

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт Кибернетики
Направление подготовки 09.04.01 Информатика и вычислительная техника
Кафедра программной инженерии

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Региональная мониторинговая система здравоохранения

УДК

Студент

Группа	ФИО	Подпись	Дата
8ВМ5В	Мохаммеджавад Бушра Джабер		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. АиКС	Пономарев Алексей Анатольевич	К.Т.Н.		

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Конотопский Владимир Юрьевич	К.Э.Н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент	Акулов Петр Анатольевич			

ДОПУСТИТЬ К ЗАЩИТЕ:

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
ПИ	Иванов Максим Анатольевич	К.Т.Н.		

Томск – 2017 г.

Планируемые результаты обучения

Код результата	Результат обучения (выпускник должен быть готов)
Общепрофессиональные компетенции	
P1	Воспринимать и самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте.
P2	Владеть и применять методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях.
P3	Демонстрировать культуру мышления, способность выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных из разных областей науки и техники, выносить суждения на основании неполных данных, анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями.
P4	Анализировать и оценивать уровни своих компетенций в сочетании со способностью и готовностью к саморегулированию дальнейшего образования и профессиональной мобильности. Владеть, по крайней мере, одним из иностранных языков на уровне социального и профессионального общения, применять специальную лексику и профессиональную терминологию языка.
Профессиональные компетенции	
P5	Выполнять инновационные инженерные проекты по разработке аппаратных и программных средств автоматизированных систем различного назначения с использованием современных методов проектирования, систем автоматизированного проектирования, передового опыта разработки конкурентно способных изделий.
P6	Планировать и проводить теоретические и экспериментальные исследования в области проектирования аппаратных и программных средств автоматизированных систем с использованием новейших достижений науки и техники, передового отечественного и зарубежного опыта. Критически оценивать полученные данные и делать выводы.
P7	Осуществлять авторское сопровождение процессов проектирования, внедрения и эксплуатации аппаратных и программных средств автоматизированных систем различного назначения.
Общекультурные компетенции	
P8	Использовать на практике умения и навыки в организации исследовательских, проектных работ и профессиональной эксплуатации

Код результата	Результат обучения (выпускник должен быть готов)
	современного оборудования и приборов, в управлении коллективом.
P9	Осуществлять коммуникации в профессиональной среде и в обществе в целом, активно владеть иностранным языком, разрабатывать документацию, презентовать и защищать результаты инновационной инженерной деятельности, в том числе на иностранном языке.
P10	Совершенствовать и развивать свой интеллектуальный и общекультурный уровень. Проявлять инициативу, в том числе в ситуациях риска, брать на себя всю полноту ответственности.
P11	Демонстрировать способность к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности, способность самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности, способность к педагогической деятельности.

Министерство образования и науки Российской Федерации

федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт кибернетики
Направление подготовки (специальность) 09.04.01 Информатика и вычислительная техника
Кафедра программной инженерии

УТВЕРЖДАЮ:
Зав. кафедрой ПИ
М. А. Иванов
(Подпись) (Дата) (Ф.И.О.)

**ЗАДАНИЕ
на выполнение выпускной квалификационной работы**

В форме:

магистерской диссертации

Студенту:

Группа	ФИО
8ВМ5В	Мохаммеджавад Бушра Джабер

Тема работы:

Региональная мониторинговая система здравоохранения	
Утверждена приказом директора (дата, номер)	1652/с от 13.03.2017

Срок сдачи студентом выполненной работы:	08 июня 2017
--	---------------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	
<i>(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).</i>	<ol style="list-style-type: none">1. Ситуационные центры здравоохранения.2. Форматы хранения медицинских данных.3. Медицинские статистические данные по Томской области.

<p>Перечень подлежащих исследованию, проектированию и разработке вопросов</p> <p><i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i></p>	<ol style="list-style-type: none"> Исследование предметной области для проектирования приложения для генерации медицинских данных в распространенных форматах Анализ статистических медицинских данных по Томской области Исследование средств визуализации и интерпретации медицинских данных Программная реализация приложения для генерации медицинских данных Финансовый менеджмент, ресурс эффективность и ресурсосбережение Социальная ответственность
<p>Перечень графического материала</p> <p><i>(с точным указанием обязательных чертежей)</i></p>	<ol style="list-style-type: none"> Презентация в программе MS PowerPoint (в формате pptx).

Консультанты по разделам выпускной квалификационной работы
(с указанием разделов)

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Конотопский Владимир Юрьевич
Социальная ответственность	Акулов Петр Анатольевич
Раздел на иностранном языке	Сидоренко Татьяна Валерьевна

Названия разделов, которые должны быть написаны на русском и иностранном языках:

Глава 3 Разработка программных решений для задач формирования и мониторинга показателей здоровья населения региона.

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	
---	--

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. АиКС	Пономарев Алексей Анатольевич	к.т.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ5В	Мохаммеджавад Бушра Джабер		

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт кибернетики
Направление подготовки Информатика и вычислительная техника
Уровень образования Магистратура
Кафедра программной инженерии
Период выполнения весенний семестр 2016/2017 учебного года

Форма представления работы:

магистерская диссертация

(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	8 июня 2017
--	-------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
27.5.2017	Основная часть	60
03.05.2017	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	15
12.05.2017	Социальная ответственность	10
27.5.2017	Обязательное приложение на иностранном языке	15

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. АиКС	Пономарев А. А.	К.Т.Н.		

СОГЛАСОВАНО:

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
ПИ	Иванов М. А.	К.Т.Н.		

РЕФЕРАТ

Выпускная квалификационная работа 145 с., 44 рисунка, 13 таблиц, 46 источников, 1 приложение.

Ключевые слова: ЭМК, CDA, ситуационный центр, алгоритм облако тегов, Data mining.

Объектом исследования являются ситуационные центры здравоохранения.

Цель работы –разработка программных решений для решения задач формирования и мониторинга показателей здоровья населения региона.

В процессе работы проводились исследования существующих стандартов для организации ЭМК, эффективных средств интерпретации данных ЭМК.

В результате исследования было разработано программное обеспечение для генерации медицинских данных в общепринятом формате, проведен анализ статистических медицинских данных Томской области, приведен пример интерпретации медицинских данных.

Эффективность приложения определяется повышением скорости работы персонала с медицинскими данными и удобством их интерпретации.

Областью применения являются медицинские учреждения и департамент здравоохранения.

Экономическая значимость работы. Выпускная квалификационная работа выполнена в текстовом редакторе Microsoft Word 2010 и представлена на диске (в конверте на обороте обложки).

В будущем планируется развитие приложения для центра обработки данных.

Определения, обозначения, сокращения, нормативные ссылки

Нормативные ссылки

В данной выпускной квалификационной работе использованы ссылки на следующие нормативные документы:

ГОСТ 12.0.003-74. Система стандартов безопасности труда. Опасные и вредные производственные факторы.

ГОСТ 12.2.032-78 ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования.

ГОСТ 22269-76. Система "Человек-машина". Рабочее место оператора. Взаимное расположение элементов рабочего места. Общие эргономические требования.

ГОСТ Р 12.1.019-2009 ССБТ. Электробезопасность. Общие требования и номенклатура видов защиты.

ГОСТ Р 50923-96. Дисплеи. Рабочее место оператора. Общие эргономические требования и требования к производственной среде. Методы измерения.

СанПиН 2.2.2/2.4.1340-03. Гигиенические требования к персональным электронно-вычислительным машинам и организации работы.

СанПиН 2.2.4.548-96. Гигиенические требования к микроклимату производственных помещений.

СНиП 21-01-97. Пожарная безопасность зданий и сооружений.

СП 52.13330.2011. Естественное и искусственное освещение.

Определения

Электронная медицинская карта (ЭМК) пациента – это комплекс данных о состоянии здоровья пациента и назначаемом ему лечении, которые хранятся и обрабатываются в электронном виде.

Clinical Document Architecture (CDA) – один из стандартов HL7, разработанный для стандартизации структуры и обеспечения семантической совместимости мед систем при обмене медицинской информацией и/или мед документами.

OpenEHR – это виртуальное сообщество работает над средствами превращения данных о состоянии здоровья от физической формы в электронную форму и обеспечения всеобщей совместимости среди всех форм электронных данных. Основное внимание в своей деятельности на электронных медицинских записей (ЭМК) и связанных с ними систем.

BPEL (англ. Business Process Execution Language) — язык на основе XML для формального описания бизнес-процессов и протоколов их взаимодействия между собой.

Облако тегов (Tag-Cloud) – это визуальное представление списка категорий (или тегов, также называемых метками, ярлыками, ключевыми словами и т. п.) Обычно используется для описания ключевых слов (тегов) на веб-сайтах, или для представления неформатированного текста. Ключевые слова чаще всего представляют собой отдельные слова, и важность каждого ключевого слова обозначается размером шрифта или цветом. Такое представление удобно для быстрого восприятия наиболее известных терминов и для распределения терминов по популярности относительно друг друга.

Data Mining – это процесс обнаружения в "сырых" данных ранее неизвестных нетривиальных практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой

деятельности. Data Mining является одним из шагов Knowledge Discovery in Databases.

Обозначения и сокращения

В данной выпускной квалификационной работе были использованы следующие обозначения и сокращения:

ЭМК – Электронная медицинская карта.

ЭИБ – Электронная история болезни.

СЦ – Ситуационный центр.

ТАП – Талон амбулаторного пациента.

CDA – Clinical Document Architecture.

СЭМД – Стандартизированный электронный медицинский документ.

BPEL – Business Process Execution Language.

МИС – Медицинская информационная система.

XML – (eXtensible Markup Language) – расширяемый язык разметки.

МКБ – Международный классификатор болезней.

БД – база данных.

Оглавление

Введение.....	14
Глава 1. Современное состояние информатизации в области здравоохранения.....	15
1.1 Анализ современного состояния здравоохранения в России и за рубежом в разрезе использования информационных технологий и организации документооборота.....	15
1.1.1 Оформление электронной медкарты пациента.....	21
1.1.2 Электронный документооборот в медицинских учреждениях: проблемы законодательства.....	21
1.2 Анализ существующих стандартов для организации ЭМК.....	22
1.2.1 Clinical Document Architecture (CDA).....	22
1.2.2 OpenEHR.....	30
1.2.3 OpenUMS.....	33
Глава 2. Организация ситуационных центров для задач здравоохранения.....	35
2.1 Понятие ситуационного центра.....	35
2.1.1 Цель СЦ.....	36
2.1.2 Задачи СЦ.....	36
2.3 Методология формирования ЭМК пациента с использованием сочетания форматов OpenUMS, CDA и BPEL.....	38
2.3.1 CDA.....	38
2.3.2 OpenEHR.....	39
2.3.3 OpenUMS.....	39
2.3.4 BPEL Business Process Execution Language.....	40
BPMN.....	46
2.4 Разработка алгоритмических и программных средств интеграции данных при создании ЭМК пациента.....	48
2.5 Разработка визуальных компонентов для задач быстрой навигации по ЭМК.....	55
2.5.1 Анализ существующих средств быстрой навигации.....	55
2.5.2 Источники данных.....	57
2.5.3 Использование алгоритма “Облако тегов” для навигации по электронной истории болезни.....	58
Глава 3. Разработка программных решений для задач формирования и мониторинга показателей здоровья населения региона.....	67
3.1 Анализ и прогнозирование показателей здоровья населения и среды обитания человека.....	67
3.2 Прогнозирование потребности в медицинской помощи на основании данных об уровне и характере заболеваемости по различным нозологическим формам, группам населения, территориальным и другим признакам.....	69
3.2.1 Линейная регрессия (частный случай анализа временных рядов).....	69
3.2.2 нейронные сети.....	72
3.2.3 кластерный анализ.....	74
Глава 4. Разработка средств интерпретации данных мониторинга.....	77

4.1. Разработка эффективных средств интерпретации данных ЭМК	77
4.2 геоинформационные системы (ГИС)	84
4.2.1 Значки	84
4.2.2 Картодиаграммы	85
4.2.3 Картограммы	86
Глава 5. Оценка эффективности разработанных методик, алгоритмов и программного комплекса	88
5.1 Оценка соответствия разработанных электронных форм и реальных документов.....	88
5.2 Тестирование разработанной системы.....	91
5.3 Оценка результатов тестирования.....	94
Глава 6. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	95
6.1. Организация и планирование работ	96
6.1.1. Продолжительность этапов работ	97
6.1.2. Расчет накопления готовности проекта	100
6.2. Расчет сметы затрат на выполнение проекта	102
6.2.1 Расчет затрат на материалы	102
6.2.1. Расчет заработной платы.....	103
6.2.2. Расчет затрат на социальный налог.....	104
6.2.3. Расчет затрат на электроэнергию	104
6.2.4. Расчет амортизационных расходов	106
6.2.5. Расчет прочих расходов.....	107
6.2.6. Расчет общей себестоимости разработки	107
6.2.7. Расчет прибыли	108
6.2.8. Расчет НДС	108
6.2.9. Цена разработки НИР	108
6.3. Оценка экономической эффективности проекта	108
Глава 7. Социальная ответственность.....	110
Введение	112
7.1. Производственная безопасность	112
7.1.1. Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности.	112
7.1.1.1 Микроклимат.....	112
7.1.1.2 недостаточное искусственное освещение.....	113
7.1.2. Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности.	115
7.1.2.1 Пожарная безопасность.....	115
7.1.2.2 Электробезопасность.....	115
7.1.3. Рекомендации по минимизации влияния.....	117
7.1.3.1 Рекомендации по улучшению микроклимата	117
7.1.3.2 Рекомендации по минимизации влияния освещения	117

7.1.3.3 Меры по обеспечению пожарной безопасности	118
7.1.3.4 Рекомендации по защите от электрического тока	118
7.2. Экологическая безопасность.....	120
7.2.1 Анализ воздействия на окружающую среду	120
7.2.2 Рекомендации по минимизации влияния на окружающую среду.....	120
7.3.1 Перечень возможных чрезвычайных ситуаций на объекте	121
7.3.2. Меры по ликвидации ЧС и их последствий	122
7.4.1 Психофизиологические аспекты	123
7.4.2. Организационные мероприятия при компоновке рабочей зоны.....	124
7.4.3. Обеспечение гарантий защиты конфиденциальных данных граждан.....	125
Заключение	126
Список публикаций студента	127
Список использованной литературы	128
Приложение А.....	135
.....	

Введение

Состояние современного мира обусловлено быстрым развитием различных технологий; особенно актуальны новые технологии, связанные со здоровьем человека. В настоящее время медицинские учреждения России переходят от бумажного документооборота к электронному, что должно упрощать работу медицинского персонала, создать базу для статистических исследований работников департамента здравоохранения и улучшить доступ к своим медицинским данным для пациентов.

С появлением электронного медицинского документооборота объем пригодного для анализа набора данных существенно увеличился. Учитывая, что пользователями каких систем являются не только IT-специалисты, возникает потребность разработки новых интерфейсов для обеспечения поиска, фильтрации, сортировки.

В работе рассматриваются наиболее популярные форматы представления медицинских документов в России и за рубежом.

Также в исследовании рассматриваются некоторые идеи по использованию подхода «*Облако тегов*» для облегчения процесса навигации по ЭМК.

Разработанные программные решения в сфере здравоохранения позволят пользователям (врачам и работникам департамента здравоохранения) быстро и эффективно найти интересующую их информацию о пациентах, их заболеваниях и диагнозах. Также одной из областей применения результатов работы является анализ состояния распространенности заболеваний в Томской области.

При использовании общего формата ЭМК облегчается решение задачи разработки программных решений для решения задач формирования и мониторинга показателей здоровья населения региона, что является целью настоящей работы.

Глава 1. Современное состояние информатизации в области здравоохранения

1.1 Анализ современного состояния здравоохранения в России и за рубежом в разрезе использования информационных технологий и организации документооборота

Повышение эффективности использования кадрового потенциала медицинских работников в контексте обозначенных реформ и модернизации сферы здравоохранения непосредственно связано с процессами информатизации отдельных подразделений медицинских учреждений.

Одним из главных направлений такой модернизации отрасли в условиях императивов общенациональной инноватизации и информатизации является ее техническое переоснащение, предполагающее постоянное обновление применяемого оборудования, использование новых технологий, в том числе информационных.

Медицина переходит от двузначной модели «врач – пациент» к трехзначной – «врач – технология – пациент». Тем не менее, рассматривая уровень информатизации современных медицинских учреждений в России, можно отметить существенное отставание показателей внедрения информационных систем и технологий по сравнению с ведущими зарубежными странами, что непосредственно сказывается на показателях эффективности реализации кадрового потенциала, а также на удовлетворенности населения качеством оказываемых услуг. [1]

По данным, предоставленным Министерством здравоохранения РФ, в 2016 году только 20% российских медицинских учреждений были оснащены современными информационными технологиями. На одно медицинское учреждение приходилось в среднем 37 компьютеров. Тем не менее, за последние пять лет расходы на информационно-коммуникационные технологии растут достаточно высокими темпами. В целом затраты всех медицинских учреждений

РФ, связанных с развитием информационных и коммуникационных технологий, в 2016 году превысили 44 млрд. руб. Причем дополнительные возможности финансирования проектов информатизации сферы здравоохранения предоставляет региональным властям наделение субъектов Федерации определенной самостоятельностью в отношении формирования и реализации финансовой политики на своей территории.

Проведенный анализ показал, что внедрение информационных технологий в деятельность медицинских учреждений происходит с различным темпом по отдельным функциональным направлениям, а также в государственной и коммерческой сфере. Так, большая часть частных медицинских учреждений крупных городов используют электронную регистратуру, ведут электронные медицинские карты. Как правило, данные медицинские системы интегрированы с финансовой системой 1С.

Наиболее прогрессивные учреждения здравоохранения уже начали внедрять решения в сфере бизнес-аналитики и удаленного мониторинга здоровья. В большинстве медицинских учреждений автоматизирована деятельность бухгалтерии, а также такие процессы, как выписка льготных рецептов и запись на приём. Однако, информатизация практически не коснулась непосредственно работы медицинского персонала. Тогда как, например, важнейшей проблемой, существующей в медицинских учреждениях (особенно государственных) является автоматизация деятельности регистратуры.

Основная цель автоматизации регистрационной деятельности медицинских учреждений – перевод части документооборота организации в электронный вид. Создание электронной картотеки позволяет упростить и ускорить процесс работы регистратуры для более быстрого и качественно обслуживания клиентов.

ЭМК каждого пациента должна содержать множество полезных сведений и заполняться всего за несколько минут. Администратор должен иметь

возможность быстро отсортировать записи в картотеке по разным параметрам или найти нужного пациента при помощи поиска по любому из полей электронной карты.

Кроме того, в условиях работы учреждения в системе медицинского страхования актуальными является вопрос верификации принадлежности пациента к той или иной страховой компании, и, следовательно, корректного оформления ЭМК.

В целом основные преимущества автоматизации в медицинских учреждениях заключаются в следующем:

- увеличение пропускной способности медицинского учреждения;
- повышение качества обслуживания;
- повышение качества ведения записей в медицинской карте пациента;
- оформление данных о пациенте в строгом соответствии с стандартами оказания медицинской помощи и рекомендациям Минздравсоцразвития РФ;
- сокращение затрат времени на поиск информации о пациенте, тем самым, увеличивается время общения врача и пациента, что позволяет оказывать медицинскую помощь на более качественном уровне;
- повышение качества планирования лечебно-профилактических мероприятий на основе наличия единой базы о пациентах;
- автоматическое формирование счетов и регистров для оплаты медицинских услуг, что позволяет избежать ошибок, связанных с человеческим фактором;
- возможность получения статистических данных, необходимых для планирования и прогнозирования различного рода мероприятий в медицинской деятельности. [1]

Для автоматизации медицинского учреждения возможно формирование модульной МИС в рамках создания автоматизированного рабочего места, возможно использование специальных Интернет-ресурсов.

Заметим, что online- сервисы гарантируют полную конфиденциальность информации, размещаемой медицинским учреждением.

Независимо от выбранного способа автоматизации функционал работы регистратуры должен включать следующее (рисунок 1.1).

Подводя итог, можно сделать вывод, что независимо от выбранного способа автоматизации деятельности медицинских учреждений, электронная регистратура обладает рядом несомненных преимуществ, обеспечивающих повышение эффективности реализации кадрового потенциала работников системы здравоохранения. Это выражается со стороны пациентов:

- в достоверном учете данных пациентов, обращающихся за медицинской помощью, эффективном контроле равномерного распределения пациентов между врачами, ведущими амбулаторный прием, согласно действующим нормам и стандартам приема;
- специалисты медицинских учреждений получают возможность планирования и контроля времени каждого приема за счет прогнозирования персон, записанных на прием, возможность подготовки и оформления документов перед приемом;
- руководитель медицинского учреждения получает возможность объективного контроля деятельности врачей, ведущих прием, оценки востребованности каждого специалиста, а также контроля финансовых потоков.



Рисунок 1.1- Автоматизации функционал работы регистратуры

Медицинские учреждения выбирают способ автоматизации регистратуры, ориентируясь в основном на финансовые возможности и масштаб деятельности организации, наличие каких-либо специфических требований регистрационного учета, лояльность персонала к внедрению новых информационных технологий.

Представляется, что с учетом концепции Правительства РФ, направленной на информатизацию сферы здравоохранения и диктующей определенные требования к способу автоматизации регистрационного учета наиболее подходящим вариантом автоматизации является комплексная модульная МИС. Это обусловлено тем, что в перспективе автоматизация и информатизация медицинских учреждений должна становиться всё «глубже».

Если медицинское учреждение выберет такой способ, как приобретение коробочной или разработанной «под себя» электронной регистратуры, то впоследствии возникнут трудности, а в худшем случае и невозможность, совмещения её с другими программными продуктами.

С другой стороны, СЭМД в медицинских учреждениях все еще не может гарантировать гражданам полную защиту их личных данных. Недостаточно разработанной является и нормативно-правовая база в части внедрения современных электронных систем. В связи с этим, любые инициативы, которые связаны с информатизацией, следует рассматривать, в первую очередь, с точки зрения норм права.

СЭМД в медицинских учреждениях в настоящее время активно проводится в рамках концепции развития здравоохранения в РФ до 2020 года.

На сегодняшний день положено начало повсеместного применения информационных технологий в отрасли здравоохранения. Главная цель новых технологий – повысить качества оказываемой гражданам медпомощи, автоматизировать процесс ведения их персональных данных.

Экономические расчеты говорят о том, что процесс информатизации неизбежно сопровождается значительными издержками, однако, в будущем приносит медучреждениям значительную экономию, связанную с внедрением ЭМК, принятием управленческих решений и т.д. Оцифровка основных медицинских документов позволяет сократить такие риски.

Кроме того, информатизация позволяет: обеспечить сохранность истории болезни в электронной форме, например, если ее оригинал будет утерян; медработники в других медучреждениях могут оперативно обмениваться информацией в условиях одной информационной базы.

Несмотря на то, что конвертация бумажных документов в электронные позволяет обеспечить их сохранность, актуальной является и проблема обеспечения их надежной защиты и валидности.

1.1.1 Оформление ЭМК пациента

Впервые стало возможным применять ЭМК, а также их архивировать, после принятия приказа Росте регулирования № 407-ст от 27.12.2006 года. Минздрав РФ в письме № 18-1/1010 от 11.11.2013 года утвердил основные разделы ЭМК, предъявил требования к стандартизации таких карт.

СЭМД в медицинских учреждениях, в том числе и ЭМК в первую очередь ориентированы на пациентов. Поэтому они должны содержать полную информацию о медицинском обеспечении, включая экстренные и вспомогательные услуги. В ЭМК, помимо прочего, указываются мнения специалистов, результаты их наблюдения, а также планы дальнейшего лечения. Зафиксированные в ЭМК пациента данные позволяют осуществлять контроль лечебно-диагностического процесса, приводить рекомендации специалистов в перспективном плане лечения и обследования, сформировать необходимые данные для контролирующих органов. [2]

1.1.2 СЭМД в медицинских учреждениях: проблемы законодательства

Ни в одном из нормативных актов в сфере информатизации здравоохранения не сказано о реализации прав больных. В законах и подзаконных актах не уточняется то, каким образом при переводе обычных медкарт в электронные можно будет осуществить право больного на получение им информации о своем состоянии. Действующие системы пока не предоставляют больным возможность знакомиться со своей ЭМК удаленно, а в связи с развитием компьютерной преступности, актуальна возможность утечки важной и конфиденциальной информации, поэтому необходимо внедрять новые методы и способы защиты существующих программ.

На сегодняшний день ФСБ Росси совместно с Роскомнадзором разработан объёмный массив регламентов, посвящённых защите прав граждан. Однако, в плане информатизации здравоохранения, защите персональных данных пациентов достаточное внимание не уделяется. В заключении стоит

отметить, что информатизация здравоохранения – задача актуальная, разумная и реализуемая. Несмотря на это, остаются многие недоработки и сложности в реализации разработанных программ на местах, что в итоге может повлечь нарушение прав больных.

Предполагаем, что развитие отрасли, а также информатизация медицины сможет улучшить доступность и качество оказываемой медпомощи, а также существенно сократит сроки принятия основных врачебных решений.[2]

1.2 Анализ существующих стандартов для организации ЭМК

1.2.1 Clinical Document Architecture (CDA)

Clinical Document Architecture (CDA) — один из стандартов HL7, разработанный для стандартизации структуры и обеспечения семантической совместимости мед систем при обмене медицинской информацией и/или мед документами. Первая версия стандарта была одобрена ANSI ещё в 2001 году. Вторая версия, которая используется и по сей день, была утверждена ANSI в 2005. Третья версия, CDA R3, находится в стадии разработки и согласования.

В соответствии с требованиями, изложенными в данном стандарте, текущий документ состоит из профиля аннотации, и профиля локализации.

CDA R2 (Release 2) гарантирует наличие следующих семи характеристик в CDA документе:

- сохранность представленной информации;
- управление представленной информацией;
- поддержка требований к аутентификации всей представленной информации;
- поддержка контекста представленной информации;
- поддержка целостности информации;
- возможность чтения представленной информации человеком;
- поддержка бинарной информации, таких как мультимедийные компоненты, PDF, изображения и прочее.[5]

Подобные характеристики делают CDA крайне гибким к использованию в различных областях.

Медицинский документ, соответствующий стандарту CDA не должен быть в нечитабельном состоянии, при использовании популярных браузеров поддерживающих XML-документацию. Для оформления документации в медицинском учреждении допускается применение внутренних (собственных) стилей отображения.

Согласно общепринятому стандарту HL7, CDA документация содержит заголовок и текст документации [6]. Заголовок CDA-документа состоит из следующих данных:

- область применения документа;
- имя документа;
- шаблон;
- дата и время создания документа;
- уровень защищенности;
- -язык;
- версия документа;
- персональные данные пациента;
- персональные данные врача и медицинского учреждения.

Текст CDA документа состоит из информации, описанной в соответствии с СЭМД.

При этом информация может быть представлена как:

- структурированные записи в соответствии шаблона XML;
- в произвольной форме (текстовый документ, изображение, видео, звук);
- в виде внешней ссылки на исходную информацию.

Как и любой документ, CDA содержит заголовок документа (CDA Header) и тело документа (CDA Body). В заголовок CDA включена некоторая административная информация на основе классов и сущностей RIM модели HL7.

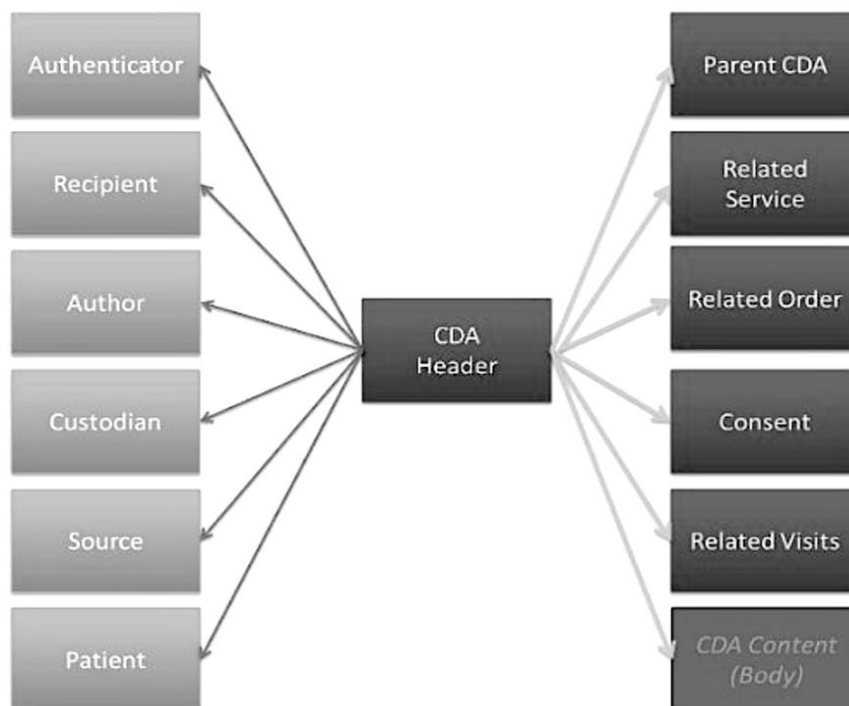


Рисунок 2.2. Структура CDA

Основной класс в заголовке CDA так и называется — Clinical Document, и, в соответствии с его названием, содержит кроме всего прочего информацию для идентификации документа, заголовок, используемый язык, версию, дату публикации и уровень конфиденциальности.

Назначение основных классов:

Authenticator

Этот класс включен в документ для идентификации лица и/или организации, которые могут использоваться для проверки подлинности всего содержимого документа CDA.

Recipient

Здесь всё просто, в заголовке документа recipient используется для идентификации получателя (лица и/или организации) документа и может включать такие данные как имя, адрес и прочую информацию. Получателей может быть множество.

Author

В данном классе кодируется информация для идентификации и проверки лица или устройства, а также организации их представляющих, которые ответственны за данные в CDA документе.

Custodian

Данный класс используется для идентификации и проверки организации ответственной за сохранность документа. Подобная организация отвечает за сохранность всех данных в документе (за весь документ).

Source

В данном классе кодируется информация о лице и/или организации поместившим данные в медицинскую систему, на основе которых и был создан документ. Подобным лицом может быть, например, Data Entry clerk.

Parent CDA

Поскольку документ CDA может быть частью другого документа CDA, подобная информация должна быть где-то представлена. В данном классе как раз кодируется информация о взаимосвязях между двумя и более документами.

Patient

Используя данный класс в заголовке CDA документа представлена информация о главном участнике всех событий – пациенте. Имя, адрес, опекуны и всякая прочая информация для полной идентификации пациента. Стоит заметить, что некоторые типы данных запрещены в некоторых странах. Например, информация о расовой и религиозной принадлежности требуется в США, но недопустима в Европе.

