Chapter 10

Exploring organization through contributions: using Activity Theory for the study of contemporary digital labour practices

David Rozas and Steven Huckle

Abstract

This chapter focuses on peer-production as a form of collaborative digital work, closely allied to crowdsourcing and other contemporary working practices that are mediated by digital platforms. Such platforms are a growing form of digital work; however, they raise complex methodological issues. First, although often a single collaborative platform coordinates groups, work can be distributed globally. Second, multimodal approaches require the researcher to transition between online and offline media. Finally, it can be challenging to identify what is 'work' as activity boundaries are blurred. It is argued that the use of Activity Theory overcomes some of these issues and its utility in an analysis of the production of the open source software, Drupal, is demonstrated, highlighting the potential for Activity Theory to enable cross-contextual comparisons and proposing the concept of 'socio-technical systems of contribution' as a way to understand interactions between networks of collaboration. The limitations of the approach and potential future developments are noted.

Keywords: Activity Theory, Commons-Based Peer Production, Collaborative Economy Digital Labour, Drupal, Free/Libre Open Source Software, Platform Economy

Introduction

Contemporary working practices and the organisational forms that sustain them are changing. Distributed and crowdsourced forms of labour are becoming increasingly important, and major corporations are becoming involved in the digital labour space (Gray and Suri 2019). Documenting and analysing the new organisational forms and practices that are dependent on distributed digital labour has become an urgent task for researchers.

Research on the organisational practices of Free/Libre and Open Source Software (FLOSS) communities provides valuable use-cases for academics interested in studying emergent forms of distributed digital labour. FLOSS features non-discriminatory behaviour and non-restrictive licenses, a practice that requires applications (and associated source code) to be freely redistributed (Gnu.org 2001). The result is that FLOSS software inhabits a space referred to as a *digital commons* (Stadler 2010), which is a freely available and collectively produced repository of code, information and knowledge.

C*ommoning practices*, however, are not exclusive to software. The extension of these commoning practices has resulted in what Benkler (2002; 2006) describes as an emergent mode of production: Commons-Based Peer Production (CBPP), which represents an alternative to the traditional hierarchical modes. Collaborative work in the commons is present in a diverse range of areas (Fuster Morell et al. 2014), including open science, urban commons, peer funding and open design, to name but a few. Hence, whereas capitalism uses the economics of commodification to transform goods into commodities that can be bought and sold on the market, CBPP communities enable counter-commodification. They revolve around an *economy of contribution*, where "people contribute to a project because they want

it to succeed" (Siefkes 2008, 9). The digital commons is inhabited by *digital commoners*, who are a self-determining, politically independent voluntary collective of skilled enthusiasts who cooperate with often ethnically and geographically diverse peers (Kelty 2008). It is a mode of practice which represents an alternative to the traditional hierarchical styles of production that feature within capitalist culture. Often, *digital commoners* become more involved with CBPP projects as they gain trust and, with greater trust, gain greater access to governance processes. Thus, the practices of CBPP create a self-reinforcing loop (Bollier, 2003).

Because of their governance and economic models, CBPP communities present significant differences to wider forms of platform[1] digital labour, such as Amazon Mechanical Turk (MTurk) and Universal Human Relevance System (Microsoft), labour marketplaces that outsource projects to globally diverse workers. However they also present similar characteristics and challenges for researchers interested in studying them. First, the researcher must adopt different types of methodological approaches that combine diverse data sets drawn from online and offline media (Hine 2015) in order to explore organisational forms that can be highly complex and difficult to capture in a conventional model of hierarchies and sub-divisions. Secondly, it can be difficult to distinguish work from recreation because blurred organisational boundaries result in flexible work-lives that reduce the distinction between the private and the professional. Finally, it is challenging to draw conclusions from globally disparate groups, especially when the platforms that mediate these groups are not necessarily open to scrutiny.

We have employed Activity Theory (AT) to explore perceptions of contribution in CBPP (Rozas et al. 2021), a setting in which activities considered as "work" and individual

contributions to that work are particularly blurred as a result of being increasingly created by crowds and communities of diverse participants (Arvidsson and Peitersen 2013). As stated above, CBPP communities focussed on digital commons, such as Drupal, typically rely on an economy of contribution (Wittel 2013) that gives value and recognition to various forms of activity as contributions to the organisation. Consequently an important focus for our AT analysis was "What does it mean to contribute in a community such as Drupal?". We also used AT to identify how two intertwined dynamics of formalisation and decentralisation of decision-making operate in this type of community (Rozas and Huckle 2021). In this chapter, we use examples from Drupal, however researchers focused on other forms of distributed digital labour can benefit from the use of AT as an organising framework for their research.

First, we introduce our research site, Drupal, then we provide an overview of AT and the main concepts of the analytical framework applied during these studies. Subsequently, we describe the application of AT to Drupal. Finally, we reflect on the benefits and limitations of applying AT to the study of organisational practices in broader forms of digital labour.

An overview of Drupal

Drupal is a FLOSS content management system that provides a framework for the development of websites. The history of the Drupal project began in 1998 at the University of Antwerp (Dolin 2011, 822), when two undergraduate students decided to establish a wireless bridge to share their Internet connection and develop a simple content management framework for exchanging messages and news between students. The system was publicly released as FLOSS in 2001 and it has since gone on to power approximately 1.5% of websites worldwide[2]. The main motto of the project, - "come for the software, stay for the community" - reflects the idea that Drupal cannot be understood without considering its

community. The community has experienced significant growth: regular Drupal communitarian events are held all over the world, and more than 1.3 million people have registered some interest in the system at its central mediating platform, drupal.org (Rozas 2017, 85-95), becoming a notable example of the phenomenon of CBPP. Our selection of Drupal as a case study was driven by our interest in understanding how a large and global CBPP community self-organises (Rozas 2017, 99). Our studies (Rozas et al. 2021; Rozas and Huckle 2021) of the Drupal community explored two key types of activities in the community: *the development of source code* and *the organisation of events*.

For the development of source code, we explored the organisational aspects of three types of projects within Drupal. The first type is Drupal's *core projects*. These are the projects that form part of a default installation of Drupal, presenting the basic set of functionalities to develop a website. The second type is *contributed projects*: those that form part of the main collaboration platform, drupal.org, and provide additional features. Core and contributed projects require communitarian peer-reviewing practices in order to become part of the platform. Drupal.org provides tools for the coordination of the development, maintenance and decision-making for these digital commons. The final type of project is *custom projects*, which are projects that have been freely shared and developed in external platforms such as github.com but have not been subject to the required peer-reviewing to include them in drupal.org. In this chapter, we focus particularly on *contributed projects*.

For the organisation of events, we explored the three main types of events present in the community. The first is *DrupalCons*, which are annual conferences attended by thousands of participants. They have a global scope and last almost a week. The second is *DrupalCamps*, which are two-day events organised regionally or nationally by local Drupal communities.

Hundreds typically attend them. The final type includes a wide range of *local events*, such as presentations or social gatherings with other Drupal members. Between 10 and 30 participants typically attend local events. While the events are face-to-face, the coordination to organise them is facilitated through online platforms. These include drupal.org, Drupal websites specifically developed to coordinate each event, or external sites, such as meetup.com. Online platforms are essential to coordinate organisation. For example, they enable peer-reviewing practices in the selection of presentations at events. Other online tools, such as Telegram and Whatsapp groups, are also key to sustain the coordination of Drupal face-to-face events.

Activity Theory

AT is the main theoretical framework which we employed to explore the organisational processes and dynamics of Drupal and which enabled us to study the digital labour practices surrounding both the development of source code and the organisation of events. Rather than a theory in the strictest sense, we used AT as an analytical tool: a lens that helped us to untangle the complexity behind the organisation in large CBPP communities, such as Drupal. As we will further detail in section 5, we found the use of AT particularly useful: firstly, for its capacity to create cross-contextual comparisons by bringing together substantially different organisational processes within the same analytical umbrella. Secondly, AT is agnostic with respect to an a priori macro/micro level of analysis. It instead provides the researcher with a set of concepts which enable analyses at both levels and, if needed, link them. Thirdly, AT incorporates the notion of tension, which is valuable to trace changes in fluid organisational settings, as is the case for CBPP communities. As such we argue that AT

offers a useful framework for researchers interested in understanding the emerging organisational forms of distributed digital labour.

