

Яндекс

Москва,
23 октября 2014

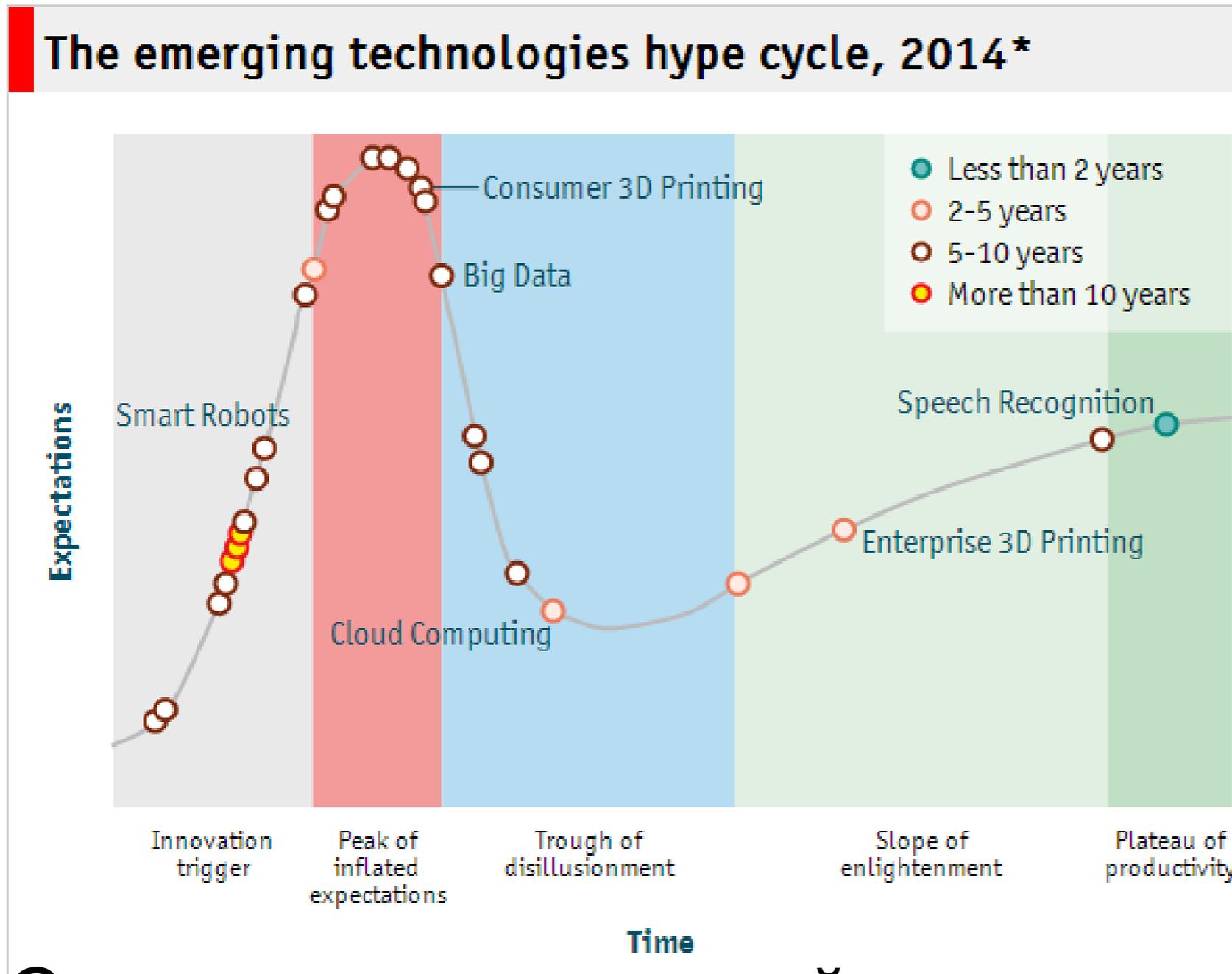


Люди и алгоритмы: вместе или вместо?

