

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
DEPARTAMENTO ACADÊMICO DE COMPUTAÇÃO  
CURSO SUPERIOR DE TECNOLOGIA EM SISTEMAS PARA INTERNET**

**SAULO FERNANDO CARNEIRO VERUSSA**

**A INFLUÊNCIA DO PAGAMENTO DE BOLSAS DE CURTO PRAZO  
NA RETENÇÃO DE VOLUNTÁRIOS DO GOOGLE SUMMER OF  
CODE**

**TRABALHO DE CONCLUSÃO DE CURSO**

**CAMPO MOURÃO - PR**

**2016**

**SAULO FERNANDO CARNEIRO VERUSSA**

**A INFLUÊNCIA DO PAGAMENTO DE BOLSAS DE CURTO PRAZO  
NA RETENÇÃO DE VOLUNTÁRIOS DO GOOGLE SUMMER OF  
CODE**

Trabalho de Conclusão de Curso apresentado ao Curso Superior de Tecnologia em Sistemas para Internet da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de Tecnólogo em Tecnologia em Sistemas para Internet.

Orientador: Igor Fábio Steinmacher, PhD.

**CAMPO MOURÃO - PR**

**2016**

## **AGRADECIMENTOS**

A minha querida esposa, Mônica Vaz de Carvalho Verussa, que muito me incentivou durante todos esses anos de estudo.

A minha pequena filha, Júlia Verussa que em muitas noites sentiu minha ausência.

Aos professores aos quais eu tive o prazer de frequentar as aulas, em especial ao meu orientador Igor Fábio Steinmacher, pela paciência e presteza nas orientações que tornaram possível este projeto.



## ATA DA DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO

Às **dezenove horas** do dia **onze de dezembro de dois mil e quinze** foi realizada no **auditório do EAD** a sessão pública da defesa do Trabalho de Conclusão do Curso Superior de Tecnologia em Sistemas para Internet do acadêmico **Saulo Fernando Carneiro Verussa** com o título **“A influência do pagamento de bolsas de curto prazo na retenção de voluntários do Google Summer of Code”**. Estavam presentes, além do acadêmico, os membros da banca examinadora composta pelo professor **Igor Fabio Steinmacher** (Orientador-Presidente), pelo professor **Rafael Liberato Roberto** e pelo professor **Marco Aurélio Graciotto Silva**. Inicialmente, o aluno fez a apresentação do seu trabalho, sendo, em seguida, arguido pela banca examinadora. Após as arguições, sem a presença do acadêmico, a banca examinadora o considerou **APROVADO** na disciplina de Trabalho de Conclusão de Curso e atribuiu, em consenso, a nota **9.0 ( NOVE )**. Este resultado foi comunicado ao acadêmico e aos presentes na sessão pública. A banca examinadora também comunicou ao acadêmico que este resultado fica condicionado à entrega da versão final dentro dos padrões e da documentação exigida pela UTFPR ao professor Responsável do TCC no prazo de **onze dias**. Em seguida foi encerrada a sessão e, para constar, foi lavrada a presente Ata que segue assinada pelos membros da banca examinadora, após lida e considerada conforme.

Observações:

---

---

---

---

---

Campo Mourão, 11 de dezembro de 2015

A folha de aprovação assinada encontra-se na coordenação do curso

## RESUMO

VERUSSA, Saulo Fernando Carneiro. A INFLUÊNCIA DO PAGAMENTO DE BOLSAS DE CURTO PRAZO NA RETENÇÃO DE VOLUNTÁRIOS DO GOOGLE SUMMER OF CODE. 28 f. Trabalho de Conclusão de Curso – Curso Superior de Tecnologia em Sistemas para Internet, Universidade Tecnológica Federal do Paraná. Campo Mourão - PR, 2016.

A participação de voluntários é fundamental para o desenvolvimento de projetos de Software Livre baseados em comunidades. A compreensão dos fatores que motivam estudantes a ingressar e contribuir com esses projetos é de extrema relevância, haja vista que o ingresso e a retenção de novatos são de vital importância para o futuro do Software Livre. Nesse sentido, tendo como base o programa global Google Summer of Code (GSoC), que oferece o pagamento de bolsas de estudo para voluntários, analisamos a eficiência da abordagem de motivação financeira utilizada pelo GSoC no que tange ao aumento da taxa de retenção de novatos para as comunidades de Software Livre. Por meio do uso de técnicas de mineração dos repositórios de software, acumulamos dados relevantes para a análise das taxas de participação dos voluntários e sua evolução após o fim do projeto. Dos voluntários que contribuíram para o LibreOffice no GSoC de 2013, mineramos dados das listas de e-mail, gerenciador de funcionalidades, do repositório e da revisão de código(Git-SCM). Consideramos os períodos de pré-inscrição, inscrição, durante, e após o estágio. A análise dos dados nos permitiu observar que o projeto foi satisfatório e cumpriu com seu papel de incentivar o ingresso de novos voluntários. Mesmo que alguns estudantes tenham uma participação anterior ao projeto, os números indicam um aumento na participação. Contudo, apenas o pagamento da bolsa de estágio, enquanto fator de incentivo inicial, não é determinante para a retenção dos voluntários novatos na comunidade de Software Livre. Provavelmente outros fatores, que nos são desconhecidos e fogem ao alcance dessa pesquisa, são influentes e determinantes para um voluntário continuar a colaborar. Acreditamos que, no tocante ao LibreOffice, esta pesquisa representa uma contribuição significativa para o estudo da eficiência do incentivo financeiro e também para a sustentabilidade dos projetos de Software Livre.

**Palavras-chave:** Software Livre, Google Summer of Code, estudante voluntário, motivação, novato

## ABSTRACT

VERUSSA, Saulo Fernando Carneiro. TITLE GOOGLE SUMMER OF CODE. 28 f. Trabalho de Conclusão de Curso – Curso Superior de Tecnologia em Sistemas para Internet, Universidade Tecnológica Federal do Paraná. Campo Mourão - PR, 2016.

The participation of volunteers is critical to the development of Free Software projects based on communities. Understanding the factors that motivate students to join and contribute to these projects is extremely important, given that the entry and retention of newcomers are of vital importance to the future of Free Software. In this sense, based on the global Google Summer of Code program (GSoC), which provides for the payment of scholarships for volunteers, we analyze the efficiency of financial motivation approach used by the GSoC with respect to increased beginners retention rate to the Free Software communities. Through the use of the software repositories mining techniques, we have accumulated data relevant to the analysis of volunteers' participation rates and their evolution after the end of the project. The volunteers who contributed to the LibreOffice in 2013 GSoC, mined data from e-mail lists, functions manager, repository and code review (Git-SCM). Consider the periods of pre-registration, registration, during, and after the stage. Data analysis allowed us to observe that the project was satisfactory and fulfilled its role of encouraging the entry of new volunteers. While some students have a previous participation to the project, the figures indicate an increase in participation. However, only the payment of the training grant, as an initial incentive factor is not decisive for the retention of novice volunteers in the Free Software community. Probably other factors that are unknown to us and are beyond the scope of this research are influential and instrumental in a volunteer continue working. We believe that, with regard to LibreOffice, this research is a significant contribution to the study of the efficiency of financial incentives and also to the sustainability of Free Software projects.

**Keywords:** Open Source, Google Summer of Code, voluntary student, motivation, beginners

## LISTA DE FIGURAS

FIGURA 1	– Método de pesquisa .....	8
FIGURA 2	– Distribuição de interações dos voluntários com outliers .....	16
FIGURA 3	– Distribuição de interações dos voluntários sem outliers .....	17
FIGURA 4	– Total de interações por voluntário .....	18
FIGURA 5	– Total de interações na ferramentas por voluntário .....	19
FIGURA 6	– Tabela da interação dos voluntários antes e após o Google Summer of Code	21
FIGURA 7	– Interação dos voluntários por ferramentas .....	22
FIGURA 8	– Tabela .....	26

## **LISTA DE TABELAS**

TABELA 1	– Total de estagiários/bolsistas distintos em cada período. ....	15
----------	--	----



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>3</b>
2.1	FUNDAMENTAÇÃO TEÓRICA	3
2.1.1	Software Livre E O GOOGLE SUMMER OF CODE	3
2.1.2	VOLUNTÁRIOS E MOTIVAÇÃO	4
2.1.3	MINERAÇÃO DE DADOS	5
2.2	TRABALHOS RELACIONADOS	5
2.3	CONSIDERAÇÕES FINAIS	7
<b>3</b>	<b>MÉTODO DE PESQUISA</b>	<b>8</b>
3.1	ESCOLHA DO PROJETO	9
3.2	RECUPERAÇÃO DOS DADOS	10
3.3	EXTRAÇÃO DE DADOS	10
3.4	ANÁLISE DOS DADOS	11
<b>4</b>	<b>RESULTADOS</b>	<b>14</b>
4.1	ESCOLHA DO PROJETO A SER ANALISADO	14
4.2	MINERAÇÃO DOS REPOSITÓRIOS	14
4.3	ANÁLISE DOS DADOS	15
4.3.1	Análise das interações dos voluntários	17
4.4	CONSIDERAÇÕES FINAIS	21
<b>5</b>	<b>CONCLUSÕES</b>	<b>23</b>
5.1	LIMITAÇÕES	24
5.2	TRABALHOS FUTUROS	25
	<b>Apêndice A – TABELA DAS PARTICIPAÇÕES DOS VOLUNTÁRIOS</b>	<b>26</b>
	<b>REFERÊNCIAS</b>	<b>27</b>

## 1 INTRODUÇÃO

Uma grande quantidade de Software Livre está disponível na Internet para qualquer usuário que tenha interesse em utilizar, modificar ou até mesmo contribuir com os projetos (KROGH et al., 2012). A facilidade de acesso e qualidade dos produtos disponibilizados sob licenças de Software Livre é um dos principais fatores de sucesso do modelo. Assim, empresas ou desenvolvedores se tornam usuários e contribuintes desse tipo de projeto (SPAETH et al., 2014).

