

Программа «Информационно-аналитические системы»

Позин Борис Аронович
д.т.н., профессор НИУ ВШЭ,
Технический директор ЗАО «ЕС-лизинг»



Группа компаний ЕС-лизинг (ЕСЛ)



- Группа ЕСЛ включает компании **ЕС-лизинг** и **ЕС-лизинг СОФТ**
- Президент группы ЕСЛ — профессор, д.т.н., лауреат премии Правительства РФ в области науки и техники **Шмид Александр Викторович**
- Системный интегратор с опытом работы около 30 лет на российском и мировом рынке
- Опыт создания десятков систем для крупных заказчиков, таких как **Банк России, Газпром, Пенсионный фонд, ВДНХ, Unilever** и многих других отечественных и зарубежных компаний
- Эксперт в области корпоративного, социального и государственного управления, автоматизации, цифровой медицины, больших данных, нейросетей
- Разработчик и правообладатель отечественной управлеченческой методологии **PerAspera Methodology**, цифровых платформ **«Peraspera Платформа»**, **«Peraspera Cardio»**, комплексного решения для автоматизации управления поиском и сбором данных в интернете и корпоративных источниках **«Peraspera Data Hunter»**
- Разработчик автоматизированной системы **«Центр Кибербезопасности»** (совместно с Институтом системного программирования им. В.П.Иванникова РАН).



Откуда что взялось?



- **Группа компаний ЕСЛ** работает на рынке с 1994 года
- Образована из числа сотрудников лидера рынка СССР – Научно-исследовательского центра электронной вычислительной техники (НИЦЭВТ), разработчика Единой Системы ЭВМ стран социалистического содружества (ЕС ЭВМ)
- **Сотрудники компаний ЕСЛ** - специалисты высочайшего класса: проектировщики систем, системные администраторы, разработчики приложений в самых различных областях, владеющие технологиями корпорации IBM и других крупнейших ИТ-компаний
- Создатели *первого отечественного* Центра компетенций по технологии IBM Big Data и около 900 инструментальных средств системы *Watson*
- Разработчик катастрофоустойчивой территориально-распределённой системы централизованной обработки платежной информации Банка России (Премия Правительства Российской Федерации)
- Компания ЕСЛ СОФТ вошла в десятку ИТ-компаний Москвы по направлению SmartCity



Цифровизация и ИИ это вызов!



- **Цифровизация** – это о том, **ЧТО** (процессы, объекты, технологии) можно описывать цифрой и обрабатывать **КАК** цифры
- **Обрабатывать** это значит **превращать** сырьё (данные) в товар (информацию, решения) для удовлетворения информационных потребностей управленацев, медиков, технологов, аналитиков, специалистов, маркетологов – разных специалистов из различных областей деятельности
- **Превращать** это значит обрабатывать данные математическими методами, выявлять закономерности и отвечать на информационную потребность специалистов – представлять информацию в удобной для понимания форме
- **ИИ и его элементы** это математические модели для обработки данных, надо учиться их применять – а при необходимости разрабатывать
- Нужно учиться совместно использовать **разные математические модели** для получения промышленно значимого результата – **информационно-аналитических систем**



Новые методы работы с Большими Данными



**Новые методы работы с большими
данными: победные стратегии
управления в бизнес-аналитике:
Научно-практический сборник.
Под редакцией доктора технических наук,
профессора А.В.Шмida. –
М.: ПАЛЬМИР, 2016. – 528с.: илл.**

Книга включает руководства по выполнению 10 лабораторных работ

Обобщен опыт 4-х лет обучения и написаны инструкций по подготовке магистров по технологии Big Data

Обеспечение ЧИСЛЕННОСТИ РАБОТАЮЩИХ

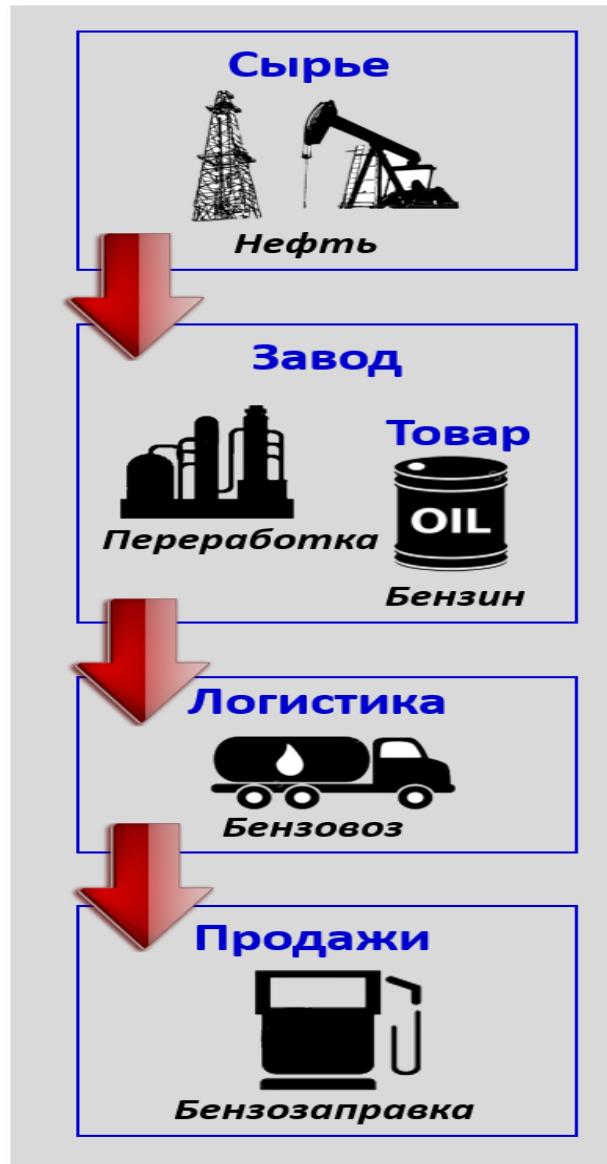
**Концепция подготовки кадров ЦИФРОВОЙ ЭКОНОМИКИ
на примерах:
от детализации к созданию (сборке)
и обучению ЭКСПЕРТНЫХ СИСТЕМ**



