Affective Computing

Olga Perepelkina HSE, 2022

План лекции

- Тема 11: Автоматическое распознавание и синтез характеристик личности
- Презентации по теме «Теории личности»
- Тема 12: Приложения, индустриальные кейсы, тренды, этические вопросы
- Презентации по теме «Этика и искусственный интеллект»

Автоматическое распознавание и синтез характеристик личности (Personality Computing)

Personality

- Личность это психологическая конструкция, направленная на объяснение широкого разнообразия поведения человека с точки зрения нескольких устойчивых и измеримых индивидуальных характеристик.
- В этом отношении любая технология, включающая понимание, прогнозирование и **синтез человеческого поведения**, вероятно, выиграет от personality computing, то есть от технологий, способных работать с человеческой личностью.

Personality

- Современные модели личности успешно предсказывают «шаблоны мышления, эмоций и поведения» [3], а также важные жизненные аспекты, в том числе «счастье, физическое и психологическое здоровье, [...] качество отношений со сверстниками, семьей и романтическим партнером». [...] выбор профессии, удовлетворение и результативность, [...] участие в жизни сообщества, преступную деятельность и политическую идеологию» [4].
- [3] D. Funder, "Personality," Annual Reviews of Psychology, vol. 52, pp. 197–221, 2001.
- [4] D. Ozer and V. Benet-Martinez, "Personality and the prediction of consequential outcomes," Annual Reviews of Psychology, vol. 57, pp. 401–421, 2006.

Personality

- Ключевое допущение психологии личности состоит в том, что стабильные индивидуальные характеристики приводят к стабильным поведенческим паттернам, которые люди склонны демонстрировать, по крайней мере, в определенной степени, независимо от ситуации.
- Цель психологии личности различать внутренние свойства человека исходя из наблюдаемого поведения и исследовать причинно-следственные связи между ними.

Рост интереса к личности с стороны НСІ

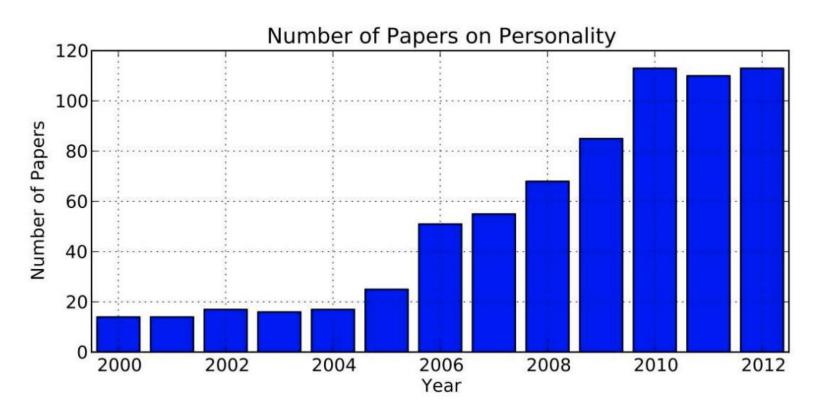


Fig. 1. The chart reports the number of papers per year with the word "personality" in their title (sum over IEEE Xplore and ACM Digital Library).

Рост интереса к личности с стороны НСІ

• Причины:

- 1) Рост объема персональной информации в социальных сетях
- 2) Упрощение сбора ежедневной информации о спонтанном поведении с помощью смартфонов
- 3) Попытка наделить машины социальным и эмоциональным интеллектом, способностью взаимодействовать с людьми, как это делают люди

Зачем автоматически определять личностные черты?

Зачем автоматически определять личностные черты?

- Предсказание поведения: распознаем личность -> более точные предсказания
- Персонализация
 - Экстраверты в интернет-магазинах могут предпочитать читать чужие обзоры, а интроверты техническое описание товара
 - Онлайн игры и групповые чаты могут подбирать участников на основе их личностных черт (рекомендуя совместимых партнеров)
- Убеждение
 - Людей проще убедить, если агент подстраивается под их личностные черты
- Обучение: компьютер может помочь формировать навыки (уверенность, эмпатия)

Модели личности

- Основаны на:
 - физиологии (the biological perspective)
 - бессознательном (the psychoanalytic perspective)
 - окружающей среде (the behaviorist perspective)
 - внутренних состояниях (the humanistic perspective)
 - мышлении (the cognitive perspective)
 - чертах (traits)

Модели личности

- Модели черт строятся на семантическом сходстве и отношениях между прилагательными, которые люди используют для описания себя и других
- Хотя термины, используемые для описания людей, многочисленны и сильно различаются, они обычно составляют лишь несколько основных измерений.
- Эти измерения, если они достаточно стабильны, затем принимаются как черты личности, то есть как факторы, способные улавливать устойчивые индивидуальные характеристики, лежащие в основе наблюдаемого поведения.

Модели личности

- Критика моделей черт: они являются описательными и не отражают реальных характеристик личности
- С другой стороны, несколько десятилетий исследований и экспериментов показали, что одни и те же черты появляются с удивительной регулярностью в широком спектре ситуаций и культур, что позволяет предположить, что они действительно соответствуют психологически значимым явлениям.

