Open data for accountability in the fight against corruption

Hulstijn, Joris; Darusalam, Darusalam; Janssen, Marijn

Publication date
2018

Document Version
Publisher’s PDF, also known as Version of record

Published in
CEUR Workshop Proceedings

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.
Open data for accountability in the fight against corruption

Joris Hulstijn1  Darusalam Darusalam2  Marijn Janssen2

1 Tilburg School of Economics and Management, Tilburg University
2 Department of Technology, Policy and Management, Delft University of Technology
j.hulstijn@uvt.nl,  d.darusalam@tudelft.nl,  M.F.W.H.A.Janssen@tudelft.nl

Abstract. There has been a lot of research on the use of open data in the fight against corruption. Although there are some promising examples, it appears that a systematic approach is lacking. What are the design principles for an architecture to open up data and thereby reduce corruption? In this paper we use theory about fraud, and about public accountability to derive design principles for an open data architecture. Crucial is the sustained presence of a specific forum: a group of people who are critical, have expertise, are free to challenge the authorities. Unlike the general public, a specific forum has an interest in reviewing the data. The architecture is motivated and illustrated by an extensive example of an E-procurement system in the context of an anticorruption program in Palembang, Indonesia.

1 Introduction

There has been a lot of attention for ‘opening up’ government data in the field of e-government [28]. One particular application area of open data is the fight against corruption. Corruption is a complex and difficult problem, also in government [12]. Corruption can harm society and result in increased poverty, reduce available money for essential public services, destroy citizens’ trust in government and undermine economic growth [26]. There are many factors that can lead to corruption in government [21]. In principle, addressing these different factors will lead to different approaches in the fight against corruption. In this paper, we will focus on the factors of transparency and accountability. Therefore we investigate a particular strategy to fight corruption: to open up government data to the public [25]. One of the claimed benefits of open data is that the government becomes more transparent and accountable [15]. This appears intuitive. Open data allows scrutiny by the public, which will reduce opportunities for corruption and increase changes of detection. Indeed, there are successful examples of e-government systems to reduce corruption, see [16] and [2] (p 265-266). There have
also been examples of systems to open up government data to reduce corruption, for example in Brazil. A recent study by Transparency International [14] shows that a lot has been achieved in Brazil. “Nonetheless, open data is still underutilized, and anti-corruption enforcement is weak” [14](p 4). In particular, there are problems with data quality. Moreover, the initiatives lack involvement: when there is a ‘lead’ there is not always a follow up. Apparently, the relationship between open data, transparency and accountability and eventually reducing corruption is not immediate, but rests on specific assumptions. For instance, in the context of corporate governance, Elia argues that opening up too much data, in too much detail, is pointless or even counterproductive [11]. Details of illicit behavior may be drowned out.

In this discussion paper we therefore propose design principles for an open data architecture that should foster accountability, and in turn, reduce opportunities for corruption. We will take a theoretical approach, based on an analogy with fraud [1, 9], and theory of public accountability by Bovens [3]. Following Day and Klein [10], Bovens characterizes accountability as a dialogue between an actor and a ‘significant other’, called the forum. The actor feels obliged to provide accounts of its conduct to the forum. In order to hold the actor accountable, the forum should be critical and powerful enough to ask for explanations, and eventually pass judgement about the conduct of the actor.

To motivate the approach and validate the design principles, we discuss a case study of an e-procurement system at the local government of Palembang in Indonesia.

The remainder of the paper is structured as follows. Section 2 provides a background on corruption and how it can be reduced. Section explains the idea of accountability as a dialogue, and derives some general principles about the actor and the forum. Section 3 applies these principles to an architecture for opening up data. Section 4, finally, discusses application of these ideas in the Palembang case study.

2 Corruption and Fraud

In this section we provide a background of the notion of corruption in government and how it can be reduced. Corruption is an important topic that deserves a proper literature review. In this discussion paper, we cannot provide a full review. Instead, we will focus on two elements.
First, we will draw an analogy with fraud, and derive three ways to address corruption. Second, we will discuss anti-corruption strategies.

The word ‘corruption’ is based on Latin: *corrumpere* means to spoil or destroy. Corruption is defined as “behavior which deviates from the normal duties of a public role because of private-regarding (family, close private clique) pecuniary or status gains” [21] (p. 4). An additional aspect of corruption in the case of government, is the misuse of power [12]. So similar to the notion of fraud, which is defined as a violation of trust [9], corruption involves the misuse of responsibilities related to a public role or duty. For instance, corruption involves “bribery (use of reward to pervert the judgments of a person in a position of trust), nepotism (bestowal of patronage because of ascriptive relationship rather than merit), and misappropriation (illegal appropriation of public resources for private-regarding uses” [21] (p. 4).

