

**Большие данные для исследования рынка труда:  
Анализ онлайн-вакансий**

Виртуальный курс обучения

Полная программа: 8, 10 и 15 июня 2021 г.

Многоязычный: английский-французский-русский

Регистрация: [ссылка](#)

**Программа:**

**День 1: 08.06.2021, 09:00–14:30 ЦЕВ (GMT+2)**

- Исследование рынка труда – на основе веб-данных: от онлайн-вакансий до данных и анализа навыков, профессий, секторов. Новые данные, новая аналитика. Обзор, методология, области применения.
- Изучение конкретного примера

**День 2: 10.06.2021, 09:00–14:30 ЦЕВ (GMT+2)**

- Исследование рынка труда – на основе веб-данных: ориентация на приеме, обработке и классификации данных с помощью процессов на базе искусственного интеллекта.
- Изучение конкретного примера

**День 3: 15.06.2021, 09:00–14:15 ЦЕВ (GMT+2)**

- Исследование рынка труда – на основе веб-данных: примеры использования, практические упражнения и извлеченные уроки.



Сотрудничество ЕФО с Burning Glass Europe




[Link to access](#)  
[Link of the Webpage](#)

## Проект ЕФО Большие данные в исследовании рынка труда

В рамках запущенного во второй половине 2018 года инновационного проекта ЕФО сначала было разработано краткое методическое руководство «[Большие данные в исследовании рынка труда: вводное руководство](#)» (2019). [Курс подготовки](#) для членов экспертной сети «Make it Match» Восточного партнерства был успешно завершён в ноябре 2019 года. В 2020 году была создана комплексная система данных для анализа онлайн-вакансий с информационными панелями, доступными для: [Туниса](#) и [Украины](#). Информационные панели дополняются данными об онлайн-вакансиях, собранными до 31 марта 2021 года. Система данных ориентирована на спрос (онлайн-вакансии).

Проект ЕФО Большие данные в исследовании рынка труда продолжит реализовываться в 2021 году с включением новой страны и продолжением сбора, классификации и анализа данных об онлайн-вакансиях для Туниса и Украины. С опорой на эту растущую базу данных онлайн-вакансий можно будет проанализировать многие новые вопросы и запросы, согласно парадигме «пусть говорят данные». Обширные массивы данных станут источником новых идей и выводов, и будет разработана «Зеленая панель».

**Трехдневное путешествие по этой виртуальной обучающей программе призвано дать ответы на эти вопросы.**

- Однако, что включает в себя аналитика больших данных для информационных систем рынка труда?
- Каким образом следует исследовать и извлекать выгоду из больших объемов онлайн-вакансий для информационных систем рынка труда в режиме реального времени?
- Какие аспекты и проблемы, связанные с навыками, квалификацией и вопросами динамики рынка труда, можно проанализировать с помощью больших данных/исследования рынка труда в реальном времени?

- Какие требования предъявляются к использованию этих новых источников данных и систем странами-партнерами ЕФО?
- Каким образом можно использовать процессы исследования рынка труда в реальном времени, основанные на классификации больших данных и искусственного интеллекта, и дополнить установленную национальную статистику?

**Участники:** представители исследовательских и аналитических отделов министерств (образования, обучения и труда) и специализированных исследовательских центров из всех стран-партнеров ЕФО. Участие во всех мероприятиях, запланированных в рамках трехдневной программы, рекомендуется (обязательно) для всех зарегистрированных участников.

Трехдневное путешествие по этой виртуальной учебной программе займет 4,5 часа в день и включает дни для обзора и подготовки тематических блоков учебной программы.

