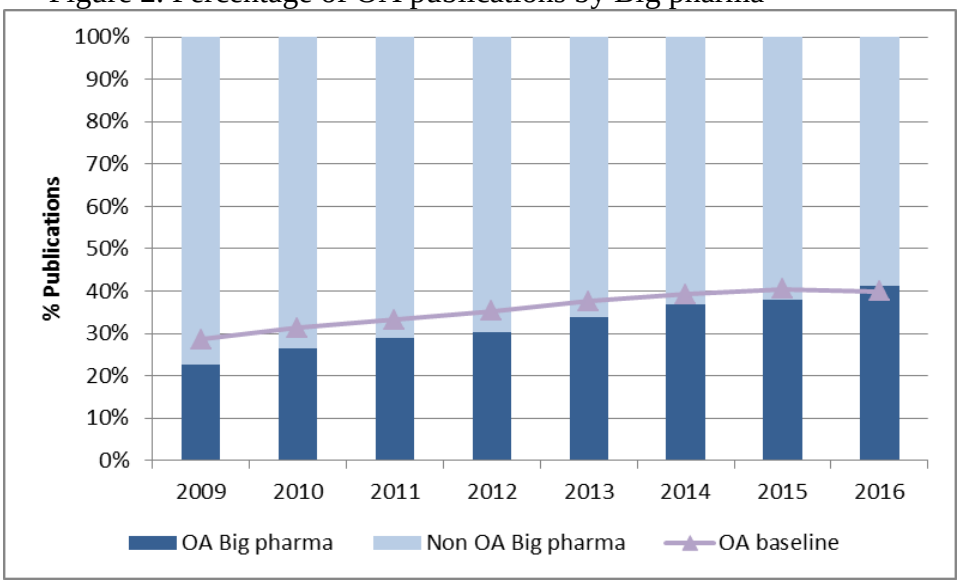
Results

Production of OA publications by big pharma

During the period 2009-2016, the 23 large pharmaceutical companies considered in this study contributed to more than 91,000 articles, reviews and letters. Around one third of these published documents (32.2%) are accessible as OA. This figure is slightly lower compared to the baseline publications (36.2%), which represents all articles, letters and reviews in the WoS subject categories where big pharma published 80% of all their publications.

As shown in Figure 2, the amount of OA publications by big pharma has been continuously increasing over time. The percentage of OA publications in 2016 is actually twice as large as the percentage of OA publications produced in 2009. This figure also shows that the growth in the percentage of OA publications has been higher in publications by Big pharma compared to the baseline. Since the beginning, and for most of the period, Big pharma has been lagging slightly behind the baseline in terms of OA. However, this gap has been closing over time and in the year 2016 the percentage of OA publications by Big pharma has surpassed the baseline.

Figure 2. Percentage of OA publications by Big pharma

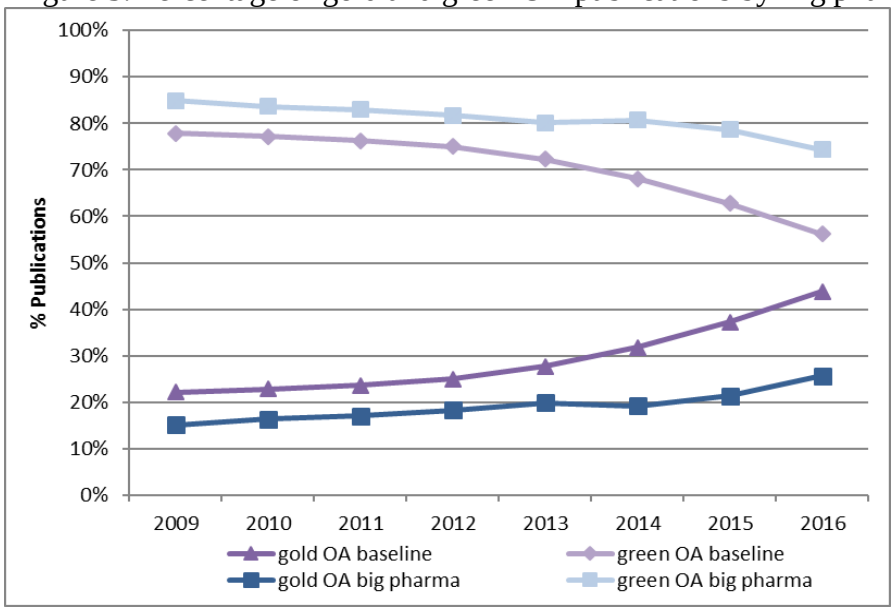


Next, we focus the analysis on the different modalities on OA to determine which is the preferred option for Big pharma companies when it comes to publish in OA. We distinguish between two main categories of OA, gold OA when papers were published in journals included in the ROAD or DOAJ lists, while green OA refers to publications made freely available by authors themselves by, for instance, self-archiving the publication in a repository. This part of the analysis is performed considering all OA publications only.

