

удк 519.68

Я. И.–о. Гулиев, В. Л. Малых  
**Архитектура ИИ-Х**

**Аннотация.** В статье предложен подход к построению архитектуры корпоративных медицинских информационных систем. В основе подхода лежит представление о необходимости поддержки информационной системой эволюционного процесса концептуализации предметной области. Архитектура характеризуется открытостью, независимостью от средств программной реализации, масштабируемостью до уровня отдельных удаленных подразделений, ориентированностью на существующие и разрабатываемые стандарты.

*Ключевые слова и фразы:* МИС Медицинская Информационная Система.

## **1. Введение в проблему построения архитектуры современной МИС**

Основные проблемы, стоящие перед разработчиками медицинских информационных систем (МИС), кратко можно охарактеризовать следующим образом:

- необходимость вводить и анализировать все больший объем медицинской информации для оказания качественной медицинской помощи;
- необходимость и актуальность поддержки единого информационного пространства (полные данные по каждому пациенту независимо от места оказания медицинской помощи);
- постоянно расширяющаяся понятийно-концептуальная база предметной области (медицины);
- недостаточная формализация (концептуализация и стандартизация) предметной области;
- быстрый прогресс информационных технологий.

Медицинская информация становится все более сложной и все более важной. По данным [1], около 15% ресурсов лечебного медицинского учреждения расходуется на сбор данных. Врачи и медсестры затрачивают 25% своего рабочего времени на поиск нужной им информации. В качестве примера можно привести статистику из университетского госпиталя в Гейдельберге (University Hospital of Heidelberg) на 1 700 коек. В течение года в госпитале создается около 400 000 новых

медицинских карт, содержащих в бумажном виде 6,3 млн. страниц. Выполняется 250 000 исследований и 20 000 процедур. По данным US Institute of Medicine Report, около 100 000 граждан США ежегодно умирают по причине медицинских ошибок (восьмое место в упорядоченных по частоте причинах смерти). Одна из основных причин ошибок — недостаточная или некачественная информация о пациенте. Объем медицинской литературы, необходимой для качественного лечения, удваивается каждые 10–15 лет уже в течение последних 300 лет. Ежегодно в области биомедицины выходят 20 000 журналов и 17 000 книг. Примерно половина концептов (500 000) из всех выработанных человечеством непосредственно относится к медицине. Многоязычный тезаурус UMLS (Унифицированная система медицинского языка) содержит 730 000 концептов и 1,5 млн. понятий.

Практически все исследователи в области МИС подчеркивают *разнородность, сложность и эволюционную природу медицинских данных*. Отсюда проблемы стандартизации медицинской информации, проблемы обмена данными между отдельными МИС. До сих пор не преодолены сложности создания единого информационного пространства медицинских данных даже для граждан развитых западных стран. Попытки жесткой шаблонной формализации ввода и представления медицинской информации врачами отвергаются, как искажающие суть самой информации.

С другой стороны, весь мировой опыт свидетельствует о необходимости стандартизации программного обеспечения. Эффективность подобного подхода в промышленности, связанной с разработкой технических изделий массового потребления, уже давно очевидна. И медицина не должна быть исключением из этого правила.

Многие исследования последних десятилетий связаны с поиском концептуальных решений проблемы стандартизации медицинской информации. Большие надежды связываются с повышением уровня абстракции предлагаемых формальных моделей. Прочитируем в качестве примера подобного подхода: “The solution is to use more abstract models with fewer, but more expressive, objects. The patient information model has to be simplified and clarified and a uniform and correct level of abstraction must be found. . . For example, items such as birth, weight or haemoglobin concentration should not be found included as named attributes of the model. The specific clinical entities should be represented

in a concept/vocabulary data base that is separate from the data model”<sup>1</sup> [2].

Во всех развитых странах интенсивно ведутся работы по стандартизации медицинской информации. По-видимому, принятие единого общеевропейского или общемирового стандарта представления медицинской информации — это вопрос времени. Архитекторы современных российских МИС должны учитывать неизбежность появления в будущем общегосударственного стандарта.

## 2. Направления стандартизации в медицинской информатике

«К настоящему моменту практически все существующие стандарты медицинской информатики носят не обязательный, а рекомендательный характер. В США разработка стандартов медицинской информатики координируется соответствующими подкомитетами Американского национального института стандартизации ANSI, в Европе — подкомитетом TC251 Европейского комитета по стандартизации CEN. Особенность стандартизации медицинской информатики ярко выражена следующей дилеммой: чем уже круг экспертов, тем сложнее сделать стандарт общепризнанным; чем он шире, тем дольше разрабатываются стандартные решения» [3].

Практически все стандарты медицинской информатики так или иначе связаны с ведением электронной истории болезни [4]. Одни стандарты описывают терминологию, которая должна быть в ней использована, другие — передачу медицинских документов и изображений в электронную историю болезни, третьи — способы организации данных в электронной истории болезни, четвертые — обеспечение доступа медицинских работников и самих пациентов к электронной истории болезни и т. д.