CDA Body

Тело документа или CDA Body может быть как неструктурированным для данных, которые не кодируются в XML, так и структурированным для представленных в формате XML данных. С особенностями такого кодирования тесно связано понятие уровней семантической совместимости CDA (в данной статье не рассматриваются).

Non-structured Body

Неструктурированное тело документа может содержать только одно вложение со ссылкой на данные вне этого CDA документа, либо Base64 кодированное бинарное вложение (PDF, изображение, аудио, видео и т.п.).

Structured Body

Структурированное тело документа разработано как крайне гибкий механизм для включения разнообразных клинических и административных данных. Именно это делает CDA настолько мощным в концептуальном плане, что теоретически его можно использовать для кодирования всей истории болезни пациента, какой бы сложности она не была.

Структурированное тело документа может содержать одну или более секций, каждая из которых может содержать мед или административную информацию, а также вложения. При этом в самом стандарте CDA ничего не сказано о том как устанавливать отношения между подобными данными, для

этого нужно обращаться к многочисленным руководствам разработчика доступным на официальном сайте HL7 либо разрабатываемых каждой организацией в отдельности.

Вся клиническая информация сгруппирована внутри секций СЭМД [7].

Элементы CDA способны описать одни и те же события различными способами. Для обеспечения однозначного понимания описываемых событий между каждым звеном электронного обмена, применяется шаблонный механизм. Шаблоны являются дополнительным ограничением CDA и определены OID-номером шаблона. Шаблоны способствуют одинаковому описанию одних и тех же событий в соответствующем контексте.

Документация автоматически анализируется на наличие указания шаблона, а также на предмет соответствия этому шаблону.

Система шаблонов имеет огромный потенциал к дальнейшему развитию и модернизации.

В информационной системе ИЭМК на сегодняшний день существуют такие типы СЭМД как:

- выписной эпикриз в стационаре;
- талон амбулаторного посещения;
- направление на консультацию и во вспомогательные кабинеты.

«Выписной эпикриз в стационаре» - это основной документ, резюмирующий пребывание пациента в МУ стационарного типа. Документ представляет собой важнейшую форму медицинской документации, которая предназначена для обмена данными между врачами различных организаций и позволяет отслеживать состояние здоровья пациента на протяжении всего периода жизни[6].



Рисунок 1.3 – Дерево шаблонов СЭМД

Примерный состав сведений Эпикриза в стационаре выписного отображен на рис 1.2



Рисунок 1.4 – Структура эпикриза в стационаре выписного

Информационная система ведения ИЭМК использует в своем контексте стандартизированные справочники, опубликованные в Реестре НСИ (OID 1.2.643.5.1.13.2.1.1.xxx). При отсутствии необходимых справочников в Реестре НСИ, используются внутренние справочники ИЭМК (OID 1.2.643.5.1.13.2.7.1.xxx).

Примерный состав внутренних справочников ИЭМК:

- вид случая заболевания;
- тип направления;
- вид стационарозамещающей помощи;
- уровень конфиденциальности для опекуна/представителя пациента;
- уровень конфиденциальности для пациента;
- система кодирования событий;

- тип диагноза;
- отношение к пациенту;
- система кодирования секций компонентов CDA;
- соответствие клинического и патологоанатомического диагнозов;
- параметр физикального осмотра;
- роль в оказании медицинской помощи;
- тип выдачи препарата;
- тип документа о временной нетрудоспособности;
- тип внешнего документа;
- вид рецепта;
- виды аллергии (непереносимости), аллергены и т.д.
- обстоятельства возникновения заболевания;
- тип внешнего идентификатора;
- способ применения медикамента;
- показания к проведению профилактических прививок;
- причины отказа от проведения процедуры;
- параметры функционального обследования [8].

1.2.2 OpenEHR

Стандарты СЭМД openEHR – открытый стандарт управления, хранения и обмена электронными историями болезни (ЭИБ).

В openEHR, все медицинские данные о пациенте хранятся в течение всей его жизни, данные не зависимы от того какая организация их создала, вся размещенная информация ориентирована на человека. Основной целью openEHR является не обмен данных между EHR-системами, а стандарт сообщений.

Ключевое новшество в openEHR структуре – оставить всю спецификацию клинической информации вне информационной модели, но наиболее важно, обеспечено мощное средство выражения, как клинических данных и медицинских данных пациента, что позволяет записать их таким образом, что

информация может пониматься и обрабатываться требованием. Основой этого стандарта является создание клинических информационных моделей на основе формирования архетипов и шаблонов. Общая схема представлена на рисунке. 1.5 Ядром системы является система базовых типов и способ построения ссылок. Это статическая неизменяемая часть.

Описание семантики описываемых сущностей является одной из важнейших задач представления данных. Для системы, которая изначально должна поддерживать различные, зачастую заранее не формализованные объекты, возможность конструировать описание новых сущностей в процессе работы важно вдвойне.

В стандарте openEHR предлагается технология двухуровневого проектирования:



Рисунок 1.5 Общая схема описания данных стандарта openEHR

- нижний уровень описывается в виде небольшого количества стабильных базовых сущностей (изложены в Эталонной модели), таких как «Композит», «Каталог», «Раздел», «Элемент», различные типы данных;
- верхний уровень описывается в терминах предметной области посредством архетипов и шаблонов, таких как «Консультация врача», «Давление крови» и т.п.

Архетипы отделены от данных и хранятся независимо, и применяются для придания данным определенной структуры, связывая их между собой семантически. Количество архетипов невелико и ими описываются только базовые сущности, заданные в максимально общем виде. Архетипы

используются для анализа данных, поскольку на их основе формируется семантика запросов к объектам. Шаблоны являются средством формирования конкретных объектов предметной области путем комбинирования нескольких архетипов. Шаблоны создаются для решения уже конкретных задач.

Для работы в МИС необходимо обеспечить ведение семантических описаний данных в качестве отдельного сервиса, который должен быть частью МИС. Это требование позволит поддержать работу МИС в условиях автономной локальной инсталляции.[9]

Структура ЭМЗ в терминах стандарта openEHR представлена на Рисунке 1.6



Рисунок 1.6 – Структура ЭМЗ в терминах стандарта openEHR

Представление ЭМК как коллекции объектов ЭМЗ, с точки зрения способа представления данных, наиболее близко к работе со структурированными документами. Вместе с тем, ограничиться только одними документами не удастся, поскольку требуется поддерживать структуру контейнеров документов, а также включать в ЭМК неструктурированные бинарные объекты.[10]

Таким образом, организация хранения ЭМК должна включать в себя все технологии хранения, но ключевой среди них должна быть система представления медицинских документов.

1.2.3 OpenUMS

OpenUMS — концепция построения, ведения, хранения и извлечения электронных медицинских документов, разработанная с целью унификации медицинской информации и облегчения информатизации медицинских учреждений, в соответствии с принятой 8 русской медицинской школе моделью синтеза и анализа медицинских данных и сопряженной с международными стандартами в области хранения и передачи данных медицинского профиля .

Стандарт подразумевает 4 уровня:

- модель базовых понятий;
- модель структуры;
- модель отображения;
- визуальный интерпретатор.

Формализованные данные используются для структурирования части ЭПМЗ 7.1.9 «Формализованные данные, прикрепленные к ЭПМЗ» (ГОСТ Р 52636-2006 Электронная история болезни). Уровни представлены на рис. 1.6.

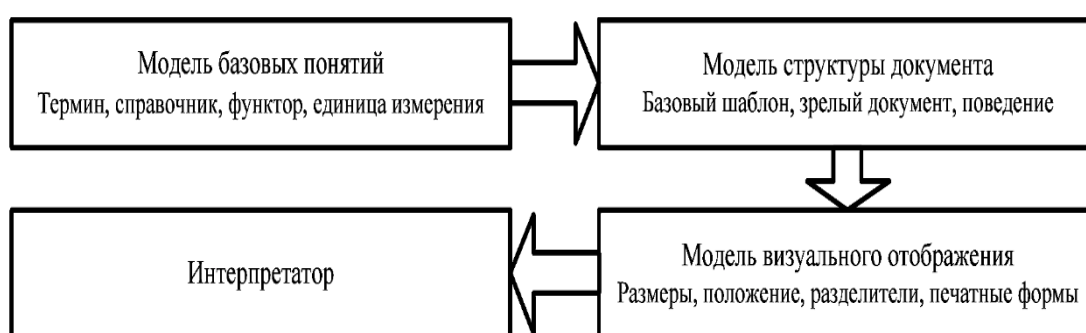


Рисунок 1.6 – Модель построения медицинского документа формата

Одной из особенностей стандарта openUMS является возможность трансформации данных из системы, использующей данный стандарт в международные форматы ISO 13606 и CDA.

Принципиальное отличие openUMS от openEHR состоит в том, что в openUMS ядром системы являются базовые термины, уже связанные с определенной онтологией и типами данных, а в openEHR ядром системы является набор правил и стандартов из которых потом строятся архетипы, описывающие процесс, привязанный к справочным системам. [9]

В отличие от развитых стран мира, где общеприняты определенные электронные форматы документов, в Российской Федерации, традиционный бумажный оборот карточек и медицинской информации. Реже разворачиваются персональные информационные системы, на основе базовых сетевых технологий и доступного свободно распространяемого программного обеспечения. Суть персональных информационных систем сводится к обработке данных из устоявшейся медицинской формы 025/12у.

В настоящее время сфера медицинского обслуживания переводится в электронный вид. Появилось большое разнообразие медицинских информационных систем, но, несмотря на это, медицинские учреждения не могут отойти от ведения документов в бумажном виде, так же существует проблема доступа пациентов к результатам собственных обследований и осмотров узкими специалистами.

Глава 2. Организация ситуационных центров для задач здравоохранения

2.1 Понятие ситуационного центра

Ситуационный центр (СЦ) – это программно-аппаратный комплекс, который обеспечивает для осуществления полного спектра деятельности медицинских учреждений и в режиме реального времени предоставляет информацию о состоянии системы здравоохранения, а также сообщает о проблемных ситуациях в регионах Российской Федерации.

СЦ работает в режиме реального времени и получает информацию о различных показателях, ситуациях, отчетах, входящих в сферу здравоохранения.

Данные, касающиеся состояния системы здравоохранения, можно увидеть в нескольких разрезах: по федеральному округу, субъекту Российской Федерации и муниципальному образованию или району.

У руководителя есть возможность сравнивать показатели нескольких регионов или выбрать регионы по заданным целевым показателям. Такие данные можно посмотреть в форме графика для того, чтобы видеть ситуацию в динамике. Такой график покажет изменения с привязкой ко времени.

По каждому субъекту такие системы имеют большую базу данных. В данных о регионе указываются руководители ключевых региональных министерств и ведомств, мощность учреждений здравоохранения, уровень заболеваемости, число случаев оказания высокотехнологичной помощи и другие. [11]

Для задач СЦ здравоохранения такая справочная информация по регионам может включать реестр медицинских услуг.

На основе внесенных в МИС ЛПУ данных СЦ позволяет осуществлять аналитическую деятельность и облегчать процесс принятия управленческих решений в сфере здравоохранения региона:

- визуализация данных по заболеваемости в регионе с координатной привязкой;
- наблюдение динамики распространения эпидемий;
- выявление и определение границ зон распространения заболеваний;
- формирование штатного состава медицинских учреждений с учетом особенностей уровня заболеваемости в конкретной области региона;
- контроль общего состояния отрасли здравоохранения в режиме реального времени;
- выявление наиболее эффективных методов оздоровления и борьбы с эпидемиями Оценка результатов профилактических мер;
- формирование программы профилактики с учетом склонности к заболеваниям за счет расположения промышленных зон, климатических особенностей и других факторов характерных для определенной территории.

2.1.1 Цель СЦ

Основной целью СЦ является предоставление консолидированной информации для использования и развития организационных возможностей работников медицинской сферы на основе внедрения новейших информационно-аналитических технологий, как для организационного строительства, так и для развития оперативного управления.

2.1.2 Задачи СЦ

Техническое обеспечение СЦ обеспечивает решение следующих задач:

- хранение данных;
- управление проекционным, аудио- видео- оборудованием ИСЦ;
- подсоединение оборудования ИСЦ по локальной сети;

- администрирование и управление оборудованием и информационными системами ИСЦ;
- взаимодействие и передача данных с внешними сетями и информационными системами, и базами данных.

Другими задачами ИСЦ являются:

- реализация оперативного мониторинга протекающих бизнес процессов;
- реализация своевременного обнаружения проблемных ситуаций;
- автоматическое создание отчетных документов и их рассылка заинтересованным лицам; [12]
- обеспечение накопления статистических данных и простой механизм их анализа;
- мониторинг обстановки и предоставления руководству обработанной и классифицированной информации определенного формата;
- повышение эффективности при коллективной работе группой экспертов и аналитиков;
- моделирование развития ситуаций, и оценка различных вариантов решения ситуации;
- комплексное информационное обеспечение руководства компании;
- контроль за качеством технологического и производственного процесса, оперативная обратная связь при выходе ключевых показателей за границы нормы.[13]

СЦ – это комплекс видеоинформационных технологий, коммуникационных средств, математических методов и моделей, обеспечивающих в режиме реального времени обработку информации и принятия коллективом специалистов научно обоснованных решений, направленных на оптимизацию и повышение эффективности работы.

Автоматизированная система управления наивысшего уровня направлена на оптимизацию и повышение эффективности работы за счет применения расчётно-прогнозного характера и предназначена для работы первых лиц отрасли на основе прогнозов развития значений целевых показателей,

полученных путём применения в модели развития отрасли фактических значений показателей. [14]

2.3 Методология формирования ЭМК пациента с использованием сочетания форматов OpenUMS, CDA и BPEL

2.3.1 CDA

Карточная система, построенная по модели CDA взаимодействует с пациентом, медицинским работником, медицинской информационной системой, карточкой пациента и карточкой медицинского работника.

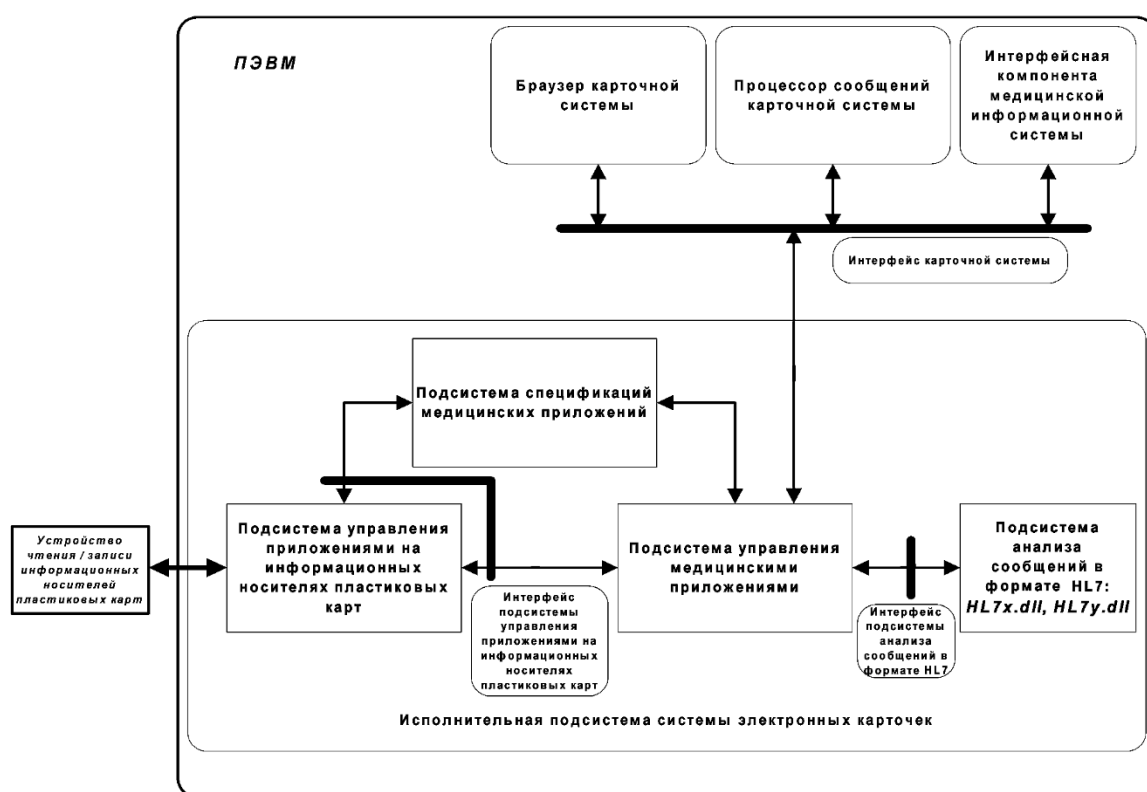


Рисунок 2.1 – Компоненты открытой медицинской карточной системы

Медицинский работник аутентифицируется карточной системой с помощью своей карточки, содержащей в том числе его идентификацию, определяющую права доступа этого медицинского работника к сведениям о пациенте.

Для предоставления пользователям возможности автономной работы с карточками в отсутствие связи с МИС имеется программа просмотра данных на

карточках - браузер. Имея ноутбук с карточной системой и браузером, врач бригады скорой помощи может читать данные с карточки пациента.

Если карточная система обнаружила карточку медицинского работника, удостоверилась в ее работоспособности и успешно провела аутентификацию, то появляется основная экранная форма браузера, предназначенная для просмотра ЭМК. Она заполнится информацией о пациенте, как только карточка пациента будет вставлена в считыватель. [14]

2.3.2 OpenEHR

EHR и PHR — комплексные архивы медицинской информации, содержащие медицинские данные пациентов, поступающие из разных источников. Именно такие архивы могут обеспечить преемственность и неразрывность лечения, осуществляемого разными специалистами в разных медицинских организациях, исключить дублирование дорогостоящих исследований, использование несовместимых лекарств и др.

EHR — архивы, организуемые на региональном или национальном уровне, либо группами медицинских организаций и врачей.

Ядром системы является система базовых типов и способ построения ссылок. Это статическая неизменяемая часть.

2.3.3 OpenUMS

Одной из особенностей стандарта openUMS является возможность трансформации данных из системы, использующей данный стандарт в международные форматы ISO 13606 и CDA.

Принципиальное отличие openUMS от openEHR состоит в том, что в openUMS ядром системы являются базовые термины, уже связанные с определенной онтологией и типами данных, а в openEHR ядром системы является набор правил и стандартов из которых потом строятся архетипы, описывающие процесс, привязанный к справочным системам.

Разработка медицинской карты с применением вышеописанных стандартов OpenEHR и OpenUMS предполагает централизованное хранение данных и доступ к ним из любой точки, где есть выход в интернет.

2.3.4 BPEL Business Process Execution Language

Ключом для решения проблемы интеграции является обеспечение общеупотребительного стандарта. Фокусирование такого стандарта на транзакционные бизнес-процессы требует описания, как именно происходят транзакции, и в каком порядке.

Язык выполнения бизнес-процессов (**Business Process Execution Language – BPEL**) предлагает общеупотребительный язык для бизнес-приложений и сервисов. BPEL является новым стандартом для интеграции гетерогенных приложений и сервисов в транзакционные бизнес-процессы. BPEL является языком потоков технологических процессов и данных (*workflow and process flow language*). Поэтому если имеется несколько стадий, которые нужно объединить в единое целое для формирования бизнес-процесса, то BPEL – это тот язык, который используется для описания того, как и в какой последовательности должны происходить события.

BPEL позволяет описывать бизнес-процесс в терминах некоторой последовательности обращения к *Веб-сервисам*. BPEL, по существу, является скриптовым языком программирования, который поддерживает синхронные и асинхронные взаимодействия, параллельное выполнение и обработку исключений. Программа (приложение), написанное на языке BPEL, является XML документом. [15]

Веб-сервисы представляют собой интерфейсы для доступа к автономным, модульным приложениям. Для того чтобы обратиться к Веб-сервису необходимо послать SOAP сообщение по определенному адресу, которое представляет собой XML документ, при этом не имеет значения каким именно образом формируются эти послания.

Язык BPEL позволяет задавать бизнес-процессы, при этом приложение, написанное на языке BPEL, можно рассматривать как Веб-сервис, и к нему можно обращаться посредством посылки SOAP-сообщений.

Система типов BPEL соответствует языку XML: это язык с выразительными управляющими конструкциями, поддержкой параллельного исполнения, обработкой исключений, поддержкой транзакций, взаимодействия процессов между собой и др. возможностями.

BPEL предоставляет такие механизмы управления вычислительным процессом, такие как присваивание значений переменным, условные операторы, возможность организации циклов, работа с событиями и др.[15]

Описание процессов в BPEL

BPEL-процесс состоит из активностей (activities). Различают базовые (basic) и структурные (structured). Базовые активности не включают в себя другие активности и выполняют такие элементарные действия как прием сообщения от партнера или выполнения элементарных действий с данными. Структурированные активности включают в себя другие активности и обеспечивают реализацию бизнес логики.

Важнейшими базовыми активностями являются активности, ориентированные на получение и отправку сообщений. Это активности типа invoke, receive и reply, которые определяют два способа взаимодействия: асинхронное и синхронное взаимодействия.

Активность типа invoke предполагает одностороннее взаимодействие. Вызывающая сторона посылает сообщение и продолжает функционирование. Получение ответа не предусматривается.

Синхронное взаимодействие реализуется с помощью пары активностей. Сервис реализует активность типа receive – находится в состоянии ожидания

запроса. Получив запрос, сервер формирует ответ, посредством реализации активности reply . До получения ответа вызывающая сторона находится в состоянии ожидания, т.е. находится в заблокированном состоянии.

Имеются и другие базовые активности, такие как активности, ориентированные на присвоение значений переменным (assign), остановка реализации сервиса (terminate), отсутствие действий (empty), также активность типа pick и Event Handler, ориентированные на поддержку работы с событиями.

В рамках BPEL определен достаточно широкий набор структурированных активностей, основными из которых являются следующие:

- задание последовательности выполнения действий (<sequence>);
- цикл (<while>);
- выбор одного из нескольких альтернативных маршрутов (<pick>);
- параллельное выполнение (<flow>);
- обработка ошибок <throw> и <catch>;
- объединение нескольких действий (<scope>);
- и др.

Пример BPEL-документа

В качестве примера рассмотрим бизнес-процесс[16]

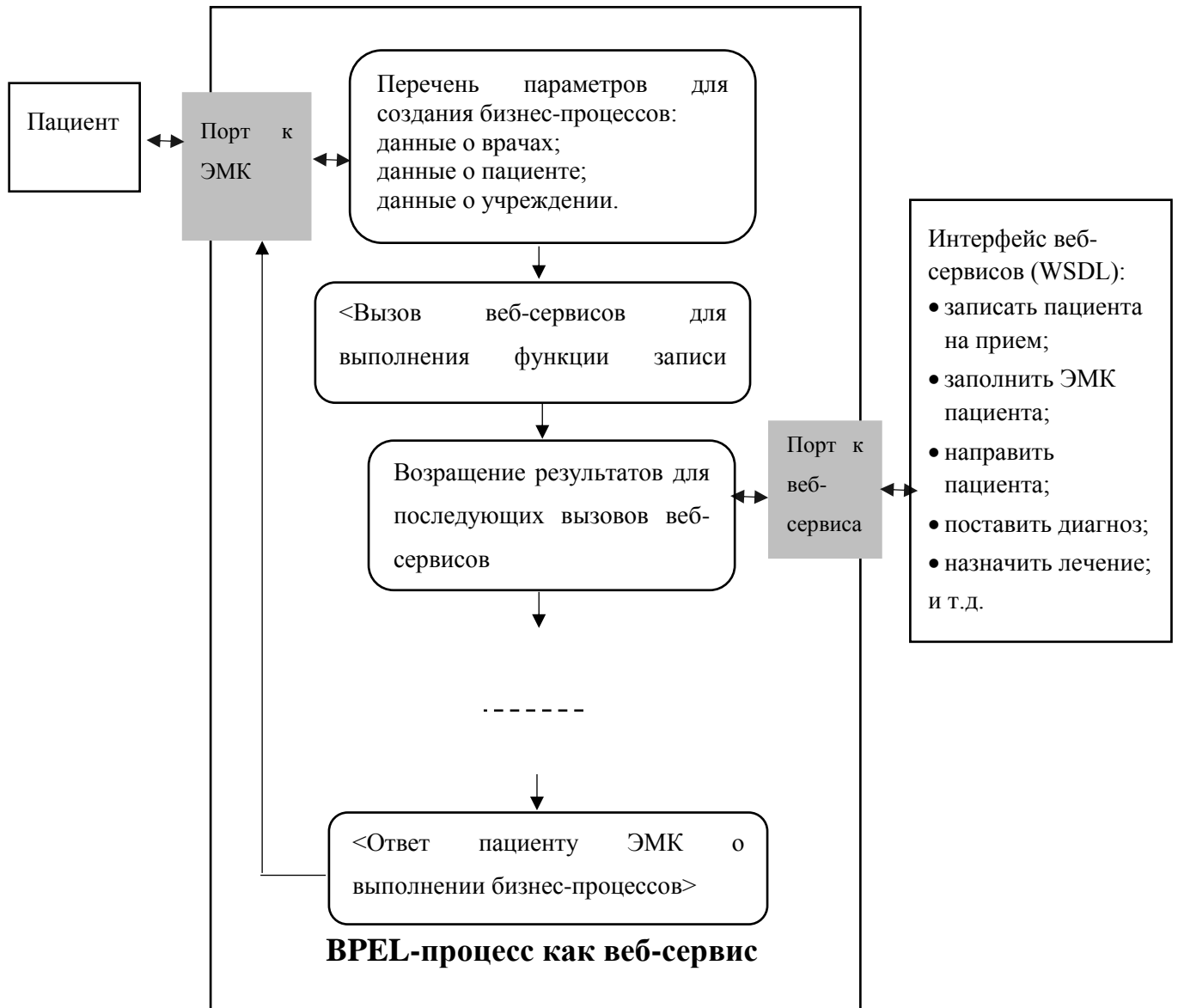


Рисунок 2.2 – Связь между BPEL-процессом и веб-сервисами

Описание процесса «записи пациента на прием» модели обслуживания пациента в текстовом виде [16]

```
<bpel:process name="AuroraProcess" ... >
...
<bpel:partnerLinks>
// Описание всех участников в процессе;
</bpel:partnerLinks>
...
<bpel:variables>
// Перечень всех переменных, которые передаются в процесс;
</bpel:variables>
<bpel:sequence name="Модель обслуживания пациента">
// Описание логики исполнения бизнес-процессов;
<bpel:sequence name="Запись пациента на прием">
<bpel:receive name="Ожидание новой заявки" partnerLink="patient" />
<bpel:assign name="Передача входных переменных для вызова веб-сервиса">
// Описание механизма передачи переменных веб-сервису;
</bpel:assign>
<bpel:invoke name="Вызов внешней функции (веб-сервиса)" ></bpel:invoke>
<bpel:assign validate="no" name="Передача полученного результата функции">
// Описание механизма передачи переменных следующему процессу;
</bpel:process>

</bpel:assign>
<bpel:reply name="Возвращение результата" />
</bpel:sequence>
...
</bpel:process>
```

В ходе работы рассмотрена типовая модель обслуживания пациента, которая описывает основные этапы плана лечения пациента с момента записи на прием до постановки диагноза и назначения лечения. На следующем рисунке показано описание алгоритма.

На рисунке 2.3 в соответствии принципа BPEL выполнения модели процесса медицинской обслуживания пациента обслуживания пациента .

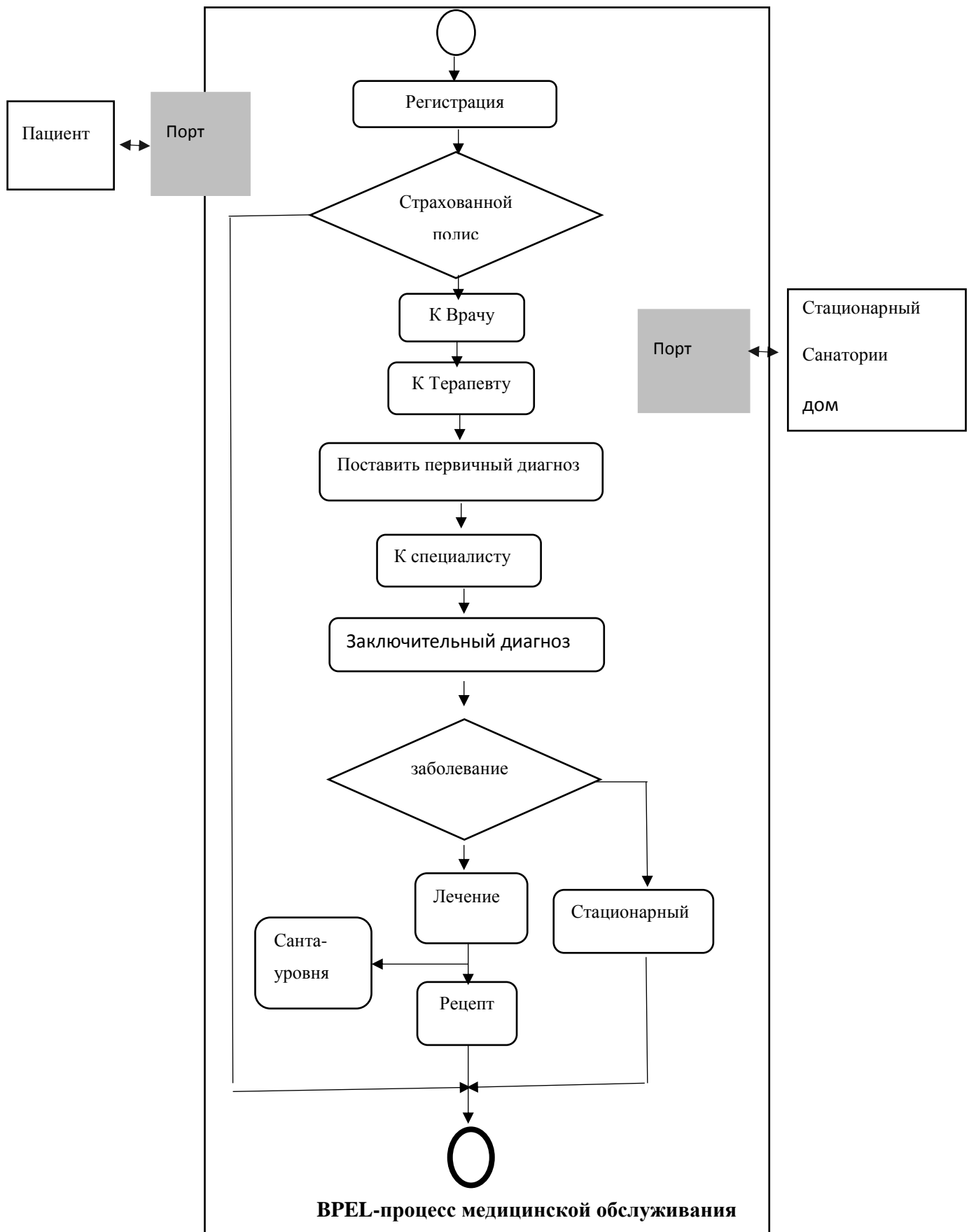


Рисунок 2.3– в соответствии принципа ВРЕL построения процесс медицинской обслуживания

BPMN

(Business Process Model and Notation, нотация и модель бизнес-процессов) — система условных обозначений (нотация) для моделирования бизнес-процессов. Разработана Business Process Management Initiative (BPMI.org) и поддерживается Object Management Group.

