

# **Лекция 4.**

## **Угрозы, риски и последствия развития искусственного интеллекта**

**Дисциплина «Цифровая культура и цифровой профессионализм»**

# Исследование АЦ при Правительстве России и ВЦИОМ (2019 г.)

## ИССЛЕДОВАНИЕ: ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ГЛАЗАМИ РОССИЯН

Жители России положительно относятся к использованию технологий искусственного интеллекта в различных сферах и согласны с тем, что государство должно поддерживать развитие этого направления.

Исследование проведено Проектным офисом национальной программы «Цифровая экономика» Аналитического центра при Правительстве Российской Федерации совместно с ВЦИОМ



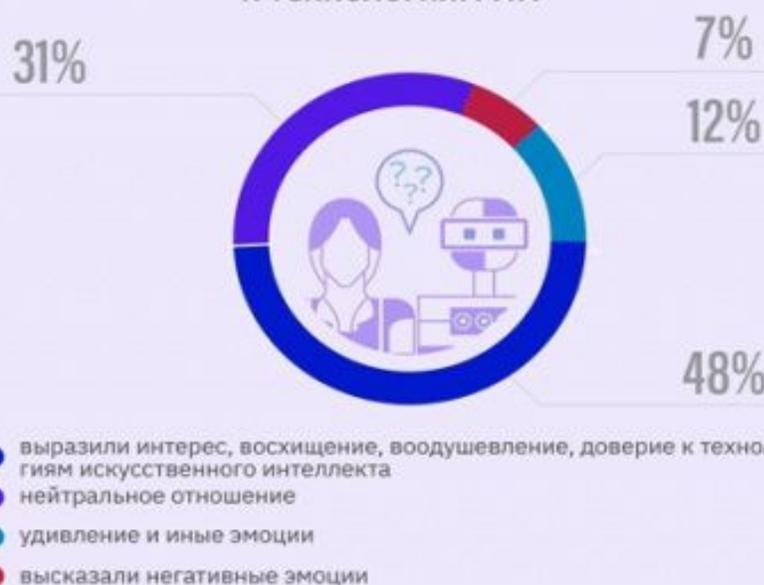
### Россияне знают о технологиях ИИ и готовы их использовать



### Положительно оценивают внедрение ИИ в сферах



### Какие эмоции испытывают россияне к технологиям ИИ



### Опасения, связанные с технологиями ИИ



### Россияне готовы применять технологии ИИ

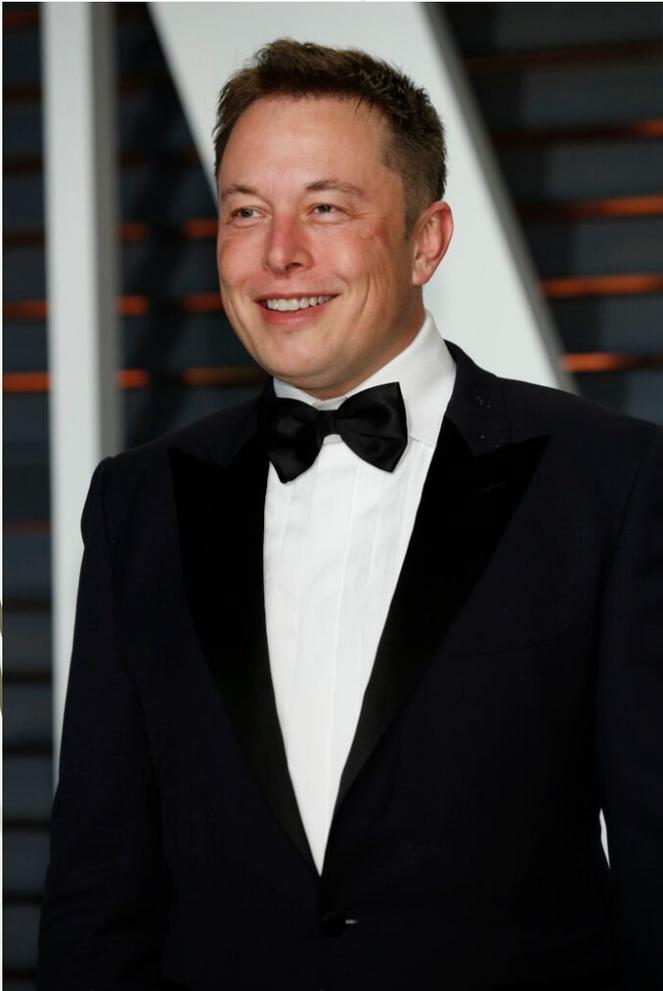


## Владимир Путин



«Мы должны ...совместными усилиями прийти к согласию в сфере регулирования, которое исключило бы потенциальные угрозы, причем с точки зрения не только военной, технологической безопасности, но и традиций, права, морали человеческого общения»