Число работающих в цифровой экономике и ее производительные силы

Тетрада нефтяной экономики

\$1.5
трлн.



Тетрада цифровой экономики

\$30
трлн.

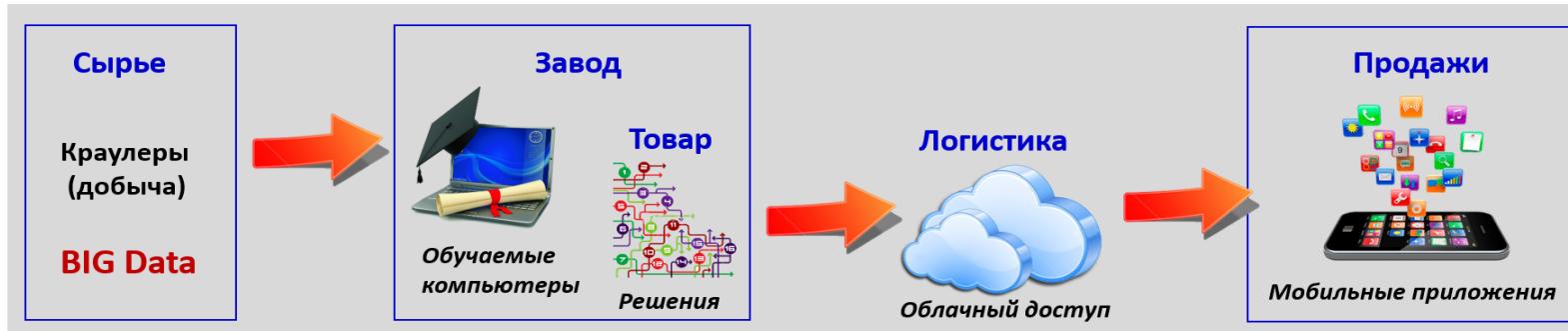


С.Ю. Витте: «По головам или по стадам?»

Как обеспечить численность работающих?

Структура подготовки кадров цифровой экономики

Тетрада
цифровой
экономики



По фирмам
или по людям?



Оценка требований к численности работающих в цифровой экономике:

\$30 трлн. / \$1млн. (Google) = **30 млн. человек**

5 % - 1.5 млн. человек 1,5% - **450 тыс. человек**



О программе



Целью программы является подготовка *специалистов по созданию, эксплуатации и сопровождению информационно-аналитических систем:*

- владеющих знаниями в области работы с большими данными, компьютерных наук, методов проектирования информационно-аналитических систем и их отдельных компонент
- практически освоивших математические и инженерные методы анализа данных для удовлетворения информационных потребностей лиц, принимающих решения
- владеющих методиками и технологиями создания информационно-аналитических систем, в том числе на основе платформенного подхода и с использованием как разрабатываемых программных решений, так и с безопасным использованием Open Source – решений
- владеющих методиками и технологиями обеспечения жизненного цикла безопасного программного обеспечения



Программа рассчитана на подготовку выпускников, *обладающих знаниями и практическими навыками*

- по системному проектированию
- по созданию новых технологий информационно-аналитических систем, развитию и сопровождению существующих технологий в их жизненном цикле
- обладающих навыками и практически владеющих технологиями, обеспечивающими при разработке информационно-аналитических систем преимущество по высокой производительности труда
- способных работать в ведущих ИТ-компаниях, исследовательских центрах и центрах по эксплуатации информационно-аналитических систем различного назначения



Основные научно-технические направления деятельности Группы компаний ЕС-лизинг



- Разработка заказных систем и коробочных решений по созданию информационно-аналитических систем различного назначения в области корпоративного, социального и государственного управления, больших данных, нейросетей – разрабатываемых на основе применения **платформы Peraspera**
- Формирование решений на основе комплексной системы для автоматизации управления поиском и сбором данных в интернете и корпоративных источниках **«Peraspera Data Hunter»**.
- Разворачивание отечественных телемедицинских систем отбора, подготовки и реабилитации спортсменов, реабилитации и отслеживания функционального состояния детей и взрослых, скрининга населения, включая неинвазивный скрининг нарушений углеводного обмена – на основе применения технологии **Peraspera Cardio**, достижений отечественных космической медицины и медицины спорта высших достижений
- Развитие, внедрение и распространение отечественных импортозамещающих решений в области обеспечения жизненного цикла доверенного ПО на основе автоматизированной системы **Центр Кибербезопасности**

Что такое и зачем платформа

- **Функциональные программы**, осуществляющие сбор, очистку, преобразование данных и реализующие математические методы обработки структурированных, полуструктурных и неструктурных данных
- **Среда выполнения** функциональных программ для получения результатов для включения в состав отчетов, удовлетворяющих информационные потребности конечного пользователя
- **Программы для представления полученных результатов** в виде, удобном для конечных пользователей – лиц, принимающих решения, прежде всего на тактическом и стратегическом уровнях управления
- Применение платформ для разработки прикладных систем позволяет повысить производительность труда и качество программных решений



