P. Mitic, Int. J. of Safety and Security Eng., Vol. 8, No. 1 (2018) 171–180

REPUTATION RISK: MEASURED

PETER MITIC
Santander UK, 2 Triton Square, Regent's Place, London.

ABSTRACT
Two principal results for reputation risk are established. First, reputation risk can be measured in terms of a single index, arising from a data mining process directed at the opinions in a complex multi-agent network. Second, the results of the measurement process, gathered over an extended period, can be expressed directly in monetary terms by finding a correlation between the daily changes in the index and in sales. Stressed periods are modelled by calculating value-at-risk using a 'loss-distribution/ scenario' approach, as for operational risk capital. The short-term effect of reputation risk events on sales and profits can be significant in absolute terms, but is small as a percentage of total sales. Negative reputation has a more significant impact than positive reputation.
Keywords: Reputation, Reputation risk, alva, sentiment analysis, correlation, Loss Distribution, Scenarios, stressed
Disclaimer: The opinions, ideas, approaches and numerical values presented are those of the author and do not necessarily reflect Santander's position. Actual Results are not given.

1 INTRODUCTION

Reputation risk (*Rep-Risk*) has its own reputation: that of being difficult, and some would say 'impossible', to quantify. On the contrary, quantification of *Rep-Risk* is certainly possible, and we report the effects of reputation risk on sales and profit. Without a formal demonstration of this link, the view that profits might suffer as a result of poor reputation is purely anecdotal. Section 2 of the paper describes how electronic data feeds may be used to condense the opinions of agents in a complex network to a single numerical index on a daily basis. In Section 3, a correlation between the daily changes in the index and in sales and profits is established. Techniques from operational risk are used to estimate sales/profit changes for normal (unstressed) business operations, and also in stressed conditions using Value-at-Risk (VaR). Issues of governance and management of reputational risk are not the subject of this paper. Instead, see, for example, the collection of papers in Klewes and Wreschniok [1].

1.1 Complexity in organisation/stakeholder networks

We envisage a network comprising multiple instances of two classes of interconnected and interdependent agents: organisation and stakeholder. Each agent interacts with at least one other agent, independent of type. A pattern emerges from the interactions, enabling measurement of the interaction extent, namely the reputation score described in Section 2.

1.2 Definitions

Loose definitions for 'reputation risk' are used frequently, but are insufficiently rigorous for quantitative use. They are often linked with phrases such as "opinion of the public towards an organisation" (Walter [2]), or "negative perception" (BCBS157 [3]). This quote from Honey [4] is useful, because it relates customer expectation to corporate performance.

ISSN: 2041-9031 (paper format), ISSN: 2041-904X (online), http://www.witpress.com/journals
DOI: 10.2495/SAFE-V8-N1-171-180

Table 1: Informal Reputation-related definitions.

Reputation	Stakeholder perception of an organisation that can affect, positively or negatively, the business relationship between the stakeholder and the organisation
Reputation Risk (*Rep-Risk*)	The difference between stakeholder expectation and organisation performance
Reputation Event	An occurrence or action that affects Reputation
Rep-Risk Measurement	Numerical assessment of Reputation
Rep-Risk Value	The mapping of Reputation Risk Measurement to physical or monetary quantities

"The reputation of an organisation is influenced by its performance, policies and people…A risk to reputation occurs where the organisation fails to meet the expectations of a specific stakeholder group."

The performance-expectation link forms the basis of the method used to link reputation measurement with sales, and is reflected in the definitions in Table 1. The element *positive reputation* has hitherto not been used in the literature.