---

<sup>1</sup>«Решение заключается в использовании более абстрактных моделей с немногими, но высоко выразительными абстрактными объектами. Модель данных о пациенте следует сделать проще и яснее, для чего необходимо найти унифицированный и корректный уровень абстракции. Например, такие элементы данных, как дату рождения, вес или концентрацию гемоглобина не следовало бы включать в модель в качестве именованных атрибутов. Специфические клинические сущности следовало бы представлять на основе словаря, независимого от модели данных». (Перев. авторов.)

В сущности, разработка стандартов медицинской информатики преследует цель воссоздания универсального языка общения медицинских работников, другими словами — воскрешения латыни на самом современном уровне информационных технологий. В целом эти стандарты нужны для того, чтобы каждая запись электронной истории болезни могла быть одинаково понята представителями различных медицинских школ. При этом компьютеры должны стать как бы переводчиками с привычного естественного медицинского языка на унифицированный электронный язык и обратно. Поэтому неудивительно, что в последние десятилетия наибольшие усилия специалистов по медицинской информатике были сосредоточены в двух основных областях: стандартизации медицинской терминологии и стандартизации передачи записей в электронную историю болезни.

«В настоящее время можно выделить два стандарта, ведущих свое происхождение из США, но получивших достаточно широкое признание и в других странах:

- стандарт обмена электронными текстовыми медицинскими документами Health Level Seven (HL7);
- стандарт обмена электронными изображениями лучевой диагностики Digital Imaging and Communication in Medicine (DICOM).

К сожалению, в России работа по стандартизации электронного обмена медицинскими документами в подобном объеме и с таким качеством не ведется» [3].

Перед создателями отечественных МИС встает непростой вопрос: как относиться к разрабатываемым на Западе стандартам представления (стандартам описания медицинских данных и терминологических словарей) и обмена медицинской информацией? Следует ли включать поддержку этих стандартов в архитектуру МИС, или следует подождать появления отечественного стандарта, или пытаться разработать собственный «внутрифирменный» стандарт?

По нашему глубокому убеждению, перейти к использованию западных стандартов можно только в том случае, если само ведение медицинской карты и оказание медицинской помощи будет построено по западному образцу. Для этого придется практически «перестроить» все отечественное здравоохранение, начиная с системы медицинского образования. Что вряд ли приемлемо для русской медицины с ее богатой историей.

Не имея возможности непосредственного заимствования стандартов, мы можем, безусловно, *поддержать саму идею стандартизации как главную стратегическую идею развития медицинской информатики.*

Разработка общероссийских стандартов — это прерогатива прежде всего государства. Разработка потребует немалых ресурсов и вряд ли окажется по силам отдельным производителям медицинских информационных систем и специализированного программного обеспечения.

Единственно верным стратегическим решением в этой ситуации нам представляется разработка архитектуры, *ориентированной на стандарты.* Чтобы ориентироваться на то, «чего еще нет», необходимо сделать некоторые предположения, интуитивно предугадать направление, в котором пойдет стандартизация в медицине. В настоящее время в развитии стандартизации на Западе просматриваются два основных течения. Первое, связанное с историей стандарта HL7, определяет жесткие нерасширяемые модели данных (концептов), пытаясь все многообразие медицинской информации свести к небольшому числу конкретных информационных сообщений. Второе, связанное с европейскими разработками прототипа стандарта GENR (The Good European Health Record project), пытается лишь определить общие правила и ограничения конструирования произвольных информационных сообщений (структур) над словарями понятий предметной области. Если рассмотреть возможный синтез этих двух подходов, то можно предположить, что *стандарт должен будет заключать в себе как жестко формализованные концепты предметной области, так и свободно конструируемые по заданным правилам информационные структуры.*

Если в МИС любые медицинские данные и их структуры будут стандартизованы и описаны на языке концептов, понятий (терминов) и контекстов их использования в базе знаний системы, перевод системы на общегосударственный стандарт будет означать переход от одного метаописания к другому, в противовес внедрению стандарта в систему, не имеющую явного стандартизованного метаописания. Архитектура современной МИС должна быть готова к работе со стандартами.

Оставшаяся часть работы будет посвящена проблеме концептуального моделирования предметной области, ориентированного на

будущую стандартизацию описания медицинских данных. Излагаемый подход получил по аналогии с HL7 условное наименование «архитектура HL-X (Health Level XML)».

### **3. Определение понятийных (терминологических), информационных, функциональных и прочих требований к документу HL-X как к основному носителю всей медицинской информации**

В основе архитектуры HL-X лежит понятие документа HL-X, его свойств и методов работы с ним. Концептуально документ HL-X может быть рассмотрен с различных точек зрения. Соответственно с каждой точкой зрения можно предложить некоторую модель, отражающую различные аспекты документа.

