Towards open government through open source software for web analytics: The case of Matomo

Jonas Gamalielsson,[1*] Björn Lundell,[2] Simon Butler,[3] Christoffer Brax,[4] Tomas Persson,[5] Anders Mattsson,[6] Tomas Gustavsson,[7] Jonas Feist,[8] Erik Lönroth[9]

[1*] ORCID Nr: 0000-0003-2700-2535
University of Skövde, Skövde, Sweden, jonas.gamalielsson@his.se

[2] ORCID Nr: 0000-0002-2825-135X
University of Skövde, Skövde, Sweden, bjorn.lundell@his.se

[3] ORCID Nr: 0000-0002-6215-3753
University of Skövde, Skövde, Sweden, simon.butler@his.se

[4] ORCID Nr: 0000-0002-2161-164X
Combitech AB, Skövde, Sweden, christoffer.brax@combitech.com

[5] ORCID Nr: 0000-0003-3959-0266
Digitalist Sweden AB, Stockholm, Sweden, tomas.persson@digitalistgroup.com

[6] ORCID Nr: 0000-0001-8986-8751
Husqvarna AB, Huskvarna, Sweden, anders.mattsson@husqvarnagroup.com

[7] ORCID Nr: 0000-0001-7844-5614
PrimeKey Solutions AB, Solna, Sweden, tomas.gustavsson@primekey.com

[8] ORCID Nr: 0000-0002-6745-0853
RedBridge AB, Stockholm, Sweden, jonas.feist@redbridge.se

[9] ORCID Nr: 0000-0002-2206-9979
Scania CV AB, Södertälje, Sweden, erik.lonroth@scania.com

Abstract: Web analytics technologies provide opportunities for organisations to obtain information about users visiting their websites in order to understand and optimise web usage. Use of such technologies often leads to issues related to data privacy and potential lock-in to specific suppliers and proprietary technologies. Use of open source software (OSS) for web analytics can create conditions for avoiding issues related to data privacy and lock-in, and thereby provides opportunities for a long-term sustainable solution for organisations both in the public

and private sectors. This paper characterises use of and engagement with OSS projects for web analytics. Specifically, we contribute a characterisation of use of OSS licensed web analytics technologies in Swedish government authorities, and a characterisation of organisational engagement with the Matomo OSS project for web analytics.

Keywords: Open source software, web analytics, government authorities, Matomo

1. Introduction

Web analytics technologies have become increasingly important for organisations that wish to obtain information about users that visit their websites, to understand and optimise web usage and support decision making (Phippen et al., 2004; Rohloff et al., 2019). Such technologies are deployed in numerous organisations both in the private and public sectors. For public sector organisations in particular, data privacy is an increasing concern at both the level of individuals and organisations, and web analytics technologies can be problematic when addressing privacy concerns (Akkus et al., 2012; Chandler and Wallace, 2016; Persson, 2020; Quintel and Wilson, 2020; Schwab 2019) and legal perspectives with respect to, for example, the General Data Protection Regulation (GDPR) Article 35 (EU, 2016), the CLOUD Act (USC, 2018), and the recent Schrems II court ruling (CJEU, 2020). It may therefore be unlawful in certain jurisdictions, for example in the European context, to use technologies that transfer personal data to a third country. In the Swedish context, the governmental use of web analytics technologies that transfer personal data to third countries has recently received significant attention in the media. The Swedish public service radio aired a news feature on 25 February 2021 revealing that a large number of Swedish government authority websites send personal data in the form of IP adresses to Google when using their proprietary solution for web analytics, despite the fact that website visitors are promised anonymity (SR, 2021). Subsequently, on 5 March 2021, it was reported that more than 100 municipalities, health care regions and government authorities in Sweden had reported themselves to the Swedish Authority for Privacy Protection for transferring personal data to a third country (SVT, 2021a).

Use of web analytics technologies and especially those that are proprietary and hosted by global companies as software, as a service can become problematic from a legal perspective (CJEU, 2020; Lundell et al., 2020; SOU, 2021) and can cause different kinds of lock-in to specific suppliers and proprietary technologies (Lundell et al., 2020). Lock-in due to dependency on specific suppliers of proprietary technologies and services can be avoided through use of open standards and open source software (OSS) (Lundell and Gamalielsson, 2018; Lundell et al., 2020). There are a number of OSS project initiatives that implement web analytics, including Matomo (matomo.org, 2020a), Open Web Analytics (openwebanalytics.com, 2020), and AWStats (awstats.org, 2020). OSS for web analytics is attractive since it empowers organisations to set up and host solutions on their own or through use of a variety of different suppliers. This approach creates conditions for avoiding the transfer of

sensitive personal data to third countries and also creates conditions for avoiding lock-in to specific suppliers and proprietary technologies and provides opportunities for a long-term sustainable solution for web analytics. Organisational engagement can be considered an important aspect in assessing the long-term sustainability and diversity of an OSS project.

The **overarching goal** in this paper is to characterise use of and engagement with OSS projects for web analytics. Specifically, there are **two principle contributions**. In the **first part**, we characterise use of OSS licensed web analytics technologies in Swedish government authorities. In the **second part**, we characterise organisational engagement with the Matomo OSS project for web analytics.

The focus on government authorities in the first part is motivated by the recognition that such organisations constitute an important subset of public sector organisations that citizens need to interact with in many different contexts and where personal data is often involved. Further, the choice of the Swedish context is one example of a country subject to European law, where third country transfers of personal data are problematic (e.g. EU, 2016). The choice of the Matomo OSS project in the second part of the study is motivated by the fact that it is a widely deployed OSS solution for web analytics (e.g. w3techs.com, 2021) and it has been reported to be a viable alternative to proprietary web analytics solutions (Quintel and Wilson, 2020). The characterisation of organisational engagement with Matomo over time, in the second part of the study contributes knowledge about the long-term sustainability of and diversity of contributors to the OSS project. This knowledge can be used to support any decision concerning the adoption of the OSS project Matomo in an organisational context. The health and long-term sustainability of OSS project communities is important to assess prior to adoption and it has been claimed that "If your assessment leaves you feeling that the community isn't right for you, be prepared to consider alternatives, no matter how attractive the code is." (Crowston and Howison, 2006).

2. Background

2.1. Open government through open source software

OSS is software that complies with the Open Source Definition (OSD, 2021). OSS is usually provided under a small set of software licences that the Open Source Initiative (OSI, 2021) has approved. Broadly, OSS licences can be categorised as permissive licences or copyleft licences. Essentially, OSS provided under a permissive licence can be distributed as part of a larger software product under almost any other licence (including a proprietary non-OSS licence) as long as attribution to the original authors is done. Further, a "copyleft license is one that requires, as a condition of distribution (or making available) of software binaries, that the distributor make the corresponding source code available under the same licensing terms." (Meeker, 2020, p. 10)

There are a number of OSS initiatives that have been initiated and led by governments in different countries. This includes initiatives for OSS-based infrastructure to promote sharing amongst institutions in the public sector (Hollman et al., 2013). There have also been a number different efforts in governmental contexts in different countries to adopt and utilise OSS (e.g. Maldonado, 2010). An example is the plan for achieving technological sovereignty by the City of Barcelona, which has a

focus on "open and efficient government that uses technology to transform and digitally innovate the public sector based on open source software and open standards" (BCC, 2017). There has also been similar initiatives in the UK: "The overarching policy and directive stemming from central government in the UK promotes the idea of open source software use, open data, and open standards." (Shaikh, 2016), and in the Netherlands (von Loon and Toshkov, 2015).