Peraspera Платформа



Применение программы "**Peraspera Платформа**" позволяет сконцентрировать усилия специалистов заказчика на решении бизнес-задач, а не на создании инфраструктуры – все системотехнические вопросы уже решены

Состав Peraspera Платформы

Peraspera Data Hunter

Поиск и сбор данных

Peraspera ESB

Передача данных

Data Storage

Хранение данных

Data Flow

Потоковая обработка данных

Data Reports

Формирование отчётов

Peraspera ML

Машинное обучение

Peraspera BI

Анализ данных и визуализация

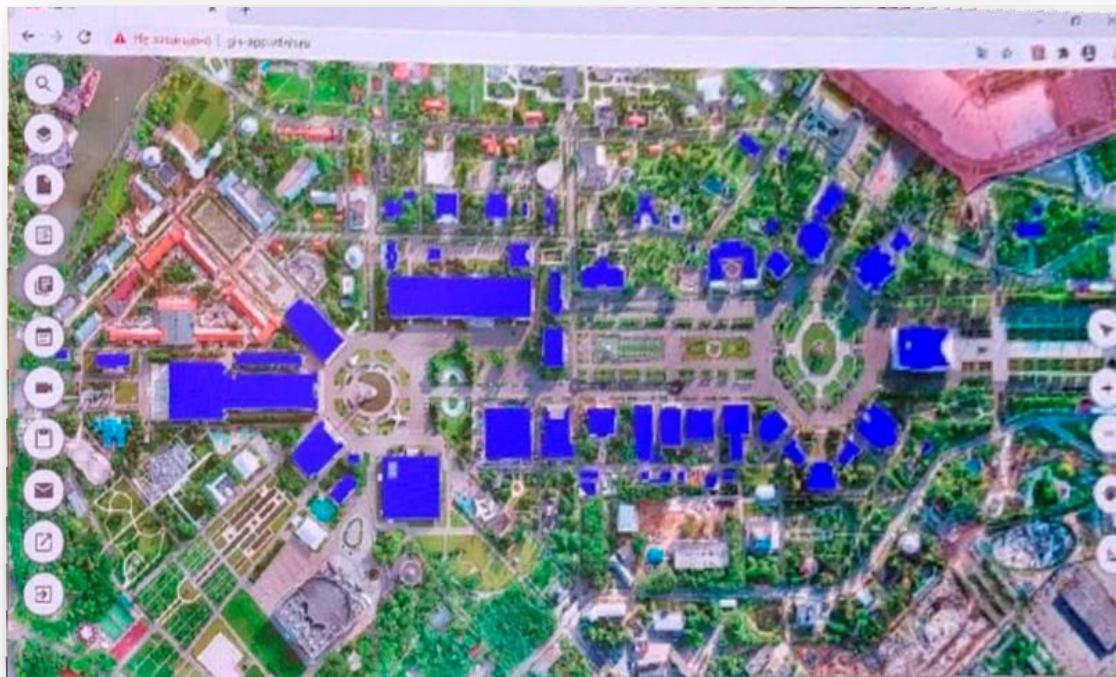
Peraspera Security

Информационная безопасность

Пример прикладной системы на основе платформы

- Система управления территориями CAFM: Computer Aided Facilities Management
- Объект автоматизации – ВДНХ
- Система находится в промышленной эксплуатации
- Развивается в течение семи лет
- Наполнение системы метаданными и настройка отдельных возможностей выполняется силами персонала ВДНХ (не ИТ – специалистами)
- Имеются личные кабинеты всех резидентов ВДНХ





