Trust and Risk Relationship Analysis on a Workflow Basis: A Use Case

Valentina Viduto, Karim Djemame, Paul Townend, Jie Xu, Sarah Fores, Lydia Lau, Vania Dimitrova
School of Computing, University of Leeds
Leeds, UK

{V.Viduto, K.Djemame, P.M.Townend, J.Xu, S.Fores, L.M.S.Lau, V.G.Dimitrova}@leeds.ac.uk

Martyn Fletcher[2], Stephen Hobson[2,]
Jim Austin[1,2], John McAvoy[2],
[1] Department of Computer Sciece, University of York,
[2] Cybula Ltd.
York, UK
{martyn, stephen, austin, mcavoy}@cybula.com

Charlie Dibsdale,
Rolls Royce PLC
Derby, UK
charlie.e.dibsdale@o-sys.com

Abstract— **Trust and risk are often seen in proportion to each other; as such, high trust may induce low risk and vice versa. However, recent research argues that trust and risk relationship is implicit rather than proportional. Considering that trust and risk are implicit, this paper proposes for the first time a novel approach to view trust and risk on a basis of a W3C PROV provenance data model applied in a healthcare domain. We argue that high trust in healthcare domain can be placed in data despite of its high risk, and low trust data can have low risk depending on data quality attributes and its provenance. This is demonstrated by our trust and risk models applied to the BII case study data. The proposed theoretical approach first calculates risk values at each workflow step considering PROV concepts and second, aggregates the final risk score for the whole provenance chain. Different from risk model, trust of a workflow is derived by applying DS/AHP method. The results prove our assumption that trust and risk relationship is implicit.**

Keywords- trust; risk model; provenance; decision support; workflow; DS/AHP.

I. INTRODUCTION

In recent years, business critical decisions heavily rely on data collected and manipulated by many distributed sources and services. To make sure that crucial, high value decisions will not put business at risk, it becomes important to put trust in information and system data outputs. Trust is one of the concepts that is used to verify the usefulness and/or criticality of data, systems, personnel and whole workflow. However, it is quite challenging to define the term because it is being used with a variety of meanings and in many different contexts, sociology, psychology, and philosophy. The common notions of trust are associated with hope, faith, belief, confidence reliance on the integrity, dependence or character of a person or thing [10]. The variety of common terms shows that there is no precise definition of trust as it largely depends on author's viewpoint. Trust is also often situation specific; in one environment trust does not directly transfer to another environment and the notion of context is necessary [10].

Recent research inherently links trust to risk. There is no reason to trust if there is no risk involved. Thus, the cooperation or interaction with the system or human is less likely with higher risk unless the benefits from such interaction are worth the risk. The SECURE project [10] has made a good attempt in demonstrating that risk and trust are inexorably linked and must both be considered when making a decision about some ambiguity whose outcome depends on another entity's action. Also, considering observations made by Solhaug et al. [2] seeing that trust is generally neither proportional nor inverse proportional to risk under various constraints, in this paper, we put a first attempt to demonstrate how trust and risk relationship can enhance trustworthiness in systems and inform decisions. Inspired by the challenge of relating trust while considering consequences of risk, the trusted digital Spaces through Timely Reliable And Personalised Provenance (STRAPP) project [17] aims to provide an approach to enable users make informative decisions by considering three notions associated with the data: risk, provenance and trust. To demonstrate the STRAPP view of trust and risk relationship we use W3C PROV Data model [11] for provenance interchange. This data model describes entities, activities and people involved in the creation of data, its operation and decision making. It allows the decision maker to see the chain of activities, processes and data inputs as well as agents who performed certain actions with regard to data.

