International
Open
Source
Network

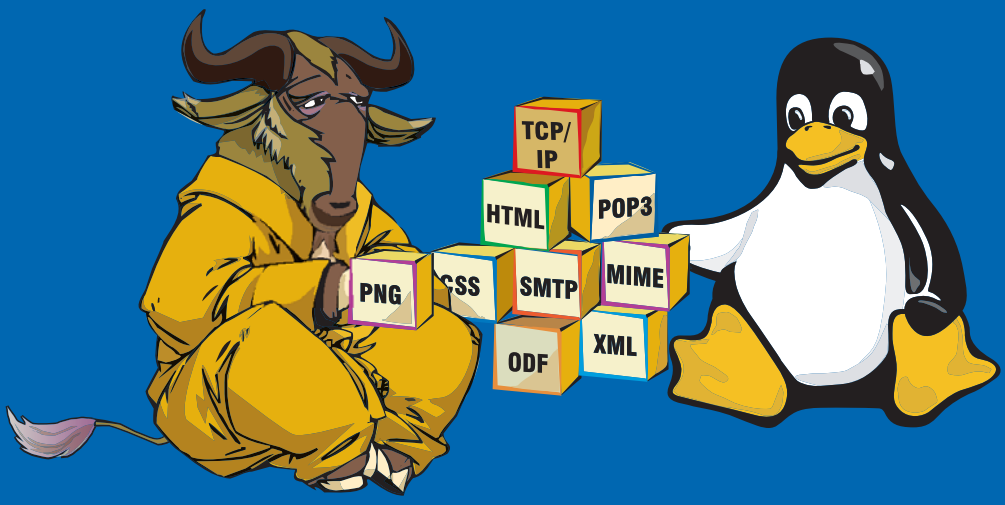
ELSEVIER

UNDP

Free/Open Source Software

Open Standards

Nah Soo Hoe



Foreword by

PETER J. QUINN

Asia-Pacific Development Information Programme
e-Primers on Free/Open Source Software

Free/Open Source Software Open Standards

Nah Soo Hoe

Asia-Pacific Development Information Programme
e-Primers on Free/Open Source Software

Web: www.apdip.net
Email: info@apdip.net

ELSEVIER
A division of
Reed Elsevier India Private Limited
17A/1, Lajpat Nagar IV,
New Delhi 110 024
Tel: +91 11 2644 7160
Fax: +91 11 2644 7156
Website: www.asiaelsevier.com

ISBN-13: 978-81-312-0538-9
ISBN-10: 81-312-0538-X

Academic Press, Butterworth-Heinemann, Digital Press, Focal Press, Morgan Kaufmann, North Holland and Pergamon are the Science and Technology imprints of Elsevier.

Printed and bound in India.

TABLE OF CONTENTS

FOREWORD

One promise of the United Nations (UN), when it was founded in 1948, was to be the voice of those less advantaged and disenfranchised. And within that promise was the thought that the UN could and would lobby on behalf of those who could not act on their own behalf. The United Nations' history is rich with successful intervention, especially for the aged and children. I am reminded of this continued work as I leave Sydney, Australia, late in March, 2006, having seen the Bears of the World on behalf of UNICEF. This decorative display brings a smile to all who move about this collage of bears representing each world nation with unique and vibrant colour schemes. Each individual or group seeks its homeland of today or the heritage of their ancestors. And it is all great fun on behalf of humanity's most vulnerable constituents, children. UNICEF remains UN's most cherished and successful endeavour.

Approaching its 60th birthday, UN has undertaken another very meaningful and important initiative that seeks to continue its original mandate of again assisting those who are in need of help. Through its International Open Source Network (IOSN) initiative via the Asia-Pacific Development Information Programme (APDIP), the United Nations Development Programme (UNDP) is again enabling those governments who cannot necessarily create their own ICT ecosystem, by helping to reduce costs.

Governments throughout the world are and will remain cash challenged. In the USA, healthcare costs are increasing 8 to 15 percent a year, with much of that burden falling on both the states and the federal government. Emerging nations and mature governments throughout the world face the same daunting challenges. The UNDP, through its IOSN initiative, will enable widespread adoption of Free/Open Source Software (FOSS) and create at least one avenue for nations to deal with the "unsustainable cost of government".

FOSS, based on Open Standards, provides governments with a sustainable and sufficiently robust software model that fosters collaboration and ever-increasing innovation. In addition to the obvious cost advantage for acquisition, the ongoing operational expense will also contribute to lowering the cost of government. By engaging the worldwide open source community, governments can benefit from each others' efforts and share applications that each have built. Remember, every government does the same things: collects taxes, provides assistance to those citizens in need, register births, deaths, marriages and motor vehicles, houses prison inmates and issues drivers' licenses. So why do individual governments continue to build applications on their own and most often do a very poor job of it? By incorporating the best tenets of FOSS, increased collaboration and innovation, governments working together can create meaningful value for their citizens and for each other.

At no time in modern history has it been more important than now for governments to band together and foster meaningful information exchange. And in this age of information and communications technology, standards and the means by which standards are set become vital. The attributes of open standards and the model for establishing open standards are what will allow for sustainable information exchange, interoperability, and flexibility. Where public funds are concerned, adopted standards should be vendor neutral and open to all to implement without royalties. Otherwise citizens will not be able to consume information produced by the government without having to purchase or pirate software.

FOSS mimics the ubiquity of the Internet and can transcend geographic and political boundaries. Its communal nature unites humanity, helps bridge our numerous divides, and can continually contribute to closing the digital divide.

UNDP's IOSN efforts through FOSS, open standards, and open content represents the best of its mission and legacy. We should all applaud and support this extremely important agenda.

Peter J. Quinn
Former Chief Information Officer
Commonwealth of Massachusetts

PREFACE

This primer is part of a series of primers on Free and Open Source Software (FOSS) from IOSN serving as introductory documents to FOSS in general, as well as covering particular topic areas that are deemed important to FOSS such as open standards. Open standards are not the same as FOSS. However, like FOSS, they can minimize the possibility of technology and vendor lock-ins and level the playing field. They can also play an important role in promoting the interoperability of FOSS and proprietary software and this is crucial in the current, mixed Information technology (IT) environment. Being a primer in the IOSN FOSS series, the issues concerning open standards are approached from the FOSS and software perspectives and emphasis is given to the relationship that some of these standards have with FOSS.

The definition of an open standard has generated much controversy with regard to whether it should contain patents licensed under reasonable and non-discriminatory (RAND) terms. The FOSS community, in general, is of the view that such RAND-encumbered standards should not be considered as open standards but most of the standards development organizations and bodies do accept patents available under RAND terms in their standards. The primer has incorporated definitions of open standards from both sides and also put into perspective the minimal characteristics that an open standard should have.

It is hoped that this primer will provide the reader with a better understanding as to why open standards are important and how they can complement FOSS in fostering a more open IT environment. As users and consumers, the readers of this primer should demand from their software, conformance to open standards as far as possible. In addition to promoting interoperability and making more choices available, this will make it easier for FOSS to co-exist and take root in environments filled with proprietary software.

Nah Soo Hoe
March 2006

LIST OF ACRONYMS

AAC	Advanced Audio Coding
ABI	Application Binary Interface
AMD	Advanced Micro Device
ANSI	American National Standards Institute
APDIP	Asia-Pacific Development Information Programme
API	Application Program Interface
ARPANET	Advanced Research Projects Agency Network
ASCII	American Standard Code for Information Interchange
ASF	Advanced Systems Format
AVC	Advanced Video Coding
AVI	Audio Video Interleave
BIS	Bureau of Indian Standards
BSA	Business Software Alliance
BSD	Berkeley Software Distribution
BSI	British Standards Institute
ccTLD	Country Code Top-Level Domain
CD	Compact Disc
CDDL	Common Development and Distribution License
CE	Consumer Electronics
CII	Computer-Implemented Invention
COTS	Common Off The Shelf
DNS	Domain Name Service
DOC	Document
DOS	Disk Operating System
DTD	Document Type Definition
DVD	Digital Video Disc
EC	European Commission
ECMA	European Computer Manufacturers Association
EEE	Embrace, Extend and Extinguish
e-GIF	e-Government Interoperability Framework
EICTA	European Information, Communications and Consumer Technology Industry Association
EIF	European Interoperability Framework
ELF	Executable and Linking Format
EPC	European Patent Convention
EPO	European Patent Office
ETRM	Enterprise Technology Reference Model
ETSI	European Telecommunications Standards Institute
EU	European Union
FAQ	Frequently Asked Questions
FAT	File Allocation Table
FHS	Filesystem Hierarchy Standard
FLAC	Free Lossless Audio Codec
FOSS	Free/Open Source Software
FSF	Free Software Foundation
FSG	Free Standards Group
G2B	Government to Business
GIF	Graphics Interchange Format
GL	Graphics Library
GNOME	GNU Network Object Model Environment
GNU	Gnu's Not Unix
GPL	General Public License (GNU)
GSM	Global System for Mobile Communications

GUI	Graphics User Interface
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
HTTPd	Hypertext Transfer Protocol Daemon
I18N	Internationalization
IA	Internet Architecture
IAB	Internet Architecture Board
IANA	Internet Assigned Numbers Authority
ICT	Information and Communications Technology
ICTU	Dutch Organization for Information and Communications Technology
IDABC	Interoperable Delivery of European e-Government Services to Public Administrations, Businesses and Citizens
IEC	International Electro-Technical Commission
IEEE	Institute of Electrical and Electronics Engineers
IEEE-SA	Institute of Electrical and Electronics Engineers Standards Association
IESG	Internet Engineering Steering Group
IETF	Internet Engineering Task Force
IJG	Independent JPEG Group
IMAP	Internet Message Access Protocol
IOSN	International Open Source Network
IP	Internet Protocol
IPR	Intellectual Property Rights
IS	Information System
ISO	International Organization for Standardization
ISOC	Internet Society
ISV	Independent Software Vendor
IT	Information Technology
ITU	International Telecommunication Union
ITU-D	International Telecommunication Union – Development Sector
ITU-R	International Telecommunication Union – Radio Communication Sector
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector
JFIF	Joint Photographic Experts Group File Interchange Format
JPEG	Joint Photographic Experts Group
JTC	Joint Technical Committee
KDE	K Desktop Environment
L10N	Localization
LAN	Local Area Network
LCID	Locale Identifier
LDAP	Lightweight Directory Access Protocol
LSB	Linux Standard Base
LZW	Lempel-Ziv-Welch
M17N-lib	Multilingualization library
MAN	Metropolitan Area Network
MIME	Multipurpose Internet Mail Extensions
MLP	Mozilla Localization Project
MNCC	Malaysian National Computer Confederation
MOV	Movie File Format
MP3	Motion Pictures Experts Group-1 Audio Layer 3
MP4	Motion Pictures Experts Group-4 Part 14
MPEG	Motion Pictures Experts Group
MPL	Mozilla Public License
MS	Microsoft
MUI	Multilingual User Interface
MyGIF	Malaysian Government Interoperability Framework
NCSA	National Center for Supercomputing Applications
NFS	Network File System
OASIS	Organization for the Advancement of Structured Information Standards
ODP	OpenDocument Presentation
ODS	OpenDocument Spreadsheet
ODT	OpenDocument Text

OOo	OpenOffice.org
OpenI18N	Open Internationalization
ORG	Organization
OS	Operating System
OSDL	Open Source Development Labs
OSI	Open Source Initiative
OSOSS	Open Standards and Open Source Software
PAG	Patent Advisory Group
PAS	Publicly Available Specification
PC	Personal Computer
PCL	Printer Control Language
PDA	Personal Digital Assistant
PDF	Portable Document Format
PIKOM	Association of the Computer and Multimedia Industry of Malaysia
PNG	Portable Network Graphics
POSIX	Portable Operating System Interface for UNIX
PPC	Power Personal Computer
PPT	PowerPoint
RAND	Reasonable and Non-Discriminatory
RF	Royalty-Free
RFC	Request for Comments
RIFF	Resource Interchange File Format
RPM	Red Hat Package Manager
RTF	Rich Text Format
SDO	Standard Development Organization
SGML	Standard Generalized Markup Language
SMPTE	Society of Motion Picture and Television Engineers
SMTP	Simple Mail Transfer Protocol
SOAP	Simple Object Access Protocol
SQL	Structured Query Language
SSL	Secure Sockets Layer
SSO	Standard-Setting Organization
STD	Standard
SVG	Scalable Vector Graphics
SVID	System V Interface Definition
SXC	StarOffice Calc
SXI	StarOffice Impression
SXW	StarOffice Writer
TCP	Transmission Control Protocol
TIFF	Tagged Image File Format
TV	Television
UCS	Universal Character Set
UDDI	Universal Description, Discovery, and Integration
UFO	Uniform Fee Only
UNDP	United Nations Development Programme
UNICODE	Unique, Universal and Uniform Character enCoding
USA	United States of America
UTF	Universal Character Set Transformation Format
VC	Video Codec
VCD	Video Compact Disc
VHS	Video Home System
VMS	Virtual Memory System
W3C	World Wide Web Consortium
WAV	WAVE form Audio Format
Wi-Fi	Wireless Fidelity
WMA	Windows Media Audio
WMV	Windows Media Video
WWW	World Wide Web
XLS	Excel
XML	Extensible Markup Language
XPath	Extensible Markup Language Path Language

XPM	XPixMap
XSL	Extensible Stylesheet Language
XSL-FO	Extensible Stylesheet Language Formatting Objects
XSLT	Extensible Stylesheet Language Transformations

INTRODUCTION

What are Standards and Why are They Important?

The word "standard" has several different meanings. Working within the context of the subject matter of this document, its meaning in everyday usage[1,2] can be taken to refer to:

- a level of quality or attainment, or
- an item or a specification against which all others may be measured.

In technical usage, a standard[3] is a framework of specifications that has been:

- approved by a recognized organization, or
- is generally accepted and widely used throughout by the industry.

For the rest of this document, unless specified otherwise, when the word standard is used the technical meaning is implied.

Standards are extremely important in modern society. They ensure that products and services are of adequate quality and that they can interoperate and work together even though they may be from different parties or entities. Ultimately, they raise levels of quality, safety, reliability, efficiency and interoperability, and provide such benefits at an economical cost.[4]

In the IT industry, standards are particularly important because they allow interoperability of products, services, hardware and software from different parties. Without standards, users may be forced to use only hardware and software or services from one party or vendor. Internationally recognized standards define common interfaces and any changes or modifications in the standards are usually carried out by common agreement. For example, the Internet would not achieve its current ubiquitous presence, where it is accessible from almost any type of computer platform and device, if it did not use widely accepted technical standards in its networking infrastructure and supported services.

Open Standards

Having defined what standards mean in general and technical usage, let us turn our attention to the main focus of this primer – open standards. There are many differing opinions on what constitutes open standards.[5,6,7,8,9,10]

Definition of Open Standards

Well-known Open Source exponent Bruce Perens argues that an open standard is more than just a specification, and that the principles underlying the standard and the practice of offering and operating the standard are what make the standard open.[5] He proposes that open standards should follow the principles of availability, and maximize end-user choice. In addition, there should be no royalty, no discrimination, extension of subset and predatory practices, and certain practices should be followed to ensure that these principles are adhered to. The Perens definition has found wide acceptance among the FOSS communities worldwide.

Principles of Open Standards – Bruce Perens

Bruce Perens has proposed the following principles for open standards. According to his definition,[5] an open standard is more than just a specification. The principles behind the standard, and the practice of offering and operating the standard, are what make the standard open. The principles

listed by Bruce Perens are reproduced below.

Principles

Availability: Open standards are available for all to read and implement.

Maximize end-user choice: Open standards create a fair, competitive market for implementations of the standard. They do not lock the customer into a particular vendor or group.

No royalty: Open standards are free for all to implement, with no royalty or fee. Certification of compliance by the standards organization may involve a fee.

No discrimination: Open standards and the organizations that administer them do not favour one implementor over another for any reason other than the technical standards compliance of a vendor's implementation. Certification organizations must provide a path for low- and zero-cost implementations to be validated, but may also provide enhanced certification services.

Extension or subset: Implementations of open standards may be extended, or offered in subset form. However, certification organizations may decline to certify subset implementations, and may place requirements upon extensions (see Predatory Practices).

Predatory practices: Open standards may employ license terms that protect against subversion of the standard by embrace-and-extend tactics. The licenses attached to the standard may require the publication of reference information for extensions, and a license for all others to create, distribute and sell software that is compatible with the extensions. An open standard may not otherwise prohibit extensions.

Practice

Recommended practices for offering and operating each of the principles above have also been discussed by Bruce Perens. (The interested reader should check the reference cited[5] for these.)

The Open Standards Policy of the State of Massachusetts, USA[6] defines it as specifications for systems that are publicly available and are developed by an open community and affirmed by a standards body. The European Commission's European Interoperability Framework (EIF)[7] adds on the requirements that open standards should be available either for free or at a nominal charge for usage, copying and distribution and any patents present are to be made irrevocably available on a royalty-free basis, and there should be no constraints on the re-use of the standard.

Thus, the Perens definition is consistent generally with those of the EIF and the State of Massachusetts, and this approach has gained currency with many policy makers as the basis of their open standards policies.

Other organizations such as the American National Standards Institute (ANSI), the International Telegraph Union Telecommunication Standardization Sector (ITU-T) and the Business Software Alliance (BSA) have also come out with their definitions and policies on open standards. While all of these recognize that open standards have to be publicly available for implementation and participation in development by interested parties, they also recognize the inclusion of essential intellectual property rights (IPR) so long as these IPR can be made available under non-discriminatory terms and for a reasonable fee or for no fee at all.

Thus, we find that while there may be numerous detailed definitions and meanings given to open standards, in general, they all satisfy the following characteristics:

- easy accessibility for all to read and use;
- developed by a process that is open and relatively easy for anyone to participate in; and
- no control or tie-in by any specific group or vendor.

Examples of open IT standards are:

- the Transmission Control Protocol/Internet Protocol (TCP/IP) suite of networking protocols from the Internet Engineering Task Force (IETF);
- the Hypertext Transfer Protocol (HTTP) service protocol from the World Wide Web Consortium (W3C) and the International Organization for Standardization (ISO);
- the Unicode coding standard from the Unicode Consortium and ISO; and
- the Portable Operating System Interface for UNIX (POSIX) portable operating system interface from the Open Group, the Institute of Electrical and Electronics Engineers (IEEE) and ISO.

Many organizations as well as governments are starting to emphasize that their IT usage follow or adhere to open standards as far as possible as they now realize that by implementing open standards they can have more flexibility in their choice of technology, vendor and solutions. In an increasingly complex and heterogeneous IT environment, no single technology or vendor can offer solutions on everything and so the ability to mix and match and to interoperate is of critical importance. Information is now exchanged and stored electronically as never before. It is only by following open standards in the exchange and storage/retrieval of the data that an organization can be assured of access to that data, both now and later when the technology or vendor may be long gone.

Some Other Definitions of Open Standards

American National Standards Institute

ANSI has described open standards[8] as those standards that are developed by a process where there is consensus by a group or "consensus body" that is open to representatives from all materially affected and interested parties and there is consideration of and response to comments submitted by voting members of the relevant consensus body as well as by the public. There should also be broad-based public review and comment on draft standards. An avenue for appeal is available for participants who feel that the ANSI open standards principles were not respected during the standards-development process.

In addition, ANSI tries to balance the interests of the implementors and users of the standard with the parties who own IPRs that are essential to implement the standard by allowing the payment of reasonable license fees and/or other reasonable and non-discriminatory license terms that may be required by the IPR holders.

Business Software Alliance

BSA, an industry trade association representing commercial software providers, has specified a set of characteristics for open standards[9]. Going by this, an open standard should be published without restriction and in sufficient detail to enable a complete understanding of the standard's scope and purpose and should be publicly available without cost or for a reasonable fee for adoption and implementation. Any patent rights necessary to implement it are to be made available by those developing the specification to all implementors on reasonable and non-discriminatory (RAND) terms (either with or without payment of a reasonable royalty or fee).

International Telegraph Union Telecommunication Standardization Sector

ITU-T has adopted a definition of open standards[10] that reflects the same key elements as ANSI. They define open standards as standards that are made available to the general public and are developed (or approved) and maintained via a collaborative and consensus driven process. These standards should be developed using a collaborative and transparent consensus driven process that is reasonably open to all interested parties.

Intellectual property essential to implement the standard should be licensed to all applicants on a worldwide, non-discriminatory basis, either for free and under other reasonable terms and conditions or on reasonable terms and conditions, which may include monetary compensation.