BPMN служит для связи между фазой дизайна бизнес-процесса и фазой его исполнения посредством стандартного набора условных обозначений для отображения бизнес-процессов (БП) в виде диаграмм БП и, следовательно, рассчитана как на IT-специалистов, так и на бизнес-пользователей.

Бизнес-пользователями BPMN являются:

- бизнес-аналитики – создают и улучшают процессы,
- технические разработчики – отвечают за реализацию процессов и
- менеджеры – следят за процессами и управляют ими.

Условные обозначения для отображения БП описаны в спецификации BPMN. Также спецификация BPMN определяет, как диаграммы, описывающие БП, могут быть преобразованы в разрабатываемые модели на языке BPEL.

Моделирование БП в BPMN реализовано с помощью диаграмм, содержащих небольшое число базовых, интуитивно понятных пользователю графических элементов, позволяющих строить простейшие диаграммы БП. Элементы подразделяются на четыре категории:

1. Объекты потока управления: события, действия и логические операторы.
2. Соединяющие объекты: поток управления, поток сообщений и ассоциации.
3. Роли: пулы и дорожки.
- 4.Arteфакты: данные, группы и текстовые аннотации.

Разработчик может повысить выразительность модели с помощью создания новых типов объектов потока управления и артефактов. Спецификация это разрешает.

Поскольку единого стандарта для моделирования БП не существует, BPMN могла бы стать инструментом унификации способов представления не только базовых, но и более сложных концепций БП.

Важным ограничением BPMN является то, что она поддерживает лишь набор концепций, необходимый для моделирования БП (например, моделирование организационных структур и моделей данных в BPMN не описывается) и не является схемой информационных потоков, хотя позволяет моделировать потоки данных и потоки сообщений, а также ассоциировать данные с действиями. [17]

Модель бизнес-процесса в BPMN представлена на Рисунке 2.2.

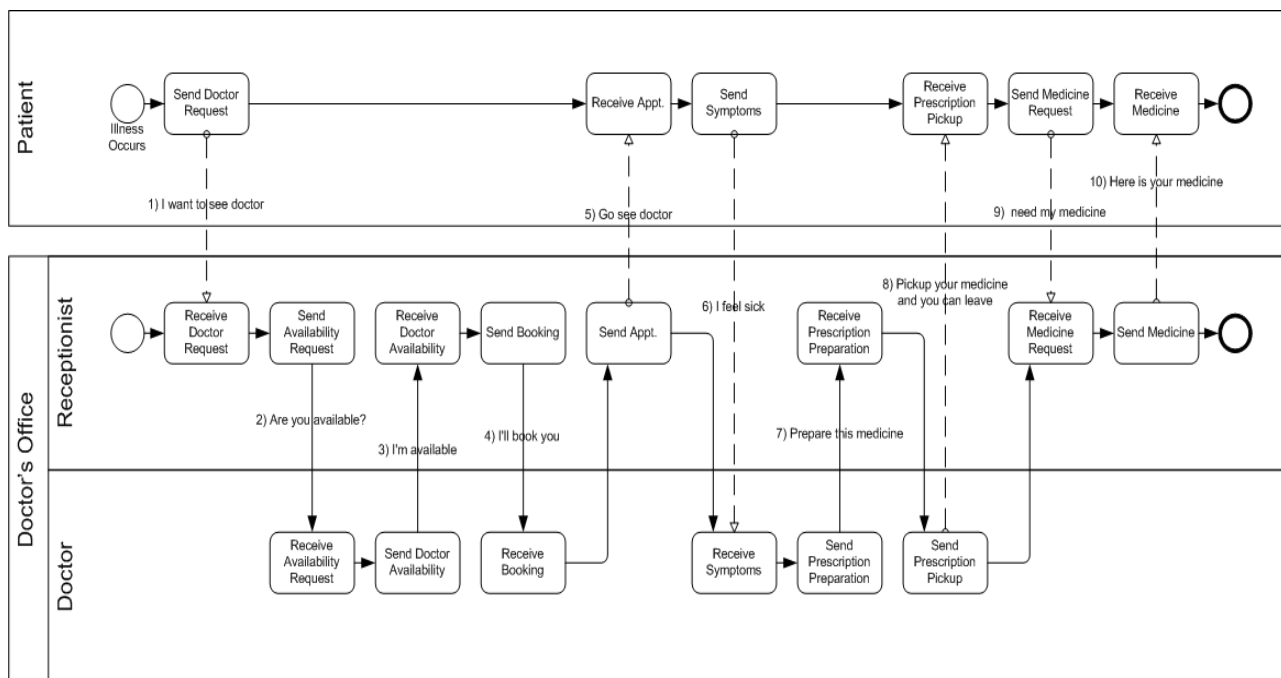


Рисунок 2.4 – Модель бизнес-процесса в BPMN обработки запрос к врачу

2.4 Разработка алгоритмических и программных средств интеграции данных при создании ЭМК пациента

Электронная медицинская карта пациента (ЭМК) – это составной, полиморфный, интерактивный электронный документ, состоящий из записей стандартной структуры документ, содержащих данные различного типа: текст, числа, коды, даты, графики, изображения, аудио, видео.

Требования к ЭМК аналогичны требованиям к традиционной бумажной медицинской карте плюс новые возможности:

- регистрация хронологически последовательных событий процесса ведения пациента (документирование);
- средство общения «врач – врач», «врач – сестра»: преемственность лечения («врач – пациент»);
- возможность записи комментариев для врача, для сестры, для руководителя, для пациента;
- информативность -> полнота, подробность, актуальность, достоверность, точность записей = адекватность, достаточность для принятия клинических решений и оценки действий врача;
- документальность: юридическая значимость (авторизация записей + защита от исправлений) электронная подпись (ЭП) [18].

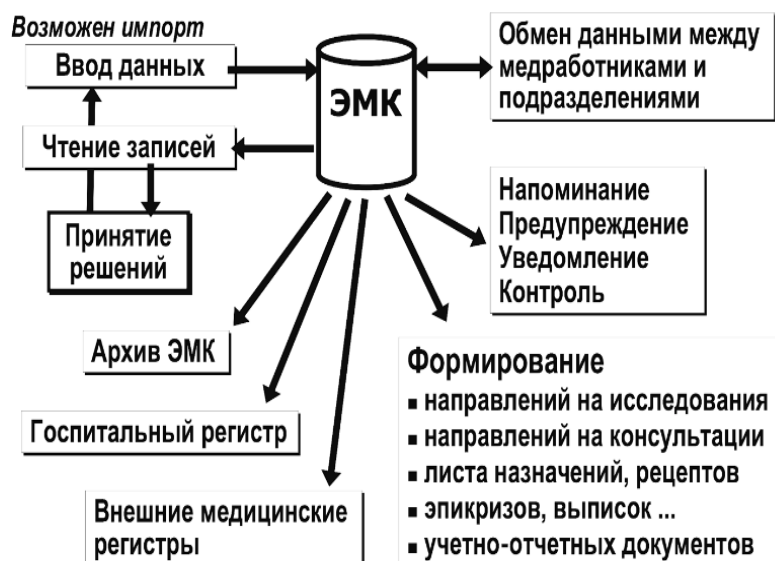


Рисунок 2.5 – Функции ЭМК

Большая часть информации, с которой работают медучреждения, поступает либо через регистратуру, либо создается в регистратуре.

Существуют два типичных сценария госпитализации пациентов: пациент поступает в больницу, отделение скорой помощи или клинику впервые, либо повторно.

Первичная регистрация. В процессе регистрации больные предоставляют медсестре или регистратору целый ряд сведений, к ним относятся:

- имя и адрес — озвучиваются в устной форме и вручную вводятся в электронную форму, или предоставляются в бумажной форме, заполненной и подписанной пациентом или его опекуном;
- страховое свидетельство — обычно это пластиковая карта, с которой снимается копия;
- удостоверение личности с фотографией — часто это пластиковые водительские права, с которых снимается копия.

Пациенты также подписывают целый ряд форм:

- разрешение на использование персональной информации;
- согласие на лечение и условия госпитализации;

- подтверждение требований HIPAA к конфиденциальности;
- предварительное уведомление бенефициара и предварительные инструкции;
- общая медицинская информация для госпитализации.

Фактически, в большинстве больниц и клиник часто существует несколько процессов регистрации пациентов, когда пациент после процедуры общей регистрации переходит в различные лечебные отделения. Таким образом, информация создается не в одной, а в нескольких точках.

Сбор такой информации до сих пор осуществляется с использованием большого количества бумажных документов. Даже при наличии электронных форм, многие формы документов распечатываются и заполняются от руки.

Любые документы, требующие подписи пациента, также необходимо распечатывать. В конечном итоге большое количество информации находится на бумажных носителях.

Как правило, информация дублируется и передается в физической форме — из регистратуры в отделение хирургии, из хирургии в отделение анестезиологии, из анестезиологии в отделение кардиологии и так далее — в буквальном смысле «бумажная волокита».

Очень часто одни и те же данные снова и снова вводятся в различные системы, что зачастую приводит к задержкам, ошибкам и утере документов.

Повторная госпитализация: процедуры ведения активного учета медицинских карт в разных больницах отличаются, но обычно период активного учета длится от 12 до 24 месяцев с момента последних процедур пациента. По окончании этого периода бумажные документы из медицинской карты зачастую перевозят на хранение во внешние архивы. В этом случае при повторной госпитализации пациента необходимо найти его бумажные документы, доставить их и скомбинировать с текущей информацией. Как минимум, данная процедура предполагает большие материальные и временные затраты. [19]

Ввод информации при помощи интеллектуальных электронных форм устраняет необходимость использования бумажных документов в точке происхождения информации. Данные формы устраняют необходимость повторного ввода и могут осуществлять автоматическую валидацию информации с использованием внешних систем и баз данных. Эти формы также предусматривают возможность динамического изменения доступных вариантов после заполнения формы.

Таблица 1.1 Данные медицинского документооборота

Группа	Наполнение
Медицинские/Клинические	Состояние пациента Витальные данные Антропометрические данные Анамнезы Назначение, услуги, рецепты Результаты исследований Исход заболеваний Эпикризы
Административные	Госпитализация, перевод, выписка Исполнители, врач, сестра Согласие, отказ № медкарты документ о временном нетрудоспособности результаты экспертизы источники оплаты стоимость
Паспортные	Идентификационные Демографические Инвалиды, льготы

После передачи данных в депозитарий извлеченные метаданные создают сообщение на языке HL7 (язык интерфейса сферы здравоохранения), направляемое через систему интерфейсов в различные информационные

системы отделений — травматологию, хирургию, радиологию или терапевтические отделения. [20]

Гиперссылки позволяют пользователям данных систем по щелчку мыши переходить из медицинской карты к вспомогательным изображениям документов.

Технологию ввода информации, идеально подходящую для процесса регистрации пациентов, можно использовать в масштабах всей организации в централизованной и распределенной среде, включая ввод:

- специализированной информации с удаленных площадок;
- больших объемов информации из кабинетов врачей больниц и клиник;
- небольших объемов отсканированной информации с настольных компьютеров, сетевых сканеров и многофункциональных периферийных устройств.

Извлеченные данные могут инициировать рабочие процессы, интегрированные с клиническими и оперативными системами. К примеру, данные формы первичного осмотра пациента могут запускать рабочий процесс, передающий информацию об аллергии пациента на определенные медикаменты и/или пищевые продукты в отделение ухода за больными. Рабочий процесс также может обеспечить выполнение надлежащих процедур для исключения отрицательной реакции на пищу или лекарства. [19]

После завершения процедуры или выписки пациента финансовая группа не всегда получает уведомление об этих фактах до момента поступления соответствующих бумажных документов. При больших объемах бумажной документации сотрудникам может потребоваться много времени для соотнесения получателей платежей с плательщиками, что увеличивает период времени до получения медучреждением платежа и затягивает проведение оплаты поставщикам.

Кроме того, в большинстве контрактов со страховыми компаниями указывается максимальный период времени между оказанием услуги и выставлением инвойса. Затягивание сроков выплат может привести к отказу страховой компании от обязательств по выплатам.

Используя электронный ввод финансовой информации посредством сканирования бумажных документов, финансовые департаменты могут быстрее находить и накапливать информацию, необходимую для управления дебиторской и кредиторской задолженностью, поддерживать функции аудита и предоставлять пациентам финансовую информацию по интерактивным каналам.

После того, как цифровые изображения финансовых документов будут сохранены в систему, становится возможным их направление в рамках рабочего процесса департамента, который может включать правила для сроков выполнения и передачи разрешения проблем на более высокий уровень, обеспечивающие исполнение требуемых процедур. [21]

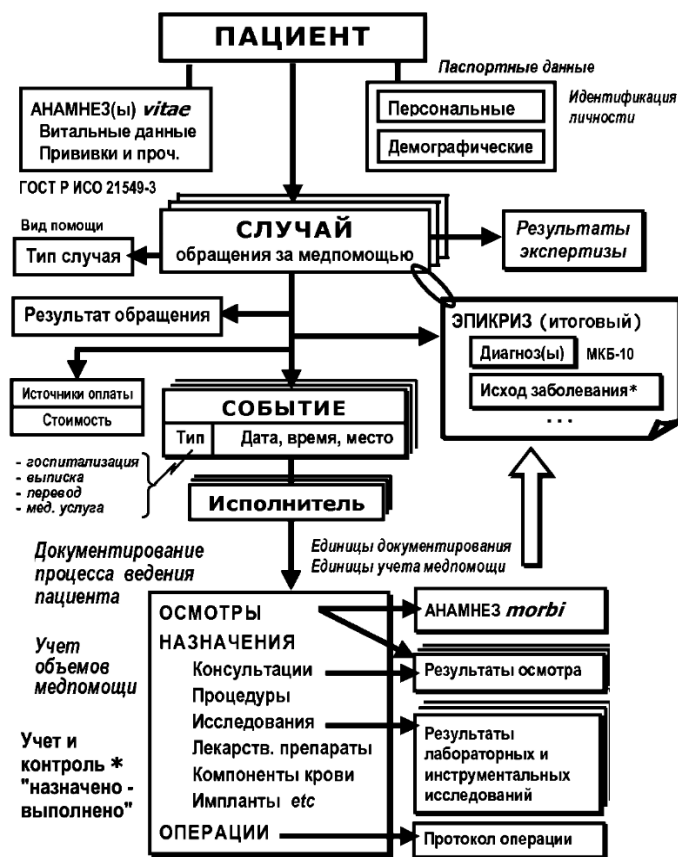


Рисунок 2.6– Алгоритм обращения пациентов ЭМК

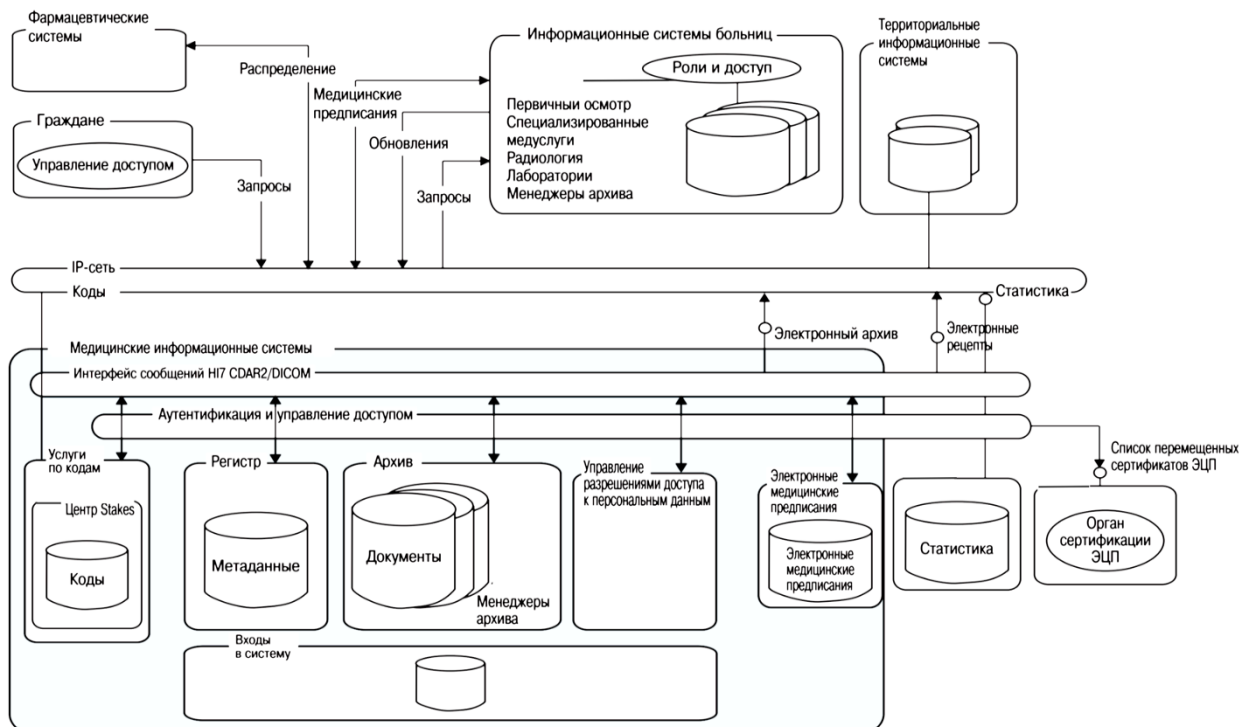


Рисунок 2.7 – Функциональная структура системы в составе системы электронного здравоохранения

2.5 Разработка визуальных компонентов для задач быстрой навигации по ЭМК

2.5.1 Анализ существующих средств быстрой навигации

Традиционно быстрая навигация осуществляется в следующих видах:

- древовидная форма;
- табличная форма;
- окно поиска;
- облако тегов.

Древовидная форма берет свое начало еще с времен активного развития ЭВМ. Впервые была реализована для быстрой навигации в файловой системе ЭВМ. Активно была внедрена в приложениях операционной системе MS-DOS и на сегодняшний день является одним из популярных средств быстрой навигации. Огромное количество сайтов, программ и различных в правой или левой части экрана имеет классификацию списков по категориям с знаком (+). При раскрытии данного списка, знак (+) который меняется на (-) и раскрывается подкатегория с новым древовидным списком.



Рисунок 2.8 – Древовидные формы быстрой навигации

Несмотря на все достоинства, данный способ навигации имеет свои недостатки, а именно: [25]

- строгая систематизация данных;
- длительный поиск нужной категории;
- низкое быстродействие.

Проработка всей архитектуры, на момент заполнения документа и сложные средства автоматизации поиска данных структуры.

За счет того что древовидная форма, зачастую определяется на этапе проектирования, то данные дерева загружаются в оперативную память при загрузке модуля или же подгружаются в процессе раскрытия списков. Это негативно влияет на быстродействие данного способа навигации. В ЭВМ типа IBM 486 процедура автоматической проработки построения дерева каталогов занимала десятки минут.

Табличная форма является привычной, в связи с тем, что, в отличии от древовидной, ее удобно использовать для анализа систематизированной по категориям (ячейкам) данных. Часто практикуется в бумажном документообороте. Данный способ быстрой навигации является одним из самых распространенных и пользуется широкой популярностью, однако тоже имеет свои недостатки:

- очень строгая систематизация данных;
- ограниченность или избыточность информации;
- большое визуальное поле для работы с информацией.

Поисковое окно появилось раньше чем древовидное представление данных. Суть данного способа заключается в том что пользователь вводит в строку поиска определенное ключевое слово в результате чего ему выдаются результаты. Отсутствие строгой систематизации расширяет возможности поиска, однако все равно ограничения касаются конкретного слова[5].



Рисунок 2.9 – Окно поиска

Чтобы конкретизировать поиск применяются различные фильтры, однако они не позволяют полностью избавиться от всех недостатков данного способа быстрой навигации.

2.5.2 Источники данных

На сегодняшний день обилие источников информации весьма ограничено в виду того что ЭМК пациента является, согласно закону «о обработке персональных данных», личным, фактически конфиденциальным документом.

Это обусловлено тем, что разглашение состояния здоровья 3-х лиц способствует изменению отношений внутри коллективов и общества, даже на международном уровне. Всемирная организация борьбы со СПИДом - является ярким тому примером.

Поэтому существуют 3 пути получения исходных медицинских данных:

- автоматическая генерация с использованием существующих стандартов и форм;
- использование тестовых режимов существующих медицинских информационных систем;
- использование реальных данных. В данном случае, программный модуль передается в медицинский статистический центр, где специалисты имеющий доступ к медицинским данным проводят обработку существующей реальной информации.

Результаты обработки передаются диссертанту. Результат содержит набор ключевых слов, что ни коем образом не влияет на распространение конфиденциальной информации.

2.5.3 Использование алгоритма “Облако тегов” для навигации по ЭМК

Предлагаемый способ поиска/навигации является новым и перспективным, но сегодняшнего момента не был реализован в существующих МИС для быстрой навигации и поиска медицинской информации.

Описываемый способ представляет собой визуальное представление текстовых данных с учетом частоты их использования в рассматриваемых документах. Ключевые слова представляют собой отдельные слова, и значение каждого отображается таким размером или цветом шрифта значение которого



Рисунок 2.10 –использования облака тегов для медицинской информации используется чаще.

Этот формат полезен для быстрого восприятия наиболее известных терминов и для определения местоположения термина в алфавитном или случайном порядке.[22]

Особенностью применения для задачи навигации в МИС является возможность ограничить период времени создания документов и перечень

авторов-врачебных специальностей. Удобство такого подхода достигается возможностью применения специальных жестов при использовании мобильных устройств. Последующая фильтрация и доступ к полнотекстовым документам достигается переходом с выбранного тега к множеству документов, на основании которого он был сформирован.

Исходные данные могут быть представлены не только в фиксированном медицинском формате, но и в качестве записей в текстовом файле.

Алгоритм формирования облака тегов представлен на рисунке 2.1

Каждая из ячеек – переход к тому или иному классу, описывающему событие.[23]

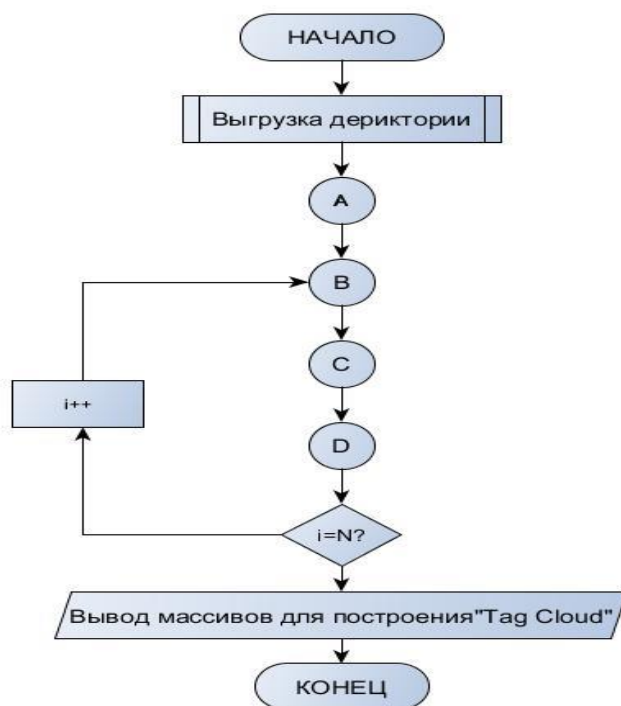


Рисунок 2.11 Общий алгоритм обработки.

загрузка директории подразумевает простую процедуру, копирование или перенос (в крайнем случае, ссылка) на место на диске где находятся все необходимые для обработки медицинские документы читаемого формата.

Блок А служит для того чтобы определить общее количество файлов предназначенных для анализа данных. На его выходе формируется массив fname,

который состоит из индекса файла и его названия, а так же две глобальные переменные N и i предназначенные для дальнейшей работы.

i – номер обрабатываемого файла.

N – общее количество файлов.

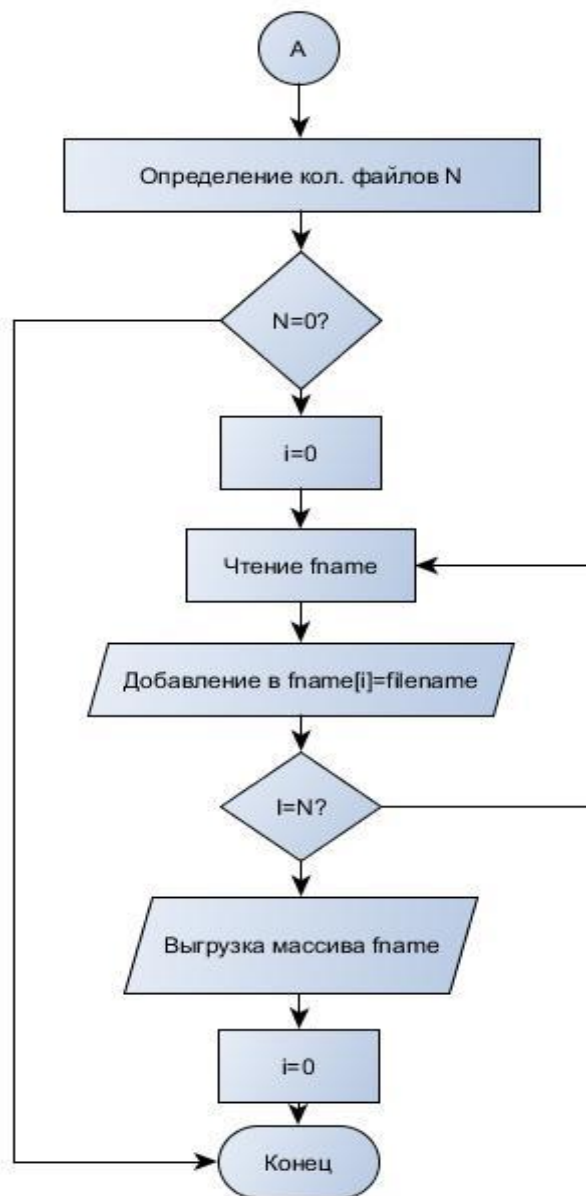


Рисунок 2.12 Общий алгоритм обработки директории.

Это может быть директория с документами текстового формата (txt), документа Microsoft Office (doc, docx, xls), Xml - формата и любого другого, текстосодержащего файла для его дальнейшего анализа и имеющего ценность с точки зрения медицинской информации Допустима так же выгрузка из базы данных SQL – запросом, но для этого требуется написать соответствующий класс.

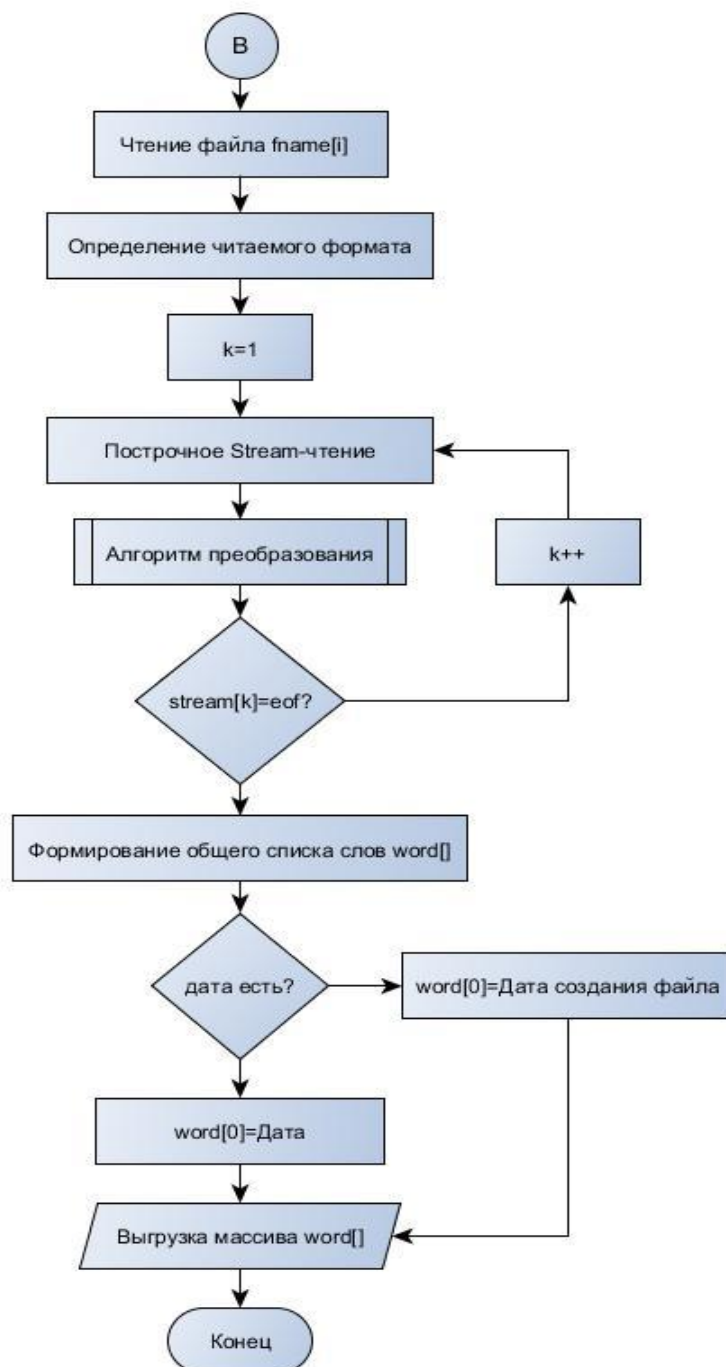


Рисунок 2.13 –Общий алгоритм одного файла.

Учитывая особенности русского языка, сформировать строгий алгоритм поиска тегов и быстрой его сортировки довольно таки проблематично в виду того что, к примеру множество слов имеют нетипичные окончания. В зависимости от рода и падежей, а так же различных морфологических и лексических особенностей (а так же грамматических ошибок) формировать облака тегов весьма проблематично, поэтому необходимо анализировать подобие слов.

За это преобразование отвечает алгоритм «С»

Тут рассматривается 3 варианта классификации

- подбор типичных окончаний в зависимости от падежа;
- корреляция слов;
- использование словаря.