Strategies for OSS have been developed both at EU-level and in its member countries. The open source software strategy 2020-2023 for Europe, published by the European Commission "sets out a vision for encouraging and leveraging the transformative, innovative and collaborative power of open source, its principles and development practices" and "promotes the sharing and reuse of software solutions, knowledge and expertise, to deliver better European services that benefit society and lower costs to that society." (EC, 2020a). Strategies for promoting use of OSS have been developed in several countries, including the Netherlands and Sweden. For example, the ministry of Economic Affairs in the Netherlands has published an action plan for the use of open standards and open source software in the public and semi-public sector. The strategy aims to increase "interoperability between and with the different building blocks and forms of service provision of eGovernment by accelerating the use of open standards", to reduce "dependence on suppliers in the use of ICT through faster introduction of open standards and open source software", and to promote "a level playing field in the software market" and "innovation and the economy by forceful stimulation of the use of open source software." (NOC, 2007). The Swedish strategy for development of software, published by the Swedish Agency for Digital Governance, promotes use of open source software in software development and procurement projects in government contexts to achieve high quality software that can be widely utilised in different government and organisational contexts (DIGG, 2019). Further, to facilitate the adoption of OSS in governmental contexts in Europe, the European eGovernment Services has released a guideline on the public procurement of Open Source Software (EES, 2010).

There have been a number of initiatives that promote open government through use of specific OSS projects. One example is the X-Road initiative that emerged in Estonia and is an open source, licensed data exchange layer solution that enables public and private sector organisations, that are potentially located in different countries, to exchange information over the Internet. (Kalja, 2002; Robles et al., 2019). The Ministry of Transport and Communications in Estonia was initially governing the development of X-Road, but in 2017 a new governance regime that involved both the Estonian and Finnish government and the creation of the Nordic Institute for Interoperability Solutions (NIIS) in a joint effort for the further development of the X-Road project. Another example is a knowledge management ecosystem based on the Drupal platform for web content management, with the aim to promote the collaboration between public administrations in Spain (García-Holgado and García-Peñalvo, 2014). A third example is promotion of OSS in the context of governmental agencies in Denmark where OS2, which is "an open source community open to all Danish public agencies" (OSOR, 2017) that includes the majority of all Danish municipalities, is maintaining a portfolio of OSS projects that have been adopted in a number of municpalities in Denmark and beyond. One of the most widely adopted OSS projects in the OS2 portfolio is OS2kitos, which is a "record system for portfolio management of systems, contracts, projects and interfaces." (OS2, 2016).

There are a number of software projects which provide software applications for web analytics as OSS relevant for government authority contexts. Such solutions can be developed and deployed in a number of different ways, including local installation in a public sector organisation (with or without support from commercial companies) and as a SaaS (Software as a Service) solution provided by an external organisation (e.g. a commercial supplier). For example, amongst solutions provided as OSS we find Matomo (matomo.org, 2020a) and AWStats (awstats.org, 2020), both provided under the GPL 3.0 licence, and Open Web Analytics (openwebanalytics.com, 2020), provided under the GPL 2.0 licence.

2.2. Earlier research

There is a lack of research on the topic of OSS for web analytics in the context of government authorities, and there is limited earlier research covering OSS for web analytics in other contexts. Examples include: a feasibility study and comparison of Matomo and Google Analytics in a library context, with a specific emphasis on functionality and privacy (Quintel and Wilson, 2020); a similar study that focused on the migration from Google Analytics to Matomo in a library context (Chandler and Wallace, 2016); research reporting on use and integration of Open Web Analytics and Matomo in the context of an online public access catalog (Mandal, 2019) based on the OSS project Koha (koha-community.org, 2020); a study with a focus on the analysis of user behaviour in the context of an online grammatical information system through the use of Matomo (Ripp and Falke, 2018); and research reporting on the utilisation of web analytics technologies (including Matomo) in the context of learning analytics for large-scale online learning (Rohloff et al., 2019).

There is earlier research related to organisational engagement with OSS, but not specifically OSS for web analytics. One such study explored organisational contributions to source code repositories over time for the OSS modelling tools Topcased and Papyrus (Gamalielsson et al., 2011) and a different study reported results on organisational contributions to mailing lists for the OSS project Nagios through the analysis of email address subdomains (Gamalielsson et al., 2010). Other studies explore organisational collaboration in specific OSS communities through approaches involving social network analysis (Linåker et al., 2016; Teixeira et al., 2015). There are also studies focused on organisational aspects, for example addressing different motivations for firms to participate in OSS projects (Bonaccorsi and Rossi, 2006), how firms engage in OSS projects under different governance models (Schaarschmidt et al., 2015), community building aspects in communities sponsored by organisations (West and O'Mahony, 2008), emerging engagement of professional and commercial organisations in OSS (Fitzgerald, 2006). Further, the question of how different OSS projects may benefit from the engagement of companies was explored in (Capiluppi et al., 2012). Community evolution in terms of inflow and retention in OSS projects with commercial engagement, through use of a case study approach was reported in (Zhou et al., 2016). A different study reports from a field study of corporate engagement with OSS communities (Germonprez et al., 2017). Similarly, the rationale behind corporate engagement in OSS development was explored in (Kendall et al., 2016). A different study focused on coopetition (co-operation on top of competition) in the context of software firms in OSS ecosystems (Nguyen et al., 2017). Further, a model has been proposed to aid understanding of relations between organisations and OSS communities (Link and Jeske, 2017). The use of data analytics techniques for understanding how companies interact with OSS communities is addressed

in (Gonzalez-Barahona et al., 2013). Another study investigates how different contributor groups associated with public and corporate interests interact in an OSS development project (the Linux kernel) from an innovation perspective (Homscheid et al., 2015). Corporate engagement in OSS projects from a political and economic perspective is addressed in (Birkinbine, 2015). The discovery of community patterns accounting for organisational engagement in OSS projects was described in (Tamburri et al., 2019). A different study empirically investigates objectives and complexities that organisations need to consider and balance between, when deciding on which software to share as OSS, when to share it, and whether to create a new OSS project or contribute to an existing OSS project community (Linåker and Regnell, 2020). Another study focuses on work practices used in public communication channels by company contributors to engage with and contribute to eight community OSS projects (Butler et al., 2021). Earlier research has also explored how and why companies engage in corporate open sourcing (Harutyunyan et al., 2020; Harutyunyan and Riehle, 2021).

There is a lack of studies that focus on actual organisational engagement in OSS projects in concrete cases. Hence, our study contributes to filling that knowledge gap and thereby contributes the first study on the use of and organisational engagement with an OSS project (Matomo), for web analytics.

3. Research approach

The sample used in the first part of the study is the complete set of 253 Swedish government administrative authorities (swe. Statliga förvaltningsmyndigheter) and government enterprises (swe. Statliga verk) according to lists provided by Statistics Sweden (SCB, 2020). The selected governmental authorities and enterprises are hereafter referred to as the "GAs". The web domains for the official websites of each of the GAs were mined for use of web analytics technologies through use of the *builtwith* tool (builtwith.com, 2020) on 5 November 2020. The output HTML file for each GA was manually inspected and used technologies were recorded and semiautomatically processed through use of custom made Perl scripts in order to derive usage statistics. A manual web search[1] was conducted to find statements about the GAs that use OSS for web analytics.

