

Estudo bibliométrico de software livre e open source

Free and open source bibliometric study

Domingos Silva

Instituto Universitario de Lisboa
(ISCTE-IUL) ISTAR-IUL

domingos_silva@iscte-iul.pt

Manuela Aparicio

Instituto Universitario de Lisboa
(ISCTE-IUL) ISTAR-IUL
Information Management School (Nova
IMS), Universidade Nova de Lisboa,

Portugal
manuela.aparicio@iscte-iul.pt

Carlos J. Costa

Advance/CSG, ISEG (Lisbon School of
Economics & Management),
Universidade de Lisboa

cjcosta@iseg.ulisboa.pt

Resumo — O software open source tem adquirido uma relevância crescente no mundo organizacional e social. Este artigo tem por propósito identificar os principais tópicos de investigação no contexto do open source e software livre. Para tal foi realizado um estudo bibliométrico no qual foram encontrados os clusters baseados em coocorrência de palavras-chave e coautoria, em duas bibliotecas digitais distintas, IEEE Explore e ACM Digital Library. Os resultados permitiram quantificar por clusters e concluir que temas são mais proeminentes no ecossistema de Software Livre. Embora se tenha identificado um crescimento quase exponencial em termos de publicações, formadas por pequenos grupos de investigadores e com poucas ligações entre elas.

Palavras Chave – *Software Livre, Código Aberto, Estudo Bibliométrico.*

Abstract — *Open source software has gained increasing relevance in the organizational and societal world. This paper aims to identify the main research topics in the context of open source and free software. To accomplish this objective, we carried out a bibliometric study in two different digital libraries, IEEE Explore and ACM Digital Library. The methodology intent was based on clusters co-occurrence of keywords and co-authorship. The results allowed to detect the clusters and themes that are more prominent in the Free Software ecosystem. Although an almost exponential growth in publications has been identified, formed by small groups of researchers and with few links between them.*

Keywords - *Free Software, Open Source, Bibliometric Study.*

I. INTRODUÇÃO

Há um conjunto de autores [1]–[5] que acredita que a adoção do Software Livre e de Código Aberto (FOSS) é uma medida para estimular o crescimento sustentável e que contribui muito para o desenvolvimento económico, social e tecnológico das organizações []. Como tal, essa suposição tem vindo a inspirar organizações públicas e privadas em todo o mundo [6] a adotar gradualmente o FOSS como alternativa viável ao software

comercial e como fonte estratégica para adoção mais rápida da tecnologia [7], [8], a um custo operacional menor comparado ao software comercial [9]. No entanto, apesar dos potenciais benefícios declarados que têm influenciado as organizações a adotar o FOSS, julga-se haver uma carência de estudos empíricos que examinem rigorosamente a participação do FOSS em diferentes domínios

Considerando o que foi exposto, este estudo tem como principal objetivo realizar uma revisão da literatura sobre software livre. De modo mais específico, mapear as principais publicações e autores de maior impacto no domínio do FOSS. Através de uma análise bibliométrica, tais como coautoria e coocorrência de autores, a fim de se identificar as principais áreas de estudo e avaliar a interlocução entre os autores inseridos nesta temática. O artigo caracterizou-se por um estudo de natureza teórica e conceitual, apoiado por uma revisão sistemática da literatura obtidos nas base de dados da IEEE Xplore Digital Library [10] e ACM DL [11], sobre Software Livre, Código Aberto, Adoção e Estudo Bibliométrico, resultando na identificação de artigos e autores inseridos ao ecossistema das principais áreas de desenvolvimento desta temática.

O estudo contribui para mapear os assuntos estudados no tópico do open source, bem como indicar novas pesquisas no tema. Este é um dos primeiros estudos bibliométricos sobre esta temática.

O presente artigo está organizado inicialmente pela secção da revisão de literatura sobre Software Livre. Na sequência é explicada a metodologia aplicada para recolha e tratamento de dados. Os resultados, discussão e por último, são expostas as considerações finais do artigo.