2 REPUTATION MEASUREMENT

This section concentrates on the measurement of *Rep-Risk*. The measure described corresponds to the informal term "Reputational Risk Measurement" in Table 1. A measure of daily sentiment, hereinafter referred to as the *alva* Reputation Index (*aRI*) is produced by the reputation consultancy *alva* (www.alva-group.com). The index is compiled daily, and provides a short-term measure of sentiment (i.e. reputation), mainly for use as a means to make informed strategic business decisions. Two levels of complexity are involved in producing it. First, there is a data mining stage of data from agents, the number of which is unknown in advance of collection. Second, the complexity of language is reduced to tokens, to which numerical values can be assigned. The *aRI* is constructed in *Process 1*, below.

1. Live electronic feeds from publicly available media sources supply articles, comments, reports etc. that are relevant to particular industry sectors. These feeds comprise news media (TV, radio, newspapers etc), social media (Twitter, Facebook, blogs), trade reports and surveys. This stage is termed 'content harvesting', and each item received is termed a 'content'. Alternatively, the term 'data mining' is appropriate.
2. Contents that mention particular target keywords (e.g. names of an organisation, prominent people in it) are filtered and retained.
3. Identify key words and phrases in each content that convey sentiment.
4. Each content received in a 24-hour period is scored on a scale 1..10 for four factors: overall sentiment, influence of the source, prominence of the organisation, and relevance.
5. The mean of the scores for the four factors is calculated. For content i on day t, call this value $m_{i,t}$.
6. Weights, reflecting the importance of the content as a whole, are then assigned to each content. National media or influential persons are weighted most highly. Regional media are weighted lower. The lowest weights are assigned to social media users with few followers. For content i on day t, call the weight $w_{i,t}$.

7. The final reputation index value on day t, R_t, is a weighted average of the scores for all factors of the content within the 24-hour period: $R_t = \sum w_{i,t} m_{i,t}$

Since scoring is always on a scale of 1..10, the median point 5.5 represents a neutral position with respect to sentiment. A score $R > 5.5$ represents 'positive' sentiment, whereas a score $R < 5.5$ represents 'negative' sentiment. Most commonly, movements of the *aRI* from one day to the next are less than 0.5. Daily movements greater than 1 are rare. Empirically, values of the *aRI* are normally distributed. The *aRI* has a 'short term memory', consistent with an auto-regressive AR(1) statistical model. The *aRI* values on successive days are significantly correlated.

The method described in *Process 1* corresponds to 'unsupervised learning', as described by Turney [5]. Step 3 in *Process 1* is a classification based on a collection of fixed syntactic phrases that are likely to be used to express opinions. An overview of the general technique is given below.

2.1 Sentiment scoring example

To get a flavour of the contents that are routinely received, consider the Twitter remark from @blogpenzance (1,053 followers) "*I'm a big fan of @XYZ-Bank*". Under Process 1, that content might be scored as in Table 2.

The result is a mean score 24.5/4 = 6.125 for this content. The source is treated as of low importance compared to the national press, so it attracts a low weight: 0.12 (on a scale 0..1). Suppose further that there are two additional contents (labelled C2 and C3), and the Table 2 content is C1. Table 3 shows how a reputation index may be compiled from them. Only the mean scores are shown.

Table 2: Content scoring example.

Category	**Sentiment**	**Score, *s***
Sentiment	Positive, qualified by 'big'	8.0
Influence	Few followers: not influential	1.0
Prominence	Neutral	5.5
Relevance	No references to other organisations	10.0

Table 3: Index compilation example.

Content		**Mean Score *m***	**Weight *w***	***mw***
C1	"*I'm a big fan of @XYZ-Bank*"	6.125	0.12	0.735
C2	"*XYZ-Bank hardly provides good service*" (Local TV consumer feature)	4.7	0.6	2.82
C3	"*XYZ-Bank's mortgage interest rate is the best available*" (Sunday Times 'Best Buy' tables)	8.62	0.9	7.758
Sum			1.62	11.313

The index value is 11.313/1.62 = 6.983. Overall it conveys positive sentiment, mainly due to content C3, which is very positive and from an influential source. Qualifiers such as ‘best’ and ‘big’ serve to shift the base score away from 5.5. The word ‘not’ is an important negation indicator. Others are ‘dis’, ‘un’ and ‘down’.

