

Open Source Mythologies and Reality in the Web-based Periodicals

Andres Baravalle

Department of ICT, The Open University

email: a.baravalle@open.ac.uk

Sarah Chambers

Department of Computer Science, The University of Sheffield

email: s.chambers@dcs.shef.ac.uk

Abstract

Large organisations have announced in the last 3 years wide Open Source migrations. Open Source is announced as the panacea that will solve cost problems. Nevertheless, the reality shows that Open Source is not accepted in some contexts, and confusion about the reality of Open Source migrations is still wide.

Italy, Spain, Germany, France and Austria are leading experimentations, while the United Kingdom seems to be unaffected by this stream. In this paper we aim to investigate the approach of the media when discussion Open Source, analysing three web-based periodicals from Italy, United Kingdom and USA.

Keywords: Open Source, Public Opinion, On-line periodicals, Press, Italy, United Kingdom, USA.

Introduction

In the recent years Open Source has been under the focus of a number of researches. Several authors focused on the analysis of the more sociological aspects of Open Source (Stallman 2004; Victor 2003; Moglen 1999; Benkler 2002), others analysed the more pragmatic aspects (Bonaccorsi & Rossi 2003; Haruvy, Prasad & Sethi 2003; Hertel, Niedner & Herrmann 2003; Lakhani & Wolf

2003; Lerner & Tirole 2001; Lerner & Tirole 2002).

This paper aims to contribute to the latter line of works by considering the motivations that are driving the adoption of Open Source, by investigating the press coverage of Open Source.

The starting point for our work is the experience of the COSPA project (Kovaks et al. 2004), a project that is investing the use of Open Source and Open Data Standards in the public administrations in Europe. The project involved partners from the academia, from the industry and several public administrations (three in Italy, two in Ireland, one in the United Kingdom, one in Denmark and one in Hungary).

Apart from the partners, a number of observers and subcontractors have been involved, at different levels, in the COSPA project. The COSPA project has led the experimentation of 12 public administrations (between partners, subcontractors and observers), that have been replacing existing proprietary software with Open Source alternatives. Nearly 2,000 users have been involved between the partner organisations (effective migrations to Open Source software), while subcontractors and observers, jointly considered, report 3,300 between target and ongoing experimentations.

Other public administrations in Europe, not linked to the COSPA project, have announced Open Source migrations, the main ones being the governments of Andalusia and Extremadura (10,000 computers migrated to Linux and a target of 100,000, using Guadalinex) (IT Analysis 2003; De Cala 2005), the municipalities of Paris (17,000 desktops) (Sayer 2005), Munich (14,000 desktops) and Schwäbisch Hall (500 desktops) (IDABC 2005), Vienna (2,500 to 7,500 desktops) (Marson 2005), a number of state schools in the Former Yugoslav Republic of Macedonia (5,000 desktops) (Angov 2005). The numbers reported are interesting: amongst the five more populated countries of the European Union (Germany, United Kingdom, France, Italy and Spain), the United Kingdom is the only one that does not have any relevant successful Open Source experimentation. While some authors (Waring & Maddocks 2005) report successful Open Source experimentations in public administrations in the United Kingdom, we have already underlined (Baravalle et al. 2006) that the software used in those migrations is not commonly considered Open Source.

Interviews and casual conversation with personnel working in the public administrations, that we have met during the months of the COSPA project, suggests that there is a different perception of Open Source across different countries – considered generally more positively in some countries than in others.

The question that arises, and the investigation line that we have been following, is: why Open Source has such different levels of acceptance and success in different countries, and which are the factors that influence it?

Decisions depend on a number of different technical factors, including technical, economical and

socio-cultural, but the information that the persons have plays a fundamental role in the evaluation of the factors (Lippmann 1950). When this research was started, we thought that the press would have been analysing Open Source in very different ways in different countries, with inevitable consequences in the definition of a public opinion on Open Source. The next sections will compare the articles featured in a selection of web-based periodicals, highlighting the more important considerations that we could infer.

Method

As a basis for our research, we have selected three web-based periodicals from the United Kingdom, Italy and the USA, analysing the articles on Open Source published in October and November 2005. The following paragraphs describe the motivations of our decision and the methodology that we have used to collect and analyse the data.

The decision of focusing on web-based periodicals was based mainly on the technical advantages, as it allowed ways to search faster for the required content. Analysing the paper copies of printed periodicals would have been far slower, but we do not exclude to analyse printed periodicals which provide subscription-based web versions of their content in the future.

We do not aim, at least at the present time, to identify global patterns or rules: our aim is to investigate in the specificity of United Kingdom, Italy and the USA, that appeared to have, in our experience, different attitudes towards Open Source.

For such scope, we selected a periodical focused on IT in each of the countries: The Register for the United Kingdom, Punto Informatico for Italy and C-Net News.com for the USA.

