A Layered Approach to Managing Risks in OSS Projects

Xavier Franch[1], Ron Kenett[2], Fabio Mancinelli[3], Angelo Susi[4], David Ameller[1], Ron Ben-Jacob[2], and Alberto Siena[4]

[1] Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
[2] KPA, Raanana, Israel
[3] XWiki, Paris, France
[4] Fondazione Bruno Kessler (FBK), Trento, Italy
{franch,dameller}@essi.upc.edu, {ron,ronb}@kpa-group.com,
fabio.mancinelli@xwiki.com, {susi,siena}@fbk.eu

Abstract. In this paper, we propose a layered approach to managing risks in OSS projects. We define three layers: the first one for defining risk drivers by collecting and summarising available data from different data sources, including human-provided contextual information; the second layer, for converting these risk drivers into risk indicators; the third layer for assessing how these indicators impact the business of the adopting organisation. The contributions are: 1) the complexity of gathering data is isolated in one layer using appropriate techniques, 2) the context needed to interpret this data is provided by expert involvement evaluating risk scenarios and answering questionnaires in a second layer, 3) a pattern-based approach and risk reasoning techniques to link risks to business goals is proposed in the third layer.

Keywords: OSS, Open Source, Risk Management, Layered Model.

1 Introduction

Translating dynamics of a complex system into focused management insights has been a challenge in various application domains [1]. In this paper we focus on organisations adopting, integrating and maintaining open source software (OSS) components in order to reduce time to market, introduce innovation and overcome development bottlenecks. Several companies have been observed to understand which are the main risks and risks indicators related to this OSS-related activities that are perceived by technical and business managers [2].

We propose a three layered approach: 1) the first layer focuses on collecting and summarising available data from different data sources, including human-provided contextual information; 2) the second layer, converts these data risk drivers into risk indicators [3] and 3) the third layer assesses how these indicators impact into the business of the adopting organisation. Key methodological as well as theoretical questions need to be answered to derive risk related insights from measurable data as described in [4][5] such as, the indicators to define for measuring risk events, how to operationalise an indicator into one or more specific metrics for measurement and the predictive ability of measurements related to risks events needs to be validated.

L. Corral et al. (Eds.): OSS 2014, IFIP AICT 427, pp. 168–171, 2014.

2 A Layered Approach to Risk Management

Here we describe the three layered approach to risk management in OSS projects that we are proposing in the EU project RISCOSS [6], see Fig. 1.

2.1 Layer 1: Raw Data and Risk Driver Measures

In this layer we deal with data collected from OSS communities and projects that determines the risk drivers. The data has a twofold nature. On the one side it refers to the characteristics of the OSS components developed by the communities, e.g. *Number of open bugs*, *Forum posts per day, Mails per day*, *Amount of documentation.* On the other hand, other measures highlight the structure of the community in terms of its evolution, e.g. changes in its roles and members and in the quality and quantity of relationships between them mainly via social network analysis techniques [7].

The data sources are community repositories, versioning systems, mail lists, bug trackers and forums, among others. Human intervention may be eventually needed because of: 1) data sources that may be unavailable for a particular component or community, 2) values that can eventually be calculated but require a dedicated activity to do so, 3) values that are not directly accessible or are very costly to compute.

2.2 Layer 2: Risk Indicators and Risk Model

In this layer we define the set of indicators of possible risks and models that allow linking these risks to the possible objectives of the adopting organisation. The indicators are variables extracted via the OSS community data analysis obtained from OSS project measurements and OSS community measurements as described before or via expert assessment. Several categories of indicators can be observed. Here we refer to three of them: 1) risk indicators related to OSS projects can be grouped following some criteria such as *Maintainability*; 2) risk indicators related to the communities coming from the aforementioned community measures, e.g. *Community activeness* or *Community cohesion*; 3) contextual risk indicators, elicited from experts, mainly depend on the objective of the organisation, e.g. *OSS business strategy*.

Here Statistical Analysis, Bayesian Networks and Social Network Analysis are exploited to determine values of risk. In particular:

Fig. 1. The 3-layered approach

- Statistical analysis of data from OSS communities allows determining the trends and distributions of data.
- Bayesian networks are used to link the community data gathered from the community data sources and the community risk metrics to the risk indicators and the community risk indicators using data generated by experts' assessment based on their experience in OSS adoption and community context.
- The community measures can be also analysed via Social Network Analysis techniques in order to understand the structure and evolution of the OSS community.

All the risk indicators will contribute to the definition of a risk model. This model allows the representation of the possible causes of risks, basically the risk indicators, and of their connection to the possible risk events for the adopter organisation. Moreover, the model also allows representing the impact that the possible risk events have on the strategic and business goals of the organisation.