The aim of the paper is to address an assumption that trust in system can be placed knowing the data source and its quality, and risk associated with some processes may be high despite of good quality data used. We model risk and trust independently on a basis of a same workflow generated using BII (Brain Injury Index) case study data. Under STRAPP context, we define risk as a *"probability of some unwanted events at every workflow process which may result in unwanted consequences to this process"*, whereas trust is assessed in the context of data quality of a particular data file, and defined as *"a degree of confidence placed in input data while considering data quality attributes: completeness, accuracy, relevance, of the data file."* Data file in the BII case study consists of several metadata input fields that are

assessed in terms of their quality and importance. The ranking of input files is performed by applying Dempster–Shafer Analytical Hierarchy Process (DS/AHP) [14].

The remainder of the paper is organised as follows: Section II gives an overview of the STRAPP project highlighting its aims and applicability to the BII case study. Section III provides the most relevant work in three research areas: trust, risk and provenance and tries to highlight how these fields can facilitate decision making process. Section IV discusses BII case study as well as presents risk and trust models on a workflow basis. Section V summarises the results, work accomplished and provides future research directions.

II.   STRAPP OVERVIEW

The STRAPP project has been established, funded by Rolls-Royce, Cybula Ltd, and the UK Technology Strategy Board to facilitate the assessment of provenance-based, personalised trusted digital spaces where timely and critical decisions should be made. The objective of STRAPP is to enable users to place increased trust on data shown by, and decisions made by a system and by allowing them to view the provenance of that data or decision, presented in a personalised manner (for example, based on their role; managers may need to view the provenance and risk of a decision at a different level than software engineers, etc.) Furthermore, the project aims to provide visualization mechanisms to ensure users understand trust and the risks associated with data and decision-making. In the short term, these mechanisms are integrated to both the Equipment Health Management (EHM) system developed by OSys - a subsidiary company of Rolls-Royce PLC - that provides customers (primarily in the aerospace, marine and energy sectors) with the ability to diagnose and predict equipment faults, and to the Brain Injury Index (BII) system developed by Cybula Ltd that assists researchers and practitioners in the healthcare industry, with a focus on neuroscience. In the longer-term, it is hoped that many other decision-support systems in a wide range of sectors will be able to take advantage of the STRAPP system.

In this paper, we are primarily concerned with the trust and risk assessment components modelled using BII case study data. The purpose is to demonstrate the implicit relationship between trust and risk, as discussed by Solhaug et al. [2] and Cahill et al. [10] and visualise this relationship on a workflow basis.

III.   RELATED WORK

Our research encompasses several research directions: trust assessment and modelling, risk analysis and its conceptual relation to trust, provenance modelling and its usability with regard to decision making process. Therefore, in this paper we will focus on trust and risk modelling on a basis of provenance data to make an attempt of demonstrating the implicit relationship between risk and trust as it was observed [2] [10], under specific use.

Trust is a widely explored topic within a variety of computer science domains. Trust is defined as a relationship between two entities, a trustor and a trustee where a trustor places some level of trust in a trustee under a specific set of contexts. Thus, trust, in literature, is used in a variety of meanings. A distinction between context-independent trust (reliability trust) and context-dependent trust (decision trust) can often be recognized among scientific community, although usually not explicitly expressed [4]. Reliability trust is interpreted as the reliability of something or somebody independent of the context. As such, according to Gambetta [1], *trust is a particular level of the subjective probability with which "an agent assesses that another agent or group of agents will perform particular action, both before he can monitor such action and in the context in which it affects his own action."* It is a crucial question then, whether or not to engage in cooperation with an agent. This cooperation depends on the extent to which the agent (trustor) believes that the trustee will behave in a certain way. Hence, the level of trust is determined subjectively based on evidences available to the trustor on trustree's behaviour and constraints by which this behaviour might be regulated.

Decision trust, when seen within a context, is defined as *the extent to which a given party is willing to depend on something or somebody in a given situation with a feeling of relative security, even though negative consequences are possible* [4]. This definition implicitly covers contextual elements, such as possible outcomes, environmental factors (existing safety/security mechanisms) and risk attitude (taking, avoiding, and transferring). Josang and Presti in [5] draw a model of trust composed of a reliability trust as the probability of a transaction success and a decision trust derived from a decision surface. With such example, authors provide a first attempt to shape the relationship between risk and trust. The model first, calculates expected gain of a possible transaction and second, introduces a fraction of the capital the agent is willing to risk. Risk, as part of the model, is taken in order to derive a more complete definition of trust, the decision trust. Therefore, the approach of including risk into the model provides more meaningful notion of trust because it combines trust with risk attitudes.

