Новые профессии, рождаемые развитием искусственного интеллекта

Александр Николаевич Горбань

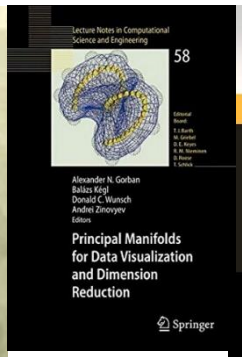
University of Leicester, UK

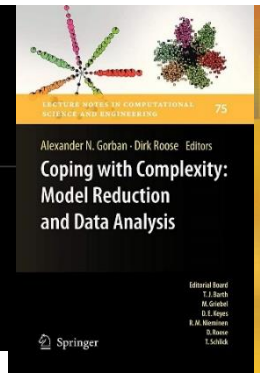
Нижегородский государственный университет им. Н.И. Лобачевского, Россия

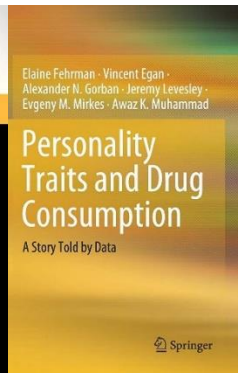
Я верю, что это так; что мы находимся на пороге эпохи, в которой будет очень велико влияние машин для решения интеллектуальных задач. Весьма возможно, что они будут доминировать.

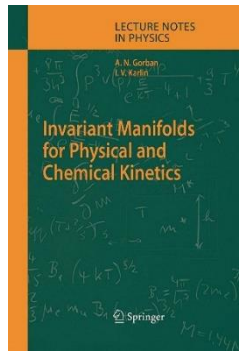
Marvin Minsky, 1961

Позвольте представиться









Alexander N. Gorban (АН Горбань)

ПОДПИСКА ОФОРМЛЕНА

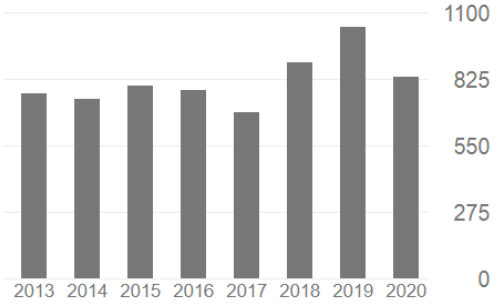
Professor of Applied Mathematics, University of Leicester, UK
Подтвержден адрес электронной почты в домене le.ac.uk - Главная страница

Multiscale Analysis Model Reduction Kinetic Equations Mathematical Chemistry Mathematical Neuroscience

НАЗВАНИЕ	ПРОЦИТИРОВАНО	ГОД
Neural networks on personal computer AN Gorban, DA Rossiev Novosibirsk: the Science, 276	1081 *	1996
Training neural networks AN Gorban Moscow: USSR-USA JV" Paragraph", Russian. English translation: AMSE Trans …	736 *	1990
Neiroinformatika AN Gorban, WL Dunin-Barkowski, AN Kirdin, EM Mirkes, ... Neuroinformatics), Novosibirsk: Nauka	659 *	1998

Процитировано ПРОСМОТРЕТЬ ВСЕ

	Все	Начиная с 2015 г.
Статистика цитирования	12408	5051
h-индекс	53	35
i10-индекс	191	112



Позвольте

Оргкомитет девятнадцати ежегодных международных конференций
Нейроинформатика 1999 - 2017 и
Правление Российской ассоциации нейроинформатики

В знак признания выдающихся успехов в области теории и приложений искусственных нейронных сетей и за неоценимый вклад в организацию и проведение конференций присвоили

ПРОФЕССОРУ
ГОРБАНЮ АЛЕКСАНДРУ НИКОЛАЕВИЧУ

почётное звание

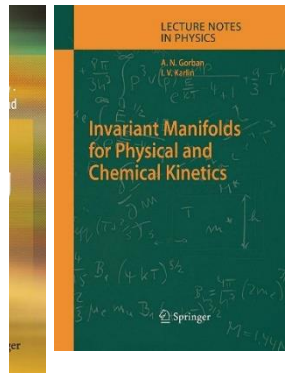
ПИОНЕРА РОССИЙСКОЙ НЕЙРОИНФОРМАТИКИ

Проректор НИЯУ МИФИ — В.В. Ужва

Председатель оргкомитета девятнадцати ежегодных международных конференций Нейроинформатика 1999-2017, профессор — Б.Н. Оныкий

Президент Российской ассоциации нейроинформатики, член-корреспондент Российской академии наук — Б.В. Крыжановский

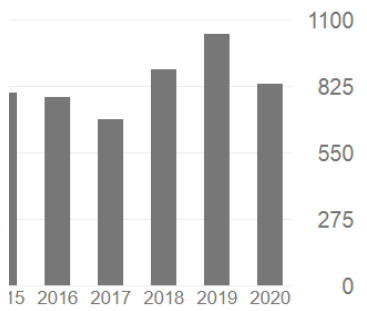
МОСКВА, 2 ОКТЯБРЯ 2017 г.



