

Master's Thesis

# **The Impact of Open Government Data in the Case of Education**

submitted to

Faculty of Business, Economics and Social Sciences  
at the University of Bern

Institute of Information Systems

Information Engineering / Information Management

**Dr. Matthias Stürmer**

submitted by

Aydin Taner

of Utzenstorf, Bern (BE)

in the 3. Semester

Matriculation Number: 09-101-569

Address

Hasenmattstrasse 2A

3427 Utzenstorf

+41 79 694 85 49

taner.aydin@hotmail.ch

Bern, 30.12.2016

### **Summary**

In this thesis, Open Government Data initiatives are analyzed. This analysis should give a brief overview of the current situation in the field of education data. The analysis involves the dataflow-chain from the data origin at the schools to the end-users. The analysis shows, which way the data goes, which outputs and which outcomes result out of the process. The impact map on Open Government Data value creation illustrates the findings. The product of this master's thesis is the holistic view of the OGD initiatives with the correlated inputs, outputs, outcomes, impacts and the relations between the different stages.

### **Zusammenfassung**

In dieser Masterarbeit werden Open Government Data Initiativen analysiert. Diese Analyse gibt eine grobe Übersicht über die aktuelle Situation im Bereich der Bildungsdaten. Die Analyse umfasst die Datenflusskette von ihrem Ursprung an den Schulen bis hin zu den Endbenutzern. Diese Analyse zeigt somit auf, welchen Weg die Daten machen und welche Outputs und Outcomes daraus Resultieren. Die Impact Map für Open Government Data Initiativen illustriert diese Erkenntnisse. Das Endprodukt dieser Masterarbeit ist ein holistisches Bild über die OGD Initiativen mit all den Inputs, Outputs, Outcomes, Impacts und deren Beziehungen zueinander zu kreieren

---

## Table of Contents

<b>1</b>	<b>Introduction</b> .....	<b>4</b>
1.1	<i>Background</i> .....	4
1.2	<i>Problem</i> .....	5
1.3	<i>Aim of the Thesis</i> .....	6
1.4	<i>Structure of the Thesis</i> .....	6
<b>2</b>	<b>Literature Review</b> .....	<b>8</b>
2.1	<i>Open Data</i> .....	10
2.2	<i>Open Government</i> .....	11
<b>3</b>	<b>Impact Monitoring Framework and the Social return on investment</b> .....	<b>13</b>
3.1	<i>Impact Monitoring Framework</i> .....	13
3.2	<i>Methodology of SROI</i> .....	13
3.3	<i>G8 Charta</i> .....	25
<b>4</b>	<b>Applying the Impact Monitoring Framework for the high value category</b>	
	<b>“Education”</b> .....	<b>27</b>
4.1	<i>Interview Partners</i> .....	28
4.2	<i>Ecosystem of Open Government Data value creation</i> .....	31
<b>5</b>	<b>Conclusions</b> .....	<b>46</b>
5.1	<i>Findings and Implications</i> .....	46
5.2	<i>Limitations and Suggestions for Further Research</i> .....	46
	<b>Appendix</b> .....	<b>48</b>
	<b>List of Figures</b> .....	<b>108</b>
	<b>List of Tables</b> .....	<b>109</b>
	<b>References</b> .....	<b>110</b>
	<b>Statement of autonomous and independent work</b> .....	<b>113</b>
	<b>Publication of the Thesis</b> .....	<b>114</b>

# 1 Introduction

## 1.1 Background

Managers today pursue the financial returns shareholders expect, without having full information about the impact of their operations on the environment and human well-being. Therefore, there was no possibility to optimize these impacts. The purpose of financial accounting is to provide an overview of the quantified value created by the firms. However, these information do not address the measure of social and environmental impact of the firm. For the financial impact, there is the Return on Investment (ROI) approach. ROI is a globally understood relative measure of a company's success and calculated to compare companies around the world relative to each other in a given industry over time (Olsen and Lingane, 2003). This approach will be extended with the social factor in this thesis to value the social impacts.

In the past couple of years, the field of Open Government Data (OGD) got more and more important. Many countries and governments do set OGD as a goal, which is pushed and promoted. The idea behind OGD is that the data, which is collected by the government, is published. Since these data is already collected by the government, this is not a substantial additional expenditure. In doing so, the transparency of government actions and the collaboration with the citizen can be improved (Gurstein, 2011). The published data can then be re-used by citizens or organizations. Making the right data accessible for the public can generate new businesses (Linders, 2012), growth (Klessmann et al., 2012; Buergi-Schmelz, 2013; Manyika et al., 2013; Carrera et al., 2015) and social value (Janssen et al., 2012), when published in the right form. According to the OECD, OGD can create public value (OECD, 2015). It's also believed that OGD can contribute to the creation of public governance (Stuermer and Ritz, 2014). Linders says that, new applications and businesses create growth, innovations, workplaces and welfare for society (Linders, 2012). Opening data to the public can therefore have various positive effects. However, the lack of legal foundations, technical barriers and cultural shifts are holding back the release in OGD, even if OGD policies

are formulated and established by many governments (Zuiderwijk and Janssen, 2014; Janssen et al., 2012).

## 1.2 Problem

Measuring the value OGD generates is very complicated and a challenging task. There is a broad consensus among scientists, politicians and other involved groups that OGD does have various positive impacts. But there is a disagreement on how to measure these impacts. Since the return on investment (ROI) does not measure the social component of value, there is a need for another tool to do so. There the social return on investment (SROI) comes into play. The goal of SROI is to measure the social benefit in a monetary form as the ROI does. Roberts Enterprise Development Fund introduced the SROI approach in the late 90's. The SROI was developed to measure the social value of philanthropic investments. Since the year 2000 the SROI is getting developed in the USA, UK, Canada and the Netherlands. (Emerson et al., 2000)

Sara Olsen and Alison Lingane made a subsequent empirical assessment with the business plans of 88 companies during the Global Social Venture Competition. They provided guidelines to extend the SROI approach on how to calculate the monetary value of social value. The SROI approach is divided into four different pieces: 1. Input, 2. Output, 3. Outcome and 4. Impact (Areon-Thomas et al., 2004). The SROI approach has not been used for OGD impact measuring so far. In the past it's been only used to measure the impact of other projects. Introducing the OGD Impact Monitoring Framework, Dapp and Stuermer (2015) suggest to apply the SROI approach to measure the value of OGD in the defined 14 high-value categories.

### **1.3 Aim of the Thesis**

The aim of the thesis is to empirically analyze a data category with the “Impact Monitoring Framework” introduced by Dapp and Stuermer (2015). In their paper, they suggest to choose a high value category of the G8 Open Data Charter for applying the Impact Monitoring Framework. For this thesis, the category “Education” was chosen. Therefore this thesis is a first attempt to value the impact of Open Government Data in the high-value category of Education for Switzerland, using the Impact Monitoring Framework.

Education is a topic, which accompanies us throughout years in our lives. There is a lot of data about education at its origin. These can be the numbers of students, or the characteristics of students, the locations of the schools, the disciplines, the annually budget and so on. Analyzing this data could significantly improve the education and therefore also the lives. A lot of the data is available on a national basis. Considered these reasons, the Open Government Data in the field of education is a reasonable category to apply the Impact Monitoring Framework. The education data was also re-used by third parties. Those developed some innovative apps or informative tools to make all this data more accessible for laymen.

### **1.4 Structure of the Thesis**

#### ***1.4.1 Methodological Approach***

The Impact Monitoring Framework developed by Dapp and Stuermer (2015) is the groundwork of this thesis. This Thesis is an empirical analysis for the high-value category “Education”, which is one of the 14-high-value data categories defined by the G8 summit (G8, 2013). The data collection is the main part of an empirical analysis. The data collection will be assured by qualitative, structured interviews with different stakeholders in the field of education. The stakeholders are defined as experts from different institutions, companies or private persons that work with Open Education Data.

---

Therefore, the methodological approach of this thesis is an empirical analysis which brings empirical insights throughout qualitative surveys in form of face-to-face interviews.

### **1.4.2 Structure of the Thesis**

Following this background introduction, there will be a short literature review. This review will give a brief overview of the actual state of affairs in the field of giving Open Government Data projects a value. Throughout this literature review we will discern that there is a positive correlation between Open Government Data projects and creation of social value, hence value in general. We will also see that the empirical research is still in its fledging stage. The third chapter is about the “Impact Monitoring Framework”. This framework from Dapp and Stuermer, which is the foundation of this thesis, will be introduced in detail. This will give the basis for the underlying valuing process, the Social Return on Investment (SROI). The SROI approach is explained very explicitly, since this is the part where the value is given to Open Government Data projects. This chapter is followed by the high value categories, which were defined by the G8 Charta (G8, 2013). The main chapter of this thesis is chapter four. This chapter shows the empirical work, which was performed using the face-to-face interview method. In a first step the interview-partners are introduced. Then the ecosystem of Open Government Data value creation is described. In this part, the four factors, Input, Output, Outcome and Impact are discussed in detail with all the finding from the interviews. The last chapter summarizes the conclusions and gives a brief suggestion for future research.

## 2 Literature Review

This part of the thesis will give an overview over the stage of literature of Open Government Data projects. Although Open Government Data is constantly growing and the available data is increasing, the theoretical research is still in its very early stages. (Dapp & Stuermer, 2015, p. 6).

A first prove that open data can have a positive impact is shown by a case study from Chicago. This case study explains how open data portals operated by the government supported citizens to engage more in public administration, leading to more transparency in terms of political lobbyism and safety in public parks (Kassen, 2013). Research in developing countries exploring the impact of open data initiatives proved the positive effect. These initiatives lead to more transparency and accountability, increased innovation as well as economic development, improved inclusion and empowerment of citizens (Davies et al., 2013). Davies et al. also delivered empirical evidence of the positive effects of open data initiatives. They analyzed case studies in developing countries like Kenya, Nepal or India showing that such initiatives can create new space for government accountability and effectiveness. Furthermore they showed that open data initiatives change the way governments create, prepare and use its data (Davies, 2014). The impact of open data initiatives is also indicated by quantitative studies. The Open Data Barometer by Davies et al. (2015) assesses the results from all around the world, analyzing the impact of open data initiatives including the identification of their outcome. The Open Data Barometer measures the entrepreneurial effects, the transparency and the accountability, the government efficiency, the economic growth, the environmental sustainability and the inclusion of marginalized groups. It shows that entrepreneurial open data use is the most observed impact from such initiatives whereas the second most observed impact is the Transparency and accountability, followed by government efficiency and economic growth (Davies et al., 2015). It is important to mention that the Open Data Barometer is a width report analyzing global trends and therefore no in-depth assessment is made. The provided insights are based on expert opinions that are used as a proxy to measure and evaluate the impact.



Additionally to the empirical studies, there have been several studies on generic indicators of open data policies and strategies. Barbara Ubaldi encouraged a set of metrics in her paper from the year 2013 on theoretical fundamentals and presented an analytical framework (Ubaldi, 2013). In a subsequent empirical paper, Ubaldi surveyed 25 OECD countries regarding “data on OGD (Open Government Data) policy making and implementation”. This second paper is the empirical implementation of the first paper and therefore does not add any new indicators. Interesting, however, is the fact that 96% of the respondents (N=16) answered with “No” to the question “Has the government adopted a methodology to measure return on investment on investment on OGD (e.g. potential cost savings, value for new services, more efficient service delivery)?” (Ubaldi, 2014).

Comparing the different Dutch Open Data directives, Zuiderwijk and Janssen found out that the goals of politicians and public servants are often not congruent when it comes to achieving an impact with the release of Open Data. Managing the risk and the complexity of releasing Open Data, often causes public servants having a hard time. (Zuiderwijk and Janssen, 2014)

There was an attempt to compare governments by their open government maturity. The use of the ecosystem metaphor is suggested in the paper of Harrison et al. The interdependency of organizations, actors and infrastructures is the source of value creation in Open Government Data initiatives. The approach of an ecosystem helps to understand this interdependency. There are several types of impact identified by Harrison et al. Financial, political, social, strategic, quality of life, ideological and stewardship just to name a few. (Harrison et al., 2012) A Study from Jetzek et al. proves that Open Government Data initiatives create economic as well as social value. But there is no suggestion given on how to value and measure the created economic and social value. (Jetzek et al., 2012)

## 2.1 Open Data

Open Data is about data, which is publicized by an organization to others for further processing (Klessmann, 2012). According to the Open Knowledge Foundation, Open Data needs to be understood as a set of data, which complies with certain principles (Open Definition, 2015). The 10 principles of Open Data are shown in the Table 1 below. When data does meet these principles it can be labeled as Open Data (Klessmann, 2012).

<b>1. <i>Completeness</i>: All public data will be published.</b>
<b>2. <i>Primacy</i>: The data is collected in its origin.</b>
<b>3. <i>Timeliness</i>: The data is published immediately.</b>
<b>4. <i>Ease of Physical and Electronical Access</i>: The data is provided for as many purposes as possible.</b>
<b>5. <i>Machine readability</i>: Data is saved in a open, structured format.</b>
<b>6. <i>Non-discrimination</i>: Data is available without registration.</b>
<b>7. <i>Use of Commonly Owned Standards</i>: There is no necessity for proprietary software for the data interpretation.</b>
<b>8. <i>Licensing</i>: Data is available under a free license.</b>
<b>9. <i>Permanence</i>: Data stock is permanently online and versioned available.</b>
<b>10. <i>Usage Costs</i>: The access to the data is free of charge.</b>

Table 1: 10 Principles of Open Data (Sunlight Foundation, 2016)

In the public sector, Open Data has a huge potential. Through the opening of these data, the public could gain a better transparency in the processes and decisions of the departments of the government (Klessmann, 2012).

In the case of this thesis, the definition of Open Data or Open Government Data is different. When referred as Open Data or Open Government Data, the underlying data is not genuinely Open Data as described above. This means solely that this data is made public by the government actively or upon request and that this data is free of charge.

## 2.2 Open Government

Since taking office of Barak Obama, the former president of the United States back in 2009, Open Data is defined through these three elements (Obama, 2009):

- Transparency
- Participation
- Collaboration

The *transparency*, which is promoted by statutory basis, represents the most important pillar of the open government. Such publicity or freedom of information laws are a part of the constitution of roughly 100 countries (Right2INFO, 2012). As long as government data does not jeopardize the safety of individuals or the safety of the country, these data has to be public. This data is available upon request. The open government approach however involves that this data is not just passively available but rather actively made public by the government. This free accessible government data is a prerequisite for all open government applications. (Stuermer and Ritz, 2014) The relationship between transparency, information and democracy is fundamental. In governmental forms in which the citizens vote and therefore delegate the power to make decisions, the transparency is essential for long-term trust. Without transparent information, it is almost impossible for the citizen to hold the elected government accountable for their actions (Harrison et al., 2012a).

The *participation* as the second pillar of the open government, gives the citizens the possibility to take part in the public processes (Stuermer and Ritz, 2014). This is not a reinvention of open government rather an attempt to increase the participation, which is a fundamental grasp of democracy (Kwak, 2012). With the increase in participation, the political decisions and public services can be improved (McDermott, 2010).

Once participation is achieved, the next step is *collaboration*. Collaboration refers to the cooperation of the citizen with the authorities in the administrative tasks of the government. There is a joint, target-oriented discussion about issues with a subsequent development of solution approaches (Hilgers, 2012).

The collaboration delimits itself from participation with complex tasks and the coproduction of results. These may include initiatives as crowd funding, crowd knowledgebase and crowd monitoring (von Lucke, 2014).

### 3 Impact Monitoring Framework and the Social return on investment

#### 3.1 Impact Monitoring Framework

In 2015, Dapp and Stuermer developed the Impact Monitoring Framework to measure the outcomes and impact of open data activities. The Impact Monitoring Framework could be used to analyze open data activities retrospectively as well as to plan the release of new Open Government Data projects to the public, indicating the intended effects and benefits. This Framework offers also a way to measure and monitor ongoing open data activities. For the measurement of the impact, the Social Return on Investment (SROI) method is used. The SROI is elaborated in detail in the following chapter (Dapp and Stuermer, 2015).

#### 3.2 Methodology of SROI

Every SROI has a theory of change, which is represented by the process from the edited data (input), to the publication (output), on to the actions of others (outcome), which finally leads to the difference, to the situation without open data (impact). (Dapp and Stuermer, 2015, p.12)

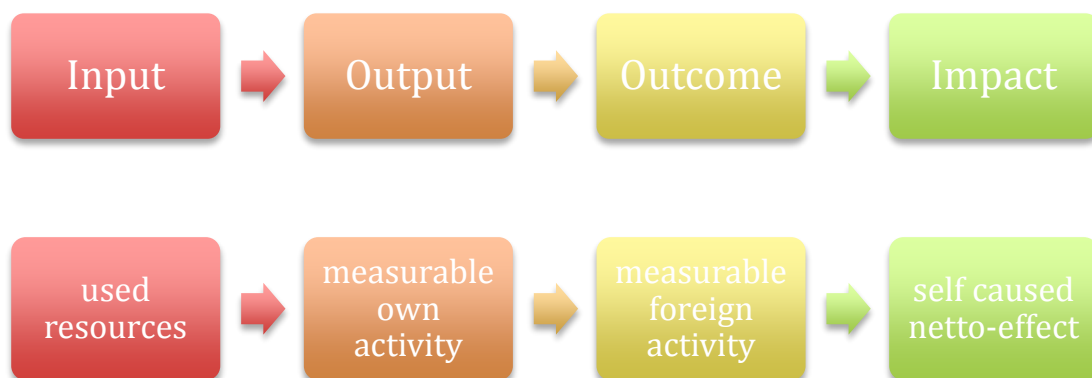


Figure 1: Theory of Change (Dapp and Stuermer, 2015)

In the SROI framework used in this paper, the approach is divided into four different steps: Input, output, outcome and impact. The institutions publishing the data directly control input and output, where outcome and impact are not. (Nicholls et al. 2009)

**Input:** “Input” in SROI describes the resources used to execute a certain action by an organization. These include, money, human capital, buildings and equipment. In open data input covers also the costs to transform native data into raw open data. Native data is the data collected by the organizations and raw open data, according to open data knowledge foundation, is data, which was processed following the 10 principles of open data. (Sunlight Foundation, 2016)

**Output:** In SROI, by “output” is meant the output, which an organization generates. In terms of an open data ecosystem, this refers to an open data portal where the published open data is available for the public under official open data licenses. There are some significant differences in those platforms concerning quality or completeness of the data and other criteria.

**Outcome:** According to SROI the “outcome” includes all direct and indirect consequences of certain actions of the organization. In the view of open data we consider all results of the users of published open data as outcome. This means everybody using this data generates outcome. This includes the engineers, entrepreneurs, citizens, journalists, scientists, artists or other groups using these open data. The outcome cannot be controlled by the organization publishing the data. They can only support and encourage re-use of the data. The outcome can result in form of organizations of hackathones, development of open data visualizations, programming of web and mobile apps, initialization of open data based business models and many others.

**Impact:** In SROI the “impact” is defined as the difference between an outcome without the publishing of open data and the outcome with the publishing and re-using open data. This means that only results arising from the re-

lease of the data is to be counted. In practice this is really hard to measure. One possible way is to use comparison groups or benchmarks.

The model of SROI is based on these seven principles (Nicholls et al., 2009):

1. *Involve stakeholders: Inform what gets measured and how this is measured and valued by involving stakeholders.* Stakeholders are those people or organizations that experience change as a result of the activity and they will be best placed to describe the change. This principle means that stakeholders need to be identified and then involved in consultation throughout the analysis, in order that the value, and the way that it is measured, is informed by those affected by or who affect the activity.
2. *Understand what changes: Articulate how change is created and evaluate this through evidence gathered, recognizing positive and negative changes as well as those that are intended and unintended.* Value is created for or by different stakeholders as a result of different types of change; changes that the stakeholders intend and do not intend, as well as changes that are positive and negative. This principle requires the theory of how these changes are created to be stated and supported by evidence. These changes are the outcomes of the activity, made possible by the contributions of stakeholders, and often thought of as social, economic or environmental outcomes. These are outcomes, that should be measured in order to provide evidence that the change has taken place.
3. *Value the things that matter: Use financial proxies in order that the value of the outcomes can be recognized. Many outcomes are not traded in markets and as a result their value is not recognized.* Financial proxies should be used in order to recognize the value of these outcomes and to give a voice to those excluded from markets but who are affected by activities. This will influence the existing balance of power between different stakeholders.

4. *Only include what is material: Determine what information and evidence must be included in the accounts to give a true and fair picture, such that stakeholders can draw reasonable conclusions about impact.* This principle requires an assessment of whether a person would make a different decision about the activity if a particular piece of information were excluded. This covers decisions about which stakeholders experience significant change, as well as the information about the outcomes. Deciding what is material, requires reference to the organization's own policies, its peers, societal norms, and short-term financial impacts. External assurance becomes important in order to give those using the account comfort that material issues have been included.
5. *Do not over-claim: Only claim the value that organizations are responsible for creating.* This principle requires reference to trends and benchmarks to help assess the change caused by the activity, as opposed to other factors, and to take account of what would have happened anyway. It also requires consideration of the contribution of other people or organizations to the reported outcomes in order to match the contributions to the outcomes.
6. *Be transparent: Demonstrate the basis on which the analysis may be considered accurate and honest, and show that it will be reported to and discussed with stakeholders.* This principle requires that each decision relating to stakeholders, outcomes, indicators and benchmarks; the sources and methods of information collection; the difference scenarios considered and the communication of the results to stakeholders should be explained and documented. This will include an account of how those responsible for the activity will change the activity as a result of the analysis. The analysis will be more credible when the reasons for the decisions are transparent.
7. *Verify the result: Ensure appropriate independent assurance.* Although an SROI analysis provides the opportunity for a more complete understanding of the value being created by an activity, it inevitably involves



subjectivity. Appropriate independent assurance is required to help stakeholders assess whether or not the decisions made by those responsible for the analysis were reasonable.

### **3.2.1 Six steps to establish a SROI valuation:**

The following six steps to establish a SROI valuation are extracted from “a guide to social return on investment” from Nicholls et al. (2009).

#### **3.2.1.1 Stage 1: Establishing scope and identifying stakeholders**

In the first stage, there are three steps:

##### **1. Establishing scope**

The scope defines the boundaries of a SROI analysis of what is being considered. It's the part where it's needed to be declared what to measure and what is feasible. In order to set the scope, there have to be considered different questions:

1. What is the purpose of the SROI analysis?
2. Who is the audience?
3. Background → How to try to make a difference?
4. What resources will be required?
5. Who will carry out the work?
6. Focused range of activities?
7. Period of time for which the SROI will be made?
8. Is it a forecast or an evaluation?

##### **2. Identifying stakeholders**

Stakeholders in the case of SROI are defined as organizations or people who experience change through the activity, which is being analyzed. The goal of SROI is to find out how much value was generated or destroyed and who did benefit. To define the stakeholders, every person or organization affected by the activities within the scope has to be listed.

Once the list is accomplished, it has to be decided which stakeholders are relevant. The focus is tending to be set on the positive outcomes for the stakeholders however all the outcomes, whether positive, negative, intended or unintended are relevant for the SROI. To help to organize the relevant stakeholders, it's helpful to build a table with the key stakeholders and the reasons for inclusion.

### **3. Deciding how to involve stakeholders**

The most important part is to collect information. It's important to involve stakeholders in the best possible way. This can be done in various ways like:

- One-2-one interview with stakeholders
- Group interviews with stakeholders
- Workshops with stakeholders
- Let the stakeholders fill out a questionnaire

The information should ideally be collected directly from the stakeholders. Information coming from existing research with stakeholders might be applicable when there is a lack of time or resources.

