

Программа кандидатского экзамена по научной специальности 2.3.8 «**Информатика и информационные процессы**» составлена в соответствии с требованиями следующих нормативных документов:

1. Федеральные государственные требования к структуре программ подготовки научных и научно-педагогических кадров в аспирантуре - приказ Минобрнауки России от 20.10.2021 г. № 951.

2. Паспорт научной специальности 2.3.8 «Информатика и информационные процессы», разработанный экспертами ВАК Минобрнауки России в рамках Номенклатуры научных специальностей, утвержденной приказом Минобрнауки России от 24.02.2021 г. № 118

Приказ Министерства образования и науки Российской Федерации от 30.07.2014 № 875 «Об утверждении федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.06.01 Информатика и вычислительная техника (уровень подготовки кадров высшей квалификации)».

Исполнители (разработчики программы):

Михеенкова Мария Анатольевна
ВИНИТИ РАН, вед.н.с. ОНИ по проблемам информатики
доктор технических наук

Программа кандидатского экзамена по научной специальности 2.3.8 «**Информатика и информационные процессы**» зарегистрирована в аспирантуре под учетным номером 9 на правах учебно-методического издания.

Зав Аспирантурой
ВИНИТИ РАН



Камнева И.Е.

«Программа-минимум - Теоретические основы информатики»

В основу настоящей программы положены следующие дисциплины: теоретические основы информатики, теория вероятностей и математическая статистика, вычислительная техника и программирование, теория информации и теория кодирования, теоретические основы методов машинного обучения и искусственного интеллекта, теоретические основы информационных систем и технологий, базы данных, проектирование информационных систем, интеллектуальные информационные системы, высокоуровневые методы информатики и программирования, управление информационными ресурсами

1. Информатика как наука, отрасль промышленности и инфраструктурная область

1.1. Информатика как наука, изучающая информацию и ее свойства в естественных, искусственных и гибридных системах. Место информатики в системе наук. Информатика как обрабатывающая информацию отрасль индустрии, ее роль и значение в ускорении научно-технического прогресса.

1.2. Понятие информации, определения информации. Виды и общие свойства информации. Кодирование информации. Информационное взаимодействие.

1.3. Предметная область информатики. Информационные проблемы современности. Информационные потребности индивидуальных и коллективных пользователей. Информационные коммуникативные процессы. Современные информационные технологии на базе широкого применения вычислительной техники. Социальные аспекты информатизации и компьютеризации общества. Организация информационных служб и отраслей народного хозяйства

1.4. Инфраструктура информационной деятельности. Понятие научной коммуникации. Системы научной информации, системы научной коммуникации, автоматизированные системы и сети, их взаимосвязь и взаимозависимость. Основная и информационная деятельность. Виды основной и информационной деятельности, их взаимозаменяемость. Система информационного обмена, организационные и функциональные элементы.

1.5. Понятие информационного продукта и информационной услуги, их классификация. Жизненный цикл информационного продукта. Экономика информационных сетей. Методы управления производством и распределением информационных продуктов. Методы анализа и оценки качества информационных продуктов и услуг. Основные секторы информационной сферы – информация; электронные коммуникации; тематическая классификация. Деловая информация, информация для специалистов, научно-техническая информация. Другие виды профессионально-ориентированной информации. Социально-значимая (правовая, социальная, политическая, экологическая, образовательная и др.) информация.

1.6. Информационные ресурсы. Принципы оценки информации как ресурса общества и объекта интеллектуальной собственности. Проблемы правового регулирования научной интеллектуальной собственности. Противоправное использование информационных ресурсов. Государственная политика в области защиты информационных ресурсов общества.

1.7. Синтаксический, семантический и прагматический аспекты в изучении информации. Способы и единицы измерения информации. Информация и управление. Соотношение понятий «данные», «знания», «информация». Язык и сигнальные системы. Виды информации и их особенности.

