

R packages from a Fedora perspective

José Matos

jamatos@{fep.up.pt, fedoraproject.org}

(Mathematics Research Center, Economics School) Porto University and Fedora
Project

12 August 2008



Overview

Introduction

(Linux) software distribution

Free/Open software languages

Fedora

General

Languages

Fedora and R

Package structure

Fedora R Guidelines

Conclusions

Common characteristics

- ▶ distribute a coherent set of packages
- ▶ have a package system that deals with software with software inter-dependencies
- ▶ follows well established rules regarding the source code
- ▶ associated with a free/open source software
- ▶ works in multiple platforms (for some in even more than 20! hardware platforms)
- ▶ worldwide mirrors network to fetch and update packages easily

Key differences

- ▶ Release cycles
- ▶ packaging policies (package names, granularity, boundaries)
- ▶ update policies
- ▶ support periods
- ▶ packaging method/tools

Examples

Linux

- ▶ Fedora
- ▶ OpenSuse
- ▶ gentoo
- ▶ Slackware
- ▶ Debian
 - ▶ Ubuntu

Other Unixes

- ▶ *BSD
 - ▶ Free, Open, Net, ...
- ▶ OpenSolaris
 - ▶ Indiana
 - ▶ Nexenta

Characteristics

- ▶ Free/Open Source license
- ▶ strong set of public available libraries
- ▶ common set of package rules
- ▶ worldwide mirrors network to fetch and update packages easily

Examples

Languages

- ▶ R (S)
- ▶ Python
- ▶ Perl
- ▶ PHP
- ▶ T_EX

Repositories

- ▶ CRAN
- ▶ pypy
- ▶ CPAN
- ▶ pear
- ▶ CTAN

Goals



- ▶ Release often/release earlier (~ 6 month release cycles)
- ▶ Advance of free software (on the front edge)
- ▶ Strong commitment to free/open software (only free licenses are accepted for software)
- ▶ Documented guidelines for building packages

Derived distributions

- ▶ RHEL - Red Hat (TM) Enterprise Linux
- ▶ CentOS
- ▶ Scientific Linux

Supported Architectures

▶ Primary Architectures

- ▶ i*86
- ▶ x86-64
- ▶ ppc/ppc64/cell (PS3)

▶ Secondary Architectures

- ▶ Alpha
- ▶ Sun Sparc 32/64
- ▶ ia64
- ▶ arm*

Some counting

- ▶ R packages in Fedora ~40
- ▶ Fedora packages for F-9 ~9000
 - ▶ Rate of new packages per week ~20
- ▶ Debian (for comparison - not) >20000

Library update

- ▶ The library defines the language (in a sense)
- ▶ There are two incompatible ways to update libraries inside Fedora
 - ▶ system
 - ▶ maintains internal coherence on the installed packages
 - ▶ resolves external packages dependencies
 - ▶ language
 - ▶ does not depend on the availability of packages for the distribution

R

▶ DESCRIPTION file

```
Package: waveslim
```

```
Version: 1.6.1
```

```
Date: 2007-10-9
```

```
Title: Basic wavelet routines for one-, two- and three-  
dimensional signal processing
```

```
Author: Brandon Whitcher <brandon.j.whitcher@gsk.com>
```

```
Maintainer: Brandon Whitcher <brandon.j.whitcher@gsk.com>
```

```
Depends: R (>= 2.0), stats, graphics, grDevices
```

```
ZipData: no
```

```
LazyLoad: yes
```

```
LazyData: yes
```

R (cont)

▶ DESCRIPTION file

Description: Basic wavelet routines for time series
(1D), image (2D)...

License: GPL (>= 2)

URL: <http://www.image.ucar.edu/~whitcher/>
<http://www.image.ucar.edu/~whitcher/book/>

Packaged: Tue Oct 9 11:55:13 2007; bjw34032

Built: R 2.6.2; i386-redhat-linux-gnu; 2008-02-14 07:51

Fedora

- ▶ rpm spec file

```
%define packname waveslim
```

```
Summary: R module, Basic wavelet routines for 1,2  
and 3-dimensional signal processing
```

```
Name: R-%{packname}
```

```
Version: 1.6.1
```

```
Release: 2%{?dist}
```

```
License: GPLv2+
```

```
Source0:ftp://cran.r-project.org/pub/R/contrib/main/  
%{packname}_%{version}.tar.gz
```

```
URL: http://www.image.ucar.edu/staff/whitcher/
```

Fedora (cont)

- ▶ rpm spec (cont)

```
Requires: R
```

```
BuildRequires: R-devel, tetex-latex, gcc-gfortran
```

```
Requires: R-stats, R-stats, R-grDevices
```

```
%description Basic wavelet routines for  
time series (1D), image (2D) and array (3D) analysis...
```

```
...
```


Naming

- ▶ All R packages are named with R- prefixed
 - ▶ R-Biobase
 - ▶ R-mAr
 - ▶ R-zoo

Note

The package name capitalization is kept!

Guidelines

- ▶ <https://fedoraproject.org/wiki/Packaging/R>
- ▶ compile the best way to create R packages for Fedora

How to create rpm spec files for Fedora R packages

- ▶ generic templates (arch and noarch) contained in fedora-devtools
- ▶ Taking advantage of the information contained in DESCRIPTION
 - ▶ cran2rpmspec: <http://jamatos.fedorapeople.org/cran2rpmspec>
 - ▶ R2spec: <https://fedorahosted.org/r2spec/>
 - ▶ the resulting spec files needs sometimes to be fine tuned
- ▶ using the Fedora R SIG mailing list:
fedora-r-devel-list@redhat.com
<http://www.redhat.com/mailman/listinfo/fedora-r-devel-list>

Future

- ▶ better tools to create and maintain spec files updated with corresponding meta-information in R package
- ▶ increased the number of R packages in Fedora
- ▶ allow rpm to automatically verify what R packages provides and requires (such as it does now for Perl and python packages)
- ▶ ease new system installations and maintenance using distribution specific tools

Question for the assistance

- ▶ What do you do with the package recommended **Suggests** when installing new R packages?
 1. Always installs.
 2. Install on a basis to basis case.
 3. Never install, do not care about it.