II. SOFTWARE LIVRE E OPEN SOURCE: CONCEITOS E ÁREAS DE INVESTIGAÇÃO

A. *Software Livre e de Código Aberto - Conceito*

Segundo a Free Software Foundation (FSF) [12], a organização responsável pelo movimento social, termos e condições de uso do Software Livre. Entende-se por Software Livre o Software que normalmente é obtido de forma gratuita e sem restrições de uso. Ou seja, qualquer pessoa ou organização que tenha as seguintes liberdades: (Liberdade 0) executar o Software como desejar, (Liberdade 1) estudar o funcionamento de um Software e adaptá-lo às suas necessidades, (Liberdade 2) redistribuir cópias do Software a outros quando desejar e (Liberdade 3) publicar as modificações feitas ao Software com o respetivo código fonte. No entanto, qualquer software que respeite as 4 liberdades fundamentais é Software Livre.

Software Livre e Software de Código Aberto são termos relacionados ao mesmo tipo de Software. No entanto, com diferentes visões e valores. Comumente, eles são representados por um único acrónimo FOSS (Free and Open Source Software). Vale aqui estabelecer uma diferença entre eles. Basicamente, o Software Livre concentra-se no movimento social e filosofia de uso e o Software de Código Aberto na metodologia de desenvolvimento e qualidade do Software [13].

Eric Raymond em seu ensaio “The Cathedral and the Bazaar” descreve a metodologia e modelo de desenvolvimento do Software Livre como “Bazaar” devido ao seu modelo descentralizado e inclusivo, assumindo que o Software Livre geralmente é mais barato, apresenta melhor qualidade, segurança e menos erros do que o Software comercial ou proprietário devido ao grande número de desenvolvedores voluntários espalhados pelo mundo que detetam e propõem melhorias ao Software Livre em menos tempo.

O mesmo autor também se refere ao Software comercial ou proprietário como “Cathedral”, porque tem um modelo de desenvolvimento fechado e é protegido contra cópia e distribuição. Consequentemente, cria aprisionamento tecnológico para utilizadores e organizações.

B. *Vantagens e Desvantagens do Software Livre*

Como pode ser visto nos conceitos acima, o uso do Software Livre evidencia várias vantagens sobre o Software comercial ou proprietário. Nosso estudo identificou que a adoção do Software Livre pode trazer vantagens significativas para organizações com economias em desenvolvimento. Pois, o Software Livre promove a economia local devido o seu baixo custo operacional [14], favorece a inclusão social e digital através dos seus valores essenciais descritos nas quatro liberdades [15] e a independência tecnológica e propriedade intelectual [16]. De fato, a utilização do Software Livre tem menor probabilidade de aprisionamento tecnológico “lock-in” e os utilizadores têm maior controle sobre seu ambiente computacional, porque qualquer um pode inspecionar o código fonte para analisar, corrigir e modificar quando for necessário (Liberdade 1) [17]. Outras vantagens estão associada a sua confiabilidade, segurança e qualidade [18], [31]-[38]. Todas estas vantagens são aspetos facilitadores para a sua adoção.

Alguns autores mencionam que por vezes o software open source, pode ter um interface pouco intuitivo [17], mencionam também a preocupação relativamente ao padrão nos aplicativos e complexidade do Software o que implica uma curva de aprendizagem bastante superior do que se este estivesse utilizando um Software proprietário. Outros estudos [19], [20] apontam a ausência de suporte e documentação pouco esclarecedora como os principais inibidores relacionados com a utilização de software livre.

C. *Principais áreas de estudo*

A decisão de adotar ou não uma tecnologia depende fundamentalmente da percepção dos potenciais utilizadores do sistema e do alinhamento aos objetivos das organizações. Este fenómeno tem sido estudado há pelo menos duas décadas por diferentes autores [12]-[18].

Entre as principais teorias utilizadas para explicar a intenção de uso e outros fatores condutores da adoção da tecnologia, a teoria da difusão da inovação (DOI) originalmente proposto por Rogers [12] define que a inovação é uma ideia, prática ou objeto que é percebido como novo por um indivíduo ou outra unidade de adoção. Rogers [12], indica cinco fatores condutores para adoção tecnológica: vantagem relativa, compatibilidade, complexidade, utilidade percebida e facilidade de uso. O mesmo autor diz que a adoção é fortemente motivada pelo aumento da eficiência e desempenho. Os mesmos fatores são também mencionados em estudos de outros autores como Dedrick e West [13], Glynn [14], Chau e Tam [15] no domínio do Software Livre, como potenciais condutores da adoção do Software Livre.

