

Use of Free/Libre Open Source Software in Sepsis "-Omics" Research

Citation for published version (APA):

Evangelatos, N., Satyamourthy, K., Levidou, G., Brand, H., Bauer, P., Kouskouti, C., & Brand, A. (2018). Use of Free/Libre Open Source Software in Sepsis "-Omics" Research: A Bibliometric, Comparative Analysis Among the United States, EU-28 Member States, and China. *OMICS-a journal of Integrative Biology*, 22(5), 365-372. <https://doi.org/10.1089/omi.2018.0032>

Document status and date:

Published: 01/05/2018

DOI:

[10.1089/omi.2018.0032](https://doi.org/10.1089/omi.2018.0032)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Use of Free/Libre Open Source Software in Sepsis “-Omics” Research: A Bibliometric, Comparative Analysis Among the United States, EU-28 Member States, and China

Nikolaos Evangelatos,¹⁻³ Kapaettu Satyamourthy,⁴ Georgia Levidou,⁵ Helmut Brand,⁶
Pia Bauer,² Christina Kouskouti,⁷ and Angela Brand^{1,8}

Abstract

“-Omics” systems sciences are at the epicenter of personalized medicine and public health, and drivers of knowledge-based biotechnology innovation. Bioinformatics, a core component of omics research, is one of the disciplines that first employed Free/Libre Open Source Software (FLOSS), and thus provided a fertile ground for its further development. Understanding the use and characteristics of FLOSS deployed in the omics field is valuable for future innovation strategies, policy and funding priorities. We conducted a bibliometric, longitudinal study of the use of FLOSS in sepsis omics research from 2011 to 2015 in the United States, EU-28 and China. Because sepsis is an interdisciplinary field at the intersection of multiple omics technologies and medical specialties, it was chosen as a model innovation ecosystem for this empirical analysis, which used publicly available data. Despite development of and competition from proprietary commercial software, scholars in omics continue to employ FLOSS routinely, and independent of the type of omics technology they work with. The number of articles using FLOSS increased significantly over time in the EU-28, as opposed to the United States and China ($R=0.96$, $p=0.004$). Furthermore, in an era where sharing of knowledge is being strongly advocated and promoted by public agencies and social institutions, we discuss possible correlations between the use of FLOSS and various funding sources in omics research. These observations and analyses provide new insights into the use of FLOSS in sepsis omics research across three (supra)national regions. Further benchmarking studies are warranted for FLOSS trends in other omics fields and geographical settings. These could, in time, lead to the development of new composite innovation and technology use metrics in omics systems sciences and bioinformatics communities.

Keywords: commons, sepsis, open source software, omics research trends, innovation and technology policy

Introduction

THE SEQUENCING AND ANALYSIS of the human genome signaled the advent of a new era in biomedicine, where “-omics” technologies and Big Data have started revolutionizing our approach to screening, prevention, diagnosis, and treatment

of diseases, making personalized medicine a reality (Lander, 2011; Obermeyer and Emanuel, 2016). The anticipated downstream benefits are not restricted to biomedicine alone, but extend to practically every aspect of public health (Brand et al., 2016).

Apart from implications for research, clinical practice, and public health, omics technologies and Big Data have been

¹Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT), Maastricht University, Maastricht, The Netherlands.

²Intensive Care Medicine Unit, Department of Respiratory Medicine, Allergology and Sleep Medicine, Paracelsus Medical University, Nuremberg, Germany.

³Dr. TMA Pai Endowment Chair in Research Policy, Research Policy in Biomedical Sciences and Public Health, Prasanna School of Public Health (PSPH), Manipal University, Manipal, India.

⁴Department of Biotechnology, School of Life Sciences, Manipal University, Manipal, India.

⁵Department of Pathology, Klinikum Nuremberg, Paracelsus Medical University, Nuremberg, Germany.

⁶Department of International Health, Faculty of Health, Medicine and Life Sciences, Department of International Health, Maastricht University, Maastricht, The Netherlands.

⁷Department of Obstetrics and Perinatal Medicine, Klinik Hallerwiese, Nuremberg, Germany.

⁸Health Genomics, Manipal University, Manipal, India.

key drivers of the knowledge-based biotechnology innovation (Jimenez-Sanchez and Philp, 2015). The latter relies, by and large, on the modular knowledge base generated by omics technologies and is evidently distinct from traditional economic models, in that, it promotes new kinds of cooperation, often in the form of commons arrangements, between the economic actors (Evangelatos et al., 2016).

Advancements in molecular biology and biochemistry that gave birth to the new era have been largely based on parallel developments in theoretical biology as well as on disruptive progresses in information and communication technologies (ICT) (Portin, 2002). To gain a sense of proportion, in the last decades, modern hardware equipment has increased our ability to sequence DNA and RNA molecules by several orders of magnitude, allowing for accurate and affordable whole-genome sequencing (Heather and Chain, 2016).

