Abstract—Free/Libre/Open Source Software (FLOSS) communities have produced a large amount of valuable software that is directly or indirectly used daily by any person with access to a computer. The field of Software Engineering studies processes, mechanisms, tools, and frameworks for the development of software artifacts. Historically, however, most of Software Engineering research and education does not benefit from the large and rich source of data and experimental testbeds offered by FLOSS projects and their hundreds of millions of lines of working code. In this paper, we discuss how Software Engineering research and education can greatly benefit from the wealth of information available in the FLOSS ecosystem. We then evaluate how FLOSS has been used, up to now, by papers published in the Brazilian Symposium on Software Engineering. Finally, we present an agenda for the future, proposing concrete ways to exploit the synergies between research and education in Software Engineering and FLOSS projects.

I. INTRODUCTION

Software Engineering (SE) research aims at advancing our understanding of the process of software development and its outcomes. As in other fields of research, knowledge must be built on top of verifiable evidence. However, it is not uncommon to see Software Engineering research based on non-public data and non-reproducible methodologies, or even based on no data at all.

Moreover, the education of future software engineers (and possibly SE researchers) should be based on the study of existing real-world software, very much like the education of future artists include extensive review of past artistic achievements and styles or the education of future architects is based on the study of classical buildings and construction styles.

Free/Libre/Open Source Software (FLOSS) offers SE researchers the opportunity of basing their research on abundant and publicly available data, freely accessible data analysis, and software development tools. These are basic building blocks for reproducible and extensible science. Accordingly, SE apprentices can apply the concepts that they learn in the classroom in real-world situations made available by FLOSS projects.

The international SE community has already recognized the potential of FLOSS in both research and education [1], [2], as well as opportunities for future research [3]. In this paper, we discuss important opportunities brought by FLOSS to the SE community, and analyze whether or not the Brazilian SE community is taking advantage of them. We performed a study over papers published in the Brazilian Symposium on Software Engineering (SBES) which, from the top of its 25 years, is the main Brazilian venue of publication for SE researchers. We analyzed research papers published between 1999 and 2010, tools papers published between 2001 and 2010 and SE education papers published between 2008 and 2010.

In this analysis, we classified SE research according to the way FLOSS was used in research reported by SBES papers and compared these categories with the discussed opportunities. Based on that, we propose an agenda to better take advantage of FLOSS in SE research and education.

The remainder of this paper is organized as follows. Section II presents definitions and basic concepts about the FLOSS ecosystem, and Section III discusses opportunities that FLOSS brings for SE researchers and educators. In Section IV, we describe our study and its results. In Section V, we discuss the results in light of the identified opportunities. We conclude the paper in Section VI by proposing an agenda for fostering the synergies between FLOSS and SE in Brazil.

II. ABOUT FLOSS

“FLOSS” is a broad acronym that refers to software that promotes user freedom, does not discriminate users or uses and, at the same time, is based on a collaborative, efficient, and open development process. FLOSS basically allows users to use, study, modify, and redistribute the software with virtually no restrictions except, sometimes, preventing users to impose restrictions on other users. In order to make this possible, access to the source code is a necessary condition.

Typically, such software exists by means of development projects that are centered around some publicly-accessible source code, where both developers and users interact with each other mainly over the Internet. The code is necessarily licensed under terms that comply with either the Free Software
definition ² or the Open Source Definition ³.

Most software recognized as “Free Software” or “Open Source” actually complies with both definitions, even if their specific terms are quite different from one another; The use of a specific terminology reflects the desire to emphasize either the ethical or the technical aspects of the same phenomena. The acronym FLOSS (Free/Libre/Open-Source Software), which we use in this text, is gaining strong acceptance, as it encompasses both major terminologies.

A. Why FLOSS?

FLOSS offers benefits from social, ethical, technical, and economic points of view [4], [5]. From a social point of view, FLOSS may be one of the many elements in easing access to technological resources to the population at large. Also, the use and support of FLOSS in e-government may offer more transparency and ease access to services and data to citizens. Finally, source code access and sharing eases learning and promotes participation in the development of software that might otherwise impose restrictions beyond those embodied in the Law.

