Open Source Software and Open Data Standards as a form of Technology Adoption: a Case Study

Bruno Rossi, Barbara Russo, and Giancarlo Succi
Free University of Bolzano-Bozen, Faculty of Computer Science,
Domenikanerplatz 3, 39100 Bolzano-Bozen, Italy
{bruno.rossi,barbara.russo,giancarlo.succi}@unibz.it
WWW home page: http://www.case.unibz.it

Abstract. The process of technology adoption has been studied for long time to give instruments to evaluate the best strategies to ease the introduction of technology. While the main research on Open Source Software focuses mainly on the development process, team collaboration and programmers' motivations, very few studies consider Open Source Software in this context. In this paper, we provide an overview of literature on technology adoption that can be useful to relate the concepts. We then provide a case study with historical data about file generation and usage across time to evaluate the adoption of Open Source Software and Open Data Standards in the specific case provided.

Keywords: Open Source Software; Data Standards; Technology Adoption.

1 Introduction

Open Source Software (OSS) has acquired recently a growing popularity in the software market. The free availability of source code and the freedom to modify and redistribute the source code are the main characteristics that are at the basis of its crescent popularity. Particularly in the governmental setting, these characteristics have increased the interest towards OSS. The Interoperable Delivery of European e-Government Services to public Administrations, Businesses and Citizens organisation (IDABC), identifies five aspects in this context, that can be of interest for organisations willing to adopt OSS [10]:

- political aspects, concepts related to governmental tasks, goals and responsibilities like freedom and equality, digital endurance, digital heritage and stimulation of innovation;
- economical aspects, related to cost reduction and market health;
- social aspects, in particular for education and team work support;

- managerial and/or technical aspects, in particular quality of the products in terms of stability and reliability, transparence, support and security;
- legal aspects, related to licensing and liability.

All these different point of views make the adoption of OSS inside organisations a very appealing option.

Furthermore, a concept sometimes overlooked, but frequently associated to OSS is the one of Open Data Standards (ODS). ODS are a subcategory of data standards. Data standards provide a standardised way to store different typologies of data, and emerge generally in two different ways, as output of an evolution of the market (so called *de facto* standards*)* or after being recognised by a standardisation committee (*de jure* standards). The distinction that is of our interest is between Open and Proprietary data standards. In this sense, many different definitions of ODS exist, we would like to propose the definition given by the Danish Board of Technology in 2002 [4]:

- An open standard is accessible to everyone free of charge (i.e. there is no discrimination between users, and no payment or other considerations are required as a condition of use of the standard);
- An open standard of necessity remains accessible and free of charge (i.e. owners renounce their options, if indeed such exist, to limit access to the standard at a later date, for example, by committing themselves to openness during the remainder of a possible patent's life);
- An open standard is accessible free of charge and documented in all its details (i.e. all aspects of the standard are transparent and documented, and both access to and use of the documentation is free);

As can be noted, the importance of open standards lies, in particular, in the avoidance of the commitment to a single supplier. In this paper, we review the adoption process of Open Source Software (OSS) and Open Data Standards (ODS) in an empirical case, by analysing the file generation and usage process in a single Public Administration. In this early work, we start to insert the empirical case studied in the context of technology adoption.

We first introduce two main concepts that have been discussed extensively in literature: technology adoption and lock-in. While the review we provide is not strictly related to OSS, it is necessary for the overview of the next section about technology adoption studies related to OSS. In the final part, we provide the details of the case study and the main results obtained.

2   Technology adoption

Technology adoption, diffusion and acceptance research bases its foundation on the early work of Everitt Rogers, in the book titled Diffusion of Innovations [11]. Rogers interest lies in studying the diffusion process that characterises technology adoption. In his seminal work, technology adopters are categorised according to the phase in which they make the adoption decision. The main distinction is among innovators,

early adopters, early majority, late majority and laggards. In particular, the author models the diffusion as an S-shaped curve characterised by an initial adoption speed and a later growth rate. The claim is that different technologies will lead to different adoption patterns.

