TOWARDS UNDERSTANDING OPEN AND COOPETITIVE PLATFORM ECOSYSTEMS: THE CASE OF TENSORFLOW

Completed Research Paper

Abstract

In this research, we address two contemporary management paradoxes in the high-tech sector: competition versus cooperation *and* open-source versus proprietary *technology development. We follow the TensorFlow open-source platform ecosystem led by Google where competitors often cooperate in the co-production of a widely-used platform for machine learning and artificial intelligence. We provide a narrative, complemented with social network visualizations, which depicts the evolution of open-source coopetition. Our findings contribute to the extant theory on motivations for open-sourcing technology and cooperating with competitors. Our results suggest that releasing a complex technological platform in an open source leads to losing intellectual property, while, on the other hand, it can also lead to an increased size of the market that in turn increases demand for complementary products and services. Furthermore, our results also suggest that firms are often 'forced' to engage in coopetition in open-source technological platforms to protect the market share of their complementary products and services.*

Keywords: Alliances, Platforms, Business Ecosystems, Coopetition, Open-Coopetition, Open-Source.

1 Introduction

Many see software as being produced by a single firm. Plenty of evidence shows however that software is often co-produced in networks (see Teixeira, 2023). Paradoxically, those networks can link rival and competing firms that cooperate with each other in an open-source way. For, instance it is known that Apple and Google cooperated in the development of open-source web browsing technologies while fighting expensive patent wars in the courts worldwide, or that Toyota, Ford, and Mitsubishi Motors cooperated in the co-production of open-source automotive software while fighting for car sales in overlapping geographical areas (Teixeira, 2023; Teixeira and Lin, 2014). Managing the production of software in an open and coopetitive modus operandi is increasingly popular (Czakon et al., 2020; Teixeira, 2023) but remains challenging as value can erode via commoditization, free-riding and unintended spillover effects (Gnyawali et al., 2011; Teixeira, Mian, et al., 2016). As pointed out in a recent mapping study by (Herbold et al., 2021), "since more and more companies contribute to open source software and/or develop their software products as open source, collaboration between developers of competing companies becomes an important issue. If developers from competing organizations contribute to the same project, this could lead to issues within a project" (p.12 Herbold et al., 2021).

As research explaining coopetition in an open-source way (aka open-coopetition) remains scarce but gathers cross-disciplinary interest in Software Engineering, Strategic Management, Innovation Studies and Information Systems (e.g., Nguyen-Duc et al., 2019; Roth et al., 2020; Roy et al., 2018; Teixeira, 2023), and as "companies use more and more open source in communities including their competitors it seems very important to know why, how and for which outcomes they follow this kind of strategy"(Czakon et al., 2020), we conducted an exploratory case study guided by two broad and open research questions: *"Why do tech giants like Google open-source advanced and complex technological platforms that started in-house?"* and *"Why are different organizations cooperating with their competitors in the co-production of those open-source platforms?"*. To do so, we take the case of TensorFlow, an advanced and complex technological platform for Machine Learning (ML) and Artificial Intelligence (AI) that started in-house at Google but was released as an open-source project in 2015, and became a key and very popular building block for the ones embedding deep learning technology into thousands of products and services worldwide with impact on everyday life. While the latest evidence points out that open-sourcing spurs entrepreneurial growth in the context of startups (Osborne et al., 2024), less is known about why high-tech giants open-source disruptive technology that started in-house.

2 Theoretical background

2.1 Cooperation among competitors

As noted in the literature on business ecosystems (Clarysse et al., 2014; Iansiti et al., 2004), (Iansiti et al., 2004), strategic cooperation among competitors (also known as coopetition) is not uncommon. This phenomenon can be observed in various industries such as automotive, pharmaceuticals, and airlines. For example, the car roadster Fiat 124 Spider and Mazda MX-5 Miata come out of the same Mazda's Hiroshima factory in Japan. Similarly, alliances in the airline industry have blurred competition between individual firms by forming airline alliances against one another (Gudmundsson et al., 2006). Additionally, companies like Apple, Google and Samsung have collaborated on open-source web-browsing technologies despite engaging in patent war across the courts worldwide (Teixeira and Lin, 2014).

Coopetition cases are not always success stories. As pointed out there is a risk of unwanted leakage of information and opportunistic behaviour (Tidström, 2014). This is especially critical in the open-source domain characterized by transparency, inclusiveness and a weak intellectual property regime (Teixeira, 2015). As far as innovation is concerned, coopetition strategies have been revealed to be more conducive to than purely cooperative or competitive strategies (Quintana-Garcia et al., 2004). There is a call for more research into the connection between innovation and coopetition (Corbo et al., 2023). Furthermore, and also regarding coopetition research, there have been multiple calls to analyze coopetition at multiple levels (Bengtsson et al., 2014; Tidström and Rajala, 2016).

Although competition and collaboration have individually been extensively investigated, researchers have given limited attention to the fundamental issue of how these two concepts interact (Chen et al., 2022). Additionally, existing literature on cooperation among competitors is mostly based on joint ventures and R&D consortia where access is restricted to a few selected members (see Lee et al., 2021). In contrast, in open source software development, third-party actors generally do not require permission to contribute. Despite recognition of the importance of understanding open-source software from both competitive and cooperative perspectives in strategic management literature (see McGaughey, 2002), there are very few empirical cases exploring cooperation among competitors in the open-source arena (see Teixeira, 2023, for a recent review). This lack of empirical research on the topic is unfortunate as "companies use more and more open source in communities including their competitors it seems very important to know why, how and for which outcomes they follow this kind of strategy" (p.6 Czakon et al., 2020).

2.2 Open-source software

Much of the groundbreaking coding that drives software applications, operating systems, cloud servers, and the Internet is a result of "open-source" code - this means code that is freely distributed rather than being kept confidential or protected with a strong intellectual property regime. Numerous individuals contribute to open-source projects for various reasons such as aiding others, gaining recognition in their field, acquiring new skills and knowledge, finding enjoyment and fulfilment from their work, and out of altruism. Additionally, some are incentivized by employment or financial compensation to participate in open-source initiatives (Gerosa et al., 2021).

From an innovation studies perspective, Lakhani et al. (2003) and von Hippel (2005) suggested that open-source software development shows that users program to solve their own as well as shared technical problems and freely reveal their innovations without appropriating private returns from selling the software. Such "free" user-to-user assistance has turned open-source into a remarkable example of user-innovation (von Hippel, 2005). It was also reported that the open-source trend has been so strong that previous, rather monolithic, organizations (e.g., SAP, Intel, Apple, Philips, Xerox, and IBM among others) decentralized research labs, opened up their proprietary technology, and increase their absorptive capacity for outside-in innovation processes within open-source ecosystems (Chesbrough et al., 2006).

