



Yandex



# Welcome to Machine Learning for HEP summer school'17

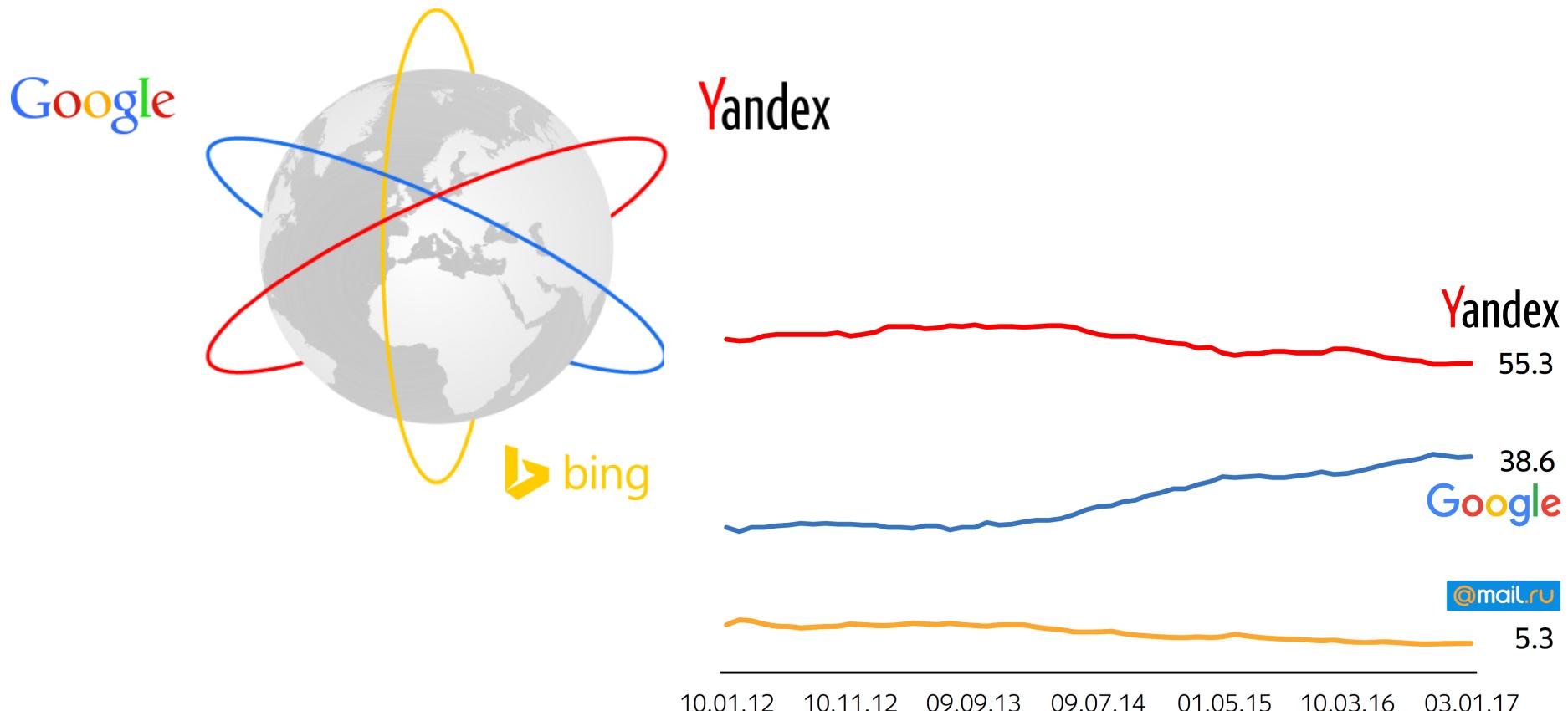
Andrey Ustyuzhanin<sup>1,2</sup>

<sup>1</sup>Yandex School of Data Analysis

<sup>2</sup> National Research University  
Higher School of Economics

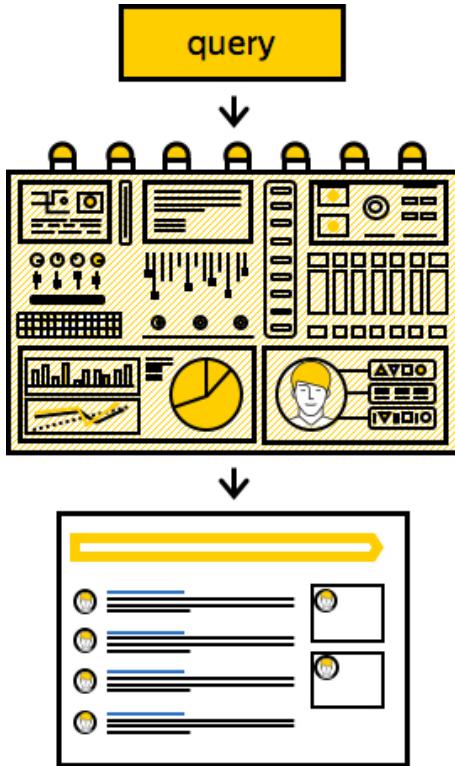
Imperial College  
London

# Global Search Engines



Source: Liveinternet.ru 2012-December, 2016;  
includes desktop and mobile

# Core Yandex ML tasks



## Search Engine tasks:

- Processing of the entire web ( $\sim 10^{10}$  pages, <http://www.worldwidewebsize.com> )
- Aggregation of users on behaviour analytics and micro-segmentation ( $\sim 10^8\text{-}10^9$ )
- Personalized delivery of relevant **URLs/Images/Ads** with the highest probability of click

## Internal Machine learning application examples:

- Navigation / maps
- Natural Language Translation
- Speech recognition/ synthesis
- Ya.Market (aggregator of Internet Webstores)
- Music streaming / recommendation
- Weather prediction (“nowcasting”)
- Ya.Taxi – demand prediction for specific space-time region, driving cars

# Yandex School of Data Analysis



Established in 2007

Best lecturers in the field of Statistics, ML, Deep Learning

Joint master's program with leading Russian universities

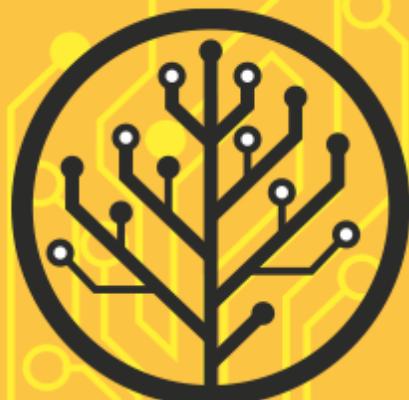
- 2 years, 12+ courses (<https://yandexdataschool.com/edu-process/program>)

- Algorithms & Data Structures
- Programming Languages (C++ and Python)
- Machine Learning
- Parallel and Distributed Computing
- Big Data Storage & Processing
- Images & Video processing
- Information Retrieval
- Deep Learning
- Reinforcement Learning

Admission

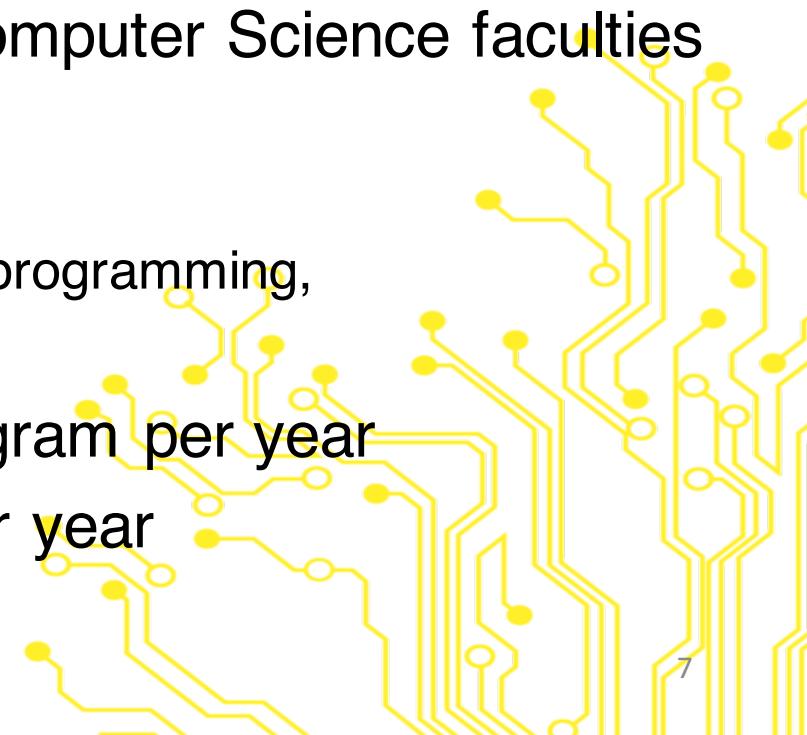
- 4000 applications per year
- 250 are accepted per year
- ~500 graduated so far

