

An information system for assessing the likelihood of child labor in supplier locations leveraging Bayesian networks and text mining

Andreas Thöni¹ · Alfred Taudes² · A Min Tjoa¹

Received: 20 June 2017 / Revised: 16 December 2017 / Accepted: 9 January 2018 /

Published online: 7 February 2018

© The Author(s) 2018. This article is an open access publication

Abstract This paper presents an expert system to monitor social sustainability compliance in supply chains. The system allows to continuously rank suppliers based on their risk of breaching sustainability standards on child labor. It uses a Bayesian network to determine the breach likelihood for each supplier location based on the integration of statistical data, audit results and public reports of child labor incidents. Publicly available statistics on the frequency of child labor in different regions and industries are used as contextual prior. The impact of audit results on the breach likelihood is calibrated based on expert input. Child labor incident observations are included automatically from publicly available news sources using text mining algorithms. The impact of an observation on the breach likelihood is determined by its relevance, credibility and frequency. Extensive tests reveal that the expert system correctly replicates the decisions of domain experts in the fields supply chain management, sustainability management, and risk management.

Keywords Social sustainability · Supply chain risk management · Child labor · Bayesian network risk model · Text mining

✉ Andreas Thöni
andreas.thoeni@gmail.com

Alfred Taudes
alfred.taudes@wu.ac.at

A Min Tjoa
amin@ifs.tuwien.ac.at

¹ Vienna University of Technology, Vienna, Austria

² Vienna University of Economics and Business, Vienna, Austria

1 Introduction

1.1 Supply chain risk child labor

Child labor deprives children of their childhood, their potential and their dignity, and is harmful to their physical and mental development. In conventions Nos. 182 and 138, the International Labour Organization (ILO) calls for the elimination of child labor (International Labour Organization 1973; International Labour Organization 1999) and the 2014 Nobel prize for Kailash Satyarthi and Malala Yousafzay recently focused global attention on the topic (Nobelprize.org 2014). Nevertheless, child labor is still common in many fields of work, especially in developing countries. In 2012, the number of children working in any form totaled 168 million, with 85 million of these performing hazardous work involving physical abuse or the handling of dangerous machinery (International Programme on the Elimination of Child Labour 2013, 13).

Practicing child labor or doing business with parties that engage in such practices is not sustainable from a social point of view. Issues with elements of social sustainability such as human rights or forced labor can also have a significant effect on a company's image (Taylor et al. 2009), even if they occur elsewhere in the supply chain (Lemke and Petersen 2013). Although the responsibility for sustainability is shared along the whole supply chain (Vermeulen and Seuring 2009), the focal, dominant company may be particularly affected. This can be seen in the cases of Nike and Gap in 2000 (Kenyon et al. 2000) or of Foxconn and Apple in 2012 (Tsukayama 2012), which involved accusations of unethical working conditions and show that current risk management practices at least partly fail. Triggered boycotts and loss in brand value on the product and labor market also pose a significant economic risk (Anderson and Anderson 2009). Consequently, sustainability risk management needs to cover the whole supply chain (Seuring and Müller 2008) and containing the likelihood of child labor in supplier locations is an important aspect of supply chain risk management (Lemke and Petersen 2013) and corporate social responsibility (Hutchins and Sutherland 2008).

"Codes of conduct" are acknowledged as the primary instrument for managing social sustainability risks in supply chains. They define what standards need to be followed, thereby regulating sustainability aspects in supply chains and guiding suppliers (Ciliberti et al. 2008; Egels-Zandén 2007; Pedersen and Andersen 2006). However, codes of conduct do not solve the problem of information asymmetry in supply chains (Sarkis et al. 2011), i.e. they do not change the situation that the supplier has knowledge of the local situation which is not available to the focal firm. Consequently, they must be monitored and enforced to guarantee compliance (Pedersen and Andersen 2006). Supplier assessments are seen as a particularly important tool to safeguard compliance with pre-defined standards (Keating et al. 2008; Miemczyk et al. 2012). For this purpose, auditing is an important approach to ensure that pre-defined standards are met (Klassen and Vereecke 2012). On-site auditing is often required to gain insights into social sustainability (Kogg and Mont 2012; Klassen and Vereecke 2012; Benoît and Vickery-Niederman 2011). Another important

instrument often associated with audits is certifications, defined at an inter-company level (Ashby et al. 2012), which help reduce control costs through sector-specific or cross-sector initiatives (Vermeulen and Seuring 2009). Typically, they standardize audit details to a certain extent (Kogg and Mont 2012). These approaches to measuring social sustainability often require a significant amount of resources together with third party input (Vermeulen and Seuring 2009). Consequently, one approach has been to form coalitions to share insights from monitoring and reduce costs (Bremer and Udovich 2001).

However, monitoring and assessment approaches based on audits and certifications are constrained, for a number of reasons: First, supplier monitoring and assessment in modern supply chains is complex. Today's supply chains have a large number of suppliers, often globally dispersed; this complexity and the associated costs (Kogg and Mont 2012) make effective ongoing verification of compliance with standards difficult. It seems impossible for companies to have all factors evaluated in depth, let alone first-hand, by an internal company employee, although the need would exist in practice (Kogg and Mont 2012). Even more resources are required if deeper levels of the supply chain are considered in case these suppliers are known (Svensson 2009; Grimm et al. 2016).

Second, timeliness is a problem. Given that internal supplier information is mostly only available through audits, certifications, or supplier communication (e.g. Klassen and Vereecke 2012), there is an associated lag in compliance verification. Infrequent certifications only provide limited defense against issues, as, for example, the child labor revelations at the OTTO group showed in 2007 (Aktiv gegen Kinderarbeit 2014; McDougall and Schmitz 2007). Moreover, suppliers nowadays tend to be overloaded with requests for certifications and audits (Ceres and Sustainalytics 2014), which is referred to as "audit fatigue" (Kogg and Mont 2012, 162). Others even describe monitoring as a way to convey an adversarial stance and that focal firms act more like supply chain 'bullies' than CSR 'champions' (Boyd et al. 2007).

Third, there is a lack of objectivity. Data gathered through monitoring systems may not reflect the truth due to the potential effects of bribery, corruption, and culture or standards differences, particularly if third parties are used to perform the ground work (Leire and Mont 2010). Locke et al. (2009) declare the information gathered through audits to be often inaccurate, biased, and incomplete. Some companies (e.g. British Telecommunications 2012) use supplier questionnaires to identify the companies to follow up. However, these selections are based on the quality of data that suppliers themselves provide (Leire and Mont 2010).

Finally, the management of the collected information and its aggregation to valid KPIs can take considerable amounts of time. Usually, managers prefer indicators that are easy and fast to calculate (Mcintyre et al. 1998).

1.2 Automated child labor risk monitoring

In order to reduce the effort of monitoring the risk of child labor in the supply chain, one could envision an expert system that automatically computes a risk score based on the geographic location and industry of a supplier and the result of the

respective audits and certifications. This score would then be used to direct on-site audits to the supplier locations with the highest risk. Given the fact that unstructured text has already been promisingly integrated into risk management approaches in non-sustainability domains such as financial risks (e.g. Groth and Muntermann 2011), general business risks (Leidner and Schilder 2010), tracking society-related sustainability indicators (Rivera et al. 2014), or employee fraud (Holton 2009), it appears fruitful to also integrate external public information into the computation of the child labor risk score in order to enhance the objectivity of monitoring results and to overcome the time lag between supplier reviews. In fact, external risk-related information collection has already been reported in the sustainability domain; however, automated, integrated IT-supported risk modeling is not specified (Koplin et al. 2007). Ongoing input from news sources or social networks may help to identify risk-relevant events in the supply chain. These events can be gathered based on geography-, sector-, and production-specific relations (see e.g. UNEP 2009, 60), or on other cause-and-effect relationships derived from literature related to child labor. First suggestions for the system have been presented in TBD [removed for peer-review/conference paper].

To check if a child labor risk management system as described above would also be useful in practice, expert interviews on information needs were conducted with five managers in charge of sustainability in the oil and gas, paper and retail industries. The requirements stated to be top priority were “Inclusion of external data (external blacklists, external platforms, social media, news media)”, “Inclusion of risk metadata (country/region data, supplier data, legal data, components/products, certifications)”, “Definition of KPI structure and aggregation logic” and “Allow for supplier selection, supplier ranking and audit triggering/prioritization”. Thus, it is worthwhile to study a risk management system for monitoring child labor risk in supplier locations which mines text sources for reports of child labor incidents and automatically combines this evidence with risk assessments based on supplier location and industry as well as results of audits and certifications.

1.3 Existing literature

Several quantitative models for supply chain risk management that reflect social sustainability have been described in literature. A model specific to textile supply chains is suggested in Rabenasolo and Zeng (2012). It uses linguistic variables and relies on a combination of performance indicators. Weber et al. (2010) include sustainability risks in a credit risk indicator template filled by credit officers. Badurdeen et al. (2014) outline a Bayesian-network-based approach to combine various risk categories into a final probabilistic score. Still, social or environmental concerns are just one input out of several, and the analysis is based on expert input only. Finally, Hadiguna (Hadiguna 2012) suggests, without offering mathematical details, a decision-support framework for risk management that addresses social elements such as labor strikes, demonstrations, and local customs to a limited extent. Other contexts for the application of Bayesian networks have been, for example, new product development (Chin et al. 2009) or early crisis warning (Dabrowski et al. 2016).

Sarkis and Dhavale (2014) introduce an initial approach using Bayesian inference (without a network). They combine a series of criteria reflecting sustainability, norm them, and apply Gibbs sampling to derive a posterior distribution from prior distributions and observations. Erol et al. (2011) present a yearly sustainability alert system based on fuzzy entropy and fuzzy multi-attribute utility that creates alerts if values for sustainability indicators are above or below certain thresholds. Fuge et al. (2013) include probabilistic measures of sustainability integrating different private and public indicators via a weighting. They replace unknown specific indicators with more general indicators while respecting the additional variance this brings. Similarly, Ahi and Searcy (2014) suggest a stochastic model to measure the sustainability performance of a supply chain. Shokravi and Kurnia (2014) include an importance parameter which is based on analyzing texts in order to derive a weight for each measure. Wu et al. (2017) use a multitude of quantitative techniques to combine qualitative, social media and quantitative data. For social media they strongly build on term frequencies and company-related sources. Mani et al. (2017) show the usefulness of big data analysis for sustainability analysis in supply chains using a case study. Summarizing, to the best of the authors' knowledge, none of the related quantitative approaches deals with ongoing automatic monitoring combining internal like audits and external data sources, including broad news. This is while text mining is recognized as one of the key techniques in big data analysis (Nguyen et al. 2017) and more tools for sustainability analysis with big data are called for (Wang et al. 2016).

1.4 Outline

Given this state of the art, a prototype of an expert system for monitoring child labor risk in supply chains was built. This prototype was then tested by comparing its decisions with those of human experts. Section 2 of the paper deals with the description of the prototype, and Sect. 3 reports the results of the tests. After the discussion in Sect. 4, a summary of the findings, limitations of the work and suggestions for further research are the subject of Sect. 5.

2 Description of the system

2.1 Overview

We propose using a Bayesian network (BN) to compute the likelihood of child labor in a supplier location based on the evidence from geography and sector, audits and news reports. BNs are a well-known probabilistic modeling technique introduced by Pearl (1988). BNs are based on Bayes' Theorem. A BN is a directed acyclic graph (DAG) in which nodes correspond to random variables of interest and directed arcs represent direct causal or influential relation between nodes. The uncertainty of the interdependence of the variables is represented locally by a conditional probability table (Wattayhu and Peng 2004).