Alexander N. G
Professor of Applied M
Подтвержден адрес э
Multiscale Analysis M
Mathematical Neurosc

вано ПРОСМОТРЕТЬ ВСЕ

	Все	Начиная с 2015 г.
	12408	5051
	53	35
	191	112



НАЗВАНИЕ

- Neural networks on personal compute
  AN Gorban, DA Rossiev
  Novosibirsk: the Science, 276
- Training neural networks
  AN Gorban
  Moscow: USSR-USA JV" Paragraph", Russiar
- Neiroinformatika
  AN Gorban, WL Dunin-Barkowski, AN Kirdin, E
  Neuroinformatics), Novosibirsk: Nauka

«Кейнот» пленарный доклад по нейронным сетям на Всемирном Конгрессе Вычислительного Интеллекта (IEEE WCCI2020)

wcci2020.org/speakers/

IEEE WCCI 2020
IEEE World Congress on Computational Intelligence
Glasgow, Scotland, United Kingdom – July 19-24, 2020

WCCI 2020 | SPEAKERS | PROGRAM SCHEDULE | PROGRAMS | REGISTRATION | CONTACT US

IJCNN Keynote Speakers

ALEXANDER N GORBAN

Monday, 20th July, 2020 12:45-13:45 Professor of

KUNIHIKO FUKUSHIMA

Tuesday, 21st July, 2020 13:00-14:00 Kunihiko Fukushima received

План

1. Что бросается в глаза в Российском университете после 15 лет в университетах Запада? Отсутствие установки на постоянное самообновление.
2. Поток новых специализаций в небольшом департаменте за 15 лет.
3. Data analysis for business intelligence (Анализ данных для бизнес-аналитики) – чему они учатся
4. Два исторических типа ИИ.
5. Производственные операции в ИИ проектах: от сбора данных до валидации конечного продукта.
6. Пирамида профессий

Что бросается в глаза в Российском университете после 15 лет в университетах Запада?

- Специальности как музей прошлых достижений.
- (Почти) Нет постоянного поиска новых дисциплин, ориентированных на **завтрашний** спрос.
- (Почти) Нет исследовательских проектов, ориентированных на **послезавтрашние приложения.**
- Великие достижения надо сохранять, но... реальность убегает.

Бразды пушистые взрывая,
Летит кибитка удалая.
Ямщик сидит на облучке
В тулупе, в красном кушаке.

Современный детский рисунок (из сети)

Бразды пушистые взрывая,
Летит кибитка удалая.
Ямщик сидит на облучке
В тулупе, в красном кушаке.

Современный детский рисунок (из сети)

Что я должен делать, чтобы моя группа, департамент выжили послезавтра?

- Представить себе завтрашний/послезавтрашний мир, возникающие потребности производства, их ожидание у потребителей образования, хайпы и тренды в спросе на образование.
- Сформировать новые дисциплины, с учетом рынка этих услуг и развития ближайших конкурентов.
- Заполнить куррикулум (учебный план) предметами, которые мы можем вести (наша научная экспертиза позволяет нам быть во многих отношения впереди впроса на образование).
- Остались дыры? Выявить дефициты в нашей экспертизе и заполнить ее за счет собственного дообразования или (если ресурсы позволят), найма соответствующих экспертов.

Поток новых специализаций в небольшом департаменте за 15 лет

- 2005 Финансовая математика
- 2007 Актуарная математика
- 2007 Прикладные вычисления и численное моделирование
- 2012 Анализ данных для бизнес-аналитики
- 2020 Математика и ИИ (с 1 годом промышленой стажировки)
- 2020 Управление данными

Анализ данных для бизнес-аналитики – чему они учатся 1

- Подготовительные дисциплины: статистика, мат. моделирование, элементы программирования, эконометрика.

Метафора производственных позиций и цепочек:

- Базовые технологичекие операции: *классификация, кластер-анализ, прогнозирование, оценка функций распределения, сокрашение описания* (по 3-5 алгоритмов выбора на каждую операцию).
- Оценка данных и основные операции предобработки (пре-процессинга).
- Анализ временных рядов.

Анализ данных для бизнес-аналитики – чему они учатся 2

Метафора производственных позиций и цепочек:

0!!! **Разработка и выполнение проектных заданий**. Этому учим на практике проектирования. Каждый студент выполняет 6-7 небольших проектов и 1 большой.

Примеры небольших проектов:

Даны накопленные данные по опухолям разной этиологии. Подготовить распознающую систему рак/не рак 3мя способами. Сравнить разные технологические решения.