From an ethical point of view, if a resource may be easily shared, preventing this should only be pursued under very specific circumstances and with very good motivation. The growth of the FLOSS model in the last decades undermines arguments in favor of such restrictions for software, as its sharing does not prevent other benefits to society.

From the technical and economical points of view, FLOSS offers a development model that may result in high quality code that gets adapted quickly to different situations with lower direct costs. It also leverages the development effort of different companies or individuals into common products, reducing the duplication of effort.

B. FLOSS Development Process

From a Software Engineering point of view, the most interesting aspect of FLOSS is its development process. A FLOSS project starts when an individual developer, or an organization, decides to make a software product publicly available on the Internet so that it can be freely used, modified and redistributed.

After an initial version is released and advertised in the appropriate communication channels, early adopters will start using the product. Some of these early adopters may be software developers themselves, who will be able to review the source code and propose changes that fix defects for their specific environments or add features for their own use cases. These changes may be sent back to the original developer(s) in the form of patches. The project leader(s) will review the proposed changes and apply them (or not) to the official version, so that when a new release is made, end users will have access to these new functionalities or bug fixes.

In the course of time, each release of the product may have more features and be more portable than its predecessor due to contributions from outside developers. The most frequent contributors may gain trust from the initial developer(s), and receive direct write access to the official source code of the project, becoming able to make changes directly to the official version of the product.

This development process is often driven by means of a version control system (VCS). While the repository is often publicly available for read access, write access is restricted to a limited group of developers. Other developers will need their patches to be reviewed by a developer with the needed privileges in order to get their contributions into the project’s official repository.

Besides the version control system, most FLOSS projects use a fairly standardized environment for collaboration between their developers: an issue tracker (or bug tracker) for organizing pending, ongoing and future activities; one or more mailing lists for coordination and communication between the team; and usually a web-based content management system (such as a wiki) for community-driven documentation writing.

The process of developers joining FLOSS projects and gaining responsibility may range from very informal to very formal, depending on the project. Small projects normally have informal procedures for that, e.g. one existing developer just offers an account in the version control system to the new contributor. Larger projects, on the other hand, may have more “bureaucratic” processes for accepting a new developer with privileged access to the project’s resources. This “bureaucracy” may involve filling forms with an application, signing terms of copyright transference and other procedures. [6], [7]

The following characteristics make FLOSS projects different enough from “conventional” software projects, making them interesting objects of study:

- **Source code availability.** Source code of FLOSS projects is always available on the Internet, since most projects have a publicly-accessible version control repository.
- **User/developer symbiosis.** In most FLOSS projects the developers are also users of the software, and they also provide requirements. Maybe because of that, several free software projects do not have explicit requirement documents, and the development proceeds at a pace adequate for the developers to satisfy their own needs.
- **Non-contractual work.** A large amount of work in FLOSS projects is done in a non-contractual fashion. This does not imply that the developers are necessarily volunteers, but only that there is no central management with control over all of the developers’ activities.
- **Work is self-assigned.** The absence of a central central management with control over the contributors’ activities promotes work self-assignment: volunteer developers tend to work on the parts of the project that most appeal to them, and employed developers will work on the parts that are of most interest to their employers.
• **Geographical Distribution.** In most FLOSS projects the developers are spread among several different locations in the world. In projects with high geographical dispersion, communication is mostly performed through electronic means.

The Software Engineering literature tends to portray FLOSS as a homogeneous phenomenon[8], but most of these characteristics do not apply to all FLOSS projects, and some of them may be manifested in different ways across projects.

III. FLOSS OPPORTUNITIES FOR ACADEMIC SOFTWARE ENGINEERING

FLOSS brings multiple benefits for Software Engineering in an academic scenario, providing benefits both for education and research.

A. Using FLOSS for Software Engineering Education

Both FLOSS development and development processes taught in traditional SE textbooks address the challenges of multi-person construction and maintenance of multi-version programs [9], but with development processes, work practices, and project forms that differ significantly and in interesting ways [3]. This raises the question of whether FLOSS development could be useful for educational purposes in SE and computer science undergraduate courses, by exposing students and teachers to different aspects of professional practice. Such exposure may encompass a wide range of issues and activities that includes problem solving, management, ethical and legal concerns, written and oral communication, working as part of a team, and remaining current in a rapidly changing discipline [10].