Recently, trust is modelled by highlighting the presence and importance of provenance data. The semantic representation of trust and provenance data is modelled through the provenance ontology. As such, Emaldi et al. [6] present a trust model for the measurement of trust value in the context of smart cities. The trust value is calculated according to each factor independently. The factors calculated are defined as trust of authority, popularity, recommendation, provenance, timeliness and geographical distance. Another method for assessing trust based on provenance information is presented by Hartig and Zhao [7]. The authors proposed an assessment method which calculates trust values based on timeliness of data quality. Ceolin et al. [8] assessed the trust by first computing reputation-based trust values and second, trust values are computed based on provenance information, represented by

means of W3C standard PROV model. By merging trust values authors claim that it can be beneficial for reliability of the estimated trust value. In trust management domain, *reputation* is used to define trust between two agents. Reputation is what generally said or believed about a person's or thing's character or standing [4]. It influences trust in two ways: firstly, it positively affects the trustor's reliability trust in the trustee and secondly, it disciplines the trustee as it is known that bad behavior will be seen. The good example of difference between trust and reputation can be seen in the following statements: *(1) I trust because of its good reputation (2) I trust despite of its bad reputation.* Statement 1 states that trust is placed based on reputation, while statement 2 reflects that a relying party has some extra knowledge about a party to trust, e.g., through direct experience or relationship that can overrule any positive or negative reputation. A fuzzy model for calculating trust based on a workflow was proposed in [9]. Rajbhandari et al. argue that provenance provides a useful way to capture information and to be used to evaluate trust and fuzzy rules enable greater degree of flexibility in assessing provenance information.

There are many forms and variations of risk and trust analysis, depending on the application domain, such as health care, finance, reliability and safety, IT security. In finance, risk analysis is concerned with balancing potential gain against risk of investment loss. In this setting risk can be both positive and negative. Within reliability, safety and IT security risk analysis is concerned with protecting existing infrastructure and assets. This paper focuses on analysing risk and trust of a health care system under specific use case. We are aiming to demonstrate that risk and trust are not necessarily proportional [2], but have an impersonal relation [3] and fulfill each other. In safety critical and health care systems, it is often stated that trust is better understood in terms of cost/benefit analysis and calculated risks, as well as by knowing provenance information. Therefore, in a situation when users should make critical decisions they users should be aware of possible outcomes and their probabilities, risks to be taken and uncertainties involved in the analysis as well as provenance of information.

As it can be seen, the research on trust often highlights importance of provenance. Moreover, the way trust is modelled depends on perspective of the domain and trust definition. We base our research on the assumption that trust can be enhanced knowing the quality of data and its provenance. Also, we make an assumption that knowing data related risks and their scale can improve the knowledge of a system, its processes and most critical data-related activities. In overall, knowing how data was processed, derived, operated, agents involved as well as associated trust and risk values provided at each stage of data processing.

IV. BII CASE STUDY

A neuroscience researcher wants to choose a set of data files on which to validate a new analysis technique. They use the BII portal to select files for appropriate patients, but want to be able to choose a subset of these files which represent the data which is the most trustworthy. For any given file, the researcher wants to see a summary which helps them understand to what extent they can trust the data and what is the level of risk associated with this data.

All files on the BII portal have associated metadata. If the metadata is not present, the data should be deemed to be less trustworthy. However, it will not necessarily mean the data is more risky, as the risk is associated with other parameters, such as threats of agent's failure, wrong data export settings and/or various bugs in software agents.