По базе данных постройте систему предсказания выборов американских президентов 2мя способами. . Сравните разные технологические решения.

Проанализировать предсказуемость 3 финансовых временных рядов и эффективность нескольких методов технического анализа.

Система связанных дисциплин

- Искусственный интеллект (ИИ)
- Интеллектуальный анализ данных
- Машинное обучение
- Статистика

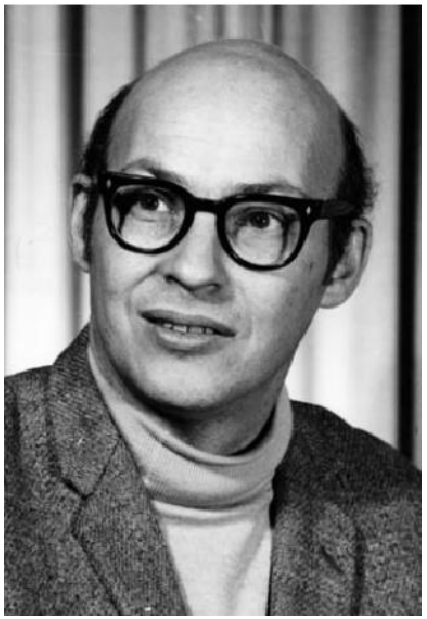
Анализ данных

Проводятся лекции и семинары о том, как связаны эти дисциплины, в чем сходство и различие.

Дартмутский семинар – два месяца летом 1956. 10 классиков

Джон Маккарти –
лидер «аккуратных»

Мы попытаемся понять, как обучить машины использовать естественные языки, формировать абстракции и концепции, решать задачи, сейчас подвластные только людям, и улучшать самих себя.



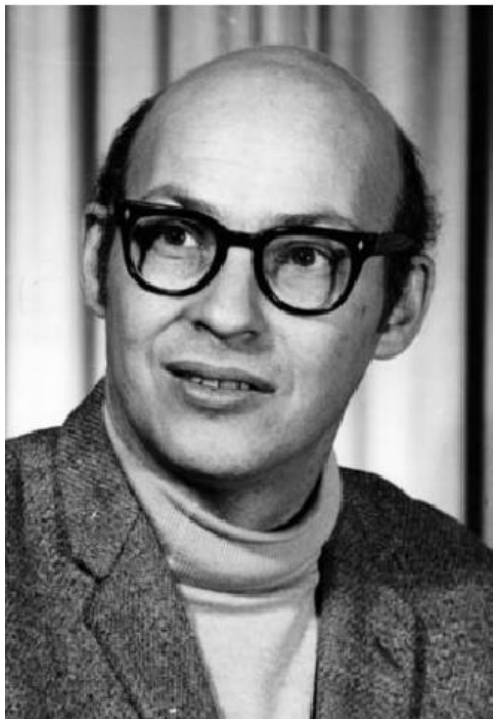
Марвин Минский –
лидер «неряшливых»

В 1959 Минский и Маккарти основали в MIT лабораторию информатики и искусственного интеллекта

Две базовых формы ИИ

- ***ИИ, основанный на знаниях*** – это очень ☺ продвинутая форма справочника.

- ***ИИ, основанный на данных*** – это очень ☺ усовершенствованная форма статистики. По определению 1833 года: «цель статистики заключается в представлении фактов в наиболее сжатой форме» (для учета и прогноза).



Типы задач для ИИ

Удобно разделить задачи на пять основных направлений: поиск, распознавание образов, самообучение, планирование и индукция (индуктивный вывод)… .

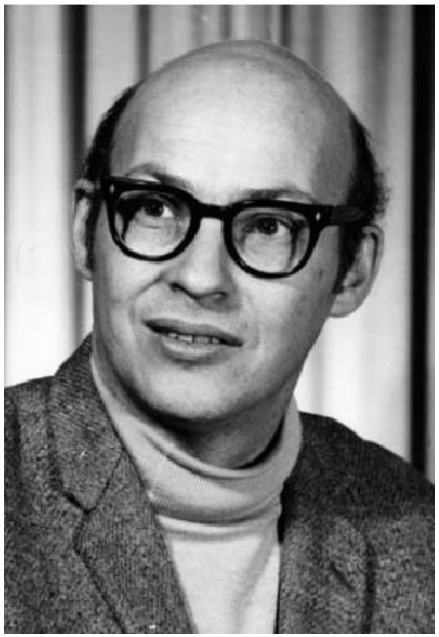
Marvin Minsky, 1961

PROCEEDINGS OF THE IRE *Janu*

Steps Toward Artificial Intelligence*

MARVIN MINSKY†, MEMBER, IRE

The work toward attaining "artificial intelligence" is the center of considerable computer research, design, and application. The field is in its starting transient, characterized by many varied and independent efforts. Marvin Minsky has been requested to draw this work together into a coherent summary, supplement it with appropriate explanatory or theoretical noncomputer information, and introduce his assessment of the state-of-



Типы задач для ИИ

Удобно разделить задачи на пять основных направлений: поиск, распознавание образов, самообучение, планирование и индукция (индуктивный вывод)… .