This issue has been exploited in different international forums [11], [12], [13], [14], [15], [16]. The perspective of using FLOSS for educational purposes has brought a series of interesting opportunities and challenges into the Software Engineering Education agenda, such as:

- Involving students and faculty in large-scale FLOSS projects to provide them with real-world experience and an understanding of the issues found in large, complex software projects [16], [13];
- Reinvigorating the CS/SE curriculum and faculty members [13]. FLOSS promotes project and problem-based learning in which developers work on projects that interest them; by working on interesting and meaningful projects, they can learn concepts, skills, and aptitudes [17];
- Exploiting successful open source projects, in which software is highly modular and APIs are well documented, for teaching principles and good practices [15];
- Providing quantitative data from real, open and freely available source code on which to perform analysis and base decisions [10].

Evidently, the use of FLOSS also brings difficulties, such as:

- Because of its intrinsic characteristics, FLOSS development is not typically representative of many traditional development methodologies. Therefore, its usefulness in SE education is restricted to only part of SE topics or to methodologies related to common FLOSS development processes;
- Academic experience on FLOSS may be seen as less important for the preparation of professionals in environments where proprietary software is the norm, especially if it would result in less academic experience with other widespread technologies;
- Involving teachers and students in FLOSS-related studies brings specific difficulties, for example finding projects that gave appropriate size and complexity (not big enough to be unbearable and not small enough to be trivial), allow exploring the required course topics and have enough documentation for starters.

B. Using data from FLOSS projects in Software Engineering Research

While data from private software development is often scarce and may impose high acquisition costs, FLOSS project data is plenty and readily available for free to anyone with an Internet connection. Private data usually cannot be shared for replicating research results due to issues such as confidentiality requirements and non-disclosure agreements, whereas FLOSS data is already public anyway [1].

FLOSS projects can serve as an abundant source of data for Software Engineering research. Hundreds of thousands of projects make available their source code, which can be analyzed either manually or by automated tools. For example, in a recent work, we studied 6773 projects using automated tools and showed how the source code metrics affected the attractiveness of the projects, in terms of number of downloads, page hits, and project members [18]. In another study, our group analyzed the impact of changes in the licenses of 756 FLOSS projects over 44 months to show how these changes affect the attractiveness of these projects [19]. In yet another study we analyzed the full history of 7 FLOSS web server projects, containing 13553 software changes, in order to investigate developers’ level of participation in the project as a predictor for variations in structural complexity [20].

Most FLOSS projects have all their communication media and development resources – version control system, bug tracking system, mailing lists, documentation websites – publicly available on the Internet, and through them researchers can monitor, track and analyze the activity of the development group working in a certain software artifact. Such data is being used extensively in contemporary Software Engineering research, and systems like bicho [21], Analizo [22] and others can help automate the acquisition and analysis of this kind of data from publicly accessible repositories on the Web.

Again, it should be noted that FLOSS does not cover all possible data relevant for SE; some methodologies and topics cannot be assessed by means of FLOSS data analysis. FLOSS-based research, however, is not limited to the previously existing software engineering research sub-areas: studying FLOSS in itself is also a promising research area. FLOSS development
groups have been producing large and complex products using methodologies and processes that are sometimes substantially different from “conventional” software development. Understanding how FLOSS works and its differences and similarities with regard to “conventional” software engineering practices can help researchers and practitioners to enhance the software development practice in general.

Research on FLOSS development has brought a series of interesting topics into the Software Engineering research agenda, such as the social structure of development communities [23], [24], [6], [7], communication and work flow patterns in FLOSS projects, [25], [26], developer evolution and participation [27], [20], and attractiveness of FLOSS projects [19], [18].

C. FLOSS software products in research activities

This opportunity applies not only to Software Engineering research, but also to Computer Science research in general, as well as to any research in which there is software development. Both researchers and society can benefit from the interplay between FLOSS and research activity.

Releasing the source of software produced in research activities meets the basic scientific principle of reproducibility. The scientific method demands that, when describing a scientific achievement in a paper, researchers must provide all the details required for an independent group of scientists to reproduce the experiment to validate it, checking whether or not the same results are achieved. When software plays an important role in a scientific study, either as subject under study or as a data analysis tool, making it publicly available under a FLOSS license will facilitate the results to be expanded and built upon.

