

# СВОБОДНОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ЗДРАВООХРАНЕНИЯ

БЛЮМ В.С., ЗАБОЛОТСКИЙ В.П.

---

УДК 502.084

*Блюм В. С., Заболотский В. П. Свободное программное обеспечение для здравоохранения*

**Аннотация.** Предложен подход к формированию полной и достоверной первичной медицинской информации. Выполнен анализ годового объема данных о здоровье нации. Указано на отсутствие иных источников информации для вычисления врачебных ошибок. В качестве метода доступа к системе медицинских документов предлагается специализированная информационно-поисковая система, реализованная на принципах иммунокомпьютинга. Система полной, достоверной и доступной первичной медицинской информации, а также соответствующее программное обеспечение рассматриваются как основа единого информационного пространства здравоохранения. Принципиальным свойством данного программного обеспечения является использование открытого исходного кода и свободной лицензии.

**Ключевые слова:** информатизация здравоохранения, иммунокомпьютинг.

UDC 502.084

*Blum V. S., Zabolotski V.P. Open Source Software for Health*

**Abstract.** An approach to the formation of complete and reliable primary health information. The analysis of annual data on the nation's ill health. Pointed to the lack of other sources of information for the calculation of medical errors. As a method of access to primary health care documents are invited to specialized information retrieval system implemented on immunocomputing. System is full, reliable and affordable basic health information, as well as related software are considered as the basis of a common information space of public health. A fundamental feature of this software is the use of open source and free license.

**Keywords:** Public Health, immunocomputing.

---

**1. Введение.** Системная сложность объекта здравоохранения – человека, обуславливает особенное понимание и обработку информационных явлений. При этом полезно помнить некоторые числовые характеристики указанной системной сложности.

Человеческий организм — это согласованная работа  $10^{14}$  живых клеток. Причем число эффекторов (нейронов и лимфоцитов) в двух интеллектуальных подсистемах организма (нервной и иммунной) имеет один порядок — около  $10^{12}$  клеток. У взрослого человека общая длина нервных волокон — информационных магистралей нервной системы, расходящихся по всему телу, составляет 75 км. Электрические сигналы проходят по ним со скоростью, превышающей 400 км/ч. А общая длина кровеносных сосудов — информационных каналов

иммунной системы, превышает 95 000 километров. Общая емкость капиллярного русла взрослого человека составляет 25-30 л, при этом объем циркулирующей крови составляет только около пяти литров. Выталкиваемая из сердца в аорту кровь в виде волны распространяется по артериям со скоростью примерно 11 м/с. Площадь непрерывной поверхности вторжений организма, границ, которые отделяют организм от внешней среды, огромна. Она включает в себя кожный покров, поверхности органов дыхания и пищеварения. Например, площадь слизистой оболочки тонкой кишки при длине 2,8–3,2 м составляет 180–200м<sup>2</sup>. Все эти километры информационных каналов и миллиарды живых клеток упакованы в минимальный объем и оснащены инструментарием гомеостаза.

Задача системы охраны здоровья состоит в оказании эффективной медицинской помощи гражданину России, включая организацию медицинской профилактики и фармацевтические услуги. При этом главной целью системы является сохранение человеческой жизни.

Однако сложность объекта вспомоществования и ряд иных факторов неминуемо приводят к возникновению тщательно скрываемых врачебных и диагностических ошибок. Эти ошибки уносят жизни сотен тысяч пациентов, что в несколько раз превышает потери от дорожно-транспортных происшествий. По официальным данным в США в год от врачебных ошибок погибает более 100 тыс. человек (исследования корпорации HealthGrade), а экономический ущерб составляет десятки миллиардов долларов. Только затраты, связанные с лекарственными осложнениями, составляют в США около 76,6 млрд долл в год [1].

По мнению академика А. Чучалина: "Процент врачебных ошибок в России очень высок, более 30 %. Так, например, из 1,5 млн заболеваний пневмонией диагностируется не более 500 тысяч. Причин тому несколько, в том числе и полное отсутствие в стране системы контроля качества оказания врачебной помощи".

Определяя первоочередные задачи информатизации здравоохранения — процесса создания, развития и массового применения информационных средств и технологий, обеспечивающего достижение и поддержание уровня информированности субъектов здравоохранения, необходимого и достаточного для кардинального улучшения охраны здоровья каждого гражданина России [2], во главу угла следует положить создание информационных предпосылок для контроля качества оказания врачебной помощи и информационной поддержки врачебных решений.