Как было рассмотрено ранее, удобно взять перечень общих ключевых слов, и на их основании обрабатывать исходный текст.

В том случае, когда идет распределение облака тегов на категории органов и возрастные критерии, то формирование облака тегов по принципу сравнения с исходными данными во многом способствует оптимизации работы. Существует так же массив слов, к которым относятся предлоги, местоимения, теги и другие символы, которые уместны только на этапе конструкции файла и предложений и его смысловой значимости, а не в формировании облака тегов.

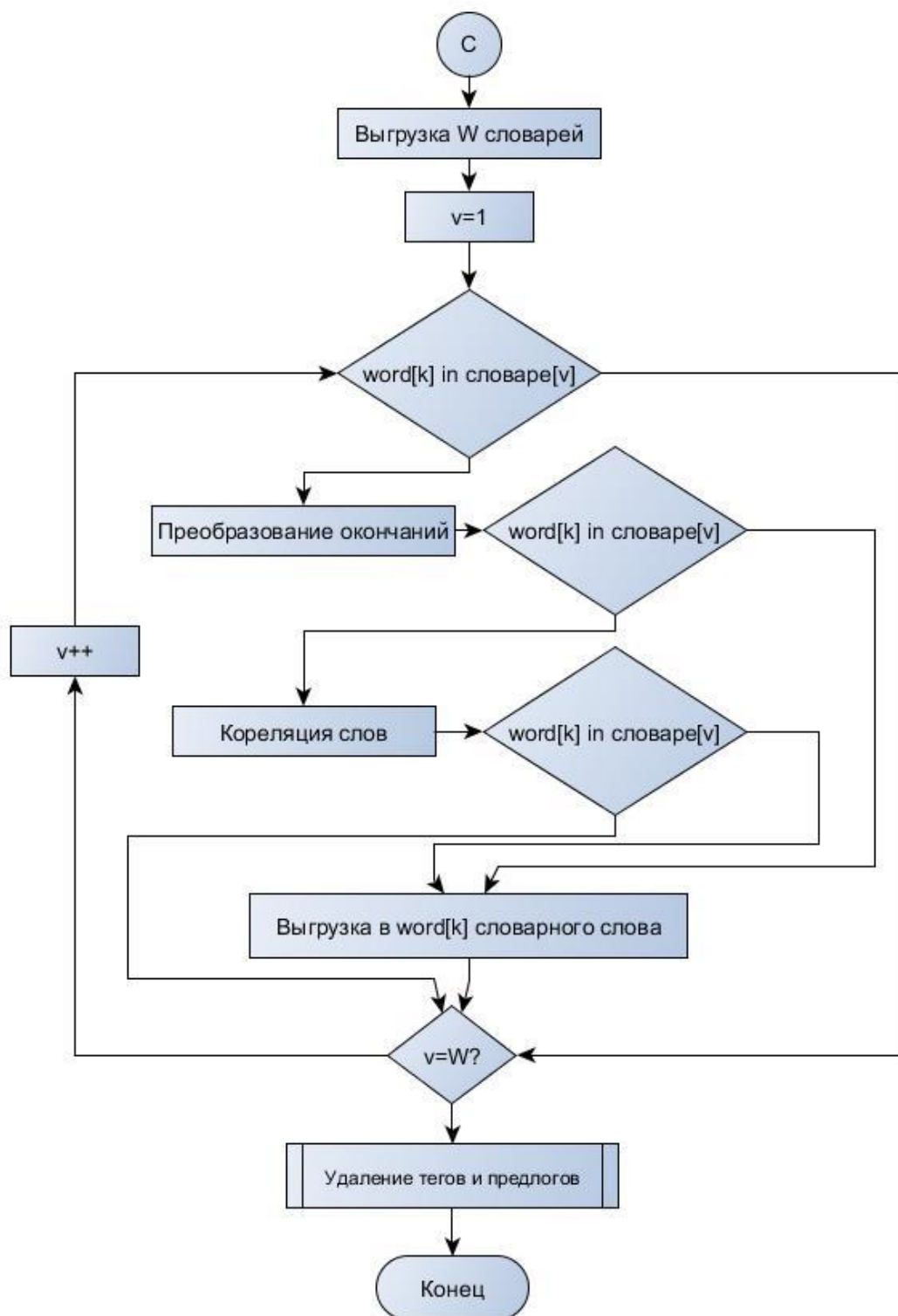


Рисунок 2.14 – Алгоритм обработки словарем.

На последнем этапе блок схемы алгоритма «С» проводится разделение слов в предложении. Предполагается, что разделительными знаками могут быть

знаки «пробела», «запятой», «точки», «двоеточия», «точки с запятой», «дефиса», «вопросительного и восклицательного знака» и «окончания строки».

В результате формируется модифицированный массив `word[k]` который поступает на вход алгоритмического блока D.

Блок D заполняет 4 выходных массива

`Tags[]` – перечень всех встречающихся тегов (согласно словарю или аналитических данных)

И четыре массива одинаковой размерности, с эквивалентными индексами:

`Cloud[]` – `Int` массив обозначающий ссылку на повторяемый тег в файле

`Freq[]` – `Int` массив обозначающий частоту повторения тегового слова в файле

`IdFile[]` – ссылка на ячейку файла из массива `fname`

`DateTag[]` – дата создания записи (файла)

Например 1-й файл содержит 5 раз упоминаний слова температура, 3 раза слова грипп, дата файла 11.12.10

Второй файл содержит 3 раза упоминаний слова температура 1 раз упоминания слова бронхит, дата файла 10.02.11 .

Следовательно выходные массивы будут следующие:

`Tags[]` = {температура , грипп, бронхит }

`Cloud[]` – {1,2,1,3}

`Freq[]` – {5,3,3,1}

`IdFile[]` – {1,1,2,2}

`DateTag[]` – {11.12.10, 11.12.10, 10.02.11, 10.02.11}

Поле формирования массивов считается общее количество повторяемости тегов за определенное время, фиксированное в массиве `DateTag[]`, что способствует ускорению навигации.

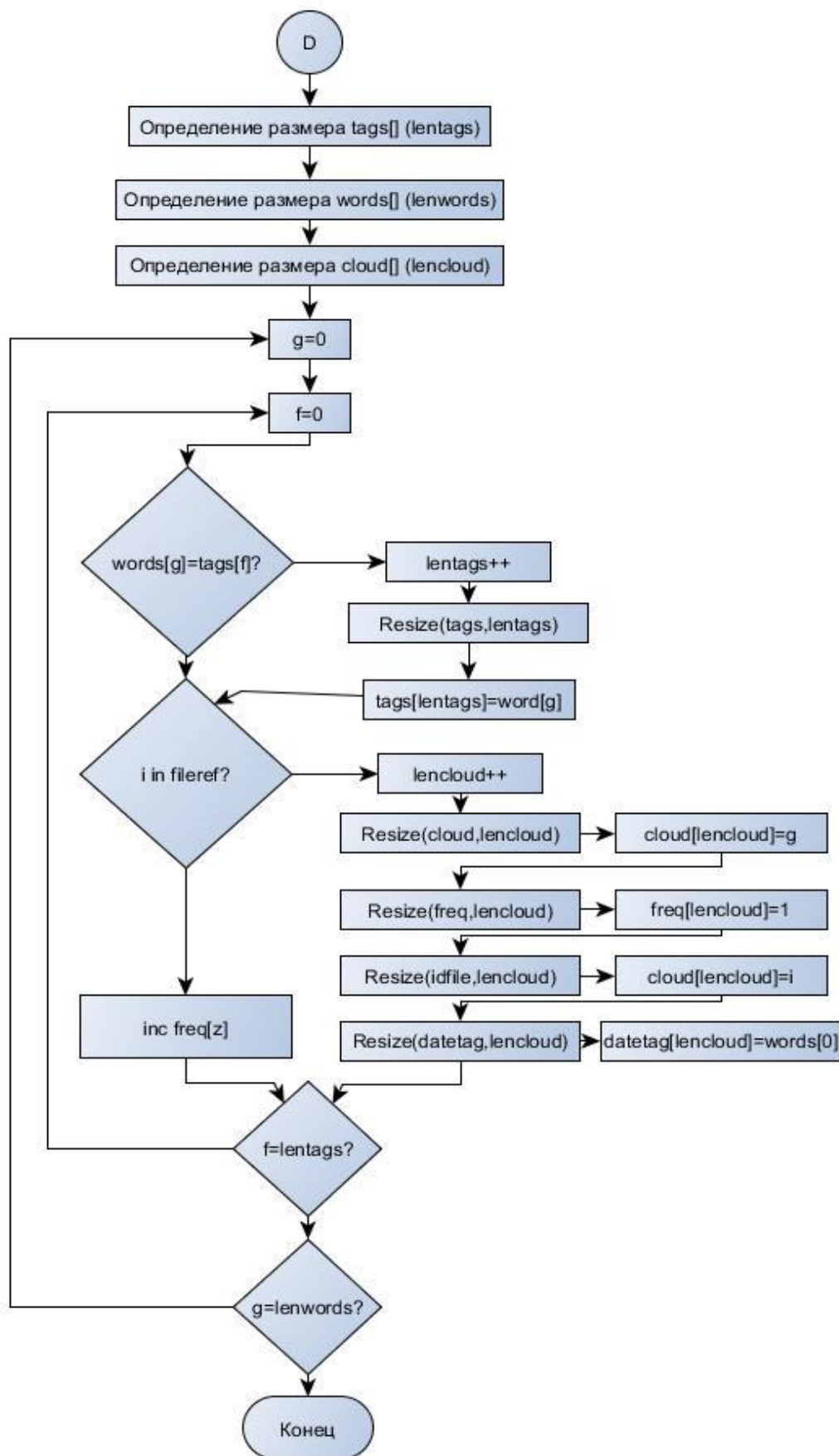


Рисунок 2.15 – Алгоритм формирования выходных массивов

Все эти особенности будут включены в дальнейшее проектирования магистерской диссертационной работы, при формировании облака тегов при навигации по текстовому файлу.

На этапе разработки предлагается изначально сгенерировать 1000 карточек со случайными значениями, которые затем можно пропустить через данный алгоритм.

Выходные массивы позволяют сформировать облако тегов и дочерние объекты или страницы, которые позволят ссылаться на конкретный файл из сформированного тега. Перспективой развития данного алгоритма, является распределения тегов по принадлежности к той или иной системе человека

Глава 3. Разработка программных решений для задач формирования и мониторинга показателей здоровья населения региона

3.1 Анализ и прогнозирование показателей здоровья населения и среды обитания человека

Многочисленные исследования показали, что факторами, обуславливающими здоровье, являются:

- биологические (наследственность, тип высшей нервной деятельности, конституция, темперамент и т. п.);
- природные;
- состояние окружающей среды;
- социально-экономические;
- уровень развития здравоохранения.

Для прогнозирования здоровья населения определенного региона особенно важно знать информацию о природных, социально-экономических факторах, о состоянии окружающей среды и уровне развития здравоохранения в регионе. Рассмотрим эти факторы поподробнее.

К природным факторам относятся: климат, ландшафт, флора, фауна и другие.

Для Томской области характерен континентальный и резко-континентальный климат с жарким летом и холодной зимой (например, для Томска минимальная зарегистрированная температура $-55\text{ }^{\circ}\text{C}$, а максимальная $+37,7\text{ }^{\circ}\text{C}$). Флору и фауну на территории Томской области представляют около 1000 видов высших растений и около 2000 видов животных, большая часть из которых представлена насекомыми. Это определяет некоторые заболевания, характерные для Томской области из-за большого количества клещей (болезнь Лайма, энцефалит).

Состояние окружающей среды также оказывает большое влияние на здоровье человека. Так, Томск характеризуется не очень благоприятной

экологической обстановкой. По данным комитета по охране окружающей среды Томска и Томской области на 2007 год, атмосферный воздух загрязнён практически над всей территорией города. В большей его части, расположенной вокруг городских лесов и вдалеке от промышленных предприятий, отмечается повышенный уровень загрязнения. Содержание в атмосфере оксида углерода также превышает предельно-допустимые нормы. В местах скопления автомобилей — перекрёстках — концентрация оксида превышает предельно допустимые значения в десятки раз. В Томске расположены 4 промышленных предприятия — Томская ТЭЦ-3, Томская ГРЭС-2, ЗАО «Метанол» и ОАО «Томский нефтехимический завод», которые давали 13,5 тыс. тонн выбросов в год. Качество воды на многих водоёмах, популярных как места купания, зачастую не соответствует санитарным нормам.[24]

Социально-экономический фактор и уровень здравоохранения в Томской области характеризуется следующим: в Томске довольно высокий уровень образованности людей (что влияет на их понимание проблем со здоровьем, своевременное обращение к специалистам и другое). Также количество медицинских учреждений и уровень специалистов в них тоже довольно высокий. Но в небольших селах Томской области ситуация намного хуже.

Прогноз состояния здоровья населения в зависимости от этих факторов очень сложен и должен учитывать каждый из этих факторов. Так, например, простудные заболевания характерны для весеннего и осеннего сезонов, когда температура воздуха понижается. Прогноз таких заболеваний, как болезнь Лайма и клещевой энцефалит характеризуются сезонностью (весна-лето). Следовательно, можно прогнозировать показатели здоровья населения в зависимости от описанных выше факторов при помощи регрессионного анализа или же при анализе временных рядов и выделении сезонных особенностей.[25]

3.2 Прогнозирование потребности в медицинской помощи на основании данных об уровне и характере заболеваемости по различным нозологическим формам, группам населения, территориальным и другим признакам

Был проведен анализ различных методов и алгоритмов Data mining для определения их применимости к решению задачи прогнозирования числа людей с различными заболеваниями по регионам в будущем. Были рассмотрены следующие методы: линейная регрессия, деревья принятия решений, метод k-ближайших соседей, нейронные сети, метод опорных векторов. В результате для решения задачи прогнозирования были выбраны следующие методы прогнозирования данных [26].

3.2.1 Линейная регрессия (частный случай анализа временных рядов)

Для прогноза количества людей с определенным диагнозом в будущем можно построить модель линейной регрессии по существующим значениям.

Математическое уравнение, которое оценивает линию простой (парной) линейной регрессии:

$$Y=a+bx.$$

x называется независимой переменной или предиктором.

Y – зависимая переменная или переменная отклика. Это значение, которое мы ожидаем для y (в среднем), если мы знаем величину x , т.е. это «предсказанное значение y »

a – свободный член (пересечение) линии оценки; это значение Y , когда $x=0$ (Рис.1).

b – угловой коэффициент или градиент оценённой линии; она представляет собой величину, на которую Y увеличивается в среднем, если мы увеличиваем x на одну единицу.

а и b называют коэффициентами регрессии оценённой линии, хотя этот термин часто используют только для b.

В нашем случае Y – прогнозируемое число людей с определенным диагнозом, x – дата прогноза. [27]

Для вычисления параметров a, b нужно использовать данные вида:

Дата	Количество человек с данным диагнозом
------	---------------------------------------

Можно разделить всех больных с исследуемым диагнозом по регионам (адресу) и провести аналогичный анализ для каждого региона.

Был проведен регрессионный анализ для прогнозирования числа больных артериальной гипертензией (АГ) по статистическим медицинским данным Томской области. Для построения прогноза выявления больных АГ использовалась зависимость влияние штатной укомплектованности медучреждений врачами первичного звена на количество выявленных случаев заболевания АГ. следовалась зависимость: «Заболеваемость АГ всего населения области (на 100 тыс.чел.)» – выходная переменная Y от фактора X – «Обеспеченность терапевтами (на 10 тыс.чел.)». Проверка значимости и адекватности модели выполнена для уровня значимости $\alpha=0,05$. Факторы, связанные с геномом, образом жизни и экологической ситуацией, отнесены к случайной составляющей ϵ и не учитываются. исходными данными для анализа служат показатели отклика и фактора Томской области.

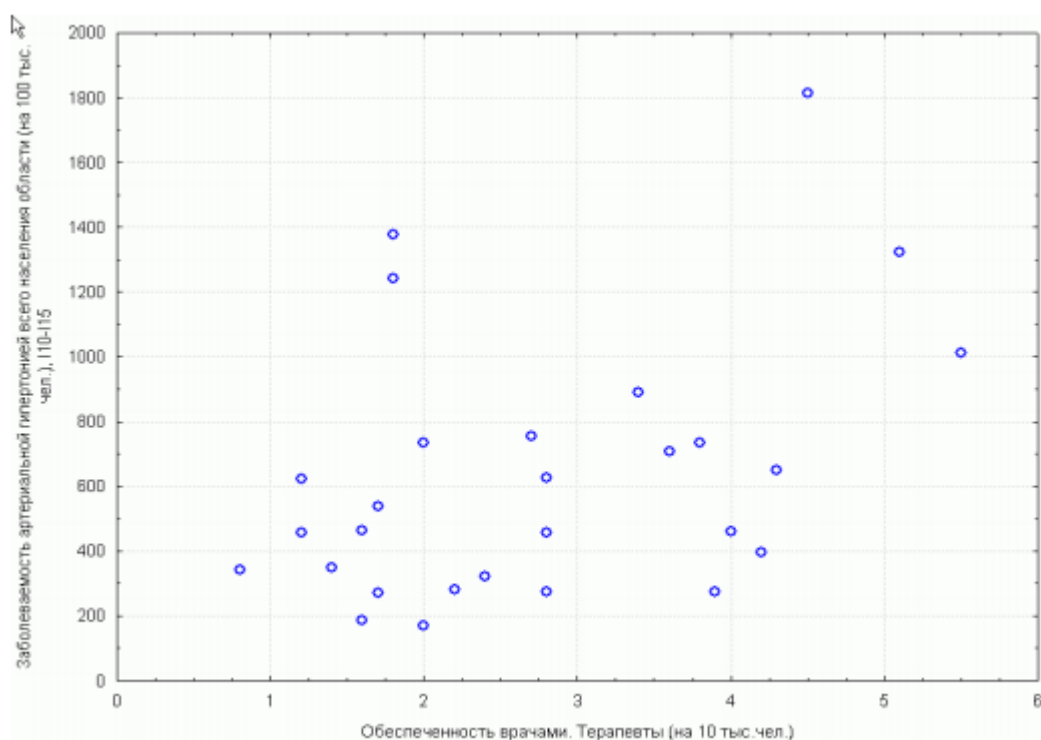


Рисунок 3.1– Распределение медицинских учреждений по обеспеченности терапевтами и количеству пациентов с артериальной гипертонией

[11] После строится корреляционная матрица отклика и фактора, которая свидетельствует, что наибольший коэффициент корреляции с Y имеет переменная X – обеспеченность врачами (0,413657).

Correlations of Vectors in Design Matrix X (Выборка(03)-2010)
Correlation matrix for the vectors
in the design matrix X

Effect	Col. 2 Обеспеченность врачами. Терапевты (на 10 тыс.чел.)	Col. 3 Заболеваемость острым инфарктом миокарда всего населения области (на 100 тыс. чел.), I21	Col. 4 Заболеваемость инсультами всего населения области (на 100 тыс. чел.), I60-I64	Col. 5 Смертность от острого инфаркта миокарда среди всего населения области (на 100 тыс. чел.), I21	Col. 6 Смертность от мозгового инсульта среди всего населения области (на 100 тыс. чел.), I60-I64	Col. 7 Заболеваемость артериальной гипертонией всего населения области (на 100 тыс.чел.), I10-I15
Intercept						
Обеспеченность врачами. Терапевты (на 10 тыс.чел.)	1,000000	0,225032	0,017760	-0,256067	0,025353	0,413657
Заболеваемость острым инфарктом миокарда всего населения области (на 100 тыс.чел.)	0,225032	1,000000	0,299255	0,206834	-0,112680	0,396039
Заболеваемость инсультами всего населения области (на 100 тыс.чел.), I60-I64	0,017760	0,299255	1,000000	0,220558	0,451370	0,360208
Смертность от острого инфаркта миокарда среди всего населения области (на 100 тыс.чел.)	-0,256067	0,206834	0,220558	1,000000	0,166654	-0,083257
Смертность от мозгового инсульта среди всего населения области (на 100 тыс.чел.)	0,025353	-0,112680	0,451370	0,166654	1,000000	0,026704
Заболеваемость артериальной гипертонией всего населения области (на 100 тыс.чел.)	0,413657	0,396039	0,360208	-0,083257	0,026704	1,000000

Рисунок 3.2 – Корреляционная матрица

Полученные коэффициенты линейной регрессии используются для построения ее модели, чтобы определить зависимости заболеваемости АГ от количества специалистов в сфере медицины.

Регрессионная модель прогноза для показателя, отобранного на уровне значимости $\alpha=0,05$, имеет следующий вид:

$$Y = 275,2280 + 130,4859 X.$$

Для проверки гипотезы о незначимости регрессионной модели, используется дисперсионный анализ, который подтвердил, что уравнение регрессии значимо.

Уравнение линейной регрессии может быть использовано для прогноза изменения выходной переменной.

3.2.2 Нейронные сети

Нейронная сеть – математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма. В настоящее время нейронные сети могут быть применены для решения различных задач, например классификация, кластеризация или прогнозирование, поэтому они были выбраны для прогнозирования временных рядов изменения числа людей с определенными заболеваниями.[28]

Нейронные сети способны делать прогнозы благодаря их способности к выделению скрытых зависимостей между входными и выходными данными. Для решения определенных задач нейронную сеть обучают на обучающей выборке, после чего обученная сеть способна предсказывать выходные значения на основе предыдущих значений. Единственной особенностью применения нейронных сетей для задач прогнозирования является наличие этой самой зависимости между предыдущими изменениями и будущими. Например, имеет смысл прогнозирование котировок акций по историческим данным (выявление периодичности, всплесков и упадков, но прогнозирование результатов лотерей

не будет иметь особых результатов, так как данные в лотереях имеют случайных характер.

Нейронные сети можно разделить на сети, требующие обучения с учителем (перцептроны) и сети со смешанным обучением (сети радиально-базисных функций). [29]

В качестве примера рассмотрим использование нейронных сетей для медицинского прогноза нахождения пациента в реанимации.

Существует проблема нехватки реанимационных палат в медицинских учреждениях, в связи с чем для кардиохирургических отделений характерны длинные очереди. При этом увеличение числа реанимационных палат требует очень больших денежных вложений.

Поэтому целесообразно более эффективно использовать имеющиеся палаты. Есть категория больных с тяжелым состоянием здоровья, которым требуется длительное пребывание в реанимационной палате (более двух суток). В связи с этим хирурги не смогут проводить некоторые операции, так как пациентов после операции некуда будет класть. Поэтому тяжелых больных следует оперировать перед выходными или праздниками, когда операции не проводятся (хирурги отдыхают, а тяжелобольные смогут восстановиться в реанимации). Следовательно, людей, имеющих лучшее состояние здоровья и несерьезные операции, можно прооперировать в начале рабочей недели, так как они будут нуждаться в реанимационной палате непродолжительное время. Тогда койки в реанимации будут освобождаться быстрее, и количество прооперированных возрастет.

Д. Ту и М. Гуэррир (университет в Торонто) предложили использовать нейронные сети для прогнозирования состояния пациента после операции, что позволило бы более оптимально использовать реанимационные палаты. В качестве исходных данных для нейронной сети использовались известные в предоперационный период сведения о пациенте.

В качестве архитектуры нейронной сети был выбран двухслойный перцептрон, который разделял больных на три группы риска. При этом учитывались такие факторы как: возраст, пол, функциональное состояние левого желудочка, степень сложности предстоящей операции и наличие сопутствующих заболеваний. Результат работы нейронной сети был следующим: только 16,3% пациентов, отнесенных к группе малого риска задержки в реанимации, провели в ней более двух дней, [а больше 60% пациентов из группы повышенного риска провели в реанимационных палатах много времени. [30]

3.2.3 Кластерный анализ

Одним из подходов Data mining является кластерный анализ (или обучение без учителя). В таком случае исходная совокупность данных о здравоохранении по заданным признакам (количество больных с определенным диагнозом, количество врачей и другие) разбивается сгруппированы в кластеры. В дальнейшем результаты кластеризации интерпретируются.

Для кластерного анализа выделим следующий набор признаков для медицинских учреждений Томской области:

- расходы на здравоохранение из всех источников финансирования. Сумма на 1 чел. (руб.);
- обеспеченность врачами (на 10 тыс.чел.);
- обеспеченность средним медицинским персоналом (на 10 тыс.чел.);
- заболеваемость психическими расстройствами (на 100 тыс. населения).

На основе нашего исследования в качестве метрики для определения расстояния в пространстве признаков между исследуемыми объектами была выбрана евклидова метрика:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}.$$

Для кластерного анализа могут быть использованы различные алгоритмы, в зависимости от особенностей исходных данных, например, K-средних, иерархический алгоритм или DBSCAN.

Был выбран иерархический алгоритм кластеризации для медицинских данных (он позволяет заранее не определять число конечных кластеров, а оценить наиболее подходящее разбиение по дендрограмме). Так, в результате была получена древовидная структура (дендрограмма), по которой определяются конечные кластеры. Предварительно была проведена нормировка – преобразование признаков к безразмерным величинам. Это особенно важно, если признаки представлены в разных единицах измерения.

Результат распределения медицинских учреждений по Томской области в зависимости выделенных признаков изображен ниже.

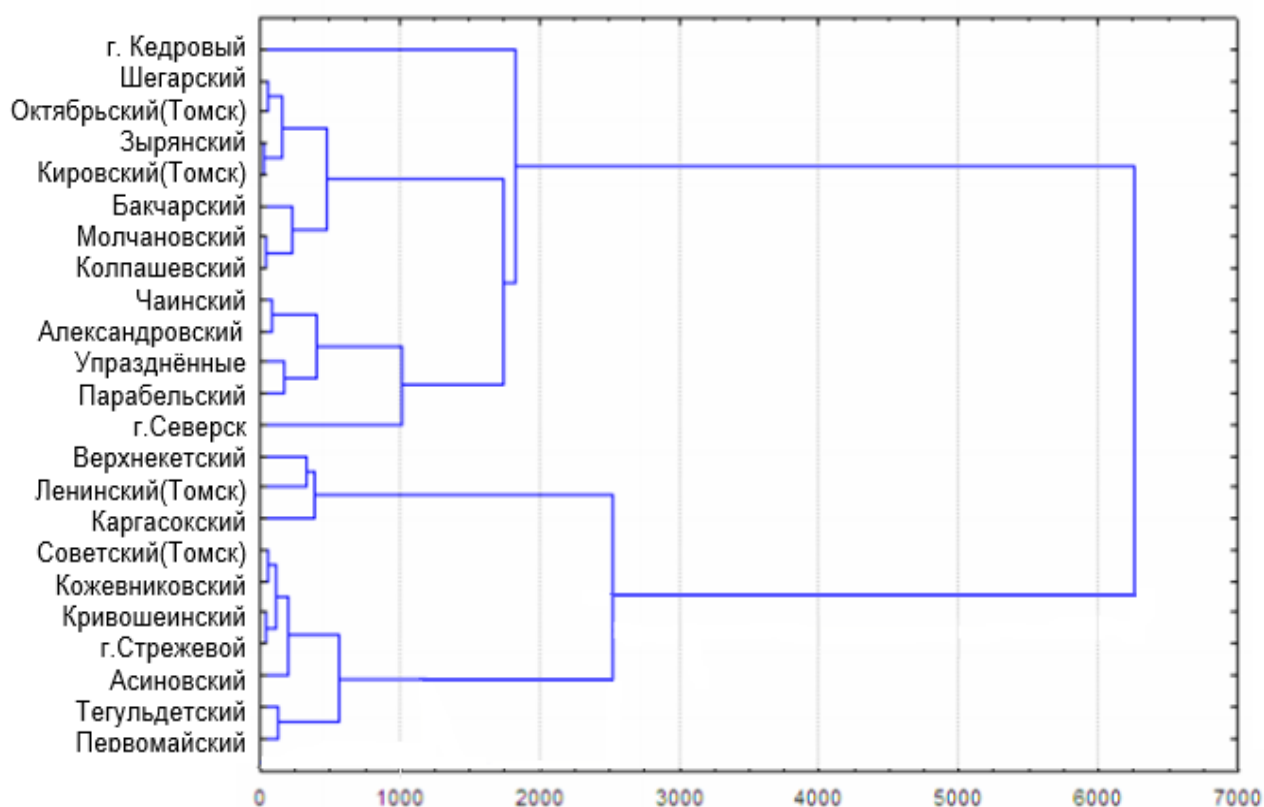


Рисунок 3.3 – Дендрограмма разбиения муниципальных образований Томской области на кластеры

Таблица 3.1 – Результаты кластерного анализа муниципальных образований Томской области

Муниципальные образования Томской области	Кластер	Муниципальные образования Томской области	Кластер
г.Кедровый	1	Молчановский	1
г.Северск	1	Ленинский (Томск)	2
Александровский	1	Верхнекетский	2
Шегарский	1	Каргасокский	2
Кировский (Томск)	1	Советский (Томск)	3
Октябрьский (Томск)	1	Первомайский	3
Зырянский	1	Кожевниковский	3
Парабельский	1	Асиновский	3
Колпашевский	1	Кривошеинский	3
Бакчарский	1	Тегульдетский	3
Упразднённые	1	г.Стрежевой	3
Чаинский	1		

Глава 4. Разработка средств интерпретации данных мониторинга

4.1. Разработка эффективных средств интерпретации данных ЭМК

В исследовательской части для разработки средств интерпретации данных была разработана модель, посредством которой создавались тестовые выборки, обеспечивающие возможность получения произвольного набора тестовых данных в формате CDA за произвольный период времени, что упростило решения задачи по отладке предлагаемых алгоритмов.

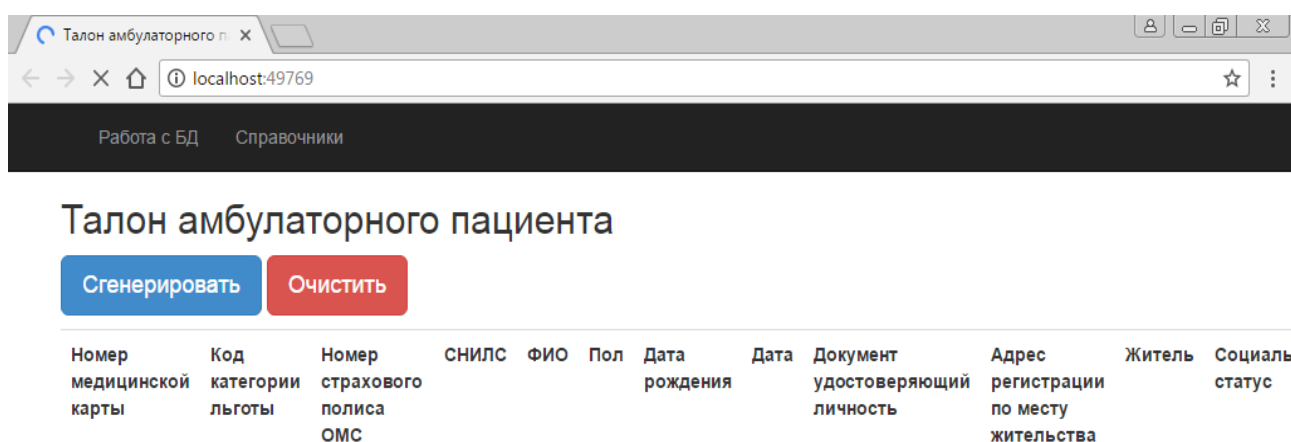


Рисунок 4.1 – главное окно программы

Для решения задачи создания тестовой выборки была обеспечена возможность работы с НСИ, по набору справочников, соответствующей учетной форме 025/12у.

