



Measuring the promise of open data: Development of the Impact Monitoring Framework

Abstract: It is a challenge to measure the impact of releasing data to the public since the effects may not be directly linked to particular open data activities or substantial impact may only occur several years after publishing the data. This paper proposes a framework to assess the impact of releasing open data by applying the Social Return on Investment (SROI) approach. SROI was developed for organizations intended to generate social and environmental benefits thus fitting the purpose of most open data initiatives. We link the four steps of SROI (input, output, outcome, impact) with the 14 high-value data categories of the G8 Open Data Charter to create a matrix of open data examples, activities, and impacts in each of the data categories. This Impact Monitoring Framework helps data providers to navigate the impact space of open data laying out the conceptual basis for further research.

Keywords: open data, impact measuring, social return on investment, SROI

1. Introduction

Releasing open data to the public is said to provide substantial economic growth (Klessmann et al., 2012; Manyika et al., 2013; Bürgi-Schmelz, 2013; Carrara et al., 2015) as well as broad social value to a society (Janssen et al., 2012). According to studies, enabling citizens and start-ups to use previously closed data in an unrestricted form new public services as well as new businesses may evolve (Linders, 2012).

Compared to other ICT related activities there is usually not much cost involved in publishing already available data. However, the release of open data by governments still faces substantial challenges in regard to cultural change involved, lack of legal foundation, or technical barriers (Janssen et al., 2012). Therefore it is justified to ask which particular value can be attributed to a specific open data initiative. Addressing this challenge of impact assessment helps not only to convince critical voices of decision makers regarding open data activities. It also supports the effective and efficient use of resources to focus on high-impact activities, thus increasing the success of new open data initiatives.

A way to assess the social value of activities is provided by the Social Return on Investment (SROI) approach. The SROI method distinguishes the use of resources (input), directly controllable results (output), indirect activities (outcome), and value-creating consequences (impact). This concept has been applied in a recent report for the Swiss government proposing an SROI-based framework to assess the impact of releasing open government data (Dapp and Stuermer, 2015).

In this article we summarize the findings of that report while deepening the academic foundation of our work. Thus the article provides a new approach by combining the SROI technique with open data impact literature leading to our Impact Monitoring Framework. The matrix in section 4 depicts practical examples on how to apply the framework for 14 distinct high-value open data categories for all 4 steps within SROI.

2. Literature review

In our review of literature we will first present a short overview of the SROI approach followed by a discussion of some of the current attempts to assess the impact of open data initiatives.

2.1. Social return on investment

The SROI approach was introduced by the Roberts Enterprise Development Fund (REDF) in the late 90's in order to measure the social value of philanthropic investments (Emerson et al., 2000). The idea was simple: While the Return on Investment (ROI) approach only measures profit, SROI includes the economic, socio-economic, and social value of an enterprise. However, "quantifying the immeasurable" proved to be difficult. REDF therefore created a sophisticated spreadsheet to calculate SROI metrics (REDF, 1999). Subsequently Sara Olsen and Alison Lingane empirically assessed the business plans of 88 companies during the Global Social Venture Competition (Olsen and Lingane, 2003; Lingane and Olsen, 2004). They extended the SROI approach by providing guidelines on how to calculate monetization of social value.

Based on the work of the REDF in the United States, The New Economics Foundation (nef) in the United Kingdom published a practitioner-oriented manual on how to apply the SROI technique (Aeron-Thomas et al., 2004). They were the first to suggest applying an impact map and indicators capturing the causal relationship between input, output, outcome, and impact. Another elaborate guide on how to apply SROI was issued by the UK Cabinet Office (Nicholls et al., 2009). It explains in detail the steps involved ranging from stakeholder identification to mapping of outcomes and, finally, calculating and reporting. Further academic publications present a rather critical view of the methodology (Ryan and Lyne, 2008), pointing out challenges and limitations of SROI (Arvidson et al., 2010) and identifying practical challenges when conducting impact assessments with SROI (Millar and Hall, 2012). Despite the critique of the SROI approach results, most authors acknowledge the method to estimate the holistic value of not-for-profit activities. Thus we consider SROI appropriate for describing the impact of open data activities.