It's important to have a representative sample. A commonly used approach is the so-called "saturation". In the case of interviews this could mean that one can stop doing new research whenever the information collected is overlapping and no new findings are found.

The part of data collection is a very time consuming work. It's important to plan ahead and use the time effectively.

#### **3.2.1.2 Stage 2: Mapping outcomes**

The Impact Map is a tool, which helps to organize the different stages of the SROI approach. It's usually a spreadsheet that is filled out with the information collected from the stakeholders. On the top section of the Impact Map there is the information about the project, which is valued and the scope of the SROI. Further it's divided into different sections, which represent the stages 1 to 5. There are four steps to fill out an Impact Map:

### **1. Identifying inputs**

The inputs are the investments of the SROI. This can be money or time spent for the activity. This refers to the financial value of the inputs.

### **2. Valuing inputs**

Inputs other than financial investments have to be valued and monetized. This is often the case with volunteer time. If the activity would have a significant change without these inputs, they have to be valued. There are two main types of non-monetized inputs relevant in the SROI: Volunteer time and provision of goods and services in any kind.

For the value of the volunteer time it's the best practice to value them according to the equivalent value of the average hourly rate for the same kind of work performed. This value has to be attached to that volunteer time even if no cash was paid.

### **3. Describing outcomes**

The approach of the SROI is an outcome based measurement tool. This means that the outcome is the source of the change for the stakeholders. For example, if the aim of education is getting better jobs than the completion of the education itself is an output not an outcome. Getting a better job is an outcome. It's very important to measure the right things, in this case the outcome, which may not always be intuitive.

#### ***3.2.1.3 Stage 3: Evidencing outcomes and giving them a value***

In stage 3 the outcome indicators are developed and used to collect evidence on the outcome that is occurred.

#### **1. Developing outcome indicators**

To measure the change there is a need for indicators. The indicators measure if a change by the outcome has occurred and if so, by how much.

## **2. Collecting outcomes data**

New data usually comes from people working in organizations, which are directly involved creating social value. The most common forms of data collection are:

- One-2-one interviews
- Record keeping
- Focus groups
- Workshops and seminars
- Questionnaires

The size of the sample should be representative and depends on the feasibility and budget. The chosen sample size should be defensible.

## **3. Establishing how long outcomes last**

Not every outcome has the same duration. Some last longer than others.

When the assumption is made that the outcome will last after the activity has stopped, then this will also continue to generate value. This can be defined in the benefit period of an outcome. For every outcome there has to be an estimation of the benefit period, which can be determined in a discussion with the stakeholders creating the social value.

## **4. Putting value on the outcome**

The next step is to give each outcome a financial value. It's important to remember that the value is given to the outcome and not to the indicator.

The value given to the outcome is an approximation because these transactions do not have a market price. The value will differ from person to person and situation to situation. The value is highly subjective and there is no "objective" or "true" value in this case.

#### **3.2.1.4 Stage 4: Establishing impact**

Stage 4 is assessing whether the outcome is a result of the given activities. These methods provide an approximation of which outcome would have occurred anyway and which outcome was due to the activities.

This is called establishing impact. As a result, risk of over claiming is reduced and the credibility increases.

#### **1. Deadweight and displacement**

*Deadweight* describes a measure of outcome that would have happened even if the activity did not take place. This means that this is a calculation of the change in percentage, which would have occurred anyway without the value creating activity. To calculate this deadweight there is a need for comparison groups or benchmarks. The perfect comparison would be a group with same characteristics, which did not benefit from the activity like the first group. But in reality this can be really difficult to achieve. Therefore, the deadweight will be an estimate.

*Displacement* is another component of impact and describes how much of the outcome has been displaced to other outcome. For example, the activity's objective is to reduce crime in a certain area. In such cases it's possible that the crime in this area might shrink but in other areas it might increase as a result of this crime-reducing activity in the first area. This is called displacement. However, displacement is not relevant in every SROI.

#### **2. Attribution**

Attribution is an evaluation of how much of the outcome was actually caused by the activity. It might be the case, that other activities from other organizations have a similar effect. Then the attribution has to be calculated, how much was due to the purposed activity and how much is due to other activities from other organizations. It is impossible to get an accurate calculation of the attribution. Attribution is rather to increase the awareness that the outcome might have other causes.

There are three main practices to estimate attribution:

- Estimation based on experience
- Estimation based on expertise of stakeholders
- Estimation based on consulting other possible attributing organizations

### **3. Drop-off**

Drop-off is used to account for the fact that as the years pass, the amount of the outcome is likely to be less or, if the same, more likely to be influenced by other factors. This decreases the attribution of the activity of the primary organization. Drop-off is only calculated for outcomes, that last longer than one year. The amount, which is dropped-off, is a fixed percentage from the remaining level of the outcome every year.

### **4. Calculating your impact**

Once all these aspects of impact are defined, they need to be calculated. For the calculation, the outcome is multiplied by the financial proxy, which gives us a total value. Then this value has to be adjusted for the deadweight, the displacement and the attribution. Since these aspects are expressed as percentages, it's easy to adjust for them. These percentages have to be deducted from the total value, which was calculated beforehand. Once this procedure is made for every outcome all the totals from each outcome have to be added up. This is then the resulting impact of the activity.

#### ***3.2.1.5 Stage 5: Calculating the SROI***

In this part the financial value of the investment and the financial value of the social costs and benefits are calculated. The results are two numbers of which the relation can be reported in several different ways.

#### **1. Projecting into the future**

The first step of the calculation is the projecting the value of the outcomes into the future. The outcome for one time period was set before. So now this outcome has to be adjusted for any drop-off. This procedure has to be made for every year, which the outcome lasts.

## 2. Calculating the net present value

Since money has a time value, the calculation of the net present value (NPV) is absolutely essential. In order to calculate the NPV, the paid or received costs and benefits have to be added up. These then have to be discounted.

The usage of NPV in SROI is a controversial area in the on going research and discussion. The main critic on NPV is the encouragement of short-termism. In environmental outcomes this can be problematic, because the value of the outcome may even increase. However to set an appropriate discount rate is very difficult. The HM Treasury's Green Book recommended a discount rate of 3.5% for the public sector.

## 3. Calculating the ratio

The initial SROI ratio is a simple sum. The discounted value of benefits is divided by the total investment.

$$SROI = \frac{\textit{Present Value}}{\textit{Value of inputs}}$$

The net SROI is an alternative calculation. This ratio divides the NPV by the value of the investment.

$$\textit{Net SROI} = \frac{\textit{Net Present Value}}{\textit{Value of inputs}}$$

Both of these ratios are applicable, but it has to be clear which one is used.

## 4. Sensitivity analysis

After the calculation of the SROI, it's possible to make a sensitivity analysis with the spreadsheet. This assesses the importance of elements of the model. Since the model is based on some assumptions, there is a need to assess how much the result would change if the assumptions change. This analysis aims to test which assumptions have the greatest effect in the model.

The sensitivity analysis has to involve changes to:

- Estimates of deadweight, attribution and drop-off
- Financial proxies
- Quantity of the outcome
- Value of inputs

The recommended approach is to find out how much change in each estimation is needed to make the ratio change from positive to negative or vice versa. The bigger the needed change in the estimation, the less sensitive is this estimation. All these findings have to be discussed in the final SROI report.

### **5. Payback period**

The payback period describes how long it takes the investment to pay off. Basically it says at what point the social value starts to exceed the investment. The payback period is reported in months. Thus, the annual impact has to be divided by 12 to get the monthly impact. Then the investment has to be divided by the monthly impact. The payback period formula is:

$$\text{Payback Period (in Months)} = \frac{\text{Investment}}{\text{Annual impact}/12}$$

#### **3.2.1.6 Stage 6: Reporting, using and embedding**

The last stage of the SROI is reporting to the stakeholders, communicating and using the results, and embedding the SROI process in the organization.

##### **1. Reporting to stakeholders**

When reporting to the audiences, which were defined before, it has to be made sure that the way of communication is relevant for them. The target group may be internal or public or any others. With the preparation of a report recommendations can be made to influence what happens in the future when things are moving forward.

To fulfill the accountability of the SROI towards the stakeholders, both the process and the results have to be published. It has to be a report where not only the social return but also an qualitative, quantitative and financial aspects have to be analyzed and assessed. The report has to include enough



information to allow others reading the report to be assured that the calculations are robust and accurate.

A SROI report should include following information:

- Information relating to the organization, work, key stakeholders and activities
- Description of the scope of the analysis, details of stakeholder involvement, methods of data collection, and any assumptions and limitations underlying the calculations
- The impact map with the relevant indicators and approximations
- Case studies or quotes from participants that illustrate particular findings
- Details of the calculations and a discussion of the assumptions and estimates including a sensitivity analysis
- An audit trail for decision making including which stakeholders, indicators and outcomes were included and which were not
- An executive summary aimed at a broad audience including the participants

It's also important to distinguish which benefits are not happening and which benefits could have happened but cannot be evidenced. There have to be also recommendations included in the report, how to improve the data collection and evidencing.

## **2. Using the results**

The most important part of an SROI analysis is using the results. If the results are not used, there was not much point in undertaking it in the first place. The SROI analysis should always result in change.

## **3.3 G8 Charta**

Dapp and Stuermer (2015) suggest conducting the Open Data Charta of the G8 Summit 2013 for the analysis through the SROI approach. The Open Da-

ta Charta defined these 14 categories to be high-value data categories to improve the democracies and encourage innovative re-use of the data. The countries following this Charta should engage in opening-up their data in those 14 high-value categories. This is a widely spread classification of the sector-specific categorization. These categories are shown in the following Table 1. (G8, 2013)

<b>Data Category</b>	<b>Example dataset (from G8 Open Data Charter)</b>
<b><i>Companies</i></b>	Company / Business register
<b><i>Crime and Justice</i></b>	Crime statistics, safety
<b><i>Earth observation</i></b>	Meteorological / weather, agriculture, forestry, fishing and hunting
<b><i>Education</i></b>	List of schools, performance of schools, digital skills
<b><i>Energy and Environment</i></b>	Pollution levels, energy consumption
<b><i>Finance and contracts</i></b>	Transaction spend, contracts let, call for tender, future tenders, local budget, national budget
<b><i>Geospatial</i></b>	Topography, postcodes, national maps, local maps
<b><i>Global development</i></b>	Aid, food security, extractives, land
<b><i>Government Accountability and Democracy</i></b>	Government contact points, election results, legislation and statutes, salaries (pay scales, hospitality / gifts)
<b><i>Health</i></b>	Prescription data, performance data
<b><i>Science and Research</i></b>	Genome data, research and educational activity, experiment results
<b><i>Statistics</i></b>	National statistics, Census, infrastructure, wealth, skills
<b><i>Social mobility and welfare</i></b>	Housing, health insurance and unemployment benefits
<b><i>Transport and Infrastructure</i></b>	Public transport timetable, access points for broadband penetration

Table 2: 14 High-value categories (G8, 2013)

## **4 Applying the Impact Monitoring Framework for the high-value category “Education”**

In this chapter, the preceding theory will be applied to the field of Open Government Data in the case of Education. The 1<sup>st</sup> stage of the theory describes the scope, the stakeholders and the involvement of the stakeholders.

The purpose of this analysis is to raise the awareness concerning this data. This thesis should provide a first attempt to value such open Education Data projects. It's important to perceive the OGD actions and the related impacts arising. There is no specific audience targeted, rather than a broad range of readers desirable. The stakeholders affected by this work are the statistical departments of Switzerland and any other institutions, organizations or private persons using openly available Education Data. These stakeholders are involved by a representing member, through a one-to-one interview, where a predefined set of questions is asked in a qualitative interview. Therefore, first the interview partners will be introduced in the subsequent chapter.

In a second step the different inputs, outputs, outcomes and the impact will be described in detail. According to the theory the next step would be stage 2 of the SROI guidelines. In this stage the inputs would be identified and valued. Giving these inputs and outputs a value is one of the main tasks in the SROI guidance of Nicholls e al. (2009). This was also the original plan when starting this thesis. After the first few interviews, there was an observation made, that the anticipated information were not available. The gathered evidence from the interviews does not provide sufficient findings to ad a proper value. Because of the missing values, the approach of SROI cannot be performed in the way the theory stipulates this. Therefore, the approach is slightly changed and the focus of this thesis is the theory of change and the ecosystem of OGD value creation. Instead of an impact map in form of a spreadsheet, an impact map in form of a graph is used to describe the relations and cohesions between inputs, outputs, outcomes and impacts. The theory of change is represents a dynamic process starting from gathering the data to the impact. Onwards from stage 2, the SROI theory from Nicholls et al. had to be discarded due to the lack of necessary information and the focus is shifted to the Ecosystem of Open Government Data value creation.

## 4.1 Interview Partners

For this thesis, there were experts and other stakeholders interviewed to gather information, expertise and future trend estimations about Open Government Data in the field of Education. Eight different stakeholders from different fields, organizations and part of Switzerland were interviewed. The interviews were structured, face-to-face interviews held with private persons or employees of the given department or organization. Below, these interview partners are shortly introduced.

### 4.1.1 Tobias Schalit

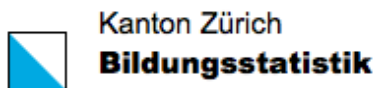


Figure 2: Logo Canton Zurich  
(Bildungsstatistik Kanton Zürich,  
2016)

Tobias Schalit is a scientific employee and statistician at the education statistics of the Canton of Zurich. The education statistics of the Canton of Zurich has the legal mission to gather statistical data of the education field and to analyze this data. The data is the foundation of political planning and decision processes for the education sector. The goal of this department is it to prepare the educational system of the Canton of Zurich for future needs.

### 4.1.2 Jürgen Allraum



Figure 3: Logo Canton Bern (Er-  
ziehungsdirektion Kanton Bern,  
2016)

Jürgen Allraum is the technical manager of the education statistics department of the Canton of Bern. The education statistics of the Canton of Bern has the legal mission to gather statistical data of the education field and to analyze this data. The data is the foundation of political planning and decision processes for the education sector. The goal of this department is it to prepare the educational system of the Canton of Bern for future needs.

### 4.1.3 Ursula Telley



Figure 4: Logo Canton Bern (Erziehungsdirektion Kanton Bern, 2016)

Ursula Telley is the statistics coordinator of the Canton of Bern. She is responsible for the coordination of the statistical tasks in the Canton of Bern. Bern does not have a statistical department as other Cantons do. The statistic coordination acts as a single point of contact in statistical concerns. The coordination is more of a organizational nature. This department does not get in touch with the gathered data.

### 4.1.4 Rolf Lischer



Figure 5: Logo Canton Solothurn & Board of School Solothurn (Volksschulamt Kanton Solothurn 2016)

Rolf Lischer is the coordinator of the education statistics of the Canton of Solothurn. He works as a scientific employee at the department of the board school of Solothurn. The department of the board school is the regulator of all board schools and is responsible for enforcement measures. The main missions are the educational, didactic, organizational and the staff administrative concerns.



Fidel Thomet  
@fidelthomet

Figure 6: Twitter Profile Fidel Thomet (Thomet, 2016)

### 4.1.5 Fidel Thomet

Fidel Thomet is a former student of the Zurich University of the Arts. He studied text design as he got in touch with data visualization. After his studies, he made an internship at Open Data Zurich. He made a visualization based on the data of SRFdata.

#### 4.1.6 Julian Schmidli



Figure 7: Logo SRFdata (SRFdata, 2016)

Julian Schmidli is a data journalist at SRFdata. SRF Data is the data team of the SRF. They work tri-medial with all editorial information offices. SRFdata is the com-

petence center for all questions regarding data. They advise editorial offices with statistics and data. A big part of their work is to gather information and data for news coverage or analyses.

#### 4.1.7 Carole Signer



Figure 8: Logo Eckhaus AG (Eckhaus AG, 2016)

Carole Signer is an employee at the Eckhaus AG in Zurich. Eckhaus AG is specialized in land-use and city planning. In association with Ubique, a software developer, they developed an app called Eckdaten App. This application, which is available for IOS user only, is a tool to observe demographic changes. Eckhaus AG makes use of this app to advise communities and cities regarding planning of the educational infrastructure.

#### 4.1.8 Matthias Kipfer

### Schulgemeinde Wallisellen **Wallisellen macht Schule**

Figure 9: School of Wallisellen (Schule Wallisellen, 2016)

Matthias Kipfer is the head of school administration of the schools in Wallisellen, Zurich. The school administration in Wallisellen was advised by Eckhaus AG, regarding the infrastructure planning of the schools in their community.

## **4.2 Creating the Ecosystem of Open Government Data Value Creation**

The ecosystem of Open Government Data value creation is a system, which describes the different inputs, outputs, outcomes and impacts. Furthermore, this system illustrates the dependencies and relations between the inputs, outputs, outcomes and impacts. Based on this ecosystem it's possible to visualize the coherences between the different components.

In a first step, the different components were made out in an interview with stakeholders of data providers and data users. In a second step, the dependencies and relations between the components were identified.

An impact model helping the stakeholders to understand the theory of change underlying their actions is the core conclusion of this ecosystem (Harrison et al., 2012).

To ensure the knowledge-transfer from research to practice, decision-makers from different administrative organizations and from civil society were directly involved in the validation of the results and findings of this ecosystem.

### **4.2.1 Input**

The input in this thesis is basically the collected data. More specific, it's the collected open education data. The cantonal authorities gather this data from the schools right at its origin. Then the data is aggregated at the cantonal level. There are big differences between the cantons. Some cantons do more than is required and some do it just to the required extends due to financial straits.

#### **4.2.1.1 Statistics**

In Switzerland the Charta of the public statistics of Switzerland defines the general conditions of government statistics. According to the Charta, the public statistics has the following mission: To cover the need of statistical relevant information, which are public interest as well as to fulfill the public duties. (Charta öffentliche Statistik der Schweiz, 2012) It's also enshrined that statistical information is a public good, which has to be published, as long as some criteria are met.

The interviewed statistic departments mainly gather data for three major statistics:

- Student Data
- Teacher Data
- Graduation Data

### **Student Data**

The gathered student data contains several information about the class and the student. The information collected about the class is the location of the class, the characteristics of the class, the education form, the language of education and the type of the class. For the student in the other hand the collected data is made of the social security number, the birthdate, the gender, the nationality, native tongue, the domicile and some information about the discipline. This statistics aims to improve the education planning and education politics. It serves also to calculate the financing flows from the cantons to the municipality and back as well as the calculation of the federal subvention in education. (Erziehungsdirektion Kanton Bern, 2016)

### **Teacher Data**

The statistics of teachers is a publication of information about the teachers, school administration staff, special education staff and the executed tasks. This statistic serves the goal of building a basis for education planning and education policy, the analysis of sociodemographic attributes and the occupational situation of the school staff on a national as well as international basis.

### **Graduation Data**

The graduation data statistics contains the annually completed graduation degrees of the secondary education and tertiary stage B. The secondary education degrees include different types of school leaving examinations, such as: “gymnasiale Maturität”, “Fachmittelausweis”, “Berufsmaturität” and “Passarelle”. The tertiary stage B includes the diplomas of the higher profes-



sional schools, diplomas from the non-government-regulated higher professional education and degrees from post-graduate studies.

#### **4.2.2 Output**

In this part the outputs resulting from the inputs are described. Here the term output means the product emerged in its original state through the use of resources.

##### **4.2.2.1 GitHub and Data-Portals**

Publishing the gathered data is the most important pillar in OGD. When the collected data is not published, it cannot generate value in any meaning. Therefore it's very important to open up this data to the public. There are several ways to do so. The easiest way is to upload the data on the web page of the organization in a spreadsheet. For the publishing organization this might be the easiest way but for the user this kind of provision is not the most convenient. In this form there is no interaction within the user and the data. To make this possible, some organizations do have GitHubs and Data-Portals. A GitHub is basically a web page, which hosts files. On those pages it's possible to give access to others so they can download the files. These sites are often used to publish data. (GitHub, 2016) SRFdata has their own GitHub page, [www.srfdata.github.io](http://www.srfdata.github.io), where they host and publish their data. SRFdata knowingly chose this approach to publish their data. They wanted the research and download of the data as convenient as possible for the users. (Interview Mr. Schmidli, 2016)

Unfortunately, SRFdata is the only organization throughout the interview-partners using this kind of approach. Even the Education Statistics of Zurich, the biggest of its kind in Switzerland, does not have a Data-Portal (Interview Mr. Schalit, 2016).

##### **4.2.2.2 Publications**

The variety of the publications differs a lot between the different organizations and authorities. Some statistic departments do only the mandatory part

due to a set of reasons, whereas other statistic departments publish a lot more. This also depends on the size of the department and the resources available to perform the work. In the following paragraphs the published data of the interviewed education statistic departments will be described and compared.

The education statistics of the Canton of Zurich is an example, which publishes more data than required by the Charta of the public statistics of Switzerland. In one hand, they publish a statistical compilation of the schools in the Canton of Zurich. In this publication the published data includes the total numbers of the students, the kindergarten, the primary school, the secondary school 1 & 2, the tertiary school, the school staff and the department of education. All this data is visualized in different forms as tables, graphs and maps. This enables any layperson to easily understand the complicated data behind these elicitations. Additionally to this statistics, the education statistics department provides data to the statistics department of the Canton of Zurich to publish a statistical Yearbook of the Canton of Zurich. These two publications are in form of a PDF or hardcopy in the case of the yearbook. But the education statistic department does also publish visualizations and tables on their homepage. There are interactive visualizations to different education related topics. The underlying data for these visualizations can be easily downloaded as an excel file. The example of the education statistic department of the Canton of Zurich is a good example when it comes to publishing data. But it has also to be taken into account that this department in Zurich is the biggest cantonal education statistic department in Switzerland. (Bildungsstatistik des Kantons Zürich, 2016)

The education statistic of the Canton of Bern does, as well as Zurich, publish more than required by the Charta. First they publish a document, which contains the basic data. It's a PDF, which includes:

- Education system of the Canton of Bern
- Information about the students listed according to the primary, secondary and tertiary school

- 
- Development of the numbers of students
  - Age of the students
  - Transfer rates
  - Cultural heterogeneity of the classes
  - Day-schools
  - Graduations
  - School staff
  - Costs of the public schools
  - Educational grant

Additionally to this document, the education statistics of the Canton of Bern publishes school and municipal tables. In these tables there are basic statistics to the number, gender, nationality and language of the students for the municipality and the school site. Besides of those publications, they also provide selected thematically analyses and tables.

The education statistics of the Canton of Solothurn is the smallest of those three examples. On the homepage of this department, there is a table with a trend-analysis published, where the numbers of students for each stage of the education, including private schools are to be seen. Furthermore, there are two PDF documents. One gives a detailed view over the allocation of those students according to the disciplines. The second published document gives a more detailed insight of the analysis of the data.

#### **4.2.2.3 Requests**

A major part of the work of open government data providing departments is to process received requests. There is the ability to request a lot of information. This is a service offered by the statistical departments. At the education statistics of the Canton of Zurich, they keep a list with all the requests obtained, so they can analyze the frequency of some requests. If there are several requests for a certain topic, they do upload the data on their website. (Interview T. Schalit, 2016) Also the other statistical departments work with these requests. At the education statistics of the Canton of Solothurn per ex-

ample, only one employee performs these tasks. Thus it's not possible to satisfy all the requesters. (Interview R. Lischer, 2016) There are also some requesters who try to misuse the ability to request data. In the interview with Jürgen Allraum, the technical manager of the education statistics of the Canton of Bern, he complained about some requestors who wanted to get a list of the schools to snow under them with advertisement.

### **4.2.3 Outcome**

The outcome describes a product, which was enabled through the use of outputs. This output is usually generated by third parties and is also responsible for the implied impacts.