It should be noted here that wide usage of a standard does not necessarily mean that it is open. Numerous examples are found in the IT industry (e.g. the Portable Document Format or PDF from Adobe Inc., the

Powerpoint presentation file format from Microsoft), where some technology or file/data format associated with a popular product is very widely used, so much so that it becomes a *de facto* standard, i.e. a standard established through widespread usage and acceptance in the industry. However, because this is very often based on a technology by a specific party (vendor or close group) and is under the control of this party, it does not qualify as an open standard. There are potential pitfalls in adopting this as a standard as there is no open mechanism for the user to participate in its development and no guarantee that the party in control will not try to lock in users into its product or technology. In some cases, the owner of the product/technology may agree to submit it to an internationally recognized standards-setting body and in so doing it may then become an open standard.

At the same time, however, governments and organizations should be open to the possibility that some such *de facto* standards, like PDF for example, are so widely and reasonably licensed and so broadly deployed that it makes sense to temporarily support such standards as part of a government's or organization's interoperability programme. To completely ignore such standards as a rule may actually impair interoperability, particularly if there is no adequate open standard substitute for such a *de facto* standard

Open Standards and FOSS

Many people are confused between the terms open standards and FOSS thinking that they are one and the same or one cannot exist without the other. To be consistent with other publications by IOSN, the term FOSS will be used in this document to refer to open-source software and/or free software. Unless otherwise stated open standards is not the same as FOSS, which refers to software that follows certain principles in its creation, modification, usage, licensing and distribution.[11] In particular, it should have the four fundamental freedoms:

- freedom to run the program, for any purpose;
- freedom to study how the program works, and adapt it to your needs;
- freedom to redistribute copies so you can help others; and
- freedom to improve the program, and release your improvements to the public.

The *FOSS: A General Introduction* primer[12] from IOSN may be referred for more background and details on FOSS.

FOSS is software whereas open standards refer to standards – two different things altogether. The processes and issues involved in developing software and a standard are also very different. It is entirely possible for a functionality in a non-FOSS software (often called proprietary software) to be implemented following an open standard. Open standards are neutral with regard to software licensing or business models and so it is equally possible for an open standard to be implemented in proprietary software as it is in FOSS. For example, proprietary software like the Microsoft Windows operating system can still implement the TCP/IP networking protocols following the open standards from IETF and be compliant with them.

Widespread usage of standards, and especially open standards, is very important to FOSS. It makes it easier for FOSS to be compatible with proprietary software. It is a practical reality that FOSS needs to co-exist with proprietary software and that compatibility with the proprietary platforms is facilitated if standards are adhered to by all. If all software products were to follow standards strictly, they should be able to interoperate and communicate among themselves well and data files could be read and written transparently. While both proprietary and open standards may allow this to happen, the latter are preferred by the FOSS community as they facilitate free access, open development and participation.

FOSS support may be difficult in cases where a proprietary specification is not publicly published but needs to be licensed. In the past, one way to work around this problem is to reverse engineer some proprietary product that implements the specification or protocol but, in recent times, more and more proprietary licenses have specifically forbidden this. In some countries, legislation has also been passed (e.g. the Digital Millennium Copyright Act in the USA) that makes it illegal to reverse engineer a product if it is deemed that the process can assist in the circumvention of measures implemented to protect against illegal copying of the product. These developments have re-enforced the important role that open standards play in ensuring that FOSS can interoperate well with proprietary software. The emergence

of FOSS and the open standards that it uses highlight the needs and benefits of open standards in a world where interoperability is required.

FOSS has also benefited much from open standards in that the current widespread usage and popularity of FOSS owes much to the Internet and the open standards that the Internet utilizes. While programmers (and many users who write their own programs) have been freely exchanging programs with source code since the early days of the computer, it was only after the Internet explosion in the 1990s that the idea and culture of FOSS became widely known and accepted by the mainstream IT industry. FOSS that implements the open standards and protocols used on the Internet like TCP/IP, HTML, Simple Mail Transfer Protocol (SMTP), etc, were easily available and many people as well as organizations started to use these. From there they became aware of FOSS and the latter grew in strength and acceptance as more and more used it and contributed towards it.

Some may argue that the freedom in FOSS for anyone to modify the software source code will allow and may even encourage the inclusion of code that does not conform to published standards. This is possible but in practice it is seldom done (i.e. modifying a FOSS mainstream product to make it non-compliant with an open standards and redistributing the modified software). Also FOSS project owners guard against this as they realize that it is to the advantage of FOSS if open standards are adhered to as much as possible. In fact, it is very natural for FOSS to promote the adoption of open standards since the ideals and development model of FOSS itself encourages availability, openness and participation by all – the very traits and characteristics of an open standard.

FOSS is Useful for Popularizing Open Standards

FOSS can play a useful role in popularizing an open standard. A FOSS implementation of a standard usually results in an open and free-working reference implementation. A lot of the benefits of open standards are negated, if its only implementation is a closed and proprietary one. The availability of a FOSS implementation will spur quicker adoption and acceptance of the standard as everyone has easy access to the implementation of the standard and so can try and test it out. A very good example of this is the Internet HTTP standard. One reason why this service became universally accepted is that very early on there were free and open implementations of both the HTTP server (e.g., National Center for Supercomputing Applications or NCSA HTTPd, Apache) and client (e.g., NCSA Mosaic).

Focus of the Primer

This primer is part of a series of primers on FOSS from IOSN serving as introductory documents to FOSS in general, as well as covering particular topics that are deemed important to FOSS in greater detail. As such, the issue of open standards is approached from a FOSS perspective and emphasis is given to the relationship that some of these standards have with FOSS. While open standards are available and important for both hardware and software, the examples and references given in this primer focus mainly on standards related to software.

IMPORTANCE AND BENEFITS OF OPEN STANDARDS

The benefits of using open standards have been alluded to in the Introduction. Here we shall delve into more details on the importance and benefits of open standards.

Benefits of Using Open Standards

Numerous benefits are obtained if an organization ensures that its technological and IT procurements and implementations follow open standards as far as possible. First and foremost, there is less chance of being locked in by a specific technology and/or vendor. Since the specifications are known and open, it is always possible to get another party to implement the same solution adhering to the standards being followed. Another major benefit is that it will be easier for systems from different parties or using different technologies to interoperate and communicate with one another. As a result, there will be improved data interchange and exchange. It will not be necessary to use the same software or software from a particular vendor to read or write data files. For example, if a multinational organization requires that all its offices worldwide use office software applications that can read and write files using the Open Document format – an open, standardized XML-based file format from the Organization for the Advancement of Structured Information Standards (OASIS).[13] An individual office will have the flexibility of using whatever office software that is best suited for it and at the same time be able to read, write and exchange documents with other offices in the organization.

Using open standards will also offer better protection of the data files created by an application against obsolescence of the application. If the data file format used is proprietary then, should the application become obsolete, the user may have a difficult time converting the data files to another format needed by a new application. However, if the data format follows an open standard and, hence, is known, either the new application will be able to use it as it is or it will be easier to convert the data so that the new application can use it.

It stands to reason that if a user demands that open standards are adhered to, there will be more choices available as more vendors can participate to offer solutions and it may be possible to even mix and match solutions from multiple vendors to provide best-of-breed solutions as far as possible.

If open standards are followed, applications are easier to port from one platform to another since the technical implementation follows known guidelines and rules, and the interfaces, both internally and externally, are known. In addition to this, the skills learned from one platform or application can be utilized with possibly less re-training needed. This can be contrasted with the usage in applications of proprietary standards that are not openly published and where there is inadequate information publicly available about them.

The benefits obtained with respect to using data and file formats whose specifications are publicly published and widely accessible cannot be over-emphasized, especially with respect to an organization that possesses huge amounts of data stored electronically. A national government is a good example of such an organization. If the data formats are not known or easily available, the organization may find it difficult to migrate or change its information systems since it can be prohibitively expensive or even impossible to convert data files.

National Considerations

From the national viewpoint, the usage of open standards by a government is even more important. In this information age, a government will need to use IT solutions to ensure that it has adequate and reliable information to enable it to govern the country effectively. It is vital that these IT implementations make use of standards that are open as far as possible. In cases where open standards are not available, the government may want to consider other standards that are freely available for usage and implementation. It should also take into consideration how open these standards are and whether they

have the possibility of becoming open standards later.

All this can help ensure that there is less likelihood of its information systems being locked in later by any single technology or product. It is also in the interests of national security that open standards are followed to guard against the possibility of over-reliance on foreign technologies/products. Imagine the implications to a sovereign nation if the electronic records of its citizens are kept in databases that can be accessed readily only by proprietary software from a foreign vendor or the documents of the government are kept in a format that belongs to a vendor who thus has total control over its accessibility both now and in the future.

e-Government Projects Specify Open Standards

Many countries have started on e-government projects or initiatives, most of which have policies stating that, as far as possible, open IT standards and specifications are to be followed. Countries that have such policies include Norway, Denmark, the United Kingdom, the Netherlands, France, Brazil, Australia, New Zealand, and Malaysia.

The European Union's EIF, a framework to facilitate the interoperability of its member countries' e-government services, recommends the use of open standards for maximum interoperability.

In addition, more and more public sector agencies all over the world have adopted or are considering adopting policies that require open standards.

Another important national benefit is that open standards will make it easier and, in some cases, the only possible means for local companies to participate as major players in supplying services and solutions to the government. This is because the local companies usually lack the strength and resources of multinationals and may be strong only in certain areas or solutions. The government can leverage open standards to mix and match solutions from different suppliers in order to give the local suppliers a chance.

It is a reality in the IT world that the main language used and supported by all mainstream software is English and because of this it is sometimes difficult to produce electronic documents in another language. The availability of an open character coding standard, Unicode,[14] designed to support the worldwide interchange, processing, and display of the written texts of diverse languages makes it feasible for the translation and localization of software and electronic office documents for nations or cultures whose language is not English.

Embrace, Extend and Extinguish Tactics

Much has been said in this document of how open standards can prevent product lock-ins by a particular vendor but users have to be aware that sometimes open standards can be taken advantage of by some vendors. There have been cases whereby particular vendors have tried to exploit open standards (e.g. standards like Kerberos, HTML, SMTP) to their own ends with a view to lock-in customers to their products and/or services by deploying what is termed "Embrace, Extend and Extinguish (EEE)" tactics.[126, 127]

Embrace

The vendor, first of all, announces that it will support a particular open standard in its products and it may even contribute resources to the development of the standard. It then implements the standard in its product and markets them.

Extend

In the implementation of the standard, the vendor adds in proprietary enhancements to the specifications of the standard, claiming that these are needed to address customer needs or to differentiate its products from the competitors. These will be made usually in areas where the standard is silent or where the specifications are not well defined. While some standards do provide some leeway for different implementations to differentiate themselves, it is important that the enhanced implementation be done such that a basic implementation can still interoperate with it. A vendor that is using EEE tactics will not ensure this and as a result, products from other sources

may not be now compatible with this vendor's products. The problem really arises if the vendor's products are widely used. If that is the case, other implementations of the standard may have to be modified so as to make them compatible with this enhanced implementation since the latter is dominant.

Extinguish

After some time, if the enhanced implementation of the standard becomes so widely used that the majority of implementations support it, this implementation effectively becomes the *de facto* standard instead. Since the enhancements are proprietary, the vendor has now essentially hijacked the open standard and made it proprietary.

Particular Benefits of Open Standards

Open standards are particularly beneficial to some IT activities or services. Some of these are examined in greater detail here.

File Formats

Modern information systems generate data (lots of it in many cases) that have to be stored in some form of electronic file formats for efficient storage, retrieval and exchange. If their specifications are not publicly known, only software and systems from the owner of these proprietary formats can readily access them. Also, the exchange of information is essential to the functioning of modern society. This exchange will be severely hampered if non-open file formats are utilized as products from one vendor may not be able to retrieve, access or store the information from the products of another vendor properly.

In some cases, while the format may be known, it may be the property of a particular party and this party may control the way the format evolves or is used. In such cases, users can have very little say or control over the format. Also it may be possible that the owner may not publish the format specifications at a later stage for a new version. So while compatible systems can be created that can access the files now, there is no guarantee of this when a newer version comes out. In addition, there have been cases where, when a proprietary format becomes popular and is widely used by the industry, the owner of the format starts to impose restrictions like charging a fee or royalty charges (if it is patented) for using the format at a later stage. The case of Microsoft attempting to charge flash drive makers and manufacturers of devices, such as digital cameras, a licensing fee for using its File Allocation Table or FAT file format[15] is a good example of this.

All this shows that it is of utmost importance that electronic file formats should follow some specifications that are accessible to all interested parties and also be developed by processes that are open and easy for any party to participate. In other words, they should be implemented using open standards. It is vital in today's information-centric society that the data from which information is derived can be stored and exchanged following standards that are open so that no single party or even group can control the access to this data.

Office Applications

This lack of complete compatibility between documents created using MS Office and the competing alternatives has prevented some users from using or migrating to the latter. This effectively results in a specific product/vendor lock-in.

This example illustrates that open and standardized file formats are needed to give users the flexibility and freedom to choose and use products from different vendors and to prevent them from being locked in to a specific product and/or vendor. The recently published OpenDocument standard[13] from OASIS for office applications offers this. Currently, applications that support this open format include StarOffice, KOffice, IBM Works, Abiword and OpenOffice.org. Microsoft does not support this but instead it has come up with its own XML-based file formats for its office suite. Again, while the MS Office XML schemas are publicly published and licensed for use royalty-free, they are owned by a single vendor (Microsoft) and hence are subject to the potential abuse discussed previously for non-open formats. In an attempt to allay fears over this and acceding to the requests of some of its biggest customers, the Microsoft

Office XML file formats have been submitted to European Computer Manufacturers Association (ECMA) International for development as formal standards.

Internet Services and Applications

The Internet is perhaps the best showcase of how when technologies are implemented using mainly open standards, there is almost universal accessibility, acceptance and benefits. Most networking infrastructure of the Internet is implemented based on open standards drawn up by IETF. In addition, many services and applications running now as well as being planned for the future are being implemented following open standards and recommendations from several bodies notably, IETF, W3C and OASIS. As a result, today, it is possible for one to access major services offered on the Internet using a multitude of environments ranging from commodity PCs, hand-held Personal Digital Assistants (PDAs) and mobile devices to proprietary set-top black boxes and TV sets. Without this adherence to open standards, the Internet would not be as ubiquitous as it is today.

STANDARD-SETTING AND OPEN STANDARDS

This section will look into standard-setting processes and the more important standards bodies in IT, and how they relate to the setting and adoption of open standards.

Standard-Setting Organizations

In this document, the term Standard-Setting Organization (SSO) is taken to refer to an organization that attempts to set standards or make recommendations which, *when widely deployed*, become *de facto* standards. There are many SSOs, national, regional as well as industry-based. A formal SSO refers to one that is recognized directly or indirectly by a government entity.[16] Very often, there will exist a formal SSO in a country that the government recognizes as the national standards body and which has the authority to designate a specification as the national standard for the country. Thus, for example, in India, the Bureau of Indian Standards (BIS) is the national standards body; in the USA, the American National Standards Institute (ANSI) is the official body; while in the United Kingdom, it is the British Standards Institute (BSI).

While any organization can come up with its own specification and call it its standard, to be an internationally acceptable standard, it has to be either set or adopted/adapted by an SSO that is recognized as an international standard-setting body. The three organizations having the highest international recognition are the International Organization for Standardization(ISO), International Electro-technical Commission (IEC) and the International Telecommunication Union (ITU).

ISO[17] is an international standard-setting body made up mainly of representation from national standards bodies. IEC[18] is a standards organization that deals mainly in setting standards for electrical, electronic and related technologies. A body that is an accredited representative to ISO or IEC is called a Standard Development Organization (SDO); most national standards bodies are SDOs. ISO produces standards in many domains, including IT. Many of its standards are also developed jointly with IEC, in particular, the ISO/IEC Joint Technical Committee 1 (JTC 1) is active in setting standards for the IT domain.

The International Organization for Standardization (ISO)

ISO is a non-governmental organization for standards with its secretariat in Geneva, Switzerland. Membership of ISO is open only to national standards institutes or similar organizations most representative of standardization in their country (one member in each country). Currently, there are over 150 members representing nations from all over the world.

ISO sets standards for a wide variety of industries ranging from agriculture to rubber and plastics and to IT. Standards approved by ISO are agreed upon (by consensus) between national delegations representing all the economic stakeholders concerned – suppliers, users and governments. ISO standards are usually regarded as international standards.

ITU,[19] one of the world's oldest international standards bodies, was established to standardize and regulate international radio and telecommunications. With the convergence of IT and telecommunications, ITU (specifically its Telecommunication Standardization Sector, ITU-T) is now also involved in specifying standards (or Recommendations as it calls them) that impact the ICT world.

The International Telecommunication Union (ITU)

ITU has its headquarters in Geneva, Switzerland, and it is an international organization within the UN System where governments and the private sector coordinate global telecom networks and

services. It started out as the International Telegraph Union in 1865 to facilitate the interoperability of the then-fledgling telegraphy system among countries. From there it has grown and evolved to the ITU of today, which is involved in the standardization and regulation of international radio and telecommunications.

Membership of the ITU is open to governments as well as to private organizations involved in the telecommunications industry, e.g. carriers, equipment manufacturers, large telecommunication organizations, research bodies, etc.

ITU is divided into three sectors: Radio Communication (ITU-R), Telecommunication Standardization (ITU-T), and Telecommunication Development (ITU-D). ITU-T is increasingly becoming an important international body for the development of IT standards due to the convergence of IT and Telecommunications.

Standard-Setting Processes

The setting or creation of new technical standards can basically follow several main processes: *de jure*, *de facto*, and industry-created standards.

De jure Standards

De jure standards are normally created by formal SSOs following procedures that have been established by these bodies. Based on a need, work on the creation of a new standard is proposed by one or more members of the organization. This is called a new work item proposal. If there is enough support, work on drafting the new standard is started by a small committee or working group. The working draft may go through several cycles of deliberation, voting and modifications by the working group members (as far as possible, a consensus among the members is usually sought) before it is released as a draft to other members of the main organization or committee for scrutiny. At this level, it may be sent back to the working group for further changes and the cycle repeated until it is accepted as a draft standard for publication by the organization. Once it is published, it becomes a formal standard from the organization.

In SSOs, like ISO, the final acceptability of the draft is determined by a formal vote from the participating national bodies. After this final round of voting, the draft document is published.

The advantage of such a process as described above is that formal and accountable procedures are followed and each step in the process is accomplished through consensus as far as possible. The members of the SSO are given an opportunity to contribute during the drafting of the document. Some SSOs also allow contributions from invited subject-matter experts. The idea is that everyone interested in the standard should participate; and the standards creation process be seen as neutral and transparent, not controlled by any particular group or party.

There are several disadvantages to the process involved in the creation of *de jure* standards. First of all, the entire standard drafting process can be quite long because of the structure and makeup of the formal SSOs. For example, in the case of ISO standards, there is commonly a time span of two to three years from the new work item proposal to the publication of a standard.

While the standard-setting process formally tries to be neutral and impartial to any group, in practice this may not be so. In some cases, vendors and commercial organizations will send their experts to participate and push their own agendas, e.g. the inclusion of the specifications of their particular technology into the standard. Also some formal SSOs, like ISO, allow participation mainly by the national standards body only, so direct participation is restricted. However, interested parties should be able to participate at the local level via their national standards body, that will then carry the so-called national viewpoints, which may or may not concur with those of the interested parties.

The publication of a *de jure* standard by no means guarantees its success in implementation and acceptance by the industry and users. Sometimes, a simpler and more practical standard from the industry may win over a more complex and difficult to implement standard simply because implementation is simpler and faster, which results in better acceptance in the industry. A classic example of this is the highly complex but more complete X.400 suite of messaging protocols which is not widely used today as compared with the simpler but more easily implemented SMTP mail protocol that forms the backbone

of Internet email. The former was developed by the formal SSOs, ISO and ITU-T, while the latter came from the industry-driven IETF body.

Examples of internationally recognized SSOs that are active in putting out *de jure* standards are ISO, IEEE, ITU-T and ANSI. Examples of widely used *de jure* standards include:

- EEE 802 – a set of standards for Local Area Networking (LAN)
- ISO 10918 – a standard for the JPEG graphics compression and file format
- ITU-T X.25 – a standard for packet switching networks

Not all standards are created from scratch. Very often, an entity (e.g. an industry forum or group) may propose that a standards body, like ISO, adopt or adapt its standard or specification as an international standard. Sometimes a *de facto* standard may also be submitted to a standards body for adopting/adapting as an international standard.

