

# Aligning Public Administrators and Citizens on and Around Open Data: An Activity Theory Approach

Jonathan Groff<sup>(✉)</sup>, Michael Baker, and Françoise Détienne

I3 (Institut Interdisciplinaire de l'Innovation), Centre National de la Recherche Scientifique,  
Telecom-ParisTech, Université Paris-Saclay, 46 rue Barrault, 75013 Paris, France  
{jonathan.groff,michael.baker,  
francoise.dettienne}@telecom-paristech.fr

**Abstract.** Open data have recently become key vectors of the implementation of open government in terms of the notion of transparency. The present study is precisely part of an EU-funded project whose aim is to develop a European platform devoted to the collective exploitation of open data. It aims to understand the processes of production and use of open data by, respectively, public administrators-(PAs) and citizen groups. On the basis of an Activity Theory analysis of focus groups involving these two sets of social actors, potential internal and external tensions are identified, with respect to objects, rules and instruments of activity. Main results showed that PAs practice a “strategically opaque transparency” policy by selecting data to open with the aim of preserving politico-economical interests, thereby limiting their reuse. We propose that interactions with citizens on the ROUTE-TA-PA platform could support PAs in publishing relevant data for users, whilst respecting these interests.

**Keywords:** Open data · Public administrations · Transparency · Activity theory · Tensions · Social platform

## 1 Introduction

Over the past ten years, open data initiatives have become decisive factors in the policy of *transparency* in governments, and the use of Information and Communication Technologies has been proposed an effective way for promoting it [1]. The research described here is being carried out within the EU-funded project “ROUTE-TO-PA” ([www.routetopa.eu](http://www.routetopa.eu)), whose aim is to improve the transparency of Public Administrations (PAs) by enabling and improving dialogue between them and citizens, on and around open data produced by such administrations. This is being done by developing two related tools, in collaboration with the administrations of several European cities (Dublin, Den Haag, Groningen, Issy-les-Moulineaux, Prato): (1) a Social Platform for Open Data (SPOD), enabling social interactions involving open data citizen-users and administrator-producers; (2) a visualisation toolset (Transparency-Enhancing-Toolset or TET), integrated with major Open Data Platforms (such as CKAN). The overall design, implementation and evaluation approach is based on agile methods in order to quickly implement, test and improve different versions of the

SPOD-TET platform. Since what is to be developed is not only a technical system, but rather a socio-technical system, a participatory design approach is being adopted.

This paper is focused on the first step of this participatory design approach, aiming to understand citizens' and civil servants' goals as well as their motivations for engaging in the production and use of Open Data, and to develop epistemic communities of practice around them. In the framework of Activity Theory [2], this means identifying the distinct activity systems of both potential users and of producers of open data, and the tensions that may emerge within and between them. Within this aim, we organised two focus groups, with PAs and with a specific group of citizens. The first focus group was carried out with seven public administrators in the Paris region; the second involved eight young entrepreneurs who were in the process of creating a company, as well as people who had created their company less than two years ago.

Analysis of the focus group data (verbal interactions) showed that these two groups share respective objects of activity, but are faced with internal tensions, principally due to contradictions between their rules and their goals, as well as external tensions, due to mismatches between users' needs and PAs' rules. We discuss these results and their implications in terms of the design of the socio-technical system.

## 2 Background: Open Data and Activity Theory

The concept of open data is defined as “the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control” [3]. In other words, giving citizens access to public data is a means to engage them in their own governance and thereby to involve them in political decision-making. In this way, citizens are not only observers in the process of governance but are also able to contribute to it.

Such participation implies that citizens visualize data but also use and transform it in order to propose collective solutions to public issues. This practice is only possible if datasets are completely accessible and available, and if they can interoperate with others datasets.

Yet, the use of public data, as currently released, is restricted to specific software, due to the various transformative processes that they undergo before publication.

The aim of the present work is precisely to study the effects of these access constraints on the reuse of open data and on their collective exploitation. We will study how this technical limit could impact the relationships between users' and producers' activity systems.

The digital portals that encourage open data reuse are generally focused on the potential applications of data rather than on the mode of data production and storage. As such, data are commonly considered as “naturally” “raw” and in no sense as political or normative agents [4]. However, data, whatever they are, must be produced before being rendered open and/or analysed.