#### **4.2.3.1 Eckdaten-App**

The Eckdaten-App, former Schulraumplanung-App, is an app developed by Eckhaus AG in association with Ubique, a software developer. Ubique is based in Zurich and developed this app for Eckhaus AG. The Eckdaten-App visualizes the answers to the questions about the birthrate, migration or the general housing situation. The source of the data used for the application is from the "Bundesamt für Statistik", the federal department of statistics. This department consolidates the data received from the cantonal statistic departments, which gather the data. The data received is in form of CSV. Eckhaus AG has a deal with the federal department of statistics. Principally the data is free of charge but since the data format needed by Eckhaus AG is specific, the federal department of statistics charges an editing fee. The city- and land-use planning firm could also edit the data by themselves but they outsourced this process. The app serves as a tool for Eckhaus AG to advise municipalities on the school infrastructure concerns. The Eckdaten-App is only available on the iTunes App Store and therefore it operates only on iOS devices. Below you have an extract of the user-surface of the app.



Figure 10: Eckdaten App (Eckhaus AG, 2016)

The app is structured into general and a professional area. The access to the professional area is subject to a charge of four Swiss francs per annum. In the general area, one has the possibility to analyze one of the predefined info graphics or the student growth in the next ten years. With this professional access, one can personalize the types of graphs needed and the scope needed.

In addition to the general and professional area there is a login area. Within this login area there are three different options. First, there is the municipality login, then there is the school infrastructure-planning login and last but not least there is the research login. (Interview C. Signer, 2016)

With the aid of this app, the Eckhaus AG supports the planning of the school infrastructure in different municipalities. Based on a school in Wallisellen, Zurich, a detailed example of this service provided by Eckhaus AG is shown in a subsequent chapter.

#### 4.2.3.2 Visualization

One part of the re-use of the data is the data-visualization. Data visualization is the presentation of data in a graphical form. It enables decision makers or interested parties to see the presented analytics visually. Thus they can see and understand difficult concepts and analytics. There are different categories of data visualization. The most used ones are the steady and the interactive visualizations. The steady visualization shows a graph, which cannot be influenced by the user, whereas the interactive visualization might be personalized in a given way. (SAS Institute Inc., 2016)

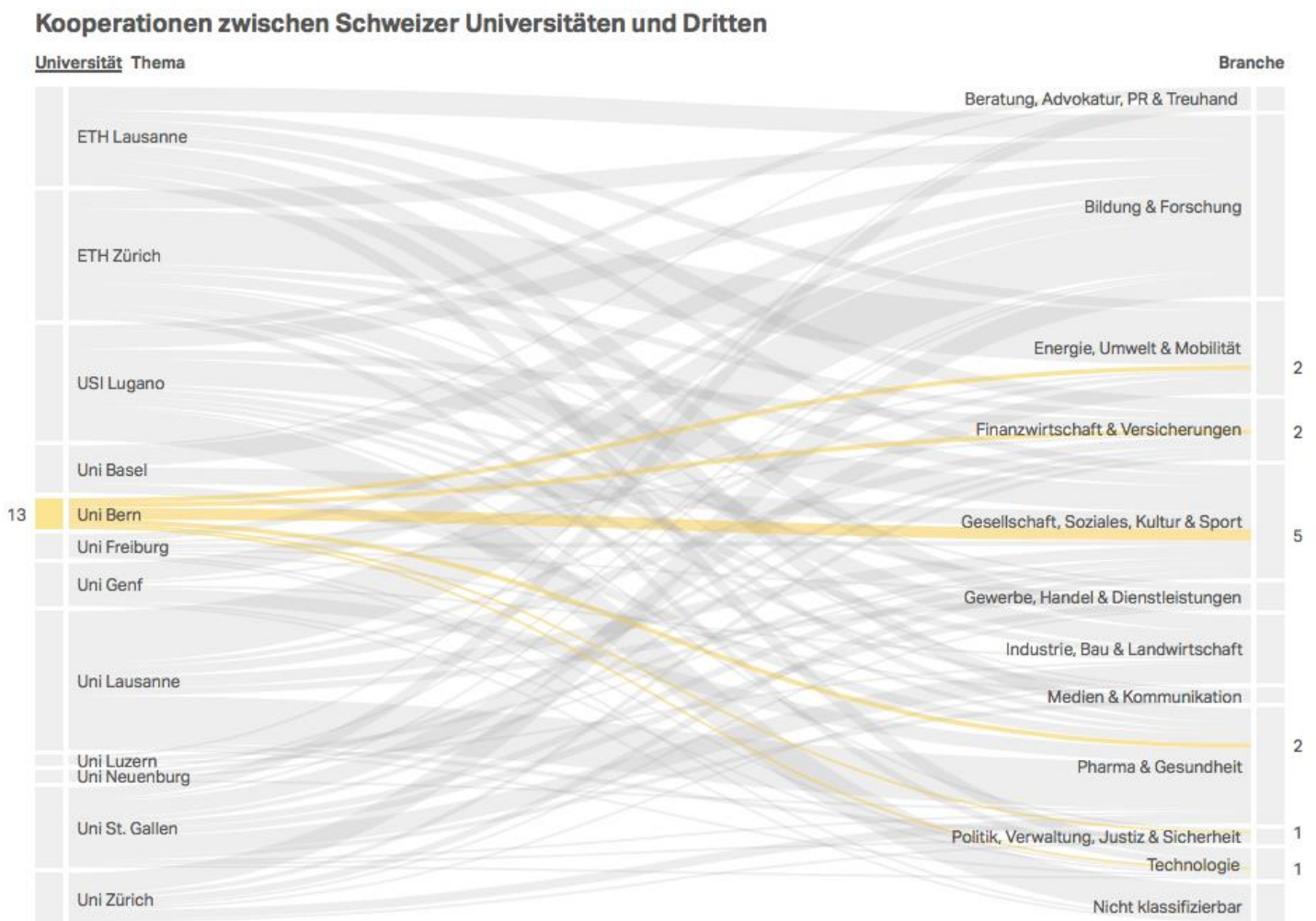


Figure 11: Cooperation between swiss universities and third parties (Thomet, 2016)

Fidel Thomet made such an interactive visualization. He is a former student of the Zurich University of Arts and worked at Open Data Zurich. The visualization, which he made, is based on data of the SRFdata portal. The data is about the professors at swiss universities. More specific, the data covers all

the additional occupations of the professors. This was a major project realized by SRFdata. Fidel Thomets visualization shows the different connections from professors to the corporate sector, thus the connection of universities to the corporate sector. This might give some indication about the intentions of the corporate sector businesses supporting universities. Above you see an extract about his visualization. The interactive part is divided into two sections. First, there is the adjustability of the view, where one can adjust the view to university or topic. The second part is the active mouse-over function, which highlights the selected university or topic. With the aid of such visualizations, laymen can better understand complex data. This can lead to a more widespread knowledge and awareness about some specific issues. (Interview Mr. Thomet, 2016)

#### **4.2.3.3 News coverage**

Newspapers, news networks, radio stations etc. are often referring to outside sources to get a data basis to apply a scientific coverage approach. Therefore, as Tobias Schalit said, they are provided with data by the statistic departments. Whether they can download it straight away from the Internet pages of the departments or they can request the data, if there is no online download possibility. Also Julian Schmidli, the data journalist of SRFdata, confirmed that these kinds of data-sources are commonly used for news coverage. He further told, that the data set they had for the additional occupations of the university professors of Switzerland involved 16 news coverage related to this topic through different media channels.

#### **4.2.4 Impact**

The goal of OGD is to create an impact. The impact can occur in different ways. In this thesis, the impact is structured into four different parts. (1) Government efficiency and effectiveness, (2) economic growth, (3) government transparency and accountability and (4) inclusion and empowerment. The different kinds of impact are described below.

#### **4.2.4.1 Infrastructure**

A good infrastructure is a necessity for appropriate education. The classrooms and the school buildings are taken for granted. But the process to get there is much more complicated and complex. The school infrastructure planning is a dynamic, ongoing process. There are dozens of calculations and trend analysis made to make sure that there is exactly as much school buildings as needed. With OGD projects, this is even more promoted. Third parties, such as urban-planning companies can analyze the published data and try to get an advisory contract with the municipalities to advise them on infrastructural planning for school infrastructure. At the end, this serves the government and creates value in form of efficiency and effectiveness.

Eckhaus AG provides these kinds of services. A Zurich based firm that does advise municipalities in urban-planning concerns. They did the infrastructural planning advisory for the municipality of Wallisellen in the Canton of Zurich. This example shows how opening up the data to the public can generate value and help the government to be more efficient. There are different organizations and departments involved in this case. First of all, there is the school of Wallisellen, which is the origin of the gathered data. The board of the school does collect the data of the students. Then they do report these figures to the Education Statistics Department of the Canton of Zurich. This department again does aggregate the data from the municipality of Wallisellen and all the other municipalities of the Canton of Zurich. Then, these data is published on the Internet page of the department. The department on the other hand, does report the aggregated figures to the Federal Statistics Department (Bundesamt für Statistik). They do the same as the Education Statistics does, just in a national wide scope. This means, the data is aggregated on a national level. Then, these data is again published at the homepage of the Federal Statistics Department. This is where the third party, Eckhaus AG, comes into play. Eckhaus AG receives the data from the Federal Statistics Department, adds other data and runs an analysis with the aid of their app, the Eckdaten App. This app was developed to analyze the evolution of the demographic figures. Through the analysis of this app, Eckhaus AG can make forecasts for student trends and other key figures. With this expertise they earn from this analysis, they do advise municipalities. This was also the



case with the municipality of Wallisellen in the Canton of Zurich. They needed a private firm to advise them on infrastructural planning in for the education infrastructure. Eckhaus AG did work hand in hand with the local authorities to achieve the best result possible. Due to the advisory serviced of Eckhaus AG, the municipality of Wallisellen has appropriate school infrastructure and thus does safe financial resources. The worst-case scenario could be that they do not have appropriate school infrastructure and thus would be forced to install temporary trailers, which would serve as classrooms etc. This again would generate unnecessary costs for the municipality. At the end of the day, the schools benefits from the OGD. In this case there is a perfect circle as shown below. The data-origin is the school in Wallisellen, which gives the Education Statistics of the Canton of Zurich the necessary basis for their inputs. The Education Statistics on the other hand does provide the statistics, which then are published. Those published statistics serve as a foundation for the analysis of Eckhaus AG, which then do provide an advisory service to the School in Wallisellen. And that's where the benefit is realized. This shows the beneficial character of the OGD process. (Interview Mrs.

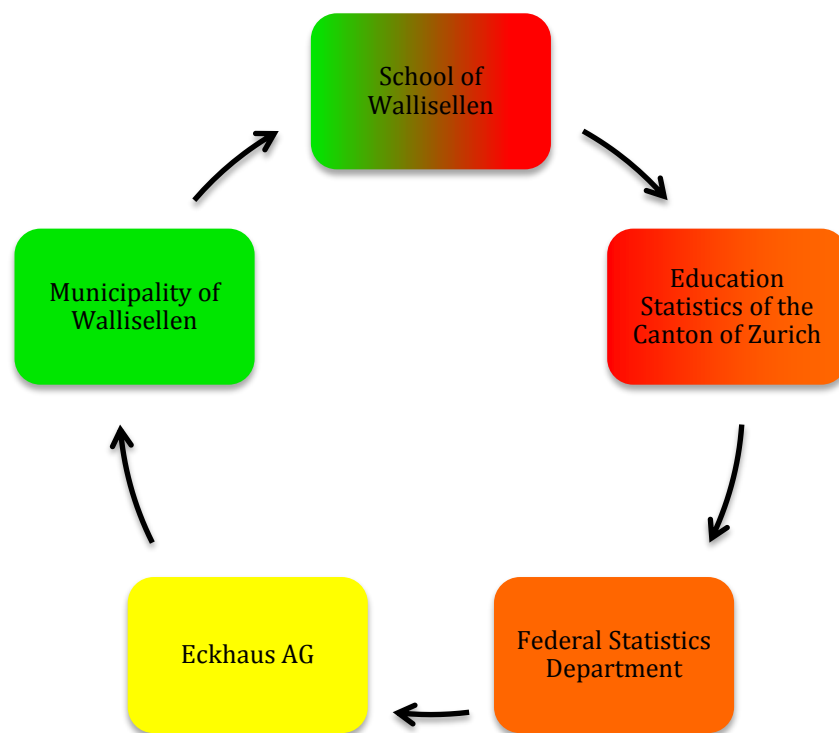


Figure 12: Circlly of OGD Impact

Signer, 2016; Interview Mr. Kipfer 2016) In this case, the circle could be made leaner. The detour via the Federal Statistics Department is not a necessary step but this is more convenient for Eckhaus AG. In doing so, they can get national wide data at a blow. However, there is still potential to improve this circle.

#### **4.2.4.2 Quality**

Opening up the collected data in an Open Government Data project may lead to a better education quality (Interview Mr. Schmidli 2016; Interview Mr. Thomet 2016). The schools among a Canton or even within the whole nation can be compared and rankings can be established. This could lead to a kind of competition between the different schools. Schools would even try harder to accomplish better results and offer better education. The increase in education quality would result in better-educated students. After the graduation, these students could make a difference in the swiss economy. Thus, better education could lead to economical growth.

#### **4.2.4.3 Awareness**

Open Government Data is in its early stages (Dapp and Stuermer, 2015, p. 6). Since a lot of people are not aware of all this data being published, the OGD movement could increase the awareness of openly available data, which in turn could lead to inclusion and empowerment. The citizens could gain a better understanding for the government and their decisions since they can access the data behind these decisions and therefore could engage in the political decision making process. The gathered data from SRFdata concerning the additional occupations of the university professors in Switzerland had a series of aftermaths. After all these published coverage's and visualizations, there were changes in the regulations of additional occupations of university professors made (Interview Mr. Schmidli 2016).

#### **4.2.4.4 Potential in the future**

Opening up government data could have a strong signaling effect. To open up the collected data would show the citizens that the government does not

want to hide anything. This could have a significant positive impact on the trust of the citizens towards the government. This increase in government transparency could have a positive effect on the accountability of the government. Mr. Schalit says that the education career lasts for almost 20 years. Therefore the utilization-interval is very limited. But nevertheless, he could imagine to have different apps based on OGD in the field of education. Per example about the commuting area of high school students (Interview Mr. Schalit, 2016) or an app to locate the next school or kindergarten (Interview Mr. Allraum, 2016). The data could also be matched with the social security number. This would allow re-users to make new kinds of analyses. But this is a major issue in the eyes of privacy protectors. (Interview Mr. Lischer, 2016)

#### **4.2.4.5 Protection of Privacy**

Throughout the interviews there was the topic protection of privacy, which was always discussed about. Since the data collected is about individuals, the debate about the protection of privacy is crucial. The Charta of the public statistics of Switzerland defines clear rules, which data can be published and which data has to be protected by the protection of privacy law. It's not allowed to publish any data that allows any conclusions about particular individuals (Charta öffentliche Statistik der Schweiz, 2012). Also the interview-partners confirmed, that the handling of sensitive data is been taken care of very seriously (Interview Mr. Schalit, 2016; Mr. Kipfer, 2016; Mr. Lischer, 2016; Mr. Allraum, 2016; Mr. Thomet, 2016). But another point is that the protection of privacy slows down the OGD movement. The statistical departments cannot operate as freely as they wish (Interview Mr. Allraum, 2016; Mr. Lischer, 2016; Mr. Schmidli, 2016). To improve the OGD, also the laws and the protection of privacy have to be adjusted (Interview Mr. Schmidli, 2016).

### 4.3 The Ecosystem of Open Government Data value creation

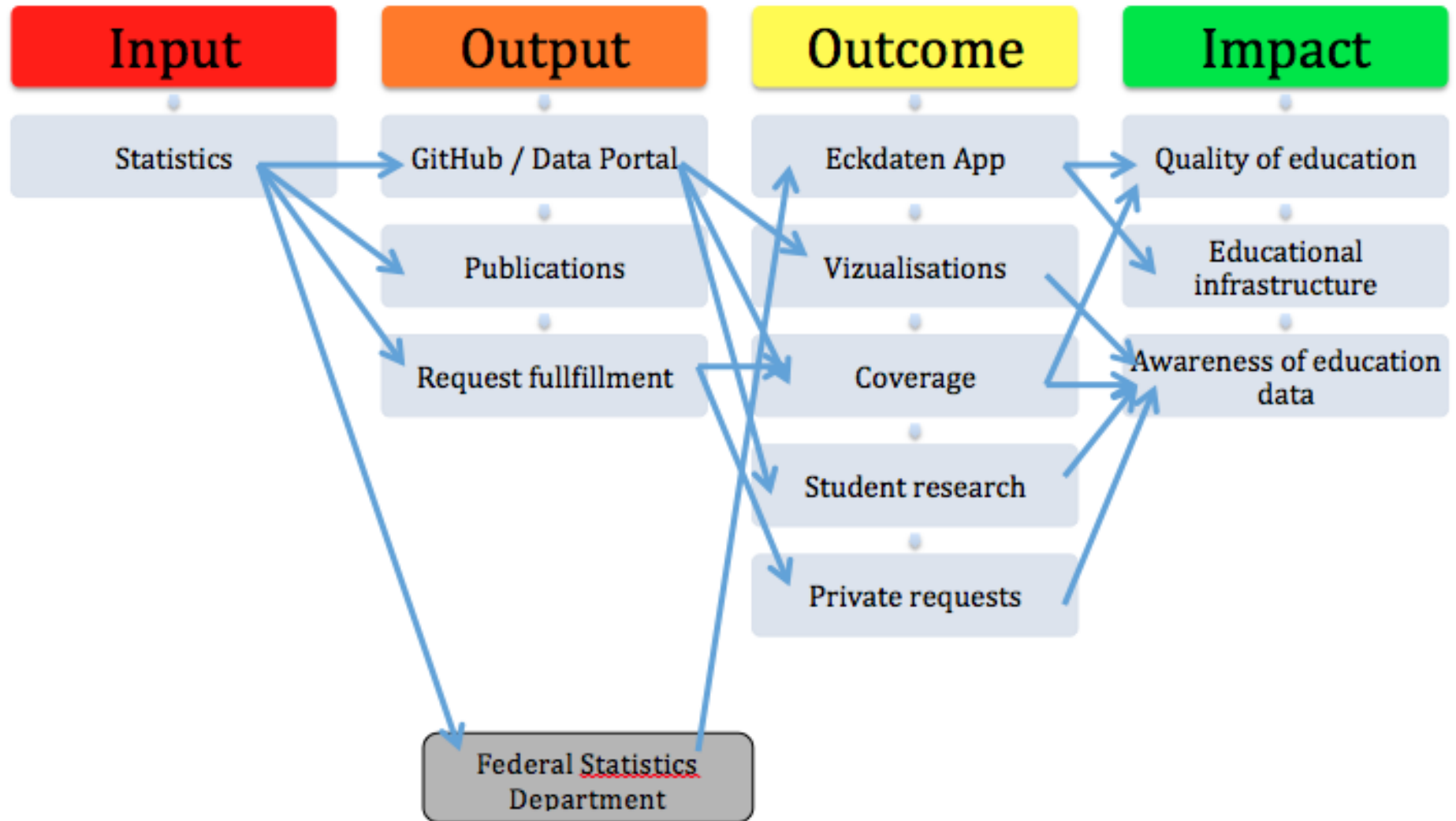


Figure 13: Ecosystem of OGD value creation

---

The Ecosystem of Open Government Data value creation shown above, gives a holistic view of the interdependencies of different stages and organizations. It's a visualization of the dataflow through the different stages input, output, outcome and impact. The blue arrows symbolize the dataflow, whereas the grey fields describe the underlying objectives. For example the input statistics go to the Federal Statistics Department and are also published on the Education Statistics Departments homepages etc. After, the data proceeds to the Eckhaus AG, which then implement this into their Eckdaten App. This app then helps Eckhaus AG to advise municipalities in educational infrastructure issues. At this stage, the Open Government Data, hence the statistics from the stage input, generate value in form of an impact in the education quality as well as the educational infrastructure.

## **5 Conclusions**

### **5.1 Findings and Implications**

The initial goal of this thesis was to evaluate a quantitative, financial value for the OGD initiatives. As described in chapter four, there was not enough quantitative information available as desired to fulfill the valuation as dictated by the theory from Nicholls et al. (2009). Instead of a quantitative valuation there was a detailed impact map established to picture the Ecosystem of Open Government Data value creation. Owing to this Ecosystem, it is possible to see and understand the relations between the actions undertaken by the government and the impacts arising for the citizens. On the basis of this impact map, the positive impact of OGD is revealed. The ecosystem of value creation shows how the different organizations are effected in which way by the OGD initiative and simultaneously prove the positive impact in economic growth, efficiency & effectiveness, inclusion & empowerment and transparency & accountability. Since the data is gathered, no matter if published or not, the step to publish these data does not pose a massive additional expenditure. Therefore the impact is considered as big, in light of the fact that the expenditures are relatively small.

### **5.2 Limitations and Suggestions for Further Research**

As explained earlier in the paper, as a result of insufficient evidence for a quantitative valuation, the approach of the originally intended SROI could not be fully implemented. Due to the fact that the organizations do not really know or at least track the costs and the impacts of these OGD projects, it is nearly impossible to give those projects a quantitative value. Further this thesis did consider only the OGD projects in the field of Education and only on a national level. The organizations interviewed were direct stakeholders of these data. There were no data publishing organizations interviewed or considered, which were not the first-hand publisher. Therefore the future research could address the other data publisher as private organizations or data re-publisher. The same analysis should also be made for the other 13 high-value categories defined by the G8 Charta. After analyzing these re-

maining categories, a comparison among these 14 high-value categories OGD could take place. Another suggestion for future research might be the analysis of these kinds of data on an international basis to take the comparison even a level higher.

## Appendix

### Transcribed Interviews

#### Interview with Mr. Schmidli, SRFdata (20.09.2016)

Wer sind sie, und was machen sie genau?

**Wir sind ein datenjournalismus-Team, vom SRF. Dass heisst wir arbeiten tri-medial mit allen Informatonsredaktionen zusammen. Diese sind die grossen Radiosendungen, TV-Sendungen und Tageschau 10vor10 usw. Einfach alles was irgendwie Informationen verarbeitet. Was wir machen sind verschiedene Sachen. Wir sind einerseits Kompetenzzentrum für alle Fragen die Daten betreffen, das heisst alles was Statistiken betrifft oder Auswertungen betrifft oder auch Redaktionen selber machen oder Redaktionen machen lassen oder bekommen von Unis von Datenauswertungsinstituten, dass wir quasi wie zur Verfügung stehen und Ihnen helfen und Tipps geben wie man was interpretieren kann, was man nicht machen kann, was gefährlich ist etc. Das zweite ist ein kleiner Teil von unserer Arbeit. Der grosse Teil unserer Arbeit ist vor allem eigenständige Recherchen zu machen, Auswertungen zu machen, Berichterstattung zu machen über verschiedene Kanäle. Es kann alles sein. Thematisch sind wir da offen. Es ist immer ein Informationsbezug. Man versucht immer Aktualität aufzunehmen oder neue Aktualitäten zu schaffen → also Agenda Setting. Unsere Themen sind sehr breit. Wir machen irgendwie über Sport wie sinnvoll ist Sportförderung der Behörden am Bsp. der Olympiade. Es kann sein dass wir auswerten wie gut sind eigentlich Jobsituation von neuen Müttern im Schweizer Job-Markt.**

Machen die Datenerhebung zu solchen Themen selber?