2.2 Sentiment analysis

Lui [6] gives a full summary of the current and past research on sentiment analysis, which dates from the 1990s. The steps in Process 1 are a particular instance of a generic sentiment analysis (aliter opinion mining) algorithm. Pang *et al*. [7] summarise it as Process 2, below.

1. Split each content into tokens: words or phrases in a standard form. See Chaudhari and Govilkar [8] for a review (step 3 in Process 1).
2. Feature extraction to derive: <holder> (who expresses the content), <target> (the aim of the content), <polarity> (the opinion expressed). Siqueira and Barros give an overview [9] (also step 3, Process 1).
3. Classify the features. In particular, assess the <polarity> as positive, negative or neutral, and the emphasis of the sentiment. Several techniques have been used, such as Naïve Bayes, Maximum Entropy and Support Vector Machines. [8, 10] (step 3 in Process 1). This step often uses an existing lexicon to gauge the extent of the sentiment polarity. Jurafsky has a good summary [10].

To give a brief example of polarity, ‘good’ and ‘increased’ indicate a positive sentiment, whereas ‘bad’ and ‘not’ indicate negative sentiment. The words ‘better’ and ‘worse’ are examples of emphasis indicators. An awkward case is ‘dreadful’, which conveys emphasised negative sentiment in a single word. Such cases have to be treated separately in the lexicon. A further problem is that some words can convey negative sentiment in some cases and positive in others. For example, ‘quiet’ is positive for cars but negative for a speakerphone. A lexicon that uses syntactic or co-occurrence patterns is required.

3 REPUTATION, SALES AND PROFIT

In this section we link the *aRI* to sales and profit, and establish the principle that ‘reputation means money’. The counterpart to every sale is a purchase, and we can regard measures such as ‘daily sales’ as a distillation of the activities of a network of agents, distinct from the network that contributed to the compilation of the *aRI*. It is not possible to link the two networks on an agent-to-agent basis, so we have to make do with a periodic average of both. The impact of linking reputation to money is that it enables a translation of a somewhat abstract *aRI* score to something entirely familiar. That enables organisations to improve forecasting and risk mitigation, and to optimise decision making. All three are highlighted in Sargut & McGrath [11] in the context of differentiation between ‘complex’ and ‘complicated’. Networks in the world of reputation are ‘complex’ because interactions are continually changing, and individual interactions are not deterministic or predictable.

Existing models for *Rep-Risk* are due to Perry and de Fontnouvelle [12], and Fiordelisi *et al.* [13, 14]. They depend on correlations of reputational events with share price, without the assumption that share price movements arise from reputation events. Share price analyses do not address other business activities of a bank, such as retail sales. In these three cases the correlation methodology starts with a 1-factor statistical model for n reputation risk events:

$$r_{it} = \alpha_i + \beta_i\, r_{mt} + e_{it} \tag{1}$$

where r_{it} is the log-return for a stock linked to event i at time t, r_{mt} is the log-return for a stock *index* containing the stock at the same time t, ε_{it} is a random term, and α_i and β_i are coefficients to be determined by Ordinary Least Squares (OLS). Then, having calculated estimates $\hat{\alpha}_i$ and $\hat{\beta}_i$ for α_i and β_i, *abnormal returns* AR_{it} (i.e. the residuals) are calculated using:

$$AR_{it} = r_{it} - \left(\hat{\alpha}_i + \hat{\beta}_i r_{mt}\right). \tag{2}$$

Summing over a time interval $[t_0, t_1]$ produces a *cumulative abnormal return* (*CAR*):

$$CAR\left(i,t_0,t_1\right) = \sum_{t=t_0}^{t_1} AR_{it} \tag{3}$$

The *CAR* is the basis for the measurement of *Rep-Risk* in terms of share price movements. In our study, we use a variant of it, detailed in the following section.