The first generation of Activity Theory

The capacity of AT to identify patterns and establish cross-contextual comparisons is built upon the idea of object-driven activities (Marx 1924, 143-145). Marxism has, therefore, a crucial influence in AT. Marx understood the processes of production and transformation as historical phenomena dependent on social practices, whereby the subject produces itself by producing the object and thereby, transforms the object's nature. Such processes might introduce several systemic inner contradictions and tensions, but these too may become a force for development (Hegel 1975). Hence, Marx (1924, 143-145) considered the relationship between objects and subjects as critical to understanding transformation. Subsequently, the first generation of AT was built on Vygotsky's (1978) notion of *mediated action*, whereby culturally meaningful artefacts condition the individual because when someone develops activities in collaboration with other humans, they internalise social norms and modes of behaviour. Leont'ev (1978) developed Vygotsky's ideas and established the concept of activity as a unit of analysis. As is the case with other socio-cultural perspectives (Kaptelinin 2012), Leont'ev assumed the social nature of the human mind, as well as its inseparability from the activity. For Leont'ev (1978), there was no activity without an object. Thus, activities influence an object's characteristics and vice versa. AT incorporates these notions of relationships between elements and tensions between them as conceptual elements for the study of activity systems.

The second generation of Activity Theory

In the 1980s, Leont'ev's ideas became known to a new generation of academic theorists, such as Engeström (1987). The first generation of AT did not present a conceptual model for the study of collective activities. Engeström (1987) proposed a new model which included them. Engeström's proposal became known as the second generation of AT (2GAT). Figure 1, the model of *human activity system*, represents the outcome of an activity through six interrelated elements that account for social relations:

1. **Subject**. The actors who perform the activity and who are subject to the internalisation processes.
2. **Tools**. The mediating artefacts employed by the actors in the system. Cultural factors influence tools and they change according to the accumulated experience.
3. **Object**. The element towards which the activity is directed. It has social and cultural properties. The object is transformed as the activity progresses.
4. **Rules**. The rules that regulate the subject's actions toward an object and their relations with other participants. They can be explicit or implicit.
5. **Community**. The totality of actors sharing an interest in the same object.
6. **Division of Labour**. A representation of the distribution of processes between actors.

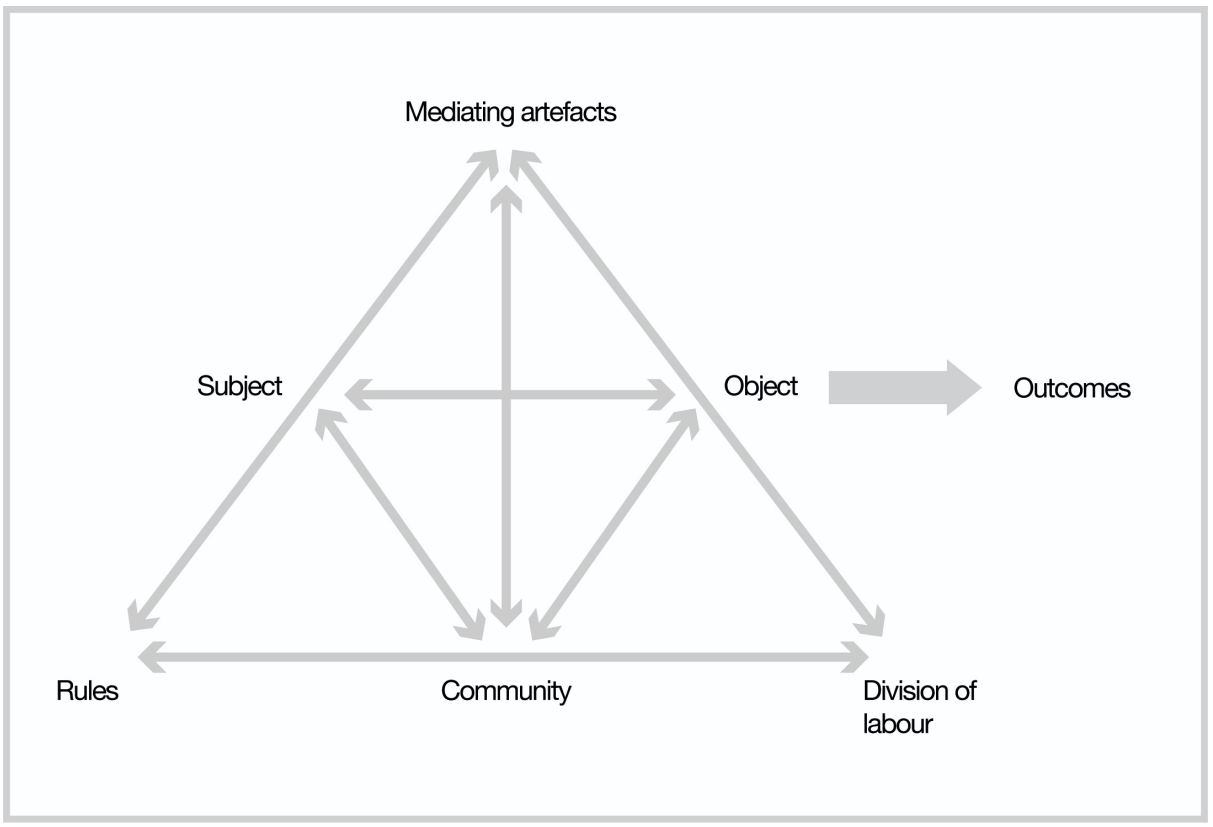

Figure 1: Model of the human activity system of the 2GAT. Adapted from Figure 2.6 in Engeström (1987, 78).

For example, imagine an activity that consists of the redesign of the User Interface (UI) of a particular software application. Here, the object of the activity is the current design of the computer interface (Kaptelinin 2012). A community forms to complete the action, which involves a division of labour, because the activity requires, for example, project managers, developers and UI designers. A UI designer might use a set of artefacts to work on the transformation of the object, from hardware capable of rendering sophisticated graphics, to software for designing those graphics. Additionally, the UI designer might interact with the community through implicit and explicit rules; for instance, they may attend project meetings and receive a salary for their efforts. Overall, the coordinated work of the team produces a set of new outcomes - in this instance, a new user interface.

The third generation of Activity Theory and beyond

During the 1990s, researchers began to realise the necessity of developing the next generation of AT. Engeström (1999) proposed a new approach which captures the interactions between several human activity systems. The minimal model, with two activities, is shown in Figure 2. The interactions between activity systems result in the sharing of (often fragmented) objects, enabling the researcher to study forces for development that result from the inherent tensions between systems (Engeström 2009). Thus, the third generation of AT (3GAT) continued to be inspired by Marx, who, as described above, also considered the development potential of processes that introduce systemic inner contradictions (Hegel 1975).

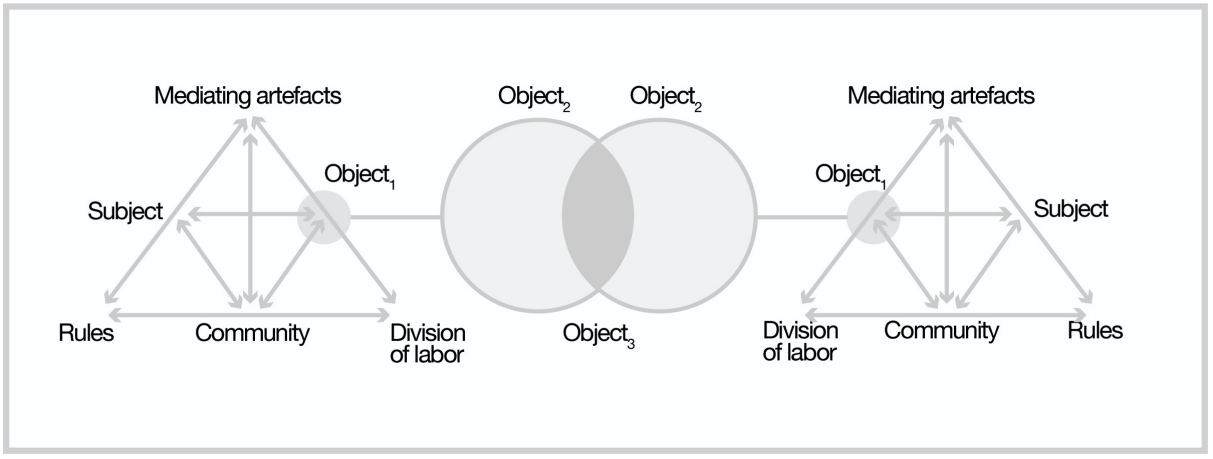

Figure 2: The minimal model of the 3GAT. Adapted from Figure 3 in Engeström (2001, 136).