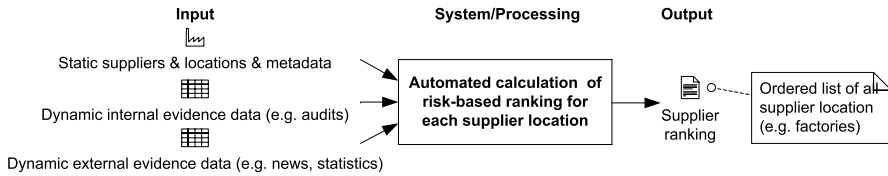


Fig. 1 Overview of risk model and system

BNs have previously been successfully applied to risk management due to their understandability and ease of information integration (Duespohl et al. 2012; Koks and Challa 2005; Wooldridge 2003). The key advantage of BNs is their explicit treatment of uncertain information supporting decision making (Reckhow 1999) and the possibility to include different types of sources into a single consistent model (Uusitalo 2007; Duespohl et al. 2012; Wooldridge 2003). A process of updates allows the inclusion of news as it comes up, continuously triggering updates of the likelihoods (Neapolitan 2003, 12–29). BNs tend to be easily communicable, fostering a common understanding (Duespohl et al. 2012; Correa et al. 2009). These features of BNs are of special interest for a quantitative risk model, as a decision-relevant information system must be understandable for company executives who have to make and defend their decisions based on the input (Hubbard 2009).

We suggest implementing a BN for each supplier location. We propose starting with the initial hypothesis that a supplier conforms to given, pre-defined social sustainability standards (e.g., a “code of conduct”). Then we calculate a relative measure for the likelihood of this hypothesis being false based on the evidence on the likelihood of a compliance breach. Hence, computing the likelihoods for individual supplier locations and relating them can provide a relative risk ranking (see Fig. 1). Figure 2 depicts the structure of a BN for a supplier location in the notation of Netica, which was used as implementation environment (Norsys 2013). Input data for the BN can come from multiple static or dynamic sources that either provide structured or unstructured data. To gather these inputs, survey data on child labor and audit scores are used. Text mining is leveraged to extract information on child labor incidents from unstructured news articles. For any new location, the prior distribution is taken. Once the location is known, the country, area type and sector are given and the priors can be updated. The priors for audits are updated anytime when the result of an audit of the respective supplier location is entered into the BN. The priors for observations are updated when information on a child labor incident relevant to the location is supplied as input to the BN.

The parameterization and testing of the system was conducted together with 28 experts, 13 of which with background in supply chain management, 6 with background in sustainability management, and 3 in general management, risk management and other, respectively. 15 of the experts consulted had more than 5 years of experience in their position while another twelve had between 1 and 5 years. Only 1 expert had less than 1 year of experience. 10 experts had their workplace in Austria, 7 in Germany, 2 each were from China, Malaysia, and the United Kingdom, and

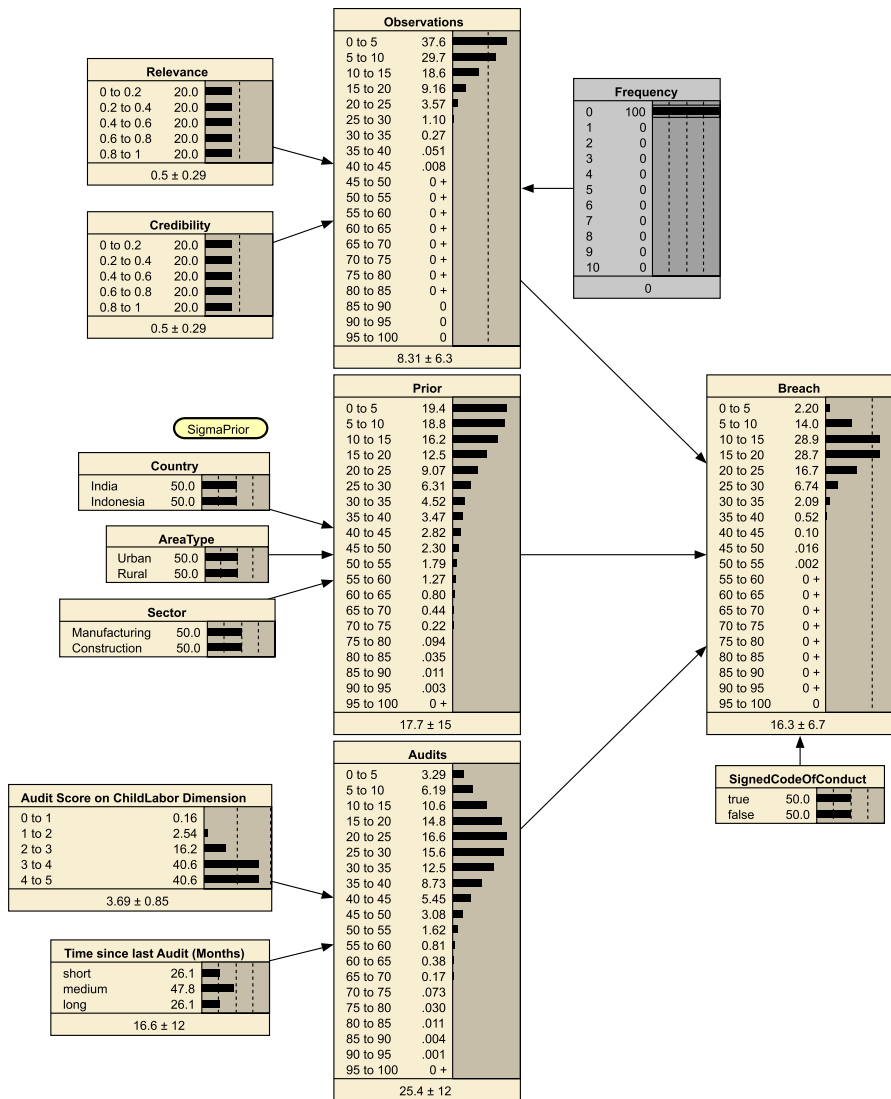


Fig. 2 Full view of Bayesian risk network

1 each from the Czech Republic, Denmark, Romania, France and Columbia. The companies the experts worked for were in a variety of industries, but with a significant spike in manufacturing and wholesale/retail trade (18 experts), which have a potentially higher exposure to child labor. According to India, National Sample Survey Round 66 (NSS-R66) 2009–2010 (Understanding Children’s Work 2010), 23.6% of children from age 5 to 14 work in commerce in urban areas. This is the highest percentage of all sectors. 23 experts worked in companies with more than 1000 employees.

Although experts are subject to at least not fully rational (bounded) behavior (for an early discussion, see e.g. Edwards 1954), experts decide using a set of decision strategies that employ a significant amount of heuristics and learning from the past (March 1994; Shanteau 1988). Therefore, a comparison with expert output provides an indication of any overlap between the system's approaches and these strategies. Moreover, it helps in discussing whether the experts' responses can be covered by the system's design.

2.2 Impact of region and sector

Let us now describe the various components of a supplier location BN in detail. Sources like UCW (Understanding Children's Work 2014) show that the frequency of child labor varies regionally, depending on the country (C) and whether the area is rural or urban (R), and between sectors (S). Therefore, suppliers located in different areas and working in different sectors have different prior probabilities of employing children. These contextual priors can be determined on the basis of publicly available statistics on the number of children and companies per context C,R,S and the fraction of children who are non-self-employed workers. As one cannot derive the number of children working at a particular company from the statistical data available it makes sense to assume that the distribution of children working to the companies in the context is random. If one assumes that all child workers are randomly assigned to the $NCOMP_{C,R,S}$ companies in the specific context the probability that a child is not linked to a specific company is

$$1 - \frac{1}{NCOMP_{C,R,S}}$$

Multiplying the frequency of children working non-self-employed in context C,R,S with the total number of children in the respective context yields the number of non-self-employed working children $ANCL_{C,R,S}$. The probability that all children working in the respective context are not working for a particular company is given by the probability that a child is not linked to a particular company to the power of $ANCL_{C,R,S}$. Then, the probability that all these children are not working at the specific company is given as

$$\left(1 - \frac{1}{NCOMP_{C,R,S}}\right)^{ANCL_{C,R,S}}$$

and the probability of having at least one child working at the respective company is given as 1 minus the probability that all children working in the respective context are not working for a particular company:

$$P_{CL} = 1 - \left(1 - \frac{1}{NCOMP_{C,R,S}}\right)^{ANCL_{C,R,S}}$$

Table 1 Comparison of example child labor incident priors for different approaches, countries, area type, and sectors

Country	Area type	Manufacturing (%)	Construction (%)
India	Urban	3.9	23.6
	Rural	5.9	44.7
Indonesia	Urban	4.9	0.0
	Rural	2.6	1.5

Table 2 Estimation of standard error of child labor rate calculations based on country comparisons

Country	Absolute difference	Adj. absolute deviation from mean per country
Bangladesh	14.9	7.45
Bolivia	9.1	4.55
Brazil	0.3	0.15
Cameroon	48.3	24.15
Ghana	26.5	13.25
Kenya	36	18
Lesotho	30.9	15.45
Sao Tome e Principe	16.7	8.35
Senegal	12.9	6.45
Average	21.73	
Standard deviation		13.32

Table 1 shows the resulting P_{CL} values for supplier locations in India and Indonesia in year 2012 calculated on the basis of BPS Statistic Indonesia (2008), Diallo et al. (2013), International Programme on the Elimination of Child Labour (2013), Ministry of Statistics and Programme and Implementation of India (2006), The World Bank Group (2013) and Understanding Children's Work (2010).

To derive the standard deviation of this contextual prior, it is assumed that the errors made when calculating the probability of a child labor incident can be compared to those made when estimating the frequency of child labor in a particular country. Table 2 depicts the differences between the results of two surveys referring to the same reference period in nine countries analyzed (see Guarcello et al. 2010, 10). We interpret these differences as 2σ intervals, and set the standard deviation of the contextual prior σ_{prior} to 13.32, the mean of the differences reported in Guarcello et al. (2010) divided by 2.

As we only have estimates for mean and variance the assumption of a normal distribution is the most parsimonious one based on entropy arguments (Cover and Thomas 1991, 409f.). We therefore assume that P_{CL} is the expected value of a normally distributed random variable. As can be seen from Fig. 2 negative values are dealt with in the BN by the categorizing the potential outcomes into intervals where the lowest one aggregates negative values. Initially, all realizations of the context are assumed as equally likely in the BN of a supplier location. Once the respective country, area and sector are known, one value is selected with 100% probability.

In order to validate the method for determining the contextual prior, the experts were asked to rank the four hypothetical supplier locations shown in Table 3 according to child labor risk. Based on the data provided, the experts suggested an initial ranking that is in many ways comparable to the one created using the model. In order to better compare the two approaches, Fig. 3 introduces a scaled measure. This measure was derived by norming the best-rated supplier to zero and the worst-rated supplier to one. For the model, relative distances were then calculated using the mean prior values. The model was set up using the values for the prior that correspond to the ones used in the questionnaire. The adjustments for the expert responses were calculated using a weighted score (response frequency f and weight w of 1 for best rank, 4 for worst rank with combination of $f \cdot w$) and norming to 0–1. It must be noted that the experts' responses have not been interval scaled, so the ordering is important. Experts tend to view the location of supplier B as worse than the one for supplier D when compared with the model under these assumptions. Nevertheless, in general the model prior and expert responses appear to have a comparable pattern.