2.2. Measuring the impact of open data

There is much anecdotal evidence indicating the positive impact of open data activities by governments and companies. For example a case study from Chicago explains how open data portals by the government supported citizen engagement leading to more transparency regarding political lobbying and safety in public parks (Kassen, 2013). Research exploring the impact of open data initiatives in developing countries found more transparency and accountability, increased innovation as well as economic development improved inclusion and empowerment of citizens

(Davies et al., 2013). Examples of the impact of open data initiatives have also been found in developing countries by analysing case studies from the Philippines, Kenya, Nepal, India and many more (Davies, 2014).

Quantitative studies also indicate the impact of open data initiatives. The Open Data Barometer assesses the results of open data initiatives around the globe including the identification of their outcome and impact (Davies et al., 2015). It measures entrepreneurial effects, transparency and accountability, government efficiency, economic growth, environmental sustainability, and inclusion of marginalized groups. The study finds a strong correlation between open data readiness (for example the availability of open data policies) and open data impact indicating the positive effect of governmental support of open data initiatives. However, while the Open Data Barometer report provides a vast amount of insight by analysing global trends and providing comparative data from different countries it is based on expert opinions as a proxy measure to evaluate the impact of open data initiatives. Due to the width of the report no in-depth assessment of particular initiatives is made.

In addition to empirical work there have been several studies on generic indicators and impact of open data policies and strategies. In 2013 Barbara Ubaldi proposed a set of metrics based on empirical analysis of open government data initiatives in the OECD (Ubaldi, 2013). In a subsequent survey among OECD countries 96% of the respondents (N=16) answered negatively if their government had “adopted a methodology to measure return on investment on open government data (e.g. potential cost savings, value for new services, more efficient service delivery)” (Ubaldi, 2014). Explaining differences between policies and their consequences is also related to impact analysis. A comparison of Dutch open data directives shows the gap between the goals of politicians praising the public value of open data activities while public servants are often challenged with the risks and complexities of actually releasing data (Zuiderwijk and Janssen, 2014). In the aim of benchmarking governments regarding their open government maturity and e-government openness a recent study illustrates the impact of open data and user involvement (Veljković et al., 2014). Finally a practitioner-oriented approach with global reach is provided by the Open Government Partnership initiative through their monitoring mechanisms of national action plans (Khan and Foti, 2015).

2.3. Need for a generic Impact Monitoring Framework

As illustrated by this literature review measuring the social impact of organizations has advanced substantially in the last decade while assessing the impact of open data is still in its beginning. The open data impact measurement approaches mentioned do successfully provide feedback on the particular topics. However, a methodologically clear structure regarding the causal relationships and interactions within the dynamics of the open data ecosystem is still missing. We therefore propose to introduce a generic open data Impact Monitoring Framework by applying the SROI technique and its terminology. Learnings and good practices of SROI can thus be used with the approach of measuring the social return of open data activities.

3. Horizontal and vertical dimensions of the framework

This section describes the two dimensions of the Impact Monitoring Framework matrix. First, the four values of the SROI model are defined in the context of open data activities and mapped horizontally. Second, 14 data categories used for the vertical framework dimension are introduced.

3.1. SROI as horizontal dimension

Following the SROI approach by Aeron-Thomas et al. (2004) four different steps are taken into account: input, output, outcome and impact. While input and output are controlled directly by the organization, outcomes and impacts are not. We use these four steps for the horizontal dimension of our Impact Monitoring Framework indicating the functional chain when releasing data. In order to receive a common understanding which open data activities are located within the corresponding step we describe each of them in general in Table 1.