De facto Standards

In the fast-moving IT industry, very often, some technology or product may become so popular that, as a result, it becomes generally accepted and widely used by a majority of users throughout the industry. As a result of this, a *de facto* standard is established that everybody seems to follow as though it was an authorized standard from a standards body. Examples of these are:

- the FAT file system from Microsoft
- the Adobe Portable Document Format (PDF)
- the Hayes command set for dial-up modem control
- the Hewlett-Packard Printer Control Language (PCL)

The main advantage of a *de facto* standard is that widespread acceptance in its implementation and usage is assured. It is unlike a *de jure* standard where the standard is just debated and agreed upon by the committee of the SSO and hence industry acceptance is by no means guaranteed.

Since a *de facto* standard does not have to wait for committee debate and approval, changes and modifications are made much faster. Indeed, very often it tends to change as and when the product is upgraded or improved.

The main disadvantage of a standard set in this way is that, very often, it starts off as part of a product implementation and as such will invariably include some technology and/or specification that is either owned or controlled by the vendor or group that produces the product. Unless that party is willing to give up control or at least share the control by allowing other stakeholders to be involved in developing and driving the *de facto* standard easily, there is a possibility of a lock-in later.

In some cases, after some time, a *de facto* standard may be submitted to a more independent standards body for adoption or adaptation whereby the proprietary control is relinquished and it may then become a real open standard. An example of this is the Network File System (NFS) that was originally introduced by Sun Microsystems as a means of allowing a user to access a file on a remote machine in a way similar to how a local file is used. Later, with the widespread usage of NFS even on other vendors' systems, it became part of the TCP/IP application standards from the IETF.

Industry-driven Standards

These are sort of intermediate between the *de jure* standards set by formal standards bodies and product-based *de facto* standards set mainly by vendors and owners of products. There is a trend nowadays in the IT industry for various consortia or groups to be formed among stakeholders in a particular segment of the industry. One of the functions of such a group may be to develop standards and/or recommendations deemed important and necessary for the progress of the sector. A good example of such a group is OASIS. OASIS is a not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards. It produces many Web services and Internet-related standards for

e-business deployment, such as Universal Description, Discovery and Integration (UDDI) and OpenDocument Format for Office Applications. The W3C is another consortium that has influence in the Web industry. It develops interoperable technologies (specifications, guidelines, software, and tools) for Web usage, e.g. HTML, XML, SOAP, etc. Although it is not a formal standard-setting body, it does come out with recommendations on Web technologies and services that are followed by many developers and/or vendors.

While the industry may adopt and support many of the standards or recommendations from these industry consortia as *de facto* standards, the established ones are eventually submitted to be adopted by traditional international standards organization like ISO to become a "legitimate" international standard. Many of these industry bodies have on-going liaisons with the technical committees of the international SSOs.

Open Standards Organizations

Bodies dealing with standards are usually non-profit and may be government-appointed, industry-backed, non-government organizations or even voluntary ones. While almost all of these claim to be "open", some are more open than others especially with respect to the free and easy accessibility and open participation criteria discussed in the Introduction. Some more active organizations that are generally perceived to be open include IETF, IEEE, OASIS, W3C and the Free Standards Group (FSG).

Note that this list is by no means an exhaustive listing of open standards bodies and indeed some may dispute the inclusion of one or more of these and/or the exclusion of other bodies if the accessibility and open participation criteria are applied strictly. However, in terms of important IT standardization activities and relative "openness" to world-wide participation and access by organizations big and small, the organizations listed earlier do stand out.

Standards and/or recommendations from these bodies account for many of the standards being deployed or developed in the IT and Internet/Web industries. Many of these standards have also been adopted as standards by international SSOs like ISO.

As noted earlier, these non-formal SSOs often have liaisons, especially at the technical working group level, with formal organizations such as ISO and ITU-T. Therefore, there is awareness and knowledge of the work and activities of the respective working groups from the various organizations working in the same area.

The Internet Engineering Task Force

Internet networking standards and protocols, like TCP/IP, became *de facto* standards when the Internet was widely embraced throughout the world. IETF is charged with developing and promoting Internet standards.[20] It is a voluntary organization with membership open to any interested individual. The actual technical work of IETF is done by its working groups which are formed, based on topics, into several key areas. Each area is overseen by an area director and the area directors, together with the IETF Chair, form the Internet Engineering Steering Group (IESG), which is responsible for the overall operation of IETF.[21] ETF is overseen by the Internet Architecture Board (IAB) which is, in turn, responsible to the Internet Society (ISOC).[22]

The drafting and setting of specifications and standards by IETF is carried out considerably faster when compared to the formal SSOs. IETF working groups do the drafting work. A new set of specifications starts off as an Internet Draft which is placed in IETF's "Internet-Drafts" directory and also replicated on a number of Internet hosts. Interested parties are encouraged to comment on this, usually through the working group's mailing lists. Based on comments and feedback, the draft undergoes several rounds of modification and then moves on to become a Requests for Comments (RFC) document and is published.

The specifications in a RFC document may be implemented by the Internet community and it can become a *de facto* standard if it receives wide acceptance. An RFC specification for which significant implementation and successful operational experience have been obtained may be elevated to the Internet standard level[23] and is assigned a number in the STD series while retaining its RFC number.[24]

The World Wide Web Consortium

W3C[25] is an international consortium that specializes in the development of protocols and guidelines for use on the World Wide Web. It is the leading body for specifications on Web technologies and applications. It calls its guidelines and specifications "Recommendations", which it considers as equivalent to Web standards. Many W3C Recommendations have been submitted to a formal standards body like ISO to become international standards.

W3C believes in complete interoperability for the Web to function and realize its full potential. Towards this end it publishes open standards for Web languages and protocols. This makes it possible for Web technologies to be compatible with one another and to allow any hardware and software used to access the Web to work together.

W3C is an independent body, membership is open to any organization and there are several categories of membership depending on the nature of the organization. W3C counts vendors of technology products and services, content providers, corporate users, research laboratories, standards bodies, and governments among its members. Individuals who are not employees of W3C member organizations can also be involved by participating in the technical discussions in its many public mailing lists.

The Organization for the Advancement of Structured Information Standards

OASIS[26] is a non-profit, international consortium that drives the development, convergence, and adoption of e-business standards. Standards produced by OASIS include those for security, Web services, conformance, business transactions, supply chain, public sector, and interoperability within and between marketplaces.

Membership of OASIS is open to both individuals and organizations all over the world. There are several types of membership and OASIS has a diverse membership base, counting users and vendors, governments and universities, trade groups and service providers among its members.

OASIS prides itself on its transparent governance and operating procedures. The members themselves set the OASIS technical agenda using a process designed to promote consensus and unite disparate efforts. Completed work is ratified by open ballot before it is published as an OASIS standard.

The Free Standards Group

FSG[27] is an independent, non-profit organization dedicated to accelerating the use of free and open-source software by developing and promoting standards. It is supported by both commercial corporations in the IT industry as well as the FOSS development community. All standards produced by FSG are available free and are distributed under open source licenses. Anyone can participate in and contribute to the FSG standards development by participating in the various FSG standards projects mailing lists.

The FSG is responsible for the important Linux Standard Base (LSB) standardization activity and the Open Internationalization (OpenI18N) initiative. Some LSB specifications have been submitted to the ISO/IEC JTC1 SC22 working group on GNU/Linux standardization.

The Institute of Electrical and Electronics Engineers

IEEE is a non-profit, technical, professional association of more than 360,000 individual members in over 175 countries. The IEEE Standards Association (IEEE-SA)[28] is active in the development of technical standards in the fields of information technology, telecommunications, and energy and power. IEEE standards development is guided by the five basic principles of due process, openness, consensus, balance and right of appeal; it is open to all and not restricted to a particular type or category of participants.

The working groups that are developing the standards are open to the public and have well-publicized procedures regarding membership, voting, officers, record-keeping and other areas. They try to be as transparent as they can, agendas for meetings are distributed beforehand and the results of a group's deliberations are publicly available, usually through meeting minutes.

When a draft standard is deemed mature enough, it goes up for balloting to become an IEEE standard. The sponsor of the standard forms a balloting group by inviting people from an "invitation pool". The

latter consists of IEEE-SA members or people who have paid a ballot fee and are interested in balloting some of the draft standards. Unlike the development stage where anyone can contribute comments, only members of the balloting group can vote in the ballot. The ballot sponsor has to take care that the balloting group is balanced with no domination by any one group or company.

Many IEEE standards have found international recognition and usage, e.g. the IEEE 802 series of LAN/MAN networking standards like 802.3 (Ethernet) and 802.11 (Wireless Fidelity (Wi-Fi)).

SOME IMPORTANT OPEN STANDARDS

This section will discuss some of the more important open standards that are either currently already available or actively being developed. The standards listed here are by no means exhaustive but they do represent those that are most widely used in the industry today.

Internet Networking and Applications/Services

The Internet is what it is today mainly because of the almost universal accessibility of the applications and services offered over it and its seamless connectivity. This is a direct result of the widespread use of open standards in the implementation of the Internet, both historically and currently. The standards mainly responsible for the Internet infrastructure and for the popular World Wide Web and Internet email services are highlighted here.

Transmission Control Protocol/Internet Protocol

The TCP/IP suite of networking standards provides the foundation for the network infrastructure of the Internet. All major services and applications on the Internet ride on top of TCP/IP. These protocols were originally developed by the pioneers of the Internet, the engineers and scientists from universities, research institutions and companies who collaborated on the US Department of Defence's Advanced Research Projects Agency Network (ARPANET) project. This evolved to become the Internet as we know it today, and TCP/IP became a *de facto* standard. It is now an IETF Standard and IETF is charged with its continued development.

TCP/IP is a two-layered packet-switching specification in which data to be communicated between two end-points on a network is first broken up into smaller data packets that are then individually routed through the network from the source to the destination points. The higher layer, Transmission Control Protocol (TCP),[29] manages the disassembling of the data into smaller packets at the source and the reassembling at the destination point upon receipt of the data packets. The lower layer, Internet Protocol (IP),[30] handles the addressing and routing of each packet so that it gets to the correct destination.

TCP/IP just provides the transport mechanism for sending data across the Internet or an IP network. In order for this to be useful, a service or application has to be specified and implemented. Again, IETF is mainly responsible in overseeing and setting the specifications for most of these services. The widespread implementation and acceptance of these specifications coupled with open standards bodies like IETF and W3C make the Internet the best showcase for open standards at work. Some of these standards are listed below.

Hypertext Transfer Protocol

HTTP is perhaps the most widely used Internet service protocol. It is the primary method used to access the WWW. Web content, in the form of HTML pages and possibly also other multimedia formats, is transferred from a Web server to a user's Web accessing agent using the HTTP protocol. HTTP was developed by W3C in co-operation with IETF working groups. The standard most widely deployed and supported on the Web today is HTTP version 1.1 or HTTP/1.1.[31]

The HTTP protocol is a request-response protocol using a client-server model in which an HTTP client, e.g. a Web browser, initiates a request by establishing a TCP connection to the server computer that will respond to the request commands sent by the client. The commands to support as well as for the behaviour of both client and server are spelt out in the HTTP specification.

It is through this universal acceptance of the HTTP protocol standard that the Web has become the ubiquitous information dissemination and exchange medium that it is now. One major factor in its wide

acceptance by all the stakeholders and players on the Internet is its open standard status.

Hypertext Markup Language

While HTTP defines how the contents of a Web page can be transmitted between a Web server and a client, HTML is an open standard specifying the structure and presentation of the content. A document composed with HTML consists of the contents intermingled with symbols and tags that tell the software needed to interpret and display the HTML document structure and presentation of the content. The HTML specification is now being maintained by W3C. It has undergone several revisions and the most current specification is HTML 4.01.[32] HTML is also available as an ISO standard,[33] which is a subset of HTML 4.

In its simplest form, an HTML document consists of the text of the document as well as tags that specify the markup needed to be performed on it. For example, in the sample below:

```
<h3>My Work Experience</h3>
 <img src="mypic.png">
 <p>
 <b>Work Experience</b><br><br>
  1990 - 1995 System engineer<br>
  1995 - 2005 Network manager<br>
 </p>
```

The tags <h3> and </h3> specify that the text enclosed within them is to be rendered as third level headings, while the tag <img src="mypic.png"> specifies the display of a graphics file. The tags , specify that the text enclosed within them should be rendered as bold, and the tag
 signifies a line break.

An HTML user agent software is needed to render a document made up of HTML and the most common agent is a Web browser. If the W3C HTML specifications are adhered to, an HTML document can be displayed properly by any user agent (which conforms to the specifications) and this can form the basis for a standard format for textual document information exchange. One major limitation of using HTML to display a document though, is that page breaks are not easily represented or controlled.

The use of HTML in email has gained popularity as it enables one to impose some simple formatting on the text as well as embed graphics and multimedia content into the message. However, it is generally considered not good practice by security-conscious users to utilize HTML in mail messages as some popular HTML-enabled email software have been known to possess vulnerabilities. This makes them open to potential exploitation by a rogue HTML email message which may result in the compromise of a user's system.

Email Protocols

Internet email has become almost as important as the telephone service. Every time we send an email, we assume that the mail will be relayed correctly by the mail server to its destination. When we send attachments or incorporate some non-textual content into our email we just assume that the attachment will be incorporated correctly and when the recipient gets it s/he will be able to get it back into its original form. All this works seamlessly irrespective of the hardware and software deployed because Internet email makes use of several important open standards in its mail transmission as well as in the encoding of email messages.

Simple Mail Transfer Protocol

SMTP[34] enables the transport and routing of email from the sender to the recipient using their email addresses. This standard is client-server based whereby the SMTP client (usually the user's email software or mail user agent) will initiate a TCP connection to the SMTP server (the mail relay host). Communications between the server and client is done using the SMTP protocol. This is a simple text-based protocol where, essentially, the client informs the server of the email addresses of the sender and the recipient(s).

After that, if all goes well and the server allows it (based on its mail relay policy), the client will transmit the mail message to the server. The server will then attempt to deliver it to the computer housing the recipient's mailbox or, if necessary, forward the email to another server for delivery to the recipient's mailbox.

The SMTP protocol started out supporting only 7-bit ASCII (American Standard Code for Information Interchange) text in the messages, effectively limiting it to the transmission of English-based text. Non-English language texts that make use of more than the 7-bit ASCII character set as well as binary file attachments have to be encoded by the email user agent software before transmission. The message format of this text-based mail is specified by another IETF standard, RFC 2822.[35] The SMTP standard has been extended to support 8-bit text,[36] permitting the transmission without encoding of text messages in more languages.

Multipurpose Internet Mail Extensions

As Internet email became more and more popular, users found it a convenient, economical and efficient way to send information to one another. Users tried to send other types of content, e.g. audio, video, images, software programs, besides text messages via email. However, since the original Internet email specifications were meant primarily for English-based text messages, some new sets of specifications had to be drawn up to allow interoperability and seamless transmission of multipurpose content. This resulted in IETF producing the Multipurpose Internet Mail Extensions (MIME) standard.[37]

MIME is an extension of the basic text-based Internet mail standard. It defines mechanisms for sending other kinds of information in email. These include non-English text using character sets beyond ASCII and binary file content such as multimedia files and computer software. To support these as well as to retain backward compatibility with the simple ASCII-based mail format, a set of email headers for specifying additional attributes of a message, e.g. content type, and a set of transfer encodings that can be used to represent 8-bit data using characters from the 7-bit ASCII set are defined. The encoding of non-ASCII characters in mail message headers is also catered for in MIME allowing the usage of non-English characters in them. The MIME standard specifies a means to register new content types and transfer encodings making it flexible for supporting new multimedia types in the future.

MIME is also an important standard for the Web as the HTTP protocol makes use of mail-like MIME formatting rules and syntax for its data formatting.

The Extensible Markup Language

The Extensible Markup Language (XML) is a Recommendation[38] from W3C that specifies a meta markup language (a meta language is a language for describing other languages) for the creation of other markup languages for use on the WWW. HTML is a single predefined markup language and hence possesses severe limitations to describe and represent all sorts of data for dissemination, exchange and interaction. XML, being a markup specification language, is capable of being used to design markups for describing many different kinds of data for storage, transmission, or processing by a program.[39] It describes the data but it does not tell you what you should do with the data.

One should note that XML and HTML were designed with different goals in that XML was designed to store, carry, and exchange data whereas HTML was designed to display data and to focus on how data looks.[40] XML was created for deployment on the Web by using a subset of an existing, widely used international standard for text document markup – the Standard Generalized Markup Language (SGML).[41]

Due to its design goals, XML is well suited for data transfer and exchange and as a format for document storage and processing. This and the fact that it is under the charge of an open specifications/standards body, W3C, has resulted in XML being used as the base for specifying many other data formats and exchange protocols. According to the community-based XML portal, XML.ORG,[42] it is now viewed as the standard way for information exchange in environments that do not share common platforms.

Special purpose languages and standards developed using XML for specific environments or activities are announced almost daily and several hundred have been adopted since XML 1.0 was released in February 1998. In particular, the e-government and e-commerce segments are very active in developing and implementing XML-based specifications.

A simple XML document is shown below:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<?xml-stylesheet type="text/xsl" href="bookcollection.xsl"?>
<BookCollection>
      <Book>
            <Title>Chronicles: Volume One</Title>
            <Author>Bob Dylan</Author>
            <Publisher>Simon and Schuster</Publisher>
            <Year>2004</Year>
      </Book>
      <Book>
            <Title>Harry Potter and the Goblet of Fire</Title>
            <Author>J.K. Rowling</Author>
            <Publisher>Bloomsbury Publishing</Publisher>
            <Year>2000</Year>
      </Book>
</BookCollection>
```

Note that while XML uses syntax tags to identify various types of data in a document file, these tags are not predefined. So the document creator has to define and describe them using what is called an XML schema and associate the document with the schema. To create the schema, an XML schema language is used, e.g. Document Type Definition (DTD), XML Schema and RELAX NG. The purpose of the schema is to define the legal building blocks of the XML document, i.e. the elements, data attributes, tags, etc., that can appear in the document. DTD has limitations with respect to its extensibility and lack of support of several useful features, e.g. data types and namespaces. XML Schema, which is also another W3C Recommendation, is more suitable for use in many practical Web applications.

While the schema may define the legal components of the XML document, it does not carry information about how to display the data. So in order for the data in an XML document to be displayed properly by say, a Web browser, a display style has to be specified. The Extensible Stylesheet Language (XSL) is used to perform this. Styling is about transforming and formatting information and the W3C specifications separate these processes. In addition, the components in an XML document have to be navigated to extract and process them. Hence, the XSL Recommendation from W3C consists of three parts:

- XSL Transformations (XSLT): a language for transforming XML documents
- XSL Formatting Objects (XSL-FO): a language for formatting XML documents
- XML Path Language (XPath): a language for navigating in XML documents

An example of an XSLT transformation of the XML example document above to a Web browser displayable HTML output is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:template match="/">

      <html>
      <body>
      <h3>Book Collection</h3>
      <table>
      <tr bgcolor="#ff0000">
            <th align="center">Title</th>
            <th align="center">Author</th>
            <th align="center">Publisher</th>
            <th align="center">Year</th>
      </tr>
```

```
        <xsl:for-each select="BookCollection/Book">
        <tr>
               <td><xsl:value-of select="Title"/></td>
               <td><xsl:value-of select="Author"/></td>
               <td><xsl:value-of select="Publisher"/></td>
               <td><xsl:value-of select="Year"/></td>
        </tr>
        </xsl:for-each>
        </table>
        </body>
        </html>

</xsl:template>
</xsl:stylesheet>
```

Computer Graphics and Multimedia

In the old days of computing, the display was predominantly text-based and any graphics displayed was, at best, line-graphics implemented using special line drawing character sets. Computer terminals that can display full-fledged graphics were expensive and used only for special purposes or applications. Today, with the proliferation of inexpensive personal computers that have the power to process and display graphics and multimedia, even the user interface is graphics-based. One of the main attractions of the Web is its widespread support and usage of graphic images and multimedia to make the content interesting and lively.

It is important that open standards are followed as much as possible in graphics and multimedia data storage, processing and retrieval to enable diverse devices and computing platforms to offer the same degree of Web experience.

