



NEW
PRO
LAB

СПЕЦИАЛИСТ

ПО БОЛЬШИМ ДАННЫМ 13.0

МАРАФОН ДЛИНОЮ В 12 НЕДЕЛЬ

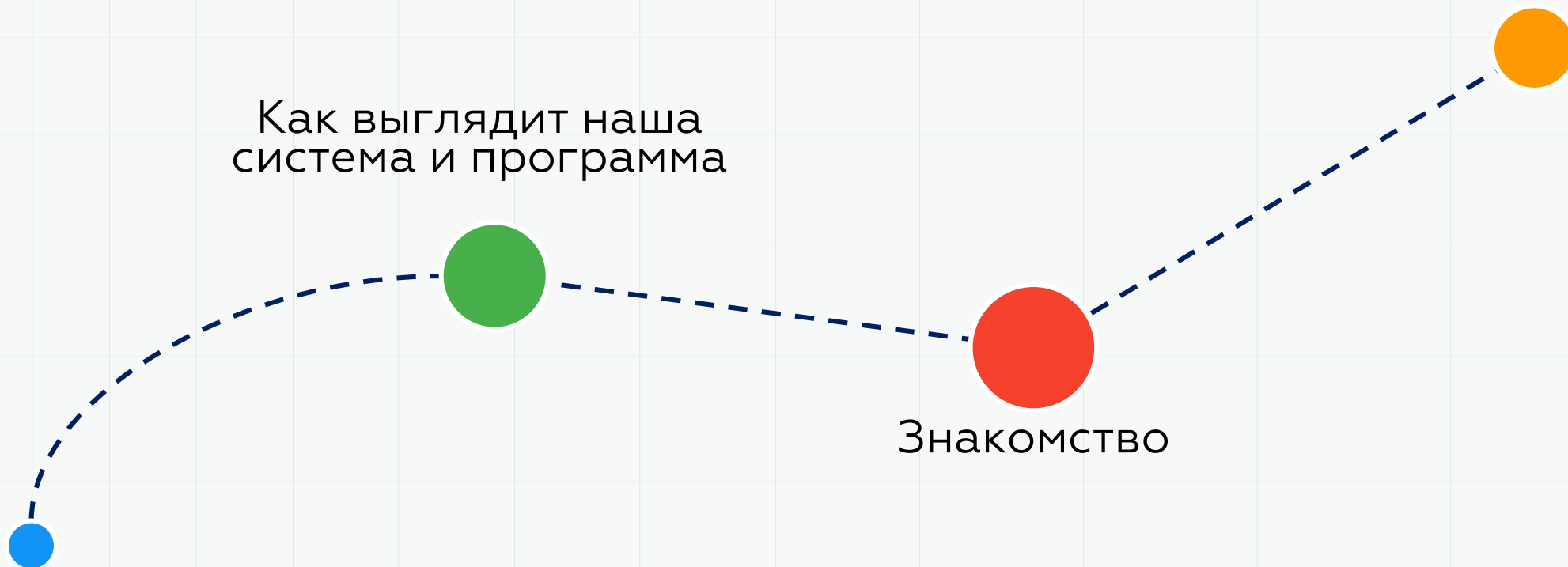
План

Кто мы такие

Как выглядит наша
система и программа

Знакомство

Введение в
большие данные





КТО МЫ?

- Занимаемся обучением работе с данными с 2015 года.
- Самая широкая линейка программ про данные на рынке.
- Используем андрагогику и hands-on подход для создания практикоориентированных программ.

История

Выпускники:

- 600+ с открытых программ
- 600 с корпоративных



Клуб **CDO**

3 Big Data

2 Big Data

2 Big Data

2 Big Data

2 Big Data

2 Big Data

1 Deep Learning

3 Deep Learning

2 Deep Learning

2 Deep Learning

2 Deep Learning

1 Data Engineer

3 for Executives

2 Data Engineer

2 for Executives

2 Data Engineer

2 for Executives

2 Data Engineer

1 for Executives

2 Scala

1 Scala



2015

2016

2017

2018

2019

2020

Линейка программ

Управление всей
цепочкой целиком

CDO

Сырые
данные

Data Engineer

Обработанные
данные

Deep Learning
Big Data Specialist

Знания

**Big Data for
Executives**

Стратегия

**Анализ данных
на Scala**

Продукт,
процесс

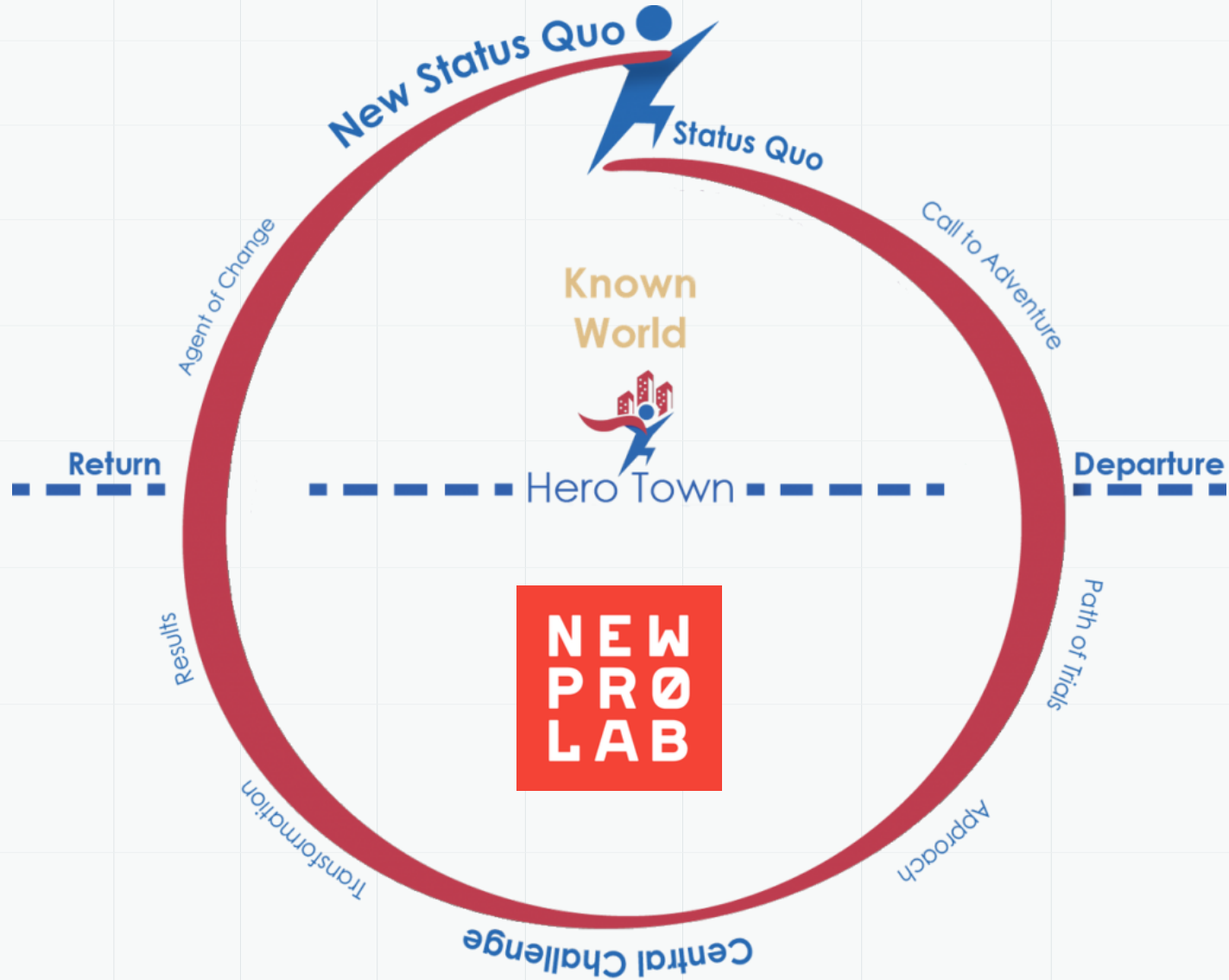
Мы - школа обучения работе с данными, но **необычная...**

Наша цель лежит **за пределами** простого покрытия тем или передачи знаний и навыков.