**2. Анализ источников информации о здоровье нации.** Информация о здоровье (точнее, о нездоровье) пациентов признается государством в том случае, когда она имеет статус информационной услуги, которую вправе оказать только две категории объектов системы охраны здоровья — это множество действующих дипломированных врачей и множество лицензированных диагностических лабораторий. Именно эти и только эти звенья системы охраны здоровья допущены государством к телу пациента по поводу его (тела) здоровья. Только они имеют право и обязаны генерировать информацию (писать тексты, строить графики, фиксировать изображения) о состоянии здоровья граждан своей страны, а также предлагать и фиксировать схемы их лечения.

Предположим, что каждый из работающих по специальности (дипломированных, лицензированных, сертифицированных, аттестованных) врачей имеет свой уникальный идентификатор. А полный перечень идентификаторов представлен в государственном реестре врачей. Этот реестр мы будем использовать для проверки правомочности предоставления информационной медицинской услуги пациенту. Такую же контрольную функцию будет выполнять и государственный реестр работающих (лицензированных) диагностических медицинских лабораторий, в котором зафиксирован полный перечень соответствующих уникальных идентификаторов.

Пусть  $n$  — число идентификаторов в реестре врачей, а  $m$  — число идентификаторов в реестре диагностических лабораторий (ДЛ). Построим неубывающую функцию  $I_{nm}(t)$ , значениями которой будет количество информации, накопленное указанными источниками системы здравоохранения с начала отсчета.

Каждая встреча пациента с врачом или контакт пациента с ДЛ приводит к возникновению информации (текста, графика, изображения), синтаксис которой имеет следующий вид:

$$\begin{aligned} \langle \text{информационная услуга ВРАЧА} \rangle ::= & \\ & \langle \text{время сеанса} \rangle \\ & \langle \text{идентификатор ВРАЧА} \rangle \\ & \langle \text{идентификатор ПАЦИЕНТА} \rangle \\ & \langle \text{документ ВРАЧА} \rangle \end{aligned} \quad (1)$$

или

$$\begin{aligned} \langle \text{информационная услуга ДЛ} \rangle ::= & \\ & \langle \text{время сеанса} \rangle \\ & \langle \text{идентификатор ДЛ} \rangle \\ & \langle \text{идентификатор ПАЦИЕНТА} \rangle \end{aligned}$$

Пусть результат каждой информационной медицинской услуги записывается на цифровой компакт-диск (CD-диск). Выберем такой тип диска, объем которого достаточен для записи самого большого из возможных медицинских информационных массивов.

В результате, каждый акт взаимодействия пациента с врачом или ДЛ будет связан с возникновением CD-диска, на котором записана информация о состоянии здоровья пациента, выявленная и зафиксированная врачом или ДЛ в конкретном сеансе обслуживания.

Положим, что материализованная таким образом информационная медицинская услуга без промедления складывается в определенном месте в две стопки CD-дисков: CD-диски врачей и CD-диски диагностических лабораторий (рис 1).



Рис. 1. Стопки CD-дисков первичной медицинской информации

Попытаемся оценить размеры «информационных» столбов и объем сохраненной в них в течение года информации.

Согласно официальной статистике, в России в течение года регистрируется около 200 млн заболеваний [3].

Если принять, что в процессе обслуживания каждого заболевания врачи пишут только две страницы текста о состоянии больного (результаты общего осмотра, анамнез, диагноз, назначения) и десяток параметров измерений (температура, давление и т. п.), что в сумме составит около пяти кбайт новой информации, то за год минимальный объем новых данных о состоянии здоровья нации составит около 1000 Гбайт. При этом, если предположить, что для обслуживания одного заболевания пациент только два раза встречается с врачом (чтобы от-

крыть и закрыть больничный лист), то число CD-дисков врачей составит 400 млн штук — столб размером 480 км (высота CD — 1,2 мм).

С другой стороны, в России работает около 700 тыс. врачей [4]. Не имея достоверной информации о количестве диагностических лабораторий (ДЛ), примем, что их число соответствует числу лечебно-профилактических учреждений (ЛПУ), т. е. около 20 тыс.