3.1 Reputation risk correlation model: index based

Our model of *Rep-Risk* comprises three distinct stages, each resulting in a measure of the monetary value of reputation. Stage 1 gives a base ‘expected value’ measure: what happens under unstressed conditions. This stage encompasses the concept that reputation is a comparison between performance and expectation. Its basis is a correlation between changes in the *aRI* and changes in product sales with a prediction of changes in sales. In stage 2, we separate upward and downward changes in the *aRI*, fit an appropriate ‘loss’ distribution to each separately, using the *Loss Distribution* approach of Frachot *et al.* [15], thereby deriving a 99.9% value-at-risk (VaR). This provides a view of what happens if there are severe reputation events. Stage 3 does the same for extremely severe circumstances, using scenarios.

3.1.1 Correlation stage

Given a fixed time interval divided into n equal intervals, let t_i ($i = 1..n$) be the time (measured from zero) at the end of the i^{th} time interval, and set $t_0 = 0$. The i^{th} time interval is therefore $[t_{i-1}, t_i]$, and the duration of all time intervals is t_n/n.

A measureable utility, $U(t_i, l)$, (in particular, sales) is observed at t_i, but refers to an earlier time period, the time lag being l periods. Let $\hat{U}_j\left(t_i,l\right)$ be the j^{th} member of a set of *significant* predictors of $U(t_i,l)$, obtained by any appropriate method. In practice, the most useful predictors were means or single exponential smoothings of previous utility values. Denoting the total *Rep-Risk* loss and gain at time t_i by $R^-\left(t_i\right)$ and $R^+\left(t_i\right)$ respectively, equation (4) gives a symbolic definition of *Rep-Risk* loss/gain. In (4), the changes in value of the reputation index over the same period are denoted by $\Delta S^-\left(t_i\right)$ and $\Delta S^+\left(t_i\right)$ respectively, and the superscripts $^-$ and $^+$ denote negative and positive extractions respectively.

$$\begin{aligned} R^-\left(t_i\right) &= \underset{j}{median}\left(\sum_{i=1}^{n}\left(\hat{U}_j\left(t_i,l\right) - U\left(t_i,l\right)\right)^-\right)\frac{1}{\Delta S^-\left(t_i\right)} \\ R^+\left(t_i\right) &= \underset{j}{median}\left(\sum_{i=1}^{n}\left(\hat{U}_j\left(t_i,l\right) - U\left(t_i,l\right)\right)^+\right)\frac{1}{\Delta S^+\left(t_i\right)} \end{aligned} \tag{4}$$

The sum of differences between predicted and observed utility is similar to the *CAR* measure in equation (3). The units of $R^-(t_i)$ and $R^+(t_i)$ are "units sold", and to translate them into money we multiply by the unit price, $f(t_i)$, to derive expressions for "Reputational Risk Value" (see Table 1). The factor $f(t_i)$ can also be used to account for a translation into profits, rather than sales value.

$$\begin{aligned} R_M^-(t_i) &= R^-(t_i) f(t_i) \\ R_M^+(t_i) &= R^+(t_i) f(t_i) \end{aligned} \tag{5}$$

A 'reasonableness test' is incorporated into the $f(t_i)$ factors. If the set of values of the *aRI* used in the correlation analysis are denoted by *R*, the quantities $r(+) = \max(R)–5.5$ and $r(–) = \min(R)+5.5$ measure the maximum and minimum deviation from the base *aRI* value, 5.5, respectively. These quantities measure what can be considered a practical 'reasonable' movement in the Index. Negative and positive movements in sales are scaled by $r(–)$ and $r(+)$ respectively.

The formulation in (5) is necessarily linked to the sales and profit profiles of a particular organisation. As a result, (5) cannot constitute an absolute expression of *Rep-Risk* in monetary terms. However, the given formulation does provide direct management information and guidance for the organisation on the potential effect of reputation events. It is therefore more applicable than a general measure.