Выделим следующие модели документа:

- понятийная модель;
- информационная модель (структурная модель);
- модель обработки данных документа;
- модель визуализации;
- функциональная модель документа;
- модель безопасности.

*Понятийная модель* представляет документ в виде структуры абстрактных и соответствующих им конкретных понятий. Она предоставляет следующие возможности:

- (1) определить, какие понятия содержатся в документе и какими понятиями «владеет» база данных (БД) или база знаний (БЗ) документов HL-X;
- (2) выполнить перевод понятийной модели на другие языки при наличии словарей;
- (3) является основой для добавления семантических связей на понятиях и использования технологий искусственного интеллекта для анализа свободных документов (тематический поиск информации, поиск «неизвестных» знаний — Data Mining и т. п.);
- (4) фактически самодокументирует документ и может быть использована как умолчательная визуализация документа.

*Информационная модель* представляет документ HL-X в виде структуры из различных *информационных объектов* (концептов предметной области). Здесь под информационными объектами в широком смысле понимаются некоторые абстрактные связанные элементы модели, аналогами которых в различных языках программирования являются объекты, записи, структуры, массивы и т. п. Наличие у документа структуры позволяет декомпозировать его на элементы и выполнять манипуляции с этими элементами при необходимости. Информационная модель служит основой для:

- конструирования документа из элементарных информационных объектов;
- конструирования документа на основе уже имеющихся документов, с учетом уже накопленного знания о контекстах использования того или иного объекта (понятия), путем заимствования, полного или частичного, уже существующих информационных моделей;
- разбора документа и обработки содержащихся в нем данных (например, для структурированного хранения в реляционной БД или для экспорта и импорта данных, для усвоения знания и пополнения БЗ непосредственно из экземпляров документов и их информационных моделей).

*Модели обработки данных* документа позволяют специфицировать обработку документа как единого целого в интересах конкретной информационной системы (ИС). Они определяют, какие именно данные нужно обрабатывать и как это делать. Модель обработки данных документа дополняет информационную модель документа конкретными, зависящими от ИС инструкциями. На основе модели выполняются все основные манипуляции с данными документа:

- разбор и хранение в БД;
- создание и редактирование документов;
- использование шаблонов и других документов в качестве источников данных.

*Модели визуализации* позволяют создавать различные визуальные представления данных документа. В основе визуальной модели лежит понятие визуальной компоненты как некоторой абстрактной элементарной формы представления данных. В различных конструкторах интерфейса аналогами визуальным компонентам служат компоненты визуальных палитр (таблицы, поля, панели, всевозможные

элементы управления и т. п.). Основное назначение модели — создать абстрактное, не привязанное к конкретному программному языку или конструктору интерфейса описание представления данных документа в интерфейсе пользователя. Модель визуализации предназначена для:

- визуализации документа в конкретной реализации пользовательского интерфейса со свободными документами;
- предоставления, совместно с функциональной моделью, интерфейса для манипуляции данными документа.

*Функциональная модель* документа определяет возможные манипуляции с данными документа и соответственно поддерживает определенные ограничения на данных (целостность данных). К области компетентности этой модели отнесем следующее:

- транзакционность — обработка документа и всех входящих в него объектов как единого целого в рамках одной транзакции;
- дисциплина коллективной работы над документом (пользовательские блокировки);
- определение возможных манипуляций над документом на основе данных самого документа (учет статуса и т. п.);
- целостность данных (обязательные элементы данных, домены возможных значений, контроль типов и форматов данных и т. п.).

*Модель безопасности* документа определяет права доступа к элементам данных и права на манипуляцию данными. В качестве структурного элемента документа, на который можно будет установить права доступа, предлагается принять элементарное понятие из понятийной модели. В информационной модели этому элементу доступа соответствует атрибут информационного объекта, или сам объект, а в визуальной модели — отдельное поле визуализации или ввода. Более элементарных целостных единиц данных, к которым можно было бы определить доступ, в документе просто нет. На уровне структуры документа предложена самая сильная модель безопасности.

Необходимо смотреть на документ HL-X как на единство всех этих моделей — их синтез. Документ HL-X — это и сами данные, и организация этих данных в структуру, и знания, заключенные в структуре данных, и правила (ограничения) манипулирования данными,



включая права доступа, и визуальное представление данных, и инструкции по обработке этих данных в ИС. Главное достоинство документа HL-X — в высоком уровне абстракции этих моделей, делающих его независимым от конкретной информационной системы, базы данных, технологических средств разработки (конструкторы интерфейса и языки программирования), каналов связи и средств доставки документа (SQL\*NET, TCP/IP, и т. п.):

- документ HL-X может быть легко доставлен из одной ИС в другую с помощью любых каналов связи и средств доставки;
- документ HL-X может быть понят и интерпретирован пользователем без использования модели визуализации на основе своей понятийной модели, благодаря которой он фактически является самодокументируемым;
- модель обработки данных документа в каждой ИС может быть своя, исходя из целей использования этих данных системой (степени интеграции этих данных системой);
- моделей визуализации может быть много, в соответствии с желаемыми представлениями данных.