Fichman & Kemerer [8] report two critical factors that influence the technology assimilation process: knowledge barriers and increasing returns.

The first effect relates to the effort necessary to acquire the necessary knowledge and skills to properly adopt a certain technology. This effect leads to what are known as knowledge barriers [2,8].

The second phenomenon, reports that the adoption of certain technologies is subject not only to supply-side benefits due to economies of scale [12], but also to a demand-side effect called increasing returns effect [1]. The effect leads to an increase of the utility in adoption for each successive adopter, based on the number of previous adopters. Arthur goes further in this analysis, claiming that economy can become locked-in to a technological path that is not necessarily efficient, not possible to predict from usual knowledge of supply and demand functions, and not easy to change by standard tax or subsidy policies [1]. In this sense, it may not be possible to easily switch from a certain technology once a certain critical level of adoption has been reached.

Considering OSS, there are not many studies that evaluate OSS from this point of view. An interesting overview is given in [5], where following the "context for change methodology" defined in [6], factors that lead adoption process are categorised in technological, organisational and environmental.

Glynn et al. [9] developed a framework for assimilation based on four categories: external environment, organisational context, technological context and individual factors. The framework is then applied to a large-scale survey.

Bitzer and Schroder [3], analyse the innovation performance of Proprietary and Open Source Software, showing the results of the competition between the two software typologies in different market settings. The focus is more on innovation that on the adoption process itself.

Economides [7] studies the incentives that lead to platform innovation. A case study between Linux and Windows is provided.

3   A case study of OSS migration

To provide some real data about a concrete case of OSS and ODS adoption process, we consider an experimentation that took place during an experimental migration from Microsoft Office to OpenOffice.org in one medium-size European Public Administration. The users involved were 100. Data have been collected by means of the PRO Metrics (PROM) software [13], software that permits to collect metrics about software usage, and FLEA (FiLe Extension Analyzer), software that allows collecting information about the data standards available on the target system.

Operations performed were the installation of OpenOffice.org in parallel with the available version of Microsoft Office, installation of the PROM agent, scan of the file-system with FLEA, training of users and support.

4   Main results

We report the main result from the analysis of data standards and software usage. In table 1, we show the total number of all the data standards collected at the beginning of the experimentation, divided per category.

Table 1. Total number of data standards detected by the data collection software

Text Documents		Graphic Format		Database		
DOC	310648	BMP	6908	DB		3361
DVI	0	GIF	36259	DBF		6865
PDF	12518	JPEG	83143	MDA		10
PS	656	PNG	2173	MDB		2170
RTF	6185	SVG	0	Music		
SXW	160	TGA	0	MP3		967
TEX	1	TIFF	11061	RA/RM/RAM		4
TXT	10422	Drawing		Movie		
Spreadsheets		DWG	36051	AVI		265
SXC	47	DXF	411	MOV		227
XLS	60267	SXD	9	MPEG		77
Presentations		Web		SWF		232
PPT	2541	CSS	1370			
SXI	27	HTML	16057			
Compression		XHTML	0			
ACE	1	Data Exchange				
ARJ	37	CSV	64			
GZ	19	DTD	6			
RAR	43	SDXF	0			
TAR	0	XML	483			
ZIP	5338					

As we can see, some data standards are largely predominant in their category, like DOC (Microsoft Word) documents, or XLS documents (Microsoft Excel). The former accounts for 91,21% of the files in the category, while the latter 99,92%. Also the ZIP format is largely dominant, with a percentage of 98,16%. If we use the Shapiro & Varian [12] categorisation of switching costs and consider the information and databases category, we can evaluate that a complete migration and adoption of the platform can be costly, due to the effort necessary required by the conversion of a large amount of documents.