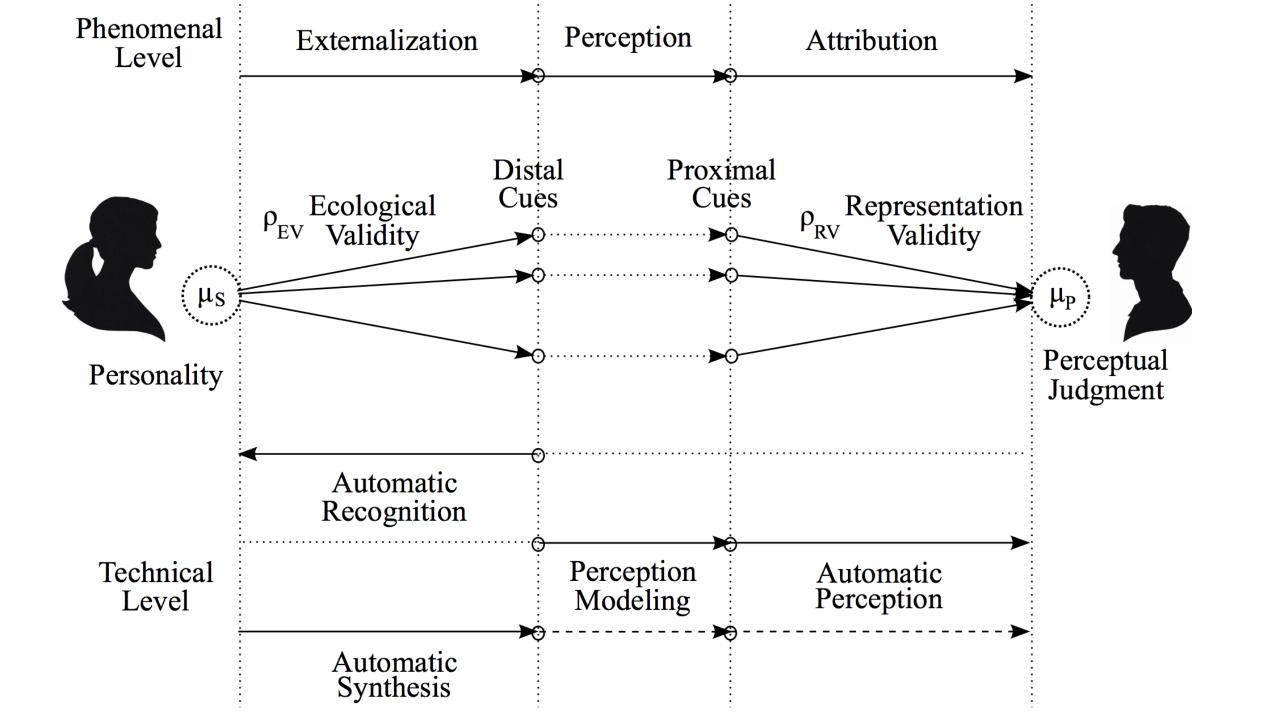
Big-5

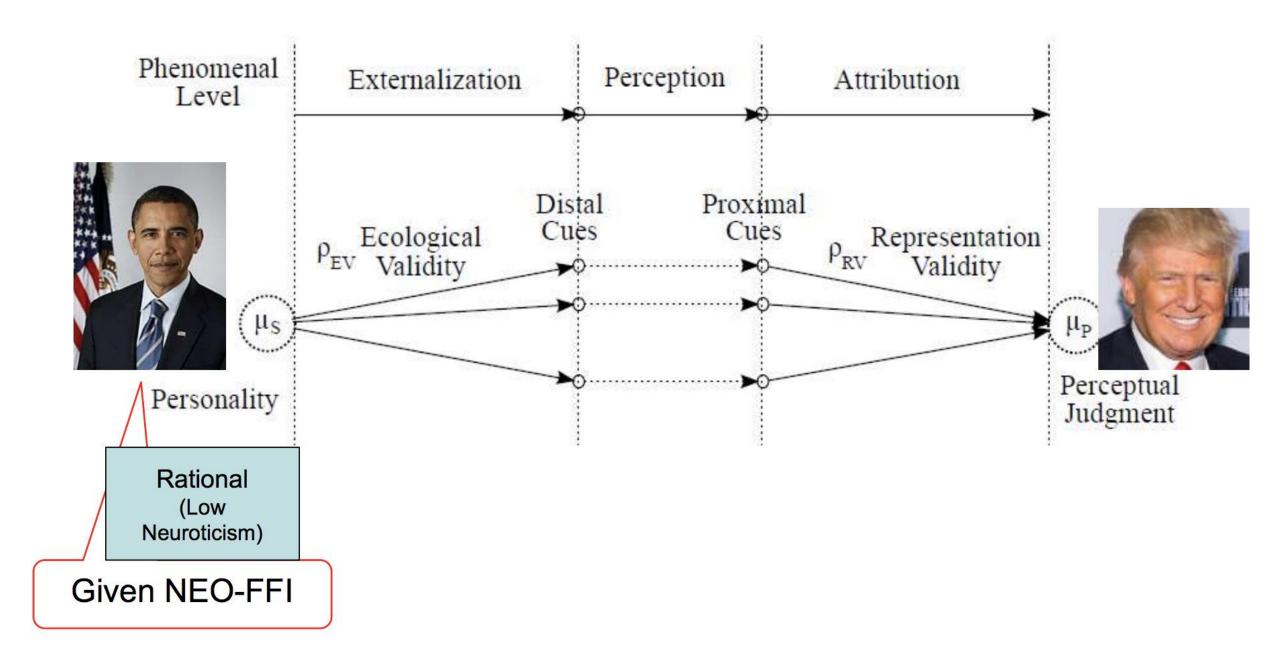
- Наиболее распространенная модель черт личности
- Экстраверсия: Active, Assertive, Energetic, Outgoing, Talkative
- Согласие (сотрудничество): Appreciative, Kind, Generous, Forgiving, Sympathetic, Trusting
- Сознательность (добросовестность): Efficient, Organized, Planful, Reliable, Responsible, Thorough
- Нейротизм: Anxious, Self-pitying, Tense, Touchy, Unstable, Worrying
- Открытость опыту: Artistic, Curious, Imaginative, Insightful, Original, Wide interests

Personality Computing

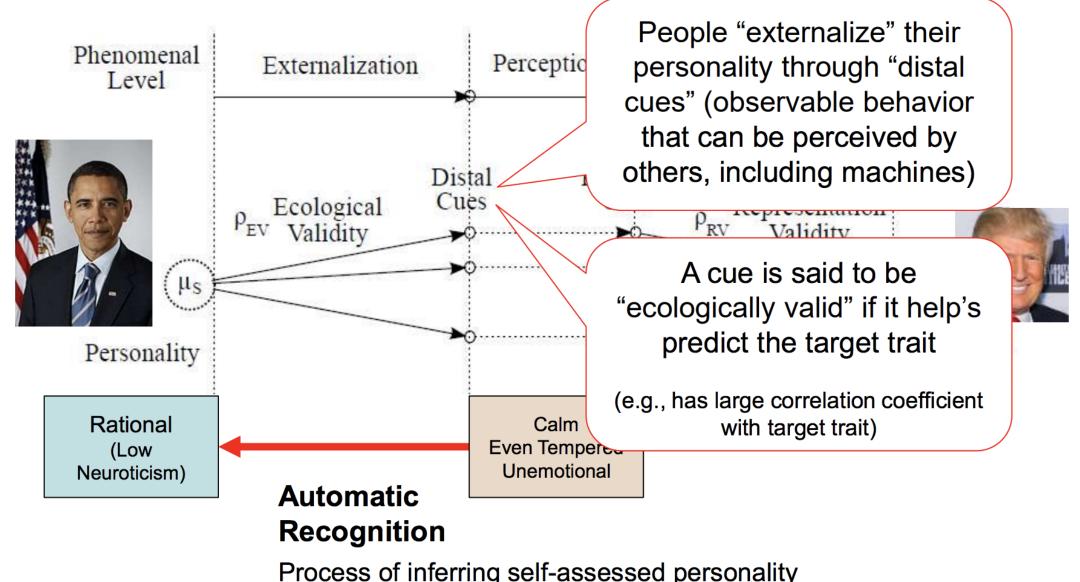
- Automatic Personality Recognition: inference of the true personality of an individual from behavioral evidence (распознавание истинной личности человека из поведения)
- Automatic Personality Perception: inference of personality others attribute to an individual based on her observable behavior (предсказание личности, которую другие приписывают индивиду, на основе его наблюдаемого поведения)
- Automatic Personality Synthesis: generation of artificial personalities via embodied agents (генерация искусственных личностей через воплощенных агентов)

- Объединяет все три термина: Personality recognition, Personality perception, Personality synthesis
- Личность недоступна для непосредственного наблюдения
- Она проявляется через «физические следы», «маркеры», то есть наблюдаемое поведение (distal cues)
- Это поведение может быть воспринято другим





