Integrating Linked Open Data for improved Social Sustainability Risk Management in Supply Chains

Andreas Thöni

Institute of Software Technology and Interactive Systems
Vienna University of Technology
Favoritenstrasse 9-11/188
A-1040 Vienna
andreas.thoeni@tuwien.ac.at

Abstract: Social sustainability risk management in supply chains has to include a multitude of factors in a broad environment. Beside, Linked open Data (LoD) has become a major element of the semantic web being potentially relevant for the topic. This paper analyses LoD's usefulness in this specific context differentiating two dimensions, methods of LoD and data published as LoD. While the usefulness of the former can be conceptually highlighted for disambiguation, publication and retrieval, the latter has been empirically analyzed for 924 LoD datasets at the LoD registry CKAN (http://datahub.io/). 4% were identified to contain relevant data for the supply chain risk management context. In a deep dive of corruption analysis data, trust and infrequent updates were identified as obstacles for LoD usage.

1 Introduction

In recent years, sustainability has become a major business topic also influencing supply chain management (SCM) (e.g. Seuring and Müller [SM08]). Different definitions of sustainability exist. Famous are the three elements of sustainability – economic, environmental and social – developed with the Brundtland report [Un87]. The economic perspective is already an integrated part of research. While companies are increasingly aware of environmental topics, the understanding of social issues is lagging [KV12]. Moreover, also the scope addressed in academia is limited. In their literature review Seuring and Müller [SM08] see a lack of research in the domain of socially sustainable SCM. Especially in international supply chains, social sustainability can be an important topic as social issues such as child labor may result in significant reputational damage to a company ([Ts12]; [KV12]).

Therefore, a company's risk management approach has also to include social sustainability factors [CR08]. Standards, such as certification guidelines, define a multitude of elements relevant for social sustainability risk management (see e.g. the GRI standard [Gl11]). However, these have to be evaluated in supply chains which can include hun-

dreds of suppliers in a first, second or further tier today [KM12]. Given this broad and complex issue, an exhaustive coverage of risks seems difficult and in-depth or in-person auditing approaches are often cost-constrained [KW12]. Nevertheless, data availability is a key for supply chain risk management [GSK12]. As information technology provides methods for improved data access and data sharing in the supply chain context (see e.g. Ghadge et al. [Ga10]), all options for leverage should be studied.

It has been suggested that an extended supply chain risk analysis covers data from diverse internal and external sources in (near) real-time [Ho11a]. In a supply chain context this includes different external organizations and suppliers. Moreover, international supply chains tend to change over time, making updates in risk analyses important. Furthermore, different stakeholders are potentially interested in the data [KM12], thus, building a case for an easy way of sharing. In recent years the Semantic Web, a machine-readable version of the current web [Be01], has gained importance in academic research. By fostering the amount of web-enabled data, Linked open Data (LoD) as a special field has nourished the content necessary for a semantic and machine-readable form of the web [BHB09]. Semantic representations combined with online and interlinked accessibility are important features of LoD [He08]. Hence, LoD allows for easy publication and integration of a diverse set of different data based on a semantic context. Therefore, LoD shows the potential of fulfilling the requirements outlined above. Moreover, it potentially contains data particularly relevant for the social sustainability context.

This paper tries to contribute to the existing stack of knowledge by analyzing LoD's usefulness in the context of social sustainability risks in supply chains from a company's perspective. Chapter 2 briefly reviews LoD, whereas chapter 3 discusses options how to leverage LoD. Chapter 4 empirically focuses on the specific aspect of how LoD can be used as an information resource based on the current state of the LoD cloud as available in March 2013. For this purpose a content analysis of LoD dataset descriptions has been applied. Chapter 5 discusses the results and chapter 6 will draw a conclusion.

2 Related work

With regard to the paper's length only the most relevant literature has been included. A detailed literature review is currently in progress. In general, LoD is a rather new concept. Only in 2006 Tim Berners-Lee defined the set of rules that together describe LoD [BHB09]. Since then, the size of the LoD has grown significantly. CKAN (http://datahub.io/; [Th13]) is a central database where dataset metadata of linked open data can be published and which has also been the source for the known LoD diagram (by Bizer et al. [BJC11]). For the five years until end 2012 the number of datasets at CKAN grew with a compound annual growth rate of roughly 109% [Th13].