Another benefit of releasing research-originated software as FLOSS is fostering technology transfer from research institutions to society: being freely licensed allows research prototypes to be enhanced into production-class products without the researchers having to be (necessarily) involved indefinitely.

There are several examples of successful FLOSS products that were initially developed in the context of research activities: the PostgreSQL relational database management system and the BSD family of operating systems were created at the University of California, Berkeley; the Xen virtualization technology was initially developed at the University of Cambridge; the Ginga platform for digital television was developed in the University of California, Berkeley; the Xen virtualization technology was initially developed at the University of Cambridge; the PostgreSQL relational database management system and the BSD family of operating systems were created at the University of California, Berkeley. Researchers may use existing FLOSS systems such as the Linux operating system, the GCC compiler, the Eclipse framework for IDEs, or the OpenNebula Cloud Computing middleware and test their ideas by changing just a small portion of them. This way researchers can concentrate on the exact point they want to explore and advance science and technology more rapidly with more reliable results.

IV. FLOSS in SBES

To evaluate where the Brazilian Software Engineering research community stands with regard to FLOSS, we carried out a study on the papers published in the Brazilian Symposium on Software Engineering (SBES). We have examined: (i) main track papers between 1999 and 2010, (ii) tools session papers between 2001 and 2010, and (iii) Software Engineering Education Forum (Fórum de Educação em Engenharia de Software – FEES) papers from 2008, its first edition, to 2010. This section describes this study and its findings.

A. Data acquisition

We obtained the full text of SBES main track, tools session, and FEES papers from the following sources:

- Main track papers from SBES 2005 were obtained contacting the program chair and from SBES 2010 were obtained from the IEEE Xplore Digital Library6.
- Tools papers from SBES 2001 and 2002 were also obtained from BDBComp. Because SBES tools session papers from most years are not available on-line, we got 2004, 2005, 2006, 2007, 2008, 2009, and 2010 papers contacting the program chair of each edition.
- FEES papers from 2008 and 2009 were obtained at the FEES homepage hosted by PUC-Rio7, and 2010 papers were obtained from CBSoft 2010 organizers.

### TABLE I

<table>
<thead>
<tr>
<th>Year</th>
<th>Analyzed</th>
<th>Mention FLOSS</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>26</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>2001</td>
<td>20</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>2002</td>
<td>18</td>
<td>2</td>
<td>0.11</td>
</tr>
<tr>
<td>2004</td>
<td>17</td>
<td>1</td>
<td>0.06</td>
</tr>
<tr>
<td>2005</td>
<td>21</td>
<td>2</td>
<td>0.10</td>
</tr>
<tr>
<td>2006</td>
<td>19</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>2007</td>
<td>23</td>
<td>2</td>
<td>0.09</td>
</tr>
<tr>
<td>2008</td>
<td>19</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>2009</td>
<td>24</td>
<td>6</td>
<td>0.25</td>
</tr>
<tr>
<td>2010</td>
<td>19</td>
<td>11</td>
<td>0.58</td>
</tr>
</tbody>
</table>

5http://www.lbd.dcc.ufmg.br/bdbcomp
6http://ieeexplore.ieee.org
7http://fees.inf.puc-rio.br/
The number of main track and tools papers analyzed by year are presented in Tables I and II, respectively. FEES papers were 14 from 2008, 8 for 2009 and 8 for 2010. Despite not being able to obtain every single paper since 1999, the papers we obtained gave us a reasonable approximation of what happened during the period. We analyzed a total amount of 378 papers: 206 from the main track, 142 from the tools session, and 30 from the Software Engineering Education Forum.

The PDF files were processed by a script that extracted their text, and matched their contents against the following terms: “software(s) livre(s)”, “ferramenta(s) livre(s)”, “ferramenta(s) aberta(s)”, “software(s) aberto(s)”, “código aberto”, “repositorio(s) de software”, “free software”, “open source”, “open software”, “libre software”, “software repository”, “OSS”, “FLOSS”, “FOSS”, and “OSSD”. To validate this identification process of the papers related to FLOSS, we checked all of them manually by reading the abstract and the sections that include one of the terms above.

