

# Debian Med

## Integrated software environment for all medical purposes based on Debian GNU/Linux

Andreas Tille, Steffen Möller

31 October 2008

### Abstract

The Debian Med project started in 2002 with the objective to bring free medical software into the focus of users. The first step was the investigation of available Free Software in this area and to accomplish the conditions for a simple and solid installation of this software in Debian GNU/Linux. It has shown that the Debian Med project has a positive effect for the cooperating upstream projects.

It has shown that fields like molecular biology and medical imaging were covered quite good with Free Software solutions while there is not really much to manage a medical practice. The Debian Med project tries to be the missing link between developers and users to support Free Software solutions for all areas in medicine.

Debian Med is intended to be useful for service providers who want to distribute their free or proprietary solutions in medical care.

The whole project is embedded into the Debian Pure Blends framework (formerly known as Custom Debian Distributions) as well as Debian Edu, Debian Science and others.

## 1 Introduction

### 1.1 Motivation and purpose of Debian Med

The advent of Debian Med is characterised by the rise of several new Free Software projects that were useful for certain tasks in medical care. Several of these projects – even if looking promising in the beginning – do not really exist any more because they did not respected the principle of Free Software: care for a solid user base and recruit qualified developers from it. This task is one of the hardest in the field of health care: there is a much smaller number of users for such specific software compared to a web browser or an office suite.

A big problem was that it was quite hard to obtain solid information about all these projects. Some engaged users tried to assemble lists of these projects and published these in the web. While those lists are an interesting start they are finally not really helpful for users: Users just need ready to run programs with no effort to obtain and install the software. The idea of Debian Med was now to provide not only a list but exactly what users need.

## 1.2 Status of Free Software in health care

Common programs like a web server, or a mail user agent are installed on most computers and have a very large user base. Knowing this, many gifted programmers feel obliged for this kind of Free Software - they just need it for their own and thus feel motivated to spend their time on it.

The fact that a piece of software is needed for the own work is often the basic motivation to write Free Software. Biological software is often developed by scientists themselves because they are the only persons who have the needed insight in the topic. Many authors of this software just realised the profit they might gain to share the code and thus the field of biology is really well covered with Free Software.

Another quite well covered part of Debian Med is medical imaging. Even if there are quite important packages like Bio Image Suite [University, 2008] not yet packaged for Debian first steps are done and there is a certain coverage of DICOM viewers and other software which is very useful in medical imaging.

It has shown that the field of patient management and medical health record applications which is considered by most people as the basic health care software has only one single representative which is GNUmed [Hilbert, 2008]. It has turned out that there are a lot of similar Free Software projects that try to solve more or less the same problem but diverge in the techniques used (programming language, database server, etc.), user interface (GUI or web application) and the basic ideas about workflow and philosophy.

The even larger task to manage a hospital is also tackled by several projects and the most famous one OpenVista[OpenVista, 2008] as a real enterprise grade health care information system is a really complex project and needs a stronger team than the current Debian Med team with a larger technical background in the specific techniques used in OpenVista. To overcome this problem the strategy of Debian Med is to try to involve the upstream authors into the packaging: They are really the experts of the software in question and we try to teach and help them in packaging.

## 2 Methods

### 2.1 Unique technology supporting Debian's principles in packaging

#### 2.1.1 Build daemons

Sites external to the Debian main distribution may offer packages only for a subset of architectures. The Debian main distribution, however, automatically compiles software for all 11 architectures that are supported by the Debian effort. To get a package into Debian, be it novel or an update of an existing package, the maintainer of a package submits the source code of the program together with his changes on the code to create the package. The build daemons (or autobuilders) compile the packages for each of the supported systems and make the resulting package publicly available for download. Logs of the build platforms are available online for everybody's inspection.

#### 2.1.2 Bug tracking system

Users should give immediate feedback about problems arising in using a package. They always have the choice of reporting these to the upstream developer, usually per email.

A particular strength of [SourceForget.net](#) is to bring users of a particular software together. Earlier than this effort was the [Debian Bug Tracking System](#) (BTS). The maintainer of a software can decide if the bug should be forwarded to the upstream developers of the package or if it is fixed by himself. All problems are made public and hence the whole community may contribute to solving a particular issue.

## 2.2 Sharing the work

The Debian Project is an association of individuals who share the intention to create the best possible free operating system. This operating system that is created is called Debian GNU/Linux, or simply Debian for short. Everybody in the internet may initiate a site and offer packages for the installation in Debian. A local administrator has to decide, if this public source may be trusted.

