Metadata for

gitenberg

Eric Hellman
Free Ebook Foundation

How ebooks are read in 2015

Project Gutenberg

46 million downloads per year
… from *48,000* texts

static files only
fix or improve manually
robots not allowed
opaque metadata creation

How software is built in 2015

GitHub

9.6 million people collaborating

... across *22.9 million* repositories

git version control

supports distributed workflow

built to be automated

rich documentation tools

9⁺ GITenber × | MARC 2 × | MARC D × | On BIBF × | MARC 2 × | ONIX Co × | edsu/pyr × | EP MoPad: × | Hello Wo × | GITenber × | GITenber ×

← → C 🔒 GitHub, Inc. [US] https://github.com/gitenberg

Apps 📰 The New York Times f Facebook | Home 🐦 Twitter / Home in LinkedIn: Home 📁 sports 📁 gluejar 📁 weather 📁 my accounts .epub » 📁 Other Bookmarks

Search GitHub

Pull requests   Issues   Gist

GITenberg ⓘ

a collaborative, trackable, scriptable digital library using Git

🔗 https://gitenberg.github.io   ✉ seth+gitenberg@sethish.com

People   3 ›

gitenjar
Eric Gitenberg Hellman

rdhyee-GITenberg
Raymond Yee (GITenberg)

sethwoodworth
Seth Woodworth

Filters ▾   🔍 Find a repository…

Little-Dorrit_963
★ 0   ⑂ 1

Little Dorrit by Dickens, Charles is a Project Gutenberg book, now on Github.

Updated 4 hours ago

A-Brief-History-of-the-Internet_250
★ 0   ⑂ 0

A Brief History of the Internet by Hart, Michael, 1947-2011 is a Project Gutenberg book, now on Github.

Updated 5 days ago

Thyrza_4302
HTML   ★ 0   ⑂ 1

Thyrza by Gissing, George is a Project Gutenberg book, now on Github.

Updated 22 days ago

Moby-Dick--Or-The-Whale_2701
HTML   ★ 4   ⑂ 3

Moby Dick; Or, The Whale by Melville, Herman is a Project Gutenberg book, now on Github.

Updated 22 days ago

How metadata could be built in 2015?

GitHub

9.6 million people collaborating

… across *22.9 million* repositories

git version control

supports distributed workflow

built to be automated

rich documentation tools

Timothy Ball on May 17 add minority_url

5 contributors

1205 lines (1204 sloc)   51.047 kB   Raw   Blame   History

```
 1   - type: house
 2     name: House Committee on Agriculture
 3     url: http://agriculture.house.gov/
 4     minority_url: http://democrats.agriculture.house.gov/
 5     thomas_id: HSAG
 6     house_committee_id: AG
 7     subcommittees:
 8     - name: Conservation and Forestry
 9       thomas_id: '15'
10       address: 1301 LHOB; Washington, DC 20515
11       phone: (202) 225-2171
12     - name: Commodity Exchanges, Energy, and Credit
13       thomas_id: '22'
14       address: 1301 LHOB; Washington, DC 20515
15       phone: (202) 225-2171
16     - name: General Farm Commodities and Risk Management
17       thomas_id: '16'
18       address: 1301 LHOB; Washington, DC 20515
19       phone: (202) 225-2171
20     - name: Livestock and Foreign Agriculture
21       thomas_id: '29'
22       address: 1301 LHOB; Washington, DC 20515
```

Most committees for 114th congress

Browse files

handlers authored on Feb 23   1 parent d1a81f2   commit 4c4c92993bd84ff8e9ecb2dfdd299fe594329bce

Showing **2 changed files** with **11,608 additions** and **14,162 deletions**.

| Unified | Split |

25,709 ■■■■□ committee-membership-current.yaml   **View**

11,570 additions, 14,139 deletions not shown

61 ■■■■□ committees-current.yaml   **View**

```
@@ -4,11 +4,11 @@
4    4       thomas_id: HSAG
5    5       house_committee_id: AG
6    6       subcommittees:
7    -    - name: Conservation, Energy, and Forestry
     7    +    - name: Conservation and Forestry
8    8         thomas_id: '15'
9    9         address: 1301 LHOB; Washington, DC 20515
10   10        phone: (202) 225-2171
11   -    - name: Department Operations, Oversight, and Nutrition
     11   +    - name: Commodity Exchanges, Energy, and Credit
12   12        thomas_id: '22'
13   13        address: 1301 LHOB; Washington, DC 20515
14   14        phone: (202) 225-2171
@@ -16,14 +16,18 @@
16   16        thomas_id: '16'
17   17        address: 1301 LHOB; Washington, DC 20515
18   18        phone: (202) 225-2171
19   -    - name: Livestock, Rural Development, and Credit
     19   +    - name: Livestock and Foreign Agriculture
20   20        thomas_id: '29'
21   21        address: 1301 LHOB; Washington, DC 20515
```

Joe is his nickname #290

💬 Conversation 1   ⦵ Commits 1   📄 Files changed 1

timball commented on May 21

Joe Crowley has a nickname , it's Joe .

--timball

⦵ 🟡 `Joe is his nickname`   ✔ 6242266

konklone commented on May 21   Owner

Okay, http://crowley.house.gov/about-me/full-biography is good enough for me!