![Figure 1. Fraud triangle [1, 9]](image)

There is an interesting analogy between corruption and the notion of fraud as studied in accounting [1] and criminology [9]. In Figure 1, fraud is depicted as a triangle, involving three necessary elements.

1. **Opportunity**: the actor has been trusted with responsibilities that involve certain powers (budget; decision rights),
2. **Pressure**: the actor has a ‘non-shareable problem’ (e.g. financial problems, status), and misuse of trust can offer a ‘private solution’
3. **Rationalization**: the actor provides a kind of explanation for crossing the line: “I am only borrowing the money” or “others do it too”.

The triangle suggests three ways of reducing fraud or corruption. These are effectively means to improve management control [19] and the system of internal controls [8], in particular the control environment.

First, one could try to reduce opportunities for corruption. That means limiting the powers of an actor in a certain role, by putting administrative restrictions in the procedures or by recording details of execution (audit trail). This is the easiest element to adjust. In modern organizations, many institutional powers are controlled by ICT, such as the ERP system
or business process management system. In accounting, there is a lot of expertise on ways to strengthen internal controls, partly by means of ICT. In particular, there are so called application controls, which are built into software applications. For instance, there are authorization tables, restrictions on the budget a person can control, on tariffs and pricing policies, etc. In addition, there are also many manual controls, that are supported by technological means. For instance, if all invoices are numbered consecutively, it is harder to make one disappear. Finally, there are the IT general controls, that help to maintain a reliable control environment. In particular, consider such measures as segregation of duties, audit trail, access controls, logging and monitoring [23]

Second, one could try to reduce pressure. In the financial sector there has been a lot of debate about remuneration. If you pay people enough, is the argument, there should be less incentives to misuse trust for personal gains. However, the evidence as to the effectiveness of this measure is inconclusive, at best [24]. Another option to reduce pressure is suggested by Cressey’s original paper [9]: foster an environment in which people feel safe to discuss problems. In this way, when people do have a problem (gambling debts, dissatisfaction), they are less likely to see misuse of trust as a way out. For example, one could implement a regular meeting structure, in which officials have an opportunity to discuss – in secret – difficult decisions with peers.

Third, one could try to raise moral standards and make it harder for people to ‘rationalize’ misuse of trust. This is perhaps the most effective but also the hardest option. It involves elements of organizational culture, such as public values, social practices and role models [13, 22]. Organizational culture can’t be managed directly. Instead, one can only try and improve side conditions. The ‘tone at the top’ makes an enormous difference. For example, in the Enron case, efforts to improve compliance were useless, as executives themselves were seen to violate the rules [24]. Public values such as quality awareness, safety or integrity, can be stimulated. For example, to stimulate quality awareness, one could routinely publish examples of products that are rejected. This will demonstrate the borderline. Practices and procedures may help to reach consensus. If it is mandatory to rate and discuss quality of a product with colleagues, this will help establish a common threshold.

We will now discuss an example from the literature, to see how these approaches to fraud are applied.
Example. Kim [16] reports about the OPEN system, as part of a long-term effort in Korea to fight corruption (1998-2007). The OPEN system “is an online system used to disclose administrative procedures (likely to be related to corruption) to citizens in various public service areas (such as housing and construction, sanitation, and urban planning, among others).” [16] (p 45). The purpose was to improve efficiency, by preventing unnecessary delays or unfair handling by civil servants, and also to improve transparency in some areas of civil administration.

First, such a system is likely to reduce opportunities for corruption. By allowing citizens to view the procedures and status of their requests, an audit trail is created. Decisions can be compared. Moreover, the system was supported by projects to clarify the procedures and design systems that simplify, standardize, and de-personalize delivery of services [16].

Second, we do not know if the system helped to reduce pressure. We do know, that the system reduced personal contact between applicants and officials [16], thus reducing opportunities to be put under pressure.

Third, the ‘tone at the top’ proved decisive. The system was supported by Kun Koh, mayor of the city of Seoul. “Mayor Koh took resolute steps to fight corruption. He launched a variety of reform programs which culminated in the declaration of a ‘War against Corruption’” [16] (p 46). By doing this, Koy set values such ‘clean’ and ‘transparent’ as norms that no one could deny.

Strategies to Combat Corruption
Activities to reduce corruption are often presented as phases in an anti-corruption cycle: prevention, detection, investigation and sanction [20].