2.1 Linked open data and supply chain risk analysis

According to Tim Berners-Lee LoD has four key rules [Be06]. First, LoD uses unique identifiers called URIs that name entities. Second, the HTTP protocol allows dereferenc-

ing these URIs and retrieving further details. Third, entities can be described in more detail based on the semantic web standards RDF and SPARQL. Fourth, links between data (referenced by URIs) form a web of interlinked datasets. Beside, the web of data relies on a set of existing vocabularies to relate things and give semantics. These can be extended by new vocabularies when needed and the respective schema can be made public by the very same principles [BHB09]. Altogether, LoD allows for an easy way of semantic-based data integration from different sources that is needed for risk analysis.

In the supply chain risk management context, Hofman ([Ho11a], [Ho11b]) proposed that companies use LoD principles to publish supply chain information needed by government custom authorities for their risk analysis. Hofman's work mainly focused on supply chain internal data to be used by custom authorities, thus, he advocated for an authority-driven crawling of LoD data along different supply chains. He further focused on the software architecture, i.e. the way of publishing and consuming data. More general, Hulstijn et al. applied LoD to supply chains in order to exchange trade documents by "publishing and connecting structured data" [Hu12, p. 352]. In a context of sustainability, Ciroth [Ci12] proposed to use the semantic web for life cycle assessments, especially for "collaboration and data sharing" [Ci12, p. 154]. Further it has been proposed to use linked data principles for publishing green house gas data [Ge10]. To the best of the author's knowledge, no paper has addressed the connection between socially sustainable supply chain management, supply chain risk management and LoD. Therefore, the next section will discuss which dimensions of LoD could be interesting for further usage.

2.2 Linked open Data as a Method and as a Resource

When dealing with LoD, a differentiation between the concept behind and the content represented with LoD is needed. Key method of LoD is the linkage between different data entities and databases based on RDF. It uses triples identified by an URI [Be01]. As a content resource LoD has already gained a significant size. According to a statistic that counts all elements within the LoD cloud (a subset of all LoD), it contained 31.6 billion triples in September 2011 [BJC11]. For comparison: the LoD version of the English Wikipedia counts "only" 0.4 billion triples [DB13]. Three quarters of the data within the LoD cloud is either government, geographic or cross-domain data [BJC11].

3 Linked open Data for social sustainability risk analysis

From a business view the key question is how LoD can be leveraged to improve social sustainability risk management in supply chains. On the one hand, the concept of LoD introduces methods whose application may facilitate and improve risk management activities (using 'LoD as a method'). On the other hand, LoD provides a large amount of accessible data that might be included in risk management processes (using 'LoD as an information resource'). Hence, Table 1 distinguishes the key concepts of LoD and presents options, how each of them may be leveraged. Each concept of LoD could be applied and will be discussed below.

LoD components	Key concepts/elements	Possibilities within context
LoD as a method (see 3.1)	Unique Resource Identifiers (URIs)	Disambiguation of entities (companies, countries, etc.)
	HTTP protocol	Well-used standard – reuse in combination with other elements
	Semantic web standards	Re-use in publication of sustainability data or for data exchange
	External linking	Representation of supply chain and automated retrieval of supply chain relevant data
LoD as an information resource (see 3.2)	Resources (subjects/objects)	Retrieval of relevant data (full triples needed to identify relevant information)
	Verbs	
	Literals (Integers, Strings, …)	

Table 1: Options for LoD usage for social sustainability risk analysis in supply chains (own repr.)

3.1 LoD as a method

Disambiguation of concepts

As described above, LoD builds on the principle of unique identification of concepts in the world. Within a certain namespace, each concept can be uniquely addressed using its identifier (URI). In the Semantic Web the ontological metalanguage OWL is used to define vocabularies of concepts and their semantic relations. Known vocabularies defined in OWL [W304] are important for LoD, because they can be used to relate resources in LoD in a meaningful way [BHB09]. Moreover, OWL also allows to state that two concepts with different URIs are actually the same. Although in currently available LoD these links are often missing, they can be necessary in the context of social sustainability and risk management. Concepts like suppliers, geographic locations or sectors might need disambiguation. An URI-based approach can help for a unique identification (see also O' Riain et al. [OCH12]). Furthermore, relations to known concepts can help for further clarification.