**Илон Маск** - американский предприниматель, изобретатель, инженер



□ «...ЭТО КАК «ВЫЗОВ ДЕМОНА», КОТОРОГО НЕЛЬЗЯ ПРИЗЫВАТЬ, ЕСЛИ НЕ МОЖЕТЕ ЕГО КОНТРОЛИРОВАТЬ...»

**Стивен Хокинг** - английский физик-теоретик, космолог и астрофизик, писатель, директор по научной работе Центра теоретической космологии Кембриджского университета



- «...любое дальнейшее исследование искусственного интеллекта может быть фатальной ошибкой...развитый ИИ будет обладать способностью переделывать, усовершенствовать себя»
- «Когда ИИ достигнет IQ 100 или 150, проблем не будет. Но когда это будет тысяча или 10 тысяч, что тогда? Это может быть величайшее достижение человечества. Но, в то же время его конец!»

**Билл Гейтс** - американский предприниматель и общественный деятель, один из создателей (совместно с Полом Алленом) и бывший крупнейший акционер компании Microsoft



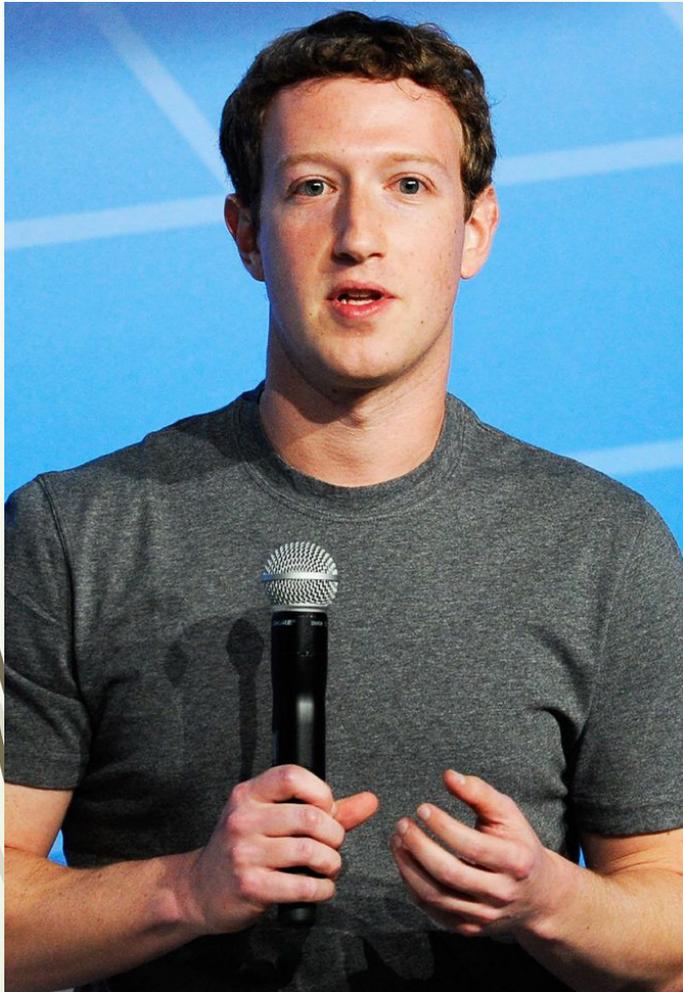
«...устройства с искусственным интеллектом поначалу будут простыми. Но со временем они будут узнавать все больше о нас и о нашем мире. Они будут становиться все более могущественными и интеллектуальными, чем человечество...»