Table 1: Four steps of the Social Return on Investment (SROI) approach in the context of open data

Step	Our understanding in the context of open data
<i>Input</i>	In SROI “input” includes all resources such as money, people, equipment, and facilities that are used by an organization to execute a certain action. Within open data activities “input”, too, covers these resources while adding the additional resource of native data produced by the organization. These are the data assets collected or measured by the organization based on existing laws and regulations. We call this “native data” because it is available e.g. in proprietary formats and software tools not yet ready to be published as “raw open data”. Much technical effort may be required in order to have it prepared for publishing on an open data portal. Nevertheless it is important to note that we consider native data to be already available as input resource. Therefore all other resources such as money and people are solely used for technical transformations, legal clarifications and other activities necessary for the publication of the data. There is no need to collect the native data but to convert it to raw open data.
<i>Output</i>	In SROI “output” refers to the direct and tangible deliverables produced by the organization. Within the open data ecosystem “output” refers mainly to the setup and operation of an open data portal. On such platforms organizations publish their open data records in appropriate technical formats ideally indexed semantically and released under official open data licenses. The quality of such open data portals differs widely today as a recent analysis of seven different country platforms has shown (Lourenço, 2015). The study considers quality of metadata, completeness of the datasets, the platform’s accessibility and visibility, usability and comprehensibility, timeliness, value and usefulness as well as granularity as important requirements of such open data portals.
<i>Outcome</i>	Following SROI, “outcome” covers all direct and indirect consequences of certain output actions by the organization. In the context of open data we consider all activities by the users of open data as “outcome”. Therefore this includes engineers, entrepreneurs, citizens, journalists, scientists, artists or administrators re-using the

	available open data in some form. Obviously the uptake of open data cannot be controlled by the organization releasing the data, it can only support and encourage data reuse. Outcome activities include for example organization of hackathons, development of open data visualizations, programming of web and mobile applications, integration and linking of open data with proprietary data, initialization of new open data based business models, publishing of data journalism reports, or creating new research projects with open data. For open government data in particular, there are abundant examples how citizens interact with public institutions by co-creating new services (Linders, 2012).
Impact	SROI defines “impact” as outcome adjusted for the effects that would have happened without the intervention by the organization. Within open data this means only the results actually caused by releasing the data are counted. For example assuming the release of an attractive set of data leads to more startups, only those will matter which are directly connected with the new open data records. In practice measuring the impact of specific output is not easy. Guidelines suggest using comparison groups or benchmarks in order to assess the net effect (Nicholls et al., 2009). By measuring and accounting for all the relevant factors deadweight and displacement can be allowed for in order to find the net impact.

Those four steps form a simplified structure of how certain data input leads to specific output results causing outcome effects and thus eventually making a difference measured by the impact. Obviously there is much interaction between the outcomes and most probably feedback loops influencing output and possibly even input factors. Nevertheless, to explain the idea of an Impact Monitoring Framework we chose to use only those four columns for the table.

3.2. Categories of data as vertical dimension

Data is most relevant if it is discussed within a certain context. Therefore as a vertical dimension we introduce a sector-specific categorization of data. This enables us to derive target groups for which such data is important. Known as a widely support classification we choose the high-value data categories defined within the Open Data Charter during the G8 summit in 2013 (G8, 2013). Action 2 states member countries should focus publishing data from those 14 areas (listed in Table 2) since they help most to improve democracy and support innovative re-use of data. This data categorization agreed upon in 2013 by the G8 countries (nowadays G7) is also applied by the Open Data Barometer (Davies et al., 2015) assessing open data progress on a country-level.

Table 2: High-value data categories from the G8 Open Data Charter including example datasets (G8, 2013)

Data Category	Example datasets (from G8 Open Data Charter)
<i>Companies</i>	Company/business register
<i>Crime and Justice</i>	Crime statistics, safety
<i>Earth observation</i>	Meteorological/weather, agriculture, forestry, fishing, and hunting

<i>Education</i>	List of schools; performance of schools, digital skills
<i>Energy and Environment</i>	Pollution levels, energy consumption
<i>Finance and contracts</i>	Transaction spend, contracts let, call for tender, future tenders, local budget, national budget (planned and spent)
<i>Geospatial</i>	Topography, postcodes, national maps, local maps
<i>Global Development</i>	Aid, food security, extractives, land
<i>Government Accountability and Democracy</i>	Government contact points, election results, legislation and statutes, salaries (pay scales), hospitality/ gifts
<i>Health</i>	Prescription data, performance data
<i>Science and Research</i>	Genome data, research and educational activity, experiment results
<i>Statistics</i>	National Statistics, Census, infrastructure, wealth, skills
<i>Social mobility and welfare</i>	Housing, health insurance and unemployment benefits
<i>Transport and Infrastructure</i>	Public transport timetables, access points for broadband penetration

4. The open data Impact Monitoring Framework

The following Table 3 represents the basic structure of our open data Impact Monitoring Framework. We have filled in some real as well as possible examples for each of the combinations to illustrate the applicability of the approach. Into the *Input* fields only native data examples are entered as the other resources such as money and people remain the same for each data category.