Marvin Minsky, 1961

- Проблемы Минского сосредоточили исследования на наиболее важных технических вопросах.
- Где-то 90% НН публикаций до сих пор состедоточено на постепенном продвижении в этих проблемах.
- В настоящее время мы видим впечатляющие успехи в этих областях.

Соревнования по отдельным задачам проводятся на всех крупных конференциях

Интеллектуальный анализ данных: технологическая цепочка

1. Определение целей, задач и структуры проекта
2. Техническое задание на сбор данных
3. Сбор данных
4. Оценка и предобработка данных. Формирование баз для обучения
5. Сокращение данных
6. Подбор последовательности базовых операций
7. Выбор алгоритмов, архитектур и технологий обучения
8. Обработка (обучение)
9. Валидация, интерпретация и возврат назад к 1-8, если надо
10. Интеграция в систему пользователя
11. Пробная эксплуатация и возврат назад к 1-10, если надо

Интеллектуальный анализ данных: технологическая цепочка

1. Определение целей, задач и структуры проекта
2. Техническое задание на сбор данных
3. **Сбор данных**
4. Оценка и предобработка данных. Формирование баз для обучения
5. Сокращение данных
6. Подбор последовательности базовых операций
7. Выбор алгоритмов, архитектур и технологий обучения
8. Обработка (обучение)
9. Валидация, интерпретация и возврат назад к 1-8, если надо
10. Интеграция в систему пользователя
11. Пробная эксплуатация и возврат назад к 1-10, если надо

Самые массовые работы – сбор данных и управление данными

- Сбор данных включает и различные работы, специфические для проблемной области, и ряд общих технологических элементов складирования и управления собранными данными.
- В различных областях возникают специализированные индустрии сбора данных для формирования ИИ решений.
- Быстро развивается общая дисциплина управления данными (data management and integration).
- *Иеархия профессий: от сборщиков нижнего уровня, до квалифицированных пользователей высокотехнологичных систем управления и интеграции данных и, наконец, разработчиков этих систем.*
- Данные собираются «впрок» для дальнейшего использования во многих потенциальных и пока неизвестных проектах. *Нужны еще инженеры-проектировщики такого сбора данных, умеющие предвидеть потребности будущих проектов.*

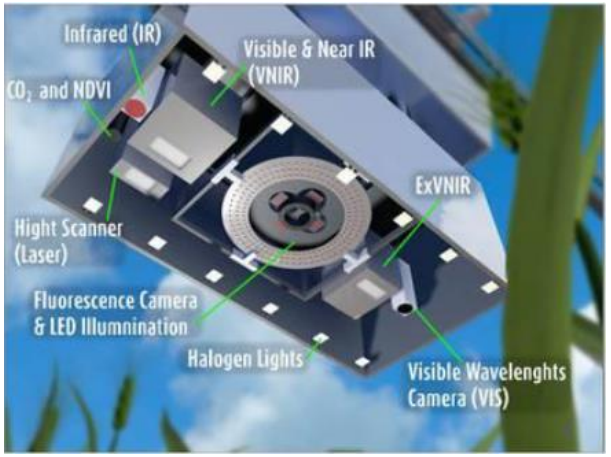
Пример сервиса сбора и разметки данных - **Яндекс.Толока**

Яндекс Карты

Команда Яндекс.Карт:

«Наша служба следит за актуальностью данных на Яндекс.Картах: мы обновляем часы работы организаций, добавляем новые, удаляем закрытые. Наши задания толокеры выполняют на улицах города — например, фотографируют таблички с графиком работы. Каждый месяц мы даём 15 миллионов подобных заданий 50 тысячам исполнителей — и только благодаря этому можем поддерживать огромную базу Яндекс.Карт в актуальном состоянии».