A more granular analysis of the result, however, shows that experts often strongly disagree in their judgment of the riskiness of the different suppliers (see Fig. 4). The rank selected by the majority of experts is only equivalent to the one calculated by the model for ranks three and four. Suppliers B and D show a particularly large spread of answers.

2.3 Impact of audits

Audits are limited in what they can measure and can only be conducted within a defined timeframe, leaving the suppliers alone before and after this timeframe (Locke et al. 2009). Also, a higher number of compliance audits does not suggest that a supplier is better than others. Rather, often the compliance level of the supplier stays the same and sometimes even worsens (Locke et al. 2007). Consequently, we only include the result of the last audit into the BN and assume that the variance of the breach likelihood increases with the time since the last audit.

In order to infer the relationship between audit score and breach likelihood we asked the experts “Which probability of having a child labor incident (if only audit data is taken into consideration) would you associate with a random supplier reaching either a minimum (worst), a medium, or a maximum (best) audit score?”. As can be seen from the high standard deviations provided in Fig. 5 and Table 4, the relation between audit scores and average probability of an incident is judged to be very ambiguous. While some experts put a lot of trust in audit scores, others see only limited value. Even if an audit attributes the best score to a supplier, experts tend to still see a certain probability of an incident. Similarly, the worst audit score does not necessarily indicate that child labor is present.

We assume that an audit yields results in the range $[a_{min}, \dots, a_{max}]$, where the minimum audit score a_{min} is assumed to be greater or equal zero. As we only have estimates for mean and variance, the assumption of a normal distribution is the most parsimonious one based on entropy arguments (Cover T, Thomas J 1991, 409f.).

Table 3 Supplier locations used for validation

Supplier	Country	Rural/urban	Sector	Est. child labor rate	Sector's share of overall child labor	Est. # of children	Est. # of companies in sector
Supplier A	Blueland	Rural	Construction	4.9	0.1	1475	95,226
Supplier B	Pinkland	Urban	Manufacturing	0.5	34.2	71,435	1,808,369
Supplier C	Blueland	Urban	Construction	1.9	0	0	85,093
Supplier D	Blueland	Urban	Manufacturing	1.9	17.6	51,103	1,027,252

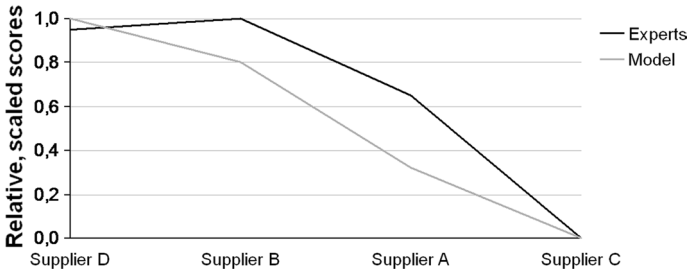


Fig. 3 Consolidated ranking comparison between all experts and model

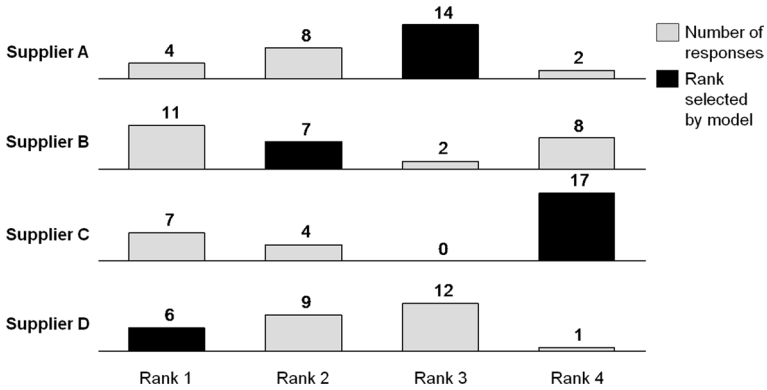


Fig. 4 Rankings per expert and model

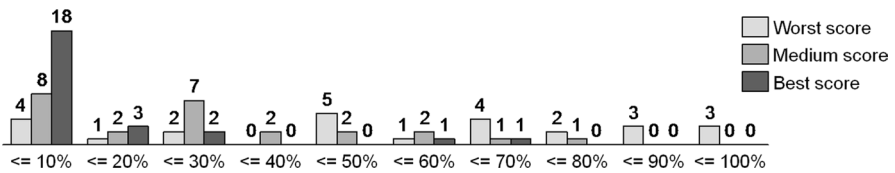


Fig. 5 Number of responses for incident probability given a supplier audit score grouped by percentage categories

Table 4 Average estimated probability values (incl. standard deviation) of incident for different audit scores

Value	Worst score	Medium score	Best score
Average (percent)	57.4	29.3	13.0
Std. deviation	30.6	22.4	17.9

Based on these judgments, it is assumed that the audit likelihood P_{audit} follows a normal distribution, whose expected value $E(P_{audit})$ is related to the audit score via

$$E(P_{audit}) = m - a \times \frac{m - n}{a_{max}}$$

Table 5 Audit probability standard deviation values depending on time since last audit

Time since last audit	Time since last audit (months)	Assumed standard deviation (percent points)
Short	$\geq 0 \ \&\& \ < 6$	5
Medium	$\geq 6 \ \&\& \ < 24$	10
Long	≥ 24	15

where $a_{min} \geq 0$ is the a minimum audit score and a_{max} the maximum audit score. For the prototype, the minimum audit score a_{min} is set to 0 and the maximum a_{max} to 5, and the parameters m and n are set to $m = 57$ and $n = 48.2$. This formulation fits the values given in Table 4 and was considered valid by the experts. Table 5 contains the time-dependent values used for the standard deviation of the audit probability.

For the prior of the audit score, a normal distribution is assumed with a mean of 4 and standard deviation of 1, while for the variance of P_{audit} we assume a prior value of a mean of 9 months with a deviation of 3 months. If an introductory audit for a supplier location is entered (i.e., marking one discretization with 100% probability), the prior for the node has no influence anymore. The respective distributions were derived from discussions with the experts and given to them for validation.

2.4 Impact of observations

Both observations of child labor incidents in related contexts and news on drivers affecting the demand and supply of child labor are candidate inputs for the determination of the breach likelihood. Empirical research has uncovered a multitude of factors influencing the extent of child labor. Only few of these factors such as socio-economic dislocation (economic crisis, political and social transition) or production peaks/labor shortages are observable from external information and have short-time impact. These are difficult to detect automatically, though, as they cover a wide array of happenings, including earthquakes, volcanic eruptions, strikes, or demand surges. Moreover, while the literature identifies connections between these events and child labor, the propensity of the effect varies by context. Also, descriptions of actual child labor incidence observations are more homogeneous than descriptions of factors that influence child labor. Thus, descriptions of child labor incidents are easier to detect and to codify automatically than descriptions of factors influencing child labor. Also, one can argue that they indirectly cover relevant influence factors, as socio-economic dislocation or production peaks/labor shortages should affect all companies operating in a similar context. We therefore choose to only include news reporting incidents of child labor into the expert system.

In order to determine the impact of various types of news on child labor the experts were confronted with the four hypothetical news articles in the Appendix 1. Then they were asked “How much does the news report influence the perceived probability of a child labor incident at Supplier B?” and a five-step Likert influence scale was used (extremely influential-not at all influential; Wigas 2006) to code the answers. As can be seen from the score depicted in Fig. 6 all articles provided have

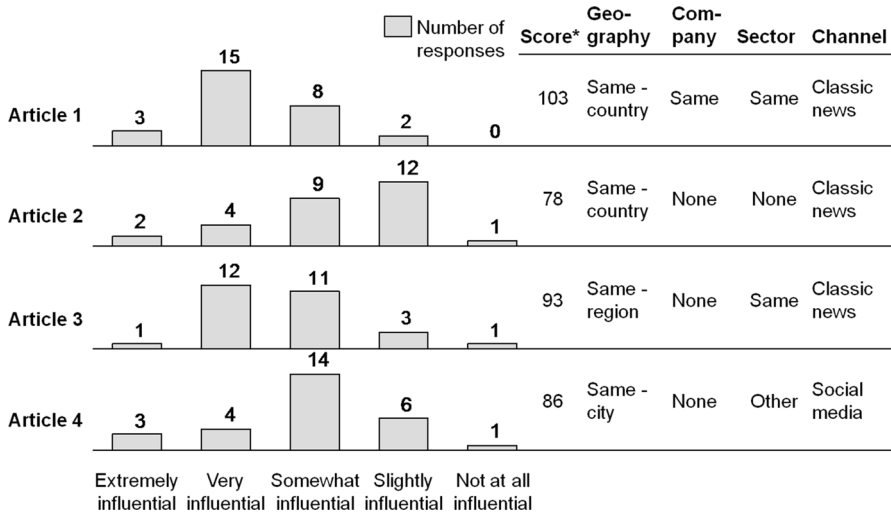


Fig. 6 Number of responses per article and influence selection (*calculation of score: extremely influential—5 points; very influential—4 points; somewhat influential—3 points; slightly influential—2 points; not at all influential—1 point)

at least a slight influence on the experts’ decisions. Comparing articles one, two and three, the additional geographic detail (region) is nearly as influential as the explicit mentioning of the company. Hence, closely related geographic proximity drives relevance. This is not the case for the article obtained through social media (the worsening could also be due to the reference to a different sector).

Consequently, we suggest considering two variables, credibility and relevance, to represent the content quality of an observation. Credibility *c* is defined as comprising the content of evidence captured by a sensor which includes veracity, objectivity, observational sensitivity, and self-confidence (Blasch et al. 2013), while relevance *r* assesses how a given uncertainty representation is able to capture whether a given input is related to the problem that was the source of the data request (Blasch et al. 2013). In other words, the model understands relevance as capturing how closely the messages used as input for a certain supplier location are in fact related to the supplier location. In order to derive a relevance measure, the availability of dimensional attributes in the news articles is used as an approximate indicator. The more an observation can be linked to a certain location in a granular and specific way, the more relevant it is. If observations with partly conflicting dimensional information are included, the relevance can only be derived based on the non-conflicting dimensional information. Credibility is suggested as being defined either at an input channel or source level in order to cover different media types as completely as possible.

Tables 6 and 7 show the particular values used for credibility *c* and relevance *r*.