Uma característica importante de diversos projetos são as comunidades de voluntários que colaboram com seu conhecimento, influências e tempo de trabalho em conjunto para o crescimento do projeto e sua gerência, formando um projeto complexo e intensamente ativo (GOLDMAN; GABRIEL, 2005). Os voluntários de projetos de Software Livre podem colaborar de diferentes formas: em discussões sobre o projeto, indicação de defeitos do produto (*bugs*), correção de códigos defeituosos e contribuindo com código de novas funcionalidades.

Apesar dos projetos necessitarem de voluntários, por vezes os novos desenvolvedores enfrentam barreiras e desistem de contribuir (STEINMACHER et al., 2015). Para manter os projetos sustentáveis, é fundamental que as comunidades cresçam com a entrada de novos membros que possam se manter ativos. Assim, a entrada e a permanência de novos participantes se tornam cruciais para o sucesso e expansão do projeto (STEINMACHER et al., 2014).

A fim de fomentar a entrada de novos participantes em seus projetos, o Google começou, no ano de 2009, o projeto Google Summer of Code (GOOGLE, 2015a). Anualmente, organizações submetem propostas de trabalhos para projetos de Software Livre, acessíveis para estudantes em todo o mundo. Havendo consonância entre a proposta da organização e a do aluno, o mesmo pode ingressar e colaborar com projetos em várias áreas. Em troca, o Google oferece recompensa financeira na forma de bolsa de estágio de três meses, paga à medida que o projeto evolui até a entrega final do trabalho (TRAINER et al., 2014).

Ainda que o objetivo seja incentivar a entrada de novos participantes e aumentar as contribuições para projetos de Software Livre, pouco se sabe sobre o real impacto dessa inicia-

tiva na prática. Foram encontrados na literatura poucos artigos que forneçam dados e pesquisas que revelem a taxa de retenção ou que relatem os benefícios do GSoC nos projetos.

Esta pesquisa teve como objetivo investigar qual o nível de atividade dos voluntários incentivados pela bolsa de estudo e quais as atividades que os mesmos realizaram na comunidade após o fim das tarefas propostas no projeto.

Como guia, utilizamos as seguintes questões de pesquisa:

- Os participantes do GSoC tinham envolvimento anterior com o projeto?
- Que atividade desenvolveram durante o GSoC?
- As contribuições realizadas pelos participantes do GSoC foram aceitas pelo projeto?
- Qual o índice de retenção dos bolsistas após o término do projeto?

O restante deste trabalho está organizado da seguinte maneira: O referencial teórico é exposto no Capítulo 2. O Capítulo 3 dispõe sobre a metodologia de pesquisa, detalhando como foram realizados os passos deste trabalho. No Capítulo 4 há uma explicação e análise sobre os resultados encontrados, e, por fim, no Capítulo 5 as conclusões.

## 2 REVISÃO BIBLIOGRÁFICA

Uma revisão da literatura foi realizada com o intuito de encontrar estudos relacionados com os conceitos deste trabalho. Foram analisados trabalhos relacionados a Software Livre e o Google Summer of Code abordado na Seção 2.1.1. E motivação de voluntários de Software Livre apresentados na Seção 2.1.2.

### 2.1 FUNDAMENTAÇÃO TEÓRICA

Nesta seção apresentamos alguns conceitos relevantes a este trabalho. Foram abordados conceitos relacionados ao Software Livre, voluntários e motivação.

#### 2.1.1 SOFTWARE LIVRE E O GOOGLE SUMMER OF CODE

Os desenvolvedores que disponibilizam seu software como livres utilizam licenças que garantem um número de direitos para seus futuros usuários. Isso é possível pelo fato do software não ter custo de aquisição e permitir acesso livre a qualquer pessoa, podendo, inclusive, ser alterado e/ ou melhorado para atender novas necessidades. Existem inúmeras licenças, podemos citar como exemplo a General Public License (GPL), que garante direitos de acesso, cópia e até mesmo obtenção do código fonte do software (HIPPEL; KROGH, 2003).

O Software Livre, com sua grande diversidade de produtos, tem coexistido com o mercado corporativo há anos. Esse modelo de colaboração estima o ganho da comunidade com a participação do voluntariado e tem despertado muito o interesse de pesquisas sobre a sua sustentabilidade. Tal interesse decorre do fato de que a entrada de novatos é fundamental para o que a comunidade se mantenha ativa e seus membros adquiram cada vez mais importância (KRISHNAMURTHY et al., 2014).

O Google Summer of Code é um programa global que ocorre anualmente desde 2009 e tem como objetivo inspirar a participação de estudantes no desenvolvimento de Software Livre, além de ajudar os projetos a adquirirem novos desenvolvedores (GOOGLE, 2015b). É oferecida

uma bolsa de incentivo no valor de US 5.500,00 (cinco e mil e quinhentos dólares). Assim como o projeto, a bolsa é dividida em três etapas pagas conforme o trabalho é desenvolvido. O GSoC somente aceita a participação de estudantes como voluntários, os quais são orientados por mentores durante o andamento do projeto (TRAINER et al., 2014).

Desde que foi criado o programa, organizações diversificadas ofertam tarefas que demandam desenvolvidas essencialmente para projetos de Software Livre. Cabe mencionar que a quantidade de organizações é limitada por edições anuais.

As comunidades de Software Livre que têm interesse de participar, entregam projetos de funcionalidades que necessitam ser desenvolvidas. Em etapa posterior, os estudantes fazem um outro projeto para o desenvolvimento das funcionalidades especificadas. Aos mentores cabe a escolha dos estudantes para se tornar voluntários, dentre esses projetos.

Nos últimos anos o GSoC teve participantes de 101 países e beneficiou mais de 500 organizações. Estima-se que houve o envolvimento de aproximadamente 8.500 estudantes e 8.300 (oito mil e trezentos) tutores, produzindo um total superior a 50 milhões de linhas de códigos (GOOGLE, 2015b).

### 2.1.2 VOLUNTÁRIOS E MOTIVAÇÃO

Como citado pelos pesquisadores Steinmacher et al. (2014), a entrada de novos voluntários é de suma importância. Todavia, há múltiplas barreiras que dificultam o ingresso de novos voluntários no projeto, prejudicando o aumento da comunidade e a retenção de novatos. O tempo de aprendizado e a distância entre os participantes ilustram essa situação e, como são comuns nesses ambientes, geram redução importante na taxa de entrada e retenção de novatos.

Dada a relevância dos voluntários para uma comunidade de Software Livre, fatores que motivam seu ingresso são constantemente alvo de interesse de pesquisadores. No tocante a motivação dos desenvolvedores de software, Krishnamurthy et al. (2014) identificam duas formas: motivação intrínseca e extrínseca. Na primeira, a satisfação por colaborar com a comunidade é originada por uma vontade que parte do próprio desenvolvedor, como quando a pessoa gosta de programar e faz isso para se sentir bem, por exemplo. Já a motivação extrínseca parte de uma força externa que motiva o voluntário a conquistar algo de seu interesse, como a recompensa financeira recebida após um trabalho ou algum benefício para sua carreira como desenvolvedor.

Alexy e Leitner (2010) apontam que voluntários movidos pela motivação intrínseca são muitas vezes impulsionados por necessidades e determinações pessoais. Este é o caso de alguns

ideologistas que consideram a sua contribuição como um presente para a comunidade. Por outro lado, quando os voluntários têm inclinações para motivações extrínsecas, é mais provável que o incentivo financeiro tenha maior impacto sobre os novatos. Assim, considerando a motivação extrínseca, várias empresas privadas nos últimos anos aumentaram consideravelmente o valor investido em Software Livre.

### 2.1.3 MINERAÇÃO DE DADOS

Para a coleta dos dados fizemos uso da base de dados de algumas ferramentas empregadas no desenvolvimento de software. Por meio destas ferramentas de mineração conseguimos gerar dados para observação. Esta técnica de coleta de dados foi utilizada no capítulo de metodologia 3.3 para extração de dados.

Atualmente a identificação de padrões em bases de dados é uma importante fonte de pesquisas. O procedimento de tornar dados em conhecimento pode ser denominado como mineração de dados. Para Fayyad et al. (1996) a mineração é um passo no processo de busca por conhecimento, realizado por uma aplicação de algoritmos específicos para extração de padrões nos dados. Sua aplicação pode ser de proveito para diferentes áreas, tais como marketing, investimentos ou esportes.

Na área de desenvolvimento de software, a mineração de dados tem auxiliado no estudos de diversas partes. Davies et al. (2010) citam como exemplo estudos de propagações de alterações, erros de código fonte, previsão e identificação de erros, padrões da dinâmica de equipes de desenvolvimento e evolução de software.

Nesta pesquisa utilizamos algumas ferramentas como exibido na Seção 3.2, para auxiliar no reconhecimento do padrão dos voluntários.

## 2.2 TRABALHOS RELACIONADOS

Analisando a literatura encontramos algumas pesquisas que revelam a taxa de retenção ou que relatam os benefícios do GSoC nos projetos. Há trabalhos que dissertam sobre Google Summer of Code e a motivação financeira para entrada em projetos de Software Livre.

Trainer et al. (2014) estudam os impactos da participação de voluntários em projetos de Software Livre, particularmente relacionados ao GSoC e *Hackathons*<sup>1</sup>. Para o estudo foi conduzida entrevista estruturada com os desenvolvedores. As informações obtidas foram cru-

---

<sup>1</sup>Evento onde programadores trabalham intensivamente <https://en.wikipedia.org/wiki/Hackathon>

zadas com dados da literatura e com os projetos *hackathons*. Os resultados da entrevista com usuários do GSoC indicam uma permanência no projeto de 18% do total de entrevistados. Os pesquisadores levantam alguns pontos de discussão, como o crescimento da comunidade e a oportunidade que o GSoC fornece para os novatos terem contato com membros mais experientes pelo sistema de mentores adotado nas tarefas.

Outro estudos sobre a entrada e retenção de novatos em projetos de Software Livre. Em seu estudo, Schilling et al. (2012) analisam dois grupos com características distintas e avaliam o comportamento de retenção de desenvolvedores. O primeiro é formado por pessoas selecionadas com perfil destinado ao trabalho, com aptidões mais compatíveis com a necessidade da atividade a ser executada. No segundo grupo, há pessoas com perfil selecionado para maior compatibilidade com a equipe existente. A pesquisa utiliza estudantes que participaram de projetos do GSoC relacionados ao KDE<sup>2</sup> e conclui que é apropriado selecionar pessoas com as características dos dois grupos citados para melhorar a retenção dos voluntários no projeto. Além disso, contrariando as expectativas, estudantes com formação acadêmica de nível superior não permanecem mais tempo no projeto em relação aos demais.