Положим, что в течение рабочего дня врач обслуживает 10 пациентов, а ДЛ обслуживает 20 пациентов. Пусть в результате каждой встречи с пациентом врач пишет одну страницу текста объемом два кбайт, а ДЛ генерирует массив данных размером 50 кбайт. Тогда в течение года врачами будет написан текст объемом 3500 Гбайт (1750 млн штук CD — столб размером 2100 км), а ДЛ сгенерируют файл объемом 5000 Гбайт (100 млн. штук CD — столб размером 120 км).

Итогом наших умозрительных построений стало создание двух информационных столпов здравоохранения — полного набора данных о здоровье нации. Этот набор соответствует современному уровню развития системы охраны здоровья и является единственным безальтернативным источником информации, по которой может быть организован автоматизированный мониторинг врачебных и диагностических ошибок. Благодаря выбранному способу нормировки (размещение данных об одной медицинской услуге на CD-диске стандартного размера) мы можем не только посчитать число оказанных медицинских услуг, но и измерить объем накопленной информации как в традиционных битах, а также в условных «метрах».

В результате мысленного эксперимента собраны две коллекции ранжированных по времени документов вида (1) и (2), сумма которых, по нашему мнению, обладает свойством полноты. Можно утверждать, что *не существует иной первичной медицинской информации о состоянии здоровья нации*, кроме собранной в указанных двух коллекциях документов, суммарный объем которой определяется значением функции  $I_{nm}(t)$  на заданном интервале времени.

Российская система обучения и специализации врачей производит специалистов по 140 специальностям [5]. Несмотря на отмеченное разнообразие, общим свойством врачебной деятельности является то, что каждый врач обязан фиксировать свои наблюдения и решения в форме текстов. Объем писательской работы занимает до 50 % рабочего времени врача.

Разнообразие методов диагностических исследований пациентов не меньше, чем разнообразие врачебных специальностей. Более того,

число методов объективного контроля быстро увеличивается, что обусловлено высокими темпами научно-технического прогресса.

Разнообразие врачебных специализаций, а также методов и результатов лабораторных исследований не может быть препятствием к формированию единого информационного пространства первичных медицинских документов, которое должно быть полным и достоверным. Очевидно, что изъятие любого документа (см. рис.1) может привести к фатальной врачебной ошибке.

На следующем шаге мысленного эксперимента выполним копирование ранее созданных CD-дисков врачей и ДЛ. Из этих копий создадим еще один многокилометровый ранжированный по времени столб застывшей информации, который станет собственностью пациентов — CD-дисков пациентов (рис. 2).

Мотивом для двукратного увеличения объема хранимой информации является наше стремление исключить всякую фальсификацию, которая может стоить жизни пациенту. Достоверными будем считать только те CD пациентов, у которых существует точная оригинальная копия.

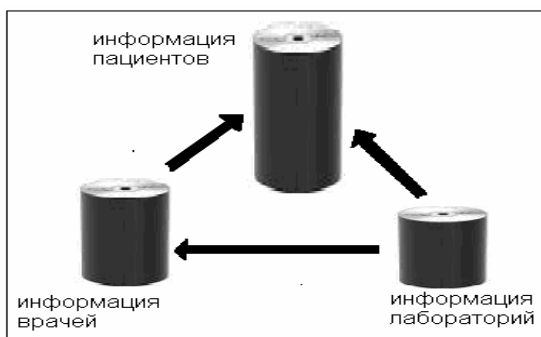


Рис. 2. Система полной и достоверной первичной медицинской информации

Пусть  $k$  — число идентификаторов различных пациентов, обслуженных на заданном интервале времени. Тогда функция количества достоверной информации:

$$I_{nmk}(t) = 2I_{nm}(t).$$

Для формирования достоверных наборов данных об информационных результатах работы врачей и ДЛ мы использовали якобы существующие реестры уникальных идентификаторов.

После того, как указанные три набора данных сформированы, над ними могут быть определены для любого заданного интервала времени функции  $N(t)$  — число идентификаторов врачей и  $M(t)$  — число идентификаторов ДЛ. Для вновь организованного набора данных пациентов, определим функцию  $K(t)$  — число идентификаторов пациентов. Функция  $K(t)$  вычисляется методом последовательного перебора CD-дисков пациентов для заданного интервала времени. Чтобы не возникало коллизий при расчете функции  $K(t)$ , позаботимся об однозначной идентификации пациентов, полагая, что идентификатор пациента соответствует номеру его полиса ОМС.

