



HAL
open science

Subverting or preserving the institution: Competing IT firm and foundation discourses about open source

Laure Muselli, Mathieu O'Neil, Fred Pailler, Stefano Zacchioli

► To cite this version:

Laure Muselli, Mathieu O'Neil, Fred Pailler, Stefano Zacchioli. Subverting or preserving the institution: Competing IT firm and foundation discourses about open source. *New Media and Society*, 2024, 10.1177/14614448231222249 . hal-04416993

HAL Id: hal-04416993

<https://hal.science/hal-04416993>

Submitted on 25 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Article

Subverting or preserving the institution: Competing IT firm and foundation discourses about open source

new media & society

1–23

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: [10.1177/14614448231222249](https://doi.org/10.1177/14614448231222249)

journals.sagepub.com/home/nms



Laure Muselli

i3, Télécom Paris, Institut Polytechnique de Paris, France

Mathieu O’Neil 

News and Media Research Centre, University of Canberra, Australia

Fred Pailler

C²DH, University of Luxembourg, Luxembourg

Stefano Zacchioli 

LTCl, Télécom Paris, Institut Polytechnique de Paris, France

Abstract

The data economy depends on digital infrastructure produced in self-managed projects and communities. To understand how information technology (IT) firms communicate to a volunteer workforce, we examine IT firm and foundation employee discourses about open source. We posit that organizations employ rhetorical strategies to advocate for or resist changing the meaning of this institution. Our analysis of discourses collected at three open source professional conferences in 2019 is complemented by computational methods, which generate semantic clusters from presentation summaries. In terms of defining digital infrastructure, business models, and the firm-community relationship, we find a clear division between the discourses of large firm and consortia foundation employees, on one hand, and small firm and non-profit foundation employees, on the other. These divisions reflect these entities’ roles in the data economy and levels of

Corresponding author:

Mathieu O’Neil, 9CI (Building 9), News and Media Research Centre, Faculty of Arts and Design, University of Canberra, ACT 2617, Australia.

Email: mathieu.oneil@canberra.edu.au

concern about predatory “Big Tech” practices, which transform common goods to be shared into proprietary assets to be sold.

Keywords

Digital infrastructure, institutions, open source software, organizational communication, voluntary labor, volunteers

Introduction

Free, libre and open source software (FLOSS) emerged in the 1990s as a social movement in which participants mobilized against the practices of proprietary software publishers. This social movement comprised new forms of organization, such as self-governed projects coordinating the work of remote developers, and new types of resources whose open status was protected by “copyleft” licenses such as the General Public License (GPL). The free software movement also challenged the values and discourses conveyed by proprietary software vendors and disseminated new values and discourses, centered around the preservation of software commons. The purpose of the free software movement was to promote “the global use of free software with the goal of eliminating proprietary software” (Elliott and Scacchi, 2008: 5). A famous article by Eric Raymond (1999) announced a significant change. The opposition it drew between a top-down and well-organized model of development (“the cathedral”), and a messy, rambunctious marketplace of ideas (“the bazaar”) was not outlined in terms of the ethical significance of freely sharing code: what mattered was that massive numbers of programmers could openly collaborate (Couture, 2021). As is well known, this re-imagining of free software in terms of technological efficiency was given the more business-friendly name “open source,” and from then on a “commercial form” of FLOSS progressively grew closer to the market sphere, while a “community form” stayed true to the movement’s original values (Coris, 2006).

In the wake of IBM’s investment in Linux in the late 1990s¹ and of Google’s integration of Linux into its Android OS in 2008, FLOSS, though still largely produced by distributed teams of volunteers, was integrated into the IT (Information Technology) firm ecosystem. Dominant firms embraced open source discourses, values, and practices. Microsoft, which once viewed Linux as a serious threat to its business model, exemplifies this trend.² IT firms invested in open source for several reasons: engaging with FLOSS is a source of legitimacy and contributes to their Human Resources strategy, as in the highly competitive skilled IT professional job market it attracts prospective employees; thanks to coordination by foundations, firms can share development costs with other firms; finally firm investment may also lead to shaping the governance and technical orientation of software projects and products.

Relationships with other firms follows a “coopetition” model (Nguyen-Duc et al., 2019) as firms cooperate to develop open digital infrastructures to reduce costs, but compete when it comes to leveraging data. Following hardware in the 1980s, and software in the 2000s, data are now a key strategic asset. The data market is controlled by Web giants

including Meta, Google, and Amazon, or by IT firms such as Microsoft, which has acquired Skype, LinkedIn, and GitHub. So-called cloud computing was introduced in 2006 (Amazon, 2006), and the data market is now dependent on distributed servers and business solutions enabling high-volume data storage and analysis. These services are provided as closed solutions, contradicting the open source social movement's values of sharing and transparency. Firm discourses about open source software will doubtless contain manifestations of this contradiction, but what form do these discourses take? Who says what, and where?

This article examines the discourses mobilized by IT firm and foundation employees when discussing open source software during three trade conferences. The first purpose of firm discourses is practical. Like any business, IT firms need to communicate their requirements to their workforce, but these firms must contend with a unique predicament: they do not only rely on paid employees, who are beholden to professional standards by contracts. In FLOSS projects volunteers play important roles, which raises operational and communication challenges. Volunteers may decide from one day to the next that they no longer wish to contribute to a project and leave. Other volunteers may step in to continue their work but (absent a handover) this could prove disruptive. This unusual characteristic of IT firms explains why open source conferences operate as privileged sites for firm-volunteer communication. We surmise that there are significant differences between conference populations, which are reflected in the types of discourses which emerge from these locations, and our first two research questions are as follows:

RQ1. To what extent does the population of attendees vary between different open source conferences?

RQ2. To what extent do IT firm and foundation discourses about FLOSS vary between different open source conferences?

In addition to the practical purpose of communicating to a volunteer workforce, discourses produced by IT firm and foundation employees during open source professional conferences can also be understood as a means to advance firm and foundation conceptions of what FLOSS is, and of how firms or foundations should relate to it. We understand these rhetorical strategies as the institutional use of language in a professional context (Suddaby and Greenwood, 2005): these discourses seek to shape the perceptions held by software developers, other IT firms, and industrial clients who rely on IT products regarding which elements of the open source "institution" (e.g. system of norms, beliefs, attitudes, and practices) are legitimate. From a neo-institutionalist perspective, the attempt to shape perceptions is defined as "institutional work," and our third and fourth research questions focus on this aspect:

RQ3. To what extent do IT firm and foundation employee discourses about and toward FLOSS projects vary and constitute distinct forms of institutional work?

RQ4. What form do these rhetorical strategies take?

To address these practical and institutional dimensions, we conduct a qualitative analysis of 155 presentations at three open source conferences (data were acquired in person in two sites and by video analysis in the third) as well as a quantitative semantic-clustering analysis of 1087 presentation summaries collected from two conference websites. We focus on presentations by employees of IT firms and foundations. In the United States, there are two main types of software foundations: 501(c)(3) such as the GNOME or Apache Foundations are non-profits charities committed to the public good; 501(c)(6) such as the Linux Foundation are commercial, representing the interests of industrial consortia. To simplify, we refer to 501(c)(3) as “non-profit” and to 501(c)(6) as “consortia.”

The article is organized as follows: we first review the literature on the use of open source software by firms and on communication between firms and projects, then outline our neo-institutionalist theoretical framework. We next present our data collection and analytical methods. We separate our qualitative and quantitative findings in two parts, addressing first firm employees’ presence at conferences, then firm and foundation employee discourses. We find a clear division between how employees of large IT firms (and consortia foundations) and small IT firms (and non-profit foundations) communicate about open source. We then discuss the implications of our findings and offer concluding remarks.

