Github Events Data

Thomas Gersdorf

December 12, 2013

1 Fundamental and Derived Measures

1.1 Terminology

Gist Gists are snippets of code to be shared with others. Can be clone, forked etc. like a repository.

Fork Forks are clones/copies of a repository or Gist into a User's own account.

1.2 Users

Event type	Description	Contents	# for all users in example dataset
IssuesEvent		Opened & Closed Issues, Assignee (API)	29368
CreateEvent			116683
PushEvent	Source code is pushed into repository.	Size, SHA Hash	298729
GistEvent			26267
Issue CommentEvent			45166
Commit CommentEvent			7827
ForkEvent			19867
WatchEvent			65193
PullRequestEvent			20818
PullRequest ReviewCommentEvent			3726
FollowEvent			17049
GollumEvent			11781
DeleteEvent			3785
MemberEvent			2764
DownloadEvent			2576
PublicEvent			553
TotalEvents	-	-	672152

Constructed measures:

Issues_Opened # of issues the user has opened

Issues_Closed # of issues the user has closed

Event_Continuity $\rightarrow$ TODO

Events_Own Events on user's repository

Events_Active_Involvement Measure how much the user engages in projects of other user or organizations, events on other repository

Events_Passive_Involvment Measures how much other users engage in projects of the user, events on the user's repository by other users

1.3   Repositories

2   Constructed Leadership Measures

2.1   Envisioning and Goal Orientation

$$LS_{\text{milestone}} = \frac{\# \text{ issue directly assigned to a milestone or release}}{\# \text{ total issues opened}} \tag{1}$$

2.2   Leading People

2.2.1   Open Issue and Assign User/Milestone

Measure:

$$LS_{\text{assign}} = \frac{\# \text{ issue directly assigned to other person}}{\# \text{ total issues opened}} \tag{2}$$

Reasoning: The more people directly assign users to issues, the more they know about the people's capabilities and empower other users to solve this issue.
Control for:

- Total number of issues

- Size of project

Assignee currently only via API from `https://api.github.com/repos/thgersdorf/github-play/issues/1`, see reference.

2.3   Empowering

2.4   Project Governance

2.4.1   Closing issues by other people's commits — Gatekeepers

Measure:

$$LS_{\text{assign}} = \frac{\# \text{ issue directly assigned to other person}}{\# \text{ total issues opened}} \tag{3}$$

Control for:

3   Performance Measures

3.1   Contributors

3.2   Release, Commits, Issue closures

$$\text{Issue closing} = \tag{4}$$

Issue User Effectiveness:

$$IUE = \frac{\Delta \frac{\# \text{ Issues Open}}{\# \text{ Total Issues}}}{\Delta \# \text{ User contributing}} \tag{5}$$

3.3   Downloads, Views etc.

4   Reference

4.1   Assignee from API

Example URL: `https://api.github.com/repos/thgersdorf/github-play/issues/1`

```json
{
  "url": "https://api.github.com/repos/kevans91/Chatters/issues/5",
  "labels_url": "https://api.github.com/repos/kevans91/Chatters/issues/5/labels{/name}",
  "comments_url": "https://api.github.com/repos/kevans91/Chatters/issues/5/comments",
  "events_url": "https://api.github.com/repos/kevans91/Chatters/issues/5/events",
  "html_url": "https://github.com/kevans91/Chatters/issues/5",
  "id": 4059710,
  "number": 5,
  "title": "Locked Menu Entries",
  "user": {
    "login": "kevans91",
    "id": 646318,
    "avatar_url": "https://gravatar.com/avatar/8d4c2928fe43bb196970867f59ab6b93?d=https%3A%2F%2Fidenticons.github.com%2F5b22ea77797a2c6e6e0859b128be4d94.png&r=x",
    "gravatar_id": "8d4c2928fe43bb196970867f59ab6b93",
    "url": "https://api.github.com/users/kevans91",
    "html_url": "https://github.com/kevans91",
    "followers_url": "https://api.github.com/users/kevans91/followers",
    "following_url": "https://api.github.com/users/kevans91/following{/other_user}",
    "gists_url": "https://api.github.com/users/kevans91/gists{/gist_id}",
    "starred_url": "https://api.github.com/users/kevans91/starred{/owner}{/repo}",
    "subscriptions_url": "https://api.github.com/users/kevans91/subscriptions",
    "organizations_url": "https://api.github.com/users/kevans91/orgs",
    "repos_url": "https://api.github.com/users/kevans91/repos",
    "events_url": "https://api.github.com/users/kevans91/events{/privacy}",
    "received_events_url": "https://api.github.com/users/kevans91/received_events",
    "type": "User",
    "site_admin": false
  },
  "labels": [
    {
      "url": "https://api.github.com/repos/kevans91/Chatters/labels/enhancement",
      "name": "enhancement",
      "color": "84b6eb"
    }
  ],
  "state": "open",
  "assignee": null,
  "milestone": null,
  "comments": 0,
  "created_at": "2012-04-11T06:43:08Z",
  "updated_at": "2012-04-11T06:43:08Z",
  "closed_at": null,
  "pull_request": {
    "html_url": null,
    "diff_url": null,
    "patch_url": null
  },
  "body": "Locked menu entries currently clutter up the menus at the top pretty badly. They either need to be removed or used.",
  "closed_by": null
}
```