Large pharmaceutical companies published 19.7% of all OA documents using the gold OA formula while the vast majority of papers were made freely available through the green OA option (80.3%). This general pattern is also observed in the baseline publications, where most of the OA publications follow the green route, however a close look to the percentage of publications in the two routes reveals that the percentage of publications following the gold route is higher in the baseline compared to that of the big pharma (30.6%) while the percentage of publications in the green OA is lower compared to that of the big pharma (69.4%).

Gold OA is less frequent compared to the green route, both in large pharma companies as well as in the baseline, however Figure 3 shows that the percentage of green OA publications is continuously decreasing over time while the percentage of gold OA has experienced a steady increase. This pattern is more evident and marked for the baseline publications. Big pharma OA publications follow the same pattern towards more gold and less green, although this pattern is apparently following a slower pace compared to the baseline.

Figure 3. Percentage of gold and green OA publications by Big pharma



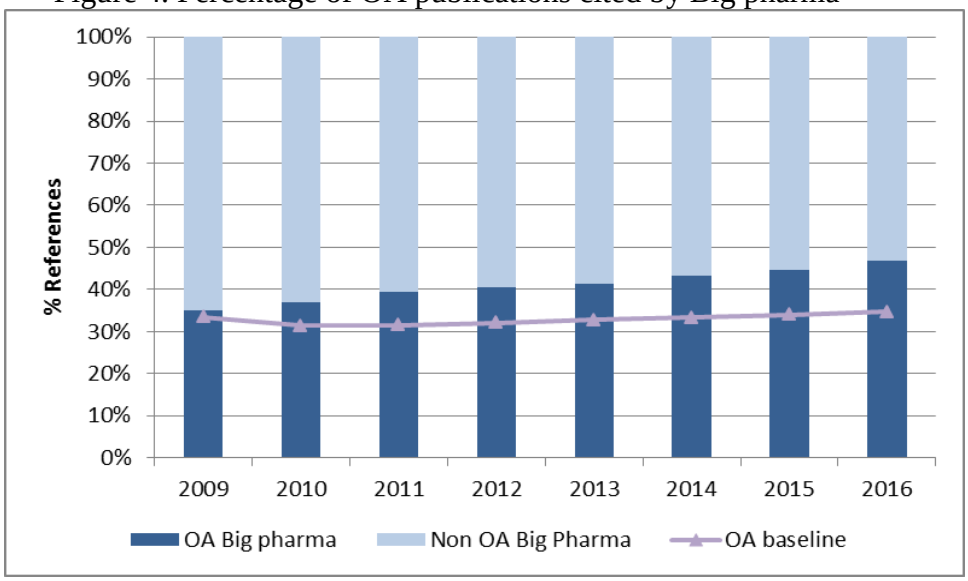
Uptake of OA publications by big pharma

We next look at the uptake of OA by Big pharma companies, and this is done by analysing the reference lists of publications (co-)authored by these companies. As explained in the methodology of our study, we consider unique cited article, reviews and letters that were published in the period 2009-2016.

The 91,000 articles, reviews and letters published by Big pharma in the period 2009-2016, cited 646,290 unique publications covered in the WoS (also articles, letters and reviews published in the period 2009-2016). Within all these cited publications, 43.2% were published in OA while 56.8% were non-OA. The percentage of OA publications cited by big pharma is higher compared to that of the cited OA publications by the baseline (33.4%).

Figure 4 shows for each citing year, the percentage of publications cited by Big pharma that were published on OA. It can be observed a gradual and positive increase in the percentage of cited OA publications (35.1% in 2009 and 46.8% in 2016), which differs from the almost stagnant trend observed in the baseline. This leads to an increasing difference over time between the percentage of cited OA publications by Big pharma and the baseline.