A BN is initialized at a particular point in time which can serve as a basic reference point. Until this time, zero or more observations of child labor incidents may have been stored and a set of observations can be retrieved as discussed above. In general, when revising the probability based on evidence from textual media

Table 6 Credibility values based on publishing channels

Attributes of child labor incident observation		Credibility <i>c</i>
Publishing channel	Examples	
Social media	Twitter, Facebook	0.2
Web 2.0 website	Blogspot	0.5
News channel	CNN, BBC, Reuters	0.8

Table 7 Relevance values based on dimensional attribute availability

Attributes of child labor incident observation			Relevance <i>r</i>
Geographic features	Company features	Sector features	
None	Company name	None	0.5
None	None	ISIC sector	0.0
None	Company name	ISIC sector	0.8
Country	None	None	0.1
Region	None	None	0.3
City	None	None	0.5
Country	None	ISIC sector	0.2
Region	None	ISIC sector	0.5
City	None	ISIC sector	0.7
Country	Company name	None	0.7
Region	Company name	None	0.8
City	Company name	None	0.9
Country	Company name	ISIC sector	0.7
Region	Company name	ISIC sector	0.8
City	Company name	ISIC sector	0.9

sources, two options may be considered. Either only the latest observation is entered as a single finding or the network is continuously updated with the evidence from new observations. If the observations are assumed to be independent, each one is likely to include valuable information. Consequently, the BN will be modeled using the latter option, allowing the inclusion of evidence from multiple reports. It is then the task of the input procedure to ensure independence between the incidence observations. Relevance and credibility can be evaluated for each observation as described above using the pre-configured values in Tables 6 and 7 given an observation's data. Even observations with low credibility or relevance are understood to increase the overall observational probability. Given these assumptions, the expected value of the observational likelihood P_{obs} needs to be a monotonically increasing function of the number of independent incident observations included: $\forall f \geq 0 : E(P_{obs}(f_1, c_1, r_1)) \leq E(P_{obs}(f_2, c_2, r_2)); f_2 \geq f_1; 0 \leq p, c_{1,2}, r_{1,2} \leq 1; f_{1,2} \in \mathbb{N}$.

Both the credibility and the relevance will be continuously updated with new evidence from observations. The evidence will then be entered into the BNs whose

context overlaps with the observations' context. This leads to a model containing frequency f , credibility c , and relevance r as variables. These are combined via the equation $x = f \cdot (c + r)$, which fulfils the requirement of monotonicity requirement formulated above and conforms to the notion that both credibility and relevance increase with frequency in a simple way. However, this monotonically increasing function x has no defined upper bound. Therefore, a scaling function is needed to return a value between 0 and 1 for the mean of the normal distribution of P_{obs} . For this purpose, a monotonically increasing function with limit 1 is suggested. This can be achieved with an inverted, shifted hyperbola. The frequency score function $s(x)$ is suggested for this (the function can be parameterized through the parameter τ , which is initially set to 5): $s(x) = 1 - \left(\frac{1}{1 + \frac{x}{\tau}} \right)$. This formula has the desired property of

limit 1 and is simple.

The node of the observational likelihood P_{obs} is represented with a normal distribution given relevance r , credibility c , and frequency f . Hence, knowing about f observations for a specific supplier location, the calculated likelihood of a child labor incident should be within a predefined confidence interval. This interval should be smaller the higher the number of observations with high reliability and credibility received. For a known standard distribution, 95% of its probability mass lies within the mean μ plus/minus 1.95994 times the standard deviation. If, as defined by the user, the area covered by the 95% interval is p' percent points if no observation has been received and p'' percent points if ten fully credible and relevant observations have been received, then the respective standard deviations in percent points can be calculated with $\sigma(p) = \frac{p/2}{1.959964}$. For example, setting p' to 40 and p'' to 10 percent points yields $\sigma' = 10.204$ and $\sigma'' = 2.551$. σ is seen dependent on the values of f, r, c and a linear functional connection is assumed.

$$\sigma(f, r, c) = \alpha - \beta \cdot f \cdot r \cdot c$$

Using σ' and σ'' , the values for α and β can be determined leading to the following function

$$\sigma(f, r, c) = \alpha - \beta \cdot f \cdot r \cdot c = 10.204 - 0.7653 \cdot f \cdot r \cdot c$$

As shown in Thoeni (2015), this specification is consistent with the above stated monotonicity condition. Besides fulfilling the monotonicity requirement, it also was considered plausible by the experts.

2.5 Text mining child labor incidence observations

2.5.1 Methodology

Manually reading, extracting and coding child labor incidents from a continuously arriving stream of text is clearly infeasible. However, automatic extraction is difficult: the examples shown in Appendix 2 give an impression of the diverse way in which child labor is depicted in various texts. First, the datasets also include broader reports. They often reference a broader array of different child labor incidents (#1,

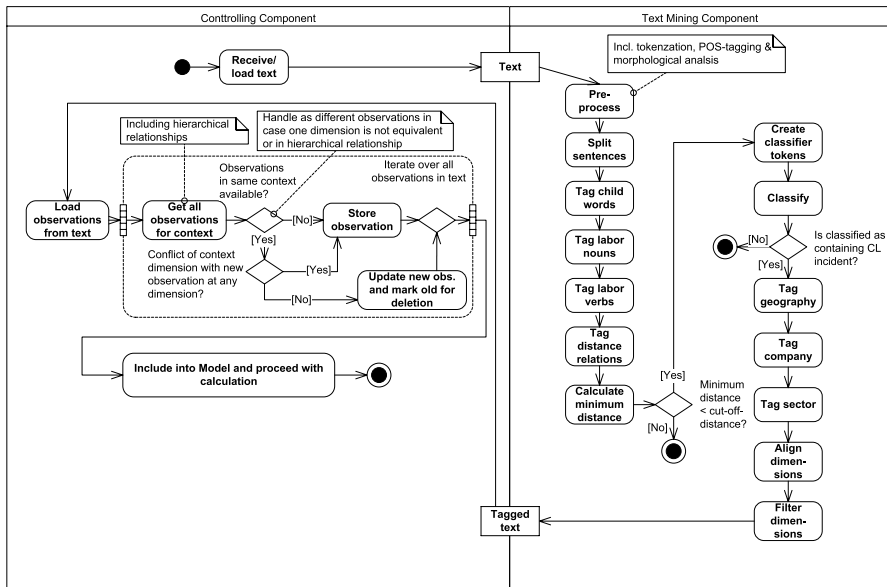


Fig. 7 Text mining and BN updating workflow for one document

#2), together with contextual references (e.g. #3). Beyond these general reports, other texts also give broader references to a combined set of multiple child labor cases (#4, #5). Incidents may be depicted in narrative fashion building on a single individual case (#6) or at least referencing it directly (#7). However, incidents are also reported directly, as can be seen in the later text excerpts (#8 to #11). There, the reports on child labor can also include child labor categories such as prostitution, begging, or domestic work that are less relevant from a company perspective (#4).

The automatic text mining and BN updating procedure depicted in Fig. 7 has been developed to cope with this challenge. Given the low frequency of child labor incidents and the large variety of forms in which these are expressed, we employ data-driven document classification with candidate set reduction and tagged event extraction. In particular, the following four text mining steps are performed:

1. *Preprocessing and Candidate Set Reduction* A tokenizer splits the words and other characters, a sentence splitter detects sentence boundaries, and a POS-tagger is used to differentiate word lemmas. As sentences close together in a text tend to be on the same topic (Zha 2002) and a direct mention of “child labor” may be seen as the most obvious trigger of a child labor incident event, a distance-based approach using a cut-off distance is suggested to prune negative cases. The distance is measured as the number of characters between “child” and a word indicating “labor.” Indicative words may be synonyms, hypernyms, or other related word sets. Stop words are eliminated between “child” and “labor”. The cut-off distance to be used is determined together with model selection and parameter estimation so as to optimize the F1 measure combining precision and recall.

2. *Classification* The thus identified text passage between “child” and “labor” could describe a child labor event. This could in turn contain several child labor observations, as demonstrated by the following example: “As many as 36 cases from Koderma and 22 from Khunti were brought to Dube’s notice when he visited these districts. Such incidents include a girl from Khunti missing since 2009 when she went to Delhi for work [...]”. Classification based on the relative word frequencies of the words within the feature is used to verify if the extracted feature actually deals with a child labor event. For training, the Reuters TRC2 corpus containing 1,800,370 news articles (Reuters, National Institute of Standards and Technology 2009) was used. 16,948 articles in Reuters TC2 contained the word “child” and manual tagging yielded 117 articles that contained a child labor event. A number of variants for feature construction (maximum or minimum distance between “child” and “labor” including or excluding leading or trailing words to complete sentences), model selection (SVM, PAUM, KNN, NB, C4.5) and cut-off values were tried out using this gold standard (Li et al. 2005; Quinlan 1993). This resulted in the choice of SVM with a cutoff distance of 80 applied to the sentences within the maximum distance as the best variant with precision 97.1%, recall 73.7% and F1 value of 83.4%.
3. *Event extraction* The result of the classification step is a list of news reports of which each should (and in the case of 100% precision actually does) contain at least one child labor incident observation referring to a child labor incident event. The goal of the next step is to extract these incident observations together with the corresponding attribute values from the text. DBpedia Spotlight and Open Calais by Reuters were used for geography tagging, and Open Calais for company tagging, yielding respective URIs. Sector tagging must yield a sector conforming to United Nations ISIC industry classification so as to conform to the statistical data. This was done via a rule-based gazetteer and an ML-based approach, where the labels and descriptions from the ISIC classes were used to train a classifier based on the lemmas of the respective tokens. Given that different taggers produce syntactically and also partly semantically different tags, they have to be aligned to a common tag set via a domain ontology (Gangemi 2013; Rizzo et al. 2012).
4. *Independent Observation Extraction* Output of event extraction is a frame with zero to several values for each of the dimensions (components of context) location (hierarchy country, region, city), company and sector hierarchy. In a next step, child labor incident observations that contain at most one value per dimension are generated using heuristics based on proximity and child labor distances. This procedure yielded an overall F1 value of 44.7%, and F1 values of 57.6% for geography, 74.4% for organization and 27.2% for sector in the best alternative. Comparing these results to those obtained for the classification step reveals that observation extraction shows weaker results. This is particularly true for the sector dimension. Consequently, a manual cleaning step appears necessary before values are eventually entered into a risk management system in a productive scenario. Finally, these child labor observations are checked for duplicates by checking if there are no conflicts with already stored observations within a given time-frame. If this is the case, the old observation is deleted and the frequency of the new one is increased (see Fig. 8).

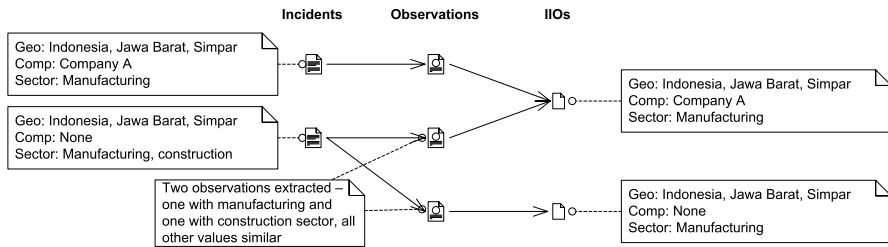


Fig. 8 Example independent observation (IIO) extraction

The output of these steps is a list of independent incident observations with fully or partly filled attributes linked to a domain ontology representing this frame. These can then be incorporated into the risk model by activating the supplier location BNs whose context variables overlap with the dimensions of the observations incorporating hierarchical relationships.

2.5.2 Validation of input data availability

A key criterion for the usefulness of the system is the availability of child labor incident observations in sufficient number and granularity. To probe into this issue, the suitability of two publicly available data sets was investigated. News was gathered through searches on Google News, the European Media Monitor, and two selected RSS feeds, resulting in a very broad coverage of news. Altogether, this dataset contains 48,339 news articles published between 15th March 2011 and 16th September 2014. Most articles were retrieved from the British Broadcasting Corporation (BBC), Times of India, The Daily Mail, The Guardian, and The Hindu.