Alternatively, the literature mentions four strategies to combat corruption: prevention, enforcement, access to information and empowerment, and capacity building [16]. Unlike the stages, which focus on officials, these strategies also assign an important role to citizens. They indicate how ICT or open data can help: by providing access to information and empowering citizens, and by helping to organize and improve expertise, i.e. capacity building. In the next section, we will try to make these ideas about empowerment more systemic, by relating them to a theory of public accountability.
3 Accountability as a dialogue

Accountability has positive connotations, and is used as a kind of synonym to good governance. However, the meaning of the term depends on the context, on a relationship with others: “Accountability can be defined as a social relationship in which an actor feels an obligation to explain and to justify his or her conduct to some significant other” [10]. According to Bovens [3, 4] an accountability relationship involves an actor and a forum (Figure 2). The actor can be a person or agency. Then there is some significant other, who can be a person or agency, but can also be a more abstract entity, such as God or ‘the public’. Bovens calls this the forum.

The accountability relationship develops in three stages. First, the actor must feel obliged to inform the forum about its conduct, including justifications in case of failure. The obligation may be both formal, i.e., required by law or by contract, or informal and self-imposed, for instance because the actor is dependent on the forum. Second, the information may be reason for the forum to challenge the actor, ask for explanations and debate the adequacy of the conduct. The actor responds, providing additional facts or motivations. Third, the forum passes judgement on the actor’s conduct. A negative judgment often leads to some kind of sanction; again this can be both formal and informal. This means that an accountability relation should provide room for discussion: it is not a one-way stream of reports, but rather a dialogue. The process can be formalized as a dialogue game [5].

Bovens [3] identifies five necessary conditions of public accountability, which characterize the dialogue game just described: (1) “public accessibility of the account giving—and not purely internal, discrete informing;
(2) explanation and justification of conduct—and not propaganda, or the provision of information or instructions to the general public;

(3) the explanation should be directed at a specific forum—and not be given at random;

(4) the actor must feel obliged to come forward—instead of being at liberty to provide any account whatsoever; and

(5) there must be a possibility for debate and judgment, including an optional imposition of (informal) sanctions, by the forum—and not a monologue without engagement” [3] (p 185); italics added.

The prototypical example of an accountability relation exists between government and parliament, where parliament derives its power from the population, and can send the government home (confidence vote). Journalists play an important role, demanding transparency and challenging the government. By informing the public, they can mobilize the underlying power of parliament.

Another example of an accountability relation is auditing [6, 5]. In that case, management is the actor, reporting the annual financial accounts. The shareholders act as a forum. The forum is assisted by an auditor, who has the expertise and authority to evaluate the financial accounts, compare them to standards and norms and provide assurance about their reliability [17]. However, the auditor as such has little power. The power derives from the shareholders, or other powerful stakeholders, such as the board of directors or regulators.

Based on these examples, we derive four additional necessary conditions for a fruitful accountability relationship, compare also [6].

(6) The organization of the actor must provide reliable reports about its conduct. Reports must be accurate (correspond to reality) and complete (contain all relevant aspects of reality). In practice this can be ascertained by internal controls.

(7) The forum must be independent of the actor, allowing it to challenge the actor, as stated in (5) above. In particular, the forum should not depend on the actor for employment. Individual members should not have family ties, or personal relationships.

(8) The forum must have, or have access to, expertise, to evaluate the reports, compare them to standards and norms, and eventually pass judgement. If necessary, the forum can be supported by experts or auditors, who possess specific expertise. That means that those experts must also be independent.
(9) The forum must wield enough *power* or influence over the actor to “make the actor feel obliged to come forward” as in (4), and to have a credible claim to sanctions, as in (5). In particular, when the actor is dependent on the forum in some way, the forum has the power to withhold whatever the actor depends upon.

### 4 An open data architecture for accountability

So how can we apply these ideas to an open data architecture?

An architecture is a set of fundamental concepts or properties of a system in its environment, embodied in its elements, relationships, and in the principles of its design and evolution [18] (p 2).

What are the components, what are their relationships, and what are the design principles? If we want to conform to conditions (1) – (9) above, we have to include some form of governance. So we will treat architecture here in a broad sense, including also the organizational structure. An open data architecture is a socio-technical system [7].