• Automatic Personality Recognition: автоматические распознавание личности – определить самооценку личности по детектируемым машиной дистальным сигналам (выявление черт согласно самооценки – *истинных* черт личности индивида)

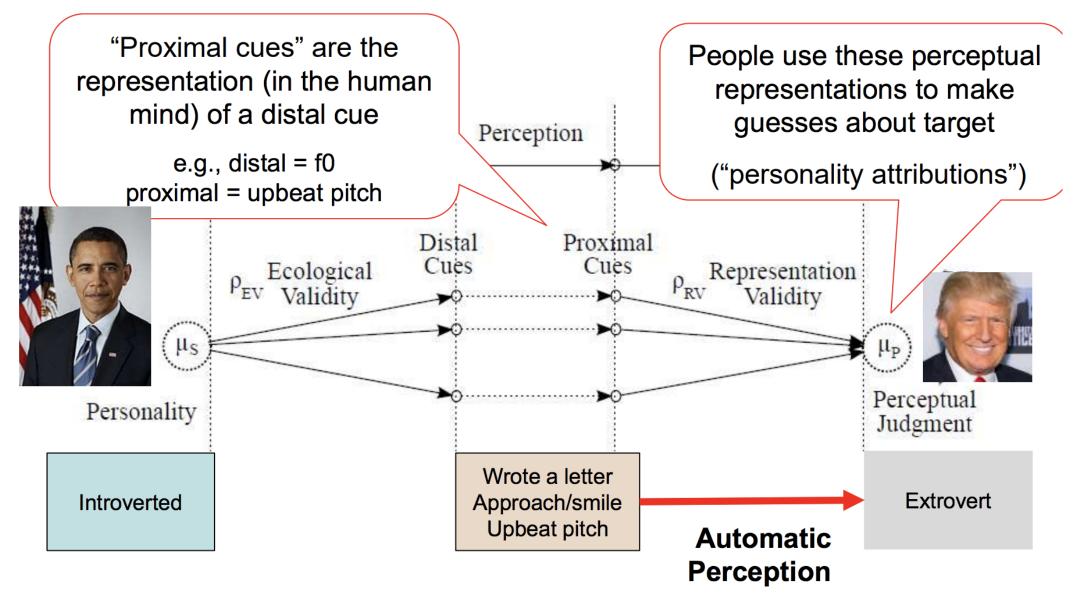


Process of inferring self-assessed personality from machine detectable cues e.g., computer predicts Obama is low in Neuroticism

- Дистальные сигналы достигают наблюдателя и вовлекаются в процесс восприятия, в результате которого формируется образ (перцепт), то есть «ментальная репрезентация того, что воспринимается»
- Наблюдатель принимает **proximal cues** проксимальные сигналы активируют процесс атрибуции, то есть выработку **перцептивного суждения**, которое определяет черты характера, которые **наблюдатель приписывает наблюдаемому человеку**.

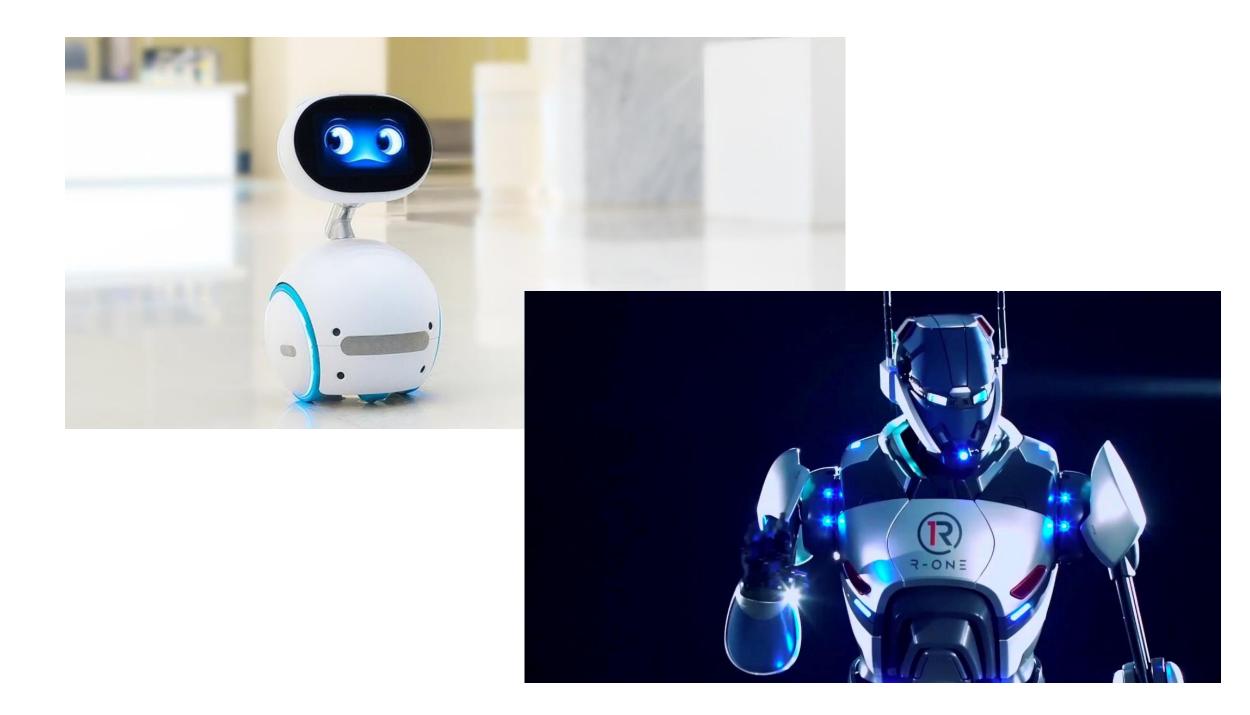
- Automatic Personality Perception (APP) задача автоматического выявления приписываемой личности по проксимальным признакам
- В отличие от Automatic Personality Recognition (APR) цель APP не истинная личность, а то, какой ее воспринимают другие
- То есть APR строится на **самооценке**, а APP на оценках **других** людей
- Методики не могут использовать проксимальные сигналы, используют **дистальные** в качестве приближения

- Цель APP обычно предсказать среднее значение черты, полученное **несколькими** наблюдателями
- Наблюдатели отличаются (социальный статус, когнитивные особенности, личность и тд) => их оценки тоже отличаются
- Индивидуальные оценки являются результатом фактического процесса атрибуции. Поэтому их можно считать настоящими приписанными личностями. Средняя оценка не является результатом процесса атрибуции и, следовательно, не может считаться реальной приписанной личностью (но все равно важно предсказывать средние оценки можно выявить общие закономерности и механизмы процесса атрибуции)