Попробуйте представить в этой роли документ, реализованный, например, на Oracle Forms (традиционное средство разработки экранных форм от корпорации Oracle). В форме есть модель визуализации, информационная модель данных, основанная на блоках, модель обработки данных и функциональная модель, основанные на триггерах, блоках и свойствах элементов модели. Но попробуйте передать документ, представленный этой формой, в другую ИС. Предложите этой форме поработать автономно без постоянной открытой сессии с сервером БД и соответствующим потреблением ресурсов. Предложите конечному пользователю сконструировать новый документ или изменить существующий. И будет видно, что все присутствующие в форме модели являются очень специфическими и не могут быть использованы вне контекста конкретной МИС, конкретной реляционной БД, сетевой программы SQL\*NET и т. п. Это можно сказать и про другие распространенные средства разработки в технологии клиент-сервер. Специфические реализации и модели стареют и умирают вместе с технологией, их породившей, и не востребуются другими технологиями.

Большую свободу предоставляют различные процессоры документов (Word, Acrobat). Они реализуют очень изощренные модели

визуализации. Однако они не сильны по части обработки содержащихся в документе данных, предоставлении языков запросов к данным документов и т. п.

Преимущества документов HL-X:

- свобода от конкретных, подверженных постоянному развитию и изменению технологий реализации ИС;
- самоценность документа HL-X как носителя информации, описывающей определенную предметную область;
- продление времени жизни модели документа в силу ее инвариантности по отношению к реализации документа HL-X в ИС;
- гибкость при реализации ИС на базе документов HL-X. К этой гибкости относится возможность произвольного выбора БД (реляционной, постреляционной или объектной). Возможность выбора различных технологий реализации интерфейса документа HL-X. Были опробованы в качестве языков реализации прототипа архитектуры: Delphi, PL/SQL, Java, JavaScript, HTML, XML;
- восприимчивость к инновациям в информационных технологиях. Возможность быстрого переноса накопленных моделей документов HL-X в новую информационную среду без их существенных переделок и развертывания ИС на базе новых технологий.

Глубина и широта реализации концепции документа HL-X в ИС определяется конкретными требованиями системы.

В простейшем случае обработка данных документов может не производиться. Система будет оперировать документами как единым целым. В этом случае могут не понадобиться модель обработки данных и, может быть, функциональная модель документа. Особенно если эти документы импортируются из другой ИС и их изменение в принимающей ИС не предполагается. Визуализация может быть сведена к понятийной модели документа и не потребует дополнительных моделей визуализации.

Полная реализация концепции документа HL-X в ИС должна, по нашему мнению, обладать следующими важными концептуальными характеристиками:

- поддерживать многоязычные словари понятий для тех или иных предметных областей;

- предоставлять наборы моделей документов HL-X (архетипы — совершенные образцы), описывающих предметную область. Ценность моделей должна быть подтверждена их реальным практическим использованием в различных ИС, они должны соответствовать государственным или отраслевым стандартам, поддерживаться мнением экспертов. Поддерживать базу знаний, «усвоенных» из архетипов документов;
- поддерживать средства разработки новых моделей документов (конструктор документов) на основе базы знаний;
- предоставлять средства импорта/экспорта словарей понятий и моделей документов;
- предоставлять возможность использования различных технологических средств работы с HL-X документами (хранилища документов, пользовательский интерфейс к документам).

**3.1. Требования терминологические.** Документ HL-X рассматривается нами как формализованная в первую очередь с понятийной точки зрения структура данных. Документ HL-X должен основываться на словаре понятий, достаточном для описания (формализации) всех документов основной предметной области (медицины), а также связанных с основной предметных областей (управление организацией и персоналом, статистика, материально-техническое снабжение, диетпитание и т. п.). Объем словаря понятий потенциально неограничен.