Consider the redesign of the UI, described above, and that the expected outcome (a new UI), is part of an effort to develop a new version of a much broader application. This requires the integration of the UI with outcomes from other activity systems, such as back-office databases. This introduces inherent tensions as the different systems attempt to provide their solution to the whole. However, that leads to a development dynamic that delivers a solution

because the interaction between the UI designer and the database administrator results in software that meets requirements satisfying all parties.

In the last few decades the emergence of new forms of organisation, such as CBPP, distributed work or crowdsourcing, opened up a debate between activity theorists on the need to rethink the shape of these activity systems. Engeström (2006) contributed the concept of *runaway object* to the debate, which he later developed (Engeström 2009). Engeström (2009, 306) cites four prerequisites for such objects:

1. They must have intrinsic properties that transcend the utilitarian profit motive
2. They must yield useful intermediate products, yet remain incomplete
3. They must be visible, accessible and cumulable so that participants return time and again
4. There must be useful feedback from and exchange among the participants

Engeström highlighted the notions of negotiation and peer review as key to understanding the coordination mechanisms and the new forms of organisation that emerge in the constant development of these runaway objects. Figure 3 depicts this new model, which acknowledges and highlights that the boundaries and structures in these new forms of organisation, such as those in CBPP, are not so clear: they are subsumed by the object, rather than the other way around.

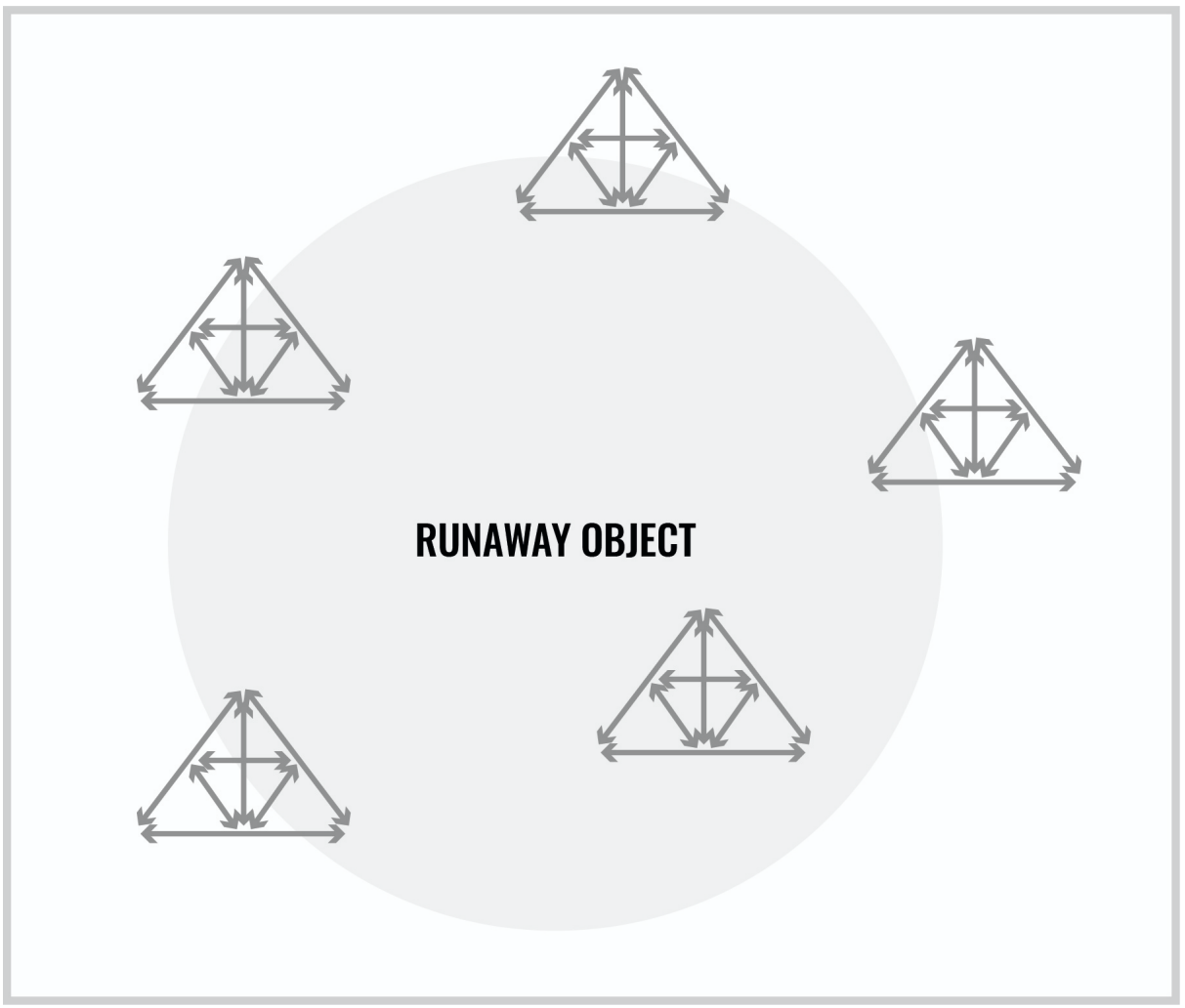

Figure 3: A runaway object with several activity systems. Adapted from Figure 19.2 in Engeström (2009, 306).

Hence, for Engeström, the challenge for the future of AT involves integrating analytical tools that capture a multitude of issues of subjectivity and their multiple interconnected human activity systems which remain valuable to study the changes in more blurred and distributed forms of organisation, such as those in Drupal.

Applying AT to the study of peer-production: the case of Drupal

In this section, we discuss the application of AT in our study of Drupal. To investigate the digital labour practices of Drupal, we used an ethnographic approach combining: three years of participant observation; documentary analysis of an archive of 8,613 documents[3]; and 15

semi-structured interviews (Rozas 2017, 125-158). The first author (Rozas 2017, 129-130) was as an "insider researcher" (Brannick and Coghlan 2007):

> "the researcher-author describes a cultural setting to which s/he has a 'natural access', is an active participant, more or less on equal terms with other participants [and] then works and/or lives in the setting and then uses the experiences, knowledge and access to empirical material for research purposes" (Alvesson 2003, 174).

Due to the digital nature of the main object sustained by the community - software - and its size and global nature, a large amount of the day-to-day activity in Drupal is, unsurprisingly, carried out through online media. For that reason, the research initially drew on virtual ethnographic methods (Hine 2000). Nevertheless, the relevance that offline activities have in the community emerged during the course of the study and it was consequently concluded that this research required immersion and participation in both online and offline activities. requiring a breakdown of the traditional dualism (Orgad 2005). This approach was congruent with that already taken in similar studies, as in Coleman's (2013) study of FLOSS communities and hacker culture.

AT was employed to constantly inform the research methodologically. The online/offline distinction emerged as blurred and continuous, rather than binary (Rozas 2017, 127-128). However, conversely, the definition of 'contribution activity' as the main unit of analysis facilitated a clearer distinction with respect to immersion and participation in the communitarian activities. Furthermore, it led us to the question: "What does it mean to contribute in a community such as Drupal?". We applied the 2GAT model for the study of such contributions. We explore this further in the next section. Additionally, AT does not

consider an a priori micro/macro divide (Miettinen 1999), which allowed us to expand the level of analysis to the study of the organisational environments in which these activities were taking place. For that purpose, we drew on the concept of runaway object from the 3GAT, and this is illustrated in the following section.

Application of AT at a micro-level: writing code and organising events

For the analysis at the micro-level, this study drew on the model of the human activity system from the 2GAT. Hence, during the first stage of the research it was necessary to study the notion of contribution within Drupal in detail and to include within that the less visible forms of contribution. Furthermore, to understand how a vast global community organises, it was also necessary to include other elements and factors (e.g. processes, dynamics and structures among others) which surround such contributions.

Using the 2GAT model as a lens, we analysed some of these contribution activities in-depth, including the relationships between the artefacts employed for collaboration, the roles played by its members (division of labour) and the implicit and explicit rules. The use of the activity system as a unit of analysis enabled the incorporation of these notions as part of a dynamic phenomenon (Uden et al. 2007), avoiding simple monocausal explanations in the study of CBPP.