**Der grösste Teil ist, dass wir selber neue Daten erfassen und publizieren. Wir haben auch eine eigene GitHub Page, [www.srfdata.github.io](http://www.srfdata.github.io), wo wir unsere eigenen Datensätze die wir haben oder die wir entwickeln veröffentlichen inkl. die Methodologie also ein bisschen mit einem wissenschaftlichen Approach. Zum teil schaffen wir neue Datensätze welche wir wollen das es sie gibt. Indem dass wir diese schaffen, versuchen wir eigentlich anzuregen, dass die Behörden das übernehmen. Ein Bsp. ist dass wir eine grössere Recherche gemacht haben zum Thema Sponsoring von Lehrstühlen an Universitäten und Nebenbeschäftigungen von Professoren. Wir haben da eigentlich eine Datenbank gebaut. Wir haben all die Daten aus den Universitäten rausgeklagt, aggregiert und dann berichtet. Es wurden 16 verschiedene Sendungen über dieses Thema gemacht. Dies hatte auch einen wesentlichen Impact. Die Unis mussten ihr Reglement ändern, es gibt bald eine schweizweite Regulierung was sie dürfen und was nicht. Und es fällt eigentlich alles auf das zurück. Es ist auch angekündigt worden, dass swiss universities, dass sie in Zukunft diese Datenbank selber erheben und selber publizieren möchten. Da ist quasi ein Progress Sichtbar. Das ziel von solchen Sachen ist, dass wir selber versuchen Datensätze zu schaffen um dann irgendwie versuchen anzuregen, dass dies eigentlich die Aufgabe der Behörden ist dies zu machen und der Öffentlichkeit zur Verfügung zu stellen. Das ist das eine. Das andere ist, Daten die es schon gibt zu nehmen und diese zu enhance oder zu kombinieren bisschen mash-up mässig. Zum Teil auch Daten dazu verwendet. Wir haben auch öffentliche Daten welche auch schon zugänglich gewesen sind von den Universitäten und haben dann quasi gemasht mit den Daten die wir selber gehabt haben und haben versucht zu schauen ob es einen zusammenhang gibt, ob es eine Struktur gibt.**

Dann kommt es auch vor dass sie Daten nehmen von der Behörden und anhand von denen eine Auswertung oder Berichterstattung macht jetzt speziell auf Bildungsdaten bezogen?

**Es gibt's schon auch, dass wir zum Teil kleinere Analysen machen und dann halt auf Daten Bezug nehmen welche bereits offen sind und dort sind unsere Quellen das BfS, Open Data ZH finden wir sehr gut von ihrem Approach her und dann auch punktuell Bundesämter welche je nach dem gute Daten veröffentlichen.**

Für mich ist es sehr interessant da ich selber nicht wusste, dass Ihr selber so viel Daten generiert, dass dies ein so grosser Teil Ihrer Arbeit ist, dass Ihr halt eher mehr Daten welche schon vorhanden sind auswertet und dann aufgrund dessen versucht eine Aussage zu machen. Die Daten welche sie bekommen, sprich nicht selber generieren, in welcher Form bekommt ihr diese? In welcher Formatierung? Sind das ganz klassische Open Data oder sind das auch Daten die spezifisch für Sie erhoben werden?

**Was wir sehr viel machen ist, dass wir Sachen rausklagen mit BGÖ. Dann geht's relativ lange, bis man die Daten bekommt. Meistens geschieht das, weil sie uns die Daten nicht freiwillig geben möchten. Das heisst, sie geben uns Daten in einer Form in welcher sie finden, dass es für uns möglichst schwierig ist diese zu verarbeiten. Meistens PDF's oder Fotos oder ausgedruckte Dokumente welche per Post geliefert werden welche dann mühsam formatiert werden müssen. Deshalb ist Datencleaning und Strukturierung ein wichtiger teil bei welchem wir grosse fortschritte machen. Wir arbeiten jetzt zunehmen mit Scrips und versuchen unsere Daten zu verbessern und so eine Art Strukturierung versuchen zu schaffen.**

Die Daten, die sie selber erheben, die werden ja auch veröffentlicht auf ihrer Datenbank. Sind diese Daten für alle öffentlich zugänglich in einer maschinenlesbaren form?

**Was wir oft machen ist, dass wir diese Daten auf GitHub veröffentlichen. Diese sind dann meistens lesbar also als SQL Datenbank je nach dem. Meistens als CSV oder als Jason. Aber wir machen das nicht mit allen Daten, weil dies ein zusätzlicher aufwand für uns ist diese Daten**

**sauber aufzubereiten und zu publizieren. Manchmal haben wir Zeit dafür und manchmal nicht.**

Was ist entscheidend ob man die Daten veröffentlicht oder nicht veröffentlicht? Hat das damit zu tun was es für Daten sind?

**Ja, es kommt drauf an welchen wert die Daten haben und welchen nutzen andere Leute aus diesen Daten schaffen. Deshalb sind wir da auch data-owner, dass andere Leute in der Open Data Szene diese Daten nehmen und mit denen etwas anderes machen. Dies ist auch in unserem Interesse vom Prinzip Service Public her das wir sagen das wir nicht nur Datenjournalismus machen sondern das wir auch versuchen neue Daten zu erheben und diese der Öffentlichkeit zur Verfügung stellen und durch das auch versuchen den Impact zu maximieren welchen wir versuchen zu erreichen.**

Veröffentlichen sie spezifisch Bildungsdaten und wenn ja, was fliessen da für Ressourcen rein um diese Daten zu veröffentlichen?

**Ich glaube wir sind ein kleines Team, ca. 2.5 FTE, und immer noch paar Praktikanten die mithelfen. Wir machen nicht viele Projekte im Jahr. Dies begrenzt sich auch 15-20 Veröffentlichungen. In Sachen Bildungsdaten ist das mit den Universitäten sicher das grösste das wir gemacht haben. Da flossen auch viele Ressourcen rein. Die Idee war halt publik zu machen, was für Verbindungen von Universitäten zu dritten, vor allem mit der Privatwirtschaft, haben. Der Professoren, ihre Nebenjobs. Wer zahlt eigentlich? In welchen Verwaltungsräten sind sie? Wer zahlt eigentlich unsere Forschung und Lehren an den Universitäten? Wer finanziert das mit über Drittmittel? Und da war halt unser Antrieb das dies nicht öffentlich gemacht wurde ausser sehr begrenzt von einigen Universitäten. Und unser Antrieb war, dass es wichtig ist, das die Öffentlichkeit weiss wer Forschung und Lehre finanziert, um gewisse Bias zu veröffentlichen, um gewisse Interessenkonflikte öffentlich zu machen. Da haben sich gewisse Universitäten sehr gewehrt. Sie wollten**

uns die Daten nicht geben und argumentiert das dies zu kompliziert sei und gar nicht richtig erfasst werde etc. Wir haben da immer wieder nachgehakt und das Öffentlichkeitsgesetzes kantonal versucht anzuwenden und sind sehr weit gegangen zum die Daten bekommen, zu bearbeiten und zu publizieren. Wir haben angefangen indem wir die Daten überall zu einem gewissen Zeitpunkt eingesammelt haben, danach die Daten welche in verschiedenen Formaten gekommen sind, zu standardisieren. Als zweites haben wir sehr viel dafür investiert, dass wir diese Daten kann vergleichbar machen. Die Daten in eine Struktur und in eine Datenbank zu bringen, dass man diese angemessen festhalten kann. Das Problem war das die kantonalen Daten andere Gesetzesgrundlagen haben und deswegen auch andere Erhebungsmethoden und andere Strukturen von den Systemen an der Universitäten und man hat auch andere Sprachen. All das führt dann dazu, dass man wahnsinnig unterschiedliche Werte hat. Dies ist ein riesen Problem für uns. Andererseits diese Unstrukturiertheit, dass wir eine Art Deduplikation machen müssen z.B. das die Firma X gleichzeitig die Firma Y ist etc. Was wir auch noch gemacht haben ist das wir ein Register von Professoren erstellt haben. Wir haben eine Homepage gescrapet. Die wird von jemand anderem betrieben und heisst [www.prof.ch](http://www.prof.ch). Da kann man alle Professoren nachschauen. Man musste halt alle diese Daten scrapen, weil sie uns die Daten nicht geben wollten. Man musste dann halt einen Roboter schreiben der alles runtergeladen hat und immer abglich ob der Professor der in den Daten war ob es ihn überhaupt gibt, ob er in der gleichen Institution tätig ist usw. Danach konnten wir die Email Adressen dieser Professoren haben. Als wir alle die Daten gesammelt hatten über die Professoren, haben wir allen eine Email geschickt mit den gesammelten Informationen über die einzelnen Professoren. Wir haben gesagt wir veröffentlichen in vier Wochen eine Datenbank mit diesen Informationen über sie. Bitte schreiben sie uns zurück falls sie Fehler finden, falls sie etwas dagegen haben usw. Wir haben alle veröffentlicht. Niemand konnte sagen ich will nicht veröffentlicht werden. Dies war ja auch das ziel. Man konnte sich nicht von dem ausschließen lassen. Man konnte nur die Angaben korrigieren jedoch nicht rausstreichen. Sie haben zum

**teil versucht gewisse Tätigkeiten als unwichtig rausstreichen zu lassen jedoch haben wir an der Vollständigkeit festgehalten und alles veröffentlicht. Wir haben nicht bewertet. Jeder konnte selber urteilen ob dies ein Interessenkonflikt sei oder nicht. Den ganzen Prozess haben wir gleich mitveröffentlicht.**

Das Ziel dieser Veröffentlichung war nur das sie dieses Register veröffentlichen?

**Genau. Wir haben da eine interaktive Datenbank gewählt wo man alles selber suchen kann. Was wir aber gemacht haben sind Analysen aufgrund dieses Registers. Über diese Analysen wurde dann auch berichtet. Dies waren Themen die wir spannend gefunden haben oder über besondere Fälle wo wir gefunden haben das sind wirklich Interessenskonflikte. Über dies wurde schon berichtet. Dies war aber unabhängig von der Struktur der Daten.**

Wäre es möglich den Aufwand für diese Veröffentlichung zu beziffern?

**Wir geben eigentlich keine Zahlen raus. Die werden auch nicht speziell erhoben. Was man aber sagen kann ist das der grösste Aufwand für die Strukturierung der Daten reingeflossen ist. Wir haben da elf Monate gebraucht von der ersten Mail bis zur Publikation. Dazwischen ist viel nichts passiert. Wir mussten viel warten auf fristen von Behörden oder aber Antworten von Mail etc. Es gab aber auch Momente in denen wir zu dritt oder zu viert zwei Monate intensiv daran gearbeitet haben. Für die Publikation waren dann auch viel mehr Leute beteiligt. Aber insgesamt ist es unmöglich zu sagen wie viele stunden da reingeflossen sind. Aber es waren im Vergleich zu anderen journalistischen Projekten doch sehr viel.**

Wo kann ich all diese Publikationen zu dem Uni-Projekt finden? Kannst du mir einen Link aufs Mail schicken?

---

**Das alles findest du auf [www.srf.ch/uni](http://www.srf.ch/uni). Da ist alles verfügbar.**

Was denkst du persönlich, was habt ihr jetzt deiner Meinung nach damit für einen Mehrwert generiert? Hatte es aus deiner Sicht einen sehr grossen Impact?

**Wir haben damit auf mehreren Ebenen einen Impact generiert. Erstens ist es wichtig dass wir im Moment in einer Zeit sind wo Wissen sehr wichtig ist. Universitäten sind das Gehirn unserer Gesellschaft. Sie müssen unabhängig lehren und forschen können. Was sollen diese Drittmittel ermöglichen? Was ist der Sinn? Was sind die Gefahren von möglichem Einfluss von Firmen, von anderen Interessen? Das man darüber diskutiert und sich eine Meinung bildet und überlegt. Das man sich auch bewusst ist was man riskiert. Es war auch wichtig einen Impact auf der politischen Ebene zu haben. Dass da eine Diskussion stattfindet über die Regulierung und Sinn und Unsinn von solchen Sponsoring und Einflussnahmen. Am Ende geht es darum die Lehre und die Forschung zu schützen vor dem Vorwurf der Korruption. Unsere Studien sind wertvoll, wenn sie unabhängig also vom Staat finanziert sind und keinen Bias haben. Dies ist wichtig. Will wollen nicht in eine Richtung gehen wo unsere Forschung angegriffen werden kann, weil andere Leute diese finanzieren und dadurch unsere Forschung an Wert verliert. Wir finden, auch wenn es für die Universitäten im ersten Moment kritisch war oder aber ein PR-Problem war, helfen wir den Universitäten ihre Unabhängigkeit zu schützen.**

Die Daten welche sie in ihrem Tool haben, führen sie Buch darüber welche Datensätze wie viel Klicks haben?

**Ja, eigentlich schon aber im Moment ist es so, dass sind wir daran ein neues System zu integrieren. Es heisst Com-score, ist von Adobe und ist wie Google Analytics. Da kann man jede Interaktion vom User messen. Aber bis jetzt ist es so gewesen, dass es gescheitert ist an mangelnden Tools und jetzt sind wir in einer Testphase. Jetzt versuchen wir**

**den Impact besser zu messen auf unterschiedlichen Ebenen. Wie wird es geshared? Wie viel User schauen was an? Wen interessiert es eigentlich? Ist es beliebt? Lohnt sich das? Dies sind alles Sachen wo wir erst am Anfang stehen.**

Das heisst bis jetzt haben Sie noch gar nicht registriert wie viele Klicks Sie erzielt haben?

**Doch wir haben schon Anhaltspunkte. Die Erfahrung ist schon so, dass wenn wir Grafiken von Mehrwert bieten, auch mit solchen Datensätzen, dann ist es schon so, dass es überdurchschnittlich oft verwendet wird. Bei spezifischen Datensätzen ist es aber eher eine Minderheit aus der Open Data Community. Das sind ein paar Superuser die das dann viel machen. Das ist aber auch für uns wichtig um auch kommunizieren zu können, dass wir eine Transparenz schaffen und unsere Arbeit reproduzierbar ist wie in der Wissenschaft. Das Ziel ist auch eine höhere Glaubwürdigkeit zu erreichen bei den Fachpersonen, welche genau hinschauen ob man Fehler gemacht hat. Das hilft uns auch.**

Wäre es da möglich dass ich konkrete klickzahlen erhalten könnte?

**Klickzahlen geben wir eigentlich auch nicht raus. Es ist noch schwierig zu sagen.**

Oder konkrete Beispiele was diese Superuser mit Bildungsdaten gemacht haben?

**Es gibt halt andere Visualisierungen welche andere Leute aufgrund unserer Daten gemacht haben.**

Genau. In diesem BSP sind sie ja diejenigen die den Input und Output machen und die anderen die es wiederverwenden generieren den Outcome.

**Ja da kann man sicher etwas schicken.**

Das wäre super wenn sie das machen könnten. Haben sie mit den Daten der Unis Feedback bekommen was für wellen das geschlagen hat?

**Ja, massiv. Es haben sich viele Professoren und Studenten bei uns gemeldet. Überwiegend positiv. Sie finden dies wichtig. Es hilft auch ihrer Arbeit, wenn sie eine kritische Auseinandersetzung über dieses Thema erhalten. Wir haben auch einen Preis gewonnen mit dieser Arbeit. Primedia Preis. Dieser ist einer von den wichtigsten Wissenschaftspreisen in der Schweiz und es ist nicht selbstverständlich dass so eine kritische Auseinandersetzung mit den Unis auch honoriert wird. Wir haben erwartet, dass es nicht honoriert wird, weil wir vielen Menschen auf die Füße gestanden sind. Was explizit erwähnt wurde ist, dass es eine sehr gründliche und aufwendige Recherche war und differenziert darüber berichtet haben. Wir haben nicht nur draufgeschlagen sondern sehr differenziert verschiedene Leute zu Wort haben kommen lassen die finden, dass solche Drittmittel gezahlt werden. Somit war eine konstruktive Diskussion das Ziel von uns und nicht verhärtete Fronten haben von Leuten die sauer sind aufeinander und keinen Austausch haben.**

Mit diesen Daten welche Sie veröffentlicht haben wissen sie ja paar Sachen die gemacht werden. Was haben sie für Vermutungen wo es in 20 Jahren, weil heute ja mit Bildungsdaten nicht viel gemacht wird, hingehen könnte?

**Ich glaube das Hauptproblem ist das Persönlichkeitsrecht welches in der Schweiz sehr hoch gehalten ist. Solange sich das nicht ändert, dann ist es schon fraglich wie sich das alles entwickeln wird. Sehr viele Entschuldigungen werden auf dieses Persönlichkeitsrecht abgebschoben. Interessant wäre auch das Stipendienwesen. Aus welchen Schichten kommen diese Leute? Werden die richtigen Leute gefördert und funktioniert unser Stipendienwesen?**

**Bei solchen Daten heisst es dann immer nein das bekommt ihr nicht, das basiert auf Persönlichkeitsschutz. Wir wollen nicht veröffentlichen welche Person wie viel Geld vom Staat bekommt. Wobei ich finde dass**



**das Interesse der Öffentlichkeit ist viel grösser. Dann könnte man genau schauen ob es überhaupt die richtigen Personen sind die gefördert sind. Wer wird nicht gefördert. Und ob es da racial bias gibt. Es hat halt solche Sachen die man unbedingt untersuchen muss.**

Für mich ist es ja gerade auch interessant ob die demographischen Unterschiede in Schulen gewisse Auswirkungen haben.

**Genau aber das Problem ist halt, dass die Schulen zum Teil da „draufsitzen“. Man ist da gerade in diesem Bereich wahnsinnig ängstlich, dass man gemessen wird. Man hat schlechte Erfahrungen gemacht mit PISA, dass man da in ein Ranking gesteckt wird, welches die Komplexität von den Anforderungen eines Schulsystems nicht gerecht wird. Gleichzeitig muss man aber auch dies der Wissenschaft und den Medien zutrauen, dass man das differenzieren kann, weiterverarbeiten und berichten kann. Es ist ja auch gerade die Messbarkeit anhand der wir feststellen kann ob unser Schulsystem funktioniert. Solange es keine öffentlichen Zahlen gibt, kann man eher nur über Anekdoten berichten was dem Thema auch nicht gerecht wird.**

Du hast gerade was über die Messbarkeit gesagt. Ich selber habe das Gefühl, dass wenn man etwas messen und vergleichen kann, z.B. die Schulen der Gemeinde X mit der Gemeinde Y. Kann das auch einen negativen Einfluss auf die Schule haben? Aus meiner Sicht könnte die Messbarkeit einen Anreiz darstellen den Unterricht zu verbessern.

**Auf jeden Fall. Es kann ein Anreiz sein aber es kann auch ein Schuss nach hinten sein. Es könnte sein, dass Gelder gestrichen werden von der Politik, weil sie finden das sind schlechte Schulen oder aber auch umgekehrt. Das ist ein bisschen das Problem. Unsere Gesellschaft geht immer sehr defensiv um mit solchen Auswertungen. Man kriecht dann eher zurück als das man erst recht gerade da investiert und weiterrecherchiert. Man sagt so Evidence-Based, das ist dass man seine Entscheidungen auf Beweise basiert, was Daten genau erlauben würden.**

---

**Dies macht man nicht und man basiert eher auf Ideen und Glauben und entscheidet intuitiv.**

In den USA geht man mit Daten viel offener um. Dies ist meiner Sicht nach auch erstrebenswert in der Schweiz und in diese Richtung geht. Diese Behördendaten werden gesammelt und es braucht ja nicht viel mehr Ressourcen, wenn man diese dann auch noch dazu veröffentlicht. Dies würde dann Privatpersonen Analysen erlauben z.B. ein Elternpaar welches ihr Kind einschulen kann analysieren wo sie das Kind einschulen wollen aufgrund der Durchfallquoten der Nachbargemeinden etc.

**Genau das ist das Problem. Man könnte dadurch auch die Segregation fördern. Die Personen, welche die Möglichkeiten haben könnten dann einfach wegziehen.**

Das ist klar. Da spielen Steuern, Mieten und andere Faktoren eine Rolle. Aber schlussendlich wäre es doch interessant wie da die Verteilung ist.

**Ich finde, die Vorteile der Veröffentlichung der Daten überwiegt die Nachteile bei Weitem. Meiner Meinung nach muss man in diesem Bereich unbedingt vorwärts kommen. Es gibt viel zu wenig Anreize um die Daten zu veröffentlichen. Da sollte man viel stärker eingreifen.**

Da Sie tagtäglich mit Open Data zu tun haben, können Sie Open Data kritisieren, das man so viele Daten öffentlich zugänglich macht? Könnte man diese missbrauchen?

**Nein. Ich finde es wichtig und völlig im Zeitgeist, dass man diese Daten veröffentlicht. Ich finde es eine gute Sache. Ich finde es auch gut, dass sie nicht aggressiv sind sondern eher motivierend. Es wäre auch die Rolle der Medien dies zu stärken bzw. rauszuklagen. Es gibt zum Beispiel ein Teil der Open Data Bewegung, der Verein öffentlichkeitsgesetz.ch. Die sind aggressiver und versuchen auch wirklich Druck zu machen auf Behörden und Verwaltung mit klagen und ziehen diese bis**

---

**vor das Bundesverwaltungsgericht. Das finde ich einen enorm wichtigen teil. Ich glaube für eine kritische Öffentlichkeit dieses Thema immer wichtiger wird.**

Du hast vorhin selber gesagt, dass falls ihr die Daten rausklagt nicht wirklich kooperieren wollen. Wie könnte man diese Organisationen dazu bringen mitzumachen, zu motivieren?

**Da ist der Change von Innen das grosse Thema. Da gibt es auch Bundesämter die vorbildlich vorpreschen, die selber merken, dass es für sie auch viel besser ist. Es gibt neue Governance Strukturen und das ist halt der punkt. Man Sieht immer nur das negative. Die kritische Berichterstattung über jemanden darf man nicht als Todesstoß sehen. Dies wäre ein mega Rückschritt**

-Ende-

## **Interview with Mrs. Signer Eckhaus AG (20.09.2016)**

Vielleicht willst du kurz erzählen wer du bist und was ihr macht.

**Ja ich hole nur noch kurz das Ipad. Dann kann ich dir gerade die App die wir entwickelt haben noch zeigen.**

Ich habe diese App auch schon mit Herrn Schalit angeschaut jedoch hatten wir nicht die Vollversion deshalb konnten wir nur beschränkt anschauen.