2. Концептуальные модели информатики. Информационное и лингвистическое обеспечение информационных процессов

2.1. Предметная область и её модели. Общие принципы моделирования окружающей среды, процессов мышления человека и человеко-машинного взаимодействия. Объекты, свойства,

отношения. Декларативное и процедурное представление знаний и данных. База знаний и база данных. Методы хранения, поиска и обработки данных, методы естественно-языкового человеко- машинного взаимодействия.

2.2. Моделирование рассуждений. Семиотический треугольник и его элементы. Понятия «экстенционал» и «интенционал». Автоматизация дедуктивных рассуждений. Поиск доказательств теорем методом резолюций.

2.3. Представление знаний

2.3.1. Классификационные системы: иерархические классификации, фасетные классификации, алфавитно-предметные классификации. Тезаурусные методы представления знаний.

2.3.2. Системы, основанные на отношениях. Объектно-характеристические таблицы. Предикатно-актантные структуры.

2.3.3. Семантические сети. Семантические отношения и их виды. Лингвистические, логические, теоретико-множественные, квантификационные отношения. Абстрактные и конкретные семантические сети.

2.3.4. Фреймы — системно-структурное описание предметной области. Принципы фрейм-представлений. Понятие «СЛОТА».

2.3.5. Продукционные системы представления знаний. Канонические системы Поста. Представление неформальных знаний.

2.3.6. Редукционные системы. Синтез плана решения задач с автоматическим построением редукционной модели.

2.4. Представление данных

2.4.1. Обработка данных. Структуры данных. Уровни представления данных. Языки описания и манипулирования данными, языки запросов.

2.4.2. Базы данных. Основные понятия. Независимость программ и данных. Интегрированное использование данных. Непротиворечивость данных. Целостность и защита данных. Структуры БД. Администрирование банков данных. Типы пользователей. Администратор БД. Понятие концептуальной, логической, физической структуры БД. Представления пользователей и подсхемы. Понятие о словарях данных, языках описания и манипулирования данными. БД и файловые системы. Документальные и фактографические базы данных, базы знаний. Полнотекстовые БД. Физическая и логическая структура. Файл полного текста. Частотный словарь, инверсный файл. Положительный и отрицательный словари. Стандартные строки и словосочетания, включаемые в частотный словарь. Описание БД. Обработка текстов при загрузке БД. Понятие экспорта-импорта документов-данных.

2.4.3. Понятие модели данных. Иерархическая и сетевая модели данных, сравнительный анализ, противоречия и парадоксы. Реляционная модель данных. Экземпляры отношений, домены, атрибуты. Операции над отношениями: селекция, проекция, естественное соединение. Понятие реляционной полноты языка манипулирования данными. Модель данных «сущность—связь».

2.4.4. Понятие «больших данных». Данные и метаданные. Непротиворечивость данных. Целостность и защита данных. Многопользовательские и распределённые системы

2.5. Базовые методы машинного обучения. Основные задачи машинного обучения: классификация, кластеризация, визуализация. Байесовский классификатор. Метод опорных векторов. Решающие деревья. Основные понятия нейронных сетей, многослойные нейронные сети, метод обратного распространения ошибки.

2.6. Методы и алгоритмы анализа текстовых данных, приложения задач анализа текстовых данных. Тематическая обработка данных.

2.7. Информационный поиск. Основные понятия и виды поиска. Информационно-поисковые языки. Понятия пертинентности, смысловой и формальной релевантности. Критерии выдачи. Модели поиска. Стратегия поиска. Функциональная эффективность

поиска. Поисковые массивы, способы их организации. Понятия об ассоциативном поиске и условиях его реализации. Оперативный и регламентный режим поиска.

2.8. Автоматизированные информационные системы (АИС), их определение, назначение. Роль и место АИС в системах информационного обеспечения управления, науки, экономики. Классификация АИС по функциональному назначению, уровню, структуре данных. Структура АИС.