Figure 4 depicts an example of this application of the model of activity for the study of the development of contributed projects in Drupal, in which the elements of the activity were defined as follows:

1. **Subject**. The maintainers of the contributed project. In other words, the Drupal members responsible for coordinating development and maintenance.

2. **Tools**. The mediating artefacts used by the maintainers and the rest of the members of the community, such as issues lists. This is where users report bugs, tasks can be assigned, or new features are requested. Thus, issues lists are artefacts that function as a coordination tool for maintainers and enable technical discussions and decision-making about how to address issues. Other types of mediating artefacts are chat channels, email, social networks, such as Twitter, or Drupal discussion groups[4].
3. **Object**. The contributed project.
4. **Rules**. Examples of explicit rules are community-agreed coding standards[5] and guidelines for contribution[6]. For example, if a project is to move from a custom project to a contributed project, developers must follow a peer-reviewing process called the *Project Application Process* (Rozas and Huckle 2021, 212). Examples of implicit rules are those employed by maintainers for the evaluation of contributions by other Drupal members who do not have direct permission to make changes to the project.
5. **Community**. All the members of the Drupal community. Their involvement in an individual project often arises because they are users of it. They can make use of tools as mediated artefacts to provide feedback, supply patches to solve bugs or extend the project's features.
6. **Division of Labour**. The different roles typically associated with contributed projects; for example, developers or UI designers. Again, tasks can be allocated by using the tools as mediating artefacts.

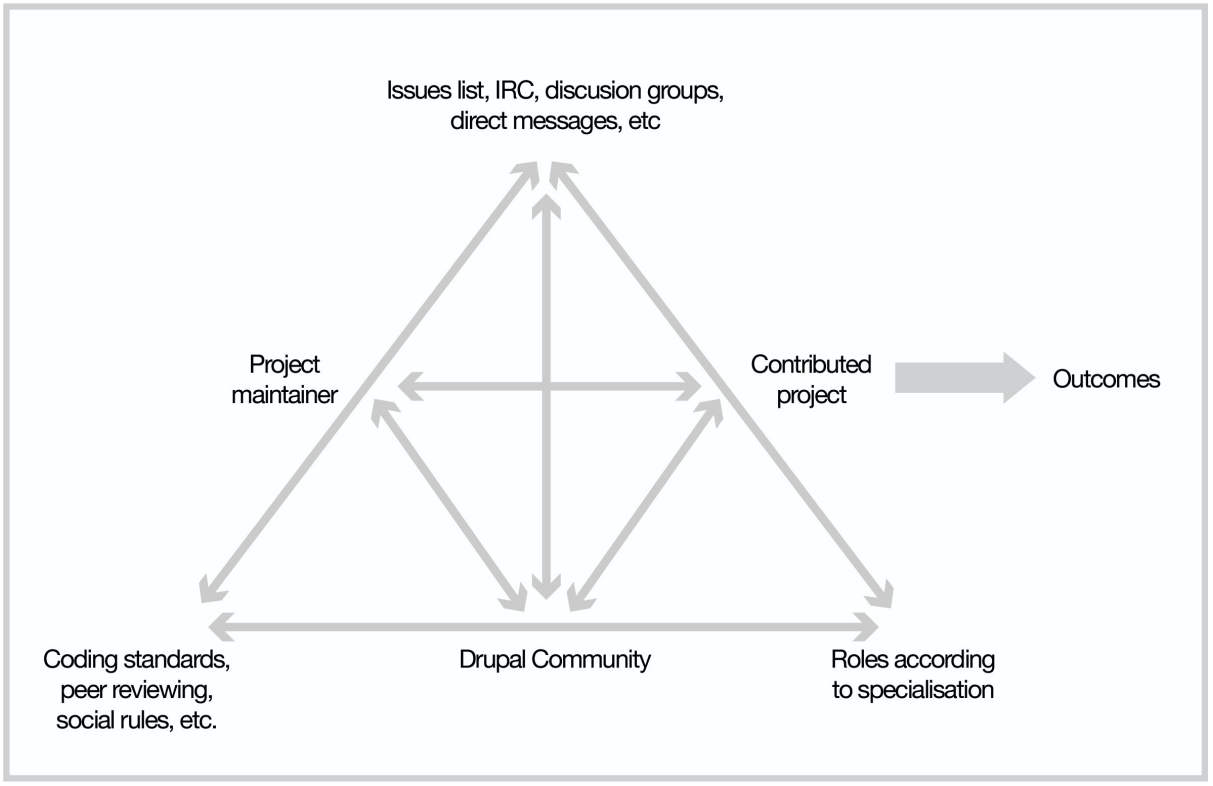

Figure 4: Conceptualisation of the development of contributed projects from an AT perspective. Adapted from Figure 3.7 in Rozas (2017, 119)

Similarly, we employed the model of the activity system for the study of a significantly different activity, the organisation of events. For example, Figure 5 provides an example of the application of the model to study the organisation of a DrupalCamp, a type of event organised by local communities consisting of a conference typically lasting two or three days:

1. **Subject**. The participants in the event.
2. **Tools**. The mediating artefacts used to coordinate the event. For example, the platforms employed to coordinate the event, mailing lists and specific discussion groups for the event at groups.drupal.org.
3. **Object**. The DrupalCamp.

4. **Rules**. Examples of explicit rules are the selection criteria for any presentations made at a DrupalCamp[7], as well as DrupalCamp codes of conduct, which outline the shared ideals and values of the community[8]. Examples of implicit rules are social norms related to the reputation of a subject in the community. For example, to be able to organise a DrupalCamp, community members require a significant degree of legitimacy.
5. **Community**. All the members of the Drupal community.
6. **Division of Labour**. The different roles of the participants of the event; for example, session reviewers, presenters and the DrupalCamp attendees themselves.

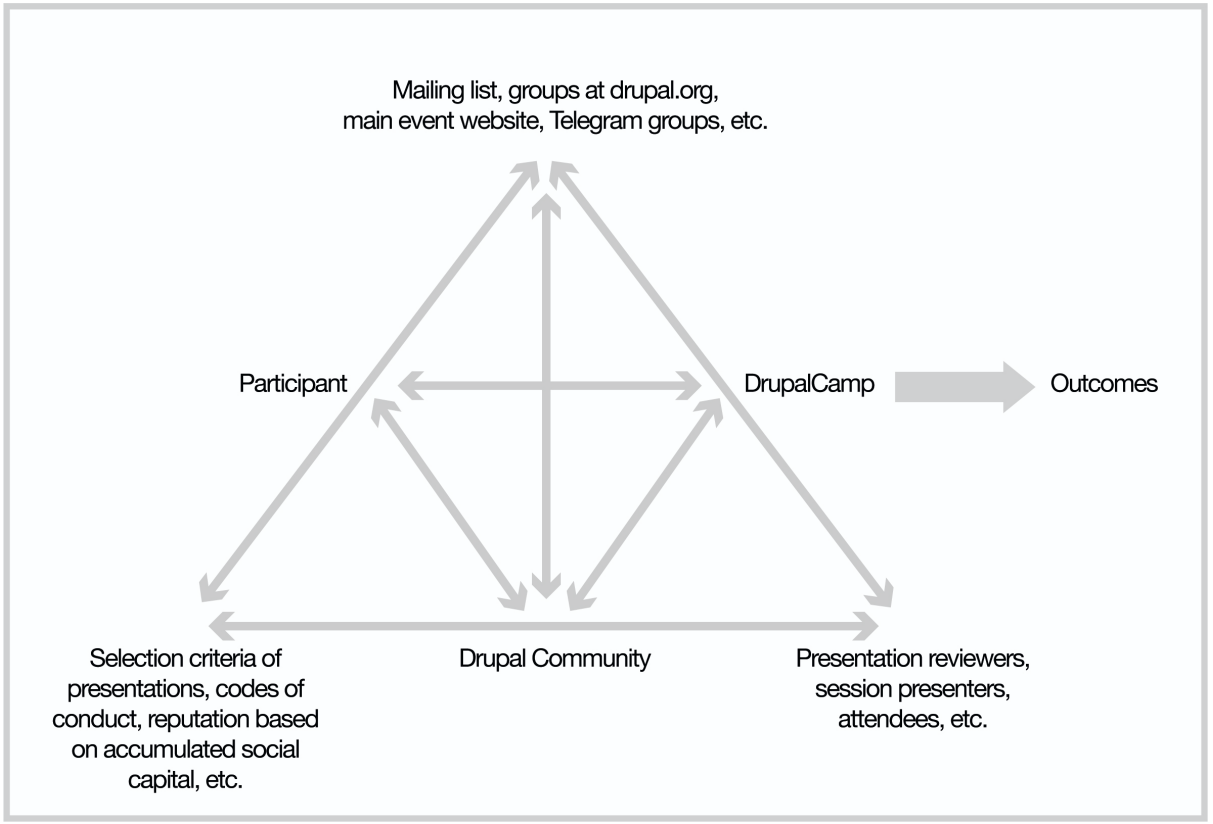

Figure 5: Conceptualisation of the organisation of DrupalCamps from an AT perspective. Adapted from Figure 3.8 in Rozas (2017, 120)