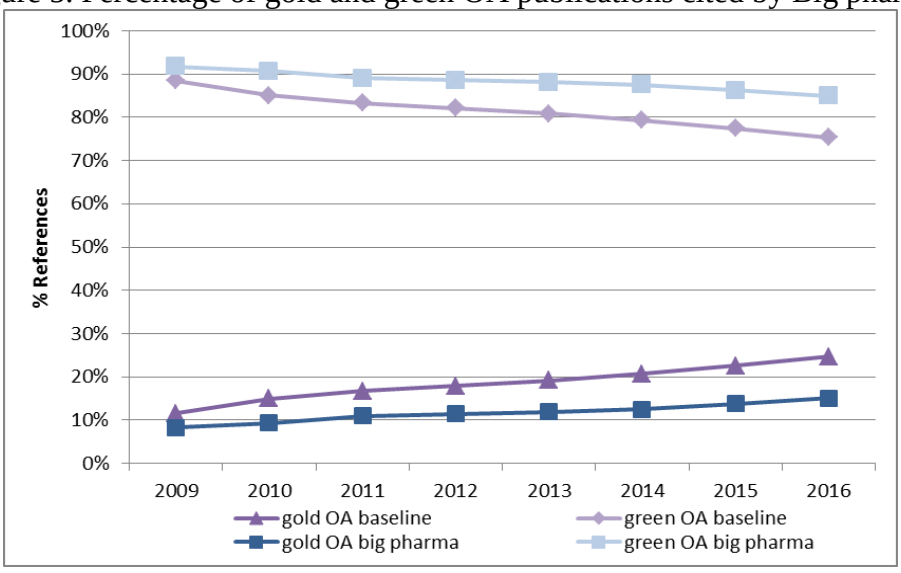
Figure 4. Percentage of OA publications cited by Big pharma



Having a closer look to these cited publications that were published as OA, we further distinguish again between gold and green OA. Most of the OA cited literature during the period 2009-2016 was made open via the green route (87.1%), while cited gold OA represents 12.9% of all OA cited literature. The picture when looking at the baseline is quite similar, with a slightly higher percentage of cited gold OA (21.1%) and therefore slightly lower percentage of cited green OA (78.9%), compared to Big pharma.

Figure 5 shows the trend in cited green and gold OA. This figure shows for both, Big pharma and the baseline, there is an increasing tendency towards the use of gold OA, although even for the most recent years of the period most of the OA cited literature was published via green OA.

Figure 5. Percentage of gold and green OA publications cited by Big pharma



Conclusions

The OA movement has gained a prominent position in the policy agenda during the past years. So far a number of studies have been analysing different aspects related to OA, such as the general degree of implementation, the effect of OA on later citation impact, etc. However, so far the analysis of the degree of adaptation and use of OA literature by the private sector has remained relatively under studied. The main contribution of our research is to provide evidence on how the private sector, with a focus in the Big pharma, is embracing the OA publishing, both as a way of disclosing their own research results and also as a mechanism to uptake previously published research.

Our findings point to a clear tendency towards the publication in OA by Big pharma. According to our results, gold OA seems to be gaining momentum when it comes to the publication of results, but the scene it is still dominated by green OA. The amount of cited OA publications is also increasing over time. Big pharma is therefore more and more actively publishing and consuming OA literature. Although at the beginning of the studied period Big pharma was benefiting relatively more from OA compared to what they were publishing as OA, this difference has been vanishing over time, and in recent years Big pharma is basically contributing to the OA pool of papers as much as they are benefiting from it.

As discussed, the study is not free of limitations. Some additional results from this study, not reported here due to the space limitations, will be presented in the STI conference.

References

Doctorow, C (2012). Why all pharmaceutical research should be made open access. The Guardian. https://www.theguardian.com/technology/2012/nov/20/pharmaceutical-research-open-access

European Commission (2017) The 2017 EU Industrial R&D Investment Scoreboard. JRC IRI – Economics of Industrial Research and Innovation

Mignani, S; Huber, S; Tomas, H; Rodrigues, J, Majoral, JP (2016) Why and how have drug discovery strategies in pharma changed? What are the new mindsets? Drug discovery today, 21(2):239-249

Rafols, I., Hopkins, M.M., Hoekman, J., Siepel, J., O'Hare, A., Perianes-Rodríguez, A., Nightingale, P. (2014) Big Pharma, Little Science? A bibliometric perspective on big pharma's R&D decline. Technological Forecasting & Social Change, 81, 22-38.

Schuhmacher, A; Gassmann, O; Hinder, M (2016) Changing R&D models in research-based pharmaceutical companies. Journal of Translational Medicine 14:105. https://doi.org/10.1186/s12967-016-0838-4

Siler, K., Haustein, S., Smith, E., Larivière, V.; Alperin, J.P (2018) Authorial and institutional stratification in open access publishing: The case of global health research. PeerJ 6:e4269; DOI 10.7717/peerj.4269

Stjernberg, H; Carlsson, H, Herbertsson LH; Kristiansson, RM, Sulg, P(2008) Open Access in Swedish private sector R&D. Project report 2008-12-01. https://gu.se/english/research/publication/?publicationId=79776

van Leeuwen, T., C. Tatum and P.F. Wouters (2018). Exploring Possibilities to Use Bibliometric Data to Monitor Gold Open Access Publishing at the National Level, accepted for publication in JASIST

van Leeuwen, T., Meijer, I., Yegros-Yegros, A., & Costas, R. (2017). Developing indicatorson Open Access by combining evidence from diverse data sources. Science and Technology Indicators conference, Paris 6-8 September. http://arxiv.org/abs/1802.02827