Table 3: The open data Impact Monitoring Framework with exemplary content

Data Category	Input <i>Native data, money, people, infrastructure, equipment</i>	Output <i>Open data portal with metadata, updated content, open format etc.</i>	Outcome <i>Hackathons, apps, new firms, linking of data, research etc.</i>	Impact <i>Net effect of output intervention</i>
Companies	Business register Data on occupations, products, services, shareholders, sector activities, bankruptcies	Updated directory of bankruptcies Directory of products and services offered in a certain region	Food production tracking app Transparency on CO ₂ and social footprint Ingredients and possible allergic reactions of consumers	Better informed consumers Higher level of corporate social responsibility Less CO ₂ emissions

<i>Crime and Justice</i>	<p>Reported crimes</p> <p>Data on police forces</p> <p>Data on court cases and convictions</p>	<p>Geo-location of crimes</p> <p>Arrests and convictions</p> <p>Court rulings with metatags</p>	<p>Visualization of crimes on a map (e.g. UK police)</p> <p>Security app with notification possibility</p> <p>Justice monitoring app</p>	<p>Public pressure for less criminality</p> <p>More successful police operations</p> <p>More transparency of court decisions</p>
<i>Earth observation</i>	<p>Data on the rise of average temperature</p> <p>Climate monitoring stations</p> <p>Data on temperature and barometric pressure</p>	<p>Data on melting of glaciers</p> <p>Historic climate data</p> <p>Real-time weather data</p> <p>Ocean rise</p> <p>Spread of aridity</p>	<p>Visualization of deglaciation</p> <p>Mashup of climate data with private weather stations</p> <p>Geo-located weather forecast</p> <p>Improved flood warnings</p>	<p>Creating awareness regarding climate change</p> <p>Better forecast on refugees and migration</p> <p>New businesses providing services on localized crowd-sourced weather forecast</p>
<i>Education</i>	<p>Data on educational institutions</p> <p>Anonymized grades of pupils and students</p> <p>Educational data of teachers</p>	<p>Number of students per institution</p> <p>Budgets of educational institutions</p> <p>Graduations per institution</p>	<p>Regional portal for assigning pupils to schools</p> <p>Transparency platform on teaching quality and graduations</p>	<p>Better distribution and integration of pupils</p> <p>Higher educational quality</p> <p>More equal opportunities for children</p>
<i>Energy and Environment</i>	<p>Solar radiation data</p> <p>Energy consumption data</p> <p>Data on pollution, emissions, and waste disposal</p>	<p>Consistent data on solar radiation, weather condition, and roof pitch</p> <p>Real-time data of air quality in cities</p>	<p>Portal for calculating ideal locations of solar panels in a town (solar cadastre)</p> <p>Portal pinpointing polluted areas and calling for action</p>	<p>Increasing the foundation of energy cooperatives (production and consumption of renewable energy)</p> <p>Less pollution</p>

<i>Finance and contracts</i>	Local, regional and federal budget data Procurement contracts and transactions Data on building applications	Spending data of all public administrations Procurements by public institutions Realized constructions with budget and actual spending	Analysis of public procurements Crowd monitoring of public spending (e.g. OpenCoesione.it and Monithon.it) Data journalism stories on public spending	More transparency, comparability and accountability of public spending Less mismanagement and corruption Better financial comprehension by citizens
<i>Geospatial</i>	All geographical data including 3D data of ground, buildings etc.	Accurate and complete geographical data API	Linking of geographical data with all other geo-located data	More start-up businesses providing geo-located services
<i>Global Development</i>	Aid data Data on direct foreign investments Income and poverty data	Providing national development aid data according to International Aid Transparency Initiative (IATI) standard	Visualizations of development aid projects Global repository of all development aid projects	More transparency of aid spending leading to higher acceptance of development aid in industrialized countries Less duplication of aid projects
<i>Government Accountability and Democracy</i>	Election data Party and candidate donations Decisions of parliamentarians	Detailed geo-located election data Central repository of all politicians of a country Standardized votes of elected representatives	Transparency platform on donations and vested interests of politicians Monitoring of campaign promises and actual actions	More independent politicians Higher integrity of politicians Less corruption and less hidden conflicts of interest

