

КРАСОВСКИЙ Р. В.
ПРОВЕРКА ОДНОРОДНОСТИ РЕЖИМА СБОРА И ДАТИРОВКА РЯДОВ
ВОСТРЕБОВАННОСТИ ИТ-НАВЫКОВ

УДК 004.62:311, ГРНТИ 28.29.51

Статья поступила в редакцию 10.09.2026

Проверка однородности режима сбора
и датировка рядов востребованности
ИТ-навыков

Checking the consistency of data
collection methods and dating time series
of IT skill demand

Р. В. Красовский

R. V. Krasovsky

Пермский национальный
исследовательский политехнический
университет; г. Пермь

Perm National Research Polytechnic
University; Perm

Предложены три процедуры проверки месячных рядов востребованности ИТ-навыков до прогнозирования: контроль целостности привязки навыков, панельный критерий структурного разрыва и поправка на предметный состав корпуса стандартизацией. На корпусе из 893 710 записей показано, что порядок процедур определяет результат: без контроля целостности критерий ложно объявляет разрывы, порождённые дефектом привязки, а не сменой режима сбора. Дрейф состава подтверждён измерением. Стандартизация по классам названий вакансий снимает один разрыв, удлиняет однородный отрезок и снижает шум оценивания. Смена датировки сама по себе отрезок не удлиняет. Сделан вывод: оценка перспективности технологий должна начинаться с проверки ряда, а не с выбора модели.

Three procedures are proposed for testing monthly series of IT skill demand before forecasting: a control of the integrity of skill attachment, a panel structural break criterion, and a correction for the subject composition of the corpus by standardisation. On a corpus of 893,710 vacancy records it is shown that the order of the procedures determines the result: without the integrity control the criterion falsely declares breaks produced by a defect of skill attachment rather than by a change of the collection regime. The drift of composition is confirmed by measurement. Standardisation by vacancy title classes removes one break, lengthens the homogeneous segment and reduces estimation noise. Re-dating alone does not lengthen the segment. It is concluded that assessing technology prospects should begin with testing the series rather than with choosing a model.

Ключевые слова: рынок труда, востребованность навыков, временные ряды, структурный разрыв, панельный критерий, датировка данных, шум оценивания, дрейф состава корпуса, качество данных, образовательные программы

Keywords: labor market, skill demand, time series, structural break, panel criterion, data dating, estimation noise, corpus composition drift, data quality, educational programs

Введение

Помесячные ряды востребованности навыков, строящиеся автоматизированными системами мониторинга рынка труда, служат исходными данными для оценки перспективности технологического стека образовательных программ [1, 2]. В цитируемых работах ряд принимается как данность: вопрос о том, отражает ли наблюдаемое изменение доли рыночное событие или изменение процедуры сбора, не ставится. Между тем ряд формируется по базе собранных вакансий, и любое изменение сбора – расширение перечня отслеживаемых профессий, смена расписания обхода, подключение новых источников – сдвигает наблюдаемые частоты, не являясь рыночным событием. Такой сдвиг неотличим от рыночной динамики средствами последующих прогнозных моделей: модель, обученная на неоднородном ряду, воспроизведёт артефакт сбора как сигнал спроса.

Методы извлечения навыков из текстов вакансий и их отображения на стандартизованные таксономии разработаны существенно полнее, чем контроль качества получаемого ряда. В [3] систематизированы архитектуры извлечения требований, в [4] и [5] извлечённые навыки отображаются на внешние классификации; предметом этих работ служит точность извлечения на отдельной вакансии, тогда как свойства агрегированного по месяцам ряда не рассматриваются. В [6] тренды навыков анализируются ретроспективно, в [7] ежемесячные ряды экстраполируются рекуррентными нейронными сетями, а в предшествующей работе автора [8] по наклонам рядов формируется перечень растущих технологий; ни в одной из них ряд предварительно не проверяется на однородность режима сбора, вследствие чего артефакт сбора неотличим от рыночной динамики.

Аппарат обнаружения структурных изменений развит и распространён на панельные данные [9], однако предложенные там критерии рассчитаны на длинные ряды с известным числом разрывов и не учитывают, что доля навыка есть выборочная оценка с биномиальным шумом, величина которого различна для редких и частотных навыков. Применительно к рядам востребованности навыков проверка однородности режима сбора, насколько известно автору, не проводилась. Близкий по постановке материал даёт практика прогнозирования по открытым статистическим рядам: в [10] модель регионального социального ресурса строится

по публикуемым показателям, в [11] отдельно разбирается аппроксимация факторов такой модели при разреженных наблюдениях, а в [12] помесечный ряд выпуска моделируется авторегрессией с явным выделением сезонной составляющей. Общим для этих работ является допущение о неизменности инструмента измерения: перечень показателей и методика их получения считаются фиксированными, вследствие чего всякое изменение ряда относится к объекту наблюдения. Ряд долей навыков этому допущению не удовлетворяет, поскольку предметный состав корпуса вакансий меняется вслед за настройками сборщика, и потому проверка однородности режима сбора здесь предшествует моделированию, а не следует за ним.