It is not easy to estimate the number of readers of the web-based periodicals without having access to reserved information. It is instead possible to have an estimate of the traffic, and for this we used the data provided by Alexa (alexa.com). In the period that we are examining, C-Net News.com is estimated to have had 4,000-11,000 millions of page views per day, The Register has a more modest 500-1,300 millions of page views per day and Punto Informatico had 80-170 millions of page views per day.

All the selected periodicals are amongst the most read IT periodicals in their respective countries. A survey in the search engines highlighted that The Register and C-Net News.com, while targeted to the respective countries, did intercept a international audience as well: it is estimated that 1 out of 5 persons in the world speak English at some level (Graddol 2000). The audience of Punto Informatico is limited to 70 millions circa of first language Italian speakers plus roughly other 60 millions of second language speakers worldwide, but it does not appear to have a relevant number of international visitors.

To harvest the information from the periodicals, we have developed a set of parsers, that can be used through a web interface. The parsers allow to easily store in a local, relational database the results on a free-text query on the archives of the periodicals.

Using the parser, we have been able to collect information about more than 13,000 articles on Open Source, from 1998 till the present, and to store them in a way suitable for further elaboration. To search for pages featuring content on Open Source, we used keywords such as “Free Software”, “Open Source”, “Mozilla”, “Linux” and “Open Office”.

While the research of this paper is focused on the articles that have been published between October and November 2005, but for a selection of topics we will refer to a wider time frame.

Results

For the period under evaluation, we retrieved a total of 469 articles on Open Source. The number of articles on Open Source in the selected time frame is not noticeably different from the yearly average for the periodical (Table 1).

	Total number of articles in Oct.-Nov. 2005	Average articles per month in Oct.-Nov. 2005	Total number of articles in Dec. 2004-Nov. 2005	Average articles per month in Dec. 2004-Nov. 2005
The Register	144	72	796	66
Punto Informatico	128	64	679	57
C-Net News.com	197	99	1250	104
Total	469			

Table 1: Open Source articles

To have an idea of the context in which the articles on Open Source appeared, we looked at the articles, in the period October-November 2005, that contained the keyword “Microsoft”. The Register had 238 articles (119 per month), C-Net News.com had 438 (219 per month) and Punto Informatico had 156 (78 per month). We can safely state that in all the periodicals the interest on Open Source was high, compared with its diffusion and with the number of articles on Microsoft. As the next step in our research, we have analysed the articles categorizing them by level of relevance (Table 2) and main topic (Table 3).

The analysis of the relevance was done to be able to select the number of articles that were discussing issues related to Open Source, compared to the number of articles that had just were just referring to Open Source en-passant, without any further development.

Periodical	Date	Level of relevance of article to Open Source		
		Mild	Some/Focus	Total
The Register	Oct-2005	24 (39%)	38 (61%)	62
	Nov-2005	32 (43%)	43 (57%)	75
Punto Informatico	Oct-2005	14 (26%)	39 (74%)	53
	Nov-2005	13 (19%)	55 (81%)	68
C-Net News.Com	Oct-2005	32 (30%)	75 (70%)	107
	Nov-2005	36 (42%)	49 (58%)	85
Total		151 (32%)	299 (68%)	450

Table 2: Relevance of articles

From Table 2 we can see how the the articles on Open Source had had a different level of relevance in the three periodicals: in Punto Informatico a higher percentage of articles that were referring to Open Source were not doing that superficially.

Periodical	Software	Community	Business	Licensing	Trivia	Security	Migrations	Total
The Register	38 (47%)	3 (4%)	14 (17%)	9 (11%)	11 (14%)	4 (5%)	2 (2%)	91
Punto Informatico	57 (61%)	8 (9%)	7 (7%)	10 (11%)	3 (3%)	7 (7%)	2 (2%)	94
C-Net News.Com	62 (49%)	7 (6%)	30 (24%)	12 (10%)	1 (1%)	11 (10%)	1 (1%)	126
Total	164	20	55	31	5	26	5	306

Table 3: Main topics of the articles that had some focus on Open Source or were focused on Open Source

In Table 3 we grouped the articles that had some focus on Open Source or were focused on Open Source in a number of exclusive categories:

- Software: including articles focusing on describing software or software standards;
- Community: including political and sociological articles;
- Business: including articles on acquisitions, funding, marketing, personnel news;
- Licensing: including articles on patents, licenses, intellectual property violations, legal cases;
- Trivia: including articles that are on Open Source but do not provide factual information to the reader (e.g. joke articles);
- Security: articles on software vulnerabilities, bugs, anonymity, malware and counter measures;
- Migrations: articles on migrations to Open Source.

Articles that might have been falling under more than one category have been classified according to the predominant classification.

Some categories have more or less the same coverage across the publications: Community, Licensing, Security, Migrations have less than a 5% difference in the numbers of featured articles. The main difference is in the categories Software and Business: Punto Informatico dedicated 14% more space to software news compared to The Register, and 12% compared to C-Net News.com. At the same time, Punto Informatico dedicated less space to business (10% less compared to The Register and 17% less compared to C-Net News.com). The Register has been featuring more intensively articles in the category Trivia, including letters and book adverts.