Hero's journey



Hero's journey



За время программы я научился:

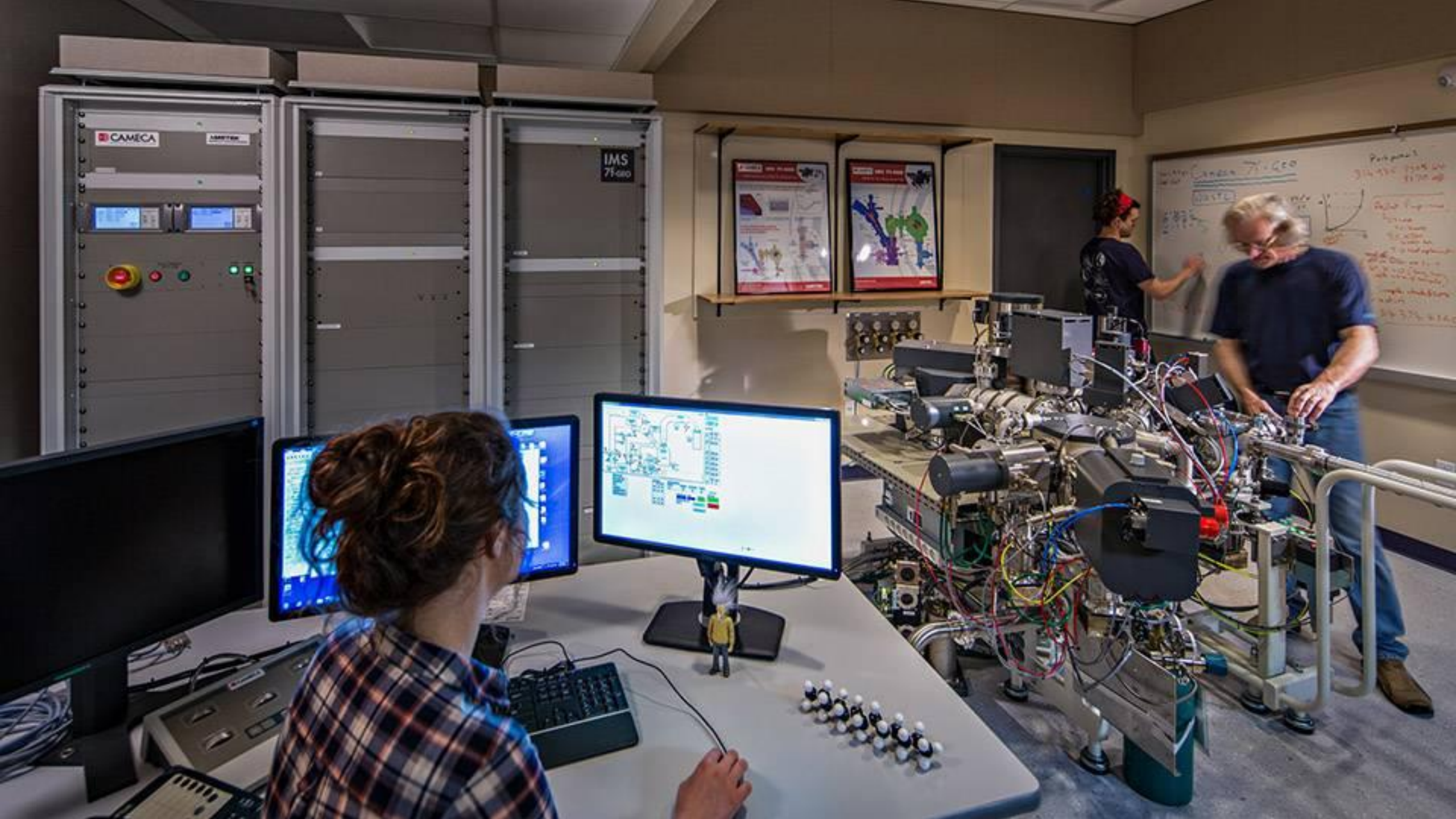
- “Терпению, усидчивости, самоконтролю, скрупулезности и веры в самого себя. Я крут, я всё могу и всё смогу.”
- “Упорству. То есть решать задачи, которые я сходу не знаю, как решить.”
- “Очень часто лабы у меня получались со 2-3 попытки. И это, как ни странно, дало больше уверенности в себе.”

КАК ВЫГЛЯДИТ НАША СИСТЕМА

Концепция

Держа в уме нашу цель, мы создали пространство – лабораторию, в которой объединили:

1. реальные задачи из индустрии,
2. современные инструменты,
3. технологии, помогающие в обучении,
4. старших товарищей (преподаватели и координатор),
5. методологию, удерживающую баланс между самостоятельностью и поддержкой.





