Towards a Reference Model on How to Utilise Open Standards in Open Source Projects: Experiences Based on Drupal

Jonas Gamalielsson, Björn Lundell, Alexander Grahn, Stefan Andersson, Jonas Feist, Tomas Gustavsson, Henrik Strindberg

HAL Id: hal-01467575

https://inria.hal.science/hal-01467575

Submitted on 14 Feb 2017

Towards a reference model on how to utilise Open Standards in Open Source projects: experiences based on Drupal

Jonas Gamalielsson[1], Björn Lundell[1], Alexander Grahn[1], Stefan Andersson[4], Jonas Feist[4], Tomas Gustavsson[3] and Henrik Strindberg[2]

[1]University of Skövde, Skövde, Sweden,
`{jonas.gamalielsson,bjorn.lundell,alexander.grahn}@his.se`
[2]Findwise AB, Göteborg, Sweden,
`henrik.strindberg@findwise.com`
[3]PrimeKey Solutions AB, Solna, Sweden,
`tomas@primekey.se`
[4]RedBridge AB, Kista, Sweden,
`{stefan.andersson,jfeist}@redbridge.se`

Abstract. It is known that standards implemented in Open Source software (OSS) can promote a competitive market, reduce the risk for lock-in and improve interoperability, whilst there is limited knowledge concerning the relationship between standards and their implementations in OSS. In this paper we report from an ongoing case study conducted in the context of the ORIOS (Open Source software Reference Implementations of Open Standards) project in which influences between OSS communities and software standard communities are investigated. The study focuses on the Drupal project and three of its implemented standards (RDFa, CMIS, and OpenID).

1   Introduction

Many organisations are currently restricted in their choice of software because of restrictions imposed by existing systems. There is a lack of interoperability and a risk of different types of lock-in. The use of Open Standards and OSS implementations of standards can reduce the risk of lock-in, improve interoperability and stimulate innovation (Lundell et al., 2012; Friedrich, 2011). Further, it is widely acknowledged that there are challenges in implementing Open Standards (FRAND, 2012) and that standardisation has significant impact in the IT market and is subject to review within the digital agenda in the EU (Europe Economics, 2012). Open Standards, especially when implemented in OSS, have the potential to address challenges such as promoting a healthy and competitive market, reducing the risk for organisations of being technologically locked-in, and creating a basis for interoperability, and offering a basis for long-term access and reuse of digital assets (Lundell et al., 2012). Open standards are especially important for small companies, something which is acknowledged in national IT policies (Gov.uk, 2012).

In this paper we consider influences between OSS communities and software standard communities. In so doing, we report on an investigation of how communities involved in reporting and handling issues related to the implementation of a standard influence and are influenced by communities involved in the development and maintenance of the standard. We focus on OSS communities for the Drupal implementations of the three software standards RDFa, OpenID and CMIS, and associated software standard communities.

2   Towards a reference model for Open Standards and OSS

OSS implementations of software standards have made significant contributions to the establishment of standards (Behlendorf, 2009). Even if a number of standards in the software domain have been adopted and implemented in OSS projects, there is limited knowledge concerning the relationship between standards and their implementations in OSS (FRAND, 2012). Such knowledge is of particular relevance to small companies. For this reason this relationship with associated issues are explored in an ongoing collaborative research project (Lundell et al., 2012). In this research project, we seek to establish a reference model to aid concrete actions for any stakeholder wishing to utilise software standards in an Open Source context. In order to achieve this, we are conducting a number of studies on specific software standards and specific OSS projects.

