

World Conference on Technology, Innovation and Entrepreneurship

Generating Digital Reputation Index: A Case Study

Sadi Evren Seker^{a*}, Enes Eryarsoy^b

^a*Istanbul Medeniyet University, Dept. of Business, Goztepe, Kadikoy, Istanbul, Turkey*

^b*Istanbul Sehir University, Dept. of International Management, Altunizade, Uskudar, Istanbul Turkey*

Abstract

Digital world is transforming wares to soft-wares, business to e-business, brick and mortar companies to online companies. With its exponentially increasing nature the transformation is continuing. One crucial value of any company, which is interrelated with all the sub divisions and operations of the company, can be considered as its reputation. This paper mainly discusses the quantitative face of reputation from the social capital perspective and its transformation by the increasing and irrepressible power of technology. We attempt to answer the question “how to represent the digital reputation of the companies in a digital world?” under the effect of transformation from web 1.0 to web 3.0. As a solution we propose a quantitative methodology for aggregating digital quantities collected from social network sites, company web pages, blogs and wikis and we propose a formulation and indexing method, built on different dimensions of digital world and by the way, the companies can be ranked respectively. As a case study, we focus on the stock market companies in Turkey and their digital reputation and we output a digital reputation index but more importantly we discuss the methodology of creating a digital index for the reputation of companies. Finally we conduct diagnostics on the output index to assess its degree of validity. We believe the research will be a guidance for the research studies in the digital reputation index, since it is one of the first index creation methodology on digital reputation studies and data source has a great variety, volume and velocity as the big data.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of Istanbul Univeristy.

Keywords: Reputation Management; Reputation Index; Digital Reputation; Big Data, Social Medai Analysis; Web 2.0

* Corresponding author. Tel.: +9-0532-446-7882.

E-mail address: academic@sadievrenseker.com

1. Introduction

Many organizations use the Internet to increase their visibility, and in return their brand equities. They mostly focus their efforts on immediate gains such as gains via increased visibility on the Internet. This visibility, in the context of marketing, is an inseparable part of share of voice of a firm/brand, and is one of the major factors affecting market shares in longer run. The firms are in need of a mirror that can shows their visibilities.

Aim of this study is to propose a methodology in order to create a recurring online digital reputation index which can mirror firms' digital visibilities. The term digital reputation index can be defined as aggregate opinion about a social entity as a result of a social evaluation given a set of criteria with a positive connotation (Seker, Cankir, & Arslan, 2014). For example any entity including, a brand, a person or a company, an institution can have digital reputation. The digital reputation of an entity may be quantified through measuring positive remarks, online visibility, public opinion/remarks about that entity. Similarly an index can be built upon measured digital reputations in one field or across many fields. For instance for textile industry, celebrities, governmental offices, sport clubs, universities, etc. (Arslan & Seker, 2014).

As we create an index for digital reputations, we investigate different approaches for creating indices, we propose different verification methods to test the validity of the indices, and we develop necessary metrics and techniques to be further included for recurring generation of index development.

This study proposes a methodology for creating a digital reputation index. The paper is organized accordingly. First we briefly review the literature for index creation, then for digital reputation. We briefly talk about the numeracy of indices creation.

Many earlier studies already addressed index creation for different domains on different subjects. For instance, Kızıl, Arslan & Seker (2013) proposed an index for BIST30 (top 30 traded companies at Istanbul Stock Exchange). This paper, however, has two novel contributions: 1) The methodology has been devised from social science indexing approach. 2) The data available for research was not available prior to this study, and it spans a one year. Therefore the study is built on a more robust and detailed dataset.

2. Literature Review

This section is organized on two subsections. The first subsection provides a literature review about index creation process. The second subsection reviews the literature on digital reputation studies.

2.1. Index creation

Real world entities are multi-dimensional and often difficult to assess. Indices contain ratings and rankings of such entities from different points of views. Although an index may not be sufficient to rank/rate entities with perfect accuracy, they are useful and are often required as a proxy to perfect ranking in order to be able to monitor and report formed opinions perceptions. For example, success of an online brand depends on its prevalence of its advertisements, online testimonies, endorsements, banners and so on. These together forms or mirrors aggregate feelings and opinions of individuals about the brand, and influences how they make their purchasing decisions. Individual decisions on aggregate level often reflect perceived reputation rankings of the entities (brands, in this example), and aggregate perceptions form the rankings. Hence reputation of products or services becomes a new form of currency (Adam & Nicolai, 2013).