Portable Network Graphics

In the early days of the Web when Internet links and connections were relatively slow, many simple images and animations displayed in Web pages made use of a graphics format called Graphics Interchange Format (GIF)[43] as this format resulted in small graphic file sizes. The GIF format included the use of the Lempel-Ziv-Welch (LZW) compression algorithm that was patented in the USA by Unisys who eventually decided to ask for royalty payments for all software that utilize GIF. This led to the creation of the Portable Network Graphics (PNG) format[44] to replace GIF for use as a single-image Web format. The PNG format later became a W3C Recommendation as well as an ISO international standard (ISO/IEC 15948).

PNG is an extensible file format for the lossless, portable, well-compressed storage of raster images. Indexed-colour, greyscale, and true colour images are supported, plus an optional alpha channel for transparency. It is fully streamable with a progressive display option making it useful for online graphics display in Web pages. It also boasts robust features, providing both full file integrity checking and simple detection of common transmission errors.[44]

The X Window System

The graphics user interface (GUI) that is now common on desktop computers uses a graphical window metaphor as the basic user interface. This window system GUI enables different programs to run simultaneously in their own individual windows and these windows can be opened, closed and resized. The windowing systems found on platforms like Microsoft Windows and Mac OS X are proprietary ones. On the other hand, UNIX and UNIX-like operating systems (e.g. GNU/Linux, FreeBSD) make use of an open window system – the X Window System.

The X Window System, or X, is an open windowing system standard led by the X.Org Foundation.[45] X provides a framework for the display and management of graphical information and on top of this a GUI may be built. X uses a client-server model. The X client is usually the application that sends graphical output for display on the X server. The X server interacts with the user using primarily the keyboard and mouse as input devices and the input is transmitted to the client to act upon.

The X client and server may be running on the same machine or they may be on different physical devices connected together over a network. The intrinsic client-server property of X constitutes the main difference between it and other well known window systems like Microsoft Windows, which simply displays graphical applications local to the device on which the application is running on.

Being an open standard, besides UNIX and UNIX-like systems, X has been implemented on a variety of hardware and operating systems, including the various generations of Macintoshes, PCs running MS-DOS and Microsoft Windows as well as OpenVMS from Hewlett-Packard (formerly Digital Equipment), etc.

Ogg Vorbis

Ogg Vorbis is a general-purpose compressed audio format for storing and playing digital music. It is comparable in quality to other formats such as the popular MP3. However, unlike MP3, it is an open format and it claims to be free from patents. The format originated from the Xiph.Org Foundation,[46] a non-profit organization dedicated to producing free and open protocols, formats and software for multimedia.

Vorbis is the name of the audio compression scheme and this is contained in Ogg, the name of Xiph.org's container format for audio, video, and meta-data – hence the name Ogg Vorbis. Vorbis is a lossy codec, i.e., it uses a compression algorithm that discards data in order to increase the compression possible. The Ogg Vorbis specification is in the public domain and is completely free for commercial or non-commercial use.[47]

Software and hardware devices that support Ogg Vorbis are steadily increasing in number and may be found on the Vorbis wiki at Xiph.Org.[48]

Office Documents

Office applications are some of the most widely used applications for personal computers in a modern-day office. These applications include word-processing, spreadsheets and presentation software. Available on the market are several office applications, e.g. Microsoft Office, WordPerfect Office, OpenOffice.org and Applixware Office. Each of these invariably used different formats for storing their files in the past.

As a result, it was difficult to convert from one file format to another and for one application to read/write a file created by another application. It is thus a real step forward in terms of office interoperability and productivity when OASIS announced that it was recommending the Open Document Format for Office Applications as a standard file format use in office applications

OpenDocument

OpenDocument[13] is a file format developed by OASIS for storing office documents created by applications such as spreadsheets, word processors, charts and presentations software. It makes use of a royalty-free, open and vendor-independent XML-based format. The format is based on the file format of OpenOffice.org, which was submitted to OASIS to form the basis for the standard. OpenDocument provides a single XML schema for text processing, spreadsheet, presentation, drawing, charting, and mathematical documents.

Office software which have announced that they will support the OpenDocument format as their primary/native format include the office suites of OpenOffice.org, StarOffice and KOffice.

Open Standards Usage

Table 1 summarizes the usage and penetration levels of the open standards described above in their respective domains. As can be seen, open standards are widely deployed on the Internet and in running Internet-related/derived services and applications. However, for the graphics, multimedia and office applications areas they are still very limited in acceptance. The limited penetration in these domains is a result of the fact that they are dominated by proprietary products like those from Apple and Microsoft that make use of their own proprietary formats and specifications (see the next section on "Comparison

of File Formats"). The incentive for these vendors to support open standards or at least make their specifications more open is not strong due to their dominant positions.

Table 1: Important Open Standards

Domain	Standard	Organization	Usage
Networking	TCP/IP	IETF	Universal
WWW	HTTP	W3C, IETF	Universal
Web content	HTML	W3C	Universal
Email	SMTP	IETF	Universal
Email, WWW	MIME	IETF	Universal
Doc exchange	XML	W3C	Universal
Graphics	PNG	W3C	Wide
Window system	X Window	X.Org	Limited
Audio	Ogg Vorbis	Xiph.Org	Limited
Office documents	OpenDocument	OASIS	Limited

COMPARISON OF FILE FORMATS

This section will list, compare and discuss the degrees of openness and/or lack of openness of several popular file formats. These include file formats for the following application areas:

- office applications
- graphics
- audio
- video

Office Applications File Formats

Microsoft Office Formats

Currently, the most popular office application is Microsoft Office (MS Office). This suite of office software comprises mainly (depending on the type of suite purchased) word processing (MS Word), spreadsheet (MS Excel) and presentation software (MS PowerPoint). Up till version 10 (MS Office 10), the file formats used were binary (i.e. non-plain text) in nature and not publicly published. MS Word, MS Excel and MS PowerPoint use the binary DOC, XLS and PPT formats, respectively, and these are proprietary formats, being owned and controlled entirely by Microsoft.

The file formats for these applications are widely used due to the popularity of MS Office. Other software not from Microsoft, e.g. OpenOffice.org or StarOffice, are able to read and write files using these proprietary formats but the compatibility is incomplete. Competing products cannot be totally compatible with MS Office unless they are provided with the file format specifications by Microsoft.

Some MS Office applications like Word and Excel can save their data in what is known as the Rich Text Format (RTF) file format. This is a non-binary file format that has been developed by Microsoft for cross-platform document interchange. Technical documentation on RTF is published by Microsoft and as many non-Microsoft software support the RTF file format well, it is widely used for document exchange between MS Office and other office applications. However, the RTF format does not completely support the more complicated and sophisticated features found in MS Office, and complex documents may not be properly represented using the RTF format.

With MS Office 11 (MS Office 2003), the option to use a new XML-based[49] file format for Word and Excel was made available. However, these XML-based formats have been criticized in some quarters for being incomplete and immature. They were not available for all the software applications in the suite and some major functionalities were not supported in those available. As a result, the traditional binary MS Office file formats remained in use mainly. In June 2005, Microsoft announced that MS Office 12, due in 2006, will deliver support for a new set of XML file formats called the "Microsoft Office Open XML Formats".[50] The applications that will use these formats by default are Word, Excel and PowerPoint.

Office XML Open Format is also being published by Microsoft on a royalty-free basis to the industry. While, potentially, this will make it possible and easier for third-party products to be compatible with MS Office, the file format will still be owned and controlled by Microsoft and, hence, is not open.

In an attempt to allay fears over this and to allow customers, notably corporations and national governments with long-term archival needs, to access the contents of their documents created with MS Office without being dependent on Microsoft, the Office XML formats have been submitted to ECMA International for standardization.

OpenOffice.org and StarOffice Formats

OpenOffice.org (OOo)[51] is a full-fledged Open Source office application suite, comprising word processor, spreadsheet, presentation software, graphics editor and a database program (available in OOo version 2 only). The original file formats used by OOo were XML-based. As there were several files associated with a single document, all the files were compressed and stored as a single zip-compressed file. OpenOffice.org is available on multi-platforms, e.g. GNU/Linux, MS-Windows, Mac OS X, etc., and offers multi-lingual support. It is compatible with all other major office suites. In particular, it is able to read and write MS Office file formats. The degree of compatibility is very good though not complete.

The OpenOffice.org file format was submitted to OASIS to form the basis for a new standard for office applications and this resulted in OASIS coming up with the Open Document Format for Office Applications (OpenDocument) v1.0 in May 2005. New versions of OOo as well as other office suites like KOffice and StarOffice will support OpenDocument as their native file formats. This will significantly improve the interoperability of office software and enhance document exchange. What is most important though is that all these office applications now use a standard open file format for storing their data. The OpenDocument format is not owned or controlled by a single vendor, instead it falls under the ambit of OASIS, an open standards body. Users can, thus, be assured that they will have access to their documents and data from a variety of software.

StarOffice shares the same code base as OOo but it is released under a proprietary commercial license. In addition to the core functionalities of OOo, it also comes with some proprietary and third-party modules, e.g. the Adabas B database and some proprietary clip art galleries and templates. StarOffice uses and supports the same file formats as OpenOffice.org.

Adobe's Portable Document Format

PDF is a file format developed by Adobe Systems, Incorporated[52] for secure and reliable electronic document distribution and exchange. The format is able to preserve the look and integrity of the original document, regardless of the application and platform used to create it even if it contains complex combinations of text, graphics and images. As such, the PDF format is very useful as a format for multi-platform document exchange and distribution and for sharing information. However, one major drawback of PDF is that it is an end-form format, i.e., it is not suitable for modifying or re-writing its contents.

Table 2: Office Document Formats

Format	Organization	Published	Non-Proprietary	International Standard
DOC (text) XLS (spreadsheet) PPT (presentation)	Microsoft Microsoft Microsoft	No No No	No No No	No No No
SXW (text) SXC (spreadsheet) SXI (presentation)	OpenOffice.org OpenOffice.org OpenOffice.org	Yes Yes Yes	Yes Yes Yes	No No No
ODT (text) ODS (spreadsheet) ODP (presentation)	OASIS OASIS OASIS	Yes Yes Yes	Yes Yes Yes	Yes Yes Yes
PDF (text and presentation)	Adobe	Yes	No	No

The specifications for the PDF format are publicly published by Adobe[53] and it can be implemented without restrictions by anyone. As a result, a variety of software on many different platforms is available that can read the PDF format, and a (smaller) number of applications that can write out the contents of a document in PDF.[54] Due to its popularity and wide support, PDF can be considered a *de facto* standard as a file format for information exchange and sharing but since it is created, owned and controlled by Adobe Corporation, it does not meet the technical definition of an open standard.

Graphics/Image File Formats

A picture is worth a thousand words, as the saying goes. It is not surprising then that, with the advent of powerful desktop systems that are able to display high resolution graphics, images are being utilized more and more to convey information. Modern computer systems use what is known as raster graphics to display an image on the video screen. A raster graphics image, digital image, or bitmap, is a data file or structure representing a generally rectangular grid of pixels, or points of colour, on a computer display monitor.[55] Each point or pixel on the screen is represented by a value denoting its colour and this bitmap is stored in memory. Using this bitmap, the entire screen is repainted 30 or more times per second by the video device resulting in the human eye seeing the image being displayed on the screen. There are many ways to create and store this raster graphics image file and so if we are to be able to exchange and share useful graphical information there is a need to have a format that is supported on multiple platforms and by various graphics software.

Many graphics file formats in use today are proprietary by nature, being derived and tied to the software used to create them. There are some formats that have gained wide acceptance as *de facto* standards and a few of these have emerged as open graphic file formats.

GIF

GIF is a bitmap image format[43] that is widely used on the World Wide Web, especially in its early days as this format resulted in small graphic file sizes. Images stored as GIF files are generally limited to 256 colours. The GIF format makes use of the LZW compression algorithm that was patented in the USA by Unisys. After the GIF format found widespread use on the Web, Unisys asked for royalty payments for all software that utilizes GIF (this patent has since expired in the USA, in 2003). This led to the diminished use of GIF and also to the creation of alternatives to it, notably the PNG format.[44]

PNG

The PNG format was created as an alternative to GIF when Unisys decided to enforce its software patent on LZW data compression that was used in the then popular GIF format. PNG is an extensible file format for the lossless, portable, well-compressed storage of raster images. It offers indexed-colour, grayscale, and true colour image support, plus an optional alpha channel for transparency. It is fully streamable with a progressive display option making it useful for online graphics display in Web pages. It also boasts robust features, providing both full file integrity checking and simple detection of common transmission errors.[44]

PNG is supported by all major graphics software and is now very widely used. It has become an open file format standard and it is a W3C recommendation as well as an ISO international standard (ISO/IEC 15948).

XPM

The XPM (XPixMap) format[56] is a *de facto* standard for creating icon pixmaps for use in GUIs based on the X Window System. It consists of an ASCII image format and a C library. The XPM format defines how to store colour images (X Pixmap) in a portable way while the associated library provides a set of functions to store and retrieve images to and from XPM format data.

TIFF

The Tagged Image File Format (TIFF) is a file format for digital images. It is a specification that is now owned by Adobe Systems, Incorporated. TIFF is widely used in image applications in the publishing

Table 3: Graphics Formats				
Format	**Organization**	**Published**	**Non-Proprietary**	**International Standard**
GIF	CompuServe	Yes	No	No
PNG	W3C	Yes	Yes	Yes
XPM	X.Org	Yes	Yes	No
TIFF	Adobe	Yes	No	TIFF/IT
JPEG	ISO	Yes	Yes	Yes
SVG	W3C	Yes	Yes	No

industry and also supported by most image scanning and editing software. The specifications for the TIFF format[57] is publicly published by Adobe and it can be implemented without restrictions by anyone. As a result, there is available software on many different platforms that can read and write the TIFF format. It has become a *de facto* standard graphics format for high colour depth (32-bit) graphics.

TIFF/IT, which is based on TIFF, is a specification for the exchange of digital advertisements and complete pages (e.g., newspapers, magazines). This has been made an ISO standard (ISO 12639) as a media-independent means for pre-press electronic data exchange.

JPEG JFIF

JPEG is a standardized image compression mechanism from the Joint Photographic Experts Group (JPEG).[58] The file format that employs this compression is JFIF (JPEG File Interchange Format) and JPEG JFIF is what people generally mean when they refer to "JPEG". The JFIF file format was created by the Independent JPEG Group (IJG) for the transport of single JPEG-compressed images.[59]

The JPEG compression uses a lossy mechanism for compressing colour or greyscale images. It works well on natural, real-world scenes like photographs, naturalistic artwork and similar material but it does not fare too well on lettering, simple cartoons or line drawings.[60] The basic JPEG format is the most common format used for storing and displaying photographic images on the Web. One reason for this popularity is that the amount of compression can be adjusted to achieve the desired trade-off between file size and visual quality. The JPEG compression is now an ISO standard – ISO/IEC 10918 Parts 1-4. There are potential patent issues with JPEG, especially with some of its optional features, namely arithmetic coding and hierarchical storage and so for this reason, these optional features are seldom used on the Web.[59]

SVG

Unlike other file formats listed above that are meant for raster graphics, the SVG (Scalable Vector Graphics) format is meant for vector graphics, i.e. the use of geometrical primitives such as points, lines, curves, and polygons to represent images in computer graphics.[61] SVG consists of an XML-based file format and a programming API for graphical applications. It is a W3C recommendation[62] and is starting to become a popular choice for including graphics in XML documents. As an SVG document can include raster images such as JPEG and PNG, it can be used to add raster and mixed vector/raster graphics to XML documents.

The SVG format is important as it offers a way based on open standards to render graphics optimally on all types of devices. While currently the usage of SVG usage on the Web is somewhat limited, this should change in due course as more Web browsers support it natively. For the mobile phone industry, it has become the basis for its graphics platform with the publication of the SVG Mobile profile targeted at resource-limited devices such as mobile handsets and PDAs.

Audio File Formats

There are two major groups of audio file formats:

- those using lossless compression, e.g. like WAV, FLAC

- those using lossy compression, e.g. MP3, Ogg Vorbis, WMA, AAC

In the lossless compression of a piece of data, nothing is lost during the compression and the original data is restored upon uncompressing. In lossy compression, some data is lost during compression and upon uncompressing the data is not identical to the original but possibly close to it. Lossy compression is used mainly in the compression of multimedia data like audio or video where the loss of some details is tolerable under certain conditions, e.g., the human eye is unable to discern the loss in certain details of an image or video.

WAV

WAVEform audio format (WAV) is a Microsoft and IBM audio file format for storing audio on PCs. It is the main format used on Microsoft Windows systems for raw audio storage. The WAV format is most commonly used with an uncompressed, lossless storage method (pulse-code modulation) resulting in comparatively large audio files. Today, the WAV audio format is no longer popular being superseded by other more efficient means of audio storage.[63]

FLAC

Free Lossless Audio Codec (FLAC) is a popular lossless audio format with compression designed specifically for audio data streams, achieving compression rates of 30-50 percent. The format specification is publicly available and forms part of the FLAC Open Source project.[64] It is supported by a growing list of audio software and devices.

MP3

MPEG-1 audio layer 3 (MP3) is a popular lossy compression audio format. The MP3 specification was set by the Motion Pictures Experts Group (MPEG), a working group of ISO/IEC charged with the development of video and audio encoding standards. The compression scheme and format for MP3 forms part of the MPEG-1 video and audio compression standard specifications and is an ISO standard, ISO/IEC 11172-3.

MP3 is one of the most popular audio file formats in use today. Music files encoded with MP3 are particularly popular on music exchange and download sites on the Internet due, in part, to the relatively small size of such files and the wide availability of free software on PCs that allow easy creation, sharing, collecting and playing of MP3 files.

MP3 makes use of patented technology and so software and devices that support it are subject to royalty payments in those countries that recognize software patents. This has led to the creation of alternatives to MP3, e.g. Ogg Vorbis and WMA.

WMA

Windows Media Audio (WMA) is a lossy compression audio file format developed by Microsoft. It is a proprietary format but is widely used and supported due to the popularity of the MS Windows platform.

Table 4: Audio Formats

Format	Organization	Published	Non-Proprietary	International Standard
WAV	Microsoft	Yes	No	No
FLAC	Xiph.Org	Yes	Yes	No
MP3	MPEG/ISO	Yes	Yes	Yes
WMA	Microsoft	No	No	No
AAC	MPEG/ISO	Yes	Yes	Yes
RealAudio	RealNetworks	Yes	No	No
Ogg Vorbis	Xiph.org	Yes	Yes	No

AAC

Advanced Audio Coding (AAC) from MPEG is a lossy data compression scheme intended for audio streams. It was designed to provide better quality at the same bit-rate than MP3, or the same quality at lower bit-rates (and hence smaller file sizes). The compression scheme and format for AAC forms part of the MPEG-2 video and audio compression standard specifications and is an ISO standard, ISO/IEC 13818-7. This MPEG-2 AAC specification makes use of patents from several companies and a patent license is needed for products that make use of this standard.

The newer MPEG-4 standard also specifies an audio compression technology that incorporates MPEG-2 AAC. This is known as MPEG-4 AAC, and is an ISO standard, ISO/IEC 14496-3.

Apple's popular iTunes service and iPod products have music available in AAC and this has led to an upsurge in the popularity of AAC despite the required patent license royalty payments.

RealAudio

RealAudio is a proprietary audio format developed by RealNetworks for low bandwidth usage. It was first introduced in 1995 and it became popular especially for streaming audio, i.e., the audio is being played in real time as it is downloaded. Many radio stations use RealAudio to stream their programmes over the Internet.

Ogg Vorbis

Ogg Vorbis is a compressed audio format that is believed to be free of patents and royalty payments. The format originated from the Xiph.Org Foundation,[46] a non-profit organization dedicated to producing free and open protocols, formats and software for multimedia.

Ogg Vorbis uses the Vorbis lossy audio compression scheme. The audio data is wrapped up in the Ogg container format, the name of Xiph.org's container format for audio, video, and meta-data – hence the name Ogg Vorbis. The Ogg Vorbis specification is in the public domain and is completely free for commercial or non-commercial use.[47] There is growing support for the Ogg Vorbis format from software and hardware devices[48] as well as online audio services.

Video Formats

In order that a multimedia experience can be enjoyed properly by all without any discrimination, it is important that there exist multi-platform and multi-software support for it. This underlies the important role that open standards play in relation to video formats and technologies.