3.1.2 "Loss distribution approach" (LDA) stage

The LDA method results in an assessment of VaR, representing stressed circumstances. Positive and negative *aRI* values are treated separately. In each case the *n aRI* values $\{a_1, a_2, ..., a_n\}$ are ordered in increasing size order, and a cumulative probability $p_i = (a_i - 0.5)/n$ is assigned to each a_i. Trial distributions are fitted to pairs $\{a_i, p_i\}$. A best-fit severity distribution is found, goodness of fit being tested using the dedicated method of Mitic [16]. This test is appropriate because it is independent of the small number of data points used: about 25 in each case. In most cases, a Lognormal distribution was the best fit. Figure 1 below shows the example of the CDF of a Lognormal distribution fitted to positive sentiment difference.

The LDA method goes on to estimate VaR at 99.9% (the *stressed* case), and the expected *aRI* value (the *unstressed* case). For details, see Frachot *et al.* [15]. The 99.9% figure is commonly used in similar operational risk calculations.

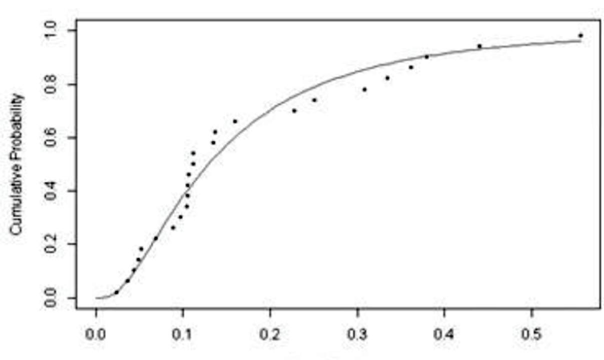

Figure 1: Lognormal fit for positive sentiment difference.

3.1.3 Scenario stage

The LDA model can be stressed by adding scenarios similar to the methods of operational risk, but with a difference. Adding extra *aRI* values less than 4.0 simulates extreme negative sentiment, and adding extra *aRI* values greater than 7.0 simulates extreme positive sentiment. The latter is missing from operational risk. A complete recalculation is then done using the revised *aRI* values. In practice, approximately 25% extra sentiments were added by generating random data from a uniform distribution centred closely on a given high or low sentiment value. The resulting VaR values are sensitive to the number of extra points added, which reflects the degree to which good or poor reputation is sustained. Alternatively, specialised scenario workshops are a lengthier process.

4 RESULTS

In this section we present the results of the correlation and LDA models. Retail sales data with corresponding *aRI* values for the period 1 January 2014 to 31 Dec 2014 were used.

4.1 Unstressed and stressed reputation

Table 4 shows the % changes in sales and profit (the results of (5) divided by total annual sales/profit) for three product lines in the *unstressed* case (i.e. expected value of the LDA process, following a correlation analysis summarised by (5)).

The entries in Table 4 represent *Rep-Risk* gains/losses in 'business as usual' circumstances. Two points are notable. First, *Rep-Risk* gains/losses expressed as percentages of total gains/losses are small, even though in absolute terms they might be significant. Second, the gains from positive *Rep-Risk* and the losses from negative *Rep-Risk* are roughly symmetric, indicating that there is as much to be gained by managing positive sentiment as there is by mitigating negative sentiment. This observation does not apply when sentiment is stressed (Table 5).

Comparing corresponding results in Tables 4 and 5, the stressed *Rep-Risk* gains/losses are approximately twice the unstressed *Rep-Risk* gains/losses.

4.2 Super-stressed reputation

The effect of adding additional *aRI*s near a given value in the range (4, 7.5) is to exaggerate the derived VaR. The exaggerated VaR % change is a measure of the impact of a period of extreme reputation stress. This variation is shown in Fig. 2, which is for Product 1. Products

Table 4: Expected Value of Unstressed annual *Rep-Risk* gains/losses.

