Githubでサイズ
の大きなファイルを扱う話

2017 10 14 mishima.syk #11
@iwatobipen

自己紹介


pen
@iwatobipen
Medicinal Chemist. Interested in Chemoinformatics, Running.
Opinions and tweets are my own. 流れるつぶやきはあくまで個人の見解です。
静岡県東部のほう。 ・ iwatobipen.wordpress.com

Gitの問題

- バイナリファイルの差分は記録できない

- 大きなファイルを使うとローカルリポジトリが肥大化する

https://help.github.com/articles/what-is-my-disk-quota/

Gitの問題

File and repository size limitations

We recommend repositories be kept under 1GB each. This limit is easy to stay within if large files are kept out of the repository. If your repository exceeds 1GB, you might receive a polite email from GitHub Support requesting that you reduce the size of the repository to bring it back down.

In addition, we place a strict limit of files exceeding 100 MB in size. For more information, see "Working with large files."

> **Note:** If you add a file to a repository via a browser, the file can be no larger than 25 MB. For more information, see "Adding a file to a repository".

https://help.github.com/articles/what-is-my-disk-quota/

100MB以上のファイルプッシュすんなよ。

リポジトリ1GB以上にすんなよ。

https://help.github.com/articles/what-is-my-disk-quota/


だが
断(ことわ)る

だって、、、

- 深層学習とか色々ファイルサイズ大きくないっすか。

- データーサイズってだんだん増えますよね、、、

そこでGit Lfs!

