Midterm demo visualizations

This document describes the visualizations that will be shown in the midterm demo on May 27. The descriptions provide both technical details, like the type of graph and the axis, and motivations to explain why these graphs are interesting to the Octopeer plugin user.

The goal of Octopeer Analytics at this point is to provide an individual peer reviewer useful visualizations to gain insight in their reviews, so the proposed visualizations all focus on individual statistics. This should make it attractive to install the Octopeer browser plugin for developers.

Glossary

- Closed pull request: a pull request that has been closed, without having been merged.
- Semantic session: a session that is indicated in the Octopeer database by semantic events (event types 401 and 402 for a start and finish). It means the pull request is being opened in the browser and closed again. A peer review may consist of multiple semantic sessions.

1. Comments on pull requests

Description

Visualization type: bar chart (histogram).
X-axis: number of comments of the user.
y-axis: number of reviewed pull requests.

The bars are vertical and splitted in a open, merged and closed part, indicating how many pull requests in the category have been successful.

Motivation

This visualization shows the connection the number of comments submitted by the user, as a peer reviewer, and the successfulness of the corresponding peer reviews. It should be interesting to see what influence your comments generally have, and whether there is a optimal number of comments indicating a pull request may be merged.

2. Duration of peer reviews

Description

Visualization type: combined bar and line chart.
x-axis: peer reviews of the user.
y-axis: duration of the peer review sessions.

The graph show 20 vertical bars, representing the most recent peer reviews of the user. It does not include pull requests that are opened by the user. A bar consist of several parts when the user has multiple semantic sessions, splitted in a subtle way. Again, the status of the pull requests are indicated with different colors. A pull request closed or merged by the user is indicated with an icon above the bar.

The line indicates the number of changed lines per pull request that is being reviewed.
In the next weeks, it maybe possible to scroll back through this graph, loading older peer reviews.

Motivation

Peer reviewers will seek to improving their reviews both qualitatively and quantitatively. This means doing as much useful reviewing in as little time as possible. This visualization will show the user's recent peer reviews. The line will provide a way to show the relation between the size of the pull request, measured in changed lines, and the time the user has spent reviewing it.

The division of the bars is created to provide insight in coming back to the pull request. After a first review session, the reviewer may come back and assess the code again. This can be because he or she didn’t finish the entire review in the first session, or to view commits that have been pushed after the first time. The bars indicate how much time is spent in each semantic session.

3. Peer review timeline

Description

Visualization type: line chart.
x-axis: time spent (either absolute or relative).
y-axis: peer reviews.

The graph consists of 10 horizontal lines, each for one of the latest peer reviews, indicating how much time per activity type (eg reading code, commenting, viewing CI results) is spent per review. A merge is indicated with an icon.

The user can switch between a representation of time in an absolute or a relative scale. In the absolute chart, the length of the lines will usually differ, because the peer review did not take (exactly) the same time. In the relative chart, all lines will be stretched to the full width of the graph.

Motivation

The previous visualization (2) provides insight in the general *length* of peer reviews and how the different review *sessions* of each pull request are divided. This visualization shows how the peer reviews are divided in *activity types*. This will show developers how many time they usually spend on reading code, reading comments, writing comments, browsing other sites, etc. This will also provide a tool to assess the user's recent reviews. For example, if a user notices he looked very shortly at a recent pull requests code, he may wonder why that was.

A switch between absolute or relative time is considered useful, because both graphs will be interesting to look at. The absolute timelines create a stronger emphasis on bigger (longer) reviews, and can be used to compare the total time spent on different review sessions. The relative timelines will provide a clearer picture of shorter reviews, and give insight in the percentage of time spent on the different aspects of the peer reviews.

4. Element types - event type

Description

Visualization type: matrix diagram (like https://bost.ocks.org/mike/miserables/)
x-axis: element types.
y-axis: event types.

The x-axis displays all semantic element types in the Octopeer database, like code blocks and comment boxes. The y-axis displays all semantic event types, like scrolling in/out, clicking, and mouse hovering.

Motivation

This graph shows the user's activity during peer reviews at an abstract level. It does not focus on specific review, but indicates which sections of the pull request pages are being viewed and interacted with.

5. Long-term trends

Description

Visualization type: line graph.
x-axis: time, from start of Octopeer plugin usage to present.
y-axis: performance indicators.

This visualisation contains multiple lines indicating the trends of the user's review behaviour, from the time he or she starts using the Octopeer plugin to the present time. The following data will be included in different lines:

- Number of comments;
- Total time spent;
- Time spent on looking at code;
- Number of semantic review sessions.

This is relative to the number of peer reviews in a given timespan; this may be grouped per month.

Motivation

A peer reviewer will be interested in his personal development over time. This graph will show his behavioural trends over the greatest available timespan, which will be the time he or she uses the Octopeer plugin. The mentioned data, distributed over the lines, should represent the most important performance indicators of the reviews.

6. (Productive) peer review times

Description

Visualization type: scatterplot (comparable with the GitHub Punch Card).
x-axis: time of each day.
y-axis: days of the week.

This plot indicates.
Points will be distributed in the following way:

- Placing a comment: 100 points;
- Looking at comments: 10 points for each comment viewed;
- Looking at code: 1 for each changed line viewed;
- Looking at CI: 1 for each line viewed.

The division of this points is based on assumptions. For the final product, this has to be adjusted to and backed with available publications and our own research.

Motivation

The goal of this graph is to give an overview of the user's productivity in peer reviews, compared to the time of the week. Because there are points awarded in a given timespan, it shows at what times the reviews are most effective. This may help the user to plan his week and optimize his productivity.