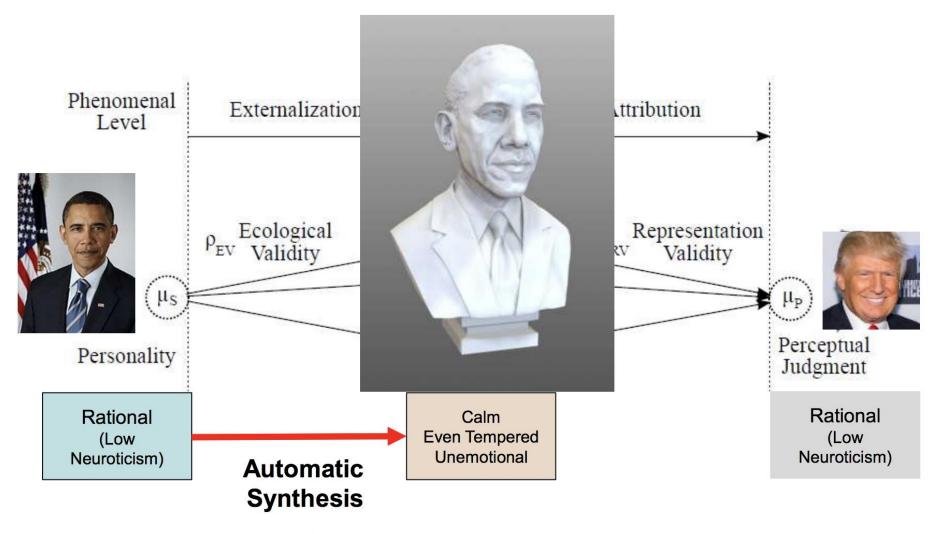


Process of predicting the personality attributions people might make Based on that exchange, computer predicts Trump thinks Obama is Extroverted

- Люди самопроизвольно и бессознательно присваивают социально значимые характеристики, в том числе черты личности, любому человеку, с которым они встречаются
- [J. S. Uleman, S. A. Saribay, and C. M. Gonzalez, "Spontaneous inferences, implicit impressions, and implicit theories," Annual Reviews of Psychology, vol. 59, pp. 329–360, 2008.]
- Это явление настолько естественное и распространенное, что оно применимо не только к людям, но и к любому устройству, которое имеет черты, подобные человеческим

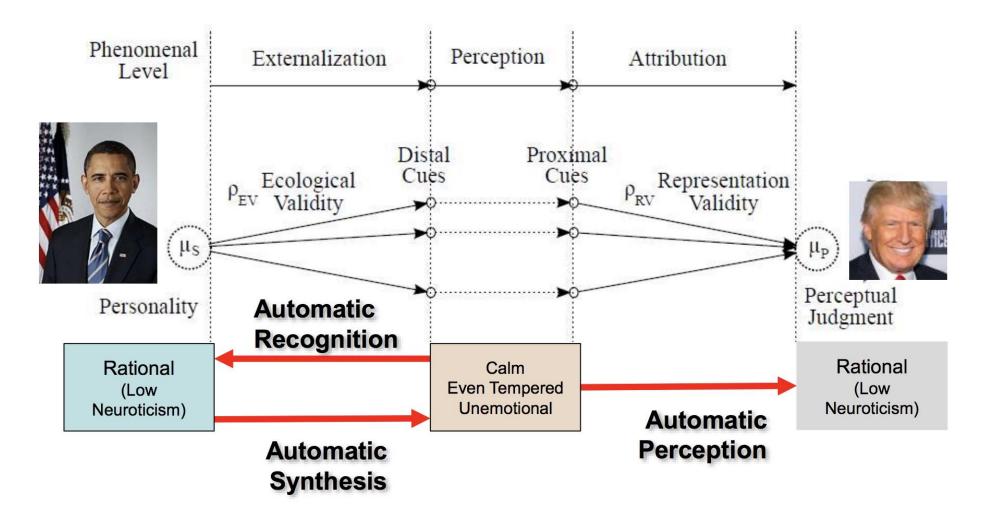


- Automatic Personality Synthesis (APS) автоматический синтез личности задача автоматической генерации дистальных сигналов, нацеленных на демонстрацию желаемых черт личности.
- Включает 2 процесса: экстернализация (генерация признаков, cues)
 машиной/роботом/искусственным агентом и тд
- И атрибуция люди, как правило, не осознанно, приписывают черты личности машине.
- Основная цель гарантировать, что признаки, генерируемые машиной, вызывают нужное восприятие черт личности наблюдателями.



Task of automatically generating distal cues to elicit the attribution of a desired trait

Personality computing summary



AUTOMATIC PERSONALITY RECOGNITION

Automatic Personality Recognition (APR)

- Подходы APR, представленные в литературе, учитывают широкий спектр дистальных сигналов, включая письменные тексты, невербальное поведение, данные, собранные с помощью мобильных или носимых устройств, онлайн-игры.
- Эксперименты основаны на **самооценках**, и в некоторых случаях это позволяет проводить эксперименты над **несколькими тысячами** субъектов.
- Используют **классификацию** (предсказывать, находится ли человек выше или ниже медианы по шкале Big-5) и **регрессию** (значение по шкале Big-5)

Personality Recognition from text

- Классическая парадигма исследований:
 - Дать испытуемым заполнить личностные опросники (стандартные, например, Big-5)
 - Дать испытуемым задачу сгенерировать текст (например, написать эссе)
 - Провести простой анализ текста: например, посчитать корреляцию между частотой разных слов с личностными чертами
- (Идеально) показать «внешнюю валидность»: продемонстрировать, что обученная модель предсказывает поведение, связанное с этими чертами
 - Например, слова, связанные с экстраверсией, предсказывают, что блогеры будут больше говорить о себе

Personality Recognition from text

- Одно из ранних исследование было проведено в 2005 г.
 - 1200 студентов написали 2263 эссе и заполнили NEO-FFI (типа Big-5)
 - Цель исследования предсказывать по тексту экстраверсию и нейротизм
 - Модель (SVM) показала точность классификации (ассuracy) 58% для обеих черт
 - В более позднем исследовании на тех же данных и всех чертах Big-5 получили от 50% от 62% (максимальная для открытости)
- S. Argamon, S. Dhawle, M. Koppel, and J. Pennbaker, "Lexical predictors of personality type," In Proceedings of Interface and the Classification Society of North America, 2005.

Personality Recognition from text

- Позже стали проводить исследования на блогерах (также просили их заполнить личностные опросники)
- 5042 posts written by 2393 bloggers
- Точность получилась низкая, но выявили интересный факт, что тревожные блогеры пишут, чтобы избавиться от напряжения, а экстраверты чтобы рассказать о своей жизни
- A. J. Gill, S. Nowson, and J. Oberlander, "What are they blogging about? personality, topic and motivation in blogs," in Proceedings of the The International AAAI Conference on Weblogs and Social Media, 2009, pp. 18–25.

Personality Recognition and Nonverbal Communication

- Исследования включают: невербальные характеристики **речи** (просодия, интонация), движения **тела**, межличностные **расстояния**, **мультимодальные сигналы**
- Определение черт личности по невербальным параметрам речи:
 - Extraversion (63% accuracy) and Conscientiousness (95% accuracy) (24 испытуемых, симулирующих телефонные звонки 119 звонков)
 - Extraversion (66%) and Neuroticism (75%) (13 испытуемых)
 - A. V. Ivanov, G. Riccardi, A. J. Sporka, and J. Franc, "Recognition of personality traits from human spoken conversations," in Proceedings of Interspeech, 2011, pp. 1549–1552.