Indices are favored due to several reasons. Foa and Tanner (2012) list four main benefits of using indices. First, indices can map multi-dimensional data into a single, and easy to comprehend dimension. Second, thanks to their simplicity indices can provide an easier interpretation for difficult concepts requiring otherwise-complex analyses. Third, creation of such indices makes sure that all stakeholders such as consumers, companies, and regulators are periodically monitored, measured, and included. Fourth, they are a good starting point for discussions and comparisons (Foa & Tanner, 2012).

Developing an index requires paying attention to several aspects. There is no one-clear way of deriving an index however wider set of variables are necessary for fuzzier concepts and for weaker datasets. Index creation in social studies follow certain steps. Four steps crucial in index creation can be summarized as below:

1. Selection of data sources and variables
2. Weighting the data sources
3. Building a statistical method for index
4. Evaluation and diagnosis of the index

First step is creation of constructs as building blocks. In our case, we identified different constructs for digital reputation. Then data sources or variables are identified, and selected. Reliability and variability of data sources affect the accuracy and credibility of an index. A limited data source or a narrower variable set may result in a questionable or incredible index.

After variable selection, the next step is to decide on weight assignment to indicator variables/constructs. Ranging from use of equal weights to variable weights, or even PCA-based or regression-based weights, different weighting schemes are possible. Avoiding complexity at the expense of reduced accuracy is often possible. Indices rarely contain complex weighting schemes mainly due to the fact that they are not easy to explain, and weights themselves may change overtime. Another problem during weighting is missing data. For example during the creation of an index, sometimes there are missing data from some sources and methods like regression analysis are susceptible to bias or even error in the case of missing data. Although some imputation methods like column/list-wise deletion may help cleaning or hot/cold deck or mean/stochastic imputation exist, these undermines the models accuracy and increases its complexity while reducing its interpretability. However, there are indices that use methods as complex as Markov Chain Monte Carlo simulations for missing value imputation, such as Environmental Sustainability Index (ESI). Deciding on the weighting, and/or imputation has serious implications from the statistical point of view.

After creating an indexing method based on the variety of data sources and weights, the index creation also requires a diagnostic phase to measure its success. Techniques such as confirmatory factor analysis, outlier detection, cluster analysis are frequently used in the literature. A key requirement is to assess construct validity (Carmines and Zeller 1979). Some major indicators that are useful for the diagnostic test are outlier variables, convergent validity and appropriate weighting removal (Foa & Tanner 2012).

2.2. Digital Reputation

Digital reputation can be defined at different levels such as Individual, Product/Brand, Application, or Company. At the individual level reputation can be defined as the trust or mistrust towards an individual. Digital reputation is a proxy to reputation of an individual. According to the European Court of Justice (EJC) Google is a reputational mirror (ruling on “right to be forgotten” on May 13th 2014).

At organizational level, on the other hand, reputation is an indirect measure of the aggregate opinion of people, social groups or other organizations about the organization. This reputation in digital world corresponds to the digital reputation which can be measured at three levels: web 1.0 with static web pages, web 2.0 with dynamic and interactive contents like wikis, blogs, social networks and web 3.0 with semantic web. Most of the studies in the field are concentrated on social networks and web page contents. Measuring digital reputation is still blurry and fuzzy.

There have been different types of studies on the topic. Some studies focus on the reputation on electronic markets, such as Amazon or eBay, and study the reputation of each member and build a selection mechanism based on the reputation of member (Ghose, Ipeirotis, & Sundararajan, 2006). Another group of researchers studied value added by reputation across the markets. Their researches show that the price of a product can be increased depending on the reputation of company in terms of quality (Kato, 2006).

Even though there is no standard way of associating digital reputation to brand equity, some 3rd party trackers, such as Reputation.com, track the online reputation as a service. But the return of investment on such a service is still ambiguous (Economist, 2013). However, it is important to note that a reputation management as new science has an increasing importance today. For example the stock market prices of companies may be linked to the

reputation of the company. Some reputation tracking companies link company related content on the Internet to company's share prices. An Australian company, Cubit Media Research, whose clients include Microsoft, SAP and Ford provides such a service (Engardio & Arndt, 2007).

We believe that these studies mainly suffer from three shortcomings:

1. Narrow Focus (only map web presences onto firm performances. The result of such approach is usually ad-hoc analysis with smaller chance of generalization).
2. Limited Locus (only one sector, or firms of similar size are taken into account. These approaches suffer from generalizability).
3. Direct-relationship assumption: studies attempt to establish a direct link between firm performances and their web presences.

