

УДК 004

О СБОРЕ ДАННЫХ ДЛЯ АНАЛИЗА РЫНКА ВАКАНСИЙ СПЕЦИАЛИСТОВ ИТ-ОТРАСЛИ

В. А. ЛИВИНСКАЯ, Е. И. ГАЙЧУКОВ

Белорусско-Российский университет

Могилев, Беларусь

Рынок труда, наряду с показателями официальной и реальной безработицы, может считаться индикатором состояния экономики любого государства. Резкое увеличение числа вакансий в той или иной области может являться поводом для образовательных учреждений корректировать учебные планы и даже открывать новые специальности. И наоборот, проблемы с распределением выпускников на производственную практику и на постоянное место работы могут свидетельствовать о недостаточно качественной подготовке специалистов. Поэтому организация регулярного мониторинга рынка вакансий и резюме, анализ динамики их соотношений является важной и актуальной задачей.

Современные информационные технологии позволяют организовать мониторинг рынка вакансий почти в реальном времени, путем сбора информации с помощью парсинга сайтов-агрегаторов с объявлениями о размещаемых вакансиях. Обычно в объявлениях работодатели указывают определенные компетенции (так называемый *skills*), которым должен соответствовать соискатель, а также опыт работы в соответствующей области и условия труда, которые обязуется обеспечить наниматель. Информация о гарантируемой заработной плате указывается довольно редко в открытом доступе.

Основными целями данного исследования являются сбор и анализ информации о вакансиях специалистов ИТ-отрасли с сайта HeadHunter (hh.ru) – одного из популярных ресурсов рунета.

Для реализации поставленной цели на языке Python была написана программа-парсер, производившая сбор данных с сайта. Сервис HeadHunter предоставляет специальный API (*application programming interface*), с помощью которого становится возможным получать информацию о соискателях и компаниях, а также использовать функциональность HeadHunter для своего сайта или приложения.

Всё API работает по протоколу HTTPS. Данные, получаемые в ответах сервиса, доступны только в формате JSON (*JavaScript Object Notation*). На базовый Uniform Resource Locator (URL) отправляются запросы приложений, использующих функциональность сервиса HeadHunter, чтобы получить ответ в виде данных, которые представлены в формате JSON. Также в запросе могут передаваться дополнительные параметры, задающие правила и критерии поиска и фильтрации.

Стоит отметить, что имеется возможность задавать такие параметры поиска, как нижняя и верхняя границы вилки оклада (параметры *salary.from* и *salary.to* соответственно); департамент, от имени которого размещается вакансия (если

данная возможность доступна для компании) (параметр *department*); информация о ключевых навыках, заявленных в вакансии (параметр *key_skills*); требуемый опыт работы (параметр *experience*).

С парсингом резюме на HeadHunters дело обстоит несколько сложнее, т. к. для этого необходима платная подписка. Поэтому приходится использовать сторонние модули языка Python, например, *re*, *requests*, *bs4* для того, чтобы получить данные и произвести дальнейший разбор HTML.

Процесс накопления данных для полноценного анализа требует много времени, т. к. каждый день на сайте появляются более 500 новых объявлений (в зависимости о выбранного количества регионов). Представляет интерес получение ответов на целый ряд вопросов: какие языки программирования пользуются наибольшим спросом; в какие предметные области требуются специалисты; как различается спрос по регионам; есть ли тенденции, связанные с пандемией.

Для специалиста IT-отрасли ключевыми навыками являются язык программирования и знание предметной области. За 10 дней сбора информации в январе 2021 г. (с 10/01 по 20/01) 82 % вакансий hh.ru в этой категории были размещены вакансии в таких сферах деятельности компаний, как системная интеграция, web-инженерия, разработка систем управления предприятием (ERP), CRM системы, создание банковского ПО.

Анализ доли различных языков программирования в общем количестве вакансий, собранных за анализируемый период, представлен на рис. 1. По имеющейся выборке наиболее распространенным оказался язык Java, хотя по международным исследованиям (TOIBE INDEX for January 2021) язык Python явился лидером 2020 г.

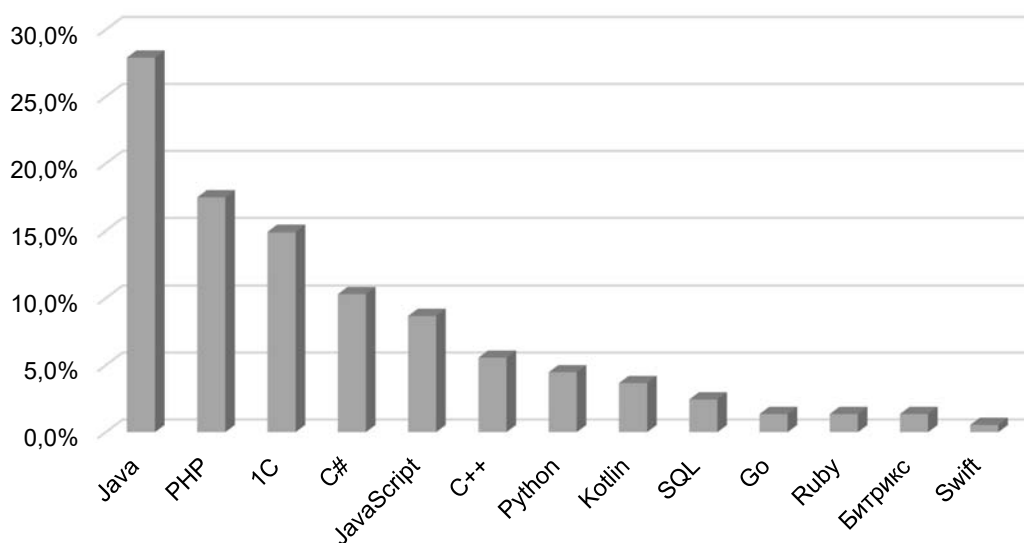


Рис. 1. Структура потребности рынка IT в языках программирования в январе 2021 г.