Background: firm-project communication

Firms and open source software

The literature about this topic is extensive, so we summarize three relevant dimensions: integrating factors, modalities of firm engagement, and firms and open source conferences.

A significant internal integrating factor is that the FLOSS movement’s opposition to proprietary software and intellectual property rights is “recursive” (Kelty, 2008): it primarily aims to consolidate and grow the material conditions for the survival of the software infrastructure. To neutralize the disruptive potential of free software, firms provided access to code through “open” licenses, thereby ensuring participants would continue to build their environments, irrespective of whether contributions had commercial or communal purposes (O’Neil et al., 2021b). External integrating factors include mediating entities such as the Linux Foundation, online platforms such as GitHub, which enables large-scale collaborative development, and Stack Overflow, which has become a key resource for mentoring and advice.

Firm involvement in open source has been examined in terms of its impact on the labor market (Dafermos and Söderberg, 2009), of ownership of projects (Gonzalez-Barahona and Robles, 2013), of which project areas firm employees most contribute to (Butler et al., 2019) and of how dominant firms have created non-overlapping “contribution territories” on GitHub (O’Neil et al., 2022). None of these studies address firm communications during conferences.

Coleman (2010) conducted a classic study of “the hacker conference” as a site where members of a FLOSS project ritually celebrate the “re-enchantment of the quotidian,”

but firms did not feature prominently in this narrative. Munir et al. (2018) found that Sony Mobile's Tools department discussed large features directly with the community at "hackathons," spaces where firms engage with participants to progress code development. In-person meetings were preferred over electronic options such as mailing lists and issue trackers because time differences resulted in discussion lags (Munir et al., 2018). Apart from hackathons, to the best of our knowledge, there are no other scientific studies of firm discourses during open source events.

Firm discourses and institutional work

The sociological branch of neo-institutionalism originates in work by Meyer and Rowan (1977), Meyer and Scott (1983) and Di Maggio and Powell (1983). For neo-institutionalists, actions and decisions taken within organizations are guided by the search for social legitimacy. Legitimacy is "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within 'institutions'" (Scott, 1995). Institutions are, socially constructed assemblages of values and beliefs, formal, and informal rules, "symbolic systems through which [people] categorize their activity and infuse it with meaning" (Friedland and Alford, 1991: 232).

Institutions not only provide a framework for the beliefs and actions of actors, but also feature contradictions likely to be challenged by these actors, and thus, lead to institutional change (Emirbayer and Mische, 1998). The notions of "institutional entrepreneurs" actively engaged in transforming institutions (Fligstein, 1997; Maguire et al., 2004; Oliver, 1991) and of "institutional work," defined as "the broad category of purposive action aimed at creating, maintaining, and disrupting institutions and businesses" (Lawrence and Suddaby, 2006: 216) suggest that institutional change is driven by strategic interests (Lawrence et al., 2010). Sahlin and Wedlin (2008) highlight the agentic, collective, and contested character of these transformation processes. Actors provide different "common meanings and identities" to justify their strategic actions (Fligstein, 1997: 398), using contradictions in institutions to induce transformations that further their interests.

To assert the legitimacy of their proposals for change and to secure shifts in institutions, individuals, and organizations may make a "strategic use of persuasive language or rhetoric" (Suddaby and Greenwood, 2005). Organizational discursive legitimation strategies have been analyzed during unit shutdowns (Vaara and Tienari, 2008), cross-border mergers (Vaara and Tienari, 2011) or during the introduction of a contested issue such as shareholder value (Meyer and Höllerer, 2010). In this article, we analyze firm rhetorical strategies as a type of institutional work seeking to affect the *open source institution*. We define this institution as a set of principles, values and practices which includes, among other features: decentralizing goal setting and execution to networks of individuals, collective decision-making, and the use of open licenses to guarantee transparency, sharing, and non-appropriation. What types of discourse are being produced within the open source industrial ecosystem? What types of actors are involved, and what do their discourses tell us about their institutional work? And finally, what impact is this institutional work seeking to have on the principles, values, and practices of the open source institution?

Methodology

Selection of sites

We analyze talks and presentations at three European open source conferences held in 2019: the Free and Open Source Software Developers' Meeting (FOSDEM, Brussels, Belgium, 2–3 February);³ the Open Source Summit Europe (OSSE, Lyon, France, 28–30 October);⁴ and the Paris Open Source Summit (POSS, Paris, France, 10–11 December).⁵ The three conferences share common features: large IT companies were the main sponsors at two events.⁶ Some speakers, including well-known technology “evangelists” have attended for several years, and it is not uncommon for developers and firm employees to participate in all three events. Yet these conferences also represent distinct professional worldviews: FOSDEM19, an international community event held on the campus of the Université Libre de Bruxelles, was developer-centric and managed by volunteers. OSSE 2019, an international corporate event, was organized by the Linux Foundation, the world's leading consortia foundation (Biddle, 2019), to promote commercially-oriented open source technology. Finally POSS, a European/French business event held in a moderately size conference center outside Paris, mainly featured smaller European IT firms.

Our original intention was for teams of three researchers to attend OSSE19, POSS19, and FOSDEM20 in person, recording presentations and taking field notes. However, this plan was adversely affected by the national strikes against pension reform that paralyzed France between December 2019 and February 2020. We were unable to attend FOSDEM20 (held 1–2 February 2020) due to the unsettled transport situation.

Instead, we elected to collect and analyze presentations from the FOSDEM19 video archive of presentations. This mode of data gathering would not have been possible in the case of OSSE19 and POSS19: while keynotes are typically broadcast live online in all three sites, this was not the case for other kinds of presentations. Out of 350 keynotes and presentations at OSSE19 in Lyon, 80 were archived on YouTube; the number was even lower at POSS19 in Paris, which only documented 30 presentations out of 250. Conversely FOSDEM is designed to allow talks to be accessed remotely and asynchronously, so all 740 speeches and talks from the 2019 edition were downloadable from the conference website. We now outline how we built our two conference discourse datasets.

Qualitative data. Dataset 1: 155 presentations in three conferences

After careful consideration of the conference programs, we selected keynotes, talks and workshops according to our main interest, the participation of firms in open source projects and the firm-volunteer community relationship, leading to a focus on management and organizational, rather than technical, presentations. On-site in Lyon and Paris, in addition to attending and recording presentations, we took pictures of presenters' slides, we identified the topics covered and took notes about how they were addressed. We also collected some ethnographic data by documenting different kinds of behavior during and outside the talks. For example, we counted the number of attendees; we evaluated the capacity of rooms; and we recorded significant interactions between speakers and audiences, such as interjections. We noted our observations by following a predetermined

Table 1. Qualitative and quantitative data collection at three open source conferences.

	OSS, Lyon	POS, Paris	FOS, Brussels
Keynotes/talks/workshops	350	400	740
Speaker/audience interactions	Yes	Yes	Yes
Collection method	On-site recording	On-site recording	Online archives
Presentations collected by	Research team	Research team	Event staff
Collection format	Audio	Audio	Video
Annotations + verbatims	55	52	48
Additional interviews	Booths + F2F	Q&A (author talk)	No
Summaries scraped from website	349	Not available	738

template (categories included occupancy rate, proportion of women present, explicit mentions of the firm-project relationship, and themes covered). We spent time at firm booths, and a team member gave a presentation in Paris, enabling us to collect professional insights during the Q&A. Table 1 summarizes the collected data. Room capacity indicated which streams or speakers were deemed important by organizers, while speaker/audience interactions served to identify overt and latent conflicts and allegiances. We organized our notes and recordings in a database: rows included speakers, their profession and organizational affiliation, as well as notes about the presentation and/or audio/video recording.