⦵ **konklone** merged commit **b7016c2** into `unitedstates:master` on May 21

edlabor.house.gov -> edworkforce.house.gov
Timothy Ball authored on May 17

49a240b

Commits on May 14, 2015

url'ing like it's 1995 fixes issue #278
Timothy Ball authored on May 14

f9a617f

Commits on Mar 16, 2015

A few new committee names
handlers authored on Mar 16

e047460

Commits on Feb 23, 2015

Most committees for 114th congress
handlers authored on Feb 23

4c4c929

Commits on Nov 13, 2014

ran committee_membership.py
JoshData authored on Nov 13, 2014

b095f5f

Commits on Sep 12, 2014

Update committee yaml files for Benghazi
handlers authored on Sep 12, 2014

00079ce

Commits on Sep 11, 2014

Adding thomas_id, fixing house id
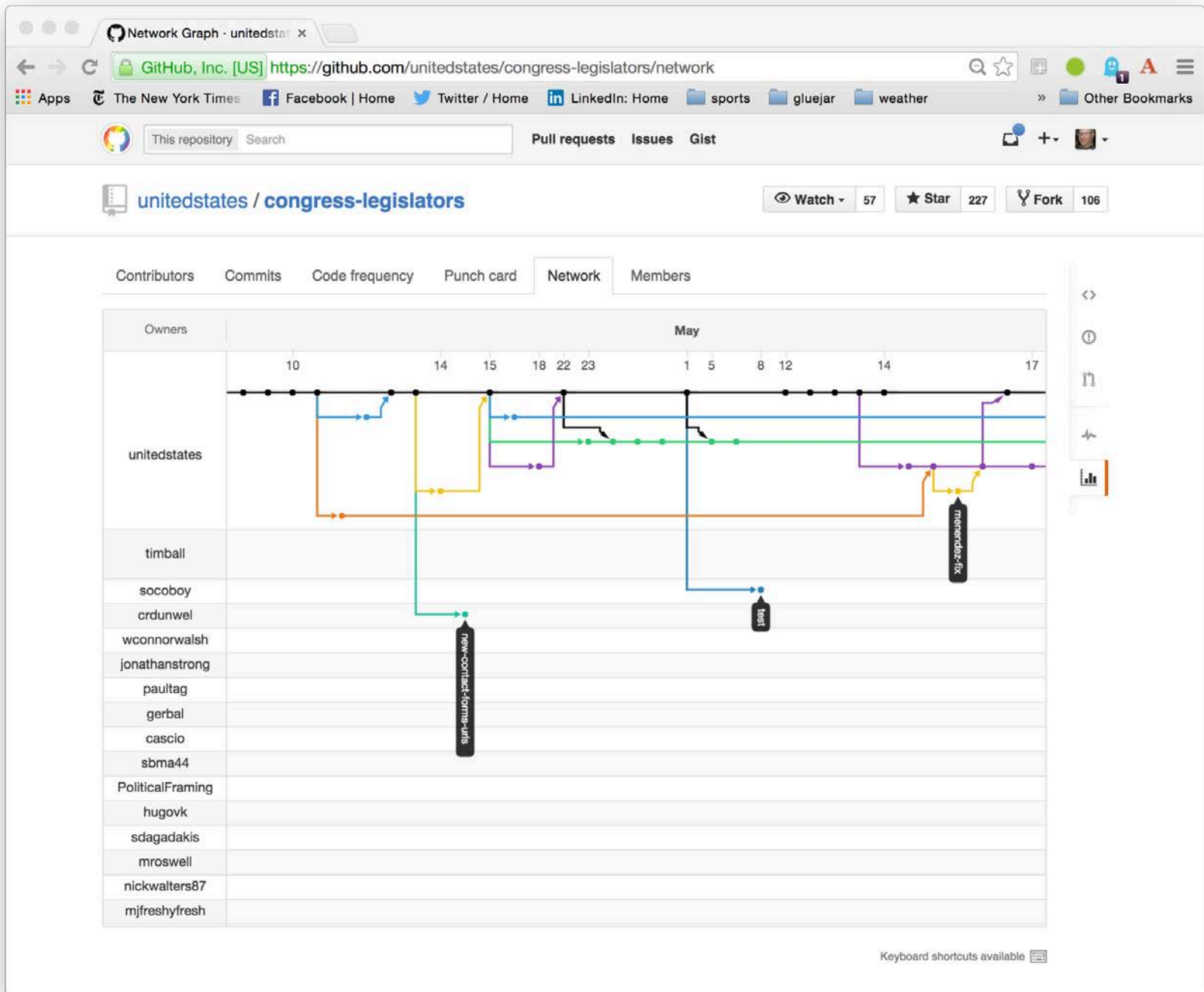LindsayYoung authored on Sep 11, 2014

4cf6a30

Benghazi committee is haveing a hearing so I am adding their House co…
LindsayYoung authored on Sep 11, 2014

73476db

Metadata for Gitenberg???

- Version Control ➜ YAML

_repo: Space-Viking_20728
authors:
 -
  author_id_gutenberg: 8301
  author_name: H. Beam Piper
  author_sortname: Piper, H. Beam
  author_wikipedia: http://en.wikipedia.org/wiki/H._Beam_Piper
language: en
publication_date_original: '1964'
publisher: Project Gutenberg
publisher_original: Ace Books
rights: PD-US
rights_url:
subjects:
 - !lcsh Science fiction
 - !lcsh Space warfare -- Fiction
 - !lcsh Revenge -- Fiction
 - !lcc PS
 - Space Opera
title: Space Viking

Metadata for Gitenberg!

- YAML software – DONE

- Initial YAML production – in progress

- Property definitions - ongoing

- YAML ➔ MARC -DONE

We need help!

www.gitenberg.org