Отдельно следует указать на сложность процесса формирования словаря понятий. Этот процесс вовсе не однозначный с точки зрения архитектуры HL-X. Поясним на примере. Пусть у нас имеются в словаре понятия: «группа», «инвалидность», «кровь». И пусть мы стоим перед выбором: следует ли включать в словарь понятия «группа инвалидности» и «группа крови»? Есть две возможности — либо включить эти понятия непосредственно в словарь, либо передать эти понятия через понятийную и информационную модель в виде соответствующих структур:

<Инвалидность>

<Группа>Первая</Группа>

</Инвалидность>

<Кровь>

<Группа>Первая</Группа>

<Резус>Rh+</Резус>

</Кровь>

В этой ситуации для правильного понимания значения высказывания «группа — первая» нам необходим контекст использования термина «группа». Самого термина уже недостаточно. Но в этом заключается основное отличие всех контекстно-зависимых языков — естественных языков — от формальных алгоритмических и математических языков [6, 7]. Именно поэтому мы сталкиваемся с непреодолимыми трудностями при попытке создать машину, «понимающую» человеческий язык. Другой вариант вышеуказанных структур будет выглядеть так:

<Инвалидность>

<Группа инвалидности>Первая</Группа инвалидности>

</Инвалидность>

<Кровь>

<Группа крови>Первая</Группа крови>

<Резус>Rh+<Резус>

</Кровь>

Кажется на первый взгляд, что мы избавились от контекстной зависимости, но это не так. Кто инвалид? Чья группа крови? Все равно для понимания этих данных необходим дополнительный контекст. Избавиться совсем от контекстной зависимости и необходимости проведения контекстного анализа нам не удастся даже в полностью формализованной системе. Единственная разница в том, что в формализованной системе все контексты заранее известны, а в открытой системе контексты возникают динамически и заранее не predeterminedены.

В архитектуре HL-X тезаурус понятий должен быть дополнен всеми моделями и контекстами использования этих понятий, зафиксированных в архетипах документов. Поэтому оба варианта описания становятся фактически равноправными. В первом приведенном примере мощность тезауруса уменьшается за счет перенесения определения и знания составных понятий в архетипы документов.

**3.2. Требования информационные.** Современное направление развития информатики как науки связано в первую очередь с интеллектуализацией информационных систем и переходом от парадигмы накопления формализованных данных (базы данных) к накоплению формализованных знаний (базы знаний), «открытию» новых неизвестных знаний (Data Mining). В работе [6] этот процесс описан

как переход от работы с обыкновенной бесконтекстной информацией (*informatio vulgaris*) к работе с объектной информацией (*informatio objecta*), и в пределе должен привести к работе с контекстно-зависимой субъектной информацией (*informatio subjecta*), которой оперирует человеческий мозг. В основе этого процесса лежит контекстно-зависимая интерпретация данных. Для нас важно учесть стратегию развития информатики и в этом направлении. Поэтому предлагается сформировать новый документальный уровень представления знаний в технологии HL-X. Каждый документ HL-X, структурированный в виде понятий и концептов (объектов) предметной области, рассматривается нами как носитель знания. Это знание в формализованном виде будет представляться в виде тезауруса понятий и в виде множества контекстов использования любого понятия. Знание должно автоматически «усваиваться» системой непосредственно из документов, сразу же после появления в системе нового архетипа и соответствующих ему экземпляров документов. На основе этого знания предполагается разработать конструктор. При выборе понятия для добавления его в конструируемый документ система сама будет предлагать разработчику документа все известные ей контексты использования этого понятия. Станет возможным контекстный поиск среди документов.

Поведем итог сказанному. Документы HL-X должны рассматриваться как носители знания, представляемого в виде контекстной зависимости понятий и концептов (объектов) предметной области. Подробнее см. работу [17].

**3.3. Требования функциональные.** Основным системообразующим функциональным свойством документа HL-X, отражающим его целостность, должна стать транзакционность. Это свойство означает, что документ и все его содержимое (структурированные формализованные данные и знания) обрабатывается как единое целое, как одна транзакция. Изменение данных документа, например, непосредственно в таблицах БД, вне рамок документа, недопустимо.

Информационная система с поддержкой документа HL-X должна обладать следующей функциональностью:

- транзакционность (целостность);
- авторизация любой информации, содержащейся в документе HL-X;
- историчность представления документа;

- шаблонируемость содержания документа;
- поддержка мультимедийных типов данных;
- возможность разметки графических данных;
- поддержка неограниченного объема данных;
- возможность получения твердой копии любого документа, в том числе с предварительной ее обработкой;
- внедрение (заимствование) информации из других документов;
- обмен информацией с другими системами;
- конструирование документов на уровне прикладного разработчика;
- поддержка коллективной работы с документами (блокировки, права на операции, документооборот, механизмы рабочего стола и т. п.);
- динамическая генерация пользовательского интерфейса для работы с документами;
- возможность работы с документами HL-X в тонком клиенте через Интранет/Интернет;
- возможность автономной работы с документами HL-X в off-line;
- возможность экспорта документа в формат HTML или XML для использования документов HL-X телемедицинскими системами.

Разъясним эти пункты:

*Транзакционность.* Документ и все входящие в него понятия обрабатываются как единое целое в рамках одной транзакции. Изменение содержания документа вне транзакции недопустимо.

*Авторизация информации.* Любая информация, сохраняющаяся в БД в виде документа HL-X, авторизована пользователем системы.