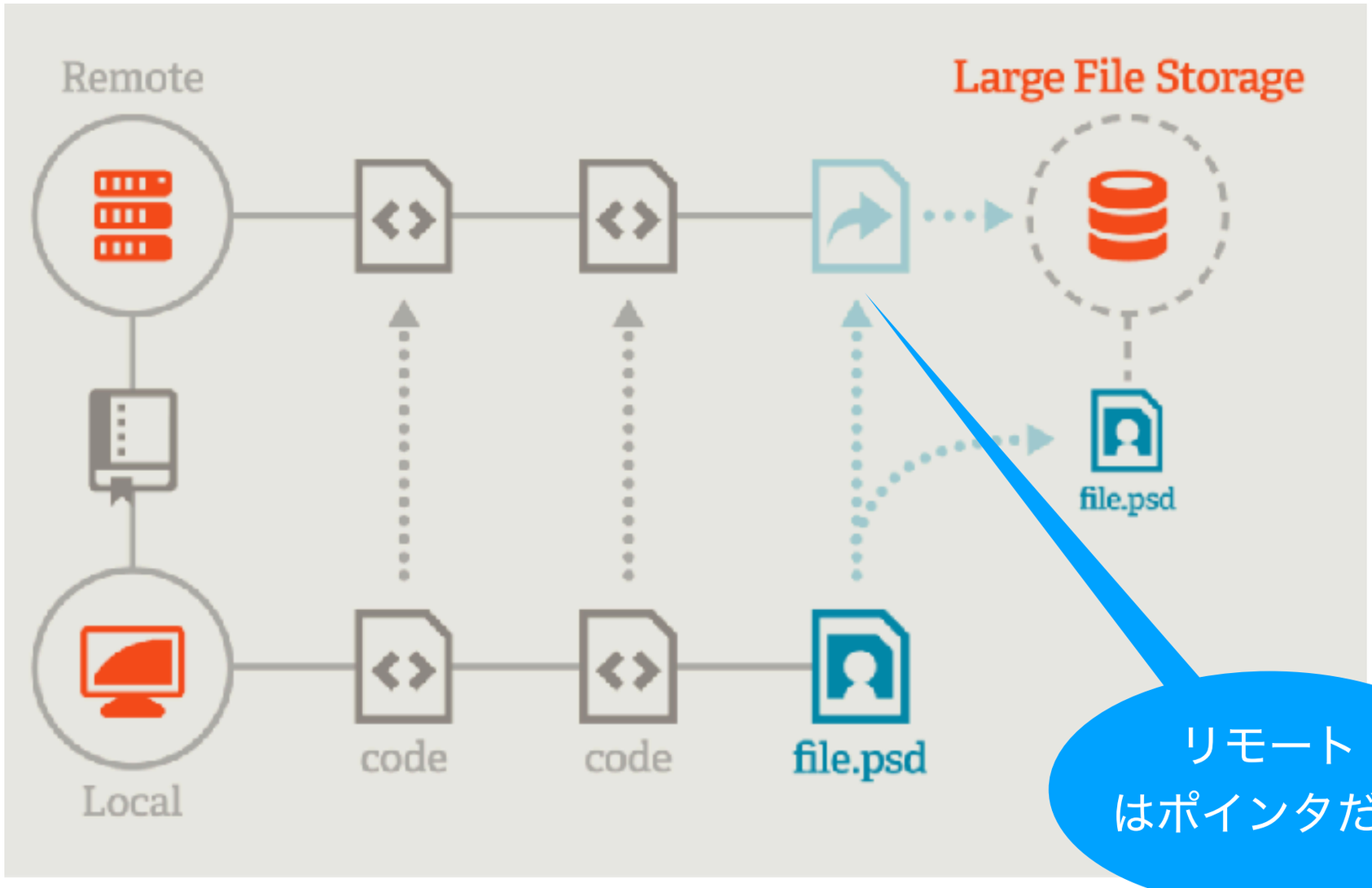
Git Lfsって？

Lfsの名前のまんま。

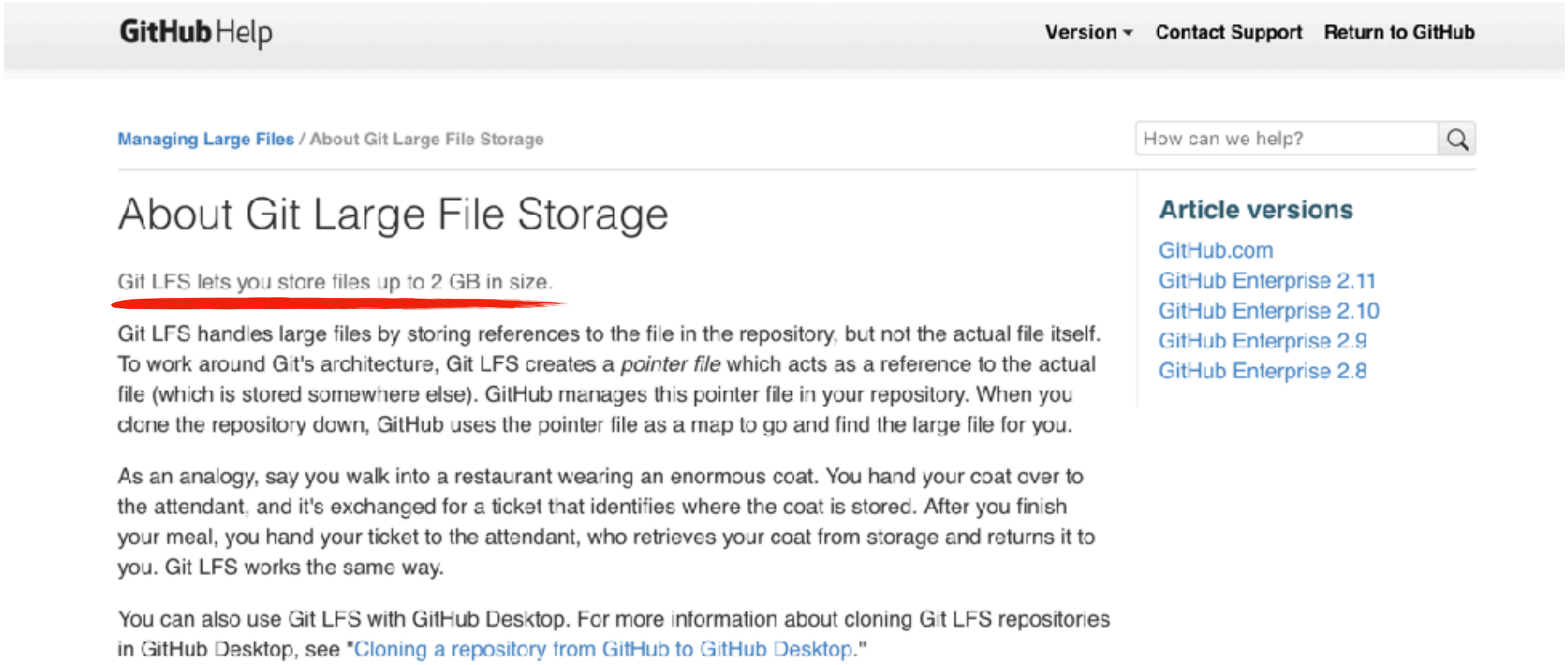
Git Lfsって？

Gitはコード（文字）を管理する目的だったのでバイナリファイルは不得手。

そこを解消するソリューソンがGit Lfs



2GBまでは無料プラン内



https://help.github.com/articles/about-git-large-file-storage/

セットアップ手順

OSx
- brew install git-lfs
Debian
- sudo apt-get install git-lfs
windows
- https://github.com/git-lfs/git-lfs/releases (σ・∀・)σｹﾞｯｿ!!