Advancements in the instrumentation, however, need to be complemented by appropriate software solutions that allow for the generation, analysis, and interpretation of Big Data. High-resolution instruments place huge strains both on the computing hardware and the instrument vendors' software, and the need for accurate data visualization and complex statistical methods sets practically new standards in research, where open collaboration is indispensable (Levine and Prietula, 2014).

Free/Libre Open Source Software (FLOSS) is a prominent example of open collaboration and has been a core element of the bioinformatics tools employed in the analysis and interpretation of biological data. Indeed, the predecessors of FLOSS in the modern thriving field of bioinformatics are traced in the post-WWII period, when pioneers in molecular biology and biochemistry were, in many instances, prompted to generate unique software scripts, using appropriate scripting languages, to produce, analyze, and eventually put complex data in context (Needleman and Wunsch, 1970). As such scripts and software programs were often made freely available to colleagues, who were able to modify and use them according to their own needs, bioinformatics may be considered one of the disciplines that first employed FLOSS and provided fertile ground for its further development (Higgins and Sharp, 1988; Searls, 2010).

The attributes of FLOSS, as a model of open collaboration in the healthcare sector, are at great extent determined by the peculiarities of the latter, with the most significant domains of employment being the fields of bioinformatics (mainly in omics research) and healthcare delivery (Paton and Karopka, 2017). Despite a considerable growth over the last years, the use of FLOSS in the healthcare sector, compared to the ICT, is still lagging behind (Karopka et al., 2014) and attempts to assess its use, especially in the field of omics research, are still missing.

In this article, we argue that collaboration and sharing of knowledge, in the form of FLOSS, are an essential, integral component of omics research and that, despite the existence of proprietary software solutions, researchers working with omics nowadays still employ FLOSS routinely and independent of the type of omics technology they work with. Furthermore, taking into consideration the importance of FLOSS for the (supra)national economic performance, we investigate the differences in the use of FLOSS in omics research along national lines as well as the level and form of change over time. Last, in an era where sharing of knowledge

is being strongly advocated and promoted by public agencies and social institutions (Archibugi and Filippetti, 2015; Taichman et al., 2017), we investigate possible correlations between the use of FLOSS and various funding sources in omics research.

To address these issues, we focused our empirical lens on omics research in the highly heterogeneous syndrome of sepsis. As a syndrome, sepsis encompasses a broad spectrum of clinical situations and the relevant studies' populations present substantial variability. Because of this heterogeneity, research in sepsis is being conducted in different settings and the omics research agenda is broad, ranging from the bench to bedside and including practically every step in the omics cascade between the genome and the phenome, such as, for example, genomics, transcriptomics, proteomics, and metabolomics (Skibsted et al., 2013). Therefore, it may be argued that the use of FLOSS in omics research in sepsis is, at least qualitatively, representative of the relevant trends in omics research in biomedicine in general.

Materials and Methods

We conducted a retrospective, bibliometrical, longitudinal study of the use of FLOSS in omics research in sepsis, along the type of omics technologies employed and the (supra)national settings. To portray both the level and form of change over time and not just provide static measurements of the various percentages, we preferred a longitudinal study to a cross-sectional one (Ployhart and Vandenberg, 2010). Intending to augment reliability and avoid measurement errors, we conducted multiwave consecutive measurements for the period between 2011 and 2015.

We chose to begin with the year 2011, that is, 10 years after the announcement of the completion of the Human Genome Project (International Human Genome Sequencing Consortium, 2001), which essentially signaled the dawn of the era of omics and Big Data. This way, we allowed enough time for the development of proprietary software that may have substituted FLOSS in omics research.

All articles in omics research in sepsis published as free full texts in the Medline database in the consecutive years between 2011 and 2015 were retrieved using an inclusive MeSH strategy. The articles were reviewed for the presence of FLOSS and the type of omics technologies used, as well as for the (supra)national setting and the funding sources, and those rendered suitable were included in the analysis.

The inclusion criteria were as follows: full free text, articles with bioinformatics analysis, and English language (see MeSH Strategy section in Supplementary Data). For the purpose of our analysis, FLOSS referred to open-source software used for statistical analysis, data visualization, or every other type of bioinformatics analysis employed in the publication.

We differentiated between publications from the United States, EU-28, and China based on the affiliation of the first author. Despite the fact that some authors may have performed their research in a different national setting than the one of her/his primary affiliation (e.g., as visiting fellows), the resulting bias has not been considered significant for our research purposes. The same holds for the inclusion criterion of the English language, since the majority of the relevant research is, indeed, published in English.

We differentiated among the various omics regimes based either on the technologies employed (such as in genomics and transcriptomics, where the use of certain methods, such as sequencing, is almost a prerequisite) (Heather and Chain, 2016) or on the conceptual framework of the research (such as in proteomics and metabolomics, where a multitude of different technologies, ranging from immunoassays to mass spectrometry, can be employed) (Altelaar et al., 2013; Holmes et al., 2008).