Researches on digital reputation highlight two major online reputation sources that are publicly available. The first source is search engines. They act as a huge gateway that directs the entities, including customers, business partners or competitors to related web pages. The second source is social networks where the people are getting news and making social searches (Madden & Smith, 2010).

Besides the publicly available reputation gateways, there are also closed or semi-closed reputation sources, which shows community-based reputation. The increasing trend and social behavior of technology brings about new online communities. Unlike the classical communities, which are collected around a shared interest, the online communities are sharing projects, working together and producing knowledge. Nowadays reputation management also includes relationship management with these communities/crowds. For example, crowd sourcing is an increasing trend, and organizations need to establish a reputation among crowds. Most of the companies support crowd sourcing with activities such as free technology allocation, early access to some online resources and promotions in order to increase their reputations (Allahbakhsh et al., 2012; Howe, 2009).

Digital reputation studies have a great variety of applications in the literature. For example some studies underlines the correlation between digital reputation and financial crisis (Seker & Georgoevski, Financial Crisis and the ICT Industry, Cross Market Research on Europe, US, Turkish and Gulf Countries, 2014), another research studies the correlation between macro economical information and communication technology (ICT) facts and digital reputation (Seker & Kulaklı, Macroeconomic ICT Facts and Mobile Telecom Operators via Social Networks and Web Pages, 2016) on another study is about the correlation between digital reputation of job web sites and the unemployment in countries (Mert, SEKER, Jamburia, & Temel, 2014) (Arslan & Seker, The Impact of Employment Web Sites' Traffic on Unemployment: A Cross Country Comparison, 2013). A preliminary study about reputation and the stock market quoted companies can be overviewed in some research (Seker, Cankir, & Okur, Strategic Competition of Internet Interfaces for XU30 Quoted Companies, 2014). The main difference in this paper from the previous studies is the index creation approach to the previous reputation based studies.

3. Methodology

The aim of this study is to create a robust, easy to measure metric as an indicative of brands online preference. The metric must be able to generalize domain expert ratings/views, and therefore, and needs to be trivial. In this section, we first describe the method we used to build digital reputation index.

After the Web 2.0, the Internet is now can be viewed as a web of i) social networks, ii) wikis, iii) blogs and iv) web pages. The data sources must measure reputation across these four. Any entity, such as an individual, brand, or an organization can have related entries on Wikipedia as wiki, related content on Facebook, Twitter or LinkedIn as social network, blog as blog or its own web site as a web page, like personal homepages or company web sites. Collection of data from all these four data sources is proposed in this study. The overview of study is demonstrated in Fig. 1.

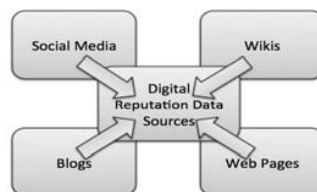


Fig. 1. Digital Reputation data sources

We then selected a wide set of variables for the research.

Table 1. Candidate variable set for measuring digital reputation.

Group	Variable	Explanation
Web Page Statistics	Alexa Country Rank	Country rank of the brand web page on Alexa
	Alexa Global Rank	Global rank of the brand web page on Alexa
	Daily Time on Site	Daily time spent on the web page in 100 hours
	Daily Visitors on Web Page	Number of incoming visitors to the web page.
	Google Pagerank	A number representing the quality of incoming links as a rough estimate of how important that web page is.
	Google Trends	A number denoting relative weight of the brand-related web searches.
	Indexed Pages by Bing	Number of pages indexed by the search engine Bing.
	Number of Backlinks	Number of inbound links to the brands web page.
	Value of the Webpage	Combined value of all web-o-metrics measured by several web pages.
Social Media	Facebook Likes	Number of likes on the official web page of the brand
	Facebook Shares	Number of news/status updates shared by facebook users
	Tweets (Hashtags)	Number of tweets including the brand in hashtags (both positive and negative)
	Linked In	Number of followers on LinkedIn
Type	SIC Number	U.S. Standard Industrial Classification number (number corresponding to the division) of that industry.
	Type	Type of the business (B2B or rB2C)
Blogs	Complaint Remarks	Number of complaints (aka HateMark) based on the most popular Consumer Opinion website.
	Number of Views	Number of page views of the blog.
	Compliment Remarks	Number of compliments (aka HateMark) based on the most popular Consumer Opinion website.
Wikis	Wikipedia Number of Lanuguages	Number of languages that Wikipedia offers content about the brand.
	Wikipedia Page Views	Number of Wikipedia page views for the brand

The list is given in Table 1. We then collected data corresponding to brand/company web pages listed in Istanbul BIST30. The variable list is given in Table 1.

3.1. Notation and Methodology

We organize firm-variables in two-blocks: emporographic (block A) and digital reputation related (block B). Emporographic variables (A) corresponding to each firm are denoted with $[A_k, k \in \{0, \dots, n\}]$. Digital reputiton variables are denoted with $[B_l, l \in \{0, \dots, r\}]$. Emporographic variables include SIC numbers, retrieved from U.S. Standard Industrial Classification number, and Type variable (whether b2b or b2c).

We then calculated Cronbach alpha's for each of the four categories. Within We Page Statistics category we discarded 3 variables due to lack of internal consistency (Alexa Country Rank, Alexa Global Rank, Google Trends, Value of the Webpage). The remaining 5 variables are taken into account. Table 2 summarizes the results.

Table 2. Internal consistency.

Group	Number of Items	Standardized Cronbach's Alpha
Web Page Statistics	5	0.638
Social Media	4	0.653
Blog	3	0.934
Wikis	2	0.844

Following Foa and Tanner (2012) we decided on simple non-weighted indexing creation. We indexed 30 BIST30 companies according to the four categories. The Index is given in Table 3.

Table 3. BIST30 Online Reputation Index Table

ISIM	SIRKET	SIC	TYPE	Web Page Statistics	Social Media	Blogs	Wikis	Index
<i>isbank.com.tr</i>	ISBANK	6000	B2C	100.0	80.9	92.0	72.4	100.0
<i>garanti.com.tr</i>	GARANTI	6000	B2C	93.0	73.4	100.0	44.8	89.2
<i>thy.com.tr</i>	THY	4000	B2C	52.3	100.0	64.4	87.9	88.8
<i>turktelekom.com.tr</i>	TURK TELEKOM	4000	B2C	58.1	90.4	86.2	69.0	88.6
<i>halkbank.com.tr</i>	HALKBANK	6000	B2C	83.7	73.4	77.0	69.0	87.1
<i>akbank.com.tr</i>	AKBANK	6000	B2C	72.1	56.4	88.5	70.7	83.9
<i>arcelik.com.tr</i>	ARCELIK	5200	B2C	66.3	61.7	73.6	77.6	81.1
<i>tofas.com.tr</i>	TOFAS	2000	B2B	36.0	51.1	56.3	100.0	72.0
<i>yapikredi.com.tr</i>	YAPI KREDI	6000	B2C	52.3	78.7	88.5	32.8	71.8
<i>vakifbank.com.tr</i>	VAKIFBANK	6000	B2C	72.1	63.8	78.2	31.0	68.4
<i>bankasya.com.tr</i>	BANK ASYA	6000	B2C	58.1	76.6	65.5	39.7	67.0
<i>migros.com.tr</i>	MIGROS	5200	B2C	22.1	76.6	62.1	55.2	61.8
<i>turkcell.com.tr</i>	TURKCELL	4000	B2C	69.8	78.7	44.8	29.3	59.7
<i>bim.com.tr</i>	BIM	5200	B2C	33.7	72.3	60.9	44.8	59.6
<i>petkim.com.tr</i>	PETKIM	2000	B2B	22.1	43.6	55.2	74.1	56.7
<i>tupras.com.tr</i>	TUPRAS	2000	B2B	52.3	34.0	11.5	84.5	50.1
<i>aselsan.com</i>	ASELSAN	2000	B2B	24.4	58.5	33.3	51.7	46.0
<i>kardemir.com</i>	KARDEMIR	2000	B2B	10.5	39.4	5.7	87.9	39.9
<i>sabanci.com</i>	SABANCI HOLDING	0	B2B	16.3	45.7	13.8	63.8	37.6
<i>doganholding.com.tr</i>	DOGAN HOLDING	0	B2B	2.3	36.2	10.3	56.9	27.5
<i>pegasus.com.tr</i>	PEGASUS	0	B2C	57.0	18.1	6.9	25.9	24.5
<i>emlakkonut.com.tr</i>	EMLAK KONUT	1500	B2B	18.6	12.8	56.3	1.7	21.9
<i>enka.com.tr</i>	ENKA INSAAT	1500	B2B	11.6	47.9	3.4	29.3	21.2
<i>koc.com.tr</i>	KOC HOLDING	0	B2B	67.4	9.6	17.2	0.0	19.3
<i>erdemir.com.tr</i>	ERDEMIR	2000	B2B	15.1	26.6	8.0	32.8	18.8
<i>tavhavalimanlari.com.tr</i>	TAV HAVA MEYDANLARI	7000	B2B	10.5	41.5	9.2	22.4	18.6
<i>sisecam.com.tr</i>	SISECAM	2000	B2B	23.3	20.2	14.9	10.3	13.7
<i>ihlas.com.tr</i>	IHLAS HOLDING	0	B2B	16.3	4.3	0.0	24.1	6.9
<i>kozaaltin.com.tr</i>	KOZA ALTIN	1000	B2B	5.8	26.6	4.6	6.9	6.1
<i>kozamadencilik.com</i>	KOZA MADENCILIK	1000	B2B	0.0	0.0	1.1	19.0	0.0

4. Conclusion and Future Directions

Data sources, where the methodology has been implemented, can be categorized as challenging since the volume, variety, velocity and veracity of the data sources of the study is high in all dimensions. The data sources require a numerical reporting on whole social media and web pages. Also the blogs and wikis are relatively less ambiguous, less in the size, less in the velocity and variety is lower, they are still beyond the computation power of a single computer. As a solution we propose utilizing the online tools as data sources like search engines, social media reporting tools, wiki and blog statistics. After collecting the statistical data sources, we propose weighting and statistical models over the weights, and output an index, we go further and propose a diagnosis test by the experts. Using sector specific indexes via homogeneous data sets can enhance the study and all methodologies including data source selection, data source weighting and statistical methodologies can be fine-tuned in the next step. The study can be considered as first and fresh output of recurring data collection for 1 year.

References

- Adam, P., & Nicolai, A. (2013). *The Ethical Economy, Rebuilding Value After the Crisis*. Columbia University Press.
- Allahbakhsh, M., Ignjatovic, A., Benatallah, B., Beheshti, S., Bertino, E., & Foo, N. (2012). Reputation management in crowdsourcing systems. *Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, 2012 8th International Conference on (pp. 664 - 671). Pittsburgh, PA: IEEE.
- Arslan, M. L., & Seker, S. E. (2013). The Impact of Employment Web Sites' Traffic on Unemployment: A Cross Country Comparison. *INTERNATIONAL JOURNAL OF SOCIAL SCIENCES AND HUMANITY STUDIES*, 5 (2), 130-138.
- Arslan, M. L., & Seker, S. E. (2014). Web Based Reputation Index of Turkish Universities. *International Journal of E-Education E-Business E-Management and E-Learning (IJEEEE)*, 4 (3), 197-203.
- Economist, T. (2013, Feb 23). The Price of Reputation, Is the market for protected personal information about to take off? *The Economist*, from the print edition .
- Engardio, P., & Arndt, M. (2007, Jul 8). What Price Reputation?
- Foa, R., & Tanner, J. C. (2012, 4). Methodology of the Indices of Social Development . *ISD Working Paper Series*, 1-66.
- Ghose, A., Ipeirotis, P. G., & Sundararajan, A. (2006, Feb.). The Dimensions of Reputation in Electronic Markets. *NYU Center for Digital Economy Research* .
- Howe, J. (2009). *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. NY, US: Crown Business.
- Kato, A. (2006). Price, Quality and Reputation: Evidence from An Online Field Experiment. *RAND Journal of Economics*, 37 (4).
- Kizil, C., Arslan, M. L., & Seker, S. E. (2013). Correlation between Intellectual Capital and Web Trends of Top 30 Companies in Turkey. *International Journal of Social Sciences and Humanity Studies*, 5 (2), 39 - 49.
- Madden, M., & Smith, A. (2010, Mat 26). Reputation Management and Social Media. *Pew Internet & American Life Project* .
- Mert, C., SEKER, S. E., Jamburia, G., & Temel, M. H. (2014). Technological Impact Of Employment Web Sites In Caucasus Region. *European Journal of Business Research*, 14 (3), 81-86.
- Seker, S. E., & Georgievski, B. (2014). Financial Crisis and the ICT Industry, Cross Market Research on Europe, US, Turkish and Gulf Countries. *Research Journal of Finance and Accounting*, 5 (24), 78-87.
- Seker, S. E., & Kulaklı, A. (2016). Macroeconomic ICT Facts and Mobile Telecom Operators via Social Networks and Web Pages. *Journal of Economics Business and Management*, 4 (2), 99-103.
- Seker, S. E., Cankir, B., & Arslan, M. L. (2014). Information and Communication Technology Reputation for XU030 Quote Companies. *International Journal of Innovation, Management and Technology*, 5 (3), 221-225.
- Seker, S. E., Cankir, B., & Okur, M. E. (2014). Strategic Competition of Internet Interfaces for XU30 Quoted Companies. *International Journal of Computer and Communication Engineering*, 3 (6), 464-468.