The open-source software phenomenon keeps evolving from the earliest purist views focusing on freedom (Stallman, 1985) to newer perspectives considering open-source as an alternative and viable way of doing business (Fitzgerald, 2006; Teixeira, Mian, et al., 2016). Moreover, it has expanded from open-source software to open-data (Gurstein, 2011), open-hardware (Maharaj et al., 2008), open-knowledge (Awazu et al., 2004), and open-access (Davis et al., 2008), among other manifestations of increasing openness in the co-production of goods. Even if the open-source phenomenon started to attract early scholarly attention in computer science and software engineering, the phenomenon is more recently capturing the largest interest from business and management scholars (Raasch et al., 2013). Therefore, as pointed out by Carillo et al. (2015) and von Krogh et al. (2007), information systems as a discipline is well positioned to be at the centre of trans-disciplinary research addressing the phenomenon.

Many scholars have used network perspectives to study the open-source phenomenon from various disciplines (Herbold et al., 2021). For example, while Crowston et al. (2005) conducted a network analysis based *who fixes bugs with who* on 120 projects hosted on SourceForge, while Teixeira, Hyrynsalmi, et al. (2020) examined networks on *who reviews who* on the Linux kernel. In our research, we also performed a longitudinal view of the cooperative network's evolution over time. Unlike most previous studies that used cross-sectional analysis of static networks and extracted quantitative indicators solely from digital artefacts, we attempted to explain the evolution of inter-organizational networks year after year while paying constant attention to the surrounding industrial environment in which those networks were embedded.

3 Empirical background

TensorFlow, was created by the Google Brain team, a deep learning AI research team under the umbrella of Google AI. It was open-sourced (i.e., released under an open-source license) in November 2015. It originated from Google's earlier framework called DistBelief, which was focused on training deep neural networks as a proprietary system. Unlike its predecessor, TensorFlow aimed to offer greater adaptability and scalability by supporting various ML algorithms and being deployable across multiple computer platforms. Its quick adoption stemmed from its capacity to handle large-scale ML tasks effectively via its graph-based computational model that was quite intuitive for representing data transformations. Since its introduction, TensorFlow has emerged as one of the primary frameworks in this field, significantly contributing to progress in AI by empowering researchers and developers to build intricate ML models with relative simplicity. Compared to other deep learning frameworks made available by academic researchers, TensorFlow was released with an effort to make deep learning more accessible. Tensorflow provided extensive documentation, tutorials, YouTube content, trained models, large training data, and even limited free access to cloud computing infrastructure to train models. Furthermore, it was mostly written in Python, and allowed the definition of AI and ML models in Python, one of the most popular languages within the open-source community.

As pointed out recently by others, to date, "prior work on open source co-opetition primarily focuses on projects that are hosted by vendor-neutral foundations" such as the OpenStack, the Linux, the Eclipse, and the Apache foundations. As we know that "due to their vendor-neutrality, foundations play a structural role in enabling collaboration between unexpected allies", it is important to research as well "projects that lack such vendor-neutral governance, such as ones that are initiated, hosted, and governed by one company" (p.1 Wright et al., 2024). Yet another reason to investigate TensorFlow as it is an ecosystem centred around one company (i.e. Google).

4 Method

Given that TensorFlow is a complex software ecosystem with thousands of developers and hundreds of companies, and given its open-source transparent nature that allows us to obtain data on who works with who in the co-production of the platform, we found network analysis to particularity suitable to study the co-production of this complex system over time. For now, we attempt to provide an overall view of the forest before zooming in to a limited sample of trees (see Bergenholtz et al., 2011).

Our case relied mostly on digital trace data publicly available on the Internet. The collected data was naturally occurring, in the sense that it did not result from researchers' actions, but rather created and maintained by the TensorFlow community in their pursuits of developing an open-source platform for ML and AI. To make sense of coopetition within TensorFlow, we followed the methodological approach by Teixeira, Robles, et al. (2015) that combines the qualitative analysis of archival data (QA), the mining software repositories (MSR), and Social Network Analysis (SNA) to reconstruct and visualize the evolution of collaboration in a sequence of visual networks (aka sociograms). The methodological approach by Teixeira, Robles, et al. (2015) was previously explored by others (e.g., Zhang et al., 2021) and according to a recent systematic mapping study by Herbold et al. (2021, p. 12), merits by not only modelling inter-developer collaboration but also considering "inter-company" collaborations. The method and the the respective tool[1] associates software developers with each other by co-edits of the software source code, and attributes an organizational affiliation to each developer by the e-mail domain they use.

We began our research qualitatively while learning about TensorFlow and its surrounding empirical background. This first stage, allowed us to review an immense amount of online information regarding the ML and AI industry. An internal web-based information system was used to keep and organize the retrieved qualitative materials. After attaining a better understanding of the industrial cooperative and competitive dynamics, we extracted and analyzed the social network of the TensorFlow open-source project by leveraging SNA (Wasserman et al., 1994). To do so, we retrieved the TensorFlow repository and its *changelog*. The commit logs were then scrapped, modelled and visualized as social networks using *Python NetworkX*(v3.3), *Python Matplotlib*(v3.8.4) and *Visone*(v2.27.1) all guided by the methodology provided by Teixeira, Robles, et al. (2015). As the resulting inter-individual social network was very dense to visually understand or explain (see Figure 1), we first segmented the analysis from year to year since TensorFlow does not follow a time-based release strategy (see Teixeira, Robles, et al., 2015).

Furthermore, we transformed the inter-individual networks where organizational affiliation is a node (i.e., developer) attribute to inter-organizational networks where organizations are the nodes, and edges are weighted with the number of unique pairs of developers affiliated with two organisations cooperating with each other. The more developers affiliated with two firms cooperate, the more we expect the two organizations to cooperate as well. Even if we lost information of who works with who at the individual level, we got inter-organizational networks of who works with who that are much easier to interpret visually (c.f., Teixeira, Robles, et al., 2015). The loss of information on who works with who at the individual level is irrelevant as we discuss coopetition among organizations in the open-source arena, not among individuals. In other words, by applying a transformation to the networks that result from the Teixeira, Robles, et al. (2015) methodology, we could better analyse and discuss open-coopetition as a "business strategy" (see Teixeira, 2023).

By mining digital traces of collaboration, and uncovering the evolution of social structure in the TensorFlow project, the computerized SNA led to novel insights on open-source and competition. The combination of methods was not only fundamental for the retrieval of inter-organizational cooperation structures but also for explaining them. We attempted to explain how the inter-organizational networks evolved year after

[1] See https://github.com/jaateixeira/ScrapLogGit2Net.

year (2023-2024). Our analysis started with the first code commit on Nov 7 2013 by Vijay Vasudevan (i.e. the birth of TensorFlow as an open-source project outside of Google) and we closed our data analysis on 12 Nov 2024 for reporting the first results of this research[2]. The inter-organizational networks were filtered to cover the top contributors to the TensorFlow core project that can be easily identified as affiliated with competing firms (i.e.,'google','microsoft','ibm','amazon','intel','amd','nvidia','arm','meta', and 'bytedance'). We could easily depict how those 10 firms plus Google market similar products and services on the same geographical areas.

5 Results

As in Teixeira, Mian, et al. (2016) and Teixeira, Robles, et al. (2015) our results are presented within a narrative and chronological format complemented with "pictures" of the evolving social structure of TensorFlow.