Figure 1. Conceptualizing open data [27] (p 157)

At first sight, an open data architecture seems to be just a website to publish data in a particular format. The website makes the data available to the forum (D0. Availability). Crucial elements are the legislation, policies and business rules that determine which data eventually gets
published (D1. Data selection). Published data should be relevant to the issues addressed by the application, and later to the challenges or inquiries made. Data should be fit to be published. In particular, state secrets, data that are recognized to be confidential (e.g. medical; military) or data about individual persons should not be published. This means that data selection should involve a careful trade-off between transparency and confidentiality. Ideally, the forum should have an advisory role on setting the roles and policies for data selection. Moreover, data sets should adhere to representation standards, that improve interoperability between various systems, and allow for comparison (D2. Representation standards). Data sets are filled with real data by interfaces with day-to-day operational systems, such as databases and ERP systems (D3. Content). A coherent system of internal controls should help guarantee reliability of the data (accuracy and completeness) (D4. Reliability). Note that completeness is hard to ensure, as it involves comparison to what is considered to be ‘all’. Once selected for publication, officials should not be able to delete data or obstruct publication. Finally, the architecture should have features to respond to challenges and inquiries with data that are relevant (D5. Relevance).

Because accountability involves a relationship, the architecture should not only focus on what governments have to do in order to reliably ‘open up’ data sets, but also on the forum. That means in particular that the architecture should have components for the forum to gather, and mobilize support (F0. Mobilize). Consider for example outlets on social media. Clearly, mobilizing a forum is not the responsibility of the government! The government should collaborate with those institutions that happen to represent the forum, such as pressure groups, branch organizations, watchdogs, regulators, parliaments, etc. Given such collaboration, the platform should allow all members of the forum easy access (F1. Access). The platform should have components for the forum, or experts and auditors, to analyze the data (F2. Analysis). The architecture should have components to pose challenges and questions, about specific data elements and get a relevant proper response (F3. Challenge and response), see also D5. Responses should be archived, and should in principle also be made public (F4. Archive). In some cases, absence of a response would also be telling. Finally, the architecture should make it possible to escalate and file complaints, in case of a suspected deviation from the policies (F5. Complaints
procedure). In addition, the architecture should have specific support for journalists, experts or auditors (F1 and F2).

The list of requirements is summarized in Table 1.

<table>
<thead>
<tr>
<th>Data publishing requirements</th>
<th>Forum requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>D0. Availability</td>
<td>F0. Mobilize</td>
</tr>
<tr>
<td>D1. Data selection policies</td>
<td>F1. Access</td>
</tr>
<tr>
<td>D2. Representation standard</td>
<td>F2. Analysis</td>
</tr>
<tr>
<td>D3. Content</td>
<td>F3. Challenge and response</td>
</tr>
<tr>
<td>D4. Reliability</td>
<td>F4. Archive</td>
</tr>
<tr>
<td>D5. Relevant</td>
<td>F5. Complaints procedure</td>
</tr>
</tbody>
</table>

Table 1. Requirements for an open data architecture, focusing both on publishing open data (D) and on the role of the forum (F)

5. **Case: E-procurement in Palembang**

**Data Collection** Data for this case study was collected by investigating various sources, including the official website, official documents, prior research and the participating observations of one of the authors.

**Case Description.** The local government in Palembang South Sumatera Indonesia uses the Systems Electronic Procurement Service (LPSE) for the procurement of government goods and services. The LPSE system has been introduced to improve the efficiency, effectiveness, quality, and transparency in the procurement of goods and services. The system matches vendors and governments. LPSE is hosted by a unit formed across ministries and other institutions. This system facilitates procurement officers and also provides services for the provider of goods and services, in the territory of the LPSE concerned.

The procurement process starts by defining the needs, which is followed by publishing the Request for Quotation (RFQ). Next, suppliers receive the RFQ and develop their quotes. The supplier sends the quotes to government, government who receives the proposals. Once the proposals are received and the deadline has passed, the proposals are evaluated and supplier(s) will be selected and a final tender will be requested and the contract is signed. Thereafter follows the execution of the contract in which the products and services are delivered and paid. Finally, the delivered products and services can be evaluated.
According to local government regulation, all working units of local government must use the LPSE system. This should prevent bypassing and the risks of fraud. All the working units are obliged to announce their planning, implementation and results of their procurement processes via the LPSE system. As such, there is a huge potential for opening data. The types of data available in the LPSE system are as follows:

- **Auction announcement**: The LPSE system provides an announcement about what types of procurement are available from working units in Palembang;
- **Information about system failure**: The system provides information in case a package cannot be generated, or if a file failed to upload, so the system can provide solutions;
- **Electronic Catalogue (EC)**: This catalogue provides a detailed listing of vendor offerings. For example, description of products, prices, delivery times.
- **Monitoring and online evaluation**: The system provides information about planning packages of procurements, financial progress, physical progress, procurement of goods and services progress;
- **Whistleblowing**: The LPSE system has a link to [https://wbs.lkpp.go.id/container.php](https://wbs.lkpp.go.id/container.php). This website provides an opportunity for a person who has information about illegal, unethical or corruption related behavior related to procurement. The person can report these activities to the corruption watch or audit board.