We draw from an ongoing case study in which we specifically explore the Drupal OSS project and its relationship with three specific standards governed by three different standardisation organisations. By choosing RDFa (a W3C standard), CMIS (an OASIS standard), and OpenID (a foundation governed standard), our study includes investigation of standards provided under different governance models. Further, it was of interest to explore both core and add-on implementations of standards. RDFa and OpenID are both core implementations in Drupal, whereas CMIS is implemented as an add-on. The selected OSS project and the three specific standards represent specific examples of projects which are of interest to the companies involved in the ORIOS (Open Source software Reference Implementations of Open Standards) project. These exemplify core technologies used in the daily business and constitute a relevant set for investigation for all stakeholders involved in the project. Further, the relationship between standards and their implementation is complex (FRAND, 2012), which potentially has significant impact for small companies wishing to use Open Source provided under a copyleft license (Bain, 2012). From this we choose to investigate an OSS project (Drupal) provided under a strong copyleft license.

As part of our approach we first establish a characterisation of the three selected standards and the Drupal project by undertaking an analysis of release history, commits to the source code repository and contributing committers over time. Second, we investigate influences between OSS communities and software standard communities with respect to participation. In so doing, we focus on the three chosen

standards using the issue tracking system of Drupal. Third, for RDFa we investigate influences between OSS communities and software standard communities with respect to common issues. The data for the Drupal project was collected from the Drupal website (Drupal.org, 2012), where all issues for Drupal core, RDFa, CMIS and OpenID were used in the analysis. Issue data for all releases of Drupal version 7 was used (including development versions), from the date of the first issue posting until 30 Nov. 2012. The issue data was collected and thereafter parsed and analysed using custom made scripts. More specifically, the timestamp and contributor ID for all issue postings was recorded. In addition, a search for issues in issue tracker, forums and mailing lists was performed by means of manual inspection of content in order to identify issues common to both the Drupal RDFa community and the W3C RDFa community.

3 Results

RDFa is a standard model for interchange of data on the web by embedding of metadata within web documents (W3.org, 2012), and is governed by W3C since 2008. CMIS (Content Management Interoperability Services) is a standard that defines a layer of abstraction for the control of various document management systems and repositories by the use of web based protocols (Oasis-open.org, 2010) and is governed by OASIS since 2010. The OpenID standard for decentralised authentication provides means for proving that an end user is in control of an identifier (Openid.net, 2007) and it is governed by the OpenID Foundation since 2007. All standards investigated in this paper are licensed under royalty-free conditions[1] which allow implementation in GPL licensed OSS projects.

Drupal is a content management platform written mainly in PHP, which is provided under the GPL Open Source license (Drupal.org, 2012). It can be used to create "broschureware" style web sites as well as web sites involving blogs, forums and other forms of collaborative environments. There are more than 600000 users and developers in the Drupal communities. Further, there have been 146 committers who have contributed a total of 88091 commits over 828174 lines of code (Ohloh.net, 2012) to Drupal core. The first commit to Drupal core was contributed in May 2000, and the most recent commit in Dec. 2012. There have been seven first level Drupal releases in the interval Jan. 2001 through Jan. 2011 (v1-3 in 2001, v4 in 2002, v5 in 2007, v6 in 2008, and v7 in 2011). In fact, there have been 113 releases (evenly distributed in time) in total since v1.0 including second and third level releases. The latest release (v7.17) was made available on 7 Nov. 2012.

Figure 1 (top diagram) shows number of issue postings (including issue creation and commenting) for Drupal core. Number of contributors over time for Drupal core

---

[1] RDF: http://www.w3.org/2012/09/rdfa-wg-charter, CMIS: https://www.oasis-open.org/news /pr/oasis-members-approve-content-management-interoperability-services-cmis-standard, OpenID: http://openid.net/intellectual-property/