In public administrations, the process of producing open data involves several steps, described by Denis and Goëta [5, 6] as “identification” (finding the services that collect public data), “extraction” (identifying data availability and collecting it) and

“rawification” (making data visible and readable by the most common tools handled by developers). This series of sociotechnical manipulations calls into question the concept of *transparency*. Indeed, transparency is defined as “the process through which public authorities make decisions should be understandable and open. . . . the information on which the decisions are based should be available to the public” [7]. However, the different steps of the open data “fabrication” process are not “open” for citizens, who do not have access to the mechanisms which transform professional data into generic data, adapted to private users (e.g. citizens, IT developers, companies). There is therefore a gap between the officially declared aim of “opening up data” — *sharing information initially reserved to PAs in order to promote a participatory democracy* — and the constraints relating to the open data production process — *proposing restricted content in a specific format*.

Our study addresses the effect of a specific rule related to the PA’s activity system on this production process: the protection of citizens’ interests. We analyse how this politico-social commitment could lead to internal and external tensions with users’ activity systems. According to Activity Theory [2, 8], human activity is articulated by dynamic and reciprocal interrelation of different entities. An activity system is represented by a systemic model that involves: the subject who carries out the action, the instrument-mediator, the object towards which activity is oriented, the division of labour, the community and its rules.

In the present study, we used this framework to analyse and compare activity systems of open data producers and open data users. Our goal is to identify the different components, the possible tensions and double binds implied within and between open data users’ and producers’ activity systems. In this way, we aim to develop the theoretical and methodological foundations for designing a socio-technical system devoted to collective exploitation of open data.

### 3 Focus Groups: Implementation and Analysis

We organised two focus groups, one with Public administrators (“PAs”) and a second with start-up company chief executive officers (or, “young entrepreneurs”). The first focus group was carried out with seven public administrators in the Paris region. Public administrators are representatives of public affairs. They apply, supervise and coordinate the policy programmes of governments at local and regional levels. The second involved eight young entrepreneurs. (By “young entrepreneurs” we mean people who are in the process of creating a company and people who have created their company less than two years ago).

These two particular groups (PAs and start-up CEOs) were chosen because they both seek to foster the development of an economic environment around the transformation of open data into new applications and services, by creating companies that promote job creation. Young entrepreneurs wish to create sustainable enterprises by using open data. PAs want not only to restore public trust, but also to improve communication between local communities and private companies in order to find a converging model of development that boosts employment.

Each focus group session lasted 2 h 30 min. Participants were interviewed on the usefulness of a community platform devoted to publication, sharing and exploitation of open data. Each participant was asked firstly to answer individually then all participants were invited to complete their answers by interacting together. Following this, participants were asked to verbalize their needs in terms of information, exchanges and functionalities on the basis of a usage scenario which describes a (fictitious) young entrepreneur (“Annie”), who wants to develop applications by using open data. Participants had to exchange and to define collectively their expectations concerning the nature of data, the type of interaction and the tools that they need if they were a start-up CEO such as Annie.

Participants’ answers were categorized according to a coding scheme based on dialogic function (e.g. question, assertion, request) and epistemic content [9].

## 4 Results

Using the conceptual framework of Activity Theory [2, 8], we analysed the activity system of PAs and Young entrepreneurs, i.e. the relations between instrument, subjects and objects. We also identified tensions that could arise within and between these activity systems. We first present each activity system then describe the potential tensions between them.

### 4.1 PAs’ Activity System

*Subjects.* Open data producers were represented by: Public Administrators in charge of technical services (e.g. information management) and business activities in the Paris Region.

*Objects.* In accordance with policy commitments, PAs generate public access to policies and financial information with the aim of restoring public trust and thus facilitate dialogue between electors and local governments. They want to make government processes and decisions open. Intrinsically, they share open data in order to drive economic growth in their Region by encouraging companies to design applications using them. They seek to create a business network, gathering entrepreneurs and local governments, focused on open data.

*Instruments.* PAs consider a collaborative platform devoted to open data as a means: (i) to identify the most relevant data to publish (with the aim of promoting the development of digital applications) and (ii) to identify data which should give entrepreneurs answers to typical initial difficulties associated with creating businesses (e.g. taxation, human resources, watching out for competitors, etc.).

*Rules.* Public Administrators are required to implement political commitments undertaken by governments. For this reason, they are reticent to publish data that might have a negative effect on the attractiveness of the city in terms of economy, ecology or safety,

or else data that would allow local pressure groups to criticise them (or at least oblige them to engage in time-consuming discussions).