Андрей Себрант

Яндекс
Директор по маркетингу сервисов

Преодолеть пропасть



Стадия разочарований опасна, выгоду получают спокойные и умные;)

Традиционное восприятие науки

Three Takeaways Regarding Predictive Analytics

Gottfried Wilhelm Leibniz

“If someone could have a **sufficient insight** into the inner parts of things, and in addition had **remembrance** and **intelligence** enough to consider all the circumstances and to take them into account, he would be a prophet and would see the future in the present as in a mirror.”



July 1, 1646 – November 14, 1716

- An originator of Calculus
- Inventor of Mechanical Calculators
- Created the first mass-produced calculators
- Refined the binary number system

Data Science != Analytics

Традиция:

Глубокое понимание

Человеческий разум

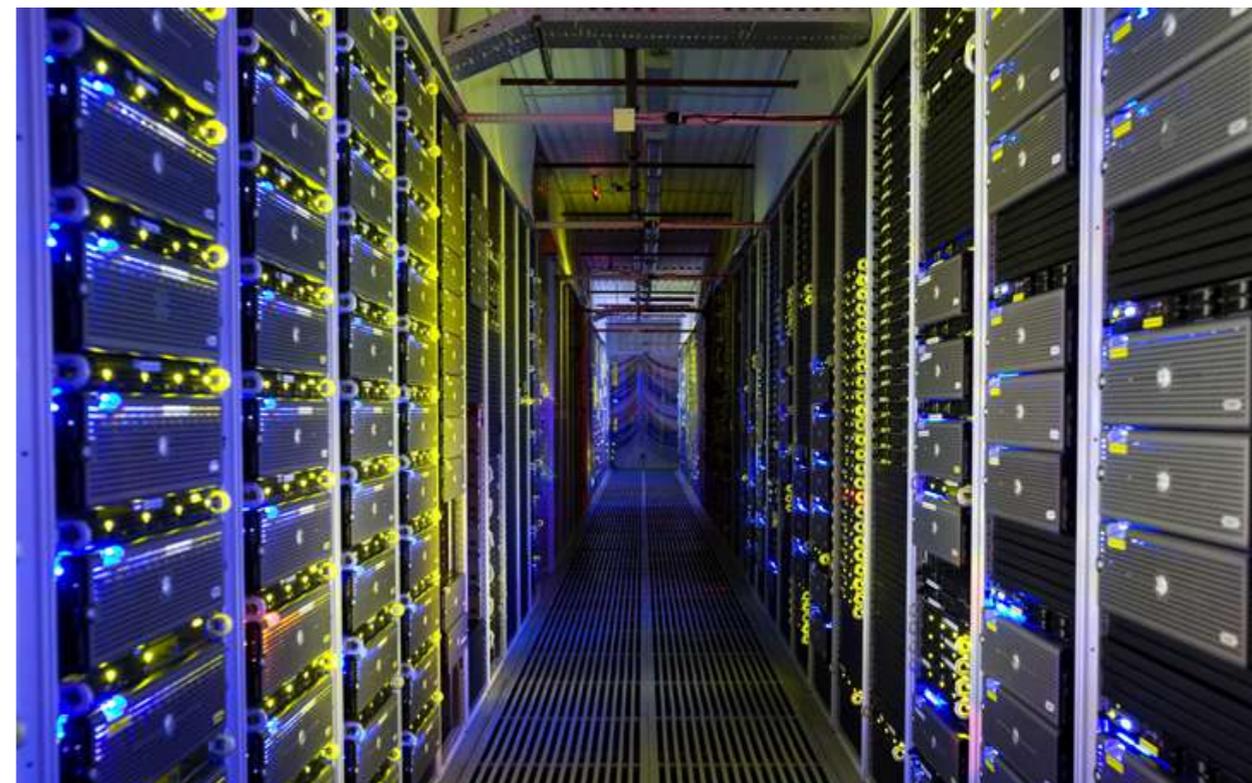
Построение моделей

Длительный процесс
(подчас месяцы и годы)



Data Science:

- Машинное обучение
- Колоссальные вычисления
- Алгоритмические предсказания
- Ответ в реальном времени
(часто за доли секунды)