Tables I and II present the raw results for the SBES main track and tools session, respectively. The tables presents the number of papers analyzed in each year, how many of them mentioned FLOSS, and the relative frequency of papers mentioning FLOSS (i.e. the ratio between the number of papers analyzing FLOSS and the total of papers). We can see that the number of papers mentioning FLOSS in both the main track and the tools session have been increasing, getting to around 50% in their 2010 editions.

The source code for the data extraction and analysis scripts we used is available together with the sources for this paper itself\(^8\), and is licensed under the GNU General Public License version 3 or any later version.

### B. Analysis of research papers mentioning FLOSS

We investigated in which way the FLOSS terms were used in these 25 papers. For that, we classified them in 6 different categories, according to our interpretation of each paper. Each paper was classified in exactly one category, as shown in Table III. The first category comprised three papers that did not mention FLOSS explicitly in their text but just referenced bibliography whose title mentioned FLOSS [28], [29], [30]. The other categories are detailed below.

\(^8\)https://gitorious.org/flosspapers/cbsoft2011-sbes25

### 1) Example or Comparison: Papers that are not focused on FLOSS but used FLOSS as an example or comparison. 

Yamaguti and Price [31] presented a web-based reflective architecture. They argued that any FLOSS CASE tool is an example of “open object”, a concept defined in their paper. Maciel et al. [32] explained that there are several FLOSS and proprietary MDD/MDA tools with different features that can use their proposed approach for model-driven process.

### 2) Running FLOSS tools: Some of the analyzed papers described the development or use of FLOSS tools to carry out their research. 

Santos and Schiel [34] presented a study about the interoperability of data between different domains on the Internet, applying the Resource Description Framework (RDF). They used PostgreSQL (a FLOSS database) in one of their case studies. Mendonça et al. [35] among their contributions presented a refactoring framework, called RefaX. They used Jikes, which is an adaptation from IBM’s FLOSS Java compiler to implement a Java version of Refax, called RefaX4Java. Maciel and Yano [36] proposed a workflow language based Web Services and a FLOSS run-time environment to this language.

### 3) FLOSS tool as target: Several of the selected papers used FLOSS tools as a target to test another tool, framework, or environment. 

Souza and Mendonça [40] defined a reverse engineering environment to extract and detect “implies scenarios” from traces. They tested the proposed environment against MyPetStore, a customized version of PetStore, which is a FLOSS tool developed by Sun Microsystems to illustrate programming resources available in the J2EE platform.

### 4) Data from FLOSS: Part of the papers used data from FLOSS projects to test their tools and theories. 

Colaço Jr.
et al. [44] analyzed 18 large Brazilian industrial projects trying to use association rules obtained from mining their source code repositories to predict pairs of files that would need to be changed together in the future. They compared their results with other results from a previous similar study with eight FLOSS projects (Eclipse, GCC, Gimp, JBOSS, JEdit, KOffice, Postgres, and Python). They found that, in that industrial environment, the prediction power was even higher than with the FLOSS projects. Ferreira et al. [45] collected source code metrics from 40 Java FLOSS projects to perform an empirical study and propose reference values for Object-Oriented software metrics. Carneiro et al. [46] conducted an experimental study to assess a multiple views approach based on concern-driven software visualization resources, arguing that visual views support code smell detection. To perform their study, they analyzed five consecutive versions of a FLOSS project called MobileMedia, an application for photo, music, and video on mobile devices. Costa and Barros [47] analyzed eight FLOSS projects (Azureus, Eclipse ANT, Jena, JMule, JUnit, JVI, Poor Man CMS, and Sweet Home 3D) to perform an experimental study to observe if the adoption of the Common-Closure principle improves a set of software design metrics, based on a proposed technique to organize the classes into packages according to this principle. Netto et al. [48] used data from Eclipse Bugzilla to test a method based on a genetic algorithm to find the closest optimal bug correction task schedules, according to the relevance of information from bug repositories. As a practical result, they suggested a better schedule to Eclipse IDE developers.