5.1 From Nov 2015 to Nov 2024 - A very dense collaborative network

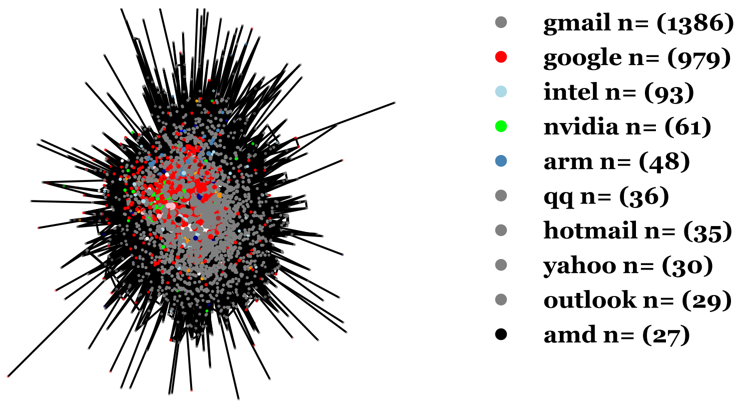

Figure 1. Sociogram capturing collaboration among developers during Nov 2014 - Nov 2024.

After identifying and filtering for 'bots' that commit code to the project[3], Figure 1 captures the network of collaboration among developers in the TensorFlow. As it covers almost 10 years of digital trace data on the co-production of TensorFlow code, the network is very dense. We could identify 4220 nodes/developers and 37831 edges/relationships. We note that many contributions were associated with email accounts for personal use (e.g., gmail, outlook, hotmail, ee and others). Google is visible in a central position, followed by chip-makers such as arm, intel, nvidia, and amd in the leaderboard of the collaborative software co-production efforts.

5.2 Later 2015 and 2016 - Microsoft, Huawei, IBM, and Intel join the project

The Figure 2 captures code-collaboration in the TensorFlow during 2016. Nodes represent organizations, and edges have weights that denote the number of unique pairs of developers affiliated with two organisations cooperating with each other. The more developers affiliated with two firms cooperate, the more we expect that the two organizations to cooperate as well.

After the initial "code bomb" by Vijay Vasudevan from Google on November 9, 2015, that officially started TensorFlow as an open-source project, we found that six developers of Google worked on the project before the Christmas holidays. During the remainder of 2015, many contributions were submitted

[2] Note we do not provide inter-organizational visualization results for 2015 and 2024 as the covered period is too small for drawing parallels with other years.

[3] Mostly the gardener@tensorflow.org bot that commits code as part of continuous integration and quality assurance non-human automation within Google.

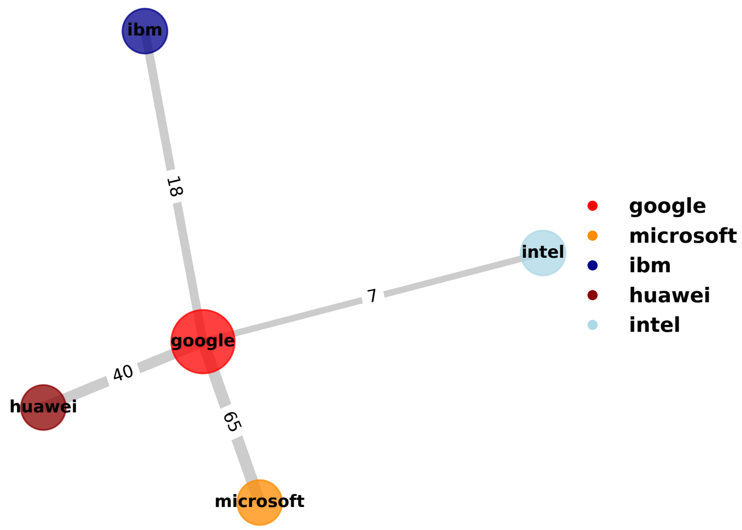

Figure 2. TOP 10 organizational contributors with Google in 2016.

using gmail accounts. The first organizational contributions came from reputed research institutions (e.g., CERN), and specialized startups. During 2016, researchers in the fields of AI and ML contributed to TensorFlow from all over the world. High-tech giants started contributing as well with Microsoft, Huawei, IBM, and Intel leading the way in terms of the number of developers within the collaborative networks of TensorFlow. By 2016 AI and ML were rapidly gaining importance in the tech industry, and the strong investment and support from Google brought many to the TensorFLow ecosystem.

During the first months, Google's decision to open-source TensorFlow gained much attention and we could witness a lot of naturally occurring qualitative material emerging on the Internet ranging from televised expert commentaries[4], to Internet forums[5]. While many wonder why giving up so much intellectual property already on the table, many outlined other explanations such as crowdsourcing innovation, fomenting open innovation, resource complementarity, sharing risk, increased cooperation with researchers, enhancing collaborative efficiency in product development, practical benefits that can lead to faster product development and maintenance, reuse of software artefacts, exposition of the technology to thousands of developers in the open-source community as well as fearing the threat of new entrants to the market.

Facebook which changed its corporate name to Meta Platforms, Inc. on October 28, 2021, took competitive action and released PyTorch - a competing open-source platform. Its initial development was led by Facebook's AI Research lab (FAIR), and it was officially made open-source in October 2016, but its popularity remained small for its first year. This besides being very friendly towards researchers by allowing rapid prototyping and experimentation. TensorFlow and PyTorch kept implementing very similar features over the next years and the latter was less centralized around a single firm.

5.3 2017 - Nvidia joins in force

Figure 3 captures code-collaboration in the TensorFlow during 2017. When compared to 2016 (see Figure 2), we can notice that IBM, Intel and Microsoft collaborated with Google much more (i.e., with more pairs of developers engaging in cooperation between the two firms). During the same year, Nvidia started contributing in force to TensorFlow in cooperation with Intel and Google. By contributing, Nvidia early optimised the performance of their GPUs to meet the market demand for TensorFlow-based AI and ML workloads.

Nvidia launched in 2017, the NVIDIA GPU Cloud (NGC), a data-centre hub for GPU-optimized software

[4] See *Why Google wants everyone to have access to TensorFlow* via Fox News https://www.foxnews.com/video/4611174773001.
[5] See https://www.quora.com/Why-did-Google-open-source-TensorFlow-Whats-in-it-for-them.

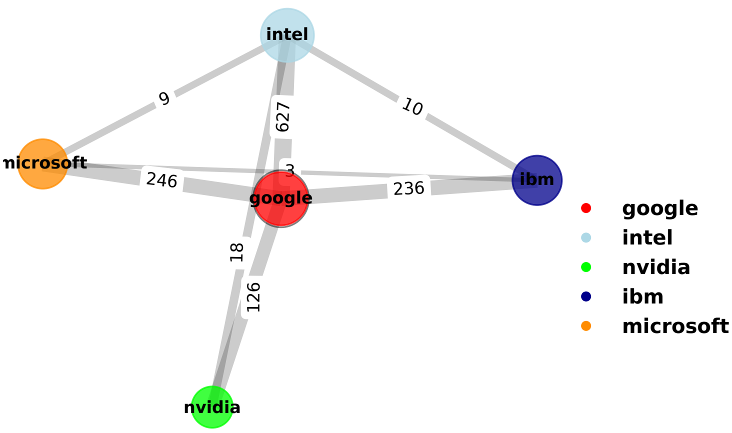

Figure 3. TOP 10 organizational contributors with Google in 2017.

for deep learning and high-performance computing. NGC included containers with TensorFlow and other frameworks optimized for Nvidia GPUs, making it easier for developers to deploy and scale their AI workloads. Nvidia became an early provider of AI and ML services on the cloud. Many could turn to Nvidia to rent the latest state-of-the-art computing power to train their models. In the meanwhile, others could not get their hands on Nvidia's best-performing GPUs at reasonable prices losing precious time for the AI race.