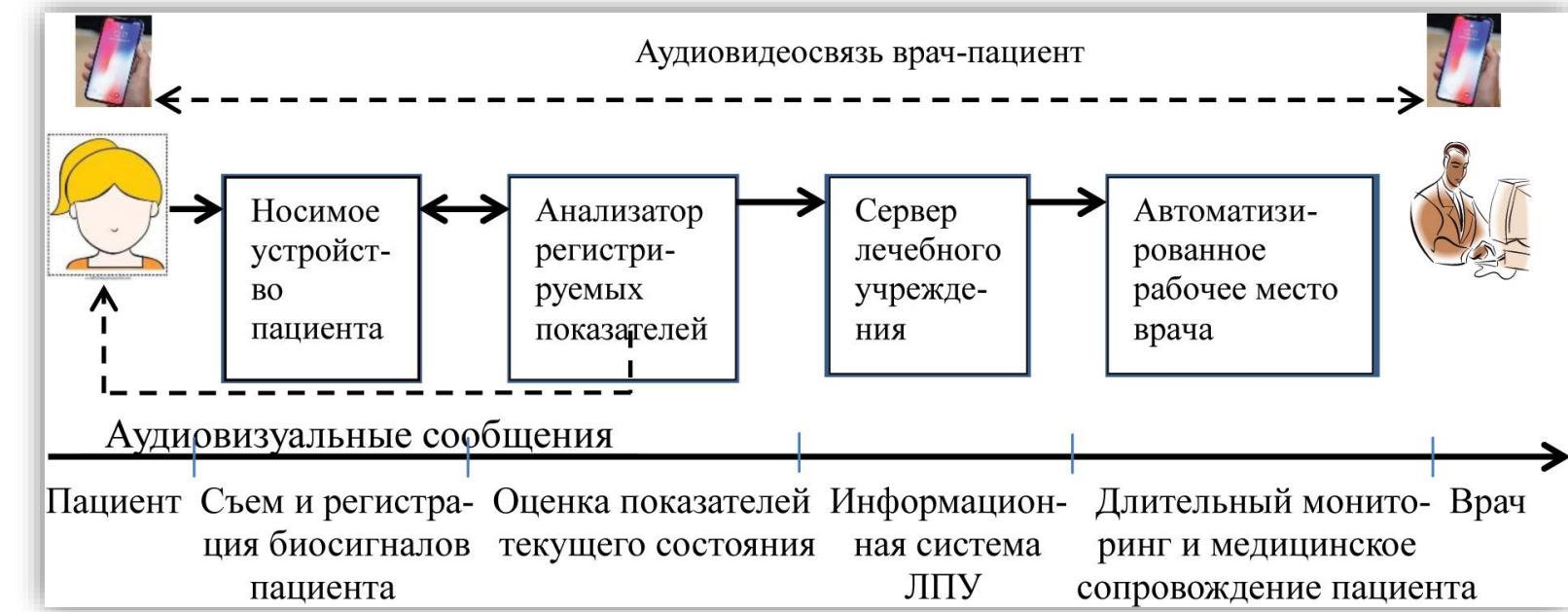
-  Парковые комплексы
 -  Заводы, фабрики, производство
 -  Музейные комплексы
 -  Зоны развлечений
 -  Особые экономические зоны
 -  Медицинские комплексы
 -  Образовательные учреждения
 -  Аэропорты
 -  Санатории, пансионаты, гостиничные комплексы
 -  ЖК, Крупные девелоперы

Телемедицинские системы

Базовая кафедра
Информационно-
аналитических
систем ЗАО "ЕС-
лизинг"

→ Практическая
реализация
современных
технологий создания
информационных
систем

**Системы, предназначенные
для проведения удаленных
медицинских консультаций и
дистанционного обучения по
различным видам
диагностики и лечения, как в
реальном времени, так и в
отложенном режиме, с
использованием любых
каналов связи.**



**Общая схема телемедицинской системы, работающей в режиме
реального времени**

Гельман В. Я., Дохов М. А. Проблемы развития домашнего мониторинга состояния
здоровья. Медицина 2020; 8(2): 50-60.

Математическая кардиология

Базовая кафедра
Информационно-
аналитических
систем ЗАО "ЕС-
лизинг"

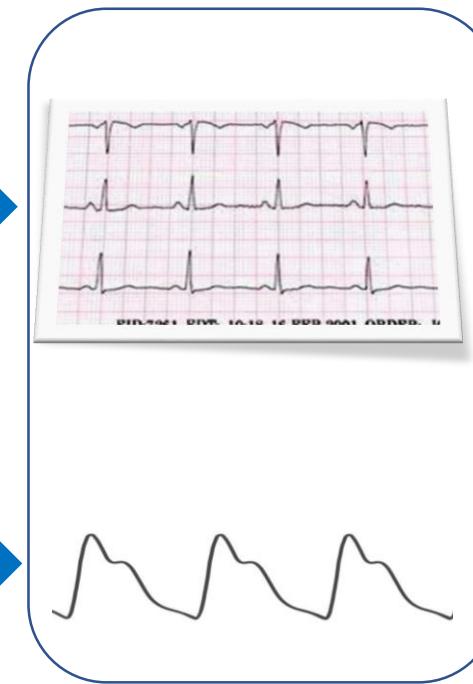
→ Практическая
реализация
современных
технологий создания
информационных
систем

→ Математическая
кардиология

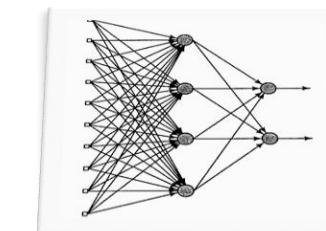
Применение математических моделей в кардиологии



Пациент



Кардиологические
данные



Математическая
модель



Выводы по пациенту



Врач-
кардиолог/
спортивный
врач



Пациент

Сфера применения

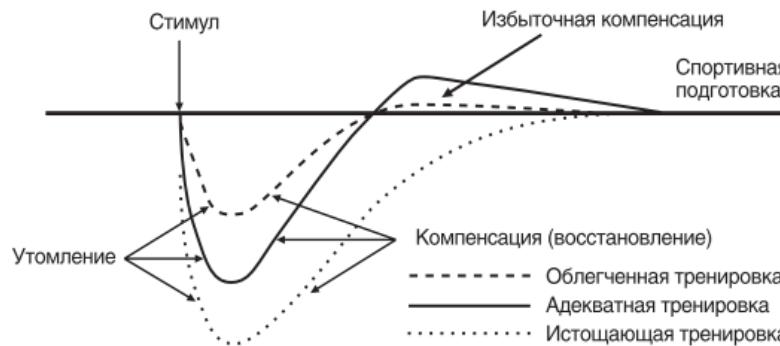


Базовая кафедра
Информационно-
аналитических
систем ЗАО "ЕС-
лизинг"

