# The Social Shaping of Open Data through Administrative Processes

Sirko Hunnius
IfG.CC – Institute for eGovernment
Kutschstall-Karree, Am Neuen Markt 9c
D-14467 Potsdam, Germany
+49(0)331.7403.6739
shunnius@ifg.cc

Bernhard Krieger
IfG.CC – Institute for eGovernment
Kutschstall-Karree, Am Neuen Markt 9c
D-14467 Potsdam, Germany
+49(0)331.7403.6739
b.krieger@cantab.net

## ABSTRACT

Many models have been provided in the last years that aim at describing an optimal open data publication process. However, they fail to explain the different outcomes of open data initiatives. Based on qualitative research this paper conceptualises the open data phenomenon as a set of techno-political arenas in which different interests of a variety of actors potentially and actually collide. The micro-political arena model constitutes an instrument to delineate the social and institutional context of open data that can be employed to explain the successes, as well as the failures of individual open data projects.

## Categories and Subject Descriptors

Management of Computing and Information Systems, Data Management Systems, Data Modeling, Information Integration, Digital Libraries and Archives, Resource Description Framework (RDF)

## General Terms

Management, Human Factors, Standardization, Theory.

## Keywords

Open data, process models, techno-political arenas, social and institutional context.

## 1. PROBLEM STATEMENT AND RESEARCH QUESTION

The topic of open data is generating considerable interest among researchers, technology developers and practitioners in public administration. The conversation so far often circles around the potentials of open data [8, 20, 26]. However, up until now there is

little evidence of any significant economic or societal impact [12]. Regarding the sheer amount of open data anglo-american governments seem to provide far more data and render the latter in a more sophisticated way than governments in continental European countries. This difference is puzzling as open data seems to be an international trend, fostered by an international community and pushed by international advocacy groups (e.g. Open Knowledge Foundation). Multinational initiatives such as the Open Government Partnership are taken up by countries as diverse as the United States of America, Chile, Austria, Russia, Kenya and Malaysia. Considering this heterogeneity of actors on the one hand and the differences in the implementation of open data activities on the other institutional factors appear to play an important role in how open data is perceived and adopted in the different public sector organisations.

This paper investigates the administrative practices involved in the provision of open data by public sector institutions. In particular it analyses the impact of the former on open data, especially on data and meta data quality. Therefore, the focus of this paper is limited on processes in the context of open data within the public administration. Such focus excludes the usage of open data for this endeavour. The research question thus concerns the relationship between institutional arrangements of the involved public sector actors and actual publication activities of the latter. An in-depth scrutiny of open data processes provides a deeper understanding of the kind of data, as well as of its structure, vocabulary, meta data, license etc. employed by governmental institutions.

## 2. OPEN DATA TRAJECTORIES

Various models of (linked) open data have been put forward under different headings. They have been termed the open data life cycle, the open data value chain or plain open data process [30]. The different terminologies illustrate different purposes – practical guidance [13] or analytical separation – and foci. Whereas value chain models focus more on the creation of value during open data usage [15], the life cycle models aim to structure the handling of the data itself. Existing process models focus on activities within public administration, such as generating, editing and publishing the data without paying too much attention on the outside-use.

Most models contain similar elements and differ only regarding semantics, granularity or the extension of the process. Hyland et al. [13] provide a six-step guidance model that contains the steps to (1) identify, (2) model, (3) name, (4) describe, (5) convert, (6) publish the data and the reverse activity to maintain it, similar to Villazon-Terrazas et al. [28]. Another model by Hausenblas et al. [11] also includes the user perspective, adding the steps "discovery", "integration" and "use cases". With the ambition to