5.4 2018 - ByteDance, AMD, and Amazon started contributing

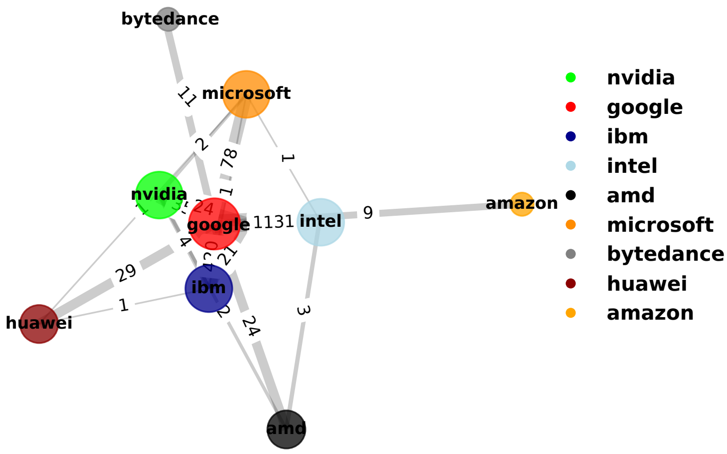

Figure 4. TOP 10 organizational contributors with Google in 2018.

Figure 4 captures code-collaboration in the TensorFlow during 2018. Compared to 2017 (see Figure 2) we can note that ByteDance, AMD and Amazon started contributing in force making it to the TOP 10 contributors to the project in terms of number of nodes (aka developers). ByteDance, the company behind popular apps like TikTok (known as Douyin in China), began contributing to TensorFlow in 2018 as its apps, particularly TikTok and CapCut, rely heavily on AI and machine learning for content recommendation, personalization, and user engagement. Contributing to TensorFlow allowed ByteDance to leverage and improve the framework to better meet its own needs. Note ByteDance cooperate mostly with Google. Both ran large research centers specialized in AI and both recruited many talented researchers from universities and research institutes around the world.

From the network, we can also note that AMD started working on the project, something that experts remark as "too little to late"[6]. During the same time, Arm and Nvidia announced in March 2018 a partnership in which Arm would use Nvidia's open-source Deep Learning Accelerator (NVDLA) architecture[7]. For our surprise, Arm developers were slow to engage in the co-production of TensorFlow, as they do not appear in the network yet. Amazon also started contributing more heavily to the project in cooperation with Intel. Microsoft early supported TensorFlow functionally on top of its Azure cloud computing services and Amazon needed do the same with its AWS cloud computing services to not lose market share.

5.5 2019 - Arm finally starts contributing to the project

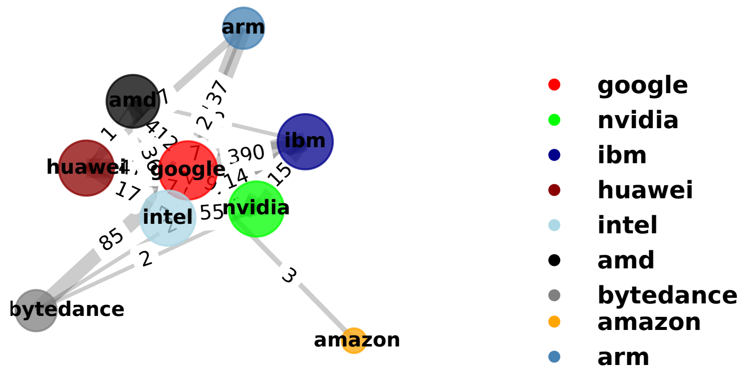

Figure 5. TOP 10 organizational contributors with Google in 2019.

Our visualizations in Figure 5 capture inter-developer collaboration during 2019. By comparing it with the prior year, we can see that Arm finally started contributing to the project and joined other chipset makers at the network core. It is also notable that ByteDance, which relies on TensorFlow to power their algorithmic recommendation systems, started cooperating intensively with Google with 85 unique pairs of developers collaborating with each other.

On Jun 2019, Lex Fridman a YouTuber and research scientist at the MIT Laboratory for Information and Decision Systems, posted a long interview (1 hour and 11 minutes) with Rajat Monga that lead the R&D of TensorFlow at Google[8]. To our surprise and benefit, the interview indirectly addressed our first research question by pointing out several key reasons how why Google open-sourced TensorFlow. First, as a large corporate research unit, they wanted to share their findings and push the state of the art in machine learning and deep learning forward. By open-sourcing TensorFlow, they aimed to foster collaboration and innovation within the broader research community, as they had been building on existing research and wanted to contribute back. Secondly, there was a recognition that existing software, primarily developed in academia, did not meet the demands of large-scale applications or the diverse hardware environments that Google was working with. By creating TensorFlow as an open-source project, Google could provide a more robust, modular and scalable solution that could be used widely, thus helping the community and setting a standard for machine learning frameworks[9]. Additionally, the decision to open source TensorFlow aligned with Google's strategy of promoting open innovation, allowing others to build on their work and contribute to the ecosystem. This approach not only benefited Google by enhancing their own technology but also encouraged a vibrant community of developers and researchers to engage with and improve TensorFlow, ultimately leading to its widespread adoption across various industries[10]. In the words of Rajat Monga, their goal was "to get machine learning on every device" and by open-sourcing TensorFlow and providing extensive documentation and community support, they "made

[6] See https://www.servethehome.com/tensorflow-1-8-with-amd-rocm-support/.
[7] See https://techcrunch.com/2018/03/27/arm-chips-will-with-nvidia-ai-could-change-the-internet-of-things/.
[8] Transcript publicly available at https://transcript.lol/read/youtube/@lexfridman/652310c5033150beacd17e4e.
[9] See [04:02 - 07:49] at https://www.youtube.com/watch?v=NERNE4UThHU.
[10] See [30:04 - 32:07] [41:58 - 43:57] from the same recorded video interview.

deep learning accessible to a broader audience, including developers who previously had no experience with machine learning". This shift allowed not only researchers but also hobbyists and enterprises to adopt and implement machine-learning solutions. The growing ecosystem around TensorFlow, including integrations with Google Cloud, further facilitated its adoption in various industries, thus expanding the overall market for machine learning technologies[11].