For the second part of the study, 61 of the 68 GitHub source repositories under the Matomo Analytics umbrella (GitHub, 2020), that given the repository description were deemed to have a clear relation to Matomo (and formerly Piwik), were cloned and git logs were extracted on 13 November 2020. Apart from Matomo core, there are repositories for plugins, software development kits, etc. The git logs for the 61 repositories were processed through use of custom made Perl scripts in order to compile a total picture of author contributions to the source code. The contributor identifiers were manually inspected to eliminate redundancy in terms of used aliases. Identified contributors were subsequently looked up on LinkedIn (a business oriented online service) for their organisational affiliation for the years they contributed to Matomo. One contributor can have several affiliations for each year. If contributors were not present on LinkedIn, a manual web search was conducted in an attempt to find information about organisational affiliation. A total picture of contributions for

---

[1]   through use of the Google search engine

authors' associated organisations, organisation types and headquarter location for organisations was compiled through use of custom made Perl scripts.

4. On use of OSS licensed web analytics technologies

In this section we characterise use of OSS licensed web analytics technologies by looking at, in total, 253 selected GAs. Four GAs did not have a website. No use of client side web analytics technology was identified for 37 GAs (representing 15%). In total, 36 different technologies were identified through the analysis of all 253 GAs, and Table 1 shows country headquarter location (HQ) according to ISO 3166-1 alpha-2, number of GAs that use a certain technology (N) and the proportion of all 253 GAs that this number represents (%). Each GA may use multiple web analytics technologies.

Table 1. Use of web analytics technologies amongst Swedish government authorities

Technology	HQ	N	%
Google Analytics	US	199	79
Google Global Site Tag	US	136	54
Hotjar	AU	44	17
Google Analytics IP Anonymization	US	20	8
Matomo	NZ	16	6
Vizzit	SE	15	6
Facebook Pixel	US	14	6
Facebook Conversion Tracking	US	13	5
Facebook Signal	US	12	5
New Relic	US	10	4
Facebook Domain Insights	US	8	3
Microsoft Application Insights	US	7	3
APSIS Lead	SE	6	2
Google Optimize 360	US	6	2
Microsoft Visual Studio Tracking	US	6	2
LinkedIn Insights	US	4	2
Matomo Tag Manager	NZ	3	1
Moat	US	3	1
Pingdom RUM	US	3	1
Usabilla	NL	3	1
WebTrends	US	3	1
Adobe Analytics	US	2	1
Adobe Marketing Cloud	US	2	1
CrazyEgg	US	2	1
eMarketeer	SE	2	1
Google Conversion Linker	US	2	1
Omniture SiteCatalyst	US	2	1
Quantcast Measure	US	2	1
Bing Universal Event Tracking	US	1	<1
Cloudflare Rocket Loader	US	1	<1
Easyresearch	NO	1	<1
Google Content Experiments	US	1	<1

Google DoubleClick Floodlight	US	1	<1
Monsido	DK	1	<1
ProspectEye	SE	1	<1
Report URI	UK	1	<1

It is evident that use of Google Analytics (199 GA domains, representing 79%) and related technologies is clearly dominating amongst the GAs. Further, 34 of the 36 (representing 94%) identified technologies for web analytics are proprietary. There is also a clear US dominance amongst the technologies used, where 25 (representing 69%) are provided by a US-based supplier. The only identified technology that is provided as open source software is Matomo (and the associated Matomo Tag Manager). The rest of the technologies are proprietary and require services provided by suppliers. The 16 (representing 6%) specific GAs that have been identified to use Matomo on the domain for the web page are, (with Swedish name in brackets): The Swedish Public Employment service (Arbetsförmedlingen), the Taxation of Research Workers Board (Forskarskattenämnden), the Swedish Chemicals Agency (Kemikalieinspektionen), the War Insurance Board (Krigsförsäkringsnämnden), the National Library of Sweden (Kungliga Biblioteket), the Swedish mapping, cadastral and land registration authority (Lantmäteriet), the Local safety committee at Ringhals (Lokala säkerhetsnämnden vid Ringhals), the Board for Board Representation Issues (Nämnden för styrelserepresentationsfrågor), the Oil Crisis Board (Oljekrisnämnden), The Swedish Government Offices (Regeringskansliet), the Swedish Tax Agency (Skatteverket), the National Agency for Special Needs Education and Schools (Specialpedagogiska skolmyndigheten), the Swedish Board for Workers' Inventions (Statens nämnd för arbetstagares uppfinningar), the Swedish FAO Committee (Svenska FAO-kommittén), the Swedish Defence Research Agency (Totalförsvarets forskningsinstitut), and the Swedish Election Authority (Valmyndigheten). Matomo was identified as the only web analytics technology for six of the domains for these 16 GAs.

The implementation of Matomo in the IT-intensive Swedish Tax Agency (one of the six Matomo-only GAs) was especially noticed in 2018 where the importance of keeping personal data in Sweden was stressed (Djurstad, 2018). Since the implementation of Matomo the Swedish Tax Agency and the Swedish Public Employment service have advertised several jobs where experience from Matomo is highlighted. Further, for two of the Matomo-only GAs and for two GAs which also use other web technologies, there is information about use of Matomo on webpages about technical details and cookies.

5. On organisational engagement with the Matomo OSS project

The Piwik OSS project was initiated in 2007 and was renamed to Matomo in 2018 (matomo.org, 2018). Matomo is provided under the GPL 3.0 licence and the source code is mainly in PHP. There have been regular stable releases since the first alpha release in 2007 (matomo.org, 2020b). Figure 1 shows the number of (mostly third level) stable releases for Matomo over time where beta releases have been excluded. On average there have been 10 releases each year.

In the remainder of this section we characterise organisational engagement with the Matomo OSS project using information about affiliations for authors contributing to the source code. A total of

47948 author contributions were made by approximately 800 contributors to the 61 selected source code repositories under the Matomo Analytics umbrella from the first contribution in 2007 until the date of the data collection (13 November 2020). There are 802 author identifiers in total and we chose to further focus on the 83 actual (and manually curated) authors (rather than identifiers) who each have provided at least 10 contributions for the remainder of this section. Those 83 authors provide 42122 (88%) of all author contributions. The top three, top 10, and top 20 authors provide 53%, 78%, and 83%, respectively.

Figure 1: Number of stable releases over time for Matomo
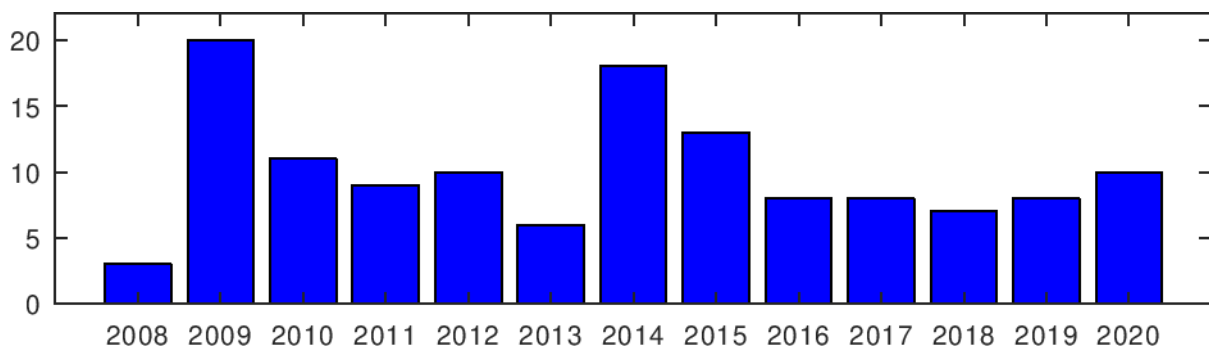
For the 83 authors, in total 96 different organisational affiliations have been identified. These affiliations account for a total of 82942 associated author contributions (since one author can have several affiliations this number is larger than the 42122 contributions at author level). A special case of affiliation in this study is "Self-employed", which occurs when an author has been identified to be self-employed working on a freelance basis where the actual name of the company could not be identified. Figure 2 shows the number of author contributions over time for the top 20 organisations (where the organisation type and country code for HQ are shown in parentheses). The colour of a cell in Figure 2 indicates number of author contributions. The top row in the figure shows the organisation that has contributed the largest number of author contributions over time. Organisation types that have been identified for author associated affiliations are defined as follows in this study: Small and Medium-sized Enterprise (SME, an enterprise with 10–200 employees), Micro Enterprise (MiE, an enterprise with 1–9 employees), Larger Enterprise (LE, an enterprise with more than 200 employees), Non-profit Organisation (NPO), Higher Education Institution (HEI), and Research Institute (RI). The 20 affiliations in Figure 2 account for 97% of the 82942 associated author contributions. We note that the company Matomo is (unsurprisingly) dominating and has been engaged continuously since the start and dominates clearly by accounting for 33% of the 82942 associated author contributions. The top three (i.e. Matomo, Piwik Pro, and Innocraft) and top 10 affiliations account for 66% and 91%, respectively. Further, it can be observed that Innocraft, Mayflower GmbH, and TDG have been engaged continuously at least since 2011. It should also be mentioned that the business offerings of the top three contributing organisations are closely connected to and based on the Matomo OSS project. Other noteworthy organisations that are major players in other contexts related to IT and the Web, but where associated contributions to Matomo are limited, include Amazon Web Services, Intel Corporation, Google, and Apple.