We draw on the previous two examples to illustrate the very different nature of these two activities: maintaining the source code of a contributed project; and organising a DrupalCamp. The application of AT to both allows us to establish cross-contextual comparisons because all activities are categorised using consistent concepts. For instance, we could make comparisons between the emergence of peer-reviewing practices to assess quality and acceptance in the main platform of source code and the practices for accepting a presentation at a community event, which are both categorised as rules within AT. We can also draw comparisons between activities in which the focus is the same, but whose organisational characteristics differ; for example, between having code officially accepted in core projects and in contributed projects.

To achieve this, we employed the previously described AT elements as initial analytical categories for each activity, and then explored the relationships between them. An example of the type of relationships explored are those between: the artefacts employed for collaboration (e.g. the issues list of a contributed project in Drupal.org or the website to coordinate the organisation of a DrupalCamp); the division of labour (e.g. Drupal roles played by participants such as being a maintainer of a contributed project or being a member of the peer-reviewing team for presentations submitted at a DrupalCamp); and the implicit and explicit rules around such activities (e.g. coding standards for contributed projects, or codes of conduct for DrupalCamps). Establishing such cross-contextual comparisons between relationships led us, for example, to find similarities in the emergence of peer-reviewing practices to assess the quality of source code and the submission of presentations to communitarian events (Rozas and Huckle 2021). Subsequently, following these practices, we found similar organisational characteristics, such as degrees of specialisation, legitimacy and

perceived value (see Tables 3 and 4 in Rozas and Huckle 2021, 209-210), despite the different nature of the activities.

The employment of the 2GAT model also proved useful for incorporating findings from previous studies of Drupal. For instance, the tensions between designers and developers described by Zilouchian-Moghaddam et al. (2012) were incorporated initially as emerging from the division of labour. As a result, we could study their impact on other entities of the 2GAT model. For example, how contributions are represented or not in the collaboration artefacts, such as the official user profiles at drupal.org. For instance, one category was "object-oriented" contributions, encompassing all activities whose main outcomes from an AT perspective are typically directed towards digital commons such as source code, documentation and translations. The second category was "community-oriented" contributions, in which the main outcomes from an AT perspective are directed towards the community. These categories helped us to identify significant differences between the indicators which measure and aggregate forms of value in CBPP (Rozas et al. 2021). These findings led us to argue (Rozas et al. 2021) for a need to broaden our understanding of contribution in CBPP communities, so that traditionally less visible forms of work are acknowledged as of value in the communities and visibilised in the online platforms that support peer production.

Application of AT at a macro-level: Drupal as a "runaway object"

AT does not establish an a priori micro/macro divide (Miettinen 1999). Instead, as above, AT provides researchers with a set of analytical concepts, such as the main elements of the 2GAT model, to foster conceptual connections in the context of their studies. In CBPP communities, such as Drupal, "the boundaries and structures of activity systems seem to fade away"

(Engeström 2009, 309). Because their simultaneous reciprocal processes are multidirectional and multilayered, their boundaries and structures are often difficult to distinguish and do not usually have a single stable centre. Instead, this mode of production requires and creates "bounded hubs of concentrated coordination efforts" (Engeström 2009, 310). In order to connect the micro and macro aspects of this case study, this research explored these bounded hubs shedding light on how a large global CBPP community such as Drupal organises itself (Rozas and Huckle 2021). To achieve this, we carried out a first step in the conceptualisation by framing the whole of Drupal as a "runaway object", as depicted in Figure 6.

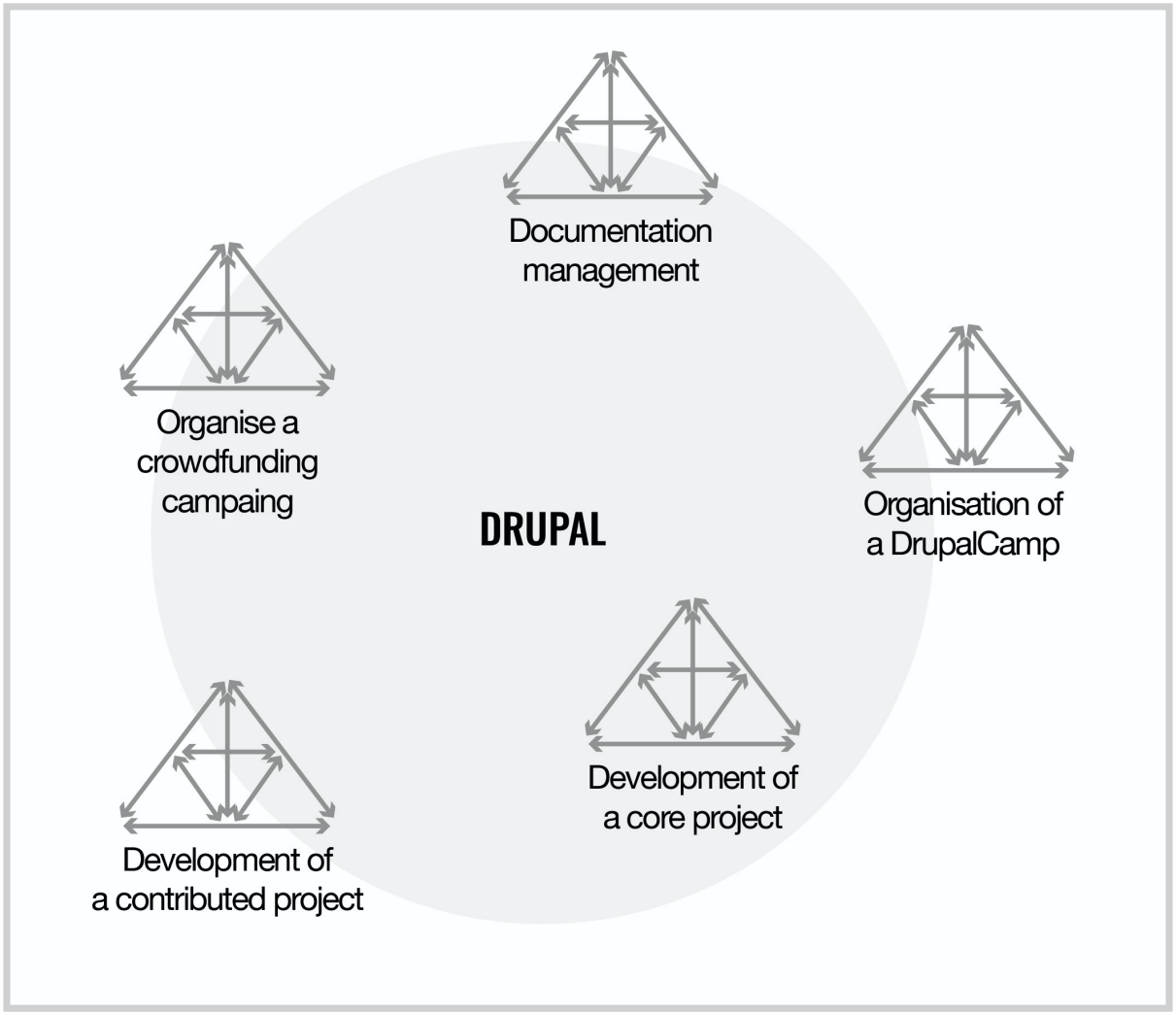

Figure 6: Interconnecting Drupal activities result in Drupal as a runaway object. Adapted from Figure 1 in Rozas et al. (2015, 5).