Полная и достоверная система медицинских данных является воплощением идеи академика РАМН и РАН А. И. Воробьева о том, что лечащий врач не может быть простым регистратором заключений лабораторий и диагностических служб; ему необходим весь объем первичной медицинской информации — и рентгеновские снимки, и фотографии гистологических препаратов и т. д.

Проведем сортировку каждого из трех столбов CD по соответствующему ключевому параметру:

- ◆ идентификатор врачей — для CD врачей,
- ◆ идентификатор ДЛ — для CD ДЛ,
- ◆ идентификатор пациентов — для CD пациентов.

В результате получим три класса персонифицированных коллекций медицинских документов: коллекции врачей, ДЛ и пациентов (рис. 3).

Отметим, что процедура сортировки не должна нарушать ранжирования документов по времени внутри каждой персональной коллекции.

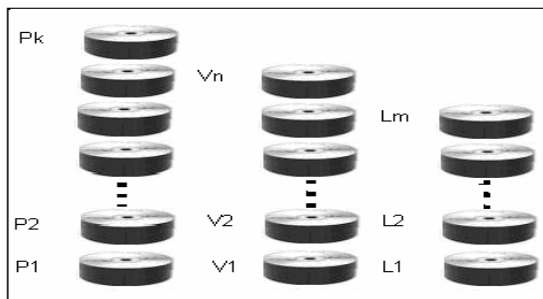


Рис. 3. Три класса персонифицированных коллекций медицинских документов

В коллекции документов, сформированных врачами  $\{V_1, V_2, \dots, V_n\}$ , где  $V_i$  — идентификатор врача (например, номер действующей лицензии) может использоваться в качестве основы адреса сетевого ресурса для доступа к этим данным, а также для формирования гиперссылок к коллекциям обслуженных пациентов.

Коллекция данных, сформированная диагностическими лабораториями  $\{L_1, L_2, \dots, L_m\}$ , имеет аналогичную структуру и смысловую нагрузку.

В коллекции документов, сформированных для пациентов  $\{P_1, P_2, \dots, P_k\}$ , где  $P_i$  — это идентификатор пациента (например, номер его полиса ОМС), который может быть использован для автоматического формирования (вычисления) имени соответствующего сетевого ресурса. Коллекция  $P_i$  — это ни что иное как медицинская карта пациента, представленная, в данном случае, в виде стопки ранжированных по времени CD-дисков, каждый из которых имеет гиперссылку к своей оригинальной коллекции. Гиперссылки строго упорядочены по времени. В этом ряду, очевидно, первым является документ о рождении пациента, а заключительным станет документ, в котором врач фиксирует смерть пациента.

Коллекция документов  $\{V_1, V_2, \dots, V_n\}$  — это собрание историй врачевания каждого допущенного к этой работе доктора. Граф врача  $V_i$  — это тот объект, к которому сегодня можно адресовать слова Н.И.Пирогова: “Усовершенствования во врачебном искусстве можно добиться только путем изучения ошибок, допущенных у постели больного.”

**3. Формирование индекса полной и достоверной медицинской информации.** По оценкам Всемирной организации здравоохранения около 20% врачебных ошибок связано с неполнотой данных или невозможностью оперативного получения необходимой информации.

Преыдушие шаги мысленного эксперимента были посвящены формированию полной и достоверной системы медицинской информации. Теперь мы можем отказаться от измерения информации в «метрах», выделить для каждой коллекции документов  $\{V_1, V_2, \dots, V_n\}$ ,  $\{L_1, L_2, \dots, L_m\}$  и  $\{P_1, P_2, \dots, P_k\}$   $n + m + k$  персональных сетевых ресурсов в распределенной сети и заняться разработкой способа оперативного получения необходимых данных.



Для оперативного доступа к информации необходима специализированная поисковая машина, в основе которой лежит постоянно корректируемый поисковый индекс здравоохранения.

*Поисковый индекс здравоохранения* — это система данных (локальная или распределенная), в которой хранится полный перечень используемых в современном здравоохранении терминов, связь этих терминов с содержащими их сетевыми ресурсами, а также локальные ссылки на документы и их весовые характеристики.

Предлагаемая структура поискового индекса близка к классической, которая описана в литературе [6, 7]. В его основе лежат инверсные списки вхождений слов, которые используются для поиска релевантных документов, а также прямые индексы термов для формирования фрагментов, возвращаемых пользователю.