Figure 2: Number of author contributions over time for the top 20 organisations
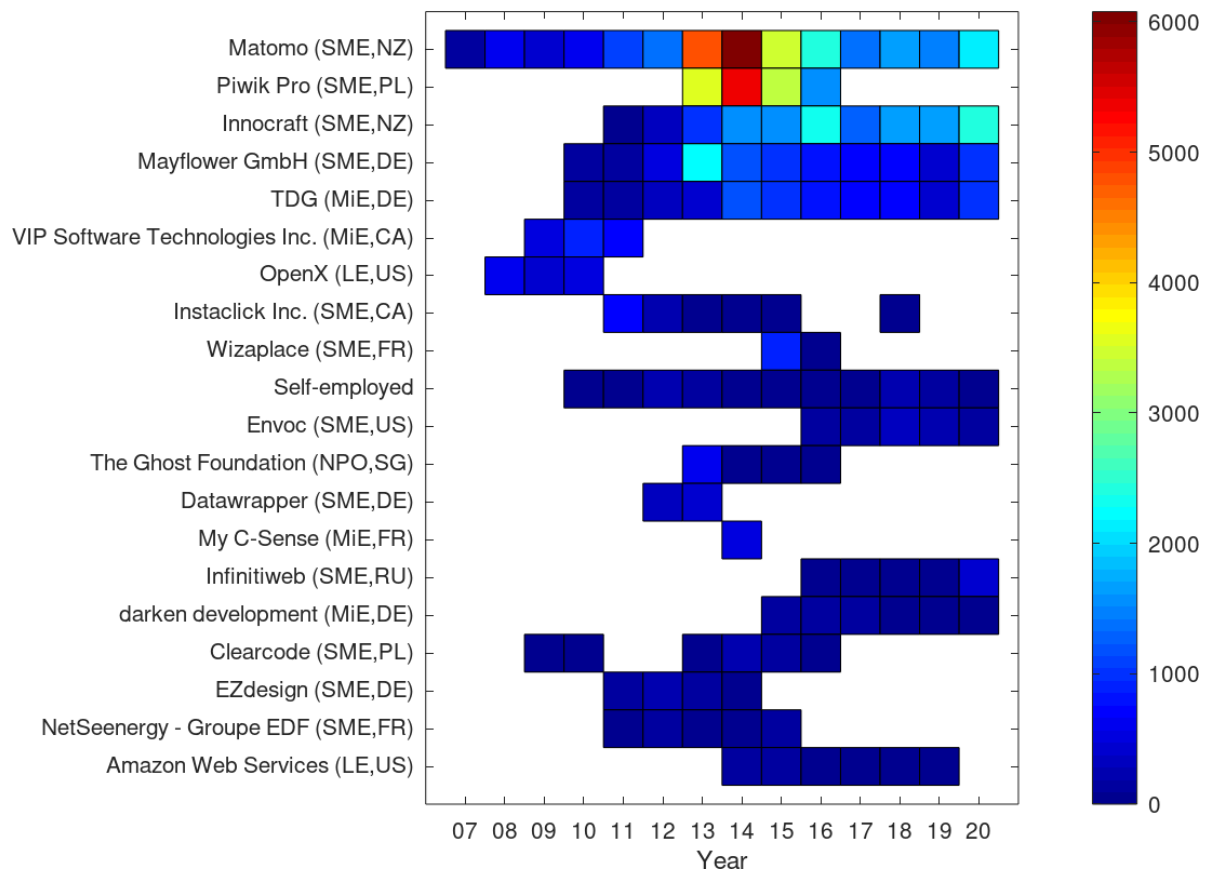

Table 2 shows a summary of the contributions in Figure 2 over different time periods. It can be observed that the number of contributing organisations (amongst the top 20 for the whole interval) that have made author contributions during the first nine years (2007-2015) is higher than during the last five years (2016-2020). In connection to this, we note that two additional organisations (Envoc and Infinitiweb) have contributed during the last five years compared to the first nine years. In terms of author contributions, 64% of, in total, 80102 contributions (for the top 20 organisations) were made during the first nine years and 36% during the last five years. The number of author contributions on average per year is approximately the same for the two time periods, which indicates that the OSS project is as active in later years as in the earlier phase of the project.

Table 2. Summary of contributions in Figure 2 over different time periods

	2007-2020	**2007-2015**	**2016-2020**
Number of organisations (amongst top 20)	20	18	14
Number of author contributions	80102	51533	28569

Figure 3 illustrates the number of author contributions over time for the six different identified organisation types. The top row in the figure shows the organisation type that has contributed the largest number of author contributions over time. It can be observed that authors representing SMEs have been engaged continuously since the start and are clearly dominant, accounting for 83% of the 82942 associated author contributions. The other affiliation types MiE, LE, NPO, HEI, and RI account in total for 13%, 3%, 1%, <1%, and <1% of the associated author contributions, respectively. Hence,

authors representing SMEs and MiEs (that have been engaged continuously since 2009) provide an overwhelming majority of the author contributions (96%). Table 3 shows a summary of the contributions in Figure 3 over different time periods.