Personality Recognition and Nonverbal Communication

- Определение черт личности **по межличностным расстояниям и скорости походки**: 13 испытуемых, Экстраверсия (66%), Неротизм (75%) [G. Zen et al., 2010]
- **Мультимодальное** определение черт личности [L. M. Batrinca et al., 2011]:
 - 89 Skype звонков
 - Признаки: голосовые (pitch, intensity, duration), направление взгляда, позы, движения головой, руками, нахмуренность и тд
 - 5 черт личности от 65% до 76% точность

Personality Recognition on Social Media

- Социальные сети (Facebook, Twitter и тд) один из основных каналов, по которым люди взаимодействуют с другими, идеальное средство для самораскрытия и, следовательно, отличная площадка для исследований в области Personality computing.
- Признаки: сами посты, информация о себе (пол, возраст, предпочтения и тд), about me etc
- [J. Golbeck et al., 2011]: анализ пользователей Twitter, 2000 твитов 279 пользователей. МАЕ (регрессия) от 0.11 до 0.18 (Big-5)
- [M. Cristani et al., 2013] Анализ любимых изображений 300 пользователей Flickr (200 изображений для каждого пользователя). Корреляция между предсказанной и реальной чертой от 0,17 до 0,22, значимая только для Openness

Personality Recognition via Mobile and Wearable Devices

- Преимущества: натуралистичные условия, экологическая валидность
- Признаки: данные с носимых устройств (PPG, акселерометр, температура и тп), параметры использования мобильного телефона (приложения, разговоры, параметры речи из разговоров, двигательная активность)
- [G. Chittaranjan et al., 2011] Анализ параметров использования телефонов (приложения, длина звонков и смс). 117 subjects over a period of 17 months, точность Big-5 от 40% до 80% (F-score)
- [J. Staiano et al., 2012] 53 испытуемых. 8 недельный мониторинг. Анализ того, кто кому звонит и длительность разговоров. Точность от 65% до 80%.

APR and Computer Games

- Преимущества анализа игр для автоматического распознавания черт личности:
 - Наблюдение за людьми в «естественных условиях» может быть более показательным, чем самоотчеты или стандартные лабораторные задания.
 - Игры, возможно, могут вызвать естественное поведение
- Однако:
 - Игры позволяют людям выражать аспекты себя, которые они обычно скрывают
 - Игры позволяют людям экспериментировать с альтернативными личностями (например, застенчивые люди могут быть экстравертами в виртуальных мирах)

AUTOMATIC PERSONALITY PERCEPTION

Automatic Personality Perception (APP)

- APP: Предсказание личности, которую другие приписывают индивиду, на основе его наблюдаемого поведения
- Подходы APP сосредоточены в основном на **невербальном поведении** (в частности, в речи) и **социальных сетях**.
- Количество испытуемых обычно **ниже**, чем в случае с APR, поскольку сбор нескольких оценок на человека ограничивает число лиц, которые могут быть вовлечены в эксперименты.

Personality Perception from Paralanguage

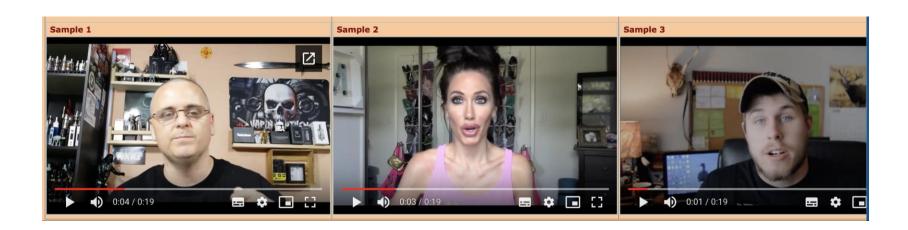
- По **речевым признакам** (просодия, вокализация, fillers) предсказывают черты личности, оцененные **внешними** наблюдателями
- На одном из датасетов провели сравнение APR и APP [см Viciarelli et al., 2014]
 - Точность получилась выше для APP— вероятно потому, что внешним наблюдателям и машине доступен один и тот же источник информации (cues)
 - В случае с APR у человека есть доступ к «внутренней» информации (например, к личностной истории)

Personality Perception and Nonverbal Behavior

- Весь спектр невербального поведения: речевое (просодия, интонации и тп), лицевые экспрессии, жесты и тп
- [J. Biel and D. Gatica-Perez, 2012] 442 YouTube видо блогов (vlogs) 50-70 сек. 5 внешних оценщиков. RMSE 0.7-1.0
- [R. Srivastava et al., 2012] Из фильмов нарезали 3907 клипов, внешними аннотаторами оценили черты личности (Big-5) для 50 персонажей. Точность от 60% до 85%

Personality Perception and Nonverbal Behavior

- 2016 First Impressions challenge
- 10000 videos (видео из интернета, отзывы, влоги и тд) http://chalearnlap.cvc.uab.es/dataset/20/description/#
- 84 аннотатора разметили личностные черты (на Amazon MTurk)
- Важно для Job interview



Personality Perception and Social Media

- В отличие от APR, проблеме восприятия личности через социальные сети уделяется лишь ограниченное внимание.
- В работах [S. Fitzgerald et al., 2009], [D. Evans et al., 2008] исследовано соответствие между фактической личностью пользователей социальных сетей и характеристиками, которые им приписываются, когда они публикуют различные типы материалов.

AUTOMATIC PERSONALITY SYNTHESIS

Speech Based Synthesis

- Эксперименты [C. Nass and S. Brave, 2005], в которых показывается, что можно синтезировать речь, которую люди-слушатели склонны воспринимать как более высокую или более низкую экстраверсию (за счет громкости, высоты основного тона, скорости речи и тп).
- Также показано, что слушатели-люди с высоким уровнем экстраверсии с большей вероятностью примут книжные рекомендации, сделанные с экстравертной речью, и наоборот.

Synthesis and Artificial Agents

- Embodied Conversational Agents: Эксперименты [M. McRorie et al., 2011], проведенные с участием 187 испытуемых, показывают, что различия во внешнем виде, поведении (движения тела, жесты и т. д.), движений лица (более или менее прямой взгляд, мимика и т. д.) и паралингвистике (длина и частота пауз и т. д.) приводят к систематическим различиям в атрибуции черт личности
- Так, например, исследуют [E. Krahmer et al., 2003] влияние комбинаций внешних черт (паттерны взгляда прямой/непрямой, стили просодии громкость, высота, движения бровей) на воспринимаемую экстраверсию агента (все черты значимо влияли)

Synthesis and Robots

- В исследованиях показывают, что людям-экстравертам больше нравится общаться с роботами, которые также демострируют экстраверсию (люди с низкой экстраверсией с роботами с низкой экстраверсией). Экстраверсия моделируется паралингвистикой (высота, громкость), вербальным контентом (выбор слов и предложений) и проксемикой (межличностная дистанция) similarity attraction effect [A. Tapus et al., 2008]
- Но этот эффект показан не во всех работах. В некоторых исследованиях показан complementarity-attraction effect (интровертам нравится робот-экстраверт и наоборот)
- Интересное наблюдение: испытуемые подходят к роботу ближе, чем к людям в аналогичных обстоятельствах [М. Walters et al., 2005]

Challenges of Personality computing

- Доступ к размеченным данным: большинство существующих работ основаны на само-опросах (напр., игроки/блогеры заполняют личностные опросники)
- Ситуация ограничивает степень, с которой могут проявляться индивидуальные отличия (напр., на красный свет светофора большво стоят, на желтый кто-то идет, кто-то стоит) -> окружение должно позволять выразить черту; некоторые черты проявляются в поведении более явно (напр., экстраверсия влияет на внешнее поведение, нейротизм на способ мышления)
- Что является Ground Truth?