Product	Positive sentiment (% change)	Negative sentiment (% change)
Product 1 sales volume	1.6	-2.3
Product 1 profit after tax	0.7	-0.9
Product 2 sales volume	1.9	-1.4
Product 2 profit after tax	0.3	-0.3
Product 3 sales volume	2.4	-2.0
Product 3 profit after tax	0.7	-2.0

Table 5: 99.9% VaR of Stressed annual *Rep-Risk* gains/losses.

Product	Positive sentiment (% change)	Negative sentiment (% change)
Product 1 sales volume	3.4	-7.9
Product 1profit after tax	1.3	-3.6
Product 2 sales volume	3.8	-2.8
Product 2 profit after tax	0.6	-0.5
Product 3 sales volume	4.9	-4.1
Product 3 profit after tax	4.9	-4.1

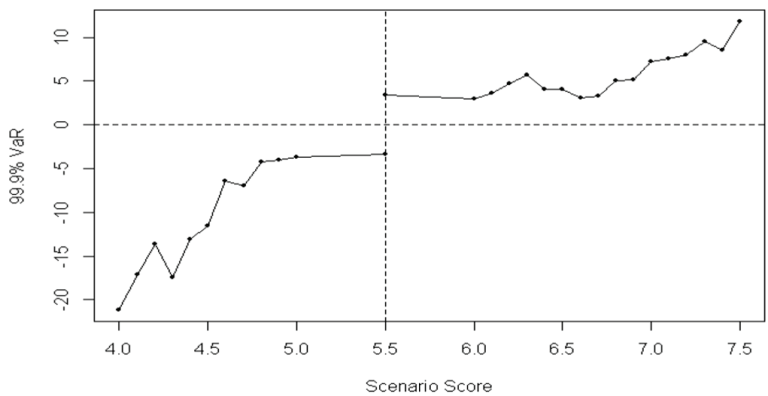

Figure 2: Varying degrees of super-stress

2 and 3 are similar, but with a less extreme range of VaR. This reflects the fact that Product 1 is the principal product line. The profile in Fig. 2 applies for both sales and profit, since one is a linear scaling of the other (through factors $f(t_i)$ in (5)).

Two features are significant. First, the *Rep-Risk* gain/loss is an increasing function of the severity of the scenario. Stochastic variation accounts for a lack of smoothness in the profiles. Second, negative scenarios have a much greater effect than positive scenarios. Such a case is not unreasonable: in February 2015 HSBC's practice of tax evasion by use of Swiss bank accounts was exposed [17]. The *aRI* remained at about 3.5 for two weeks as the extent of the deception became apparent. A 17% fall in profits was attributed to this deception.

4.3 Dependence of results on the predictor

The dependence of equation (4) on predictors introduces a potentially large source of error in the calculation using (4). We seek to capture this error by calculating the maximum and minimum sales and profit instead of the median of those quantities. They show a significant variation from the median results. This is a warning that a wide error band should be expected. The results for maximum and minimum sales and profit for Product 1 (the major product line) are shown in Fig. 3. The chart indicates the percentage variation from the results of Table 5 for both positive ("Up") and negative "Down" sentiment. The median results are on the zero

Figure 3: Dependence minimum and maximum stressed loss

line. The results indicate a possible underestimate of the effect of positive sentiment, the effect of which could be greater than might appear from using the median indicator.

5 DISCUSSION AND SUMMARY

In this paper we have reduced a complex network of agents, who are engaged in expressing opinions on an organisation, to a single daily index value (the *aRI*) which represents group sentiment towards that organisation. Using this index, a formal definition has been given for *Rep-Risk*, along with a summary of related informal terms. The *aRI* uses an arbitrary scale, and we express *aRI* in monetary terms by linking it to sales and profit. Indeed, this is the first formal demonstration of a sentiment-money link. In contrast to all previous studies, the monetary gain arising from positive sentiment has been assessed as well as the loss arising from negative sentiment.