For Free Software development to work it requires a critical mass of supporters. Development without feedback prior to the submission of the final product is disadvantageous. The development of programs is not the main concern of a regular Linux Distribution. However, with the focus on Free Software and smooth local compilation, Debian considerably eases the contribution of comments and concise feedback of the technically skilled early adopters. Debian such helps to bring developers and users of applications together.

### 2.2.1 Debian Policy

All GNU/Linux distributions have a certain amount of common ground, and the [Linux Standard Base](#) (LSB)[[Workgroup, 2005](#)] is attempting to develop and promote a set of standards that will increase compatibility among Linux distributions, hereby enabling software applications to run on any compliant system. The very essence of any distribution is the choice of *policy statements*.

While every single maintainer of a Debian package has to build the package in compliance with the policy he has the ability and the right to decide which software is worth packaging. Normally maintainers choose the software which is used in their own work and they are free to move the development of Debian in a certain direction (as long as they follow the rules of the policy). This is referred to as *Do-o-cracy* in Debian which means: The doer decides what is done.

## 2.3 Selection of packages

Debian contains nearly 20000 binary packages, and this number is constantly increasing. There is no single user who needs all these packages. The regular user is interested in a subset of these packages. To specify packages of one's particular interest, several options are provided by Debian:

**tasksel** Provision of a reasonable selection of rather general tasks that can be accomplished using a set of packages installed on a Debian GNU/Linux system. However, these are not yet covering scientific applications. The Blend toolkit which is currently developed will also support `tasksel` to enable selecting for instance Debian Med right after a fresh installation of a general Debian system.

**command line package management** `apt` provide means to search for packages of particular interest by its name or words in the package's description. Every package also indicates, as set by its maintainer, references to other packages of potential interest.

**GUI** There are several graphical user interfaces to manipulate the installation of packages on a Debian installation. The most popular is currently `synaptic` which enables users to seek for certain packages and displays detailed information about each package.

Debian officially maintains 11 different architectures with many more not officially supported ports to other operating systems, which includes some that run another flavour of UNIX. Its technology for package management has been adopted for other operating systems, i.e. Fink on MacOSX ([fink.sourceforge.net](http://fink.sourceforge.net)).

A *distribution* is a collection of software packages around the GNU Linux operating system that satisfies the needs of the target user group. There are general distributions, which try to support all users, and there are several specialised distributions, which each target a special group of users.

*Distributors* are those companies that are building these collections of software around the GNU Linux operating system. Since the software is Free, the user who buys a distribution pays for the service that the distributor is providing. These services might be:

- Preparing a useful collection of software around GNU Linux.
- Caring for smooth installation that the target user is able to manage.
- Providing software updates and security fixes.
- Writing documentation and translations to enable the user to use the distribution with maximum effect.
- Selling Boxes with ready to install CDs and printed documentation.
- Offering training and qualification.

## 3 Results

### 3.1 Comparable Debian-associated repositories

**Bio-Linux Bioinformatics package repository** The Bio-Linux Bioinformatics package repository contains the [Bio-Linux 4 bioinformatics software](#) and can be installed from a centralised repository located on the EGTDC server. The packages available from this site have been created by the EGTDC specifically for the Bio-Linux project and are in deb format.

The packages are not, however, policy compliant Debian packages because they install files into `/usr/local` hierarchy in contrast to the Debian policy which does not allow files inside packages at this location because `/usr/local` is reserved for locally installed files that do not fall under responsibility of the Debian package manager.

Besides this technical fact Bio-Linux authors were not that strict regarding licensing and copyright of packaged projects. Every official Debian package has to comply to the Debian Free Software Guidelines (DFSG) and the copyright information has to be shipped with the file `/usr/share/doc/<packagename>/copyright`. Moreover the source of a binary Debian package has to be provided next to the binary. All these very important requirements are not fulfilled in most cases in Bio-Linux.

However, the authors did a great job in collecting a certain amount of very useful software for biologists and the Debian Med project is seeking for possibilities for cooperation.

**BioLinux-BR Project** A similar project is the [BioLinux-BR Project](#) which is a project

directed to the scientific community. Their goal is to create a Linux distribution for people with little familiarity with the installation of the operational system and mainly for people who do not know to proceed unpacking a program, compile and install it correctly.

In fact, this project has assembled a huge amount of packages, probably the most complete collection of Free Software in biology. Packages for multiple distributions are provided, which includes Debian, and a live CD.

### 3.2 Other repositories of biology related software

Looking beyond Debian and related distributions which share more or less the same technique we find similar efforts to deliver sets of ready to install software

**FreeBSD Ports: Biology** The Free Software world does not only know Linux as free operating system. There are others out there like several BSD derivatives, OpenSolaris, Hurd and others. The FreeBSD project has a really nice [collection of biological software](#).

### 3.3 Bioinformatics Live CDs

The concept of a live CD allows to create a CD or DVD that boots a computer, starts a defined set of application without a user's intervention and has all tools in place that suits a particular community. Such provide fully featured Linux workstations without additional installations of access to local disk space, alternatively booting via the network is supported by Debian, which particularly appeals to Blades or large clusters. The most successful such LiveCD is the Debian-derived Knoppix[Knopper, 2005].

**The Quantian Scientific Computing Environment** Quantian is a remastering of a well established effort (Knoppix). The interesting part for biologists is that Quantian contains in addition all interesting packages of Debian Med. The author Dirk Eddelbuettel, who is a Debian developer himself, just used the simply to install biological software feature we provide and thus made a great profit from Debian Med.

**Vigyaan - the biochemical software workbench** Vigyaan is an electronic workbench for bioinformatics, computational biology and computational chemistry. It has been designed to meet the needs of both beginners and experts. VigyaanCD is a live Linux CD containing all the required software to boot the computer with ready to use modelling software. VigyaanCD v0.1 is based on Knoppix v3.3.

Vigyann contains some programs which are not yet contained in Debian. It might be mutually beneficial to include these provided that the license fits the DFSG.

**BioKnoppix** BioKnoppix is a customised distribution of Knoppix Linux Live CD. It is a very similar project to the previous which specialises Knoppix for computational biology and chemistry.

**VLinux Bioinformatics Workbench** Also VLinux is at the time of writing a Live CD based on the same outdated Knoppix version 3.3 as Vigyann and includes a slightly changed software selection and surely a different background layout.

These are too many different initiatives that could all well do much more in order to share the burden of maintenance and updates. With Debian they have the right basic infrastructure. The time will show, whose packages will gain most momentum.

## 3.4 Comparison with Debian Med

### 3.4.1 Other fields than only biology

Looking at all the projects above it becomes evident that they are all dealing only with biological software. Above it was stated that one strong column of Debian Med is this specific field and it is for a reason: The amount of free biological software is large and most of these projects are relatively easy to turn into packages - so the amount of work per package is much smaller compared to for instance medical record applications with preparation of databases, dedicated user management, etc.

So there is one major difference between the projects mentioned above and Debian Med: While the biological part is really interesting for medical care Debian Med tries to cover all other fields of medical care as well. This goal is not yet reached but continuous work is done into this direction and some important steps are done.

### 3.4.2 Debian Pure Blend

An even more important difference than the more general approach compared to the other repositories is the fact that Debian Med is not only about just packaging software. The Debian Med project is one of the earliest so called *Debian Pure Blends* (formerly known as Custom Debian Distributions) and just wants to do more for the comfort of their users than adding binary packages to the Debian package pool. The main goal is to turn Debian into the distribution of choice for people working in the field of medicine and to make Debian an operating system that is particularly well fit for the requirements for medical practice and research. The goal of Debian Med is a complete system for all tasks in medical care which is build completely on free software.

### 3.4.3 Flexibility in supporting small user groups

On the organisational side the project tries to attract people working in the field of Free Software in medicine to share the effort of building a common platform which reduces the amount of work for developers and users. Sharing the effort in publishing free medical software to a large user base is a main advantage of Debian Med.

The strength of Debian is the huge number of developers (more than 1000) all over the world working in different fields. Some of them are working in the field of biology or medicine and thus have a natural interest in developing a rock solid system they can relay on for their own work (not only commercial interest to sell service per accident). So sometimes the chances to realise specific support for small user groups are better inside a community driven distribution than in a commercial distribution: You just need some developers who have a specific interest and they will realise and publish an environment for their needs and will share it with other users. A company that has to gain a certain market share is not flexible enough in this regard to cover very specific interests.

The underlying principle that those things will be done if there is somebody who just does the work is called *Do-O-Cracy* – which just means the doer decides what gets done.

That is the reason why Debian is often the platform of choice for researcher in the field of biology: Some biologists are Debian maintainers and so they added support for biological packages. The more the Debian user in the field of biology report back about problems or wishes the more Debian maintainers are able to enhance their system for their own and their users profit.

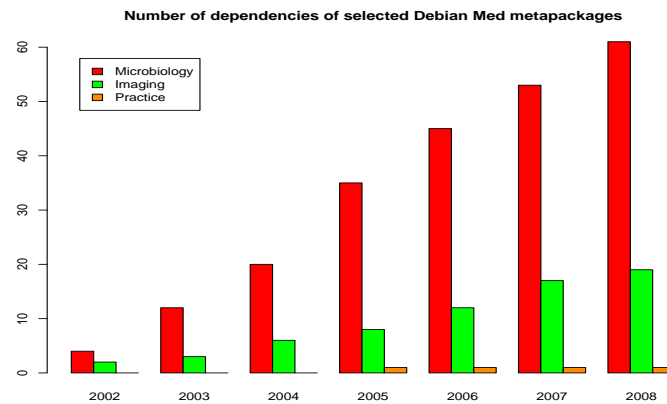


Figure 1: Number of dependencies of selected metapackages

### 3.4.4 Metapackages

On the technical side Debian Med contains a set of metapackages that declare dependencies on other Debian packages, and that way the complete system is prepared for solving particular tasks. So the user has not to deal with the large number of package descriptions of 20000 packages inside the Debian distribution – it is just enough to seek for metapackages starting with prefix `med-` and install the metapackage of choice. The package management system will care for the installation of all packages that are in the list of dependencies of this metapackage - so the user can be sure that all packages he might need for the job will be installed on his system. Once one of the metapackages is installed a special user menu will be created to enhance usability for the user working in the field of medicine.

Currently inside Debian Med applications are provided in certain categories: medical practice and patient management, medical research, hospital information systems, medical imaging, documentation, molecular biology and medical genetics and others.

There are two so called metapackages which are named `med-bio` and `med-bio-dev`. The sense of a meta package is that you have to install only one single package using a package management software inside Debian to get all interesting packages which are necessary for a single task. For instance if a user types in:

```
apt-get install med-bio
```

all applications inside Debian which are related to the field of molecular biology and medical genetics will be installed. The `med-bio-dev` package just installs programming libraries and tools which are interesting for users who want to develop biological applications.

### 3.4.5 Continuous growth

Several Free Software projects which try to deal with small user group software started with a lot of enthusiasm but at some point in time developers had other interests or just were unable to maintain the project because of lack of man power. The strategy of Debian Med is to stay strictly inside Debian – so even if manpower is a problem the whole infrastructure around will stay solid and does not drain extra resources. So

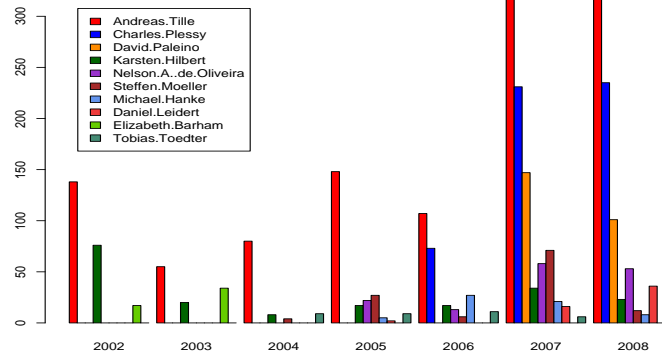


Figure 2: Activity of most active authors on the Debian Med mailing list

nobody of the Debian Med team has to care about writing installers, running an online repository and mirrors work on a bug tracking system etc. All this infrastructure is just there.

The success of this strategy can be proven by a continuous growth for instance if the number of packages inside Debian which is interesting for health care. Taking the number of dependencies of some metapackages into account (see figure 1) at the beginning of the project in 2002 a quite low number of packages useful for medical care was available. A nearly linear growth with a gradient that perfectly reflects the availability of programs in this field can be observed.

This growth of the output of a project is an important part but we also try to measure the commitment of the people involved in the project. It has to be ensured that fresh blood is flooding into the project to make sure it can cope with the normal loss of supporters which always happens in Free Software projects (people find new jobs with different orientation or less spare time for private reasons etc.) A raw measure for the activity of members might be their mails to the project mailing list. Figure 2) shows the number of mails of the ten most active posters of the Debian Med mailing list. This graph shows perfectly that the number of active supporters is growing solidly. In the last two years three people can be considered as very active and there is no dramatic loss of people in the project. The general activity on the mailing list is constantly growing.

### 3.4.6 Role inside Debian

The Debian Pure Blends framework was mainly inspired by the Debian Med needs. Regarding implementation the Debian Edu project – another Blend – is much more advanced. There are several reasons for this fact. The main reasons are the availability of software for education and the fact that a Norwegian company paid developers to work full day on this project. Debian Med tried to gain profit from common technologies and generalised some tools of Debian Edu for all Blends. Moreover inside Debian Med some new Blend tools were developed which enhance the easy build of metapackages. The latest contribution to the Blend framework is a set of automatically generated web pages which can be used by any Blend. For instance there is an overview about all



the so called tasks (fields like biology, medical imaging, practice management) which provides a nice list of all the packages including some metainformation and the description of the packages in this task. So the output of the project can be easily viewed at [the tasks page of Debian Med](#).

By providing this kind of tools for other Blends as well Debian Med has set some cornerstone in the technical cooperation between user oriented projects inside Debian. This in turn awakes the interest of other Debian developers who might provide other tools for Debian Med.

## 4 Discussion

The Debian Med project serves as a common platform for all Free Software that may be utilised in medical care. Tools developed in computational biology is just a part of it because it is an important brick in medical science. With Debian Med's ambition to become the platform of choice for medical work, conform with the principles of the [Debian Policy](#) [Jackson et al., 2005], by the means of the distribution of development within the Debian Society, a well established reference platform for bioinformatics research and its medical applications has evolved and will continue to improve. The organisation is open, both to new members and to external sites offering packages for installations.

### 4.1 Differences from other distributions

The Debian GNU/Linux distribution differs from others in several ways. Firstly, Debian is a non-commercial organisation of volunteers, that does not sell anything. The second and most appealing difference is the peer review and continuous pressure among the members to provide a high quality of packages. The Debian society has a constitution, elects its leader, and transparently describes policies for the creation of packages utilising specific technologies.

With these principles, Debian achieved the largest collection of ready-to-install Free Software on the Internet.

#### 4.1.1 Importance of community support

That strong support within the community of users is essential for the development of software, for quality assurance, feedback on features, and not at least for the motivation of staff, all commercial distributors are well aware of. E.g., RedHat has initiated Fedora as a free supplement to their commercial distribution. It is this reason why Debian Med is part of Debian and why groups external to the Debian society, like BioLinux, are also keen on close collaborations with the community.

#### 4.1.2 Road map to come closer to Debian Med

1. Join the [Debian Med mailing list](#).
2. Check what projects are missing and ask Debian maintainers for official inclusion. There is a sponsoring program by which even non Debian developers can provide packages which are checked and uploaded by official maintainers. There is no point in keeping good quality softwares outside of Debian.

3. [The tasks page of Debian Med](#) mentioned above does not only contain the work that was done – it contains also a nicely formatted list which projects would be interesting to reach the final goal to cover each task in medical care with Free Software. This list inside each section might give some idea where help might be needed.
4. Verify whether one needs special configuration for your project. If yes, verify which possibilities are given in the Debian Pure Blends effort. It is more than collecting software but bringing the software to your target users while taking the burden from any configuration issues from his back.
5. The only reason to keep things outside of Debian are licenses which are not compatible with DFSG. All other parts of your projects can be included and your time for everyday package building tasks can be saved and the workload shared with other people following the same road.

## 5 Conclusions

We have shown that there is a considerable heterogeneous shape of Free Software for medical care even if some fields like micro biology are better covered than others. The continuous updates of data and the addition of novel important tools for a general medical environment cannot be performed by a single maintainer. The adherence to a policy and the sharing of maintenance are basic technologies to allow inter-institutional software projects of different kind in health care.

Debian and its special dedication to medical software in Debian Med, but also the technical infrastructure behind this community project renders a comfortable solution. The volunteers behind Debian Med strive to support everybody's specific projects as best as they possibly can. It is the particular challenge of users of Free Software, to determine together with the community the available packages that already serve their needs or may be adapted respectively.

For Debian GNU/Linux to become the race-horse for Free Software in health care, further important software which is listed at [Wikipedia list of open source healthcare software](#) [Wikipedia, 2008] like for instance OpenVista[OpenVista, 2008] and other enterprise grade health care information systems has to be packaged for Debian.

## References

- [Hilbert, 2008] Hilbert, K. (2008). Gnumed.
- [Jackson et al., 2005] Jackson, I., Schwarz, C., and Debian (2005). Debian policy manual.
- [Knopper, 2005] Knopper, K. (2005). Knoppix live cd.
- [OpenVista, 2008] OpenVista (2008). Enterprise grade health care information system developed by the u.s. department of veterans affairs (va).
- [University, 2008] University, Y. (2008). Bioimagesuite an integrated image analysis software suite.
- [Wikipedia, 2008] Wikipedia (2008). List of open source healthcare software.
- [Workgroup, 2005] Workgroup, L. (2005). Linux standard base.