We used publicly available data (free full texts) from the Medline database. Despite an increasing number of governmental and social initiatives that require the provision of free access to de-identified patient data from publicly and industry-funded research, these requirements remain to be fulfilled. Therefore, with no access to all relevant articles, a selection bias may confound the results of an analysis. However, our study focuses on comparisons between (supra)national entities. Therefore, the selection bias of relying on free full texts from Medline is systematic (affects equally the [supra]national entities under comparison) and does not affect the methodological integrity of the study.

Furthermore, working with publicly available data allows for control of result reproducibility by external reviewers, who have access to the same free data sets. In an era where knowledge has been increasingly conceptualized as a public good, bibliometric studies based on open access material reinforce the confidence of the academic community in the latter and prompt researchers to make their findings available to others with obvious benefits.

In our case, it may be legitimately argued that researchers who use FLOSS tend to publish their work in journals that provide open access to their articles more frequently than those who do not use FLOSS. This hypothesis rejects the null hypothesis upfront; nevertheless, it still raises a methodological concern. A possible solution would be to conduct a series of case studies to gain insight into the relevant behavior of the academic community. However, to the best of our knowledge, such studies are still missing and conclusion as to the magnitude of the effect cannot be drawn. Moreover, research has indicated that the penetration of the so-called open-source ethos beyond traditional software markets is not as substantial as initially thought (Lerner and Tirole, 2005). Therefore, it may be argued that the effect of the relevant bias does not have a substantial impact on our conclusions.

Results

In all, 1426 articles were retrieved for the years 2011–2015. Of those, 513 were unsuitable for inclusion in the analysis (336 not related, 117 reviews, 2 case reports, and 37 others, e.g., double entries). Of the rest 913 articles, information regarding the software and bioinformatics tools used were available for 833 of them. The results from the analysis of these 833 articles are depicted in Figures 1–3 as well as Tables 1 and 2.

The percentage of publications with omics research in sepsis that use FLOSS does not seem to change significantly over time, with the exception of 2013, where there is a significant worldwide decrease (Fig. 1).

The EU-28 shows a higher frequency of FLOSS in publications with omics research in sepsis when compared to the United States and China (Wilcoxon matched-pairs signed-

ranks test, $p=0.0431$ for both comparisons). Furthermore, and in contrary to the United States and China, the number of articles using FLOSS increases significantly over time in the EU-28 ($R=0.9568$ $p=0.004$, Table 1).

As to the frequency of FLOSS across different omics platforms, the highest frequency of use of FLOSS in total is observed in publications on genomics and the lowest in publications on metabolomics. However, the differences across different omics platforms did not reach statistical significance (total 60.36%, 47.83% in metabolomics, 50% in epigenomics, 53.85% in transcriptomics, 60.61% in proteomics, and 62% in genomics, Fischer's exact test, $p=0.290$, Fig. 2).

Only in China, this relatively higher frequency of FLOSS in the publications on genomics (81.25%), when compared to other omics platforms, achieved statistical significance (42.86% in transcriptomics, 60% in proteomics, and 25% in metabolomics, Fischer's exact test, $p=0.018$, Fig. 2).

Figure 3 depicts the differences in the use of FLOSS between single layer analysis (i.e., research within only one of the omics regimes that extend between the genome and the phenome) and multiomics approaches (i.e., research that employs more than one omics technologies and is being conducted in more than one omics layers). In this study, only the EU-28 shows a higher frequency of use of FLOSS in publications with multiomics when compared to publications with only one omics technology (67.61% vs. 61.88% respectively), a difference which, however, failed to attain statistical significance (Fischer's exact test, $p=0.468$, Fig. 3).

On the contrary, in China, the use of FLOSS is mostly observed in publications within a single omics regime (67.35%) and less frequently in multiomics research efforts (29.63%) (Fischer's exact test, $p=0.002$, Fig. 3).

When it comes to funding, university-funded research efforts in the United States and in EU-28 seem to show a lower frequency of FLOSS employment when compared with publications based on only government/state or nongovernment/state and nonuniversity funding (Table 2). This difference, however, did not reach statistical significance (Fischer's exact test, $p>0.10$ for all comparisons).

Discussion

Context of the study

Over the past several decades, omics technologies have been reshaping both public health and the economic landscape, facilitating the emergence of knowledge-based biotechnology innovations (Jimenez-Sanchez and Philp, 2015). The interdisciplinary field of bioinformatics is a core element of omics research and, at the same time, one of the disciplines that has provided fertile ground for the development of FLOSS. The Bioinformatics Open Source Conference (BOSC, 2018), and the Open Bioinformatics Foundation (OBF, 2018) are examples of resources dedicated to promoting and facilitating the open source development of bioinformatics tools within the biomedical research community.