2.9. Языковые средства АИС. Входные и внутренние языки. Структура входных языков. Языковые средства для ввода и обновления информации, для поиска, обобщения и выдачи информации. Языковые средства общения с БД. Анкетный язык. Языковые средства документальных (в том числе полнотекстовых) ИПС: три уровня грамматики информационно-поисковых языков (теоретико-множественный, линейный, сетевой). Информационно-поисковый язык. Язык информационно-логический. Язык процедурно ориентированный. Непроцедурный язык концептуального уровня. Язык диалога. Естественный язык. Словарный комплекс АИС. Классификаторы. Кодификаторы. Тезаурусы – состав и структура. Языки описания данных и словарь данных. Языки запросов SQL и QBE.

3. Математические основы информатики.

3.1. Теория вероятностей и математическая статистика. Основные понятия и аксиомы теории вероятностей. Случайные величины, их распределение и моменты. Закон больших чисел. Центральная предельная теорема. Понятие случайного процесса. Цепи Маркова. Стационарные и эргодические процессы. Обработка опытов. Понятие о случайной выборке. Проверка статистических гипотез. Оценка параметров по независимым наблюдениям. Оценки максимального правдоподобия. Метод наименьших квадратов. Многомерный статистический анализ. Множественный корреляционно-регрессионный анализ. Компонентный анализ. Факторный анализ. Кластерный анализ.

3.2. Теория информации и теория кодирования. Энтропия стационарного случайного процесса. Количество информации и его свойства. Избыточность сообщений. Сжатие данных. Ошибки при передаче и хранении данных. Дискретные каналы связи с памятью и без памяти. Помехозащитное кодирование. Теорема Шеннона о пропускной способности канала связи. Линейные помехозащитные коды. Блочные коды. Сверточные коды. Обнаружение и исправление ошибок помехозащитными кодами.

3.3 Теория вычислимости. Частично-вычислимые функции. Формальные математические системы (канонические системы Поста, алгоритмы Маркова, машины Тьюринга). Эквивалентность различных формальных систем. Тезис Черча. Универсальные системы. Существование неразрешимых множеств. Временная сложность алгоритмов, классы P, NP, NP-полнота, примеры NP-полных задач. Квантовые вычисления, проблемы реализации квантовых компьютеров.

3.4 Теория программирования

3.4.1. Программный принцип работы ЭВМ. Программное обеспечение, его структура. Операционная система, ее основные компоненты, примеры ОС.

3.4.2. Классификация языков программирования. Базовые конструкции языков программирования. Структурное программирование. Объектно-ориентированное и компонентное программирование. Алгоритмы и структуры данных для сортировки, поиска, хэширования. Компиляция и интерпретация программ.

3.4.3. Математические методы принятия решений. Исследование операций. Дискретная оптимизация. Многокритериальная оптимизация, оптимальность по Парето. Методы оптимизации: градиентный спуск, покоординатный спуск, метод штрафных функций. Экономические, математические, статистические модели для задач принятия решений и управления. Постановка задач оптимального управления, их классификация.

3.4.4. Динамическое программирование, линейное программирование. Симплекс-метод. Транспортная задача, задача распределения ресурсов. Использование линейного программирования для решения оптимизационных задач.

3.4.5. Системное программирование. Управление программами, данными и заданиями. Трансляция, редактирование и загрузка. Управление процессами и памятью. Восстановление вычислительного процесса. Пакеты прикладных программ, их архитектура. Базы данных и работа с ними. Надежность программного обеспечения.

3.5. Математические модели информационных систем и ресурсов – описание, оценка, оптимизация. Модели описания информационных процессов и ресурсов. Теоретико-множественное описание сообщений, запросов, массивов документов. Универсальный информационный поток. Линейная модель. Матрица информационного потока. Ассоциативные матрицы информационного потока.