Пример – сбор цифровой информации о растениях по мультиспектральным наблюдениям для будущих ИИ проектов



Интеллектуальный анализ данных: технологическая цепочка

1. Определение целей, задач и структуры проекта
2. Техническое задание на сбор данных
3. Сбор данных
4. **Оценка и предобработка данных. Формирование баз для обучения**
5. Сокращение данных
6. Подбор последовательности базовых операций
7. Выбор алгоритмов, архитектур и технологий обучения
8. Обработка (обучение)
9. Валидация, интерпретация и возврат назад к 1-8, если надо
10. Интеграция в систему пользователя
11. Пробная эксплуатация и возврат назад к 1-10, если надо

Спрос на управление данными доминирует

ITJobsWatch

Database & Business Intelligence Skills

Home | UK Job Vacancies | WFH | CV Upload

Permanent IT Jobs | Contract IT Jobs

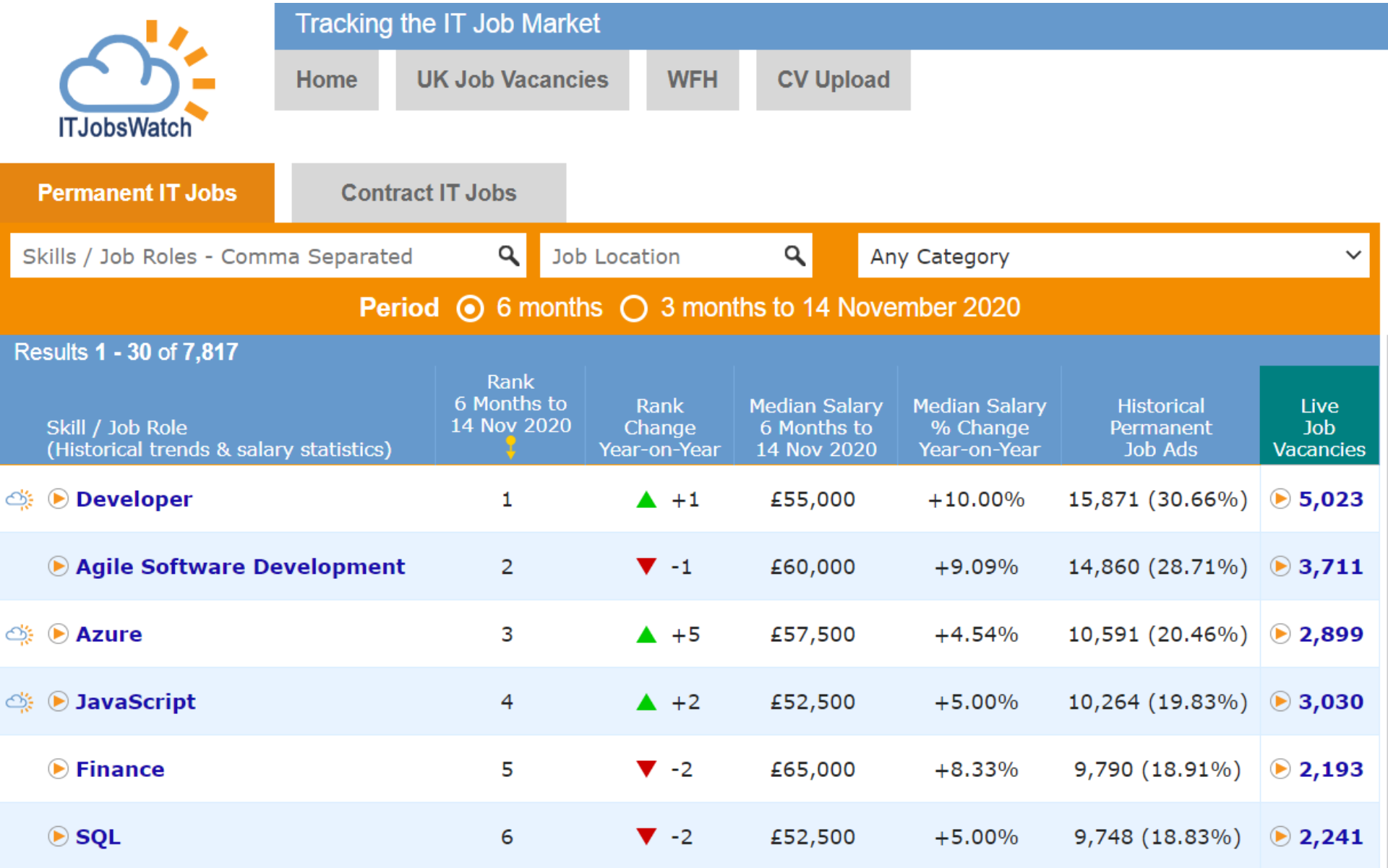
Skills / Job Roles - Comma Separated | Job Location | Database & Business Intelligence

Period ◉ 6 months ○ 3 months to 14 November 2020

Results **1 - 30** of **173**

Skill / Job Role (Historical trends & salary statistics)	Rank 6 Months to 14 Nov 2020	Rank Change Year-on-Year	Median Salary 6 Months to 14 Nov 2020	Median Salary % Change Year-on-Year	Historical Permanent Job Ads	Live Job Vacancies
SQL Server	18	▼ -3	£50,000	+5.26%	5,230 (10.10%)	**1,358**
MySQL	67	▲ +6	£50,000	+8.69%	2,211 (4.27%)	**759**
NoSQL	85	▲ +50	£60,000	-	1,941 (3.75%)	**372**
Data Warehouse	105	▲ +6	£55,000	-	1,546 (2.99%)	**279**
Big Data	107	▼ -12	£65,000	-7.14%	1,534 (2.96%)	**345**
Power BI	116	▲ +27	£50,000	-	1,463 (2.83%)	**363**