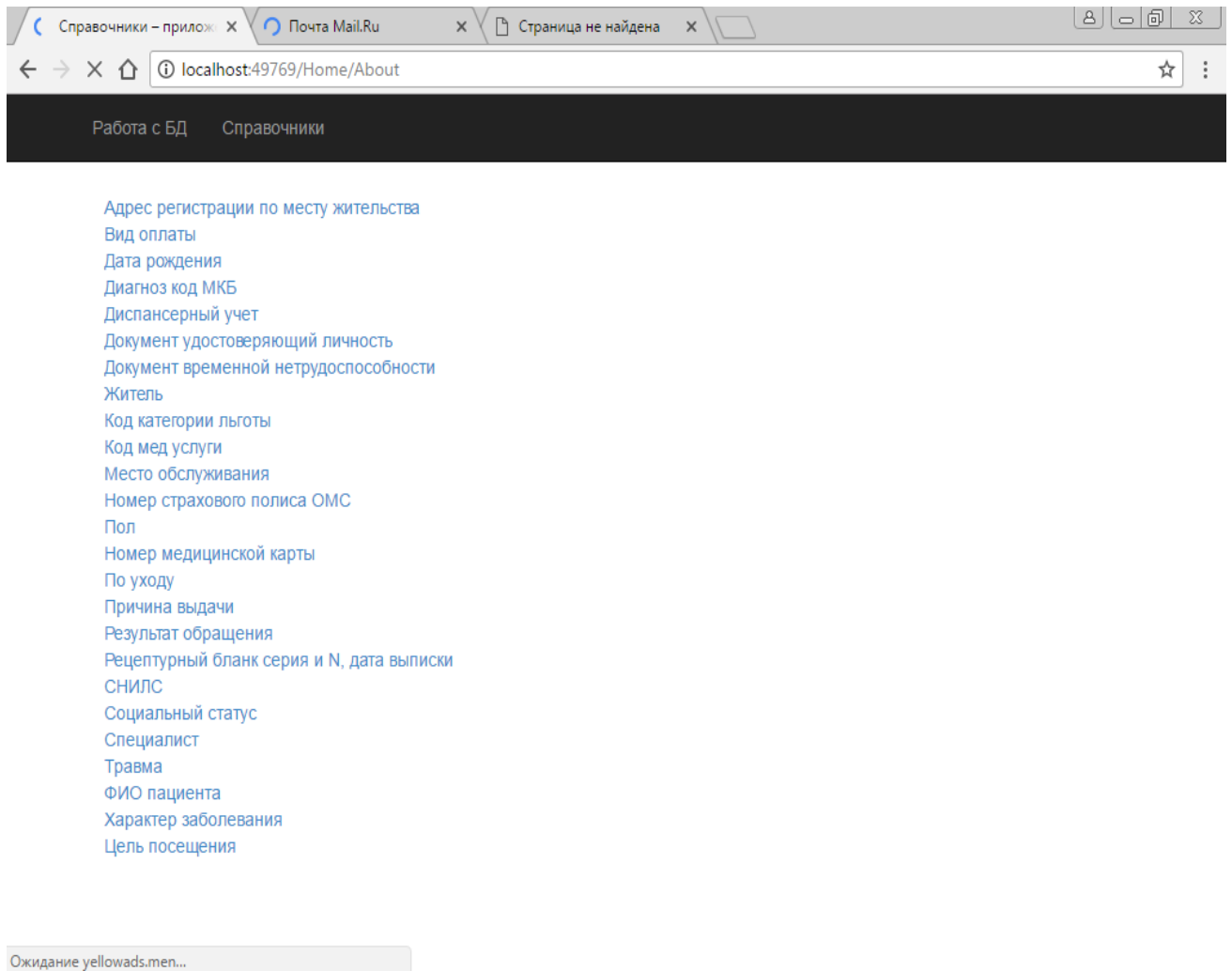


Рисунок 4.2 – Списки справочников

На следующем рисунке пример справочника «диагноз» МКБ10

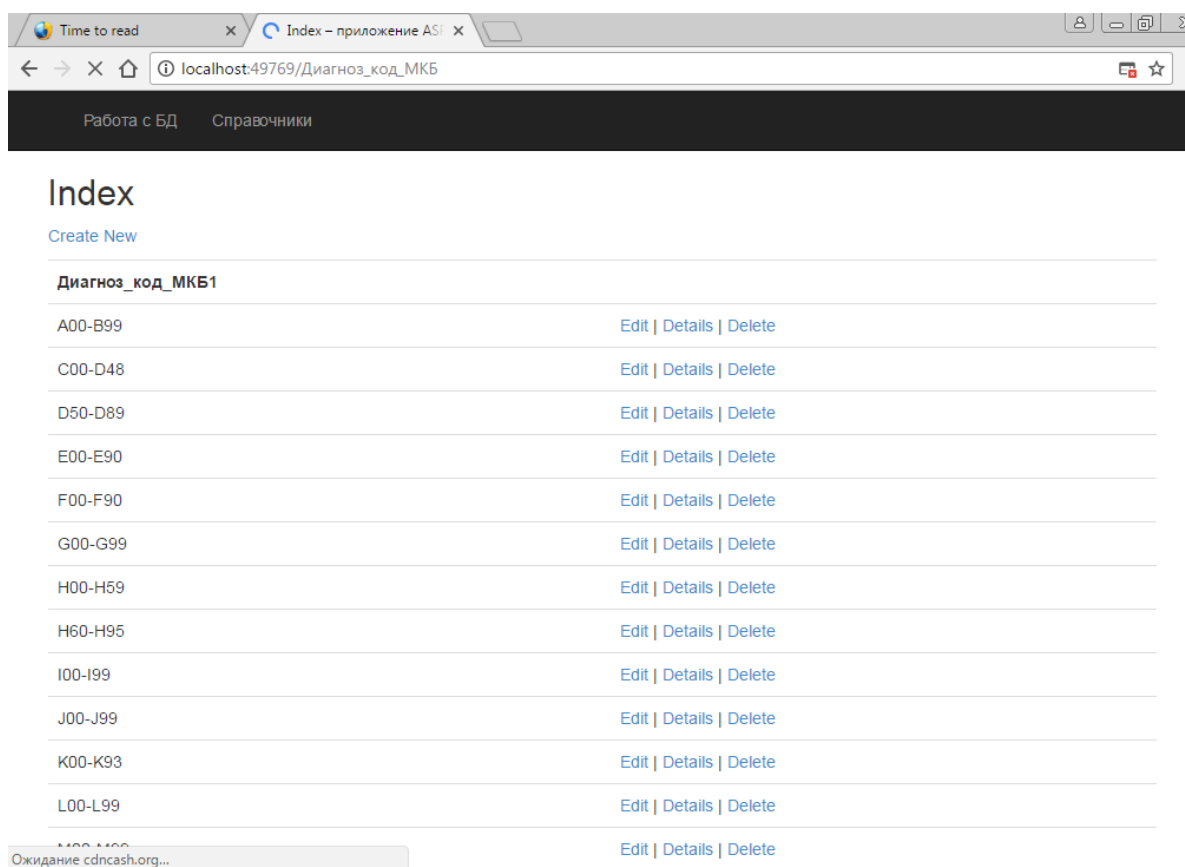


Рисунок 4.3 – Окно пример справочника

Были сгенерированы данные за месяц (апрель 2017). Ниже представлены результаты визуализации полученных данных.

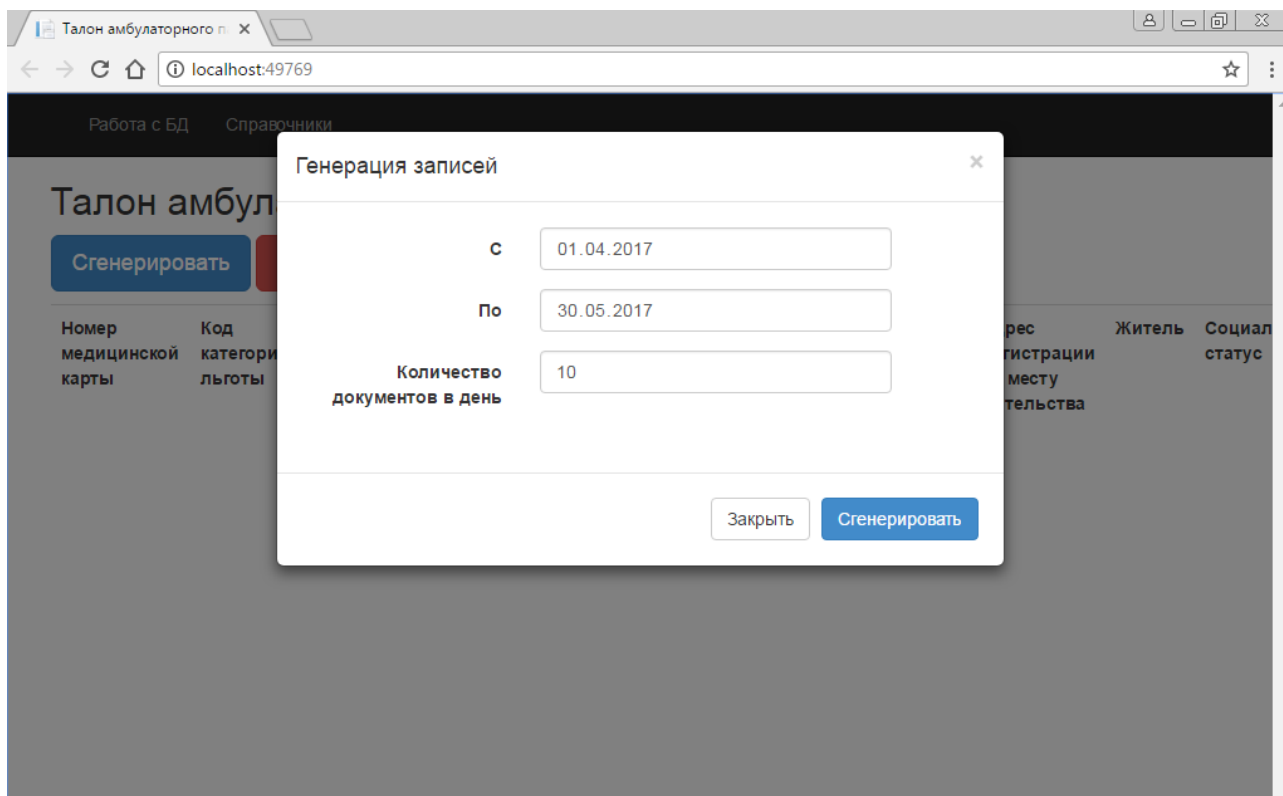


Рисунок 4.4 – Окно генераций записей

После генерирования записей можно сохраниваться «Экспорт» в формате CSV и открывается в Excel, также можно отображать каждую запись отдельно в формате «xml» CDA «clinical document architecture».

Код мед. услуги	Характер заболевания	Диспансерный учет	Травма	Документ временной нетрудоспособности	Причина выдачи	Рецептурный бланк серия и N, дата выписки	Export to Csv
A25.18.001	впервые в жизни установленное хроническое	снят	ДТП	закрыт	карантин	1234 4321 08.09.2014 0:00:00	Сохранить как XML
B01.067.002	острое	по выздоровлению	полученная в результате террористических действий	закрыт	санаторно-курортное лечение	963963 852852 03.02.2017 0:00:00	Сохранить как XML
B01.046.005	острое	состоит	промышленная	открыт	заболевание	21342432 1235432 17.01.2017 0:00:00	Сохранить как XML

Рисунок 4.5 – Окно перевести информации в Excel или сохранить их как xml

Для построения этих диаграмм был использован Excel, сводные диаграммы.



Рисунок 4.6 – Распределение пациентов за апрель по половому признаку

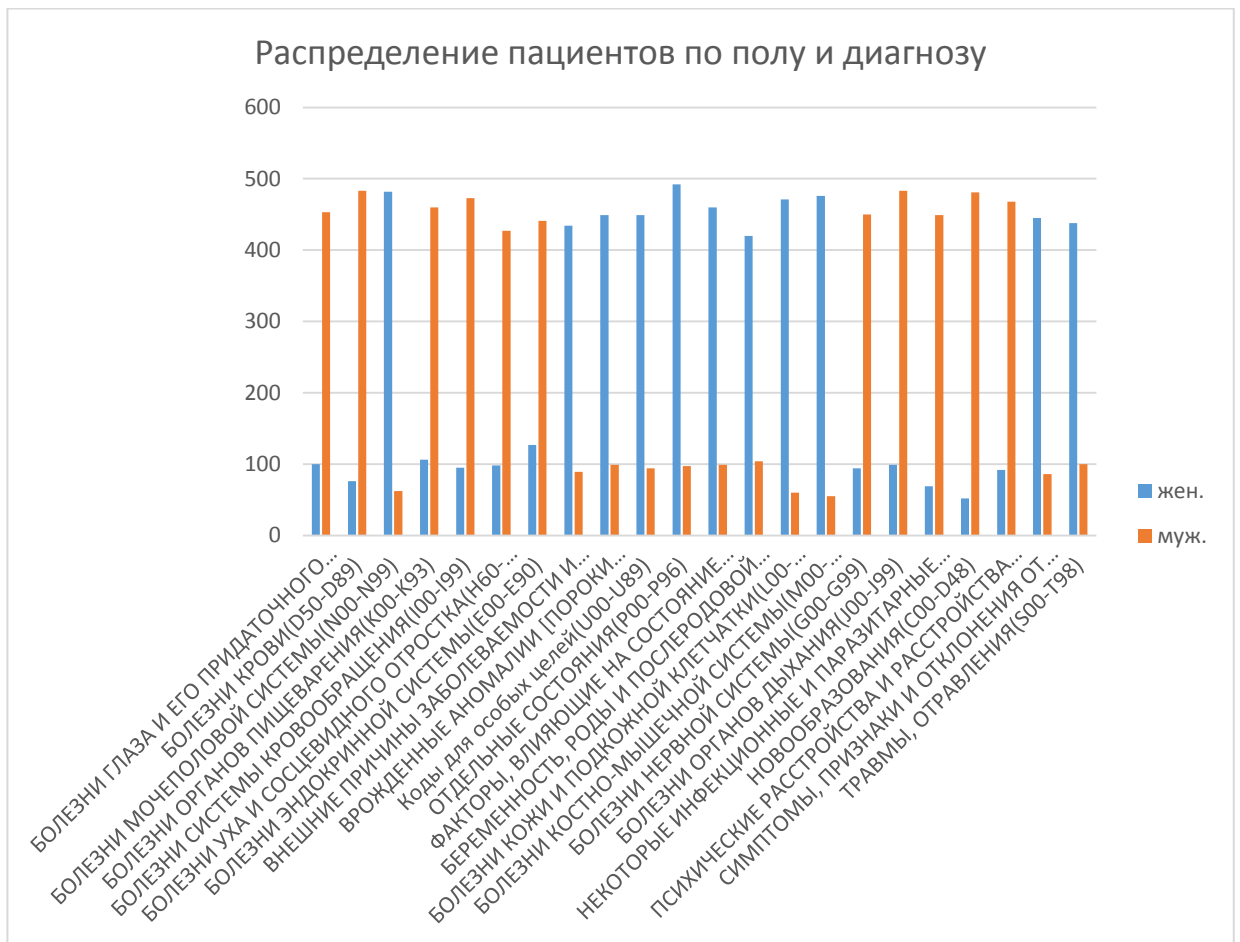


Рисунок 4.7 – Распределение пациентов по полу и диагнозу

Этот график показывает, пациенты какого пола чаще болеют конкретным заболеванием.



Рисунок 4.8 – Распределение числа пациентов по диагнозам

По такой диаграмме можно судить о том, какое заболевание чаще встречается в определенный период времени.

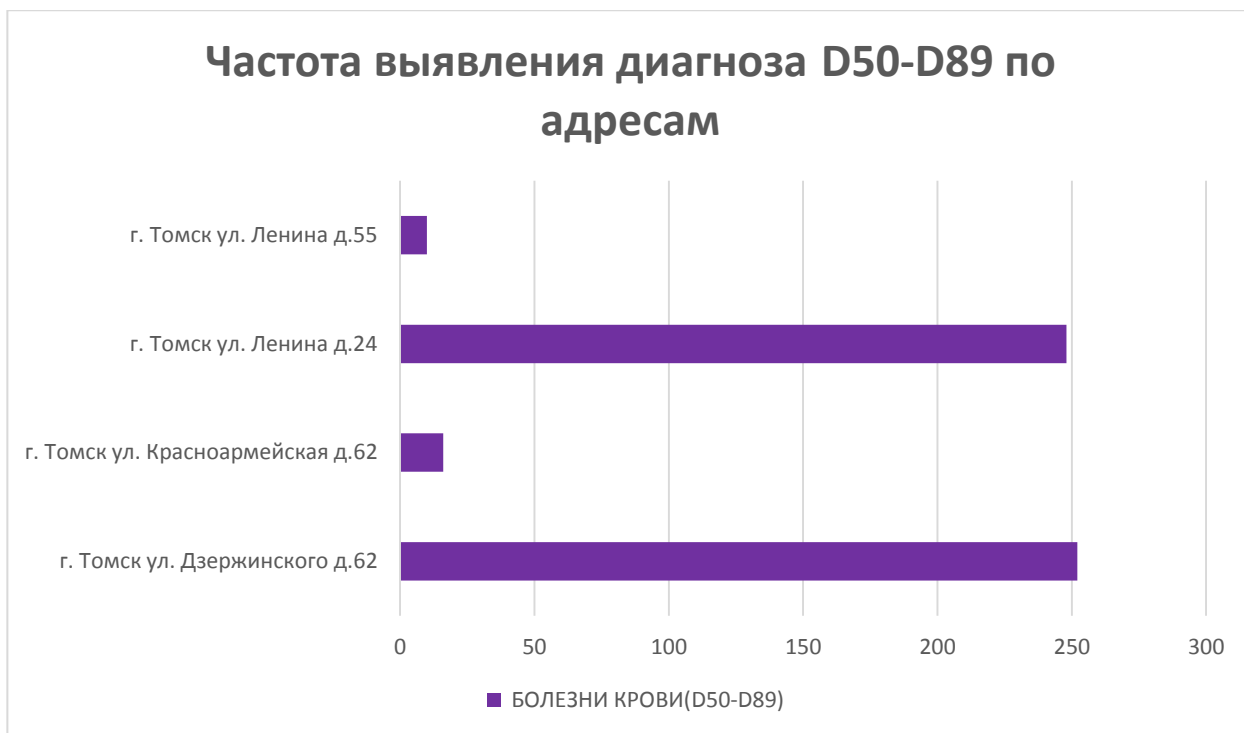


Рисунок 4.9 – Распределение пациентов по диагнозу и по адресам

По данной диаграмме можно отслеживать вспышки заболевания по территориальному признаку.

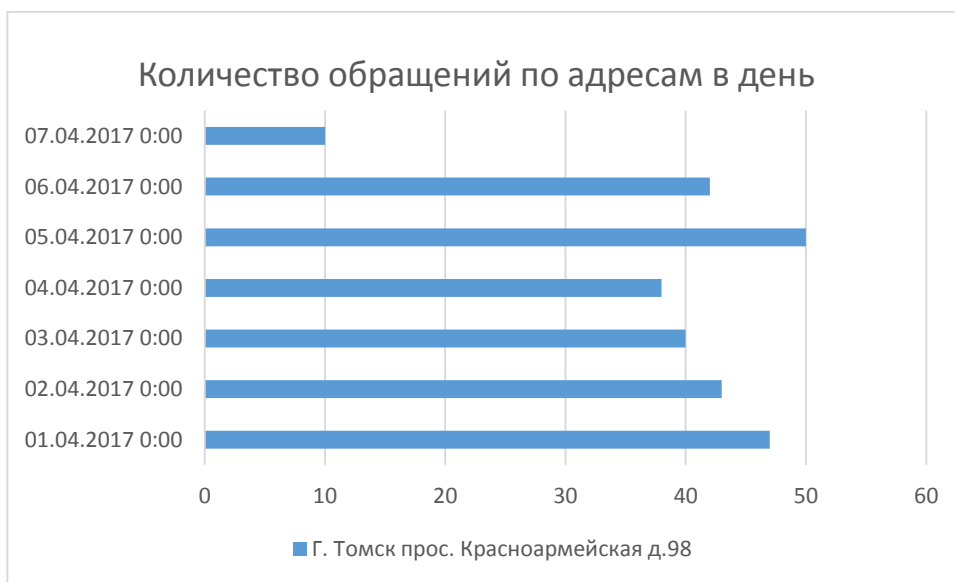


Рисунок 4.10 – Распределение пациентов по обращению по адресам за рассматриваемый период

Данная диаграмма показывает, сколько обращений по конкретному адресу было в день (если посещений в конкретный день не было, то он пропущен в диаграмме).

4.2 Геоинформационные системы (ГИС)

Для интерпретации и визуализации данных здравоохранения по Томской области, полученных в результате интеллектуального анализа, можно использовать такой класс информационных систем как геоинформационные системы. Геоинформационная система (ГИС) — система сбора, хранения, анализа и графической визуализации пространственных (географических) данных и связанной с ними информации о необходимых объектах. Таким образом, если пространственные данные имеют привязанные к ним данные, характеризующие этот субъект, то ГИС – один из лучших инструментов для их визуализации.[31]

ГИС позволяют создавать тематические карты – это карты для отображения узкого (тематического) круга явлений, которые распределены по поверхности и привязаны к точкам, областям, административным образованиям и т. п. Визуализация этих явлений осуществляется в виде: значков, картодиаграмм и картограмм.

4.2.1 Значки

Данный способ применяется для данных, привязанных к некоторой точке. Так обычно отмечаются точки интересов (POI, Point of interest) на карте. В случае со здравоохранительными данными, можно отображать конкретные медицинские учреждения в виде таких POI – тогда можно легко визуально оценить расстояния между медицинскими учреждениями, искать ближайшее



Рисунок 4.12 – Тематическая карта в виде картодиаграмм с пропорциональными символами

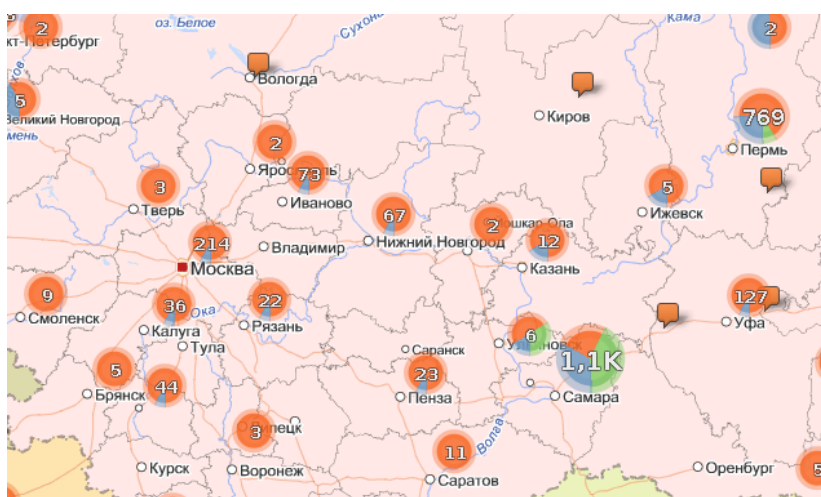


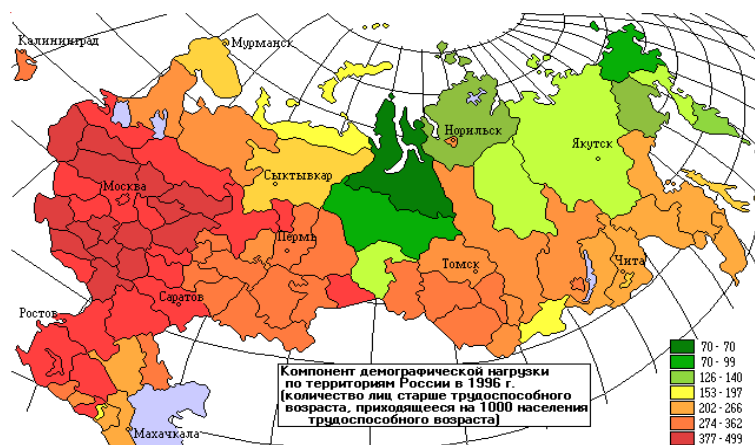
Рисунок 4.13 – Тематическая карта с использованием круговых (секторных) диаграмм

В случае с медицинскими данными, можно отображать в виде картодиаграмм соотношение больных с определенным заболеванием по полам (мужскому и женскому).

4.2.3 Картограммы

Картограмма – это карта, на которой различные области имеют различные заливки (с изменением степени насыщенности) или штриховку, в зависимости от

интенсивности какого-либо показателя. В качестве таких областей обычно используются объекты административно-территориального деления (районы,



города, страны и т. д.).

Рисунок 4.14 – Тематическая карта с использованием картограмм

С помощью этого подхода можно разделить карту Томской области по районам и отображать разнообразные медицинские данные, например, число зарегистрированных случаев клещевого энцефалита.

Таким образом, ГИС является удобным средством для визуализации результатов анализа медицинских данных. Наиболее популярными и распространенными ГИС для построения таких тематических карт являются ArcGIS (платная), MapInfo (платная) и QGIS (свободно распространяемая).

Единственной особенностью, которая не позволила нам использовать ГИС для интерпретации результатов анализа является то, что данные ГИС работают с картами в определенном формате (Shapefile), а в свободном доступе такой карты с разбиением на те же районы, что и в нашем исследовании, не оказалось. Но если такая карта будет разработана, то в атрибутивные базы данных можно будет заносить информацию из результатов анализа. [32]

Глава 5. Оценка эффективности разработанных методик, алгоритмов и программного комплекса

В данной главе проведена оценка соответствия генерируемых данных формам реальных документов, используемых в медицинских учреждениях; проведено тестирование разработанной системы и оценка полученных результатов.

5.1 Оценка соответствия разработанных электронных форм и реальных документов

Существует множество форм и форматов хранения медицинских данных, но наиболее популярным и распространенным по всему миру является CDA.

```
▼<ClinicalDocument xmlns="urn:hl7-org:v3" xmlns:voc="urn:hl7-org:v3/voc" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  templateId="2.16.840.1.113883.3.27.1776">
  ▼<!--
    *****
    CDA Header
    *****
  -->
  <id extension="c266" root="2.16.840.1.113883.3.933"/>
  <code code="11488-4" codeSystem="2.16.840.1.113883.6.1" displayName="Consultation note"/>
  <title>Good Health Clinic Consultation Note</title>
  <effectiveTime value="20000407"/>
  <confidentialityCode code="N" codeSystem="2.16.840.1.113883.5.25"/>
  <setId extension="BB35" root="2.16.840.1.113883.3.933"/>
  <versionNumber value="2"/>
  ▼<legalAuthenticator>
    <time value="20000408"/>
    <signatureCode code="S"/>
    ▼<assignedEntity>
      <id extension="KP00017" root="2.16.840.1.113883.3.933"/>
      ▼<assignedPerson>
        ▼<name>
          <given>Robert</given>
          <family>Dolin</family>
          <suffix>MD</suffix>
        </name>
      </assignedPerson>
      ▼<representedOrganization>
        <id extension="M345" root="2.16.840.1.113883.3.933"/>
      </representedOrganization>
    </assignedEntity>
  </legalAuthenticator>
  ▼<author>
    <time value="20000407"/>
    ▼<assignedAuthor>
```

Рисунок 5.1 А – документ в формате CDA


```

*****
CDA Body
*****
-->
▼<component>
  ▼<StructuredBody>
    ▼<!--
      *****
      History of Present Illness section
      *****
    -->
    ▼<component>
      ▼<section>
        <code code="10164-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC"/>
        <title>History of Present Illness</title>
        ▼<text>
          Henry Levin, the 7th is a 67 year old male referred for further asthma management. Onset of asthma in his
          <content revised="delete">twenties</content>
          <content revised="insert">teens</content>
          . He was hospitalized twice last year, and already twice this year. He has not been able to be weaned off stero:
          several months.
        </text>
      </section>
    </component>
  ▼<!--
    *****
    Past Medical History section
    *****
  -->
  ▼<component>
    ▼<section>
      <code code="10153-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC"/>
      <title>Past Medical History</title>
      ▼<text>
        ▼<list>

```

Рисунок 5.1 В– документ в формате CDA

В Российской системе здравоохранения пока формат CDA не поддерживается полностью (но в ближайшем будущем это планируется). Поэтому для перехода к электронному документообороту в медицинских учреждениях ввели Талон амбулаторного пациента (ТАП).

ФОРМА N 025-12/У ТАЛОН АМБУЛАТОРНОГО ПАЦИЕНТА

N медицинской карты _____ Дата _____			
1. Код категории льготы			
2. Номер страхового полиса ОМС			
3. СНИЛС			
4. Пациент: код <1>			ф.и.о.
5. Пол <4>: 1 - муж.; 2 - жен.	6. Дата рождения		
7. Документ, удостоверяющий личность (название, серия и номер) <4>:			
8. Адрес регистрации по месту жительства <4>:			
			9. Житель <4>: 1 - город; 2 - сел
10. Социальный статус, в т.ч. занятость: 1. - дошкольник: 1.1. - организован, 1.2. неорганизован, 2 - учащийся, 3 - работающий,			
4 - неработающий; 5 - пенсионер; 6 - военнослужащий, код		7 - член семьи военнослужащего; 8 - без определенного места жительства	
12. Специалист: код			
13. Специалист: код <2>			ф.и.о.
14. Вид оплаты: 1 - ОМС; 2 - бюджет, 3 - платные услуги, в т.ч. 4 - ДМС; 5 - другое			
14 Место обслуживания: 1 - поликлиника, 2 - на дому, в т.ч. 3 - актив.			
16. Цель посещения: 1 - заболевание, 2 - профосмотр; 3 - патронаж; 4 - другое			
17. Результат обращения <5>: случай закончен: 1 - выздоровл.; 2 - улучшение; 3 - динамическое набл., направлен: 4 - на госпитализацию,			
5 - в дневной стационар, 6 - стационар на дому, 7 - на консультацию, 8 - на консультацию в др. ЛПУ, 9 - справка для получения путевки,			
10 - санаторно-курортная карта			

- <1> При использовании кода, принятого в ЛПУ.
 <2> Заполняется при учете работы среднего мед. персонала.
 <3> При оплате: по посещению проставляется код посещения или стандарта медицинской помощи.
 <4> Заполняются при разовом обращении пациента (например, иногородний).
 <5> Заполняется при последнем посещении по данному случаю.