Disambiguation gets especially easy if a central storage of identifiers exists. The DBpedia database currently consists of 3.8 million concepts of which 62% percent are classified with an established ontology [DB13]. Many different LoD resources relate to DBpedia and its large pool of objects and only some sets in the LoD cloud show comparable sizes [CJ11]. Even though DBpedia has a strong potential for disambiguation, also some smaller datasets could be of particular relevance. These are for example the LoD version of the Common Procurement Vocabulary (CPV; [Ro13a]), the GeoNames database [Ge13] or the International Standard Industrial Classification [Ro13b].

Publication and dissemination

LoD principles that ease the exchange of information can be used to publish risk related information [Ho11b]. In that sense LoD has several advantages. First of all, the RDF standard [Rd04] has been widely accepted as a way of knowledge representation. Using the standard to store publicly available data allows for an easy integration into other applications that need to use the data. Also data consumers can integrate the data more easily into their system environments [OCH12]. The SPARQL query language [W308],

which is applied for LoD, allows for queries directly on the dataset. Moreover, it allows combining multiple datasets in one query in a non-document-centric way [GG09]. Thereby, LoD does not need to conform to a particular schema [He08]. The context setting as well as the embedding of LoD into vocabularies makes it easier to use for any consumer as the information provided can be better understood and compared.

In the sustainability supply chain environment, various company-specific sustainability information can be made publicly retrievable. This can e.g. be sustainability reports such as GRI ratings, background data or product-specific carbon footprints or certificates.

Automated Supply Chain retrieval

Generally, LoD should always link to other datasets [Ha12]. These links allow for the discovery of related sets and define a context. Hofman ([Ho11a], [Ho11b]) proposed to use the approach of publishing data linked to other sets in the supply chain, especially in order to exchange data between companies and government agencies. The basic idea can be seen in a LoD-based model of the supply chain. If data sets in a supply chain point to other supply chain partners' sets, the whole supply chain can be retrieved by following linkages. This allows including additional data. An example is given in Figure 1.

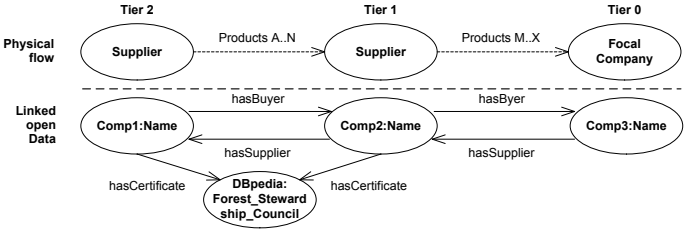

Figure 1: Possible LoD approach for supply chain modeling (own representation)

The online links are related to the physical flow of goods. Hofman [Ho11b] saw an important use case in automated retrieval of information from the whole supply chain by crawling along links. This thought can be extended to focal companies, who have incomplete supplier data after the first tier. Publication needs and linkages with unique references could allow gathering a more complete picture in less time. Moreover, it could be augmented with extra external LoD (also García and Gil [GG09]). Especially large focal companies could use their dominance to push the dissemination in a LoD. However, data sensitivity constrains the publication of sensitive sourcing or customer data. Moreover, conflict resolution procedures will be necessary to ensure data quality.

3.2 LoD as an information resource

Central to the "web of data" is the availability of machine-readable data. Therefore, also the possibility of using it as an information resource needs to be considered. Generally, LoD can be an internal information resource in a supply chain when provided through supply chain members or an external when provided by third parties to the supply chain. Hofman [Ho11a] suggested combining information from the internal logistic domain

with information from external resources to assess the risk of a specific supply chain not focusing on sustainability. From the perspective of a focal supply chain company, multiple strategies may be possible to leverage LoD as an information resource for social sustainability risk management. Two dimensions need to be considered (see Figure 2): (a) where is the data's origin? (b) is the data already available? For supply chain internal data, a focal company can have significant influence. In the case of internally available data, the company's IT systems can fetch the information by crawling across linkages and integrate it into risk algorithms. If internal data is not available, it can push partners to publish necessary data. External data can be directly collected and incorporated. If external data is not available, large companies could make publication suggestions.

