Managing community backlog with the help of metrics

CHAOSScon Europe 2020
Alberto Pérez García-Plaza (Bitergia): @alpgarcia
Ray Paik (GitLab): @rspaik
Agenda

• Responding to community contributions
• How are we looking at the community backlog for Merge Requests (MRs) at GitLab?
• GrimoireLab Backlog implementation
• GrimoireLab: Use case analysis
• Evolution of GrimoireLab dashboard
Responding to community contributions
One way of looking at responsiveness:
Number of days that merge requests are open

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Merge Requests</th>
<th>Submitters</th>
<th>Repositories</th>
<th>Median Open Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.3</td>
<td>236</td>
<td>112</td>
<td>18</td>
<td>3.58</td>
</tr>
<tr>
<td>12.4</td>
<td>134</td>
<td>78</td>
<td>12</td>
<td>2.9</td>
</tr>
<tr>
<td>12.5</td>
<td>182</td>
<td>94</td>
<td>22</td>
<td>2.855</td>
</tr>
<tr>
<td>12.6</td>
<td>248</td>
<td>100</td>
<td>17</td>
<td>4.4</td>
</tr>
<tr>
<td>12.7</td>
<td>82</td>
<td>49</td>
<td>18</td>
<td>3.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Repository</th>
<th>Merge Requests</th>
<th>Submitters</th>
<th>Reviewers</th>
<th>Avg. Open Days</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="https://gitlab.com/www-gitlab-com">https://gitlab.com/www-gitlab-com</a></td>
<td>37,848</td>
<td>2,084</td>
<td>274</td>
<td>13.046</td>
</tr>
<tr>
<td><a href="https://gitlab.com/gitlab-org/gitlab">https://gitlab.com/gitlab-org/gitlab</a></td>
<td>18,361</td>
<td>1,026</td>
<td>111</td>
<td>14.364</td>
</tr>
<tr>
<td><a href="https://gitlab.com/gitlab-org/omnibus-gitlab">https://gitlab.com/gitlab-org/omnibus-gitlab</a></td>
<td>3,842</td>
<td>563</td>
<td>29</td>
<td>29.254</td>
</tr>
<tr>
<td><a href="https://gitlab.com/gitlab-org/gitlab-runner">https://gitlab.com/gitlab-org/gitlab-runner</a></td>
<td>1,751</td>
<td>591</td>
<td>14</td>
<td>113.999</td>
</tr>
<tr>
<td><a href="https://gitlab.com/gitlab-org/gitlab">https://gitlab.com/gitlab-org/gitlab</a></td>
<td>1,739</td>
<td>83</td>
<td>17</td>
<td>10.117</td>
</tr>
<tr>
<td><a href="https://gitlab.com/meltano/meltano">https://gitlab.com/meltano/meltano</a></td>
<td>1,312</td>
<td>38</td>
<td>12</td>
<td>2.763</td>
</tr>
<tr>
<td><a href="https://gitlab.com/gitlab-org/charts/gitlab">https://gitlab.com/gitlab-org/charts/gitlab</a></td>
<td>1,098</td>
<td>149</td>
<td>11</td>
<td>9.315</td>
</tr>
<tr>
<td><a href="https://gitlab.com/gitlab-org/release-tools">https://gitlab.com/gitlab-org/release-tools</a></td>
<td>802</td>
<td>71</td>
<td>15</td>
<td>11.714</td>
</tr>
</tbody>
</table>
We also wanted to look at the backlog as the community grew...
Looking at community backlog
You may remember me from talks such as "Creating a Collection of Panels" or "I have a GrimoireLab dashboard. Now, what?"

Member of Bitergia Consultancy Team
@alpgarcia
GrimoireLab Backlog Implementation

GL Backlog dashboards are based on what remains open.
GrimoireLab Backlog Implementation

How many items created during the last 90 days are still open?

[Image of a dashboard showing backlog details]
GrimoireLab Backlog Implementation

These items were created on these dates...and remain open.

1,426 Open Merge Requests

21,400.81 Accumulated Open Days

15,008 Avg. Time Open Days per Merge Request
GrimoireLab: Use Case Analysis

Limitations. It doesn’t allow analyzing the backlog:

- **Backlog evolution over time.**
- **Backlog status at specific time frames.**
GrimoireLab: Use Case Analysis

In short, offers a view based on the present status:

How many items created during the last 90 days are still open?
GrimoireLab: Use Case Analysis

GitLab user #1
I saw we are progressing closing tasks

GitLab user #2
But the filter only shows open tasks at the moment
GrimoireLab: Use Case Analysis

You told us you were closing tasks

But not when

GitLab users

Me (hard wip)
GrimoireLab: Use Case Analysis

Then, how can we look for past MRs state?
GrimoireLab: Use Case Analysis

- Current approaches:
  - We relied on Timelion to solve this with current data model:
    - Each Item contains **creation** and **closing/merging** dates.
    - We need to count it between both dates.
  - A recent community proposed approach is based on a new study and index:
    - [https://github.com/chaoss/grimoirelab-elk/issues/761](https://github.com/chaoss/grimoirelab-elk/issues/761)
    - New index with number of opened issues per project, origin and labels.
GrimoireLab: Use Case Analysis

- Current approaches:
  - **We relied on Timelion** to solve this with *current data model*:
    - Each Item contains **creation** and **closing/merging** dates.
    - We need to count it between both dates.
  