We conceptualised Drupal (Rozas et al. 2015) in line with Engeström's (2009, 306) set of prerequisites for runaway objects, discussed earlier. As with any other FLOSS project, Drupal transcends utilitarian motives and bases its sustainability on collaborative production (requisites 1 and 4). The nature of the project is dynamic, and it is in a constant process of change (requisite 2). Additionally, the main production processes and project outcomes are also visible, cumulable and accessible all of the time (requisite 3).

However, while the concept of runaway object operates as a nexus allowing an initial connection between organisational micro and macro aspects, this study required a more precise definition of these "bounded hubs of concentrated coordination efforts". This reflects a critique of the 3GAT (Spinuzzi 2011). Spinuzzi (2011) proposes instead to corral the *runaway object* in ways which allow it to be enriched appropriately. In other words, to attune the definition of what the runaway object is in order to contextualise it to the case study. Overall, this need to provide a more accurate definition is in line with the ongoing efforts of activity theorists to rethink the 3GAT to accommodate the changes in newer forms of organisation, characterised by a distributed workforce and the predominance of knowledge work, as is the case in peer-production.

Being aware of this conceptual issue as described by Spinuzzi in the application of the 3GAT, we attuned to it by providing a more accurate definition of what Engeström's (2009) bounded hubs of coordination were in the context of Drupal. As a result, we brought together the 3GAT's concept of a runaway object and the concept of socio-technical systems from organisational theory (Trist 1981). The result was the development of the concept of a '*socio-technical system of contribution'*, in the context of peer-production, defining it as (Rozas 2017, 122):

"A set of interacting parts, including people, software, hardware, procedures or rules among others, which form a complex whole that revolves around networks of human activity systems which are perceived contributions within the community and share a similar main focus of action."

For example, while we previously presented our conceptualisation of the development of a contributed Drupal project as a human activity system from a 2GAT perspective, the network of thousands of contributed modules in Drupal.org can be conceptualised as a socio-technical

system of contribution within the community. Similarly, while we employed the model of the human activity system for the analysis of the organisation of a DrupalCamp, the network of DrupalCamps was framed as a socio-technical system of contribution. Figure 7 provides an illustration of the application of this concept, in which the human activity systems are grouped according to the socio-technical system they belong to.

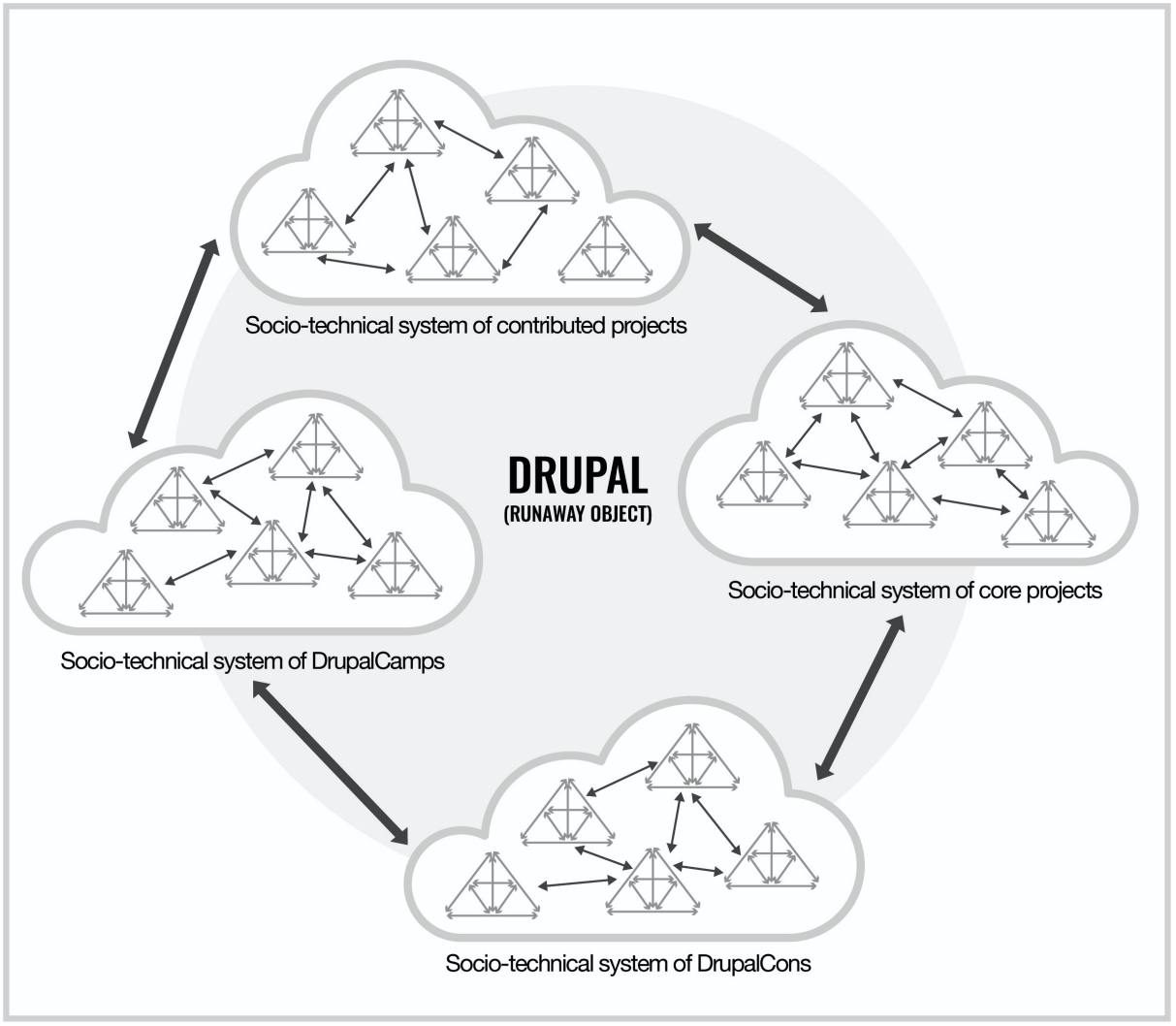

Figure 7: Conceptualisation of Drupal (runaway object) as a set of socio-technical systems of contribution. Adapted from Figure 3.9 in Rozas (2017, 123).

The illustration also shows how the human activities within these groups are interconnected and displays the interactions between different socio-technical systems of contribution within the runaway object of Drupal.

The notion of socio-technical systems of contribution enabled them to be connected to the macro levels at which they occur, as well as the tensions between different systems. The result was to enable an analysis not exclusively focussed on the workings of contribution activities themselves (micro-level), but also on the interactions between the networks they form as socio-technical systems of contribution (macro-level). This led us to study how these socio-technical systems of contribution emerge, evolve, interact with each other, and are shaped by different organisational dynamics (Rozas and Huckle 2021).

Reflections

Here, we discuss the main insights concerning the application of AT to broader forms of digital labour beyond CBPP. We subsequently provide an evaluation of the general challenges tackled during the undertaking of this research and discuss specific limitations concerning the application of AT.

Insights from applying Activity Theory to digital work

We found three main benefits of the application of AT to the study of digital labour. Firstly, AT facilitates the identification and deconstruction of activities into several components in ways that allow strategic and cross-contextual comparisons. This study used 2GAT to describe the elements of contribution activities within the development of source code and the organisation of face-to-face communitarian events. Although they are substantially different organisational activities, the use of AT provided a useful lens to compare them. This

capacity to facilitate cross-contextual comparisons of AT can also be useful when researching broader forms of digital labour.

Secondly, the use of AT as an analytical framework to explore collaboration was valuable in order to connect the study of micro and macro organisational aspects of peer-production practices. AT's set of analytical concepts and its tools' flexibility proved to be useful in this case study to connect actions at different levels, from carrying out a code commit or submitting a presentation for an event (micro-level), to the whole socio-technical systems of contribution for the development of software projects and the organisation of communitarian events (macro-level) and their different peer-reviewing practices. This helped us to trace the evolution of organisational structures, and to identify two intertwined dynamics of formalisation and decentralisation in the case study (Rozas and Huckle 2021). In this respect, AT provides a useful lens to cope with complexity and untangle the dense and multidirectional dynamics which lie within a broader range of contemporary working practices which are increasingly mediated by digital platforms. Similarly to this case study, a researcher exploring a platform such as MTurk could employ AT to connect actions such as "getting a task" with the socio-technical systems and the dynamics to distribute value which operate in such scenarios.