*Историчность.* Историчность позволяет поддерживать неизменными представления документа в интерфейсе пользователя и в виде твердой копии документа (печать документа), хотя описания архетипов документов могут быть в любой момент изменены. Например, в специализированный осмотр врача могут добавляться новые разделы и поля, но «старые» осмотры будут сохранять свой вид при вызове на просмотр или на редактирование.

*Шаблонируемость.* Любой документ может быть сохранен как шаблон с различными уровнями доступа: общедоступные шаблоны, шаблоны подразделений, личные шаблоны. Шаблоны применяются

для быстрого начального заполнения нового экземпляра документа. Допускается возможность применения шаблона одного типа документа к экземпляру другого (близкого по структуре) типа. Это означает, что шаблоны старых осмотров могут быть применены и к новым, отличающимся от старых осмотрам.

*Поддержка мультимедийных типов данных.* В документы HL-X можно включать мультимедийные типы данных (снимки, видео, звукозапись и т. п.).

*Возможность разметки графических данных.* Включаемые в документ графические объекты (например, снимки или отсканированные документы) могут быть размечены с помощью специальной графической компоненты, без изменения оригинального графического объекта. Это позволяет выполнять разметку и комментировать визуальные диагностические данные, например рентгеновские снимки.

*Поддержка неограниченного объема данных.* Поля документа при необходимости могут содержать практически неограниченные объемы данных. Актуально для документов с варьируемым и заранее неограниченным размером (выписка, протоколы исследования вместе с данными).

*Возможность получения твердой копии документа с предварительной ее обработкой.* Документы могут содержать множество опциональных, не обязательных для заполнения полей. Любой документ может быть просмотрен и распечатан на принтере с подавлением пустых полей и разделов. Это улучшает восприятие документа и снижает расход бумажных носителей.

*Внедрение (заимствование) информации.* В качестве динамических шаблонов при создании документа HL-X могут выступать данные из любых других документов системы. Широко используется при формировании выписки и различных эпикризов.

*Обмен информацией с другими системами.* Открытые для редактирования документы HL-X могут получать данные из других систем через специализированный интерфейс обмена данными и инкорпорировать их в документ. Например, можно включать непосредственно в документ рентгеновские диагностические снимки, получаемые через HTTP-сервер диагностической системы.

*Конструирование документов.* Позволяет расширять и модифицировать документы системы. Должно быть основано на документальном уровне представления знаний в виде контекстной зависимости понятий и концептов.

*Поддержка коллективной работы с документами.* Предоставляет возможность коллективной работы над документами, в том числе и через Интернет.

*Динамическая генерация пользовательского интерфейса.* Обеспечивает возможность вводить в работу новые типы документов без необходимости изменения каких-либо модулей системы. Решает задачу динамической генерации интерфейса по запросам тонкого клиента, работающего через локальные или глобальные сети.

*Возможность работы с документами HL-X в тонком клиенте через Интранет/Интернет.* Позволяет организовать удаленные рабочие места без необходимости инсталляции на них модулей системы. Может потребоваться установка программ шифрования данных для безопасного обмена данными через глобальную сеть.

*Возможность автономной работы с документами HL-X в режиме off-line.* Позволит работать с документами на удаленных от центрального ядра системы участках (здравпункты, амбулатории), характеризующихся невозможностью или экономической нецелесообразностью организации связи on-line с центральным ядром системы.

*Возможность выгрузки документа в HTML-или XML-формате.* Предоставляет механизм для экспорта документов в другие системы и использования документов HL-X в телемедицинских системах.

**3.4. Требования архитектурные.** Как было отмечено выше, документ HL-X должен быть реализован в открытой архитектуре и должен обладать известной инвариантностью и независимостью от быстро развивающихся информационных технологий. Это требование приводит к необходимости создания высоко абстрактного модельного представления документов и средств работы с моделями.

Документы HL-X должны «уметь» работать как в сосредоточенном информационном пространстве в рамках одного информационного узла, так и в распределенном информационном пространстве в рамках различных информационных узлов. Особо следует рассмотреть возможность автономной работы документов HL-X в узлах информационного пространства, не обеспечивающих связь on-line с центральным узлом. Речь идет об удаленных от центрального ядра системы участках (здравпункты, амбулатории), в которых невозможна или экономически нецелесообразна организации связи on-line с центральным ядром системы. Отсюда требования автономной работы,



пересылки данных в центральный узел, играющий роль интегрированного хранилища данных, экономическая нецелесообразность установки и сопровождения на удаленных узлах мощных СУБД, таких как Oracle.

Это означает, что архитектура HL-X должна уметь разворачиваться в информационных пространствах самой разной конфигурации и на самых разных технических средствах, начиная с мощного многосерверного центрального узла и вплоть до единственного персонального компьютера (ПК), совмещающего функции клиента и сервера.