### **Программа обучения**

**День 1: 08.06.2021: 09:00–14:30 (СЕТ) – 07:00–12:30 GMT**

<b>Время (СЕТ)</b>	<b>Сессия</b>	<b>Докладчик</b>
<b>09:00–09:30</b>	Открытие, приветственное слово. Задачи обучающего семинара. Представление участников	Эдуарда Кастель-Бранко (Eduarda Castel-Branco), ЕФО
<b>09:30–09:45</b>	Краткий обзор проекта ЕФО «Большие данные для исследования рынка труда»	Эдуарда Кастель-Бранко (Eduarda Castel-Branco)
<b>09:45–11:15</b> <b>Сессия 1</b>	<b>Большие данные для исследования рынка труда</b> Презентация системы и результаты <ul style="list-style-type: none"> <li>• Общий обзор</li> <li>• Проблемные моменты</li> <li>• Обзор потока данных</li> </ul>	Алессандро Ваккарино (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)
<b>11:15–11:45</b>	Опрос. Вопросы и ответы. Обсуждение	Эдуарда Кастель-Бранко (Eduarda Castel-Branco) Алессандро Ваккарино (Alessandro Vaccarino), Мауро Пелуччи (Mauro Pelucchi)
<b>11:45–12:00</b>	<b>Перерыв на кофе</b>	
<b>12:00–13:30</b> <b>Сессия 2</b>	<b>Большие данные для исследования рынка труда</b> Описание системы <ul style="list-style-type: none"> <li>• Получение данных</li> <li>• Предварительная обработка</li> <li>• Классификация</li> </ul>	Алессандро Ваккарино (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)
<b>13:30–13:45</b>	Вопросы и ответы. Обсуждение	Алессандро Ваккарино (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)
<b>13:45–14:30</b>	Изучение конкретного примера: Евростат – Умная статистика. Ориентирование на данные из онлайн-объявлений о вакансиях	Евростат

**День 2: 10.06.2021: 09:00–14:30 (CET) – 07:00–12:30 GMT**

Время (CET)	Сессия	Докладчик
09:00–09:15	Открытие. Впечатление от дня 1.	Эдуарда Кастель-Бранко (Euarda Castel-Branco), ЕФО
09:15–11:00 Сессия 1	<b>Большие данные для исследования рынка труда</b> Компоненты системы <ul style="list-style-type: none"> <li>• Резюме и практическое применение (Процесс получения, внесения и обработки данных для последующего их использования или хранения в базе данных)</li> <li>• Открытая кухня: сбор данных с корпоративного веб-сайта</li> <li>• Вопросы</li> </ul>	Алессандро Ваккарينو (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)
11:00–11:30	Вопросы, разъяснения. Опрос.	
11:30–12:00	<b>Перерыв на кофе</b>	
12:00–13:30 Сессия 2	<b>Большие данные для исследования рынка труда</b> Лаборатория <ul style="list-style-type: none"> <li>• Открытая кухня: создание классификационного дерева</li> </ul>	Алессандро Ваккарينو (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)
13:00–13:15	Вопросы и ответы. Обсуждение	Алессандро Ваккарينو (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)
13:45–14:30	<b>Изучение конкретного примера:</b>  Сферы применения исследований рынка труда в реальном времени: смежность навыков, продвижение по карьерной лестнице и осознанный выбор.	Докладчик: Давор Мискулин (Davor Miskulin), Burning Glass Technologies

**День 3: 15.06.2021: 09:00–14:15 (CET) – 07:00–12:15 GMT**

**Технические требования: Tableau Public, установленный на рабочем месте ([ссылка для загрузки](#))**

Время (CET)	Сессия	Докладчик
09:00–09:15	Открытие. Подведение итогов дней 1 и 2.	Эдуарда Кастель-Бранко (Euarda Castel-Branco), ЕФО
09:15–11:00 Сессия 1 (англ.)	<b>Большие данные для исследования рынка труда</b> Компоненты системы: презентация данных <ul style="list-style-type: none"> <li>• Визуализация данных</li> <li>• Открытая кухня: визуализация сходства профессий</li> </ul>	Алессандро Ваккарينو (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)
11:00–11:15	<b>Перерыв на кофе</b>	
11.15-12.00 <b>Special</b>	<b>ESCO</b> – the taxonomy of skills and occupations that supports labour and learning mobility, and links with qualifications and	Dimitrios Pikos ESCO – Directorate General

<b>session</b>	employment.	Employment Social Affairs and Inclusion – European Commission
<b>12:00–13:15</b> <b>Сессия 2</b>	<b>Большие данные для исследования рынка труда</b> <ul style="list-style-type: none"> <li>Открытая кухня: использование Tableau Public и упражнения по презентации</li> <li>Практическое занятие: анализ данных онлайн-рынка труда</li> </ul>	Алессандро Ваккарينو (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)
<b>13:15–14:15</b>	Обсуждение результатов и выводов сессии 2. Финальная обратная связь Закрытие обучающего семинара	Эдуарда Кастель-Бранко (Euarda Castel-Branco) Алессандро Ваккарينو (Alessandro Vaccarino) Мауро Пелуччи (Mauro Pelucchi)