<i>Health</i>	Data on hospital treatments Data on patients, doctors and medication Data on nutrition and ingredients	Quality indicators of hospitals (e.g. Swiss Inpatient Quality Indicators) Cost of actual treatments and success rates Chances of recovery regarding certain treatments and hospitals	Comparability platform of hospitals and doctors Ranking of medication and reported effects	Better treatments More know-how exchange between hospitals and doctors regarding best practices Cost savings
<i>Science and Research</i>	Data on funding programs and successful proposals Original research/science data	Funding success per institution Open science platform	Ranking of funding applications, directory of best practices Replication of research studies Peer-Review of science data	Better research institutions, more competition among universities Higher research quality, more rigorous results
<i>Statistics</i>	Data on population, economy, social environment, employment etc.	Linked open data platform for using the data connected to other data	Mashup of various data sets, e.g. unemployment and crimes	Better use of available statistical data More insights based on statistical data
<i>Social mobility and welfare</i>	Data about housing, labor market, education level, migration, income etc.	Data enhanced with geo-locations, metadata, financial data etc.	Visualization of housing rents, vendor pricing, demand etc. Visualization of allocation of refugees within a certain region.	More transparency in housing market leading to better pricing of properties Higher acceptance of refugee centers

<i>Transport and Infrastructure</i>	Data on routes of railway tracks, stations, tunnels, cableways etc. Data on traffic accidents	Public transportation data platform with actual traffic Anonymous geo-located accidents	Innovative real-time applications for connecting with public transportation Application warning when approaching a dangerous location	More citizens using public transportation, less use of private cars, less CO ₂ emissions Less accidents
-------------------------------------	--	--	--	---

5. Conclusions

In this article we have introduced the SROI approach as a device to measure the social impact of open data initiatives. By defining two dimensions we have developed an Impact Monitoring Framework providing a method to structure open data activities and connect their input and output with their intended outcomes and areas of impact. The result may be applied in different forms:

First, this framework may be used as a tool to analyze open data activities in retrospect. By identifying existing open data applications and mapping them to the needed open data records (input and output) as well as to their social effects (impact) an assessment of past and current open data activities is facilitated. Thus success stories but also failures where e.g. the intended impact was not achieved may be communicated in a systematic way.

Second, the framework presents an approach to plan the release of data to the public through indicating the intended effects and benefits. By showing the potential impact of open data initiatives decision makers may better understand why they invest resources into open data activities in the first place. Thus the framework acts as tool to implement e.g. a certain open data strategy and to plan future open data activities.

Third, the Impact Monitoring Framework offers a way to monitor the ongoing open data activities of e.g. a public institution or a national government. By tracking the effects of the release of public sector information the Impact Monitoring Framework acts as a tool for portfolio management of open data initiatives leading to a better overview and supporting the use of synergies between different initiatives.

While the framework presents a first proposal to measure and monitor open data activities it is far from being complete in its current form. There are several open issues and further steps to be taken in order to get a thorough monitoring tool. Next, an empirical analysis of each data category reporting the actual and future input, output, outcome and impact areas should to be implemented. Also the influence of cross-category open data initiatives such as e.g. trainings on data literacy or developing an open data strategy is not yet included in the current version of the framework. And, finally, the SROI approach intends to actually put a monetary value to each of

the steps. Thus it should be further investigated how the outcome and impact open data initiatives may be measured in terms of financial value.