Thirdly, the use of the model of activity as a unit of analysis also helped in reconsidering the notion of contribution in CBPP communities, leading to the exploration of certain contribution activities which have traditionally remained less visible in FLOSS and CBPP literature, such as the training and mentoring of community members (Rozas et al., 2021). These perceptions of what can be considered contribution contrast with those recognised in the main collaboration platform. This issue was identified while exploring a tension between

the division of labour (the different roles present in the Drupal community) and the artefacts (in this case how drupal.org profiles record some of these contributions). The result was a call to broaden our understanding of the notion of contribution in CBPP communities, incorporating new kinds of contributions customarily left invisible (Rozas et al. 2018). In this respect, AT offers a powerful lens for the study of similar, increasingly blurred boundaries on a broader spectrum of digital labour for what is or is not considered work. Researchers interested in the study of such boundaries in crowdsourcing platforms, such as MTurk or Upwork, might similarly identify tensions and interactions from which to explore the practices behind more extensive forms of digital labour. Examples of these tensions and interactions could be those between the division of labour (e.g. owners and workers of the platform) and the rules (e.g. those embedded in the algorithms employed to allocate work). From these interactions between AT elements, researchers can trace the emergence of organisational structures associated with them, or perhaps experiment with a gradual shifting of the power of decision-making about the rules that determine the distribution of value to the members of such platforms[9].

Challenges

An important source of challenges derived from the first author's position as a researcher, since he was already an active member of the Drupal community for over three and a half years before embarking in this research. This previous experience proved valuable for more rapid access to the community: from a faster understanding of the meanings around the software and the community to practicalities for entering the field site and gaining access to certain activities. This previous experience came at the cost, however, of having to address challenges related to the dynamics of insider research, such as role duality and preunderstanding (Brannick and Coghlan, 67-71). Regarding the latter, for example, the fact

that the first author's previous experience within the Drupal community was mainly as a software developer was identified as a potential source of partiality. Consequently, an effort was made to have a wider understanding from the perspectives of Drupalistas with different roles (the division of labour from an AT perspective) during participant observation, the selection of interviewees and the documentary analysis.

Another source of challenges derived from the ethical aspects arising from the use of ethnographic methods. Overall, this involved a constant assessment of the possibility of new ethical issues arising during its course. When new issues were discovered, actions were designed and implemented. For instance, with regards to the type of access while conducting participant observation, there was a constant effort to undertake it in the most overt way possible, but being aware of the limitations. Examples of these actions include for the first author to present himself at local events as "a Drupalista who was currently studying the Drupal community for his PhD research" (Rozas 2017, 137-143, 153-157), and making his role as researcher visible in the digital platforms employed by the community[10]. Furthermore, efforts were made to expose his role as a researcher to the global community to the highest degree possible. Examples include: participation in a podcast and interviews in communitarian channels[11], and dissemination of the research findings at local events, DrupalCamps and a DrupalCon[12]. Notwithstanding, limitations existed: the role could not be qualified as that of a completely known observer, since some attendees were not aware of the first author's role, especially while participating in large international events, such as DrupalCons, with thousands of attendees.

Limitations from the application of AT

As with any other analytical lens, the application of AT is not free from limitations (Nardi 1996, 63-64). In the context of its application for this case study, we identified two limitations of particular relevance when reflecting on a more generic application to studies on digital labour.

Firstly, defining human activity as the primary unit of analysis and observation has an impact, as any other choice would, on the emergence of relevant thematic areas explored during the overall processes of data collection and content analysis. For example, a deeper understanding of certain organisational aspects, such as the role of private companies to influence the direction of Drupal, emerged only tangentially. A different choice of unit of analysis would have highlighted this type of organisational aspect. The study could have focused on participants carrying out changes to contributed projects under company sponsorship. That may have helped uncover the role of companies in shaping the direction of large FLOSS projects (and CBPP projects in general), thereby framing Drupal as a *community of companies* (González-Barahona et al. 2013) rather than a community of individuals. Researchers drawing on AT for the study of other forms of digital labour should be carrying out a continuous process of review and self-reflection in order to identify the impact of the choice of unit of analysis and observation on their data to implement adaptations in the research design accordingly.

Secondly, the flexibility offered by AT as an analytical framework can, at times, be a "double-edged sword". As we have seen, the notion of runaway object can lead to a need for methodological and theoretical contractions of the object (Spinuzzi 2011). Researchers studying broader forms of digital labour might face similar challenges. Overall, this type of

issue connects with the efforts (e.g. Spinuzzi 2020) of activity theorists who acknowledge that the ongoing changes in new forms of organisation, such as CBPP and broader forms of digital labour, are so profound that they require a radical rethinking of the models. The identification of limitations, drawing on case studies of these novel forms of organisation, is essential to sustain the path to an emerging fourth generation of AT (Spinuzzi and Guile 2019). The application of AT to more comprehensive cases of digital labour could contribute, in that respect, to better identifying such limitations which would be useful to reflect and refine a new generation of AT conceptually into the current context, as has happened in previous periods (e.g. Kuutti 1996).

Conclusion

In this chapter, we have found three benefits of particular relevance for the use of Activity Theory for research on digital labour. First, AT is useful to identify, deconstruct and compare collaborative activities mediated by digital platforms. Secondly, AT possesses a high degree of flexibility for the study of organisational practices at different levels. Rather than imposing a predefined micro/macro divide, AT provides a series of concepts which researchers can adapt according to specific digitally-mediated contexts. Thirdly, AT offers a suitable lens to deal with the blurred organisational characteristics present in emerging forms of digital labour. Key notions of AT, such as its models, entities and tensions, provide useful categories from which to explore practices in digital labour.

Given the benefits we demonstrate, we hope other digital labour researchers employ AT, contribute to its development and identify its limitations. In this way, AT will continue to be an invaluable analytical lens for the study of organisational practices.

Acknowledgements

This is a draft of a chapter that has been accepted for publication by Oxford University Press in the book "Research Methods for Digital Work and Organization" edited by Gillian Symon, Katrina Pritchard and Christine Hine published in 2021 (ISBN: 9780198860686 - https://global.oup.com/academic/product/research-methods-for-digital-work-and-organization-9780198860686?cc=us&lang=en&).

This work was partially supported by the project P2P Models (https://p2pmodels.eu) funded by the European Research Council ERC-2017-STG (grant no.: 759207) and by the project Chain Community funded by the Spanish Ministry of Science, Innovation and Universities [grant no.: RTI2018-096820-A-100].

We would like to thank Christine Hine, Gillian Symon and Juan Pavón for their helpful comments and suggestions. We also thank Elena Martínez Vicente for her help editing the diagrams. Finally, we would like to thank Tabitha Whittall for her help in copy-editing and proofreading this chapter.

Endnotes

1. See Gray and Suri (2019) for an in-depth account of the increasing "under-the-hood" task-based and content-driven work which is mediated by platforms such as MTurk (Amazon), Universal Human Relevance System (Microsoft) and UpWork, among others.

2. Usage statistics and market share of Drupal - https://w3techs.com/technologies/details/cm-drupal/all/all, accessed on 19th October 2020. This percentage includes well-known websites with complex architectures and high loads of traffic, such as mtv.co.uk and economist.com.

3. The 8,613 documents included a significant amount of material collected from an open Drupal archive. They were processed using scripts available at https://davidrozas.cc/lab/drupal_planet_archive.php.

4. Drupal discussion groups are available at https://groups.drupal.org/.

5. Drupal coding standards are available at https://drupal.org/coding-standards.

6. The guidelines for contribution are available at https://drupal.org/contribute/development.

7. For an example of speaker guidelines used at *DrupalCamp Spain 2012*, see http://2012.drupalcamp.es/en/node/23.html.

8. For an example of a code of conduct used at *DrupalCamp Brighton 2015*, see http://www.drupalcampbrighton.co.uk/content/code-conduct.

9. See https://p2pmodels.eu/exploring-models-for-a-more-cooperative-distribution-of-tasks/ for an example of these types of initiatives in the context of undergoing research with blockchain technologies.

10. For examples, see https://www.drupal.org/u/drozas (profile in Drupal's main platform), https://events.drupal.org/u/drozas (profile in platform that supports the organisation of DrupalCons), and https://www.meetup.com/London-Drupal-Pub-Meet/members/122334662/ (profile in platform employed to organise events in London).