Индекс формируется специальной программой — сетевым роботом ("пауком"). Паук не копирует страницы сайтов в индекс поисковой машины, а сохраняет информацию о структуре каждого документа сайта — например, какие слова встречаются в документе и в каком порядке, адреса гиперссылок страницы сайта, размер документа в килобайтах, дата его создания и многое другое.

Попутно необходимо решить достаточно много мелких, неприятных, но необходимых, задач "очистки документов": ликвидацию переноса слов и фиксацию таблиц, проверку алфавита и словосочетаний, проверку структуры, форматирование текста, ликвидацию лишних пустых строк и строк разметки и т. п.

Построение предметного индекса на однородной коллекции документов это, по существу, процедура семантического сжатия информации.

На первом этапе выполняется последовательно несколько известных алгоритмов семантического сжатия индекса слов коллекции, полученного после стандартной операции индексирования, а именно:

- ◆ извлечение стоп-слов (наречий, союзов, предлогов и т. д.);
- ◆ учет статистических закономерностей текстов (ранговые распределения как системное свойство текстов);
- ◆ морфологическая коррекция терминов и коррекция по частям речи;
- ◆ определение словосочетаний в полученном наборе терминов.

Структуры данных должны быть выбраны так, чтобы они допускали подключение таких критериев оценки терминов, как позиционное взвешивание и взвешивание по инверсной частоте.

На втором этапе вычисляются соотношения терминов в виде статистических данных, определяются семантические связи между терминами и выполняется усечение словаря терминов. Критерием усечения является малое число связей термина с другими терминами. Результатом второго этапа является минимизированный словарь связанных терминов, по существу представляющий собой предметный индекс коллекции медицинских документов.

Пусть задано множество из  $n$  документов.

На его основе можно построить множество из  $m$  терминов, которые хоть раз встречались в каком-либо документе.

Тогда можно утверждать, что в основе индекса здравоохранения лежит матрица сопряженности «документ–термин» размерностью  $(n \times m)$ , которая имеет вид:

$$C = \begin{array}{c|cccc} & t_1 & t_2 & \cdots & t_m \\ d_1 & c_{11} & c_{12} & \cdots & c_{1m} \\ d_2 & c_{21} & c_{22} & \cdots & c_{2m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ d_n & c_{n1} & c_{n2} & \cdots & c_{nm} \end{array}$$

Элемент  $c_{ij}$  это тройка  $\langle c, T_0, \Delta_t \rangle$ , где  $c$  указывает на наличие  $j$ -го термина в  $i$ -м документе и вычисляется по методу TF-IDF для данной коллекции документов,  $T_0$  – время первого появления термина  $t_j$  в документе  $d_i$ ,  $\Delta_t$  – интервал времени от первого до последнего появления термина  $t_j$  в документе  $d_i$ . Параметры  $T_0$  и  $\Delta_t$  позволяют учитывать специфику медицинской информации при формировании ответа на запрос. Обработка этой матрицы выполняется высокопроизводительными алгоритмами иммунокомпьютинга [8].

Следует отметить, что борьба за оперативный доступ к медицинской информации имеет свою цену — объем базы данных индекса имеет тот же порядок, что и исходный объем хранимых данных. Обозначим функцию количества полной, достоверной и индексированной информации на заданном интервале времени —  $I_{[nmk]}(t)$ .

Тогда

$$I_{[nmk]}(t) \approx 3I_{nm}(t).$$

Таким образом, в результате мысленного эксперимента мы определили полную и достоверную систему первичных медицинских до-

кументов, к которой гарантирован оперативный доступ. Этот результат следует рассматривать как основу и необходимое условие на пути от современной информационной системы здравоохранения к социальной иммунной системе (СИМ), в которой цена компьютерной памяти стремительно падает, а цена человеческой жизни неизменно растет.

**4. Состав базового программного обеспечения охраны здоровья.** Очевидным выводом из проведенного мысленного эксперимента является необходимость в рамках информатизации здравоохранения предоставить каждому врачу вместо авторучки клавиатуру компьютера и сетевой ресурс для хранения данных с тем, чтобы ни одно слово из истории болезни нации не было потеряно для государственной системы контроля и раннего обнаружения врачебных ошибок. Эксперимент также указывает на состав базовых комплексов программного обеспечения, способных формировать и поддерживать единое информационное пространство первичной медицинской документации.