Ainda sobre a entrada de novatos, alguns estudos concentram-se unicamente na motivação que leva os desenvolvedores a entrarem em projetos de Software Livre. Alexy e Leitner (2010) confrontam as motivações extrínsecas e intrínsecas com outros fatores, propondo que o efeito de motivação dos desenvolvedores por recompensas financeiras é muito mais complexa. A coleta dos dados foi realizada por meio de um questionário online com cenários simulados e questões direcionadas a desenvolvedores. Com a obtenção de 229 respostas foi possível observar que os efeitos do pagamento influenciam na motivação intrínseca e que há a necessidade de considerar essa complexidade em pesquisas futuras.

Krishnamurthy et al. (2014) apresentam um estudo sobre a aceitação do pagamento por desenvolvedores de Software Livre, utilizando uma tabela proposta para aferir a intenção de aceitar incentivos financeiros. A mineração envolveu repositórios do Sourceforge<sup>3</sup>, nos quais foram escolhidos os 15 projetos mais ativos. A pesquisa constatou que desenvolvedores com tendências motivacionais extrínsecas e intrínsecas têm maior intenção de aceitar recompensas financeiras, enquanto que motivações do tipo ideológicas e comunitárias levam a uma redução no nível de aceitação.

Já os pesquisadores Steinmacher et al. (2012) investigaram os motivos da desistência de novatos. Pelo estudo dos repositórios do *framework* Hadoop Common<sup>4</sup> foram extraídos

---

<sup>2</sup>Projeto de Interface Gráfica Livre para Sistema Linux

<sup>3</sup><http://sourceforge.net/>

<sup>4</sup><https://hadoop.apache.org/>

dados dos e-mails do projeto e analisadas as informações referentes aos novatos, com base no tipo de respostas que recebiam dos membros mais experientes. Concluíram que 18% dos novatos continuaram a participar do projeto e que uma grande barreira é conseguir sanar as dúvidas iniciais sobre o mesmo.

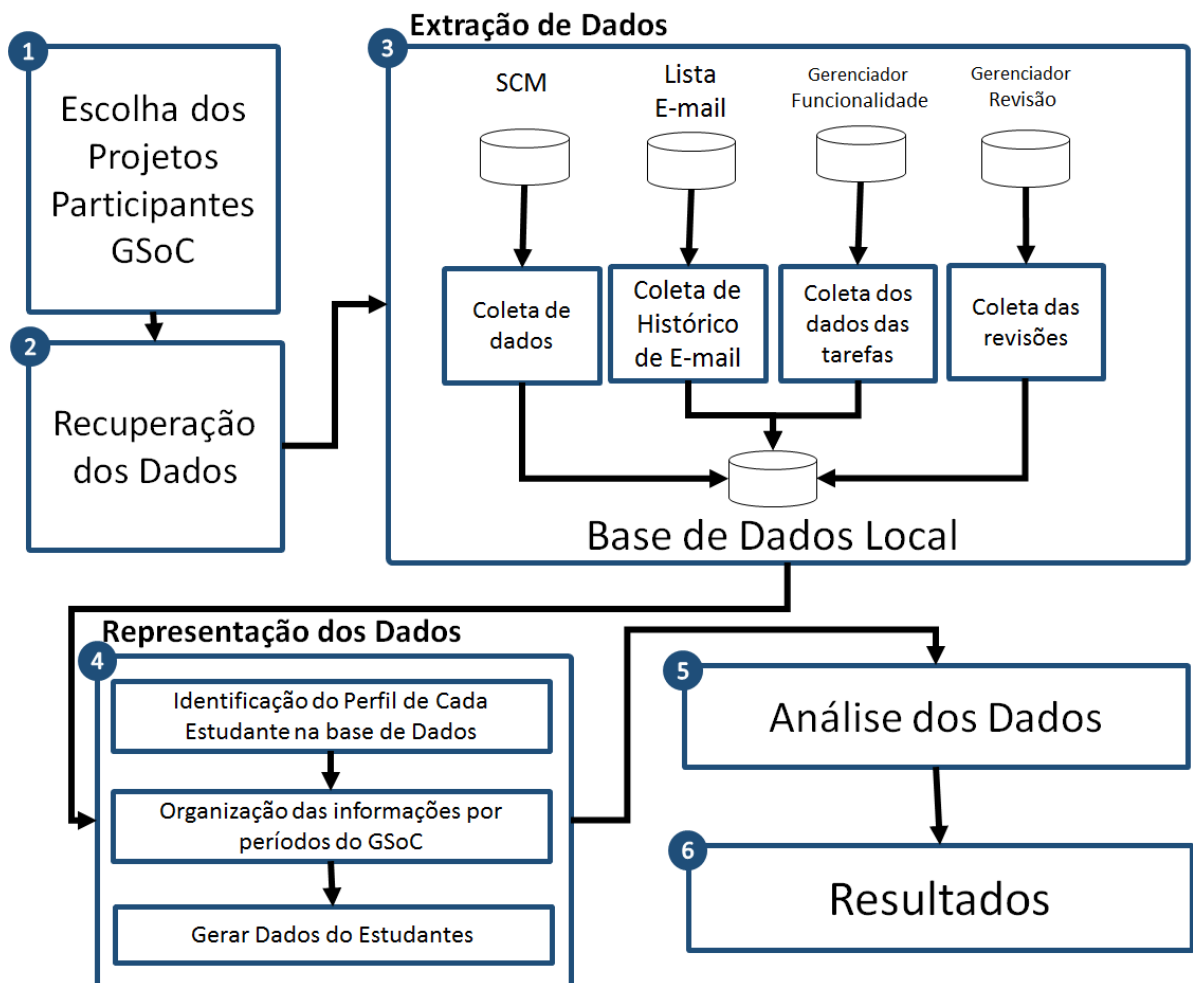
### 2.3 CONSIDERAÇÕES FINAIS

Tendo como foco a motivação extrínseca buscamos verificar se a abordagem de motivação financeira utilizada pelo GSoC, por meio de pagamento de bolsas para alunos foi eficiente ou não, no que diz respeito a aumentar as taxas de retenção de voluntários para as comunidades de Software Livre.

Considerando a importância do ingresso e retenção de novatos para o futuro dessas comunidades, entendemos que pesquisas como a que propomos - através da mineração dos repositórios, análise das taxas de participação e evolução após o fim do projeto - também se apresentam como uma contribuição significativa para a sustentabilidade dos projetos de Software Livre.



### 3 MÉTODO DE PESQUISA



**Figura 1: Método de pesquisa**

A Figura 1 ilustra o processo. Para alcançar o objetivo proposto, o método de pesquisa foi dividido em cinco passos distintos. O primeiro passo da pesquisa (Escolha dos Projetos) foi definir qual projeto participante do Google Summer of Code seria analisado. O passo seguinte (Recuperação dos Dados) foi recolher informações sobre o projeto alvo, como o repositório de código fonte, *issue tracking* e lista de e-mails. O passo três (Mineração dos Repositórios) consistiu em extrair as informações da interação com a comunidade do projeto no período proposto.

Em seguida, no passo quatro (Representação dos Dados), as informações foram representadas em participações, quantificadas e organizadas para análise. O perfil do usuário foi identificado e suas interações foram filtradas e quantificadas. Por fim, o quinto passo (Análise dos Dados), os dados dos voluntários foram convertidos em tabelas e gráficos para que pudessem ser analisados.

### 3.1 ESCOLHA DO PROJETO

Os projetos das comunidades que são aceitos no Google Summer of Code estão todos documentados e armazenados. Os seus arquivos, separados por edição anual, estão disponíveis para consulta no seguinte endereço de livre acesso da plataforma de gerencia do GSoC (GOOGLE, 2015a). A lista de projetos aceitos conta com informações como o nome do voluntário que participou do projeto, quem foi o seu tutor, para qual organização o trabalho foi destinado e qual era o seu objetivo para desenvolver.

Organizações de áreas diversas distintas estão presentes no GSoC, constituindo uma base de dados ampla e variada. Como explorá-la mais aprofundadamente seria inviável para a presente pesquisa, optamos pela delimitação de um grupo específico de voluntários.

Escolhemos os voluntários participantes do projeto LibreOffice<sup>1</sup> no período da edição do GSoC 2013. Alguns fatores foram determinantes na definição do grupo como a disponibilidade e facilidade de acesso aos repositórios, a estabilidade do projeto, bem como a consistência dos dados.

As ferramentas de gerenciamento utilizadas pela comunidade também foram determinantes para a escolha do projeto. Elas asseguram o acesso a dados do histórico do desenvolvimento do software. Como os projetos participantes são livres tivemos acesso aos gerenciadores de código fonte (GIT-SCM), de funcionalidades (Bugzilla), de revisão de código (Gerrit) e lista de e-mail. Todas são ferramentas comumente utilizadas por comunidades de Software Livre.

A cada edição do GSoC são estipuladas previamente datas, para conclusão de cada etapa do projeto, isso é importante para esta pesquisa. Pois com os períodos definidos, é possível fazer uma observação das atividades dos voluntários e identificar padrões nas suas contribuições.

---

<sup>1</sup><https://libreoffice.org/>

### 3.2 RECUPERAÇÃO DOS DADOS

Para estudar o projeto definido foi necessário recolher dados do histórico do desenvolvimento referentes aos voluntários em foco. Em projetos de Software Livre há certas características que tornam possível a recuperação de tais informações para estudo. Assim, após a definição dos projetos e voluntários, algumas ferramentas foram analisadas, como gerenciador de funcionalidades, listas de e-mail, revisão de código e repositório de código fonte (SCM).

O gerenciador de funcionalidades (Bugzilla), é um ambiente no qual a comunidade relata problemas no software e também faz solicitações de novas funcionalidades. Ele possibilita a troca de informações, sugestões e relatos de experiências, comentando-se nas requisições criadas.