5) Research about FLOSS: Finally, there are papers that focus on FLOSS in itself. Costa et al. [49] described Transflow, a tool aimed at analyzing data about the co-evolution of the source code and the developers’ participation and social interaction in FLOSS projects. Cavalcanti et al. [50] discussed about the bug report duplication problem that can have a negative influence on software maintenance, in particular, because it increases time spent on report analysis and validation. They conducted a study about that using eight FLOSS projects (Bugzilla, Eclipse Epiphany, Evolution, Firefox, GCC, Thunderbird, and Tomcat). Terceiro et al. [20] studied how the participation of developers in FLOSS projects can be associated with the structural complexity of the project source code. Meirelles et al. [18] analyzed source code attributes as predictors for the attractiveness of FLOSS projects, i.e., its ability to attract and retain users and developers, which is crucial for the success of the project.

C. Analysis of tools papers mentioning FLOSS

We analyzed 30 tools papers in which FLOSS terms are mentioned explicitly to identify if they described actual FLOSS tools. We separated them in four different groups according to how these tools papers mentioned FLOSS and whether they explicitly mentioned a distribution license and provided the location of a source code repository. Table IV shows the distribution of these tools papers according to our classification.

Two tools are not FLOSS, but their respective papers cited FLOSS in their bibliography: DDE by Garcia et al. [51] and X-CORE by Oliveira et al. [52]. Other four tools are not FLOSS but their authors promised that they would be FLOSS in the future: ControlPro by Moro et al. [53], ReqODE by Martins et al. [54], CRISTA by Porto et al. [55], and StArt by Zamboni et al. [56]. Another group of four tools are not FLOSS but mentioned FLOSS in their papers because they depend on FLOSS products: Mudelegen, by Simão et al. [57], depends on Bison and Flex, which are FLOSS compiler development tools. GAW, by Mangan et al. [58], is a Plug-in for Eclipse IDE. BAST, by Cavalcanti et al. [59], uses a MySQL database. ModelT2, by Albuquerque et al. [60], is related to the FLOSS UML tool BOUML.

The majority of the papers mentioning FLOSS present their respective tool as FLOSS, but 14 of them did not specify a license or repository. Additionally, their respective websites were either unavailable or did not contain this information. Although the authors had the intention to make their tools available as FLOSS, they did not succeed in doing so. Even in the best cases when the tool is actually available, they cannot be considered FLOSS because they either do not make the source code available or did not specify a license that complies with either the Free Software definition or the Open Source definition. This way C&L by Felicíssimo et al. [61], SearchEngine by Sales et al. [62], Merlin by Mrack et al. [63], WebAPSEE by Lima et al. [64], Codispe-Req by Brito and Vasconcelos [65], Captor by Shimabukuro Jr. et al. [66], BSsmart by Gomes et al. [67], Batcave by Marinho et al. [68], TeTooDS by Araujo and Delamaro et al. [69], FastInterface by Oliveira and Lula Jr. et al. [70], RBBTool by Venâncio et al. [71], Ferminie by Almeida et al. [72], AssistME

9Copyright law in most countries assumes that, unless the author explicitly states otherwise, every work of creation such as software is distributed under "All Rights Reserved" terms, so that third parties are prevented from making derived works or even redistributing the work. That is why, whenever a software product does not specify a distribution license, we must assume it is not FLOSS.
by Queiroz et al. [73], ComSCIId & DMAsp by Parreira Jr. et al. [74], cannot be actually be considered as FLOSS tools.

Finally, only six tool papers provided all the information needed to verify that they described actual FLOSS tools. Reis [75] described in details the concepts and features of Bugzilla, a well-known FLOSS issue tracking tool. Duarte et al. [76] described GridUnit, a tool to execute software automated tests in grid environments. It is available under the LGPL 2.1 license and its source code is available in a CVS repository on SourceForge. Paiva et al. [77] developed MVCASE to support their model for design rationale capture and implementation. It is available under the BSD license and its source code can be accessed from a Subversion repository. Meirelles et al. [78] presented Crab, the first version of a tool for configuration and interpretation of source code metrics. It was released under the BSD license in a Subversion repository. It was later rewritten, renamed to Kalibro, relicensed under the LGPL license, and made available from from a Git repository. Terceiro et al. [22] described Analizo, a free, multi-language, extensible source code analysis and visualization toolkit that calculates a fair number of source code metrics, generates dependency graphs, and makes software evolution analysis. Analizo source code can be accessed from a Git repository and was released under the GPL version 3 license. Ferreira et al. [79] created TaRGeT to reduce the testing costs via a systematic approach that generates test suites from use case specifications. Its source code can be accessed from a Subversion repository and is licensed under the MIT license.

