

УДК 004.912

ПОДХОД К ПРЕДОСТАВЛЕНИЮ АКТУАЛЬНЫХ ДАННЫХ ДЛЯ ПРОЕКТИРОВАНИЯ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ

Холошня В.Д. (Университет ИТМО)

Научный руководитель – к.т.н. Болдырева Е.А.
(Университет ИТМО)

Данная работа посвящена вопросу создания информационной системы для получения данных, необходимых при разработке образовательных программ. Разрабатываемая система реализована в виде программного интерфейса, предоставляющего данные в формализованном виде, что позволяет использовать ее не только как отдельный источник данных, но также, например, встраивать в информационные системы университетов. Новизна результатов заключается в том, что, в отличие от существующих информационных систем, разрабатываемая система объединяет данные из большого количества источников, загружает, обрабатывает и предоставляет их в формализованном виде. Полученные данные позволяют формировать актуальный контент учебных задач практикумов и подстраивать реализацию образовательного процесса под потребности рынка труда, что повышает актуальность навыков, получаемых при прохождении образовательных программ.

Введение. В настоящее время, в период высокой конкуренции на рынке труда основной задачей образовательных учреждений является подготовка конкурентоспособных на рынке труда выпускников. Постоянная переработка образовательных программ, с целью повышения актуальности получаемых студентами навыков становится не просто требованием нормативных актов, а необходимым условием обеспечения качественной подготовки выпускников. На уровне федеральных государственных стандартов предпринимаются меры по усилению влияния работодателей на процесс обучения в вузах: привлечение их к разработке образовательных стандартов, создание базовых кафедр на предприятиях, обеспечение практической направленности обучения.

Подобные меры являются особенно актуальными, когда речь идет о подготовке специалистов в наукоемких областях. Особенностью наукоемких областей является постоянное обновление технологической и инструментальной базы, и, как следствие, постоянное изменение требований работодателей.

Данное исследование посвящено вопросу создания информационной системы для получения актуальных данных, необходимых при разработке образовательных программ. Полученные данные позволяют формировать актуальный контент учебных задач практикумов и подстраивать реализацию образовательного процесса под потребности рынка труда, что повышает актуальность навыков, получаемых при прохождении образовательных программ.

В настоящее время не существует централизованной информационной системы, предоставляющей подобные актуальные данные, необходимые при разработке образовательных программ. В данном исследовании была проведена разработка и частичное внедрение в реальный образовательный процесс информационной системы анализа профессиональной и образовательной областей. Тренды рынка труда обновляются быстрее, чем студент успевает закончить обучение. Поэтому разрабатываемая система — это попытка не только “успеть” за рынком, но и спрогнозировать его развитие, чтобы действовать на опережение.

Основная часть. Разрабатываемая система предлагает подход для решения поставленной задачи получения актуальных данных для разработки образовательных стандартов. Система разделена на модули, связанные с внешним представлением, базой данных, а также друг с другом.

Программная реализация системы описана на языке программирования Python. Программный интерфейс информационной системы разработан с помощью библиотеки

Django. Текстовые данные обрабатываются с использованием библиотеки natural language toolkit, кластеризация данных осуществляется с помощью библиотеки scikit-learn, для хранения промежуточных данных в виде матриц и массивов используется библиотека pandas, а также стандартные библиотеки для обработки и получения данных, такие, как re и request. Система хранения систематизированных данных реализована с помощью реляционной базы данных PostgreSQL.

Данные предоставляются на внешний слой программного интерфейса в формате JSON, что позволяет использовать их не только для отправки на пользовательские интерфейсы, но и любым другим приложениям, например, информационным системам университетов. Предлагаемая система предоставляет следующие формализованные данные: информация об образовательных и профессиональных стандартах, требования работодателей и ключевые навыки соискателей в соответствии со стандартами, мероприятия, связанные с профессиональной сферой. Образовательные и профессиональные стандарты загружаются из открытых государственных источников. Данные о требованиях работодателей и ключевых навыках соискателя извлекаются из вакансий, получаемых с помощью программного интерфейса hh.ru в соответствии со стандартами. Данные обрабатываются с помощью технологий обработки естественного языка и кластерного анализа. Текстовые данные проходят нормализацию с помощью стеммера Портера, создается матрица весов TF-IDF и проводится кластерный анализ с помощью алгоритма DBSCAN для того, чтобы определить наиболее необходимые и актуальные навыки в той или иной профессиональной сфере. Дополнительной функцией является система подбора мероприятий, связанных с профессиональной областью.

Выводы. Предложенный подход и реализованные инструментальные средства были использованы для формирования списка обязанностей, требований и ключевых навыков по профилю “Встраиваемые и киберфизические системы”. В качестве базового набора данных для формирования фрагмента, посвященного программному обеспечению встраиваемых систем, были взяты и проанализированы требования и ключевые навыки для связанных с программным обеспечением встраиваемых систем профессиональных стандартов 6 разных специализаций - всего 1340 вакансии и 3610 требований.

В статье рассмотрен вопрос создания информационной системы для получения актуальных данных, необходимых при разработке образовательных программ. Предоставляемые системой данные позволяют повысить актуальность получаемых, после прохождения образовательных программ, навыков, а также ускорить процесс их разработки. Полученные данные показывают взаимосвязь требований, обязанностей и ключевых навыков с профессиональными стандартами, рабочими функциями и специализациями. Результаты анализа полученных данных позволяют формировать актуальный контент учебных задач практикума и подстраивать реализацию образовательного процесса под потребности рынка труда.

В качестве направления дальнейших исследований решено рассмотреть создание системы для автоматического поиска и формирования учебных задач.