Alguns outros fatores, como organização, tecnologia e ambiente, são apresentados por Tornatzky e Fleischer [16]. Para os autores, o contexto organizacional refere-se às características e recursos da organização que restringem ou facilitam a adoção e implementação de inovações. Varian e Shapiro [19], identificam o baixo custo operacional do Software Livre como fator relevante para a organização.

Para Glynn [14], o contexto tecnológico está relacionado com às tecnologias interna e externa relevantes para a organização. Com base na ampla gama de produtos disponíveis em Software Livre, capaz de atender aos diferentes desafios tecnológicos de uma organização, o mesmo autor afirma que as vantagens do Software Livre se identificam como potenciais influenciadores da adoção.

Tornatzky e Fleischer concluem que o contexto ambiental está relacionado com o tamanho, estrutura, concorrentes e o ambiente macroeconómico que se insere a organização. Segundo Dedrick e West [13] fatores ambientais constituem oportunidades de presença no mercado como prestador de suporte e serviços, consequentemente a adoção de Software Livre. O Software Livre desencadeou estudos sobre uma variedade de vetores. Estes incluem estudos que destacam aspetos relacionados com a política [21] para estímulo e inovação a nível do governo. Isto é, estudos que analisam a viabilidade de governos adotarem Software Livre para assegurar um ambiente viável de desenvolvimento das TICs.

Outros estudos [22] procuram mostrar como a utilização do Software Livre proporciona vantagens económicas; outros buscam enfatizar o vetor técnico [23], isto é, a confiabilidade,

A tabela 1 resultou da análise das palavras-chave dos clusters na IEEE Xplore. Tentou-se de certa forma, designar um termo comum que melhor identificasse os clusters.

Clusters	Palavras	Designação
Clusters01	Computer Architecture, Operating Systems, Linux, Servers, Security, Cloud Computing	Infraestrutura de Comunicação e Segurança
Clusters02	Open Source Software, Public Domain Software, Costs, Adoption, Innovation	Sistemas de Informação e Inovação
Clusters03	Software, Java, Data Mining, Tools, Github, Software Testing	Prospecção de dados
Clusters04	Educational Institutions, Teaching, Education Course, Computer Science Education	Sistema Colaborativo
Clusters05	Application Software, Cryptography, Grid Computer, Usability, Privacy	Sistemas de Alta Escala
Clusters06	Middleware, Ubiquitous Computing, Transport Protocols, Open Source Tools, Web Services	Internet
Clusters07	Android, Testing, Social Network Services, Mobile Device, Humanoid Robots	Aplicações Moveis

C. Resultados da pesquisa na ACM DL

di penta, massimiliano

ven, kris

baik, jongmoon

niazi, mahmood

lützenberger, marco

küster, robias

kemer, chris f.

fichman, robert g.

hassan, ahmed e.

roebuck, kevin

spinellis, diomidis

lassenius, casper

mei, hong

kauffman, robert j.

janssen, marijn

tiley, scott

murphy, hill

witzschky, jeni

jones, lawrence g.

hammouda, imed

martinez, jabier

concas, giulio

bissiyandé, tchawendé f.

aparicio, manuela

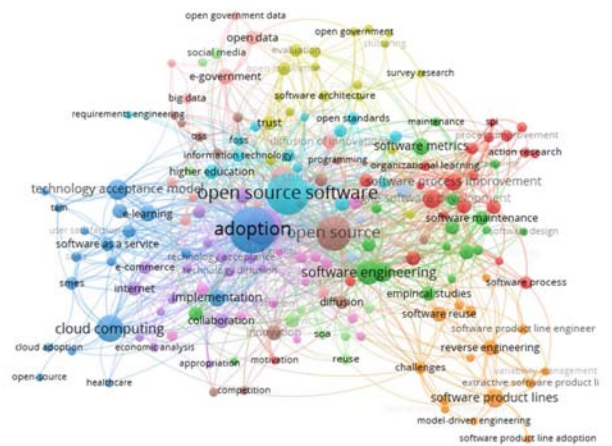
fitzgerald, brian

stol, klaus-jan

lundell, björn

ling, chuan

O mapa com as palavras-chaves totalizaram 4752 com 5 ocorrências mínimas e geraram 176 palavras-chaves em limiar, resultando em 12 clusters (Fig. 5.).