The increasing importance of FLOSS is highlighted by the fact that, as of February 2018, the website SourceForge.net, an open source community resource providing free services to open source software projects, listed over 430,000 projects and hosted over 3.7 million registered users (Sourceforge,

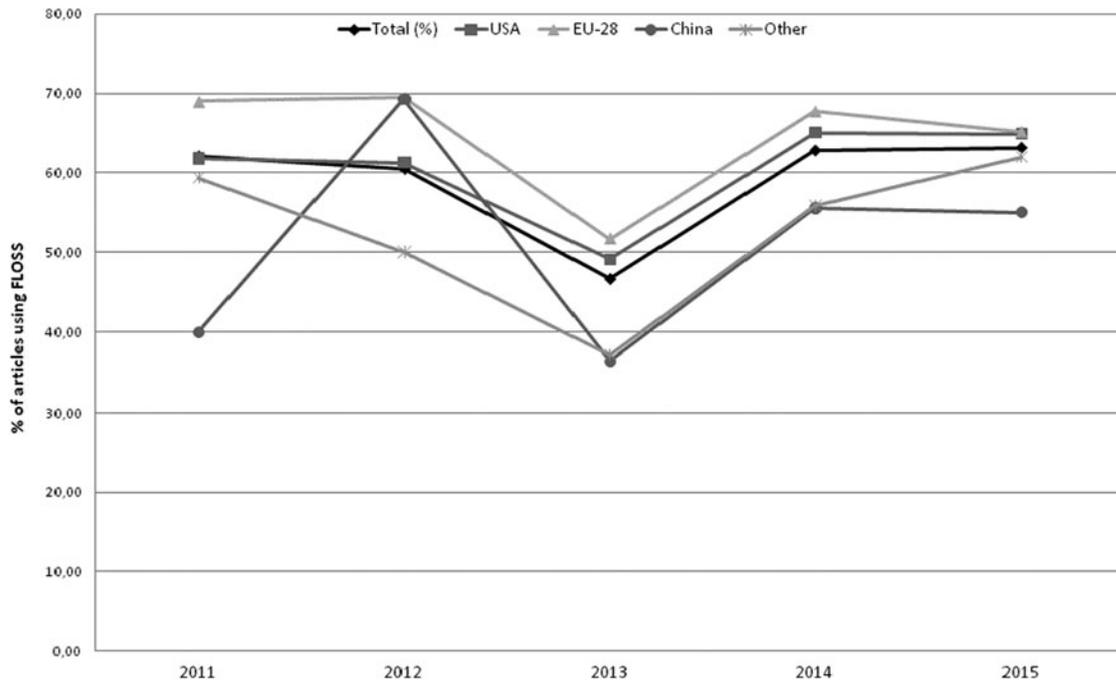


FIG. 1. Percentages of publications using FLOSS according to their origin 2011–2015. FLOSS, Free/Libre Open Source Software.

2018). Attempts to assess the economic impact of open source software on the innovation capacity and the competitiveness of the ICT sector have identified the direct and indirect economic relevance of FLOSS and the relative strengths and weaknesses in the United States and EU-28 (UMIC, 2006). However, similar studies for the healthcare sector are still missing.

In this study, we sought to investigate the characteristics of FLOSS use in omics research and its differences among (supra)national entities using sepsis as our empirical unit of analysis. FLOSS may be preferable over proprietary platforms in omics research for a variety of reasons. The active community of developers (most often consisting of the researchers who are users themselves) built around a source

code allows for open inspection and provides technical support and prompt fix of bugs. Furthermore, lack of costs associated with proprietary software lowers the entry barrier in omics research, especially in low-resource environments.

However, lack of software validation and educational support as well as difficult to use interfaces, and abandonment and fragmentation of projects render the use of FLOSS on occasion problematic (Earl, 2012). Coupled with the dissemination and maturation of omics technologies, these facts have given rise to the development of proprietary, user-friendly, stand-alone packages, which offer multiple bioinformatics tools together (Smith, 2015). Consensus to the relative virtues of FLOSS and proprietary software remains elusive, and studies that have sought to investigate the

