



ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

HAROKOPIO UNIVERSITY

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

“ΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ ΤΗΛΕΜΑΤΙΚΗ”

ΚΑΤΕΥΘΥΝΣΗ

“ΤΕΧΝΟΛΟΓΙΕΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΙΣΤΟΥ”

Ιωάννης Παπαδόπουλος

ΑΜ: 20126

Ακαδημαϊκό έτος: 2021-2022

Τμήμα Πληροφορικής και Τηλεματικής

Department of Informatics and Telematics



**Μεταπτυχιακή εργασία**

*“Αποθετήριο πακέτων λογισμικού διανομών Linux”*

**Master Thesis**

*“Software repository aggregator for Linux distributions”*

- a web application
- periodically collects Linux packages info
- stores it in a database
- provides a website with search filters
- extendable for more Linux distros
- fully containerized



## Supported Linux OS

- ❖ Ubuntu
- ❖ Debian
- ❖ Kali
- ❖ Fedora
- ❖ CentOS

## Collected Package Info

- ❖ name
- ❖ Linux distribution
- ❖ type
- ❖ category
- ❖ license
- ❖ maintainer
- ❖ description
- ❖ homepage website URL
- ❖ code repository URL

## Collected Package Version Info

- ❖ name
- ❖ architecture
- ❖ size
- ❖ binary file URL

- ❖ open source project
- ❖ easy package search
- ❖ package compare
- ❖ no need to have multiple Linux OS installed
- ❖ Dockerfile generation
- ❖ package info extraction to multiple formats
- ❖ direct package download
- ❖ direct access to package source code
- ❖ promoting quality packages via rating system

```
# Generated by PKGman at 2022-02-25 22:19:00
FROM debian:11

RUN apt-get update

RUN apt-get install 0install-core=2.16-1 -y && \
    apt-get install sqlacodegen -y
```

*Generated Dockerfile*

```
{
  "id": 6,
  "versions": [
    {
      "id": 32,
      "user_rating": {
        "rating_id": 1,
        "rate": 4
      },
      "rating": 3.5,
      "version": "2.16-1",
      "architecture": "x86_64",
      "swid": "",
      "swid_exists": null,
      "size": 38719,
      "binary_URL": "http://ftp.debian.org/debian/pool/main/z/zeroinstall-injector/0install-core_2.16-1_amd64.deb",
      "package": 6
    }
  ],
  "rating": 3.5,
  "name": "0install-core",
  "distro": "Debian:11",
  "type": "deb",
  "section": "admin",
  "license": "No license",
  "maintainer": "Thomas Leonard <talex5@gmail.com>",
  "description": "cross-distribution packaging system (non-GUI parts)",
  "homepage": "http://0install.net/",
  "repo_URL": "https://github.com/0install/0install-debian",
  "selectedVersion": "2.16-1"
},
```