Figure 3: Number of author contributions over time for different organisation types
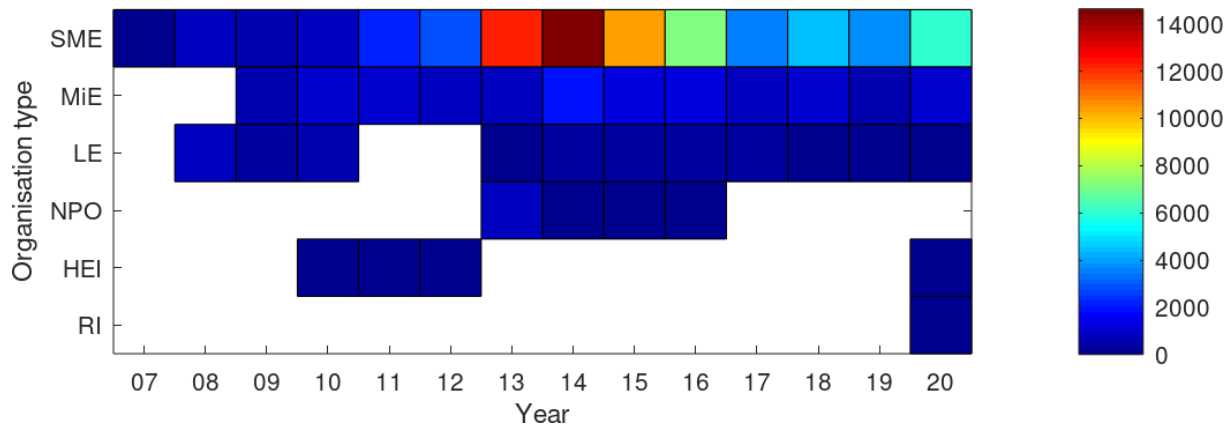

Table 3 shows a summary of the contributions in Figure 3 over different time periods. It can be observed that the number of organisation types for which there are author contributions during the first nine years is lower than during the last five years. In connection to this, we note that one additional organisation type (Research Institute) has emerged during the last five years compared to the first nine years, indicating an increase in organisation type diversity in later years. In terms of author contributions, 64% of the total 82942 contributions were made during the first nine years and 36% during the last five years. Hence, like for Table 2, this indicates that the OSS project is as active in later years as in the earlier phase of the project.

Table 3. Summary of contributions in Figure 3 over different time periods

	2007-2020	2007-2015	2016-2020
Number of organisation types	6	5	6
Number of author contributions	82942	53333	29609

Figure 4 shows the number of author contributions over time for the (total of 23) different identified organisation HQ locations. The top row in the figure shows the HQ location for which the largest number of author contributions, over time, have been contributed. It can be noted that authors representing New Zealand based organisations (which are also SMEs) have been engaged continuously since the start and dominate clearly by accounting for 49% of the 82942 associated author contributions. The top three (i.e. New Zealand, Germany and Poland) and top 10 HQ locations account for 86% and 99% of the associated author contributions, respectively. The organisation HQ named "?" in table 3 is for the self-employed contributors, for which the location of their business is unknown. Analysis at the level of continents shows that Oceania, Europe, North America, and Asia account for 50%, 41%, 7%, and 1% of the author contributions, respectively. Organisations where the location (country and continent) of HQ is unknown, account for 1% of the author contributions.

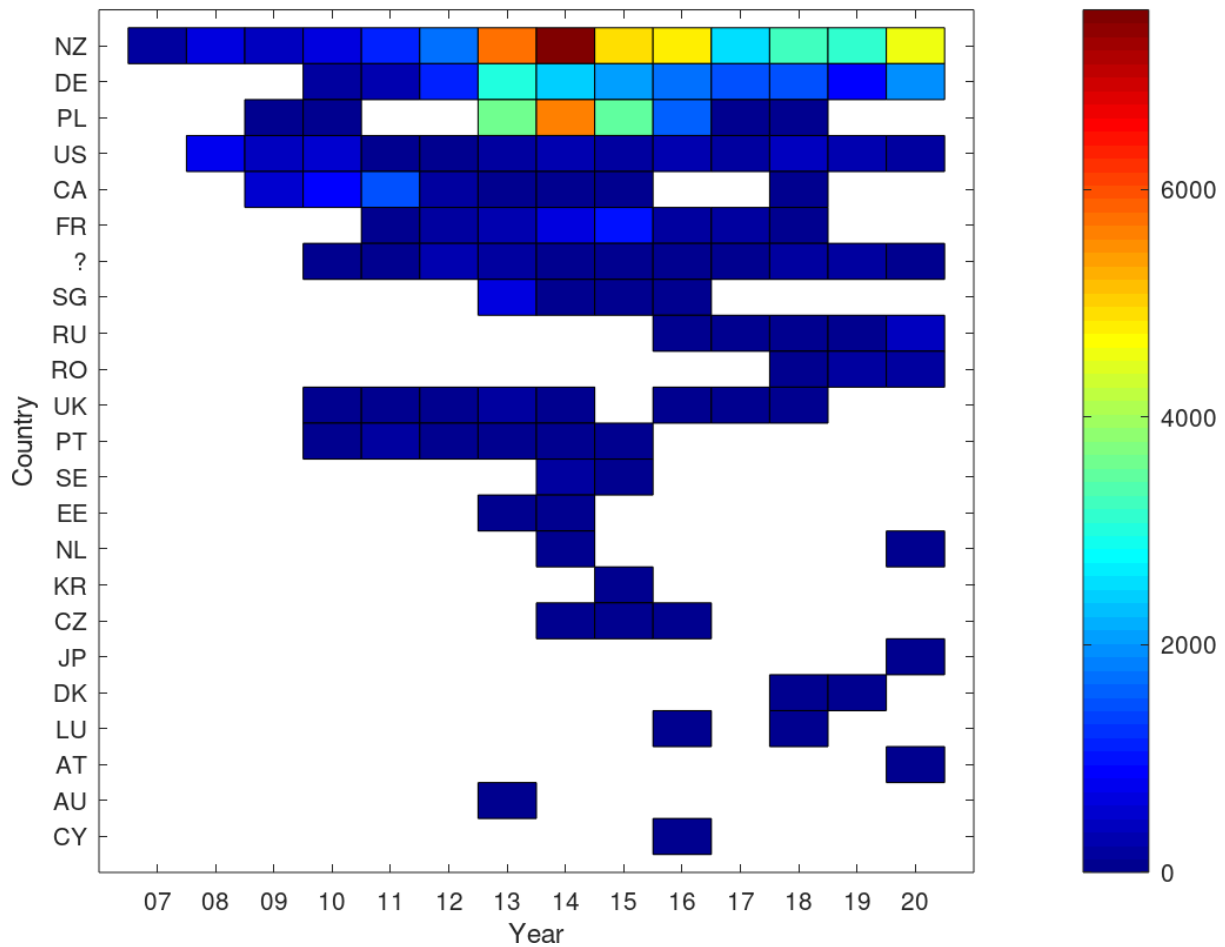*Figure 4: Number of author contributions over time for different organisation HQ locations*

Table 4 shows a summary of the contributions in Figure 4 over different time periods. It can be observed that the number of organisation HQ locations, for which there are author contributions during the first nine years, is lower than during the last five years. In connection to this, we note that seven additional organisation HQ locations (Russia, Romania, Japan, Denmark, Luxembourg, Austria, and Cyprus) have emerged during the last five years compared to the first nine years. Hence, this suggests that the geographical diversity of HQ locations for contributing organisations, has increased during later years.

Table 4. Summary of contributions in Figure 4 over different time periods

	2007-2020	2007-2015	2016-2020
Number of HQ locations	23	16	18
Number of author contributions	82942	53333	29609

6. Discussion and conclusions

The study contributes a characterisation of the use of OSS licensed web analytics technologies in Swedish government authorities and organisational engagement with the Matomo OSS project for web analytics. In summary, **findings from the first part** of the study show that an overwhelming majority of the selected Swedish governmental agencies use proprietary technologies hosted by specific global suppliers, where Google Analytics clearly dominates and where OSS (Matomo) is used by a limited number of governmental agencies. **Findings from the second part** show that the Matomo OSS project for web analytics can be considered to be long-term sustainable since it has had regular releases since its inception, and four of the top five associated organisations in terms of author contributions have been engaged continually at least since 2011. The business offerings of the top three contributing organisations are closely connected to the Matomo OSS project but there is also a fair amount of contributions associated with a number of external organisations. The project is dominated by small and medium sized enterprises and micro enterprises with headquarters in Oceania (predominantly New Zealand) and Europe (predominantly Germany and Poland).