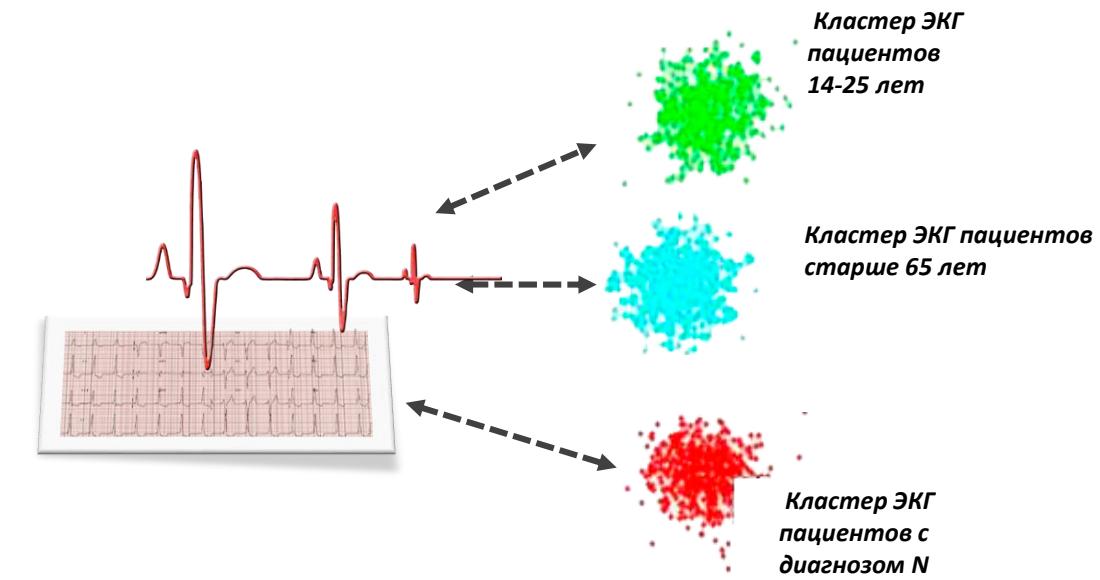
Практическая
реализация
современных
технологий создания
информационных
систем

Математическая
кардиология

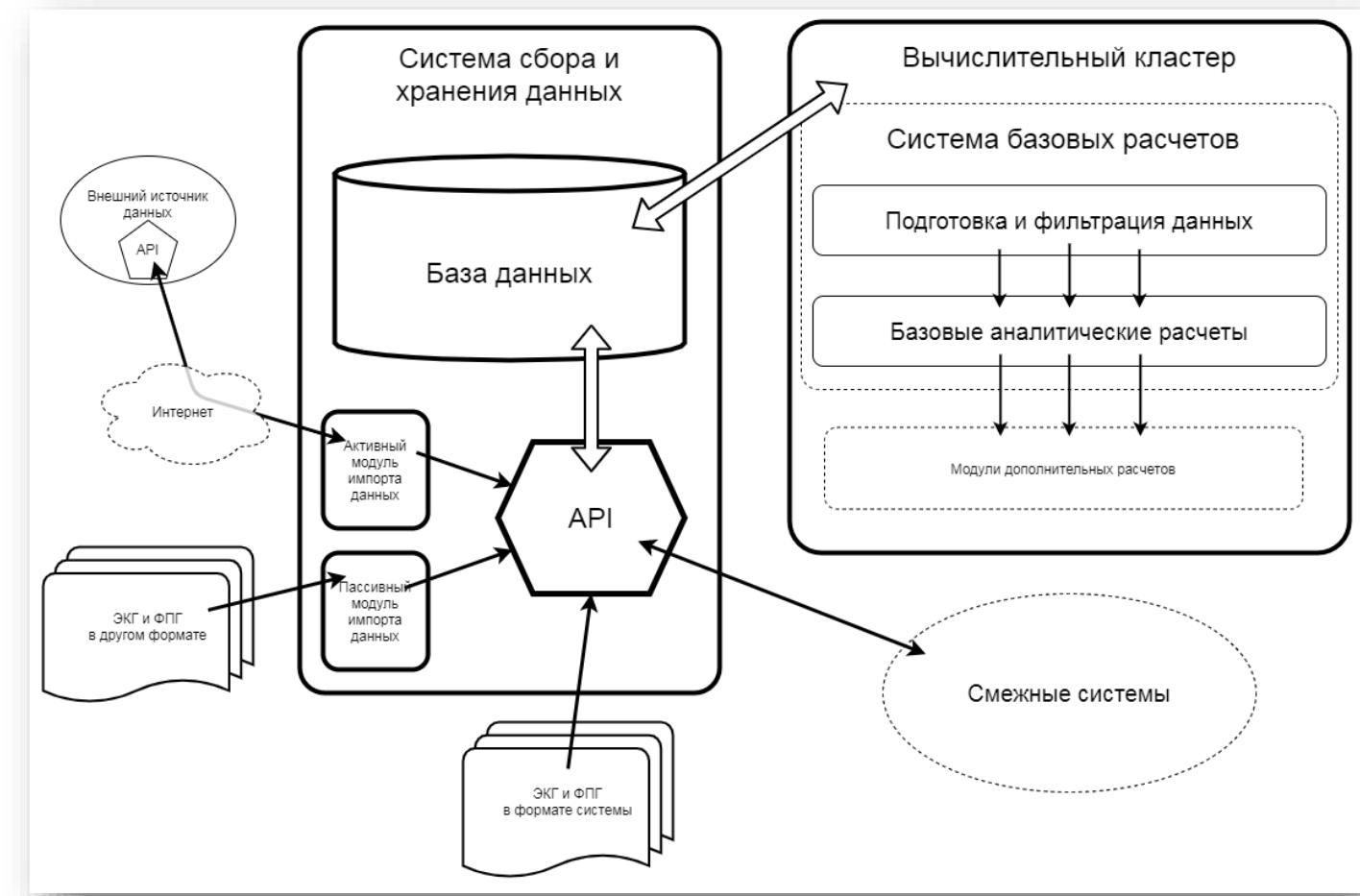
Спортивная медицина: спорт, фитнес, реабилитация