Also to our surprise, the interview also touched into the second research question by pointing out why large corporation engage in open-coopetition. For example, many large corporations (e.g., AMD, Intel, Arm, and Nvidia) produce hardware that is widely used in machine learning and deep learning applications. By contributing to TensorFlow, they can optimize the platform to better leverage their specific architectures, ensuring that TensorFlow runs efficiently on their processors and GPUs. In addition, they can gain insights into emerging trends and technologies, which can inform their own product development strategies. Finally, that can also support the developer community by "providing resources, tools, and libraries that enhance the overall experience of using their own hardware for machine learning tasks". The same could be said for companies like Microsoft, Amazon, and IBM that offer their own cloud services and technologies. By contributing to TensorFlow, they can optimize the platform to integrate better with their own "stack", enhancing the overall user experience for their customers. By contributing code, and engaging in special interest groups in the project, they can influence the direction of the project to better suit their needs and the needs of their customers. The interviewee also remarked that some companies like IBM contribute to TensorFlow because their customers demand it. They ask for customized solutions that force them to conribute to the TensorFlow platform[12].

5.6 2020

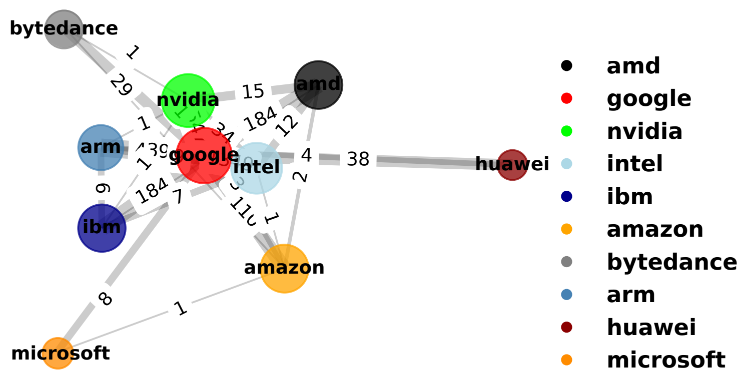

Figure 6. TOP 10 organizational contributors with Google in 2020.

Our visualizations in Figure 6 capture inter-developer collaboration during 2020. By comparing it with the prior year, we note the re-entering of Microsoft as a notable contributor, signalling their increased involvement in the project. Microsoft collaborated mostly with Google (eight developer-developer links) and to a smaller extent with Amazon (just one developer-developer link). Microsoft continued optimizing TensorFlow for its Azure cloud platform to compete more effectively with Google Cloud and Amazon Web Services. In addition Microsoft integrated TensorFlow into some of its products such as Office, Bing and Cortana.

Despite being competitors, large tech companies often collaborated on the project to ensure interoperability and progress. Others feared the fragmentation of AI platforms. In 2020, the concern about fragmentation in AI platforms was a significant issue for major tech companies, including Microsoft. To address this challenge, they increased collaboration on open-source projects and developed tools and standards that

[11]See [02:08 - 07:49] [14:07 - 18:07] [41:58 - 43:57]from the same recorded video interview.
[12]See [05:56 - 07:49] [10:00 - 12:03] [30:04 - 32:07] [41:58 - 43:57] [43:58 - 45:56] from the same recorded video interview.

supported multiple platforms. These efforts aimed to promote interoperability, reduce the burden on developers, and accelerate innovation and adoption[13].

5.7 2021 - Chipset makers at the network core

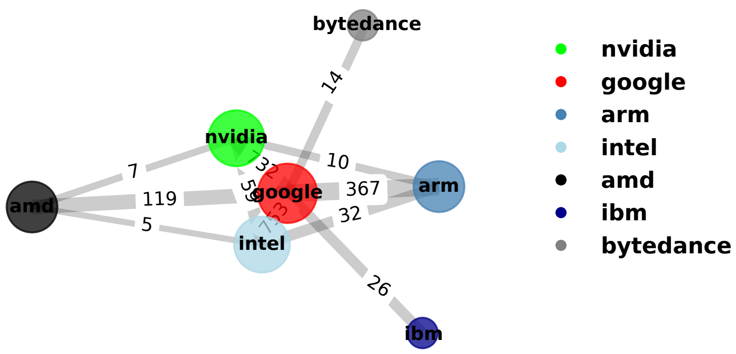

Figure 7. TOP 10 organizational contributors with Google in 2021.

Our visualizations in Figure 7 capture inter-developer collaboration during 2021. By comparing it with the prior year, we note that Huawei, Amazon and Microsoft did not make it to the top 10 contributors in terms of number of developers that engaged in cooperation to co-produce TensorFlow. We could say that their commitment to the TensorFlow open-source decreased regarding source-code contributions. The drop in Huawei could be explained by the ongoing trade war between the United States and China. Rather than following the lead of Google, Huawei started MindSpore, an open-source AI computing framework that was in many ways similar to TensorFlow and PyTorch. MindSpore became a third open-source platform for AI and ML - but a less popular one out of China. Amazon also started diverting from TensorFlow and started developing alternative open-source technologies like MXNet and Gluon, and proprietary machine learning frameworks and services like SageMaker. Amazon prioritised its technologies and integrated them tightly with AWS - at the time the leading cloud computing provider in the world exploiting network effects.

In the meanwhile, Microsoft did something similar to Amazon. On the one hand, it went proprietary and invested heavily in its own machine learning and AI technologies (e.g., Azure Machine Learning) while, on the other hand, it continued contributing to other open-source projects besides TensorFlow such as the ONNX project, PyTorch and even started the CNTK (Microsoft Cognitive Toolkit) that also become popular. Note here that it is known that the Azure Machine Learning products and services commercialized by Microsoft are not open-source even if it is known that they they integrate plenty of open-source components. Fare enough Microsoft remained a top contributor to many of the open-source projects they integrated. Those who remained firm at the inter-organization cooperative network core were the Chipset makers Nvidia, Intel and Arm—all with dozens of developers working together with Google.

Open-coopetition is not easy and conflict can emerge. In 2021, Google and the TensorFlow community also adopted more strict codes of conduct and conflict resolution mechanisms common in other open-source ecosystems[14]. Reflecting that many individuals and organizations have different interests in the TensorFlow community, many Special Interest Groups (SIGs) were also created during 2021[15]. Furthermore, as developers complained about lack of transparency on how contributions are reviewed and integrated into the TensorFLow core, Google set up some initiatives aiming at increasing transparency in the up-stream code integration processes[16]. TensorFlow as a high-performance platform for AI and ML

[13]See https://thenewstack.io/open-neural-network-exchange-brings-interoperability-to-machine-learning-frameworks/.
[14]See https://github.com/tensorflow/tensorflow/blob/master/.
[15]See https://www.tensorflow.org/community/.
[16]See https://transcript.lol/read/youtube/@lexfridman/652310c5033150beacd17e4e.

had now more alternatives on the market (e.g., PyTorch and Amazon SageMaker among others).