А вообще в IT skills лидирует *Agile*



Skill / Job Role (Historical trends & salary statistics)	Rank 6 Months to 14 Nov 2020	Rank Change Year-on-Year	Median Salary 6 Months to 14 Nov 2020	Median Salary % Change Year-on-Year	Historical Permanent Job Ads	Live Job Vacancies
Developer	1	▲ +1	£55,000	+10.00%	15,871 (30.66%)	5,023
Agile Software Development	2	▼ -1	£60,000	+9.09%	14,860 (28.71%)	3,711
Azure	3	▲ +5	£57,500	+4.54%	10,591 (20.46%)	2,899
JavaScript	4	▲ +2	£52,500	+5.00%	10,264 (19.83%)	3,030
Finance	5	▼ -2	£65,000	+8.33%	9,790 (18.91%)	2,193
SQL	6	▼ -2	£52,500	+5.00%	9,748 (18.83%)	2,241

Agile - гибкая методология разработки

- Разработка сводится к серии коротких циклов, называемых итерациями (две-три недели). Каждая итерация - программный проект в миниатюре.
- По окончании каждой итерации команда выполняет переоценку приоритетов разработки.
- Основные идеи:
  люди и взаимодействие важнее процессов и инструментов;
  работающий продукт важнее исчерпывающей документации;
  сотрудничество с заказчиком важнее согласования условий контракта;
  готовность к изменениям важнее следования первоначальному плану.

Azure – облачная платформа компании Microsoft

- Для разработки, выполнения приложений и хранения данных на серверах, расположенных в распределённых дата-центрах.
- **Опять управление данными!**

Интеллектуальный анализ данных: технологическая цепочка

1. Определение целей, задач и структуры проекта
2. Техническое задание на сбор данных
3. Сбор данных
4. **Оценка и предобработка данных. Формирование баз для обучения**
5. **Сокращение данных**
6. **Подбор последовательности базовых операций**
7. **Выбор алгоритмов, архитектур и технологий обучения**
8. **Обработка (обучение)**
9. **Валидация, интерпретация** и возврат назад к 1-8, если надо
10. Интеграция в систему пользователя
11. Пробная эксплуатация и возврат назад к 1-10, если надо

Специалист по машинному обучению (технический)

Основные компетенции

- **Оценка и предобработка данных. Формирование баз для обучения**
- **Сокращение данных**
- **Подбор последовательности базовых операций**
- **Выбор алгоритмов, архитектур и технологий обучения**
- **Обработка (обучение)**
- **Валидация, интерпретация**

Первое (известное мне) состязание по машинному обучению: 1я (и последняя) Всесоюзная Олимпиада студентов и школьников по нейроинформатике, 1990

Специалист по машинному обучению (технический)

В командах наблюдается дальнейшее разделение труда.

Разрез идет или по типам работ или по группам алгоритмов и методов. (*Например, Валидация требует компетенции в статистике*.)

Но даже чаще – по используемым техническим средствам. Здесь рынок захватывают универсальные инструментарии.

Дополнительные компетенции нужны для взаимопонимания и гибкости.