Целью работы является построение процедур, позволяющих до применения прогнозных моделей установить, пригоден ли помесечный ряд долей навыков как прокси рыночного спроса. Предлагаются три такие процедуры: панельный критерий структурного разрыва, разложение вариации ряда и контроль дрейфа предметного состава корпуса. Приводится их проверка на реальном корпусе вакансий.

Методология исследования

Ряд востребованности навыка строится следующим образом. Пусть D – корпус вакансий, $N(t)$ – число вакансий месяца t , а $n(s, t)$ – число вакансий этого месяца, в требованиях которых указан навык s . Востребованность измеряется долей

$$p(s, t) = \frac{n(s, t)}{N(t)}, \quad (1)$$

а не абсолютной частотой: объём собираемых вакансий меняется от месяца к месяцу в разы, и абсолютные частоты отражают прежде всего интенсивность сбора. Вакансии дедулицируются по адресу объявления с сохранением первого появления.

Изменение процедуры сбора сдвигает долю (1) одновременно у значительной части навыков, тогда как рыночная динамика перераспределяет доли постепенно и разнонаправленно. Это различие положено в основу критерия. Для каждой пары смежных месяцев и каждого навыка панели проверяется значимость изменения доли по двухвыборочному z -критерию

$$z = \frac{p_1 - p_2}{\sqrt{\bar{p}(1 - \bar{p})\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}, \quad \bar{p} = \frac{n_1 + n_2}{N_1 + N_2}, \quad (2)$$

где индексы 1 и 2 относятся к смежным месяцам. О разрыве судят по доле навыков, изменение у которых значимо при $\alpha = 0,01$. Порог, отделяющий разрывы от обычных переходов, не назначается вручную: наблюдаемые доли упорядочиваются, и порогом берётся среднее геометрическое соседней пары

значений с наибольшим относительным зазором. Разбиение поэтому воспроизводимо.

Дальнейший анализ идёт только внутри однородных сегментов: ряд, проложенный через разрыв, измерял бы смену процедуры сбора, а не рынок.

Наблюдаемые колебания доли складываются из динамики и случайности месячной выборки. Доля оценивается по $n(s, t)$ успехам в $N(t)$ испытаниях, поэтому дисперсия оценки равна $p(1 - p)/N(t)$. Доля вариации ряда, объяснимая шумом оценивания, определяется как

$$\rho(s) = \frac{\overline{p(s, t) (1 - p(s, t)) / N(t)}}{\text{Var}_t p(s, t)}. \quad (3)$$

Величина $\rho(s)$ используется как критерий пригодности навыка к прогнозированию: при $\rho \rightarrow 1$ наблюдаемые колебания ряда неотличимы от случайности выборки, и прогнозировать в нём нечего.

Третья процедура отвечает на то, к чему панельный критерий слеп по построению. Он сравнивает соседние месяцы и обнаруживает скачкообразную смену режима, но постепенный дрейф предметного состава корпуса, растянутый на несколько месяцев, для него невидим. Дрейф при этом не диагностируется, а устраняется: доля навыка считается при фиксированном составе прямой помесечной стандартизацией

$$p_{\text{ст}}(s, t) = \sum_g w_g \frac{n(s, g, t)}{N(g, t)}, \quad (4)$$

где g – страта, $N(g, t)$ – число вакансий страты в месяце, w_g – средняя доля страты за окно, а страты, пустые в месяце, выпадают из суммы с перенормировкой весов. Существенно, по чему берутся страты. Пометка сборщика о поисковом запросе для этого непригодна: доля модального названия вакансии внутри такой пометки опускается до 6 %, а у двух групп из восьми модальным названием оказывается «руководитель отдела продаж». Страты берутся по классу нормализованного названия самой вакансии, то есть по признаку, не зависящему от того, как сборщик её нашёл. Сегментация выполняется после поправки: разрыв, который поправка снимает, был сменой состава; разрыв, который не снимает, есть граница однородного отрезка.

Экспериментальные результаты

Данные взяты из корпуса, который накапливает система мониторинга (её эксплуатационный контроль описан в [13]): 893 710 записей, из них 719 089 уникальны по адресу объявления. Дата первого появления вакансии у источника восстановлена для 809 183 записей, то есть для 90,5 % корпуса: у 8,2 % вакансия к моменту обращения уже была недоступна, у 1,2 % не удалось разобрать

идентификатор. В панель входят навыки, встречающиеся во все месяцы отрезка и набравшие в нём не менее 500 упоминаний.

Сначала сборщик сохранял только дату записи вакансии, а не дату публикации работодателем. Насколько велико расхождение, показала случайная выборка из 950 вакансий: интерфейс источника отдал 89,4 % записей, причём для старых вакансий не реже, чем для свежих. Поэтому дату удалось восстановить для всего корпуса.

Задержка между появлением вакансии и её записью сильно скошена вправо: медиана нулевая, 75-й процентиль 1 день, 95-й 28 дней, а среднее 6,4 дня. Месяц записи совпадает с месяцем появления у 88,1 % вакансий, сдвинут на месяц у 8,9 % и больше чем на месяц у 3,0 %.