Findings from the study show that the Matomo OSS project for web analytics has been chosen as a viable and long-term sustainable alternative to proprietary software solutions by a number of Swedish government agencies. Through deployment of an OSS solution, such as Matomo, government authorities are empowered to make decisions concerning how solutions should be deployed and thereby promote "technological sovereignty" (EC, 2020b). Based on previous research (Lundell et al., 2021) which conceptualises development and deployment of software (see Figure 5), it has been clarified that government authorities have different options for how to realise the deployment of software. In Figure 5, examples of software for web analytics (both provided as OSS and under proprietary conditions) are shown. From the perspective of a procuring government authority, the development of software which is provided as OSS may be internally deployed (lower left quadrant of Figure 5) or be externally deployed (lower right quadrant of Figure 5). It has been observed in the results in the first part of the study, that several government authorities use internal deployment of OSS (Matomo), but the licensing model for OSS also allows an authority to choose external deployment through the use of an external supplier. Similarly, the development of software provided as closed (proprietary) software may be internally or externally deployed. In the case of external deployment, the government authority may choose between deployment in or outside the EU. Choice of deployment may imply specific challenges related to the GDPR, privacy, and transfer of data to third countries (CJEU, 2020; Shurson, 2020; Wagner, 2018). For example, the Swedish Tax Agency and the Swedish Enforcement Agency have decided not to deploy and use Microsoft Teams due to such challenges (Skatteverket/Kronofogden, 2021a, 2021b). Further, challenges related to deployment of software and privacy have been scrutinised and discussed by an investigative journalist in Swedish national broadcasting, where the focus was on third country transfer of patient data (SVT, 2021b). To address these challenges, deployment of OSS solutions within the EU by European organisations may create conditions for avoiding transfer of sensitive personal data to third countries, and also create conditions for avoidance of lock-in to specific suppliers and proprietary technologies (Lundell et al., 2020).

The different options for how an OSS project for web analytics can be deployed are particularly important for any organisation concerned with privacy issues, such as many public sector organisations in the EU that are bound by GDPR. For example, OSS from the Matomo OSS project can be deployed and used in a single organisation (perhaps with support from external consultants) even though it may also be used as a service (e.g. through provision of Matomo as a service that is provided by an external company, where both organisations are bound by the same legislation). Hence, we conjecture that the variety of different deployment options may be especially appealing when legal conditions related to privacy may impose new challenges.

Figure 5. Conceptualisation of principal ways for development and deployment of software

In light of findings in the first part, we recommend that governmental agencies consider the reasons for implementing web analytics. Unlike companies, public sector organisations are not competitive and the goal should not be to maximise traffic to their web pages. A relevant goal may be to collect data in order to optimise the experience for citizens so that the best service can be provided. Another aspect is something which can be considered as competence lock-in, where employees from professional web consultancy firms, hired by public sector organisations, often are trained to implement web analytics through use of widely deployed (and often proprietary) technologies like Google Analytics. In such a scenario it becomes important that public sector organisations carefully specify requirements for their solutions for web analytics during public procurement in order to provide opportunities for a long-term sustainable solution that avoids lock-in and privacy related issues.

We acknowledge that only source code repositories have been considered in the study. Other data sources (including issue trackers for Matomo on GitHub and the Matomo forum) could contribute to further detail organisational engagement with Matomo.

A topic for future research is to analyse how government authorities' use of web technologies changes over time, with a view to also understand reasons for change. Another relevant topic for future research is to investigate perceptions and views concerning organisational implementation of open source software for web analytics in different contexts in order to identify and understand challenges and opportunities related to such implementations.

References

Akkus, I. E., & Chen, R., & Hardt, M., & Francis, P., & Gehrke, J. (2012). Non-tracking web analytics. In Proceedings of the 2012 ACM conference on Computer and communications security (CCS '12). Association for Computing Machinery, New York, NY, USA, 687–698.
DOI: https://doi.org/10.1145/2382196.2382268

awstats.org (2020). AWStats official web site (2020). https://www.awstats.org/

BCC (2017). Barcelona City Council technological sovereignty guide, Barcelona City Council, September, https://www.barcelona.cat/digitalstandards/en/tech-sovereignty/0.1/_attachments/barcelona_tech_sovereignty_0.1.en.pdf

Birkinbine, B. J. (2015). Conflict in the Commons: Towards a Political Economy of Corporate Involvement in Free and Open Source Software. The Political Economy of Communication 2(2), ISSN 2357-1705.

Bonaccorsi, A., & Rossi, C. (2006). Comparing motivations of individual programmers and firms to take part in the open source movement: from community to business. Knowledge, Technology & Policy, 18(4), 40–64.

Builtwith (2020). Find out what websites are Built With. https://builtwith.com/

Butler, S., & Gamalielsson, J., & Lundell, B., & Brax, C., & Sjöberg, J., & Mattsson, A., & Gustavsson, T., & Feist, J., & Lönroth, E. (2021). On Company Contributions to Community Open Source Software Projects. IEEE Transactions on Software Engineering, 47 (7), 1381-1401.
DOI: https://doi.org/10.1109/TSE.2019.2919305

Capiluppi, A., & Stol, K.-J., & Boldyreff, C. (2012). Exploring the Role of Commercial Stakeholders in Open Source Software Evolution. In: Hammouda, I., Lundell, B., Mikkonen, T., Scacchi, W. (Eds.) Open Source Systems: Long-Term Sustainability. OSS 2012. IFIP Advances in Information and Communication Technology, vol. 378. Springer, Berlin, Heidelberg.

Chandler, A., & Wallace, M. (2016). Using Piwik Instead of Google Analytics at the Cornell University Library. The Serials Librarian, 71(3-4), 173-179.

CJEU (2020). Judgment of the Court in Case C-311/18. Court of Justice of the European Union, 16 July, http://curia.europa.eu/juris/document/document.jsf?text=&docid=228677&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=12312155

Crowston, K. and Howison, J. (2006). Assessing the Health of Open Source Communities. IEEE Computer, 39(5), 89-91.

DIGG (2019). Policy for utveckling av programvara, Myndigheten för Digital förvaltning (Agency for Digital Government), Dnr. 2019-136, 8 May. https://www.digg.se/4a3a3e/globalassets/dokument/om-oss/nyheter/policy-for-utveckling-av-programvara.pdf

Djurstad, J. A. (2018). Svensk Openstack-pionjär petar Google hos Skatteverket, Computer Sweden, 16 March, https://computersweden.idg.se/2.2683/1.699368/elastx-openstack-skatteverket

EC (2020a). Open Source Software Strategy 2020-2023, European Commission, C(2020) 7149 final, 21 October, https://ec.europa.eu/info/departments/informatics/open-source-software-strategy_en#opensourcesoftwarestrategy

EC (2020b). Shaping Europe's Digital Future, Communication from the Commission to the Euro-pean Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, European Commission, Communication, COM(2020) 67 final, 19 February.

EES (2010). Guideline on public procurement of OpenSource Software, IDABC - European eGovernment Services, European Communities, March, https://joinup.ec.europa.eu/sites/default/files/document/2011-12/OSS-procurement-guideline%20-final.pdf

EU (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679 &from=EN

Fitzgerald, B. (2006). The transformation of open source software. MIS Quarterly, 30(3), 587-598.

Gamalielsson, J., & Lundell, B., & Lings, B. (2010). The Nagios community: an extended quantitative analysis. In: Ågerfalk, P. et al. (eds.) Open source software: New Horizons. Berlin, Springer, 85–96.