5.8 2022 - IBM Watson and Amazon AWS embrace TensorFLow

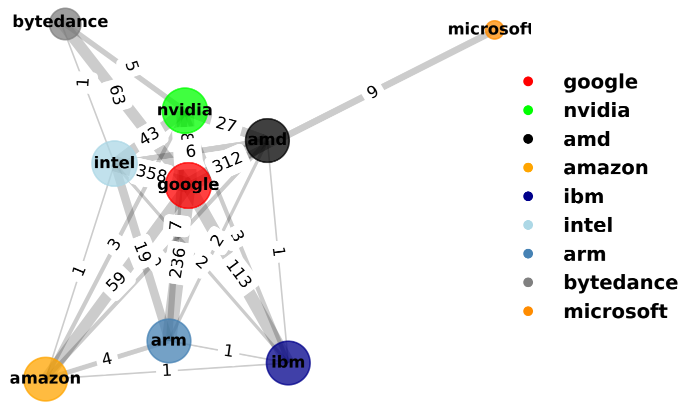

Figure 8. TOP 10 organizational contributors with Google in 2022.

Our visualizations in Figure 8 capture inter-developer collaboration during 2022. By comparing it with the prior year, we note that Amazon and Microsoft made it again to become among the top 10 contributors in terms of number of developers that engaged in cooperation to co-produce TensorFlow. They were interested in ensuring compatibility of TensorFlow to their in-house platforms (i.e., Azure Machine Learning and Amazon SageMaker). Contributing to TensorFlow can help ensure that their platforms are compatible with a widely-used framework, making it easier for 3rd party developers (e.g., customers of Microsoft and Amazon) to integrate TensorFlow models into their ecosystems. As depicted in the visualization, Microsoft worked mostly with AMD. By looking at their code contributions, we can say that they were working towards improving the performance of TensorFlow workloads on AMD hardware that powered many of Microsoft's cloud services. Here we must point out that by that time the Nvidia GPU Cloud (NGC) was already one of the most profitable business units of Nvidia

Chipset markers (i.e., Nvidia, Intel, AMD and Arm) remained in the network core. From the network periphery, IBM was active in insuring that TensorFlow runs well in the powerful servers they sell[17] as well in their IBM Watson computing services[18]. Amazon also took TensorFLow very serious and even gifted high-quality data-sets to the project[19]. If IBM was tuning TensorFlow for their IBM Watson computing services, Amazon was doing the same for their computing services as well[20].

By November 30, 2022, Large Language Models (LLMs) started to become popular. ChatGPT, developed by OpenAI, was released to the public and the model quickly gained attention for its ability to generate human-like text and engage in conversational interactions. While receiving venture capital from Microsoft OpenAI used both TensorFlow and PyTorch to train their models. The exact physical locations of the training infrastructure are not publicly disclosed, it is known that OpenAI utilizes Microsoft Azure's cloud computing resources to train and deploy its models.

5.9 2023 - The rise of LLMs and the unmatched performance of Nvidia

Our visualizations in Figure 9 capture inter-developer collaboration during 2023. By comparing it with the prior year, we note that Microsoft failed again to rank among the top 10 contributors in terms of the

[17] See https://developer.ibm.com/blogs/run-ai-inferencing-on-power10-leveraging-mma/.
[18] See https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/ml-manage-frame-and-specs.html?context=cpdaas.
[19] See https://www.tensorflow.org/datasets/catalog/
[20] See https://aws.amazon.com/blogs/machine-learning/best-practices-for-tensorflow-1-x-acceleration-training-on-amazon-sagemaker/.

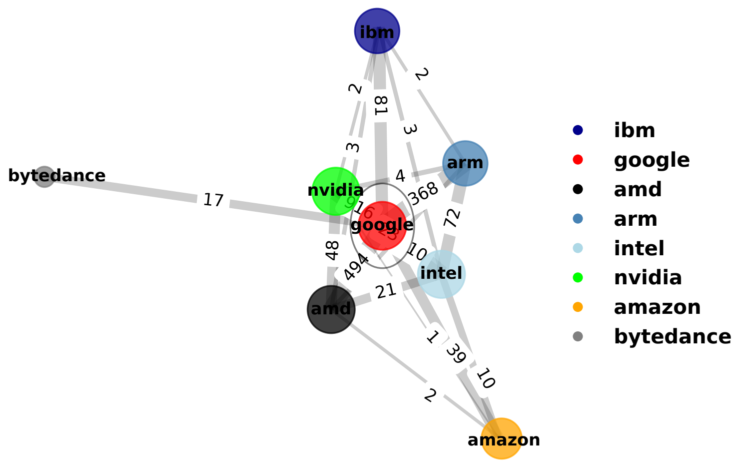

Figure 9. TOP 10 organizational contributors with Google in 2023.

number of developers. This could be perhaps explained by Microsoft's strategic alignment with OpenAI and its popular ChatGPT model.

In 2023, we witnessed the rise of open-source and commercial LLMs. Large2 by Mistral AI, GPT-4 by OpenAI, PaLM 2 by Google, Llama 2 by Meta and Claude by Anthropic further accelerated the adoption and development of these models. These models are being used in a wide range of applications, from chatbots to content generation. All the firms we cover in this research paper contributed to the rise of LLMs on multiple fronts (e.g., doing research, developing software, funding AI startups, providing infrastructure to train models).

During the same year TensorFlow re-aligned its strategy to align with the rise of Large Language Models (LLMs) as well as to recent research advancements in quantization (i.e., a group of techniques designed to make models faster, smaller, and less resourceful to train and serve)[21]. By 2023, Nvidia became a star on the stock exchange markets as its GPUs led the technical benchmarks for deep learning workloads[22]. On the meanwhile, players that used to cooperate with Google in the co-production of TensorFlow like Meta and Microsoft started drifting towards the other competing deep learning framework that was growing in popularity: PyTorch[23]. By the end of 2023, Google and Nvidia were the biggest beneficiaries of the popularity of TensorFlow. A story that has only been 10 years so far. So far TensorFlow is the most deployed deep learning platform, while PyTorch is gaining popularity among researchers, hobbits and professionals who want to start into AI and ML.

6 Discussion

Our research addressed a direct call by Czakon et al. (2020) that stated the importance of knowing "why and for which outcomes companies use more and more open source in communities including their competitors" (Czakon et al., 2020). By addressing the research questions *"Why do tech giants like Google open-source advanced and complex technological platforms that started in-house?"* and *"Why are different organizations cooperating with their competitors in the co-production of those open-source platforms?"*, we aim to increase our understanding of the open-coopetition phenomena. Guided by Teixeira, Robles, et al. (2015) methodological approach, we narrated the first circa ten years of TensorFlow with complementary visualizations of its social structure. Furthermore, and to be able to explain the retrieved 'computerized' social network visualizations over time we digested plenty of naturally occurring qualitative material that explained why Google open-sourced TensorFlow and why many other high-tech

[21]See https://blog.tensorflow.org/2023/05/google-io-2023-whats-new-in-tensorflow-and-keras.html.
[22]See https://www.statmuse.com/money/ask/nvda-stock-price-2023.
[23]See https://about.fb.com/news/2023/07/llama-2/.

giants engage in open-coopetition.

