

Distr.
GENERAL

CES/AC.71/2004/18 (Summary)
1 March 2004

Original: RUSSIAN

**СТАТИСТИЧЕСКАЯ КОМИССИЯ и
ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ
ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ**

**ЕВРОПЕЙСКАЯ КОМИССИЯ
СТАТИСТИЧЕСКОЕ УПРАВЛЕНИЕ
ЕВРОПЕЙСКИХ СООБЩЕСТВ
(ЕВРОСТАТ)**

КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ

**ОРГАНИЗАЦИЯ ЭКОНОМИЧЕСКОГО
СОТРУДНИЧЕСТВА И РАЗВИТИЯ (ОЭСР)
СТАТИСТИЧЕСКИЙ ДИРЕКТОРАТ**

**Совместное совещание ЕЭК/Евростата/ОЭСР по вопросам управления статистическими
информационными системами (УСИС)**
(Женева, 17-19 мая 2004 года)

Тема III: Использование ПО с открытым исходным кодом (open source) в статистике

**Open Sources Software и управление информационными технологиями учреждения
статистики**

Вспомогательный документ

Представлен Государственным Комитетом по Статистике Азербайджана¹

Резюме

Введение

1. Движение Open Sources Software это направление коллективной разработки программ, при котором программы распространяются свободно всем желающим вместе с исходными кодами. В отличие от классического закрытого метода разработки, метод Open Sources – это такой способ, при котором:

- Программное обеспечение создается без единого подробно разработанного и утвержденного проекта;
- Программное обеспечение создается виртуальным коллективом свободных программистов, открытым для любого желающего;
- Документы разработки и тексты программ полностью открыты всем вообще, не только участникам разработки.

¹ Автор > Файг Жалиов (ssc@azstat.org)

2. Целью данного документа не является анализ движения Open Sources, его психологических, моральных, социальных и коммерческих аспектов. Об этом написано достаточно много, в частности такими лидерами движения, как Richard Stallman, Eric Raymond, Tim O'Reilly и другими.

3. Ограничимся лишь вопросами использования Open Sources Software в статистических учреждениях и перспективами применения самого метода для разработки прикладных программ обработки статистической информации.

Опыт использования Open Sources Software

4. Во первых, это использование системных продуктов для построения полноценной корпоративной сети Госкомстата. К ним относятся:

- **Операционная система Linux.** С 2001 года мы полностью перешли от ОС Novell и Windows NT на Linux. В настоящее время все сервера нашей корпоративной сети работают под управлением этой системы;
- **Компоненты Интернет-Инtranет.** Это служба DNS домена на базе BIND, файловый сервер Samba, Web-сервер Apache интегрированный с Perl, PHP и MySQL (Наш сайт www.azstat.org распространяется с этого сервера), почтовый сервер Sendmail, Proxy-сервер Squid, шлюз и файрвол iptable и другие.

5. Во вторых – это использование свободно распространяемых языков программирования и MySQL для разработки прикладных клиент-серверных статистических задач. В этих задачах используется также большое количество свободно распространяемых программ и модулей. Конечно, пока бессмысленно полностью отказываться от коммерческих продуктов. Для разработки приложений мы в основном используем средства и системы разработки программ компаний Microsoft, Borland, Oracle и других.

Перспективы разработки прикладных программ для статистики методом Open Sources

6. Можно ли разрабатывать задачи обработки статистической информации методом Open Sources? Если да, то какой класс задач и каким образом организовать работу? Обычно персонал программистов учреждения статистики невелик и не способен охватить разработку всех задач учреждения. Приходится либо увеличивать состав, либо заказывать часть задач сторонним фирмам. То и другое требует значительных затрат и имеет существенные недостатки. Открытый метод разработки задач путем привлечения виртуального коллектива единомышленников могло бы частично, а может и полностью решить проблему. Рассмотрим некоторые сопутствующие вопросы:

- **Класс задач.** Задачи, предлагаемые статистическим учреждением для открытой разработки должны быть достаточно универсальны и предлагать оригинальные и интересные решения. Это необходимо для привлечения большого коллектива свободных программистов для ее реализации. Следует иметь в виду, что программисты подключаются к проекту вовсе не из альтруистических соображений. Это потенциальные пользователи, которые также как и мы заинтересованы конечным результатом. Нельзя рассчитывать на привлечение большого виртуального коллектива, если предлагаемая задача имеет очень узкую отраслевую сферу применения.
- **Управление проектом.** Проект может быть объявлен, на рабочем сайте. Но нельзя начинать его с пустого предложения. Важно с самого начала предложить разработанную программу с исходными текстами. Она может быть на начальной стадии, содержать ошибки и плохо работать. Это не важно. Главное идея. Проект должен управляться высококвалифицированным специалистом учреждения, который имеет качества общения с коллективом.