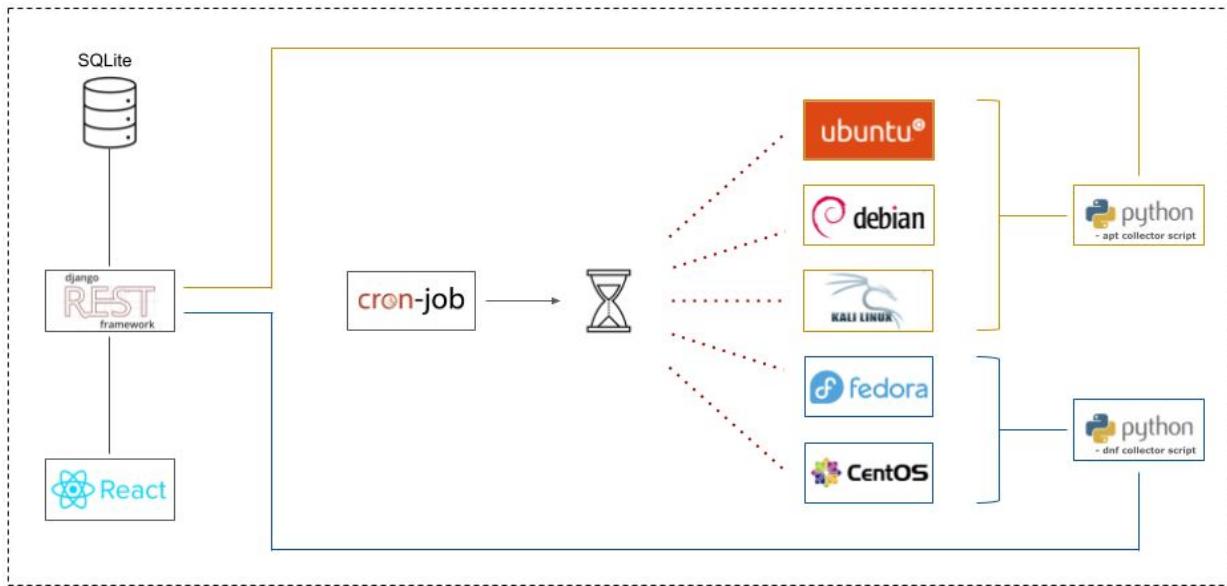
*Extracted Package Info in JSON format*

## dnf package manager

```
$ dnf info tilix
Last metadata expiration check: 27 days, 10:00:23
Installed Packages
Name           : tilix
Version        : 1.6.4
Release        : 1.fc26
Arch           : x86_64
Size           : 3.6 M
Source         : tilix-1.6.4-1.fc26.src.rpm
Repo           : @System
From repo      : @commandline
Summary        : Tiling terminal emulator
URL            : https://github.com/gnunn1/tilix
License        : MPLv2.0 and GPLv3+ and CC-BY-SA
Description    : Tilix is a tiling terminal emulator
```

## apt package manager

```
$ apt show nano
Package: nano
Version: 2.8.6-3
Priority: standard
Section: editors
Origin: Ubuntu
Maintainer: Ubuntu Developers <ubuntu-devel-discuss@lists.ubuntu.com>
Original-Maintainer: Jordi Mallach <jordi@debian.org>
Bugs: https://bugs.launchpad.net/ubuntu/+filebug
Installed-Size: 766 kB
Depends: libc6 (>= 2.14), libncursesw5 (>= 6), libtinfo5 (>= 6)
Suggests: spell
Conflicts: pico
Breaks: nano-tiny (<< 2.8.6-2)
Replaces: nano-tiny (<< 2.8.6-2), pico
Homepage: https://www.nano-editor.org/
Task: standard, ubuntu-touch-core, ubuntu-touch
Supported: 9m
Download-Size: 222 kB
APT-Manual-Installed: yes
APT-Sources: http://in.archive.ubuntu.com/ubuntu artful/main amd64
Description: small, friendly text editor inspired by Pico
```



## Main services

> orchestrated by Docker-Compose

❖ the information collection mechanism

❖ the back-end

❖ the front-end

- apt\_collector.py
- dnf\_collector.py

## Command line args

- ❖ max-concurrency
- ❖ distro
- ❖ API-URL
- ❖ archives-url
- ❖ repos-url (only for rpm distros)

```
1 for pkg in list_distro_pkgs(): # parallel processing
2     exist_pkg_info = fetch_pkg_info(pkg) # HTTP request to backend
3     if exist_pkg_info is None: # new package
4         pkg_info = extract_pkg_info(pkg)
5         extra_pkg_info = get_pkg_info_from_github(pkg) or \
6             get_pkg_info_from_distro_code_repo(pkg)
7
8         save_pkg_info(pkg_info, extra_pkg_info) # HTTP request to backend
9     else: # package already exists check for new versions
10        new_versions = []
11        for version in extract_pkg_versions(pkg):
12            if version not in exist_pkg_info["versions"]: # new version
13                new_versions.append(extract_version_info(pkg, version))
14
15        save_new_pkg_version_info(pkg, new_versions) # HTTP request to backend
16
```

*Collector Scripts Pseudocode*

```
docker run --net host --shm-size=5.07gb --env GITHUB-TOKEN=<token> ubuntu_distro python3
apt_collector.py -d Ubuntu -u http://localhost:8000/api/v1 -a http://archive.ubuntu.com/ubuntu
```

*apt\_collector script execution command*

- built with Django REST framework
- utilizes a SQLite database
- queries the database
- feeds the frontend
- uses OAuth 2.0 for authorization
- provides pagination

## User Categories

- ❖ visitors
- ❖ regular users
- ❖ package collectors
- ❖ admins

packages			⌵
GET	/packages/	packages_list	🔒
POST	/packages/	packages_create	🔒
POST	/packages/dockerfile/	packages_dockerfile	🔒
GET	/packages/versions/	packages_versions_list	🔒
POST	/packages/versions/	packages_versions_create	🔒
GET	/packages/versions/ratings/	packages_versions_ratings_list	🔒
POST	/packages/versions/ratings/	packages_versions_ratings_create	🔒
GET	/packages/versions/ratings/{id}/	packages_versions_ratings_read	🔒
PUT	/packages/versions/ratings/{id}/	packages_versions_ratings_update	🔒
PATCH	/packages/versions/ratings/{id}/	packages_versions_ratings_partial_update	🔒
DELETE	/packages/versions/ratings/{id}/	packages_versions_ratings_delete	🔒
GET	/packages/versions/{id}/	packages_versions_read	🔒
PUT	/packages/versions/{id}/	packages_versions_update	🔒
PATCH	/packages/versions/{id}/	packages_versions_partial_update	🔒
DELETE	/packages/versions/{id}/	packages_versions_delete	🔒
GET	/packages/{id}/	packages_read	🔒
PUT	/packages/{id}/	packages_update	🔒
PATCH	/packages/{id}/	packages_partial_update	🔒
DELETE	/packages/{id}/	packages_delete	🔒
POST	/packages/{id}/versions/	packages_versions	🔒
users			⌵
POST	/users/login/	users_login_create	🔒
POST	/users/login/token/refresh/	users_login_token_refresh_create	🔒
POST	/users/logout/	users_logout_create	🔒
POST	/users/register/	users_register_create	🔒