**Ich bin im Büro Eckhaus seit dem Anfang dabei, seit es entstanden ist. Mein Hauptbereich ist die Schulraumplanung. Das ist auch der Hauptbereich in dem wir Open Data brauchen. In gewissen anderen Projekten brauchen wir auch z.B. Geo-Daten aber Bildungsdaten sind meistens eher so in der Schulraumplanung. Wie ich schon vorhin gesagt habe, haben wir eine App entwickelt. Die Idee ist so entstanden, dass wir erstens die Daten von verschiedenen Gemeinden gesammelt haben. Wir hatten diverse Ordner voll mit Excel-Files. Bei jeder Anfrage von einer Gemeinde musste man diese Graphen anhand dieser Daten neu erstellen. Jetzt in dieser App haben wir schweizweit eine Auswahl von allen Gemeinden mit ersten Auswertungen welche alle auf Open Data basieren. Die Daten sind alle vom BfS, von den Kantonen oder den BiSta. (Zeigt die App). Da wir in Zürich ansässig sind, und auch viele Aufträge in ZH haben, haben wir von ZH mehr Daten als von anderen Kantonen. Aber die Daten vom BfS haben wir aber schweizweit. Dies sind dann einerseits die Graphen die wir aufbereiten, wo wir jetzt neu auch Auswertungen für Regionen oder Kantone machen können, welche auf Gemeindedaten basieren. Hier haben wir noch eine Rubrik wo wir die Schweizerkarte haben wo man zu verschiedenen Themen die Gemeinden eingefärbt hat. Wir haben auch eigene Berechnungen gemacht → z.B. die Schülerentwicklung. Da haben wir verschiedene Daten kombiniert und das Verglichen und somit eine erste Einschätzung gemacht. Oder auch eine ganz banale absolute Geburtenentwicklung. Das ist al-**

les in der Gratisversion verfügbar. Das brauchen wir für unsere Berichte für die Beobachtung einer Gemeinde wenn wir einen Auftrag haben. Man kann aber auch als App User diese Vektorbasiert als PDF schicken und dann wieder selber in eine Präsentation einbauen oder als z.B. als Gemeinderat schauen wo liegt meine Gemeinde, brauchen wir Eckhaus AG für eine Schulraumplanung oder gehen unsere Schulen sowieso den Bach ab. Dann kann man sich da eine eigene Sammlung zusammenstellen, wo man die Graphen sammeln kann. Dann gibt es da noch den Profi Bereich. Den muss man sich für 4 CHF einen Jahreszugang kaufen und da kann man dann selber auswählen was für Graphen-Typ man will, will man ein Linien-Diagramm mit Balken. Dann kann man noch auswählen ob man es für einen Kanton, Region, Gemeinde will. Für welchen Zeitraum etc. Da hat man dann nicht nur die standardisierte Auswahl welche man da vorne hat sondern man kann halt eigene Kombinationen zusammenstellen. Jetzt ganz neu haben wir hier noch ein Login Bereich. Dieser ist immer nur Gemeindespezifisch und ist auch aus der Idee der Schulraumplanung entstanden. Da haben wir noch für Teilgebiete der Gemeinden Auswertungen. Man kann die Bevölkerungszahlen nach Jahren nach verschiedenen Grössen anschauen aber auch Bevölkerungsbäume für Teilgebiete der Gemeinden machen. Dieser ist eigentlich der einzige Bereich welcher nicht Open Data ist sondern, dass müssen uns Gemeinden direkt geben weil dies heikle Daten sind. Das sind Daten mit Geburtsdaten und Wohnadressen und dies ist dann halt ein Schritt weiter.

Meine Arbeit würde sich jetzt halt nur auf diese Bereiche beschränken die nur Open Data sind und wirklich nur spezifisch Bildungsdaten. Weil es gibt verschiedene High Value Bereiche in denen man diese SROI Analyse machen kann. Ich mache dies jetzt spezifisch für Bildungsdaten. Ich würde dich sonst mal ein paar Sachen fragen wie Ihr arbeitet und was Ihr mit denen Daten macht etc. Den Fragebogen hast du ja schon durchgelesen. Wenn Ihr so ein Projekt annehmt von einer Gemeinde, wie geht man da genau vor? Verlangt ihr die Daten direkt von den Gemeinden oder holt ihr die irgendwo anders?

**Dies ist sehr unterschiedlich. Je nach dem wie viel Daten wir schon da drin haben aber das hat sich seit dieser App sehr gewandelt. Früher mussten wir viele Daten von den Gemeinden haben und jetzt geht es zum Teil in die Richtung, dass sie uns sagen sie hätten diese Daten gar nicht mehr abgelegt und verweisen auf das BfS. Diese Daten haben wir aber natürlich schon. Vielmals versuchen wir noch genauere Daten zu bekommen aber oftmals greifen wir dann auf die Open Data zurück.**

Zu was genau habt ihr da Daten? Schülerzahlen? Herkunftsland?

**Bildungsspezifisch haben wir da nur zusammengefasst die Anzahl Schüler nach Stufen. Dies ist Open Data welche wir so beziehen. Dann haben wir von ihnen noch die Anzahl Klassen und Aufteilungen nach Schulhaus. Dies ist dann aber nicht mehr Open Data.**

Wie sieht es mit den Niveaus aus?

**Genau dann haben wir noch die Unterteilung in Niveaus und Spezialklassen. Das ist dann eigentlich so wie hier in der App wo man grob sieht so viele Schüler haben sie im Kindergarten, so viele in der Primar und in der Sek. Diese sind dann aber noch nicht unterteilt in Klassen. Bei den Gemeinden ist das halt dann detaillierter.**

Die Daten welche in dieser App drin sind, kommen ja eigentlich primär vom BfS?

**Ja genau und noch die Schülerzahlen vom BiSta da das BfS in Sachen Schülerzahlen immer bisschen hinterher hinkt.**

In welcher Form bekommt ihr die Daten? Bekommt ihr diese in einem Excel oder wie?

**Meistens kommen die Daten in einem CSV**

Dann müsst ihr da die Daten gar nicht mehr gross umformatieren?

**Ja das ist eigentlich auch genau der spezial-deal zwischen ihnen und uns. Die Daten sind zwar grundsätzlich gratis jedoch müssen wir noch eine Aufbereitungsgebühr bezahlen, ca. eine Stunde, so dass wir es genau in der gewünschten Struktur bekommen. Das ist auch für uns gut, weil wenn wir ein Excel-File von der ganzen Schweiz mit allen Daten bekommen, wären wir auch mind. eine Stunde dran um diese zu filtern und sortieren.**

Das wäre dann ja quasi eine Investition in ihren Output den die Eckhaus AG macht. Können sie über die finanziellen Zahlen oder Arbeitsaufwand sprechen? Was der Aufwand ist zum diese Datensätze in die App zu implementieren.

**Im Idealfall haben wir den neuen Datensatz in einer Stunde implementiert. Manchmal haben wir haben aber auch Daten von einigen Gemeinden die das nicht richtig filtern können und da sind wir dann halt schon einen halben Tag dran um diese Daten zu aktualisieren. Wir können von ca. einem halben Tag pro Datensatz ausgehen. (ca. 30 Datensätze)**

Also ist der Aufwand im Grossen und Ganzen eigentlich relativ unwesentlich. Nachdem man diese Daten eingelesen hat ist ja das Endprodukt das Tool was ihr anbietet. Damit arbeitet man ja für die Gemeinden die sie anfragen. Wie läuft das genau ab? Werden sie da von den Gemeinden angefragt und gehen dann zu den Gemeinden und beraten diese wie ihre Schülerentwicklung sein wird? Wird aufgrund dessen eine Schulraumplanung gemacht?

**(Zeigt etwas in der App) Wir haben hier eigentlich wie die Vergangenheit von den Schülerzahlen. Dann machen wir mit verschiedenen zusätzlichen Input-Daten die wir haben, wie z.B. Einwohnerwachstum, Wohnbautätigkeit etc. prognostizieren wir eine Entwicklung, wie sich ihre Schülerzahlen entwickeln werden. Das ist eigentlich der Teil der auf Daten basiert die es schon gibt. Es ist lediglich eine Hochrechnung**

von uns. Zum dann Trends zu erstellen wie sich das entwickeln wird oder was typisch für diese Gemeinde ist, für das brauchen wir dann eigentlich all diese Raumberechnungsdaten. Diese sind dann mehr allgemeine Gemeindedaten welche einen Einfluss auf die Bildungsdaten haben. Man kann nicht nur an den Bildungsdaten wie sich die Gemeinde entwickeln wird. Danach machen wir Empfehlungen in einem Gespräch mit der Gemeinde oder der Stadt oder machen Workshops und versuchen das Verständnis für die Datengrundlage zu gewinnen.

Ich habe mal auf der Homepage gesehen ob ihr da aktuelle Projekte habt. Mal angenommen ihr habt ein Projekt und macht da eine Schulraumplanung. Wie sieht es nachher aus? Überwacht ihr die Kennzahlen nach Abschluss des Projekts immer noch? Überprüfen sie ihre Prognosen?

**Ja, in der Stadt Winterthur sind wir schon seit 10-12 Jahren Prognosen am machen. Diese Prognosen aktualisieren wir alle 2-3 Jahre. Gerade die Stadt Winterthur oder aber auch die Stadt Zürich ist ja explodiert mit der Geburtenrate. Diese beobachten wir dann aber auch laufend. Wir haben gerade zu diesem Zeitpunkt begonnen mit ihnen zu arbeiten. Jedoch hat damals niemand erwartet, dass dieser Anstieg anhalten wird. Dies war dann natürlich auch in unsere Prognose falsch. Aber genau deswegen werden diese Prognosen auch immer aktualisiert. Deswegen haben wir immer die Kontrolle und machen diese Hoch und Tief Szenarien damit wir den Spielraum aufzeigen können. Für uns ist Open Data als Datengrundlage, auch wenn sie manchmal ein bisschen hinterher hinkt, ideal. Man sollte nicht versuchen eine Genauigkeit zu suggerieren welche man nicht hat.**

Was würden sie sagen was der Mehrwert der Eckhaus AG ist für die Gemeinden? Was haben diese durch die App und durch die Firma für einen Mehrwert im Vergleich zu vorher?

**Sie können mit der App sicher ihre Gemeinde selber beobachten. Im Moment ist es vielfach so, dass Gemeinden zu uns kommen und sagen:**



**„ui! Die Kinder stehen alle schon vor der Türe, wir haben das gar nicht bemerkt!“ Und mit dieser App, und paar banalen Graphen können sie vorbereitet sein auf das, so dass sie nicht mehr so überrollt werden. Ich denke der zusätzliche Mehrwert ist das Verständnis und die Kommunikation wie sie diese Daten einbetten müssen etc. da wir ihnen auch beratend zur Seite stehen können da wir den Vergleich mit anderen Gemeinden haben. Es geht auch oftmals um Gespräche das erst mal alle abgeholt werden und alle in der Gemeinde realisieren dass es eine Veränderung gibt. Und wenn sogar das BfS das mit seinen Zahlen bestätigt, dann stimmt das schon.**

Hat man zu dieser App auch konkrete Nutzungszahlen? Wie viele Klicks, welche queries am meisten Downloads hatte, welche Graphen am interessantesten waren oder aber welche Gemeinden das am meisten genutzt haben? Oder hat man da seit Jahren immer noch die gleichen Nutzerzahlen?

**Es gäbe eigentlich schon Zahlen, aber es ist bis jetzt sehr schwer Aussagen darüber zu machen da es bis jetzt als Schulraumapp sehr spezifisch für Schulraumplanung lief. Das war für die Verbreitung recht schwer. Falls jemand eine Gemeinde beobachten wollte ging er dafür nicht in eine Schulraumapp. Jetzt erst seit 09.09.2016 heisst die App neu EckdatenApp und damit öffnet sie sich für die breite Nutzung für die allgemeine Raumbesichtigung der Gemeinden. Dies ist jetzt halt erst so kurzfristig, dass man da jetzt noch nicht gross publik gemacht haben. Seither nutzen es nur einzelne User. Der grösste Wert im Moment ist unser eigener, dass wir diese App selber brauchen können für Gemeinden. Die Gemeinden sind immer sehr überrascht, was wir da alles zusammentragen können. Sie bekommen lieber alles präsentiert über. Ausgedruckt und schön zusammengestellt mit einer kurzen Beschreibung zu den Graphen.**

Es ist ja interessant, da diese Daten eigentlich von den Gemeinden generiert werden. Für mich sehr interessant zu sehen dass eigentlich der Datenursprung bei den Gemeinden ist und über das BfS zu ihnen kommt und dann

damit wiederum die Gemeinden profitieren und das ganze einen Kreis macht und man durch diesen Prozess am Schluss einen Mehrwert hat. Obwohl man dies auch von Anfang an hätte machen können.

**Wir konnten auch bei den Kantonen recht spannende Entwicklungen beobachten, dass sie eigentlich sehr gute Datengrundlagen hätten aber auch auf der Kippe stehen mit all diesen Sparprogrammen, ob sie weiterhin die Daten führen. Wenn wir mit unserer App kommen, sind sie froh, dass ein Nutzer der Daten auf diese Daten angewiesen ist. Dies ist für die Behörden auch ein Bestätigung die Daten weiterzuführen. Die Kontinuität der Datensätze ist für uns extrem wichtig. Wenn dies mal eingestellt wird für zwei Jahre, dann sind diese Daten für uns nutzlos.**

In der Schweiz stecken wir in diesen Themen ja noch in den Kinderschuhen im Vergleich zu den USA.

**Auch bei den Datennutzungsrechten kämpfen wir zum Teil extrem gegen Windmühlen an weil die Gesetzesgrundlage von Kanton zu Kanton verschieden ist. Wir mussten auch schon Nutzungsverträge unterschreiben oder Daten nach dem Auftrag wieder löschen.**

Ich habe auch schon gehört, dass die Datenbeschaffung unter Umständen sehr schwierig sein kann und dass man diese zum Teil sogar erklagen muss.

**Die Daten werden mit enormen Kosten erhoben und gepflegt aber man darf diese nicht problemlos nutzen. Es ist ja gut wenn man Datenschutz gross schreibt. Jedoch gerade ein Beispiel mit Adressen und Geburtstagen bei so einem Bevölkerungsbaum kann man ja eigentlich nicht viel herausfinden.**

Wenn wir gerade bei dem Thema sind, Open Data hat ja sicherlich ein grosses Potential. Könnte es aber auch ein Risiko darstellen? Z.B. wegen Nutzungsrechten, Privatsphäre o.ä. Könnte man Open Data auch kritisieren oder aber missbrauchen?

**Es kommt natürlich darauf an. Bei diesen Datensätzen auf die ich gestossen bin, da sehe ich eigentlich kein Problem. Es ist vielleicht gerade in Teilgebietsanalysen ein bisschen schwierig, wo man die Grenze zieht. Es sind ja nicht die Einzeldaten aber man könnte ja ein Kreis um drei Häuser ziehen und diese auch analysieren. Da würde man ja genau wissen wie alt die sind. Das ist dann halt auch die Frage ob es tragisch ist, wenn man das weiss. Da geht es halt mehr um den Verwendungszweck. Wir machen jetzt Schulraumplanung. Es könnte aber ja auch sein, dass wenn diese Daten öffentlich wären, eine aggressive Werbeagentur diese verwenden würde.**

Man muss halt immer mögliche negative Einflüsse der Open Data Politik thematisieren. Personalisierte Werbung könnte ja etwas positives sein aber es ist natürlich auch so, dass manche Personen dies gar nicht begrüßen.

**Wir haben noch oftmals, wenn wir etwas verwenden dürfen Verträge wofür genau diese Daten verwendet werden dürfen. z.B. vom Bfs mit Adressdaten inkl. den Koordinaten. Da ist es ganz klar eingegrenzt für Schulraumplanung, für Zwecke die der Gemeinde dienen und nicht für private Zwecke.**

Dies kann ja gerade auf der gesetzlichen Seite eine Rolle spielen, dass die Gemeinde oder das Volk sich absichern will gegen solche Missbräuche. Wenn ich nochmals zu den Userzahlen zurückkommen kann, wie viele User sind es denn die diese App momentan verwenden?

**Momentan sind es vielleicht 10-20 User, also eigentlich sehr wenige. Die meisten Klicks sind von uns in-house da wir selber die App zum arbeiten brauchen. Deshalb kann man nicht genauere Angaben machen was auswärtige Nutzer abrufen. Wir haben jetzt gerade im Kanton Zürich viele Städte und Gemeinden die wir betreuen. Genauso im angrenzenden Kanton Aargau. Aber es ist eben nicht so, dass diese die App selber nutzen.**

Machen sie irgendwelche Anstrengungen diese App in den Gemeinden zu verbreiten oder ist diese eher für die interne Verwendung vorgesehen?

**Es ist eigentlich schon die Idee auch kleine Gemeinden sich selber beobachten können. Und erst wenn es dann zu grösseren Veränderungen kommt brauchen sie dann uns oder aber andere Schulraumplaner. Da wir ein kleines Büro sind können wir ja sowieso nicht die ganze Schweiz abdecken. Man bringt die App schon immer wieder ein aber man verbreitet sie auch nicht mit einem offensiven Marketing in allen Zürcher Gemeinden.**

Was könnte man in der Zukunft noch spezifisch mit Bildungsdaten machen?

**Was wir beobachten ist, dass es recht strikte Trennung gibt zwischen der politischen und der Schulgemeinde. Aufgrund dieser Trennung planen diese auch eigenständig. Zum Beispiel weiss die politische Gemeinde nicht wie es mit den Schülerzahlen oder aber den Finanzen aussieht. Sie wissen nicht ob es ein neues Schulhaus braucht etc. Wenn sie in einer App zum Beispiel ihre Bevölkerungszahlen und Wohnbautätigkeit neben den Schülerzahlen sehen würden, könnte es sein, dass sie sich mehr Gedanken über die Schulraumplanung machen. Da ist es natürlich wichtig, dass diese Daten möglichst einfach Verfügbar sind. Gerade in unserem Fall ist halt die schweizweite Verfügbarkeit extrem praktisch. Dann müssen wir nicht bei allen Kantonen einzeln diese Daten verlangen. Der grösste Aufwand für uns ist an diese Daten heranzukommen. Bekommen wir diese Daten überhaupt? Wer ist die Kontaktperson? Wie viele Jahre zurück gibt es diese Daten in dieser Struktur? Mit der ganzen Open Data Bewegung wäre diese ganze Sache einfacher da wir dann alles direkt zur Verfügung hätten. Wir wurden halt auch immer vertröstet dass wir diese Daten in 2-3 Jahren ja sowieso bekämen. Der Weg dahin ist halt schwierig. Aber wenn man diese Daten mal in der App hat dann hat man schon eigentlich geschafft.**

Haben Sie grosse Unterschiede bei den Daten die sie bekommen von unterschiedlichen Gemeinden Kantonen etc.?

**Ja wir haben immer sehr unterschiedliche Daten. Dies ist sehr Personenabhängig. Zu einigen statistischen Ämtern haben wir kontakt. Dann müssen wir denen erklären was wir machen etc. und dann bekommen wir es. Beim BfS kann man zum Teil zwar eigene Query's zusammenstellen aber halt z.B. nur für 100 Gemeinden und dann muss man dem entsprechend verschiedene Files generieren bis man die ganze Schweiz hat. Wieso macht man es nicht, das man nicht direkt alles gerade bekommt? Es fühlt sich fast so an als wären es extra Hürden damit man es nicht so offensiv braucht. Wiederum will man ja, dass man diese Daten verwendet.**

Haben Sie schon mal die Erfahrung gemacht, dass Behörden gewisse Daten nicht geben wollten oder aber sich gewehrt haben?

**Es ist eigentlich eher die Verfügbarkeit die sehr kritisch ist. Wir gehen auch immer mehr in die Richtung, dass die Gemeinden direkt selber in unserer App Informationen abrufen können und Prognosen durchspielen könnten. Für die fundierten Erkenntnisse ist es aber zu schwierig alles herauszugeben.**

**Eine Gefahr von Open Data könnte sein, dass man in der Statistik falsche Zusammenhänge sehen könnte.**

Können Sie mir Personen vermitteln die einen Nutzen von Offenen Bildungsdaten hatten? Z.B. Eine Schule wo eine Schulraumplanung stattfand.

**Ja das würde ich noch intern mit dem Büro absprechen und Ihnen danach eine Mail schreiben.**

Die Daten die Sie in Ihrer App verarbeiten, kann man diese auch als CSV rausziehen oder gibt es diese nur in Form dieser Graphen?

**Das ist eigentlich eine Bedingung die wir einhalten müssen. Wir dürfen die Daten nicht weiterverkaufen. Wir dürfen nur die Aufbereitung weiterverkaufen. Deshalb gibt es nur die Graphen. Wir müssen natürlich auch den Datenschutz einhalten und dürfen keine Rückschlüsse auf Personen zulassen. Weiter dürfen wir mit diesen Daten auch nur für Aufträge arbeiten die für Gemeinden bestimmt sind. Wir dürfen diese Daten nicht für Arbeiten mit Privaten verwenden.**

-Ende-

---

## **Interview with Mrs. Telley, Statistic Coordination of the Canton of Bern (22.09.2016)**

Wer sind sie, und was machen sie genau?

**Ich bin Ursula Telley und bin Statistikkoordinatorin vom Kanton Bern. Dies ist ein bisschen speziell da unser Kanton anders Organisiert ist als andere. Wir haben kein statistisches Amt. Wir hatten früher auch eins, das grösste der Schweiz. Jedoch wurde das abgeschafft.**

Was wären die Gründe für die Abschaffung?

**Das war irgendwie in den 80er Jahren. Die Entwicklung war in die Richtung, dass dies ein sehr grossen Amt war. Aber sich immer wie weiter vom Nutzer entfernt hat. Das ist auch ein grosses Thema zur Zeit beim BfS. Sie versuchen sich immer wie besser am Nutzer zu orientieren. Wir sind sehr froh über diese Open Data Bewegung. Wir wollen, dass die Leute diese Daten brauchen.**

**Danach hat man diese Koordinationsstelle eingesetzt die für den Bund die Ansprechperson Seitens Kanton war. Wir haben dann Statistikkonferenzen eingesetzt. Jede Direktion hat einen Statistikverantwortlichen und jede von diesen ist die Aufgabe die Statistiken erheben. Wir koordinieren hier lediglich. Alle Anfragen kommen zu uns hier und wir leiten dann die Anfragen weiter. Wir machen die Stellungnahmen, wir beraten die Regierung und wir beraten nach aussen.**

**Wenn ich uns mit anderen Kantonen vergleiche dann haben wir weniger Daten die wir publizieren, weniger Ressourcen, sehr wenig Ressourcen, wir können nicht gleich professionell arbeiten wie andere weil wir Dezentral organisiert sind. Wir können nicht die gleichen Tools haben. Wir stehen in Konkurrenz, was Ressourcen anbelangt, zu anderen Stellen. Wenn es ein Sparpaket gibt und dann bei uns gespart wird dann kann man da halt nichts machen. Weiter kommt dazu, dass nicht alle Daten publizieren wollen. Bildung wäre eigentlich sehr interessant. Z.B. waren**

---

**es die Übertrittszahlen von einzelnen Schulen. Dies konnte man nicht herausgeben. Da wird zum Teil vom Grossen Rat verboten gewisse Daten zu veröffentlichen. Die einzelne Person ist bei uns immer geschützt aus Datenschutz technischen Gründen. Bei Einzeldaten sind Personen dahinter. Diese sind geschützt.**

Sie sagten ja Sie sind die Koordination hier. Wie ist das zu verstehen?

**Meine Koordination ist organisatorischer Art und nicht was Daten angeht. Die Daten werden von den Gemeinden direkt zum BfS geschickt. Diese Daten werden ja sowieso in elektronischen Tools gesammelt. Früher war das natürlich manuell. Da haben die Gemeinden das ganze erhoben und an mich geschickt, ich habe dann diese Erhebungen kontrolliert und ans BfS weitergeleitet. Dies ist jetzt alles elektronisch. Das heisst die Daten gehen alle direkt von den Direktionen zum BfS.**

-Ende, da das Interview nicht relevant für die Thesis-



---

## **Interview with Mr. Allraum, Education Statistics of the Canton of Bern (29.09.2016)**

Können Sie mir erzählen wer Sie genau sind und was Sie genau machen?

Ich bin hier verantwortlich für die Bildungsstatistik des Kantons Bern. Wir sind hier 3 Leute im Team. Wir sind mit 3 Datenerhebungen beschäftigt. Das ist einerseits die Kernerhebung wo wir die Schülerdaten von allen Berner Schüler sammeln. Inkl. Tertiärstufe B aber ohne Tertiärstufe A also ohne Hochschulen. Wir erheben auch Lehrkräftedaten und Bildungsabschlüsse. Wir erheben diese einerseits für uns damit der Kanton Bern eine gute Datengrundlage hat für die Steuerung des Bildungssystems. Dies ist der Hauptgrund. Das man halt einfach Informationen hat wie das System läuft. Die Volksschulen sind ja Gemeindegemeinschaften mit einer Ausnahme. Wir erheben einerseits diese Daten andererseits liefern wir diese drei Erhebungen dem BfS. Zwei davon würden wir ansonsten gar nicht machen, wenn wir diese nicht liefern müssten. Wir selber brauchen eigentlich nur die Lerndaten. Lehrkräftedaten sind eigentlich schon im Personalverwaltungssystem des Kantons Bern drin. Diese müssen wir nicht extra erheben. Mir diesen kann man auch schon direkt alle Auswertungen machen die interessant sind. Aber für das BfS müssen wir auch noch diese von den Privatschulen erheben. Obwohl wir noch nie eine Auswertung damit gemacht haben. Kostet aber den Kanton Bern sprich den Steuerzahler. Wir bekommen viele Anfragen von Studierenden die Daten möchten. Am liebsten gerade 20 Jährige Zeitreihen, was wir in der Regel nicht liefern können weil das momentan zu aufwändig ist, wie wir organisiert sind. Der wichtigste Nutzen dieser Daten ist das wir die Lerndaten brauchen für den Lastenausgleich zwischen Kanton und Gemeinden. Der Kanton zahlt alle Lehrer in öffentlichen Volksschulen. Die Gemeinden welche zuständig sind für diese Schulen müssen dann dieses Geld dem Kanton zurückzahlen. Dies richtet sich nach diesen Schülerzahlen. Wir erstellen unsere eigenen Publikationen wo wir natürlich hoffen dass diese beachtet werden.

**Es ist vor allem eine Broschüre. Der grosse Aufwand bei diesen Datenerhebungen ist ja die Datenbereinigung. Die Datenanalyse machen wir eher zu wenig da wir zu wenig Ressourcen dafür haben. Es ist sogar eher weniger geworden als früher. Wir müssen immer mehr Ressourcen in die Datenqualität reinbuttern, da dies vom BfS so verlangt wird. Sie haben eine Onlineplattform wo wir die Daten abliefern. Dies macht dann tausende Plausibilisierungen und da kommen danach zehn tausend Fehler dabei raus. Die müssen wir dann alle irgendwie bereinigen. Das ist irrsinnig, weil wie gesagt das ist eine Vollerhebung wo 160'000 lernenden. Da kann man viel überprüfen und viel Fehler finden. Die Datenbereinigung ist nicht einfach. Aber wir machen trotzdem die wichtigsten Auswertungen selber da es überhaupt nicht einfach ist diese auszuwerten. Es haben immer alle das Gefühl das ist ein easy peace of cake aber ist es nicht. Man muss die Daten wirklich kennen damit man gescheite Auswertungen damit machen kann. Sonst gibt es Unsinn.**

Die Veröffentlichungen die Sie machen sind in diesem Fall nicht die Rohdaten aber eher die Auswertungen?

**Ja, dies wäre schon rein aus Datenschutz Gründen gar nicht erlaubt. Das sind Individualdaten. Diese werden mit AHV-Nummern erhoben. Was wir publizieren sind aggregierte Daten. Dies ist unser einziger Beitrag an Open Data momentan.**

Das ist eben interessant für mich. Was die Leute die gerade selber direkt damit beschäftigt sind, was sie sich wünschen etc. Ich habe gemerkt das man momentan gar nicht da ist wo viele der Direktbetroffenen sein möchten mit dem Fortschritt. Sie haben ja auch Datenschutz genannt. Dies ist natürlich immer eine sehr grosse Barriere die wir hier haben. Sie haben vorher noch die Studenten genannt, welche gerne Auskunft über die letzten 20 Jahre haben möchten. Sie haben ja auch gesagt das da der Aufwand enorm wäre, aber die Daten sind ja eigentlich vorhanden?

Die Erhebung ist ausgelagert und wird von einer externen Firma gemacht. Dies wird sich in den nächsten 2-3 Jahren wieder verändern und in-house erhoben. Die Ansprüche an die Daten sind enorm gestiegen. Deshalb macht dies nicht mehr Sinn, dass dies eine externe Firma macht. Wir haben gemerkt, dass uns diese Daten wichtiger geworden sind als früher, sie sind gehaltvoller und qualitativ hochwertiger. Deshalb möchten wir diese auch selber erheben.

Wieso sind diese Daten wichtiger geworden?

Der einte Punkt ist tatsächlich dass wir seit ein paar Jahren mit der AHV-Nummer erheben. Dies hat zur Folge dass wir ganz andere Analysen möglich sind. Jetzt können wir Längsschnitt Analysen machen. Jetzt können wir schauen was für Typen von Schülern das Schulsystem so durchlaufen wie wir uns das vorstellen. Schön ein Jahr nach dem anderen. Dies kann man nun analysieren. Früher konnte man den Schüler im Kanton Bern nicht identifizieren. Andere Kantone haben die Namen miterhoben und konnten dies auch schon vorher machen. Dies ist ein Grund wieso die Daten wertvoller geworden sind. Aktuell habe ich eher das Gefühl das es vielerorts (nicht nur im Kanton Bern) zurück geht. Wieder in die Richtung, dass nur noch über den Wert der Daten gesprochen wird aber weniger genutzt wird als auch schon.

Gerade bei der politischen Machbarkeit. Hat man da grosse Barrieren, dass man gar nicht die Sachen machen kann, welche mein eigentlich möchte bei den Bildungsdaten?

Es gibt halt Zahlen die sind ein gewisses Politikum. Zum Beispiel Klassengrösse. Dies ist extrem heikel. Ich gebe keine Zahlen zu Klassengrössen raus. Dies überlasse ich dem Amt. Die bekommen unsere Daten über und werten dann selber aus. Es gibt einen Auftrag vom Grossen Rat wie gross das Klassengrössen sein sollen. Eine andere heikle Zahl ist die Übertrittsquote. Wie viele gehen in die Realschule wie viel in die Sek? Dies ist weniger für den Kanton ein Problem aber eher für die

**Gemeinde. Z.B bei einer kleinen Gemeinde welche eine Sek als eine Real liefert und schon haben wir eine komplett falsche Einschätzung. Da muss man vorsichtig sein. Diese Zahlen sind schon in der Bildungsstatistik-Broschüre drin. Aber halt nicht nach Gemeinde aufgelistet. Das Problem ist, dass man die Daten sehr gut verstehen muss um diese richtig zu interpretieren. Diese Zahlen sind nicht der Leistungsausweis einer Schule. Dies ist nicht so. z.B. eine kleine Gemeinde ist zu klein um eine Real und eine Sek zu haben. Dann gehen diese Schüler in eine andere Schule. Zwar ist unsere Auswertung nach Wohngemeinde aber es ist unwahrscheinlich, dass die Sekundarschülerquote höher ist weil dann die Schüler nicht in eine andere Gemeinde in die Schule müssen. Danach ist auch die Frage ob dies gut oder schlecht ist. Z.B. Haben sie einen langen Schulweg aber dafür machen sie die Sek. Ist dies für sie ein Nachteil? In der Stadt Bern ist dies durchaus ein Nachteil, da man mit der Realschule fast keine Lehrstelle findet und somit gezwungen ist, die längere Strecke in die Sek in kauf zu nehmen. Auf dem Land ist dies aber unter Umständen anders. Dort könnte es heissen man kennt ja die Schüler. Die bekommen alle eine Lehrstelle über. Es ist nicht so einfach zu interpretieren. Das grösste Problem ist bei kleinen Gemeinden eine solche Übertrittsquote zufällig und variiert jedes Jahr. Es gab mal eine Anfrage vom grossen Rat und dann wurden diese Daten für alle Gemeinden publiziert. Dann haben die Medien gefunden das sei ein interessantes Thema. Ein paar Jahre später sind sie wieder gekommen und wir haben diese Daten geliefert und sind vorsichtiger geworden und haben nur noch bei denen Gemeinden Angaben geliefert wo die Grundgesamtheit gross genug war. Danach kam der Vorwurf der Medien, dass wir etwas verstecken. Obwohl wir sogar noch eine Begründung geliefert haben. Deshalb sind wir auf den Schluss gekommen das es nicht korrekt ist diese Daten so zu publizieren.**

Die Daten werden hier ja erhoben. Klassengrössen, Übertrittsquoten etc. Dann ist es ja eher so, dass man Angst hat das diese Daten falsch verstanden werden, weil man die Daten nicht gut genug kennt. Ist dies auch ein Fak-

tor das man diese Daten nicht veröffentlicht sondern aber eine Broschüre rausgibt mit einer Interpretation etc.?

**Wir haben es nützlich gefunden eine kleine Publikation zu haben mit einigen Angaben. Früher haben wir viel dickere Handbücher publiziert. Da war aber der Qualität nicht sehr zu trauen. Die Broschüre wird den Schulen geliefert und dies ist auch eine Rechtfertigung der Datenerhebung und dann ist es für sie auch interessant und sie sehen das damit etwas gemacht wird. Die anderen Sachen haben wir eher Online aber wie gesagt das ist nicht wahnsinnig viel. Natürlich könnte man alles in eine Online Form bringen. Aber dies ist nicht weniger aufwändig.**

Die Sachen die sie online stellen, sind diese auch als Broschüre oder aber als Rohdaten verfügbar?

**Es ist nur in einem Fall online. In allen anderen Fällen ist es ein PDF.**

Spielen da auch noch andere Faktoren als die eben besprochenen oder aber Datenschutz eine Rolle das diese Daten nicht veröffentlicht werden?

**Das File wo ich hochgeladen habe, damit kann man fast alles machen. Damit kann man sehr viel machen.**

Sie haben gesagt Sie haben 3 Datensätze, welche sie an das BfS weiterleiten und 2 davon würden sie selber gar nicht machen, wenn das BfS diese nicht brauchen würde. Welche zwei sind das?

**Statistik von den Lehrpersonen in Privatschulen und Bildungs- und Abschlüsse. Da erheben wir aber nicht alle Abschlüsse. Z.B. die Eidg. Maturität erfasst der Bund direkt selber. Die berufliche Grundbildung auch. Die wichtigsten Abschlüsse erheben sowieso nicht wir. Deshalb ist eine Auswertung nicht aussagekräftig.**

Wieso wurde die Erhebung der Daten ausgelagert?

**Dies wurde vor 20 Jahren gemacht. Als die Behörde welche diese Erhebung machte verselbstständigt wurde, wurde damit auch die Durchführung dieser Erhebung ausgelagert. Aber da die Daten sich geändert haben ist die Informatik nicht mehr die beste Lösung für diese Erhebung.**

Die Daten welche mit dieser Broschüre oder aber im Internet veröffentlicht werden. Haben sie dazu Nutzungszahlen? Sehen Sie welche Datensätze wie oft heruntergeladen werden?

**Wir drucken nur diese eine. Jede Schule und jede Gemeinde bekommt ein Exemplar über. Das sind ca. 500 Stück. Zu den Online-Daten habe ich keine Nutzerzahlen.**

Wissen sie Beispiele von Drittpersonen welche Ihre Daten weiterverwendet haben für irgendetwas?

**Es merkt schon mal niemand das es diese Daten online gibt. Viele recherchieren gar nicht erst selber. Das ärgert mich manchmal. Über Google ist es echt nicht schwierig diese Homepage zu finden. Es hat sich noch nie jemand bei mir gemeldet und gesagt ich habe die Daten da gefunden und das und das gemacht damit. Deshalb weiss ich auch nicht genau wieso ich noch viel mehr Ressourcen da rein stecken sollte. Man könnte viel damit machen aber es wird nicht gemacht. Die Daten sind nicht selbsterklärend. Diese zu verstehen und zu interpretieren ist zeitaufwändig. Ich kenne viele Beispiele von Personen die die Daten falsch interpretieren. Sie haben dann zwar die Daten aber was sie daraus ziehen ist schlicht falsch. Die Daten werden als CSV und Excel publiziert.**

Was würden sie sich wünschen was in der Zukunft mit diesen Daten gemacht wird?

**Diese Daten werden ja schon gebraucht. Sie werden viel gebraucht aber nicht via diesen Publikationen die wir machen sondern mehr über die-**

sen Auswertungen welche wir her selber anstellen für diverse Leute. Z.B. für Schulen oder letzte Woche für einen Gymnasialschulleiter der wissen wollte wie die Demographische Entwicklung bei Ihnen sei. Solche Auswertungen werden genutzt. Für wen es sonst noch nützlich sein könnte , könnte ich mir höchstens vorstellen, dass so für ein Rating genutzt werden könnte. Attraktivität vom Wohnort. Das könnte man damit machen. Wir bekommen immer Anfragen über von Leuten die Schullisten möchten um die Schulen mit Werbung einzudecken. Deshalb können wir diese Schullisten nicht herausgeben. Die Adressen dürfen auf diesen Listen nicht ersichtlich sein. Für die Vorlesung von Herrn Stürmer wollte eine Gruppe diese Schulen geographisch lokalisieren. Da wurden z.B. die Adressen der Schulen zufällig variiert im 100m Umkreis. Aber ist natürlich wieder ein Aufwand etc. Das wäre eine Anwendung welche Leute interessieren könnte. Wo ist der nächste Kindergarten die nächste Sek das nächste Gymnasium etc.

Es wäre nützlich für eine Mobilitätsmatrix. Wir haben den Wohnort und wissen den Schulort der Schüler. Dann haben wir die Mobilitätsmatrix. Dies wurde von uns auch schon gemacht. Das Bundesamt etc. verwendet diese für ihre Mobilitätsstudien. Z.B für technical skills. Wenn es irgendwo viele Ingenieure gibt wäre es da evtl. spannender für eine technische Firma etc.

Ich würde halt mit diesen Daten mehr machen also selber mehr machen als aktuell gemacht wird.

Wäre es evtl. erstrebenswert diese Daten zu vereinfachen, um diese dem Laien verständlicher zu machen damit diese selber Auswertungen damit machen können?

Ich denke wir liefern alle Informationen die es braucht. Man muss sie halt einfach lesen. Dies machen die meisten nicht. Das ist das Erhebungsfragebogen. (zeigt etwas)

Wäre es wünschenswert, wenn man mehr Freiheiten bezüglich Datenschutz hätte?

---

**Ja dies wäre sicher auch wünschenswert. Die Politik könnte uns da schon sehr viel helfen. Kanton Bern hat kein Statistikamt. Das ist nicht sehr effizient. Die Politik hat seinerzeit dieses Amt unter anderem aus Spargründen abgeschafft.**

Was fliesst eigentlich Ressourcenmässig herein?

**Die grössere Erhebung kostet ca. 300'000 pro Jahr.**

Sind das auch die drei die man weiter ans BfS gibt?

**Genau. Auswertungen zu diesen Zahlen die wir haben bekommt man auch via BfS. Sobald man kantonale Vergleiche machen will, geht man sowieso besser zum BfS. Jedoch geben wir nicht alle Daten in der gleichen Auflösung an das BfS weiter.**

Die, welche 300k kostet, machen sie diese nur für das BfS oder brauchen sie diese selber

**Die brauchen wir selber.**

Vielen Dank für Ihre Zeit und die Auskunft.

-Ende-



## **Interview with Mr. Thomet (01.10.2016)**

Ich bin der Fidel Thomet hab in Zürich an der Zürcher Hochschule für Künste textdesign studiert und da war irgendwie ein Modul Daten Visualisierung. Das war in Kooperation mit Open Data Zürich. Darüber habe ich den Open Data Bereich und die Daten Visualisierung reingefunden. Hab dann auch Mitte letztes Jahres bis Mitte dieses Jahres das Praktikum bei Open Data Zürich gemacht einfach um noch ein bisschen stärker da rein zu tauchen und es ist ja beim statistischen Amt angegliedert. Da konnte ich auch noch mehr wissen zu Daten Visualisierung und Daten Analyse sammeln. Die Visualisierung die ich gemacht habe basiert auf Daten von SRFdata die diese sehr aufwändige Recherche gemacht haben. Die sind dann auch immer so freundlich dann, dass sie wenn es irgendwie möglich ist die Daten auf GitHub veröffentlichen. Dann habe ich halt irgendwie geschaut, wie kann ich diese Daten anders darstellen und aber auch selber ein bisschen so experimentieren. Für mich war auch der Prozess interessant, wie kriege ich die Rohdaten in die Visualisierung rein.

Die Daten die du gebraucht hast, hast du gesagt sind von GitHub sprich von SRFdata? Hast du nur diese Daten verwendet für deine Visualisierung?

**Ja genau es sind die Daten aus der Datenbank. Ich schaue noch kurz ob ich das Projekt finde.**

Für mich ist es so in meiner Arbeit. Open Data ist ja ganz klar definiert was es ist und was nicht mit der Open Data Charta. Die 10 Kriterien die Daten erfüllen müssen. Jedoch ist es im Moment so, dass sehr viele Daten nicht all diese Kriterien erfüllen. Und da geht es mir auch noch drum was ist jetzt in diesem Kontext im rahmen meiner Arbeit Open Data und was nicht. Ich denke wir werden da eine breitere Definition von Open Data haben als Open Data dies selber hat da wir sonst sehr wenige Open Data haben.

**Ich hab vorher einmal geschaut und von der Stadt Zürich ist vielleicht nur so ein Datensatz oder so verfügbar. Es gibt jetzt schon relativ wenig in diesem Bereich.**

Die Daten welche du da von GitHub hast, in welcher form hast du diese downloaden können?

**Dies war keine fertige Datenbank es war eher so eine SQL die ich dann in der Datenbank dann eigentlich einpflegen konnte. Das ist ja dann schon fast ähnlich wie das CSV nur das halt die Tabellen schon untereinander verknüpft sind.**

Was hast du danach noch alles verwendet um die Daten zu konvertieren?

**Ich habe erst mit einem R Projekt die Datenbank dann halt wirklich importiert und dann eigentlich da die Daten so aufbereitet wie ich sie dann brauche und dann halt so die CSV's exportiert**

Was hast du da für einen aufwand gehabt? Von wie vielen Stunden sprechen wir hier?

**Es war insgesamt nicht so viel da die Daten schon in einem sehr guten zustand waren. Das was ich...die Analyse war relativ schnell also in ein zwei stunden gemacht. Genau. Die Visualisierung habe ich nochmals 3-4 stunden aufgewendet. Insgesamt so etwa einen Arbeitstag.**

Das was du produziert hast, ist ja diese Visualisierung. Mit den Nebenjobs der Professoren. Kannst du ein bisschen mehr über diese Visualisierung erzählen? Was war deine Motivation und wieso genau diese Visualisierung?

**Das hat so mehrere Komponenten. Die Visualisierung die SRF gemacht hat war auch eigentlich sehr schön nur hat mir da so die Vergleichbarkeit gefehlt. Ist ja so von den Universitäten her aufgeschlüsselt. Ich hab mich dann halt gefragt gibt es vielleicht unternehmen die ganz spezi-**

**fisch einzelne Universitäten besonders in bestimmten Bereichen finanzieren. Was ich dann auch sehr spannend fand war, dass auch recht viel Geld dann für Administration irgendwie aus der Gesundheitsbranche gar nicht so erwartet. Genau. Ich fand es auch noch von der form her schön dieses Sankee Diagramm zu wählen was dann nochmals schön diese Verbindung abbildet. Es hat die 2 Seiten und die sind miteinander verknüpft. Dies ist in diesem Diagramm auch sehr gut gestalterisch so von der Metapher her.**

Ich finde jetzt auch, dass deine Darstellung viel übersichtlicher und einfacher verständlich ist.

**Bei der SRF geht es natürlich viel tiefer und ist sehr detailliert aufgeschlüsselt.**

Wieso hast du genau diesen Datensatz ausgesucht? Wieso genau Bildungsdaten?

**Dies war eigentlich sehr spontan. Ich mache gerne so Seitenprojekte aber schaffe es dann eher selten etwas fertigzubekommen. Dies war so der Zeitpunkt wo ich so mit R angefangen habe in die Datenvisualisierung und Datenanalyse einzutauchen. Für mich war es auch interessant wie ich dann so ein SQL eine Datenbank integrieren kann. Es war natürlich auch das Thema was ich recht interessant fand. Das SRF Daten zu ihren Artikel veröffentlicht war mir auch vorher bewusst. Aber ich hatte immer das Gefühl das diese Daten nicht wirklich komplex waren. Ich hatte das Gefühl man konnte da nicht viel mehr rausholen. Bei diesem Projekt war halt das Thema sehr interessant und die Rechercharbeit von SRF riesig. Und ich dachte diese Visualisierung könnte eventuell noch von der SRF gepusht werden.**

Ich habe auch deine Kontaktdaten von Julian Schmidli. Er hatte auch schon erzählt dass diese Recherche sehr aufwändig gewesen sei. Was hast du für

Ressourcen verwendet für von den Rohdaten bis zum Endprodukt zu gelangen?

**Wie gesagt habe ich für die Datenanalyse das R Stats verwendet. Das ist so eine Statistikprogrammiersprache. Das ist Open Source. Es gibt auch eine kommerzielle Version für Unternehmen und so aber eigentlich ist es gratis. Es ist auch bei Datenjournalismus eigentlich so das Standard Tool. Es ist unglaublich schnell mit großen Datenmengen zu hantieren. Ich finde auch den Einstieg erstaunlich einfach. In JavaScript ist vieles anders aber es ist halt sehr einfach und hat viele Schnittstellen um zum Beispiel auch so eine SQL Datenbank einzulesen. Dann habe ich da die CSV exportiert. Und die Visualisierung ist, dann web basiert also mit D3, was auch da der Standard ist. Ich habe schon auch in einem Grafikprogramm eine Skizze gemacht vom Projekt.**

Was hast du das Gefühl was mit deiner Visualisierung deinem Outcome erreicht werden kann?

**Ich glaube der Output von meiner Visualisierung ist jetzt glaube ich gar nicht gross. Ich habe es über Twitter publiziert und SRFdata hat es retweeted. Aber es ist natürlich extrem im Schatten der SRFdata Reportage gestanden. Ich habe auch keine Klickzahlen oder so. Ich gehe davon aus dass dies nicht wahnsinnig viel bewegt hat. Es gäbe halt schon irgendwie das Potential diese einzusetzen in einem Unterricht, weil es halt ein ziemlich zugängliches Bild über die Verstrickung gibt.**

Ich habe selber mal...Kennst du Matthias Stürmer von der Uni Bern?

**Der Name kommt mir bekannt vor.**

Ich habe mal die Vorlesung Open Data besucht bei ihm welche um Visualisierung von Open Data ging. Ich habe diese Visualisierung erstmals an der Uni in dieser Vorlesung gesehen. Deine Kontaktdaten habe ich von Julian

Schmidli von SRF Data. Er hat mir gesagt, dass du mit ihren Daten eine Visualisierung gemacht hast.

**Es ist ja jetzt nicht etwas was man in der Schule lernt. Man schneidet es höchstens an aber man studiert es ja nicht.**

Es ist sicher sehr schön dass du das macht es und veröffentlichst. Genau deine Arbeit ist ja das was Open Data legitimiert. Du zeigst das die Daten gebraucht werden und nicht auf irgendwelchen Datenfriedhöfen liegen. Darum ist das sicher etwas sehr positives, dass du eine Visualisierung gemacht hast.

**Ich finde jetzt so eine Visualisierung ist eigentlich wie ein relativ kleiner...also ist ein Tool für die Vermittlung von Informationen. Aber eben mit Open Data kann man eben noch den Schritt weitergehen und Anwendungen entwickeln welche Tools jetzt vielleicht bei dem Beispiel schwierig weil es statische Daten waren. Es geht auch viel mehr darum auf ein Thema aufmerksam zu machen. Man kann jetzt die Daten schwer weiterverwenden um wirklich einen...man hat keine App, wo man einen aktiven Nutzen draus ziehen kann. Natürlich bei andere Open Data Sachen hat man durchaus die Möglichkeit vielleicht einen Dienst der Bevölkerung zur Verfügung zu stellen der sonst vielleicht von den Behörden besetzt worden wäre evtl. viel schlechter oder aber teurere gewesen wäre.**

Du hast vorher erwähnt das du keine Klick zahlen hast. Hast du es sonst wie noch veröffentlicht als auf Twitter?