*Tensions.* The main tension is between the PAs’ objective of “serving the public and involving citizens in political decision-making” and the rules of their community requiring them to “preserve economical and political interests”. These rules lead to processes of selecting which data should be published or not (e.g. PAs are reticent to publish data — such as on air pollution — that might have a negative economic effect on the attractiveness of the city, or data — such as specific subsidies — that would allow local pressure groups to criticise the PA).

So, on one hand PAs intend to involve citizens in political decision-making to restore public trust and to improve accountability of policy makers (European E-Government Action Plan 2011–2015), whilst, on the other hand, they do not provide full access to public data. In sum, they adopt what we term *strategically opaque transparency*, restricting the available data, or else spreading information across disparate data-sets, which renders understanding more arduous. This calls into question the nature of their collaboration with citizens, and thus disrupts the elaboration of collective solutions to societal problems.

Figure 1 shows the activity system of PAs, with potential tensions within it (thick grey double arrows).

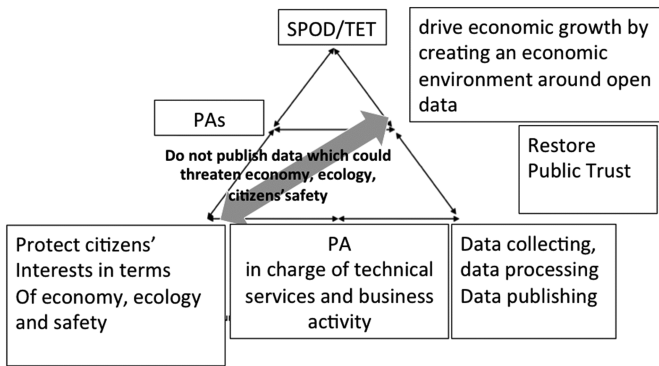


Fig. 1. Activity system of public administrators

#### 4.2 Users/Citizens’ Activity System

*Subjects.* Open data users were represented by: start-up entrepreneurs who develop innovative systems, principally in Information and Communication Technologies (ICT).

*Objects.* Start-up CEOs want to build sustainable enterprises around open data. They are also interested in participating in a network of private and public organizations with the aim of developing their companies.

*Instruments.* From the entrepreneurs’ points of view, the platform is seen as a tool which gives them possibilities to be informed about public politics in their domain, to share knowledge, experiences and of course to give them the opportunity to create applications and services.

*Rules.* Start-up CEOs develop innovative products in a highly competitive sector: ICT. So they have to respect privacy policies. These refer to information about design process of new products, financial and economical data, business strategies and organisational frameworks.

*Tensions.* The most important ones are as follows: (1) entrepreneurs need to interact around updated data in a synchronous way, in order to co-design products within a short time period, but the current tools do not enable them to do so; (2) they need to collaborate with other companies but they do not wish to disclose confidential information with potential competitors.

Firstly, entrepreneurs want to obtain quick answers to their questions. Yet, collective exploitation of open data on a social network is a long process that involves asynchronous interactions.

Secondly, entrepreneurs do not wish to disclose confidential information. At the same time they need personalized answers adapted to their individual problems. In summary, they seek to collaborate with other companies although they could be in competition with them in some lines of business.

Figure 2 shows the activity system of young entrepreneurs with tensions that could appear within it (thick grey double arrows).

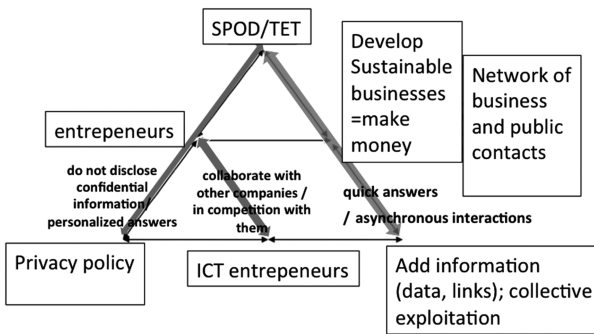


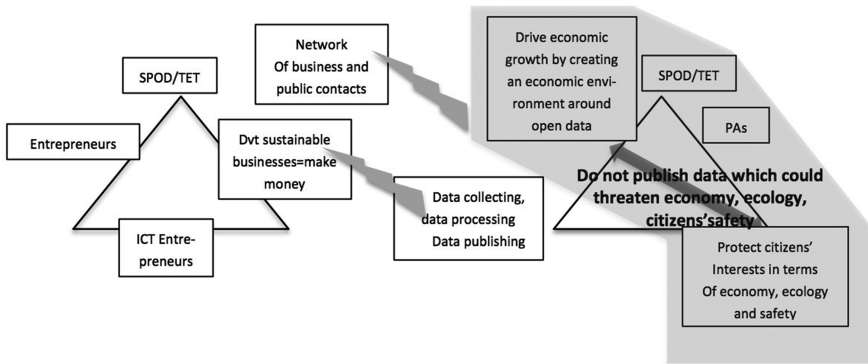
Fig. 2. Activity system of young entrepreneurs