Nas listas de e-mail são geralmente disponibilizados arquivos com informações sobre discussões entre os integrantes da comunidade. Analisando esses dados é possível descobrir quando e quem são as pessoas que criaram essas mensagens.

Outra plataforma de interação dos voluntários é a de revisão de código, a qual é utilizada para solicitar revisões das submissões de código fonte realizadas no projeto. Também permite revisar códigos de outros voluntários. Cria-se assim um sistema cooperativo entre os desenvolvedores, sendo possível identificar se o voluntário fez alguma solicitação de revisão ou se participou de alguma atividade como revisor.

O repositório de código fonte armazena o histórico do desenvolvimento do sistema. Analisando a submissões dos voluntários é possível resgatar a sua atividade no período proposto.

### 3.3 EXTRAÇÃO DE DADOS

O passo seguinte foi a extração de informações da participação dos voluntários com ajuda de recursos de pesquisa próprios da plataforma. Foram utilizadas também ferramentas de mineração de dados disponíveis para extração das tarefas realizadas. Para a assistência na recuperação dos dados utilizamos o conjunto de software de linha de comando da MetricsGrimoire<sup>2</sup>, que permite exportar informações dos repositórios e armazená-las em um banco de dados.

Para reunir os dados, fizemos uso de algumas ferramentas MetricsGrimoire, a saber:

---

<sup>2</sup><http://metricsgrimoire.github.io/>

- Bicho: utilizada para recuperar e organizar as informações de sistemas gerenciadores de tarefas e também gerenciadores de revisão, como, o Bugzilla e Gerrit;
- MailingListStats: recupera os dados de listas de e-mail. Para esta pesquisa, foi utilizada para obter e armazenar os dados históricos das listas de e-mail que são arquivados pelo projeto no formato mbox;
- CVSanaly: recupera os dados de sistemas de controle de versão e configuração (controle de versão). Utilizado para coletar os dados do repositório Git-SCM do projeto.

### 3.4 ANÁLISE DOS DADOS

Após a recuperação dos dados, armazenamos as informações dos repositórios em banco de dados local e procedemos à identificação dos voluntários conforme o nome cadastrado no site do GSoC. Encontramos também o endereço eletrônico (e-mail) explorando os arquivos de código fonte submetidos no projeto final dos voluntários.

Utilizamos uma combinação entre nome, sobrenome e endereço eletrônico (e-mail) aplicados em consultas realizadas no banco de dados:

- <NomeCompleto>
- <%Nome%>
- <%Sobrenome%>
- <%E-mail%>

Por meio das consultas nos locais já mencionados encontramos 11 voluntários participantes do projeto, representando a totalidade da nossa amostra.

Fizemos um detalhamento dos dados, observando os períodos e as ferramentas trabalhadas. Isso permitiu uma visualização minuciosa e facilitou a etapa posterior de análise das informações mais relevantes.

Assim, os dados foram organizados em seis períodos distintos para ajudar no entendimento da atividade dos voluntários. Essa organização nos permitiu visualizar como é a interação com a comunidade. Os períodos correspondem às etapas do GSoC 2013 e são apresentados a seguir:

- Pré-Inscrição: antes de 22-04-2013 (período que antecede a submissão de proposta do estudante)
- Inscrição: de 22-04-2013 até 03-05-2013 (entrega de proposta de projetos)
- Hiato: 03-05-2013 até 27-05-2013 (período de inatividade dos voluntários, esperando a aceitação do projeto)
- Integração: 27-05-2013 até 17-06-2013 (integração do voluntário com a comunidade)
- Codificação: 17-05-2013 até 30-09-2013 (desenvolvimento e entrega da funcionalidade do projeto)
- Após GSoC: após 30-09-2013 (término do projeto)

Considerando a importância de se analisar as ferramentas separadamente, em cada período especificado, as informações individuais dos voluntários foram separados por fonte, tal como segue:

- SCM
  - Submissão feita pelo voluntário
  - Submissão do voluntário enviada ao repositório por outro membro da comunidade
- Lista de E-mail
  - E-mail enviado pelo voluntário
- Gerenciador de Funcionalidade
  - Publicação submetida pelo voluntário
  - Publicação atribuída ao voluntário
  - Publicação comentada pelo voluntário
  - Publicação assistida pelo voluntário
- Gerenciador de Revisão
  - Voluntário é proprietário de revisão
  - Voluntário figura como revisor do código

Baseado nesses períodos e meios de interação, as contribuições dos voluntários com a comunidade foram quantificadas e analisadas. Isso foi feito observando a variação no decorrer do tempo. Para tal, comparamos a quantidade de interações de cada voluntário entre os períodos listados e geramos gráficos para facilitar a visualização das informações.

## 4 RESULTADOS

Aplicamos o método proposto no capítulo anterior nos repositórios do projeto LibreOffice. Identificamos onze voluntários participantes do Google Summer of Code do ano de 2013, como já mencionado.

### 4.1 ESCOLHA DO PROJETO A SER ANALISADO

Conforme apresentado na seção da metodologia de pesquisa 3.1, diversas características favorecem a escolha de um projeto para estudo por mineração. O que faz do GSoC um candidato interessante para esse tipo de pesquisa. Porque os voluntários são incentivados por motivação extrínseca. Como também as datas das etapas de desenvolvimento, que facilitam o estudo do comportamento dos voluntários pelas suas atividades nas ferramentas.

O projeto escolhido é da edição de 2013 do GSoC, essa escolha teve o intuito de facilitar a observação da atividade dos voluntários nas edições prévias e posteriores. Dessa maneira caso existisse mais de uma participação por parte do voluntário no GSoC anterior ou posterior, poderia ser identificada com mais facilidade.

No total 11 voluntários do GSoC foram estudados na pesquisa. Todos participantes de projetos de desenvolvimento para a comunidade do LibreOffice.

### 4.2 MINERAÇÃO DOS REPOSITÓRIOS

As informações foram mineradas a partir da lista de e-mail, software gerenciador de tarefas (Bugzilla), sistema de revisão de código (Gerrit) e os registros do repositório de código-fonte (Git-SCM), localizados nos endereços de internet listados abaixo, na ordem que segue:

- <http://lists.freedesktop.org/archives/libreoffice/>
- <https://bugs.documentfoundation.org/>

- <https://gerrit.libreoffice.org>
- <git://anongit.freedesktop.org/libreoffice/core>

**Tabela 1: Total de estagiários/bolsistas distintos em cada período.**

	<b>Período</b>	<b>Gerenciador Funcionalidade (Bugzilla)</b>	<b>Lista de E-mail</b>	<b>Revisor Código (Gerrit)</b>	<b>SCM (Git-SCM)</b>
TOTAL DE ESTUDANTES INTERAGINDO POR FERRAMENTA	Pré-inscrição	9	10	9	10
	Inscrição	4	5	1	2
	Hiato	4	1	0	2
	Integração	4	11	5	7
	Codificação	6	11	11	11
	Após GSoC	8	7	8	6

Após a extração dos dados da comunidade e armazenamento no banco de dados, as consultas foram realizadas para separar a participação dos voluntários pesquisados.

Os resultados foram organizados entre as diferentes plataformas, separados por atividades executadas e quantificados. A Tabela com todos os detalhes é apresentada para consulta no Apêndice A.

A Tabela 1 apresenta a quantidade de voluntários que interagiu em cada uma das ferramentas, separadas pelos períodos analisados. Verificamos, por exemplo, que durante o período de codificação todos o voluntários interagiram em três ferramentas. São elas: lista de e-mail, o revisor de código e o SCM. Observamos ainda que 10 voluntários interagiram no SCM, previamente ao início do GSoC.

A quantidade de voluntários que interagiram com as ferramentas entre os diferentes períodos do GSoC 2013. Destacamos que no período de pré-inscrição todas as ferramentas foram utilizadas por pelo menos 9 voluntários.

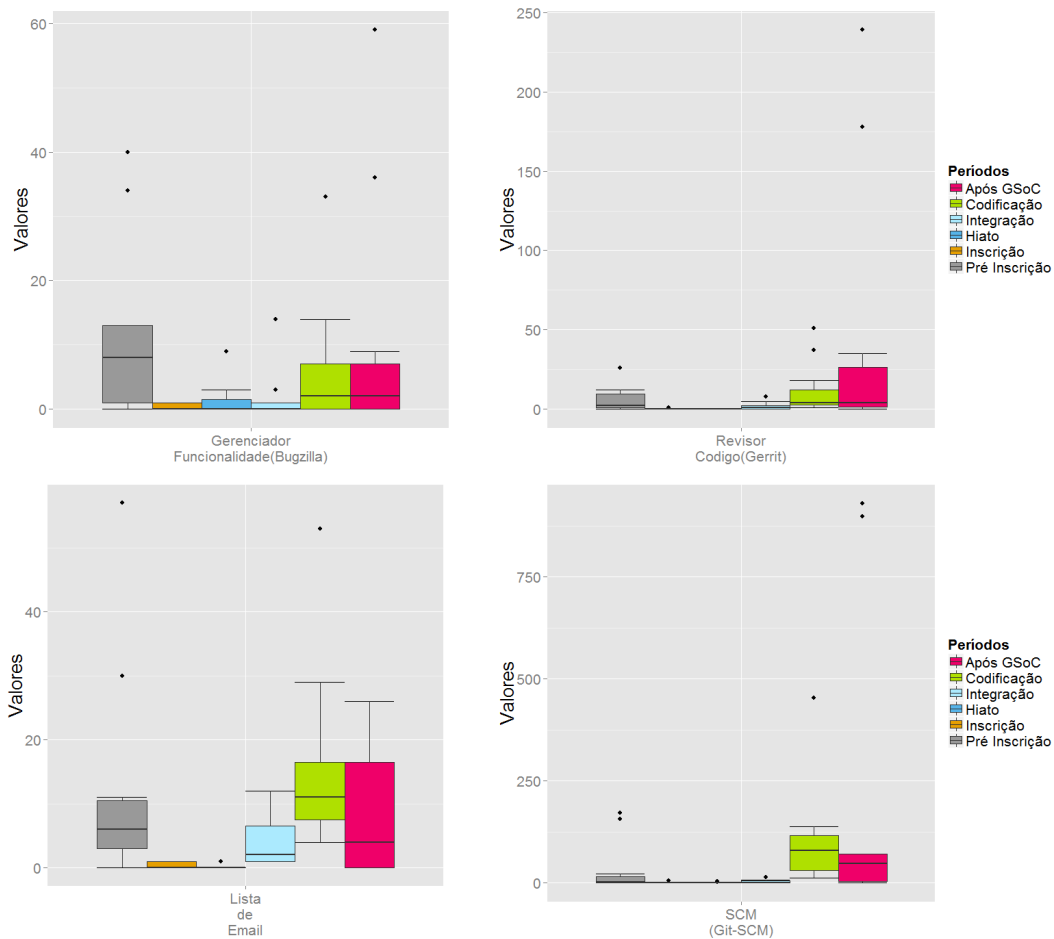
#### 4.3 ANÁLISE DOS DADOS

Os gráficos das Figuras 2 e 3, apresentam a distribuição do total das interações dos voluntários entre as ferramentas. Para cada ferramenta temos um grupo de períodos, e podemos observar em cada coluna os limites superiores e inferiores, que são indicados pelas arestas. As caixas indicam a variabilidade entre os períodos e as medianas. Os pontos fora dos limites indicam valores discrepantes.