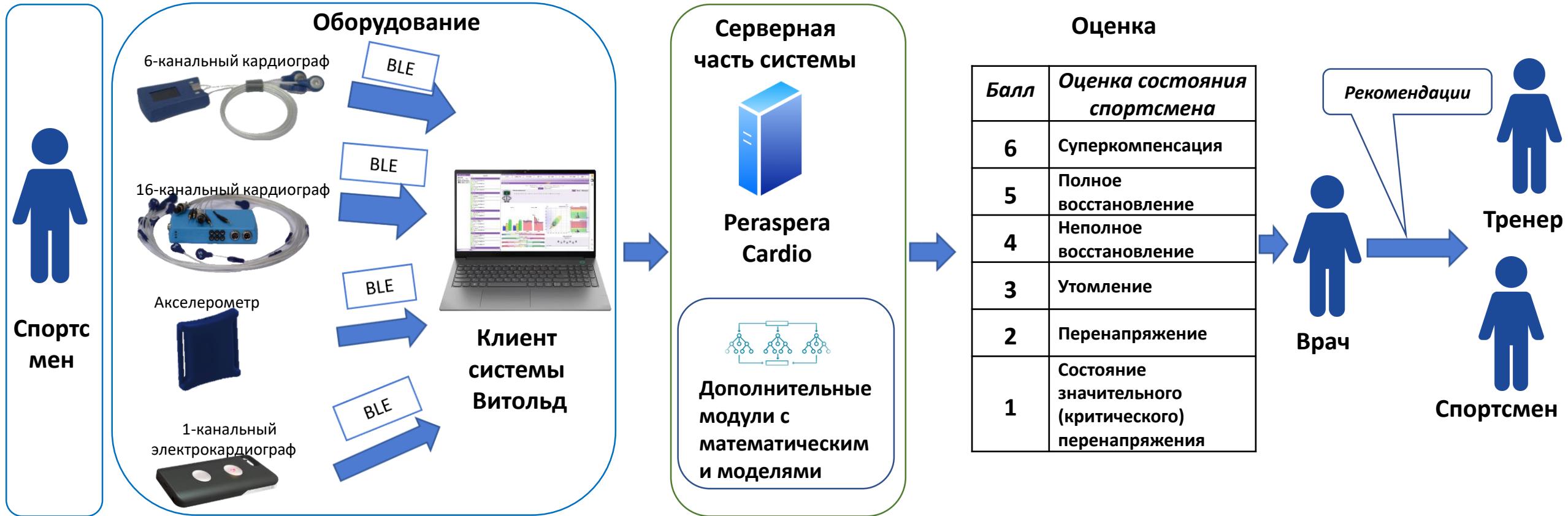
Анализ состояния пациента вне спорта



Платформа для оценки состояния сердечно-сосудистой системы человека на основе обработки ЭКГ от 1 до 16 отведений и фотоплазмограмм (ФПГ) одного или нескольких спектров