We further studied the evolution of file generation across time, in figure 1 we show the generation of DOC and XLS documents, from the data collected the more representative for proprietary formats.
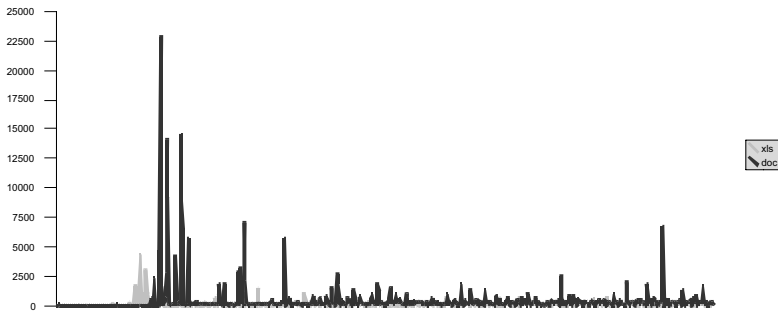
Fig. 1. Evolution of DOC and XLS files created by users during years, on x-axis time in days, on y-axis number of files generated

While the figures are given for representational purposes, we may note an interesting phenomenon that emerges by analysing data of many data standards. The creation process of documents is very consistent once a critical level of adoption has been reached. Once a large set of base documents has been constituted, the creation process somehow fades, as the activity of users will be constituted in small part also by the modification of already available documents.

5   Conclusions

While OSS research community is concerned mostly with studying the team and collaboration dynamics of the development process, OSS and ODS have still to be well studied as a form technological innovation. We overviewed some of the technology adoption literature that may be useful in this sense and some recent works that manage to insert OSS in this context. We further considered ODS as an important and often overlooked instrument that has to be associated to OSS when considering its adoption. We studied as a case study, the evolution of a migration to OSS in the office automation field, considering data standards as a sign of the presence of possible lock-in phenomena. The data analysed show the commitment of the organisation under study to proprietary data formats, in particular in the office automation category.

Acknowledgements

References

[1] Arthur, W.B. (1989). Competing Technologies, Increasing returns, and lock-in by historical events. Economic Journal, 99, 116-131.

[2] Attewell, P. (1992). Technology Diffusion and Organizational Learning: The Case of Business Computing. Organization Science. 3(1), 1-19.

[3] Bitzer, J. and P. J. H. Schroder, (2003). Competition and Innovation in a Technology Setting Software Duopoly. DIW Discussion Paper No. 363.

[4] Danish Board of Technology. (2002). Definition of open standards. Retrieved, 14th January 2007, from http://www.oio.dk/files/040622_Definition_of_open_standards.pdf

[5] Dedrick, J., West, J., (2004). An Exploratory Study into Open Source Platform Adoption, HICSS, p. 80265b, Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 8, 2004.

[6] Depietro, Rocco, Edith Wiarda and Mitschell Fleischer, (1990). The Context for Change: Organization, Technology and Environment, in Tornatzky, Louis G. and Mitchell Fleischer, The processes of technological innovation. Lexington, Mass.: Lexington Books, 1990, pp. 151-175.

[7] Economides, N., Katsamakas, E. (2006). Linux vs. Windows: A comparison of application and platform innovation incentives for open source and proprietary software platforms, in Juergen Bitzer and Philipp J.H. Schroeder (eds.) The Economics of Open Source Software Development, Elsevier Publishers, 2006

[8] Fichman, R. G., & Kemerer, C. F. (1997). The Assimilation of Software Process Innovations: An Organizational Learning Perspective. Management Science. 43(10), 1345-1363.

[9] Glynn, G, Fitzgerald, B and Exton, C. (2005) Commercial adoption of open source software: an empirical study, Proceedings of International Conference on Empirical Software Engineering, Noosa Heads, Australia, Nov 2005.

[10] IDABC, The Many Aspects of Open Source. Retrieved, 14th January 2007, from http://ec.europa.eu/idabc/en/document/1744

[11] Rogers, E. (1995). Diffusion of Innovations. N.Y.: The Free Press.

[12] Shapiro, C., & Varian H.R. (1999). Information Rules: A Strategic Guide to the Network Economy. Harvard Business School Press.

[13] Sillitti A., Janes A., Succi G., Vernazza T. (2003). Collecting, Integrating and Analyzing Software Metrics and Personal Software Process Data, in proceedings of EUROMICRO 2003, Belek-Antalya, Turkey, 1 – 6 September 2003