Video data's storage involves more than just finding an efficient means to store raw data; other data like tags, menus and possible media manipulation information need to be stored too. There may also be a need to store audio data as video frequently has sound associated with it. Also, the data stream is usually not stored in its raw form, it is transformed into a form more suitable for storage or transmission. A type of file called a container is used to store the data and associated information and a codec is utilized for encoding and decoding the data stream. It is important that the format of the container file as well as the codec that is supported by it follow open standards.

Almost all video containers popular today are proprietary. This is due to the popularity of Apple's QuickTime and Microsoft's Windows Media framework multimedia technologies. Some of these formats, through widespread usage, have emerged as *de facto* standards but remain proprietary formats all the same.

Video Containers

AVI

Audio Video Interleave (AVI) is a video container format from Microsoft containing both audio and video data. It is a Resource Interchange File Format (RIFF) file specification used with applications that capture,

Table 5: Video Containers				
Container	**Organization**	**Published**	**Non-Proprietary**	**International Standard**
AVI	Microsoft	Yes	No	No
ASF	Microsoft	No	No	No
MOV	Apple Computer	Yes	No	No
MP4	MPEG/ISO	Yes	Yes	Yes
Ogg	Xiph.Org	Yes	Yes	No

edit, and play back audio-video sequences.[65] It enjoys widespread support and it is the most common container format for audio/video data on the PC.

ASF

Advanced Systems Format (ASF) is Microsoft's proprietary container designed for streaming. The codec is not specified in ASF but the most common ones are Windows Media Audio (WMA) and Windows Media Video (WMV). The ASF container structure is patented in the United States.

MOV

The MOV container is from Apple Computer's QuickTime multimedia architecture and technology. This video file format is openly documented and available for anyone to use royalty-free. As a result, there are several non-Apple video player software available which can play QuickTime video files. The proprietary Sorenson codec is usually used with QuickTime. The QuickTime format was used as the basis of the MPEG-4 MP4 container standard (see entry on MP4 below).

MP4

MPEG-4 Part 14 (MP4) is a container specified as part of the MPEG-4 international standard, ISO/IEC 14496-14. MP4 is designed to support streaming, editing, local playback, and interchange of content. Its design is based on the QuickTime format.[66]

Ogg

The Ogg container uses a bitstream format to encapsulate data from one or more sources. It can handle both audio and video data and while the codecs are not specified,[67] there are several open codecs associated with the Ogg project: Vorbis and FLAC for audio and Theora for video.[68]

The Ogg format has been published as an IETF document, RFC 3533.

Video Codecs

MPEG Codecs

MPEG has developed several standards pertaining to video technology that are used by many digital video products on the market. The MPEG video codecs are specified in the following ISO standards:

- MPEG-1 Part 2 (ISO/IEC 11172-2)
- MPEG-2 Part 2 (ISO/IEC 13818-2)
- MPEG-4 Part 2 (ISO/IEC 14496-2)
- MPEG-4 Part 10 (ISO/IEC 14496-10)

The MPEG-2 and MPEG-4 standards make use of numerous patented technologies and the vendors of commercial products and services that use them are expected to pay patent licensing royalties.

MPEG-1 Part 2

The MPEG-1 standard that specifies the MP3 audio codec also specifies a video codec for non-interlaced video signals. This codec can be used for compressing video sequences, both 625-line and 525-lines, to bit rates of about 1.5 Mbit/s. It is used in the Video CD (VCD) specifications and the picture quality is comparable to that found for the VHS video cassette recorder.

MPEG-2 Part 2

The MPEG-2 standard specifies a video codec for interlaced and non-interlaced video signals. MPEG-2 video is not optimized for low bit-rates (less than 1 Mbit/s), but outperforms MPEG-1 at 3 Mbit/s and above. The MPEG-2 video codec is backward compatible with the MPEG-1 codec. MPEG-2 is widely adopted for video broadcasting (e.g., direct broadcast satellite and cable TV), filmmaking, and DVD discs. MPEG-2 has a lot of market acceptance and a very large installed base.

MPEG-4 Part 10 (H.264/AVC)

This video coding standard is the same as the ITU-T H.264 recommendation and the technology is also known as Advanced Video Coding (AVC). It contains several innovative features that allow it to compress video more efficiently than earlier MPEG codecs. It also possesses more flexibility, which allows it to accommodate applications in a wide variety of environments.

This is a new standard and it represents the current state-of-the-art in the series of MPEG video compression standards. It is rapidly gaining adoption in a wide variety of applications and digital broadcasting and TV systems. Apple Computer has integrated H.264 into Mac OS X version 10.4 (Tiger), as well as QuickTime version 7 while x264 is a FOSS free library for encoding H.264/AVC video streams. H.264 decoders for Windows, GNU/Linux and Macintosh as well as video servers and authoring tools are available from a number of vendors.[69]

Table 6: Video Codecs

Container	Organization	Published	Non-Proprietary	International Standard
MPEG-1	MPEG/ISO	Yes	Yes	Yes
MPEG-2	MPEG/ISO	Yes	Yes	Yes
MPEG-4	MPEG/ISO/ITU	Yes	Yes	Yes
Sorenson	Sorenson Vision	No	No	No
WMV	Microsoft	No	No	No
Theora	Xiph.org	Yes	Yes	No

Table 7: Video Formats

Container - Codec Commonly Used	Usage	Open / Close
AVI - WMV	Wide	Close
ASF - WMV	Wide	Close
MOV - Sorenson	Wide	Close
MP4 - MPEG-1,2,4	Wide	Open
Ogg - Theora	Limited	Open

To establish whether a video format is open or close, we have to look at the status of the specifications for both the container and the codec. The table above lists the popular containers as well as the codec's that are most commonly associated with them.

Sorensen

The Sorenson codec is a proprietary video codec from Sorenson Media and used by Apple's QuickTime.

Windows Media Video

This is a set of proprietary streaming video technologies developed by Microsoft as part of its Windows Media framework. It is the codec usually used in an AVI or ASF container and has support for digital rights management facilities. Microsoft has submitted WMV Version 9 to the Society of Motion Picture and Television Engineers (SMPTE) for approval as a standard under the name "VC-1".[70]

Theora

This is a video codec from Xiph.org Foundation as part of the Ogg project. It is based on patented technology but it has been irrevocably given a royalty-free license to use the patents in the codec. The Theora codec is released under a Berkley Software Distribution (BSD) FOSS license and it is available freely for commercial or non-commercial use.

STANDARDS AND INTERNATIONALIZATION/ LOCALIZATION OF SOFTWARE

Internationalization and Localization of Software

The internationalization of a product, such as software, is not the same as its localization although they may address many similar issues. Internationalization refers to the process whereby a product is made or adapted so that it can be used internationally (i.e., in different countries or regions all over the world with different cultures and conventions) without redesign. On the other hand, localization addresses how a product may be tailored for a specific country, region or culture by making it linguistically and culturally appropriate. Internationalization is often referred to using the abbreviation "I18N" or "i18n", where the number 18 refers to the number of letters omitted. Similarly, the abbreviation "L10N" or "l10n" is used for localization.

It is important that application software that is meant for deployment in many different countries with different cultures and languages be designed with internationalization in mind, to be able to accommodate possibly different ways of expressing an item of information or peculiarities of a different language. Some of the issues that internationalization needs to grapple with include:[71]

- Date and time formats
- Currency format
- Language peculiarities (e.g., alphabets, numerals and left-to-right script vs. right-to-left)
- Language character coding sets for textual display
- Names and titles
- Sorting of names and text
- Identification numbers, e.g. social security and passport numbers
- Telephone numbers, addresses and international postal codes
- Weights and measures

While the cultural and linguistic demands may change from country to country, the core program dealing with the functionalities of a software product do not change and so it is common practice to separate text and other environment-dependent data from the program code itself. This makes it easier to support internationalization as changes only need to be made to the environment-dependent resources. Minimal code changes are required.

The better internationalized an application is, the easier it is to localize. This is because a well-internationalized application will have built-in support to cater to items that are needed for localization. These may include:[71]

- Language translation
- Hardware support for certain languages, e.g. input devices and methods
- Local customs
- Local content
- Aesthetics
- Cultural values and social context.

The major work of localization is in translating the user interface and documentation but it involves more than just translating the language used. It also needs to cater to other relevant changes such as the usage of appropriate cultural and social values, symbols peculiar to the language, display of numbers,

dates, currency, appropriate input methods, etc.

In software internationalization and localization, a set of parameters, termed a locale, is used to define the user's language, country and any special variant preferences that the user wants to see in the user interface.[72] A locale identifier usually contains at least a language and a region/country identifier. Depending on the operating platform/system used, locale identifiers can be defined in several ways. Most systems utilize the two- and three-letter language codes defined by ISO 639-1 and 639-2, respectively, for the language identifier and the two-letter country codes from ISO 3166-1 for the country identifier. However, MS Windows uses a numeric Locale Identifier (LCID) that specifies the language and sort identifier.[73]

Standards Important to I18N and L10N

In this section we shall look at some important standards which are used in i18n and l10n.

Unicode and ISO/IEC 10646

Proper rendering and display as well as practical input methods for multilingual text on a computer system are essential if efforts to make software available in multiple languages are to be successful. Standards are needed for character code tables and character encoding methods. Character code tables assign integer numbers to characters while character encoding is a method by which characters or their respective integer values can be represented as a sequence of bytes for use by the software.

The international standards ISO/IEC 10646[74] and the Unicode Standard (Unicode)[14] describe and define the Universal Character Set (UCS), which is a superset of all other character set standards. It guarantees round-trip compatibility to other character sets. This means simply that no information is lost in the conversion of any text string to UCS and then back to its original encoding.[75]

The Unicode Standard Version 4.0 and ISO/IEC 10646:2003 make use of the same character set tables and character encoding methods, but the Unicode Standard additionally provides details of character properties, processing algorithms, and definitions that are useful to implementers.[74]

ISO/IEC 10646 and Unicode define several encoding forms, UCS Transformation format 8 (UTF-8), UCS-2, UTF-16, UCS-4 and UTF-32. In an encoding form, each character is represented as one or more encoding units and apart from UTF-8, all other encoding forms have an encoding unit larger than one octet (an 8-bit byte), making them hard to use in many current applications and protocols that assume 8- or even 7-bit characters.[76] UTF-8 uses all bits of an octet for its encoding and it preserves the full US-ASCII range, the latter being encoded in one octet having the normal US-ASCII value. This is important and very useful since it is backwardly compatible with the large existing volume of software that predominantly uses US-ASCII encoding. UTF-8 encodes UCS characters as a varying number of octets, where the number of octets, and the value of each, depend on the integer value assigned to the character in the Unicode character code table.

Unicode has become the dominant encoding scheme in software internationalization and usage in multilingual environments. Many other standards such as XML have adopted Unicode as the underlying scheme to represent text. Modern operating environments like those under GNU/Linux, Mac OS X and MS Windows XP have support for Unicode.[77]

ISO 639

The international standard, ISO 639-1, provides a two-letter code identifier (alpha-2) for the representation of names of languages while ISO 639-2 provides a three-letter identifier (alpha-3) for the languages.[78] Locale language identifiers make use of the ISO 3166 country codes to identify the language to use.

ISO 639-1 was devised mainly for use in terminology. It provides identifiers for those languages that are responsible for a major proportion of the world's literature and which also possess specialized vocabulary and terminology.

ISO 639-2 tries to provide a representation to the world's languages, for use in bibliography as well as terminology, but it is not as restrictive in scope as ISO 639-1. It was devised to include languages that are

Table 8: Sample ISO 639-1 and 639-2 Language Codes

639-2*	639-1	Language Name
apa		Apache languages
ara	ar	Arabic
bur/mya	my	Burmese
chi/zho	zh	Chinese
dut/nld	nl	Dutch; Flemish
eng	en	English
hin	hi	Hindi
kar		Karen
kin	rw	Kinyarwanda
tlh		Klingon; tlhIngan-Hol
may/msa	ms	Malay
nep	ne	Nepali
swa	sw	Swahili
tam	ta	Tamil
tha	th	Thai
ton	to	Tonga (Tonga Islands)

** For the 639-2 codes, where two codes are provided, the bibliographic code is given first and the terminology code is given second.*

most frequently represented in the total body of the world's literature, regardless of whether specialized terminologies exist in those languages or not. The three-letter code for ISO 639-2 means that it can accommodate more languages. So, while it limits coverage of individual languages to those for which at least modest bodies of literature have been developed, other languages are still accommodated by means of identifiers for collections of languages, such as language families.[78]

Under ISO 639-2, some languages have different codes for bibliography and terminology (see Table 8).

ISO 3166-1

ISO 3166-1 provides two (alpha-2) and three-character (alpha-3) codes for representing names of countries. It thus provides a table of country codes just as ISO 639 provides a table of language codes. However, these two standards were developed independently, and there was no attempt to use the same code for a language as that for the country in which it is spoken, and codes from each list should be used independently. Locale country identifiers make use of the ISO 3166 codes to identify the country or region location.

The ISO 3166-1 alpha-2 code is probably best known in its usage for the country code top-level domain (ccTLD) of the Internet Domain Name Service (DNS) system. However, there are several ccTLDs in use which are not part of the ISO 3166-1 two-letter codes, e.g., "uk" for the United Kingdom (the corresponding ISO 3166-1 alpha-2 code is "gb").

RFC 3066

The IETF's RFC 3066[79] describes a language tag for use in cases where it is desired to indicate the language used in an information object, how to register values for use in this language tag, and a construct for matching such language tags. RFC 3066 specifies use of a two-character language code from ISO 639-1 when it exists and when a language does not have a two-character code assigned, the three-character code is used.

The RFC also specifies the use of optional subtags (e.g., a country code from ISO 3166) and how to

Table 9: Sample ISO 3166-1 Alpha-2 Country Codes	
ISO 3166-1 (Alpha-2)	**Country/Region**
CA	Canada
DE	Germany
GB	United Kingdom
KE	Kenya
NG	Nigeria
TH	Thailand
TN	Tunisia
VE	Venezuela

register a dialect or variant information with the Internet Assigned Numbers Authority (IANA) when there is no available ISO 639 code.

Internationalization and Localization Software Initiatives

In the past, the language supported in software was very much dependent on where the authors were from. So many common off the shelf (COTS) software were written mainly for the English language due to the dominance of countries like the USA in this area. In recent times, with the emergence of the Internet and globalization, this predominantly single language-centric support for popular software is changing. There is growing awareness among software developers and authors that many softwares can be and will be deployed worldwide and it is important to be able to adapt the software to the local environment. As a result, there is much better support for internationalization and localization on modern software platforms.

For commercial proprietary software, experience has shown that any localization effort has to be considered in the light of economical viability and/or other benefits that the effort may bring to the vendor. This means that, in many cases, versions of popular commercial proprietary software are not available for languages or cultures where commercial returns are not justified. Since FOSS can be freely modified and redistributed, at times all that is needed is some interested party to take the initiative to localize software that is released as FOSS. This has resulted in many popular FOSS being localized (e.g., the Mozilla.org family of products, GNOME, KDE, OpenOffice.org) and made available in many languages, including some rather obscure ones.

The Open Internationalization Initiative

The Open Internationalization (OpenI18N) Initiative[80] is a key initiative under the Free Standards Group.[27] This initiative has several active projects under it. One of them is the OpenI18N Specification which is concerned with the specification for interfaces and functionalities that must be supported by GNU/Linux-like operating systems to run internationalized application software, as well as recommendations for such operating systems to facilitate the development of internationalized application software.[81] Other active projects include:

- Linux Internationalization Locale Name Guideline
- Common XML Locale Repository
- Internet Intranet Input Method Framework
- OpenI18N Certification Test Suite
- Multilingualization library (m17n-lib)

All the standards, publications and documentation from the OpenI18N Initiative are freely available.

Some FOSS I18n and L10n Initiatives

Most of the FOSS I18N and/or L10N projects are community-driven. Almost all major FOSS have good support and tools for I18N and L10N. Local users of the software are encouraged to contribute to the L10N projects.

Mozilla Family

The Mozilla Localization Project (MLP)[82] relies mainly on the FOSS community to make the products from the Mozilla Foundation available for different world cultures and languages. The project is focused towards software localization making use of the underlying internationalization support available in the products.

The software localization projects under MLP include:

- Mozilla (aka project Seamonkey) with over 100 languages registered
- Mozilla Firefox with over 30 languages registered
- Mozilla Thunderbird with over 50 languages registered

GNOME

The aim of the GNOME Translation Project[83] is to translate GNOME applications and documentation to every language in existence. This community-based effort currently boasts of translation projects covering well over 100 languages.

K Desktop Environment

The popular K Desktop Environment (KDE) software also has wide support for its internationalization and localization initiatives.[84] There are good guides and documentation available, and again community-driven projects for localization are well supported and received. As a result KDE is currently available in over 100 languages.

OpenOffice.org

OOo has a framework and tools for both I18N and L10N.[85] OOo is now available in over 70 languages covering all major languages and cultures of the world and also some minor ones.

Microsoft Software

The newer versions of software from Microsoft, e.g., MS Windows XP, MS Office 2003 have good internationalization support and are also available in many localized native versions.

MS Windows XP

Localized versions of MS Windows XP are available in 24 languages and the Multilingual User Interface (MUI) Pack offers more localized user interface languages. The MUI Pack is a set of language-specific resource files that can be added to the English version of MS Windows. Microsoft claims that the total number of languages supported in MS Windows XP is in excess of 140.[86]

MS Office

Localized versions of MS Office 2003 are available in over 35 languages.[87] In addition, the MS Office MUI offers support for other languages for which a localized version is not available.

PATENTS IN STANDARDS

A patent is a set of exclusive rights given by a government to a patent applicant in which the patent holder is granted the right to prevent others from making, using, selling, offering to sell or importing the invention for a specific period of time. Patents are usually granted for inventions that are considered to be non-trivial, new and novel. Patent grants are territorial in nature in that patents applied for and granted in one country are not automatically recognized in another country. Examples of patents are:

- Wankel rotary engine
- Hume concrete pipes
- Design of the Coca-Cola bottle

Software Patents

Traditionally, patents are given mainly to physical inventions but in recent times many countries have begun to grant patents for non-physical items such as business methods and computer programs (software). Software has become patentable in countries like the USA and Japan. The issues on the patentability of software and the way patent offices process software patent applications are very controversial.[88,89,90,91,92] In countries where software patents are recognized, patents may be granted to functional aspects of software that are considered to be innovative and non-obvious. The expressive elements of code are not patentable. Instead, they are covered by copyright to which almost all the countries in the world subscribe to. While many countries still do not recognize software patents, most are re-examining this issue and trying to decide whether they should change their positions. The inclusion of software patents in IT-related specifications and standards has attracted a great deal of discussion.

Policies on Patents

In the case of technical standards, it is not uncommon for patented items to be proposed to be included as part of the specifications. The standards development body has to decide on whether it should use such an item or look for an alternative. In the past, the development of standards related to software and IT has proceeded using mainly a 'reasonable and non-discriminatory (RAND)' terms policy whenever patents are included in a standard. Under RAND, the patent holder must be willing to negotiate rights to use the essential patent on reasonable and non-discriminatory terms. The intent of RAND was to prevent patent issues from hindering the adoption of a standard and to ensure that the cost of any necessary licenses needed, arising from the patent, are affordable. This has proved adequate in the past but, in recent times, the increasing proliferation of patents granted to software-based innovations (including software patents) has led standards developing and setting bodies all over the world to clearly state their patent policies to ensure that they are adequate and will continue to support the development of highly successful and widely used standards as they have in the past. Since, in general, a standard is targeted for use by all in the world, it is vital that the terms of usage of any patent that is included in the standard are clearly specified.

Standards have been produced that include patented technologies and all the main standards bodies have policies with regard to the treatment of patents in the documents that they produce.[93]

ISO

ISO has published directives on the issue of patents in its standards development process.[94,95] There is a strong recommendation to avoid references to patented items in ISO publications. Nevertheless, ISO recognizes that for technical reasons, sometimes this may not be possible and, in such exceptional situations, it does not object in principle to the inclusion of items covered by patent rights even if the

terms of the standard are such that there are no alternative means of compliance. During the preparation of the ISO document, a basic text for the identification of patent rights is to be inserted into the draft documents in those cases where compliance with an ISO document may involve the use of a patent.

IETF

RFC 3979[96] is the main document dealing with the IETF's stand on patents. In general, IETF prefers technologies with no known patent claims or patents that offer royalty-free licensing. However, the IETF working groups have the discretion to adopt technology with a commitment of RAND terms, or even with no licensing commitment, if they feel that the technology is superior enough to alternatives with no such patents or licensing encumbrances.