The first surprising empirical results are captured in Figure 1. Many of the code contributions to TensorFlow were identified with email accounts commonly associated with personal use (i.e., gmail, hotmail, yahoo, outlook and Chinese equivalent qq). We noted that many researchers associated with universities, research institutions and startup founders contributed using personal email accounts. In this sense, Google was successful at opening innovation processes by putting together an open-source community. Like in other cases of open-coopetition in the automotive industry (see Teixeira, 2023), they open up to get third-party contributions from enthusiasts, students, hackers and academics among others. In the words of Teixeira, "no one should need a permit to innovate on top of their product platforms" (p. 6 Teixeira, 2023). In this case, we noted that Google got many contributions from researchers in Germany, Russia, Cyprus, Portugal, France, Japan, Korea, South Africa, Australia, and many other geographies miles away from the Google office where TensorFlow was first developed. Innovation was not confined to a regional cluster.

The second surprising empirical result is visible in many of the figures capturing inter-organizational collaboration among the top 10 contributors to TensorFLow. The visualizations evidence that chipset makers (i.e., amd, nvidia, arm, and intel), and cloud computing vendors (i.e., microsoft, amazon, and ibm) collaborate directly with each other besides competing with each other (i.e., marketing similar products and services in the same geographical markets). In this sense, these results align with prior studies that found low levels of homophily by company affiliation in the production of complex open and coopetitive software ecosystems (see Nguyen-Duc et al., 2019; Teixeira, Hyrynsalmi, et al., 2020; Teixeira, Robles, et al., 2015). It seems that open-source communities inherently characterized by transparency and inclusiveness are a neutral environment for competitors to engage in coperatation with each other.

While the set explanations from the previously mentioned qualitative material were already covered by literature in open-source, open-innovation, user-innovation, coopetition, and open-coopetition (see e.g., Bengtsson et al., 2014; Roy et al., 2018; Teixeira, 2023; von Hippel, 2005), we did find another novel explanation (i.e., expected outcome) for releasing software in an open-source way: increasing the market size for the technology. As Google released TensorFlow under an open-source license, it became much easier and more efficient to deploy deep learning. As time passed, barriers to entry decreased, technology evolved, and people with skills to train and deploy efficient deep learning models passed from a few hundred (often researchers with advanced doctoral education) to many thousands. The range of products and services embedding deep learning increased exponentially and all that created additional demand for the computing and data services commercialized by Google. Furthermore, as a bonus, Google started also selling TPUs (i.e., integrated circuits for neural network ML) that had been specifically designed for TensorFlow and started deploying them on their Google Pixel Phones and many other products and services at google (e.g., speech and image recognition). As pointed out by Rajat Monga in the 2019 interview (see Section 5.5), "deep learning used to be done by researchers, but we can now see high schoolers building, training and deploying deep learning models".

Based on those observations we contribute to the further theorize the phenomenon of open-coopetition by laying out the following proposition:

Theoretical Proposition 1 – *Within a high tech context, releasing a complex technological platform in an open-source way will lead to a loss of intellectual property, but it can also lead to an increased size of the market that in turn can increase demand for complementary products and services.*

Regarding the second research question *Why are different organizations cooperating with their competitors in the co-production of those open-source platforms*, we found it easier to explain *what would happen if they fail to do so*. First, let's take the example of Microsoft. If Microsoft would ignore TensorFlow, we would likely see the demand for its Azure cloud computing services and for its Windows operating system decrease to other players that put the effort in steering, customizing and optimizing TensorFlow according to their services and product needs (e.g., Amazon cloud services and Linux). As a second example, if Nvidia would ignore TensorFlow, we would likely see the demand for its GPUs chipset sales decreasing to other chipset vendors performing better at running TensorFlow (e.g., Arm, Google, Intel, AMD and Samsung). Furthermore, if ByteDance would also ignore TensorFlow, it would be in a much harder position to maintain its distributed deep learning infrastructure that supports several apps such as TikTok and CapCut. Finally, companies like IBM that help their customers run customized deep learning models, could also lose businesses by not contributing to TensorFlow. Contributing to TensorFlow can bring good reputation to the firm to win projects that deploy TensorFlow deep learning in production and can also help them to meet their customer requirements. Out theoretical propositions call for the theorizing of open-coopetition (see Teixeira, 2023) as a managerial dilemma dealing with many inherent tensions. In our view, it is also clear that key market players were "forced" to engage in open-competition to not lose market share as both the TensorFlow platform ecosystem and the AI market kept growing.

While we suggest that open sourcing might increase the size of the market, failing to engage in open-coopetition can negatively impact the market share. Based on our observations we further theorize the phenomenon of open-coopetition by laying out the second and last proposition of this study:

Theoretical Proposition 2 – *Within a high tech context, firms might be forced to engage in open-coopetition in the co-production of complex technological platforms to protect the market share of their complementary products and services.*

From a practical point of view, we warn managers to not fall into the trap of being overfocused on protecting intellectual property and market share[24]. By being so concentrated on growing intellectual property assets, licensing, sales and market share, they can neglect the potential size of the market that a technology can bring. The open-sourcing of TensorFlow is now perceived as a success story, the value that TensorFlow brought to Google as an open-source project in terms of innovation and complementary demand is much higher than the value that TensorFlow could bring if kept locked and protected within the organization. In certain conditions, increasing the market size can offset the loss of intellectual property, licensing revenues, sales and market share. Managers must conciliate advice and interests from the legal and marketing teams with the advice and interests from the R&D and engineering teams, to not miss up opportunities to scale the market. While in this research we took a view at the overall inter-organizational network and its evolution, our future research will zoom in and gain a more in-depth understanding by interviewing participants from the involved firms - the dominant methodological approach in coopetition research.

7 Conclusion

In this research, we followed the TensorFlow open-source software ecosystem where competitors often cooperate in the co-production of a widely-used platform for ML and AI. We contributed to an enhanced understanding of (1) why open-sourcing technology that was made in-house, and (2) why cooperate with competitions in an open-source way.

We found out that on the one hand, by open-sourcing Tensorflow, Google benefited upstream by continuously developing the state of the art of deep learning. This as it counted with contributions from researchers from all over the world in the form of ideas, models, methods, patches, pull requests, bug fixes, security patches, etc. On the other hand, Google passed some of those benefits downstream by making it more easy and more efficient to deploy deep learning into products and services that impact real people. That in turn increased the size of the market and created complementary demand for the computing and data services that Google commercializes. After all, the best-performing deep learning models depend on large amounts of quality data and computational power that Google have at hand.

[24]Note that many organizations have incentives that reward employees and managers based on intellectual property and market share.

The main message of the paper is that Google increased the size of the market by open-sourcing TensorFlow and created complementary demand for its computing and data services. Also, Google brought innovation from outside (o.e., from researchers, specialized startups, users, competitors, deployers, etc). Many competitors were "forced" to cooperate in co-production of TensorFlow otherwise the demand for their complementary products and services (e.g., chipsets, computing services, and operating systems among others) would decrease to other players that steered, nudged, customized, and optimized TensorFlow to their own needs.

References

Awazu, Y. and Desouza, K. 2004. "Open knowledge management: Lessons from the open source revolution," *Journal of the American Society for Information Science and Technology* 55 (11), 1016–1019.