11. For examples, see https://www.drupaleasy.com/podcast/2015/10/drupaleasy-podcast-163-drupal-potato-david-rozas-open-source-contributing (podcast) and https://www.youtube.com/watch?v=DrbJ9xwSstE (interview in DrupalCamp London 2016).

12. For examples for each type of event, see https://vimeo.com/131301737 (Drupal Show and Tell in London, May 2015), https://drupalcampnorth.org/session/keynote-talk-silver-code-gold-contribution-beyond-source-code-drupal (keynote in DrupalCamp North 2015), and https://www.youtube.com/watch?v=TdEVaOjL20s&t=15m37s (keynote in DrupalCon Barcelona 2015).

References

Alvesson, M. 2003. Methodology for close up studies – struggling with closeness and closure. Higher Education, 46(2): pp.167–193. doi:10.1023/A:1024716513774

Arvidsson, A., and Peitersen, N. 2013. *The ethical economy: Rebuilding value after the crisis*. Columbia University Press.

Benkler, Y. 2002. Coase's Penguin, or, Linux and "The Nature of the Firm". The Yale Law Journal, 112(3): 369–446.

Benkler, Y. 2006. *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.

Bollier, D. 2003. The Rediscovery of the Commons. *UPGRADE*, *4*(3). Retrieved from http://dlc.dlib.indiana.edu/dlc/bitstream/handle/10535/4979/up4-3Bollier.pdf?sequence=1&isAllowed=y

Brannick, T. and Coghlan, D. 2007. In defense of being "native": the case for insider academic research. Organizational Research Methods, 10(1): pp.59–74. https://doi.org/10.1177/1094428106289253

Coleman, E. G. 2013.Coding freedom: The ethics and aesthetics of hacking. Princeton University Press.

Dolin, K. Q. 2011. Drupal's Story: A Chain of Many Unexpected Events. In B.Melançon (Ed.), The Definitive Guide to Drupal 7 (pp. 821–833). Apress.

Engeström, Y. 1987. *Learning by Expanding. An Activity-Theoretical Approach to Developmental Research*. Cambridge University Press.

Engeström, Y. 1999. Activity Theory and Individual and Social Transformation. In Y. Engeström, R. Miettinen, and R.-L. Punamäki (Eds.), *Perspectives on Activity Theory* (pp. 19–38). Cambridge University Press.

Engeström, Y. 2001. Expansive Learning at Work: Toward an activity theoretical reconceptualization. Journal of Education and Work, 14(1): pp.133–156. https://doi.org/10.1080/13639080020028747

Engeström, Y. 2006. From Well-Bounded Ethnographies to Intervening in Mycorrhizae Activities. Organization Studies, 27(12): pp.1783–1793. https://doi.org/10.1177/0170840606071898

Engeström, Y. 2009. The Future of Activity Theory: A Rough Draft. In A. Sannino, H. Daniels, and K. D. Gutierrez (Eds.), *Learning and Expanding with Activity Theory* (pp. 303–328). https://doi.org/10.1017/CBO9780511809989.020

Fuster Morell, M., Martínez, R., and Maldonado, J. 2014. Mapping the common based peer production: A crowd-sourcing experiment. *The internet, policy & politics conferences*.

Gnu.org. 2001. What is free software? Retrieved February 28, 2020, from https://www.gnu.org/philosophy/free-sw.en.html

González-Barahona, J. M., Izquierdo-Cortázar, D., Maffulli, S., and Robles, G. 2013. Understanding How Companies Interact with Free Software Communities. IEEE Software, *30*(5): pp.38–45. https://doi.org/10.1109/MS.2013.95

Gray, M. L., and Suri, S. 2019. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Eamon Dolan Books.

Hegel, G. W. F. 1975. *Lectures on the Philosophy of World History*. Cambridge University Press.

Hine, C. 2000.Virtual ethnography. Sage.

Hine, C. M. 2015. *Mixed Methods and Multimodal Research and Internet Technologies* (S.N. Hesse-Biber and R. B. Johnson, Eds.). https://doi.org/10.1093/oxfordhb/9780199933624.013.31

Kaptelinin, V. 2012. Activity Theory. In M. Soegaard and R. F. Dam (Eds.), *The Encyclopedia of Human-Computer Interaction*. Retrieved from https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/activity-theory

Kelty, C. M. 2008. *Two bits: The cultural significance of free software*. Retrieved from http://twobits.net/download/index.html

Kuutti, K. 1996. Activity theory as a potential framework for human-computer interaction research. In B. A. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp. 17–45).

Leont'ev, A. N. 1978. *Activity, consciousness, and personality*. Prentice-Hall.

Marx, K. 1924. Theses on Feuerbach. In Marx-Engels Archives (Vol. 1). Marx-Engels Institute, Moscow.

Miettinen, R. 1999. The riddle of things: Activity theory and actor-network theory as approaches to studying innovations. Mind, Culture, and Activity, 6(3): pp.170–195. https://doi.org/10.1080/10749039909524725

Nardi, B. A. 1996. Studying context: A comparison of activity theory, situated action models, and distributed cognition. In B. A. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp. 69–102).

Orgad, S. 2005. From online to offline and back: Moving from online to offline relationships with research informants. In C. Hine (Ed.),Virtual methods (Chap. 4, pp. 51–66). Berg.

Rozas, D. 2017. *Self-organisation in Commons-Based Peer Production. Drupal: "The drop is always moving"*. https://openresearch.surrey.ac.uk/esploro/outputs/doctoral/Self-organisation-in-commons-based-peer-production--Drupal--the-drop-is-always-moving/99512497702346 PhD. University of Surrey.

Rozas, D., Gilbert, N., and Hodkinson, P. 2015. Drupal as a runaway object: Conceptualisation of peer production activities through activity theory. *Proceedings of the*

European Group for Organizational Studies 2015. Subtheme 17: "Activity theory and organizations", 1–17. Athens, Greece.

Rozas, D., Gilbert, N., Hodkinson, P., and Hassan, S. 2021. Talk Is Silver, Code Is Gold? Beyond Traditional Notions of Contribution in Peer Production: The Case of Drupal. *Frontiers in Human Dynamics*, 3, 618207, 1-16. https://doi.org/10.3389/fhumd.2021.618207.

Rozas, D., and Huckle, S. 2021. Loosen control without losing control: Formalization and Decentralization within commons-based peer production. Journal of the Association for Information Science and Technology, 72(2), 204-223. https://doi.org/10.1002/asi.24393.

Siefkes, C., 2008. *From exchange to contributions. Generalizing peer production into the physical world*. Berlin: Ed. Siefkes, 2008.

Spinuzzi, C. 2011. Losing by expanding: Corralling the runaway object. Journal of Business and Technical Communication, 25(4): pp.449–486. https://doi.org/10.1177/1050651911411040

Spinuzzi, C. 2020. “Trying to predict the future”: Third-generation activity theory’s codesign orientation. Mind, Culture, and Activity, 27(1): pp.4-18. https://doi.org/10.1080/10749039.2019.1660790

Spinuzzi, C., and Guile, D. 2019. Fourth-generation activity theory: An integrative literature review and implications for professional communication. *2019 IEEE international professional communication conference (procomm)*, 37–45. https://doi.org/10.1109/ProComm.2019.00012

Stadler, F. (2010, April 22). Digital Commons: A dictionary entry. Retrieved October 10, 2017, from http://felix.openflows.com/node/137

Trist, E. 1981. The evolution of socio-technical systems: A conceptual framework and an action research program. In W. F. Joyce and A. H. Van de Ven (Eds.), *Perspectives on organization design and behavior* (pp. 19–75). Wiley.

Uden, L., Damiani, E., Gianini, G., and Ceravolo, P. 2007. *Activity Theory for OSS Ecosystems*.

Vygotsky, L. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.

Wittel, A. 2013. Counter-Commodification: The Economy of Contribution in the Digital Commons. Culture and Organization, 19(4): pp.314–331. https://doi.org/10.1080/14759551.2013.827422

Zilouchian-Moghaddam, R., Bailey, B., and Fu, W.-T. 2012. *Consensus Building in Open Source User Interface Design Discussions* (pp. 1491–1500). In *CHI '12* (pp. 1491–1500). https://doi.org/10.1145/2207676.2208611