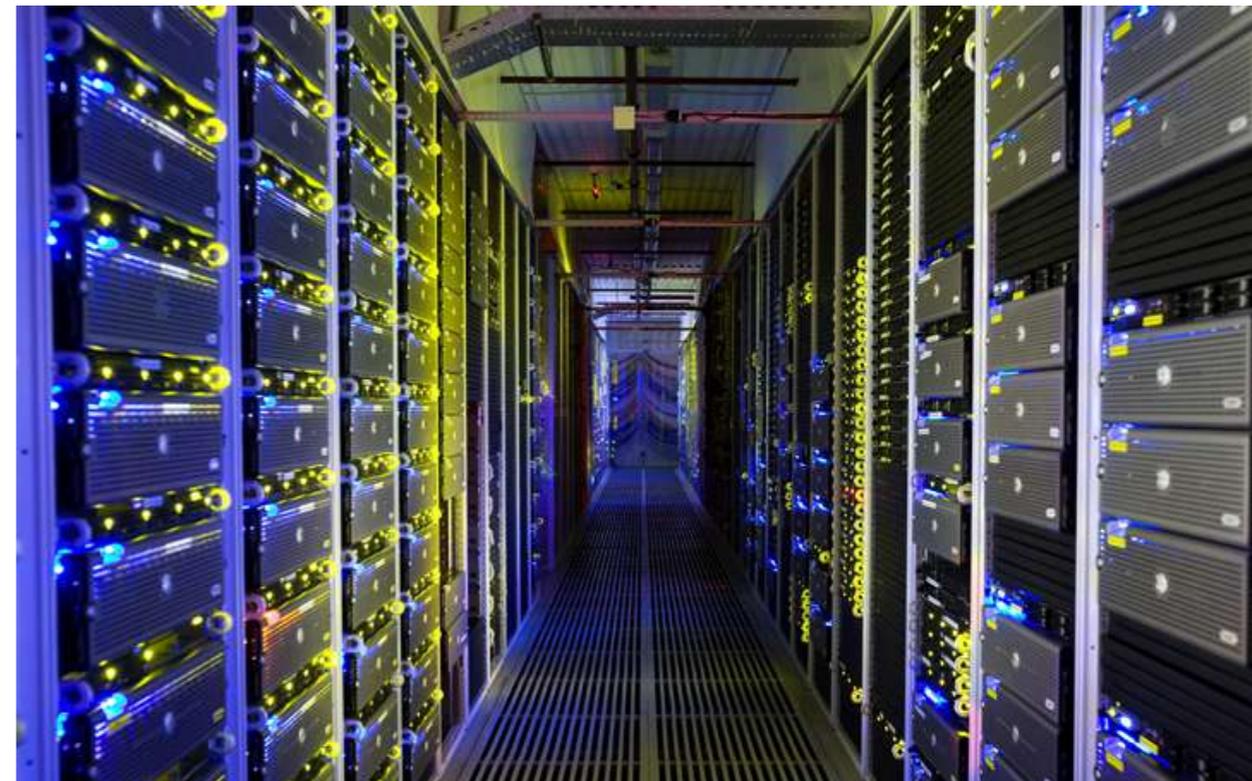
Data Science != Analytics

Традиция:

- Вместе

Data Science:

- Вместо



Data Science != Analytics

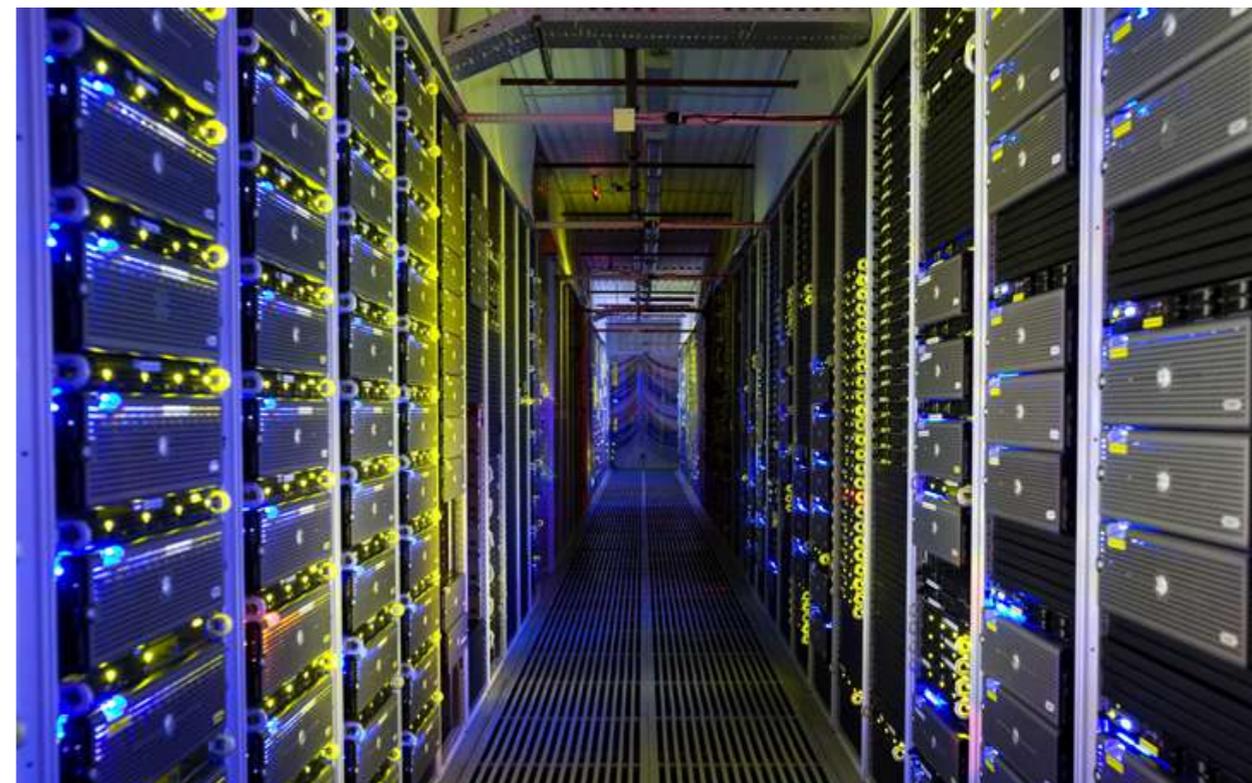
Традиция:

- Human readable output

Data Science:

- Machine readable output

<http://blogs.hbr.org/2014/08/the-question-to-ask-before-hiring-a-data-scientist/>



Есть и хорошие новости

Полезно анализировать принимаемые машиной решения:

- Проверка на здравый смысл
- Новые инсайты

Case study Yandex Crypta

Look-A-Like in Ad Targeting (and what do they search)