### 4.3 Tensions Between PAs’ and Users’ Activity Systems

Our analysis revealed tensions between the two activity systems. They are mainly generated by: (i) the data production mechanism, which involves several tasks among different actors in the division of labour (ii) and the selective publication process, with its underlying the rules relating to the mission of PAs.

- (i) PAs have the aim of boosting employment by encouraging the development of sustainable enterprises. Notwithstanding this aim, Open data undergo several stages of “formatting and normalization” before publication [5, 6], involving different services that transform and standardise data. This may slow down the development of start-ups, because they have to recruit persons having specific competency profiles who are able to analyse these standardized formats, which could restrict open data reuse to a limited number of companies.
- (ii) Furthermore, PAs have to respect rules related to the implementation of government policy on the ground level. So, even though PAs wish to generate new businesses and stimulate growth by transforming open data into new applications and services, the selective publication process, making some data confidential, does not allow companies to create all the useful and operable tools that they could. In this way, we can see a reduction in the scope of an economic environment, involving business and public actors, developed around the transformation of open data into solutions for all citizens.

Figure 3 shows the external tensions between the two activity systems. The ‘lightning bolts’ represent tensions.



**Fig. 3.** Tensions between users’ and producers’ activity systems

## 5 Conclusion

The tensions within and between public administrations’ and young entrepreneurs’ activity systems that we have identified would need to be taken into account in the specification of a socio-technical system oriented towards the productive exploitation of open data. The main barrier to transparent publication of open data by PAs, and its productive economic use, resides in PAs’ attempt to preserve environmental and economic interests of their municipalities, together with their communication strategies with a political orientation. We propose that in order to address these tensions, they could either be ‘relaxed’, or else circumnavigated. The internal tensions of the entrepreneurs’ activity system could be partially relaxed by interaction in real time around open data visualisations, using the SPOD-TET platform, including the use of private or

public discussion spaces. However, deeper relaxing of tensions between the two activity systems would need to rely on changes in the socio-political environment itself. For this project, we propose to navigate around these tensions, in particular those related to the “selective publication process”, by developing scenarios around open data which are not “economically or politically sensitive”. This last issue calls into question the notion of “transparency” as the straightforward provision and transmission of information, and leads to a more strategic vision of communication between the social actors concerned by open data.

**Acknowledgements.** This research was carried out within the EU-funded H2020 project “ROUTE-TO-PA” (route-to-pa.eu), contract number 645860. We would like to thank Issy-Média for their cooperation in this research.

## References

1. Bertot, J.C., Jaeger, P.T., Grimes, J.M.: Using ICTs to create a culture of transparency: e-government and social media as openness and anti-corruption tools for societies. *Gov. Inf. Q.* **27**, 264–271 (2010)
2. Engeström, Y.: *Perspectives on Activity Theory*. Cambridge University Press, Cambridge (1999)
3. Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.G.: DBpedia: a nucleus for a web of open data. In: Aberer, K., et al. (eds.) *ASWC 2007 and ISWC 2007*. LNCS, vol. 4825, pp. 722–735. Springer, Heidelberg (2007)
4. Ruppert, E.: Doing the transparent state: open government data as performance indicators. In: *A World of Indicators: The Production of Knowledge and Justice in an Interconnected world*, pp. 51–78. Cambridge University Press, Cambridge (2013)
5. Denis J., Goëta S.: La fabrique des données brutes. Le travail en coulisses de l’open data. In: *Journée d’études SACRED «Penser l’écosystème des données. Les enjeux scientifiques et politiques des données numériques»*. HAL, Paris (2013)
6. Denis J., Goëta S.: Exploration, extraction and ‘rawification’ the shaping of transparency in the back rooms of open data. In: *Neil Postman Graduate Conference*. HAL, New-York (2014)
7. Soderman, J.: *The citizen, the administration and community law*. General report for the 1998 Fide Congress (1998)
8. Engeström, Y.: *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Cambridge University Press, Helsinki (1987)
9. Grosz, B., Candace, S.: Attention, intentions, and the structure of discourse. *Comput. Linguist.* **12**(3), 175–204 (1986)