Observando os dados, notamos que alguns voluntários têm uma quantidade muito maior de interações, gerando dados discrepantes (*outliers*) para a pesquisa. Geramos dois



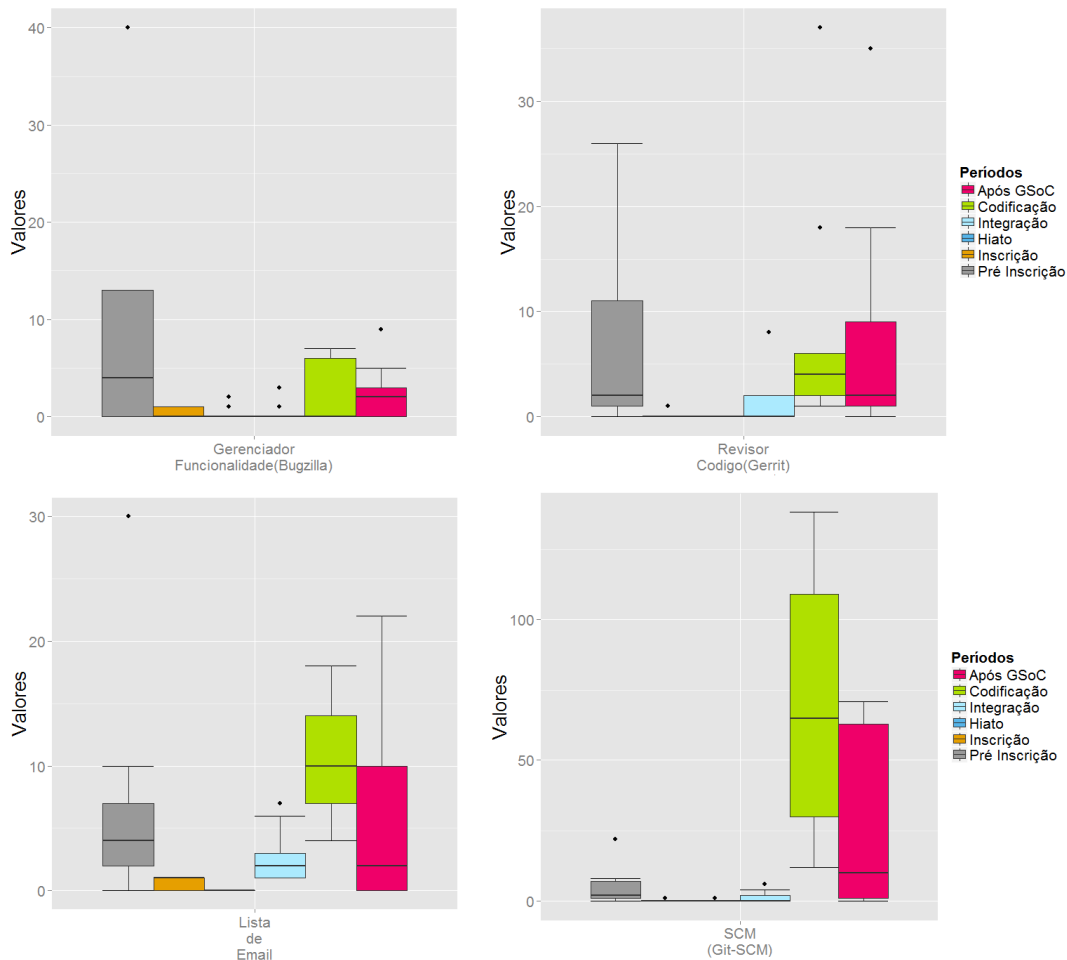
gráficos para comparação e observação. No gráfico de distribuição de interações dos voluntários com *outliers* na Figura 2 temos os dados da participação dos *outliers*. Já no gráfico de distribuição de interações dos voluntários, na Figura 3, os *outliers* foram removidos.



**Figura 2: Distribuição de interações dos voluntários com outliers**

Para analisar os números discrepantes apresentamos o gráfico de caixa (*boxplot*) da Figura 2, no qual dois voluntários aparecem com uma participação muito superior aos demais. A impressão inicial é de que os dados coletados dos outros voluntários não são expressivos para a pesquisa. No entanto, quando observamos apenas os participantes dentro da média, constatamos que tiveram uma participação relevante para a pesquisa, como representado no gráfico da Figura 3.

Pesquisando um pouco mais sobre esses dois voluntários em buscadores eletrônicos, como Google e LinkedIn, foi possível descobrir que na época ambos possuíam vínculo empregatício com uma empresa privada. A empresa contribui para o projeto LibreOffice e tem lucro fornecendo serviços especializados com ferramentas de Software Livre. Isso faz com que a interação com a comunidade seja a principal atividade de ambos os voluntários.



**Figura 3: Distribuição de interações dos voluntários sem outliers**

Diante dessas informações, optamos por não considerar os dados desses dois participantes nas análises. Caso contrário, a visualização dos gráficos acabaria mostrando uma participação ofuscada dos outros voluntários.

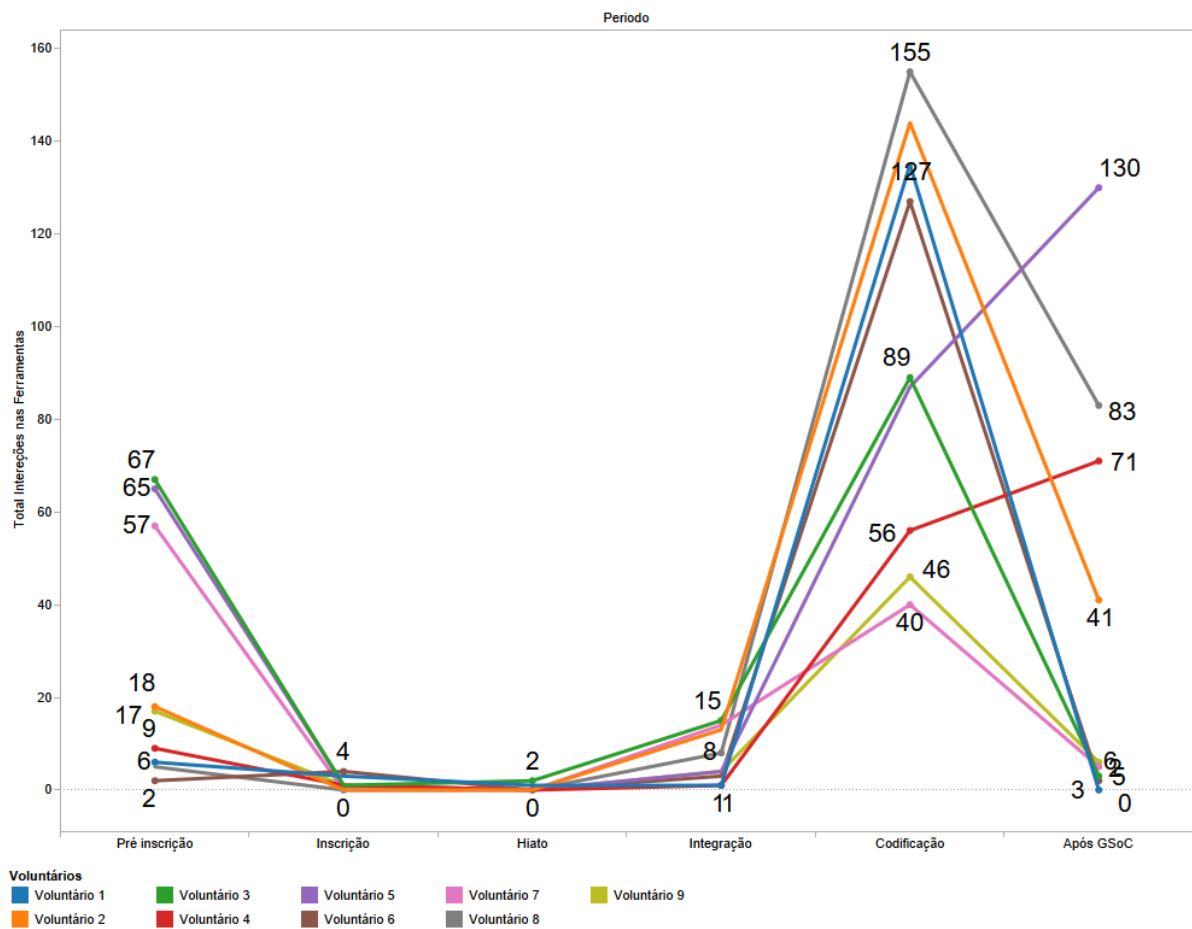
#### 4.3.1 ANÁLISE DAS INTERAÇÕES DOS VOLUNTÁRIOS

Para analisar o volume de interações de cada voluntário, consideramos cada participação como uma interação independente da plataforma utilizada. A Figura 4 ilustra o gráfico das interações dos voluntários nos diferentes períodos. Cada linha representa um voluntário que oscila entre os períodos do GSoC.