**Ich glaube nicht. Auf Facebook bin ich so wenig aktiv da ich es da wahrscheinlich nicht gemacht habe. Es ist halt eine Community-Frage. Ist halt mehr so Mund-zu-Mund Propaganda.**

Was denkst du was könnte der potentielle Nutzen sein von so Visualisierungen oder aber Open Data fürs Volk? Was könntest du dir noch vorstellen was

man mit Daten machen könnte? Das Ganze ist ja noch in den Kinderschuhen. Was hast du für Prognosen wo das Ganze hingehen könnte?

**Ich finde es hängt extrem von den Daten ab. Ich habe mich schon früher gefragt was sind das eigentlich für Daten die zur Verfügung stehen? Ein sehr cooler aber sicher auch heikler Datensatz wäre die Qualität, also die Durchfallquoten der Schulen bei der Matura oder so etwas. Es wäre vielleicht auch spannend zu sehen welche Schulen irgendwie Wechsel einerseits bei den Lehrer aber andererseits auch bei den Schülern haben. Ich glaube da gäbe es sicherlich eine Möglichkeit...Ist halt ein heikles Thema aber es könnte halt helfen die Qualität von Schulen zu erhöhen und irgendwie so eine Vergleichbarkeit zu bringen. Eigentlich ist es ja eine schreckliche Vorstellung, dass vielleicht irgendwie die Lebenswege von irgendwelchen jungen Menschen durch schlechte Schulen irgendwie so eingeschränkt und blockiert werden. Auch im Universitären Bereich finde ich es wichtig woher die Finanzierung stammt und wo wie viel Geld wofür ausgegeben wird. Es ist aber sicherlich auch noch gut wenn man die Vergleichbarkeit hat zwischen den Unis, so dass man kritisch sein kann gegenüber des Ausgangsfazits. Wenn irgendwo halt auffällt, die eine Universität braucht deutlich mehr Ressourcen in einem bestimmten Bereich, dann gibt es bestimmt all das zu hinterfragen wozu wird das genutzt und so. Ich hab auch schon häufig das Gefühl, dass es auch genug Leute gibt, die irgendwie das Gefühl haben...oder die Politik ist ja der Bildungsbereich dann auch mal so, dass man eher Geld einsparen kann usw. Und wenn man das schafft da eine Allgemein grösseres Bewusstsein zu schaffen was so der Input und der Output in der Bildung ist dann glaube ich schon, dass man eine grössere Akzeptanz hat für hohe Bildungsausgaben.**

Das war ja auch ein Punkt von SRF Data. Sie wollten eine grössere Transparenz schaffen bei den Universitäten und sind ja da sehr fest auf Widerstand gestossen. Die Universitäten haben sich dagegen gewehrt. Dies kann durchaus eine Chance sein diese Vergleichbarkeit zwischen den Universitäten oder Zwischen den Schulen mit Übertrittsdaten zur Matura usw. Dies kann

eine sehr gute Chance sein. Findest du das kann man auch kritisch beurteilen? Sei das aus Datenschutzgründen oder aber aus Gründen der Spaltung der Gesellschaft. Z.B. sagen wir das Schule X eine schlechtere Übertrittsquote hat als Schule Y. Dann könnte es ja durchaus sein, dass sich Leute die es sich leisten können, aus Mietzintechnischen oder Steuertechnischen Gründen, in die Gemeinde zu ziehen wo sich die Schule Y befindet. Somit würde den sozial schwächeren nichts anderes übrig bleiben als in der Gemeinde zu bleiben wo sich die Schule X befindet und auch diese Schule zu besuchen. Im Extremfall könnte dies ja zu Zuständen wie in zum Teil in den USA führen, wo einige Schulen sehr schlecht sind.

**Das eine ist sicher der Datenschutz. Man sollte nicht irgendwie Daten veröffentlichen die Rückschluss auf eine einzelne Person zulassen. Aber das ist ja auch schon die Idee von Open Data, das die Daten nicht auf einzelne Personen zurückzuführen sind. Der andere Punkt ist man die Schulen versuchen sollte zu verbessern, transparenter zu machen und vielleicht auch unter Druck zu setzen, sei es von den Gemeinden. Das ist glaube ich schon eher der richtige Weg. Ich sehe das nicht so sehr realistisch, dass dann massenweise Leute umziehen wegen den Schulen. Es ist ja sicherlich jetzt schon so, dass in den sozial besser gestellten Gemeinden in der Regel besser sind als andere. Ich glaube nicht, dass sich da wahnsinnig viel verändern würde.**

Vielleicht noch eine abschliessende Frage. Findest du, dass bei den Bildungsdaten noch viel Potential wäre zum noch viel mehr zu veröffentlichen als das man das jetzt schon macht? Was wäre für dich von seitens Behörden wünschenswert? Was würdest du dir noch für Daten wünschen? Wünschst du dir eine bessere Qualität?

**Ich glaube jetzt in den vorhandenen Portalen gibt es recht wenig was echtes Open Data ist. Es gibt sicher sehr viele Punkte wo man halt Daten veröffentlichen könnte weil sie halt eh schon sowieso existieren. Dies ist zwar nicht so der Fall in der Schweiz aber z.B. die Zustände der Schulgebäuden wäre noch interessant. Und was ich auch noch span-**

**nend finden würde wären die Lehrpläne. Was sind eigentlich so die Inhalte die vermittelt werden und wo sind die Unterschiede. Dies könnte noch Erkenntnisse liefern und die Vergleichbarkeit erleichtern.**

Das wäre eigentlich alles was ich soweit fragen wollte. Ich denke falls ich weitere Fragen habe kann ich dir eine Mail schreiben? Vielen Dank und dir noch einen schönen Tag.

-Ende-



---

## **Interview with Mr. Kipfer, School Administration Wallisellen (29.09.2016)**

Können sie kurz erzählen wer sie sind und was sie machen?

**Ich bin der Leiter der Schulverwaltung hier von den Schulen in Wallisellen. Das heisst von den Volksschulen sprich vom Kindergarten bis zur 9. Klasse. Ebenfalls zu den Schulen gehören die Familienunterstützenden Strukturen wie z.B. Kinderhorte, Tagesstätten etc. Das ist das was wir anbieten. Wir haben einerseits die pädagogische Seite und andererseits die betriebswirtschaftliche Seite. Ich decke die betriebswirtschaftliche Seite ab. Dazu gehört das ganze administrative Wesen dazu.**

Sie als Datenquelle, was für Daten geben sie weiter an die Bildungsstatistik?

**Wir bekommen einen ganz klaren Auftrag von der Bildungsdirektion (Bildungsstatistik) gewisse Daten zu liefern. Da werden die Schülerdaten geliefert. Schülerbestand, Sonderschulquoten etc. Es sind sehr viele Angaben welche sie diesbezüglich möchten (BiSta). Wir liefern natürlich sehr viele Daten. Dadurch ist es natürlich nicht positiv zu hören, dass damit nicht wirklich viel gemacht wird. Die Eckhaus AG geht natürlich direkt die Daten holen um uns dann den Service zu bieten. Deshalb habe ich das Gefühl es ist nicht nur ein Leerlauf. Das wesentlichste Element der Bildungsstatistik ist die Ressourcensprechung. Der Kanton gibt den Gemeinden vor, wie viele Lehrpersonen da angestellt werden dürfen. Dies ist auf Basis der Schülerzahlen, welche über die Bildungsstatistik gemessen werden.**

Was genau verlangt das BiSta von Ihnen?

**In der Regel sind das Tools, welche sie uns zur Verfügung stellen, wo wir dann die Daten einlesen und liefern.**

Stellen sie auch hier von den Schulen irgendwas auf ihrer Homepage online?

**Wir veröffentlichen höchstens die Schülerzahlenentwicklung. Weiter gehen wir nicht. Dies wird in Form einer Grafik veröffentlicht. Es wird auch als Spreadsheet online gestellt mit allen Daten von den letzten vielleicht 10 Jahren.**

Haben sie Anhaltspunkte wer ihre Daten anschaut? Sind das Eltern die Daten anschauen oder aber andere Zielgruppen?

**Die Nutzerzahlen sind überblickbar aber wer das diese Personen sind wissen wir nicht. Aber dem Feedback nach zu urteilen sind das mehrheitlich interne Leute.**

Werden die Schülerzahlen hier erhoben?

**Wir haben eine Software, eine Verwaltungssoftware, wo all diese Daten erfasst und verwaltet werden. Aus dieser Software können wir dann die Daten ziehen. Dies wird zentral von den Schulverwaltungen aufbereitet.**

Weiss man hier wie gross der Aufwand ist?

**Das müsste man runterbrechen auf die Aufgaben. Das sind sicher ein Pensum von 5% von einer Stelle.**

Haben sie schon einmal Feedback bezüglich diesen Daten bekommen von Drittpersonen?

**Es ist mehr in einem Gesamtkontext zu betrachten. Da kann man das von der Eckhaus AG nehmen. Die Schulraumplanung. Dies ist entscheidend. Die Entwicklung der Kinderzahlen werden hier sehr stark beobachtet von allen Interessensvertretern. Wann welcher Schulraum geplant und gebaut wird etc. und da bekommt man dann natürlich Feedback. Für Eltern etc. ist es eventuell besser sich direkt über das**

**BiSta zu informieren da diese für mehrere Gemeinden die Daten haben. Diese Vergleiche könnte man aber auch über Gemeindevergleich anstellen. Gemeindevergleiche.zh.ch.**

Was ist ihre Motivation und ihr Ziel der Veröffentlichung dieser Daten auf ihrer Homepage?

**Dies ist ganz klar die Transparenz für den Bürger zu schaffen. Dies ist das Öffentlichkeitsprinzip der Daten. Dies ist aber beschränkt da der Aufwand gross ist. Zum anderen werden diese Daten sowieso zentral an den Kanton abgeliefert und dort veröffentlicht.**

Was wäre aus ihrer Sicht wünschenswert was mit diesen Daten weiter geschieht welche sie online zur Verfügung stellen?

**Wenn sich Leute objektiv informieren. Informationen beschaffen natürlich auch für Beurteilungen von Situationen. Es geht ja mehr um die kritischen Zeitgeister, dass sich diese im Vorfeld einer Stellungnahme die Informationen beschaffen und daraus gewisse Schlussfolgerungen ziehen können.**

Was würde man eigentlich brauchen um die Arbeit dadurch noch mehr zu würdigen, dass man sagen könnte ja diese Daten braucht man online, die Leute machen dies und das damit?

**Die, welche wirklich was damit machen müssen, die haben diese Daten auch wenn diese nicht ganz Transparent sind, zur Verfügung. Der Punkt ist das die mehr oder weniger Rohdaten welche man zur Verfügung stellen und die Interpretation von diesen. Das ist die hohe Kunst. Das ist sehr wichtig. Was macht man mit diesen Daten?**

Haben sie Kommentare welche sie zu ihren Daten mit veröffentlichen?

**Wir veröffentlichen gar nicht so tiefgehende Daten. Wenn wir solche Daten veröffentlichen, dann ist das bezüglich eines konkreten Vorhabens. Z.B. für die Schulraumplanung. Dann können wir z.B. sagen aufgrund der Entwicklung der Schülerzahlen etc. brauchen wir ein neues Schulhaus etc. Genau da kommt die Eckhaus AG ins Spiel, welche mit diesen Daten eine Trendanalyse macht und auf Engpässe aufmerksam macht. Danach kommt der politische Prozess. Es kommen Gegenspieler welche das ganze kritisch hinterfragen etc. Da geht es halt um die Interpretation, dass man das richtig versteht. Da braucht es das mündliche Wort oder den Zeitungsartikel der erläutert was da alles für Aspekte berücksichtigt wurden. Deshalb ist es schwierig und deshalb stellen wir nicht alle Rohdaten ins Internet und sind dadurch nicht so transparent. Da könnte man noch mehr Aufwand betreiben aber das machen wir momentan nicht.**

Wird die App Eckdaten von der Eckhaus AG von euch genutzt?

**Nein diese App wird von uns nicht benutzt.**

Was war der ausschlaggebende Punkt dass sie Massnahmen ergriffen haben im Bereich Schulplanung?

**Schulraumplanung ist ein kontinuierlicher Prozess welcher über Jahrzehnte hin geführt wird. Die Behörden (Schulpflege) sind beauftragt Schulraum zur Verfügung zu stellen und die Situation im Auge zu behalten. Dies macht man dann oft mit externen Partnern (in diesem Fall die Eckhaus AG).**

Was hatten sie für einen Effekt bei der Schulraumplanung der Eckhaus AG. Was war ihr Nutzen davon im Vergleich wenn man die Eckhaus AG nicht gehabt hätte?

**Der Nutzen war das uns die Basis die Datenbasis gefehlt hätte um diese Planung durchzuführen. Die Analyse die wir machen können ist nicht**

**so genau wie die von der Eckhaus AG. Sie haben die bessere Übersicht. Wir sehen hier nur die Gemeinde bzw. Schulsicht.**

Dann ist das so zu verstehen, dass dank den Daten welche die Eckhaus AG geliefert hat, ihre Infrastruktur optimiert werden konnte? Und wie wurde das vorher gemacht?

**Genau. Dahatte die Eckhaus AG positiven Einfluss darauf. Vorher wurde das von einer anderen Firma gemacht. (Wüst & Partner).**

Hat diese Firma die Berechnungen auf ähnliche Arten gemacht?

**Ich kann ihnen diese Frage nicht beantworten. Ich war damals nicht dabei. Aber grundsätzlich haben sie auch die gleichen Aspekte berücksichtigt. Jedoch ist sicherlich die Gewichtung anders gewesen.**

Für mich ist es wichtig zu wissen, dass diese Daten für die Schulraumplanung zur Verfügung gestellt werden und somit für die Öffentlichkeit ein Nutzen daraus resultieren kann. Man könnte diese Analysen ja auch auf kantonaler Ebene auch selber machen. Aber dies wäre dann auch ein andere politischer Entscheid.

**Wie sieht es aus wenn man mit der Schulraumplanung daneben liegt? Ist es schon einmal so vorgekommen, dass man dann Notlösungen finden musste?**

Ja das ist schon so vorgekommen. Wir mussten Not-Container aufstellen. Was aber nicht unbedingt heissen muss, dass man die Schulraumplanung verschlafen hat. Sondern aber dass es auch politischen Widerstand gegen die Bauvorhaben gegeben haben könnte.

**Das wären dann ja auch verhinderbare Kosten?**

Das sind natürlich zusätzliche Kosten. Eine genaue Schulraumplanung hilft natürlich extrem um Kosten zu sparen. Im Schulwesen sind die zwei grossen Kostenpunkte das Personalwesen und die Infrastruktur. Deshalb ist es natürlich schon gut, wenn man da genaue Prognosen hat.

**Vielen Dank Herr Kipfer. Ich werde Ihnen nun kurz noch erklären wie ich das Ganze versuche zu bewerten.**

-Ende-

---

## **Interview with Mr. Schalit, Education Statistics of the Canton of Zurich (29.09.2016)**

Können sie kurz erklären wer sie sind und was ihre Arbeit ist?

**Ich arbeite hier als wissenschaftlicher Mitarbeiter und Statistiker. Und bin in meiner Funktionen in verschiedenen Projekten involviert. Erstens bin ich in Auswertungen, komplexeren Auswertungen und Visualisierungen involviert, die wir vom Kanton her machen. Dann gibt es noch ein Projekt mit diesen Daten, den Internetauftritt und Open Data. Wie können wir diese Daten in einer Benutzerfreundlichen, Bürgernahen oder aber auch für Journalisten zur Verfügung stellen. Wir veröffentlichen schon relativ viel aber es ist noch zu wenig verknüpft und wir haben auch noch kein Datawarehouse oder so. Dies generiert viele Anfragen bei uns. Das ist auch ein gewisser Teil den ich mache. Meine Tätigkeit ist im Bildungsmonitor drin und in der Bildungsstatistik. Da sind es oft Anfragen oder aber Systemerweiterungen für Datensammlung aber mehr auf konzeptioneller Ebene. Auch Visualisierungen und der Internetauftritt gehören dazu.**

Diese Daten die hier gesammelt werden sind ja primär für das BfS. Sammelt man hier auch Daten die nicht weitergegeben werden sondern welche hier gebraucht und veröffentlicht werden?

**Das es überhaupt diese Bildungsstatistik Zürich gibt, war auch genau diese Motivation, dass man diese Daten selber haben und auswerten wollte. Man wollte selber Auswertungen und Analysen machen. Es gibt viele Daten welche wir zusätzlich erfassen und nicht weitergeben (ans BfS). Die ganzen Bildungsverläufe...Das BfS hat vor ein paar Jahren den persönlichen Identifikator eingeführt. Die AHV Nummer. Wir haben das schon seit 1999. Wir sind da am BfS voraus und können diese Analysen auf längere Zeitspannen machen.**

Was veröffentlichen sie für Daten hier? Was wird im Bereich Bildung online gestellt?

**Sehr viel wird hier auf das Netz geladen. Grundsätzlich natürlich das welches Erhoben wird. Da wird aber immer geschaut, dass man nicht den Datenschutz verletzt. Man schaut auch was Gesellschaftlich relevante Themen sind. Eine grosse Dienstleistung ist die Anfragenbearbeitung. Man kann sehr vieles Anfragen. Wir schauen, dass wir den gesellschaftlichen Themen gerecht werden können.**

Ich habe auch schon auf ihrer Homepage recherchiert und bin auf sehr viele Datensätze gestossen.

**Aber zwischen den Daten und dem Verständnis der Daten besteht ein Gap. Die Daten sind reicher an Information und wertvoller als man im Moment das Gefühl hat.**

Hat man zwischendurch Interessenskonflikte mit dem Datenschutz? Zum Beispiel mit den Datensätzen wo auch noch die AHV Nummer dazu erhoben wird.

**Das ist natürlich klar, personenbezogene Daten geben wir natürlich gar nicht heraus. Gruppen kleiner als fünf, was auch noch sehr heikel ist sind so die sonderpädagogischen Massnahmen etc. Wir sind meines Wissens noch nie in so eine Situation in so einen Konflikt geraten. Wir sind nahe an dem Schulumfeld und auch dementsprechend vorsichtig. Es ist aber auch so das Journalisten gerne Daten hätten wo man dann halt sehr vorsichtig sein muss. Aber da kann man dann immer eine Lösung finden, dass es für den Journalist immer noch spannend ist und der Datenschutz immer noch gewährt ist. Da wir aber noch nicht die richtigen Open Data veröffentlichen, sind wir noch nie in so Situationen gekommen wo man an die Grenzen des Datenschutzes ausloten musste. Aber Datenschutz wird sicher ein Thema sein. Wir beschäftigen uns mit diesem Thema. Wir ziehen auch die Datenschützer mit ein.**



Die Daten welche sie auf ihrer Homepage veröffentlichen, was für Formate werden da verwendet?

**Die Daten welche jetzt runtergeladen werden können sind im CSV Format. Im Excel. Es gibt eine Formatierung (eine Kopfzeile) etc. Diese könnte man rauskopieren. Wenn wir diese Daten auf Anfrage verschicken, dann ist das meistens auch Excel oder aber CSV.**

Die Daten werden ja von ihnen gesammelt und hier konsolidiert. Ihr sammelt sie ja einerseits für ihre Auswertungen und andererseits für das BfS. Es ist ja sicherlich sehr schwierig den Aufwand für die erfassten Daten zu bestimmen. Einerseits hat man den inneren Auftrag den man für das BfS macht und andererseits die eigenen Interessen und Veröffentlichungen. Dadurch ist die Linie zwischen Open Data und innere Auftrag sicher sehr verschwommen. Aber kann man nichtsdestotrotz den Aufwand bestimmen für die Erhebungen, Datenbereinigung oder aber Datenbewirtschaftung?

**Es ist sehr schwierig da die Grenzlinie zu ziehen. Ich bin da auch die falsche Person und habe diese Informationen nicht. Aber ich denke das Open Data eher ein Zusatznutzen sein sollte zu dem normalen Aufwand den man sowieso schon für den inneren Auftrag hat. Ich denke aber auch das wir uns mit dem veröffentlichen von Daten viel Arbeit ersparen. Wenn man eine gute Datenbasis auf dem Web verfügbar hat, kommen auch weniger Anfragen und somit hat man weniger Arbeit. Mit einem Datawarehouse könnte es ja auch sein das man im Internet ein Login hat für die Gemeinden damit diese eine detailliertere Ansicht haben (mit AHV Nummer etc.), als die normalen Bürger. Open Data sollte kein wesentlicher Mehraufwand sein.**

Haben sie Zahlen zu den Nutzerzahlen der Daten die sie veröffentlichen?

**Ja wir haben natürlich die Webseitenbesucher. Aber nicht spezifisch für welche Daten. Wir können lediglich die Anfragen die wir bekommen**

**auswerten. Wir haben viele Anfragen für Daten. Das ist auch ein grosser Teil unsere Arbeit, diesen Anfragen gerecht zu werden.**

Für mich wäre es natürlich interessant zu sehen welche Datensätze wie oft gebraucht werden und so weiter.

**Wir führen eine Liste wo man sehen kann welche Daten wie oft angefragt werden, ob es intern (Verwaltung) oder aber extern ist. Es sind viele Studenten, Journalisten etc. Dies ist dann natürlich die Bestätigung, dass diese Daten genutzt werden.**

Kennen sie spezifische Beispiele was mit diesen herausgegebenen Daten gemacht wurde?

**Das sind natürlich viele Journalisten die diese Daten nutzen. Dann gibt es auch die Heimatsbücher. Diese Leute die diese Bücher erstellen greifen auch auf unsere Daten zurück. Selbstverständlich sind es auch viele Studenten.**

Sie als Daten-Herausgeber, was wäre für sie wünschenswert in der Zukunft, was mit den Daten die sie veröffentlichen gemacht wird?

**Für mich wäre es sehr wünschenswert, wenn aufgrund dieser Daten diese Themen diskutiert werden. Ich könnte mir auch Applikationen vorstellen. Zum Beispiel das Einzugsgebiet einer Mittelschule etc.**

Was sind ihre die sie mit der Veröffentlichung der Daten anstreben?

**An erster Stelle steht natürlich der Auftrag. Der Auftrag die Öffentlichkeit zu informieren, ihnen die Grundlage zu geben für Entscheidungen zu treffen etc.**

Die Anfragen welche sie bekommen, werden da Datensätze welche x mal nachgefragt werden dann auch automatisch online gestellt?

**Ja wir versuchen natürlich uns stetig zu verbessern und sind auch motiviert eine bessere Dienstleistung zu erbringen. Die Datensätze welche oft angefragt werden versuchen wir dann auch online zu stellen.**

Wieso denken sie braucht man die Bildungsdaten noch nicht so oft wie z.B. Meteodaten?

**Bildung ist ja ein öffentlicher Auftrag. Man hat ja keine grosse Auswahlmöglichkeit. Da wo man wohnt, geht man meistens auch in die Schule. Bildung ist etwas sehr komplexes und langlebiges. Eine Bildungslaufbahn geht 20 Jahre. Meteodaten sind eher kurzfristig, regnet es Morgen etc.? Der Nutzerkreis und der Nutzungsintervall ist natürlich bei Bildungsdaten sehr beschränkt.**

Vielen Dank ihnen Herr Schalit für das Interview. Das wären alle Fragen gewesen von meiner Seite.