Para a tabela 2 utilizou-se a mesma abordagem aplicada na tabela.

TABELA II PALAVRAS CARACTERIZADORAS DOS CLUSTERS_ACM

Clusters	Palavras	Designação
Clusters01	Agile, Agile Adoption, Action Research, Scrum,	Desenvolvimento e Melhoria do Processos de Software
Clusters02	Empirical Studies, Software Designer, Software Tools, Software Testing, Software Development	Engenharia de Software
Clusters03	Adoption, Cloud Computing, Healthcare, ERP	Aceitação de Tecnologia
Clusters04	Semantic Web, Survey Research, Qualitative Research,	Agrupamento e Análise de Dados
Clusters05	Collaboratin, Diffusion of Innovation, Economic Analysis, Internet, It Adoption, Open Systems	Gestão do Conhecimento
Clusters06	Higher Education, Proprietary Software, Open Standards, Interoperability	Free Open Source Software
Clusters07	Adroid, Challenges, Industrial Adoption, Reverse Engineering	Linha de produtos de software
Clusters08	Communication, Developing Countries, Diffusion, Linux, OSS, Economics	Desenvolvimento Socioeconómico
Clusters09	Benefits, Education, Documentation,	Desenvolvimento Social
Clusters10	Framework, Open Government Data, Open Data, Machine Learning, E-government	Big Data
Clusters11	Architecture, Web 2.0, Reuse, Social Media, Management	Arquitetura Orientado a Serviço
Clusters12	Code Review, Security, Technology Transfer	Qualidade de Software

Nota: As cores constantes na tabela correspondem as cores dos clusters (Fig. 5).

V. DISCUSSÃO

Neste trabalho de investigação realizamos uma revisão literatura que possibilitou mapear os principais autores, palavras-chave e conceitos referentes ao Software Livre. Os nossos resultados são coincidentes com estudos anteriores. Primeiramente, a nível dos conceitos essenciais do Software Livre o estudo mostrou um uso generalizado em diversas áreas do conhecimento relacionada com a engenharia e gestão entre outras. No que diz respeito as áreas que mais se destacam a no ecossistema de Software Livre na IEEE Xplore [10] observa-se com grande frequência temas a volta de “Desenvolvimento e Melhoria do Processos de Software” e no que diz respeito a ACM DL [11] verificou-se maior destaque à temas relacionados com “Infraestrutura de Comunicação e Segurança”. O nosso artigo também mostra quem existem muitos autores que conduziram estudos sobre Software Livre, embora de maneira muito fragmentada com pequenos grupos de investigadores e poucas ligações entre eles.

Mascarenhas [28], Wang [29] e García-Lillo [30] conduziram também estudos bibliométricos utilizando o mesmo método e ferramenta [25] chegando a resultado semelhantes que demonstram características interessantes da VOS mas em contextos diferentes. Num dos estudos [28], buscou-se explorar e analisar a literatura relacionada à cooperação universidade-indústria, utilizando uma análise de coautoria. Outro estudo [29], faz uma análise bibliométrica sobre proliferação territorial de

autores. E por fim, o estudo [30] mostra os resultados de uma análise da literatura existente sobre um dos tópicos que mais despertou interesse entre estudiosos e pesquisadores nas áreas de gestão internacional e empreendedorismo.

Todos esses autores tiveram a sua metodologia baseada principalmente nas técnicas bibliométricas de análise de citações e co-citações de autores.

VI. CONCLUSÕES

O resultado do nosso estudo revela que os benefícios do Software Livre influenciam positivamente organizações e países com economias em desenvolvimento. Também verificamos a possibilidade e a capacidade de o Software Livre atender diferentes desafios tecnológicos com qualidade igual ou superior em comparação com Software comercial ou proprietário garantida por uma comunidade de desenvolvedores espalhados por todo o mundo.