**Эндрю Ын** - американский учёный в области информатики, доцент Стэнфордского университета, исследователь робототехники и машинного обучения



- говоря о проблеме безопасности ИИ, сравнивает её с проблемой перенаселения Марса:
- «... конечно, мы когда-нибудь колонизируем Марс, и тогда, возможно, в какой-то момент там возникнет проблема перенаселения. Но стоит ли заниматься ей сегодня?»

**Марк Цукерберг** - американский программист, предприниматель в области интернет-технологий



«Искусственный интеллект сделает в будущем нашу жизнь лучше, а предсказывать конец света очень безответственно»

## Александр Кулешов – академик, глава Сколковского института науки и технологий



- «Компьютерные программы...облегчают работу сложных предприятий, снижают шансы возникновения ошибок и проблем»
- Дальнейшее развитие этих технологий радикально поменяет жизнь человека...»

# Вячеслав Демин - директор-координатор НИЦ «Курчатовский институт»



- «Уже сейчас учёные не до конца понимают, как происходит донастройка нейросетевых алгоритмов. При этом существуют методы их обмана, **искусственный интеллект совершает ошибки.**
- Однако **эта область, конечно, развивается»**

# Отчет «Вредоносное использование искусственного интеллекта» (2018 год) подготовили:

-специалисты Оксфордского и Кембриджского университетов,

-правозащитная организация

«Фонд электронных рубежей»,

- научная лаборатории «OpenAI»

[Read the Report](#)

[Read HTML Version](#)

[Download the Full Report \(pdf\)](#)

[Download](#)

Риски ИИ

Цифровая сфера

Политическая сфера

Физическая сфера

## About the Report

### Authors

The report was written by 26 authors from 14 institutions, spanning academia, civil society, and industry.

### Process

The report builds on a 2 day workshop held in Oxford, UK, in February 2017. More information can be found in Appendix A of the report.

### Feedback

We welcome questions and comments about the report. Please see the author page for email addresses of the corresponding authors.