Разработка прикладных программ для статистики корпоративным методом

7. Это метод совместной разработки прикладных задач несколькими организациями, заинтересованными в ее результатах. Такая работа может быть организована путем договоренности между этими организациями. Сами задачи могут разрабатываться как классическим методом, так и

методом открытых текстов. Таким методом могут разрабатываться задачи, которые не являются достаточно универсальными и имеют узкую отраслевую сферу применения. Поэтому они не могут рассчитывать на привлечение большого виртуального коллектива свободных программистов. Однако они интересны для нескольких статистических учреждений, которые и могут заняться реализацией совместно.

8. Ввиду ограничений объема документа, мы здесь лишь отметили определенные вопросы и направления. Более полный анализ проблемы мы надеемся дать вскоре в полной версии документа.

9. Open Sources - это направление коллективной разработки программ, при котором программы распространяются всем желающим вместе с исходными кодами.

Движение Open Sources возникло не на пустом месте. Еще на ранних стадиях развития Unix и Internet программисты обменивались своими исходными текстами. Когда фирме AT&T было запрещено вести коммерческую деятельность в области компьютерных технологий Ken Thompson и Dennis Ritchie стали рассылать всем желающим исходный код UNIX. Благодаря этому тысячи программистов получали возможность свободно развивать эту ОС, добавляя в нее свои идеи и программы. Затем создание Unix перешло на коммерческую основу, что не лучшим образом отразилось на ее дальнейшей судьбе.

Примерно с конца 70-х годов в среде свободных разработчиков и пользователей Unix и некоторых других программ начала спонтанно развиваться сеть обмена сообщениями, файлами и новостями, превратившаяся в дальнейшем в Internet.

10. В настоящее время движение Open Sources Software получило широкий размах. Оно означает такой способ разработки, при котором:

- Программное обеспечение создается без единого подробно разработанного и утвержденного проекта;
- Программное обеспечение создается виртуальным коллективом, открытым для любого желающего;
- Документы разработки и тексты программ полностью открыты всем вообще, не только участникам.

11. Linux. Поставки Linux базируются на ядре, написанном Линусом Торвальдсом (Linus Torvalds), и обычно содержат сотни других пакетов открытых текстов, причем большинство последних работоспособно также на других платформах, включая всевозможные диалекты Unix, Windows, Windows NT, MacOS и др.

- FreeBSD, OpenBSD и другие производные от Берклевского Unix. Основное внимание сейчас привлечено к Linux, однако развитие BSD Unix также имеет важные продолжения.
- Средства программирования. В рамках проекта GNU, развиваемого фондом Free Software, создан набор высококачественных средств программирования, в том числе компилятор gcc с языка C, компиляторы g++ и C++, редактор emacs и отладчик gdb. Кроме этого вне рамок проекта GNU созданы два других абсолютно неопенимых элемента культуры открытых текстов, а именно программа patch Лэри Уолла (Larry Wall), позволяющая разработчикам обмениваться небольшими исправлениями в программах вместо того, чтобы пересылать друг другу полные тексты, и система CVS работы с разными версиями исходных текстов.
- Языки. Бесспорное первенство среди языковых средств, разработанных по принципу открытых текстов, принадлежит языку Perl Лэри Уолла, однако языки Tcl Джона Аустерхаута (John Ousterhout) и Python Гидо ван Россума (Guido van Rossum) также имеют большое число поклонников.
- Apache. Как отмечалось, Группа Apache доминирует на рынке серверов Web. Согласно самому последнему обзору фирмы Netcraft (www.netcraft.co.uk/survey), 54% всех наблюдаемых в Сети Web-серверов функционирует с помощью Apache; за ними идут 23,5% серверов, разработанных с помощью средств Microsoft, и за ними Netscape с 6,5%³. Apache продолжает разрабатываться и поддерживаться группой из примерно 12 основных разработчиков, плюс широким и активным сообществом пользователей.

- Samba. Этот продукт был разработан международной группой, возглавляемой Эндрю Триджелом (Andrew Tridgell) из Австралии. Он позволяет системам Unix и Linux служить в качестве файл-сервера и сервера печати в сетях Windows 95/98/NT. Он использует технологию «невидимки», позволяя администраторам встраивать в сеть Linux без ведома управления.
- Sendmail. Эта система первоначально была разработана как часть BSD Unix, а сейчас она является собой «становой хребет» инфраструктуры почтовой системы в Internet.