18. Диагноз код МКБ			
19. Код мед. услуги (посещения, СМП, КЭС) <3>:			
20. Характер заболевания: 1 - острое (+), впервые в жизни установленное хроническое (+); 2 - диагноз установлен в предыдущ. году или ранее (-)			
21. Диспансерный учет: 1 - состоит, 2 - взят, 3 - снят, в т.ч. 4 - по выздоровлению			
22. Травма: - производственная: 1 - промышленная; 2 - транспортная, в т.ч. 3 - ДТП; 4 - сельскохозяйственная; 5 - прочие			
не производственная: 6 - бытовая; 7 - уличная; 8 - транспортная, в т.ч. 9 - ДТП; 10 - школьная; 11 - спортивная; 12 - прочие;			
13 - полученная в результате террористических действий			
23. Диагноз код			
24. Код мед. услуги (посещения, СМП, КЭС) <3>:			
25. Характер заболевания: 1 - острое (+), впервые в жизни установленное хроническое (+); 2 - диагноз установлен в предыдущ. году или ранее (-)			
26. Диспансерный учет 1 - состоит, 2 - взят, 3 - снят, в т.ч. 4 - по выздоровлению			
27. Заполняется только при изменении диагноза: ранее зарегистрированный диагноз			
Код МКБ - 10			
Дата регистрации изменяемого диагноза:			
28. Документ временной нетрудоспособности: 1 - открыт; 2 - закрыт;			
29. Причина выдачи: 1 - заболевание; 2 - по уходу; 3 - карантин; 4 - прерывание беременности;			
5 - отпуск по беременности и родам; 6 - санаторно-курортное лечение,			
29.1 по уходу: пол 1 - муж; 2 - жен. (возраст лица, получившего документ в/н)			
30. Рецепт бланк серия и N, дата выписки:			
30.1 _____;			
30.2 _____;			
30.3 _____;			
30.4 _____.			

Рисунок 5.2 – Талон амбулаторного пациента Форма 025-12/у

Разработанная программа для генерации медицинских данных поддерживаются все поля, характерные для ТАП.

Что касается формата CDA, то так как он не имеет обязательных полей и хранится в виде xml, но каждому полю из ТАП можно поставить в соответствие поля из CDA. Таким образом, в разработанной программе был добавлен функционал для того, чтобы сохранять генерируемые данные в формате CDA (xml с пропущенными значениями для не определенных в ТАП полей).

5.2 Тестирование разработанной системы

Для тестирования работы разработанной системы были заполнены все справочники, характерные для ТАП. Генерация данных осуществляется путем выбора случайных значений для каждого из полей ТАП из этих справочников.

Были протестированы следующие тестовые сценарии:

1) Наполнение справочников

Для наполнения справочников нужно перейти на соответствующую страницу, выбрать справочник из списка справочников и нажать «Create New». Также доступны инструменты для редактирования и удаления записей из справочника. В результате запись добавляется в справочник.

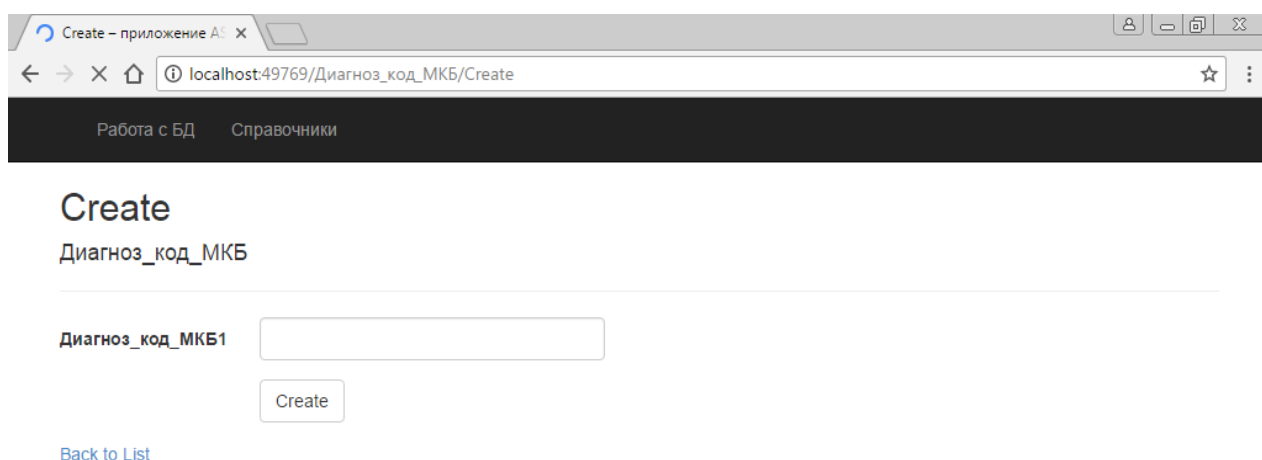


Рисунок 5.3 – Добавление записи в справочник

2) Генерация данных

Для генерации данных нужно нажать кнопку «Сгенерировать», в результате чего пользователю будет предложено выбрать дату начала и дату окончания генерации данных, количество генерируемых документов за день. В результате по данным настройкам должны быть сгенерированы документы в формате ТАП.

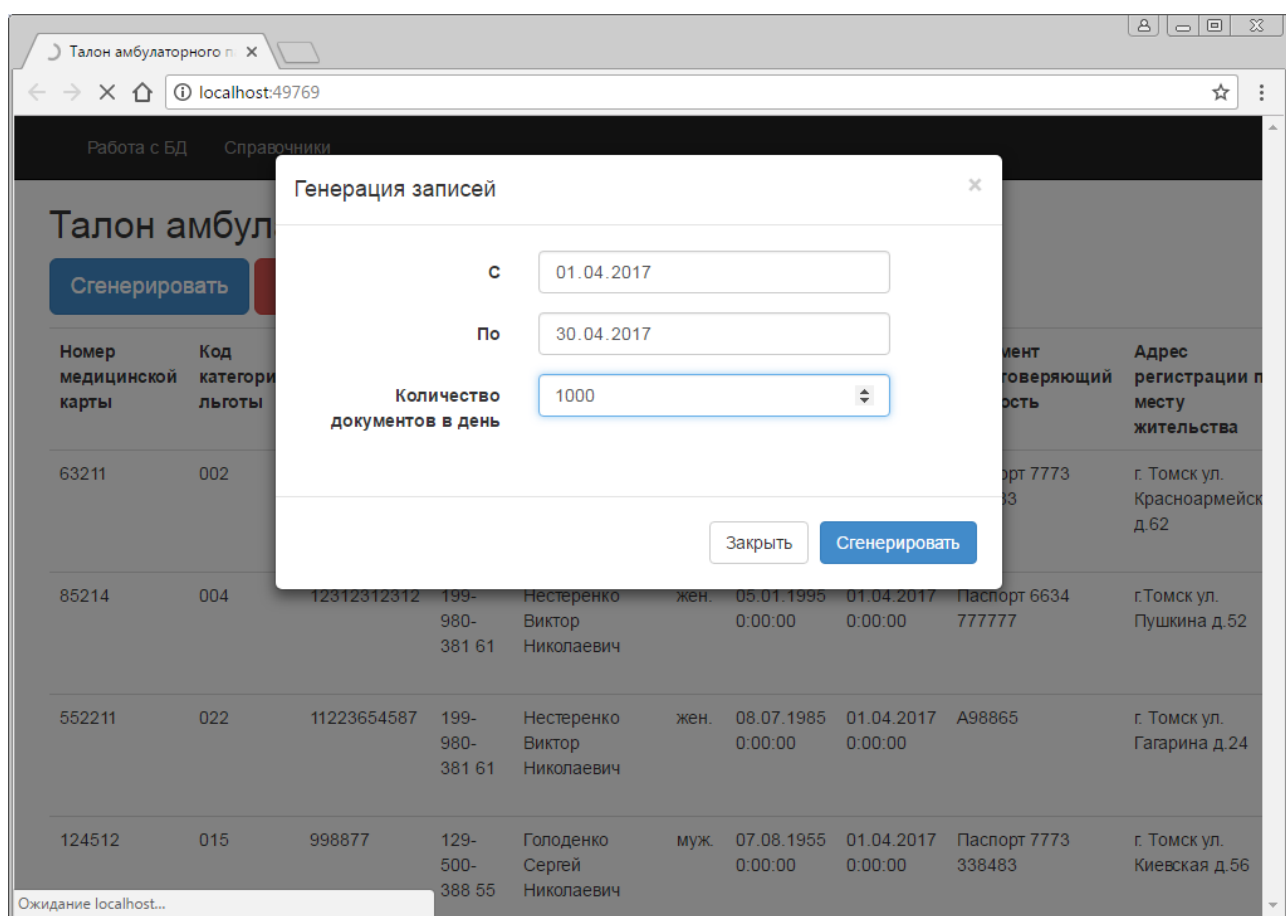


Рисунок 5.4 – Генерация документов

3) Экспорт в формат .csv

После генерации данных их можно экспортировать в формат .csv. Для этого нужно нажать на надпись «Export to Csv». В результате файл сохраняется на диск.

Номер медиц	Код категории	Номер страхового СНИЛС	ФИО	Пол	Дата рождения	Дата	Доку
96524	19	125487 129-500-388 55	Бутов Иван Михайлович	муж.	04.05.1979 0:00	01.04.2017 0:00	A988
100	130	123456789 111-800-381 88	Ягупов Анатолий Павлович	жен.	08.09.1961 0:00	01.04.2017 0:00	посп
654321	30	123123123 199-900-381 81	Болдырев Игорь Александрович	муж.	18.10.1971 0:00	01.04.2017 0:00	Пасп
23568	8	125487 129-500-388 55	Белецкий Константин Владимирович	муж.	03.10.1978 0:00	01.04.2017 0:00	Пасп
99652	26	87536545 102-200-381 31	Ходаковский Вадим Валерьевич	жен.	09.12.1966 0:00	01.04.2017 0:00	A988
1234	26	12312312 199-900-381 81	Иванов Иван Иванович	муж.	09.12.1966 0:00	01.04.2017 0:00	A988
1234 010		12312312 199-900-381 81	Ермоленко Андрей Валентинович	муж.	30.11.1993 0:00	01.04.2017 0:00	Пасп
2214	120	65274898 499-652-381 22	Ермоленко Андрей Валентинович	жен.	02.04.2005 0:00	01.04.2017 0:00	посп
85214	4	12312312312 116-973-385 89	Капустин Сергей Иванович	муж.	05.01.1995 0:00	01.04.2017 0:00	Пасп
85214	4	12312312312 199-900-381 81	Медведева Светлана Андреевна	муж.	05.01.1995 0:00	01.04.2017 0:00	Пасп
3214	12	12312312 199-900-381 81	Медведева Светлана Андреевна	муж.	30.11.1993 0:00	02.04.2017 0:00	Пасп
123321	15	998877 129-500-388 55	Голоденко Сергей Николаевич	муж.	07.08.1955 0:00	02.04.2017 0:00	Пасп
995541	15	112547236 499-652-381 22	Попова Ольга Константиновна	жен.	07.08.1955 0:00	02.04.2017 0:00	Пасп
66221	24	11223654587 102-200-381 31	Плотникова Татьяна Николаевна	жен.	12.12.2000 0:00	02.04.2017 0:00	A988
66221	24	11223654587 102-200-381 31	Плотникова Татьяна Николаевна	жен.	08.07.1985 0:00	02.04.2017 0:00	A988
4520	173	123456789 111-800-381 88	Сергеева Вероника Николаевна	жен.	06.08.2001 0:00	02.04.2017 0:00	посп
854	99	112547236 499-652-381 22	Колесникова Ирина Сергеевна	жен.	12.12.2000 0:00	02.04.2017 0:00	посп
52463	2	123123123 116-973-385 89	Сотникова Надежда Андреевна	муж.	13.04.1973 0:00	02.04.2017 0:00	Пасп
225541	2	112236545 199-980-381 61	Криворучкин Михаил Тихонович	жен.	13.04.1973 0:00	02.04.2017 0:00	Пасп
225541	19	112236545 199-980-381 61	Криворучкин Михаил Тихонович	жен.	04.05.1979 0:00	02.04.2017 0:00	A988
77441	26	87536545 102-200-381 31	Шевцов Алексей Васильевич	жен.	09.12.1966 0:00	03.04.2017 0:00	A988
77441	26	87536545 199-900-381 81	Иванов Иван Иванович	жен.	09.12.1966 0:00	03.04.2017 0:00	A988
1234 010		12312312 199-900-381 81	Ермоленко Андрей Валентинович	муж.	30.11.1993 0:00	03.04.2017 0:00	Пасп
2214	120	65274898 499-652-381 22	Капустин Сергей Иванович	жен.	02.04.2005 0:00	03.04.2017 0:00	посп
85214	4	12312312312 199-900-381 81	Медведева Светлана Андреевна	муж.	05.01.1995 0:00	03.04.2017 0:00	Пасп
3214	11	12312312 199-900-381 81	Медведева Светлана Андреевна	муж.	30.11.1993 0:00	03.04.2017 0:00	Пасп
124512	15	998877 129-500-388 55	Голоденко Сергей Николаевич	муж.	07.08.1955 0:00	03.04.2017 0:00	Пасп

Рисунок 5.5 – Экспортированные данные

4) Экспорт в формат .xml

Также можно экспортировать каждую запись о посещении пациента в формат .xml. Для этого нужно нажать на надпись напротив выбранной записи «Сохранить как XML». В результате файл сохраняется на диск.

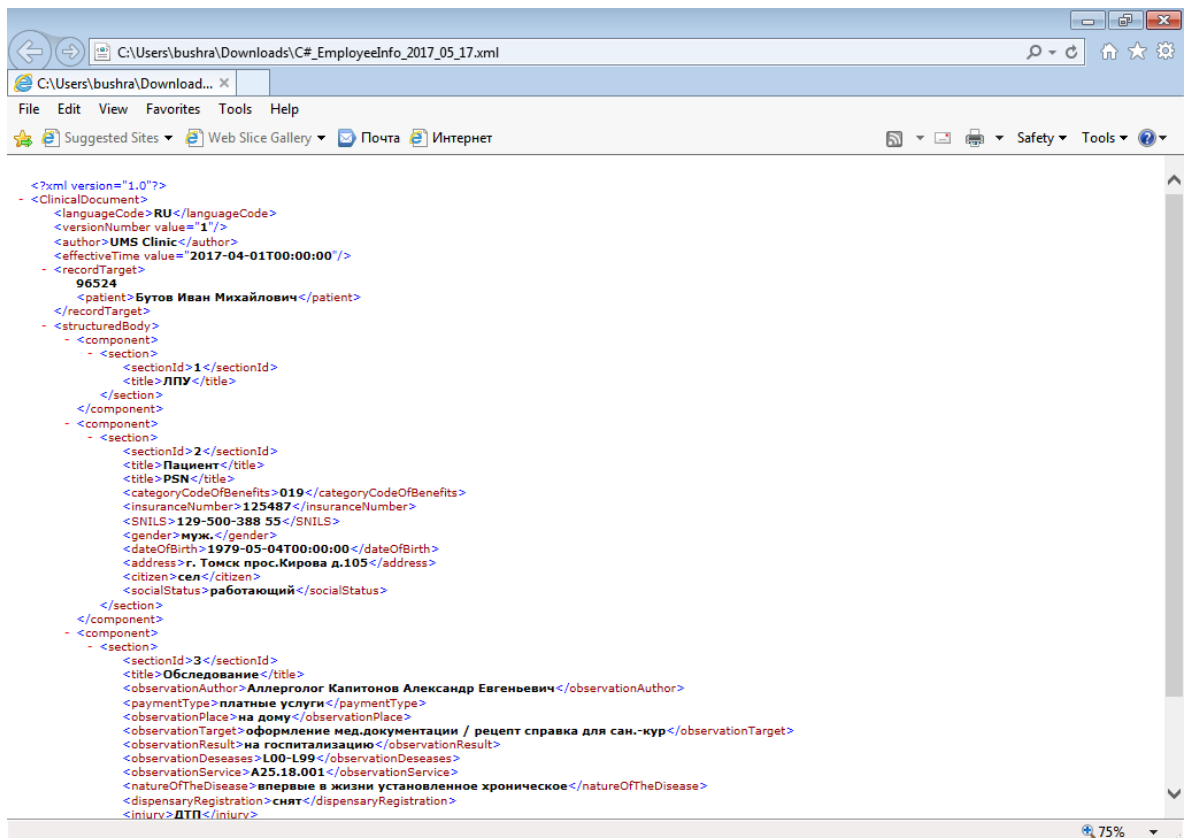


Рисунок 5.6 – Сгенерированный документ в формате xml

5.3 Оценка результатов тестирования

Результаты работы системы, полученные в ходе тестирования, полностью отвечают ожиданиям. Так, сгенерированные данные пригодны для визуализации и проведения дальнейшего анализа по ним.

Глава 6. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»

Студенту:

Группа 8ВМ5В	ФИО Мохаммеджавад Бушра Джабер
-----------------	-----------------------------------

Институт	Кибернетики	Каф едра	Программной инженерии
Уровень образования	Магистрант	Направление/ специальность	09.04.01 Информатика и вычислительная техника

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих
2. Нормы и нормативы расходования ресурсов	
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого и инновационного потенциала НТИ	
2. Разработка устава научно-технического проекта	
3. Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок	
4. Определение ресурсной, финансовой, экономической эффективности	

Перечень графического материала (с точным указанием обязательных чертежей):

1. «Портрет» потребителя результатов НТИ	
2. Сегментирование рынка	
3. Оценка конкурентоспособности технических решений	
4. Диаграмма FAST	
5. Матрица SWOT	
6. График проведения и бюджет НТИ	
7. Оценка ресурсной, финансовой и экономической эффективности НТИ	
8. Потенциальные риски	

Дата выдачи задания для раздела по линейному графику	
--	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. МЕН	Конотопский Владимир Юрьевич	К.Э.Н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ5В	Мохаммеджавад Бушра Джабер		

В этом пункте Выпускной квалификационной работе посвящается экономическому обоснованию данной работы. В ней идет речь об оценке затрат на исследование и внедрение. Цель раздела – оценка целесообразности выполненной работы с помощью показателей эффективности.

6.1. Организация и планирование работ

На данном разделе осуществляется полный перечень проводимых работ, определяются их исполнители и оптимальная продолжительность.

Данную работу вели 2 человека: исполнитель (И) Мохаммеджавад Б.Д. и научный руководитель (НР) – Пономарев А.А. доцент кафедры автоматики и компьютерных систем.

В таблице 6.1 приведен перечень работ и продолжительность их выполнения.

Таблица 6.1 – Перечень этапов, работ и распределение исполнителей

Этапы работы	Исполнители	Загрузка исполнителей
Постановка целей и задач, сбор первоначальной информации	НР И	НР – 70% И – 30 %
Планирование и технико-экономическое обоснование ВКР	И	И – 100%
Анализ опасных и вредных производственных факторов	И	И – 100%
Составление и согласование технического задания	НР, И	НР – 30% И – 70%
Анализ предметной области и сбор необходимой информации	И	И – 100%
Разработка модуля генерирования исходных данных	НР, И	НР – 10% И – 90%
Проектирование интерфейсов	И	И – 100%
Разработка визуальных компонентов для задач быстрой навигации по ЭМК	И	И – 100%
Разработка алгоритма прогнозирования потребности в медицинской помощи	И	И – 100%
Тестирование разработанной системы	НР, И	НР – 30% И – 70%
Составление и оформление пояснительной записки	И	И-100%

6.1.1. Продолжительность этапов работ

Благодаря таблице 1 можно определить значения продолжительности выполняемых работ ($t_{ож}$) экспертным путем с учетом загруженности исполнителей в рабочих днях по формуле:

$$t_{ож} = \frac{3 \cdot t_{\min} + 2 \cdot t_{\max}}{5} \quad (1)$$

где t_{\min} – минимальная продолжительность работы, дн.;

t_{\max} – максимальная продолжительность работы, дн.;

Чтобы сделать линейный график нужно рассчитать длительность этапов в рабочих днях, а потом перевести ее в календарные дни. Расчет продолжительности выполнения каждого этапа в рабочих днях ($T_{РД}$) ведется по формуле:

$$T_{РД} = \frac{t_{ож}}{K_{ВН}} \cdot K_{Д} \quad (2)$$

где $t_{ож}$ – продолжительность работы, дн.;

$K_{ВН}$ – коэффициент выполнения работ, учитывающий влияние внешних факторов на соблюдение предварительно определенных длительностей, в частности, возможно $K_{ВН} = 1$;

$K_{Д}$ – коэффициент, учитывающий дополнительное время на компенсацию непредвиденных задержек и согласование работ ($K_{Д} = 1-1,2$; в этих границах конкретное значение принимает сам исполнитель). Примем $K_{Д} = 1$.

Расчет продолжительности этапа в календарных днях ведется по формуле:

$$T_{КД} = T_{РД} \cdot T_{К} \quad (3)$$

где $T_{КД}$ – продолжительность выполнения этапа в календарных днях;

ТК – коэффициент календарности, позволяющий перейти от длительности работ в рабочих днях к их аналогам в календарных днях, и рассчитываемый по формуле 6.4

$$T_K = \frac{T_{\text{КАЛ}}}{T_{\text{КАЛ}} - T_{\text{ВД}} - T_{\text{ПД}}} \quad (4)$$

где $T_{\text{КАЛ}}$ – календарные дни ($T_{\text{КАЛ}} = 365$);

$T_{\text{ВД}}$ – выходные дни ($T_{\text{ВД}} = 53$);

$T_{\text{ПД}}$ – праздничные дни ($T_{\text{ПД}} = 13$).

$$T_K = \frac{365}{365 - 53 - 13} = 1,22$$

В таблице 6.3 приведены длительность этапов работ и число исполнителей, занятых на каждом этапе.

Таблица 6.2 – Длительность этапов работ и число исполнителей, занятых на каждом этапе.

Этапы работы	Исполнители	Загрузка исполнителей	Продолжительность работ, дни			Трудоемкость, чел/дн			
			t _{мин}	t _{мах}	t _{ож}	Трд		Ткд	
						НР	И	НР	И
Постановка целей и задач, сбор первоначальной информации	НР И	НР – 70% И – 30 %	2	4	2,8	1,96	0,84	2,39	1,02
Планирование и технико-экономическое обоснование ВКР	И	И – 100%	2	4	2,8		2,8		3,42
Анализ опасных и вредных производственных факторов	И	И – 100%	1	2	1,4		1,4		1,71
Составление и согласование	НР, И	НР – 30% И – 70%	3	5	3,8	1,14	2,66	1,39	3,25

технического задания									
Анализ предметной области и сбор необходимой информации	И	И – 100%	7	9	7,8		7,8		9,52
Разработка модуля генерирования исходных данных	НР, И	НР – 10% И – 90%	10	15	12	1,2	10,8	1,46	13,18
Проектирование интерфейсов	И	И – 100%	5	8	6,2		6,2		7,56
Разработка визуальных компонентов для задач быстрой навигации по ЭМК	И	И – 100%	12	15	13,2		13,2		16,1
Разработка алгоритма прогнозирования потребности в медицинской помощи	И	И – 100%	10	15	12		12		14,64
Тестирование разработанной системы	НР, И	НР – 30% И – 70%	3	6	4,2	1,26	2,94	1,54	3,59
Составление и оформление пояснительной записки	И	И-100%	13	15	13,8		13,8		16,84
Итого					80	5,56	74,44	6,78	90,82

6.1.2. Расчет накопления готовности проекта

В данном разделе оцениваются результаты над проектом. Этот показатель дает точно понять, на каком этапе находится данная работа.

Таблица 6.3 – Линейный график работ

Этап	НР	И	Февраль			Март			Апрель			Май		
			10	20	30	40	50	60	70	80	90	100	110	120
1	2,39	1,02	■											
2		3,42	■											
3		1,71		■										
4	1,39	3,25		■										
5		9,52		■	■									
6	1,46	13,1 8			■	■								
7		7,56				■	■							
8		16,1					■	■						
9		14,6 4							■	■				
10	1,54	3,59								■	■			
11		16,8 4									■	■		

Введем следующие обозначения:

- ТР_{общ.} – общая трудоемкость проекта;
- ТР_і (ТР_к) – трудоемкость і-го (к-го) этапа проекта, ;
- ТР_{іН} – накопленная трудоемкость і-го этапа проекта по его завершении;
- ТР_{іj} (ТР_{kj}) – трудоемкость работ, выполняемых j-м участником на і-м этапе, здесь j=1,m – индекс исполнителя (m = 2).

Степень готовности определяется формулой (5.5)

$$СГ_i = \frac{ТР_i^H}{ТР_{общ.}} = \frac{\sum_{k=1}^i ТР_k}{ТР_{общ.}} = \frac{\sum_{k=1}^i \sum_{j=1}^m ТР_{km}}{\sum_{k=1}^I \sum_{j=1}^m ТР_{km}}$$

Таблица 6.4 – Нарастание технической готовности работы и удельный вес каждого этапа.

Этапы работы	ТР _і , %	СГ _і , %
Постановка целей и задач, сбор первоначальной информации	3.49	3.49
Планирование и технико-экономическое обоснование ВКР	3.50	6.99
Анализ опасных и вредных производственных факторов	1.75	8.74
Составление и согласование технического задания	4.75	13.49
Анализ предметной области и сбор необходимой информации	9.75	23.24
Разработка модуля генерирования исходных данных	15.00	38.24
Проектирование интерфейсов	7.75	45.99
Разработка визуальных компонентов для задач быстрой навигации по ЭМК	16.50	62.49
Разработка алгоритма прогнозирования потребности в медицинской помощи	15.00	77.49
Тестирование разработанной системы	5.26	82.75
Составление и оформление пояснительной записки	17.25	100.00

6.2. Расчет сметы затрат на выполнение проекта

Для реализации проекта, включается стоимость всех расходов, а смета производится от следующих затрат:

- затраты на материалы;
- заработная плата;
- социальный налог;
- расходы на электроэнергию (без освещения);
- амортизационные отчисления;
- командировочные расходы;
- оплата услуг связи;
- арендная плата за пользование имуществом;
- прочие услуги (сторонних организаций);
- прочие (накладные расходы) расходы.

6.2.1 Расчет затрат на материалы

В этом разделе можно увидеть стоимость материалов, покупных изделий, полуфабрикатов и других материальных ценностей, расходуемых непосредственно в процессе выполнения работ над объектом проектирования. Также статья включает так называемые транспортно-заготовительные расходы, связанные с транспортировкой от поставщика к потребителю, хранением и прочими процессами, обеспечивающими движение (доставку) материальных ресурсов от поставщиков к потребителю. Сюда же включаются расходы на совершение сделки купли-продажи (т.н. транзакции). Приблизительно они оцениваются в процентах к отпускной цене закупаемых материалов, как правило, это 5 , 20 %. Исполнитель работы самостоятельно выбирает их величину в указанных границах.

Таблица 6.5 – Расчет затрат на материалы

Наименование материалов	Цена за ед., руб.	Кол-во	Сумма, руб.
Бумага для принтера формата А4	200	1 уп.	200
Картридж для принтера	2000	1 шт.	2000
Итого:			2200

Допустим, что ТЗР составляют 5 % от отпускной цены материалов, тогда расходы на материалы с учетом ТЗР равны $S_{\text{мат}} = 2200 * 1,05 = 2310$ руб.

6.2.1. Расчет заработной платы

Расчет заработной платы руководителя и исполнителя выполняется на основе трудоемкости выполнения каждого этапа и величины месячного оклада исполнителя, а также премии, входящие в фонд заработной платы

Среднедневная тарифная заработная плата (ЗПдн-т) рассчитывается по формуле:

$$\text{ЗПдн-т} = \text{МО} / 24,83 \quad (6)$$

где МО – месячный оклад.

Расчеты затрат на полную заработную плату приведены в таблице 5.4. Из таблицы 6.2. взяты затраты времени по исполнителям в рабочих днях с округлением до целого. Для учета в ее составе премий, дополнительной зарплаты и районной надбавки используется следующий ряд коэффициентов: $K_{\text{ПР}} = 1,1$; $K_{\text{доп.ЗП}} = 1,188$; $K_{\text{р}} = 1,3$.

Таким образом, для перехода от тарифной (базовой) суммы заработка исполнителя, связанной с участием в проекте, к соответствующему полному заработку (зарплатной части сметы) необходимо первую умножить на интегральный коэффициент $K_{\text{и}} = 1,1 * 1,188 * 1,3 = 1,699$. Вышеуказанное

значение Кдоп.ЗП применяется при шестидневной рабочей неделе, при пятидневной оно равно 1,113, соответственно в этом случае $K_i = 1,62$.

Из методических указаний за 2013 год был выбран оклад доцента, кандидата наук – 23264,86 руб./мес. по таблице окладов ППС и НС.

Его заработная плата была оценена приблизительно и соответствует заработной плате младшего научного сотрудника без степени – 14874 руб./мес.

Таблица 6.6 – Затраты на заработную плату

Исполнитель	Оклад, руб./мес.	Среднедневная ставка, руб./раб.день	Затраты времени, раб.дни	Коэффициент	Фонд з/платы, руб.
НР	23 264,86	936,97	6	1,699	9551,47
И	14 874	599,05	74	1,62	71814,11
Итого:			81365,58		

6.2.2. Расчет затрат на социальный налог

Затраты на единый социальный налог (ЕСН), которые включают в себя отчисления в пенсионный фонд, на социальное и медицинское страхование, составляют 30 % от полной заработной платы по проекту, т.е. $S_{соц.} = S_{зп} * 0,3$. Итак, в нашем случае $S_{соц.} = 81365,58 * 0,3 = 24409,67$ руб.

6.2.3. Расчет затрат на электроэнергию

В данном вид расходов производятся расчеты затрат на электроэнергию, потраченную в ходе выполнения проекта на работу используемого оборудования, рассчитываемые по формуле:

$$S_{эл.об.} = P_{об} \cdot t_{об} \cdot ЦЭ \quad (7)$$

где $P_{об}$ – мощность, потребляемая оборудованием, кВт;

ЦЭ – тариф на 1 кВт·час;

$t_{об}$ – время работы оборудования, час.

Для ТПУ ЦЭ = 5,782 руб./кВт·час (с НДС).

