

01.07.2016 10:31 Проект Software Heritage нацелен на создание полного архива открытых исходных текстов

☆ opennet.ru/opennews/art.shtml



**ПРОЕКТ IP-AS.RU
100% РЕЗУЛЬТАТ**

Навигация

Каналы: [XML](#) [t](#)
[B](#)

[Разделы новостей |](#)
[Сводные](#)

[Что нового на](#)
[OpenNet](#)

[Поиск в новостях](#)

[Новые комментарии](#)

[Добавить свою](#)
[новость](#)

Лента

[Важное](#)

Французский институт исследований в информатике и автоматике [INRIA](#), при поддержке правительства Франции, [ввёл в строй](#) проект [Software Heritage](#), задачей которого является сбор, сохранение и предоставление доступа к исходным текстам, собранным в одном месте из различных доступных источников. На базе Software Heritage создан крупнейший архив кода, ставящий перед собой задачу охватить все существующие открытые проекты и защитить доступный открытый код от возможной потери после закрытия сайтов проектов и репозиториях, а также дающий возможность организовать научные исследования над крупнейшей коллекцией кода.

В настоящий момент [архив](#) включает более 2.6 миллиарда файлов, охватывающих 23 миллиона открытых проектов и почти 600 млн коммитов. Архив включает код, размещённый в GitHub, в репозиториях Debian, коллекциях GNU и в других публичных архивах. Налажен процесс автоматической синхронизации изменений с GitHub, который скоро будет реализован и для других источников. Для запроса файлов пока можно использовать только контрольные суммы SHA-1.

Из планов на будущее отмечается расширение базы источников, реализация удобной навигации по коду и системы полнотекстового поиска, создание средств для отслеживания происхождения кода и цепочек заимствования, расширение метаданных и проведение статистических исследований, подготовка интерфейса для загрузки кода и истории изменений в форматах современных систем контроля версий.

Дополнительно можно отметить совместную инициативу GitHub и Google, в рамках которого [подготовлено](#) новое хранилище, расширяющее возможности [GitHub Archive](#). Хранилище включает данные об активности в более чем 2.8 млн репозиториях на GitHub, 145 млн коммитах и 2 миллиардах файлов. Предоставлен интерфейс для поиска данных с использованием регулярных выражений, а также для формирования SQL-подобных запросов (например, [можно определить](#) наиболее часто используемые в коде модули, наиболее активных разработчиков, использование Emoji в коммитах и т.п.).

Тип: К сведению

Ключевые слова: [opensource](#), [code](#), [archive](#)

При перепечатке указание ссылки на [orennet.ru](#) обязательно

- [1.1, Аноним](#), 11:13, 01/07/2016 [[ответить](#)] [[смотреть все](#)]

+5
+/-

Чтобы было натуральное наследие, надо все исходники перенести на глиняные таблич... весь текст скрыт [[показать](#)]

- 29.06 [Google](#) открыл код [SwiftShader](#), программной реализации [OpenGL ES](#) и [Direct3D 9](#) (23 +20)

Советы

- 12.05 [Загрузка](#) ядра [Linux](#), собранного в конфигурации [allmodconfig](#)

- 11.05 [Хранение](#) конфиденциальных данных в [Git-репозитории](#)

- 07.04 [Включение](#) [bash](#) и инструментария [Ubuntu](#) в [Windows 10](#)

- 03.04 [Оценка](#) работы [Unity 8](#) в изолированном контейнере, запускаемом в обычном [Ubuntu](#)

- 26.03 [Расширение](#) программного [RAID1](#) в [Linux](#) до зеркала из трёх дисков

Обсуждаемые новости

- 00:19 [Во FreeBSD](#) появились ветки [11-STABLE](#) и [12-CURRENT](#) (96)

- 00:08 [Релиз](#) библиотеки [LibVirt](#) 2.0.0 (5)

-
- [2.4](#), [hr4tklo45rhjy89o6y](#), 11:41, 01/07/2016 [^] +5
[ответить] [смотреть все] [показать ветку] +/-

>глиняные таблички

теРРабайты

- [3.47](#), [Аноним](#), 11:26, 04/07/2016 [^] +/-
[ответить] [смотреть все]

Гигатонны терабайтов.