Os resultados do nosso estudo mostram também grande participação do Software Livre em diversas áreas do conhecimento. Os resultados dessa análise nos levaram a entender a tendência de adoção do Software Livre em diversos vetores ou segmentos. Identificando-se um crescimento quase exponencial da literatura nos últimos anos. No entanto, foi identificado que as publicações estão muito fragmentadas e que os autores estão pouco ligados entre si.

Estamos, portanto, a analisar a possibilidade de explorar cada cluster de forma isolada, de maneira que possa fornecer uma explicação dos facilitadores da adoção do Software Livre de cada vetor ou segmento. Preferencialmente, aos clusters da ACM DL por serem os que mais se identificam com o nosso propósito final de estudo “Adoção de Software Livre no Setor Público”. Para tal, comecemos com os clusters menos destacados nos resultados: Clusters03, Clusters08 e Clusters09.

AGRADECIMENTOS

Este trabalho é parcialmente financiado por Fundos Nacionais através da FCT - Fundação para a Ciência e a Tecnologia no âmbito do projeto de financiamento com a Referência UID/SOC/04521/2019; e UID/MULTI/4466/2019.

REFERÊNCIAS

- [1] J. A. Lewis, “Government open source policies,” *Cent. Strateg. Int. Stud.*, 2010.
- [2] University at Albany, SUNY, I. Chengalur-Smith, S. Nevo, University at Albany, SUNY, P. Demertzoglou, and Rensselaer Polytechnic Institute, “An Empirical Analysis of the Business Value of Open Source Infrastructure Technologies,” *J. Assoc. Inf. Syst.*, vol. 11, no. 11, pp. 708–729, Nov. 2010.
- [3] N. M. Bacic, O software livre como alternativa ao aprisionamento tecnológico imposto pelo software proprietário. 2003.
- [4] P. Krakowski, “ICT and Free Open Source Software in Developing Countries,” in *Social Informatics: An Information Society for all? In Remembrance of Rob Kling*, vol. 223, J. Berleur, M. I. Nurminen, and J. Impagliazzo, Eds. Boston, MA: Springer US, 2006, pp. 319–330.
- [5] N. Bezroukov, “Open source software development as a special type of academic research: Critique of vulgar Raymondism,” *First Monday*, vol. 4, no. 10, 1999.