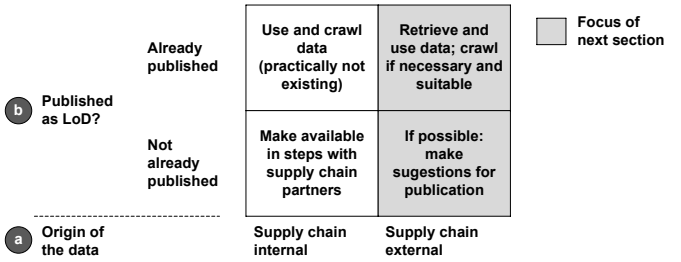

Figure 2: Focal companies strategies if considering LoD as an information resource (own repr.)

To better analyze the possibilities of using LoD as information resource, the next section will evaluate the currently available LoD if it is already "ready to use".

4 Evaluation of LoD as an external information resource

Methodologically, the evaluation follows the six-step process depicted in Figure 3. More details to the different steps are provided in the sections below.

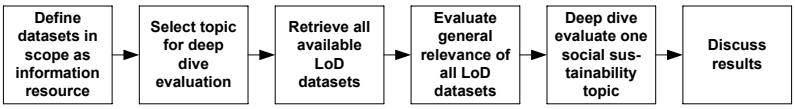

Figure 3: Evaluation process of LoD as an information resource (own, self-explanatory repr.)

4.1 Scope definition and deep dive topic

To define the datasets in scope for social sustainability risk management, a consistent view on social risk is necessary. A widely accepted sustainability standard is the Global Reporting Initiative (GRI) available in version 3.1 [Gl11]. It has become the de-facto standard for sustainability reporting [Kp11]. At a higher level, it splits social factors into four areas (labor practices and decent work, human rights, society and product responsibility) with 25 sub-dimensions. Based on a content analysis of the LoD dataset descriptions datasets were identified that with best judgment contain information relevant for a

GRI-dimensions-based social sustainability supply chain risk assessment. These were marked relevant and in scope. The datasets included geographical, sector or company-specific content that can be related to a certain supplier or datasets that possibly include secondary input such as economical data explaining causes of elements above.

Issues typically considered as especially strong are discrimination/harassment/abuse, child labor, forced or compulsory work, legal compliance and corruption. Corruption is a topic particularly discussed and on the agenda top worldwide. In 22 out of 144 global economies [Wo13], corruption is the major obstacle to business. Due to its importance and as part of GRI, this paper uses corruption as the deep dive to evaluate the extent to which LoD is already useful for retrieving social sustainability risk management data.

Corruption can be understood as "the misuse of public office for private gains" [Tr00, p. 399]. For the purpose of risk assessment, corruption could be either measured directly (i.e. corruption as dependent variable) or with the help of secondary indicators that have been published in academia (i.e. independent variables proved to be significantly correlated). Typically corruption is very hard to measure. An often used indicator is the subjective measure of a "perceived level of corruption" (e.g. You and Khagram [YK05]). Indirect reasons for corruption have been analyzed by multiple authors. Table 2 lists some significant indicators that can be used for risk analysis. Although not complete, it also includes recently highlighted indicators of corruption [AA11].

Secondary indicators	Variables suggested significant	Scope	Source
British history	Former British colony	World	[Tr00]
Democracy	Long exposure to democracy	World	[Tr00]
Economic development	GDP per capita	World	[Tr00]
	GDP per capita in PPS	EU-25	[AA11]
	GDP per capita	World	[Pa02]
Economic freedom	Economic freedom score	EU-25	[AA11]
Federal state	Unitary vs. Federalist	World	[Tr00]
Income inequality	Gini coefficient	EU-25	[AA11]
	Household net expenditure based GINI	World	[YK05]
Inflation	Harmonized Indices of Consumer Prices	EU-25	[AA11]
	Inflation rate	World	[Pa02]
Protestantism	Percent Protestant	World	[Tr00]

Table 2: Secondary indicators for corruption based on four studies (own representation)

Hence, the deep dive evaluation focused on retrieving data for these direct and indirect indicators in the LoD datasets identified as relevant for the use case. Conclusions were drawn on a more generalized level to minimize the impact of the indicator selection.