D. Analysis of the Software Engineering Education Forum papers

The Software Engineering Education Forum had no published paper since 2008 and 2010 that addresses the use of FLOSS in SEE. We had a paper accepted for the 2011 edition [80], in which we discuss preliminary results of our experience with the use of FLOSS in SEE.

V. DISCUSSION

The international SE community promotes important conferences in which FLOSS is one of the main topics of interest. For example, the ACM/IEEE International Conference on Software Engineering (ICSE) has a “Software Engineering in Practice” track that explicitly solicits quality research papers addressing new development methods that “depart from traditional software engineering, such as free and open software”.

In addition, the International Federation for Information Processing (IFIP) has a dedicated working group (WG 2.13) to promote research on FLOSS. Between 2001 and 2005, IFIP organized the Open Source Software Engineering workshop series at ICSE. Later, IFIP supported the Emerging Trends in FLOSS Research and Development workshop series from 2007 to 2010 at ICSE. Since 2005, IFIP WG 2.13 promotes the International Conference on Open Source Systems (the OSS conferences). OSS’2011 is being held for the first time in Brazil at Salvador, just one week after SBES’2011.

Brazil, on the other hand, has still little tradition in FLOSS-centered research. The only partial exception is the Free Software Workshop at FISL (International Forum on Free Software) that happens every year since 2000 in Porto Alegre, Brazil.

In this study, we analyzed the majority of papers published over the last 10 years in SBES to understand how the interest on FLOSS evolved during this period. Such interest was first noticed in 2002, when two research papers we identified as being related to FLOSS were either mentioning FLOSS superficially or reporting on the usage of a FLOSS tool. Reporting on usage of FLOSS tools was already positive, since the results obtained are easier to reproduce than when private, unavailable tools are used.

Recently, the interest in FLOSS not only grew in the numbers, but also changed in nature. FLOSS moved to a more central role in the SBES community research work. We can observe three significant trends in the last two editions of SBES, which are described below.

First, we have seen a significant increase in the usage of FLOSS project data for validating general Software Engineering hypotheses, techniques, and tools [44], [45], [46], [47], [48]. This benefits Brazilian Software Engineering research by showcasing the possibility of studying data from real, active software development projects.

Second, we have witnessed the growing interest in the study of the development processes used by FLOSS projects [49], [50], [20], [18]. By understanding how FLOSS projects work, which challenges they face and how they succeed (or not succeed), we can improve the general body of knowledge on software development.

Third, we also found out that an increasing number of Brazilian researchers are taking advantage of the opportunities in releasing their tools as FLOSS. Two papers in the main track reported that they released research that produced FLOSS products [38], [39], as well as five papers in the tools session [75], [77], [78], [22], [79]. Not only their research is easier to reproduce, but those researchers can also benefit from outside collaboration in the improvement of their tools.

### TABLE IV

<table>
<thead>
<tr>
<th>Context</th>
<th>2002</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Promise to be FLOSS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>use FLOSS</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>FLOSS – no license and/or no repository</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>FLOSS</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>30</td>
</tr>
</tbody>
</table>
Unfortunately, another 14 tool paper authors were willing to make their tools available as FLOSS, but, as far as we could verify, they did not succeed on that.

These evolution trends match the opportunities we presented in Sections III-B and III-C. We believe, however, that the Brazilian Software Engineering community could be taking more advantage of these opportunities, and that we are very late in comparison with the international scenario. Nonetheless, we also believe that this is a good start.

With regard to Software Engineering Education, there are innovative experiences in using FLOSS explicitly as an educational tool for software development in both undergraduate and graduate levels in universities such as USP and UFBA[80]. We expect these experiences to become more widely spread in Brazil and their results to be presented in academic forums such as FEES, which has not been the case up to the present date.