По времени совпадение разложено крайне неравномерно (Таблица 1): в первые месяцы наблюдения 52–64 %, начиная с июня 2025 г. уже 90–98 %, за единственным исключением ноября 2025 г. (75 %), на который приходится всплеск объёма сбора. Значит, поначалу сборщик выбирал накопленный архив источника, а не текущий поток вакансий, и только потом перешёл на сбор в темпе поступления. На этот же период приходятся все обнаруженные структурные разрывы: они порождены сменой не режима извлечения, а самого охвата по времени.

Таблица 1. Доля вакансий, у которых месяц записи совпадает с месяцем появления

Месяц записи	2025-02	2025-03	2025-04	2025-05	2025-06	2025-07	2025-08	2025-09	2025-10
Доля совпадения	0,52	0,64	0,53	0,56	0,96	0,91	0,97	0,96	0,97
Месяц записи	2025-11	2025-12	2026-01	2026-02	2026-03	2026-04	2026-05	2026-06	2026-07
Доля совпадения	0,75	0,93	0,98	0,98	0,98	0,96	0,96	0,92	0,91

Ограничение снято: ряд перестроен по восстановленной дате первого появления. Сравнение парное: один и тот же набор из 627 007 вакансий, одно окно, две датировки (Таблица 2). Само по себе исправление датировки задачу не решает, и это главный отрицательный результат работы: разрывов становится не меньше, а больше (три вместо двух), а наибольший однородный отрезок укорачивается с девяти месяцев до восьми. Передатировка снимает конкретный артефакт расписания обхода (в ноябре 2025 г. объём *сбора* подскочил с 29 538 до 49 882 вакансий, тогда как объём *публикаций* за те же месяцы сдвинулся лишь с 37 345 до 39 064), но одновременно обнажает разрывы, которые прежде были размазаны

задержкой записи. Пригодным ряд делает лишь третья процедура, поправка на предметный состав.

Таблица 2. Что даёт каждая из трёх процедур: один корпус, три способа построить ряд

Показатель	по дате записи	по дате появления	с поправкой на состав
Структурных разрывов	2	3	1
Наибольший однородный отрезок	9 мес.	8 мес.	17 мес.
Вакансий в отрезке	245 614	334 752	558 950
Навыков в панели отрезка	395	542	757
Медианная доля шума оценивания ρ	0,391	0,397	0,212
Навыков с сигналом выше шума	59,5 %	61,3 %	83,1 %

Критерий (2) на ряде по дате появления нашёл три структурных разрыва (Таблица 3); порог выбрался сам: 0,144 при наибольшем зазоре в упорядоченном наборе долей в 1,46 раза. Все три приходятся на начальный период, когда сборщик выгребал накопленный архив.

Обнаруженные разрывы имеют содержательное объяснение (Рисунок 1): на раннем этапе сборщик отслеживал единственную профессию и не сохранял тексты описаний вакансий. Отметим, что число профессий, отнесённое к месяцу публикации, почти постоянно (26–28) и потому индикатором не является: вакансии, опубликованные в начале периода, собирались позже, когда отслеживались уже все профессии.

Таблица 3. Панельный критерий до и после контроля целостности привязки навыков

Переход между месяцами	без контроля	с контролем	Вердикт
2025-02 → 2025-03	0,978	0,059	не разрыв
2025-03 → 2025-04	0,577	0,071	не разрыв
2025-04 → 2025-05	0,990	0,061	не разрыв
2025-09 → 2025-10	0,044	0,173	разрыв
2025-10 → 2025-11	0,090	0,283	разрыв
2026-06 → 2026-07	–	0,224	разрыв
остальные переходы	0,017...0,121	0,016...0,121	однородно

Отдельно установлен дефект привязки данных: в феврале–апреле 2025 г. навыки нескольких вакансий записывались на одну, вследствие чего медианное число различных навыков на вакансию достигало 132 при 6 в мае и максимуме 773,

тогда как поле источника, из которого они извлекаются, такого количества не вмещает. Доля вакансий более чем с сорока навыками составляла в эти месяцы 2–16 % против 0,1 % в мае и нуля в дальнейшем; порог отсечения применён единообразно ко всему ряду.

Третий столбец Таблицы 2 совмещает обе процедуры: ряд там построен по дате появления *и* со стандартизацией по составу, поэтому измерен вклад поправки поверх передатировки, а не поправки вместо неё. Именно такой порядок и предлагается, поскольку по отдельности передатировка отрезка не удлиняет. Тем самым непрерывный ряд за весь период наблюдения корректным прокси рыночного спроса не является. Наибольший однородный отрезок составил **17 месяцев** (февраль 2025 г. – июнь 2026 г., 757 навыков в панели, 558 950 вакансий). Прочих отрезков длиной не менее трёх месяцев не образуется: все разрывы сосредоточены в начале периода наблюдения, и после них ряд однороден до конца.

Отдельно проверено, устойчив ли сам критерий к объёму выборки, и обнаружено, что он чувствителен к нему с обеих сторон. Сверху: критерий опирается на значимость скачка доли, а значимость растёт с числом наблюдений, поэтому на синтетическом корпусе *без* разрывов их находится от нуля при 500 вакансиях в месяц до тринадцати при 40 000. Снизу: на синтетических панелях с заданным разрывом при месячной выборке не более 200 вакансий разрыв кратности 1,6 неотличим от шума ни при каком пороге, поскольку доли значимых расхождений на панелях с разрывом и без него перекрываются. Критерий поэтому применим в диапазоне примерно от 500 вакансий в месяц, а ниже этой границы отсутствие обнаруженных разрывов следует приводить как неразрешимый результат, а не как отрицательный.