Яндекс



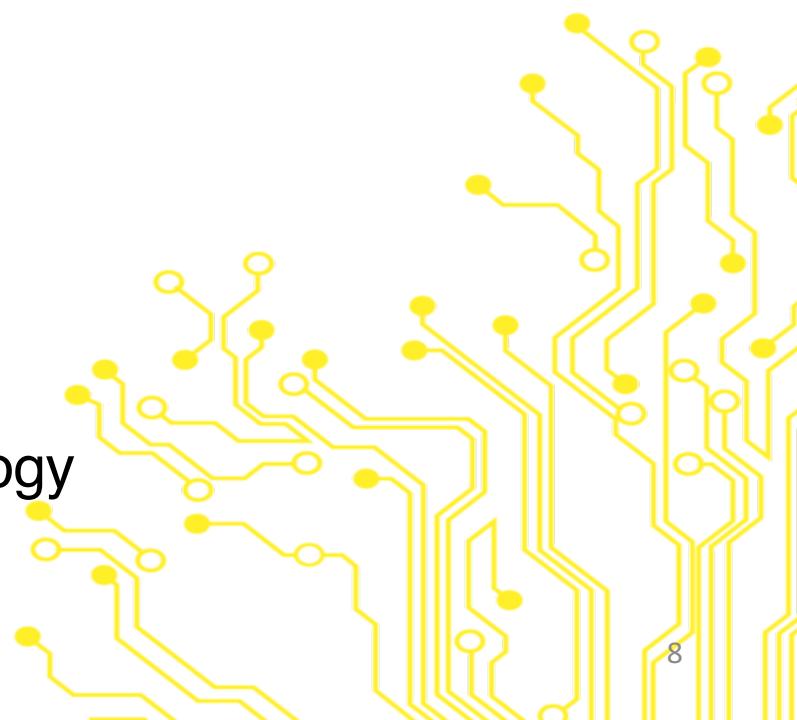
# Faculty of Computer Science,

- established in March 2014,
- active support from Yandex (more than 50 lecturers/seminarists are Yandex employee)
- Also cooperates with Kaspersky, WorldQuant, Huawei, ...
- Goal: enter the world's top 30 Computer Science faculties by 2020
  - big data storing and processing,
  - software development and system programming,
  - machine learning
- 370 students enter bachelor program per year
- ~100 finish master's program per year



# International Partner Universities

- Humboldt-Universität zu Berlin
- Technische Universiteit Eindhoven
- Technische Universität Dresden
- Université Grenoble Alpes
- Université Blaise Pascal
- Radboud University, Nijmegen
- Aalto University, Helsinki
- Lappeenranta University of Technology



# Research Laboratories

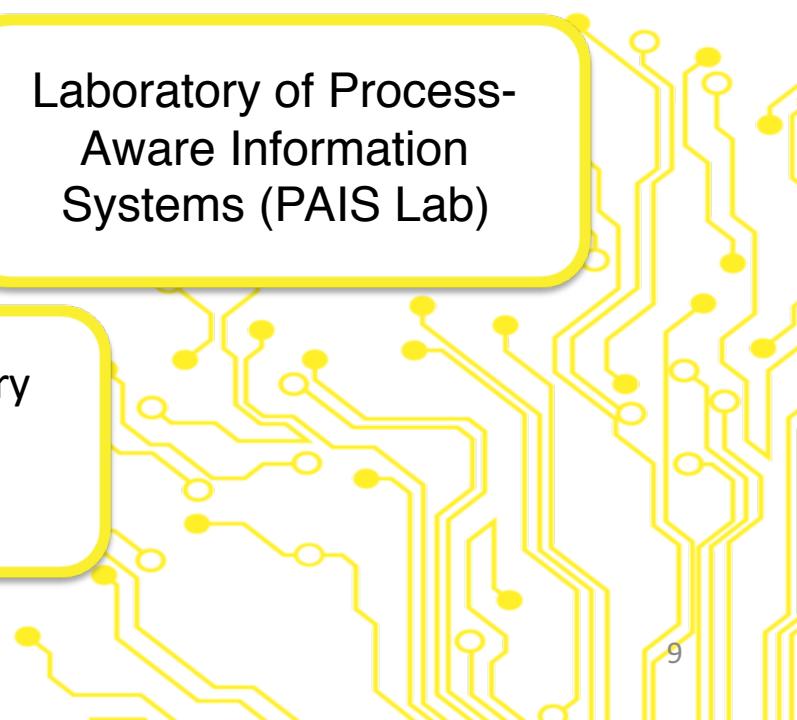
International Laboratory for Intelligent Systems and Structural Analysis

