

Visualização de Dados – Relatório Trabalho Final

Danilo Ferreira, Guilherme Avelino, Hudson Borges

¹Departamento de Ciência da Computação, UFMG

{danilofs, gaa, hsborges}@dcc.ufmg.br

Abstract. *Nesse trabalho desenvolvemos um conjunto de visualizações com o intuito de auxiliar na identificação de especialistas de código em sistemas de software. A identificação do autor/especialista em artefato de código é uma atividade importante para auxílio na manutenção e evolução de software.*

1. Introdução

Nos últimos anos o GitHub se tornou uma poderosa ferramenta de colaboração para desenvolvimento de software. Atualmente é o maior repositório de código do mundo, com mais de 6.8 milhões de colaboradores e 15.2 milhões de repositórios armazenados. Muito de sua popularidade se deve a características como controle de versão distribuído e foco em *social coding* [1].

No trabalho aqui apresentado, buscamos desenvolver visualizações que auxiliem na identificação de especialistas em artefatos de software. As visualizações aqui apresentadas foram construídas com objetivo de analisar um conjunto de dados extraídas de sistemas Open Source extraídos do GitHub¹.

2. Interface da página

A primeira visualização a ser destacada neste trabalho é o *layout* adotado para a apresentação de todas as visualizações. O *layout* adotado por todas as páginas foi desenvolvido para se adaptar a diferentes tamanhos de telas (*responsive design*), possui uma interface limpa com conteúdos bem distribuídos e tem como base novas tecnologias *web* (e.g., as páginas não são recarregadas ao navegar no *website*). Mais especificamente, a interface da página possui três componentes principais:

Menu de Navegação: O menu de navegação encontra-se posicionado à esquerda na página e lista todas as páginas acessíveis. Cada opção do menu apresenta um título e um ícone indicativo, contudo os títulos podem ser escondidos para aumentar o espaço de apresentação das visualizações. Essa funcionalidade é bastante útil em dispositivos móveis e/ou computadores que possuem telas menores, pois possibilita um melhor aproveitamento na área de apresentação das visualizações.

Menu superior: O menu superior consiste de uma pequena área localizada na parte superior da página que permite ao usuário se localizar e também apresenta a última data de atualização dos dados. Esse último detalhe é bastante importante pois as visualizações

¹<https://github.com/>

apresentadas são resultados de análises estáticas e que podem sofrer alterações durante o tempo, logo é importante deixar claro aos usuário à quando os resultados se referem.

Área de apresentação: Por fim, a área de apresentação dos dados é a maior região da página e apresenta o conteúdo da página visitada, sendo atualizada (sem necessidade de recarregamento completo) cada vez que o usuário navega pelas páginas disponíveis. Um detalhe importante está relacionada à biblioteca de desenho de gráficos utilizada, o D3.JS. Os gráficos desenhados por tal biblioteca não apresentam funcionalidades responsivas por padrão, logo, apesar da página possuir tais funcionalidades, quando um gráfico é desenhado na tela ele não é redimensionado automaticamente, para isso é necessário que o próprio usuário recarregue a página para que a biblioteca identifique e desene o gráfico de acordo com as novas dimensões.

3. Visualizações

Nas próximas subseções são apresentados detalhes sobre as principais visualizações implementadas.

3.1. Dashboard

O dashboard é a primeira visualização apresentada quando os usuários visitam o *website*. Nessa interface são apresentadas informações de alto nível que tem por objetivo fazer com o que os usuários tenham uma visão geral do nosso objetivo e os dados que utilizamos como base em nossa pesquisa. Partindo da premissa que a leitura é feita da esquerda para a direita e de cima para baixo, inicialmente apresentamos os dados referentes ao número de repositórios, desenvolvedores, *commits* e arquivos que foram analisados. Em seguida apresentamos uma tabela um pouco mais detalhada destes mesmos dados mas por linguagem. Por fim, apresentamos um gráfico de barras com a participação dos melhores autores por repositório.

3.2. Best Authors

...

3.3. Dev Multi authors

...

3.4. Workload

...

3.5. Workload - Multi authors

...

3.6. Distribution

...

Referências

- [1] L. Dabbish, C. Stuart, J. Tsay, and J. Herbsleb. Social Coding in GitHub: Transparency and Collaboration in an Open Software Repository. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, pages 1277–1286, 2012.