Bengtsson, M. and Kock, S. 2014. "Coopetition—Quo vadis? Past accomplishments and future challenges," *Industrial Marketing Management* 43 (2), 180–188.

Bergenholtz, C. and Waldstrøm, C. 2011. "Inter-Organizational Network Studies—A Literature Review," *Industry and Innovation* 18 (6), 539–562.

Carillo, K. and Bernard, J.-G. 2015. "How Many Penguins Can Hide Under an, Umbrella? An Examination of How Lay Conceptions Conceal the Contexts of Free/Open Source Software," in *Proceedings of the International Conference on Information Systems (ICIS 2015),* AIS.

Chen, H. et al. 2022. "Network Dynamics and Organizations: A Review and Research Agenda," *Journal of Management* 48 (6), 1602–1660.

Chesbrough, H. et al. 2006. *Open innovation: Researching a New Paradigm,* Oxford University Press.

Clarysse, B. et al. 2014. "Creating value in ecosystems: Crossing the chasm between knowledge and business ecosystems," *Research Policy* 43 (7), 1164–1176.

Corbo, L. et al. 2023. "Coopetition and innovation: A review and research agenda," *Technovation* 122, 102624. ISSN: 0166-4972.

Crowston, K. and Howison, J. 2005. "The social structure of free and open source software development," *First Monday* 10 (2). ISSN: 13960466.

Czakon, W. et al. 2020. "Coopetition strategies: Critical issues and research directions," *Long Range Planning* 53 (1). ISSN: 0024-6301.

Davis, P. M. et al. 2008. "Open access publishing, article downloads, and citations: randomised controlled trial," *BMJ* 337, a568.

Fitzgerald, B. 2006. "The Transformation of Open Source Software," *MIS Quarterly* 30 (3), 587–598. ISSN: 02767783.

Gerosa, M. et al. 2021. "The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source," in *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE),* pp. 1046–1058.

Gnyawali, D. R. and Park, B.-J. 2011. "Co-opetition between giants: Collaboration with competitors for technological innovation," *Research Policy* 40 (5), 650–663. ISSN: 0048-7333.

Gudmundsson, S. V. and Lechner, C. 2006. "Multilateral airline alliances: Balancing strategic constraints and opportunities," *Journal of Air Transport Management* 12 (3), 153–158. ISSN: 0969-6997.

Gurstein, M. B. 2011. "Open data: Empowering the empowered or effective data use for everyone?" *First Monday* 16 (2).

Herbold, S. et al. 2021. "A systematic mapping study of developer social network research," *Journal of Systems and Software* 171, 110802. ISSN: 0164-1212.

Iansiti, M. and Levien, R. 2004. "Strategy as Ecology," *Harvard Business Review* 82 (3), 68–81.

Lakhani, K. R. and von Hippel, E. 2003. "How open source software works:"free" user-to-user assistance," *Research Policy* 32 (6), 923–943.

Lee, J. Y.-H. et al. 2021. "Managing information sharing: Interorganizational communication in collaborations with competitors," *Information and Organization* 31 (2), 100354. ISSN: 1471-7727.

Maharaj, B. T. et al. 2008. "A low-cost open-hardware wideband multiple-input–multiple-output (MIMO) wireless channel sounder," *IEEE Transactions on Instrumentation and Measurement* 57 (10), 2283–2289.

McGaughey, S. L. 2002. "Strategic Interventions in Intellectual Asset Flows," *The Academy of Management Review* 27 (2), 248–274. ISSN: 03637425.

Nguyen-Duc, A. et al. 2019. "Do Software Firms Collaborate or Compete? A Model of Coopetition in Community-initiated OSS Projects," *e-Informatica Software Engineering Journal* 13 (1), 37–62.

Osborne, C. et al. 2024. "Characterising Open Source Co-opetition in Company-hosted Open Source Software Projects: The Cases of PyTorch, TensorFlow, and Transformers," (available online at `https://arxiv.org/abs/2410.18241`;).

Quintana-Garcia, C. and Benavides-Velasco, C. A. 2004. "Cooperation, competition, and innovative capability: a panel data of European dedicated biotechnology firms," *Technovation* 24 (12), 927–938.

Raasch, C. et al. 2013. "The rise and fall of interdisciplinary research: The case of open source innovation," *Research Policy* 42 (5), 1138–1151.

Roth, S. et al. 2020. "Open coopetition: when multiple players and rivals team up," *Journal of Business Strategy* 41 (6), 31–38.

Roy, F. L. et al. 2018. "Open coopetition: a research program," in. European Academy of Management (EURAM).

Stallman, R. 1985. *The GNU manifesto,* Free Software Foundation, Inc.

Teixeira, J. 2015. "On the openness of digital platforms/ecosystems," in *Proceedings of the 11th International Symposium on Open Collaboration (OpenSym 2015),* ACM, Aug. 2015.

— 2023. "Towards understanding open-coopetition – Lessons from the automotive industry," in *Proceedings of the 44th International Conference on Information Systems (ICIS 2023),* AIS.

Teixeira, J., Hyrynsalmi, S., et al. 2020. "Network Science, Homophily and Who Reviews Who in the Linux Kernel?" in *Proceedings of the 28th European Conference on Information Systems (ECIS 2020),* AIS.

Teixeira, J. and Lin, T. 2014. "Collaboration in the Open-source Arena: The WebKit Case," in *Proceedings of the ACM Conference on Computers and People Research (SIGSIM-CPR '14),* ISBN: 978-1-4503-2625-4.

Teixeira, J., Mian, S., et al. 2016. "Cooperation among competitors in the open-source arena: The case of OpenStack," in *Proceedings of the International Conference on Information Systems (ICIS 2016),* AIS.

Teixeira, J., Robles, G., et al. 2015. "Lessons learned from applying social network analysis on an industrial Free/Libre/Open Source Software ecosystem," *Journal of Internet Services and Applications* 6 (1) July 2015, 14 July 2015. ISSN: 1867-4828.

Tidström, A. 2014. "Managing tensions in coopetition," *Industrial Marketing Management* 43 (2), 261–271. ISSN: 0019-8501.

Tidström, A. and Rajala, A. 2016. "Coopetition strategy as interrelated praxis and practices on multiple levels," *Industrial Marketing Management* 58, 35–44. ISSN: 0019-8501.

von Hippel, E. 2005. *Democratizing Innovation,* Massachusetts, USA: MIT press.

von Krogh, G. and Spaeth, S. 2007. "The open source software phenomenon: Characteristics that promote research," *The Journal of Strategic Information Systems* 16 (3), 236–253.

Wasserman, S. and Faust, K. 1994. *Social Network Analysis: Methods and Applications,* vol. 8. Cambridge, UK: Cambridge University Press. ISBN: 0521387078.

Wright, N. et al. 2024. "Contributing to Growth? The Strategic Role of Open Source Software for Global Startups," (available online at `https://ssrn.com/abstract=4699182`;).

Zhang, Y. et al. 2021. "Companies' Participation in OSS Development–An Empirical Study of OpenStack," *IEEE Transactions on Software Engineering* 47 (10), 2242–2259.