In order for the working group and the rest of the IETF to have the information needed to make an informed decision about the use of a particular technology, a person contributing to the working group's discussions must disclose the existence of any patent claims that the individual is reasonably and personally aware of and that he (or his employer) owns or controls.

W3C

W3C has a very clear policy with regard to patent usage in its Recommendations. It seeks to issue Recommendations that can be implemented on a Royalty-Free (RF) basis. This has arisen from the experience it had with the WWW.

Many early standards (Recommendations) from W3C paid scant attention to patents. Later, as the Web became more commercial and software and business process patents increased, patent infringement issues surfaced as several patent holders, including some who had participated in the development of the standards themselves, sought license payments. As a result, W3C decided to have a clear patent policy governing the Recommendations that it develops.[97]

The key position of W3C with regard to patents that are deemed essential to a Recommendation (it calls them "essential claims") is that they have to be available for implementation in accordance with the W3C RF License requirements. An "essential claim" refers to a patent for which there is no known alternative and, therefore, it is essential to the implementation of a normative part of a Recommendation.[98]

The policy generally requires that a participating organization in a W3C working group formally commits to the RF requirements for "essential claims". The participants are not required to disclose known patents as long as their participating organization commits to licensing those patents according to RF requirements. In the event that a working group participant holding a patent does not want the patent to come under RF requirements, there is some flexibility in the policy in that it allows the participant to exclude specific patent claims from the RF commitment, provided the working goup is informed within a well-defined time limit. In this manner, a participant can still participate while specifying that strategic technology be excluded from the RF process and the working group is made aware of a potential patent conflict. As far as possible, the working group will try to resolve this conflict. However, in the event that it cannot be resolved, the matter is referred to the Patent Advisory Group (PAG) task force which will attempt to resolve the conflict. Ultimately after exhausting all other options, if the PAG does indeed recommend that an alternative to the RF licensing requirements be used, it has to go through several levels of review and consensus before W3C accepts the alternative.

W3C policy requiring commitment to the RF requirements by default is a stricter policy as compared with the RAND policy of ISO and IETF.

OASIS

OASIS has a published policy which governs the treatment of patents that are considered as "essential claims" (patents that are deemed essential for the implementation of a normative part of an OASIS standard), in the production of specifications and other works by OASIS.[99]

Unlike the W3C, OASIS does not have a single licensing agreement for "essential claims"; instead it uses three types: "Reasonable And Non-Discriminatory (RAND)", "Royalty-Free (RF) on RAND Terms" and "RF on Limited Terms".[100]

RAND defines a basic set of minimal terms that a patent holder is obliged to offer (such as granting a license that is worldwide, non-exclusive, perpetual, reasonable, and non-discriminatory, etc.) and leaves all other non-specified terms to negotiations between the patent holder and the implementor seeking a license.

RF on RAND Terms is the same as RAND with the exception that no fees or royalties are to be charged.

RF on Limited Terms specifies the exact royalty-free licensing terms that may be included in a patent holder's license and that must be granted upon request without further negotiations.

Summary of Patent Policies of Standards Organizations

As can be seen from the discussion in the previous section, most standards bodies do allow the inclusion of patents in their standards although patent-free ones are preferred. Their patent policies all revolve around allowing a RAND policy, either with some form of royalty payment or royalty-free or a mixture of both. This practice is based on the view that RAND licensing appropriately balances the legitimate rights of patent owners, who contribute innovative technology to the standard, with the interests of implementors who wish to obtain access to essential patents on reasonable terms.

RAND Licensing Terms and FOSS Licenses

The possibility that patents under RAND terms can be included in standards has very important implications for software that is released under a FOSS license. FOSS licenses usually include terms that satisfy the following clauses of the Open Source Initiative's Open Source definition.[101]

- Free Redistribution

  The license shall not restrict any party from selling or giving away the software as a component of an aggregate software distribution containing programs from several different sources. The license shall not require a royalty or other fee for such sale.

- Derived Works

  The license must allow modifications and derived works, and must allow them to be distributed under the same terms as the license of the original software.

- Distribution of License

  The rights attached to the program must apply to all to whom the program is redistributed without the need for execution of an additional license by those parties.

Although all FOSS licenses share these characteristics, the actual requirements and obligations imposed can vary from license to license. For example, the BSD license requires only copyright attribution and license reproduction, and redistributions of the software may be made under any other license. However, the Mozilla Public License (MPL) impose moderate obligations in that they require that specific files containing MPL code be distributed in source-code form and under the terms of the MPL. GNU GPL requires that any work that includes GPL code, if distributed at all, be distributed under the terms of the GPL. It also clearly states that any patent must be licensed for everyone's free use or not licensed at all. FOSS licenses, then, do differ with regard to the nature and degree of rights and obligations described. Consequently, licenses like the BSD allow the usage of technology available under RAND terms but GPL does not allow any GPL-based public distribution to include any technology available under a RAND license that is not royalty-free. The licenses cited above are the most commonly used FOSS licenses with GPL by far being the most popular one. This has the implication that a large number of FOSS products may be incompatible with RAND licensing. In connection with this issue, the Free Software Foundation has stated that RAND licensing discriminates against free software[102] as it is generally not possible for software to be freely modified and redistributed under RAND licensing terms.

Patent Offerings to the FOSS Community

To waylay the concerns of FOSS developers and users, and to reduce the fears of software patents infringement by FOSS developers, several commercial companies have recently offered all or part of

their portfolio of software patents on a no cost basis to the FOSS community for use. IBM has announced that, for a start, it will allow royalty-free use of 500 of its software patents[103] in any software that is released under an Open Source license (as recognized by the Open Source Initiative). Red Hat, a company well known for the development and commercial distribution of GNU/Linux, has offered unfettered use of its own software patents portfolio to Linux developers. Novell has said that it will use its existing patent portfolio to protect the Linux kernel and other Open Source programs included in Novell's offerings against potential third-party patent challenges. Sun Microsystems has released over 1,600 patents for use with software that is licensed under the Open Source Common Development and Distribution License (CDDL).

A Patent Commons Project[104] has been started by the Open Source Development Labs (OSDL). This initiative is aimed at the creation of a central depository where software patents and patent pledges can be housed for the benefit of the open source development community and industry. Companies that have contributed and pledged patents to this project include Computer Associates, IBM, Nokia, Novell, Red Hat and Sun Microsystems.

There is much controversy and debate over patents in software development and RAND in standardization (see Annexure: Comments on RAND, as Seen from Both Sides).

THE LINUX STANDARD BASE

The GNU/Linux operating system consists of the Linux kernel itself and, together, the rest of the system software and tools/utilities make up the operating system. Most of the system software is from the GNU Project.[105] In addition, for an operating system to be useful to most people, it has to be made available with support for some application software. The strong community-based history and support of GNU/Linux together with the nature of the licensing of the Linux kernel and GNU software resulted in many people taking the kernel, system software from GNU and possibly other FOSS utilities/tools, adding in some application software which they deem useful, and putting all these together to form a working package. This working package is termed a GNU/Linux distribution or distro. Consequently, the GNU/Linux operating system comes in very many distros.[106] The large number of distros available, coupled with the fact that since most software, if not all, included in a distro are FOSS and, hence, can be customizable to suit the requirements of a particular distro, have resulted in a fair measure of binary and configuration incompatibilities among distros. Some incompatibility problems include different library versions, package formats and differences in directory and file layouts. It has been recognized that if GNU/Linux is to be fully embraced and supported by mainstream computing as a legitimate alternative to proprietary operating systems, there is a need to cut down on these incompatibilities so that a software package with source can compile cleanly across distros and a binary version can run properly across all distros. The Linux Standard Base Project (LSB)[107] tries to do this by specifying a standard for GNU/Linux.

What is the Linux Standard Base?

The Linux Standard Base is a project under the Free Standards Group. It attempts to develop and promote a set of binary standards that will increase compatibility among GNU/Linux and other similar systems. These standards will also enable software applications to run on any conforming system.

While the main goal of the LSB project is to increase compatibility among GNU/Linux distributions by specifying and promoting standards for their use, it does not limit the applicability of the specification to only the GNU/Linux environment. The LSB specification has been written so that it can be readily implemented on any UNIX-like operating system, natively or as a compatibility layer. With some more work, it can also be implemented on other operating systems.

The LSB is a community-based project and anyone can contribute to it by participating in the various LSB mailing lists. There is considerably good support for the LSB standard among commercial software vendors like Mandrakesoft, Miracle Linux, Novell, Progeny, Red Flag, Red Hat, IBM, Oracle, Veritas, MySQL, etc.

The Linux Standard Base Specification

The LSB comprises a single common (generic) specification and architecture specific specifications. The complete specification for a particular platform consists of the generic specification plus one of the architecture specifications. Architectures supported currently are IA32 and IA64 (Intel 32- and 64-bit processors), PPC32 and PPC64 (IBM's 32- and 64-bit PowerPC family), S390 (IBM's S390 processors) and S390X (IBM zSeries processors), and AMD64 (Advanced Micro Devices 64-bit processors).

The LSB defines both a set of Application Program Interfaces (APIs) for source code and Application Binary Interfaces (ABIs) for compiled binaries. A conforming implementation has to support all of the ABIs in the LSB but not all of the source-level APIs.

The LSB is divided into specification modules in which a specification module refers to a unique collection of one or more functions that have value for a certain group of runtime implementations. The modules currently available are LSB-Core, LSB-C++, LSB-Graphics and LSB-I18N. Both LSB-Core and LSB-C++ have generic and architecture-specific specifications while the LSB-Graphics and the LSB-I18N have only the generic specification. Table 10 summarizes the currently available modules.[108]

Table 10: LSB Modules		
Module	**Functional Area**	**Architectures Available**
	ELF	Generic, Processor-specific
LSB-Core	LSB	Generic, Processor-specific
	Packaging	Generic, Processor-specific
LSB-CXX	LSB-C++	Generic, Processor-specific
LSB-Graphics	Graphics	Generic
LSB-I18n	OpenI18n	Generic

The latest version of the LSB is 3.0.0. LSB 2.0.1 had been submitted to ISO to become an international standard for GNU/Linux.

LSB-Core Specification

This is the Core module of the Linux Standard Base. This module provides the fundamental system interfaces, libraries, and runtime environment upon which all conforming applications and libraries depend. It provides specifications for the following areas:

- Executable and Linking Format (ELF)
- Base libraries
- Utility libraries
- Command and utilities
- Execution environment
- System initialization
- Users and groups
- Package format and installation

The specifications make extensive use of existing standardized APIs and ABIs from other bodies. Some normative references include those from ISO POSIX, the System V Interface Definition (SVID) and the Filesystem Hierarchy Standard (FHS).

In particular, the LSB-Core specification includes many interfaces described in ISO POSIX (ISO/IEC 9945)[109] and it specifies that such interfaces should behave exactly as specified in the POSIX standard. It is also the long-term plan of the LSB to converge with ISO/IEC 9945.

One of the problems plaguing the many different GNU/Linux distros has been the various formats used in software package distribution. The LSB addresses this by specifying that applications shall be packaged in the RPM packaging format as defined in the LSB, or supply an installer which is LSB conforming (for example, by invoking LSB commands and utilities). This means that while packages are encouraged to be supplied in Red Hat Package Manager (RPM) format the LSB does not mandate the use of the RPM program or database.

LSB-C++ Module

This is the C++ module of the LSB. It supplements the core interfaces by providing system interfaces, libraries, and a runtime environment for applications built using the C++ programming language.

Normative references include the LSB-Core, IOS POSIX and the ISO/IEC 14882 C++ Language standard.

It provides specifications for the following areas:

- Low level system information
- Base libraries
- Package information

The LSB-Graphics Module

This specification defines the graphical interface found on an LSB conforming system. Normative references include the LSB-Core and graphic libraries and specifications from the The X.Org Foundation.

It provides specifications for the following areas:

- Graphic libraries
- OpenGL libraries
- Package information

The LSB-I18N Module

This module corresponds to the OpenI18N Global Specification from the OpenI18N Project.

Linux Standard Base as an ISO Standard

LSB 2.0.1 had been submitted to ISO for use as an international standard for GNU/Linux through the ISO PAS (Publicly Available Specification) process and this was recently approved as the standard ISO 23360.

The availability of an ISO GNU/Linux standard is an important milestone, symbolically, in the development of GNU/Linux. It signifies that the GNU/Linux operating environment has come of age and is now officially recognized as a full-fledged mainstream computing platform. As a result, corporations and governments, that so far have been reluctant to use GNU/Linux due to uncertainty regarding its long-term viability and international acceptance, now have the confidence to consider it on an equal footing with other more established operating systems. An ISO GNU/Linux standard will also help the acceptance and usage of FOSS in general as many FOSS products are implemented on GNU/Linux and it is arguably the most well-known FOSS product.

Linux Standard Base Certification

GNU/Linux distributions that conform to the LSB can be certified as such. The LSB certification scheme is run on behalf of the Free Standards Group by the Open Group,[110] a vendor- and technology-neutral consortium, to ensure neutrality and confidentiality. Certification charges are kept to a minimum to encourage developers, Independent Software Vendors (ISVs) and GNU/Linux distributions to become LSB certified.

LSB certification is currently available for the following:

- LSB Runtime Environment
- LSB Application
- LSB Internationalized Runtime Environment

Developers and vendors are granted a license to use the LSB Certified trademark in connection with a particular product, once it has passed the applicable certification test suites.

GOVERNMENT/NATIONAL OPEN STANDARDS POLICIES AND INITIATIVES

Many governments all over the world have begun to realize the importance of open technologies and technical specifications, and the ability to participate in the development of these technologies and specifications. In parallel to this, they have also taken notice of FOSS and the benefits that it can bring to a nation. As a result, in many countries, the government has come out with policies and/or initiatives which advocate and favour open standards in order to bring about increased independence from specific vendors and technologies and at the same time to accommodate both FOSS and proprietary software. This is especially true for most e-government projects and initiatives all over the world.

Embracing open standards can also assist FOSS to flourish in a country. In the Introduction, it was mentioned that open standards helped to popularize FOSS to a large extent. FOSS can interoperate well with established proprietary software and technologies with the aid of open standards, thereby making its implementation more feasible. As such, countries that are looking to FOSS should also look at specifying open standards too.

Open Standards Favoured by Governments

e-Government initiatives as well as many government agencies now favour the use of open standards where possible. Some open standards that are frequently specified are:

Networking protocols – TCP/IP

Networking services – HTTP, SSL, SMTP, MIME, IMAP, LDAP

Document exchange – XML and XML-based specifications

Web services – UDDI, SOAP

Database – SQL

Internationalization - UNICODE

Some Government Open Standards Policies and Initiatives

This section will list out some of the government policies and initiatives of several countries with respect to open standards.

The European Union

The European Union (EU) comprises many nation states with many diverse cultures and languages at varying states of technology/technical development. For it to be able to function effectively, especially in the area of information exchange, the governments concerned have to establish a proper interoperability framework and standards on data interchange. The development of the European Interoperability Framework (EIF),[111] a framework for the e-government services of the member states to facilitate the interoperability of these services at pan-European level, is taking place under the European Commission's Interoperable Delivery of European eGovernment services to public Administrations, Business and Citizens (IDABC) Programme. The EIF version 1.0 recommends the use of open standards for maximum interoperability among e-government services. It defines the minimal characteristics for open standards as the following:

- the standard is adopted and maintained by a not-for-profit organization;
- the development of the standard occurs using an open decision-making process and does not

preclude any party from it;

- the standard is published and is available either free of charge or for a nominal fee;
- the published standard must be available for all to copy and distribute, either free of charge or for a nominal fee; and
- any patents present in the standard are to be irrevocably available on a royalty-free basis.

The definition above has attracted controversy even though the area of validity is confined to pan-European projects carried out in the context of the IDABC Programme. FOSS groups and advocates have welcomed it but other groups, including ANSI, BSA and EICTA, have criticized it, particularly with respect to the last two criteria. These parties point out that they are inconsistent with the approach taken by other standards development organizations that acknowledge the right of patent holders to charge reasonable royalties and to place reasonable restrictions on the licensing of their essential technology covering an open standard.

The European Commission's IDA expert group on open document formats has recommended that the European Union's public sector use open formats in their electronic documents.[112] For revisable documents, XML-based formats like the Open Document format from OASIS and Microsoft's new XML-based MS Office formats are recommended.

United Kingdom

The United Kingdom's e-government initiative places a lot of emphasis on open technical standards to achieve seamless information flow across the public sector and to provide citizens and business with better access to government services.[113] Its e-Government Interoperability Framework (e-GIF) defines the technical policies and specifications governing information flows across government and the public sector. Complying with e-GIF at the highest level includes the use of open standards like XML as the primary means for data integration and the implementation of Internet and WWW standards.[114]

Denmark

The Danish e-Government Interoperability Framework includes recommendations and status assessments for more than 450 selected standards, specifications and technologies used in its e-government solutions.[115] In general, the Framework recommends the use of open standards and centrally agreed XML schemas (which may be provided free of charge throughout the public sector) for data interchange.

As part of the Interoperability Framework, the policy on data and document exchange specifies that documents should be published in generally available formats for which free readers exist and the use of proprietary word processing formats such as MS Word or formats that do not have widely available readers should be avoided for publicly available documents.[116]

The Netherlands

The Netherlands has its OSOSS – the programme for Open Standards and Open Source Software in government.[117] This programme encourages the use of open standards and provides information on open-source software. The Dutch ICTU, the organization for ICT and government programme, runs OSOSS. While the programme targets the public sector, its results will be available for the private sector and individuals too. The programme provides information and advice to the public sector on open standards. It has set up a catalogue of recommended open standards[118] for use in the public sector.

Norway

The Norwegian Government has declared that proprietary formats will no longer be acceptable in communication between citizens and government.[119] As part of its "eNorge 2009 – the digital leap" master plan for IT – all public sector bodies in Norway are to have in place a plan for the use of open source software and open standards by the end of 2006.

Massachusetts, USA

The Commonwealth of Massachusetts, USA, has announced an IT Policy that emphasizes the importance of open standards compliance for IT investments in Massachusetts.[120] The Policy states that all prospective IT investments will have to comply with open standards referenced in the current version of the Enterprise Technology Reference Model (ETRM) of the Commonwealth. It further says that existing IT systems will be reviewed for open standards compatibility and will be enhanced to achieve open standards compatibility where appropriate. In addition, open standards solutions will be selected when existing systems are retired off or need major enhancements.

New Zealand

As part of its e-government vision, New Zealand has come up with a supporting Information Systems (IS) Policies and Standards document. The guiding principles state that the IS Policies and Standards are to be based on open standards, wherever possible.[121] New Zealand also has an e-Government Interoperability Framework (NZ e-GIF)[122] which lists the mandatory use of many open standards for compliance.

Malaysia

The Malaysian Government Interoperability Framework (MyGIF)[123] defines the minimum set of IT standards and technical specifications for use in government ministries, agencies and departments. These cover the areas of interconnection, data integration, information access, security and meta-data. Instead of creating new standards or specifications, MyGIF adopts internationally recognized open and *de facto* IT standards and specifications for all the interoperability areas mentioned.

Chile

The Government of Chile, in 2004, issued Decree 81[128] whereby all public agencies and services are required to format documents in XML. A three phase roll-out deployment plan is being implemented with the final stage scheduled for completion by 2009.

India

The Government of India has started its eBiz initiative[129] – a project to build a framework for Government to Business (G2B) services where services from the federal, state and local government agencies will be made available through a single portal. The eBiz architecture is to be built on the principles of interoperability and open standards.

Others

Many other countries, e.g. South Africa, Viet Nam, Brazil and Peru, have started initiatives and/or policies to address the digital divide and to improve their government's IT implementations by leveraging FOSS and open standards together.

Challenges in Implementing Open Standards in Government Procurement Policies

In line with their national IT policies favouring open standards, many government IT software procurement policies have specified that products and solutions should support and implement open standards before they can be considered. However, there are several challenges to overcome if this is to be put into practice. The reality is that, sometimes, open standards may not be available or are not mature enough for a required technology. Also, in some cases, the usage of a *de facto* standard is so entrenched that it is not practical to ignore it.

One way to overcome some of the problems highlighted above is to recommend, in cases where open standards are either not available or mature enough, some other specification (e.g., a *de facto* standard) for use in the meantime. This interim specification should be chosen with care. It should be one that is widely used and, as far as possible, the specification should be publicly published and available for

implementation so that software that can support it is available from different sources or vendors. Examples of these include the zip data compression file format and the PDF format for office documents and brochures.