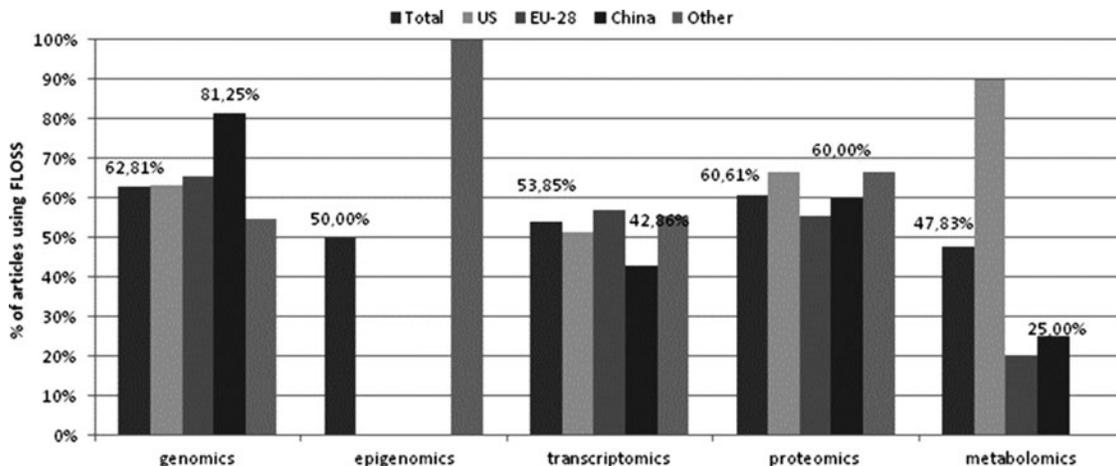


FIG. 2. Percentages of publications using FLOSS across different “-omics” platforms 2011–2015.

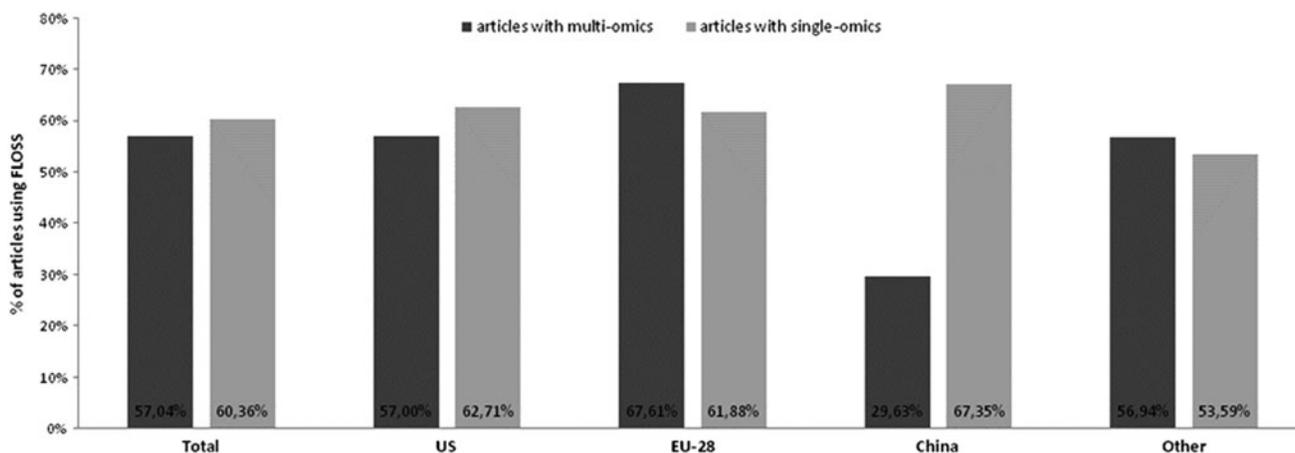


FIG. 3. Percentages of articles using FLOSS in articles using one “-omics” technology compared to those using multi-omics 2011–2015. Results of Fischer’s exact test.

advantages and market share of each platform have not focused on the use of FLOSS in omics research. Furthermore, while studies have shown that academic users are more likely to use FLOSS, when compared to private firms, possible correlations between the use of FLOSS and funding sources have not been investigated (Lemley and Shafir, 2011).

Key findings

Our results show that the percentage of FLOSS use in omics research in sepsis is high and does not change significantly over time (Fig. 1). This suggests that, despite the development of commercial software platforms, the relative virtues of FLOSS are indispensable in omics research. Accordingly, with the exception of China, the frequency of FLOSS use does not differ significantly across different omics platforms (Fig. 2). Moreover, the use of FLOSS does not differ significantly between research efforts within a certain omics regime and research efforts that span multiple layers in the omics cascade that extends from the genome to the phenome (i.e., multiomics). These findings suggest that the use of FLOSS does not depend on the peculiarities of the different omics regimes and further reinforce the conviction that FLOSS is an innate part of omics research.

In total, the highest frequency of FLOSS use is observed in genomics research and the lowest in metabolomics. Although this may be, to a certain extent, the result of the later emergence of metabolomics studies, the explanation for this dis-

crepancy may also lie with the different technological platforms used in each of these omics regimes. More specifically, the technological prerequisites for genomics research (e.g., sequencing) are more affordable and much more accessible compared to the ones for metabolomics, where advanced instrumentation, such as mass spectrometers, is needed. Instrument-specific, as in the case of mass spectrometers, built-in vendors’ software saves data in proprietary format and export of data in open, common formats, even when possible, may lead to loss of data.

Taken together, these facts may lead many researchers to use compatible commercial software for the further processing of metabolomic or proteomic data, whereas they can more heavily rely on FLOSS for the processing of genomic data. This explanation is in line with the fact that in 2013, where the use of FLOSS decreases significantly worldwide, there is the highest percentage of publications with proteomics and/or metabolomics in sepsis (data not shown). It also raises interesting questions as to the relationship of hardware and software and how the former may determine the possibilities for development of FLOSS.