1. Реальные задачи



Модуль 1. Построение
DMP-системы

Модуль 2. Разработка
рекомендательной
системы

Модуль 1

Лаба 1. Деплой кластера
в облаке и запуск
MapReduce

Лаба 2. Фильтрация
логов из HDFS в HBase

Лаба 3. Классификация
пользователей по логам

Лаба 4.
Прогнозирование
оттока клиентов банка

Лаба 5. Классификация
отзывов в интернете

Проект 1.
**Прогнозирование пола
и возраста по логам**

Модуль 2

Лаба 6.
Неперсонализированные
рек. системы

Лаба 7. Content-based рек.
системы

Лаба 8. Коллаборативная
фильтрация

Лаба 9. Гибридная рек.
система фильмов

Лаба 10. Рекомендации
фильмов по
телесмотрению

Проект 2. Рек. система для
e-commerce

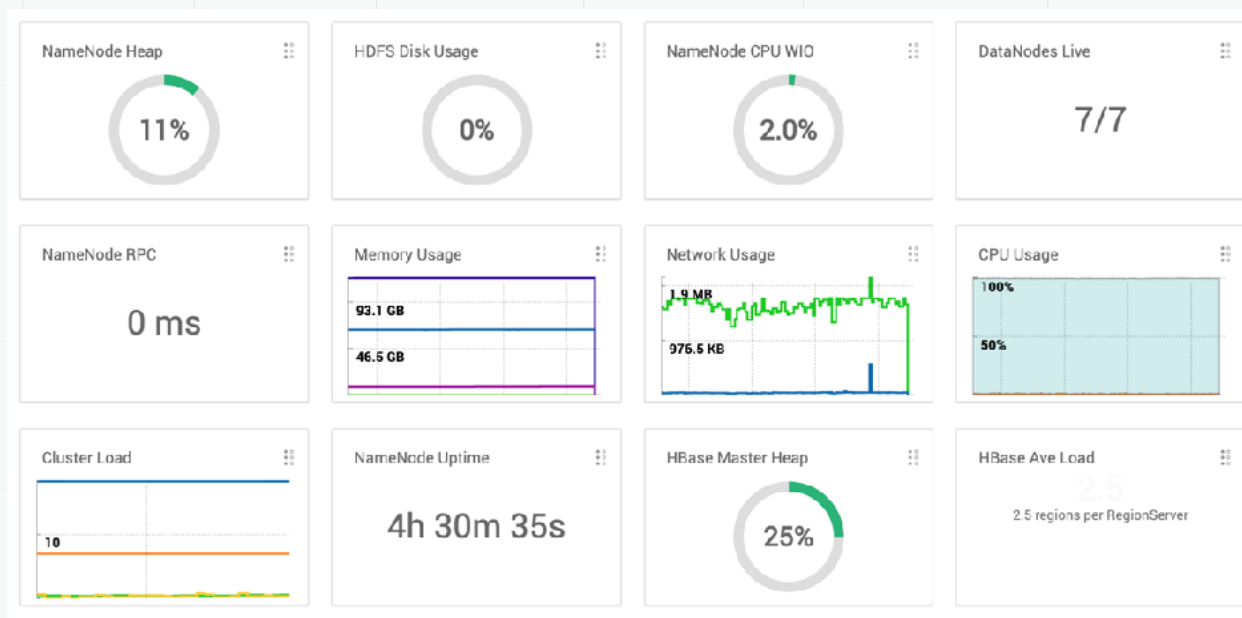
2. Инструменты



2. Инструменты

Кластер из:

- 8 нод с 16 vCPU и 80GB
- 1 нода с 32 vCPU и 256GB



3. Технологии

1. Чекеры с автоматической проверкой лаб.
2. Записи занятий.
3. Slack.
4. GitHub (**инвайт!**).
5. Google Calendar.