Before starting the research, we have been under the impression that one of the causes of the less propension to Open Source in the United Kingdom could have been caused by a lack of cover in the press on the ongoing migrations, compared to other countries like Italy. The data of our sample period shows that this is not true, and a more in-depth analysis shows that it is neither true using a longer time frame.

Periodical	Munich	Extremadura	Paris	Vienna
The Register	27/05/2003 (5 articles in total)	16/6/2003 (1 article in total)	02/07/2004 (2 articles in total)	22/11/2005 (1 article in total)

Punto Informatico	18/06/2004 (3 articles in total)	12/05/2005 (1 sentence in 1 article on Ubuntu)	22/11/2005 (1 article in total)	07/07/2005 (1 article in total)
C-Net News.Com	28/05/2003 (3 articles in total)	no references	12/02/2004 (2 articles in total)	25/01/2005 (1 article in total)

Table 4: Date of the first article on Open Source migrations

It is worth mentioning that some of the more important Open Source migrations in Europe in the past 3 years have been ignored by all the periodicals we are evaluating. This applies to the migration of Schwäbisch Hall and to the migrations in the state schools in the Former Yugoslav Republic of Macedonia. COSPA is referred 2 times in Punto Informatico, but with almost no information on the migrations, and none of the COSPA migrations (some of which have important numbers) has been featured in The Register and C-Net News.com.

As it is evident from Table 4, both The Register and C-Net News.com have been featuring more articles on Open Source migrations compared with Punto Informatico. Moreover, Punto Informatico has been always the last periodical to report on the migrations.

Discussion

In the analysis we have showed how:

1. During the period December 2004-November 2005, The Register and Punto Informatico featured a comparable average number of articles per month (67 for The Register, 57 for Punto Informatico), while C-Net News.com featured an average of 104 articles per month on Open Source;
2. In all the periodicals Open Source is well represented – there is not a cloud of information overload, with articles on Open Source in a ocean of articles on proprietary software;
3. Punto Informatico had a higher percentage of articles focused on Open Source.
4. Punto Informatico had a higher ratio of articles on Open Source compared with the number of articles including the keyword “Microsoft”;
5. Most of the articles on Open Source in Punto Informatico were focusing on Open Source itself, while both The Register and C-Net News.com had a higher number of articles on business aspects;
6. Punto Informatico had a the lowest cover on Open Source migrations as total number of articles.

On the other hand, a qualitative insight of the articles in Punto Informatico, shows that in most

cases the authors sympathize with Open Source. This appears to be less common in the other periodicals.

The Register shows less sympathy towards Open Source and several articles attack Open Source not from a technical nor from a commercial point of view. Extreme examples border on triviality:

- “The iBuzz does come complete with an attachment for the chaps which will enable Linux programmers to go 'hands-free' while poring over a Carrie Anne Moss tribute site, Matrix” (http://www.theregister.co.uk/2005/11/23/ipod_accessory/)
- “The 'open source community' - whatever that means – smells and looks worse than any other software community on the planet” (http://www.theregister.co.uk/2005/11/29/otto_fly_open)
- “The only thing worse than flying is open source code” and “Finns are socialists. Linus Torvalds is Finnish.” (http://www.theregister.co.uk/2005/10/25/otto_clothes/)

C-Net News.com does not feature similar articles, but many articles about Open Source have alongside banners for the Microsoft's “Get The Facts about Linux” campaign...

Conclusions

The core analysis of this paper is based on almost 5,00 articles, and for some issues we referred to the full data set that we gathered, which includes more than 13,000 articles.

Our experience working on the COSPA project in the United Kingdom and the failure in finding test locations for migrations to Open Source lead us to consider whether the press in the United Kingdom (and in the USA) was biased against Open Source and/or whether there was a lack in coverage.

However, according to our research, Open Source has been covered in the United Kingdom and in the USA as well as in Italy. The Register and C-Net News.com had not a worse cover of Open Source events, compared to Punto Informatico. Indeed, Punto Informatico was the one that “missed” more news on Open Source migrations.

For what regards the content, Punto Informatico appeared to sympathize with Open Source issues – while The Register had several hostile articles and C-Net News.com paired articles on Open Source with Microsoft advertisement.

Possible future directions for our research include expansion in the diversity of sources analysed (more countries) and/or expansion on the number of articles analysed. We foresee a number of difficulties in effectively categorising a high number of articles, and a increase in the number of

sources is likely to imply consequences in the type of analysis on the corpora.

Future work might focus more on the topics that have been covered in the web-based periodicals. This would allow the use of NLP techniques to automatise part of the work – including the categorising of the articles.

More work is needed to investigate the role of other types of media to seek for factors that might have influenced negatively the acceptance of Open Source. At the same time, we foresee to continue our research looking at different causes, to try to be able to define more clearly which are the factors that influence more, in the real world, the adoption of Open Source.

References