## Общая информация

### Программа обучения: темы и результаты

День 1: 8 июня	День 2: 10 июня	День 3: 15 июня
<p><b>Название:</b> Исследование рынка труда в реальном времени: Обзор, цели, преимущества и примеры применения.</p> <p><b>Тип сессии:</b> Лекция (презентация)</p> <p><b>Изучение конкретного примера:</b> дополнительная сессия (30 мин)</p> <p>Евростат и «умная» статистика, использование веб-данных (с акцентом на онлайн-вакансии) — перспективы, объем и требования.</p> <p>Потенциальные возможности для прочих национальных агентств статистики?</p> <p>Докладчик: ЕВРОСТАТ</p>	<p><b>Название:</b> Исследование рынка труда в реальном времени: Ориентация на приеме, обработке и классификации данных</p> <p><b>Тип сессии:</b> Открытая кухня</p> <p><b>Альтернативные требования:</b> Tableau Public, Google Colab, Databricks Community Edition</p> <p><b>Изучение конкретных примеров:</b> дополнительная сессия (30 мин)</p> <p>Сферы применения исследований рынка труда в реальном времени: смежность навыков, профессии и осознанный выбор.</p> <p>Докладчик: Давор Мискулин (Davor Miskulin), Burning Glass Technologies</p>	<p><b>Название:</b> Исследование навыков: презентация результатов; примеры использования, практические упражнения и извлеченные уроки</p> <p><b>Special session:</b> ESCO – the European taxonomy of skills and occupations supporting mobility.</p> <p><b>Тип сессии:</b> Семинар</p> <p>Участники анализируют данные для более глубокого понимания того, какие идеи/информация предоставляются системой исследования рынка труда, основанной на больших массивах данных.</p> <p><b>Описание:</b> на основе практических примеров участникам предлагается провести краткий анализ данных с помощью Tableau или предпочтительного инструмента анализа.</p> <p>Извлеченные уроки и выводы в группе.</p> <p><b>Требования:</b> Tableau Public</p>
<p><b>Результат:</b> участники получают полный обзор системы исследования данных, основанной на больших данных</p>	<p><b>Результат:</b> понимание системы исследования рынка труда среди участников становится глубже, с акцентом на</p>	<p><b>Результат:</b> участники смогут:</p> <ul style="list-style-type: none"> <li>- обсудить и изучить требования к большим данным для систем исследования рынка труда;</li> </ul>

(онлайн-вакансии). От теории Открытия первоначально неизвестных зависимостей в больших базах данных — к этапам сбора и обработки данных.	функционирование ее компонентов.	<ul style="list-style-type: none"> <li>- интерпретировать и подвергать сомнению информационные панели</li> <li>- выдвигать вопросы для изучения всех возможностей больших данных для исследования рынка труда;</li> <li>- анализировать навыки и профессиональные особенности и тенденции на основе больших данных.</li> </ul>
--	----------------------------------	--

## Большие данные для исследования рынка труда

### Новый мир аналитики данных...

Анализ навыков как обычный вид деятельности недостаточен для понимания направления и степени трансформации задач, рабочих мест, навыков и квалификаций, вызванных волной движущих сил изменений, которые ускорили цифровизацию большинства процессов в наших обществах. Появились новые инструменты анализа данных для повышения уровня исследования навыков и дополнения традиционных статистических данных, опросов и административных сведений.

Об информации говорят сейчас как о новой нефти. Цифровизация процессов, услуг, бизнеса, личных и социальных взаимодействий способствует генерации растущего массива данных по всему миру. Получение знаний из больших объемов данных, передаваемых с высокой скоростью и в различных масштабах, является основной целью анализа больших данных.

### ...может применяться для информирования рынка труда.

Искусственный интеллект (ИИ) и машинное обучение не только меняют рынок труда, но и дают нам новые инструменты для анализа рабочей силы. Вакансии или объявления о вакансиях публикуются и обновляются в большом количестве через веб-сайты разного типа, размера и охвата. Исследовательские центры и государственные органы в ряде стран поставили перед собой задачу по изучению внутренней информации такого большого источника данных. Эти источники массивных данных необходимы для понимания динамики и функционирования веб-рынков труда, а также изменения выбора работодателей при найме на работу.

Аналитика больших данных может использоваться для сопоставления навыков по профессиям, для выявления устаревших навыков, для выполнения прогнозного анализа спроса на новые профессии и навыки, а также для лучшего отслеживания взаимодействия навыков — на основе детализации данных и квази в реальном времени.

