

**Аннотация проекта (ПНИЭР), выполняемого в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014 – 2020 годы»**

**Номер соглашения о предоставлении субсидии (государственного контракта)**  
14.607.21.0123

**Название проекта**

Разработка интернет-технологии для персонализированной поддержки здоровьесбережения

**Тематическое направление**

Информационно-телекоммуникационные системы

**Исполнитель**

Федеральное государственное учреждение "Федеральный исследовательский центр "Информатика и управление" Российской академии наук"

**Цели и задачи исследования**

Цель исследования.

Исследование научно-технических решений в области новой информационной интернет-технологии для персонализированной поддержки процессов здоровьесбережения на основе анализа больших массивов данных о здоровье и определяющих его факторах.

Задачи исследования:

Разработка принципа формирования интегрированного паспорта здоровья, учитывающего данные о показателях здоровья и определяющих его факторах.

Разработка модели интегрированного паспорта здоровья.

Создание метода получения и занесения в интегрированный паспорт здоровья данных из источников разного типа: социальных сетей; историй болезни; баз данных Росстата; информации, предоставляемой непосредственно пациентом.

Создание метода интеллектуальной обработки медицинских данных и текстов для выявления факторов, детерминирующих здоровье.

Создание метода извлечения из интернет-пространства и структуризации информации о психологических и мотивационных характеристиках здоровьесбережения, о стандартах применения и эффективности здоровьесберегающих технологий.

Разработка интернет-технологии персонализированной поддержки здоровьесбережения.

Разработка архитектуры интернет-системы персонализированной поддержки здоровьесбережения.

Создание опытного образца банка данных интегрированных паспортов здоровья.

Создание опытного образца интернет-системы персонализированной поддержки здоровьесбережения.

**Актуальность и новизна исследования**

По имеющимся оценкам Всемирного Банка и прогнозам Минэкономразвития медико-демографическая ситуация в России является одним из главных препятствий эффективному социально-экономическому развитию страны в настоящем и будущем. Несмотря на наблюдающиеся в последние годы позитивные сдвиги, отставание России по показателям здоровья не только от

развитых, но и от и многих развивающихся стран мира в настоящее время очень велико и продолжает возрастать. По данным ВОЗ, в 2013 г. ожидаемая продолжительность жизни (ОПЖ – главный показатель здоровья нации) была, по сравнению со странами-рекордсменами, ниже на 19 лет у мужчин и на 12 лет у женщин. Среди 194 стран – членов ВОЗ Россия занимала 144 место по ОПЖ мужчин и 106 по ОПЖ женщин. Это значительно хуже показателей 1990 г. (106 место у мужчин и 63 - у женщин). По ожидаемой продолжительности здоровой жизни (ОПЗЖ) мужчин в 2013 г. Россия разделила места с 144-го по 146-е с Руандой и Сенегалом. По смертности мужчин в возрастной группе 60-64 года Россия в 2013 г. заняла первое место среди всех 194 стран – членов ВОЗ, т.е. имела самую высокую смертность в мире в данной возрастной группе! Проблемы со здоровьем обуславливают огромные потери, которые несет как страна в целом, так и каждый отдельно взятый россиянин. Увеличение ОПЖ на 1 год дает прирост ВВП на 4%. Таким образом, весьма актуальной представляется задача создания на основе современных информационных технологий средства практического обеспечения здоровьесбережения.

## **Описание исследования**

Объектом исследования является совокупность разноплановых и взаимодополняющих массивов данных, характеризующих персональное здоровье человека (медицинские карты, специализированные сайты, социальные сети, информация, непосредственно предоставляемая пациентом и др.), а также совокупность технологий здоровьесбережения.

Предметом исследования является: разработка интегрированного информационного портрета здоровья человека и методов извлечения информации для его создания из разнородных источников информации; разработка и наполнение баз данных, содержащих эти портреты, опираясь на высокопроизводительные вычислительные технологии; разработка методов структурирования, анализа и классификации этих данных, а также методов персонализированной оптимизации предлагаемых пациенту технологий здоровьесбережения; разработка программных средств, реализующих данные методы, а также специализированного Web-портала; разработка на этой основе интернет-технологий для персонализированной поддержки здоровьесбережения и экспериментальное исследование эффективности данных технологий; подготовка технической документации для промышленных образцов созданных технологий.

Основными планируемыми функциями создаваемой интернет-системы персонализированной поддержки здоровьесбережения являются: сбор информации из различных источников, включая интернет-пространство; интеллектуальная обработка медицинских данных и текстов; анализ эффективности технологий здоровьесбережения; оценка проблемных зон в здоровье конкретного человека; персонализированная оптимизация технологий

его здоровьесбережения; помощь в их применении и мониторинг их эффективности.

Для их обеспечения в рамках проекта решаются следующие главные задачи:

Разработка модели и программная реализация банка данных о показателях здоровья человека и детерминирующих здоровье факторах разных классов. Новизна данной задачи заключается в не имеющей аналогов системности и полноте описания клиента - в базе данных будет представлена максимально полная и всесторонняя характеристика здоровья человека, а также максимально полная характеристика влияющих на его здоровье факторов разного типа.

Разработка метода извлечения из источников разного типа, распознавания, стандартизации и занесения в банк данных о показателях здоровья и детерминирующих его факторах разных классов. Данный метод должен базироваться на когнитивных технологиях и позволять извлекать данные из источников разного типа: социальных сетей, баз данных Росстата и др. Новизна этой задачи связана также с не имеющей аналогов полнотой, специфичностью и разнообразием типов обрабатываемых данных.