In an era where research is closely related to the knowledge-based biotechnology innovation (Azoulay et al., 2015), the differential use of FLOSS in omics research may have implications for the economic performance of the various (supra)national economic actors. Interestingly, not only is the use of FLOSS in sepsis higher in the EU-28, compared to the United States and China, but it also increases over time

TABLE 1. NUMBER OF ARTICLES AND PERCENTAGES OF ARTICLES USING FREE/LIBRE OPEN SOURCE SOFTWARE ACCORDING TO ARTICLE ORIGATION FOR THE YEARS 2011–2015

N, (%)	2011 (%)	2012 (%)	2013 (%)	2014 (%)	2015 (%)	Total (%)	R
US	21 (61.76)	30 (61.22)	28 (49.12)	52 (65.00)	37 (64.91)	168 (60.64)	R=0.3695; p=0.165
EU-28	20 (68.97)	25 (69.44)	31 (51.67)	42 (67.74)	43 (65.15)	161 (63.64)	R=0.9568; p=0.004 ^a
China	2 (40.00)	9 (69.23)	4 (36.36)	15 (55.56)	11 (55.00)	41 (53.94)	R=0.3598; p=0.169
Other	19 (59.38)	23 (50.00)	13 (37.14)	24 (55.81)	44 (61.97)	123 (54.19)	R=0.4771; p=0.197
Total	62 (62.00)	87 (60.42)	76 (46.63)	133 (62.74)	135 (63.08)	493 (59.18)	R=0.7606; p=0.034 ^a

The bibliometric analysis is based on the Medline database and focused on the open-access (free full text) publicly available research studies in English language. Results of the Spearman’s correlation coefficient.

^aStatistical significance.

TABLE 2. PERCENTAGES OF ARTICLES USING FREE/LIBRE OPEN SOURCE SOFTWARE ACCORDING TO THE SOURCE OF FUNDING 2011–2015

Source of funding	Total	US	EU-28	China	Other
Only University	12/31 (38.71%)	1/3 (33.33%)	2/10 (20.00%)	1/1 (100%)	8/17 (47.06%)
Only government/state	176/310 (56.77%)	80/134 (59.7%)	38/67 (56.72%)	18/32 (56.25%)	40/77 (51.95%)
Nongovernment/state and nonuniversity	25/44 (56.82%)	5/10 (50%)	15/25 (60%)	0/1 (0%)	5/8 (62.5%)

The bibliometric analysis is based on the Medline database and focused on the open-access (free full text) publicly available research studies in English language.

at a statistically significant rate. Moreover, the EU-28 shows a higher frequency of FLOSS in publications with multiomics when compared to single omics research efforts, most probably a result of the increased use of FLOSS among the European research community.

These findings may be of relevance to the placement of the EU-28 and its performance regarding knowledge-based biotechnology innovation. As mentioned earlier, based on the modularity of omics data, the knowledge-based biotechnology innovation promotes the formation of new forms of collaboration, often in the form of commons arrangements. The modular base of Big Data generated by omics technologies allows for the formation of many modified commons arrangements with social and economic relevance (Hafen et al., 2014). Commons arrangements typically consist of a shared resource and a community built around it that operates on the ground of a self-regulation model, relying on a moral code. The institution of the commons has been mainly studied in the setting of nonexcludable, but rivalrous common-pool natural resources (Ostrom, 2010).

On the other hand, despite the fact that many consider knowledge a global public good (Archibugi and Filippetti, 2015), it has been argued that knowledge goods are non-rivalrous, but excludable, suggesting that knowledge goods, such as FLOSS, are in essence club goods (Evangelatos et al., 2016). FLOSS is a prominent example of open collaborations, where the source code, along with the right to copy, modify, and distribute it at will, is freely available under one of the many existing public licenses (Laurent, 2004). In that sense, researchers who develop or modify FLOSS for use in omics research are essentially forming a modified commons arrangement (in the sense of a club good), even if unaware of it or unknown to each other. In this study, commoners are the researchers/developers of the FLOSS, the latter is the shared resource, and the operating principles are those determined by the public license used, which essentially embodies the moral code on the basis of the so-called open source ethos.

The use of FLOSS in research produces knowledge spillovers and, therefore, added value to the research community, in ways that the use of commercial software cannot. More specifically, being freely available after development and modification, FLOSS can be further used and customized by researchers elsewhere for their own scopes. This way, the added value conferred on the software every time that is used and modified can be utilized by others in an ever enriching circle with obvious benefits.