- built with React.js library
- consumes the queried results
- intuitive UI
- search filters
- extract options
- sortable columns
- dynamic info presentation
- expandable rows
- selectable results

PKGman

Log in Sign up

### Useful info for Linux distributions packages

Example: grep Search

Show: type category rating license maintainer website repo description All

Package	Distribution	Type	Category	Rating	License	Maintainer	Website	Code Repository	Description
<input type="checkbox"/> > sqldcogen	Debian:11	deb	python	5.0		Python Applications Packaging Team python-apps-team@lists.alioth.debian.org	ovgs.ovh.com/ovgs...	salda.debian.org/ov...	Automatic model code generator for SQLAlchemy
<input type="checkbox"/> > Oad	Debian:11	deb	games	4.0	Other	Debian Games Team <pkg-games-devel@lists.alioth.debian.org>	elavied.com/	salda.debian.org/ga...	Real-time strategy game of ancient warfare
<input type="checkbox"/> > R-clisymbols	Fedora:34	rpm		4.0	MIT	src.fedoraproject.org/user/ryhys	CRAN.R-project.org/...	src.fedoraproject.org/...	Unicode Symbols at the R Prompt
<input type="checkbox"/> > ca-certificates	Fedora:34	rpm		3.8	Public Domain	src.fedoraproject.org/user/ryhys	fedoraproject.org/w...	src.fedoraproject.org/...	The Mozilla CA root certificate bundle
<input type="checkbox"/> > Oinstall-core	Debian:11	deb	admin	3.5	No license	Thomas Leonard <tales5@gmail.com>	0install.net/	qithub.com/0install/...	cross-distribution packaging system (non-GUI parts)
<input type="checkbox"/> > cups-libs	Fedora:34	rpm		3.5	LGPLv2 and zlib	src.fedoraproject.org/user/fedohal	openprinting.uitbhu...	src.fedoraproject.org/...	CUPS printing system - libraries
<input type="checkbox"/> > Oad	Fedora:34	rpm		3.0	GPLv2+ and BSD and MIT and IBM and MPLv2.0	src.fedoraproject.org/user/bruno	elavied.com	src.fedoraproject.org/...	Cross-Platform RTS Game of Ancient Warfare
<input type="checkbox"/> > acl	Fedora:34	rpm		3.0	GPLv2+	src.fedoraproject.org/user/fedohal	seannah.nongnu.org...	src.fedoraproject.org/...	Access control list utilities
<input type="checkbox"/> > binutils-gold	Fedora:34	rpm		3.0	GPLv3+	src.fedoraproject.org/user/hnick	sourceware.org/bin...	src.fedoraproject.org/...	The GOLD linker, a faster alternative to the bfd linker
<input type="checkbox"/> > compat-golang-github-redis-dev	Fedora:34	rpm		3.0	BSD and ASL 2.0	qithub.com/qc-redis	qithub.com/qc-redis	qithub.com/qc-redis...	Type-safe Redis client for Golang

1-10 of 238810 < > >

PKGman

Log out

### Useful info for Linux distributions packages

Example: grep Search

Show: type category rating license maintainer website repo description All

2 packages have been selected

- Oinstall-core, Debian:11
- sqldcogen, Debian:11

Deinstall All Create Dockerfile Export to CSV Export to JSON

Package	Distribution	Type	Category	Rating	License	Maintainer	Website	Code Repository	Description
<input checked="" type="checkbox"/> > sqldcogen	Debian:11	deb	python	5.0		Python Applications Packaging Team python-apps-team@lists.alioth.debian.org	ovgs.ovh.com/ovgs...	salda.debian.org/ov...	Automatic model code generator for SQLAlchemy
<input type="checkbox"/> > Oad	Debian:11	deb	games	4.0	Other	Debian Games Team <pkg-games-devel@lists.alioth.debian.org>	elavied.com/	salda.debian.org/ga...	Real-time strategy game of ancient warfare
<input type="checkbox"/> > R-clisymbols	Fedora:34	rpm		4.0	MIT	src.fedoraproject.org/user/ryhys	CRAN.R-project.org/...	src.fedoraproject.org/...	Unicode Symbols at the R Prompt
<input type="checkbox"/> > ca-certificates	Fedora:34	rpm		3.8	Public Domain	src.fedoraproject.org/user/ryhys	fedoraproject.org/w...	src.fedoraproject.org/...	The Mozilla CA root certificate bundle
<input checked="" type="checkbox"/> > Oinstall-core	Debian:11	deb	admin	3.5	No license	Thomas Leonard <tales5@gmail.com>	0install.net/	qithub.com/0install/...	cross-distribution packaging system (non-GUI parts)