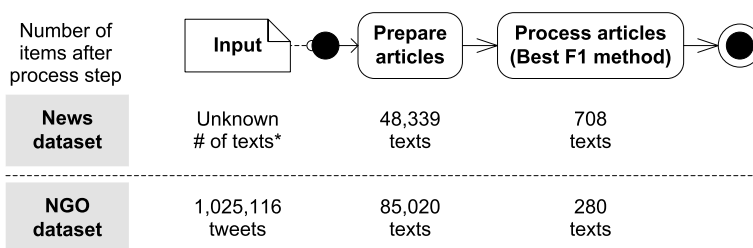
Also, a list of NGOs that potentially post on Twitter was built in order to retrieve the related content. In order to cope with the amount of data, we restricted ourselves to India, given the importance of English (Crystal 2004) and prevalence of child labor (Understanding Children's Work 2010), and Indonesia, due to its high Twitter use (Bennett 2012) and the presence of child labor (Understanding Children's Work 2009). 5138 unique NGO websites (predominantly in India) were automatically parsed to determine whether a link to Twitter was provided on the first page that opened by following the website link (see Table 8). Altogether, this resulted in a set of 778 unique twitter accounts. Using the Twitter search API iteratively, produced a set of roughly one million tweets, published between 11th July 2007 and 10th March 2014. When downloading each tweet, external links (included in the tweet) were followed. This website data is stored together with the tweet. The tweets have been reduced to a set where each linked text contains "child" at any place in the text, similar to the assumption used in the text mining methodology. Consequently, 85,020 texts were then stored as a new set for further processing.

The two sources were input into the text mining procedure described above. This resulted in 708 texts from the news data set and 280 texts from the NGO data set (see Fig. 9). The results of the analysis of the random selection of 100 articles from

Table 8 Overview of number of NGOs collected with and without websites, including sources (author's representation)

ID	Source	Description	NGOs retrieved	NGOs with website link provided	
				# Total	# Total # Unique Websites
1	ILO	NGO on Child Labor in India	13	12	12
2	Forum: Karmayog	NGOs in India	163	70	65
3	DevDir	NGOs in India	2,747	2086	1825
4	Choike	NGOs on Human Rights	10	10	10
5	IdeaList Search	Organizations in different countries	470	349	266
6	Tigweb	NGOs worldwide per country	108	63	58
7	UN	NGOs worldwide per country	671	454	393
8	NGOs of India	NGOs in India	6423	1771	1468
9	US Aid	US NGOs worldwide per country	124	124	71
10	DevDir	NGOs in Indonesia	628	506	374
11	Semru	NGOs in Indonesia	117	117	90
12	ProPoor	NGOs in India	856	599	506
Sum			12,330	6161	5139

Sources (referenced by ID): (1) <http://www.ilo.org/legacy/english/regions/asro/newdelhi/ipcc/responses/india/other.htm>, (2) http://www.karmayog.org/childlabour/childlabour_18034.htm, (3) <http://www.devdir.org/files/India.pdf>, (4) <http://www.choike.org/2009/eng/5/1/links.html>, (5) <http://www.idealists.org/search/v2/>, (6) <http://www.tigweb.org/resources/orgs/>, (7) <http://esango.un.org/civilsociety/displayAdvancedSearch.do>, (8) <http://www.ngosindia.com/ngos.php>, (9) <http://www.pvo.net/usaid/>, (10) <http://www.devdir.org/files/Indonesia.PDF>, (11) <http://www.smeru.or.id/ngolist.php>, (12) <http://www.propoor.org/search.php>; dates of collection: 9.11.2013–18.11.2013 and 16.2.2014 (ProPoor). An ascending order of IDs was used to determine unique websites—"unique website" is referenced with the first source mentioned in this order

**Fig. 9** Number of articles in different steps of input data analysis for news and NGO datasets (*unknown due to IR approach through news aggregators)

the news and NGO datasets are presented in Fig. 10. Manual inspection shows that the large majority of articles in the sets (96 and 89% respectively) do in fact include business-related child labor incidents. Only five cases had no dimension, i.e., the text contains a reference which can be classified as a child labor incident under the

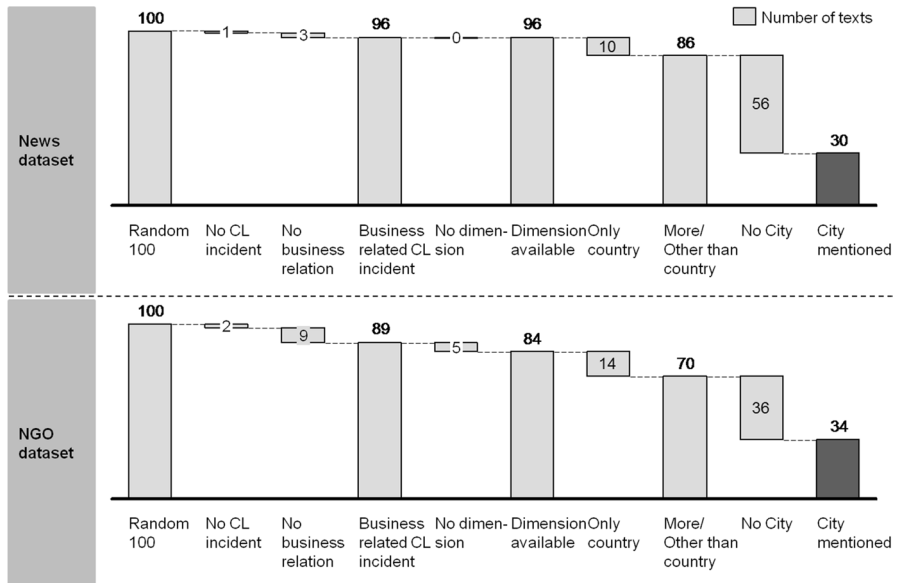


Fig. 10 Analysis of news and NGO dataset based on a random sample of 100 items

definition used in this thesis but which is too broad to be considered a dimension for text mining. Furthermore, only 24 cases mention only the country, but in many cases the additional detail does not go significantly beyond this. In fact, most articles also provide the sector (not shown above) without giving any details as granular as a geographic reference to the city level. However, the sample from the NGO dataset has more geographically detailed cases. Analyzing the types of links in the NGO dataset random sample reveals that a large share of NGO posts redirect to classic news pages such as The Guardian¹ when the tiny URLs² in the posts are expanded. Nevertheless, many unique references (41 in total) still link to non-classic news pages such as blogs, NGO websites or videos (with descriptions), and special news pages or special websites. Thus, one can state that publicly available sources provide child labor incident observations in sufficient number and granularity.

2.6 Determination of breach likelihood

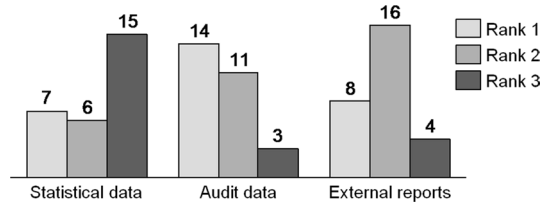
The final node of the BN models is the likelihood of a breach of child labor compliance standards P_{breach} . It combines the contextual prior with the audit and observational likelihoods, thus revising its prior. Its distribution is modeled via a sampling process. Netica creates this using a Monte Carlo sampling based on the model

¹ <http://www.theguardian.com/>.

² Tiny URLs provide an ID-based redirect to longer URLs in order to be able to better post URLs in social media settings with character size constraints.

Table 9 Average estimated relative importance of independently found probability values for overall supplier risk judgment

Value	Probability based on statistical data (e.g., child labor statistics)	Probability based on audit data (e.g., last audit result/time)	Probability based on evidence from external reports (e.g., news)
Average (percent)	23.8	42.9	33.3
Std. deviation	15.7	18.4	14.7

Fig. 11 Frequency of ranks calculated from weights attributed to different sources of evidence

equations, i.e. calculating the result for each of the nodes of the BN based on the equations outlined above. In order to determine the weight to be used for the three contributing nodes, the experts were asked “Which weight would you give the following three probabilities if they are combined in order to calculate an overall probability of a child labor incident at a supplier location? The sum should equal to 100%.” Table 9 summarizes the answers. It turns out that audits are still seen as providing the most important source of information, being most frequently weighted highest (Fig. 11). In contrast, statistics are seen as the least important for a ranking. All three are significantly different from zero (0.000 level).

Additionally to these values, the final node incorporates the distinction between supplier locations that have signed the “code of conduct” and (potential) supplier location that have not. It is assumed that suppliers need to comply with the code of conduct irrespective of a signature. However, not signing it increases the breach risk significantly. Thus, for locations that have signed “codes of conducts”, the prior probability’s mean is shifted by a user-defined factor, for which 0.25 was used. This factor is modeled via a discrete node with two states representing whether the supplier has signed a code of conduct or not. The mean of the resulting final node can be used to establish the prioritization of the supplier locations.

3 System test

Pitchforth and Mengersen (2013) outlined seven validity tests that can be performed in the context of a BN. Out of these, nomological, face and content validity have been discussed during the description of the sub-parts of the system above. Given that the network depicted here does not reuse parts of other networks, concurrent

Table 10 Supplier locations used for validation

Supplier	Country	Sector	Region	City
Supplier A	Indonesia	Construction	Provinsi Jawa Tengah	Simpang
Supplier B	India	Manufacturing	State of Andhra Pradesh	Tirupati
Supplier C	Indonesia	Construction	Provinsi Jawa Tengah	Purwokerto
Supplier D	Indonesia	Manufacturing	Jawa Barat	Badung

Table 11 Audit scores used as ranking test input

Supplier	Audit score (best score: 5.0)	Time since audit (categories acc. to BN) (months)
Supplier A	4.5	6–24
Supplier B	4.5	> 24
Supplier C	3.5	6–24
Supplier D	4.5	> 24

Table 12 Dimensions of news sources used as ranking test input

News ID	Timestamp	Max. relevance	Credibility	Lowest geography	Supplier	Sector
News 1	1 Sep 2014	0.7	0.9	India	Supplier B	Manufacturing
News 2	1 Sep 2014	0.1	0.9	Indonesia		
News 3	1 Sep 2014	0.5	0.9	Provinsi Jawa Tengah		Construction
News 4	4 Sep 2014	0.7	0.3	Tirupati		Manufacturing

validity is not tested. As there is no other BN available for child labor risk management, we cannot test convergent validity. Consequently, also discriminant validity has not been tested. This paragraph relates to testing predictive validity, particularly focusing on model sensitivity.

In order to test the whole system, we will start out by assuming that four hypothetical supplier locations in Table 10 have equal breach likelihood. Subsequently, the locations are audited, with results depicted in Table 11. After that, the four news on child labor incidents described in Table 12 are entered into the BN. This yields the development of the breach likelihoods and rankings depicted in Fig. 12. Given that the third news item details the second, the update leads to an increase in its maximum relevance. Although they have a good audit result, inputting the additional news sources causes the rank of supplier B to worsen twofold (as can be seen in 2). At the same time, the ranking of supplier D improves to first place after the third news item. These results were presented to the expert panel for discussion. The overall evaluation was positive, and the experts agreed that the system's conclusions were plausible and found the resulting derivation of the ranking of the supplier locations valid.

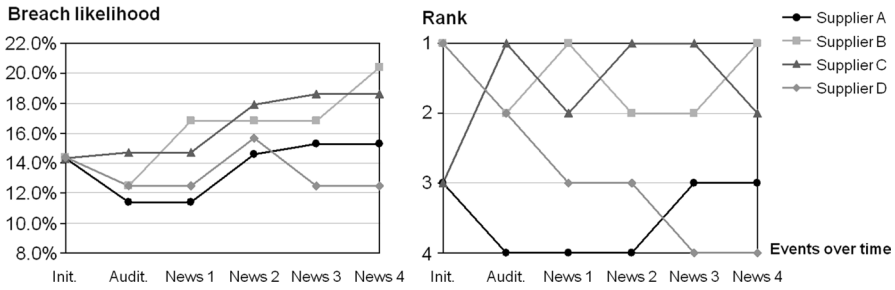


Fig. 12 Development of breach likelihood and supplier rank after audit and news inputs

4 Discussion

With the concept of a BN that estimates the likelihood of a breach in child labor standards at a given supplier location and allows the integration of evidence using relevance and credibility scores, this paper introduces this aspect with a focus on social sustainability. Continuous model updates through new evidence lead to an increase in the likelihood that a supplier will breach a code of conduct. The evidence has to conflict with the basic assumption that suppliers comply with a company’s standards. Additionally, the update process requires that the items entered as evidence are independent of each other and, thus, each update carries additional information that leads to an increase in the breach likelihood. This behavior also has been shown mathematically ([reference deleted for blind review]). Thereby, relevance and credibility are pertinent factors when differentiating the quality of news inputs. Parameters were proposed to attribute a relevance to a news text based on its level of detail. The idea that influence on the overall breach likelihood increases as articles become more detailed was supported by the expert questionnaire. This also underlines the fact that cases which do not occur directly at a supplier location should still affect the breach likelihood if they can be related to a supplier location through an article’s content.

Besides ongoing input from news sources, the network is also updated with new audit results. The more recent audit score is integrated together with the time since the last audit. Concentrating on the last known audit result means the best in-depth data on a supplier is used. This is supported by the expert questionnaire; the experts view audits as the most important source of information when needing to estimate the risk level of supplier locations with regard to child labor. Nonetheless, the experts still assign a residual child labor risk, even if a supplier achieves the best possible audit score. Instead of audit data, this input can also come from certification processes, partner companies, or through platform-based exchange. Focusing on the most recent audit result ignores other possible input from earlier data, which could provide additional conclusions such as diminishing performance.

The prior integrates data from statistical sources to mathematically determine the probability of child labor per country, sector, and area type. The cases tested together with the experts show comparable rankings for supplier locations. However, the

experts provided a wide variety of different answers and rankings, which decreases the interpretability of the results. This variation may be due to the difficulty in estimating the probability value based on the numbers provided without an additional calculation framework, or it may be due to the different heuristics experts use to determine risk probabilities given limited information. Nevertheless, the prior value provides a mathematically derived quantified number, and its components have been agreed on by the experts. In a different context than child labor, another structure for the prior might be necessary given the data availability and underlying driving forces of a different social sustainability factor.

BNs have the advantage that they are more easily understandable than other probabilistic frameworks (Duespohl et al. 2012; Koks and Challa 2005; Wooldridge 2003). The nodes of the BN proposed in this paper can be directly explained to sustainability managers. Moreover, as the expert questionnaire highlights, the requirements incorporated into the BN (together with the surrounding system) are strongly supported by experts in supply chain, sustainability, and risk management. Only one requirement did not see significant agreement—experts want to be able to manually change the final input into the risk model. But as the current system design (apart from the initial configuration) allows unbiased input into further processing steps, the amount of user input, if permitted, has to be discussed in detail. Experts do not necessarily make the correct judgments, and biases may affect the manual input, leading to a questionable ranking of supplier locations. There are also numerous factors that can unconsciously influence decisions (Bazerman 2006). Consequently, also the configuration needs to be cross-checked. Nevertheless, experts nearly significantly attribute job relevance to a model and system incorporating the chosen requirements (at the 6% level, but true for large companies), supporting the overall assumption that the system has a high likelihood of being adopted by an organization.

Altogether, the findings lead to several recommendations for practical application. Using a quantified risk model that is continuously updated based on observations can help to focus a company's auditing resources and other activities such as supplier development where they are most effective. Designing a model and system that fulfills the key requirements and is understandable is feasible. Consequently, its implementation may also allow a more unbiased and objective discussion of social sustainability activities across the supply chain. To be effective, companies would need to have the resources (internal or external) and processes to be able to work with the outcome of the risk model. They need to analyze potential issues more deeply, potentially also performing on-site visits. While the focus here was on child labor, this process can also be adapted for other social sustainability risk sources such as forced labor. A key question when performing a risk-based analysis of suppliers or when a social sustainability issue is detected at a supplier location is how to deal with the particular supplier. The appropriate reaction strongly depends on the type of issue. Besides financial aid for the child or children [as suggested for example by Social Accountability International (2008)], supplier development can offer further opportunities (Harms et al. 2013). Increased commitment, collaboration, and supportive treatment of suppliers can be steps that may already be taken in order to mitigate issues, particularly for high-risk suppliers (Locke et al. 2009).

5 Conclusion

Issues with child labor and other social sustainability themes can cause severe reputational damage to companies, even if found in very remote areas of the supply chain. Moreover, the societal impact of global supply chains has come under particular scrutiny in recent years. This paper suggests using a Bayesian network (BN) continuously fed by reports on child labor observations to estimate the risk of a breach of corporate sustainability standards at a particular supplier location. The BN risk model builds on a statistically derived prior and is updated using the most recent audit results for the location and news items containing child labor issues that can be related to the location. The latter makes use of Bayesian updating to incorporate credibility and relevance of news items as well as their number, yielding an observational probability. Through its probabilistic nature, the BN provides a quantified ranking of supplier locations based on their level of risk, which may be used for further mitigating actions in the supply chain. Domain experts have been asked to provide their input on different requirements and on the calibration of the model.

Clearly, the prototype developed has a number of limitations. Due to the limited availability of data, the nodes of a BN are assumed to be conditionally independent. One could argue that an audit score or news report might be partly related to statistical data depending on the sector or region and given data collected in practice one could build conditional probability tables connecting context, audit results and news. However, the effect of the assumption of independence should be very limited for audits as they mainly leverage company internal information. For news reports, statistical information is typically significantly older than the evidence that can be gathered from news articles, which is often affected by recent events. In addition, as discussed above, news should only influence the breach likelihood strongly if it embodies sufficiently detailed information. Verification has been limited, too, both in terms of the number of experts involved and the number of cases, which were restricted to India and Indonesia.

Thus, one can conclude that interesting further research opportunities exist in all areas covered. While the improvement of event/observation extraction and the extension of the text sources to other languages can be tackled independently from the other issues, the refinement of the BN can only be done on the basis of an enlarged set of real-world data and concurrently with an extended verification process. In order to extend the relevant data basis, it would be advisable that several firms cooperate, maybe also via NGOs and public organizations, who could offer a child labor risk management as software as a service.

Overall, this paper provides a first quantitative risk model for social sustainability monitoring in supply chains based on a Bayesian network and text mining. In organizations, a system building on the techniques suggested would still need some level of manual interaction and adaptation beyond IT, and organizational processes would need to be established to trigger appropriate responses to changes in risk measures. Still, it may be seen as a step towards greater supply chain responsibility.

Acknowledgements Open access funding provided by TU Wien (TUW).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use,

distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Hypothetical news articles

<p>B.2' 2. External data from news and social media^a In the following only Supplier B will be considered. You will be presented several news reports that reveal information that may relate to the supplier. You will be asked how the new report would affect your rating of Supplier B with regard to the probability of a child labor incident (Pinkland, Pinkregion, Pinkcity, urban, Manufacturing)</p>	
<p>B.2.1 2.1 First article received online through classic news channel Greencountry's supplier in CHILD STAFF SHAME: Factory used 'underage workers'—report "Greencountry will do 30 per cent less business with a Pinkcountry's parts maker reportedly caught with underage workers at its factory" despite the supplier's promises to comply with labour laws in future. Greencountry suspended business with Supplier B when US charity Yelloworg claimed it had found at least five child workers without contracts at the supplier. Pinkcountry's authorities discovered that Supplier B was not directly responsible for employing the children, finding that a subcontractor had hired them through a labour agency. As a result, Greencountry said it would resume business with Supplier B, albeit about a third less than before. [...] How much does the news report influence your perceived probability of a child labor incident at Supplier B</p>	<p>Extremely influential; very influential; somewhat influential; slightly influential; not at all influential</p>
<p>B.2.2 2.2 Second article received online through classic news channel Bluecountry, ILO boost co-op The government of Pinkcountry and the International Labor Organization (ILO, the UN specialized agency) have signed a program of technical cooperation for 2014–2016. The document entitled 'Decent Work Agenda' was signed by the Minister of Labour and Social Protection, the chairman of the Chamber of Commerce and the chairman of the Federation of Trade Unions. [...] Under the program, the ILO will support the implementation of the national action plan on elimination of child labor. The program also aims to improve working conditions and employment in agriculture. [...] How much does the news report influence your perceived probability of a child labor incident at Supplier B</p>	<p>Extremely influential; very influential; somewhat influential; slightly influential; not at all influential</p>

- | | |
|--|--|
| <p>B.2.3 2.3 Third article received online through classic news channel</p> <p>Over 10,000 kids working in Pinkregion</p> <p>As per the statistics, of the 10,009 child workers in Pinkregion, 4738 children are available in the job market “seeking work”</p> <p>Pinkregion may pride itself for its high human development indicators, but the latest statistics of Census 2011 released on Wednesday do not bear good news on this front. A total of 10,009 children in Pinkregion, aged between 5 and 14 are employed and shockingly 6920 of these children are working as ‘main workers’ or have been employed constantly over a prolonged period of time. Despite implementation of the Right of Child to Free and Compulsory Education (RTE) Act, which makes schooling mandatory for children up to the age of 14, there are 5351 male children and 4658 female children who have seemingly become victims of child labour in Pinkregion. [...] Of these, around 1300 have worked as manual workers. [...] Experts in the field of economics, urban development and statistics present at the session analyzing the latest Census data, said that though statistics are unable to reveal the nature of employment, generally a majority of these children appear to be employed as domestic help in households, as waiters in small eateries and as help along with parents in manual labour jobs such as manufacturing</p> <p>How much does the news report influence your perceived probability of a child labor incident at Supplier B</p> | <p>Extremely influential; very influential; somewhat influential; slightly influential; not at all influential</p> |
| <p>B.2.4 2.4 Fourth article received online through social media posting</p> <p>Six minors rescued</p> <p>NGO Greyorg’s staff have rescued six minors from a State Transport bus in Greencity. They were being trafficked to Pinkcity to work as child labourers, an official said Tuesday. A man who was taking the children along was arrested and is being interrogated. The children, in the age group of 12–16 years, were sent to a shelter Monday night after the Child Welfare Committee (CWC) chairperson visited them and requested the police to take appropriate action</p> <p>The coordinator of Greyorg’s special anti-trafficking project in Greencity said he and his colleague saw a middle-aged man with six children boarding a bus at Greencity’s district. [...]</p> <p>Children from the state are trafficked to Pinkcity to work in construction. Though the situation has improved over the years, cases of trafficking still surface</p> <p>How much does the news report influence your perceived probability of a child labor incident at Supplier B</p> | <p>Extremely influential; very influential; somewhat influential; slightly influential; not at all influential</p> |

^aThe actual survey data has been presented with visual emphasis and color coding