К базовым программным комплексам следует отнести:

1. Пакет программ ведения государственного реестра врачей.
2. Пакет программ ведения государственного реестра диагностических лабораторий.
3. Типовой пакет программ информационной поддержки российского врача.
4. Типовой пакет программ информационной поддержки российской диагностической лаборатории.
5. Типовой пакет программ информационной поддержки российского пациента.
6. Программный комплекс специализированной информационно-поисковой системы (ИПС) здравоохранения.

В задачи пакетов программ ведения государственных реестров входит не только формирование и поддержание актуальных списков действующих врачей и диагностических лабораторий, но и учет соответствующих персонифицированных сетевых ресурсов, список которых необходим для эффективной работы ИПС.

Функции типового пакета программ информационной поддержки российского врача включают в себя:

- ◆ выполнение роли авторучки врача в любом месте, где он выполняет свои обязанности,
- ◆ репликацию накопленных данных в сетевой персональный ресурс врача,
- ◆ транзакцию данных в сетевой персональный ресурс пациента,

- ◆ обеспечение доступа к истории врачевания для “паука” ИПС,
- ◆ а также ряд сервисных функций.

Функции типового пакета программ информационной поддержки российской диагностической лаборатории подобны функциям пакета врача, но учитывают специфику и стандарты представления диагностических данных.

Типовой пакет программ информационной поддержки российского пациента представляет собой Web-сервис, который обеспечивает доступ пациента к его персональным медицинским данным, создает условия для работы “паука” ИПС с историей болезни пациента, а также возможность ведения личного дневника здоровья и ряд дополнительных сервисов.

Ключевым звеном в предлагаемой системе медицинских данных и программ является информационно-поисковая система, которая обеспечивает оперативный доступ к любому фрагменту распределенных первичных документов и, тем самым, создает материальную основу единства информационного пространства здравоохранения. В задачи специализированной ИПС входит не только обеспечение доступа к конкретной истории болезни или истории врачевания, но и предоставление широкого спектра фактических медицинских данных для любого заданного интервала времени.

**5. Почему базовое программное обеспечение здравоохранения должно быть свободным?** Рассмотренная система данных и базовый комплекс программных средств определяют интерфейс взаимодействия населения с системой здравоохранения, который обслуживает поток данных о здоровье граждан государства. В основу реализации такого интерфейса должны быть положены принципы, которые обеспечивают:

- ◆ безопасность и надежность программных средств;
- ◆ высокий уровень управляемости развитием программного обеспечения;
- ◆ масштабируемость и кроссплатформенность для всего разнообразия участников информационного процесса;
- ◆ юридическую защиту прав потребителей программных средств (врачей, диагностических лабораторий и пациентов);
- ◆ гарантию качества программных инструментов;
- ◆ экономию средств на внедрение и сопровождение данной информационной технологии;
- ◆ поддержка и развитие открытых стандартов.

Все эти свойства присущи свободному программному обеспечению (ПО). Свободные программы появились в академической среде – множество утилит было разработано и распространено как научные публикации. Государство постоянно стимулирует подобную работу, получая взамен развитие технологий и качественных специалистов, экономя финансовые средства.

Осознание преимуществ свободного ПО произошло и в отношении проблем информатизации здравоохранения.

На прошедшей 7–8 мая на Филиппинских островах конференции, посвященной электронной системе здравоохранения и телемедицине (*eHealth & Telemedicine Philippines 2009 Conference and Exhibition*), участники обратили особое внимание на важную роль свободного ПО и ПО с открытым кодом (*FOSS*) в развитии электронного медицинского документооборота.

В сенате США рассматривается предложение по созданию национальной медицинской информационной системы на базе открытого ПО. Предложен новый законопроект "Health Information Technology Public Utility Act of 2009", который предусматривает выделение федеральных грантов на разработку open source информационной системы и создание набора открытых стандартов для унифицированного учета медицинских карточек во всех медицинских учреждениях страны.

Учитывая, что затраты на здравоохранение в США превышают два трлн долларов, а в России только — 1 трлн рублей, говорить о равных возможностях не приходится. Сокращение этой дистанции состоит не только в необходимом увеличении объема финансирования этой сферы услуг, но в первую очередь в более интенсивном процессе информатизации отрасли.