# Угрозы для человечества, вызванные развитием ИИ

Угроза 1

- Автономное оружие, запрограммированная угроза

Угроза 2

- Программная зависимость, ошибки ИИ, пассивные и случайные риски

Угроза 3

- Социальные манипуляции. Чипирование

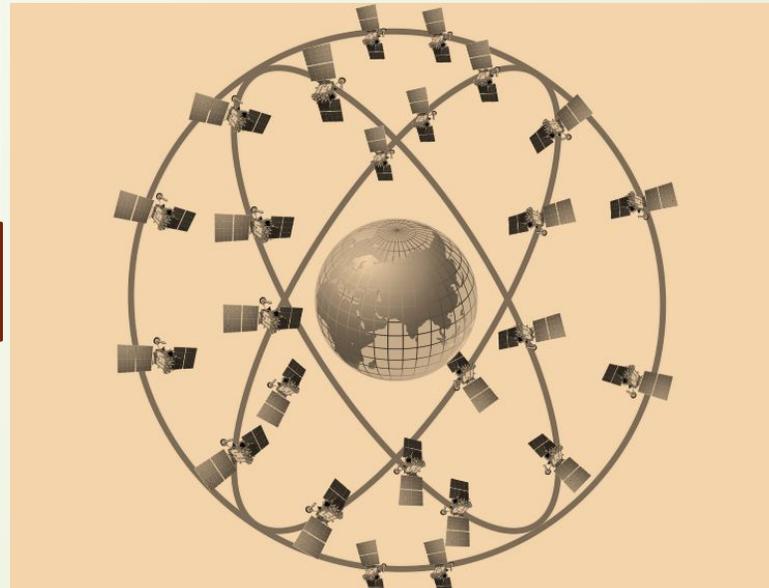
Угроза 4

- Вторжение в личную жизнь

# Угроза 1. Автономное оружие и программная враждебность



**Самостоятельные системы вооружения**



**Группировки спутников слежения**



**Военные беспилотники**



**Военные тепловизоры**



**Солдаты-роботы**

# Угроза 2. Пассивные и случайные риски, вызванные программной зависимостью, ошибки

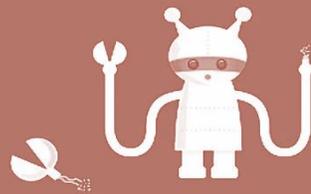
ИИ

```
1000 rows successfully bulk-copied to host-file. Total received: 78000
1000 rows successfully bulk-copied to host-file. Total received: 79000
1000 rows successfully bulk-copied to host-file. Total received: 80000
1000 rows successfully bulk-copied to host-file. Total received: 81000
1000 rows successfully bulk-copied to host-file. Total received: 82000
1000 rows successfully bulk-copied to host-file. Total received: 83000
1000 rows successfully bulk-copied to host-file. Total received: 84000
1000 rows successfully bulk-copied to host-file. Total received: 85000
1000 rows successfully bulk-copied to host-file. Total received: 86000
1000 rows successfully bulk-copied to host-file. Total received: 87000
1000 rows successfully bulk-copied to host-file. Total received: 88000
1000 rows successfully bulk-copied to host-file. Total received: 89000
1000 rows successfully bulk-copied to host-file. Total received: 90000
1000 rows successfully bulk-copied to host-file. Total received: 91000
1000 rows successfully bulk-copied to host-file. Total received: 92000
1000 rows successfully bulk-copied to host-file. Total received: 93000
1000 rows successfully bulk-copied to host-file. Total received: 94000
1000 rows successfully bulk-copied to host-file. Total received: 95000
1000 rows successfully bulk-copied to host-file. Total received: 96000
1000 rows successfully bulk-copied to host-file. Total received: 97000
1000 rows successfully bulk-copied to host-file. Total received: 98000
1000 rows successfully bulk-copied to host-file. Total received: 99000
1000 rows successfully bulk-copied to host-file. Total received: 100000
100884 rows copied.
Network packet size (bytes): 4096
Clock Time (ms.) Total : 5369125 Average : (18.72 rows per sec.)
C:\SAS\ICONSULTING\DWH_ANPAL_PACKAGE_20190603_v1.0.0_
```

**Неполные, неточные данные**



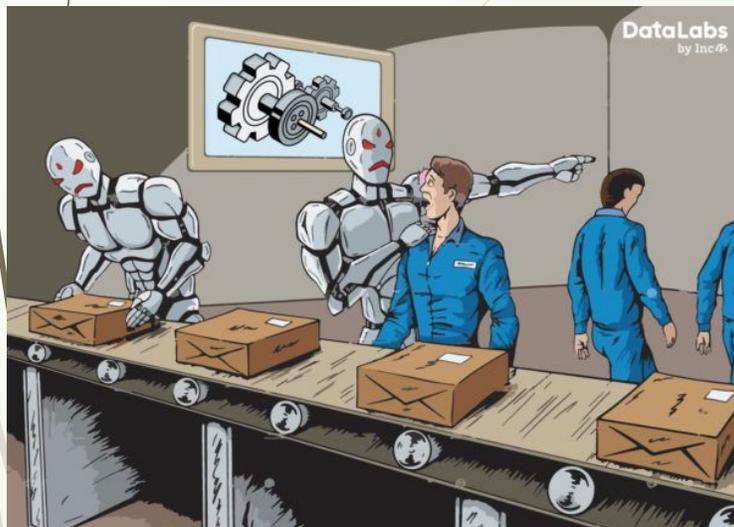
**Отсутствие установок на  
обеспечение безопасности  
человека**



**Что-то пошло не так.**

**Сбой в системе**

# Угроза 3. Социальное манипулирование

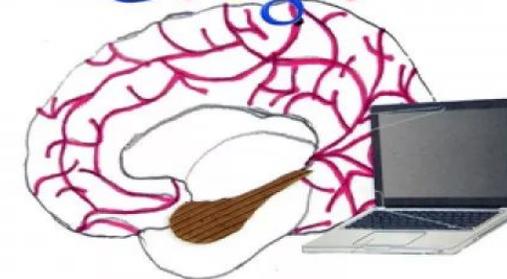


Потеря рабочих мест



Боты-манипуляторы

The Google Effect



Цифровая амнезия



Чипирование



«Фэйковая» реальность

# Некоторые факты о чипировании...

- Швеция — страна, где больше всего развито чипирование. Такая возможность зафиксирована законодательно
- В 2019 году в мире насчитывалось около 12 тысяч человек, которые добровольно пошли на вживление чипа
- В России около 200 человек чипированы
- Чипы не требуют подзарядки
- Все микросхемы помещены в крошечные капсулы из стекла, поэтому не причиняют вреда человеку
- Стоимость внедрения RFID-чипа составляет около \$30