Appendix 2: Hypothetical news articles

ID	Data-set	Date	Type	Text
1	News	Oct 09 2013	General report	Title: "ILO will miss 2016 target to end worst forms of child labor"
2	News	Jul 28 2014	General report	Title: "Explainer: Child Labor Legislation in South America"
3	NGO	Jan 16 2013	General report	"More than 1.3 million children in Yemen are involved in child labour, including 469,000 in the 5–11 age group, according to a new study. This means 17 per cent of Yemen's 7.7 million children in the 5–17 age group and 11 per cent of those aged 5–11 are involved in child labour, according to the study which is Yemen's first-ever national child labour survey carried out with support from the International Labour Organization, the Social Development Fund and UNICEF. [...]"
4	NGO	Aug 28 2013	Multiple incidents	"As many as 36 cases from Koderma and 22 from Khunti were brought to Dube's notice when he visited these districts. Such incidents include a girl from Khunti missing since 2009 when she went to Delhi for work. [...] The survivors also discussed their fate and the torture they were subjected to while they worked in Delhi and other big cities"
5	NGO	Aug 26 2013	Multiple incidents	"He and 11 other children from his village were taken by bus on a 2100 km (1300 miles) journey and put to work in Ho Chi Minh City (formerly Saigon), south Vietnam. [...] The charity helps children forced into a variety of jobs from prostitution to begging, but in the past year just over a quarter of that number have been rescued from garment factories in Ho Chi Minh City, the country's largest metropolis and industrial centre"
6	NGO	Dec 22 2012	Incident report, narratives	"Aslam, twelve, is a native of Sipur village, Azam Nagar in the Katihar district. Despite being promised a good education he too ended up in the same dingy sweatshop in Delhi. Like the other rescued children he worked very long days, sleeping in the same room where he worked"
7	NGO	Jun 13 2012	Incident report	"A young Indian bonded child laborer cries as he is walked away after being rescued during a raid by workers from Bachpan Bachao Andolan, or Save the Childhood Movement, at a garment factory in New Delhi, India, Tuesday, June 12, 2012"

ID	Data-set	Date	Type	Text
8	News	Aug 29 2014	Incident report (Samsung)	“New York-based China Labor Watch (CLW) claimed in a report Thursday that Samsung supplier HEG Technology had hired people under the age of 16 at its facility in Huizhou, China. The report said the factory, where components for Lenovo are also built, had hired new workers, [...]”
9	News	Jun 19 2014	Incident report	“Twelve minors, working as labourers in various garages, hotels, bars and dhabas in and around Chhatrapur in Ganjam district, were rescued on Wednesday”
10	News	Jul 14 2014	Incident report	“Samsung’s audit was conducted after the campaign group China Labor Watch and found evidences of illegal hiring process that took place on June 29. According to the campaign group: “During the busy season and in urgent need of labor, Shinyang hires child labor and underage student workers”
11	NGO	Jun 13 2012	Incident report	“Social activists, with the support of the police, have raided a New Delhi factory employing children aged between six and 13. The children, who had been working there for more than a year, came from other regions of India like West Bengal, Bihar and Uttar Pradesh, according to the activists”

Sources (referenced by ID): (1) <http://www.trust.org/item/20131009102946-p4t5j/>, (2) <http://www.as-coa.org/articles/explainer-child-labor-legislation-south-america>, (3) Tweet ID: 291591675367797000, <http://www.unmultimedia.org/radio/english/2013/01/more-than-1-3-million-child-labourers-in-yemen-ilo/>, (4) Tweet ID: 3726461851162378e17, <http://timesofindia.indiatimes.com/city/ranchi/NCPCR-pledges-action-against-child-trafficking/articleshow/22105876.cms>, (5) Tweet ID: 371912419372986000, <http://www.bbc.co.uk/news/world-asia-23631923>, (6) Tweet ID: 282409517118001000, http://www.huffingtonpost.com/gordon-brown/india-child-labor-laws_b_2345756.html, (7) Tweet ID: 212881625070051000, <http://www.utsandiego.com/news/2012/jun/12/factory-raids-reveal-child-labor-persists-in-india/>, (8) <http://www.bangkokpost.com/tech/computer/429443/samsung-denies-child-labour-at-chinese-supplier>, (9) <http://timesofindia.indiatimes.com/city/bhubaneswar/12-child-labourers-rescued-in-Ganjam/articleshow/36808587.cms>, (10) <http://www.digitaljournal.com/news/world/samsung-has-used-child-labor-in-china/article/389632>, (11) Tweet ID: 212825335249641000, <http://www.aljazeera.com/video/asia/2012/06/2012612164116177802.html>

References

- Ahi P, Searcy C (2014) Assessing sustainability in the supply chain: a triple bottom line approach. *Appl Math Model*. <https://doi.org/10.1016/j.apm.2014.10.055>
- Aktiv Gegen Kinderarbeit (2014) OTTO group. http://www.aktiv-gegen-kinderarbeit.de/firma/otto-group/#footnote_8_3871. Accessed 13 Jan 2015
- Anderson DR, Anderson KE (2009) Sustainability risk management. *Risk Manag Insur Rev* 12(1):25–38
- Ashby A, Leat M, Hudson-Smith M (2012) Making connections: a review of supply chain management and sustainability literature. *Supply Chain Manag Int J* 17(5):497–516
- Badurdeen F, Shuaib M, Wijekoon K, Brown A, Faulkner W, Amundson J et al (2014) Quantitative modeling and analysis of supply chain risks using Bayesian theory. *J Manuf Technol Manag* 25:631–654. <https://doi.org/10.1108/JMTM-10-2012-0097>
- Bazerman MH (2006) *Judgment in managerial decision making*, 6th edn. Wiley, New York

- Bennett S (2012) REVEALED: the top 20 countries and cities on Twitter [STATS], SocialTimes. SocialTimes. Aug 13. <http://www.adweek.com/socialtimes/twitter-top-countries/468210>. Accessed 11 Feb 2015
- Benoît C, Vickery-Niederman G (2011) Social sustainability assessment literature review. White paper #102. The sustainability consortium, Arizona State University, and University of Arkansas. http://www.sustainabilityconsortium.org/wp-content/themes/sustainability/assets/pdf/whitepapers/Social_Sustainability_Assessment.pdf. Accessed 04 Nov 2012
- Blasch E, Laskey KB, Joussemle AL, Dragos V, Costa PCG, Dezert J (2013) URREF reliability versus credibility in information fusion (STANAG 2511). In: Proceedings of 16th international conference on information fusion. FUSION'13. IEEE, pp 1600–1607
- Boyd DE, Spekman RE, Kamauff JW, Werhane P (2007) Corporate social responsibility in global supply chains: a procedural justice perspective. *Long Range Plan* 40(3):341–356
- BPS Statistic Indonesia (2008) The Indonesian 2006 economic census. The 12th East Asian statistical conference. Tokyo. <http://www.stat.go.jp/english/info/meetings/eastasia/pdf/t2indpa.pdf>. Accessed 20 Oct 2014
- Bremer J, Udovich J (2001) Alternative approaches to supply chain compliance monitoring. *J Fash Mark Manag* 5(4):333–352
- British Telecommunications (2012) Assessing environmental and social risks. <http://www.btplc.com/Responsiblebusiness/Ourstory/Sustainabilityreport/section/index.aspx?sectionid=68d95d30-36ed-4f1f-a138-6344031a2bbe>. Accessed 4 Dec 12
- Ceres and Sustainalytics (2014) Gaining ground: corporate progress on the ceres roadmap for sustainability. <http://www.ceres.org/resources/reports/gaining-ground-corporate-progress-on-the-ceres-roadmap-for-sustainability/view>. Accessed 15 Jan 15
- Chin KS, Tang TW, Yang JB, Wong SY, Wang H (2009) Assessing new product development project risk by Bayesian network with a systematic probability generation methodology. *Expert Syst Appl* 36:9879–9890
- Ciliberti F, Pontrandolfo P, Scozzi B (2008) Investigating corporate social responsibility in supply chains: a SME perspective. *J Clean Prod* 16(15):1579–1588
- Correa M, Bielza C, Pamies-Teixeira J (2009) Comparison of Bayesian networks and artificial neural net-works for quality detection in a machining process. *Expert Syst Appl* 36(3, Part 2):7270–7279
- Cover T, Thomas J (1991) Elements of information theory. Wiley, New York
- Crystal D (2004) Subcontinent raises its voice. *The guardian*. November 19. <http://www.theguardian.com/education/2004/nov/19/tefl>. Accessed 11 Feb 2015
- Dabrowski JJ, Beyers C, Villiers JP (2016) Systemic banking crisis early warning systems using dynamic Bayesian networks. *Expert Syst Appl* 62:225–242
- Diallo Y, Etienne A, Mehran F, International Labour Office, International Programme on the Elimination of Child Labour (2013) Global child labour trends 2008 to 2012. ILO, Geneva
- Duespohl M, Frank S, Doell P (2012) A review of Bayesian networks as a participatory modeling approach in support of sustainable environmental management. *J Sustain Dev* 5(12):1–18
- Edwards W (1954) The theory of decision making. *Psychol Bull* 51(4):380–417
- Egels-Zandén N (2007) Suppliers' compliance with MNCs' codes of conduct: behind the scenes at Chinese toy suppliers. *J Bus Ethics* 75(1):45–62
- Erol I, Sencer S, Sari R (2011) A new fuzzy multi-criteria framework for measuring sustainability performance of a supply chain. *Ecol Econ* 70:1088–1100. <https://doi.org/10.1016/j.ecolecon.2011.01.001>
- Fuge M, McKinstry K, Ninomiya K (2013) Impactmap: designing sustainable supply chains by incorporating data uncertainty. In: Bilec M, Choi J (eds) In: Proceedings of the international symposium on sustainable systems and technologies. Sustainable Conoscente Network, Oakland. http://www.markfuge.com/papers/fuge_mckinstry_ninomiya_ISSST2013_preprint.pdf. Accessed 21 Jan 2015
- Gangemi A (2013) A comparison of knowledge extraction tools for the semantic web. In: Cimiano P et al (eds) *The semantic web: semantics and big data*. Lecture notes in computer science, vol 7882. Springer, Berlin, pp 351–366
- Grimm JH, Hofstetter JS, Sarkis J (2016) Exploring sub-suppliers' compliance with corporate sustainability standards. *J Clean Prod* 112:1971–1984
- Groth SS, Muntermann J (2011) An intraday market risk management approach based on textual analysis. *Decision Support Systems*. *Enterprise Risk and Security Management: Data, Text and Web Mining* 50(4):680–691
- Guarcello L, Kovrova I, Lyon S, Manacorda M, Rosati FC (2010) Towards consistency in child labour measurement: assessing the comparability of estimates generated by different survey instruments.