Gamalielsson, J., & Lundell, B., & Mattsson, A. (2011). Open source software for model driven development: a case study. In: Hissam, S. (Ed) Open source systems: grounding research, IFIP advances in information and communication technology, vol. 365, Boston, Springer, p. 348–67. ISBN: 978-3-642-24417-9.

García-Holgado, A., & García-Peñalvo, F. J. (2014). Knowledge management ecosystem based on drupal platform for promoting the collaboration between public administrations. In Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM '14), Association for Computing Machinery, New York, NY, USA, 619–624. DOI: https://doi.org/10.1145/2669711.2669964

Germonprez, M., & Kendall, J. E., & Kendall, K. E., & Mathiassen, L., & Young, B., & Warner, B. (2017). A Theory of Responsive Design: A Field Study of Corporate Engagement with Open Source Communities. Information Systems Research, 28(1), 64-83.

GitHub (2020). Matomo Analytics, https://github.com/matomo-org

Gonzalez-Barahona, J. M., & Izquierdo-Cortazar, D., & Maffulli, S., & Robles, G. (2013). Understanding How Companies Interact with Free Software Communities. IEEE Software, 30(5), 38-45.

Harutyunyan, N., & Riehle, D. (2021). Getting Started with Corporate Open Source Governance: A Case Study Evaluation of Industry Best Practices. In Proceedings of the 54th Hawaii International Conference on System Sciences, HICSS '21, Manoa, US.

Harutyunyan, N., & Riehle, D., & Sathya, G. (2020). Industry Best Practices for Corporate Open Sourcing. In Proceedings of the 53rd Hawaii International Conference on System Sciences, HICSS '20, Wailea, US.

Hollmann, V., & Lee, H., & Zo, H., & Ciganek, A. P. (2013). Examining success factors of open source software repositories: the case of OSOR.eu portal. Intl Journal of Business Information Systems, 14(1), 1-20.

Homscheid, D., & Kunegis, J., & Schaarschmidt, M. (2015). Private-Collective Innovation and Open Source Software: Longitudinal Insights from Linux Kernel Development. In: Janssen, M. et al. (Eds.) Open and Big Data Management and Innovation. I3E 2015. Lecture Notes in Computer Science, vol. 9373, Springer, Cham.

Kalja, A. (2002). The X-Road project - A project to modernize Estonia's national databases. Baltic IT&T review, 24, 47-48.

Kendall, J. E., & Kendall, K. E., & Germonprez, M. (2016). Game theory and open source contribution: Rationale behind corporate participation in open source software development. Journal of Organizational Computing and Electronic Commerce, 26(4), 323-343.

koha-community.org (2020). Koha Library Software, https://koha-community.org/

Link, G. J. P., & Jeske, D. (2017). Understanding Organization and Open Source Community Relations through the Attraction-Selection-Attrition Model. In Proceedings of the 13th International Symposium on Open Collaboration (OpenSym '17). ACM, New York, NY, USA, Article 17, 8 pages.

Linåker, J., & Regnell, B. (2020). What to share, when, and where: balancing the objectives and complexities of open source software contributions. Empirical Software Engineering, 25, 3799-3840.

Linåker, J., & Rempel, P., & Regnell, B., & Mäder, P. (2016). How firms adapt and interact in open source ecosystems: analyzing stakeholder influence and collaboration patterns. In: Daneva, P., Pastor, O. (Eds.) Requirements engineering: foundation for software quality, lecture notes in computer science, vol. 9619. New York: Springer-Verlag New York, Inc., 63–81. ISBN: 978-3-319-30281-2

von Loon, A., & Toshkov, D. (2015). Adopting open source software in public administration: The importance of boundary spanners and political commitment. Government Information Quarterly, 32(2), 207-215. DOI: https://doi.org/10.1016/j.giq.2015.01.004

Lundell, B., & Gamalielsson, J. (2018). Sustainable digitalisation through different dimensions of openness: how can lock-in, interoperability, and long-term maintenance of IT systems be addressed?, In Proceedings of the 14th International Symposium on Open Collaboration (OpenSym '18), ACM, New York, ISBN: 978-1-4503-5936-8, Article 3, 10p.

Lundell, B., & Gamalielsson, J., & Katz, A. (2020). Addressing lock-in effects in the public sector: how can organisations deploy a SaaS solution while maintaining control of their digital assets?, In: Virkar, S. et al. (Eds.) CEUR Workshop Proceedings: EGOV-CeDEM-ePart 2020, vol. 2797, pp. 289-296. ISSN: 1613-0073

Lundell B., & Gamalielsson, J., & Butler, S., & Brax, C., & Persson, T., & Mattsson, A., & Gustavsson, T., & Feist, J., & Öberg, J. (2021). Enabling OSS usage through procurement projects: How can lock-in effects be avoided?, In Taibi, D. et al. (Eds.), The 13th International Conference on Open Source Systems (OSS 2021), IFIP Advances in Information and Communication Technology, Vol. 624, Springer, Cham, pp. 16-27. DOI: https://doi.org/10.1007/978-3-030-75251-4_2

Maldonado, E. (2010). The process of introducing FLOSS in the public administration: the case of Venezuela. Journal Association for Inf. Systems, 11(11), 756.

Mandal, S. (2019). Site Metrics Study of Koha OPAC through Open Web Analytics and Piwik Tools. Library Philosophy and Practice (e-journal), 2835. ISSN: 1522-0222

matomo.org (2018). Piwik is now Matomo – Announcement. 9 January, https://matomo.org/blog/2018/01/piwik-is-now-matomo/

matomo.org (2020a). Matomo, https://matomo.org/

matomo.org (2020b). Matomo Changelog, https://matomo.org/changelog/

Meeker, H. (2020). Open (Source) for Business: A Practical Guide to Open Source Software Licensing. Third edition, Kindle Direct Publishing Platform, Seattle, ISBN-13: 979-8618201773

Nguyen, D. A., & Cruzes, D. S., & Hanssen, G. K., & Snarby, T., & Abrahamsson, P. (2017). Coopetition of Software Firms in Open Source Software Ecosystems. In: Ojala, A. et al. (Eds.) Software Business. ICSOB 2017. Lecture Notes in Business Information Processing, vol. 304. Springer, Cham.

NOC (2007). The Netherlands in Open Connection: An action plan for the use of Open Standards and Open Source Software in the public and semi-public sector, Ministry of Economic Affairs, The Hague, November. https://www.ossdirectory.com/knowhow/2007_TheNetherlandsInOpenConnection.pdf

openwebanalytics.com (2020). Open Web Analytics, http://www.openwebanalytics.com/

OS2 (2016). OS2 public digitalisation network, OS2, 4 july, https://os2.eu/sites/default/files/media/presentation_os2_in_english_compressed.pdf

OSD (2021). The Open Source Definition. Open Source Initiative, https://opensource.org/osd

OSI (2021). Open Source Initiative, http://opensource.org/

OSOR (2017). Danish OS2 community for open source is professionalising. Open Source Observatory (OSOR), 2 May, https://joinup.ec.europa.eu/collection/open-source-observatory-osor/news/danish-os2-community-open

Persson, T. (2020). Svenska myndigheter måste sluta att läcka data om svenska medborgares surfvanor, https://www.digitalist.se/blogg/svenska-myndigheter-maste-sluta-att-lacka-data-om-svenska-medborgares-surfvanor

Phippen, A., & Sheppard, L., & Furnell, S. (2004). A practical evaluation of Web analytics. Internet Research, 14(4), 284-293.

Quintel, D. F., & Wilson, R. (2020). Analytics and Privacy - Using Matomo in EBSCO's Discovery Service. Information Technology and Libraries, 39(3).