4.2 Retrieval and evaluation of datasets

LoD datasets were retrieved from http://datahub.io/ (CKAN; briefly described above). The database allows for a publication of data and allows tagging and grouping for filtering specific datasets. LoD is commonly tagged with "lod". Specifically, datasets from the LoD cloud (thus, also included in the LoD diagram) are part of a special group named "Linking Open Data Cloud". Except for two datasets in the group, all are also

tagged with “lod”. Nevertheless, to analyze all sets that have been considered as LoD by their authors, all data tagged “lod” was analyzed.

Technically, a JSON dump of the CKAN database was retrieved at 23rd of March 2013. 924 entries tagged “lod” and 338 entries in the LoD cloud group (of which 336 also tagged “lod”) were identified. Datasets were split into two priority levels: Priority 1 datasets include direct and general information that does not need any further processing and where underlying data sources are clearly stated and can be controlled. Priority 2 datasets include e.g. local data (e.g. Great Britain specific), alternative datasets, information needing further processing (e.g. text/opinions) and uncontrollable sources and thus, were not considered fully in scope. For the deep dive of corruption, the documentations of the identified relevant datasets were screened for the indicators in Table 2. SPARQL queries were used to retrieve RDF-based test data to check the availability.

5 Results

5.1 General results

Based on the total 924 datasets 39 datasets (~4%) were identified as potentially relevant for social sustainability risk identification. Only eight datasets (~1%) were identified to contain data of stronger relevance and were classified as priority 1 datasets. Table 5 in the appendix shows the corresponding split into priority 1 and priority 2 datasets. The majority, four of the priority 1 and 19 of the priority 2 datasets, were also part of the LoD cloud group. All priority 1 datasets were considered as relevant for the deep dive topic of corruption and are discussed in the next section.

5.2 Deep dive results – data on corruption

In order to identify those priority 1 datasets that contain data relevant for social sustainability risk management in supply chains with a focus on corruption, all sets were explored in-depth. A time series with the inflation rates of countries worldwide can e.g. be a piece of relevant data. The resulting selection is presented in Table 3. All datasets in Table 3 are free of charge and can be classified as 5 star LoD [Ha12] as they link to other datasets such as DBpedia. However, several obstacles remain: All databases depend on third-parties, meaning that they are not controlled by the organization creating the source data. Moreover, the update frequency of four of the six databases seems to be unclear as it is either not stated or the actual cycle is not in line with the stated.

Finally, the datasets of Table 3 were screened for the indicators in Table 2 or close indicators. Altogether 9 direct and 15 indirect indicators for corruption were found. Table 4 in the appendix contains details. If multiple comparable indirect indicators were available from a source (e.g. GDP deflator, consumer price index, etc.), only the most relevant was selected. All data was available through SPARQL endpoints[1]. Direct indicators

[1] Eurostat datasets needed to be contacted via a 3rd party SPARQL endpoint targeting the RDF files.

measuring mainly the corruption perception described above were available from Worldbank and Transparency International. Indirect indicators covered GDP per capita, government type, GINI index, inflation rates, Democracy and economic freedom. Not covered by an indicator were British history and Protestantism. However, these can for example be retrieved mining implicit facts from text which is not the focus of this paper.

Set	Free	Update frequency	Third-party	Type[2]	Comment
World Bank Linked Data	Yes	Monthly	Yes	5 star	Last update 2012-08-10
Transparency International	Yes	No regularity	Yes	5 star	Last update 2012-10-03; 2012 data not yet included in LoD
Eurostat LoD	Yes	Monthly	Yes	5 star	Monthly newsletters with changes. Weekly runs stated in statistics
World Factbook (FU Berlin)	Yes	Unknown	Yes	5 star	Only links to foaf included
European Central Bank (ECB)	Yes	Unknown	Yes	5 star	Limited documentation of data. Last update of website stated on 2013-03-04
OECD Linked Data	Yes	Unknown	Yes	5 star	Limited documentation of data. Last update of website stated on 2013-02-10

Table 3: LoD sets with indicators useful for corruption identification (as of 23rd March 2013; own representation) [3]

Moreover, two other issues appeared: (1) Datasets varied strongly in the number of values – the smallest set of direct indicators contained 4, the largest 207 countries in 2010. (2) One set did not contain indicators that were already available at the source's website.