A Figura 4 ilustra a participação dos voluntários em cada ferramenta. Sugere um padrão nos diferentes períodos, começando com todos o voluntários da amostra participando na pré-inscrição. Nesta etapa, apenas um voluntário não participou interagindo com repositório. Isso sugere que os estudantes precisam se informar e interagir com a comunidade antes de

realizar as inscrições. Tal interação se mostra benéfica para que o projeto de desenvolvimento seja aceito, ou ainda, pode ser um requisito na escolha do voluntário pelos mentores.

Nas etapas de hiato e inscrição as interações dos voluntários caíram em todas as plataformas. Voltaram a subir somente na integração, quando os voluntários começaram a se comunicar com os mentores e a comunidade. Essa queda da participação pode ter ocorrido porque os voluntários ainda não sabiam se seriam aceitos pelos tutores, e aguardavam a resposta. Logo após a confirmação, começou o período de integração, no qual notamos o início das atividades.



**Figura 4: Total de interações por voluntário**

No período de codificação, quando os voluntários trabalharam desenvolvendo o código fonte do projeto proposto, observamos que todos exerceram participação ativa nas ferramentas de e-mail, revisão e no repositório de código fonte. Todos os voluntários atuaram de forma intensa nas três plataformas.

Por fim, o período de encerramento das atividades do GSoC, no qual notamos uma queda na participação nas ferramentas SCM, lista de e-mail e revisão de código. A este respeito, há alguns aspectos que analisaremos adiante.

Analisando o gráfico da Figura 5 notamos que os voluntários participaram ativamente nas lista de e-mail, possivelmente motivados pela necessidade de troca de informações iniciais necessárias para a integração e o desenvolvimento do projeto.

Quanto à interação por meio do repositório de código fonte, é interessante destacar que durante todos os períodos houve atividade. Aproximadamente 8 dos voluntários tinham uma participação efetiva anterior ao GSoC.

Destacamos os voluntários 3, 5 e 7, que no período de pré-inscrição têm uma interação de 67, 65 e 57, respectivamente. Números elevados para a média dos voluntários. Todavia o voluntário 5 manteve-se atuante no projeto, em contrapartida que os outros durante a evolução, podemos verificar que as interações foram abandonadas e as participações são mínimas após a conclusão do projeto. Isso indica que poderia haver um grande interesse em contribuir inicialmente, que não persistiu após o fim do GSoC.

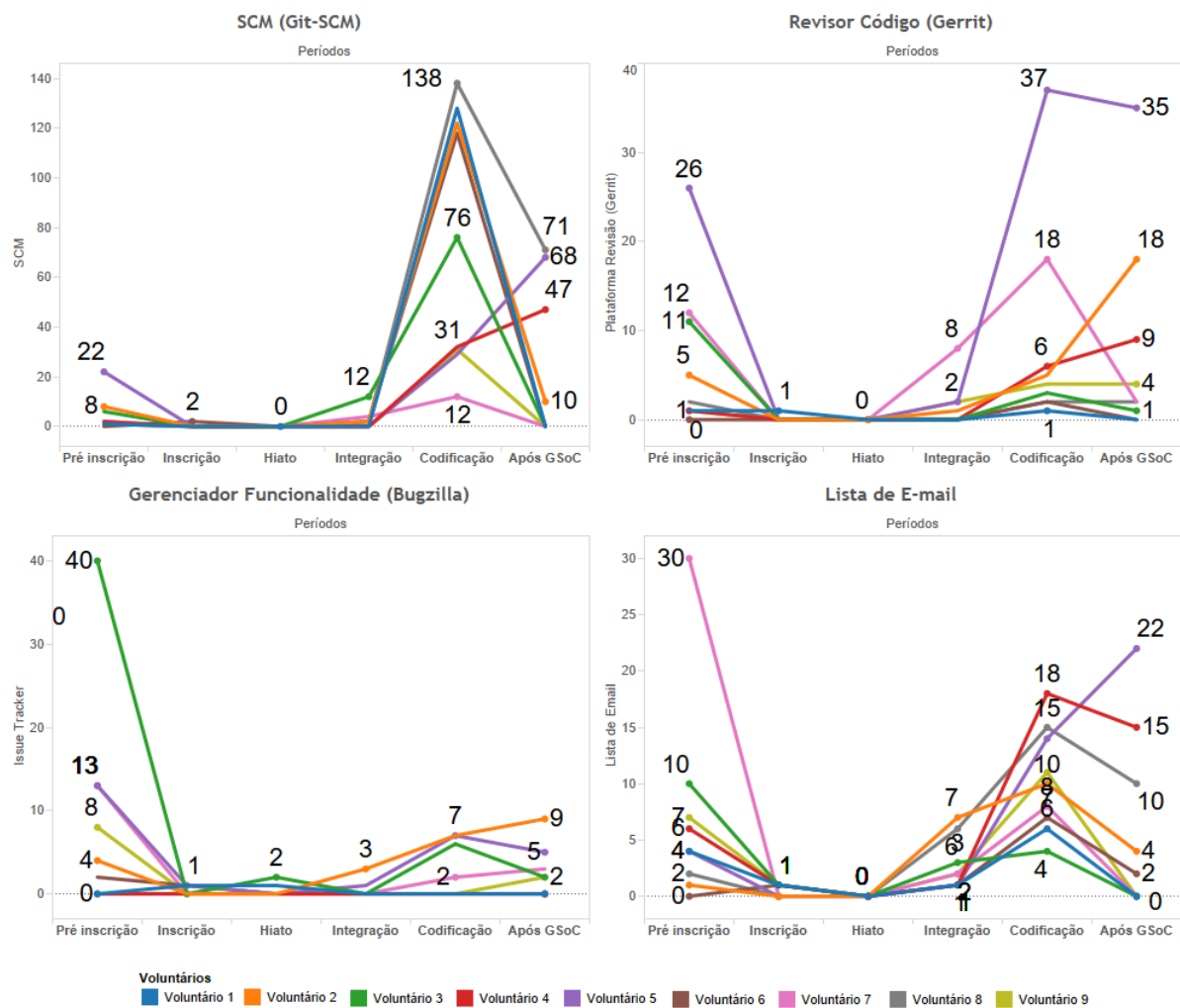


Figura 5: Total de interações na ferramentas por voluntário

O gráfico da Figura 5 é composto por quatro gráficos, cada um representando uma plataforma. O que torna possível observarmos a distribuição, entre as plataformas pesquisadas, da participação dos voluntários. As linhas representam os voluntários e assim podemos notar qual foi a distribuição das interações dos voluntários, nas diferentes plataformas.

Tivemos um período de queda na participação durante o hiato e um crescimento nas etapas subsequentes. Isso foi observado no que diz respeito a lista de e-mail, ao revisor de código, e principalmente, ao repositório de código fonte (SCM). Durante o período de codificação podemos notar um pico de participação, no qual os voluntários interagiram ativamente na comunidade por meio das suas ferramentas.

No período após o GSoC destacamos os voluntários 4 e 5, cujas interações mantêm uma crescente. Isso é um indicativo de que ambos demonstraram interesse em continuar a interagir com a comunidade, colaborando com a mesma. O voluntário 5 teve maior destaque por manter um padrão de interação mais elevado nas ferramentas de código fonte (SCM), revisor de código e lista de e-mail. Os índices demonstraram uma atuação efetiva deste voluntário.

Notamos que mesmo após o término do GSoC, os voluntários continuaram a participar de alguma forma do projeto. Executando uma busca nas mensagens das interações, constatamos que cinco voluntários da amostra ingressaram no período seguinte do GSoC, referente a 2014. Esse comportamento pode ser indicativo da existência de um interesse em contribuir. Contudo, observamos que os voluntários reingressantes voltaram a contribuir apenas na época do projeto seguinte.

Destacamos o voluntário 8, que no início do projeto teve uma atuação baixa: apenas uma vez no SCM, duas vezes na lista de email e na revisão de código. Verificamos que este voluntário teve um pico nas contribuições, sendo o que mais interagiu na ferramenta SCM. Continuou a contribuir mesmo após o fim da codificação, mantendo-se ativo na comunidade. Notamos aqui que esse voluntário começou com uma participação tímida e tornou-se um colaborador muito atuante. Inferimos que a sua contribuição tenha demandado algum tipo de revisão para melhoria ou correção de problemas.

Apuramos que a maior participação, em toda a amostra, concentrou-se nas plataformas de gerenciamento do código fonte (SCM) e lista de e-mail. Vale destacar que a forma como os voluntários trabalharam com a ferramenta SCM pode ter ajudado a inflar esses números. Por exemplo, a contribuição de um voluntário que apresentou um grande número de submissões, mas com adição e alteração de poucas linhas de código, teve o mesmo "valor" das contribuições dos demais voluntários. Isso porque a nossa proposta foi de analisar a quantidade de interações, deixando o estudo qualitativo para oportunidades futuras.

Dentre os cinco voluntários que não reingressaram no GSoC, somente o n<sup>o</sup> 8 voltou a interagir com o repositório e código fonte. O que nos indica que ele voltou a contribuir de alguma forma para a comunidade, sem o incentivo financeiro. Os demais voluntários fizeram uma quantidade pequena de interações após o fim do projeto, e nenhuma delas aconteceu no SCM.

#### 4.4 CONSIDERAÇÕES FINAIS

Na Tabela da Figura 6, estão dispostos em tabela os totais das interações de cada voluntário, antes e depois do GSoC de 2013. Deste modo, observamos o progresso da contribuição dos voluntários nas ferramentas. Nos gráficos da Figura 7 podemos visualizar em barras empilhadas os totais da interação, onde as barras inferiores representam o período de pré-inscrição e as barras superiores, o período após o GSoC. Os voluntários foram divididos em dois grupos: os que voltaram a participar do GSoC na edição de 2014 e os que não voltaram a participar.