We sorted the speakers according to the category of organization to which they belonged. To identify firms, we consulted online resources such as wikipedia.org and crunchbase.com, as well as the homepages and Twitter, GitHub, and Mastodon profiles of participants. We also queried firm websites and national or federal firm databases.⁷

In accordance with usual practice in qualitative data analysis in organizational studies (Gioia et al., 1994, 2012), as well as in neo-institutional studies (Gawer and Phillips, 2013), we analyzed these data using open coding. *Concepts* emerged as first-order codes (e.g. “documentation”), which we systematically cross-checked with the category of organizations that mentioned them. This enabled connecting discourse first-order concepts and employer types. Next, an iterative process of sorting and grouping generated a reduced number of second-order *themes* (e.g. “professionalization”). Second-order themes were finally sorted into three aggregated *dimensions*.

Quantitative data. Dataset 2: 1087 presentation summaries in two conferences

We used the *rvest* R web scraping package to collect professional affiliations and presentation summaries of the OSSE19 and FOSDEM19 conferences. To categorize employer types, we used the same query methods as for Dataset 1. We created another database in which rows included the speaker’s name, their organization, their position, the title of their presentation, the type of presentation (keynote, workshop, talk), and its summary.

Table 2. Employer categories and distribution in two conferences (D2: quantitative dataset) and three conferences (D1: qualitative dataset).

		D2, N.1087				D1, N.155	
		N. FOS	N. OSS	% FOS	% OSS	N.	%
IT-big	Mainly “Big Tech” giants, e.g. Microsoft, Google, Amazon	219	171	29.7	49.0	45	29.0
bigUSR	Firms using OS software in their products, e.g. car manufacturers	14	22	1.9	6.3	13	8.4
DEV	Independent developers (or developers who chose not to identify as firm employees)	95	16	12.9	4.6	6	3.9
IT-small	Less than 500 employees	231	79	31.3	22.6	43	27.7
EXP	Experts, academics, lawyers, and hacktivists	78	28	10.6	8.0	17	11.0
FND3	501(c)(3) foundations (non-profit), e.g. GNOME Foundation	44	6	6.0	1.7	19	12.3
FND6	501(c)(6) foundations (industrial consortia), e.g. Linux Foundation	7	24	0.9	6.9	11	7.1
Other		50	3	6.8	0.9	1	0.6
Total		738	349	100	100	155	100

The analysis of conference summaries was carried out using the iRamuteQ software, based on the Reinert method (Beaudouin, 2016; Van Meter et al., 1991). Starting from a text corpus, the software performs a detailed analysis of its vocabulary and constitutes a dictionary of words and their inflections (a process referred to in linguistics as *lemmatization*), together with their frequency. By successive splits, it then divides the text into homogeneous segments containing a sufficient number of words, and classifies these segments on the basis of co-occurrence and distribution patterns. This method enables the extraction of clusters, constituted by the most statistically significant words and sentences, representing the dominant ideas and themes or “lexical worlds” of the corpus (Schonhardt-Bailey et al., 2012).

Findings: employer categories

Employer categories: quantitative dataset

Employer categories show which type of actor engages in institutional work. Table 2 (D2, N.1087) presents employer categories and their distribution in FOSDEM19 (FOS) and OSSE19 (OSS). The combined number of large and small firm employees is by far the largest component, in roughly similar proportions across events: 61% in FOSDEM19 and 71.6% in OSSE19. However when it comes to the split between large and small firms within each event, there are clear variations.

Presenters at the community-oriented FOSDEM19, which took place in a public institution, comprised an equal number of large and small IT firm employees (around 30%).

In contrast presenters at the Linux Foundation's OSSE19 were mainly large IT firm employees (49%) with small IT firm employees only numbering 22.6%. This accords with the event's location, Lyon's largest conference center, and with the Linux Foundation's mission of acting as an interface between industrial and communal worlds.

The distribution of other types of speakers also differs between the conferences. In the case of independent developers (DEVs), 12.9% of FOSDEM19 speakers defined themselves as such, while OSSE19 only comprised 4.6%. There were twice as many presentations at FOSDEM19 (N:738) than OSSE19 (N:349) so the difference in the number of DEVs is significant: 95 individuals at FOSDEM19, against 16 at OSSE19. It is likely that this disparity is due to the greater physical presence of freelance developers at FOSDEM19, but also to the fact that some large firm employees may attend grassroots conferences under their own name, or as project representatives, not as employees. Similar differences concern employees of firms which purchase open source products (bigUSR)—1.9% at FOSDEM 19, 6.3% at OSSE19—as well as foundation representatives. FOSDEM19 mainly featured presentations by employees or representatives of non-profit foundations: 6% (N:44) with only 0.9% (N:7) representatives of consortia foundations. OSSE19 mainly hosted representatives of consortia foundations: 6.9% (N:24), with only 1.7% (N:6) for non-profits.

Employer categories: qualitative dataset

Table 2 (D1, N.155) provides a breakdown of the employers of the authors of the presentations we collected qualitatively (D1). Comparing the distribution of speakers in this smaller corpus with the larger dataset (D2) identified similarities, such as the proportion of small IT firms (IT-small, D1: 27.7%, D2: 29.1%) and experts (EXP, D1: 11%, D2: 9.7%). We also found differences: there are slightly less large-firm employees (IT-big, D1: 29%, D2: 35.3%) and independent developers (DEV, D1: 3.9%, D2: 9.7%) in the smaller dataset. In the case of large firms, this could be because of the higher representation at POSS of small and medium firms. As for developers, their smaller number in the qualitative dataset—just like the higher proportion of foundation employees or representatives (D1: 19.4%, D2: 7.6%)—stems from the fact that we mainly attended presentations addressing the regulation of open source during the qualitative phase: foundations are by definition involved in such discussions, while developers were more likely to present during technical sessions.

Findings: firm discourses

Firm discourses: quantitative dataset

We sought to identify the types of discourse conveyed by the different categories of actors, to understand the type of institutional work they engage in. Basic content analysis of D2 (1087 presentations in two conferences) determined that certain employee categories predominantly used specific terms: small and large IT firm employees were the most frequent users of “data,” “platform,” and “infrastructure.” Only small IT firm employees referred to “funding,” and large IT firm employees were much more likely to refer to the

“cloud” than other categories. Non-profit foundation representatives were more likely to use “privacy” and “community” whereas consortia foundation spokespersons were the top users of “documentation” and “safety.” These findings show that employee discourses accord with employer interests. We now present more detailed analysis using semantic cluster mapping.

Figure 1 shows a division between two lexical clusters. The red cluster (left) is used primarily by employees of non-profit and consortia foundations (31% of the total corpus). It contains terms relating to the social and practical dimensions of open source, starting with organizational and regulatory principles (“open source,” “project,” “software,” “license”). This lexicon includes ethical values (“collaboration,” “inclusion,” “diversity,” “commitment”). It also contains terms describing actors (“company,” “foundation,” “team,” “contributors”).

A second lexicon, on the right side (gray, and pink/gold clusters), relates to technical objects, services and operations (69% of the total corpus). Included terms, which relate to cloud computing infrastructure and its open source components, are almost exclusively used by people working for very large IT firms. The lexicon also describes “embedded computing”: the use of open source code in both industrial machinery and in IT-dependent consumer objects such as air conditioners, cars, connected toothbrushes, and so on. Finally, it contains terms used by small IT firm employees, independent developers and experts, which relate to a higher level in the open source technology and software stack, for example, databases and programming languages (“python,” “java,” “rust,” “javascript”) as well as algorithms, search engines, and other software enabling the manipulation of large databases (“tensorflow,” “lucene,” “spark”) and operations performed on data (“query,” “collect,” “semantics,” “visualization”).