6 Discussion and limitations

To sum up, using LoD for social sustainability risk analysis in supply chains has advantages and disadvantages. As a method LoD can be very useful for disambiguation. Due to the availability of large LoD datasets such as DBpedia, linkages can be used to identify concepts such as countries when relating risks in supply chains. Moreover, LoD can be used for publication purposes enabling stakeholders (e.g. investors) to collect company information more easily. Therefore, also a risk assessment could happen more simply if information is published in this way. Finally, automated crawling and retrieval in supply chains (as also discussed by Hofman [Ho11a]) would allow for a more detailed retrieval of information and thus for a broader analysis of risks. Especially crawling into "deeper" (beyond first tier) levels of the supply chain would support this.

Beside, the current usefulness of LoD as an information resource was evaluated based on the actual state of registered LoD. Although LoD already reached a significant size, only around 4% of the datasets seem to be particularly relevant in the risk management context discussed. Generally, especially business, industry and sustainability specific statis-

[2] 5 star LoD is open, machine-readable, based on an open standard, and linked with other sets [Be06].
[3] http://worldbank.270a.info/; http://transparency.270a.info/.html; http://eurostat.linked-statistics.org/; http://www4.wiwiss.fu-berlin.de/factbook/; http://ecb.270a.info/; http://oecd.270a.info/

tics and data appear to be missing. The deep dive into possible data for corruption risks evaluation in supply chains showed three obstacles. First of all, many datasets are third-party-driven leading to a problem of trust in the data and limitations for usage in enterprise software. Second, also likely due to third parties, the LoD datasets seem to be infrequently updated. Third, the LoD datasets are often not very well documented.

This study has four important limitations. The LoD datasets were selected from the Data Hub (CKAN; [Th13]) – although it is considered as the important directory for LoD, further sets might be available. Further, datasets were selected based on descriptions and partly also websites. In case of ambiguities in the description or a lack of a description, relevant datasets might have been missed. Moreover, only one deep dive was conducted. For a broader picture, also other indicators or other GRI categories could be considered. Finally, future studies could include open data that is not linked to other sets.

7 Conclusion

This paper discussed how Linked open Data (LoD) can be leveraged for the assessment of social sustainability risks in supply chains. As social sustainability gets more important especially for large companies and issues such as corruption or child labor have significant reputational impact, companies need to consider the topic. The striving LoD cloud can be an interesting opportunity as it matches sustainability risk management needs such as easy and fast data integration or broad publication possibilities. However, the actual usefulness depends both, on the methods of LoD and on the current state of information published as LoD. The results of the analysis performed are ambiguous. On the one hand, LoD shows potential as a method, particularly in multi-tiered supply chains. Nevertheless, this requires further analysis. On the other hand, the usefulness of LoD as an information resource is limited due to the current size of the LoD cloud and the involvement of third parties. Therefore LoD should probably not be considered as the sole information resource. Key will be an intelligent combination of the data.

8 Appendix

Factor	Indicator/variable name [underlying database as stated in section 5.2]
Democracy	Institutionalized democracy [Worldbank]
Economic development	GDP per capita (current US$) [Worldbank], BIP pro Kopf in KKS [Eurostat], GDP per capita PPP [CIA World Factbook], GDP per capita [OECD]
Econ. freedom	Ease of doing business index (1=most business-friendly regulations) [Worldbank]
Federal state	Government Type [CIA World Factbook]
Income inequality	GINI index [Worldbank], Gini-Koeffizient des verfügbaren Äquivalenzeinkommens (Quelle: SILC) [Eurostat], Income inequality: Gini coefficient, level [OECD]
Inflation	Inflation, GDP deflator (annual %) [Worldbank], HVPI - Inflationsrate [Eurostat], inflationrate_consumerprices [CIA World Factbook], HICP – Overall index [European Central Bank], CPI: all items [OECD]
Perceived corruption	Control of Corruption (estimate) - Corruption Perceptions Index (score) - Corruption (% of mgrs. surveyed ranking this as a major constraint) - Percent of firms identifying corruption as a major constraint - Bribery index (% of gift or informal payment requests during public transactions), Informal payments to public officials (% of firms) - Percent of firms choosing corruption as their biggest obstacle - CPIA transparency, accountability, and corruption in the public sector rating [all Worldbank]
	Transparency International Corruption Perceptions Index [Transp. International]

Table 4: Possible indicators for social sustainability risk rating with regard to corruption