International Laboratory of Theoretical Computer Science

Laboratory of Methods for Big Data Analysis (LAMBDA)

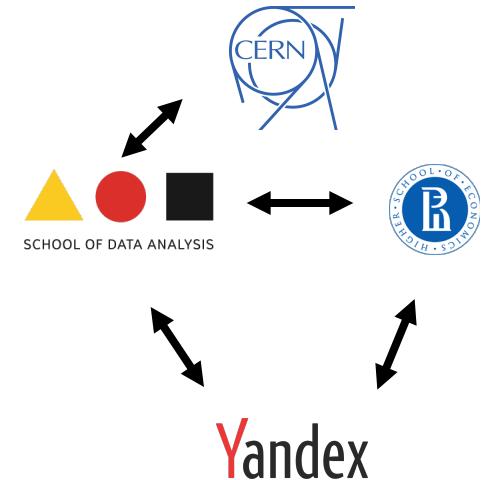
Laboratory of Process-Aware Information Systems (PAIS Lab)

International Laboratory of Deep Learning and Bayesian Methods



# Laboratory of Methods for Big Data Analysis

- Group working on data analyses in Natural sciences
- 2 physicists and 7 computer scientists
- Part of Yandex School of Data Analysis
- Members of the LHCb, SHiP, OPERA, CRAYFIS collaborations
- Project examples:
  - Storage/Speed optimization for triggers
  - Particle Identification algorithms
  - Brain cognitive studies
  - Ultra-High Energy Cosmic Ray Detection



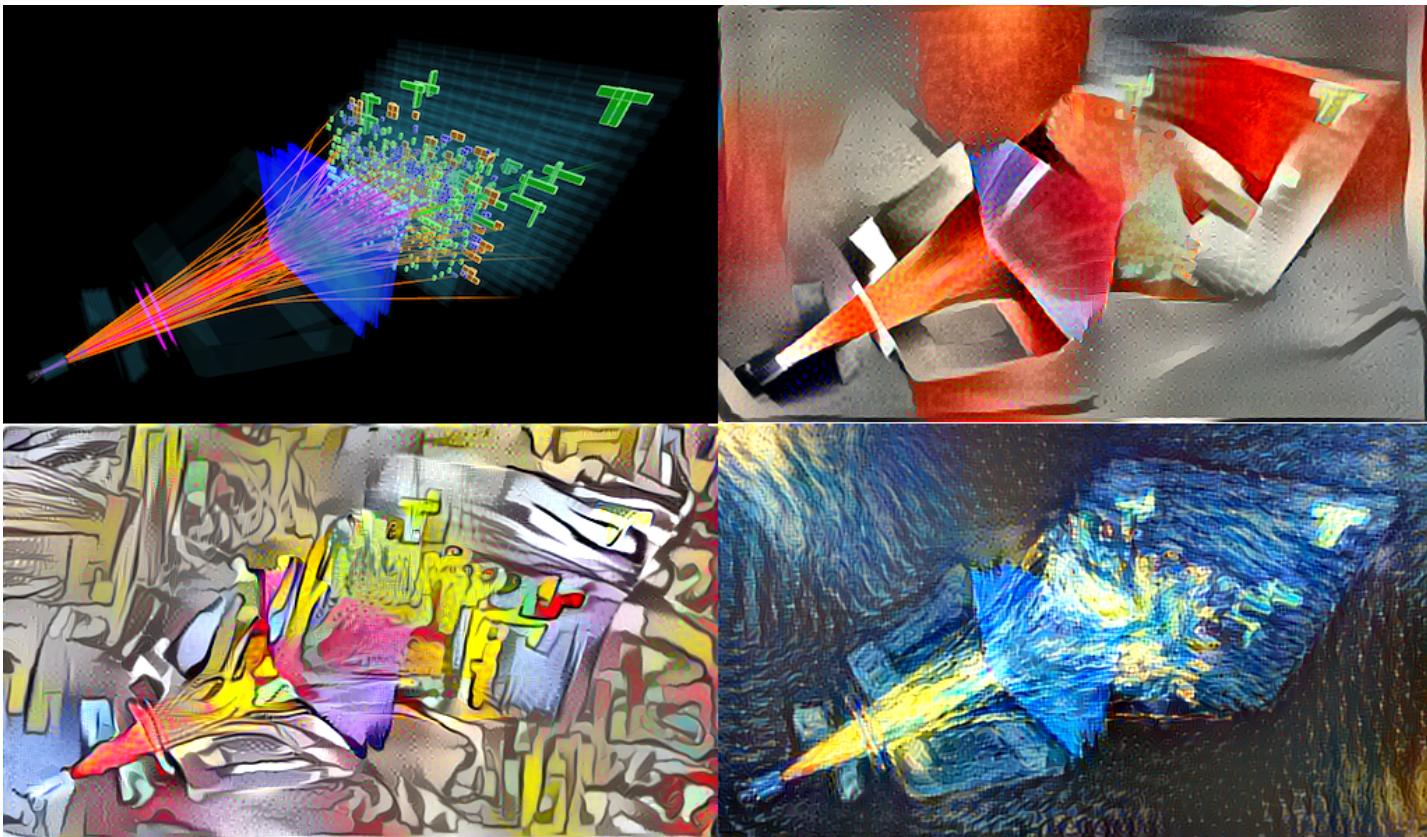
# The 2<sup>nd</sup> Summer School on Machine Learning in High Energy Physics 2016, Lund, Sweden