- Un-understanding children's work programme working paper series. http://www.ucw-project.org/attachment/st_Towards_consistency_in_child_labour_measurement20110517_123409.pdf. Accessed 20 Oct 2014
- Hadiguna RA (2012) Decision support framework for risk assessment of sustainable supply chain. *Int J Logist Econ Glob*. <https://doi.org/10.1504/ijleg.2012.047213>
- Harms D, Hansen EG, Schaltegger S (2013) Strategies in sustainable supply chain management: an empirical investigation of large german companies. *Corp Soc Responsib Environ Manag* 20:205–218. <https://doi.org/10.1002/csr.1293>
- Holton C (2009) Identifying disgruntled employee systems fraud risk through text mining: a simple solution for a multi-billion dollar problem. *Decis Support Syst* 46(4):853–864
- Hubbard G (2009) Measuring organizational performance: beyond the triple bottom line. *Bus Strategy Environ* 18(3):177–191
- Hutchins MJ, Sutherland JW (2008) An exploration of measures of social sustainability and their application to supply chain decisions. *J Clean Prod* 16(15):1688–1698
- International Labour Organization (1973) Convention C138—minimum age convention, 1973 (no. 138). NORMLEX. http://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO:12100:P12100_ILO_CODE:C138. Accessed 2 Dec 2013
- International Labour Organization (1999) Convention C182—worst forms of child labour convention, 1999 (no. 182). NORMLEX. http://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_ILO_CODE:C182. Accessed 2 Dec 13
- International Programme on the Elimination of Child Labour (2013) Marking progress against child labour—global estimates and trends 2000–2012. ILO, Geneva
- Keating B, Quazi A, Kriz A, Coltman T (2008) In pursuit of a sustainable supply chain: insights from Westpac Banking corporation. *Supply Chain Manag Int J* 13(3):175–179
- Kenyon P, Campbell F, Hawkey E (2000) Gap and nike: no sweat? October 15 2000. BBC, October 15. <http://news.bbc.co.uk/2/hi/programmes/panorama/970385.stm>. Accessed 24 July 2013
- Klassen RD, Vereecke A (2012) Social issues in supply chains: capabilities link responsibility, risk (opportunity), and performance. *Int J Prod Econ* 140(1):103–115
- Kogg B, Mont O (2012) Environmental and social responsibility in supply chains: the practise of choice and inter-organisational management. *Ecol Econ* 83:154–163
- Koks D, Challa S (2005) An introduction to Bayesian and Dempster-Shafer data fusion. DSTO Systems Sciences Laboratory, Edinburgh. <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA417895>. Accessed 13 Feb 2013
- Koplin J, Seuring S, Mesterharm M (2007) Incorporating sustainability into supply management in the automotive industry—the case of the Volkswagen AG. *J Clean Prod* 15(11–12):1053–1062
- Leidner JL, Schilder F (2010) Hunting for the black swan: risk mining from text. In: Proceedings of the ACL 2010 system demonstrations, pp 54–59
- Leire C, Mont O (2010) The implementation of socially responsible purchasing. *Corp Soc Responsib Environ Manag* 17(1):27–39
- Lemke F, Petersen HL (2013) Teaching reputational risk management in the supply chain. *Supply Chain Manag* 18(4):413–429
- Li Y, Bontcheva K, Cunningham H (2005) SVM based learning system for information extraction. In: Winkler J, Niranjana M, Lawrence N (eds) Deterministic and statistical methods in machine learning. Lecture notes in computer science, vol 3635. Springer, Berlin, pp 319–339
- Locke R, Qin F, Brause A (2007) Does monitoring improve labor standards? Lessons from Nike. *ILR Rev* 61(1):3–31
- Locke R, Amengual M, Mangla A (2009) Virtue out of necessity? Compliance, commitment, and the improvement of labor conditions in global supply chains. *Politics Soc* 37(3):319–351
- Mani V, Delgado C, Hazen BT, Patel P (2017) Mitigating supply chain risk via sustainability using big data analytics: evidence from the manufacturing supply chain. *Sustainability* 9:608. <https://doi.org/10.3390/su9040608>
- March JG (1994) Premier on decision making: how decisions happen. Simon and Schuster, New York
- McDougall D, Schmitz S (2007) Otto-Konzern: Kinderarbeit Für Den Heine-Versand. *Stern*. February 11. http://www.stern.de/wirtschaft/news/otto-konzern-kinderarbeit-fuer-den-heine-versand-581923.html?p=2&nv=ct_cb. Accessed 13 Jan 2015
- Mcintyre K, Smith HA, Henham A, Pretlove J (1998) Logistics performance measurement and greening supply chains: diverging mindsets. *Int J Logist Manag* 9(1):57–68

- Miemińczyk J, Johnsen TE, Macquet M (2012) Sustainable purchasing and supply management: a structured literature review of definitions and measures at the dyad, chain and network levels. *Supply Chain Manag* 17(5):478–496
- Ministry of Statistics and Programme and Implementation of India (2006) Provisional results of economic census 2005—all India report. New Delhi. http://mospi.nic.in/mospi_new/upload/economic_census_prov_results_2005.pdf. Accessed 12 April 2014
- Neapolitan RE (2003) *Learning Bayesian networks*. Prentice Hall, Chicago
- Nguyen T, Zhou L, Spiegler V, Ieromonachou P, Lin Y (2017) Big data analytics in supply chain management: a state-of-the-art literature review. *Comput Oper Res*. <https://doi.org/10.1016/j.cor.2017.07.004>
- Nobelprize.org (2014) Kailash Satyarthi—nobel lecture: let us globalise compassion, and set our children free. Nobel media AB 2014. http://www.nobelprize.org/nobel_prizes/peace/laureates/2014/satyarthi-lecture_en.html. Accessed 14 April 2015
- Norsys (2013) Netica application. <http://www.norsys.com/netica.html>. Accessed 19 Oct 2013
- Pearl J (1988) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers Inc., San Francisco
- Pedersen ER, Andersen M (2006) Safeguarding corporate social responsibility (CSR) in global supply chains: how codes of conduct are managed in buyer–supplier relationships. *J Public Aff* 6(3–4):228–240
- Pitchforth J, Mengersen K (2013) A proposed validation framework for expert elicited Bayesian Networks. *Expert Syst Appl* 40:162–167
- Quinlan JR (1993) *Programs for machine learning*. Morgan Kaufmann Publishers, San Francisco
- Rabenasolo B, Zeng X (2012) A risk-based multi-criteria decision support system for sustainable development in the textile supply chain. In: Lu J, Jain LC, Zhang G (eds) *Handbook of decision making*. Springer, Berlin, pp 141–170
- Reckhow KH (1999) Water quality prediction and probability network models. *Can J Fish Aquat Sci* 56(7):1150–1158
- Reuters, National Institute of Standards and Technology (2009) Reuters TRC2 corpus. <http://trec.nist.gov/data/reuters/reuters.html>. Accessed 23 Dec 2013
- Rivera SJ, Minsker BS, Work DB, Roth D (2014) A text mining framework for advancing sustainability indicators. *Environ Model Softw* 62:128–138
- Rizzo G, Troncy R, Hellmann S, Bruemmer M (2012) NERD meets NIF: lifting NLP extraction results to the linked data cloud. LDOW 937. <http://ceur-ws.org/Vol-937/>
- Sarkis J, Dhavale DG (2014) Supplier selection for sustainable operations: a triple-bottom-line approach using a Bayesian framework. *Int J Prod Econ*. <https://doi.org/10.1016/j.ijpe.2014.11.007>
- Sarkis J, Zhu Q, Lai K (2011) An organizational theoretic review of green supply chain management literature. *Int J Prod Econ* 130(1):1–15
- Seuring S, Müller M (2008) Core issues in sustainable supply chain management—a Delphi study. *Bus Strategy Environ* 17(8):455–466
- Shanteau J (1988) Psychological characteristics and strategies of expert decision makers. *Acta Psychol* 68:203–215. [https://doi.org/10.1016/0001-6918\(88\)90056-X](https://doi.org/10.1016/0001-6918(88)90056-X)
- Shokravi S, Kurnia S (2014) A step towards developing a sustainability performance measure within industrial networks. *Sustainability* 6:2201–2222. <https://doi.org/10.3390/su6042201>
- Social Accountability International (2008) Social accountability 8000:2008. http://www.sa-intl.org/_data/n_0001/resources/live/2008StdEnglishFinal.pdf. Accessed 11 Jan 2013
- Svensson G (2009) The transparency of SCM ethics: conceptual framework and empirical illustrations. *Supply Chain Manag Int J* 14(4):259–269
- Taylor MB, Zandvliet L, Forouhar M (2009) Due diligence for human rights: a risk-based approach. Working paper no. 53. Corporate social responsibility initiative. http://www.hks.harvard.edu/mrcbg/CSRI/publications/workingpaper_53_taylor_etal.pdf. Accessed 25 Oct 2013
- Thoeni A (2015) Sustainability risk monitoring in supply chains: ranking suppliers using text mining and Bayesian networks with a focus on child labor. Dissertation: Vienna University of Technology
- The World Bank Group (2013) World development indicators, The World Bank. <http://wdi.worldbank.org/table/2.6>. Accessed 5 March 2014
- Tsukayama H (2012) Foxconn and apple supply: what are the effects? *The Washington Post*. September 24. http://www.washingtonpost.com/business/economy/foxconn-and-apple-supply-what-are-the-effects/2012/09/24/126b9966-0670-11e2-aff-d6c7f20a83bf_story.html. Accessed 04 Dec 2012

- Understanding Children's Work (2009) Indonesia, child labour survey 2009. <http://www.ucw-project.org/Pages/Tables.aspx?id=1372>. Accessed 11 Feb 2015
- Understanding Children's Work (2010) India, national sample survey round 66 (NSS-R66). <http://www.ucw-project.org/Pages/Tables.aspx?id=1540>. Accessed 10 April 2014
- Understanding Children's Work (2014) UCW project—youth employment outcomes, youth employment statistics, child labour legal standards, understanding child labour, understanding employment outcomes, domestic chores. <http://www.ucw-project.org/pages/childdlabindicator.aspx>. Accessed 8 Aug 2014
- Uusitalo L (2007) Advantages and challenges of Bayesian networks in environmental modelling. *Ecol Model* 203(3–4):312–318
- UNEP (2009) Guidelines for social life cycle assessment of products. United Nations Environment Programme. <http://socialhotspot.org/userfiles/guidelines-sLCA.pdf>. Accessed 04 Nov 12
- Vermeulen WJV, Seuring S (2009) Sustainability through the market—the impacts of sustainable supply chain management: introduction. *Sustain Dev* 17(5):269–273
- Wang G, Gunasekaran A, Ngai EWT, Papadopoulos T (2016) Big data analytics in logistics and supply chain management: certain investigations for research and applications. *Int J Prod Econ* 176(2016):98–110. <https://doi.org/10.1016/j.ijpe.2016.03.014>
- Wathayu W, Peng Y (2004) A Bayesian network based framework for multi-criteria decision making. In: Proceedings of the 17th international conference on multiple criteria decision analysis. MCDM'04. Whistler
- Weber O, Scholz RW, Michalik G (2010) Incorporating sustainability criteria into credit risk management. *Bus Strategy Environ* 19:39–50. <https://doi.org/10.1002/bse.636>
- Wigas WM (2006) Likert-type scale response anchors. Clemson International Institute for Tourism & Research Development. <http://www.clemson.edu/centers-institutes/tourism/documents/sample-scales.pdf>. Accessed 27 Oct 2014
- Wooldridge S (2003) Bayesian belief networks. CSIRO Centre for Complex Systems Science. <http://cites.eerx.ist.psu.edu/viewdoc/download?doi=10.1.1.112.6230&rep=rep1&type=pdf>. Accessed 13 Dec 12
- Wu KJ, Liao CJ, Tseng ML, Lim MK, Hu J, Tan K (2017) Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties. *J Clean Prod* 142(2017):663–676. <https://doi.org/10.1016/j.jclepro.2016.04.040>
- Zha H (2002) Generic summarization and keyphrase extraction using mutual reinforcement principle and sentence clustering. In: Proceedings of the 25th annual international ACM SIGIR conference on research and development in information retrieval. SIGIR'02. ACM, New York, pp. 113–120