Начиная с 2016 года, Европейский центр профессиональной подготовки возглавляет инновационный проект в этой области в Европейском Союзе, создав обширную систему данных, основанную на анализе и классификации миллионов онлайн-вакансий в странах-членах Европейского Союза (ЕС). На платформе [OVATE](#) результаты представлены в интерактивных информационных панелях с комбинированными переменными и различным географическим охватом. В 2021 году Европейский центр профессиональной подготовки обновил информационные панели OVATE и продолжает анализ взаимосвязи между профессией, навыками и сектором. С другой стороны, Евростат берет на себя часть инфраструктуры данных, чтобы сосредоточиться на статистике и подробных временных тенденциях, а также на территориальном расположении. Этим

новым этапом проекта руководят совместно Европейский центр профессиональной подготовки и Евростат.

### Проект ЕФО Большие данные в исследовании рынка труда

Практическая реализация проекта ЕФО началась в 2019 году с технико-экономического анализа онлайн-рынков труда Марокко и Туниса, в результате чего был составлен всеобъемлющий отчет с оценкой и рейтингом **онлайн-порталов вакансий**. Основной деятельностью в 2020 году в двух пилотных странах (Украине и Тунисе) было создание интегрированной системы сбора, обработки, классификации, анализа и визуализации.

### Схематический обзор рабочего процесса и метода больших данных для исследования рынка труда



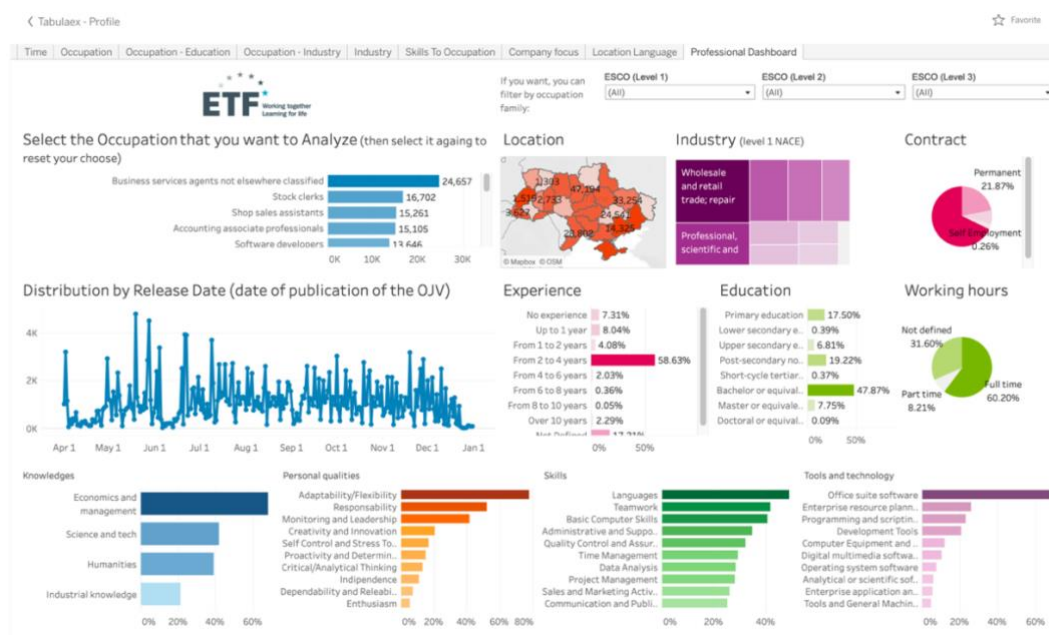
Ключевые результаты использования мероприятий ЕФО в рамках проекта Большие данные для исследования рынка труда, включая **информационные панели, обучающие программы, методическое руководство и аналитические материалы**, были опубликованы:

- Веб-страница Open Space: [Исследование рынка труда – Ориентация на большие данные](#)
- Информационные панели: [Украина](#) и [Тунис](#)
- Краткое методическое пособие (на английском языке) «[Большие данные в исследовании рынка труда: вводное руководство](#)» (2019).
- Краткое методическое пособие, 2019 ([на русском языке](#))
- Специальная [программа обучения](#) для аналитиков данных и экспертов, ноябрь 2019 г.
- Доклад: [Большие данные для исследования рынка труда, Определение контуров веб-рынка труда Украины](#), 2020.
- Доклад: [Большие данные для исследования рынка труда в Украине: аналитический отчет](#). 2020.
- [Веб-страница веб-семинара по большим данным для информационных систем рынка труда](#), 10 декабря 2020 г.