A. Provenance-Based Risk Model of a Domain-Based Workflow

In order to assess risk associated with making critical, high-value health decisions based on evidence presented by a system, it is essential to know how the data was derived, processed and transformed. For this purpose, we build on a workflow generated and associated provenance meta-data which is unique for each system under observation and contains the linking between system personnel, processes and documents along with configuration management information as a connected directed graph. The provenance modeling builds upon the W3C's de-facto ontological representation of PROV named PROV-O [15], which is defined using the W3C's Web Ontology Language OWL2 [16]. The provenance data consists of a list of entities from the workflow graph as well as provenance specific meta-data: software version, training data for software systems, personnel associated with system processes. Within STRAPP, we apply a quantitative risk assessment approach to estimate the level of risk possessed by the provenance data recorded within the PROV data model. Therefore, an identification of the elements of risk within the provenance chain becomes important. It should be noted, that the nature of risks may differ thus, the quantitative risk estimation too.

In order for a risk model to be applied to the BII use case, STRAPP first is used to generate a provenance chain. Based on a provenance chain risk model can be applied and relevant queries are made. As such, STRAPP performs a number of queries to the target system, where risk data is stored and dynamically monitored. Table 1 shows risk attributes generated by the BII system and risk matching combinations. A Domain expert usually is responsible for estimating the probability of such combinations and their impact. These data are then passed to STRAPP, which performs necessary calculations and risk aggregation as well as presents risk output on a scale from 1 to 7, where 1 is low risk and 7 is considered as high. Risk is calculated based on an Activity_ID, Entity used by and Agent associated with this Activity_ID. Fig. 1 shows an output from STRAPP

system based on BII use case data. The workflow illustrates a chain of processes starting from its initial data source (Patient) and finishing by an Entity "Diagnosis" made to the patient.

Threats and vulnerabilities shown in Table I are specific to the activities, entities and agents involved in the chain. The list can change depending upon the domain. Risks in BII domain are clearly associated with data completeness, relevance, accuracy (e.g., V2, V3, V4, V5, etc).

TABLE I. RISK COMBINATIONS

Vulnerability (Vi)	Threat (Tj)	Matching Combinations
Poor signal quality (V1)	Electrical Interference (T1)	V1T1, V1T3, V1T4
Incomplete Data (V2)	Software Agent Failure (T2)	V2T2, V2T8
Inaccurate values (V3)	Incorrect Calibration (T3)	V3T3
Incorrect data exported (V4)	Poor Electrode Contact (T4)	V4T5, V4T6, V4T7
Malfunction in a training model (V5)	Software agent Export failure (T5)	V5T5
Incorrect data set (V6)	Incorrectly labelled units (T6)	V6T6
Data set conversion failure (V7)	Wrong Export Settings (T7)	V7T10
Undetected event (V8)	Human agent error (T8)	V8T12, V8T6, V8T12
Detection routine failure (V9)	Human agent malicious intent (T9)	V9T11
Incorrect parameters chosen (V10)	Bug in conversion software (T10)	V10T6,
	Bug in detection software (T11)	
	Unseen event type (T12)	

From Fig. 1, risk is calculated per block. The block is defined in terms of an entity, activity and associated agent:

$$R_{block} \in (R_{ent}, R_{act}, R_{Ag}); \qquad (1)$$

where R_{ent}, R_{act}, R_{Ag} is risk of an entity, activity and agent respectively.

STRAPP is querying target system for an activity ID and string of risks with regard to this activity. The system should respond with a string of risks of an entity, activity and agent:

$$R_{ent}, R_{act} = \{R_1 ... R_n \}; \qquad (2)$$

Risk for an agent is defined in terms of agents' years of experience and assigned a factor from a scale of 0 to 1, where 1 is very experienced (e.g., more than 10 years experience, and 0 – no experience at all). As such, risk for an agent can be scaled as follows:

$$R_{Ag} \in [0.33, 0.66, 0.99]; \qquad (3)$$

Risk per block is aggregated as follows:

$$R_{agg_{act\_ID}} = 1 - (1 - R_1) * (1 - R_n) * R_{Ag}; \qquad (4)$$

Overall aggregated risk of a chain under analysis is calculated as follows:

$$R_{total} = 1 - \left(1 - R_{agg_{act_{ID_1}}}\right) * \left(1 - R_{agg_{act_{ID_2}}}\right) ... (1 - R_{agg_{actID_n}}) \qquad (5)$$
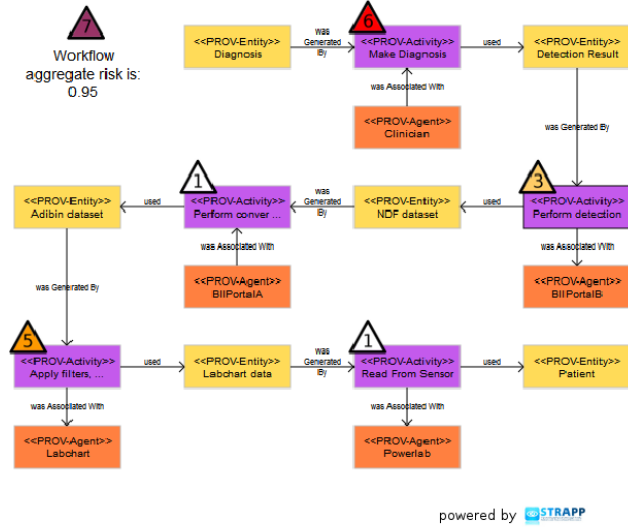

Figure 1. Risk output

Activity "*Make Diagnosis*" and agent "*Clinician*" has got high risk level. This is because agent's risk is defined in terms of its years of experience. Therefore, inexperienced clinician could make an incorrect diagnosis and result in a high aggregated workflow risk. More years of experience would dramatically reduce the overall risk of a final "*Diagnosis*".

B. Provenance-Based Trust of a Domain-Based Workflow

Our trust model is concerned with the ranking of decision alternatives over a number of attributes. Based on a case study data, some of the attributes can be incomplete. There are numerous methods to aid decision makers solve multi-attribute decision making (MADM) problems with incomplete information, amongst these methods the analytic hierarchy process (AHP) has been widely used, originally proposed by Saaty [13].

Our trust algorithm first identifies all possible focal elements from incomplete decision matrix, then it calculates the Basic Probability Assignment (BPA) of each focal element. Second, belief interval of each decision alternative is evaluated according Dempster-Shafer theory (DS). Third, by applying the ranking method decision, alternatives are determined by comparing their belief intervals. More details on DS/AHP and its application can be found in [14].

The following metadata fields shown in Table II contribute to the trust decision matrix:

TABLE II. TRUST METADATA

Field	Example value	Trust implications
Patient Identifier	KCH116	Conforms to expected format. These are a 3 letter centre ID concatenated to a three digit patient number. Trust is high or low based on presence/absence.
Centre	King's College Hospital	Reputation of centre. Should match with the patient identifier given above.
Sensor fitted by	John McAvoy	Experience/reputation of clinician. Initially based on the number of procedures carried out over the previous two years.
Data Administrator	Martyn Fletcher	Experience/training of data administrator. Each administrator is registered to upload data for a given centre- trust is reduced if data is uploaded for a different centre. Trust also based on number of files uploaded by the administrator (i.e. experience)
Data Channels	LPF0, LPF1,LPF2,HPF0,HPF1, HPF2,BP	Expected channels are present. Trust is reduced if channel names are not recognised as standard.
Recording frequency	200Hz	Is a standard recording frequency (200Hz and 400Hz are the current standards). Trust is reduced for other recording frequencies.
Start Date	21/01/2013	Date should be valid and in the past.
Recording Setup	Depth and strip probes through PowerLab	Trust is reduced in a less tried and tested setup. Where a large number of recordings have been made with a certain setup, the trust is increased.

Data on the BII portal contains provenance information about the services which were used to generate it, and the inputs to those services. This information is crucial in the determination of the level of trust which can be placed in the data. The following pieces of information shown in Table III are pertinent to the initial trust model, and will apply to all pieces of data/services in the provenance chain:

TABLE III. DATA PROVENANCE/SERVICE INFORMATION

Provenance Information	Example Value	Trust Implications
Service version	1.2	Should be the latest version of the service. Trust is reduced if an older service version was used. Additionally, some service versions may have known problems. Trust is greatly reduced where this is applicable.
Service creator	Stephen Hobson	Trust will be higher in service developers with more experience/better reputations
Number of service executions	1000	Trust will be higher in services which have been used a larger number of times
Data trustworthiness	0.9	Trust in this piece of data is partially defined by the trustworthiness of the data which was used to create it. This will have been defined using the trust model.

Some data, after analysis, will have some results associated with it, such as event detections. As part of this analysis, some measures may be available which would help determine the trustworthiness of the data. Initially these are limited, but could be increased in future. Table IV shows an example of such data:

TABLE IV. DATA

Data quality measure	Sample Value	Trust Implications
Channel Uptime	96.4%	This is a measure of the percentage of time that the channels in the file were providing "good quality data". Trust is higher where this value is higher.

Fig. 2 shows the trust levels derived by applying DS/AHP to input data shown in Tables II, III, IV. For every PROV element, trust level is estimated taking as an input a set of files with relevant data entries and applying DS/AHP algorithm the ranking is performed. As such, we have applied DS/AHP to rank the trust level at the source: Entity "Patient".
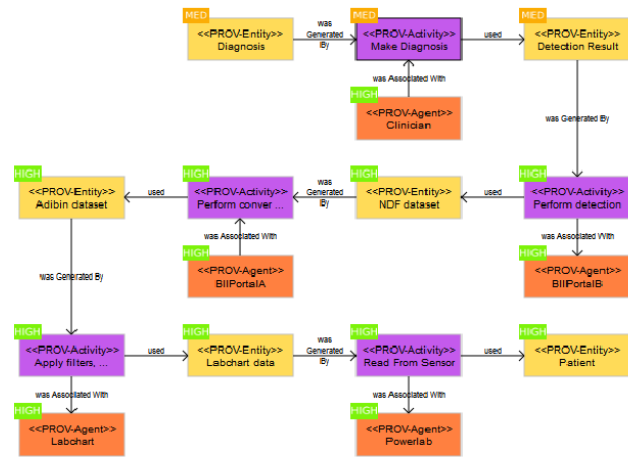
Figure 2 Trust output

The input to DS/AHP consists of 10 files, each with 8 data fields. As it can be seen in Table V, some of these fields are missing. Data fields such as patient_ID, center, sensor fitted by, administrator, data channels, recording frequency, and recording setup are treated equally, without emphasizing on importance. After running DS/AHP, it was derived that some of the files have low trust, e.g., "sample.ps". This is because most of the data fields are empty, missing or incomplete. Medium trust files have several empty fields. Similarly, the set of data files relevant to activities within a workflow can be analysed and ranked according to DS/AHP as shown in Table V. The user of a system can then see at what stage data might get lost, corrupted or tempered with. Therefore, somebody knowing such situation would be interested in knowing possible consequences or risks associated with the decision trust.

Risk and trust can be seen implicitly. As such, we have demonstrated risk view on a basis of a workflow taking as an input risks relevant to data completeness, accuracy, relevance. It was seen that high risk activities may also result in high trust, if the data is of a high quality. As such, we can compare risk and trust of an activity "Apply Filters" from Fig. 1 and Fig. 2. In terms of risk – "Apply Filters" risk level is 5 (out of 7) and trust is high. Risk was calculated knowing that a number of threats and vulnerabilities are present and may harm the data quality of a data set. However, trust algorithm when applied on this activity has shown high trust in data set, as most of the data fields were complete. Therefore, we have made an assumption, that knowing that trust level in data is high does not necessarily mean it has low risk. Risk in our context is more associated with external factors which are not considered by the trust algorithm, e.g., software bug, software agent export failure.

TABLE V. TRUST DECISION MATRIX RANKING RESULTS

Filename	Patient Identifier (2.0)	Center (2.0)	Sensor Fitted By (5.0)	Data Administrator (5.0)	Data Channels (2.0)	Recording Frequency (2.0)	Start Date (2.0)	Recording Setup (2.0)	Trust / Distance Metric
KCH101.txt	KCH101 (2.0)	KCH (2.0)	John Smith (2.0)	John Smith (2.0)	BP (2.0)	200Hz (2.0)	2013-11-29T14:04:34 (2.0)	Standard (2.0)	HIGH 0.1039
KCH116 Section.ndf	KCH116 (2.0)	KCH (2.0)	(0.0)	Stephen Hobson (2.0)	LPF 0, LPF 1, LPF 2, LPF 3, LPF 4, LPF 5, HPF 0, HPF 1, HPF 2, HPF 3, HPF 4, HPF 5, ICP (2.0)	200Hz (2.0)	2012-01-08T19:23:29 (2.0)	(0.0)	HIGH 0.16
KCH116 Section.mat	KCH116 (2.0)	KCH (2.0)	(0.0)	Stephen Hobson (2.0)	LPF 0, LPF 1, LPF 2, LPF 3, LPF 4, LPF 5, HPF 0, HPF 1, HPF 2, HPF 3, HPF 4, HPF 5, BP, ICP (2.0)	200Hz (2.0)	2012-01-08T19:23:29 (2.0)	(0.0)	HIGH 0.16
KCH011.ndf	KCH011 (2.0)	KCH (2.0)	(0.0)	Stephen Hobson (2.0)	BP (2.0)	200Hz (2.0)	2004-02-01T18:37:18 (2.0)	(0.0)	HIGH 0.16
KCH011.mat	KCH011 (2.0)	KCH (2.0)	(0.0)	Stephen Hobson (2.0)	BP (2.0)	200Hz (2.0)	2004-02-01T18:37:18 (2.0)	(0.0)	HIGH 0.16
KCH11522.ndf	KCH115 (2.0)	KCH (2.0)	(0.0)	Stephen Hobson (2.0)	BP (2.0)	(0.0)	2011-12-21T17:40:49 (2.0)	(0.0)	HIGH 0.187
KCH11522.mat	KCH115 (2.0)	KCH (2.0)	(0.0)	Stephen Hobson (2.0)	BP (2.0)	(0.0)	2011-12-21T17:40:49 (2.0)	(0.0)	HIGH 0.187
PG 29Apr_2ver1.ndf	PG29 (2.0)	PG (0.0)	(0.0)	(0.0)	(0.0)	200Hz (2.0)	2009-04-29T19:52:41 (2.0)	(0.0)	MED 0.2887
PG 29Apr_2ver1.mat	PG29 (2.0)	PG (0.0)	(0.0)	(0.0)	(0.0)	200Hz (2.0)	2009-04-29T19:52:41 (2.0)	(0.0)	MED 0.2887
sample.ps	Ps (2.0)	Ps (0.0)	(0.0)	(0.0)	(0.0)	(0.0)	Ps (0.0)	(0.0)	LOW 0.5299

V. CONCLUSION AND FUTURE WORK

Considering that trust and risk are implicit, this paper proposes for the first time a novel approach to view trust and risk on a basis of a W3C PROV provenance data model applied in the healthcare domain. We have made an assumption that high trust in data does not necessarily mean low risk, as these factors fulfill each other rather than being seen independently. This is demonstrated by our trust and risk models applied to the Brain Injury Index (BII) case study data. We first, present the risk model, which first starts by calculating risk values at each workflow step considering PROV concepts and second, aggregates the final risk score for the whole provenance chain. Different from risk model, trust of a workflow is derived by applying DS/AHP method. In situation when user should make a critical decision, users should be aware of possible outcomes and their probabilities, risks to be taken and uncertainties involved in the analysis as well as provenance of information. The system is trustworthy when these aspects are open to the system user. The evaluation of such system will be performed under the STRAPP context in the medical domain. We make a hypothesis that if user is aware of risks and trust levels involved in the PROV chain the trustworthiness in a system can be improved. Therefore, more analysis needs to be done in the area of risk and trust. Nevertheless, our first attempt of visualizing risk and trust concepts on a workflow basis and making a relational comparison of derived results proved our assumption that risk and trust are implicit, not proportional.

REFERENCES

[1] D. Gambetta, "Can We Trust Trust?", in D. Gambetta (Ed.) Trust: Making and Breaking Cooperative Relations, Electronic edition, Department of Sociology, University of Oxford, chapter 13, pp. 213-237, 2000

[2] B. Solhaug, D. Elgesem, and K. Stølen. "Why trust is not proportional to risk", 2nd International Conference on Availability, Reliability and Security (AReS'07), pp. 11-18, IEEE Computer Society, 2007.

[3] O.E Williamson, "Calculativeness, Trust, and Economic Organization", Journal of Law and Economics, University of Chicago Press, vol. 36(1), pp. 453-86, April 1993

[4] A.Jøsang, C. Keser, and T. Dimitrakos, "Can We Manage Trust? ", in Trust Management, P. Herrmann, V. Issarny, and S. Shiu (Eds.), Springer, Berlin Heidelberg. pp. 93-107, 2005

[5] A. Josang and S. L. Presti, "Analysing the Relationship Between Risk and Trust", 2nd International Conference on Trust Management (iTrust'2004), Oxford, UK, Springer, pp. 135-145, April 2004.

[6] M. Emaldi et al., "To trust, or not to trust: Highlighting the need for data provenance in mobile apps for smart cities", Computer Vol. 15, pp. 26-32, 2013

[7] O. Hartig and J. Zhao, "Using web data provenance for quality assessment", In: Proc. of the Workshop on Semantic Web and Provenance Management at ISWC, 2009

[8] D. Ceolin, P. Groth, W. R. van Hage, A. Nottamkandath, and W. J. Fokkink, "Trust evaluation through user reputation and provenance analysis", 8th Workshop on Uncertainty Reasoning for the Semantic Web (URSW'2012), Boston, Massachussetts, pp. 15-26, Nov. 2012

[9] S. Rajbhandari, O. F. Rana, and I. Wootten. "A fuzzy model for calculating workflow trust using provenance data", 15th ACM Mardi Gras conference, New York, NY, USA, pp. 1–8, 2008. ACM.

[10] V. Cahill et al. ,"Using trust for secure collaboration in uncertain environments", Pervasive Computing, IEEE , Vol. 2, No.3, pp.52-61, July 2003

[11] L. Moreau and P. Missier, "PROV-dm: The prov data model". W3C Recommendation, April 2013

[12] D. Ceolin, P. Groth, and W.R.van Hage, "Calculating the Trust of Event Descriptions using Provenance", Second International Workshop on the role of Semantic Web in Provenance Management (SWPM'10)

[13] T.L.Saaty, "The Analytic Hierarchy Process", New York:McGraw-Hill, 1980

[14] Z. Hua, B. Gong, and X. Xu, "A DS–AHP approach for multi-attribute decision making problem with incomplete information", Expert Systems with Applications, Vol. 34, iss. 3, pp. 2221-2227, ISSN 0957-4174, April 2008

[15] T. Lebo, S. Sahoo, and D. McGuinness, "PROV-O: The PROV Ontology", W3C Recommendation, April 2013

[16] W3C OWL Working Group, "OWL2 Web Ontology Language", W3C Recommendation, December 2012

[17] STRAPP: https://www.engineering.leeds.ac.uk/strapp