It is important that these points are considered to minimize the possibility of a lock-in by using a proprietary specification when an open one is not available or practical. The organizations and/or vendors responsible for the specification should be encouraged to submit it to an open standards body for adoption/adaptation as a standard. An example of this line of approach are the developments[124, 125] leading to the adoption of the OpenDocument format as the standard for document exchange by the EC's IDABC Programme.

In environments where the usage of a proprietary standard is so widespread and ingrained that it is not practical to ignore it or replace it immediately, in the interim, some means of interoperability and/or file format conversions may have to be found so that the proprietary standard can interoperate with installations using open standards. A careful implementation and operation plan has to be worked out to ensure adequate interoperability. It may also be necessary to implement the applications using the open standard in phases, so as to allow adequate time to phase out the older applications of the proprietary standard.

Summary

As can be seen from the discussion above, most of the world's governments are asking for the adoption of open standards and specifications as much as possible in their country's IT usage and/or e-government projects and initiatives. This is a good move since, if more and more countries are to insist on open standards, more software vendors (both FOSS and proprietary software) will be forced to open up their file formats and technology specifications, and adhere to open standards in their products as much as possible, thereby, further enhancing the interoperability of disparate products and systems.

CONCLUSION

This primer has tried to explain what technical standards are and the key characteristics of what may be termed as open standards in the field of information technology. Specifications that satisfy these characteristics can be viewed as open ones and those that are in widespread use and acceptance may be regarded as open standards.

Open IT standards are even more important in this present information age of IT and communications convergence and the Internet. No single technology, group or vendor can provide for everything and, therefore, interoperability in a heterogeneous environment is required more than ever. It is only by strict adherence to standards and specifications that a high degree of interoperability can be achieved. Standards that are open and non-discriminatory are preferred because there is no dependence on any single entity, all types of products can implement them and all interested parties can partake in their development.

XML and related technologies are expected to play an important role in setting new standards for better interoperability and information exchange in the areas of Web applications, services and e-commerce, as well as in office applications. It is crucial that these standards are steered and developed by open standards bodies. Towards this end, it is very important that bodies like W3C, OASIS, IETF, ISO, IEEE remain open and support non-discriminatory policies especially with regard to intellectual property rights issues.

In many environments, the demand and usage of open standards go hand-in-hand with FOSS. There have been many successful FOSS implementations of open standards and so it is not surprising that many see them as working in tandem. FOSS has much to gain from open standards and wide spread adoption of the latter will help FOSS proliferate as the Internet has demonstrated. However, as pointed out in the primer, FOSS and open standards are two distinct and different domains and it is possible to have a proprietary software product implement open standards and a FOSS product make use of a proprietary specification.

The software localization initiatives of many countries will benefit from the setting and availability of more open standards in the relevant areas. The easy and free access to open standards related to internationalization and localization will encourage more local people to participate in these initiatives.

More and more governments are asking for open standards now and this is a very good sign as they are the biggest buyers and consumers of IT products and software. The vendors will have to comply with open standards and open up any proprietary file formats or specifications in response to these demands. In conjunction with this, it is hoped that more and more users too will follow suit.

It is the aim of this primer to help educate and make the reader aware of the benefits of open standards in terms of enhancing interoperability in an increasingly heterogeneous environment. It should be the ultimate objective of users to be able to access and use applications and services using any device, platform or interface of their choice. At the same time, they should be able to exchange information and data from these applications/services with other users without suffering any degradation in content. Open standards represent one important possible way to achieve this objective.

ANNEXURE : COMMENTS ON RAND, AS SEEN FROM BOTH SIDES

There is much controversy and debate over patents in software development and Reasonable And Non-Discriminatory licensing or RAND in standardization. The following is a collection of comments by leading experts from the FOSS community, industry, and academia on the issues relating to patents in software development and standards, and in particular on the issues surrounding RAND.

Peter Chong, Corporate Attorney, Asia-Pacific Region, Microsoft Legal and Corporate Affairs (www.microsoft.com)

Proprietary does not imply "closed" or the opposite of interoperability; it simply indicates a different – often smaller and quicker – process was used to develop the standard at issue. Even when a standard is produced and controlled by a single entity and does not undergo a process of review and ratification through a standards body, as long as the standard is broadly accessible to all for use and implementation on reasonable terms, such a standard can still facilitate interoperability between products. PDF and ZIP (which is used in the OpenOffice.org file format submitted to OASIS) are two examples of standards that have achieved widespread acceptance and interoperability without being open standards, and while remaining under the ownership of a single entity. Other examples of broadly deployed proprietary standards that are widely licensed in the IT industry with a significant positive impact on interoperability include Adobe's PDF, Hewlett Packard's Printer Control Language (PCL), Microsoft's Office Open XML File Formats and the Java and Win32 APIs. Many well-known and useful open standards today also originated from a single entity before formal adoption as an open standard, e.g. ISA (a specification by IBM, later standardized by IEEE), and NFS (Network File System, introduced by Sun Microsystems, later standardized by IETF).

While adopting open standards is one good approach to achieving interoperability, it is not the only approach, and it is not the case that any other approach is without merit. Further, it is not the case that open standards exist in every field of technology. In many emerging areas, standards are still being developed by the industry. The reality faced by many governments and organizations is that there are numerous areas where open standards do not yet exist, and they have to work with other types of standards until a strong contender emerges. Premature governmental directives on the use of standards that are not yet ready and proven will stifle the industry and inhibit growth.

Governments and organizations should be, and are often, open to the possibility that some de facto standards, like PDF for example, are so widely and reasonably licensed and so broadly deployed and demanded that it makes sense to embrace such standards as part of a government's or organization's interoperability framework, rather than insisting on an open standard such as X.400. This is a much better approach than suggesting at the outset that all de facto standards, even those available royalty-free, should entirely be avoided by governments and other organizations. To completely ignore such standards as a rule may make little sense and may actually impair interoperability, particularly if there is no adequate open standard substitute for such a *de facto* standard. This is how governments (e.g. Australia) and other organizations (e.g. Asia-Pacific Economic Cooperation or APEC) continue to view and use such popular and accessible de facto standards. Governments and organizations should realize that proprietary standards and open standards are complementary ways to achieve interoperability.

Governments should avoid mandating standards themselves or mandating those that have not achieved broad industry support. Regardless of whether open standards and proprietary standards are at issue, voluntary, industry-led industry led processes have proven in all cases to be the most effective means of fueling innovation through standards. Most of the best known, widely adopted open and proprietary standards in existence today, e.g. TCP/IP, MPEG, HTML, 802.11, XML, SMTP, were developed through industry-led, voluntary efforts. This reflects the fact that the marketplace continues to be best situated to determine and develop the appropriate timing and levels of interoperability for products and services.

Microsoft has a responsibility to develop technology that is based on commonly used standards and to make our products work well with those of other companies – including our competitors. Microsoft enables interoperability between its products and thousands of third-party products and services in

various ways, including active participation in hundreds of standards organizations, support of open standards in its products, and publication and broad licensing of Microsoft's proprietary technologies. Go to www.microsoft.com/interop for more information.

Seow Hiong Goh, Director of Software Policy - Asia, Business Software Alliance (www.bsa.org)

There is no special relationship between open source and open standards – the two concepts are separate and distinct. Just as FOSS does not mean software "free-of-charge", open standards do not mean standards "free-of-charge" or "royalty-free". Openness in a standard relates to the process in which the standard is defined, ratified and maintained, and not with its cost. FOSS and commercial software are both equally capable of implementation that meets the requirements and specifications of open standards.

Key aspects of BSA's position on the EIF definition are shared by numerous other bodies such as EICTA, ITU-T, GSC, ANSI, ETSI, CompTIA and many others. We do not believe that these organizations are attempting to undermine FOSS. We do not think that many critical and well-established technologies that are implemented on the basis of accepted open standards and maintained by many existing standards bodies should no longer be considered so.

Common licensing provisions governing matters such as royalties, defensive suspension, reciprocity, field of use and sublicensing do not preclude developers from distributing standards implementations under the vast majority of open source licenses.

Certain FOSS licenses, particularly the GPL (see note below), contain restrictive terms concerning free redistribution, derived works and distribution of license that create a direct conflict with RAND terms. Even where there are no royalties or other fees associated with a RAND patent license, the GPL is still at odds with the field-of-use limitation, restriction on sublicensing and reciprocity requirement, the three common terms in standards-related patent licenses.

We believe that a licensing restriction that is unique to the GPL should not be generalized as being applicable to all open source licenses that may not face such restrictions. Most FOSS licenses are flexible enough to be used to redistribute software covered by an independent RAND patent license (assuming redistributors have taken steps to mitigate infringement risks).

Fortunately, while some conflict exists, given the increasing success and importance of FOSS, the proprietary software industry and the standards industry continue to explore ways to resolve any such conflicts and to forge an environment in which all software – commercial and FOSS alike – can co-exist and thrive, all with the ultimate goal of enhancing interoperability and consumer choice.

For a further discussion of the issues of technology standards and interoperability, please see: http://www.bsa.org/asia-eng/policy/upload/Technology-Standards-Interoperability1.pdf

Note: Unlike the vast majority of FOSS licenses, the GPL prohibits the distribution of software under the GPL if the software includes any patented technology that is licensed under terms at odds with the GPL. Hence, the requirement that the software be freely distributable, for example, is at odds with any RAND license which permits reasonable royalties, and accordingly the patented technology associated with such a RAND license may not be included in FOSS distributed under the GPL. A developer of FOSS distributed under the GPL may be unable to implement open standards covering RAND-licensed patented technology in its products, thereby impairing such software's interoperability potential.

Pamela Jones, Growlaw, USA (www.groklaw.net)

RAND terms at one time made good sense. Now RAND is a weapon to discriminate. For software, RAND terms no longer suit. Something new has happened in the world of software. The General Public License, like most other Open Source licenses, is incompatible with RAND unless the terms specify that the patents are 100 percent royalty-free. Because Linux and GNU/Linux are made available under the GPL, the dominant Open Source license, and because so many now wish to use GNU/Linux operating systems, including large corporations and government agencies, standards organizations have reached a fork in the road. They will have to adjust to include this new player if they wish to remain relevant.

The purpose of standards is to make the playing field fair to all and to enable interoperability, but RAND no longer accomplishes that purpose for software, and so an adjustment must be made to include GNU/Linux and the GPL.

Jeff Kaplan, Founder & Director, Open ePolicy Group, Berkman Centre for Internet and Society, Harvard University, USA (cyber.law.harvard.edu/epolicy)

Open standards are the subject of increasing public attention and debate. But what makes a standard open? Open standards must be royalty-free, but the discussion does not end there. The harder question is whether open standards can have other conditions if they are RAND. Despite its appealing rhetoric, RAND is not an objective benchmark. "Reasonable" is in the eye of the beholder; it is an undefined criteria only a lawyer could love. Worse, RAND can be a wolf in sheep's clothing, bringing new forms of lock-in under the guise of "openness."

Standards have degrees of openness, mainly due to restrictions and encumbrances placed upon them by vendors. The fact that a standards organization labels a standard open is not determinative. Its effect in the marketplace is a better guide.

At a minimum, open standards must allow all possible competitors to operate on a basis of equal access to the ability to implement the standard. They should not drive others to follow any specific proprietary path or effectively foreclose any software development model. A standard that in effect blocks open source developers from its implementation is not an open standard.

Any conditions (RAND or otherwise) that have the effect of limiting competition, leaving control in the hands of a single vendor, or hindering interoperability – for example, proprietary extensions of a standard– are incompatible with open standards. Ultimately, open standards must allow for self-directed innovation.

Michael R K Mudd, Director of Public Policy - Asia-Pacific, CompTIA (www.comptia.org)

The greatest fear of patents in the software standardization process is the misconception that they would somehow hinder the interoperability of different types of computer programs and file formats. CompTIA would like to make it absolutely clear that robust intellectual property protection – particularly the patentability of computer-implemented inventions – is not a deterrent to interoperability, but may in fact enhance it by providing industry with the right incentives to further invest in the development of such standards in that they can be traded between innovators on RAND terms.

Unlike trade secrets, patents also encourage transparency – in cases where software is distributed in object code form only, a patent provides the only public disclosure of the technical concepts embedded in it. As in the case of the EU Community Patent, which would create a single patent across the European Union, the European Commission has gone to great effort over recent years to harmonize computer-implemented inventions (CII's) despite many public misconceptions of what this would entail. The European Parliament has been cautious to proceed with any directive that would dilute the importance of this intellectual property right. EU Member States and a number of third countries subscribe to the European Patent Convention (EPC) which has enabled the European Patent Office (EPO) to grant patent rights for software-related inventions under intergovernmental – as opposed to supranational – agreement, for decades.

CompTIA also supports the industry position that open standards cannot be defined to the exclusion of potential patent-based royalties such as RAND. To adopt the proposed European Interoperability Framework (EIF) definition would create a preference for a particular business model and automatically exclude other models in the standards-setting process. To cite the EIF definition as a valid EU law or norm is not necessarily correct in that it would discriminate against particular forms of software development in public procurement and other markets

There are some who believe that open standards are standards that either completely "patent-free" or "royalty-free," or that are not subject to any licensing restrictions. This view is squarely at odds with the longstanding approach taken in the IPR policies of leading open standards organizations. For example, ANSI, ECMA, ETSI, IEEE, IETF, ISO/IEC, and ITU expressly acknowledge the right of patent holders to charge reasonable royalties, and to place reasonable restrictions – such as field-of-use restrictions, reciprocity requirements, and restrictions on sub-licensing – on the licensing of their essential technology covering an open standard.

Dan Ravicher, Legal Director, Software Freedom Law Centre, USA (www.softwarefreedom.org)

RAND, to the extent that it has the historical meaning of allowing an owner of a patent that covers a standard to charge a license fee to any person that adopts the standard, is an oxymoron for free and open source software, because any requirement of the charging of a license fee is – by definition –

neither reasonable nor non-discriminatory. Thus, most RAND licensing is – in fact – unreasonable to and severely discriminatory of free and open source software because it precludes such software from existing. As such, standards covered by patents licensed on a RAND basis generally cannot be implemented into free and open source software, unless the RAND license is actually royalty free.

Andre Rebentisch and Gavin Hill, Foundation for a Free Information Infrastructure, European Union (www.ffii.org)

In our opinion the basic foundation of Open Standards is the absence of control over the creation and use of standards. The openness and management of the standard development process plays a minor but crucial role. The notion of facultative use, i.e. occurring optionally in response to circumstances rather than naturally, of Open Standards and the absence of control procedures like forced certification further add to business affinity. A standardization process is always under the threat of abuse by the interests of those parties who set the standard, institutional interests of the standard-setting body and exclusive enforcement of the standard. Often standards put an artificially high cost burden on businesses. This is especially important as standards are classical means of non-tariff trade barriers and anti-competitive market control. Open Standards, correctly defined and implemented, prevent these inherent problems of standard policy. Patents as exclusive monopoly rights over such possible means of undesirable control. RAND licensing limits the exclusive nature of the patent rights. However, only royalty-free or patent-free technology does not bear a welfare burden to the free competitive market. Open Standards may not include measures to exclude competitors. Therefore, RAND licensed technology is incompatible with Open Standards. As mentioned by some contributing parties RAND provisions exclude the use of the currently predominant FOSS license, the GNU GPL.

Despite common postulates there is no indication that the patent system is appropriate to the software market, its market dynamics and common use in professional software development. As the arguments to software patenting have been driven by institutional interests of the patent system rather than market demand, we further stress that there is no legitimacy for the application of the instrument in the software market. We do not see a useful role for patents in standardization. The burden of proof lies upon the proponents of a certain incentive system which works by restricting the freedom of market in first place. The pro arguments mentioned by the FOSS community are not useful to provide such evidence as they are not accompanied by empirical data to prove the claims made. In fact they are postulates which could be applied to justify the use of the patent system in any possible field but fail to make the case. Open Standards and standards in general are especially interesting from our point of view as they are known to be areas where patents are very harmful to competition.

Richard Stallman, Founder, Free Software Movement (www.fsf.org)

Standards bodies that promulgate patent-restricted standards that prohibit free software typically have a policy of obtaining patent licenses that require a fixed fee per copy of a conforming program. They often refer to such licenses by the term "RAND," which stands for "reasonable and non-discriminatory."

That term white-washes a class of patent licenses that are normally neither reasonable nor non-discriminatory. It is true that these licenses do not discriminate against any specific person, but they do discriminate against the free software community, and that makes them unreasonable. Thus, half of "RAND" is deceptive and the other half is prejudiced.

Standards bodies should recognize that these licenses are discriminatory, and drop the use of the term "reasonable and non-discriminatory" or "RAND" to describe them. Until they do so, other writers who do not wish to join in the white-washing would do well to reject that term. To accept and use it merely because patent-wielding companies have made it widespread is to let those companies dictate the views you express.

We suggest the term "uniform fee only," or "UFO" for short, as a replacement. It is accurate because the only condition in these licenses is a uniform royalty fee.

REFERENCES

1. Merriam-Webster Online Dictionary
http://www.m-w.com

2. Wiktionary - a wiki-based Open Content dictionary
http://en.wiktionary.org/wiki/Main_Page

3. Webopedia - the online computer dictionary for computer and Internet technology definitions
http://www.webopedia.com/TERM/s/standard.html

4. Overview of the ISO system - Why standards matter
http://www.iso.org/iso/en/aboutiso/introduction/index.html

5. Perens, B., "Open Standards Principles and Practice"
http://perens.com/OpenStandards/Definition.html

6. Open Standards Policy, IT Division, Massachusetts, USA
http://www.state.ma.us/itd/openstandards.htm

7. Euporean Interoperability Framework v1.0
http://europa.eu.int/idabc/en/document/3473/5585#finalEIF

8. ANSI, "Current Attempts to Change Established Definition of Open Standards"
http://public.ansi.org/ansionline/Documents/Standards%20Activities/Critical%20Issues%20Papers/Open-Stds.pdf

9. BSA, "BSA Statement on Technology Standards"
http://www.bsa.org/usa/policy/loader.cfm?url=/commonspot/security/getfile.cfm&pageid=22407&hitboxdone=yes

10. ITU-T, TSB Director's Ad Hoc IPR Group, "Definition of Open Standards"
http://www.itu.int/ITU-T/othergroups/ipr-adhoc/openstandards.html

11. Free Software Foundation, "Free Software"
http://www.fsf.org/philosophy/free-sw.html

12. Wong, K., "FOSS General Introduction Primer"
http://www.iosn.net/foss/foss-general-primer/

13. OASIS Standards, "OpenDocument Format for Office Applications (OpenDocument) v1.0"
http://www.oasis-open.org/specs/index.php#opendocumentv1.0

14. The Unicode Standard
http://www.unicode.org/standard/standard.html

15. CNET News, "Microsoft FAT patent falls flat"
http://news.com.com/Microsoft+FAT+patent+falls+flat/2100-1014_3-5390138.html

16. Krechmer, K., "The Meaning of Open Standards",
http://www.csrstds.com/openstds.html

17. The International Organization for Standardization (ISO)
http://www.iso.org

18. The International Electrotechnical Commission (IEC)
http://www.iec.ch

19. The International Telecommunication Union (ITU)
http://www.itu.int/
20. The Internet Engineering Task Force (IETF)
http://www.ietf.org

21. IETF, "Overview of the IETF"
http://www.ietf.org/overview.html

22. The Internet Society (ISOC)
http://www.isoc.org

23. RFC 2026, "The Internet Standards Process, Revision 3"
http://www.ietf.org/rfc/rfc2026.txt

24. Official Internet Protocol Standards
http://www.rfc-editor.org/rfcxx00.html

25. The World Wide Web Consortium (W3C)
http://www.w3c.org

26. Organization for the Advancement of Structured Information Standards (OASIS)
http://www.oasis-open.org

27. The Free Standards Group
http://www.freestandards.org

28. IEEE Standards Association
http://standards.ieee.org

29. RFC 793, "Transmission Control Protocol"
http://www.ietf.org/rfc/rfc793.txt

30. RFC 791, "Internet Protocol"
http://www.ietf.org/rfc/rfc791.txt

31. RFC 2616, "Hypertext Transfer Protocol — HTTP/1.1"
http://www.ietf.org/rfc/rfc2616.txt

32. HTML 4.01 Specification
http://www.w3.org/TR/html401/

33. ISO/IEC 15445:2000
http://purl.org/NET/ISO+IEC.15445/15445.html

34. RFC 2821, "Simple Mail Transfer Protocol"
http://www.ietf.org/rfc/rfc2821.txt

35. RFC 2822, "Internet Message Format"
http://www.ietf.org/rfc/rfc2822.txt

36. RFC 1652, "SMTP Service Extension for 8bit-MIMEtransport"
http://www.ietf.org/rfc/rfc1652.txt

37. RFC 2045, 2046, 2047, 2048, 2049, Multipurpose Internet Mail Extensions (MIME) Parts One - Five
http://www.ietf.org/rfc/rfc2045.txt
http://www.ietf.org/rfc/rfc2046.txt
http://www.ietf.org/rfc/rfc2047.txt
http://www.ietf.org/rfc/rfc2048.txt
http://www.ietf.org/rfc/rfc2049.txt