Ripp, S., & Falke, S. (2018). Analyzing User Behavior with Matomo in the Online Information System Grammis. In Proceedings of the XVIII EURALEX International Congress Lexicography in Global Contexts, 17-21 July, Ljubljana, pp. 87-100.

Robles, G., & Gamalielsson, J. & Lundell, B. (2019). Setting Up Government 3.0 Solutions Based on Open Source Software: The Case of X-Road, In: Lindgren, I. et al., (Eds.) Electronic Government, Springer, Cham, pp. 69-81. DOI: https://doi.org/10.1007/978-3-030-27325-5_6

Rohloff, T., & Oldag, S., & Renz, J., & Meinel, C. (2019). Utilizing Web Analytics in the Context of Learning Analytics for Large-Scale Online Learning. In Proceedings of the 2019 IEEE Global Engineering Education Conference (EDUCON), Dubai, pp. 296-305.

SCB (2020). Myndighetsregistret, http://www.myndighetsregistret.scb.se

Schaarschmidt, M., & Walsh, G., & von Kortzfleisch, H. (2015). How do firms influence open source software communities? A framework and empirical analysis of different governance modes. Information and Organization, 25(2), 99–114.

Schwab, K.. (2019). It's time to ditch Google Analytics, https://www.fastcompany.com/90300072/its-time-to-ditch-google-analytics

Shaikh, M. (2016). Negotiating open source software adoption in the UK public sector. Government Information Quarterly, 33(1), 115-132. DOI: https://doi.org/10.1016/j.giq.2015.11.001

Shurson, J. (2020). Data protection and law enforcement access to digital evidence: resolving the reciprocal conflicts between EU and US law. International Journal of Law and Information Technology, 28.

Skatteverket/Kronofogden (2021a). Beslut: Promemoria om ersättning av Skype i Skatteverkets och Kronofogdens verksamhet, Memorandum regarding the replacement of Skype in the Swedish Tax Agency's and Swedish Enforcement Authority's operations, Decision, 3 May, The Swedish Tax Agency, Dnr. 8-958696, The Swedish Enforcement Authority, Dnr. KFM 10419–2021.

Skatteverket/Kronofogden (2021b). Decision: Memorandum regarding the replacement of Skype in the Swedish Tax Agency's and Swedish Enforcement Authority's operations, 3 May, The Swedish Tax Agency, Reference no.: 8-958696, The Swedish Enforcement Authority, Reference no.: KFM 10419–2021.

SOU (2021). Säker och kostnadseffektiv it-drift : rättsliga förutsättningar för utkontraktering - delbetänkande. Statens offentliga utredningar, SOU 2021:1, January, http://www.sou.gov.se/wp-content/uploads/2021/01/SOU_2021_1_webb.pdf

SR (2021). Myndigheter lovar anonymitet – men skickar ip-adresser till Google, Sveriges Radio, 25 February, https://sverigesradio.se/artikel/myndigheter-lovar-anonymitet-men-skickar-ip-adresser-till-google

SVT (2021a). Myndigheter och kommuner anmäler sig själva. SVT nyheter, 5 March, https://www.svt.se/nyheter/snabbkollen/myndigheter-och-kommuner-anmaler-sig-sjalva

SVT (2021b). Miljardavtalet i vården kan stoppas – patientuppgifter på amerikanskägda servrar. Uppdrag granskning, 26 February, https://www.svt.se/nyheter/granskning/ug/miljardavtalet-i-varden-kan-stoppas-patientuppgifter-pa-amerikanskagda-servrar

Tamburri, D. A., & Palomba, F., & Serebrenik, A., & Zaidman, A. (2019). Discovering community patterns in open-source: a systematic approach and its evaluation. Empirical Software Engineering, 24(3), 1369–1417.

Teixeira, J., & Robles, G., & González-Barahona, J. M. (2015). Lessons learned from applying social network analysis on an industrial Free/Libre/Open Source Software ecosystem. Journal of Internet Services and Applications, 6(14), 1-37.

USC (2018). S.2383 - CLOUD Act, United States Congress, 6 February, https://www.congress.gov/bill/115th-congress/senate-bill/2383/text

w3techs.com (2021). Usage statistics and market share of Matomo for websites, https://w3techs.com/technologies/details/ta-matomo

Wagner, J. (2018). The transfer of personal data to third countries under the GDPR: when does a recipient country provide an adequate level of protection?. International Data Privacy Law, 8.

West, J., & O'Mahony, S. (2008). The role of participation architecture in growing sponsored open source communities. Industry and Innovation, 15(2), 145–168.

Zhou, M., & Mockus, A., & Ma, X., & Zhang, L., & Mei, H. (2016). Inflow and Retention in OSS Communities with Commercial Involvement: A Case Study of Three Hybrid Projects. ACM Transactions on Software Engineering and Methodology, 25(2), Article 13.

About the Authors

Jonas Gamalielsson

Jonas Gamalielsson is a researcher and senior lecturer at the University of Skövde, Skövde, Sweden, where he is a member of the Software Systems Research Group. His research interests include open source software and open standards for addressing challenges related to lock-in, interoperability, and longevity of systems. Gamalielsson received his Ph.D. from Heriot Watt University in 2009.

Björn Lundell

Björn Lundell is a professor at the University of Skövde, Skövde, Sweden, where he leads the Software Systems Research Group. His research interests include fundamental sociotechnical challenges concerning software systems, focusing on different aspects of lock-in, interoperability, and longevity of systems. Lundell received his Ph.D. from the University of Exeter in 2001.

Simon Butler

Simon Butler is a researcher in the Software Systems Research Group, University of Skövde, Skövde, Sweden. His research interests include software engineering, open source software, program comprehension, software development tools and practices, and software maintenance. Butler received his Ph.D. in computing from the Open University in 2016.

Christoffer Brax

Christoffer Brax is a consultant at Combitech AB, Skövde, Sweden, working in systems engineering, requirements management, systems design and architecture, and IT security. Brax received his Ph.D. from Örebro University in 2011.

Tomas Persson

Tomas Persson has a long background in the IT and Web industry and started around 1995. Tomas is part of Digitalist's thought leadership and a frequent speaker at Digitalist's events. In 2005 Tomas started to use Drupal and joined the Drupal community in 2008. Tomas was among the first in Sweden to sell Open Source web solutions to larger companies and has been an ambassador for Drupal in particular since then. Tomas has been a speaker at Webbdagarna, several DrupalCamps and DrupalCons.

Anders Mattsson

Anders Mattsson is the lead architect for Internet of Things systems at Husqvarna, Huskvarna, Sweden. His research interests include strengthening software engineering practices in organizations, software architecture, and model-driven development of embedded real-time systems. Mattsson received his Ph.D. in software engineering from the University of Limerick in 2012.

Tomas Gustavsson

Tomas Gustavsson is a cofounder and the chief technology officer of PrimeKey Solutions, Solna, Sweden. Gustavsson received his M.Sc. in electrical and computer engineering from KTH Royal Institute of Technology

in 1994. He is the founder of the open source enterprise public key infrastructure project EJBCA, and contributes to numerous open source projects and is an Open Source Sweden board member.

Jonas Feist

Jonas Feist is a cofounder of RedBridge, Stockholm, Sweden. His research interests include business models related to OSS, and in cloud computing, particularly the development and application of container technologies. Feist received his M.Sc. in computer science from the Institute of Technology, Linköping University, in 1988.

Erik Lönroth

Erik Lönroth leads supercomputing technical initiatives at Scania CV, Södertälje, Sweden. His research interests include the development of supercomputer environments for industry, open source software governance, and high-performance computing. Lönroth received his M.Sc. in computer science from Linköping University.