Figure 2, which “zooms in” on Figure 1’s regulation lexicon, has three main semantic clusters. First, an *open source industry sub-lexicon* (top left, 34% of the regulation lexicon), used by speakers whose employers are large IT companies and consortia foundations. It could be described as paradigmatic, in that it contains terms that define key entities (“platform,” “foundation,” “consortium,” “client”), objects (“software,” “license,” “linux,” “certification”) and actions (“development,” “compliance,” “implementation,” “automate,” “embed,” “integration”) in the open source world. This *open source industry sub-lexicon* exemplifies the move toward professionalization, both as the hegemonic deployment of open source software throughout the IT industry, and as a “corporate” discourse now integrated by the FLOSS world, which had preserved its relative autonomy from commercial values and maintained its connection to communal values during the 1990s/2000s.

Next, a *community/diversity sub-lexicon* (bottom left, 44% of the regulation lexicon) contains elements describing work in firms and projects. It is mainly used by small IT firm employees and independent developers. This sub-lexicon relates to individual self-fulfillment both in terms of success (“opportunity,” “foster,” “individual,” “hope”) and difficulties (“struggle,” “impostor syndrome,” “underrepresented,” “barrier”). This sub-lexicon also comprises management terms (“challenge,” “team,” “organizational,” “Open Source Program Office”), which overlap with the social justice vocabulary which characterizes diversity policies (“inclusive,” “represent,” “woman,” “mentor”). This sub-lexicon is more concerned with the regulation of systemic power relations in the

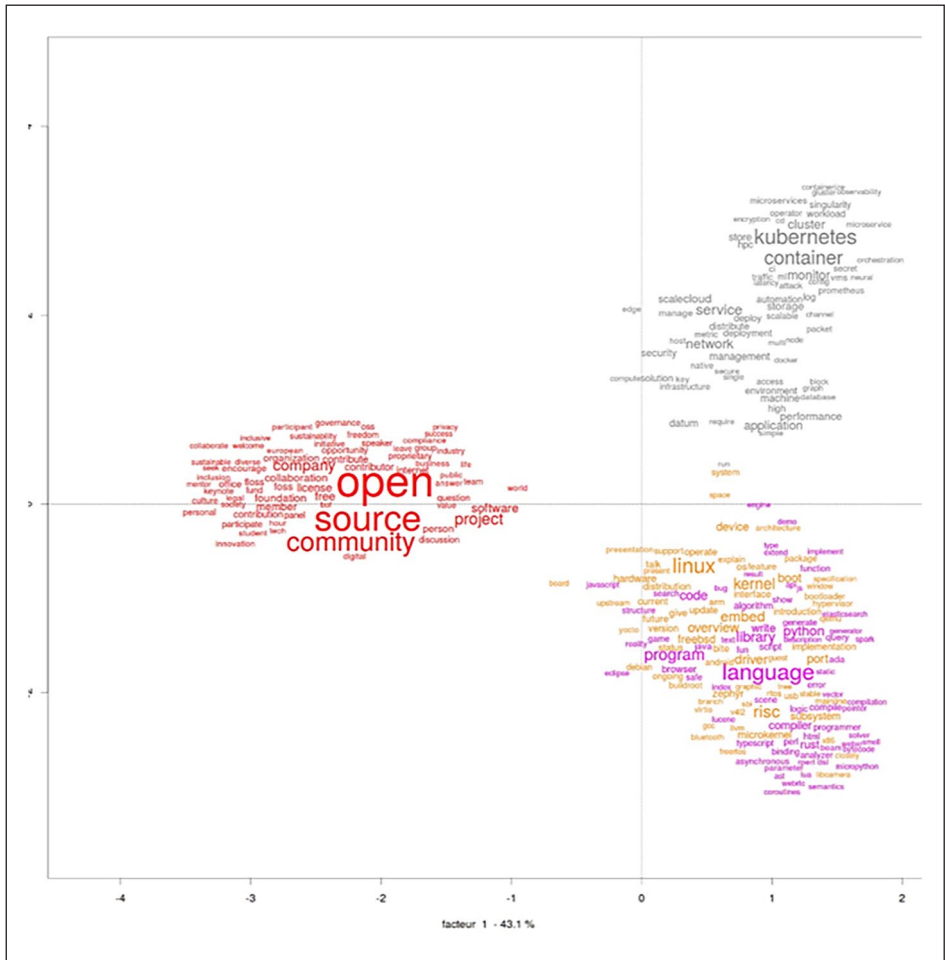


Figure 1. Two major lexicons: regulation of OS (red, left of vertical axis), and technical objects and operations (left, gray, orange, purple).

corporate workplace than in open source projects. “Community” operates as a catch-all term, being equally applicable to designate corporate teams, project teams, and communities of users.

A third *control/privacy sub-lexicon* (right, 22% of the regulation lexicon) addresses issues such as Internet control, privacy, and the regulation of technological relations of power in society. This sub-lexicon is almost exclusively present in summaries written by people employed in non-profit foundations and small IT firms which produce decentralized infrastructure and services, competing with the centralized infrastructure and data silos of large IT companies. It includes brand or service provider names: services situated furthest from the center, at the right of the graph, are specifically related to privacy and Internet control issues. Legal-political terminology is prevalent (“dominate,” “regulation,”

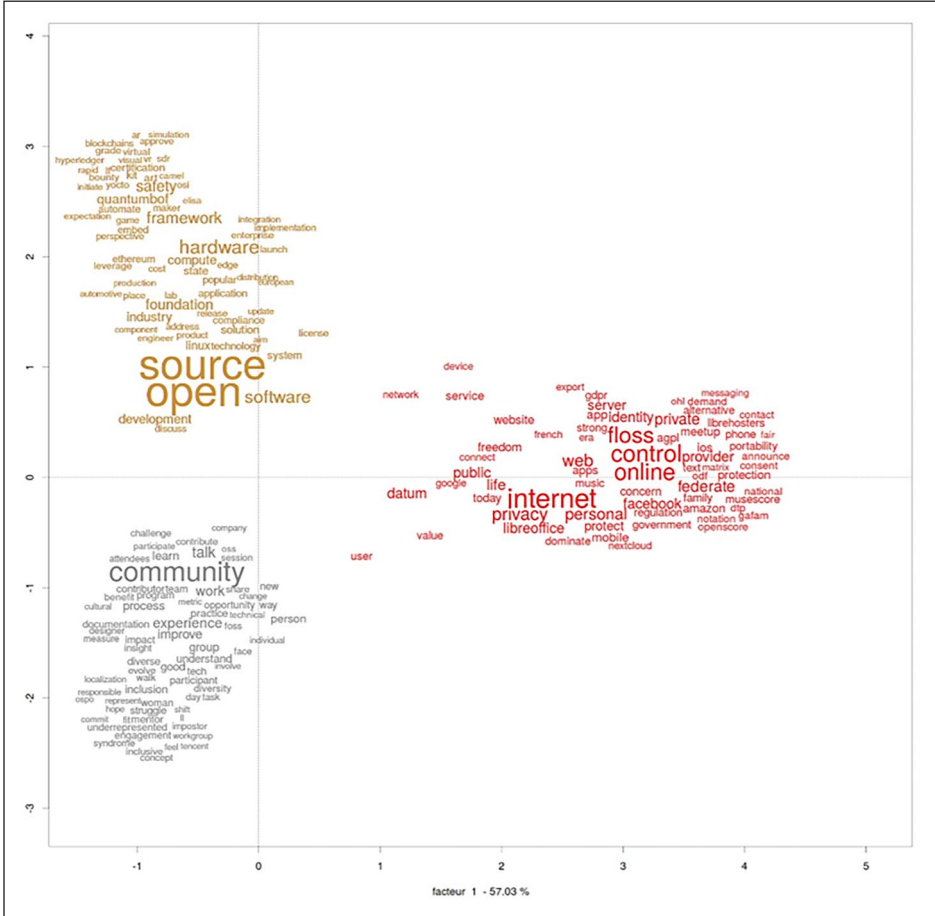


Figure 2. The OS/community lexical universe is composed of three sub-lexicons: (1) open source (brown, top left), (2) community/diversity (gray, bottom left), and (3) control/privacy (red, right of vertical axis).

“consent,” “GDPR,” “freedom,” “government”).⁸ Basic technical terminology near the center of the graph, such as “data,” “network,” and “user,” is also present, as well as vocabulary relating to personal data and their monetization (“federate,” “mobile,” “phone,” “server”). Finally, “FOSS / FLOSS” is used by non-profit foundations and small firms who criticize large IT firm monopolies and advocate for decentralized approaches. Part of the software mentioned in this group contains the term “Libre” (e.g. “libreoffice,” “librehosters”). In contrast “open source” appears in the abovementioned *open source industry sub-lexicon* which describes large IT firm involvement in projects.

Figure 2 shows that the *open source industry* and *community/diversity sub-lexicons* which appear on either side of the horizontal axis on the left side of the graph stand in opposition to each other: speakers use one of these sub-lexicons, rarely both. However,

the *open source industry* and *community/diversity sub-lexicons* enter into dialogue to some extent. They are aligned in relation to the vertical axis, share similar scores (between 0 and -2) and can be defined as forms of corporate discourse (the firm perspective on open source) and counter-discourse (individual career issues). In contrast, the *control/privacy sub-lexicon* on the right side, which obtains distinct scores (2-4), deals with the regulation of technology and data use, not internal firm dynamics.

In sum, our analysis of discourses in presentation summaries shows a significant link between these discourses and the presenters' employers. Technical discourses are used by diverse speakers, with the exception of the "cloud," solely used by large IT firm employees. Large corporations feature in open source regulatory discourses, either as the source of the professional and industrial definitions of open source terminology, or as the destination of critiques about project management and access to work. The distance and autonomy of the *control/privacy sub-lexicon* from the corporate *open source industry* and *community/diversity sub-lexicons* in Figure 2 suggests that the small IT firms and non-profit foundations which use it are adopting an oppositional position in regard to large IT firms.

Firm discourses: qualitative dataset

Our qualitative analyses confirm the existence of a division between the rhetorical strategies of large IT firms and consortia foundation employees, on one hand, and those of small IT firms and non-profit foundation employees, on the other. We sorted discourses into three *dimensions*, each containing second-order *themes*: business models, the definition of digital infrastructure, and firm-community relations.

Business models. The second-order themes of the business model dimension are revenue models, development cost reduction, software developer payment and the open source ecosystem's sustainability. According to large IT firm employees, there is no "open source business model," no means of generating revenue from the development of open source software per se. Open source represents the means to reduce the development costs of digital infrastructure. Accordingly large IT firm employees never discuss employees being paid for their contributions to open source projects. Such payments do not deserve mention, being offset by large IT firms' profitable data-storage and processing activities. In contrast, small IT firm employees, whose employers' activity is mainly centered on the production of open source software and services, frequently evoke how difficult it is to find sustainable sources of revenue. Several business models are mentioned, from Open Core, to services, to dual licensing; charging users remains a key challenge. This issue takes on added urgency as cloud computing service models such as Software as a Service (SaaS) facilitate the appropriation of open source code by third parties such as Amazon, denying revenue from firms which produce open source.⁹ The sustainability of open source and the survival of its ecosystem are a persistent theme in the small IT firm employee discourse on business models, leading to calls for "resistance" to the hegemony of a handful of large IT firms. The issue of software developer payment is explicitly raised within small IT firm employee discourses. These firms work with projects

whose survival depends on paying core software developers, after which a community of volunteer developers augments this work.

Digital infrastructure. Digital infrastructure-related second-order themes are technical versus political understandings of the data market, centralization and neutrality versus decentralization and independence, and the professionalization of open source. Small and large IT firm employees advance sharply divergent positions in respect to the data market: large IT firm employees advocate for the centralized control of data; small IT firm employees and non-profit foundation spokespersons critique centralization. For large IT firm employees and consortia foundation spokespersons, the key aim is to develop technically neutral infrastructure. In this discourse, “foundations” are open, agnostic cocoons where the rules governing collaboration between IT end-users and suppliers can be elaborated; they help to create open technical standards and foster innovation and growth by pooling the costs of developing underlying infrastructure; their actions will enhance the confidence of large user firms in open source projects. Frequent references are also made to concepts connected to the professionalization theme, which aim to boost commercial users’ acceptance of open source technologies. These concepts encompass improvements in software performance, quality, and safety (requiring better documentation), as well as the importance for projects of respecting ethical standards—for example, diversity when it comes to recruiting and managing a volunteer workforce. Tellingly, the question of data and of its market value was wholly absent from the discourses of large IT firm employees.

Conversely, the political economy of the data market was addressed by some small IT firms employees and non-profit foundations spokespersons, with the concept of “data capitalism” frequently mentioned. This critical discourse questions the hegemonic and monopolistic management of data on centralized infrastructure controlled by large IT firms. Ethical issues for small IT firm and non-profit foundation employees include equality of access, data protection, the sustainability of open source, and democratic control: technology must be efficient, but it must serve independence. Data must not be controlled exclusively by “Big Tech” firms so there must be interoperability between platforms, servers, languages, and clients (e.g. only representatives of entities such as the Mozilla Foundation discuss the Web in terms of technological interoperability and access to content). In this discourse, non-profit foundations play a key role: they protect projects, and prevent large firms from imposing models that will reinforce their dominant position in the data market. This critical discourse is often accompanied by support for decentralized or federated alternatives, implemented in FLOSS products such as Matrix, Mastodon, or Nextcloud,¹⁰ and promoted by activist collectives such as Framasoft, Riseup, and CHATONS.¹¹

The firm-community relationship. A *project* is an initiative to collaboratively develop a software solution, which may originate with volunteers, a foundation, or be sponsored by a firm. Projects usually have membership and governance rules. A *community* is an informal group of developers that share the project’s goals and values and hence support it. Our analysis revealed striking differences in the way firm and foundation employees consider relationships between firms and communities. Second-order themes in this

dimension are: diversity and breadth of the community, community management, licensing, and values. For large IT firm/consortia foundation employees, “project” and “community” have equivalent meanings. The firm-community relationship is mutually beneficial: there is a single “open source community,” which firms are naturally part of. Non-firm contributors are employees of competitors and volunteers: these cooperative relationships need to be managed. “Community management” means ensuring professional best practice is observed, so behavioral guidelines, clear internal project governance, and mentoring are important. In the context of obtaining internal firm buy-in for open source involvement, there are frequent mentions of the creation of Open Source Program Offices (OSPOs) within firms: OSPOs advocate for open source solutions and familiarize firm employees with open source culture, for example, by convincing managers to let their team members work on open source projects, or by sending top firm contributors to open source conferences. Open source values and licenses are not central themes in large IT firm employee discourses. Issues of cost and remuneration associated with managing the community are also absent.

In direct contrast, for small IT firm and non-profit foundation employees, “project” and “community” are clearly distinguished. Presenters do not refer to “the community,” but rather to distinct communities associated to specific projects, which are differentiated by their values. If the values conveyed by a project grow apart from those of its community, these communities can dissolve, or migrate to another project. Small IT firm employees raise the question of how meeting the costs of development (e.g. paying core developers) can be articulated with community values, and with a compatible license. Small IT firm/non-profit foundation employee community management discourses are much less characterized by definitions of “best practice” than by the need to respect “open source principles.” In addition, the understanding that these employees have of a community is less extensive than that of their large IT firm counterparts: community members only include developers, not contributors in charge of non-coding activities. Given documentation work and other tasks seen as less technical in FLOSS are more likely to be performed by women (Nafus, 2012; Terrell et al., 2017), this restrictive understanding of who “belongs” in projects may reflect the sexism which has long permeated FLOSS, stemming from these project’s patriarchal roots (see Couture, 2021; O’Neil et al., 2021a).¹²

Discussion

A contested open source institution

We found clear differences in the attendance (RQ1) and types of discourses produced in our three collection sites (RQ2): a number of issues were addressed in all three events (e.g. compliance, security, governance, infrastructure, languages, community), but some conferences featured little discourse about privacy issues in data processing (OSSE19), or diversity management (POSS19). Firm and foundation employee discourses at OSSE19 were not critical of large IT firms, whereas at POSS19 and FOSDEM19 critiques emanated from small IT firms and non-profit foundation employees. Some critiques were couched in economic terms, for example

“Big Tech” domination prevents small IT firms from maintaining a sustainable business model. Others were political, such as the rejection of capitalism, or of specific aspects of capitalism, for example surveillance or the monopolistic exploitation of personal data.

Our quantitative analysis of the discourses produced by IT firm and foundation employees during open source conferences identified lexical clusters of keywords, which aggregate into a coherent organizational rhetoric opposing two distinct positions; this opposition was confirmed by our qualitative analysis (RQ3). Large IT firm employee discourses challenge the open source institution from the inside, whereas small IT firm employees attempt to preserve it. Lawrence and Suddaby (2006) defined these types of institutional work as “disruption-oriented” and “maintenance-oriented.” In the present case, institutional work stems from the position of firms in the data market. Large IT firms’ accumulation and circulation of proprietary data contradict the open source ethos of sharing and transparency. However, these business goals, which underpin the discourses of large IT firm employees, are never explicitly mentioned. “Quantified Self” (self-tracking) platforms’ deliberate use of obfuscation and ambiguity when it comes to the intermingling of community and commercial values has been identified as the means to allay the concerns of users and contributors (Barta and Neff, 2016). Similarly for the “Big IT” employees in our study, data appear to be a taboo topic. We now explore firm and foundation employees’ rhetorical strategies (RQ4).

Subversion rhetorical strategies: cheezification, neutralization, and inversion

How can common resources be preyed upon without raising the open source community’s ire? Therein lies the import of large firms “institutional work” which seeks to change the open source institution, via the strategic use of rhetoric. We identify three subversion rhetorical strategies aiming to undermine or “disrupt” (Lawrence and Suddaby, 2006) this institution. The first, “cheezification” (after the French “ringardisation”) appears when progressively, conference after conference, the community’s values are redefined, so that caring about open source values begins to appear somewhat quaint, almost ridiculous—an unhip relic of a bygone era.

To create common meanings and identities, large IT firm employees also employ a “neutralization technique” (Meyer and Höllerer, 2010). They use discourses that are technical, practical, and goal-oriented, rather than normative, so they do not challenge fundamental open source principles head-on. Instead, they tend to downplay the importance of licensing—suggesting it is much more important to produce useful and robust code—or continually advocate for the establishment of common platforms and standards. This use of technical or managerial efficiency discourses accords with organizational literature examinations of professionalization in non-profit organizations (Hwang and Powell, 2009).

Finally, the use of “openness” itself constitutes a discursive device. Lund and Zukerfeld (2020) argue that for-profit firms extolling the benefits of collaboration constitutes a new “openness ideology.” Zuboff (2019) noted surveillance capitalism’s Orwellian use of positive terms such as “community,” “collaboration,” and “openness”

to describe an inverted reality. The notion that terms such as “openness” can be used in multiple ways, some of which are contradictory, resonates with Schlagwein et al.’s (2017) examination of the multifaceted nature of “openness”: they describe how the modalities of a project’s openness can evolve over time, depending on the strategy and business model of the organizations which support it; this evolution affects the support of contributors, and the community’s very existence. The necessity for firms of always portraying projects as “open” is borne out by our analysis of the discourses produced by employees of large IT firms and of the consortia foundations which represent their interests: they emphasize the existence of a single “open source community,” harmonious and free of conflict.

Preservation rhetorical strategies: maintaining the open source institution

A mirror image is provided by some small IT firm employees and non-profit foundation spokespersons’ discourses, which portray a diversity of open source communities and projects, and a specific business model: open source code-related services are a direct source of revenue, since these firms do not rely on data capture and analysis (for large IT firms open source code is an instrument of data capture and analysis which is not intrinsically profitable). Small IT firms and non-profit foundation employees therefore adopt preservation strategies, with discourses aiming to “maintain” the open source institution. They focus on normative aspects such as respecting open source principles of sharing, openness, and transparency. They also champion technological alternatives. While large IT firms pursue digital infrastructural hegemony, smaller firms and non-profit foundations advance alternative decentralized infrastructural solutions proposed by diverse actors with strong ethical values. This critical stance should be considered when addressing the issue of open source sustainability. Though these smaller IT firms may appear puny compared with Microsoft or Google, their economic weight is not insignificant. Discounting these players and only focusing on large IT firms could be equivalent to ignoring open source players in the 1990s in favor of then-dominant IT firms.

In addition to the analysis of large IT firms’ attempts to subvert open source values and practices, our neo-institutional conceptual framework provides insights into the institutional evolution of FLOSS over time. After proprietary software companies failed to deny the legitimacy of the emergent ideas and practices mobilized by FLOSS developers in the late 1990s, open source practices and organizational models became institutionalized: large IT firms embraced FLOSS practices such as open licenses. This represented a shift from an informational capitalism relying on intellectual property to a new digital capitalism which integrates “openness” into its business model (Birkinbine, 2020; Broca, 2013; Lund and Zukerfeld, 2020).

Though our semantic clusters feature the classic opposition between “free software” and “open source” licenses (Coris, 2006), this opposition is less meaningful than in the past. The reason is that large IT firms have become an essential part of the open source industrial ecosystem (O’Neil et al., 2022), thereby enhancing their institutional work capacity and impact. The rhetoric of large IT firm employees now questions the legitimacy of the debate about whether software licenses, the cornerstone of the protection of digital commons, are more or less “open”: this debate is depicted as obsolete. This profoundly

transformative rhetoric must be understood within the context of the emergence of cloud-based technical models such as SaaS which invalidate FLOSS's foundational "four freedoms" to run, study, copy, and modify software, guaranteed by "copyleft" licenses such as the GPL;¹³ of large IT firms' new business objectives, which are no longer centered on the distribution of software, but on the exploitation of user data by developing infrastructure to collect, store and sell data (as well as to train AI systems); and of a range of related "predatory" corporate practices.¹⁴

Conclusion

This article contributes to the understanding of how IT firms communicate with a workforce comprising non-employees. It also shows that the struggle over the legitimate definition of the open source institution is far from over. Indeed, the current predatory appropriation of common-pool resources by large IT firms imbues this struggle with an existential quality. Our combination of discourse and semantic analyses has allowed us to demonstrate the existence of links between different open source conferences, categories of employers of presenters, and the discursive resources mobilized by employees. This brings to mind research into both social movements and institutions, which adds political contestation to the institutional change toolkit (Schneiberg and Lounsbury, 2008). Adopting a neo-institutional approach enabled us to analyze the ongoing process of transforming and contesting the values and practices associated with the open source institution, through the mobilization of cultural beliefs and discourses.

Our study highlighted two distinct types of institutional work: large IT firm employees favor technical and managerial discourses, while smaller firm employees use normative or values-based discourses. This is particularly evident when it comes to the role of foundations. For one side, consortia foundations provide an indispensable space and governance mechanism which facilitates inter-firm collaboration; for the other, non-profit foundations are the protectors of communal open source values. In general terms, large IT firm employee discourses have interlocked aims, centered around the attempt to legitimize new values—for example, equating open source development to a purely managerial process which needs to be refined. This effectively gets rid of the foundational notion that open source development has legal, moral, and political dimensions. Related aims might include enrolling new allies such as small firms or seeking to prevent the defection of historical actors.