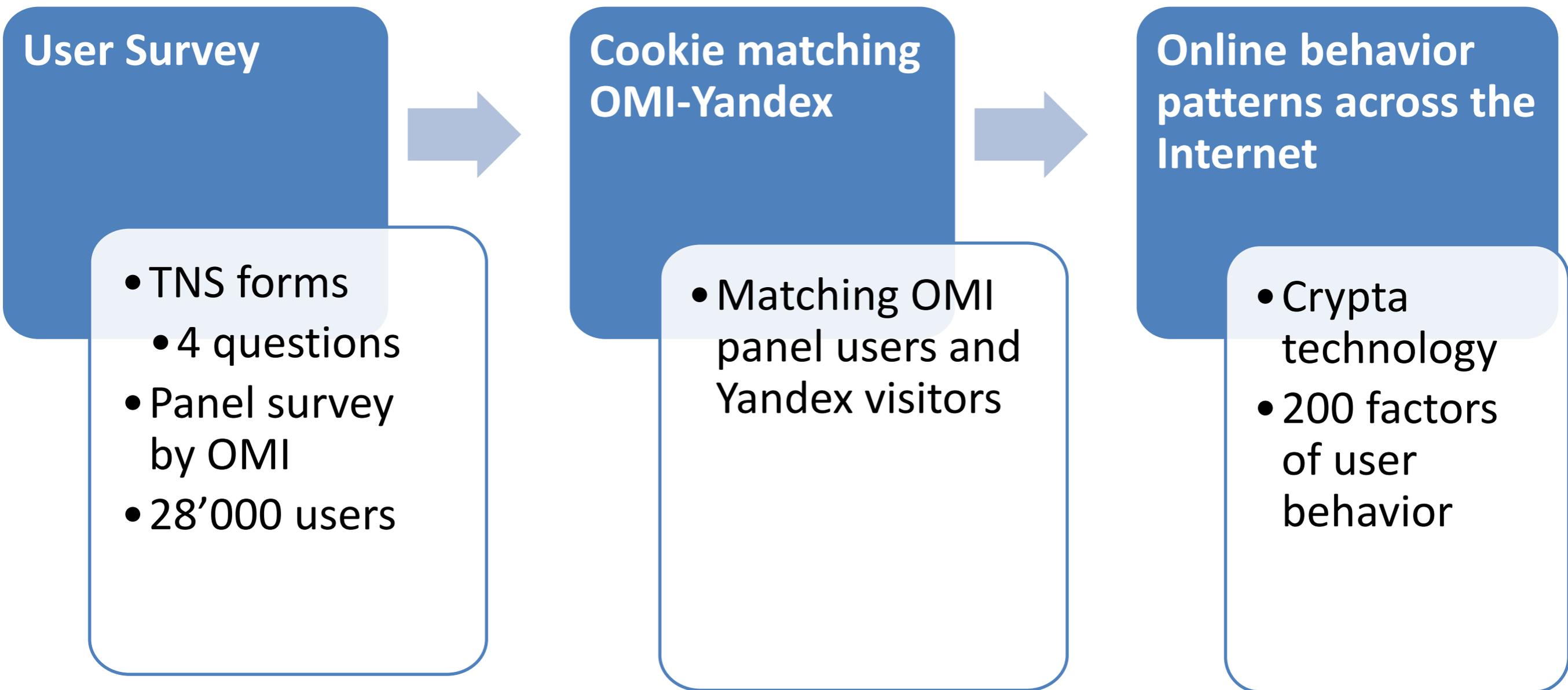


LOOK-A-LIKE



ПОИСК ПОХОЖИХ

Light TV-viewers: methodology



Heavy TV viewers



«сбербанк», «коммунальный»,
«шарлотка», «выкройка»,
«биглион», «irr», «заработать»

Больше запросов кириллицей

Light TV viewers



«книга», «переводчик»,
«словарь», «формула»,
«японский», «французский»,
«немецкий», «такси»

Много запросов латиницей

Heavy TV viewers



«ТНТ», «дом-2»,
«телепрограмма», «СТС»

Light TV viewers



«C++», «wi-fi»,
«фотошоп», «torrent»,
«adobe»

Heavy TV viewers



«спартак», «цска», «пиво»

Light TV viewers



«загранпаспорт», «авиабилет»,
«виза», «самолет»,
«аэропорт», «ржд»

Где алгоритмы уже успешней?

Есть две области помимо рекламы, где хорошо обученные алгоритмы уже доказали, что работают лучше людей.

Обе основаны на использовании алгоритмических рекомендаций.

- Электронная торговля
- Музыкальное интернет-радио

Вместо выводов

- Везде, где нужно принимать много однотипных решений, алгоритмы справляются лучше
- Инсайты, получаемые из человеческого анализа алгоритмических решений, дают новое уникальное знание
- Научитесь отличать: какие решения лучше принимает человек, а какие- машина. И никогда не путайте два этих класса ;)

Очень важный момент:

Алгоритмы, даже великолепно
отвечающие на вопросы,
не умеют вопросы задавать

Спасибо! ;)

Андрей Себрант



+7 (495) 739-7000



@asebrant



asebrant@yandex-team.ru



<https://www.facebook.com/asebrant>