- [2.8](#), [Аноним](#), 12:13, 01/07/2016 [^] +/-
[ответить] [смотреть все] [показать ветку]

У глиняных табличек плотность хранения маловата, лучше на фторопластовую перфоле... весь текст скрыт [показать]
[показать ветку]

- [3.11](#), [none_first](#), 12:29, 01/07/2016 [^] +/-
[ответить] [смотреть все]

легкий пожар и ... весь текст скрыт
[показать]

- 23:56 [Релиз IPython 5.0](#) (16)

- 23:52 [В Chrome добавлены средства шифрования, стойкие к подбору на](#) (94)

- 23:28 [Для Flatpak подготовлена технология управляемого доступа к р](#) (20)

- 23:20 [Выпуск Wine 1.9.14](#) (30)

- 23:18 [Выпуск системы сборки CMake 3.6](#) (42)

- 22:06 [Подготовленный сообществом Blender мультфильм победил на фес](#) (35)

- [4.14](#), [омномномнимус](#), 12:42, 01/07/2016 [[^](#)] [[ответить](#)] [[смотреть все](#)] +3
+/-

тогда чугуниевую

- [5.18](#), [тоже Аноним](#), 13:19, 01/07/2016 [[^](#)] [[ответить](#)] [[смотреть все](#)] +4
+/-

Один колоссальный датацентр с высокой надежностью хранения стоит в Гизе уже четы...
весь текст скрыт
[\[показать\]](#)

avatar не загружен

- [4.17](#), [FK](#), 13:07, 01/07/2016 [[^](#)] [[ответить](#)] [[смотреть все](#)] +/-

Фторопласт не горит Разве что расплавить можно ... весь текст скрыт [\[показать\]](#)

- [5.22](#), [Аноним](#), 14:29, 01/07/2016 [[^](#)] [[ответить](#)] [[смотреть все](#)] -1
+/-

А в клубок такую ленту можно замесить?

- [2.35](#), [Вареник](#), 20:28, 01/07/2016 [[^](#)]
[[ответить](#)] [[смотреть все](#)] [[показать ветку](#)]

+1
+/-

>> а то какой-то ридонли гитхаб получается.

Write-Only?

- [2.54](#), [Аноним](#), 17:20, 07/07/2016 [[^](#)]
[[ответить](#)] [[смотреть все](#)] [[показать ветку](#)]
[[к модератору](#)]

+/-

В Европе есть проект по переносу знаний на топазные таблички Между двух пласт... весь текст скрыт [[показать](#)] [[показать ветку](#)]

- [2.55](#), [Аноним](#), 18:10, 07/07/2016 [[^](#)]
[[ответить](#)] [[смотреть все](#)] [[показать ветку](#)]
[[к модератору](#)]

+/-

Есть M-Disk способен хранить информацию 1000лет, читаться и пишется недорогоими ... весь текст скрыт [[показать](#)] [[показать ветку](#)]

- [1.2](#), [Аноним](#), 11:16, 01/07/2016 [[ответить](#)]
[[смотреть все](#)]

+/-

Сто тысяч миллионов файлов средним размером 60 килобайт У них Штцву-ов не хва... весь текст скрыт [[показать](#)]

- [1.3](#), [Аноним](#), 11:36, 01/07/2016 [[ответить](#)]
[[смотреть все](#)]

+5
+/-

Молодцы, французы, шарят То золото заберут, то разживутся копией сорцов на собс... весь текст скрыт [[показать](#)]

- 1.5, Аноним, 11:42, 01/07/2016 [ответить] [смотреть все]

+2
+/-

information must be free!

- 1.6, hr4tklo45rhjy89o6y, 11:43, 01/07/2016 [ответить] [смотреть все]

+1
+/-

Хочу торренты, чтобы у Я%&#@й пригорало.

- 2.20, jgdhbfctbf5ceybebc5, 13:25, 01/07/2016 [^] [ответить] [смотреть все] [показать ветку]

+1
+/-

Вот для начала <https://www.debian.org> CD
torrent-cd <http://linuxtracker.org> h... весь текст
скрыт [показать] [показать ветку]

- 2.48, Аноним, 11:28, 04/07/2016 [^] [ответить] [смотреть все] [показать ветку]

+/-

> Хочу торренты, чтобы у Я%&#@й пригорало.

Тебе не нравится пшеница?!

- 1.7, Какаянахренразница, 11:44, 01/07/2016 [ответить] [смотреть все]

+6
+/-

> нацелен на создание полного архива открытых исходных текстов

Открывай меня полностью.
<https://www.opennet.ru/search.shtml?words=%D0%CF%CC%CE>

- [1.9, jOKer](#), 12:18, 01/07/2016 [[ответить](#)] [[смотреть все](#)] +12
+/-

"Кинутся тогда люди искать свои мебели, а где они, мебели? Вот они где! Здесь они! В шкафу. А кто сохранил, кто уберег? Коробейников. Вот господа спасибо и скажут старичку, помогут на старости лет... А мне много не нужно — по десяточке за ордерок подадут — и на том спасибо... А то иди попробуй, ищи ветра в поле. Без меня не найдут!"
(с) Двенадцать стульев



- [1.12, robux](#), 12:39, 01/07/2016 [[ответить](#)] [[смотреть все](#)] +/-

"Создано централизованное хранилище".
- вот как я прочитал новость. А между тем шёл 2016 год.