# The 3<sup>rd</sup> Summer School on Machine Learning in HEP, Reading, 2017

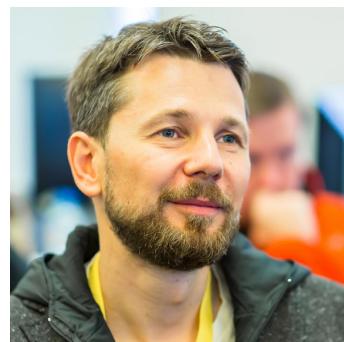
**Imperial College  
London**

Professor Ulrik Egede, head of LHCb group



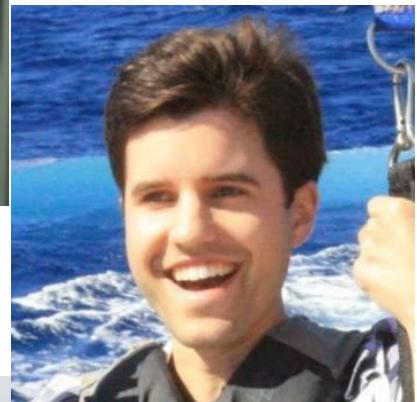
# Core Team People

- Dr. Alexey Artemov,  
Yandex, HSE
- Alexander Panin  
YSDA, Yandex
- Maxim Borisyak  
Yandex, HSE
- Nikita Kazeev  
Yandex, YSDA, HSE
- Dr. Andrey Ustyuzhanin  
Yandex, HSE



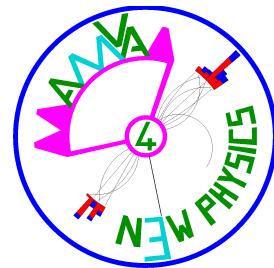
# Guest Lecturers

- Prof. Mike Williams,  
MIT
- Dr. Noel Dawe,  
University of Melbourne
- Dr. Tim Head,  
Wild Tree Tech



# Sponsors, Partners

- Winton
- AMVA4NP
- Microsoft



# School highlights

- A bit of Machine Learning theory
- Practical Seminars
- Invited talks on ML applications to HEP challenges
- ML Competitions with HEP flavour (Dark matter signals search)
- Industry ML examples

# School Communications

- Social Media
  - #MLHEP
- Wednesday: cultural day
- Thursday: Episode-1, winner's solutions
- Friday: own project round table
- Sunday: Episode-2, winner's solutions
- Reception this evening, Dinner on Friday
- Questions / Answers - anytime
  - <https://gitter.im/yandexdataschool/mlhep2017>

# Thank you!

Let's start the party lecture...