VI. PROPOSED AGENDA

Brazil has been an important player in the FLOSS International scenario. GNU/Linux has been used as a platform for research and education in several Brazilian universities since the beginning of the 1990s. Brazil hosts one of the largest and most important FLOSS events in the world: FISL. Finally, Brazilian researchers and developers have developed FLOSS systems such as Lua, Ginga/NCL, and Noofero, which are used and well known internationally.

However, the use of FLOSS as a significant subject of Software Engineering research in Brazil is still incipient. FLOSS offers hundreds of millions of lines of code to be analyzed by researchers. Hundreds of thousands of FLOSS projects with their tens of thousands of development teams post to the web detailed information about their work and the way they interact. Nevertheless, few researchers in Brazil look at this rich set of information to base their work; meanwhile, many works rely on unrealistic toy prototypes to assess their research results.

With regard to education, university courses throughout the country usually face difficulties such as tight schedules, unbalanced student bodies, and lack of local expertise on the FLOSS ecosystem. As a result, instructors often feel pressured to choose an easier solution: either adopting a more theoretical approach in which software is not developed or focusing on the development of small proof-of-concept systems. The opportunity to learn by having hands-on experience on real-world, existing FLOSS projects is often missed.

To change this situation, we propose an agenda and invite members of the Brazilian Software Engineering community to collaborate on refining, improving, and implementing it. We propose the following actions:

1) Encouraging that software research uses public data and FLOSS tools to process and analyze data. This will promote openness and reproducibility of experimental studies, enabling them to be replicated or refined by other groups, which is a fundamental principle of science.

2) Increase the number of venues welcoming results from FLOSS research. This can be accomplished, for example, by (i) reshaping the FISL Workshop on Free Software to become a venue for publication of research results with original scientific-technological contributions and/or (ii) explicitly mentioning FLOSS as a topic of interest for traditional SE venues such as SBES itself, ESELAW (Experimental Software Engineering Latin American Workshop), and SBQS (Brazilian Symposium on Software Quality).

3) Creating a repository for publishing the FLOSS tools presented in the tools sessions of Brazilian software engineering conferences and workshops. This repository must support long-term storage of files for public downloading as well as documentation about how to publish them as FLOSS properly.

4) Fostering the use of FLOSS in SE Education. This can be achieved by (i) creating a repository of educational material, like MIT’s open courseware, focused on FLOSS education, development, and exploitation. This material should be created collaboratively and shared by educators at no cost; (ii) creating a forum where SE educators can discuss and share experiences in the use of FLOSS for SE education.

5) Fostering the creation of FLOSS Competence Centers throughout the country and having them join the international network of FLOSS Competence Centers[10]. The Brazilian Computer Society (SBC) could host a special interest group about FLOSS to coordinate the pursuit of such agenda. Despite being of particular interest to the Software Engineering special commission, FLOSS is also a topic transversal to all Computer Science areas. It would be interesting to have a forum, in the context of SBC, where the several researchers with interest in FLOSS could share experiences and pursue this agenda.

The authors intend to push this agenda forward and invite all interested colleagues to get in touch. To achieve that, we created an online community in the Brazilian Free Software social network[11], where we already started working on items 3 and 4 of the proposed agenda. We can also use this space to discuss how to accomplish the other items.

VII. FUTURE WORK

In this work, we performed a literature review on the available digital data (SBES research and tools papers). However, this review was not complete nor systematic.

Therefore, as future work, we plan to conduct a systematic review [81] to establish a more controlled process of conducting the research review, as well as to gather and analyze all existing research and tool papers from all SBES editions, available on the BDBComp and other media as well.

The development of such a systematic approach may bring improvements in precision of the data and the reliability of

---

11http://softwarelivre.org/sbc
the information, and allow other researchers to reproduce the review in different time and contexts, being able to judge the adequacy and modify the chosen criteria and standards.

ACKNOWLEDGMENTS

This research was supported by the Brazilian National Research Council (CNPq), grants 304804/2009-6 and 566164/2008-6, and partially supported by the National Institute of Science and Technology for Software Engineering (INES), CNPq grant 573964/2008-4. The authors are grateful to former SBES Tools program chairs, who provided the papers of the events they have coordinated.

REFERENCES


[38] A. Dantas, M. Gaudencio, F. Brasiliero, and W. Cirne, “Obtaining...