Plataforma	Voluntários / Períodos																	
	Voluntário 1		Voluntário 2		Voluntário 3		Voluntário 4		Voluntário 5		Voluntário 6		Voluntário 7		Voluntário 8		Voluntário 9	
	Pré inscrição	Após GSoC	Pré inscrição	Após GSoC	Pré inscrição	Após GSoC	Pré inscrição	Após GSoC	Pré inscrição	Após GSoC	Pré inscrição	Após GSoC	Pré inscrição	Após GSoC	Pré inscrição	Após GSoC	Pré inscrição	Após GSoC
Issue Tracker	0	0	4	9	40	2	0	0	13	5	2	0	13	3	0	0	8	2
Lista de Email	4	0	1	4	10	0	6	15	4	22	0	2	30	0	2	10	7	0
Plataforma Revisão (Gerrit)	1	0	5	18	11	1	1	9	26	35	0	0	12	2	2	2	1	4
SCM	1	0	8	10	6	0	2	47	22	68	0	0	2	0	1	71	1	0

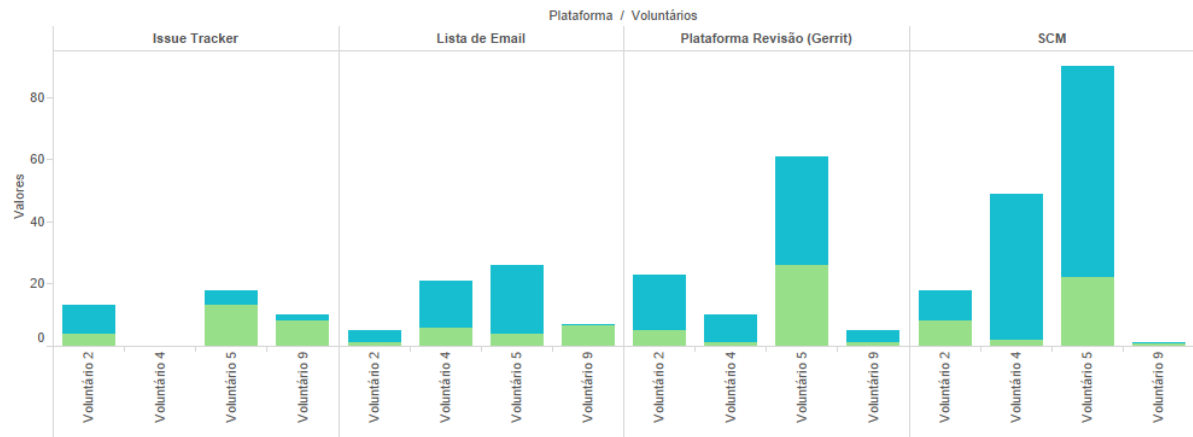
**Figura 6: Tabela da interação dos voluntários antes e após o Google Summer of Code**

Aduzimos pesquisando nos dados do GSoC e nas mensagens contidas nas ferramentas que 4 voluntários participaram do GSoC de 2014 e 5 não voltaram a participar. Percebemos que a maioria das interações após o GSoC é dos reingressantes. Isso nos leva a inferir que os voluntários que fizeram contribuições estavam interagindo devido a influência do GSoC 2014.

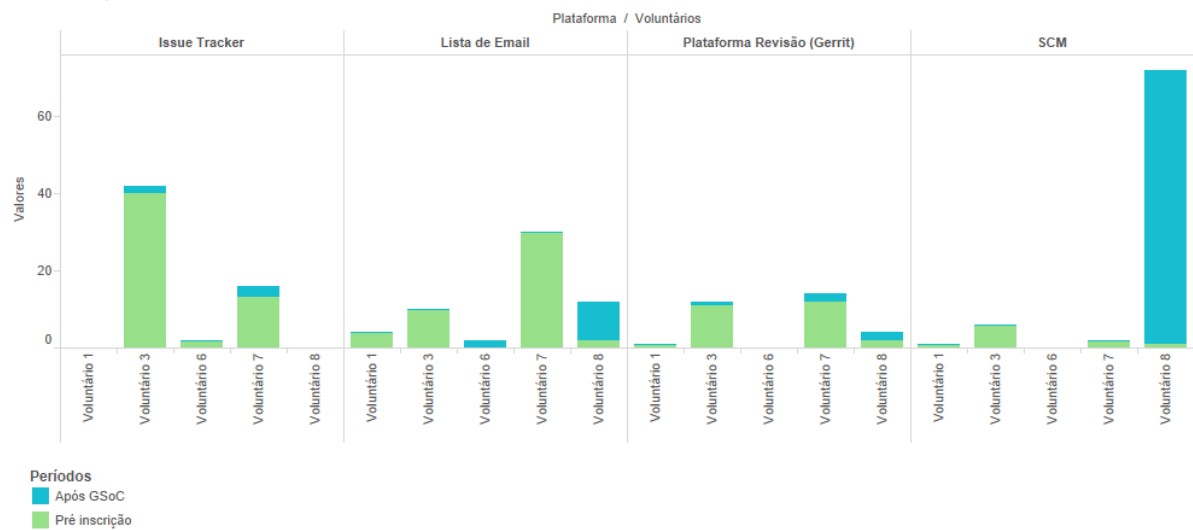
Destacamos novamente o voluntário 8, que mesmo não sendo reingressante, interagiu com as ferramentas. Principalmente no SCM e na lista de e-mail, como já mencionamos. Os demais voluntários não contribuíram, ou, se o fizeram, foi muito pouco se comparado ao grupo dos reingressantes.

Os dados indicam que o programa GSoC conseguiu, na comunidade do LibreOffice, recrutar voluntários e gerar contribuições de funcionalidades novas, desenvolvidas em um espaço

**PARTICIPAÇÃO DOS VOLUNTÁRIOS REINGRESSANTES NO GSoC 2014**



**PARTICIPAÇÃO DOS VOLUNTÁRIOS NÃO REINGRESSANTES NO GSoC 2014**



**Figura 7: Interação dos voluntários por ferramentas**

curto de tempo. Observamos que a maioria dos voluntários continuou a colaborar com a comunidade e que muitas dessas contribuições têm relação direta com a edição subsequente do GSoC, em 2014. O saldo de contribuições dos voluntários da amostra cresceu, quando comparamos antes e depois do GSoC.

## 5 CONCLUSÕES

Acreditamos que, no tocante ao projeto de Software Livre LibreOffice, esta pesquisa representa uma contribuição significativa para o estudo da eficiência do incentivo financeiro.

Os dados sugerem que o projeto foi satisfatório e cumpriu com seu papel de incentivar o ingresso de novos voluntários na amostra estudada. Mesmo que alguns tenham uma participação anterior ao projeto, os números indicam um aumento na participação dos voluntários. Dos 9 voluntários considerados na amostra, 5 aumentaram sua contribuição. Destes, 4 não voltaram a contribuir com o projeto ou fizeram pequenas interações.

Podemos notar um comportamento que se repete entre os voluntários, durante todo o GSoC. Antes do início das atividades, os voluntários já interagem fazendo colaborações, possivelmente para ganhar credibilidade com os mentores ou ainda, para adquirirem alguma experiência sobre o projeto. Na sequência, durante o período de escolha os projetos para a bolsa de estágio, ficam aguardando o resultado e quase não interagem nas ferramentas. Somente após a divulgação do resultado ocorre uma grande interação por parte dos voluntários, que atuam em todas as ferramentas, principalmente no SCM e lista de e-mail.

Observamos que alguns voluntários interagiram após o fim do GSoC e continuaram na comunidade. Da amostra total de 11 voluntários (considerando os *outliers*), 5 participaram das edições de 2013 e 2014. O que nos leva a acreditar que essas interações pós projeto, podem ter sido motivadas pelo interesse em pleitear as bolsas do ano seguinte, e que a experiência adquirida atuou como facilitador para o reingresso do voluntário. Também evidencia um comportamento centrado na recompensa. Do todo, apenas o voluntário 8 apresentou índices elevados de contribuição sem ter participado da edição seguinte.

Porem temos que considerar que apenas o pagamento da bolsa de estágio, enquanto fator de incentivo inicial para uma colaboração de longo prazo na comunidade, não é determinante para os voluntários novatos continuarem a contribuir sem o incentivo extrínseco na comunidade de Software Livre. Provavelmente, outros fatores que nos são desconhecidos e fogem ao alcance dessa pesquisa, são influentes e determinantes para um voluntário continuar



a colaborar.

Para finalizar, as considerações a respeito da comunidade estudada, LibreOffice, temos algumas considerações em relação as questões de pesquisa:

Os participantes do GSoC tinham envolvimento anterior com o projeto?

Foi notado que ocorreu uma participação antes das inscrições, denotando um comportamento que se repetiu na maioria dos participantes. Presumimos que os mentores tem uma inclinação em escolher para o estágio aqueles que já colaboram com o projeto.

Que atividade desenvolveram durante o GSoC?

Todas as ferramentas foram utilizadas pelos voluntários. O período de codificação foi o que teve maior atividade em todas as ferramentas. A ferramenta com maior atividade dentre as quatro, foi a gerencia de código fonte Git-SCM, mostrando que todos os voluntários atuaram codificando para desenvolver o projeto.

As contribuições realizadas pelos participantes do GSoC foram aceitas pelo projeto?

Como denotamos todos os voluntário atuaram nas ferramentas, e nos arquivos do projeto do GSoC todos os projeto tiveram seus códigos submetidos as mentores, embora essas duas plataformas nos confirmem que os voluntários interagiram desenvolvendo o código fonte, só poderíamos confirmar se a contribuição foi aceita na versão final do LibreOffice, executando uma pesquisa qualitativa no repositório Git-SCM.

Qual o índice de retenção dos bolsistas após o término do projeto?

Dentro a amostra de 11 estudantes, 5 interagiram com a comunidade e também participaram da próxima edição do GSoC 2014. Dentre os outros 6 estudantes, apenas um voluntário interagiu significativa como o código fonte. Notamos que a maioria dos voluntários que contribuíram com o projeto após o fim das ativadas foi decorrente da sua experiência com o GSoC, vindo a se repetir na próxima edição.

## 5.1 LIMITAÇÕES

Estamos cientes que esta pesquisa tem algumas limitações, tão importantes que devem ser consideradas.

Analizamos apenas voluntários do projeto LibreOffice, uma comunidade que tem as suas próprias características. Talvez uma amostra mais ampla entre projetos diversificados revele números diferentes dos que encontramos.