- [2.26, Crazy Alex](#), 16:37, 01/07/2016 [[^](#)] [[ответить](#)] [[смотреть все](#)] [[показать ветку](#)] +/-

Как ни крути - а это ещё один бэкап Хотя если б они к своей копии дали какую-то... весь текст скрыт [[показать](#)] [[показать ветку](#)]

- [1.15, Аноним](#), 12:53, 01/07/2016 [[ответить](#)] [[смотреть все](#)] +/-

Срезы архива будут?

- [1.16, Андрей](#), 12:59, 01/07/2016 [[ответить](#)] [[смотреть все](#)] +3
+/-

А ведь точно! Как раз в соседней новости про DMCA <https://www.opennet.ru/opennews/art.shtml?num=44700> выяснилось, что свободный код безвозвратно удаляют. Интересно, они его успели заархивировать? Или они там архивируют только после цензуры?

А вообще с историей всё выглядит печально: посмотрите на первые коммиты проектов: `initial import from cvs`. Один коммит! А в cvs же была целая история развития! Потеряна. Или они связываются с авторами на предмет, не сохранилась ли у них копия cvs? Вот это был бы действительно бесценный труд IT-археологов. Есть уже официально такая профессия?

-
- [2.19](#), тоже [Аноним](#), 13:23, 01/07/2016 [[^](#)] [[ответить](#)] [[смотреть все](#)] [[показать ветку](#)]

+/-

Судя по тонким знатокам фекалий мамонта, все больше IT-палеонтологи.

-
- [2.38](#), [Лев Исаакиевич Гуревич](#), 22:28, 01/07/2016 [[^](#)] [[ответить](#)] [[смотреть все](#)] [[показать ветку](#)]

+/-

Про CVS точно подмечено, я один раз пол дня искал откуда и зачем взялся этот код... весь текст скрыт [[показать](#)] [[показать ветку](#)]

-
- [3.45](#), [Аноним](#), 18:36, 03/07/2016 [[^](#)] [[ответить](#)] [[смотреть все](#)]

+/-

А комментировать код эти уебаны так и не научились ... весь текст скрыт [[показать](#)]

- [1.23](#), [Аноним](#), 15:48, 01/07/2016 [[ответить](#)]
[[смотреть все](#)]

+/-

Сколько весит архивчик?

- [1.24](#), [Kodir](#), 15:59, 01/07/2016 [[ответить](#)]
[[смотреть все](#)]

+/-

Фильтрация на индусокод будет?

[.....нить скрыта, показать \(8\)](#)

- [1.34](#), [Вареник](#), 20:24, 01/07/2016 [[ответить](#)]
[[смотреть все](#)]

+/-

Молодцы. Хоть кто-то этим начал заниматься.

>> при поддержке правительства Франции

Ну не Сколково же. У нас только подслушку инновационно обязать хранить в голову чиновника может прийти.

- [2.39](#), [shlnkcbu4y43uqg](#), 23:51, 01/07/2016 [[^](#)] +/-
[[ответить](#)] [[смотреть все](#)] [[показать ветку](#)]

У нас write-only нанобэкап, это для кого надо бэкап.

- [2.42](#), [DmA](#), 08:25, 03/07/2016 [[^](#)] [[ответить](#)] +2
[[смотреть все](#)] [[показать ветку](#)] +/-

ну это больше не чиновникам в голову пришло, а чекисткой опг защищающая честно ... весь текст скрыт [[показать](#)]
[[показать ветку](#)]

- [3.43](#), [Аноним](#), 17:58, 03/07/2016 [[^](#)] -1
[[ответить](#)] [[смотреть все](#)] +/-

клоун на важные события выпускной, свадьба, юбилей, похороны люди часто аренд... весь текст скрыт
[[показать](#)]

- [1.53](#), [Mna](#), 13:19, 05/07/2016 [[ответить](#)] [[смотреть все](#)] +/-
-

Хорошая новость. Конечно это должны были сделать французы, хорошо, что само правительство Франции взялось. Ведь Франция - сердце культуры. Посмотрите на список проектов [ingia](#).

Если посмотреть кто автор CiteSeerX поисковика по научным работам, то это тоже француз, который, правда, перебрался для этого в Пенсильванию, штат, мотто которого "учить и обучать"

Конечно это стоило бы сделать раньше: [archive.org](#) например существует давно. Но сейчас уже просто необходимо: в свете того что деструкторы принялись уже и это уничтожать

Выше ДжОкерт шутит что собирать архивы сравнимо с работой старьевщика-коробейника, однако сорс-коды это другая материя, это даже не материя а явления из мира идей, и оно не гниет в этом смысле, а знания сохранить важно. (Когда говорят про "гниение кода" имеют ввиду, в организационном смысле, потерю уровня понимания работы кода в активном проекте. Соответственно неактивный проект гнить в организационном плане не может)

Добавить комментарий

Имя:

E-Mail:

Заголовок:

Текст:
