

SortingHat

Wizardry on Software Project Members

D. Moreno, S. Dueñas, V. Cosentino, M.A. Fernández, A. Zerouali, G. Robles, J.M. Gonzalez-Barahona
ICSE, Montreal, May 30th, 2019 (Room: Centre-Ville, 11:20-11:40)

Jeff's tools



jalbert@shop.com
jeff@market.com



jeffy

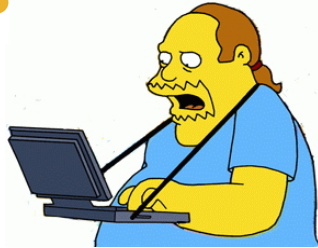


jalbertson



comicbookguy

Context



Jeff's companies



Now

06/2008



05/2008

01/2001

Description

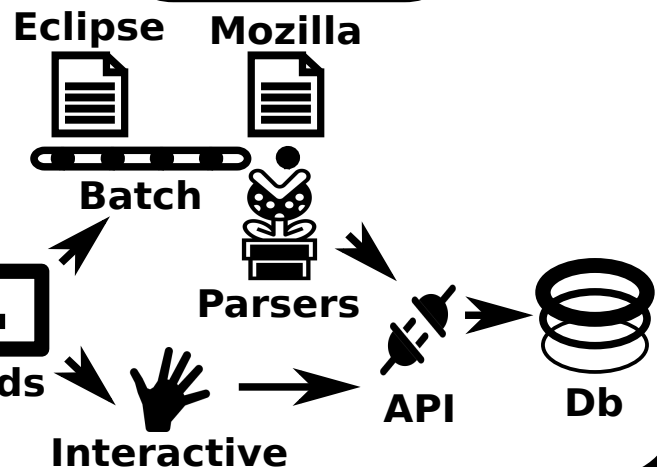
Nowadays, software projects and in particular open source ones heavily rely on a plethora of tools (e.g., Git, GitHub, Slack) to support and coordinate their activities. Despite their paramount value, they foster to fragment members' contributions, since members can access them with different identities (e.g., email and username).

SortingHat helps to track unique identities of project members and their related information such as gender, country and organization enrollments. It allows to manipulate identities interactively as well as to load bulks of identities via batch files (useful for projects with large communities).

SortingHat is a component of GrimoireLab, an industry strong free platform developed by Bitergia, which offers commercial software analytics and is part of the CHAOSS project of the Linux Foundation.



Approach



Use

Name: Miguel Ángel Fernández

e-mail: None

Bot? ☐

Country:

Enrollments

Organization	Dates	Update	enroll
Bitergia	2017-10-02 to 2100-01-01	<input type="button" value="update"/>	<input type="button" value="enroll"/>
Universidad Rey Juan Carlos	1900-01-01 to 2017-10-01	<input type="button" value="update"/>	<input type="button" value="enroll"/>

Profile Identities

Name	e-mail	Username	Source
Miguel Ángel Fernández	mfernan@bitergia.com	mfernan	github
Miguel Ángel Fernández	None	mgf@bit	github
None	mfernan@bitergia.com	None	chaoss-misc

QUERY VARIABLES

QUERY

```
1 query identities {
2   "email": "mfernan@bitergia.com",
3   "username": "mfernan",
4   "source": "github",
5   "bot": "false",
6   "country": "es",
7   "profile": {
8     "name": "Miguel Ángel Fernández",
9     "email": "mfernan@bitergia.com",
10    "username": "mfernan",
11    "source": "github",
12    "bot": "false",
13    "country": "es"
14   }
15 }
16 }
```

RESULTS

```
1 {
2   "unique_identities": 1,
3   "identities": [
4     {
5       "name": "Miguel Ángel Fernández",
6       "email": "mfernan@bitergia.com",
7       "username": "mfernan",
8       "source": "github",
9       "bot": "false",
10      "country": "es",
11      "profile": {
12        "name": "Miguel Ángel Fernández",
13        "email": "mfernan@bitergia.com",
14        "username": "mfernan",
15        "source": "github",
16        "bot": "false",
17        "country": "es"
18      }
19     }
20   ]
21 }
```

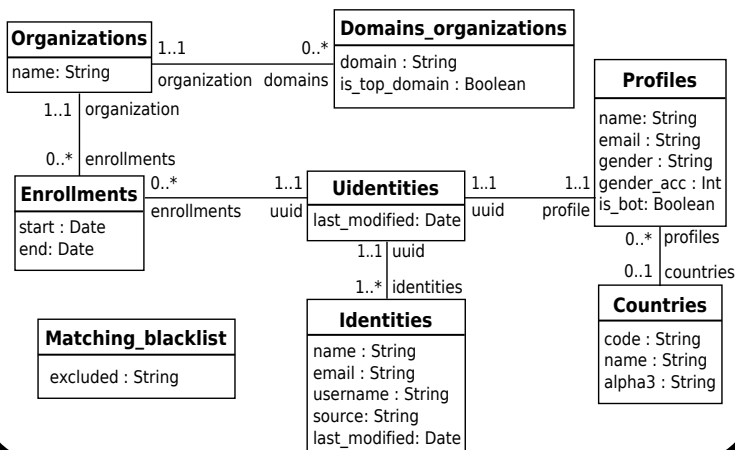
```
def load_bulk_identities(items_count, new_identities, sh_db, connector_name):
    identities_count = len(new_identities)

    SortingHat.add_identities(sh_db, new_identities, connector_name)

    logger.debug("Processed %i items identities (%i identities) from %s",
                 items_count, len(new_identities), connector_name)

    return identities_count
```

Schema



chaoss/grimoirelab-hatstall
chaoss/grimoirelab-sortinghat
chaoss.github.io/grimoirelab



@bitergia
@grimoirelab
@CHAOSSproj