Os dados que investigamos foram obtidos somente das ferramentas e dos buscadores disponíveis. Avaliamos que uma abordagem mais direta aos voluntários, por meio de questionários, seja interessante para revelar qual o seu interesse pessoal no âmbito das comunidades de Software Livre.

## 5.2 TRABALHOS FUTUROS

Para futuras pesquisas, entendemos que seja possível diversificar as comunidades participantes do GSoC. Isso poderá conferir um perfil mais amplo e variado de colaboradores. Outro ponto a se considerar é estender o período de pesquisa do GSoC, englobando outras edições.

Entendemos que a comunidade do LibreOffice poderia ter todos os membros participantes do GSoC como foco de pesquisa semelhante, haja vista que muitos participam mais de uma vez. Além disso, como os repositórios possuem grandes quantidade de informações armazenadas, tornam-se uma rica fonte de dados. Apesar desta tarefa demandar um árduo trabalho de mineração, poderá integrar projetos futuros, dada a sua relevância para a sustentabilidade das comunidades de Software Livre.

**APÊNDICE A – TABELA DAS PARTICIPAÇÕES DOS VOLUNTÁRIOS**

Voluntários	Período	Gerenciador Funcionalidade (Bugzilla)					Lista de Email	Revisor Código (Gerrit)			SCM (Git-SCM)		
		Issues Submetida	Issues Atribuídas	Issues Comentada	Issues Assistidas	Total		Owner	Reviewer	Total	Committer	Autor	Total
1	Pré inscrição	0	0	0	0	0	4	1	0	1	0	1	
	Inscrição	0	0	1	0	1	1	1	0	1	0	0	
	Hiato	0	0	1	0	1	0	0	0	0	0	0	
	Integração	0	0	0	0	0	1	0	0	0	0	0	
	Codificação	0	0	0	0	0	6	1	0	1	65	63	128
Após GSoC	0	0	0	0	0	0	0	0	0	0	0	0	
2	Pré inscrição	0	0	0	4	4	1	5	0	5	0	8	8
	Inscrição	0	0	0	0	0	0	0	0	0	0	0	0
	Hiato	0	0	0	0	0	0	0	0	0	0	0	0
	Integração	0	0	2	1	3	7	0	1	1	0	2	2
	Codificação	0	0	5	2	7	10	3	2	5	77	45	122
Após GSoC	0	0	9	0	9	4	4	14	18	9	1	10	
3	Pré inscrição	2	6	20	12	40	10	10	1	11	0	6	6
	Inscrição	0	0	0	0	0	1	0	0	0	0	0	0
	Hiato	0	0	2	0	2	0	0	0	0	0	0	0
	Integração	0	0	0	0	0	3	0	0	0	6	6	12
	Codificação	0	2	2	2	6	4	1	2	3	40	36	76
Após GSoC	0	0	0	2	2	0	0	1	1	0	0	0	
4	Pré inscrição	0	0	0	0	0	6	1	0	1	0	2	2
	Inscrição	0	0	0	0	0	1	0	0	0	0	0	0
	Hiato	0	0	0	0	0	0	0	0	0	0	0	0
	Integração	0	0	0	0	0	1	0	0	0	0	0	0
	Codificação	0	0	0	0	0	18	4	2	6	0	32	32
Após GSoC	0	0	0	0	0	15	6	3	9	29	18	47	
5	Pré inscrição	7	3	2	1	13	4	23	3	26	0	22	22
	Inscrição	0	0	1	0	1	0	0	0	0	0	0	0
	Hiato	0	0	0	0	0	0	0	0	0	0	0	0
	Integração	0	0	0	1	1	1	2	0	2	0	0	0
	Codificação	0	0	2	5	7	14	26	11	37	12	17	29
Após GSoC	0	0	0	5	5	22	27	8	35	29	39	68	
6	Pré inscrição	0	1	0	1	2	0	0	0	0	0	0	0
	Inscrição	0	0	1	0	1	1	0	0	0	1	1	2
	Hiato	0	0	0	0	0	0	0	0	0	0	0	0
	Integração	0	0	0	0	0	1	0	0	0	0	2	2
	Codificação	0	0	0	0	0	7	2	0	2	60	58	118
Após GSoC	0	0	0	0	0	0	2	0	0	0	0	0	
7	Pré inscrição	0	3	10	0	13	30	10	2	12	0	2	2
	Inscrição	0	0	0	0	0	0	0	0	0	0	0	0
	Hiato	0	0	0	0	0	0	0	0	0	0	0	0
	Integração	0	0	0	0	0	2	8	0	8	0	4	4
	Codificação	0	0	1	1	2	8	10	8	18	0	12	12
Após GSoC	0	1	0	2	3	0	1	1	2	0	0	0	
8	Pré inscrição	0	0	0	0	0	2	2	0	2	0	1	1
	Inscrição	0	0	0	0	0	0	0	0	0	0	0	0
	Hiato	0	0	0	0	0	0	0	0	0	0	0	0
	Integração	0	0	0	0	0	6	0	0	0	0	2	2
	Codificação	0	0	0	0	0	15	2	0	2	135	3	138
Após GSoC	0	0	0	0	0	10	1	1	2	33	38	71	
9	Pré inscrição	0	2	6	0	8	7	1	0	1	0	1	1
	Inscrição	0	0	0	0	0	1	0	0	0	0	0	0
	Hiato	0	0	0	0	0	0	0	0	0	0	0	0
	Integração	0	0	0	0	0	2	1	1	2	0	0	0
	Codificação	0	0	0	0	0	11	2	2	4	0	31	31
Após GSoC	0	0	2	0	2	0	2	2	4	0	0	0	
10	Pré inscrição	0	4	11	19	34	57	0	0	0	95	76	171
	Inscrição	0	0	0	1	1	0	0	0	0	0	0	0
	Hiato	0	2	3	4	9	1	0	0	0	1	4	5
	Integração	0	0	1	0	1	12	5	0	5	4	2	6
	Codificação	0	2	20	11	33	53	29	22	51	433	66	499
Após GSoC	0	6	26	27	59	18	116	62	178	687	242	929	
11	Pré inscrição	0	7	2	3	12	11	3	5	8	107	50	157
	Inscrição	0	0	0	0	0	0	0	0	0	5	0	5
	Hiato	0	1	0	2	3	0	0	0	0	2	2	4
	Integração	0	2	8	4	14	10	0	1	1	12	1	13
	Codificação	0	0	9	5	14	29	4	0	4	83	8	91
Após GSoC	0	5	26	5	36	26	130	109	239	596	300	896	

**Figura 8: Tabela**

## REFERÊNCIAS

- ALEXU, O.; LEITNER, M. A fistful of dollars: Financial rewards, payment norms, and motivation crowding in open source software development. **Payment Norms, and Motivation Crowding in Open Source Software Development**, 2010.
- DAVIES, J.; ZHANG, H.; NUSSBAUM, L.; GERMAN, D. M. Perspectives on bugs in the debian bug tracking system. In: IEEE. **Mining Software Repositories (MSR), 2010 7th IEEE Working Conference on**. Cape Town, 2010. p. 86–89.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery in databases. **AI magazine**, v. 17, n. 3, p. 37, 1996.
- GOLDMAN, R.; GABRIEL, R. P. San Francisco: Morgan Kaufmann, 2005.
- GOOGLE. **Google Summer Of Codes Melange**. @ONLINE. 2015. Data último acesso em 26/11/2015. Disponível em: <<http://www.google-melange.com/>>.
- GOOGLE. **Google Summer Of Codes**. @ONLINE. 2015. Data último acesso em 26/11/2015. Disponível em: <<https://developers.google.com/open-source/soc/?csw=1>>.
- HIPPEL, E. v.; KROGH, G. v. Open source software and the “private-collective” innovation model: Issues for organization science. **Organization science**, INFORMS, v. 14, n. 2, p. 209–223, 2003.
- KRISHNAMURTHY, S.; OU, S.; TRIPATHI, A. K. Acceptance of monetary rewards in open source software development. **Research Policy**, Elsevier, v. 43, n. 4, p. 632–644, 2014.
- KROGH, G. V.; HAEFLIGER, S.; SPAETH, S.; WALLIN, M. W. Carrots and rainbows: Motivation and social practice in open source software development. **Mis Quarterly**, v. 36, n. 2, p. 649–676, 2012.
- SCHILLING, A.; LAUMER, S.; WEITZEL, T. Who will remain? an evaluation of actual person-job and person-team fit to predict developer retention in floss projects. In: IEEE. **System Science (HICSS), 45th Hawaii International Conference on**. Maui, HI, 2012. p. 3446–3455.
- SPAETH, S.; KROGH, G. von; HE, F. Research note—perceived firm attributes and intrinsic motivation in sponsored open source software projects. **Information Systems Research**, INFORMS, v. 26, n. 1, p. 224–237, 2014.
- STEINMACHER, I.; CHAVES, A. P.; CONTE, T. U.; GEROSA, M. A. Preliminary empirical identification of barriers faced by newcomers to open source software projects. In: IEEE. **Software Engineering (SBES), Brazilian Symposium on**. Maceio, 2014. p. 51–60.
- STEINMACHER, I.; CONTE, T.; GEROSA, M. A.; REDMILES, D. Social barriers faced by newcomers placing their first contribution in open source software projects. In: **Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & #38; Social**

**Computing**. New York, NY, USA: ACM, 2015. (CSCW '15), p. 1379–1392. ISBN 978-1-4503-2922-4. Disponível em: <<http://doi.acm.org/10.1145/2675133.2675215>>.

STEINMACHER, I.; WIESE, I. S.; CHAVES, A. P.; GEROSA, M. A. Newcomers withdrawal in open source software projects: Analysis of hadoop common project. In: IEEE. **Collaborative Systems (SBSC), 2012 Brazilian Symposium on**. São Paulo, 2012. p. 65–74.

TRAINER, E. H.; CHAIHIRUNKARN, C.; KALYANASUNDARAM, A.; HERBSLEB, J. D. Community code engagements: Summer of code & hackathons for community building in scientific software. In: ACM. **Proceedings of the 18th International Conference on Supporting Group Work**. New York, NY, USA, 2014. p. 111–121.