Prio.	Datasets
Prio. 1	European Central Bank (ECB) Linked Data, Federal Reserve Economic Data RDF, Organization for Economic Co-operation and Development (OECD) Linked Data, The Eurostat Linked Data, Transparency International Linked Data, UNODC - Statistics on criminal justice, World Bank Linked Data, World Factbook (FU Berlin)
Prio. 2	business.data.gov.uk, crowdsourcing-fb, DataGovIE - Irish Government Data, DBpedia, DBpedia in Greek, DBpedia in Spanish, DBpedia in Basque, DBpedia in French, DBpedia in Japanese, dbpedia lite, DBpedia in Dutch, DBpedia in Portuguese, OpenUpLabs DCLG, European Central Bank Statistics (PublicData.eu), Feed Wrapper, Eurostat in RDF (FU Berlin), Freebase, kdata, Linked EDGAR (OntologyCentral), Linked Eurostat (OntologyCentral), Linked User Feedback, MLSA - A Multi-layered Reference Corpus for German Sentiment Analysis, Ontos News Portal, OpenCalais, ProductDB, RDFizing and Interlinking the EuroStat Data Set Effort, statistics.data.gov.uk, Twarql, Twitter Linked Data Service, YAGO, Zhishi.me

Table 5: Sets identified as a resources for social sust. risk identification (compiled from [Th13])

9 References

[Ap12] Apple: Apple - Supplier Responsibility - Auditing. 2012. *http://www.apple.com/supplier responsibility/auditing.html*. Accessed 11 January 2013.

[AA11] Ata, A. Y.; Arvas, M. A.: Determinants of Economic Corruption: a Cross-country Data Analysis. In: *Int. Journal of Business and Social Science*, 2 (13): 161–169. 2011.

[Be01] Berners-Lee, T.; Hendler, J.; Lassila, O.; and others: The Semantic Web. In: *Scientific American,* 284 (5): 28–37. 2001.

[Be06] Berners-Lee, T.: Linked Data - Design Issues. 2006. http://www.w3.org/DesignIssues/LinkedData.html. Accessed 19 March 2013.

[BJC11] Bizer, C.; Jentzsch, A.; Cyganiak, R.: State of the LOD Cloud. 2011. http://lod-cloud.net/state/. Accessed 14 March 2013.

[BHB09] Bizer, C.; Heath, T.; Berners-Lee, T.: Linked Data-the Story so Far. In: *International Journal on Semantic Web and Information Systems*, 5 (3): 1–22. 2009.

[Ci12] Ciroth, A.: Software for Life Cycle Assessment. In (Curran, A. Eds.): Life Cycle Assessment Handbook, John Wiley & Sons, New Jersey, 2012; pp. 143–157.

[CJ11] Cyganiak, R.; Jentzsch, A.: The Linking Open Data Cloud Diagram. 2011. http://lod-cloud.net/. Accessed 12 March 2013.

[CR08] Carter, C.; Rogers, D.: A framework of sustainable supply chain management: moving toward new theory. In: *International Journal of Physical Distribution & Logistics Management*, 38 (5): 360-387. 2008.

[DB13] DBpedia: Wiki.dbpedia.org. 2013. http://dbpedia.org/About. Accessed 12 March 2013.

[GSK12] Ghadge, A.; Samir, D.; Kalawsky, R.: Supply chain risk management: present and future scope. In: *International Journal of Logistics Management*, 23 (3): 313-339. 2012.

[Ge10] General Services Administration: Executive Order 13514 Section 13: Recommendations for Vendor and Contractor Emissions. 2012. http://www.gsa.gov/graphics/admin/GSA_Section13_FinalReport_040510_v2.pdf. Accessed 11 January 2013.

[Ge13] Geonames: GeoNames - Datasets - the Data Hub. 2013. http://datahub.io/dataset/geonames. 20 March 2013.

[GG09] García, R; Gil, R.: Publishing xbrl as linked open data. In: *CEUR Workshop Proceedings*, 538. 2009.

[Gl11] Global Reporting Initiative: G3.1 Sustainability Reporting Guidelines. 2011. https://www.globalreporting.org/reporting/latest-guidelines/g3-1-guidelines/Pages/default.aspx/. Accessed 28 December 2012.