Computational analyses of flows of firm employee contributions on GitHub open source code repositories focus on large IT firms (see O'Neil et al., 2022; Riehle et al., 2014). Our analysis complements these studies by providing a broader overview, which includes small IT firms' and non-profit foundations' definition of the open source institution, which is critical of large IT firms' business analytics. Furthermore, our comparison of large-scale data and more detailed qualitative analysis generated interesting findings. In contrast to our quantitative analysis, our qualitative analysis found that "data" were absent from large IT firm employee discourses. This can be explained by the fact that these employees' references to "data" took place in technical sessions, which we did not attend; its regulatory dimensions were seldom discussed. This shows that the creation of the new data market does not only occur through the production of discourse, but also

through its removal: “data” were never mentioned in a regulatory context to avoid drawing attention to the fact that changes in values and practices are needed to take control of this market. Since large IT firm employees strive to present their firms as members of the wider open source “community,” they cannot acknowledge how their employers intend to use this software during open source conferences.

Acknowledgements

We wish to thank the anonymous colleagues who reviewed this article for their valuable contributions to improving our argument.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Alfred P. Sloan Foundation and the Ford Foundation’s Critical Digital Infrastructure Fund (2019-2020).

ORCID iDs

Mathieu O’Neil  <https://orcid.org/0000-0002-3537-9036>

Stefano Zacchiroli  <https://orcid.org/0000-0002-4576-136X>

Notes

1. <https://www.ibm.com/opensource/story/>.
2. In 2018 Microsoft acquired the GitHub open source development platform for \$7.5 billion and joined the Open Innovation Network, a “defensive patent pool and community of patent non-aggression” aiming to protect Linux, signifying that the firm has renounced extracting value from violations of their patents that may occur in open source products (see <https://www.openinventionnetwork.com/>).
3. <https://archive.fosdem.org/2019/>.
4. <https://events19.linuxfoundation.org/events/open-source-summit-europe-2019/>.
5. https://www.ow2.org/view/Events/Paris_Open_Source_Summit_2019.
6. At the Paris Open Source Summit 2019 major sponsors were Inria, AlterWay, Smile, Microsoft, and Red Hat. The Open Source Summit Europe also had large IT firm sponsorship, either directly (Intel, Google, IBM/Red Hat), or through the Linux Foundation, itself funded by Microsoft, Meta, Google, Intel, Huawei, and Tencent. At FOSDEM, large IT firms (Google, Red Hat, AWS) provided limited sponsorship, as this event also drew on T-shirt sales, donations, and the free use of facilities granted by the Université Libre de Bruxelles. In our view sponsors influence discourses indirectly, by impacting who attends and speak at conferences (e.g. student FLOSS activists are more likely to attend and speak at FOSDEM than at the OSSE).
7. For example, societe.com, unternehmensregister.de, tecfindr.com, craft.co, droitbelge.be/infos-entreprises.asp, kbopub.economie.fgov.be, e-justice.europa.eu/content_find_a_company-489-en.do, globaldatabase.com, owler.com, and cbr.lv/en/company-search.
8. The General Data Protection Regulation (GDPR) regulates data protection and privacy in the European Union (EU), as well as the transfer of personal data outside the EU.
9. In a SaaS mode, a software program is never downloaded and executed onto the customers’ machine, but executed remotely on the provider’s hardware. A subscription to a service is bought, rather than a user licensing agreement being accepted. This creates a SaaS “loophole”

in the FLOSS principle, effectively negating the reciprocity obligation of copyleft licenses such as the GPL, as the service provider is no longer obliged to offer access to the code.

10. See <https://matrix.org/>; <https://joinmastodon.org/>; <https://nextcloud.com/>.
11. See <https://framasoftware.org/en/>; <https://riseup.net/>; <https://chatons.org/en/>.
12. That might be why firms, which are more diverse, and more familiar with diversity policies, promote diversity in open source projects. Speakers presenting about open source governance, on one hand, and about diversity and inclusion in projects, on the other, are not identical, and the latter tend to oppose the former, as shown in Figure 2: diversity policies operate both as a source of conflict within firms, and as a corporate attribute injected into projects.
13. See Note 9.
14. For example, Google, Amazon, Meta, Apple, and Microsoft (aka “GAMAM”) establish scientific collaborations with research institutions while seldom sharing intellectual property: 78.3% of Microsoft’s 17,405 scientific publications between 2014 and 2019 were co-authored with university researchers; during the same period Microsoft applied and was granted 76,109 patents, 0.2% of which were co-owned (Rikap and Lundvall, 2022). Firms such as Google and Meta also routinely use offers of R&D support to emerging developers or start-ups to gain knowledge of their inventions, then break off ties and launch similar innovations (Vasudevan, 2023).

References

- Amazon (2006) Announcing Amazon elastic compute cloud—beta, 4 August. Available at: [https://aws.amazon.com/about-aws/whats-new/2006/08/24/announcing-amazon-elastic-compute-cloud-amazon-ec2—beta/#:~:text=Announcing%20Amazon%20Elastic%20Compute%20Cloud%20\(Amazon%20EC2\)%20%2D%20beta,-Posted%20On%3A%20Aug&text=Amazon%20EC2’s%20simple%20web%20service,on%20Amazon’s%20proven%20computing%20environment](https://aws.amazon.com/about-aws/whats-new/2006/08/24/announcing-amazon-elastic-compute-cloud-amazon-ec2—beta/#:~:text=Announcing%20Amazon%20Elastic%20Compute%20Cloud%20(Amazon%20EC2)%20%2D%20beta,-Posted%20On%3A%20Aug&text=Amazon%20EC2’s%20simple%20web%20service,on%20Amazon’s%20proven%20computing%20environment)
- Barta K and Neff G (2016) Technologies for sharing: lessons from quantified self about the political economy of platforms. *Information, Communication & Society* 19(4): 518–531.
- Beaudouin V (2016) Statistical analysis of textual data: Benzécri and the French school of data analysis. *Glottometrics* 33: 56–72.
- Biddle CB (2019) Linux foundation is eating the world. *Journal of Open Law, Technology & Society* 11(1): 57–73.
- Birkinbine BJ (2020) *Incorporating the Digital Commons: Corporate Involvement in Free and Open Source Software*. London: University of Westminster Press.
- Broca S (2013) *Utopie du Logiciel Libre. Du bricolage technologique à la réinvention sociale*. Bernay-Neuvy-en-Champagne: Le passager clandestin.
- Butler S, Gamalielsson J, Lundell B, et al. (2019) On company contributions to community open source software projects. *IEEE Transactions on Software Engineering* 47(7): 1381–1401.
- Coleman G (2010) The hacker conference: a ritual condensation and celebration of a lifeworld. *Anthropological Quarterly* 83(1): 47–72.
- Coris M (2006) Chronique d’une absorption par la sphère marchande: Les sociétés de services en logiciels libres. *Gérer & Comprendre* 84: 12–24.
- Couture S (2021) Free & open source software. In: O’Neil M, Pentzold C and Toupin S (eds) *The Handbook of Peer Production*. Malden, MA: Wiley-Blackwell, pp. 155–168.
- Dafermos G and Söderberg J (2009) The hacker movement as a continuation of labour struggle. *Capital & Class* 33: 53–73.
- Di Maggio PJ and Powell WW (1983) The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review* 48(2): 147–160.