2.3 Layer 3: Business Goals

Business goals describe which are the aims of the organization that adopts OSS. They are impacted by several kind of risks we summarise into four categories: 1) Strategic risks, mainly related to the company's strategy and plan, such as *Pricing Pressure*, *failures in comply Regulation*, *Industry or sector downturn*, or *Partner issues*; 2) Operational risks such as *poor capacity management* or *cost overrun*; 3) Financial risks such as *assets lost*, *debts* or *accounting problems*; 4) Hazard risks related to, for example, *macroeconomic conditions* or to *political issues*. Also in this case Bayesian networks may be used in order to link concepts from the two layers.

2.4 Modelling the Layers

Business goals are included in models that represent the ecosystem that blends together communities, OSS adopting organizations and other key actors. The key relationships between these actors are represented through dependencies in goal-oriented models expressed in the *i** language [8], which allow representing, and reasoning about, business goals and business processes. Reasoning is based upon different techniques, and in our layered context, we are particularly interested in bottom-up evaluation, since the leaves are directly linked to the risk model.

A typical model will include the two fundamental actors of the OSS ecosystems, the Community and the Adopter, and how they depend on each other; some of their internal goals and activities, and their further AND/OR decompositions. We have then the risk model with the risk event (e.g., *Risk of difficulty in code refinement*) that "impacts" one of the activities of the Adopter (e.g., *Bug Report*) and that is propa-gated up to the higher level activities. The risk event is identified via the measurement and statistical analysis of the behaviour of the community and on the expert intervention that can rate the evidence of a risk indicator via the Bayesian Networks.

3 Conclusions and Future Work

The RISCOSS framework is designed to face with the problem of risk management in OSS related projects in a holistic way, allowing to pass smoothly from the dimension of the measures to those related to the decision-making in contexts where several technical and business constraints are present.

We believe that the approach can give an effective way of overcoming problems related the adoption phase. In particular, the huge volume and potential heterogeneity of the data is isolated into a layer collecting the available and potential new techniques suited for this problem; the correct interpretation in the context of the adopting organization is made also with the help of experts that can evaluate specific scenarios of risks; a pattern-based approach and risk reasoning techniques is proposed in the third layer that can help in linking risks to business goals.

Several points have to be addressed in the following years of the project. We plan to refine the approach clearly defining the boundaries of the layers and adding to each one of the layers the suitable techniques for data reasoning. An important point here is that of developing the approach in such a way to be adapted to the needs of the particular organisation that should be able to also feed in a contextual way the necessary data to effectively exploit the approach. Also we plan to integrate better our results to those coming from projects with related aims, as FLOSSMetrics (http://flossmetrics.org/), QualiPSO project (http://qualipso.org/), QualOSS (http://www.qualoss.eu/) and OSSMETER (http://www.ossmeter.eu/).

Acknowledgments. This work is a result of the RISCOSS project, funded by the EC 7th Framework Programme FP7/2007-2013 under the agreement number 318249.

References

1. Harel, A., Kenett, R.S., Ruggeri, F.: Modeling Web Usability Diagnostics on the basis of Usage Statistics. In: Statistical Methods in eCommerce Research, Wiley (2009)
2. Li, J., Conradi, R., Slyngstad, O., Torchiano, M., Morisio, M., Bunse, C.: A State-of-the-Practice Survey of Risk Management in Development with Off-the-Shelf Software Components. IEEE Trans. Software Eng. 34(2) (2008)
3. Kenett, R.S., Baker, E.: Process Improvement and CMMI for Systems and Software: Planning, Implementation, and Management. Taylor and Francis, Auerbach Pub. (2010)
4. Ligaarden, O.S., Refsdal, A., Stolen, K.: ValidKI: A Method for Designing Key Indicators to Monitor the Fulfillment of Business Objectives. In: BUSTECH 2011 (2011)
5. Wallace, L., Keil, M.: Understanding software project risk: a cluster analysis. Inf. Manage. 42(1) (2004)
6. Franch, X., Susi, A., Annosi, M.C., Ayala, C., Glott, R., Gross, D., Kenett, R., Mancinelli, F., Ramsamy, P., Thomas, C., Ameller, D., Bannier, S., Nili Bergida, N., Blumenfeld, Y., Bouzereau, O., Costal, D., Dominguez, M., Haaland, K., López, L., Morandini, M., Siena, A.: Managing Risk in Open Source Software Adoption. In: ICSOFT 2013 (2013)
7. Salter-Townshend, M., White, A., Gollini, I., Murphy, T.B.: Review of statistical network analysis: models, algorithms and software. Statistical Analysis & Data Mining 5(4) (2012)
8. Yu, E.S.K.: Modelling strategic relationships for process reengineering. PhD thesis, University of Toronto, Toronto, Ont., Canada (1995)