38. Extensible Markup Language (XML) 1.0 (Third Edition)
http://www.w3.org/TR/REC-xml

39. Flynn, P (Ed.), The XML FAQ v4.1, Cork, May 2005
http://xml.silmaril.ie

40. W3Schools, "Introduction to XML"
http://www.w3schools.com/xml/xml_whatis.asp

41. W3C, Extensible Markup Language (XML)
http://www.w3.org/XML/

42. XML.ORG
http://www.xml.org

43. Graphics Interchange Format Version 89a
http://www.w3.org/Graphics/GIF/spec-gif89a.txt

44. Portable Network Graphics (PNG) Recommendation
http://www.w3.org/TR/PNG/

45. The X.Org Foundation

http://www.x.org

46. The Xiph.Org Foundation
http://www.xiph.org

47. Ogg Vorbis General FAQ
http://www.vorbis.com/faq.psp

48. Vorbis Wiki at Xiph.org
http://wiki.xiph.org/index.php/Vorbis

49. Office 2003 XML Reference Schemas
http://www.microsoft.com/Office/xml/default.mspx

50. CNET News, 1 June 2005, "Microsoft adding XML files to Office 12"
http://news.com.com/Microsoft+adding+XML+files+to+Office+12/2100-7344_3-5728536.html?tag=st.ref.goo

51. The OpenOffice.org Project
http://www.openoffice.org

52. Adobe Inc., "What is Adobe PDF?"
http://www.adobe.com/products/acrobat/adobepdf.html

53. Adobe Inc., "PDF Specificatons"
http://partners.adobe.com/public/developer/pdf/index_reference.html

54. Wikipedia (the free-content encyclopedia) entry on "Portable Document Format"
http://en.wikipedia.org/wiki/Pdf

55. Wikipedia (the free-content encyclopedia) entry on "Raster graphics"
http://en.wikipedia.org/wiki/Raster_graphics

56. The XPM Format and Library
http://koala.ilog.fr/lehors/xpm.html

57. Adobe Inc., "TIFF Specifications"
http://partners.adobe.com/public/developer/tiff/index.html

58. The JPEG Homepage
http://www.jpeg.org/jpeg/index.html

59. JPEG JFIF
http://www.w3.org/Graphics/JPEG/

60. JPEG image compression FAQ, part 1
http://www.faqs.org/faqs/jpeg-faq/part1/

61. Wikipedia (the free-content encyclopedia) entry on "Vector graphics"
http://en.wikipedia.org/wiki/Vector_graphics

62. Scalable Vector Graphics (SVG)
http://www.w3.org/Graphics/SVG/

63. Wikipedia (the free-content encyclopedia) entry on "WAV"
http://en.wikipedia.org/wiki/WAV

64. The FLAC Project Page
http://flac.sourceforge.net

65. Microsoft Developer Network, "AVI RIFF File Reference"
http://msdn.microsoft.com/archive/default.asp?
url=/archive/en-us/dx81_c/directx_cpp/htm/aviriifffilereference.asp

66. Overview of the MPEG-4 Standard
http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm

67. The Ogg Encapsulation Format Version 0
http://www.faqs.org/rfcs/rfc3533.html

68. Xiph.org Wiki, "Projects/Formats"
http://wiki.xiph.org/index.php/Main_Page

69. Wikipedia (the free-content encyclopedia) entry on "H.264/MPEG-4 AVC"
http://en.wikipedia.org/wiki/H.264

70. Wikipedia (the free-content encyclopedia) entry on "WMV"
http://en.wikipedia.org/wiki/WMV

71. Wikipedia (the free-content encyclopedia) entry on "Internationalization and localization"
http://en.wikipedia.org/wiki/Internationalization_and_localization

72. Wikipedia (the free-content encyclopedia) entry on "Locale"
http://en.wikipedia.org/wiki/Locale

73. The Microsoft Developer Network (MSDN), "Locale Identifiers"
http://msdn.microsoft.com/library/default.asp?url=/library/en-us/intl/nls_8sj7.asp

74. ISO/IEC 10646:2003, "Information technology — Universal Multiple-Octet Coded Character Set (UCS)"
http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=39921&ICS1=35&ICS2=40&ICS3=

75. Kuhn, M., "UTF-8 and Unicode FAQ for Unix/Linux"
http://www.cl.cam.ac.uk/~mgk25/unicode.html

76. RFC 3629, "UTF-8, a transformation format of ISO 10646"
http://www.ietf.org/rfc/rfc3629.txt

77. Wikipedia (the free-content encyclopedia) entry on "Unicode"
http://en.wikipedia.org/wiki/Unicode

78. ISO 639 Frequently Asked Questions (FAQ)
http://www.loc.gov/standards/iso639-2/faq.html

79. RFC 3066, "Tags for the Identification of Languages"
http://www.ietf.org/rfc/rfc3066.txt

80. The Open Internationalization Initiative
http://www.openi18n.org

81. OpenI18N 1.3 Globalization Specification
http://www.openi18n.org/docs/pdf/OpenI18N1.3.pdf

82. The Mozilla Localization Project
http://www.mozilla.org/projects/l10n

83. The GNOME Translation Project
http://developer.gnome.org/projects/gtp

84. KDE Internationalization
http://i18n.kde.org

85. OpenOffice.org L10N and I18N Projects
http://l10n.openoffice.org

86. Windows XP LIP FAQ
http://www.microsoft.com/globaldev/DrIntl/faqs/winxp.mspx

87. Office 2003 Editions Localized Versions
http://www.microsoft.com/office/editions/prodinfo/language/localized.mspx

88. No Software Patents!
http://www.nosoftwarepatents.com

89. Richard Stallman, "Patent Absurdity"
http://www.guardian.co.uk/online/comment/story/0,12449,1510566,00.html

90. Patents for Innovation
http://www.patents4innovation.org

91. World Intellectual Property Organization, "Patenting Software"
http://www.wipo.int/sme/en/documents/software_patents.htm

92. World Intellectual Property Organization, "Business Method and Computer Software Patents"
http://www.wipo.int/sme/en/e_commerce/computer_software.htm

93. Priscilla Caplan, "Patents and Open Standards"
http://www.niso.org/press/whitepapers/Patents_Caplan.pdf

94. ISOTC Portal, "Intellectual Property Rights (IPR)"
http://isotc.iso.org/livelink/livelink.exe/fetch/2000/2122/3146825/4229629/sds_ipr.htm

95. ISO/IEC Directives, Part 1, Procedures for the technical work (Ed.5)
http://isotc.iso.org/livelink/livelink.exe?func=ll&objId=4230455&objAction=browse&sort=subtype

96. RFC 3979, "Intellectual Property Rights in IETF Technology"
http://www.ietf.org/rfc/rfc3979.txt

97. W3C Patent Policy
http://www.w3.org/Consortium/Patent-Policy/

98. Overview and Summary of W3C Patent Policy
http://www.w3.org/2004/02/05-patentsummary.html

99. OASIS Intellectual Property Rights (IPR) Policy
http://www.oasis-open.org/who/intellectualproperty.php

100. OASIS Intellectual Property Rights (IPR) Policy FAQ
http://www.oasis-open.org/who/ipr/ipr_faq.php

101. The Open Source Definition
http://www.opensource.org/docs/definition.php

102. Philosophy of the GNU Project, 'Some Confusing or Loaded Words and Phrases that are Worth Avoiding - "RAND"'
http://www.gnu.org/philosophy/words-to-avoid.html#RAND

103. CNET News, "IBM offers 500 patents for open-source use"
http://news.com.com/IBM+offers+500+patents+for+open-source+use/2100-7344_3-5524680.html

104. Cover Pages, "Open Source Development Labs (OSDL) Announces Patent Commons Project"
http://xml.coverpages.org/ni2005-08-10-a.html

105. The GNU Operating System
http://www.gnu.org

106. DistroWatch
http://distrowatch.com

107. The Linux Standard Base Project
http://www.linuxbase.org

108. LSB, "Getting Started With LSB 3.0"
http://www.linuxbase.org/build/lsb30.html

109. ISO POSIX (2003)
http://www.unix.org/version3/

110. The Open Group
http://www.opengroup.org

111. European Interoperability Framework for pan-European eGovernment services
http://europa.eu.int/idabc/en/document/2319/5644

112. Documentation on the Promotion of Open Document Exchange Format
http://europa.eu.int/idabc/en/document/3439/5585

113. "Schemas and Standards"
http://www.govtalk.gov.uk/interoperability/schemasstandards.asp

114. e-Government Interoperability Framework Version 6.1
http://www.govtalk.gov.uk/schemasstandards/egif_document.asp?docnum=949

115. Danish e-Government Interoperability Framework
http://standarder.oio.dk/English/

116. The Interoperability Framework: Technical Standards: Document and data exchange
http://standarder.oio.dk/English/Tekniske_standarder/Dokument-_og_dataudveksling/index.html

117. Programme for Open Standards and Open Source Software in Government
http://www.ososs.nl/index.jsp?alias=english

118. The Dutch Government Open Standards Catalogue
http://www.ososs.nl/matrix/matrix.jsp?id=10927

119. http://www.andwest.com/blojsom/blog/tatle/agenda/2005/06/27/

Norwegian_Minister_Proprietary_Standards_No_Longer_Acceptable_in_Communication_with_Government.html

120. Commonwealth of Massachusetts - Enterprise Open Standards Policy
http://www.mass.gov/Aitd/docs/policies_standards/openstandards.pdf

121. New Zealand Government Information Systems Policies and Standards
http://www.e-government.govt.nz/docs/is%2Dpolicies%2Dstandards/

122. New Zealand E-government Interoperability Framework (NZ e-GIF)
http://www.e-government.govt.nz/docs/e-gif-v-2-1/index.html

123. Standards, Policies and Guidelines - Malaysian Government Interoperability Framework (MyGIF)
http://www.mampu.gov.my/mampu/bm/program/ICT/ISPlan/ispdoc/Interoperability%20Framework.pdf

124. IDABC, "Promotion of Open Document Exchange Format"
http://europa.eu.int/idabc/en/document/3428/5644

125. IDABC, "TAC approval on conclusions and recommendations on open document formats"
http://europa.eu.int/idabc/en/document/2592/5588

126. David Adams, "Power Plays: The Phenomenon of Vendor Lock-in"
http://www.osnews.com/story.php?news_id=11029&page=2

127. Wikipedia (the free-content encyclopedia) entry on "Embrace, extend and extinguish"
http://en.wikipedia.org/wiki/Embrace,_extend_and_extinguish

128. CristiAn Andrés Fuenzalida Miranda, "E-Government in Chile and the adoption of XML as standard: Experiences, Challenges and Perspectives"
http://www.dcc.uchile.cl/~cfuenzal/recursos/chileXML.pdf

129. Department of Industrial Policy and Promotion, Government of India, "eBiz – a G2B Portal for Govt. of India"
http://www.nisg.org/ebiz_docs/poc/Approach%20Paper%20on%20PoC%20-22%20Apr%2004.doc

GLOSSARY

- **American National Standards Institute (ANSI)**

  The American National Standards Institute is a private, non-profit organization that administers and coordinates the US voluntary standardization and conformity assessment system.

- **Business Software Alliance (BSA)**

  The Business Software Alliance is a trade group representing some of the world's largest computer software and hardware manufacturing companies. The BSA is involved in programmes that promote copyright protection, cyber security, trade and e-commerce.

- **British Standards Institute (BSI)**

  This is the National Standards Body of the United Kingdom.

- **European Interoperability Framework (EIF)**

  The European Interoperability Framework for pan-European e-Government services provides a framework to facilitate the interoperability of the e-Government services of the European Union member states.

- **European Computer Manufacturers Association (ECMA) International**

  ECMA International is an industry association dedicated to the standardization of Information and Communication Technology (ICT) and Consumer Electronics (CE). ECMA, in co-operation with the appropriate National, European and International organizations, develops standards and technical reports to facilitate and standardize the use of ICT and CE.

- **European Information, Communications and Consumer Technology Industry Association (EICTA)**

  The European Information, Communications and Consumer Electronics Technology Industry Associations (EICTA) was formed by a consolidation of two former European federations of the information and telecommunications industries, the European Association of Consumer Electronics Manufacturers and the European Information & Communications Technology Industry Association. EICTA states that it is dedicated to improving the business environment for the European information and communications technology and consumer electronics sector, and to promoting the industry's contribution to economic growth and social progress in the European Union.

- **Free and Open Source Software (FOSS)**

  Free and Open Source Software is a term used to collectively refer to software that conforms to the definitions produced by either the Free Software Foundation or the Open Source Initiative. FOSS is usually released under at least one of the software licenses recognized by these organizations.

- **Free Software Foundation (FSF)**

  The Free Software Foundation is a non-profit organization based in the USA. Its mission is to preserve, protect and promote the freedom to use, study, copy, modify, and redistribute computer software, and to defend the rights of all Free Software users.

- **GNU's Not UNIX (GNU)**

  The GNU Project was launched in 1984 to develop a complete UNIX-like operating system which

is free software – the GNU system. Variants of the GNU operating system, which use the Linux kernel, are now widely used. GNU is a recursive acronym for "GNU's Not UNIX".

- **GNU General Public License (GPL)**

  The GNU General Public License (GPL) is a free software license originally written by Richard Stallman for software under the GNU Project. Under the GPL, the software can be freely re-distributed, source code is made available and modification and re-distribution of the modified software is permitted. The GPL incorporates the concept of "copyleft" in which derivative works of a GPL-licensed software must be licensed under the GPL also. The GPL is the most popular of the FOSS licenses.

- **Interoperable Delivery of European eGovernment services to public Administrations, Businesses and Citizens (IDABC)**

  The Interoperable Delivery of European eGovernment services to public Administrations, Businesses and Citizens (IDABC) is a community programme managed by the European Commission's Enterprise and Industry Directorate General. It uses the opportunities offered by information and communication technologies to encourage and support the delivery of cross-border public sector services to citizens and enterprises in Europe, to improve efficiency and collaboration between European public administrations and to contribute to making Europe an attractive place to live, work and invest.

- **Joint Photographic Experts Group (JPEG)**

  The Joint Photographic Experts Group (JPEG) is the working group of ISO that defined the popular JPEG Imaging Standard and more recently the JPEG 2000 family of Imaging Standards.

- **Motion Pictures Experts Group (MPEG)**

  The Motion Pictures Experts Group (MPEG) is a working group of ISO/IEC charged with the development of video and audio encoding standards. It is responsible for the family of standards used for coding audio-visual information (e.g., movies, video, and music) in a digital compressed format.

- **Open Source Initiative (OSI)**

  The Open Source Initiative (OSI) is a non-profit organization dedicated to managing and promoting the Open Source Definition, specifically through the OSI Certified Open Source Software certification mark and programme. A piece of software is recognized as Open Source software if it is released under a license certified by the OSI.

- **Open Source Development Labs (OSDL)**

  The Open Source Development Labs (OSDL) is a non-profit organization that is dedicated to accelerating the growth and adoption of GNU/Linux in the enterprise. Its membership comprises most of the prominent commercial players in the open-source industry as well as some academic institutions. It provides state-of the-art computing and test facilities in the United States and Japan available to developers around the world.

- **Portable Operating System Interface for UNIX (POSIX)**

  POSIX is an acronym for Portable Operating System Interface for UNIX, a set of IEEE and ISO standards that define an interface between programs and operating systems. Programs that conform to POSIX developed on one system can be ported more easily to other POSIX-compliant operating systems (this includes most variants of UNIX and UNIX-like operating systems).

ABOUT THE AUTHOR

The author has been in the ICT industry for over 16 years and is experienced in networking protocols, internetworking and information security issues. He participates actively in Malaysian ICT organizations, in particular, MNCC (Malaysian National Computer Confederation) and PIKOM (Association of the Computer and Multimedia Industry of Malaysia). He has been involved in numerous MNCC and PIKOM events and initiatives and has represented both bodies in various Government committees, working groups and task forces on subject matters like the Internet, ICT security, IT standards and FOSS, at both the national and international levels.

The author participates actively in local standardization activities being a member of several technical committees and working groups for IT standards in the Malaysia. He currently works as an independent consultant in the areas of FOSS deployment and information systems security.

He is also the co-author of the *Linux Desktop User Guide* published by the International Open Source Network (IOSN), an initiative under the United Nations Development Programme (UNDP).

ACKNOWLEDGEMENTS

The author would like to thank all those who have helped and contributed to the preparation of this primer, including Shahid Akhtar and Phet Sayo of UNDP-APDIP, and Sunil Abraham and Khairil Yusof of IOSN. He would also like to thank the organizations and people who have provided useful feedback and constructive criticisms to the primer; they include BSA, CompTIA, Microsoft, Richard Stallman, Tony Di Cenzo and Raul Zambrano.

APDIP

The Asia-Pacific Development Information Programme (APDIP) is an initiative of the United Nations Development Programme (UNDP) that aims to promote the development and application of information and communication technologies for sustainable human development in the Asia-Pacific region. APDIP aims to meet its goals by focusing on three inter-related core areas: (i) policy development and dialogue; (ii) access; and (iii) content development and knowledge management.

APDIP collaborates with national governments, regional, international and multi-lateral development organizations, UN agencies, educational and research organizations, civil society groups, and the private sector in integrating ICTs in the development process. It does so by employing a dynamic mix of strategies – awareness raising, capacity building, technical assistance and advice, research and development, knowledge sharing and partnership building.

www.apdip.net

IOSN

The International Open Source Network (IOSN) is an initiative of APDIP and supported by the International Development Research Centre of Canada. IOSN is a Centre of Excellence for Free/ Open Source Software (FOSS), Open Content and Open Standards in the Asia-Pacific region. It is a network with a small secretariat based at the UNDP Regional Centre in Bangkok and three centres of excellence – IOSN ASEAN+3, IOSN PIC (Pacific Island Countries), and IOSN South Asia, based in Manila, Suva and Chennai respectively.

IOSN provides policy and technical advice on FOSS to governments, civil society and the private sector. It produces FOSS awareness and training materials and distributes them under open content licenses. It also organizes awareness raising, training, research and networking initiatives to assist countries in developing a pool of human resources skilled in the use and development of FOSS. IOSN works primarily through its web portal www.iosn.net that is collectively managed by the FOSS community. The web portal serves as a clearinghouse and a platform for knowledge sharing and collaborations.

www.iosn.net

Also available from UNDP Asia-Pacific Development Information Programme

e-Primers on Free/Open Source Software

- Free/Open Source Software – A General Introduction
- Free/Open Source Software – Education
- Free/Open Source Software – Government Policy
- Free/Open Source Software – Localization

www.iosn.net

e-Primers for the Information Economy, Society and Polity

- The Information Age
- Legal and Regulatory Issues in the Information Economy
- Nets, Webs and the Information Infrastructure
- Information and Communication Technologies for Poverty Alleviation
- Internet Governance
- e-Commerce and e-Business
- e-Government
- ICT in Education
- Genes, Technology and Policy

www.apdip.net/elibrary

International Open Source Network

An initiative of the UNDP Asia-Pacific Development Information Programme and supported by the International Development Research Centre, Canada

UNDP

with support from

IDRC CRDI

IOSN

c/o Asia-Pacific Development Information Programme
UNDP Regional Centre in Bangkok
3rd Floor, United Nations Service Building
Rajdamnern Nok Avenue
Bangkok 10200, Thailand
Tel: +66 2 288 1234; 288 2129

IOSN ASEAN+3

National Telehealth Center
University of the Philippines Manila
Taft Avenue, Manila, Philippines 1000
Tel: +63 2 525 6501

IOSN PIC (Pacific Island Countries)

The University of the South Pacific
Private Bag
Suva, Fiji
Tel: +67 9 323 1000

IOSN South Asia

Centre for Development of Advanced Computing
Block-II, 6/13, Park Avenue, Keshava Perumal Puram
Chennai-600 028, India
Tel: +91 44 2461 0880