Презентации по теме «Теории личности»

- 1. APR and Computer Games
- 2. APR and Keystroke dynamics
- 3. Big5 personality & facial images

Sum Up

- Personality Computing область Affective Computing
- Из многочисленных моделей личности чаще всего пользуются моделью черт (Big-5)
- Выделяют 3 основных понятия: Automatic Personality Recognition (true personality), Automatic Personality Perception (attributed personality), Automatic Personality Synthesis
- Все три понятия укладываются в модель линзы Брунсвика
- И APR, и APP работает с текстами, аудио, видео, компьютерными играми, социальными сетями и тд

Приложения, индустриальные кейсы, тренды, этические вопросы

24 AFFECTIVE COMPUTING APPLICATIONS AND USE CASES

Marketing

- Marketing communications
- Market research
- Content optimization

Customer Service

- Intelligent call routing
- Call recommendations
- Continuous improvement

Human Resources

- Recruitment
- Employee training
- Tracking employee satisfaction

Healthcare

- · Patient care
- Medical diagnosis
- Counseling

nsurance

Fraud detection

Retail

 In-store shopping experience

Driving assistance

- Safety
- Autonomous driving performance

Education

- Measuring effectiveness
- Supporting autistic children

aming

- Testing
- Adaptive games

Government

- Understanding population
- Tracking citizen reactions

Technology

• Integration with IoT

Other

Workplace design

Продукт!= технология

- Продукт готов к использованию, решает проблему/задачу, относительно прост в использовании
- **Технология** часть (основа) продукта, благодаря которой возможен функционал. Сама по себе может не решать задачу пользователя.

Продукт!= технология

- Продукт готов к использованию, решает проблему/задачу, относительно прост в использовании
- Технология часть (основа) продукта, благодаря которой возможен функционал. Сама по себе может не решать задачу пользователя.

- Продукт это то, что компания предлагает на рынке
- То, что удовлетворяет какуюлибо потребность, решает какую-либо проблему пользователя
- Имеет ценность для пользователя (может участвовать в обмене на рынке)

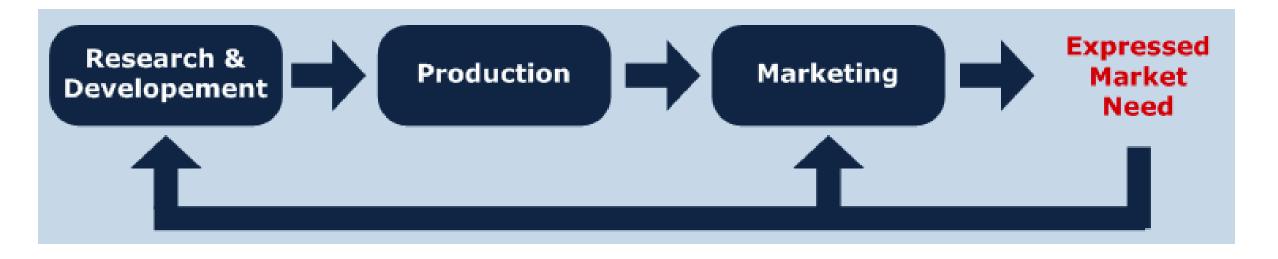
Создание продукта

- Продуктовая команда: продукт возникает на стыке технологий, маркетинга, дизайна, исследования пользователей, финансов
- Снятие рисков и неизвестностей, постепенное наращивание подтвержденного знания (тестирование гипотез)
- Делается итеративно, в тесном контакте с рынком и пользователями

Product driven (technology push)



Customer driven (market pull)



Product driven (technology push)

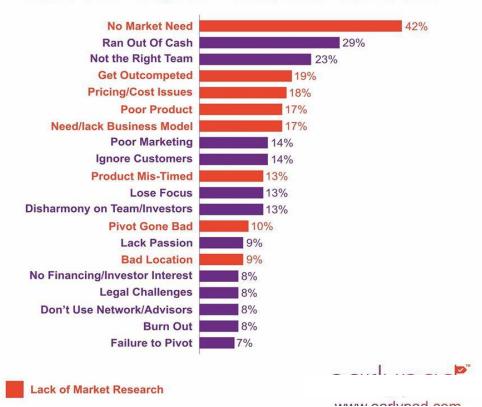
- Разработка продукта -> поиск рынка
- Предположение: отличный продукт привлечет много клиентов и денег
- **Массовый маркетинг**: предполагается, что есть много клиентов, которые ищут продукт, но пока не знают, что он есть. Цель массового маркетинга информирование как можно большего числа потенциальных клиентов.
- Уникальность продукта: клиенты не должны иметь других источников продукта.
- Позиционирование продукта: если нельзя создать уникальный продукт, нужно его позиционировать как лучший среди конкурентов (например, Coca-Cola).
- Примеры Product driven компаний: Apple, Facebook, Google

Product driven (technology push)

- Можно не угадать с идеей продукта, он окажется никому не нужен
- Разработка полноценного продукта и маркетинг - дорого

Top 20 Reasons Startups Fail

Based on an Analysis of 101 Startups Post-Mortems



Customer driven (market pull)