Время работы оборудования вычисляется на основе итоговых данных таблицы 5.2 для инженера ($T_{РД}$) из расчета, что продолжительность рабочего дня равна 8 часов.

$$t_{об} = T_{РД} * K_t, \quad (5.8)$$

где $K_t \leq 1$ – коэффициент использования оборудования по времени, равный отношению времени его работы в процессе выполнения проекта к $T_{РД}$, определяется исполнителем самостоятельно. В ряде случаев возможно определение $t_{об}$ путем прямого учета, особенно при ограниченном использовании соответствующего оборудования. Примем $K_t = 0,7$.

Мощность, потребляемая оборудованием, определяется по формуле:

$$P_{ОБ} = P_{ном.} * K_C \quad (5.9)$$

где $P_{ном.}$ – номинальная мощность оборудования, кВт;

$K_C \leq 1$ – коэффициент загрузки, зависящий от средней степени использования номинальной мощности. Для технологического оборудования малой мощности $K_C = 1$.

Пример расчета затраты на электроэнергию для технологических целей приведен в таблице 6.6.

Таблица 6.7 – Затраты на электроэнергию технологическую

Наименование оборудования	Время работы оборудования $t_{\text{ОБ}}$, час	Потребляемая мощность $P_{\text{ОБ}}$, кВт	Затраты $\text{Э}_{\text{ОБ}}$, руб.
Персональный компьютер	$(5,56 + 74,4) * 8 * 0,7 = 447,78$	0,3	776,71
Струйный принтер	10	0,1	5,78
Итого:			782,496

6.2.4. Расчет амортизационных расходов

В данной главе ведутся расчеты амортизации используемого оборудования за время выполнения проекта.

Амортизационные отчисления рассчитываются на время использования ЭВМ по формуле:

$$C_{\text{АМ}} = \frac{N_{\text{А}} * C_{\text{ОБ}} * t_{\text{рф}} * n}{F_{\text{Д}}},$$

где $N_{\text{А}}$ – годовая норма амортизации, $N_{\text{А}} = 40\%$;

$C_{\text{ОБ}}$ – цена оборудования, $C_{\text{ОБ}} = 45000$ руб.;

$F_{\text{Д}}$ – действительный годовой фонд рабочего времени, $F_{\text{Д}} = 298 * 8 = 2384$ часа;

Срок реализации программного проекта – 4 месяца (февраль, март, апрель, май) или 80 дней. Количество рабочих дней, согласно календарю на 2017 год: февраль – 24, март – 26, апрель – 24, май – 25.

$t_{\text{ВТ}}$ – время работы вычислительной техники при создании программного продукта, $t_{\text{ВТ}} = 99 * 8 = 792$ часа;

n – число задействованных ПЭВМ, $n = 1$.

Стоимость ПК 45000 руб., время использования 639.68 часа, тогда для него $C_{AM}(ПК) = (0,4*45000* 639,68 *1)/2384 = 4829,8$ руб. Стоимость принтера 8000 руб., его $F_d = 500$ час.; $HA = 0,5$; тогда его $CAM(Пр) = (0,5*8000*10*1)/500 = 80$ руб. Итого начислено амортизации 4 909,8 руб.

6.2.5. Расчет прочих расходов

В пункте «Прочие расходы» показаны расходы на реализацию проекта, которые не учтены в предыдущих статьях, их следует принять равными 10% от суммы всех предыдущих расходов, т.е.

$$C_{\text{проч.}} = (C_{\text{мат}} + C_{\text{зп}} + C_{\text{соц}} + C_{\text{эл.об.}} + C_{\text{ам}}) \cdot 0,1$$

Для нашего примера это

$$C_{\text{проч.}} = (2310 + 84867,63 + 25460,29 + 782,496 + 4\,909,8) \cdot 0,1 = 11833 \text{ руб.}$$

6.2.6. Расчет общей себестоимости разработки

Благодаря проведению ранее подсчетов по отдельным статьям затрат, мы можем вычислить общую плановую себестоимость проекта.

Таблица 6.8 – Смета затрат на разработку проекта

Статья затрат	Условное обозначение	Сумма, руб.
Расходы на материалы	$C_{\text{мат}}$	2310
Основная заработная плата	$C_{\text{зп}}$	84867,63
Отчисления в социальные фонды	$C_{\text{соц}}$	25460,29
Расходы на электроэнергию	$C_{\text{эл.}}$	782,496
Амортизационные отчисления	$C_{\text{ам}}$	4 909,8
Прочие расходы	$C_{\text{проч}}$	11833
Итого:		130163,24

Таким образом, затраты на разработку составили $C = 130163,24$ руб.

6.2.7. Расчет прибыли

Прибыль от реализации проекта в зависимости от конкретной ситуации (масштаб и характер получаемого результата, степень его определенности и коммерциализации, специфика целевого сегмента рынка и т.д.) может определяться различными способами. Если исполнитель работы не располагает данными для применения «сложных» методов, то прибыль следует принять в размере 5 - 20 % от полной себестоимости проекта. В нашем примере она составляет 13016,32 руб. (10 %) от расходов на разработку проекта

6.2.8. Расчет НДС

НДС составляет 18% от суммы затрат на разработку и прибыли. В нашем случае это $(130163.24 + 13016,32) * 0,18 = 25772.32$ руб.

6.2.9. Цена разработки НИР

Цена равна сумме полной себестоимости, прибыли и НДС, в нашем случае $ЦНИР(КР) = 130163.24 + 13016,32 + 25772.32 = 168951.88$ руб.

6.3. Оценка экономической эффективности проекта

Для того, чтобы проанализировать срез медицинских данных за месяц и построить прогноз на будущее требуется 50 рабочих часов. В 50 часов входит: 10 часов аналитика на выгрузку данных, 40 часов аналитика на подготовку данных (фильтрация, нормализация) и непосредственно анализ (построение моделей, прогнозирование, визуализация результатов). После разработки нашего программного обеспечения предполагается сокращение времени работы до 20 часов.

Средние затраты организации на аналитика возьмем равными 50000 руб. в месяц. Среднее количество часов в месяц 170. Час работы аналитика – 294 руб. Значит на выгрузку реестра тратится $294 * 50 = 14700$ руб. В год 176400 рублей. Так как наша задача сократить время работы аналитика до 20 часов, значит в год будет тратиться не 176400 рублей, а 70560 рублей.

На проект было потрачено 168951.88 рублей. Разработанный проект позволит сократить расходы на $(176400 - 70560) = 105840$ рублей. Проект может окупиться за 1,6 года.

Связан с повышением оперативности предоставления данных менеджерам актуального управляемого процесса.

Однако оценка данного эффекта выходит за рамки представлено ВКР.

Глава 7. Социальная ответственность

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8BM5B	Мохаммеджавад Бушра Джабер

Институт	Кибернетики	Кафедра	Программной инженерии
Уровень образования	Магистратура	Направление/специальность	09.04.01 Информатика и вычислительная техника

Исходные данные к разделу «Социальная ответственность»:

<p>1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения</p>	<p>Программа разрабатывается для работников в сфере медицины. Она будет полезна как врачам (для приведения данных из Талона амбулаторного пациента к формату CDA), так и работникам департамента здравоохранения (для анализа медицинских статистических данных).</p>
---	---

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<p>1. Производственная безопасность 1.1. Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности. 1.2. Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности. 1.3. Рекомендации по минимизации влияния</p>	<p>1.1. В качестве вредных факторов выделены :нарушение параметров микроклимата. 1.2. В качестве опасных факторов выделены: возможность поражения электрическим током и возникновение пожара. 1.3. Приведены рекомендации по улучшению микроклимата в офисном помещении, а также рекомендации по освещению, меры по обеспечению пожарной безопасности, способы защиты от электрического тока.</p>
<p>2. Экологическая безопасность: 2.1. Анализ воздействия на окружающую среду 2.2. Рекомендации по минимизации влияния на окружающую среду</p>	<p>2.1. Влияние объекта исследования на окружающую среду: программное обеспечение не может оказать влияния на окружающую среду, но оно разрабатывается на персональном компьютере ,который такое влияние оказать может. 2.2. Рассмотрена утилизация бумажных отходов и неисправных комплектующих ПК.</p>
<p>3. Безопасность в чрезвычайных ситуациях: 3.1. Перечень возможных ЧС на объекте 3.2. Меры по ликвидации ЧС и последствий</p>	<p>3.1. Основные ЧС в офисном помещении является возникновение пожара, также при хранении конфиденциальных данных в электронных таблицах можно говорить о</p>

	<p>возможности возникновения кибертерроризма.</p> <p>3.2.Приведены способы защиты от пожара и кибератак.</p>
<p>4. Правовые и организационные вопросы обеспечения безопасности:</p> <p>4.1. Психофизиологические факторы.</p> <p>4.2. Организационные мероприятия при компоновке рабочей зоны.</p> <p>4.3. Обеспечение гарантий защиты конфиденциальных данных граждан</p>	<p>4.1.Рассмотрены психофизиологические факторы,</p> <p>4.2.Организационные мероприятия при компоновке рабочей зоны</p> <p>4.3.Обеспечение гарантий защиты конфиденциальных данных граждан с помощью комплекса технических и юридических мер.</p>

Дата выдачи задания для раздела по линейному графику	01.04.2017
---	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент каф. ЭБЖ	Акулов Петр Анатольевич			

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ5В	Мохаммеджавад Бушра Джабер		

Введение

В ходе написания магистерской диссертации был разработан проект, который является математическим и программным обеспечением, предназначенным для анализа и обработки медицинских данных. Разработка программы велась исключительно при помощи компьютера. Сферы применения разработки очень широки: от нужд врачей до проведения анализа статистических медицинских данных работниками департамента здравоохранения. Независимо от конкретного применения, взаимодействие пользователя с разработанной программой в любом случае производится с помощью программных и аппаратных средств ПЭВМ, а также с помощью периферийных устройств, подключенных к ПЭВМ.

Данный раздел посвящен анализу вредных и опасных факторов производственной среды для операторов ПЭВМ и, в частности, для врачей, которые будут использовать данную программу; разработке программ по минимизации воздействия вредоносного и опасного влияния выявленных факторов, а также программ по снижению вредных воздействий на окружающую среду, экономии невос-

полнимых ресурсов и защите в чрезвычайных ситуациях.

7.1. Производственная безопасность

7.1.1. Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности.

7.1.1.1 Микроклимат

Микроклиматом производственных помещений называют климат внутренней среды помещений, определяющийся действующими на организм человека сочетаниями температуры, влажности и скорости движения воздуха. Наибольшее воздействие на микроклимат оказывают источники теплоты, которые находятся в помещении. Основными источниками теплоты в

описываемом помещении являются: персональные компьютеры, вспомогательное оборудование, приборы освещения, обслуживающий персонал.

В санитарных нормах установлены величины параметров микроклимата, которые создают комфортные условия. Нормы устанавливаются в зависимости от времени года, характера трудового процесса и характера производственного помещения.

Таблица 7.1. Оптимальные нормы параметров микроклимата

Период года	Параметр микроклимата	Величина
Холодный	Температура воздуха в помещении	22...24 °С
	Относительная влажность	40...60 %
	Скорость движения воздуха	до 0,1 м/с
Теплый	Температура воздуха в помещении	23...25 °С
	Относительная влажность	40...60 %
	Скорость движения воздуха	0,1...0,2 м/с

7.1.1.2 недостаточное искусственное освещение

Верно спроектированное и выполненное освещение должно обеспечивать высокую степень работоспособности, оказывать положительное влияние на работающий персонал, способствовать увеличению производительности труда.

К системе освещения необходимо предъявить требования:

- должный уровень освещенности рабочего места характеру выполняемой работы;
- достаточная равномерная яркость на рабочей поверхности и в окружающем пространстве;
- отсутствие резкой тени, отраженной и прямой блескости;
- постоянное освещение во времени;
- оптимальное направление светового потока;
- долговечное, экономичное, электро и пожаробезопасность, эстетичность, удобство в использовании.

По СанПиНу 2.2.1/2.1.1.1278-03 «Гигиенические требования к естественному, искусственному и совмещенному освещению жилых и общественных зданий» для того, что бы обеспечить освещенность рабочего места применяют уличное естественное освещение и при его недостатке подключают внутреннюю систему освещения.

По действующему Санитарному правилу и нормам 2.2.2/2.4.1340-03, искусственный свет в помещении для применения ПЭВМ необходимо осуществлять при помощи системы общего равномерного освещения. В производственном и административно-общественном помещении, в случае постоянной работы с документами, нужно использовать систему комбинированного освещения.

Уровень освещенности на поверхности стола в зонах размещения рабочих документов должна составлять 300 — 500 лк. Освещение не должно образовывать блики на поверхности экранов.

Необходимо ограничить неравномерное распределение яркости в сфере зрения пользователей ПЭВМ, при этом степень яркости между рабочей поверхностью не должен превышать соотношение 3:1 — 5:1, а между рабочей поверхностью и поверхностью стен и аппаратуры 10:1.

В дисплейном зале обычно, используют одностороннее естественное освещение. Для того, что бы снизить солнечную инсоляцию светопроем делают с северной, северо-восточной или северо-западной стороны. Монитор необходимо расположить как можно дальше от окна, и таким образом, чтобы окно располагалось сбоку.

Если экран дисплея располагается к окну, нужны определенные экранирующие устройства (светорассеивающая штора, регулируемые жалюзи, солнцезащитные пленки с металлизированной поверхностью)

О важности вопроса производственного освещения свидетельствует то, что условие деятельности оператора в системах "человек – машина" связывают с преобладанием зрительных данных – до 90% от всего объема.

Показатели соответствуют нормам СанПиН 2.2.2/2.4.1340–03 [34].

7.1.2. Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности.

7.1.2.1 Пожарная безопасность

Пожар приводит к полной потере данных и большим затратам для восстановления всех данных в прежнем объеме.

Источником воспламенения может быть: неисправность электропроводки, розетки и выключателя который может привести к короткому замыканию; применение поврежденного (неисправного) электроприбора; эксплуатация в помещениях электронагревательного прибора с открытым нагревательным элементом; появление пожара, после попадания молнии; не бережное обращение с огнем и пренебрежение мерам пожарной безопасности [43].

7.1.2.2 Электробезопасность

В связи с наличием электрооборудования для данного объекта характерным является возможность поражения электрическим током. Для снижения данного риска необходимо соблюдать нормы электробезопасности.

Электробезопасность является системой организационного и технического мероприятия и средства, которые обеспечивают защиту людей от влияния электрического тока, электрических дуг и статического электричества

Электрическая установка, к которой относят практически все оборудование ЭВМ, представляет для человека огромную потенциальную опасность, так как при эксплуатации человек может затронуть части, которые находятся под напряжением.

К опасным и вредным производственным факторам относят увеличенное значение напряжения в электрических цепях, замыкание которых может возникнуть через тело человека, увеличенное статическое электричество, электромагнитное излучение, повышенная напряженность электрических и магнитных полей. [41]

Помещение без повышенной степени опасности (сухое, хорошо отапливаемое, помещение с токонепроводящими полами, с температурой 18–20

С°, с влажностью 40–50%. Электробезопасность является опасным фактором и обычно она связана со следующими источниками:

- поражение электрическим током;
- статическое электричество;
- молниезащита.

Мероприятия защиты при электробезопасности следующие:

- отключать электрооборудование при его ремонте;
- периодически снимать электростатическое напряжение, касаясь пальцами рук, заземленных поверхностей;
- для безопасности во время гроз необходимо удостовериться о наличие молниеотвода, и того факта, что все розетки в кабинете заземлены.

В отличие от другого источника опасности, наличие электрического тока невозможно определить без необходимого оборудования и приборов, поэтому влияние его на человека зачастую происходит внезапно.

Следствия воздействия электрического тока на тело человека приведены в таблице 2.

Таблица 7.2 – Воздействие электрического тока на тело человека

Вид воздействия	Следствие	Виды электротравм
Термическое	Ожоги отдельных участков тела, нагрев внутренних органов	Электрический ожог, электрический знак, металлизация кожи.
Биологическое	Разложение и возбуждение живых тканей, судорожное сокращение мышц	Механические повреждения
Электролитическое	Разложение крови и других жидкостей, нарушение их физико-химического состава	Электрический удар

Каждое помещение, в зависимости от условия, делят на помещения:

- особо опасное;
- с увеличенной опасностью воздействия электричеством;

- без увеличенной опасности воздействия электрическим током.

В помещении, где осуществлялась работа, применяют прибор, потребляющий напряжение 220 В переменного тока с частотой в 50 Гц. Данное напряжение является опасным для жизни.

По классификации помещений в зависимости от опасности поражения электричеством, разработка ВКР проводится в помещении, в котором отсутствует повышенная опасность: то есть оно не характеризуется присутствием таких условий, как увеличенная влажность; повышенная температура; токопроводящая пыль; вероятность одновременного прикосновения к металлическим элементам, которые имеют соединение с землей и металлически корпусом электрооборудования.

7.1.3. Рекомендации по минимизации влияния

7.1.3.1 Рекомендации по улучшению микроклимата

К мероприятию по оздоровлению воздушной среды в производственных помещениях относят правильную организацию вентиляции и кондиционирования воздуха, а так же отопление помещения. Вентиляция осуществляется как естественным, так и механическим путём. В холодное время в помещении нужно предусмотреть систему отопления.

Объем помещения с ЭВМ не должен быть менее, чем 20 м³/человека.

Для того, что бы обеспечить комфортные условия, применяют как организационный метод, так и техническое средство.

7.1.3.2 Рекомендации по минимизации влияния освещения

Для того, что бы обеспечить требуемый уровень освещения в помещении применяют лампы дневного освещения. Такие лампы необходимо равномерно распределить по потолку помещения. Для освещения выбраны наиболее широко применяемые лампы типа ЛБ.

7.1.3.3 Меры по обеспечению пожарной безопасности

Одним из условий для обеспечения пожаробезопасности любого из производственных процессов является устранение возможного источника воспламенения. [44]

Для того, что бы устранить возможность пожара в помещении необходимо соблюдать данные противопожарные меры:

- ограничить объем горючего вещества;
- устранить возможные источники возгорания (электрические искры, нагрев оболочки оборудования);
- применять средств для тушения пожара;
- использовать пожарную сигнализацию;
- содержать электрооборудование в должном состоянии, использовать плавкие предохранители и автоматические выключатели в аппаратуре, по окончанию работы каждая установка должна обесточиться;
- присутствие в помещении средства пожаротушения (огнетушителя типа ОУ-3, пожарные инструменты, песок) и содержание их в должном виде;
- разрешать курения только в специальном для этого месте;
- содержать пути и проходы эвакуации людей в свободном виде;
- производить раз в год тренировки по пожарной безопасности;
- назначить ответственных за пожарную безопасность данного помещения.

7.1.3.4 Рекомендации по защите от электрического тока

Вредное и опасное влияние на людей электрического тока, электрических дуг и электромагнитного поля протекает в качестве электротравмы и профессионального заболевания.

Кроме того, при неисправном состоянии блока компьютера корпус может оказываться под воздействием тока, что ведет к электрической травме или электрическому удару. Для того, что бы устранить это, необходимо подсоединить металлический корпус оборудования к заземляющей жиле.

При появлении несчастных случаев необходимо немедленно освободить пострадавших от воздействия электрического тока и, вызвав врача, оказать ему первую помощь.

Для того, чтобы защититься от поражения электрическим током, каждую токоведущую часть нужно защитить от случайного прикосновения кожухом, корпус устройства необходимо заземлить. Заземление выполняют изолированными медными проводами с сечением в 1,5 мм². Его нужно присоединить к объединенной шине заземления с равным сечением в 5,4 мм² с помощью сварки. Общую шину нужно присоединить к заземлению. Его сопротивление не должно быть выше 4 Ом. Питание устройства необходимо осуществлять от силовых щитов посредством автоматического предохранителя, который работает при коротком замыкании нагрузки.

При использовании ЭВМ запрещено:

- эксплуатировать ЭВМ с неисправной защитой электропитания;
- отключать и подключать разъем кабеля электропитания и блока вентиляции при напряжении электросети;
- заменять съемный элемент под током;
- выполнять пайку аппаратуры, которая находится под током;
- убирать щиту, закрывающую доступ к токоведущей части;
- применять электроинструмент с напряжением 36В и выше с незаземленным корпусом.

При верном применении электроустановки и применении соответствующего средства защиты, сводится к минимуму риск поражения электрическим током.

Для того, чтобы предотвратить поражение электрическим током на предприятии нужно проводить мероприятия:

- компьютер подключать к сети только при помощи трехполюсной вилки, при этом центральный контакт необходимо надежно заземлить.
- при работе электрооборудования, рабочее место необходимо оборудовать так, что бы исключить возможность прикосновения служащего к

токоведущему устройству, шине заземления, батарее отопления, водопроводным трубам.

- обслуживающему персоналу необходимо пройти обучение технике безопасности на рабочем месте.

производится профилактическая проверка снятия напряжения, отключение неисправного электрооборудования и заземление

7.2. Экологическая безопасность

7.2.1 Анализ воздействия на окружающую среду

Охрана окружающей среды – это работа органов гос. власти РФ, органов гос. власти субъектов РФ, органов районного самоуправления, некоммерческих организаций и социальных объединений, нацеленная на восстановление и сбережение природной среды, рациональное внедрение и воспроизводство природных ресурсов, предотвращение неблагоприятного влияния домашней и другой деятельности на среду и ликвидацию ее результатов. Более интенсивной формой обороны окружающей среды от вредоносного влияния выбросов промышленных компаний считается абсолютный переход к безотходным, малоотходным и сохраняющим энергию технологиям и производствам.

При разработке каждой автоматической системы появляется надобность утилизировать производственные отходы, в качестве коих в предоставленном случае выступают бумажные отходы (макулатура) и неисправные подробности индивидуальных компов, платок, контроллеров.

7.2.2 Рекомендации по минимизации влияния на окружающую среду

Бумажные отходы обязаны переходить в надлежащие организации для последующей переработки во вторичные бумажные изделия. Неисправные комплектующие индивидуальных ПК обязаны переходить или муниципальным организациям, осуществляющим экспорт и устранение домашних и производственных отходов, или организациям, занимающимся переработкой отходов. Наиболее важными этапами работы с отходами считается их сбор, а в последующем переработка, утилизация и захоронение .[39]

Ещё одним из способов понижения бумажных отходов считается хранение данных на электронных носителях. Данное ПО позволяет хранить данные на жестком диске в электронном виде.

7.3. Безопасность в чрезвычайных ситуациях

7.3.1 Перечень возможных чрезвычайных ситуаций на объекте

При работе в кабинете могут возникнуть следующие классификации чрезвычайных ситуаций:

- преднамеренные/непреднамеренные;
- техногенные: взрывы, пожары, обрушение помещений, аварии на системах жизнеобеспечения/природные – связанные с проявлением стихийных сил природы.
 - экологические – это аномальные изменения состояния природной среды, такие как загрязнения биосферы, разрушение озонового слоя, кислотные дожди/ антропогенные – являются следствием ошибочных действий людей.
 - биологические – различные эпидемии, эпизоотии, эпифитотии;
 - комбинированные.

По скорости развития чрезвычайные ситуации могут быть: взрывные, внезапные, скоротечные, плавные. По масштабам распространения последствий: локальные, местные, территориальные. По возможности предотвращения: неизбежные (например, природные) и предотвращаемые (например, техногенные, социальные).

Одной из наиболее возможных ЧС в офисном здании является возникновение пожара. Причинами возникновения данного вида ЧС могут являться:

- возникновением короткого замыкания в электропроводке;
- возгоранием устройств ПЭВМ из-за неисправности аппаратуры;
- возгоранием устройств искусственного освещения;
- возгоранием мебели по причине нарушения правил пожарной безопасности, а также неправильного использования дополнительных бытовых электроприборов и электроустановок.[44]

7.3.2. Меры по ликвидации ЧС и их последствий

Пожарная безопасность предполагает надлежащее состояние объекта с исключением возможности возникновения очага возгорания (пожара) и его распространения в пространстве. Обеспечение пожарной безопасности — приоритетная цель для любого предприятия. Создание системы защиты регламентировано законодательством и нормативными документами различных ведомств.

Противопожарный режим для каждого предприятия считается индивидуальным. Он обязан отвечать общим требованиям, установленным законодательными актами, но с учетом специфичности работы.

Критерии хранения конфиденциальных данных в электронных таблицах. В киберпространстве могут быть использованы применены всевозможные методы для совершения кибертеракта:

- получение несанкционированного доступа к государственным и военным секретам, банковской и индивидуальной информации;
- нанесение ущерба отдельным физическим составляющим информационного пространства, к примеру, разрушение сетей электропитания, создание помех,
- использование специальных программ для разрушения аппаратных средств;
- кража или уничтожение информации, программ и технических ресурсов путем преодоления систем защиты, внедрения вирусов, программных закладок;
- воздействие на программное обеспечение и информацию;
- раскрытие и угроза публикации закрытой информации;
- захват каналов СМИ с целью распространения дезинформации, слухов, демонстрации мощи террористической организации и объявления своих требований;
- уничтожение или активное подавление линий связи, некорректная адресация, перегрузка узлов коммуникации;

- проведение информационно-психологических операций.

Направления для защиты от кибератак :

- мониторинг и реагирование на инциденты (включая поведенческий анализ пользователей);
- анализ исходных кодов ПО;
- сетевая безопасность;
- шифрование.

Основными причинами возникновения ЧС считаются: сложность технологий, недостаточная квалификация персонала, проектно-конструкторские недоработки, низкая технологическая дисциплина, а также внешние чрезвычайные ситуации: это стихийные бедствия.

7.4 Правовые и организационные вопросы обеспечения безопасности

7.4.1 Психофизиологические аспекты

К психофизиологическим вредным факторам относятся статические физические перегрузки, умственное перенапряжение, монотонность труда, эмоциональные перегрузки.

В ходе рабочего дня (смены) работнику должен быть предоставлен перерыв для отдыха и питания длительностью не более двух часов и не менее 30 минут, который в рабочее время не включается.

Время предоставления перерыва и его конкретная продолжительность устанавливаются правилами внутреннего трудового распорядка или по соглашению между работником и работодателем.

На работах, где по условиям производства (работы) предоставление перерыва для отдыха и питания невозможно, работодатель обязан обеспечить работнику возможность отдыха и приема пищи в рабочее время. Перечень таких работ, а также места для отдыха и приема пищи устанавливаются правилами внутреннего трудового распорядка.

Организация работы с ПЭВМ должна осуществляться в зависимости от вида и категории трудовой деятельности. Для предупреждения

преждевременной утомляемости пользователей ПЭВМ рекомендовано организовывать рабочую смену путем чередования работ с использованием ПЭВМ и без него. Во время регламентированных перерывов с целью снижения нервно-эмоционального напряжения, утомления зрительного анализатора, устранения влияния гиподинамии и гипокинезии, предотвращения утомления целесообразно выполнять комплексы упражнений.

7.4.2. Организационные мероприятия при компоновке рабочей зоны

Большое значение для профилактики статических физических перегрузок имеет правильная организация рабочего места человека, работающего с ПЭВМ. Рабочее место должно быть организовано в соответствии с требованиями стандартов, технических условий и (или) методических указаний по безопасности труда. Конструкция рабочего места и взаимное расположение всех его элементов (сиденье, органы управления, средства отображения информации и т.д.) должны соответствовать антропометрическим, физиологическим и психологическим требованиям, а также характеру работы.

Рабочие места с ПЭВМ при выполнении творческой работы, требующей значительного умственного напряжения или высокой концентрации внимания, рекомендуется изолировать друг от друга перегородками высотой 1,5 - 2,0 м. Экран видеомонитора должен находиться от глаз пользователя на расстоянии 600 - 700 мм, но не ближе 500 мм с учетом размеров алфавитно-цифровых знаков и символов.

Конструкция рабочего стула (кресла) должна обеспечивать поддержание рациональной рабочей позы при работе на ПЭВМ, позволять изменять позу с целью снижения статического напряжения мышц шейно-плечевой области и спины для предупреждения развития утомления.

Так же стоит принимать во внимание к размещению средств отображения информации ([Таблица 7.2](#)~~Таблица 7.2~~).

Таблица 7.2 – Расположение средств отображения информации

Тип средств отображения информации	Угол, градусы	
	В вертикальной плоскости	В горизонтальной плоскости
Очень часто используемые	± 15	±15
Часто используемые	±30	±30
Редко используемые	±60	±60

7.4.3. Обеспечение гарантий защиты конфиденциальных данных граждан

При заключении договора, подписывается «Соглашение о конфиденциальности и неразглашении информации». Его целью является урегулирование вопроса хранения, использования, разглашения и передачи конфиденциальной информации [37].

В документе «Соглашение о конфиденциальности» можно выделить следующие разделы:

- Предмет соглашения. Содержит цель и предназначение документа.
- Права и прямые обязанности сторон.
- Ответственность сторон. В данном пункте указаны факторы, при которых раскрывающая сторона несет юридическую ответственность за разглашение, использование либо несоблюдение условий хранения конфиденциальной информации. Учитываются пункты как умышленного, так и непреднамеренного пользования конфиденциальной информацией. А также ряд юридических санкций, в случае нарушения вышеуказанных положений.
- Условия раскрытия информации. Указываются строгие критерии и условия, при которых данная информация может быть раскрыта. Первым условием обязано стоять «обоюдное согласие сторон». Без данного фактора информация к разглашению должна быть запрещена.

Данный документ составлен на основе Федеральных законов[38][46].

Заключение

В рамках магистерской диссертации было проведено исследование состояния информатизации России, выявлен наиболее удобный и универсальный формат для хранения таких документов, проведен анализ статистических медицинских данных по Томской области.

Рассмотрены механизмы навигации по ЭМК. В результате проведенных исследований были предложены способы, основанные на представлении облака тегов, которое способствует более быстрому обеспечению врачей актуальной информацией о заболеваниях и систематизировать исходные данные.

По результатам исследования также было разработано и протестировано программное обеспечение для генерации медицинских документов ТАП, которые можно сохранить в формате .xml (по стандарту CDA). Для полученных данных были рассмотрены подходы и средства для их анализа и интерпретации (ГИС и Excel).

Для решения задачи формирования и мониторинга показателей здоровья населения региона был проведен анализ различных методов и алгоритмов Data mining для определения их применимости к решению задачи прогнозирования по медицинским данным.

В результате работ по накопленным статистическим данным по Томской области были разработаны алгоритмы прогнозирования числа больных артериальной гипертензией в зависимости от обеспеченности медицинских учреждений врачами (регрессионный анализ).

Также был проведен кластерный анализ методом K-means для разбиения медицинских учреждений Томской области по количеству пациентов с психическими заболеваниями, расходами на здравоохранение и обеспеченностью врачами.