https://github.com/git-lfs/git-lfs/wiki/Installation

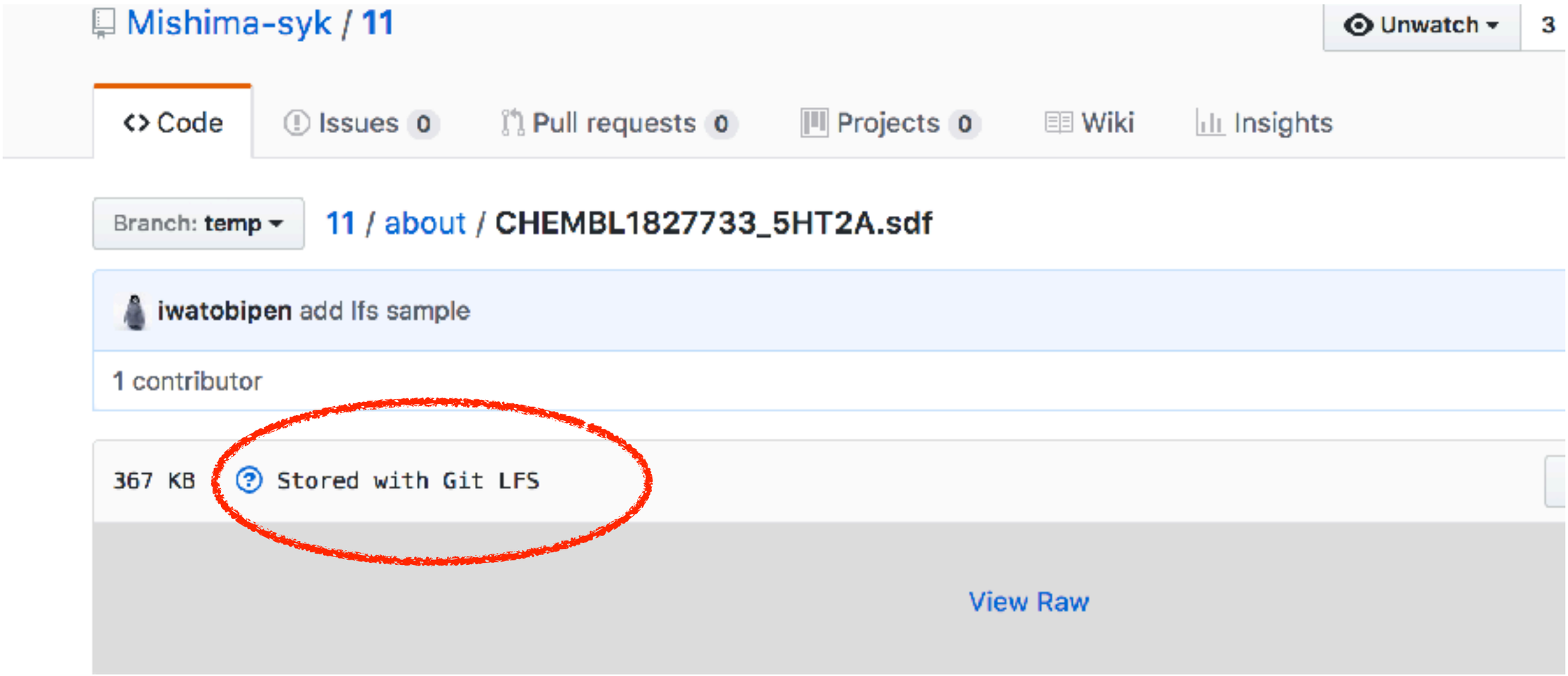
使い方

```
iwatobipen$ git lfs track '*.sdf'
    Tracking “*.sdf”
    # 上のコマンドでトラッキングするファイルを指定するだけ

iwatobipen$ cat .gitattributes
    *.sdf filter=lfs diff=lfs merge=lfs -text

iwatobipen$ git lfs track
    Listing tracked patterns
        about/*.sdf (about/.gitattributes)
```

push push push

```
iwatobipen$ git push origin temp
    Git LFS: (1 of 1 files) 367.39 KB / 367.39 KB
    Counting objects: 5, done.
    Delta compression using up to 8 threads.
Compressing objects: 100% (4/4), done.
Writing objects: 100% (5/5), 577 bytes | 577.00 KiB/s, done.
Total 5 (delta 1), reused 0 (delta 0)
remote: Resolving deltas: 100% (1/1), completed with 1 local object.
To https://github.com/Mishima-syk/11.git
   4e5b2ed..024c70e  temp -> temp
```

LFSに保存されている



まとめ

注意点

　Git LFS を利用しているリポジトリは、一見すると普通の Git リポジトリと同じに見えます。

　誤って Git LFS 非インストール環境下でリポジトリが clone され、ラージファイルの追加が push された場合、ポインタ管理が機能せずに実体がそのままリポジトリ管理されてしまいます。

　さらに、Git LFS インストール環境からは、ポインタとして見えるはずのラージファイルが実体として見えることで不整合が発生してしまい、ニッチもサッチもいかない状況になります。