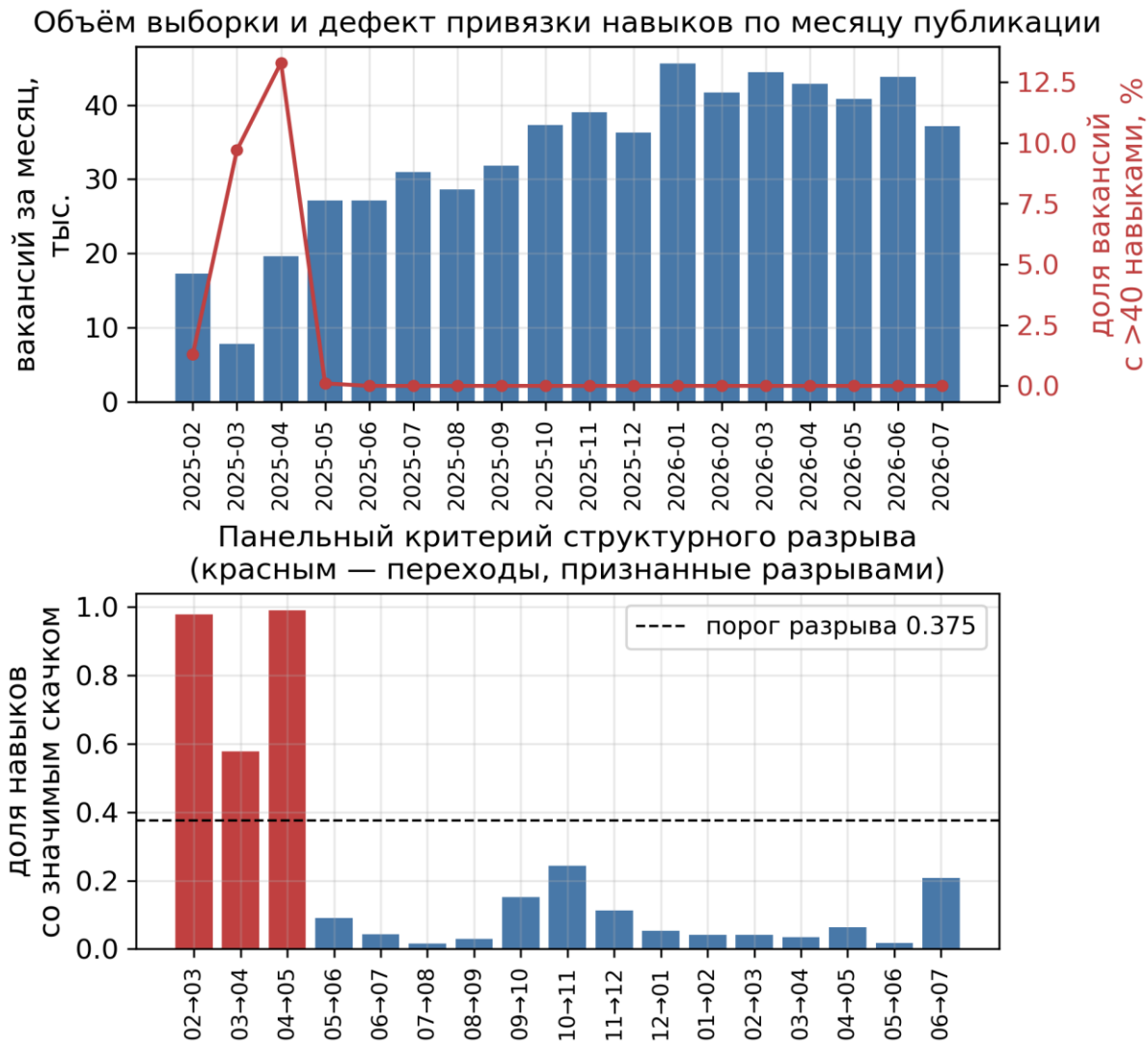


Рисунок 1. Объём выборки по месяцу публикации, доля вакансий с неправдоподобным числом навыков и результаты панельного критерия структурного разрыва

Исправленный вариант, сравнивающий скачок с полной нулевой изменчивостью $1/N_1 + 1/N_2 + (\sigma\sqrt{p/(1-p)})^2$, где рыночная составляющая с ростом выборки не убывает, выделяет в ряду два разрыва вместо трёх. Надёжно установлено лишь то, что разрывов меньше четырёх и все они приходятся на начало наблюдения: точное число зависит от выбора критерия и не выходит за пределы его разрешающей способности.

Разложение вариации (3) показывает, где именно теряется пригодность ряда (Рисунок 2). На сырых рядах медианная доля шума оценивания составляет 0,397 при датировке по дате появления и 0,391 при датировке по дате записи: сама по себе передатировка эту долю не меняет. После поправки на состав она падает до 0,212, а доля навыков, у которых содержательная динамика превышает шум оценивания,

растёт с 61,3 до 83,1 %. Причина в том, что колебания доли, порождённые сменой состава собираемых вакансий, входят в вариацию ряда наравне с рыночными и учитываются как динамика, хотя динамикой рынка не являются.