[Ha12] Hausenblas, M: 5 Star Open Data. 2012. http://5stardata.info/. Accessed 19 March 2013.

[He08] Heath, T: How Will We Interact with the Web of Data?. In: *IEEE Internet Computing*, 12 (5): 88–91. 2008.

[Ho11a] Hofman, W.: Supply Chain Visibility with Linked Open Data for Supply Chain Risk Analysis. In (Overbeek, S.; Tan, Y-H; Zomer, G. Eds.): *Workshop on IT Innovations Enabling Seamless and Secure Supply Chains*, 20–31. 2011.

[Ho11b] Hofman, W.: Supply Chain Risk Analysis with Linked Open Data. In (Vermaas, P.E.; Dignum, V. Eds.): *Frontiers in Artificial Intelligence and Applications*, 229: 77-87. 2011

[Hu12] Hulstijn, J.; Overbeek, S.; Aldewereld, H.; Christiaanse, R: Integrity of Supply Chain Visibility: Linking Information to the Physical World. In (Bajec, M.; Eder, J. Eds.): *Lecture Notes in Business Information Processing*, 112: 351–365. 2012.

[KM12] Kogg, B.; Mont, O.: Environmental and Social Responsibility in Supply Chains: The Practise of Choice and Interorganisational Management. In: *Ecological Economics*, 83: 154-163. 2012.

[KV12] Klassen, R.; Vereecke, A.: Social issues in supply chains: Capabilities link responsibility, risk (opportunity), and performance. In: *International Journal of Production Economics*, 140 (1): 103-115. 2012.

[Kp11] KPMG: KPMG International Survey of Corporate Responsibility Reporting 2011. 2011. http://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/corporate-responsibility/Documents/2011-survey.pdf. Accessed on 3 April 2013.

[OCH12] O'Riain, S.; Curry, E.; Harth, A.: XBRL and open data for global financial ecosystems: A linked data approach. In: *International Journal of Accounting Information Systems*, 13 (2): 141-162. 2012.

[Pa02] Paldam, M.: The Cross-country Pattern of Corruption: Economics, Culture and the Seesaw Dynamics. In: *European Journal of Political Economy*, 18 (2): 215–240. June 2002.

[Rd04] RDF Working Group: RDF - Semantic Web Standards. 2004. http://www.w3.org/RDF/. Accessed 20 March 2013.

[Ro13a] Rodríguez, J. M. A.: CPV 2008 - Datasets - the Data Hub. 2013. http://datahub.io/dataset/cpv-2008. Accessed 20 March 2013.

[Ro13b] Rodríguez, J. M. A.: ISIC V4 - Datasets - the Data Hub. 2013. http://datahub.io/dataset/isic-v4. Accessed 20 March 2013.

[SM08] Seuring, S.; Müller, M.: From a Literature Review to a Conceptual Framework for Sustainable Supply Chain Management. In: *Journal of Cleaner Production*, 16 (15): 1699–1710. October 2008.

[Th13] The Open Knowledge Foundation: Welcome - the Data Hub. 2013. http://datahub.io/. Accessed 21 March 2013.

[Tr00] Treisman, D.: The Causes of Corruption: a Cross-national Study. In: Journal of Public Economics, 76 (3): 399–457. 2000.

[Ts12] Tsukayama, H.: Foxconn and Apple supply: What are the effects? In: The Washington Post. 24 September 2012. http://www.washingtonpost.com/.../126b9966-0670-11e2-afff-d6c7f20a83bf_story.html. Accessed 4 December 2012.

[Un87] United Nations: Report of the World Commission on Environment and Development: Our Common Future. 1987. http://www.un-documents.net/our-common-future.pdf. Accessed 11 June 2013

[W304] W3C: Web Ontology Language OWL. 2004. http://www.w3.org/2004/OWL/. Accessed 27 March 2013.

[W308] W3C: SPARQL Query Language for RDF. 2008. http://www.w3.org/TR/rdf-sparql-query/. Accessed 27 March 2013.

[Wo13] World Economic Forum: Partnering Against Corruption Initiative. 2013. http://www.weforum.org/issues/partnering-against-corruption-initiative. Accessed 21 March 2013.

[YK05] You, J.-S.; Khagram, S.: A Comparative Study of Inequality and Corruption. In: *American Sociological Review*, 70 (1): 136–157. 2005.