«Витольд» – система оценки состояния тренировочного процесса



Решения по этому направлению запатентованы





Обеспечение жизненного цикла безопасного (доверенного) ПО



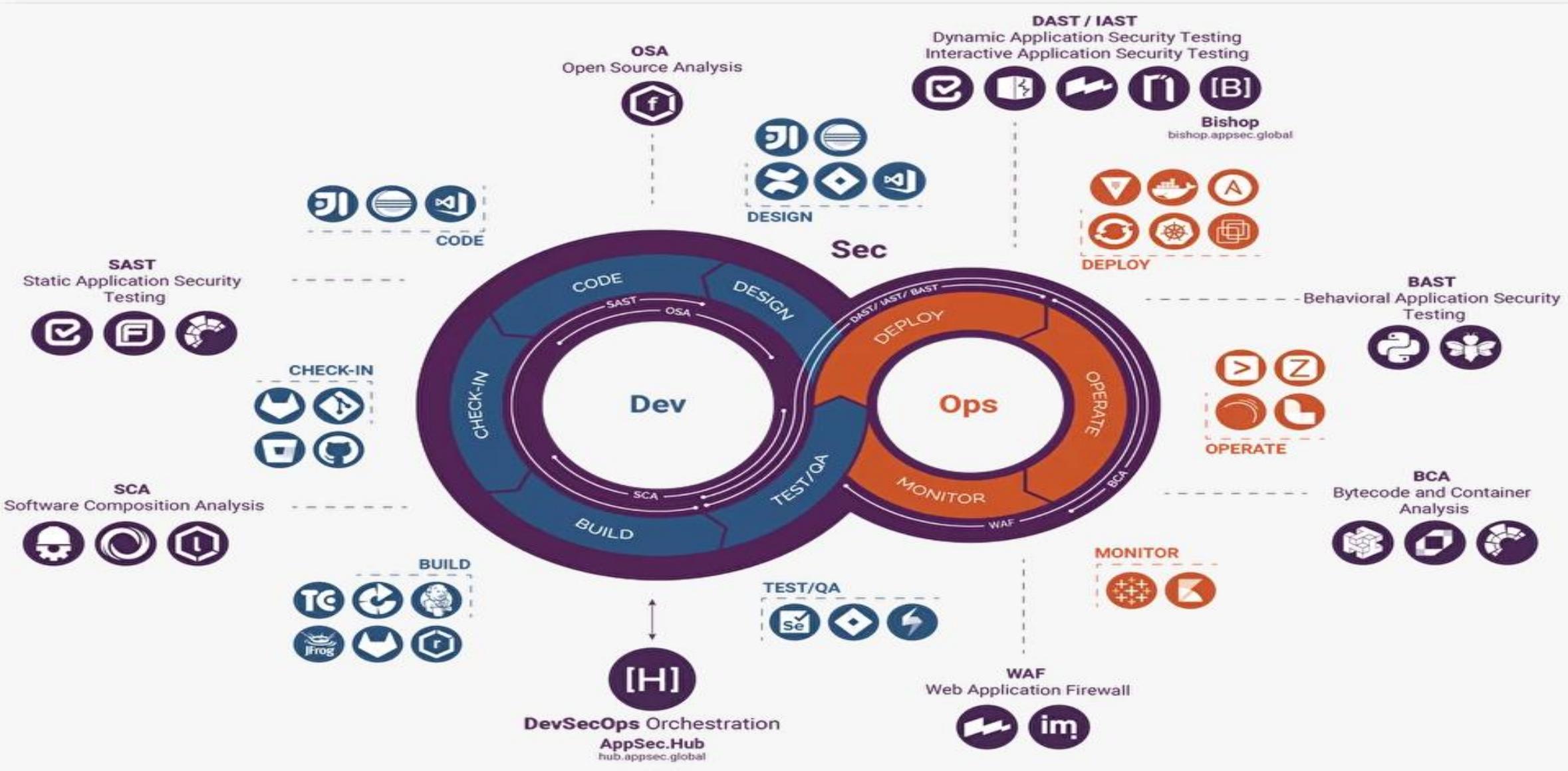
Доля ПО с открытым кодом в составе прикладного ПО современных автоматизированных систем достигает **80-85%**

А какова доля уязвимостей и недекларированных возможностей, которые привносит это ПО с открытым кодом?

ФСТЭК – госорган, отвечающий за регулирование в области создания и применения доверенного ПО

Проблема – обеспечение поиска и устранения уязвимостей разными эффективными методами для различных классов систем и архитектур ПО написанного на языках программирования, используемых при создании ПО систем

Концепция DevSeqOps



Применяемые практики по разработке безопасного ПО



ИТ-архитекторы и ИБ-специалисты,
127 ответов



Несколько
вариантов ответов

44%

SAST
(Static Application
Security Testing)

43%

DAST
(Dynamic Application
Security Testing)

33%

SCA
(Software Composition
Analysis)

19%

IAST
(Interactive Application
Security Testing)

6%

FUZZ testing
(fuzzing)

2%

Другое



Состав участников проекта АС ЦКБ



Участник	Функции	Продукты	Технологии
Институт системного программирования РАН	ТЗ, идеология, инструментальные средства	Инструментальные средства для статического (SAST), динамического (DAST) и фаззинг (Fuzzing) тестирования	SAST, DAST, Fuzzing
ЕС-лизинг	ТЗ, системная интеграция, регламентация процессов, методики и документация, Внедрение и сопровождение АС ЦКБ	Автоматизированная система АС ЦКБ, Бизнес-процесс и регламенты жизненного цикла доверенного ПО, Документация на систему	Комплексный стенд для отработки и обучения
Базальт	Сертифицированная операционная система Альт-Линукс	Альт-Линукс	Alt Linux
Астра	Сертифицированная операционная система Астра-Линукс	Астра-Линукс	Astra Linux
Postgres Professional	Сертифицированная СУБД Postgres Pro Enterprise	Postgres Pro Enterprise	СУБД
Беллсофт	Сертифицированное окружение Java	Java (Axiom JDK Certified)	Java JDK
Профископ (*)	Композиционный анализ Open Source приложений	CodeScoring	SCA
Рт-Солар (*)	Статический, динамический и композиционный анализ, включая web-приложения	Solar appScreener	SAST, DAST, SCA

Выпускник программы должен

- уметь решать поставленные задачи в выбранной профессиональной области инженерии построения информационно-аналитических систем;
- знать и уметь пользоваться математическими понятиями и методами, востребованными в области аналитической обработки данных;
- уметь решать задачи сбора, накопления, очистки и хранения данных, их аналитической обработки и представления в форме, соответствующей информационной потребности конечного пользователя информационно-аналитической системы – с применением современных методов и инструментальных средств и платформ для повышения производительности труда при создании и функциональном развитии информационно-аналитических систем;
- уметь разрабатывать самому и в составе команды доверенное программное обеспечение информационно-аналитических систем и аппаратно-программные решения;
- уметь разрабатывать документацию на систему и программное средство.