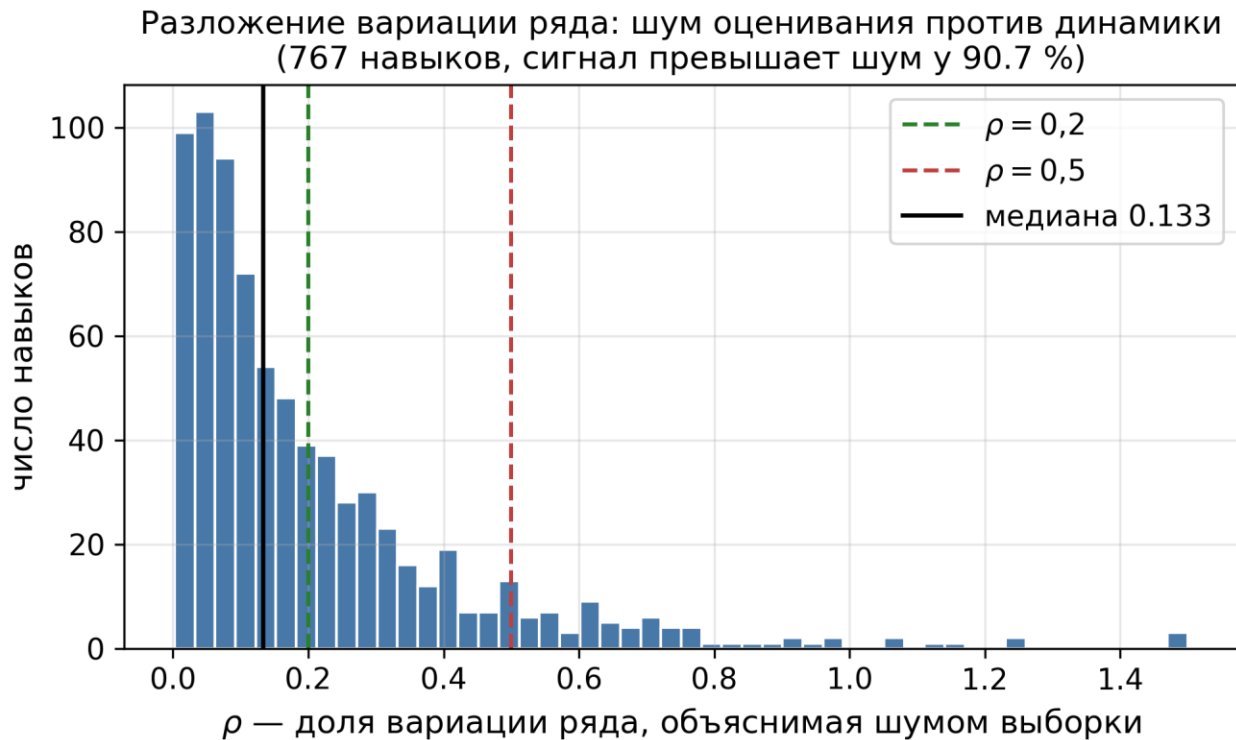


Рисунок 2. Распределение доли вариации ряда, объяснимой шумом оценивания

Внутри формально однородного отрезка обнаружен эффект, которого панельный критерий не улавливает по построению. Ранги долей наиболее частотных навыков осенью 2025 г. переставляются согласованно: управленческие навыки поднимаются, технологические опускаются. Измерение показывает, что это смена состава собираемых вакансий, а не рынка: доли двух предметных корзин расходятся в противоположных направлениях вслед за составом поисковых запросов сборщика (Рисунок 3, Таблица 4).

Сдвиг затрагивает предметные группы целиком, а не отдельные навыки. По сопоставлению первых пяти месяцев отрезка с последними пятью наибольший прирост доли получили «управление командой» (+4,87 процентного пункта), «организаторские навыки» (+4,31) и «руководство коллективом» (+4,02), а наибольшее снижение получили javascript (−1,91), sql (−1,87) и git (−1,70).

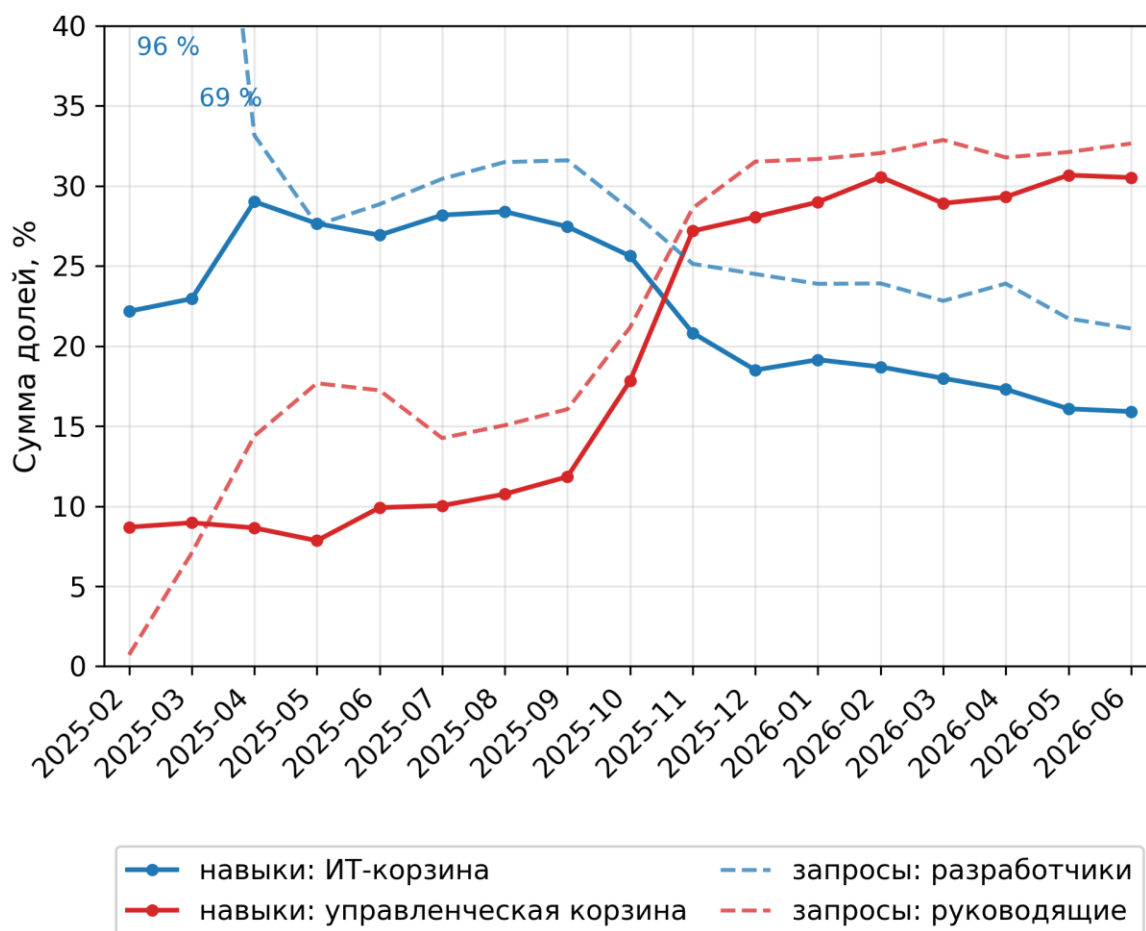


Рисунок 3. Суммарные доли предметных корзин навыков и доли поисковых запросов сборщика внутри однородного отрезка 2025-02...2026-06. Ось ординат обрезана на 40 %: в феврале и марте 2025 г. доля запросов на разработку составляла 96 и 69 %, эти значения подписаны

Таблица 4. Дрейф состава корпуса внутри однородного отрезка

Показатель	2025-05	2025-09	2026-01	2026-06
Корзина ИТ-навыков (sql, git, python, javascript, postgresql, linux, docker, java), %	27,7	27,5	19,1	15,9
Корзина управленческих навыков (управление командой, организаторские навыки, руководство коллективом и др.), %	7,8	11,8	29,0	30,5
Доля запросов сборщика руководящего профиля, %	17,7	16,1	31,7	32,7
Доля запросов сборщика на разработку и тестирование, %	27,6	31,6	23,9	21,1

Одновременное падение всех технологических навыков при одновременном росте всех управленческих объясняется составом поисковых запросов сборщика: доля запросов руководящего профиля выросла с 13,1 до 32,3 %, доля запросов на разработку снизилась с 44,3 до 22,7 %, а отдельные запросы изменились кратно («руководитель разработки» с 4,3 до 16,1 %, «frontend разработчик» с 3,9 до 0,6 %). Связь с долями навыков близка к функциональной: по четырнадцати месяцам, начиная с мая 2025 г., когда сборщик уже обходил все профессии, коэффициент корреляции Пирсона равен +0,982 между долей руководящих запросов и суммарной долей управленческой корзины, +0,960 между долей запросов на разработку и суммарной долей ИТ-корзины и –0,982 между долей руководящих запросов и долей ИТ-навыков. По всем семнадцати месяцам первая и третья связи сохраняются (+0,930 и –0,720), а вторая ослабевает до +0,175: в феврале и марте 2025 г. на один запрос приходилось от 69 до 96 % выборки, и доля запросов там состава корпуса не описывает.

Панельный критерий этого сдвига не обнаруживает по построению: он сравнивает соседние месяцы, тогда как сдвиг растянут на несколько. Поправка на состав решает задачу прямо. Доля дисперсии панели логитов, лежащая в подпространстве факторов состава, составляет $R^2 = 0,703$ при медиане суррогатов 0,514 и $p = 0,006$: дрейф состава установлен как факт, а не как предположение. После помесечной стандартизации по классам названий вакансий разрыв 2025-10 → 2025-11 исчезает по обоим критериям, а разрыв 2026-06 → 2026-07 остаётся даже при 249 стратах пересечения «запрос × класс»: внутри одного класса и внутри тех же запросов конструкторские навыки падают вдвое, а монтажно-телекоммуникационные вдвое растут. Первый разрыв, таким образом, был сменой состава сбора, а второй сменой содержания рынка, и различить их удаётся только поправкой.

Обсуждение и выводы

Три процедуры проверки, предложенные в работе, отвечают на разные вопросы и не заменяют друг друга. Панельный критерий отвечает, менялся ли режим сбора скачкообразно; разложение вариации показывает, какая часть наблюдаемых колебаний вообще может нести содержательную информацию; контроль корзин показывает, не сместился ли предметный состав корпуса постепенно. Ряд, прошедший все три, пригоден к дальнейшему анализу; ряд, не прошедший хотя бы одну, содержит артефакты, неотличимые от рыночной динамики средствами последующих моделей.

Практическое следствие для систем рассматриваемого класса состоит в том, что датировка данных не является техническим вопросом. Дата записи вакансии сборщиком и дата её появления у источника у большинства вакансий совпадают, но распределение задержки скошено вправо: 95-й перцентиль составляет 28 дней. Этого достаточно, чтобы расписание обхода превратилось в наблюдаемый

«структурный сдвиг рынка»: в ноябре 2025 г. объём сбора подскочил с 29 538 до 49 882 вакансий, тогда как число публикаций за те же месяцы изменилось с 37 345 до 39 064. Восстановление даты появления осуществимо и для уже накопленного корпуса: интерфейс источника отдал её для 89,4 % сохранённых идентификаторов независимо от возраста вакансии.

Ограничения работы следующие. Записи без восстановленной даты (9,5 % корпуса) исключены, а смещение от такого исключения не измерено, хотя доступность даты не убывает с возрастом вакансии. Границы применимости панельного критерия установлены на синтетических панелях с одним заданным разрывом; поведение критерия при нескольких близко расположенных разрывах не проверялось. Состав предметных корзин задан вручную; выбор корзин влияет на чувствительность контроля дрейфа, и автоматическое их построение составляет предмет отдельной работы. Наконец, все проверки выполнены на данных одного источника и одного сегмента рынка труда, поэтому конкретные пороги переносить на другие корпуса без повторной калибровки нельзя, тогда как сами процедуры переносятся.

Результаты означают, что оценка перспективности технологий в системах поддержки принятия решений в образовании должна начинаться не с выбора прогнозной модели, а с проверки исходного ряда. Разрывы, обнаруженные в исследованном корпусе, затрагивали до 100 % навыков панели, а дрейф состава изменил суммарную долю управленческой корзины в 3,9 раза, а ИТ-корзины в 1,8 раза за семнадцать месяцев. Ни одно из этих изменений не является рыночным событием, и любая модель, обученная на таком ряду, воспроизведёт их как сигнал.

Список использованных источников и литературы

1. Затонский А. В., Шаклеина П. А., Красовский Р. В. Современные методы сбора и анализа требований работодателей для подготовки ИТ-специалистов // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. 2025. Т. 25, № 1. С. 69–82. DOI: [10.14529/ctcr250106](https://doi.org/10.14529/ctcr250106).
2. Яруллин Д. В. Информационная система оценки соответствия образовательных программ требованиям рынка труда на основе нечетких множеств и денотатного анализа: дис. ... канд. техн. наук. Пермь: Пермский национальный исследовательский политехнический университет, 2024. 180 с.
3. Senger E. и др. Deep Learning-based Computational Job Market Analysis: A Survey on Skill Extraction and Classification from Job Postings // Proceedings of the First Workshop on Natural Language Processing for Human Resources (NLP4HR 2024). 2024. С. 1–15. DOI: [10.18653/v1/2024.nlp4hr-1.1](https://doi.org/10.18653/v1/2024.nlp4hr-1.1).
4. Colombo E., Marcato A. Skill demand and labour market concentration: evidence from Italian vacancies // International Journal of Manpower. 2023. Т. 44, № 9. С. 156–198. DOI: [10.1108/IJM-04-2023-0181](https://doi.org/10.1108/IJM-04-2023-0181).

5. Ojan M.-P., Lara-Navarra P., Sanz-Martos S. Bridging Communication Studies and Employability: ESCO-Based Curriculum Mapping and Job-Vacancy Skill Signals // Education Sciences. 2026. Т. 16, № 4. С. 606. DOI: [10.3390/educsci16040606](https://doi.org/10.3390/educsci16040606).

6. Alibasic A. и др. Evaluation of the trends in jobs and skill-sets using data analytics: a case study // Journal of Big Data. 2022. Т. 9, № 1. С. 32. DOI: [10.1186/s40537-022-00576-5](https://doi.org/10.1186/s40537-022-00576-5).

7. Garcia de Macedo M. M. и др. Practical Skills Demand Forecasting via Representation Learning of Temporal Dynamics // Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22). New York: ACM, 2022. С. 285–294. DOI: [10.1145/3514094.3534183](https://doi.org/10.1145/3514094.3534183).

8. Красовский Р. В. Предиктивная аналитика ИТ-компетенций: выявление «восходящих звезд» и категоризация трендов рынка труда на основе 400 000 вакансий // Молодежная наука в развитии регионов: материалы Всерос. науч.-практ. конф. студентов и молодых ученых (Березники, 22 апреля 2026 г.). Пермь: Изд-во Перм. нац. исслед. политехн. ун-та, 2026. С. 204–211.

9. Ditzen J., Karavias Y., Westerlund J. Testing and estimating structural breaks in time series and panel data in Stata // The Stata Journal. 2025. Т. 25, № 3. С. 526–560. DOI: [10.1177/1536867X251365449](https://doi.org/10.1177/1536867X251365449).

10. Затонский А. В., Янченко Т. В. Метод управления развитием социального ресурса региона на основе регрессионно-дифференциального моделирования // Управление большими системами: сборник трудов. 2015. № 54. С. 86–113.

11. Затонский А. В., Сиротина Н. А., Янченко Т. В. Об аппроксимации факторов дифференциальной модели социально-экономической системы // Современные исследования социальных проблем. 2012. № 11. С. 6.

12. Тугашова Л. Г., Затонский А. В. Моделирование и прогнозирование производства нефтепродуктов с учетом сезонности на основе авторегрессионных моделей // Известия Томского политехнического университета. Инжиниринг георесурсов. 2020. Т. 331, № 5. С. 109–119.

13. Красовский Р. В. Применение бота в мессенджере для обеспечения наблюдаемости, надежности и управляемости системы анализа компетенций // Материалы XVII Международной интернет-конференции молодых ученых, аспирантов, студентов. Пермь: Изд-во Перм. нац. исслед. политехн. ун-та, 2026. С. 145–148.

List of references

1. Zatonsky A.V., Shakleina P.A., Krasovsky R.V. Modern methods of collecting and analyzing employer requirements for training IT specialists. Bulletin of the South Ural State University. Series: Computer Technologies, Automatic Control, Radioelectronics. 2025. Vol. 25, No. 1. Pp. 69–82. DOI: [10.14529/ctcr250106](https://doi.org/10.14529/ctcr250106).

2. Yarullin D.V. Information system for assessing the compliance of educational programs with labor market requirements based on fuzzy sets and denotative analysis:

Cand. Tech. Sci. dissertation. Perm: Perm National Research Polytechnic University, 2024. 180 p.

3. Senger E. et al. Deep Learning-based Computational Job Market Analysis: A Survey on Skill Extraction and Classification from Job Postings. *Proceedings of the First Workshop on Natural Language Processing for Human Resources (NLP4HR 2024)*. 2024. Pp. 1–15. DOI: 10.18653/v1/2024.nlp4hr-1.1.

4. Colombo E., Marcato A. Skill demand and labour market concentration: evidence from Italian vacancies. *International Journal of Manpower*. 2023. Vol. 44, No. 9. Pp. 156–198. DOI: 10.1108/IJM-04-2023-0181.

5. Ojan M.-P., Lara-Navarra P., Sanz-Martos S. Bridging Communication Studies and Employability: ESCO-Based Curriculum Mapping and Job-Vacancy Skill Signals. *Education Sciences*. 2026. Vol. 16, No. 4. Pp. 606. DOI: 10.3390/educsci16040606.

6. Alibasic A. et al. Evaluation of the trends in jobs and skill-sets using data analytics: a case study. *Journal of Big Data*. 2022. Vol. 9, No. 1. Pp. 32. DOI: 10.1186/s40537-022-00576-5.

7. Garcia de Macedo M.M. et al. Practical Skills Demand Forecasting via Representation Learning of Temporal Dynamics. *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22)*. New York: ACM, 2022. Pp. 285–294. DOI: 10.1145/3514094.3534183.

8. Krasovsky R.V. Predictive analytics of IT competencies: identifying "rising stars" and categorizing labor market trends based on 400,000 vacancies. *Youth Science in the Development of Regions: Proceedings of the All-Russian Scientific and Practical Conference of Students and Young Scientists (Berezniki, April 22, 2026)*. Perm: Perm National Research Polytechnic University Publishing House, 2026. Pp. 204–211.

9. Ditzen J., Karavias Y., Westerlund J. Testing and estimating structural breaks in time series and panel data in Stata. *The Stata Journal*. 2025. Vol. 25, No. 3. Pp. 526–560. DOI: 10.1177/1536867X251365449.

10. Zatonsky A.V., Yanchenko T.V. Method of managing the development of the social resource of a region based on regression-differential modeling. *Large-Scale Systems Control: Collected Papers*. 2015. No. 54. Pp. 86–113.

11. Zatonsky A.V., Sirotina N.A., Yanchenko T.V. On the approximation of factors of a differential model of a socio-economic system. *Modern Studies of Social Problems*. 2012. No. 11. Pp. 6.

12. Tugashova L.G., Zatonsky A.V. Modeling and forecasting of petroleum product production taking into account seasonality based on autoregressive models. *Bulletin of the Tomsk Polytechnic University. Engineering of Geo-Resources*. 2020. Vol. 331, No. 5. Pp. 109–119.

13. Krasovsky R.V. Application of a messenger bot to ensure observability, reliability and manageability of a competency analysis system. *Proceedings of the XVII International Internet Conference of Young Scientists, Postgraduates, and Students*. Perm: Perm National Research Polytechnic University Publishing House, 2026. Pp. 145–148.