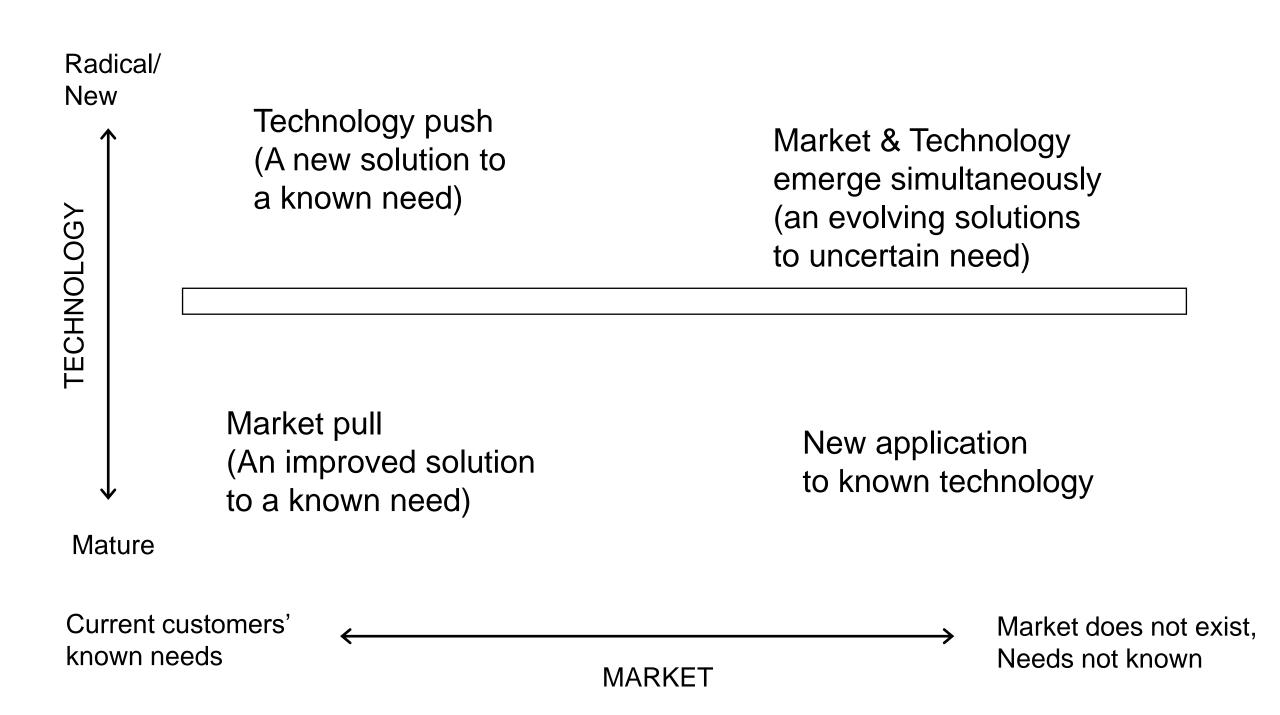
- Общение с клиентами -> создание продукта
- Упор на скорость: быстро узнавать потребности, быстро прототипировать, быстро тестировать гипотезы продуктов
- Максимизация контактов с клиентами: специализированный отдел обслуживания, сбор и анализ отзывов (через социальные сети и другие каналы)
- Примеры: Samsung

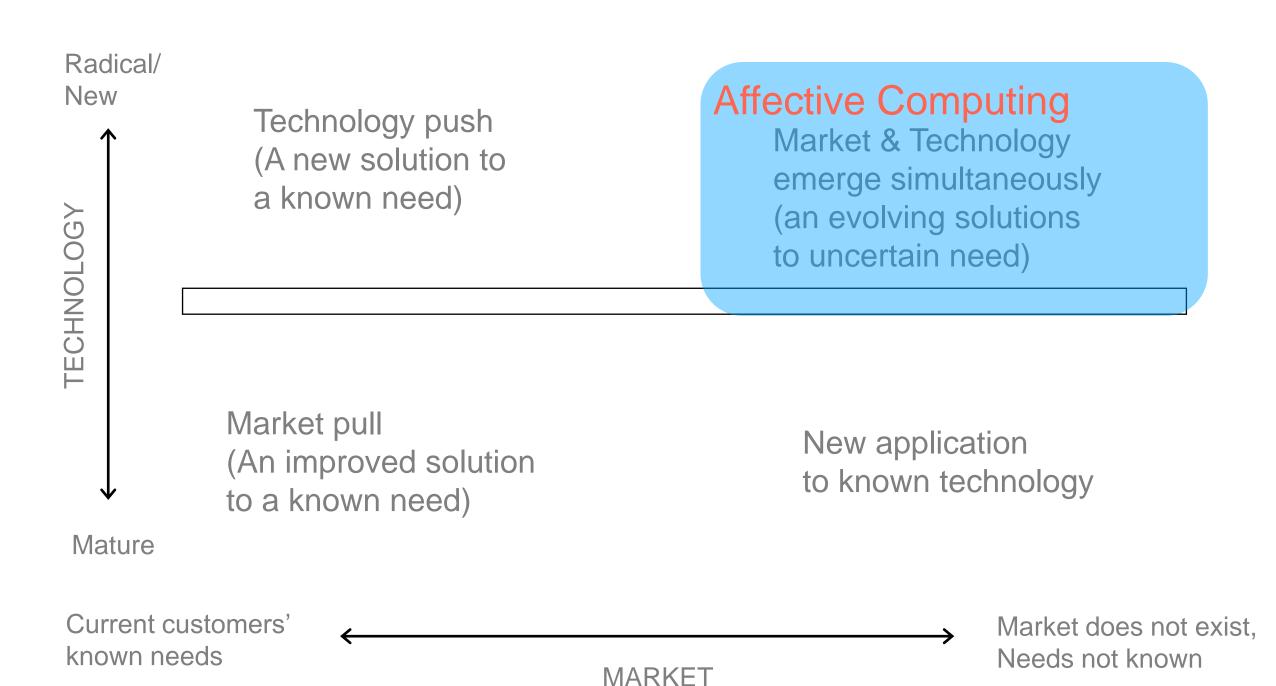
Customer driven (market pull)

- Люди не всегда знают, что хотят
- Во время интервью могут ответить одно, а действовать иначе
- Потребности можно формировать

IT'S REALLY HARD TO DESIGN PRODUCTS
BY FOCUS GROUPS. A LOT OF TIMES,
PEOPLE DON'T KNOW WHAT THEY WANT
UNTIL YOU SHOW IT TO THEM.