Разработка метода извлечения из интернет пространства и анализа данных о мотивациях и отношении людей к тем или иным детерминирующим здоровье средствам и факторам, в зависимости от характеристик окружающей среды и образа жизни, а также свойств личности человека. Результаты, полученные с помощью данного метода необходимы для решения задач выработки персонализированных рекомендаций по оптимизации личных программ профилактики и здоровьесбережения на основании результатов анализа эффективности действия детерминирующих здоровье факторов на продолжительность периода активной жизни человека.

Разработка метода интеллектуальной обработки больших данных о показателях здоровья и определяющих его факторах разного типа для выявления факторов, существенно влияющих на здоровье, оценки интенсивности данного влияния и достоверности данной оценки. Данный метод должен обеспечивать с помощью алгоритмов машинного обучения выявление зависимостей между характеристиками образа жизни человека и др. влияющими на здоровье факторами и изменениями показателей здоровья, а также определять величину долговременных эффектов изменений здоровья под действием этих факторов. В качестве методов лингвистического анализа текстовой информации должен использоваться реляционно-ситуационный анализ текстов, хорошо зарекомендовавший себя в решении задач автоматической обработки текстов.

## **Результаты исследования**

В настоящее время осуществляется реализация задач 2-го этапа проекта. Наиболее значимые результаты, полученные к данному моменту:

- аналитический обзор современной научно-технической, нормативной, методической литературы, затрагивающей научно-техническую проблему, исследуемую в рамках ПНИЭР, включающий методы машинного обучения, применяемые в области здоровьесбережения, принципы формирования и модели паспорта здоровья, методы извлечения, сбора и обработки данных о здоровье пациентов;
- результаты патентных исследований;
- модель интегрированного паспорта здоровья;
- результаты анализа существующих программных и аппаратных платформ, соответствующих требованиям к испытательному стенду для экспериментальных исследований;
- выявление потенциальных пользователей разрабатываемой интернет-системы персонализированной поддержки здоровьесбережения и анализ их потребностей;
- реестр источников тестовых данных для экспериментальных исследований;
- метод оценки эффективности задач машинного обучения в области здоровьесбережения;
- метод получения и занесения в интегрированный паспорт здоровья данных из историй болезни;
- структура панели психологических характеристик, извлекаемых из интернет-пространства;

- метод дистанционного получения и занесения в интегрированный паспорт здоровья данных от пациента;
- метод интеллектуальной обработки медицинских данных и текстов;
- метод анализа эффективности здоровьесберегающих технологий на основе данных, извлекаемых из интернет-пространства;
- метод извлечения из интернет-пространства и структуризации информации о психологических и мотивационных характеристиках здоровьесбережения;
- интернет-технология поддержки здоровьесбережения, основанная на получении он-лайн информации о показателях здоровья и детерминирующих его факторах и автоматизированного построения оптимальных программ здоровьесбережения;
- архитектура интернет-системы персонализированной поддержки здоровьесбережения.

Новизна полученных результатов Проекта определяется преодолеваемыми им недостатками как стратегического подхода государства к решению проблемы здоровьесбережения, так и недостатками имеющихся на рынке технологий в данной области. Проект направлен на преодоление таких проблем и недостатков, как:

- слабая персонализация (государственные программы, центры здоровья, интернет-порталы и т.п. предлагают унифицированные подходы и средства здоровьесбережения);
- слабая актуализация (предлагаются устаревшие методы и подходы);
- ненадежность предлагаемых средств и методов оздоровления (за счет ориентации фирм прежде всего на коммерческий эффект);
- избыточность и «замусоренность» информации о здоровье (пользователь тонет в море информации и не в состоянии оценить ее достоверность и выбрать то, что нужно лично для него).

## **Практическая значимость исследования**

Несомненно, главная область применения результатов работы – это использование разработанных методов и Системы в целом по прямому назначению для улучшения здоровья населения России, поскольку созданный инструментарий предоставит более широкие возможности для постановки задач профилактики и здоровьесбережения на более эффективную оптимизированную, персонализированную доказательную базу. Создаваемая система сможет применяться как для страны в целом, так и для отдельных организованных по отраслям или регионам контингентов населения. Алгоритмы извлечения информации из социальных сетей смогут быть использованы социальными службами, органами здравоохранения, страховыми компаниями для получения данных об отношении людей к предлагаемым на рынке и в государственных структурах технологиям здоровья, об их предпочтениях, об эффектах применения предлагаемых технологий.

Возможны следующие пути использования результатов проекта в реальном секторе экономики, а также вовлечения их в хозяйственный оборот:

- совместная деятельность с предприятиями реального сектора экономики по адаптации Системы в целях укрепления здоровья и работоспособности персонала;

- совместная деятельность с производителями товаров и услуг для рынка здоровья по совершенствованию оценок эффективности продукта;

- продажа обезличенной информации, содержащейся в Системе, заинтересованным коммерческим структурам, например фармкомпаниям, производящим продукты и услуги для рынка здоровья;

- продажа патентов или самой Системы за рубежом, что вполне возможно, поскольку рынок здоровья очень чувствителен и восприимчив ко всему новому.

Рынок продуктов и сервисов для здоровья в настоящее время является одним из самых быстроразвивающихся в России и в мире. Для России проекты в области здравоохранения на 2013-2020 годы оцениваются в 857 млрд. рублей или 13,5% от расходов на приоритетные проекты.