4. Старшие товарищи



Антон Пилипенко

Data Engineer,
Lamoda



Наталья Притыковская

Руководитель отдела
рекомендательных систем,
СберМаркет

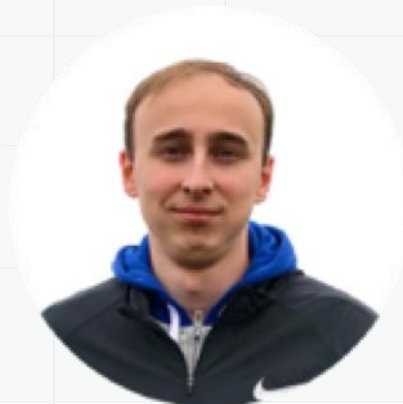
Технологии

4. Старшие товарищи



Владимир Опанасенко

Начальник управления
алгоритмов машинного обучения,
Газпром банк



Андрей Зимовнов

Старший разработчик,
Яндекс.Дзен



Петр Ермаков

Руководитель школы машинного
обучения DataGym,
Senior Data Scientist, Lamoda

Алгоритмы

4. Старшие товарищи



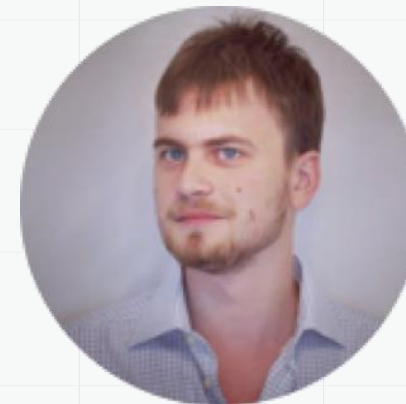
Олег Хомюк

Руководитель R&D,
и.о. CDO
Lamoda



Александр Ульянов

Data Science
Executive Director,
Сбербанк



Александр Филатов

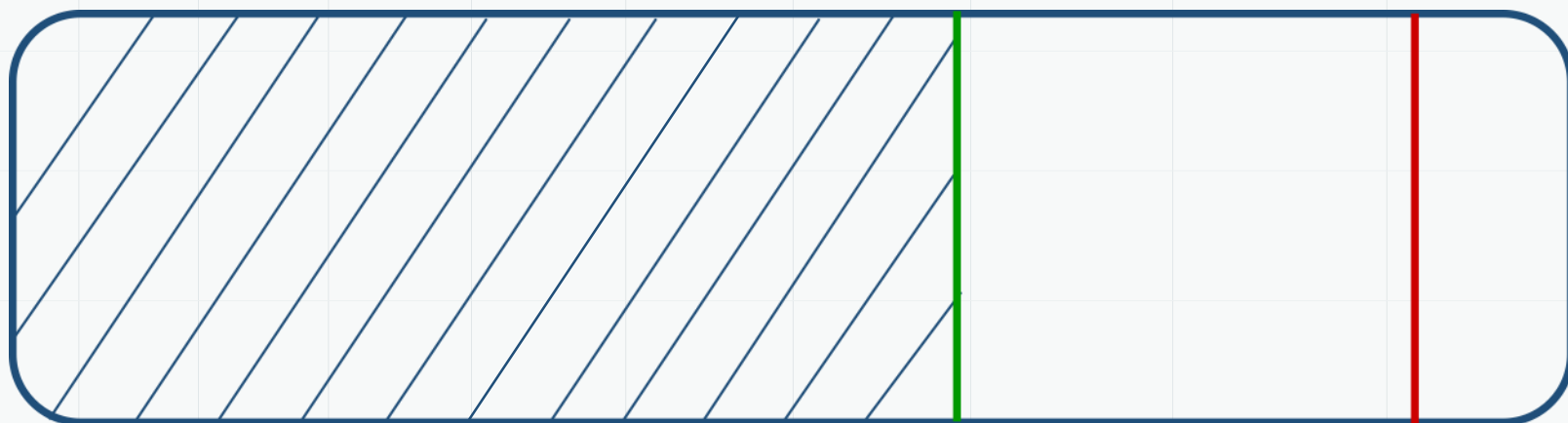
Product Analytics Manager,
VISA

Бизнес

5. Методология

1. Есть лабы. Есть суперачивки.
2. В лабах не все будет разжевываться. Часто придется гуглить.
3. Будет много самостоятельной работы.
4. Лабы будут часто выкладываться до занятий.
5. В Slack создаем возможность для p2p-обучения.

Сертификат



60% сертификат

90% с отличием

- Проект – 20 баллов (2)
- Лаба – 10 баллов (10)
- Суперачивка – 5 баллов (10)
- Код ревью – 1 балл (5)
- Тест – 1 балл (10)

Всего: 205 баллов

Тесты

- Сформировать контекст занятия.
- Проверить свое усвоение материала.
- Вскрыть неявные вещи.
- Получить дополнительные знания.

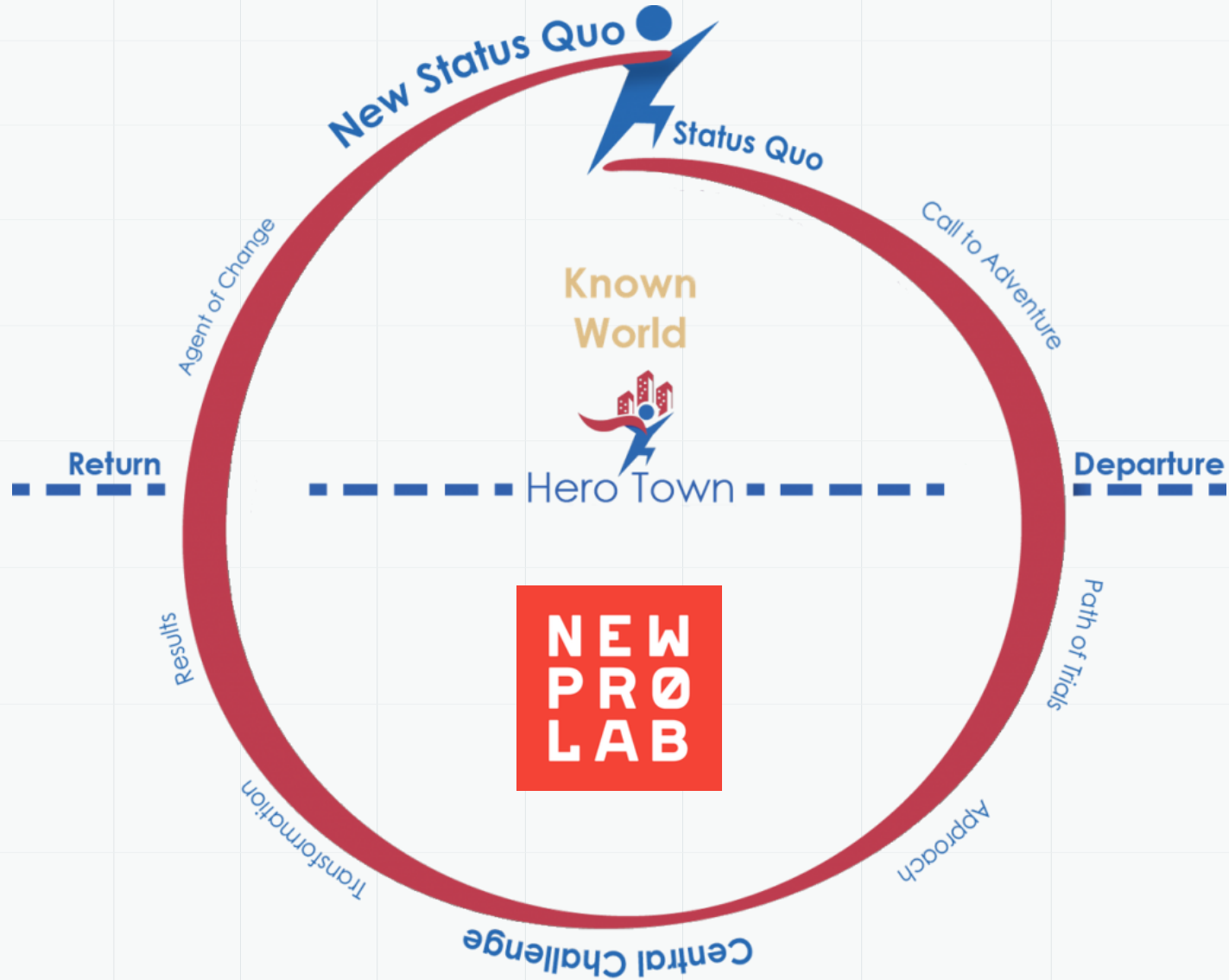
Код ревью

- Получить обратную связь на свой код.
- Записать на подкорку лучшие практики написания кода.
- Нетворкинг.

Обратная связь

Нам **действительно** важно ваше мнение.