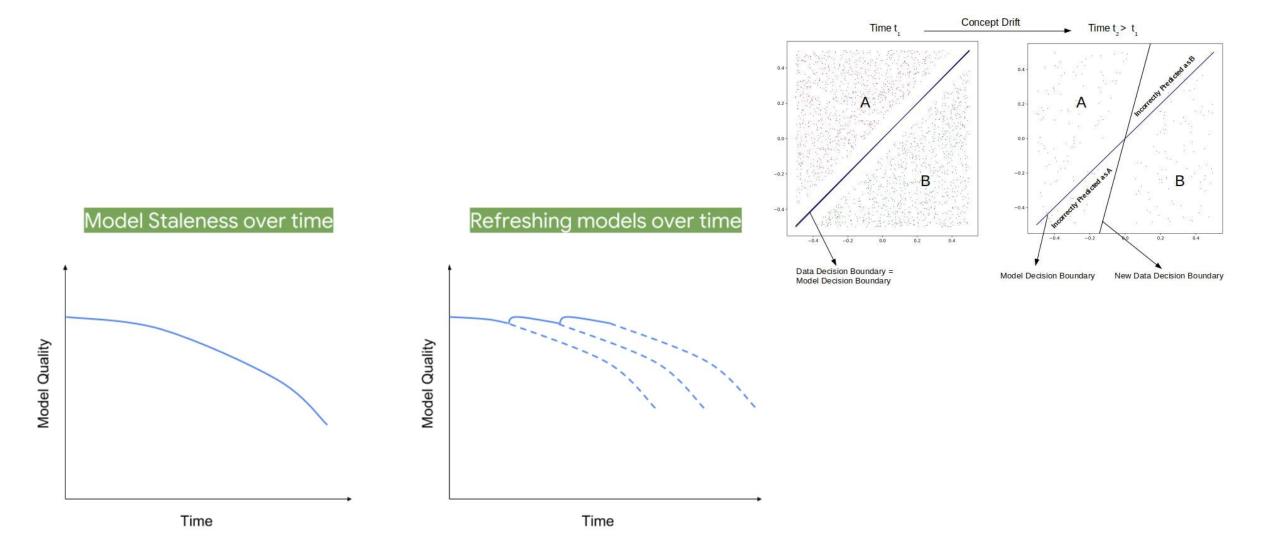
- Steve Jobs





Технологические проблемы

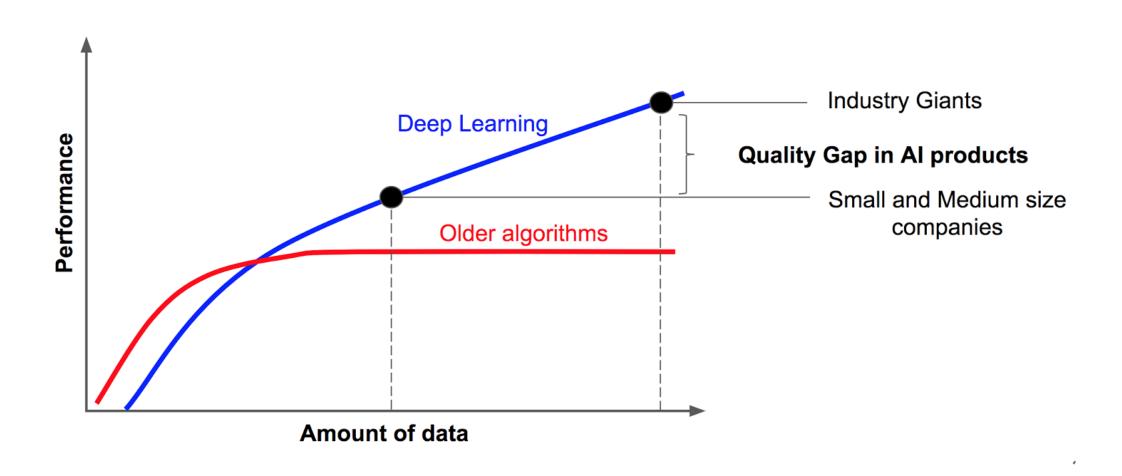
1. Модели "устаревают"



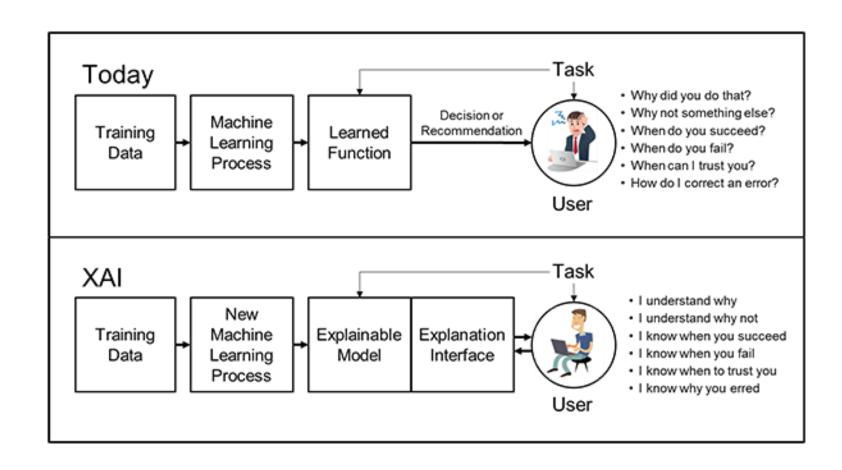
Model drift

- После того, как модель попала в prod, она начинает «устаревать».
- Concept drift статистические свойства таргета изменяются со временем
- ML алгоритм не взаимодействует с реальным миром
- Сам таргет может не измениться, но поменяется то, как результаты интерпретируют люди
- Нужно регулярно мониторить качество модели и вносить изменения

2. Много данных = дорого



3. Unexplainable results

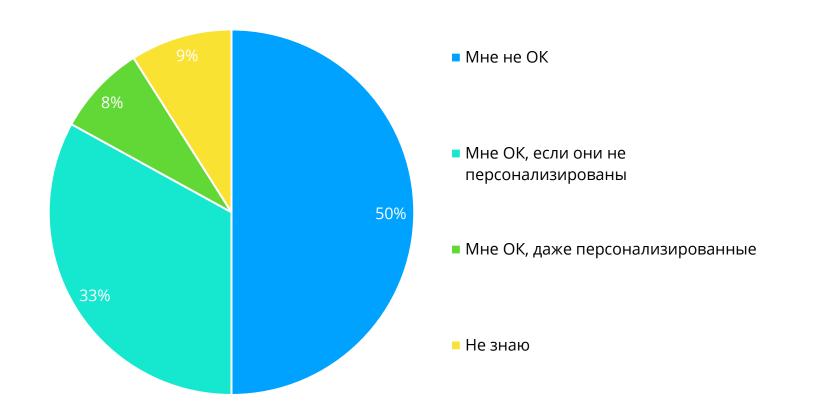


4. Biases & Unfair results

- Google Photos создает автоматические подписи для картинок. Пользователи заметили случаи, где чернокожих людей подписывали как горилл.
- Программное обеспечение камеры Nikon неправильно определило изображения некоторых азиатских людей как «моргнувших».
- Программное обеспечение для веб-камер Hewlett-Packard с трудом распознает людей с темной кожей.
- Компьютерная программа в области предупреждения преступности ошибочно приписывает высокую вероятность преступной деятельности темнокожим обвиняемым. В то же время программа ошибочно дала слишком низкие оценки риска белым подсудимым.
- Алгоритм показа рекламы в Google показывал высокооплачиваемую работу чаще мужчинам, чем женщинам.

Этические проблемы

Вопрос: рекламное агентство внедрило в камеру внешнего наблюдения распознавание эмоций, что вы испытываете?



Исследование в Англии на 2000 испытуемых — отношение к технологии считывания эмоций, 2015

https://www.ohchr.org/Documents/Issues/DigitalAge/ReportPrivacyinDigitalAge/AndrewMcStayProfessor%20of%20Digital%20Life,%20BangorUniversityWalesUK.pdf



Китайское правительство внедряет распознавание эмоций на уйгурах, 2019

https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html

nature

Explore content >

About the journal ∨

Publish with us ∨

Subscribe

nature > world view > article

WORLD VIEW | 06 April 2021

Time to regulate AI that interprets human emotions



The pandemic is being used as a pretext to push unproven artificial-intelligence tools into workplaces and schools.

Научное сообщество требует ввести регуляцию области автоматического распознавания эмоций

Nature, 2021

https://www.nature.com/articles/d41586-021-00868-5

Kate Crawford

Презентации по теме «Этика и искусственный интеллект»