

Computer systems

# An open pathology computer system

A Rubin

## Suitable for pathology laboratories?

The open source software movement is an initiative for the collaborative development of software whereby the source code can be distributed and altered freely. It is suggested that open source could provide a method for the development of cheap, robust, secure, and up to date pathology computer systems. Development of pathology systems would be aided by having open, agreed database structures

### THE PROBLEM

Departments of pathology are becoming ever more reliant on computer systems. These systems are used to record, store, transmit, and print pathology reports; calculate results, sometimes using interfaces with the machines used for analysis; and produce statistical analyses of results. Many pathologists may be satisfied with the performance of their computers, but others are frustrated by the inability of their systems to carry out straightforward tasks or to provide functions that should be well within the capability of modern systems. Core functions usually work reliably, but it may be very difficult to take forward developments, even though they ought to be well within the capability of modern systems. Users may be unable to make full use of the data that they have entered on to their systems.

*“Users often prefer sticking with outdated unsatisfactory systems to the considerable risks of plunging into an untried new system”*

There is a risk of departments becoming tied to expensive systems, which, either through deficiencies in software, or in local implementation, do not provide the functionality that would be expected. They become dependent on vendors of systems and find themselves unable to disentangle themselves without considerable expense and disruption. Changes of vendor are extremely difficult, with potential legal disputes regarding ownership of databases, difficulties in loading old data on to new systems, and the expense of maintaining crumbling legacy systems. Support may be limited by the lack of staff with the knowledge to deal with specific

computer systems. Attempts to replace systems go through lengthy procedures of production of specifications and business cases only to be frustrated by changes in local or national health strategy and shortage of funds. Users often prefer sticking with outdated unsatisfactory systems to the considerable risks of plunging into an untried new system. Is there another way of doing things?

### OPEN SOURCE

It is proposed that the open source model would provide an alternative way forward. Open source software is defined by its conformation to the open source licence.<sup>1</sup> The source code for the software and the compiled programs must be available freely, with no royalty or fee for its sale. There is no restriction to the use of the software or its use with other software, and it should not be limited to a particular technology or style of interface.

Open source software is often produced in a collaborative manner. The open nature of the project allows independent peer review with rapid evolution of the source code. Software can be developed that is more reliable and secure than if the source code is closed. Production of software can utilise a greater number of software developers, with a greater knowledge of the requirements of software users. This allows a rapid evolutionary development of software. Because the source is open, development and maintenance are not tied to a particular company, and support can be forthcoming from the entire open source community or any company with the appropriate expertise. Open source does not prevent companies from making money from providing software and services, but they are unable to monopolise systems and a true market can develop.

Open source software can be entirely hardware independent. The software can be distributed as source code, which can be configured and compiled by various hardware systems. Software is often written in the C or C++ languages, which can be compiled on a large variety of machines, or distributed as platform

independent Java classes or using scripting languages such as Perl.

Perhaps the best known example of open source is GNU-Linux. This started in 1991 as a personal project by Linus Torvalds to produce a Unix-like system for use on a personal computer.<sup>2</sup> This has now developed to run on a variety of computer architectures. Linux is being developed by hundreds of programmers all over the world. It is used in software running internet servers such as Sendmail and Apache. For personal computers, Linux offers an alternative system to Windows. It is free, with a large choice of software such as Open Office, GIMP (image manipulation), and Mozilla (web browser). It is stable and at smaller risk of virus and worm attack than other operating systems.

### IS THIS MODEL POSSIBLE FOR PATHOLOGY?

The open source model offers the possibility of longterm, stable, low cost, future proof development, and should be considered for the production of a pathology system, both for the National Health system (NHS) and international use. Such a project probably requires a critical mass of programmers interested in taking it on. Taking into account the costs to the NHS of purchasing computer systems, employment of programmers to develop open source software may be a worthwhile investment. The project would have the advantage of being carried out by, or with the close collaboration of, pathologists who understand their needs best. Development on an open source pathology system would not require the creation of an entire system from scratch. The functions required by pathology systems such as use of databases, interfaces with users, communication with other machines, messaging, and graphical display are common to other systems, and open source software including source code is freely available. This would facilitate the rapid development of a working system.

*“The open source model offers the possibility of longterm, stable, low cost, future proof development”*

The project would require sufficient interest by the managers of pathology departments in adopting such systems, without which sufficient momentum would not be maintained to take such a project forward to a successful conclusion. There would also be a requirement for support from local information technology departments in setting up, customising, and running open software systems. Local compilation of open code enables optimisation for local

circumstances but requires expertise. Where there is insufficient expertise within hospital trusts it should be possible to purchase support from commercial companies. Thus, the role of companies in pathology computer systems would change from one of supplying and running their own copywrite protected systems to configuring, compiling, installing, and running open source systems. Because the system would be open, it would be easier for trusts to replace a support company whose performance was unsatisfactory, and a true market in system support could develop.

## SECURITY

A major concern in all computer systems is that of security. Pathology systems contain confidential and sensitive information, and systems must be protected against attempts to extract information by those unauthorised to do so. In addition, there is a potential for those attacking systems to run malicious code, which could damage valuable data or otherwise disable them. It has been suggested that it would be easier to break security when system code is freely available, or even that damaging code can be deliberately inserted at source. However, computer hackers do not require a knowledge of source code to attack systems. Open source may indeed enhance security by allowing more people to examine the code for vulnerabilities and by promoting a more open and critical approach to potential dangers.<sup>3</sup> Studies suggest that open source operating systems have less vulnerability to attack than do closed systems.<sup>4</sup> This is one reason why such systems are used increasingly for web servers. The closed nature of Windows code has not protected it from attack, whereas it is alleged that secretiveness may have delayed the recognition and

admission of the presence of security flaws.

**“Studies suggest that open source operating systems have less vulnerability to attack than do closed systems”**

Of more relevance to security is not the availability of source code, but rather the ability of at least some of those developing and checking software to understand security issues and write secure programs. Security and all other aspects of the functioning of these systems would require thorough testing before they were adopted. Such testing should be all the more transparent for the software being open.

## THE IMPORTANCE OF DATA STANDARDS

One problem that plagues commercial computer pathology computer systems is the variety of ways in which data are stored. Lack of conformity in methods of data storage is a prime reason for departments finding themselves trapped in a relationship with system providers as if in an unhappy marriage. It is difficult to port data from one system to another, and users are frustrated that they are unable to use data that they know are on the system. A valuable early aim in the production of an open source pathology system could be the development of open and preferably standardised database structures. With this agreed, programmers can develop alternative software solutions in different languages, allowing evolution of the fittest software for the problem posed. Users may be able to choose the software that they prefer without issues regarding porting of data. Even if open source software is not developed, agreeing an open database structure would

assist in the development and portability of pathology computer systems.

Of particular note in regard to the development of an open database structure is the availability of open database servers such as MySQL.<sup>5</sup> MySQL is freely available under the open source licence, and is supported by a commercial company founded by its developers. MySQL is a secure, fast, multi-threaded, multi-user, and robust SQL system. It works successfully on a large number of operating systems including DEC Unix, FreeBSD, Linux, Mac OS, Solaris, SunOS, and Windows. Servers such as MySQL offer the possibility of developing a hardware independent pathology database to be used by an open source project.

## CONCLUSION

Open source software development offers the possibility of the rapid development of robust, well supported, secure, hardware independent software systems for pathology laboratories. Whether this can be taken up would depend on the degree of interest in collaboration in its development and in the willingness for pathology departments to adopt its use.

*J Clin Pathol* 2004;**57**:1252–1253.  
doi: 10.1136/jcp.2004.020826

Correspondence to: Dr A Rubin, West Hertfordshire Hospitals NHS Trust, Department of Cytology, Watford General Hospital, Vicarage Road, Watford, Hertfordshire WD18 0HB, UK; alan.rubin@whht.nhs.uk

## REFERENCES

- 1 **The Open Source Initiative** (<http://www.opensource.org>).
- 2 **Cantrell D, Johnson L, Lumens C.** Slackware Linux essentials (<http://www.slackware.com/book/>).
- 3 **Wheeler DA.** Secure programming for Linux and Unix HOWTO (<http://www.dwheeler.com/secure-programs/>).
- 4 **Wheeler DA.** Why open source software/free software (OSS/FS)? Look at the numbers! ([http://www.dwheeler.com/oss\\_fs\\_why.html](http://www.dwheeler.com/oss_fs_why.html)).
- 5 **MySQL** (<http://www.mysql.com>).