Hero's journey



Ближайшие шаги

- 30.09 – выкладываем Лабу 1 и суперачивку
(дедлайн 12.10)
- 01.10 – занятие по Linux
- 03.10 – введение в Python

A group of graduates in black gowns and yellow stoles are seen from behind, standing on a grassy field and throwing their black mortarboard caps into the air. In the background, a city skyline is visible under a dramatic, cloudy sky at sunset or sunrise. The Marina Bay Sands hotel is prominent on the left side of the skyline.

Напутствия и лайфхаки от выпускников

ЗНАКОМСТВО

Команда



Олег Агапов

координирует программу



Анастасия Клименко

отвечает за площадку
и трансляцию



Андрей Булатов

отвечает за инфраструктуру

Команда



Александра Нестерова

отвечает за документальное
сопровождение



Артем Пичугин

отвечает за развитие
программ о данных



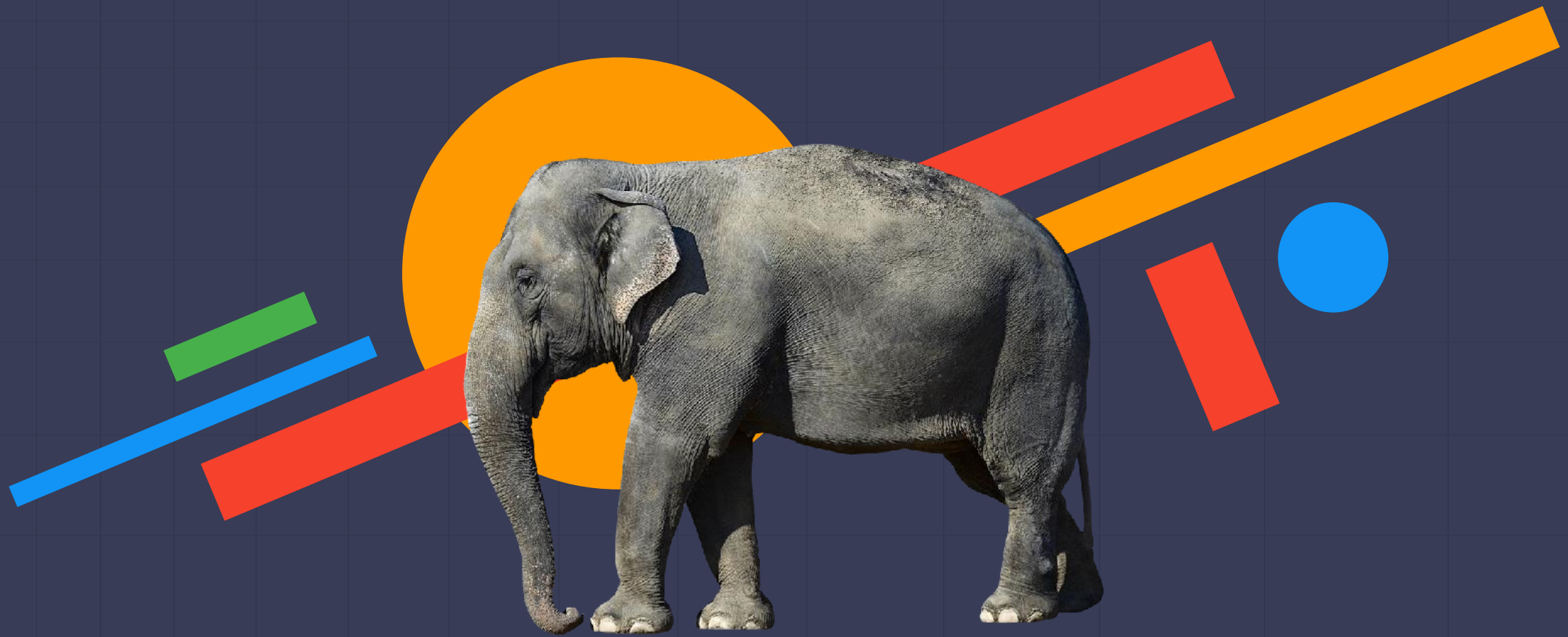
Елена Третьякова

отвечает за всё

Ваша очередь

- Имя, фамилия, компания
- Профессиональное достижение
- Любопытный факт о себе

1 минута



Big Data is Love

NEWPROLAB.COM