3.6. Критерии оценки информационных систем. Оценки качества поиска (полнота, точность и др.). Скалярные и векторные оценки. Смешанные критерии (полезная работа, корреляционный критерий, свертки и пр.). Рабочие характеристики информационно-поисковых систем (ИПС) в различных координатах. Вероятностная модель ИПС. Теоретико-множественная модель ИПС. Оптимизация режима ИПС.

3.7. Линейное представление документов, запросов, тезауруса, индексирования, поиска. Оценка структуры тезауруса. Понятие лексической совместимости и тезаурусной согласованности. Определение различительной силы термина, его различные варианты. Модели динамической корректировки запроса.

3.8. Теоретико-множественные макромоделли больших систем информационных ресурсов. Информационная и основная деятельность. Теоретико-множественные представления операций над информационными ресурсами. Операторы формирования информационных потоков. Количественная форма операторов. Линеаризованная форма операторов. Операции над операторами.

4. Интеллектуальные системы

4.1. Назначение и задачи интеллектуальных систем. Проблемы интеллектуальных систем. Воспроизведение интеллектуальной деятельности человека. Особенности неформальных знаний. Обобщение данных и правил. Уместность правил. Разрешение проблемы коллизий.

4.2. Общая структура интеллектуальной системы и её основные компоненты. Состав и структура знаний. Диалог. Анализ входных и синтез выходных сообщений. Управление решением задач. Объяснительная компонента. Пополнение знаний. Компонента воздействия на внешний мир. Классификация интеллектуальных систем. Вопросно-ответные системы. Информационный поиск. Универсальные решатели задач. Машинный перевод. Обработка текстов естественного языка. Распознавание образов. Робототехнические системы.

4.3. Понятие человеко-машинного взаимодействия. Общая схема кооперативного решения задач в диалоге человека с интеллектуальной системой. Функции пользователя при человеко-машинном решении задачи. Функции системы при решении задачи. Структуры диалогов и их влияние на успешное решение задачи. Пользовательский интерфейс, проектирование пользовательских интерфейсов. Инструменты создания интерактивных прототипов.

4.4. Экспертные системы. Общие принципы построения экспертных систем. Структура экспертных систем.

4.5. Прямые методы приобретения знаний. Когнитивные карты, концептуальные карты. Классификация онтологий. Онтологии верхнего уровня. Языки описания онтологий. Языки запросов к онтологиям.

5. Техническое обеспечение информационных систем и процессов

5.1. Физические основы вычислительных процессов. Основы построения и функционирования вычислительных машин: общие принципы построения и архитектуры вычислительных машин, информационно-логические основы вычислительных машин, их функциональная и структурная организация, память, процессоры, каналы и интерфейсы ввода-вывода, периферийные устройства.

Элементы вычислительной техники. Счетно-решающие механические и электромеханические устройства. Аналоговые и цифровые вычислительные машины. Понятие фон-неймановской машины. Процессор. Главная память. Система команд. Машинное слово. Разрядность и адресность. Программы и данные. Траектория данных в ЭВМ. Элементная база.

5.2. Архитектурные особенности и организация функционирования вычислительных машин различных классов: многомашинные и многопроцессорные вычислительные системы, типовые вычислительные структуры и программное обеспечение, режимы работы. Иерархическая структура ЭВМ. Главный процессор, каналные процессоры, контроллеры устройств. Накопители данных и внешние устройства ЭВМ.

5.3. Классификация и архитектура вычислительных сетей, техническое, информационное и программное обеспечение сетей, структура и организация функционирования сетей (глобальных, региональных, локальных).

5.4. Структура и характеристики систем телекоммуникаций: коммутация и маршрутизация телекоммуникационных систем, цифровые сети связи, электронная почта. Эффективность функционирования вычислительных машин, систем и сетей телекоммуникаций; пути ее повышения. Перспективы развития вычислительных средств. Технические средства человеко-машинного интерфейса.

6. Программное обеспечение информационных систем и процессов

6.1. Классы программных средств. Операционные системы. Системы программирования. Программные продукты.

6.2. Операционные системы. Функции операционной системы (ОС): управление задачами; управление данными; связь с оператором. Системное внешнее устройство и загрузка ОС. Резидентные модули и утилиты ОС. Управляющие программы (драйверы) внешних устройств. Запуск и остановка резидентных задач. Запуск и прекращение нерезидентных задач. Управление прохождением задачи и использованием памяти. Понятие тома и файла данных. Сообщения операционной системы. Команды и директивы оператора.

6.3. Системы программирования. Понятие разработки приложений. Состав системы программирования: язык программирования (ЯП); обработчик программ; библиотека программ и функций. История развития и сравнительный анализ ЯП. Типы данных. Элементарные данные, агрегаты данных, массивы, структуры, повторяющиеся структуры. Вычислительные данные, символьные данные, логические, адресные (метки и пойнтеры), прочие (битовые строки). Понятие блока и процедуры. Операторы ЯП: управления (организация циклов, ветвления процесса, перехода), присваивания, вычисления арифметических, логических, строчных выражений. Стандартные арифметические, логические, строчные функции.

6.4. Программные продукты (приложения). Оболочки операционной системы. Программные пакеты информационного поиска. Оболочки экспертных систем. Понятие открытого и закрытого программного продукта. Понятие генератора приложений. Системы управления базами данных (СУБД), состав и структура. Типовые функции СУБД: хранение, поиск данных; обеспечение доступа из прикладных программ и с терминала конечного пользователя; преобразование данных; словарное обеспечение БД; импорт и экспорт данных из/в файлы ОС ЭВМ. Типовая структура СУБД: ядро, обрамление, утилиты, интерпретатор/компилятор пользовательского языка манипулирования данными. Среда

конечного пользователя. Front-end процессор. Back-end-процессор. Обеспечение безопасности в СУБД. Распределение ресурсов в СУБД.

6.5. Лингвистическое обеспечение информационных систем и процессов. Методы и средства проектирования словарей данных, словарей индексирования и поиска информации, тезаурусов и иных лексических комплексов. Методы семантического, синтаксического и прагматического анализа текстовой информации для представления в базах данных и организации интерфейсов информационных систем с пользователями. Технологии извлечения и анализа информации в больших базах данных. Концепции многомерного представления (OLAP) и интеллектуального анализа данных (Data Mining).

6.6. Принципы организации и функционирования распределенных информационных систем и баз данных, прикладных протоколов информационных сетей, форматов представления данных и языков информационного поиска в распределенных информационных ресурсах.

6.7. Новейшие направления в области создания технологий программирования. Программирование в средах современных информационных систем: создание модульных программ, элементы теории модульного программирования, объектно-ориентированное проектирование и программирование.

7. Сетевые информационные технологии и ресурсы

7.1. Глобальные информационные сети, принципы создания. Общие характеристики, основные понятия, структура, организация, основные программные средства, информационные ресурсы (адрес в сети, имя в сети). Основные информационные средства и ресурсы сети. Удаленный доступ к ресурсам сети. Эмуляция удаленного терминала. Настройки на определенный тип терминала.

7.2. Машиночитаемые информационные ресурсы и их классификация. Генераторы БД. Операторы/арендаторы БД. Центры коммутации сообщений. Конечные пользователи. Генераторы и распространители (операторы) БД, классификация. Обзор состояния информационного рынка. Некоторые экономические характеристики информационных потоков генераторов БД, сравнительный анализ. Сравнительный анализ экономических характеристик продуктов и услуг операторов БД.

7.3. Обмен файлами. Архитектура взаимодействия программ. Настройка программы-сервера. Анонимный доступ к удаленной файловой системе. Организация каталогов на удаленной системе и защита от несанкционированного доступа. Электронная почта. Принципы организации системы электронной почты. Программа-сервер сообщений. Организация почтовых ящиков. Программы подготовки сообщений и рассылки. Формат почтового сообщения. Телеконференции. Принципы организации программного обеспечения телеконференции. Подписка. Сервер телеконференции. Структура почтового сообщения. Стиль диалога. Почтовые файловые серверы. Почтовый сервер – назначение и принципы работы. Команды сервера. Система приоритетов в системе электронной почты.