Одним из путей интенсификации является использование и разработка свободного ПО медицинских информационных систем, положенных в основу единого информационного пространства здравоохранения. Этот шаг приведет не только к достижению названных преимуществ ПО, но и вовлечет в процесс его создания гораздо больший круг специалистов, как российских, так и зарубежных.

Считая принципиальным использование открытого исходного кода и свободной лицензии для создания предложенной информационной системы первичной медицинской документации, следует отдавать себе отчет в необходимости расширения списка компонентов базового ПО.

В частности, необходимо создание репозитория. Одним из отличий свободной модели разработки от проприетарной (закрытой, ком-

мерческой, собственнической) является наличие репозитория – хранилища файлов для дальнейшего распространения в сети. Репозиторий используется также в системах управления версиями программных пакетов. Кроме того, необходимым является создание системы регистрации ошибок для распределенной системы разработки и эксплуатации программных продуктов.

**6. Заключение.** Социальная иммунная система (СИС), роль которой стремится выполнять современная система здравоохранения, также как ее прообраз — иммунная система человека — страдает собственными недугами. Одной из хронических болезней СИС являются врачебные ошибки. Для лечения этой болезни и минимизация потерь от ее проявлений необходимо формирование единого информационного пространства здравоохранения на основе сбора и обработки полной, достоверной и доступной первичной медицинской информации. Альтернативного информационного источника для вычисления врачебных ошибок просто не существует.

### Литература

1. Johnson J. A., Bootman J. L. Drug-related morbidity and mortality. A cost-of-illness model // *J. Arch. Intern. Med.*, 1995, Oct. 9; 155 (18), 1949-1956.
2. Юсупов Р. М., Заболотский В. П. Научно-методологические основы информатизации. СПб.: Наука, 2000. 455 с.
3. Официальная медицинская статистика. [Электронный ресурс]. — Режим доступа: [http://www.mzrd.ru/?f=np\\_about\\_pro](http://www.mzrd.ru/?f=np_about_pro)
4. Медицинская статистика. [Электронный ресурс]. — Режим доступа: <http://www.medlinks.ru/article.php?sid=30795>
5. Приказ Минздравсоцразвития России от 11.03.2008 г. № 112н «О номенклатуре специальностей специалистов с высшим и послевузовским медицинским и фармацевтическим образованием в сфере здравоохранения Российской Федерации»
6. *G. Salton, C. Buckley.* Term-weighting approaches in automatic text retrieval. // *Information Processing and Management: an International Journal*, Volume 24, Issue 5, pages: 513 – 523, 1988.
7. *S. Brin, L. Page.* The Anatomy of a Large-Scale Hypertextual Web Search Engine. // *Computer Networks and ISDN Systems*, Volume 30, number 1-7, pages 107-117, 1998.
8. Tarakanov A. O., Skormin V. A., Sokolova S. P. *Immunocomputing: Principles and Applications.* New York.: Springer, 2003, 230 p.

**Блюм Владислав Станиславович** — к.т.н., старший научный сотрудник лаборатории прикладной информатики Учреждения Российской академии наук С.-Петербургский институт информатики и автоматизации РАН (СПИИРАН). Область научных интересов: информатизация здравоохранения, иммунокомпьютинг. Число научных публикаций — 45. vlad@blum.spb.su; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-1919, факс +7(812)328-4450.

**Заболотский Вадим Петрович** — д.т.н., профессор; ведущий научный сотрудник лаборатории прикладной информатики Учреждения Российской академии наук С.-Петербургский институт информатики и автоматизации РАН (СПИИРАН). Область научных интересов: математическое моделирование, методология информатизации. Число научных публикаций — 240. lai@iias.spb.su; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-1919, факс +7(812)328-4450.

**Blum Vladislav Stanislavovich** — PhD, associate professor; senior researcher, Laboratory of Applied Informatics, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: Public Health, immunocomputing. The number of publications — 45. lai@iias.spb.su; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-1919, fax +7(812)328-4450.

**Zabolotski Vadim Petrovich** — PhD, professor; senior researcher officer, Laboratory of Applied Informatics, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: mathematical modeling, the methodology of Informatisation. The number of publications — 240. vlad@blum.spb.su; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-1919, fax +7(812)328-4450.

Рекомендовано группой проблем информатизации общества, руководитель группы Заболотский В.П., д. т. н., проф.  
Статья поступила в редакцию 24.06.2009.