  - A recent **community proposed approach** is based on a *new study and index*:
    - [https://github.com/chaoss/grimoirelab-elk/issues/761](https://github.com/chaoss/grimoirelab-elk/issues/761)
    - New index with number of opened issues per project, origin and labels.
GrimoireLab: Use Case Analysis

- Challenges:
  - We relied on Timelion to solve this with current data model:
    - Each Item contains **creation** and **closing/merging** dates.
    - We need to count it between both dates.

<table>
<thead>
<tr>
<th>Time</th>
<th>merged_at</th>
<th>closed_at</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 15th 2020, 14:23:59.538</td>
<td>January 15th 2020, 14:56:22.563</td>
<td>-</td>
</tr>
<tr>
<td>January 15th 2020, 14:04:33.270</td>
<td>January 15th 2020, 14:33:00.331</td>
<td>-</td>
</tr>
<tr>
<td>January 15th 2020, 14:01:18.951</td>
<td>January 15th 2020, 14:55:59.688</td>
<td>-</td>
</tr>
<tr>
<td>January 15th 2020, 13:47:31.656</td>
<td>January 15th 2020, 14:09:52.229</td>
<td>-</td>
</tr>
</tbody>
</table>
GrimoireLab: Use Case Analysis

- Proposed solution:
  - Cumulative sum of:
    - Items as they are created.
    - Subtract closed/merged from created ones.
    - Some items have no merged date: they were merged before the field even existed. We need to subtract them too.
GrimoireLab: Use Case Analysis

- **Cumulative sum** of items as they are created.

```
es(index=gitlab_merge_requests, timefield=created_at)
cusum()
bars().color(#338984).label("Open")
```

...we are still ignoring when they were closed/merged
GrimoireLab: Use Case Analysis

- Subtract **closed/merged** from created ones.

  ```
  .es(index=gitlab_merge_requests, timefield=created_at)
  .subtract(.es(index=gitlab_merge_requests, timefield=closed_at))
  .subtract(.es(index=gitlab_merge_requests, timefield=merged_at))
  .cusum().bars().color(#338984).label("Open")
  ```
GrimoireLab: Use Case Analysis

- Subtract **closed/merged** from created ones.

<table>
<thead>
<tr>
<th>MR State</th>
<th>t₁</th>
<th>t₂</th>
<th>t₃</th>
<th>t₄</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>open</strong></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>closed/merged</strong></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>subtraction</strong></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>cusum</strong></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GrimoireLab: Use Case Analysis

- Subtract **closed/merged** from created ones.

<table>
<thead>
<tr>
<th>MR State</th>
<th>t₁</th>
<th>t₂</th>
<th>t₃</th>
<th>t₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>open</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>closed/merged</td>
<td>0</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subtraction</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cusum</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GrimoireLab: Use Case Analysis

- Subtract closed/merged from created ones.

<table>
<thead>
<tr>
<th>MR State</th>
<th>t₁</th>
<th>t₂</th>
<th>t₃</th>
<th>t₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>open</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>closed/merged</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>subtraction</td>
<td>4</td>
<td>1</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>cusum</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
GrimoireLab: Use Case Analysis

- Subtract **items with no merged_at date** from created ones.

```
.es(index=gitlab_merge_requests, timefield=created_at)
.subtract(.es(index=gitlab_merge_requests, timefield=closed_at))
.subtract(.es(index=gitlab_merge_requests, timefield=merged_at))
.subtract(.es(index=gitlab_merge_requests, timefield=updated_at, q="merged:true AND -merged_at:*"))
cusum().bars().color(#338984).label("Open")
```
GrimoireLab: Use Case Analysis
GrimoireLab: Use Case Analysis
GrimoireLab: Use Case Analysis
Evolution of GrimoireLab Dashboard

- We started by solving the use case with custom visualizations.
- All information is publicly available\(^1\). Anyone can reproduce the visualization on top of GrimoireLab data.
- **WIP**: Sigils project will include specific dashboards\(^2\).
- This is the ideal case we would like to see:
  - GrimoireLab evolving towards metrics requested by the community.

\(^1\) [https://gitlab.com/Bitergia/c/gitlab/support/issues/31](https://gitlab.com/Bitergia/c/gitlab/support/issues/31)
\(^2\) [https://github.com/chaoss/grimoirelab-sigils/issues/410](https://github.com/chaoss/grimoirelab-sigils/issues/410)
Thank You