В сотрудничестве с группой по анализу данных из *Миланского университета Бикокка* и *Burning Glass Europe* (Италия) Европейский фонд обучения (ЕФО) завершил в декабре 2020 года решающую фазу своего инновационного проекта «Большие данные в исследованиях рынка труда». Сотни тысяч онлайн-вакансий, собранных в течение 8 последовательных месяцев в 2020 году (апрель-декабрь), обработанных и автоматически классифицированных в соответствии с такими международными классификациями/систематизациями, как ISCED 2011, NACE, NUTS / ISO и ESCO, предоставляют уникальную информацию о навыках, профессиональных характеристиках и динамике [тунисского](#) и [украинского](#) рынков труда. Мы заявляем об уникальности результатов в силу их детализации и характера в реальном времени. Некоторые из множества возможных углов анализа визуализируются на информационных панелях двух стран.



**Рисунок 1: Информационная панель Украины о вакансиях: многомерное представление переменных в одном снимке**



Система данных использует ESCO в качестве эталона для машинной классификации навыков, выявленных в сотнях тысяч онлайн-вакансиях. В случае сведений в отношении онлайн-вакансий Туниса мы использовали ESCO — во французской и английской версиях. Но в случае с Украиной был необходим дополнительный шаг: перевод навыков ESCO на русский и английский языки (более 4000 терминов).

Отдельное преимущество онлайн-вакансий в качестве источников информации для исследования рынка труда заключается в том, что они являются выражением определения работодателя в отношении профилей, которые им необходимы для целей бизнеса или деятельности в определенный период. Машинная обработка и классификация терминов и описаний навыков, представленных работодателями, дают понятие о навыках, требуемых в рамках размещенных онлайн-вакансий, без прямой отсылки к Европейской классификации навыков, компетенций и профессий (ESCO). Не удивительно: технологическая и цифровая трансформация работы и навыков намного быстрее, чем темпы согласования и обновления ESCO. Каковы следующие действия? Команда специалистов по анализу данных, участвовавшая в проекте, применила методы машинного обучения (например, Word2Vec) для улучшения навыков ESCO, создав соответствие между новым навыком, не входящим в ESCO, с близким (приблизительным) навыком из ESCO. Этот процесс и сопоставление, предлагаемое машинами, обсуждается и подтверждается профессионалами в данном секторе, профессии и технологии. Может ли этот метод и подход найти более широкое применение в контексте обновлений ESCO?

27.04.2021. Контактные данные: Эдуарда Кастель-Бранко (Eduarda Castel-Branco) – [ecb@etf.europa.eu](mailto:ecb@etf.europa.eu)

**Технические требования**

# Процесс доступа к Databricks

Выполните следующие действия, чтобы создать бесплатную учетную запись для доступа к Databricks *Community Edition*:

1. Перейдите по ссылке <https://databricks.com/try-databricks>
2. Скомпилируйте форму. Укажите действительный почтовый адрес, на который будет отправлена ссылка для активации

**Try Databricks**

An open and unified data analytics platform for data engineering, machine learning, and analytics

From the original creators of Apache Spark™, Delta Lake, MLflow, and Koalas

Tell us a little about yourself to get started.

* First Name: Mauro	* Last Name: Pelucchi
* Company Name: Unimb	* Work Email: mauro.pelucchi@unimb.it
* How would you describe your role?: Data Scientist	* What is your intended use case?: Education - Learning & Teaching

Phone Number:  
+

By Clicking "Sign Up", you agree to the [Privacy Policy](#)

Keep me informed with occasional updates about Databricks and Apache Spark™.

**SIGN UP**

3. Выберите «Community Edition»

**COMMUNITY EDITION**

For students and educational institutions

---

- Single cluster limited to 6GB and no worker nodes
- Basic notebooks without collaboration
- Limited to 3 max users
- Public environment to share your work

**GET STARTED**

By clicking "Get Started" for the Community Edition, you agree to the [Databricks Community Edition Terms of Service](#).

4. Проверьте свой почтовый ящик на наличие письма с ссылкой для активации. Откройте его, перейдите по ссылке для активации и следуйте инструкциям.
5. Проверьте доступ к Databricks, используя следующий URL:



<https://community.cloud.databricks.com/>

Войдите в систему, указав свою почту и предоставленный пароль

6. Убедитесь, что пользователь правильно завершил активацию, и система работает, как показано ниже:

