



**ІІТМО**

# **Принципы разработки компетентностной модели в сфере ИИ**

Санкт-Петербург, 2023

01.04.02 Глубокое обучение и генеративный искусственный интеллект

01.04.02 Большие данные и машинное обучение

01.04.02 Программирование и искусственный интеллект

01.04.02, 09.04.03 Искусственный интеллект и поведенческая экономика

02.04.01 Аналитика данных

09.04.01, 11.04.02, 27.04.05 Искусственный интеллект

09.04.02 Речевые технологии и машинное обучение

09.04.02 Искусственный интеллект в промышленности

09.04.04 Проектирование и разработка систем искусственного интеллекта

09.04.04 Проектирование и разработка систем больших данных

10.04.01 Безопасность систем искусственного интеллекта

15.04.06 Робототехника и искусственный интеллект

18.04.02 Химия и искусственный интеллект

# ИИ-образование в ИТМО (2/2): типизация



Категория программ	Абитуриент	ИП	Примеры программ
Мегапрограмма «Искусственный интеллект» (динамические траектории)	Работающие ИТ-шники	ИТ-компании	09.04.01, 11.04.02, 27.04.05 Искусственный интеллект
Гибридные программы «ИИ в ... определенной области» (динамические траектории)	Работающие отраслевики, выпускники отр. программ	Отраслевые компании	09.04.02 Искусственный интеллект в промышленности 09.04.03 Искусственный интеллект и поведенческая экономика
Программы-бутики	Выпускники ИТ-программ	Профильные компании по теме	01.04.02 Глубокое обучение и генеративный искусственный интеллект 01.04.02 Большие данные и машинное обучение 02.04.01 Аналитика данных
Доменные программы «Предметная область и ИИ»	Выпускники отраслевых программ	Профильные отраслевые компании	01.04.02. Программирование и ИИ 15.04.06. Робототехника и ИИ 18.04.02. Химия и ИИ

# Структура компетенций в ИТМО: свой СУОС

- ❖ Ключевые (социально-личностные и общекультурные) компетенции
- ❖ Надпрофессиональные компетенции
- ❖ Общепрофессиональные компетенции
- ❖ Профессиональные компетенции

Стандартные (СУОС ИТМО)

Руководители ОП (профстандарты, форсайт, универсальная модель компетенций в области ИИ Плеханова)

## Сборка компетентностной модели по программа ИИ:

- Ключевые, надпрофессиональные и общепрофессиональные компетенции выбираются из стандарта университета ИТМО (без изменений)
- РОП может добавить при необходимости дополнительные индикаторы к уже существующим индикаторам ОПК (?)
- Выбор профессиональных компетенций из банка

Создание собственного банка профессиональных компетенций в области ИИ с рекомендациями по категориям ОП

# Какой должна быть компетентностная модель:



- ❖ Актуальной (предусматривает возможности актуализации компетенций)
- ❖ Удобной (максимально учитывать многообразие компетенций, формируемых разными ОП в сфере ИИ)
- ❖ Гибкой (чтобы ее могла использовать любая образовательная программа по ИИ в вузе)
- ❖ Унифицированной (компетенции близкие по содержанию нужно унифицировать с возможностью добавлять индикаторы)
- ❖ Предусмотреть “пакеты компетенций”, описывающие различные профессии и профессиональные роли
- ❖ Обладать «взглядом в будущее» (место для новых компетенций)

Шаг 1. Опрос представителей ОП по их видению портрета выпускника:

**Отказываемся от формализма, работаем в терминах рынка труда**

*Максимально подробно описать портрет выпускника в терминах рынка труда (например, в формулировках требований к скиллам в вакансиях или в формулировках, которые встречаются в CV выпускника).*

Например:

- Знание Python и основных библиотек (numpy, pandas, scipy, sklearn); умение писать тесты на ds-компоненты и приложения, понимание дизайна и структуры ds-проектов.
- Уверенное владение популярными алгоритмами машинного обучения и специализированными библиотеками (LightGBM, CatBoost и др.)

# Ход работы:

## Итог:

- ❖ **213 формулировок** результатов обучения
- ❖ Покрытие большого стека технологий
- ❖ Похожие формулировки, которые отличались детализацией
- ❖ Несколько ярковыраженных “ядер” - результатов, которые встречаются в нескольких программах

Знание Python и основных библиотек

Знание и владение Python, в том числе основными библиотеками

Владение Python и основных библиотек (numpy, pandas, scipy, sklearn)

	В	Д	Е	Ф	Г
	ИИ в пром				
	Data Scientist	AI Analyst	AI PM	Data Scientist	
Знание истории развития и понимание главных трендов ИИ					
Понимание основных проблем цифровой трансформации предприятий в различных отраслях экономики	1				
Знание основных подходов к цифровой трансформации предприятий в различных отраслях экономики	1				
Знание Python и основных библиотек					
Знание и владение Python, в том числе основными библиотеками					
Владение Python и основных библиотек (numpy, pandas, scipy, sklearn)	2				
Знание основных библиотек Python для задач машинного обучения (numpy, pandas, scipy, sklearn)					
Знание основных библиотек Python для задач глубокого обучения (pytorch, tensorflow, keras)					
Знание основных библиотек Python для байесовской оптимизации (optuna и hyperopt)					
Знание основных библиотек Python для построения комбинаций алгоритмов машинного обучения (autosklearn, autoweke, tpot)					
Знание основных библиотек Python для					

# Ход работы:

## Шаг 2. Кластеризация полученных результатов обучения:

- ❖ Программирование
- ❖ Базы данных (исключая инфраструктуру БД)
- ❖ Инфраструктура БД
- ❖ Машинное обучение (вкл. Deep learning, NLP, без статистического анализа)
- ❖ Статистический анализ ?
- ❖ AI PM

**ІТМО**

Языки программирования и стандартные библиотеки для разработки ПО

Инструменты промышленной разработки программных систем

Работа с большими данными

Архитектура ПО



Структуры и алгоритмы обработки данных

Серверное программирование



# Ход работы:

## Шаг 3. Приведение к единой иерархической структуре, выделение следующих уровней:

- ❖ Домены (математика, разработка)
- ❖ Области (математические основы ИИ, Машинное обучение, Глубокое обучение, другие направления ИИ, языки программирования, инфраструктура БД, жизненный цикл систем ИИ)
- ❖ Направления  компетенции
- ❖ Темы (знания)  индикаторы:  
характеристики сущностей:  
по грейдам (уровням владения)



## Роль Data Engineer

1. Разрабатывает, поддерживает и развивает ETL-процессов
2. Проводит поиск и исследование систем-источников данных, а также выявляет возможности и способы интеграции
3. Работает с данными: сбор данных из разрозненных источников, проверка данных на корректность, очистка и приведение данных к виду, пригодному для дальнейшей обработки и анализа.
4. Взаимодействует с инженерными командами, специалистами по данным и другими заинтересованными сторонами для понимания как данные можно использовать для удовлетворения потребностей бизнеса.
5. Обеспечивает доступность данных в любое время и выполняет реализацию политик данных компании в отношении конфиденциальности и конфиденциальности данных.
6. Понимает потребности пользователей, предлагает решения."

# Ход работы:



компетенции

сущности

характеристики сущностей

C	D	E	F	G	H	I
Код компетенции	Направление	Формулировка компетенции (проект)	Темы (знания)	Умения, уровень Базовый	Умения, уровень Продвинутый	Умения, уровень Экспертный
	Основы теории вероятностей и математической статистики	Способен применять основы теории вероятностей и математической статистики для решения задач искусственного интеллекта	Оценка распределений. Оценка статистических зависимостей.	Применяет базовые методы непараметрической оценки распределений, базовые методы оценки параметров типовых распределений и методы оценки статистических линейных зависимостей случайных величин, владеет терминологией.	Применяет подходы к оценке и конструированию распределений сложной формы, методы оценки нелинейной статистической зависимости, в т.ч. для случайных процессов.	
			Проверка статистических гипотез.	Применяет популярные статистические критерии для проверки базовых гипотез, владеет терминологией.	Формулирует отличия в постановке задачи о проверке гипотезы от постановки для популярных критериев, применяет специализированные критерии.	Формулирует нестандартные гипотезы и самостоятельно разрабатывает статистические критерии на основе приближенных генеративных методов.
			Многомерные вероятностные модели.	Строит непараметрические и параметрические модели для низкоразмерных распределений (двумерных, трехмерных).	Понимает теоретические основы графических вероятностных моделей и знает их основные виды, формализует связь между вероятностными моделями и генеративными моделями машинного обучения, обучает и применяет многомерные графовые вероятностные модели на практике.	Понимает теоретические основы многомерных вероятностных моделей случайных процессов, математически формализует вероятностную постановку задачи для анализа многомерного случайного процесса, строит и применяет вероятностные модели многомерных случайных процессов.
			Марковские процессы.	Идентифицирует и применяет модели на основе марковских цепей и марковские скрытые модели, понимает основы теории марковских процессов.	Понимает теоретические основы марковских процессов принятия решений, математически формализует связь алгоритмов обучения с подкреплением и марковских процессов принятия решений.	Понимает углубленную теорию марковских процессов, разрабатывает новые методы и алгоритмы на основе марковских процессов.

индикаторы

функции ролей

Основы теории вероятностей и математической статистики



Способен применять основы теории вероятностей и математической статистики для решения задач искусственного интеллекта

Современные архитектуры генеративных глубоких сетей



Способен применять и (или) разрабатывать современные архитектуры генеративных глубоких сетей

# Проект формулировок компетенций:

## Математические основы ИИ (MF):

- ❖ Способен применять основы теории вероятностей и математической статистики для решения задач искусственного интеллекта
- ❖ Способен применять аппарат байесовской статистики и моделирования
- ❖ Способен применять численные методы для решения задач искусственного интеллекта
- ❖ Способен применять статистические методы машинного обучения
- ❖ Способен применять средства современной математики для моделирования объектов сложной структуры

## Работа с данными (WD):

- ❖ Способен осуществлять предварительный анализ данных
- ❖ Способен применять методы и инструменты сбора и разметки данных.

## Машинное обучение (ML):

- ❖ Способен анализировать и учитывать историю развития и основные тренды ИИ, проблематику больших данных, в том числе развитие технологий больших данных и их особенности
- ❖ Способен применять классические алгоритмы машинного обучения
- ❖ Способен разрабатывать и (или) применять методы повышения устойчивости, надежности, безопасности алгоритмов МО
- ❖ Способен применять алгоритмы обучения с подкреплением
- ❖ Способен применять автоматическое машинное обучение
- ❖ Способен применять алгоритмы обучения на нестандартных объемах данных



## Глубокое обучение (DL):

- ❖ Способен применять и (или) разрабатывать архитектуры глубоких нейронных сетей
- ❖ Способен применять и (или) разрабатывать современные архитектуры генеративных глубоких сетей
- ❖ Способен применять и (или) разрабатывать алгоритмы компьютерного зрения
- ❖ Способен применять и (или) разрабатывать алгоритмы обработки естественного языка
- ❖ Способен применять и (или) разрабатывать алгоритмы распознавания и генерации речи

## Другие направления ИИ:

- ❖ Способен осуществлять управление знаниями, в том числе с применением алгоритмов интеллектуального поиска решений и формирования стратегий
- ❖ Способен применять и (или) разрабатывать мульти-агентные алгоритмы
- ❖ Способен применять и (или) разрабатывать интеллектуальные методы оптимизации

## Языки программирования (PL):

- ❖ Способен применять язык программирования Python для решения задач в области ИИ
- ❖ Способен применять языки программирования Java/Scala для решения задач в области ИИ
- ❖ Способен применять языки программирования C/C++ для решения задач в области ИИ

# Проект формулировок компетенций:



## Инфраструктура больших данных (BD):

- ❖ Способен применять различные модели и (или) технологии хранения данных
- ❖ Способен применять различные модели и (или) технологии обработки данных
- ❖ Способен применять технологии организации инфраструктуры БД

## Жизненный цикл систем ИИ (LC):

- ❖ Способен проводить анализ бизнес-проблем с оценкой перспективности применения ИИ для их решения, осуществлять постановку задачи машинного обучения, формулировать требования к системе ИИ
- ❖ Способен определять требования к наборам данных для решения задач машинного обучения, проводить поиск, сбор, разметку и анализ наборов данных, оценивать качество данных, обеспечивать непрерывную интеграцию данных
- ❖ Способен проводить эксперименты на данных, формулировать гипотезы исследования, строить (обучать, дообучать) модели машинного обучения с оценкой их качества и анализом ошибок, обеспечивать воспроизводимость и масштабируемость исследований на данных
- ❖ Способен разрабатывать архитектуры систем искусственного интеллекта
- ❖ Способен управлять процессом жизненного цикла ИИ-продукта, в том числе организовывать и руководить работой команды
- ❖ Способен применять и (или) проектировать различные инструменты промышленной разработки ПС

# Чем мы отличаемся от существующих моделей:



- ❖ **От модели Плеханова:** формулировки в терминах рынка
- ❖ **От модели Альянса:** в нашем случае это модель компетенций, а не перечень тем
- ❖ Детальная ролевая модель



Спасибо за внимание!

it's **MO**re than a  
**UNIVERSITY**