Commons arrangements of this kind transcend conventional national, academic, and industrial barriers. This is especially relevant in low-resource environments, such as the developing countries, where the absence of license fees and

adaptability to local needs have proven the utility of FLOSS as a sustainable technology and knowledge transfer tool, and effectively lower entry barriers for omics research (Alkhatib et al., 2008). Especially in the case of sepsis, FLOSS may be proven an indispensable element of Research and Development (R&D) reshaping as well as technology transfer reorientation, next to other open innovation initiatives, such as innovative funding designs (e.g., crowdsourcing) and new R&D modes (Özdemir and Hekim, 2016).

Our results indicate that with omics research at the center of the knowledge-based biotechnology innovation and FLOSS being a core element of the former, the lead of the EU-28 may result in differences in its relative performance in this field, in terms of both commercial and academic productivity, compared to the United States and China. This could suggest that researchers in the EU-28 capitalize on previous efforts by using, modifying, and improving FLOSS to their benefit, whereas researchers from the United States and China rely more on commercial vendors.

Our results are in line with previous studies, which have shown that Europe, followed by the United States, has the lead in terms of FLOSS developers, whereas Asian countries lag behind (Crowston et al., 2012). It could be thus argued that, although with great discrepancies between the member states, the mentality of the European research community, regarding open collaboration, is closer to the basic premises of the new knowledge-based biotechnology innovation. In that sense, our approach could provide the basis for the development of future performance indicators for the new economic environment. In this study, future research is needed to verify the validity of these assumptions.

On the other hand, when researchers use commercial bioinformatics tools, the main beneficiary is the commercial vendor of the software platform. Coupled with the expected benefits from the shared knowledge in the form of FLOSS, this has prompted many public agencies to promote the adoption of FLOSS by the research community (Comino and Manenti, 2005). Indeed, taking into consideration the expected benefits, one may expect that the state would heavily subsidize the use of FLOSS in research by setting it, for example, as a funding requirement for research efforts.

However, in our study, we have not found any statistically significant correlations between the different sources of funding and the use of FLOSS (Table 2). Interestingly, the increased use of FLOSS in the EU-28 does not correlate with an increased public funding of omics research. These results may reflect reluctance, from the side of the states, to subsidize the use of FLOSS due to the free-rider problem (Hardin, 1968).

Responding to the challenge posed by FLOSS platforms and their support by many public agencies, commercial

developers have been interacting with FLOSS projects in a number of ways. Although improvements of FLOSS are not appropriate, to be fully updated, many firms offer complementary services to FLOSS projects that are not fully supplied by the FLOSS community and they even subsidize and participate in FLOSS projects. As a result, FLOSS and commercial platforms coexist in a selective ecosystem with relationships that range from competitive to symbiotic (Lerner and Tirole, 2005). Our data show that researchers use simultaneously FLOSS and commercial platforms in their work, irrespective of the (supra)national setting.

In all, our results suggest that with “-omics” research at the center of the knowledge-based biotechnology innovation and FLOSS being a core element of the former, the lead of the EU-28 may result in improvements in its relative performance in this field, compared to the United States and China. Further research, building on our present benchmarking findings, is needed to shed light on this evolving situation.

Conclusions

FLOSS is an integral part of omics research in sepsis and hence plays a significant role in the new knowledge-based biotechnology innovation. Our data indicate that research conducted in the EU-28 relies on FLOSS at a statistically significant higher rate compared to other (supra)national settings. These observations provide new insights on use of FLOSS in sepsis omics research across three (supra)national regions. Further benchmarking studies are warranted for FLOSS trends in other omics fields and geographical settings. This might eventually lead to the development of composite indicators for open source software use performance across omics innovation ecosystems.

Author Disclosure Statement

The authors declare that no competing financial interests exist.