Для этого предлагается *отделить хранилище и модели документов от базы данных, поместив их в базу знаний*. База знаний должна обладать свойствами переносимости на узлы произвольной конфигурации и требуемой масштабируемостью.

Что будет выступать в роли хранилища документов локально узла информационного пространства — совершенно безразлично. Можно хранить экземпляры документов HL-X целиком в поле таблицы БД Oracle, или в файловой системе, или на любом другом физическом ресурсе. В качестве оригинальной среды хранения документов для удаленных узлов предлагается архитектура расширенной PACS (Picture Archiving and Communication System) [18].

База знаний документов должна поддерживать автономную, без связи с БД, работу с документами, включая конструирование их моделей в рамках определенных нами ограничений на модели документов на основе накопленных знаний. Локальное хранилище экземпляров документов должно иметь каталог документов и язык поиска (запросов) каталогизированных документов. Домены значений полей документов небольшой мощности могут быть включены в модели документа. Справочники (МКБ10, Классификатор лекарственных средств, Справочник услуг и т. п.) могут быть представлены как отдельные реплицируемые из БД в БЗ документы. Заимствование данных из документа в документ и выбор значений из документов-справочников предусмотрено в функциональных требованиях к документам.

Локальное хранилище документов может поддерживать администрирование пользователей и их прав на работу с документами.

База данных центрального узла информационного пространства играет роль интегрирующего хранилища данных из импортированных документов. Соответственно из БД получают всю требуемую

статистику. Физическое месторасположение БД и ее конкретная реализация (Oracle DB, Microsoft SQL Server, ...) не должны иметь большого значения с точки зрения архитектуры. Достаточно заложить возможность работы с произвольным SQL-сервером. Базы знаний различных узлов информационного пространства могут обмениваться между собой моделями документов HL-X — архетипами документов.

Вся функциональные элементы метауровня знаний — база знаний, конструктор документов и бизнес-логика экземпляров документов — должны быть реализованы на серверах приложений. В качестве подходящего языка реализации можно назвать Java. Возможно использование других языков в рамках технологии Corba (Common Object Request Broker Architecture) или в рамках технологии DotNet (платформозависимая технология Microsoft).

Основная концептуальная работа при разработке архитектуры HL-X — это построение множества архетипов документов над понятиями и концептами предметной области (проблемно-ориентированных моделей документов). Необходимо будет разработать переносимую логическую модель данных и ее отображения в конкретные реляционные БД. В архитектуру закладываются возможности расширения модели предметной области (ПО) и логической модели БД, возможности саморасширения логической модели БД под влиянием расширения модели ПО.

Еще одна основная концептуальная идея архитектуры HL-X — *введение процесса концептуализации предметной области непосредственно в саму информационную систему*. Это ответ на отмеченные выше трудности, связанные с недостаточной формализацией и динамичностью предметной области. Иллюстрацией к сказанному служит Рис. 1.

**3.5. Экономическая характеристика архитектуры.** В архитектуру HL-X заложены возможности развертывания на узлах произвольной мощности информационного пространства любой конфигурации. Узел может состоять даже из отдельного ПК. Архитектура HL-X может работать в режиме off-line, обмениваясь данными с другими узлами в пакетном режиме. Архитектура HL-X должна приспосабливаться к имеющейся конфигурации сети и технических средств и не должна требовать покупки и развертывания дополнительного лицензионного программного обеспечения.

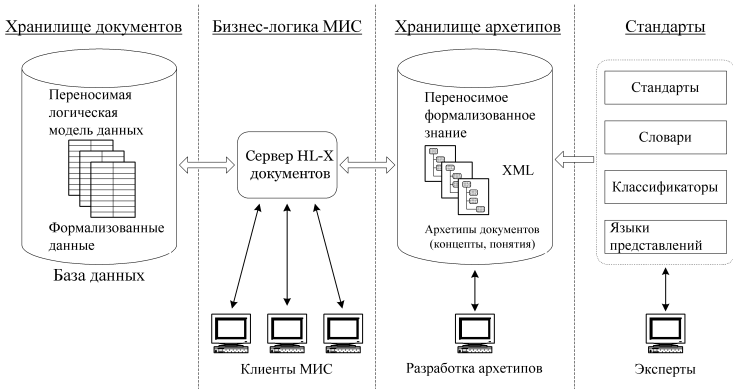


Рис. 1. Процесс концептуализации предметной области в архитектуре HL-X

Другой ожидаемый экономический выигрыш от внедрения архитектуры заключается в снижении стоимости сопровождения системы в целом. По данным [6], основные расходы по сопровождению ИС масштаба предприятия — 90% — ложатся на поддержку актуального состояния модели данных ИС. При этом внесение изменений в структуру БД и в соответствующие бизнес-процессы обработки переструктурированных данных сопряжено со значительными затратами ресурсов и обычной в таких случаях значительной временной задержкой в реализации новых требований. Предлагаемая архитектура должна в первую очередь упростить создание новых и модификацию имеющихся в системе медицинских документов. Система станет более адаптивной и начнет развиваться самими пользователями.

