

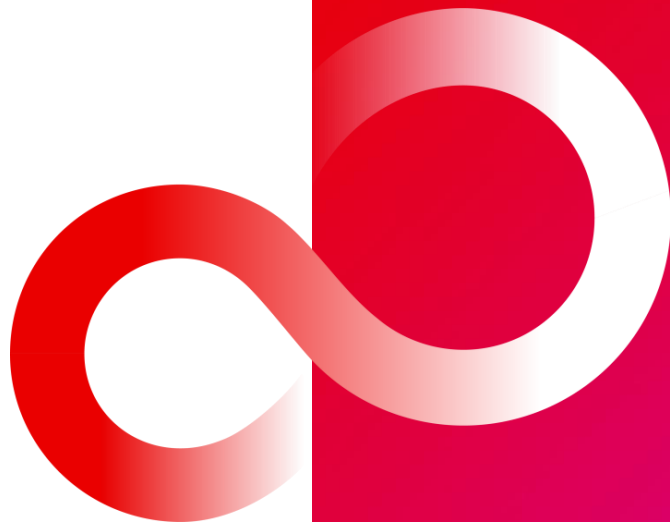
Software Heritageについて

2021年11月10日

富士通 FCTグループ

ソフトウェア事業部 ソリューション開発部

安倍昌輝



- ユネスコ（UNESCO）とフランス国立情報学自動制御研究所（INRIA）が立ち上げたプロジェクト
- 全てのソフトウェアのソースコードの収集・保存・アクセスのための基盤構築が目的
 - <https://www.softwareheritage.org/>
- ライセンスは以下から選択可能
 - AGPLv3+ for public-facing services on the Internet
 - GPLv3+ for backend code and services
 - Apache 2 for Puppet code and recipes
- 2021/11/10時点で収集された情報
 - Source files : 11,522,325,744
 - Projects : 167,702,974



- <https://archive.softwareheritage.org/>

- アーカイブされたリポジトリへアクセス
- プロジェクトの特定バージョンへアクセス
- ソースコードファイルの特定行へアクセス

- アーカイブされたリポジトリへアクセス

- アクセスしたい任意のリポジトリのURLに、以下のURLを先頭に追加

<https://archive.softwareheritage.org/browse/origin>

例: <https://github.com/mirror/busybox>

→

<https://archive.softwareheritage.org/browse/origin/https://github.com/mirror/busybox>

○プロジェクトの特定バージョンへアクセス

- リポジトリへのアクセス方法に加えて `swh:1:rev:XXX` を追記する

- SWHID(XXXの部分) は以下の方法で確認

1. ブラウザでアクセスして“permalinks”ボタンから確認
2. Gitリポジトリの場合、commit hashがSWHIDになる

- 例:

<https://archive.softwareheritage.org/browse/origin/https://github.com/mirror/busybox>

→

<https://archive.softwareheritage.org/browse/swh:1:rev:e9f8b2f830bf0273d0d51f56d24da84b955e0df5;origin=https://github.com/mirror/busybox>

○ソースコードファイルの特定行へアクセス

○SWHID(XXXの部分) は以下の方法で確認

1. ブラウザでアクセスして、任意の行を選択して“permalinks”ボタンから確認

○例:

```
https://archive.softwareheritage.org/swh:1:cnt:efab5dcb44c01387a2ee30e0a98a085c17ebdfba;origin=https://github.com/mirror/busybox;visit=swh:1:snp:1ce05af92fd28bcceb8ad22e128d2dc2ae99cbfc;anchor=swh:1:rev:3b411ebbfc749f9f12b0eb739cb5ba3ec052197e;path=/init/init.c;lines=100-110
```

○ 資料

- <https://archive.softwareheritage.org/api/>
- <https://docs.softwareheritage.org/devel/swl-web/uri-scheme-api.html>

○ 例: <https://github.com/mirror/busybox>

- Get information about directory objects.
<https://archive.softwareheritage.org/api/1/directory/da81f1546b78d25f4ac63e612e76cb27aa4c2db3/>
- Get information about a content (aka a “blob”) object.
https://archive.softwareheritage.org/api/1/content/sha1_git:becd9bf6dbfd629027b802fb585b746abf0260ee/
- Get information about the license of a content object.
https://archive.softwareheritage.org/api/1/content/sha1_git:6f50a7164838c362837b72ca2c87600ebe090039/license/

Thank you