Список публикаций студента

1. Разработка визуальных компонентов для задач быстрой навигации по ЭМК [Электронный ресурс] // Молодежь и современные информационные технологии. Сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии». Томск, 7-11 ноября 2016 г. – Томск: ТПУ, 2016 – Т. 2 – С. 262-263.
2. Современные способы организации электронной медицинской карты. Электронные средства и системы управления: материалы докладов XII Международной научно-практической конференции (16–18 ноября 2016 г.): в 2 ч. – Ч. 2. – Томск: В-Спектр, 2016. – С. 12-14.

Список использованной литературы

1. Барковская, Г. Ю. Информатизация деятельности работников сферы здравоохранения в целях повышения эффективности использования кадрового потенциала [Электронный ресурс] // Инженерный вестник Дона. – 2016. – Т. 41, № 2. – URL: http://ivdon.ru/uploads/article/pdf/IVD_37_Barkovskaya.pdf_589e2fb76a.pdf, свободный. – Загл. с экрана (дата обращения: 03.01.2017).
2. Электронный документооборот в медицине [Электронный ресурс] // Zdrav.ru : портал [информ. поддержки специалистов ЛПУ](http://www.zdrav.ru/articles/4293652235-qqq-16-m11-02-11-2016-elektronnyy-dokumentoorot-v-meditzine). – М., 2006–2017. – URL: <https://www.zdrav.ru/articles/4293652235-qqq-16-m11-02-11-2016-elektronnyy-dokumentoorot-v-meditzine>, свободный. – Загл. с экрана (дата обращения: 03.01.2017).
3. О персональных данных : Федеральный закон Рос. Федерации от 27 июля 2006 г. № 152-ФЗ // Российская газета. – 2006. – 29 июля.
4. Положение об обработке и защите персональных данных пациентов ООО «Фортуна» [Электронный ресурс] // Медицинский центр «Фортуна» : сайт. – Самара, 2012–2015. – URL: fortuna-mc.ru/docs/Polozhenie.doc, свободный. – Загл. с экрана (дата обращения: 11.03.2017).
5. Clinical Document Architecture (CDA) [Электронный ресурс] / Wayfarer1516 // Хабрахабр : сайт. – Опубл. 15 апр. 2015. – URL: <https://habrahabr.ru/post/255879/>, свободный. – Загл. с экрана (дата обращения: 22.05.2017).
6. Емелин, И. В. О стандартах электронного обмена медицинскими документами // Компьютерные технологии в медицине. 1996. – № 1. – С. 44–48.
7. Интегрированная распределенная информационная система лечебного учреждения [Электронный ресурс] / Я. И. Гулиев [и др.] // Программные продукты и системы. – 1997. – № 3. – URL:

<http://www.swsys.ru/index.php?page=article&id=1046>, свободный. – Загл. с экрана (дата обращения: 03.05.2017).

8. Интегрированная электронная медицинская карта [Электронный ресурс]. – Версия 24 – М., 2014. – 24 с. – (Описание стандартизованных электронных медицинских документов (СЭМД). – URL: portal-dev.rt-eu.ru/files/Описание%20структуры%20документов%20v1.2.odt, свободный. – Загл. с экрана (дата обращения:).

9. Николаев, Ю. А. Разработка SaaS решений для задач программы ГТО [Электронный ресурс] / Ю. А. Николаев, А. А. Пономарев; науч. рук. А. А. Пономарев // Технологии Microsoft в теории и практике программирования : сб. тр. XII Всерос. науч.-практ. конф. студентов, аспирантов и молодых ученых, г. Томск, 25–26 марта 2015 г. – Томск : Изд-во ТПУ, 2015. – С. 210–211. – URL: <http://www.lib.tpu.ru/fulltext/c/2015/C28/094.pdf>, свободный. – Загл. с экрана (дата обращения: 02.02.2017).

10. Бельшев, Д. В. Анализ методов хранения данных в современных медицинских информационных системах [Электронный ресурс] / Д. В. Бельшев, Е. В. Кочуров // Програм. системы: теория и приложения. – 2016. – № 2. – С. 85–103. – URL: http://psta.psisras.ru/read/psta2016_2_85-103.pdf, свободный. – Загл. с экрана (дата обращения: 08.02.2017).

11. В Минздраве заработал ситуационный центр [Электронный ресурс] (02.05.2014) // Моск. город. ун-т упр. Правительства Москвы : сайт. – М., 1994–2017. – URL: <http://mguu.ru/v-minzdrave-zarabotal-situatsionnyj-tsentr>, свободный. – Загл. с экрана (дата обращения: 08.02.2017).

12. Агеева, С. А. Формирование статистической отчетности с помощью OLAP-систем [Электронный ресурс] : магист. дис. по напр. подгот. (специальность) 09.04.01 Информатика и вычислительная техника / Агеева Софья Андреевна ; Томск. политехн. ун-т, каф. оптимизации систем упр. – Томск, 2016. – 129 с. URL: <http://portal.tpu.ru/cs/TPU181773.pdf>, свободный. – Загл. с экрана (дата обращения: 08.02.2017).

13. Кузьминых, Т. В. Формирование компетентности в области информационных и коммуникационных технологий у будущих государственных гражданских и муниципальных служащих : дис. ... канд. пед. наук : 13.00.08 / Кузьминых Татьяна Владиславовна. – Чебоксары, 2012. – 217 с.
14. Перлин, Ю. В. Карты в здравоохранении [Электронный ресурс] / Ю. В. Перлин, Ю. С. Товб // Пластиковые карты. – 5-е изд. – М. : БДЦ-пресс, 2005. – [13 с.]. – URL: http://www.pronit.ru/images/stories/pronit/documents/ru/products/Pronit_CMP_Publication_200501.pdf, свободный. – Загл. с экрана (дата обращения: 12.03.2017).
15. Соа. ITIL [Электронный ресурс] // StudFiles : сайт. – 2007–2017. – URL: <http://www.studfiles.ru/preview/1174781/page:38/>, свободный. – Загл. с экрана (дата обращения: 02.04.2017).
16. Фам, Ван Тап. Алгоритмические и программные средства интеграции данных при создании электронных медицинских карт : дис. ... кан. техн. наук : 05.13.11 / Фам Ван Тап. – Томск, 2011. – 186 с.
17. BPMN [Электронный ресурс] // Википедия. – 2001–2017. – URL: <https://ru.wikipedia.org/wiki/BPMN>, свободный. – Загл. с экрана (дата обращения: 01.03.2017).
18. Столбов, А. П. Современные проблемы применения ИКТ в здравоохранении. Медицинские информационные системы. Электронная медицинская карта [Электронный ресурс] : презентация докл. на XV Ежегод. специализир. конф., Москва, 9–10 окт. 2014 г. – [М., 2014]. –57 с. – http://www.med.ru/userfiles/files/ICT_3.pdf, свободный. – Загл. с экрана (дата обращения: 05.03.2017).
19. Дрожжинов, В. Применение продуктов и технологий EMC Documentum в сфере здравоохранения [Электронный ресурс]. Белая книга / В. Дрожжинов ; Центр компетенций по электронному правительству. – М., 2009. – 44 с. – URL: [130](http://www.egov-</p></div><div data-bbox=)

center.ru/files/Drozhhinov EMC Documentum health 2009.pdf/c17-19,

свободный. – Загл. с экрана (дата обращения: 08.03.2017).

20. Технологии текстового поиска [Электронный ресурс] // Зайцева И. В. Информационные технологии в государственном и муниципальном управлении : учеб. пособие. – Ставрополь : [Б. и.], 2011. – С. 79–91. – URL: <http://textarchive.ru/c-2931485-pall.html/>, свободный. – Загл. с экрана (дата обращения: 02.03.2017).

21. Пат. 2456661 Российская Федерация, МПК G06F 3/048; G09G 5/32. Эффективная навигация результатов поиска [Электронный ресурс] / Джхавери Вивек, Рэгно Роберт ; Майкрософт Корпорейшн. – № 2009109199/08 ; Заявл. 20.08.2007; Опубл. 20.07.2012. – Доступ с сайта ФГУ ФИПС, свободный. – Загл. с тит. экрана. (дата обращения: 08.03.2017).

22. Облако тегов [Электронный ресурс] // Википедия. – 2001–2017. – URL: https://ru.wikipedia.org/wiki/Облако_тегов, свободный. – Загл. с экрана (дата обращения: 01.03.2017).

23. Чедвик Д. ASP.NET MVC 4: разработка реальных веб-приложений с помощью ASP.NET MVC / Джесс Чедвик, Тодд Снайдер, Хришикеш Панда. – М. : Вильямс, 2013. – 432 с.

24. Томск [Электронный ресурс] // Академик : сайт. – 2000–2016. – URL: <http://dic.academic.ru/dic.nsf/ruwiki/2413#1>, свободный. – Загл. с экрана (дата обращения: 01.03.2017).

25. Томская область [Электронный ресурс] // Википедия. – 2001–2017. – URL: https://ru.wikipedia.org/wiki/Томская_область, свободный. – Загл. с экрана (дата обращения: 01.03.2017).

26. Data Mining [Электронный ресурс] : информация // Интуит : сайт. – 2003–2017. – URL: <http://www.intuit.ru/studies/courses/6/6/info>, свободный. – Загл. с экрана (дата обращения: 01.03.2017).

27. Основы линейной регрессии // Портал знаний : сайт. – 2017. – URL: <http://statistica.ru/theory/osnovy-lineynoy-regressii/>, свободный. – Загл. с экрана (дата обращения: 29.03.2017).

28. Искусственная нейронная сеть [Электронный ресурс] // Википедия. – 2001–2017. – URL: https://ru.wikipedia.org/wiki/Искусственная_нейронная_сеть, свободный. – Загл. с экрана (дата обращения: 22.04.2017).
29. Чучуева, И. Создаем нейронную сеть для прогнозирования временного ряда [Электронный ресурс] // Математическое бюро : сайт. – 2010–2017. – URL: <http://www.mbureau.ru/blog/sozdaem-neyronnuyu-set-dlya-prognozirovaniya-vremennogo-ryada>, свободный. – Загл. с экрана (дата обращения: 03.05.2017).
30. Золин, А. Г. Применение нейронных сетей в медицине / А. Г. Золин, А. Ю. Силаева // Актуальные проблемы науки, экономики и образования XXI века : материалы II Междунар. науч.-практ. конф. – Самара : Сам. ин-т (фил.) РГТЭУ, 2012 – Ч. 2. – С. 264–271.
31. Геоинформационная система [Электронный ресурс] // Википедия. – 2001–2017. – URL: https://ru.wikipedia.org/wiki/Геоинформационная_система, свободный. – Загл. с экрана (дата обращения: 01.05.2017).
32. Тематические карты [Электронный ресурс] / streetjournal // Geektimes : сайт. – Оpubл. 26 апр. 2011. – URL: <https://geektimes.ru/post/118169/>, свободный. – Загл. с экрана (дата обращения: 22.05.2017).
33. ГОСТ 12.1.005-88. Общие санитарно-гигиенические требования к воздуху рабочей зоны. – М.: Издательство стандартов, 1998. // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/1200003608> (дата обращения: 11.03.2017).
34. ГОСТ 12.0.002-80. Система стандартов безопасности труда. Термины и определения. – М.: ИПК Издательство стандартов, 2002 .
35. СНиП 23-05-95. Естественное и искусственное освещение. – М.: Минстрой России, 1995.
36. Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 3.07.2016) // Электронный фонд правовой и нормативно-технической

документации. URL: <http://docs.cntd.ru/document/901807664> (дата обращения: 11.03.2017).

37. СанПиН 2.2.2/2.4.1340-03. Санитарно-эпидемиологические правила и нормы. Гигиенические требования к персональным электронно-вычислительным машинам и организации работы // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/901865498> (дата обращения: 10.03.2017).

38. Федеральный закон от 29.07.2004 №98-ФЗ (ред. от 11.07.2011) «О коммерческой тайне». Федеральный закон «О защите персональных данных» ФЗ-152.

39. ГОСТ 12.0.003-74. ССБТ Опасные и вредные производственные факторы. Классификация. – М.: ИПК Издательство стандартов, 2002.

40. СН 2.2.4/2.1.8.566-96. Производственная вибрация, вибрация в помещениях жилых и общественных зданий. – М.: Минстрой России, 1997.

41. ГОСТ Р 12.1.019-2009 ССБТ. Электробезопасность. Общие требования и номенклатура видов защиты // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/1200080203> (дата обращения: 11.03.2017).

42. СНиП 21-01-97. Пожарная безопасность зданий и сооружений // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/871001022> (дата обращения: 11.03.2017).

43. Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 3.07.2016) // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/901807664> (дата обращения: 11.03.2017).

44. ГОСТ 22.0.07-95. Безопасность в чрезвычайных ситуациях. Источники техногенных чрезвычайных ситуаций. Классификация и номенклатура поражающих факторов и их параметров. – М.: Госстандарт России, 1995.

45. ГОСТ Р 55090-2012 Ресурсосбережение. Обращение с отходами. Рекомендации по утилизации отходов бумаги. // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/1200103182> (дата обращения: 11.03.2017).

46. Федеральный закон «Об информации, информационных технологиях и о защите информации».

Приложение А

(обязательное)

Раздел ВКР, выполненный на иностранном языке

ANALYTICAL REVIEW

Студент:

Группа	ФИО	Подпись	Дата
8ВМ5В	Мохаммеджавад Бушра Джабер		

Консультант кафедры ПИ _____:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент каф. ПИ	Чердынцев Е. С.	к.п.н.		

Консультант – лингвист кафедры ИЯ _____:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
зав. каф. ИЯ	Сидоренко Т. В.	к.п.н.		

Chapter 3 Development of software solutions for the formation and monitoring of health indicators in the region

3.1 The analysis and forecasting of health indicators of the population and human environment

Numerous studies have shown that the factors that determine health are:[10]

- biological (heredity, type of higher nervous activity, constitution, temperament, etc.);
- natural;
- the state of the environment;
- socio-economic;
- level of development of health care.

To predict the health of the population of a particular region, it is especially important to know information about natural, socio-economic factors, the state of the environment and the level of health development in the region. Consider these factors in more detail.

Natural factors include: climate, landscape, flora, fauna and others.

For Tomsk region is characterized by a continental and sharp-continental climate with hot summers and cold winters (for example, for Tomsk the minimum recorded temperature is -55°C , and the maximum is $+37,7^{\circ}\text{C}$).

Flora and fauna on the territory of the Tomsk region is represented by about 1,000 species of higher plants and about 2,000 species of animals, most of which are insects.

This determines some diseases characteristic of the Tomsk region due to the large number of ticks (Lyme disease, encephalitis).

The state of the environment also has a great impact on human health. Thus, Tomsk is characterized by a not very favorable environmental situation.

According to the Committee for Environmental Protection of Tomsk and Tomsk Region for 2007, atmospheric air is polluted almost over the entire territory of the city.

In most of it, located around urban forests and away from industrial enterprises, there is an increased level of pollution. The atmospheric content of carbon monoxide also exceeds the maximum permissible standards.

In the places of car clustering - intersections - the oxide concentration exceeds the maximum permissible values by dozens of times.

In Tomsk, there are 4 industrial enterprises - Tomsk TPP-3, Tomsk GRES-2, CJSC "Methanol" and JSC "Tomsk Petrochemical Plant", which produced 13.5 thousand tons of emissions per year.

The quality of water in many bodies of water, popular as bathing places, often does not meet sanitary standards.

The social and economic factor and the level of health in the Tomsk region is characterized by the following:

In Tomsk, there is a fairly high level of educated people (which influence their understanding of health problems, timely access to specialists and others).

Also the number of medical institutions and the level of specialists in them are also quite high. But in small villages of the Tomsk region the situation is much worse.

The prediction of the health status of the population, depending on these factors, are very complex and must take into account each of these factors.

So, for example, cold diseases are characteristic for the spring and autumn seasons, when the air temperature drops.

Forecast of diseases such as Lyme disease and tick-borne encephalitis are seasonal (spring-summer).

Consequently, it is possible to predict health indicators of the population depending on the factors described above with the help of regression analysis or in the analysis of time series and the allocation of seasonal features.[11]

3.2 Predicting the need for medical care on the basis of data on the level and nature of the incidence of various nosological forms, population groups, territorial and other characteristics

An analysis was made of various methods and algorithms of Data mining to determine their applicability to the task of predicting the number of people with different diseases by region in the future.

The following methods were considered:

- Linear regression,
- Decision trees,
- The method of k-nearest neighbors,
- neural networks,
- Method of support vectors.

As a result, the following data prediction methods were chosen to solve the prediction problem:

3.2.1 Linear regression (a special case of time series analysis)

To predict the number of people with a certain diagnosis in the future, you can build a linear regression model based on existing values.

A mathematical equation that estimates the line of simple (pairwise) linear regression:[12]

$$Y=a+bx.$$

x Is called an independent variable or a predictor.

Y – Dependent variable or response variable. This is the value that we expect for y (on average), if we know the value » those. This is the "predicted value of y"

Dependent variable

A – Free member (intersection) of the evaluation line; This value of Y, when

b – угловой коэффициент или градиент оценённой линии; Angular coefficient or gradient of the estimated line ‘It is the amount by which Y increases on average, if we increase x by one unit

a и b Are called regression coefficients of the estimated line, although this term is often used only for b .

In our case, Y is the predicted number of people with a certain diagnosis, x is the forecast date.

To calculate the parameters a , b , you need to use data of the form:

Data	Number of people with this diagnosis
------	--------------------------------------

It is possible to divide all patients with the diagnosed diagnosis by region (address) and carry out a similar analysis for each region.

As an example, let us consider the application of regression analysis to the problem of predicting the number of patients with Arterial Hypertension (AH).

To determine the predictions for identifying patients with AH the dependence of the number of cases of AH detection on the completeness of medical facilities by primary care physicians was used.

The analysis was carried out for the output variable Y - " AH of the whole population of the region (per 100 thousand people)" depending on factor X - "Provision of therapists (for 10 thousand people)."

The validity and adequacy of the model is checked for the significance level $\alpha = 0.05$.

Factors associated with the genome, lifestyle and environmental situation are attributed to the random component of ε and are not taken into account.

As initial data, response and factor factors for Tomsk Oblast were used.

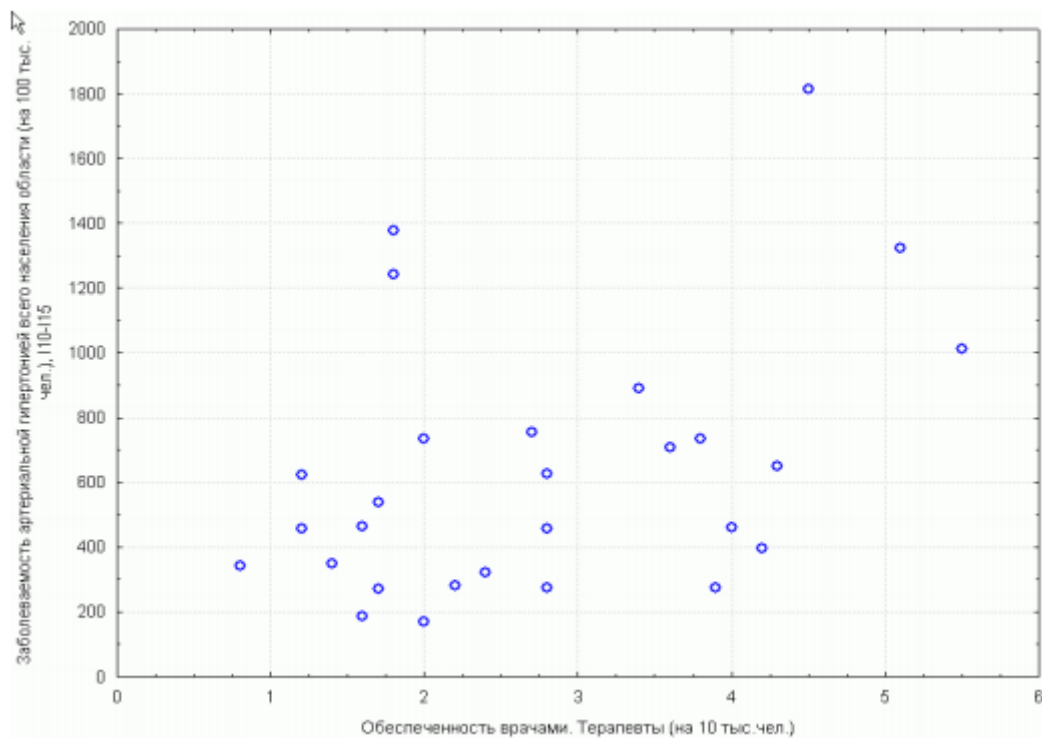


Figure 3.1 – Distribution of medical institutions by the availability of therapists and the number of patients with Arterial Hypertension

After that, the correlation matrix of the response and the factor is constructed, which indicates that the greatest correlation coefficient with Y is the variable X-provision by doctors (0.413657).

Correlations of Vectors in Design Matrix X (Выборка(03)-2010)
Correlation matrix for the vectors in the design matrix X

Effect	Col. 2 Обеспеченность врачами. Терапевты (на 10 тыс.чел.)	Col. 3 Заболееваемость острым инфарктом миокарда всего населения области (на 100 тыс. чел.), 121	Col. 4 Заболееваемость инсультами всего населения области (на 100 тыс.чел.), 160-164	Col. 5 Смертность от острого инфаркта миокарда среди всего населения области (на 100 тыс. чел.), 121	Col. 6 Смертность от мозгового инсульта среди всего населения области (на 100 тыс. чел.), 160-164	Col. 7 Заболееваемость артериальной гипертонией всего населения области (на 100 тыс.чел.), 110-115
Intercept						
Обеспеченность врачами. Терапевты (на 10 тыс.чел.)	1,000000	0,225032	0,017760	-0,256067	0,025353	0,413657
Заболееваемость острым инфарктом миокарда всего населения области (на 100 тыс.чел.), 121	0,225032	1,000000	0,299255	0,206834	-0,112680	0,396039
Заболееваемость инсультами всего населения области (на 100 тыс.чел.), 160-164	0,017760	0,299255	1,000000	0,220558	0,451370	0,360208
Смертность от острого инфаркта миокарда среди всего населения области (на 100 тыс.чел.), 121	-0,256067	0,206834	0,220558	1,000000	0,166654	-0,083257
Смертность от мозгового инсульта среди всего населения области (на 100 тыс.чел.), 160-164	0,025353	-0,112680	0,451370	0,166654	1,000000	0,026704
Заболееваемость артериальной гипертонией всего населения области (на 100 тыс.чел.), 110-115	0,413657	0,396039	0,360208	-0,083257	0,026704	1,000000

Figure 3.2 – The correlation matrix

The obtained coefficients of linear regression are used to construct its model to determine the dependence of the incidence of AΓ on the number of specialists in the field of medicine.

The regression forecast model for the indicator selected at the significance level $\alpha = 0.05$ has the following form:

$$Y = 275,2280 + 130,4859 X .$$

To test the hypothesis of insignificance of the regression model, an analysis of variance is used, which confirmed that the regression equation is significant.

The linear regression equation can be used to predict the change in the output variable.

3.2.2 Neural networks

The neural network is a mathematical model, as well as its software or hardware embodiment, built on the principle of the organization and functioning of biological neural networks - nerve cell networks of a living organism.

Currently, neural networks can be used to solve various tasks, for example classification, clustering or forecasting, so they were chosen to predict time series of changes in the number of people with certain diseases.

Neural networks are able to make predictions because of their ability to isolate the hidden dependencies between input and output data.

To solve certain tasks, the neural network is trained on a training sample, after which the trained network is able to predict the output values based on previous values. The only feature of using neural networks for prediction tasks is the presence of this very dependence between previous changes and future ones.

For example, it makes sense to predict stock prices based on historical data (revealing periodicity, bursts and decay, but predicting the results of lotteries will not have any particular results, since the data in the lotteries are random.

Neural networks can be divided into networks that require training with the teacher (perceptions) and networks with mixed learning (networks of radial-basis functions) [13].

As an example, let us consider the use of neural networks for the medical prognosis of finding a patient in intensive care.

There is a problem of shortage of resuscitation chambers in medical institutions, in connection with which long queues are characteristic for cardiosurgery departments. At the same time, an increase in the number of resuscitation chambers requires a very large amount of money.

Therefore, it is advisable to use the existing chambers more effectively. There is a category of patients with a severe state of health who require a long stay in the intensive care ward (more than two days).

In this regard, surgeons will not be able to perform some operations, since patients after surgery will have nowhere to put.

Therefore, severe patients should be operated before weekends or holidays, when operations are not performed (surgeons rest, and seriously ill patients can recover in intensive care).

Consequently, people with better health and frivolous operations can be operated at the beginning of the workweek, since they will need a resuscitation ward for a short time.

Then beds in intensive care will be released faster, and the number of operated will increase.

D. Tu and M. Guerir (University of Toronto) suggested using neural networks to predict the patient's condition after surgery, which would allow for more optimal use of resuscitation chambers.

As the initial data for the neural network, the patient's known information about the patient was used in the preoperative period

As neuron network architecture, a two-layer perceptron was selected, which divided patients into three risk groups. At the same time, factors such as age, sex, functional condition of the left ventricle, the degree of complexity of the forthcoming operation and the presence of concomitant diseases were taken into account.

The result of the work of the neural network was as follows: only 16.3% of patients assigned to the group of low-risk delay in intensive care spent more than two days, and more than 60% of patients in the high-risk group spent much time in intensive care units. [14]

3.2.3 Cluster analysis

One of the approaches of Data mining is cluster analysis (or training without a teacher).

In this case, the initial set of health data by the given characteristics (the number of patients with a definite diagnosis, the number of doctors and others) is divided into groups of similar, closely related objects - clusters.

In the future, the results of clustering are interpreted.

For cluster analysis, we will select the following set of characteristics for medical institutions in Tomsk Oblast:

Expenditure on health care from all sources of financing. Amount for 1 person. (rub.);

- Provision of doctors (for 10 thousand people);
- Provision of average medical personnel (per 10 thousand people);
- Incidence of mental disorders (per 100 thousand population).

On the basis of our research, the Euclidean metric was used as the metric for determining the distance in the feature space between the objects under study:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}.$$

For the cluster analysis, various algorithms can be used, depending on the characteristics of the input data, for example, K-means, hierarchical algorithm or DBSCAN.[15]

A hierarchical clustering algorithm for medical data was chosen (it allows us not to determine in advance the number of finite clusters, but rather to estimate the most suitable partition in the dendrogram).

Thus, as a result, a tree-like structure (dendrogram) was obtained, along which finite clusters are determined.

Preliminarily, the normalization was carried out-the conversion of characteristics to dimensionless quantities.

This is especially important if the characteristics are presented in different units of measurement.

The result of the distribution of medical institutions in the Tomsk region, depending on the features identified, is shown below.

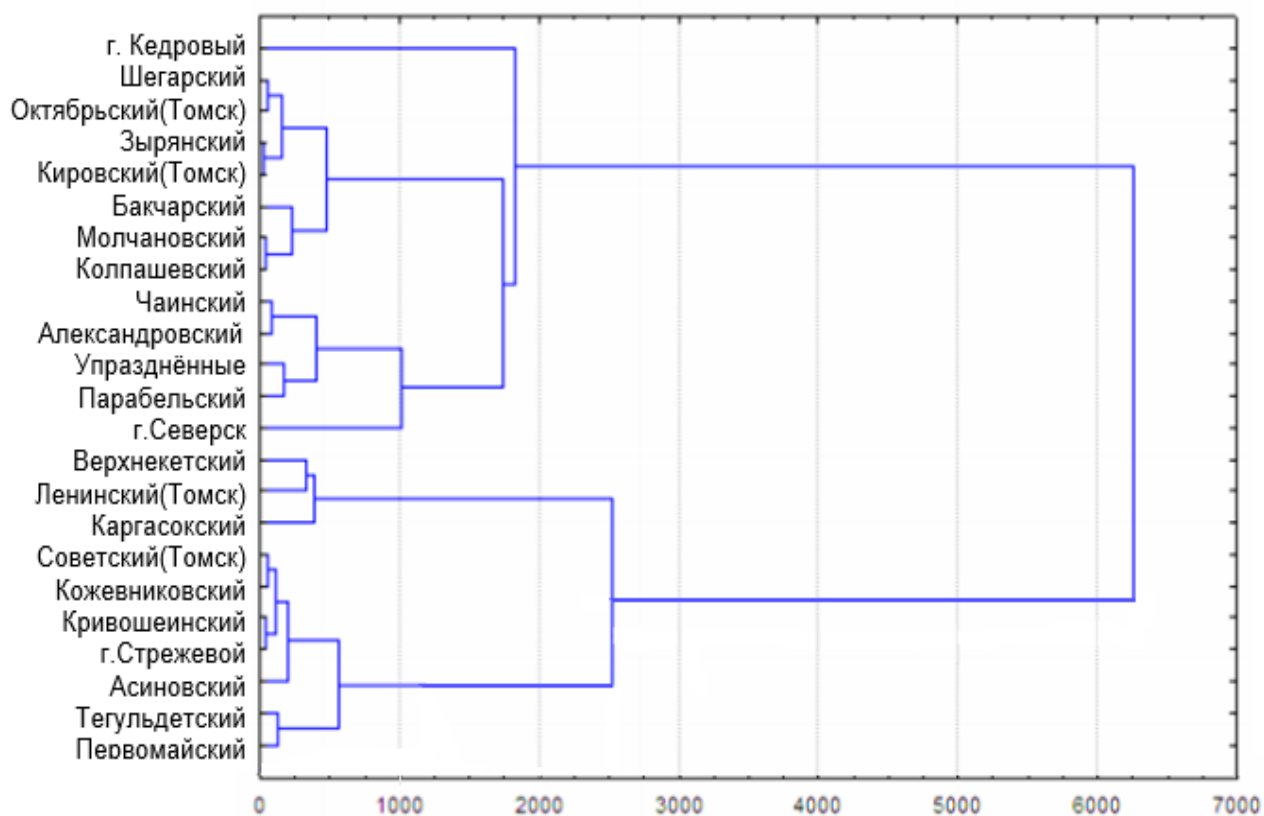


Figure 3.2 – Dendrogram of the division of municipal formations of the Tomsk Region into clusters

Table 3.1 Results of cluster analysis of municipalities of the Tomsk region

Муниципальные образования Томской области	Кластер	Муниципальные образования Томской области	Кластер
г.Кедровый	1	Молчановский	1
г.Северск	1	Ленинский (Томск)	2
Александровский	1	Верхнекетский	2
Шегарский	1	Каргасокский	2
Кировский (Томск)	1	Советский (Томск)	3
Октябрьский (Томск)	1	Первомайский	3
Зырянский	1	Кожевниковский	3
Парабельский	1	Асиновский	3
Колпашевский	1	Кривошеинский	3
Бакчарский	1	Тегульдетский	3
Упразднённые	1	г.Стрежевой	3
Чаинский	1		