Учебная лаборатория кафедры «Информационно-аналитические системы»



- Приобретение навыков по специальности в современных информационных технологиях невозможно без участия в проектных работах, в разработке информационно-аналитических систем или их составных частей, без использования современных языков программирования, инструментальных средств, систем управления базами данных
- Компания «ЕС-лизинг» на своем вычислительном центре выделила ресурсы для студентов базовой кафедры. При реализации проектов кафедры студенты получают удаленный защищенный доступ к вычислительным ресурсам и инструментальным средствам, которые использует компания в своих проектах. В том числе при необходимости студентам предоставляется доступ к платформам Peraspera, Peraspera Cardio и автоматизированной системе Центр Кибербезопасности (система для автоматизации выявления уязвимостей в ПО систем инструментальными средствами от лидеров рынка)
- Учебная лаборатория оснащена также продуктами отечественных лидеров рынка – партнеров компании: Института Системного программирования им. В.П. Иванникова РАН, Группы компаний Астра, Базальт СПО, Профископ, Postgres Professional, Беллсофт (разработчик продукта Axiom JDK), RT Solar (продукт AppScreener). Эти компании – производители сертифицированного отечественного ПО. Продукты описанных компаний, как и инструментальные средства Open Source, используются в учебных проектах кафедры



Учебные проекты кафедры



- Продукты описанных компаний, как и инструментальные средства Open Source, используются в учебных проектах кафедры.
- По сложившейся практике все проекты кафедры направлены на решение студентами научно-исследовательских, экспериментальных, учебно-методических и практических задач проектирования, разработки информационно-аналитических систем и их составных частей, а также на тестирование, поиск и устранение уязвимостей в компонентах информационно-аналитических систем, создаваемых компанией ЕС-лизинг.
- Подход к реализации проектов на средствах учебной лаборатории ничем не отличается от подхода к производственной деятельности сотрудников компании. Это дает возможность студентам получить опыт работы в команде, в производственном коллективе при решении инженерных задач.

Характеристики сервера

- Сервер **Huawei 2288H V5**
- Процессор Intel Xeon Gold **6248(2.5GHz/20-core/27.5MB/150W)**
- Память DDR4 RDIMM Memory, **32GB**
- Жесткий диск **HDD,600GB, SAS 12Gb/s,10K rpm,128MB**



Системное ПО и инструментальные средства сервера (пример)



Наименование	Версия	Описание	Единый Реестр российского ПО
Альт Сервер	5.10.166-std-def-alt1	Версия ядра ОС	Реестровая запись №1541 от 05.09.2016
SVACE	3.4.240109	Статический анализатор	Реестровая запись №4047 от 11.12.2017
Svacer	8.0-0	Сервер харения и управления результатами анализа	нет
Crusher	2.16.0	Программный комплекс, комбинирующий несколько методов динамического и статического анализа, в частности, фазинг (ИСП Fuzzer) и т.д.	Реестровая запись №10468 от 06.05.2021
Postgres Pro Enterprise	14.8	СУБД	Реестровая запись №104 от 18.03.2016
Docker		Автоматизация развёртывания и управления приложениями в средах с поддержкой контейнеризации	нет
Git	2.33.8	Распределённая система управления версиями	нет
Jenkins	2.426.1	Система для обеспечения процесса непрерывной интеграции.	нет
Gitea	built with go1.21.3	ПО совместной разработки на базе Git	нет
GCC	10.3.1	Компилятор	нет
Clang		Компилятор	нет
GNU Binutils	2.35.2.20210110	Набор инструментального ПО для обращения с объектным кодом в объектных файлах различного формата	нет
Python	3.9.6	Компилятор	нет
Java (Axiom)	17.0.7+7-LTS	Среда разработки и исполнения JAVA	Реестровая запись №16685 от 20.02.2023 35



Проекты кафедры ИАС



№	Тема проекта	Руководитель проекта	Проект №
1	Продолжение проекта «Разработка функциональной и системной архитектуры системы обеспечения жизненного цикла доверенного ПО»	Позин Б.А.	1309
2	Экспериментальная отработка методики поиска уязвимостей в комплексе программ на Java	Позин Б.А. (Кадышев)	1589
3	Экспериментальная отработка методики поиска уязвимостей в комплексе программ на Python	Позин Б.А. (Кадышев С.)	1590(**)
4	Экспериментальная отработка методики поиска уязвимостей в комплексе программ на Java с использованием инструментальных средств динамического анализа		
5	Экспериментальная отработка методики поиска уязвимостей в комплексе программ на Python с использованием инструментальных средств динамического анализа		
6	Разработка лабораторных работ по сбору данных из открытых источников данных.	Лычагин К.А.	
7	Разработка типовой модели данных для хранения данных из открытых источников в сети интернет	Лычагин К.А.	
8	Автоматизация процесса создания ETL-процессов на базе Apache Airflow	Лычагин К.А.	
9	Разработка практических заданий по моделированию данных для хранилищ данных	Лычагин К.А.	
10	"Анализ фотоплетизмограмм для оценки наличия нарушений углеводного обмена"	Новиков Р.С.	1576
11	"Планирование нагружочного тестирования для системы проведения скрининга нарушений углеводного обмена по кардиологическим данным"	Новиков Р.С.	1577
12	Исследование нейросетевых методов поиска именованных сущностей в тексте с возможностью дообучения и верификации найденных данных	Белякова Е.С.	1820 (*)

Спасибо за внимание!



Анкета для собеседования ФИО _____ Группа _____ адрес отправки анкеты rshnovikov@hse.ru



Цель обучения (для вас – что вы хотите получить в результате обучения)		
Средний балл за 2 курса		
Знания и навыки	Мат. Анализ Теория вероятностей/мат. Статистика Data Science Программирование/ Алгоритмизация Структуры данных/ Базы данных Языки программирования: Python C,C++, Java Участие в проектах/ название проекта, роль	
Английский язык (технический)		
Какими инструментами работы на компьютере владеете		
Место жительства (нас. пункт по паспорту, если в Москве – ближайшее метро/ЖД)		
Почта для связи		