4.2 Raw JSONs

4.2.1 IssuesEvent

```json
"repository":{"homepage":"http://tdurand.github.com/mapafortaleza",
"has_downloads":true,"has_issues":true,"master_branch":"gh-pages","forks":5,"language":"JavaScript","fork":false,"has_wiki":true,"url":"https://github.com/tdurand/mapafortaleza","created_at":"2011/09/29 03:27:42 -0700","size":132,"
```

```
    private":false,"name":"mapafortaleza","description":"Mapa de Fortaleza","owner"←
        :"tdurand","open_issues":9,"watchers":13,"pushed_at":"2012/02/29 16:43:30 -0800←
        "},
3  "actor_attributes":{"name":"Thibault Durand","company":"@tibbb","gravatar_id":"←
        cedb88f236afaeed4caea39917ebd0a7","blog":"http://www.thibault-durand.fr","type"←
        :"User","login":"tdurand"}
4  ,"created_at":"2012/04/11 16:01:27 -0700","public":true,"actor":"tdurand","payload"←
        :{"number":12,"action":"opened","issue":4075341},"url":"https://github.com/←
        tdurand/mapafortaleza/issues/12","type":"
5  IssuesEvent"}
```

4.2.2 PushEvent

```
1  {"repository":{"homepage":"","has_downloads":true,"has_issues":true,"
2  forks":1,"fork":false,"has_wiki":true,"url":"https://github.com/brunolopesjn/←
        JConvexHull","created_at":"2012/04/10 18:37:34 -0700","size":208,"private":←
        false,"name":"JConvexHull","description":"Programa Java Swing que resolve o ←
        problema do caixeiro viajante (TSP) utilizando o Convexo de Hull","owner":"←
        brunolopesjn","open_issues":0,"watchers":1,"pushed_at":"2012/04/11 15:01:36 ←
        -0700"},"actor_attributes":{"name":"Bruno Lopes Alcantara Batista","gravatar_id←
        ":"854bf813e2a302dcc4a059b89c502424","blog":"kariridev.blogspot.com","type":"←
        User","login":"brunolopesjn","email":"brunolopesjn@gmail.com"},"created_at":"←
        2012/04/11 15:01:37 -0700","public":true,"actor":"brunolopesjn","payload":{"←
        head":"8469f67ff919f2a776336bfe61f5ae7f40e76c72","size":1,"shas":[["8469←
        f67ff919f2a776336bfe61f5ae7f40e76c72","brunolopesjn@gmail.com","Criado a ←
        interface grafica base, a  rea  de desenho dos pontos, o desenho\ndo plano ←
        cartesiano e a l  gica de abertura de arquivos.","Bruno Lopes Alcantara ←
        Batista",true]],"ref":"refs/heads/master"},"
3  url":"https://github.com/brunolopesjn/JConvexHull/compare/ea77edbd70...8469f67ff9",←
        "type":"PushEvent"}
```

4.2.3 IssueCommentEvent

4.2.4 CommitCommentEvent

4.2.5 ForkEvent

4.2.6 FollowEvent

4.3 SQL Queries Temp

4.3.1 Generic Event Queries

```
1   # total events involvements:
2   SELECT * FROM `events` WHERE `actor` = 'mitar' AND `repository_owner` != 'mitar'
3  SELECT * FROM `events` WHERE `actor` = 'mitar' AND `repository_owner` = 'mitar'
4
5   # specific Issues properties
6   SELECT payload_number, payload_action, payload_issue from events where type  = '←
        IssuesEvent' LIMIT 30;
```