Some general conclusions are apparent. First, there is approximate symmetry between the effects of positive and negative sentiment in unstressed cases. However, in stressed cases, negative sentiment has a much greater impact than positive sentiment. It is therefore much more important to mitigate potential negative reputational events than to generate positive sentiment. Second, the effect of sentiment on sales is a small percentage of total sales. Corresponding profits depend on relationships derived from balance sheet data and are necessarily even smaller. Third, *Rep-Risk* gains/losses are higher for major business lines. This is essentially a volume effect, but may be affected by factors such as product liquidity and perceived 'value for money'. The latter factor suggests the following conjecture, which has not yet been investigated. Retail customers ignore reputation events unless they are affected personally. They are swayed by personal monetary gain and the quality of service they receive, not by general issues of conduct.

REFERENCES

[1] Klewes, J. & Wreschniok, R., (eds.) *Reputation Capital: Building and Maintaining Trust in the 21st Century,* Springer, 2013.

[2] Walter, I., Reputational risk *(Chapter 6) Finance Ethics,* eds. J.R. Boatright & R.W. Kolb, John Wiley & Sons, 2013.

[3] BCBS157 Bank for International Settlements. Enhancements to the Basel II framework, available at www.bis.org/publ/bcbs157.pdf, 2009.

[4] Honey, G., *Corporate reputation: perspectives of measuring and managing a principal risk*. CIMA, 2007.

[5] Turney, P., Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews, *Proceedings of the Association for Computational Linguistics (ACL)*, pp. 417–424, 2002.
https://doi.org/10.3115/1073083.1073153

[6] Lui, B., *Sentiment Analysis: Mining Opinions, Sentiments and Emotions*. Cambridge University Press, 2015.

[7] Pang, B., Lee, L. & Vaithyanathan, S., *Thumbs up? sentiment classification using machine learning techniques*. (eds. Haji, J. & Matsumoto, Y.) Proceedings of the EMNLP, Philadelphia, pp. 79–86, 2002

[8] Chaudhari, M. & Govilkar, S., A survey of machine learning techniques for sentiment classification. *International Journal of Computational Science & Applications*, **5**(3), pp. 13–23, 2015.
https://doi.org/10.5121/ijcsa.2015.5302

[9] Siqueira, H. & Barros, F., *A feature extraction process for sentiment analysis of opinions on services*, Proc. Int. Workshop on Web and Text Intelligence. 2010

[10] Jurafsky, D. Sentiment analysis, available at: https://web.stanford.edu/class/cs124/lec/sentiment.pdf

[11] Sargut, G. & McGrath, R., Learning to live with complexity. *Harvard Business Review*, available at https://hbr.org/2011/09/learning-to-live-with-complexity, 2011.

[12] Perry, J. & de Fontnouvelle, P., Measuring reputational risk: the market reaction to operational loss announcements, available at http://ssrn.com/abstract=861364, 2005.

[13] Fiordelisi, F., Soana, M-G. & Schwizer, P., The determinants of reputational risk in the banking sector. *Journal of Banking & Finance,* **37**(5), pp. 1359–1371, 2013.
https://doi.org/10.1016/j.jbankfin.2012.04.021

[14] Fiordelisi, F., Schwizer, P. & Soana, M-G., Reputational losses and operational risk in banking. *European Journal of Finance,* **20**(2), pp. 105–124, 2014.
https://doi.org/10.2139/ssrn.1032523

[15] Frachot, A., Georges, P. & Roncalli, T., *Loss distribution approach for operational risk,* Gropue de recherche Operationnelle, Credit Lyonnais, 2001.

[16] Mitic, P., Improved goodness-of-fit measures. *Journal of Operational Risk,* **10**(1), pp. 77–126, 2015.
https://doi.org/10.21314/jop.2015.159

[17] Reuters, available at www.reuters.com/article/2015/02/23/us-hsbc-results-idUSKBN0LR0NE20150223, 2015.