In the current system, all this data is not yet open to the public. In the next section we present patterns for opening data.

**Case Problem.** How could opening up such data, help to reduce corruption?

**Case Solution.** In earlier research on the case, we found six types of patterns for corruption detection, which are presented in Table 2. These patterns are all based on internal control measures [23].

**Evaluation.** The patterns show that not only the resulting data should be opened, but also information about the operation of the administrative processes and the implemented internal control measures to prevent corruption should be opened. However, these internal control principles are not commonly known. Accounting expertise is needed to evaluate
which segregations of duties, ought to be met. This puts additional requirements on the kind of audience that the Palembang authorities are trying to reach. Without proper guidance, providing such meta-data may distract users from viewing the actual procurements.

<table>
<thead>
<tr>
<th>Pattern name</th>
<th>Description</th>
<th>Categorization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Storing and opening</td>
<td>Opening of documents generated in several activities</td>
<td>Process data</td>
</tr>
<tr>
<td>documentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Cross-data comparison</td>
<td>Comparison of data collected in different phase to detect discrepancies</td>
<td>Process data</td>
</tr>
<tr>
<td>3. Four-eyes-principle</td>
<td>Process data, should demonstrate that decisions are made by at least two independent persons</td>
<td>Control data</td>
</tr>
<tr>
<td>4. Segregations of Duties</td>
<td>Process data, should demonstrate that a single individual or department is not allowed to process a transaction in its entirety</td>
<td>Control data</td>
</tr>
<tr>
<td>5. Authorization</td>
<td>Process data, should demonstrate that only people who are authorized to approve an activity, have done so</td>
<td>Control data</td>
</tr>
<tr>
<td>6. Application controls</td>
<td>Opening of data about built in measures to avoid the making of mistakes and the availability of alerts in the system</td>
<td>Control data</td>
</tr>
</tbody>
</table>

Table 2. Opening data, based on patterns to detect corruption

A possible ‘forum’ for this case is the pool of potential vendors. It is likely that competitors of the vendor who received the procurement contract, have enough interest to actually analyze the data, and challenge the local government in case of irregularities. Collectively, the pool of vendors is large enough to wield some power or influence over the local government, as they form a substantial portion of the voting population.

The most promising pattern seems to be number 2: cross data comparison. This may help to identify whether awarded contracts actually conform to the stated contract requirements, and whether there are any oddities in awarding behavior.
6 Conclusions

Opening up government data to the public, seems a good strategy to try and fight corruption. The claim is that opening up government data will somehow improve transparency and accountability. Current examples of systems that attempt to open up data in the fight against corruption, therefore make an ad hoc impression. In this paper, we have tried to provide some theoretical underpinning to these attempts, in two ways.

First, we have used an analogy with fraud, to identify three ways of reducing fraud or corruption: (1) by reducing opportunities for misuse of trust, essentially by strengthening the internal controls, (2) by fostering an environment in which people talk about their problems, before they become insurmountable, and (3) by stimulating an organizational culture, in which normative decisions are being discussed and recorded, establishing consensus on norms and values.

Second, we have used the theory of public accountability of Bovens, with five necessary conditions for a fruitful accountability relationship. Based on an analogy with accounting, we derived another four necessary conditions: reliable reporting (for the actor), and independence, expertise and power (for the forum). These conditions can be translated into a set of five requirements for an architecture and governance procedure for opening up data (Availability, Data selection policies, Representation standard, Content, Reliability, Relevance) and five requirements focusing on establishing a critical forum (Mobilize, Access, Analysis, Challenge and response, Archive and Complaints procedure).

Some of these requirements were discussed in the context of a case study of an e-procurements system in Palembang, that makes it possible to open up crucial data about the procurement process. The case shows that indeed it is crucial to identify a specific forum. This is contrary to current practice in the open data community, where data is being opened up for a general public, in the hope they will find some purpose for it (e.g. hackatons). Lack of a specific audience and purpose for the data being shared, means in particular, that reliability and relevance are hard to establish. This may partly explain the current problems with data quality that are reported for many open data initiatives, e.g. [28].

In the Palembang e-Procurement case, it is likely that the forum will include vendors, in particular the competitors of those vendors who received the contract, because they have an incentive to protect their interests. The case also shows, that it makes sense to open up meta-data,
to reveal if the internal controls that ensure reliable reporting, are effective. However, this will require the forum to involve experts, such as accountants, who can interpret and evaluate such data.

Future research will have to show whether also the dialogue-related aspects of the theory are valid in this case.

References