Version	Architecture	Size	Average Rating	Personal Rating	Binary Download Link
<input checked="" type="checkbox"/> 2.16-1	x86_64	37.8 kB	3.5	4	file.debian.org/debian/boon/main/vcs/installer/0install-core_2.16-1_amd64.deb

<input type="checkbox"/> > cups-libs	Fedora:34	rpm		3.5	LGPLv2 and zlib	src.fedoraproject.org/user/fedohal	openprinting.uitbhu...	src.fedoraproject.org/...	CUPS printing system - libraries
<input type="checkbox"/> > Oad	Fedora:34	rpm		3.0	GPLv2+ and BSD and MIT and IBM and MPLv2.0	src.fedoraproject.org/user/bruno	elavied.com	src.fedoraproject.org/...	Cross-Platform RTS Game of Ancient Warfare
<input type="checkbox"/> > acl	Fedora:34	rpm		3.0	GPLv2+	src.fedoraproject.org/user/fedohal	seannah.nongnu.org...	src.fedoraproject.org/...	Access control list utilities
<input type="checkbox"/> > binutils-gold	Fedora:34	rpm		3.0	GPLv3+	src.fedoraproject.org/user/hnick	sourceware.org/bin...	src.fedoraproject.org/...	The GOLD linker, a faster alternative to the

Search Filters

Type: deb X Architecture: x86\_64 X Distribution: Debian:11 X Category: database X Code Repository: qithub.com/0install/...

Show only exact search matches

Search

PKGman

Log out

Example: grep Search

Show: type category rating license maintainer website repo description All

2 packages have been selected

- Oinstall-core, Debian:11
- sqldcogen, Debian:11

Deinstall All Create Dockerfile Export to CSV Export to JSON

Package	Distribution	Type	Category	Rating	License	Maintainer	Website	Code Repository	Description
<input checked="" type="checkbox"/> > sqldcogen	Debian:11	deb	python	5.0		Python Applications Packaging Team python-apps-team@lists.alioth.debian.org	ovgs.ovh.com/ovgs...	salda.debian.org/ov...	Automatic model code generator for SQLAlchemy
<input type="checkbox"/> > Oad	Debian:11	deb	games	4.0	Other	Debian Games Team <pkg-games-devel@lists.alioth.debian.org>	elavied.com/	salda.debian.org/ga...	Real-time strategy game of ancient warfare
<input type="checkbox"/> > R-clisymbols	Fedora:34	rpm		4.0	MIT	src.fedoraproject.org/user/ryhys	CRAN.R-project.org/...	src.fedoraproject.org/...	Unicode Symbols at the R Prompt
<input type="checkbox"/> > ca-certificates	Fedora:34	rpm		3.8	Public Domain	src.fedoraproject.org/user/ryhys	fedoraproject.org/w...	src.fedoraproject.org/...	The Mozilla CA root certificate bundle
<input checked="" type="checkbox"/> > Oinstall-core	Debian:11	deb	admin	3.5	No license	Thomas Leonard <tales5@gmail.com>	0install.net/	qithub.com/0install/...	cross-distribution packaging system (non-GUI parts)

Version	Architecture	Size	Average Rating	Personal Rating	Binary Download Link
<input checked="" type="checkbox"/> 2.16-1	x86_64	37.8 kB	3.5	4	file.debian.org/debian/boon/main/vcs/installer/0install-core_2.16-1_amd64.deb

<input type="checkbox"/> > cups-libs	Fedora:34	rpm		3.5	LGPLv2 and zlib	src.fedoraproject.org/user/fedohal	openprinting.uitbhu...	src.fedoraproject.org/...	CUPS printing system - libraries
<input type="checkbox"/> > Oad	Fedora:34	rpm		3.0	GPLv2+ and BSD and MIT and IBM and MPLv2.0	src.fedoraproject.org/user/bruno	elavied.com	src.fedoraproject.org/...	Cross-Platform RTS Game of Ancient Warfare
<input type="checkbox"/> > acl	Fedora:34	rpm		3.0	GPLv2+	src.fedoraproject.org/user/fedohal	seannah.nongnu.org...	src.fedoraproject.org/...	Access control list utilities
<input type="checkbox"/> > binutils-gold	Fedora:34	rpm		3.0	GPLv3+	src.fedoraproject.org/user/hnick	sourceware.org/bin...	src.fedoraproject.org/...	The GOLD linker, a faster alternative to the