is shown in the middle diagram in the same figure. An observation is that issues relevant for version 7 of Drupal are raised more than 9 years before the first stable release in Jan. 2011. Further, it is evident that significantly elevated activity precedes the first stable release of Drupal 7, whereafter the activity drops considerably. It can also be noted that there is an increasing trend in terms of number of monthly contributors until the time of the stable release after which the number of contributors stabilizes. Figure 1 (bottom diagram) illustrates number of issue postings over time for the parts of Drupal implementing the RDFa standard (green), the CMIS standard (red) and the OpenID standard (blue). Overall, it can be noted that the issue activity level is more modest for the implementations of the standards when comparing with all parts of the core. We also note that issue postings for OpenID go back to Jul. 2007, whereas issue activity for RDFa and CMIS started considerably later (Oct. 2009 and Mar. 2010, respectively).
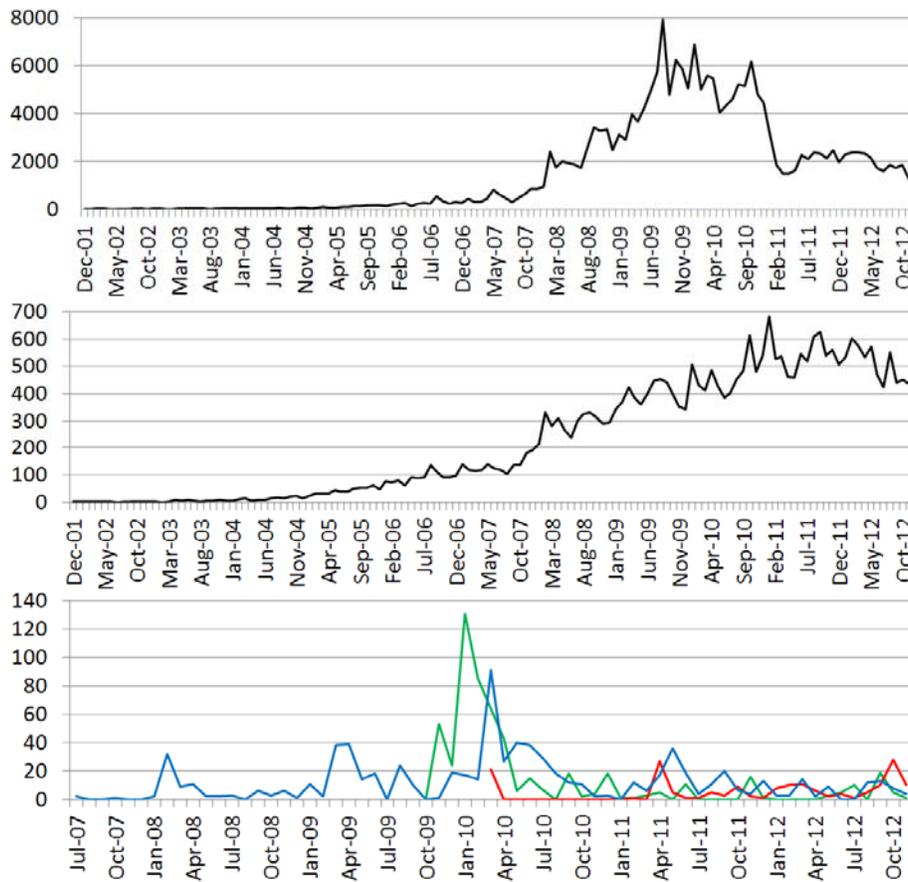

Fig. 1. Number of issue postings (top) and contributors (middle) in Drupal core.
Bottom: Number of issue postings for RDFa (green), CMIS (red), and OpenID (blue)

The Venn diagram in Figure 2 provides an overview of the contributions to the Drupal issue tracker. The total number of contributors and postings for RDFa (green circle), CMIS (red circle) and OpenID (blue circle), is shown. Further, the figure illustrates the number of contributors who have contributed to the issue tracker for the seven possible (and mutually exclusive) combinations of the three standard implementations, and also shows (in brackets) the proportion of all postings for the different standards that the contributors in the different project combinations have contributed. Proportions for RDFa, CMIS and OpenID are coloured in green, red, and blue, respectively. It can be observed that the 17 contributors who have been been active in the issue tracker for both RDFa and OpenID have contributed the majority of the postings for RDFa (81,3%) and 30,1% of the postings for OpenID. Hence, there is significant influence between the RDFa and OpenID communities that contribute to the issue tracker. Further, there is very limited influence between CMIS and the other two standards since 98,2% of all CMIS postings are provided by 45 contributors who are only active in CMIS.
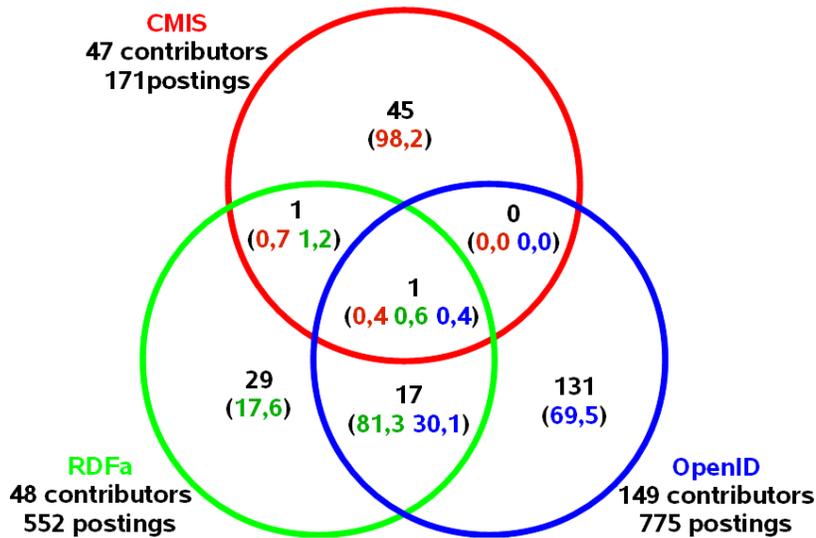

Fig. 2. Contributions to Drupal issue tracker for RDFa (green), CMIS (red) and OpenID (blue)

For the rest of this section we focus on the RDFa standard. When comparing the set of Drupal RDFa issue contributors with the set of participants in the meetings of the RDFa standard working group, there is one individual who is active in both communities. This person has contributed 199 (36%) of all 552 RDFa issue postings and has participated in 33 (31%) of all 104 RDFa working group meetings in the time period Feb. 2010 to Nov. 2012.

Further, RDFa related issues have been identified which appear in both the issue tracker and in the context of the W3C community involved in RDFa standardisation. One example is an issue on how some RDFa content is interpreted in Drupal (and

other RDFa parsers) due to the use of XML literals that was raised in Jan. 2010 on the issue tracker. Deep processing of XML literals would be required to solve the problem, something which was currently not supported in the specification of RDFa. The same issue was subsequently acknowledged and discussed at a W3C meeting in Oct. 2010. Another example is an issue from Jan. 2010 concerning the validity of Drupal generated RDFa code, where one RDF validator reports invalid Drupal code for a specific attribute whereas another validator does not (i.e. two different conformance tests produce different results). This issue was afterwards discussed on the public RDFa mailing list at W3C in Apr. 2010. Further, the issue of support for Drupal specific needs and work practices concerning RDFa has been a recurring topic during several W3C meetings.

4   Discussion and Conclusion

Based on our experiences from participation in standardisation processes in different contexts (including OASIS), we acknowledge the inherent complexity of standardisation. It is not uncommon that development of standards is initially driven by needs stemming from specific usage scenarios involving stakeholders in different organisations. Therefore, standards are often subject to subsequent refinement, adaptation, and generalisation to fit new needs and usage scenarios. Our results illustrate this in the identified issue on support for Drupal specific needs and work practices concerning RDFa. Further, standards are often not detailed enough to enable for unambiguous software implementations. This, in turn, may imply that whoever makes the first implementation has precedence in terms of how to interpret a standard, and in particular when such an implementation becomes widely adopted.

Further, our results show that two conformance tests have been used within the Drupal project, which may impact on any business agreement involving Drupal. This imposes further complexity. For example, if a public sector organisation wishes to procure a system based on Drupal for which there is a need to have interoperability with a specific legacy system, it is essential to address both conformance and interoperability between systems. In such a scenario a conformance test used earlier may now be outdated as the specification of the standard evolves. This would increase the complexity even further since requirements on a specific standard are expressed through requirements for interoperability with its legacy system.

In scenarios when there is a lack of interoperability, despite conformance to the standard according to a specific conformance test, there is of course a limit to what a customer can realistically expect from its supplier. Minor issues concerning lack of conformance and interoperability may perhaps be solved in kind through a constructive dialogue between stakeholders involved. However, further development efforts involving additional cost may be necessary in cases when customer expectations significantly exceed what the business agreement advocates.

In addition, for standards which initially address specific usage scenarios, it is often the case that only a subset of the standard is addressed in conformance tests. For some standards, there is also a lack of conformance tests.

In conclusion, our study shows an inherent complexity concerning issues stemming from different stakeholder groups involved in standardisation and software implementation of standards. Further, we observe different kinds of influence both within the Drupal community (between different sub-communities implementing different standards) and between the RDFa community in Drupal and the RDFa standardisation community in W3C. The findings from our analysis of the Drupal project make an important contribution towards a deeper understanding of challenges concerning relationships between OSS software communities and software standard communities.

References

Bain, M.: Scene Setting - Licensing models for standards and for open source. EC Workshop: Implementing FRAND standards in Open Source: mission impossible?, Brussels, Belgium, 22 Nov. (2012)

Behlendorf, B.: How Open Source Can Still Save the World. Keynote Presentation, In 5th IFIP WG 2.13 International Conference on Open Source Systems, OSS 2009, Skövde, Sweden, 5 Jun. (2009)

Drupal.org: Drupal – Open Source CMS. http://drupal.org/, Accessed 12 Dec. 2012 (2012)

Europe Economics: Guidelines for Public Procurement of ICT Goods and Services: SMART 2011/0044, D2 - Overview of Procurement Practices, Final Report (2012)

FRAND: EC Workshop: Implementing FRAND standards in Open Source: mission impossible?. Brussels, Belgium, 22 Nov. (2012)

Friedrich, J.: Making innovation happen: The role of standards and openness in an innovation-friendly ecosystem. In Proc. of the 7th International Conference on Standardization and Innovation in Information Technology (SIIT 2011), pp. 1-8 (2011)

Gov.uk: Open Standards Consultation documents. http://www.cabinetoffice.gov.uk/resource-library/open-standards-consultation-documents, Accessed 12 Dec. 2012 (2012)

Lundell, B., Abdurahmanovic, A., Andersson, S., Bergström, E., Feist, J., Gamalielsson, J., Gustavsson, T., Kahlbom, R. and Papaxanthis, K.: How can Open Standards be effectively implemented in Open Source? Challenges and the ORIOS project. In Hammouda, I. et al. (Eds.) Open Source Systems: Long-Term Sustainability, IFIP Advances in Information and Communication Technology, Vol. 378, Springer, Heidelberg, pp. 383-388 (2012)

Oasis-open.org: The CMIS v1.0 OASIS Standard Specification. http://docs.oasis-open.org/cmis/CMIS/v1.0/os/cmis-spec-v1.0.pdf, Accessed 12 Dec. 2012 (2012)

Ohloh.net: The Drupal (core) Open Source Project on Ohloh. http://www.ohloh.net/p/drupal, Accessed 12 Dec. 2012 (2012)

Openid.net: OpenID Authentication 2.0 – Final. http://openid.net/specs/openid-authentication-2_0.html, Accessed 12 Dec. 2012 (2007)

W3.org: RDFa 1.1 Primer. http://www.w3.org/TR/xhtml-rdfa-primer/, Accessed 12 Dec. 2012 (2012)