- Elliott MS and Scacchi W (2008) Mobilization of software developers: the free software movement. *Information Technology & People* 21(1): 4–33.
- Emirbayer M and Mische A (1998) What is agency? *American Journal of Sociology* 103: 962–1023.
- Fligstein N (1997) Social skill and institutional theory. *American Behavioral Scientist* 40(4): 397–405.
- Friedland R and Alford R (1991) Bringing society back in: symbols, practices, and institutional contradiction. In: Powell W and DiMaggio PJ (eds) *The New Institutionalism in Organizational Analysis*. Chicago, IL: University of Chicago Press, pp. 232–263.
- Gawer A and Phillips N (2013) Institutional work as logics shift: the case of Intel’s transformation to platform leader. *Organization Studies* 34(8): 1035–1071.
- Gioia DA, Corley KG and Hamilton AL (2012) Seeking qualitative rigor in inductive research: notes on the Gioia methodology. *Organizational Research Methods* 16(1): 15–31.
- Gioia DA, Thomas JB, Clark SM, et al. (1994) Symbolism and strategic change in academia: the dynamics of sensemaking and influence. *Organization Science* 5(3): 363–383.
- Gonzalez-Barahona JM and Robles G (2013) Trends in free, libre, open source software communities: from volunteers to companies. *Information Technology* 55(5): 173–180.
- Hwang H and Powell W (2009) The rationalization of charity: the influences of professionalism in the nonprofit sector. *Administrative Science Quarterly* 54(2): 268–298.
- Kelty CM (2008) *Two Bits: The Cultural Significance of Free Software*. Durham, NC: Duke University Press.
- Lawrence TB and Suddaby R (2006) Institutions and institutional work. In: Clegg SR, Hardy C, Lawrence TB, et al. (eds) *Sage Handbook of Organization Studies*. London: Sage, pp. 215–254.
- Lawrence TB, Suddaby R and Leca B (2010) Institutional work: refocusing institutional studies of organization. *Journal of Management Inquiry* 20(1): 52–58.
- Lund A and Zukerfeld M (2020) *Corporate Capitalism’s Use of Openness*. Basingstoke: Palgrave Macmillan.
- Maguire S, Hardy C and Lawrence TB (2004) Institutional entrepreneurship in emerging fields: HIV/AIDS treatment advocacy in Canada. *Academy of Management Journal* 47(5): 657–679.
- Meyer JW and Rowan B (1977) Institutionalized organizations: formal structure as myth and ceremony. *American Journal of Sociology* 83: 340–363.
- Meyer JW and Scott WR (1983) *Organizational Environments: Ritual and Rationality*. Beverly Hills, CA: Sage.
- Meyer R and Höllerer M (2010) Meaning structures in a contested issue field: a topographic map of shareholder value in Austria. *Academy of Management Journal* 53(6): 1241–1262.
- Munir H, Linåker J, Wnuk K, et al. (2018) Open innovation using open source tools: a case study at Sony mobile. *Empirical Software Engineering* 23: 186–223.
- Nafus D (2012) ‘Patches don’t have gender’: what is not open in open source software. *New Media & Society* 14(4): 669–683.
- Nguyen-Duc A, Cruzes D, Terje S, et al. (2019) Do software firms collaborate or compete? A model of cooperation in community-initiated OSS projects. *e-Informatica Software Engineering Journal* 13(1): 37–62.
- Oliver C (1991) Strategic responses to institutional processes. *Academy of Management Review* 16(1): 145–179.
- O’Neil M, Cai X, Muselli L, et al. (2022) Co-producing industrial public goods on GitHub: selective firm cooperation, volunteer-employee labour and participation inequality. *New Media & Society*. Epub ahead of print 27 April. DOI: 10.1177/14614448221090474.

- O'Neil M, Muselli L, Raissi M, et al. (2021a) 'Open Source has won and lost the war': justifying commercial-communal hybridisation in a FOSS project. *New Media & Society* 23(5): 1157–1180.
- O'Neil M, Toupin S and Pentzold C (2021b) The duality of peer production: infrastructure for the digital commons, free labor for free-riding firms. In: O'Neil M, Pentzold C and Toupin S (eds) *The Handbook of Peer Production*. Malden, MA: Wiley-Blackwell, pp. 3–17.
- Raymond E (1999) *The Cathedral and the Bazaar*. Sebastopol, CA: O'Reilly Media.
- Riehle, D., Riemer, P., Kolassa, C., & Schmidt, M. (2014). Paid vs. volunteer work in Open Source. 47th Hawaii International Conference on System Sciences (HICSS), USA, 3286–3295. <https://doi.org/10.1109/HICSS.2014.407>
- Rikap C and Lundvall B-A (2022) Big tech, knowledge predation and the implications for development. *Innovation and Development* 12(3): 389–416.
- Sahlin K and Wedlin L (2008). Circulating ideas: imitation, translation and editing. In: Greenwood R, Oliver C, Suddaby R, and Sahlin K (eds) *The Sage Handbook of Organizational Institutionalism*. London: Sage, pp. 218–242.
- Schlagwein D, Conboy K, Feller J, et al. (2017) "Openness" with and without information technology: a framework and a brief history. *Journal of Information Technology* 32(4): 297–305.
- Schneiberg M and Lounsbury M (2008) Social movements and institutional analysis. In: Greenwood R, Oliver C, Suddaby R, et al. (eds) *The Sage Handbook of Organizational Institutionalism*. London: Sage, pp. 650–672.
- Schonhardt-Bailey C, Yager E and Lahlou S (2012) Yes, Ronald Reagan's rhetoric was unique. But statistically, how unique? *Presidential Studies Quarterly* 42(3): 482–513.
- Scott RW (1995) *Institutions and organizations*. Sage, London.
- Suddaby R and Greenwood R (2005) Rhetorical strategies of legitimacy. *Administrative Science Quarterly* 50(1): 35–67.
- Terrell J, Kofink A, Middleton J, et al. (2017) Gender bias in open source: pull request acceptance of women versus men. *PeerJ Computer Science* 3: e111.
- Vaara E and Tienari J (2008) A discursive perspective on legitimization strategies in multinational corporations. *Academy of Management Review* 33: 985–993.
- Vaara E and Tienari J (2011) On the narrative construction of multinational corporations: an antenarrative analysis of legitimization and resistance in a cross-border merger. *Organization Science* 22(2): 370–390.
- Van Meter KM, Mounier L, Chartron G, et al. (1991) Multimethod analysis: official biographies of members of the central committee of the Soviet union communist party. *Bulletin of Sociological Methodology* 33(1): 20–37.
- Vasudevan R (2023) Share and share unlike: reciprocity, corporate power, and the open source ethos. *New Media & Society* 25: 2981–3001.
- Zuboff, S. (2019) *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York, NY: PublicAffairs.

Author biographies

Laure Muselli is Associate Professor of Information Systems Management at Télécom Paris, Polytechnic Institute of Paris. Her research interests cover open source and digital platforms infrastructure, organizational and work-related issues in the context of digital transformation, as well as digital sustainability.

Mathieu O'Neil is Professor of Communication at the University of Canberra and Honorary Associate Professor of Sociology at the Australian National University. He researches information resilience, issue networks, and the sustainability of digital commons. He founded the Journal of

Peer Production in 2011, and the Digital Commons Policy Council think tank in 2021. See <https://dpc.info/>

Fred Pailler holds a PhD in Media Studies and has a background in Sociology and Ethnography, as well as extensive experience in digital content analysis. His research focuses on infrastructures and uses of digital platforms, as well as on controversies related to work, health, gender and sexuality in the digital context.

Stefano Zacchiroli is full Professor of Computer Science at Télécom Paris, Polytechnic Institute of Paris. His current research interests span digital commons, open source software engineering, computer security, and the software supply chain. He is co-founder and CTO of Software Heritage, the largest public archive of software source code. See <https://www.softwareheritage.org/>