7.4. Конкретные информационные и файловые системы в INTERNET. Gopher, WAIS (Wide Area Information Servers), WWW (World Wide Web), особенности Web-приложений. Принципы организации. Архитектура информационных массивов. Языки запросов. Средства отображения информации. Организация гипертекстового документа. Язык разметки HTML. Встроенные графические образы. Программы отображения и воспроизведения нетекстовой информации. Протокол обмена HTTP. Организация глобальной гипертекстовой сети.

7.5. Семантическая паутина (Semantic Web). Введение в Семантическую паутину. Краткая история развития Всемирной паутины. Основные технологии Semantic Web. Структура SW-приложения. Проект Linking Open Data. Перспективы развития Semantic Web.

7.6 Защита информационных систем и ресурсов и прав на них. Информационное оружие в информационной войне. Особенности правовой охраны и защиты прав на информационные системы и ресурсы. Виды противников или «нарушителей». Три вида возможных

нарушений информационной системы. Основные положения теории информационной безопасности информационных систем. Модели безопасности и их применение. Таксономия нарушений информационной безопасности вычислительной системы и причины, обуславливающие их существование. Анализ способов нарушений информационной безопасности. Использование защищенных компьютерных систем. Методы криптографии. Основные технологии построения защищенных информационных систем.

Основная литература

1. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложения на C++. 2-е изд. М.: Бинوم; СПб: Невский диалект, 1999.
2. Гаврилова Т.А. Хорошевский В.Ф. Представление знаний в системах искусственного интеллекта. – СПб.: Питер-пресс, 2000.
3. Гэри М. Джонсон Д. Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982.
4. Дейт К.Дж. Введение в системы баз данных. – СПб.: Диалектика, 2005.
5. Ковалёв М.М. Дискретная оптимизация (целочисленное программирование) – М.: Эдиториал УРСС, 2003.
6. Лукашевич Н.В. Тезаурусы в задачах информационного поиска. – М.: Издательство Московского университета, 2011.
7. Мартин Дж. Организация баз данных в вычислительных системах, М.: Мир, 2000.
8. Михайлов А.И., Черный А.И., Гиляревский Р.Э. Основы информатики. – М.: Наука, 1978.
9. Нортон П. справочное руководство по MS DOS. Программные средства информационных систем. М., 1993. (Гл.7).
10. Попов И.И. Информационные ресурсы и системы: реализация, моделирование, управление: монография М.: ТПК Альянс, 1996.
11. Попов И.И., Максимов Н.В., Храмцов П.Б. Введение в сетевые информационные ресурсы и технологии. – М.: РГГУ, 2001.
12. Солодовников С.А. Бабайцев В.А., Браилов А.В, Математика в экономике. Учеб. М.; Финансы и статистика, 2002.

Дополнительная литература

1. Бережная Е.В., Бережной В.И. математические методы моделирования экономических систем. М: Финансы и статистика, 2002.
2. Волькенштейн М.В. Энтропия и информация. – М.: Наука, 2006.
3. Месарович М., Мако Д., Такахара И. Теория иерархических многоуровневых систем. М: Мир, 1973.
4. Николенко С.И., Кадурын А.А., Архангельская Е.О. Глубокое обучение. Погружение в мир нейронных сетей. – СПб.: Питер, 2022.
5. Попов И.И. автоматизированные информационные системы (по областям применения): Учеб. пособие М.: РЭА, 1999
6. Рассел С., Норвиг П. Искусственный интеллект: современный подход. – М.: Издательский дом «Вильямс», 2006.
7. Феллер В. Введение в теорию вероятностей и ее приложение в двух томах. М.: Мир 1982.