#### 4. Заключение

Основные тезисы, отражающие суть архитектуры HL-X:

- (1) *Архитектура* признает неизбежность появления в будущем стандартов представления медицинской информации и предполагает, что в основу будущих открытых для расширения стандартов лягут словари понятий и концептов. Стандарты

- будут определять жесткие модели данных (понятия и концепты предметной области) и правила (ограничения) конструирования информационных структур на основе понятийного и концептуального тезауруса предметной области;
- (2) *Архитектура* характеризуется достаточно высоким уровнем абстракции, что делает ее независимой от быстро изменяющихся информационных технологий. Документ HL-X является синтезом нескольких высоко абстрактных моделей, описывающих его структурные и функциональные свойства;
  - (3) *Архитектура* предлагает разделить хранилище архетипов документов HL-X и базу данных. Предполагается возможность распределенного хранения архетипов между узлами единого информационного пространства системы, с возможностью обмена знаниями об архетипах между узлами;
  - (4) процесс концептуализации предметной области, заключающийся во вводе в систему новых понятий, концептов и архетипов документов, должен быть непосредственно встроен в архитектуру системы.

Авторы статьи сознательно постарались избежать лишних технических подробностей, не затрагивающих концептуального уровня описываемой архитектуры. Некоторые детали можно найти в [15–18]. В настоящее время архитектура находится в состоянии прототипирования и опытной эксплуатации.

### Список литературы

- [1] Kalra D., Beale T., Heard S., Lloyd D. Tutorial. — <http://www.openehr.org>, 2003. ↑1
- [2] McDonald, Overhage // International Journal of Medical Informatics. ↑1
- [3] Емелин И. В., Перов Ю. Л., Серегин Ю. С., Эльчиан Р. А. *Концепция построения открытых медицинских информационных систем* // Кремлевская медицина. Клинический вестник, № 1. ↑2
- [4] Емелин И. В. *Интеграция стандартов медицинской информатики* // Кремлевская медицина. Клинический вестник, № 4. ↑2
- [5] Waegemann P., Tessier C. A REPORT ON INFORMATION CAPTURE AND REPORT GENERATION. — Newton, MA: Medical Records Institute, 2002. ↑

- [6] Поляков А. О. Информационная общность систем. — СПб: Изд-во СПбГТУ, 2002. ↑3.1, 3.2, 3.5
- [7] Лачинов В. М., Поляков А. О. Информодинамика или Путь к Миру открытых систем. — СПб: Изд-во СПбГТУ, 1999. ↑3.1
- [8] Kalra D., Beale T., Heard S., Lloyd D. An EHR Architecture for Archetyped Health Information Systems. — <http://www.openehr.org>, 2003. ↑
- [9] European Health Record Project, Deliverable No. 19, 1995. ↑
- [10] CEN Interim Report. — 1.1, 1998, See also <http://www.chime.ucl.ac.uk/HealthI/EHCR-SupA/>. ↑
- [11] Electronic Healthcare Record Support Action. — <http://www.chime.ucl.ac.uk/HealthI/EHCR-SupA/>, 2002. ↑
- [12] The Synapses Project. — <http://www.chime.ucl.ac.uk/HealthI/SYNAPSES/>, 2002. ↑
- [13] The Synapses Project. — <http://www.cs.tcd.ie/synapses/public/>. ↑
- [14] CEN/TC 251. — <http://www.centc251.org/>, 2002. ↑
- [15] Концепция интеграции системы Интернет с другими информационными системами: Отчет по НИР/ ИПС РАН. — Переславль-Залесский, 2001. ↑4
- [16] Исследование механизмов представления медицинской информации на основе понятий и моделей документов (Технология HL-X): Отчет по НИР/ ИПС РАН. — Переславль-Залесский, 2003. ↑
- [17] Малых В. Л., Юрченко С. Г. Документальный уровень представления знаний в интегрированной медицинской информационной системе ИПС РАН. — Переславль-Залесский, 2004. ↑3.2
- [18] Лапшин М. А. Расширенная архитектура PACS в распределённой медицинской информационной системе ИПС РАН. — Переславль-Залесский, 2004. ↑3.4, 4

ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР МЕДИЦИНСКОЙ ИНФОРМАТИКИ ИПС РАН

Ya. I.-o. Guliev, V. L. Malykh. *HL-X Architecture*. (in russian.)

АБСТРАКТ. In this article the authors propose an approach for design of architecture of corporate medical information systems. A conception, that underlies this approach, talks about necessity of support of evolutionary process of conceptualization of data domain in information systems. The architecture is open, independent from programming aids, oriented at standards that exist or are in the process of development, scalable to the level of separate distant subdivisions.