References

- Alkhatib J, Mohab A, and Noori H. (2008). Open source: The next big thing in technology transfer to developing nations. IAMOT 2008: International Association for Management of Technology Proceedings.
- Altelaar AF, Munoz J, and Heck AJ. (2013). Next-generation proteomics: Towards an integrative view of proteome dynamics. *Nat Rev Genet* 14, 35–48.
- Archibugi D, and Filippetti A. (2015). Knowledge as global public good. In: *Handbook of Global Science, Technology and Innovation*. Archibugi D, Filippetti A, eds. Oxford, UK: Wiley-Blackwell, 483–597.
- Azoulay P, Graff-Zivin JS, Li D, and Sampat BN. (2015). Public R&D investments and private sector patenting: Evidence from NIH funding rules. NBER working paper no. 20883. Cambridge, MA: National Bureau of Economic Research.
- BOSC. (2018). www.open-bio.org/wiki/BOSC, Accessed February 18, 2018.
- Brand A, Evangelatos N, and Satyamoorthy K. (2016). Public health genomics: The essential part for good governance in public health. *Int J Public Health* 61, 401–403.
- Comino S, and Manenti FM. (2005). Government policies supporting open source software for the mass market. *Rev Ind Organ* 26, 217–240.
- Crowston K, Wei K, Howsion J, and Wiggins A. (2012). Free/libre open source software development: What we know and what we do not know. *ACM Comput Surv* 44, Article 7.
- Earl M. (2012). Open source software for mass spectrometry and metabolomics. In: *Open Source Software in Life Science Research. Practical Solutions to Common Challenges in the Pharmaceutical Industry and Beyond*. Harland L, Forster M, eds. Amsterdam: Elsevier.
- Evangelatos N, Reumann M, Lehrach H, and Brand A. (2016). Clinical trial data as public goods: Fair trade and the virtual knowledge bank as a solution to the free rider problem—A framework for the promotion of innovation by facilitation of clinical trial data sharing among biopharmaceutical companies in the era of omics and big data. *Public Health Genomics* 19, 211–219.
- Hafen E, Kossmann D, and Brand A. (2014). Health data cooperatives—Citizen empowerment. *Methods Inf Med* 53, 82–86.
- Hardin G. (1968). The tragedy of the commons. The population problem has no technical solution; it requires a fundamental extension in morality. *Science* 162, 1243–1248.
- Heather JM, and Chain B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics* 107, 1–8.
- Higgins DG, and Sharp PM. (1988). CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* 73, 237–244.
- Holmes E, Wilson ID, and Nicholson JK. (2008). Metabolic phenotyping in health and disease. *Cell* 134, 714–717.
- International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.
- Jimenez-Sanchez G, and Philp J. (2015). Omics and the bioeconomy: Applications of genomics hold great potential for a future bio-based economy and sustainable development. *EMBO Rep* 16, 17–20.
- Karopka T, Schmuhl H, and Demski H. (2014). Free/Libre open source software in health care: A review. *Health Inform Res* 20, 11–22.
- Lander ES. (2011). Initial impact of the sequencing of the human genome. *Nature* 470, 187–197.
- Laurent AMSt. (2014). *Understanding Open Source and Free Software Licensing*. 1st ed. O’Reilly Media, Inc., 1–224.
- Lemley MA, and Shafir Z. (2011). Who chooses open-source software? *Univ Chicago Law Rev* 78, 139–164.
- Lerner J, and Tirole J. (2005). The economics of technology sharing: Open source and beyond. *J Econ Perspect* 19, 99–120.
- Levine SS, and Prietula MJ. (2014). Open collaboration for innovation: Principles and performance. *Organ Sci* 25, 1414–1433.
- Needleman SB, and Wunsch CD. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 48, 443–453.
- Obermeyer Z, and Emanuel EJ. (2016). Predicting the future—Big data, machine learning, and clinical medicine. *N Engl J Med* 375, 1216–1219.
- OBf (Open Bioinformatics Foundation). (2018). www.open-bio.org/wiki/Main_Page, Accessed February 18, 2018.
- Ostrom E. (2010). The institutional analysis and development framework and the commons. *Cornell L Rev* 95, 807–816.
- Özdemir V, and Hekim N. (2016). Innovation management? Orienting sepsis R&D and technology transfer towards stratified medicine. *EBioMedicine* 6, 8–9.

- Paton C, and Karopka T. (2017). The role of free/libre and open source software in learning health systems. *Yearb Med Inform* 8, 26.
- Ployhart RE, and Vandenberg RJ. (2010). Longitudinal research: The theory, design, and analysis of change. *J Manag* 36, 94–120.
- Portin P. (2002). Historical development of the concept of the gene. *J Med Philos* 27, 257–286.
- Searls DB. (2010). The roots of bioinformatics. *PLoS Comput Biol* 6, e1000809.
- Skibsted S, Bhasin MK, Aird WC, and Shapiro NI. (2013). Bench-to bedside review: Future novel diagnostics for sepsis—A systems biology approach. *Crit Care* 17, 231.
- Smith DR. (2015). Buying in to bioinformatics: An introduction to commercial sequence analysis software. *Brief Bioinform* 16, 700–709.
- Sourceforge. (2018). <https://sourceforge.net/about>, Accessed February 18, 2018.
- Taichman DB, Sahni P, Pinborg A, et al. (2017). Data sharing statements for clinical trials: A requirement of the International Committee of Medical Journal Editors. *JAMA* 317, 2491–2492.
- UMIC. (2006). <http://arquivo.pt/wayback/20170901141650/http://www.english.unic.pt/images/stories/2006-11-20-flossim pact.pdf>, Accessed February 18, 2018.

Address correspondence to:
Nikolaos Evangelatos, MD, PhD
Intensive Care Medicine Unit
Department of Respiratory Medicine
Allergology and Sleep Medicine
Paracelsus Medical University
Prof. Ernst-Nathan-Strasse 1
Nuremberg 90419
Germany

E-mail: nikos.evangelatos@gmail.com

Abbreviations Used

EU-28 = European Union-28 Member States
 FLOSS = Free/Libre Open Source Software
 ICT = information and communication technologies
 MeSH = Medical Subject Headings
 R&D = Research and Development
 TT = technology transfer