# Шприц с иглой-штифтом для введения RFID-чипа

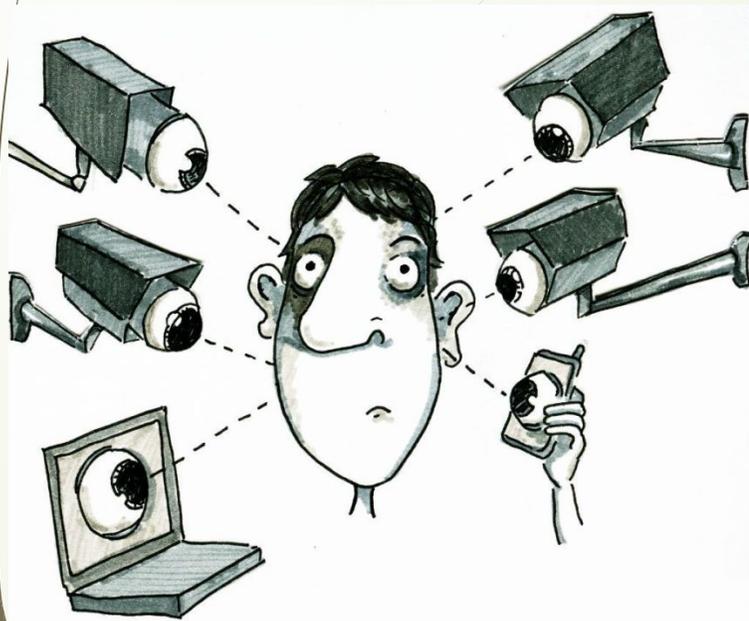


# Предполагаемые угрозы чипирования

Воздействуя на чип, можно

- оказывать серьезное влияние на нервную систему человека
- подвергать его настоящему психотеррору,
- удаленно управляя его настроением, самочувствием, состоянием психики и здоровьем внутренних органов, влияя на мыслительные процессы, вызывая зрительные, слуховые, вкусовые галлюцинации, различные боли, и т. п.

# Угроза 4. Вторжение ИИ в личную жизнь



Тотальный контроль



Манипулирование  
личными данными



Подделка личных данных

Повсеместный цифровой контроль



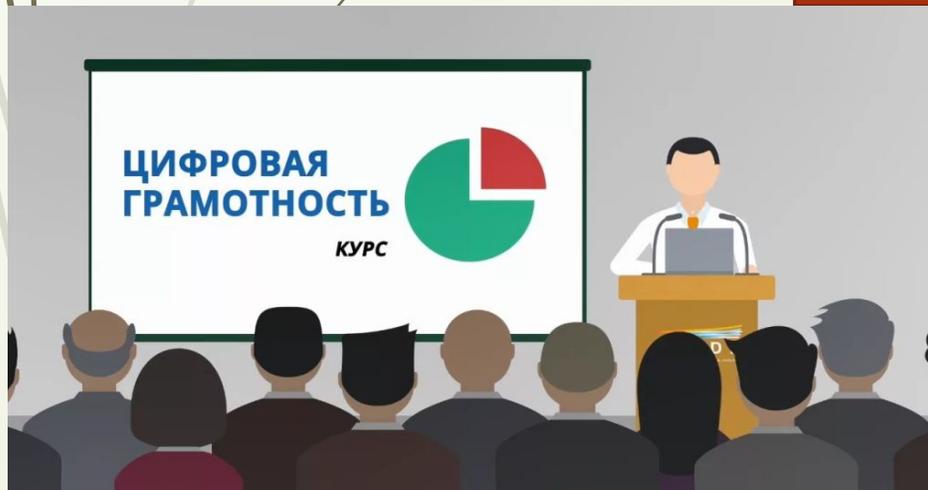
**Законодательная база ИИ**



**Регламенты безопасности**



**Подготовка  
профессионалов  
кибербезопасности**



**Цифровая грамотность  
граждан**

**Пути устранения угроз  
искусственного  
интеллекта**

Спасибо за внимание!