- [6] P. J. Agerfalk, A. Deverell, B. Fitzgerald, and L. Morgan, "Assessing the role of open source software in the European secondary software sector: a voice from industry," 2005.
- [7] M. Driver, "Drivers and incentives for the wide adoption of open-source software," *Gart. Rep. Sept.*, vol. 13, 2012.
- [8] L. Morgan and P. Finnegan, "Benefits and drawbacks of open source software: an exploratory study of secondary software firms," in *IFIP International Conference on Open Source Systems*, 2007, pp. 307–312.
- [9] B. Fitzgerald, "The transformation of open source software," *Mis Q.*, pp. 587–598, 2006.
- [10] "IEEE Xplore: Advanced Search." [Online]. Available: <https://ieeexplore.ieee.org/search/advsearch.jsp>. [Accessed: 28-Jan-2019].
- [11] "ACM Digital Library." [Online]. Available: <https://dl.acm.org/advsearch.cfm?coll=DL&dl=ACM>. [Accessed: 28-Jan-2019].
- [12] R. M. Stallman, "gnu.org." [Online]. Available: <https://www.gnu.org/philosophy/free-sw.en.html>. [Accessed: 20-Jan-2019].
- [13] R. Stallman, "Viewpoint Why open source misses the point of free software," *Commun. ACM*, vol. 52, no. 6, pp. 31–33, 2009.
- [14] "The Open Source Definition (Annotated) | Open Source Initiative." [Online]. Available: <https://opensource.org/docs/definition.php>. [Accessed: 02-Feb-2019].
- [15] R. Stallman, "Is digital inclusion a good thing? How can we make sure it is?," *IEEE Commun. Mag.*, vol. 48, no. 2, 2010.
- [16] "2006: Time to plan your company's escape from Microsoft," *SearchDataCenter*. [Online]. Available: <https://searchdatacenter.techtarget.com/news/1163576/2006-Time-to-plan-your-companys-escape-from-Microsoft>. [Accessed: 02-Feb-2019].
- [17] R. A. Hexsel, "Propostas de ações de governo para incentivar o uso de software livre," *Relatório Téc. Dep. Informática UFPR*, vol. 4, p. 2002, 2002.
- [18] H. R. Varian and C. Shapiro, "Linux adoption in the public sector: An economic analysis," *Manuscr. Univ. Calif. Berkeley*, 2003.
- [19] M. N. Garcia, S. M. B. dos Santos, R. da Silva Pereira, and G. B. Rossi, "Software livre em relação ao software proprietário: aspectos favoráveis e desfavoráveis percebidos por especialistas," *Gest. Reg.*, vol. 26, no. 78, 2011.
- [20] V. Sabino and F. Kon, "Licenças de software livre história e características," 2009, Centro de Competência em Software Livre Departamento de Ciência da Computação Instituto de Matemática e Estatística, Universidade de São Paulo, 2009.
- [21] S. Dutta, B. Lanvin, and F. Paua, *The global information technology report 2002-2003: Readiness for the networked world*. Oxford University Press, USA, 2003.
- [22] K. Wong and P. Sayo, *Free/Open Source Software-A general introduction*. UNDP-APDIP, Kuala Lumpur, MY, 2004.
- [23] D. A. Wheeler, *Why open source software/free software (OSS/FS, FLOSS, or FOSS)? Look at the numbers*. 2007.
- [24] R. M. Stallman, "The Free Software Definition," 1985.
- [25] "VOSviewer - Visualizing scientific landscapes," *VOSviewer*. [Online]. Available: <http://www.vosviewer.com/>. [Accessed: 05-Feb-2019].
- [26] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, Aug. 2010.
- [27] N. J. Van Eck and L. Waltman, "VOS: A new method for visualizing similarities between objects," in *Advances in data analysis*, Springer, 2007, pp. 299–306.
- [28] C. Mascarenhas, J. J. Ferreira, and C. Marques, "University-industry cooperation: A systematic literature review and research agenda," *Sci. Public Policy*, vol. 45, no. 5, pp. 708–718, Jan. 2018.
- [29] P. Wang, F. Zhu, H. Song, and J. Hou, "A Bibliometric Retrospective of the Journal Eurasia Journal of Mathematics, Science and Technology Education between 2012 and 2017," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 14, no. 3, pp. 765–775, Nov. 2017.
- [30] F. García-Lillo, E. Claver-Cortés, B. Marco-Lajara, and M. Úbeda-García, "Mapping the Intellectual Structure of Research on 'Born Global' Firms and INVs: A Citation/Co-citation Analysis," *Manag. Int. Rev.*, vol. 57, no. 4, pp. 631–652, Aug. 2017.
- [31] C. J. Costa, «Testing usability of ERP open source systems», apresentado na Proceedings of the Workshop on Open Source and Design of Communication, 2010, pp. 25–30.
- [32] C. Costa, «ERP Open Source», *Information Technology, Organizations and Teams*, C.J Costa, Ed. press. itml. org, pp. 159–170, 2007.
- [33] M. Aparicio e C. J. Costa, «Macroeconomics leverage trough open source», em Proceedings of Open Source and Design of Communication, New York, NY, USA, 2012, pp. 19–24..
- [34] M. Delgado, M. Aparicio, e C. Costa, «Using open source for forensic purposes», em Proceedings of the Workshop on Open Source and Design of Communication, New York, NY, USA, 2012, pp. 31–37.
- [35] C. Costa e M. Aparicio, «Organizational Tools in the Web: ERP Open Source», 2006, vol. In Proceedings of the IADIS International Conference on WWW/Internet., pp. 401–408.
- [36] M. Batista, C. J. Costa, e M. Aparicio, «ERP OS localization framework», apresentado na Proceedings of the Workshop on Open Source and Design of Communication, 2013, pp. 1–8.
- [37] C. J. Costa, M. Aparicio, e J. P. Figueiredo, «Health Portal: An Alternative Using Open Source Technology», *International Journal of Web Portals*, vol. 4, n. 4, pp. 1–18, 2012.
- [38] J. T. Aparicio e C. J. Costa, «A virtual robot solution to support programming learning an open source approach», em 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), 2018, pp. 1–6.