Спрос на Машинное обучение раз в 5-7 меньше, чем на управление данными

ITJobsWatch

Tracking the IT Job Market

Home | UK Job Vacancies | WFH | CV Upload

Permanent IT Jobs | Contract IT Jobs

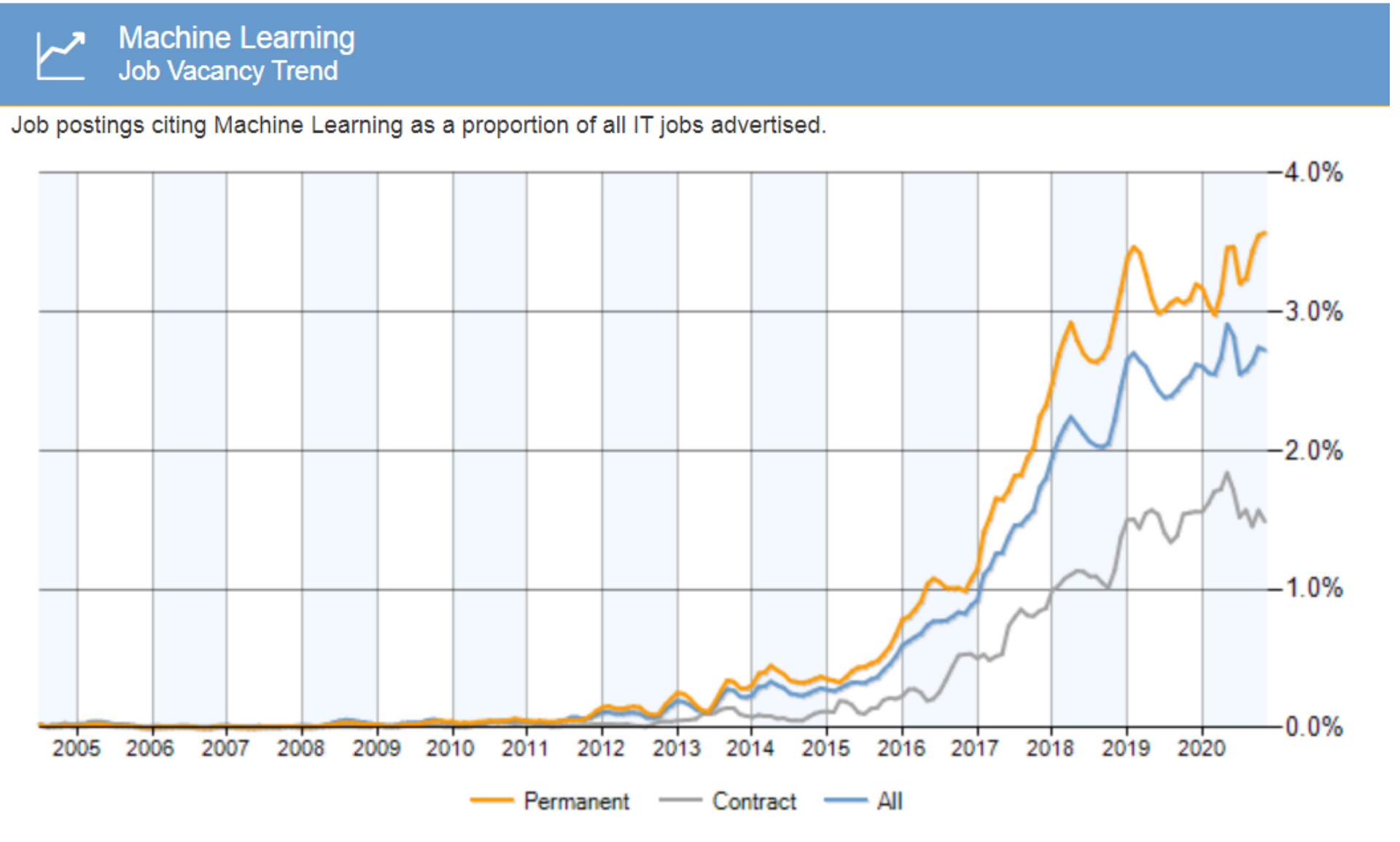
Machine Learning | Job Location | Any Category

Period ◉ 6 months ○ 3 months to 14 November 2020

Results **1 - 29** of **29**

Skill / Job Role (Historical trends & salary statistics)	Rank 6 Months to 14 Nov 2020	Rank Change Year-on-Year	Median Salary 6 Months to 14 Nov 2020	Median Salary % Change Year-on-Year	Historical Permanent Job Ads	Live Job Vacancies
Machine Learning	90	▲ +12	£62,500	-3.84%	1,759 (3.40%)	**429**
Machine Learning Engineer	602	▲ +258	£70,000	-	112 (0.22%)	**29**
Machine Learning Analyst	713	▲ +343	£45,000	-5.26%	1 (0.002%)	
Azure Machine Learning	698	▲ +243	£55,768	-14.20%	16 (0.031%)	**2**
Machine Learning Scientist	677	▲ +319	£70,000	+7.69%	37 (0.072%)	**4**

Доля вакансий с машинным обучением в IT. Насыщение лет через 5?



Интеллектуальный анализ данных: технологическая цепочка

1. Определение целей, задач и структуры проекта
2. Техническое задание на сбор данных
3. Сбор данных
4. Оценка и предобработка данных. Формирование баз для обучения
5. Сокращение данных
6. Подбор последовательности базовых операций
7. Выбор алгоритмов, архитектур и технологий обучения
8. Обработка (обучение)
9. Валидация, интерпретация и возврат назад к 1-8, если надо
10. Интеграция в систему пользователя
11. Пробная эксплуатация и возврат назад к 1-10, если надо

Но кто все это придумывает?

- Кто создает проекты сбора данных и бизнес услуг по сбору данных?
- Кто создает новые методы и подходы?
- Кто разрабатывает проекты ИИ для нужд заказчика?
- Это особая профессиональность.
- Я не знаю как ей учить. Но если дать хорошее образование и выполнить несколько хороших проектов, то это иногда происходит.

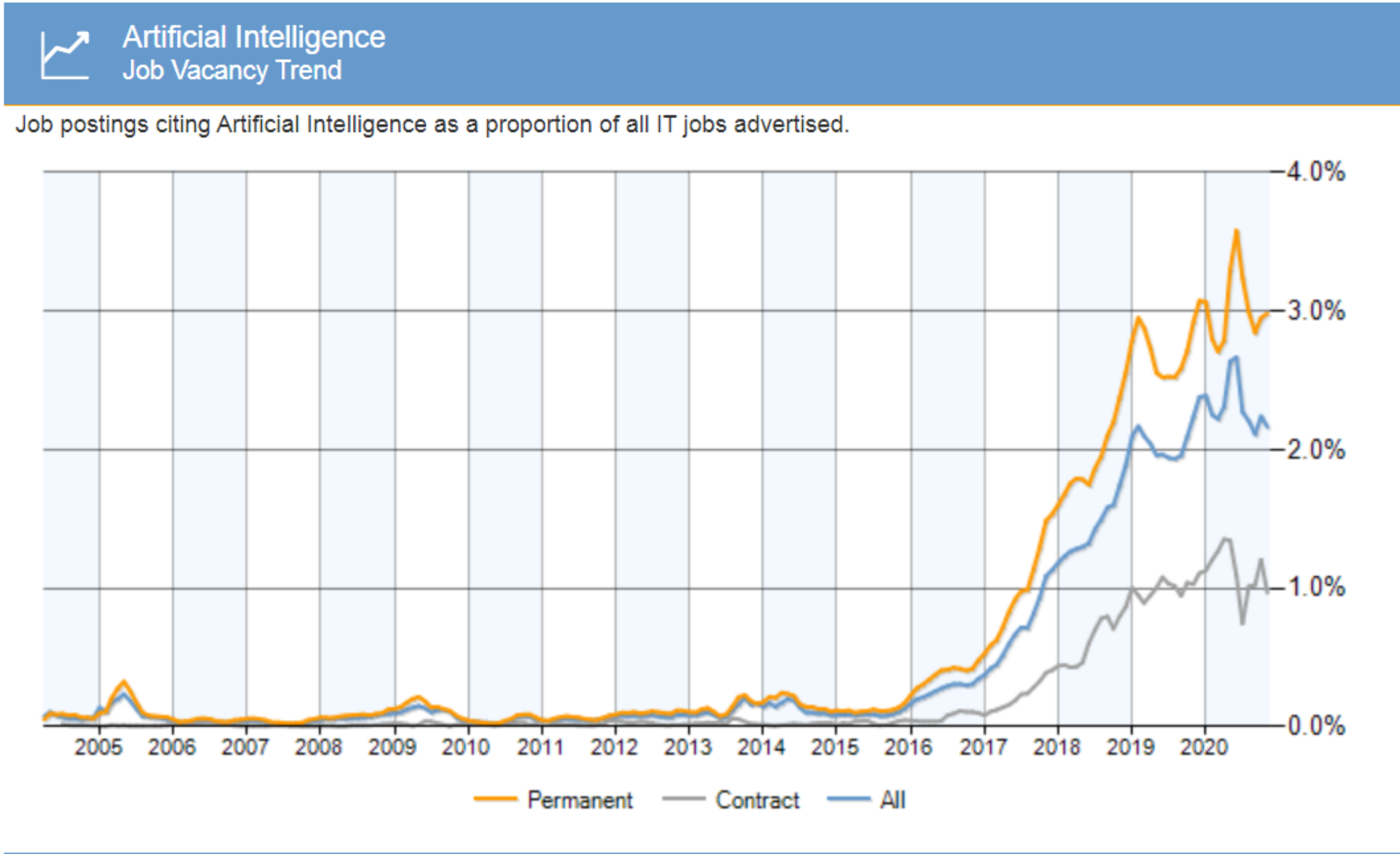
Профессии, обслуживающие создание ИИ

1. Сбор данных
2. Управление данными (включая интеграцию)
3. Машинное обучение
4. Проектирование ИИ систем
5. Создание универсальных ИИ платформ и инструментариев
6. Создание компьютерных систем, ориентированных на ИИ приложения

И, кроме того, нужны просто программисты широкого профиля для подержки.

Профессии, обслуживающие создание ИИ

1. Сбор данных
2. Управление данными (включая интегр
3. Машинное обучение
4. Проектирование ИИ систем
5. Создание универсальных ИИ платф инструментариев
6. Создание компьютерных систем, ориентированных на ИИ прилож

И, кроме того, нужны просто прогр широкого профиля для подержки.

Доля вакансий с Искусственным Интеллектом в IT. Насыщение приближается?