-Ende-

---

## **Interview with Mr. Lischer, Department of the Board School of Solothurn (30.09.2016)**

**Ich bin Rolf Lischer, ich arbeite hier im Volksschulamt als wissenschaftlicher Mitarbeiter in einem Gebilde welches der Stab ist. Ich mache hier vor allem die ganze Statistik. Ein sehr grosser Teil meiner Arbeit ist es diese Daten zu generieren.**

Wie ist der Kanton Solothurn aufgebaut? Gibt es nebst dem Volksschulamt noch ein Statistikamt?

**Ja es gibt ein Statistikamt, welches ca. 1.5 FTE beschäftigt. Sie machen aber keine eigenen Erhebungen sondern koordinieren alles.**

Was für Daten sammeln sie und werden diese auf ihrer Homepage veröffentlicht? Werden diese Daten auch weitergegeben?

**Grundsätzlich erheben wir die gleichen Daten wie die anderen Bildungsstatistikstellen von anderen Kantonen. Das sind die Anforderungen die der Bund stellt. Darüber hinaus gehen wir nicht. Wir erheben genau das was das BfS vorgibt plus Vor- und Nachname. Dies machen wir für unsere Plausibilisierung. (technische Gründe) Dies ist für die Statistik der Lernenden. Für die Statistik des Schulpersonals machen wir das was die anderen auch machen. Mit der Lehrvertragsstatistik, Statistik von der beruflichen Grundbildung habe ich nicht viel zu tun. Das machen meine Kollegen direkt im Berufsamt. Die haben eine Verbundlösung mit dem Kompass (Deutschschweizer Kantone und Liechtenstein). Was wir auch nicht machen ist das schweizerische Hochschulinformationssystem. Dies gibt es auch schon lange. Was wir noch zusätzlich machen, ist eine Schulabgängerbefragung in der zweitletzten Schulwoche. Wir nehmen aus den Erhebungen im Herbst diejenigen die ihr letztes Schuljahr haben. Da sind Sonderschulen, Privatschulen, Kantonsschulen dabei. Diese werden befragt was sie nach der Schule**

**machen. Dies ist die schnellste und für uns die vollständigste Befragung. Da sehen wir wie viele in eine Berufsbildung, ins Gymnasium oder in ein Praktikum gehen oder einen Überbrückung machen. Man sieht auch wer noch nichts hat für nach der Schule. Wenn alle Kantone die Daten an den Bund geliefert haben, bekommen wir Daten vom Bund und zwar alle welche vom Kindergarten bis und mit Kantonsschule die Ausbildung in Solothurn machen oder aber in Solothurn wohnen. Bis wir aber zu diesen Daten kommen geht es ca. 2 Jahre. Da sind wir viel schneller mit der Abgängerbefragung.**

Die Daten die sie sammeln, werden diese bei ihnen auf der Homepage veröffentlicht?

**Ja aber sehr bescheiden. Wir haben nur wenige Daten die wir veröffentlichen.**

Was ist der Grund wieso man so wenige von all diesen gesammelten Daten veröffentlicht?

**Diesen Vorwurf könnte man dem BsF auch machen. Die Bildungsverwaltung ist sich nicht gewöhnt zu publizieren. Dies mag bei anderen Statistikämtern anders sein. Aber hier in der Verwaltung in Solothurn ist es etwas was nicht geläufig ist. Dazu kommt, das in Solothurn früher alles mit Papier gemacht wurde. Erst in den letzten Jahren hat sich das Ganze auf elektronische Datenübermittlung geändert.**

Die Daten die sie auf dem Netz haben sind ja vor allem als Tabellen hochgeladen. Dazu hat man diese Daten ja auch noch im PDF Format verfügbar. Hat man diese auch als CSV oder aber Excel auf dem Internet?

**Nein das haben wir nicht. Was sie können, was auch gemacht wird, sind Anfragen. Diesen gehen wir natürlich gerne nach. Das einzige Problem da ist halt der Datenschutz. Wir geben jede Information weiter welche keine Rückschlüsse auf Personen ermöglicht. Das hier ist leider**

eine One Man Show. Deshalb ist hier nicht viel mehr möglich. Ich versuche möglichst viel zu standardisieren. Hier in der Verwaltung sind wir ja ein politisches Amt. Ich kann mir auch vorstellen, dass ein Statistiker an sich freier ist. Hier haben Leute schnell Angst vor Resultaten. Ich stehe immer wieder in dieser Auseinandersetzung wo die Verwaltung kontrollieren möchte was die Resultate sind. Die statistischen Datensammlungen sind Obrigkeiten Handlungen. Dies ist immer mit Macht und Einfluss verbunden. Diese Arbeit wird oft Verhindert. Da ist der Respekt vor den Daten und der Statistik wichtig.

Behindert sie der Datenschutz in ihrer Arbeit wie sie diese gerne machen würden?

Nein, gar nicht. Wenn die Einheiten gross genug sind kann man sehr gut damit arbeiten. Wenn sie zu klein sind ist die Aussagekraft sowieso fragwürdig. Ich könnte mir vieles vorstellen. Manchmal habe ich auch 2 Herzen in der Brust. Mir der AHV-Nummer könnten wir noch ganz andere Sachen verknüpfen wo sehr interessant wären. Wir haben nicht sehr viele Daten aber können damit schon relativ viel machen. Wenn wir diese jedoch noch mit anderen Aspekten verbinden könnten würde der Erklärungswert massiv steigen. Aber dann haben immer alle Angst von dem gläsernen Bürger. Die Gefahr vom Missbrauch ist natürlich da. Es kommt darauf an, dass es sauber gemacht wird.

Sie haben vorhin gesagt, dass man Angst vor den Resultaten hat. Wie darf ich das verstehen? Hat man da Angst, dass man als Verwaltung schlecht da steht?

Das klingt jetzt völlig absurd. Angenommen das Resultat ist 3. Ohne weitere Informationen, ist dieses Resultat gut oder schlecht? Für manche ist 3 gut, für andere ist 3 schlecht. Was für mich wichtig ist, dass es klar ist weshalb dies eine 3 ist. Für mich ist die Reliabilität sehr wichtig. Die Methodik und die Definition muss stimmen. Danach kommt es auf die Fragestellung an. Da ist natürlich die Ideologie ein ausschlagge-

**bender Punkt. Für verschiedene politische Parteien kann eine 3 verschiedene Bedeutungen haben.**

Wie sieht es mit Vergleichen aus? Angenommen sie haben eine Gemeinde die sehr gut ist im Bereich Bildung. Kann man diese Gemeinde, unter Berücksichtigung aller relevanten Parameter, als Benchmark nehmen?

**Das wir heute schon gemacht. Ich denke dass die Statistik, wie wir sie hier machen, lässt solche Vergleiche nicht wirklich zu. Diese Art von Analysen ist auf grosser Ebene viel aussagekräftiger.**

Bei der elektronischen Datenerhebung, wird diese von den Schulen oder den Gemeinden ausgefüllt?

**In Solothurn ist die Gemeindeautonomie sehr gross.**

Was benötigen sie für Ressourcen um diese Erhebungen zu machen?

**Ich habe eine 90% Anstellung. Ich arbeite ca. 60% von diesen 90% konkret für diese Statistiken. Die Daten bekommen wir von den Schulen bzw. von den Schulleitern. Nicht jede Schule hat ein Sekretariat. Vielerorts ist es auch die Verwaltung der Gemeinde. Ihren Aufwand kann ich nicht abschätzen. Es sind momentan etwa 100 Schulen von denen wir eine Antwort bekommen. Und manche Schulen haben einen grösseren und manche einen kleineren Aufwand.**

Was haben sie für einen Mehraufwand für die Daten welche sie auf ihre Internetseite stellen?

**Wir geben die Rohdaten an das BfS. Die exakt selben Daten werte ich dann noch selber aus und stelle diese bei uns auf die Homepage.**

Ist dies nur für die Homepage?

**Nein dies wird dann auch für allfällige Zeitungsberichte oder aber für Publikationen mit Aargau, Basel-Land und Basel-Stadt (Bildungsbericht) genutzt. Die Auswertungen sind sehr Zeitintensiv. Der grösste Teil der Arbeit den ich mache geht für die Sauberkeit der Daten drauf. Wir machen mit den Daten immer gerade die Auswertungen für uns welche wir für spannend halten und welche uns in denn Sinn kommen. Letztens habe ich eine Verlaufsanalyse gemacht. Die Parlamentarier waren schon am streiten obwohl man noch gar keine Auswertung hatte. Als ich ihnen dann diese zeigte waren sie platt, dass man diese jetzt machen kann. Die Qualität der Daten ist sehr wichtig. Das mit dem Quantifizieren des Aufwands ist sehr schwierig.**

Können sie Auswertungen machen zu welchen Datensätzen oder Themen oft Daten nachgefragt werden?

**Nein das kann ich leider nicht machen. Diese Woche hatte ich eine sehr interessante Anfrage. Da hat jemand gefragt was aus diesen Schülerinnen und Schülern wird welche ein Brückenangebot gemacht haben.**

Wird dies überhaupt erfasst?

**Ja dies wird erfasst. Ich weiss es von diesen welche im Kanton in solchen Gebilden drinstecken. Aber danach gehen diese evtl. ausser Kanton und da habe ich halt keine Daten vom Bund. Die bekomme ich erst in 2 Jahren. Dies ist dann sehr kompliziert.**

Für mich wäre es spannend gewesen zu sehen ob man einen Trend feststellen kann oder nicht.

**Nein. Wobei ich sagen muss es sind oft relativ simple Anfragen. Das Bildungswesen ist eines der grössten Ausgaben in der Schweiz. Aber das Geld das man dafür einsetzt für die Bildungsevaluation sprich die Erkenntnisse sind verschwindend klein. Da gibt es tatsächlich einen absoluten Mangel an Kultur und Interesse solche Auswertungen zu ma-**



**chen. Ich komme sehr selten hoch spannende Anfragen über. Diese geben dann aber auch sehr viel zu tun und muss diese dann aufgrund Zeitmangels ablehnen.**

Wissen sie von Dritten welche Auswertungen mit ihren Daten machen?

**Ja es gibt's schon. Hochschulen bekommen immer wieder von uns oder vom BfS Daten über für eigene Auswertungen machen zu können. Mit diesen Vereinbaren wir dann einen Datenschutzvertrag, um das Handling der Daten festzulegen. Da wird unter anderem die Aufbewahrung und die Vernichtung etc. vereinbart.**

Haben sie da auch konkrete Beispiele für so Analysen?

**Es gibt eine ganz berühmte. Diese ist von der SVP. Sie sind vom Namen her sehr wahrscheinlich türkischer Abstammung?**

Ja genau.

**Sie sind dann auch sehr wahrscheinlich Moslem?**

Ja, bin ich.

**Ich weiss leider nicht mehr ob es um Ausländer oder aber Muslimen ging. Jedenfalls hat es eine Zunahme gegeben und dann hat die SVP Polemik betrieben. Sie haben die Wachstumsraten der Muslime/Ausländer verglichen mit den schweizer Wachstumsraten in der Schweiz. Diese haben sie dann fortgesetzt. Das waren halt einfach „Deppen“ man kann es fast nicht anders sagen. Ich äussere mich ansonsten nicht zur Politik, aber wer solchen Mist macht ist nur ein Depp oder Polemiker. Ihnen fehlte noch ein Jahr. Wenn sie diese Analyse noch ein Jahr weitergezogen hätten, dann wären sie über 100% gekommen. Dies ist eine völlig absurde Auswertung. Es ging darum auszusagen, dass die Ausländer oder Muslime den Schweizern alles weg-**

**nehmen. Der allererste Wert war ja auch richtig. Dieser kam aus der Statistik. Aber danach haben sie eine lineare Wachstumsrate angewendet. Dies resultiert natürlich in einem exponentiellem Wachstum. Das ist die dümmste Weiterverwendung von Statistikdaten die ich kenne die je einer gemacht hat. Dies war halt so offensichtlich dumm, dass es sofort durchschaut wurde.**

Könnte genau dies ein Problem sein von so Auswertungen, dass man nicht genau weiss wie diese durchgeführt wurden?

**Das können sie natürlich jedes Mal nachfragen. Ich habe bei meinen Auswertungen auch schon diese Angaben dazugegeben. Das ist kein Problem. Ich gehe davon aus, dass wir hier Methodisch auch korrekt arbeiten. Das kann man nicht publizieren. Wir publizieren zwar die Auswertungen aber jedoch nicht das methodische Vorgehen dazu. Aber wenn sie kommen und fragen, gebe ich diese Informationen natürlich gerne.**

Kommen wir zurück zum Thema Daten verknüpfen. Was bringt die Zukunft? Was könnte man noch machen? Was könnte man mit den Verknüpfungen mit den AHV-Nummern machen?

**Da könnte ich mir Sachen vorstellen wo mir unheimlich wird, wo ich froh bin dass man nicht alles macht. Ich bin viel in Dänemark. Das ist eines von den nordischen Ländern wo Leute ihre AHV-Nummer kennen. Die haben auch eine Nummer wie wir die man von Geburt an bekommt. Die kennt jeder. Die braucht man da auch permanent. Hier kennt niemand seine AHV-Nummer. Überall wo diese Nummer gebracht wird, kann man diese Verknüpfungen machen. Dies ist bei uns noch sehr restriktiv eingesetzt. Dies könnte sich in Zukunft ändern. Es war schon eine grosse Herausforderung bis man die AHV-Nummer in der Bildungsstatistik brauchen durfte. Da waren richtige Kämpfe in Bern. Es ging darum dass die Schulen die Erlaubnis bekommen sollten die AHV-Nummer zu führen. Die Statistikämter hatten jedoch diese Erlaubnis**

nicht. Dies war ein Problem. Wir haben Daten bekommen welche die AHV-Nummer enthielten, durften diese aber eigentlich nicht haben und mussten diese aber wiederum weitergeben. Dies war sehr absurd. Dies hat Monatelang alles behindert dieses Modernisierungsprojekt. Ich würde es z.B. gerne sehen dass man diese Verknüpfungen mit Ausländern macht. Dann würde man evtl. sehen, dass die Gründe nicht in der Herkunft sondern aber im sozialen Stand liegen könnte. Im Moment können wir das nicht. Wir haben nur die Angabe zum Pass. Dies sagt natürlich sehr wenig aus.

Haben sie noch andere konkrete Beispiele was man noch machen könnte?

Es ist vieles abhängig von Bildung. Zum Beispiel bei pädagogischen Fördermassnahmen. Aber dies kann man nicht richtig beobachten da man keine Vergleichsgruppe hat wie es wäre wenn man keine Massnahmen gehabt hätte. Das ist alles nicht im Labor, dass sind keine Physikexperimente. Was ich noch gut fände aber hier leider nicht möglich ist allein schon von dem Arbeitsaufwand her, wäre ein Tool online zu stellen wo man selber Abfragen stellen könnte.

-Ende-

---

## List of Figures

Figure 1: Theory of Change (Dapp and Stuermer, 2015).....	13
Figure 2: Logo Canton Zurich (Bildungsstatistik Kanton Zürich, 2016) .....	28
Figure 3: Logo Canton Bern (Erziehungsdirektion Kanton Bern, 2016) .....	28
Figure 4: Logo Canton Bern (Erziehungsdirektion Kanton Bern, 2016) .....	29
Figure 5: Logo Canton Solothurn & Board of School Solothurn (Volksschulamt Kanton Solothurn 2016).....	29
Figure 6: Twitter Profile Fidel Thomet (Thomet, 2016).....	29
Figure 7: Logo SRFdata (SRFdata, 2016) .....	30
Figure 8: Logo Eckhaus AG (Eckhaus AG, 2016).....	30
Figure 9: School of Wallisellen (Schule Wallisellen, 2016).....	30
Figure 10: Eckdaten App (Eckhaus AG, 2016).....	37
Figure 11: Cooperation between swiss universities and third parties (Thomet, 2016).....	38
Figure 12: Circly of OGD Impact .....	41
Figure 13: Ecosystem of OGD value creation .....	44

---

## List of Tables

Table 1: 10 Principles of Open Data (Sunlight Foundation, 2016) .....	10
Table 2: 14 High-value categories (G8, 2013).....	26

---

## References

- Aeron-Thomas, D., Nicholls, J., Forster, S. and Andrea Westall, 2004. Social Return on Investment: Valuing what matters - Findings and recommendations from a pilot study, New Economics Foundation (NEF).
- Bildungsstatistik des Kantons Zürich, 2016. URL: [https://www.bista.zh.ch/\\_zf/zf\\_main.aspx](https://www.bista.zh.ch/_zf/zf_main.aspx) [Accessed: 2016-11-02]
- Bürgi-Schmelz, A., 2013. Wirtschaftliche Auswirkungen von Open Government Data," Schweizerisches Bundesarchiv.
- Carrara, W., Chan, W. S., Fische, S. and van Steenberg, E., 2015. Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources," European Commission.
- Charta öffentliche Statistik der Schweiz, 2012. Neuchatel: Bundesamt für Statistik.
- Dapp, M. and Stuermer, M., 2015. Schlussbericht Impact Monitoring Framework für OGD Schweiz.
- Davies, T., 2014. Open Data in Developing Countries - Emerging Insights from Phase I, The World Wide Web Foundation.
- Davies, T., Perini, F. and Alonso, J. M., 2013. Researching the emerging impacts of open data - ODDC conceptual framework, World Wide Web Foundation.
- Davies, T., Raed, Sharif, M. and Jose M. Alonso, 2015. Open Data Barometer 2015 Global Report - 2nd Edition, The World Wide Web Foundation.
- Eckhaus AG, 2016. URL: <http://www.eckhaus.ch> [Accessed: 2016-08-13]
- Emerson, J., Wachowicz, J. and Chun, S., 2000. Social Return on Investment: Exploring Aspects of Value Creation in the Nonprofit Sector, *Investor Perspectives*.
- Erziehungsdirektion Kanton Bern, 2016. URL: <http://www.erz.be.ch/erz/de/index.html> [Accessed: 2016-10-02]
- G8, 2103. G8 Open Data Charter, Policy paper.
- GitHub, 2016. URL: <https://github.com> [Accessed: 2016-10-16]
- Gurtstein, M., 2011. Open Data: Empowering the empowered of effective data use for everyone?. URL: <http://firstmonday.org/article/view/3316/2764> [Accessed: 2016-09-02]
- Harrison, T.M., Pardo, T.A., Cook, M., 2012. Creating Open Government Ecosystems: A Research and Development Agenda Future Internet 4, 900-928. doi:10.3390/fi4040900.

- Harrison, T. M., et al., 2012a. Open government and e-government: Democratic challenges from a public value perspective. In: *Information Polity* 17, 2, S. 83-97.
- Hilgers, D., 2012. Open Government: Theoretische Bezüge und konzeptionelle Grundlagen einer neuen Entwicklung in Staat und öffentlichen Verwaltungen, in: *Zeitschrift für Betriebswirtschaft* 82, 6, S. 631-660.
- Janssen, M., Charalabidis, Y. and Zuiderwijk, A., 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government, *Information Systems Management*, vol. 29, no. 4, pp. 258–268.
- Jetzek, T., Avital, M., Bjorn-Andersen, N., 2012. The Value of Open Government Data: A Strategic Analysis Framework. SIG eGovernment pre-ICIS Workshop, Orlando 2012.
- Kassen, M., 2013. A promising phenomenon of open data: A case study of the Chicago open data project, *Government Information Quarterly*, vol. 30, no. 4, pp. 508–513.
- Klessmann, J., Denker, P., Schieferdecker, I. and Schulz, S. E., 2012. Open Government Data Deutschland - Eine Studie zu Open Government in Deutschland im Auftrag des Bundesministerium des Innern Auftraggeber.
- Lee, G., Kwak, Y., H., 2012. An Open Government Maturity Model for social media-based public engagement, in: *Government Information Quarterly* 29, 4, S. 492-503.
- Linders, D., 2012. From e-government to we-government: Defining a typology for citizen coproduction in the age of social media, *Government Information Quarterly*, vol. 29, no. 4, pp. 446–454.
- Manyika, J., Chui, M., Groves, P., Farrell, D., Kuiken, S. V. and Doshi, E. A., 2013. Open data: Unlocking innovation and performance with liquid information.
- McDermott, P., 2010. Building open government, in: *Government Information Quarterly* 27, 4, S. 401-413.
- Nicholls, J., Lawlor, E., Neitzert, E. and Goodspeed, T., 2009. A guide to social return on investment, Office of the Third Sector, The Cabinet Office, London.
- Obama, B., 2014. Memorandum for the Heads of Executive Departments and Agencies, SUBJECT: Transparency and Open Government URL: [http://www.whitehouse.gov/the\\_press\\_office/TransparencyandOpenGovernment/](http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment/) [Accessed: 2016-19-14]
- OECD, 2015. Data-Driven Innovation. OECD Publishing, Paris.

- Olsen, S. and Lingane, A., 2003. Social Return on Investment: Standard Guidelines, Center for Responsible Business UC Berkeley.
- Open Definition, 2015. Open Definition, URL: <http://opendefinition.org/od/> [Accessed: 2016-09-09]
- Right2INFO.org, 2012. Access to information laws: Overview and statutory goals, URL: <http://www.right2info.org/access-to-information-laws> [Accessed: 2015-01-03]
- SAS Institute Inc., 2016. URL: [http://www.sas.com/en\\_us/insights/big-data/data-visualization.html](http://www.sas.com/en_us/insights/big-data/data-visualization.html) [Accessed: 2016-11-02]
- Schule Wallisellen, 2016. URL: <http://www.schule.wallisellen.ch> [Accessed: 2016-09-26]
- SRFdata, 2016. URL: <http://www.srf.ch/news/srf-data> [Accessed: 2016-09-25]
- Stuermer, M., Ritz, A., 2014. Public Governance durch Open Government: Zwei sich ergänzende Ansätze für die staatliche Aufgabenerfüllung der Zukunft, in: SGVW Jahrbuch 2014, Kompetenzzentrum Public Management Universität Bern
- Sunlight Foundation, 2016. URL: <http://sunlightfoundation.com/policy/documents/ten-open-data-principles/> [Accessed: 2016-10-13]
- Thomet, F., 2016. URL: <https://twitter.com/fidelthomet> [Accessed: 2016-08-30]
- Ubaldi, B., 2014. Open government data: implementing to deliver results Highlights from OECD data collection and analysis, Organisation for Economic Co-operation and Development OECD.
- Ubaldi, B., 2013. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives, Organisation for Economic Co-operation and Development OECD, OECD Publishing 22.
- Volksschulamt Kanton Solothurn, 2016. URL: <https://www.so.ch/verwaltung/departement-fuer-bildung-und-kultur/volksschulamt/> [Accessed: 2016-08-05]
- von Lucke, J., 2014. Open Government Collaboration: Offene Formen der Zusammenarbeit beim Regieren und Verwalten, in: von Lucke, J. (Hrsg.), Das offene Regierungs- und Verwaltungshandeln und seine Perspektiven im Wahljahr 2013: Open Government, Open Government Collaboration und die Bundestagswahl 2013, Band 6, Berlin: epubli, S. 19-78.
- Zuiderwijk, A. and Janssen, M., 2014. Open data policies, their implementation and impact: A framework for comparison, *Government Information Quarterly*, vol. 31, no. 1, pp. 17–29.



## Statement of autonomous and independent work

Ich erkläre hiermit, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.

Bern, 30.12.2016

  
Taner Aydin

## **Publication of the Thesis**

Hereby I give permission to make my thesis publically available at the Library of Information Systems.

Bern, 30.12.2016

  
Taner Aydin