References

- Aeron-Thomas, D., Nicholls, J., Forster, S., Andrea Westall, 2004. Social Return on Investment: Valuing what matters - Findings and recommendations from a pilot study. New Economics Foundation (NEF).
- Arvidson, M., Lyon, F., McKay, S., Moro, D., 2010. The ambitions and challenges of SROI. Working Paper. University of Birmingham, Birmingham, UK.
- Bürgi-Schmelz, A., 2013. Wirtschaftliche Auswirkungen von Open Government Data. Schweizerisches Bundesarchiv.
- Carrara, W., Chan, W.S., Fische, S., Steenbergen, E. van, 2015. Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources. European Commission.
- Dapp, M., Stuermer, M., 2015. Schlussbericht Impact Monitoring Framework für OGD Schweiz.
- Davies, T., 2014. Open Data in Developing Countries - Emerging Insights from Phase I, Exploring the Emerging Impacts of Open Data in Developing Countries (ODDC). The World Wide Web Foundation.
- Davies, T., Perini, F., Alonso, J.M., 2013. Researching the emerging impacts of open data - ODDC conceptual framework, Exploring the Emerging Impacts of Open Data in Developing Countries (ODDC). World Wide Web Foundation.
- Davies, T., Raed M. Sharif, Jose M. Alonso, 2015. Open Data Barometer 2015 Global Report - 2nd Edition. The World Wide Web Foundation.
- Emerson, J., Wachowicz, J., Chun, S., 2000. Social Return on Investment: Exploring Aspects of Value Creation in the Nonprofit Sector. Investor Perspectives.
- G8, 2013. G8 Open Data Charter (Policy paper).
- Janssen, M., Charalabidis, Y., Zuiderwijk, A., 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management 29, 258-268. doi:10.1080/10580530.2012.716740
- Kassen, M., 2013. A promising phenomenon of open data: A case study of the Chicago open data project. Government Information Quarterly 30, 508-513. doi:10.1016/j.giq.2013.05.012
- Khan, S., Foti, J., 2015. Independent reporting mechanism - Aligning supply and demand for better governance. Open Government Partnership.
- Klessmann, J., Denker, P., Schieferdecker, I., Schulz, S.E., 2012. Open Government Data Deutschland - Eine Studie zu Open Government in Deutschland im Auftrag des Bundesministerium des Innern Auftraggeber.

- Linders, D., 2012. From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly, Social Media in Government - Selections from the 12th Annual International Conference on Digital Government Research (dg.o2011)* 29, 446–454. doi:10.1016/j.giq.2012.06.003
- Lingane, A., Olsen, S., 2004. Guidelines for Social Return on Investment. *California Management Review* 46, 116–135. doi:10.2307/41166224
- Lourenço, R.P., 2015. An analysis of open government portals: A perspective of transparency for accountability. *Government Information Quarterly* 32, 323–332. doi:10.1016/j.giq.2015.05.006
- Manyika, J., Chui, M., Groves, P., Farrell, D., Kuiken, S.V., Doshi, E.A., 2013. Open data: Unlocking innovation and performance with liquid information.
- Millar, R., Hall, K., 2012. Social return on investment (SROI) and performance measurement: The opportunities and barriers for social enterprises in health and social care. *Public Management Review* 15, 923–941. doi:10.1080/14719037.2012.698857
- Nicholls, J., Lawlor, E., Neitzert, E., Goodspeed, T., 2009. A guide to social return on investment. Office of the Third Sector, The Cabinet Office, London.
- Olsen, S., Lingane, A., 2003. Social Return on Investment: Standard Guidelines.
- REDF, 1999. SROI Excel Model Tool.
- Ryan, P.W., Lyne, I., 2008. Social enterprise and the measurement of social value: methodological issues with the calculation and application of the social return on investment. *Education, Knowledge and Economy* 2, 223–237. doi:10.1080/17496890802426253
- Ubaldi, B., 2014. Open government data: implementing to deliver results Highlights from OECD data collection and analysis, OECD E-Government Project. Organisation for Economic Co-operation and Development OECD.
- Ubaldi, B., 2013. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives (OECD Publishing No. 22), OECD Working Papers on Public Governance. Organisation for Economic Co-operation and Development OECD.
- Veljković, N., Bogdanović-Dinić, S., Stoimenov, L., 2014. Benchmarking open government: An open data perspective. *Government Information Quarterly* 31, 278–290. doi:10.1016/j.giq.2013.10.011
- Zuiderwijk, A., Janssen, M., 2014. Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly* 31, 17–29. doi:10.1016/j.giq.2013.04.003