

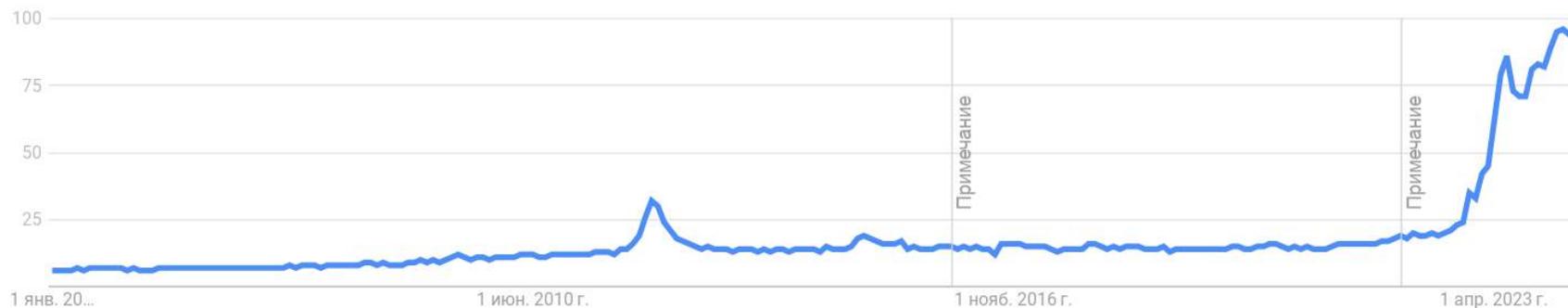
О необходимости разработки открытого стандарта маркировки контента, генерируемого нейросетями

Чайкин Павел Павлович

ООО «Базальт СПО»
отдел образовательных
проектов

ГАОУ ВО МГПУ
Департамент философии и
социальных наук



Динамика популярности ?

Искусственный интеллект — наука и инженерная деятельность, направленная на создание умных (intelligent) машин»

«Слабый» и «сильный» ИИ



ANI

Artificial Narrow Intelligence

better than humans in one specific task
e.g. autonomous driving



AGI

Artificial General Intelligence

capable like humans in every task



ASI

Artificial Super Intelligence

better than humans in every task

Sora AI



Deepfake



Генеративные ИИ и высшая школа

В среде высшего образования ИИ используется для написания как домашних работ, так и научных статей.

По данным сравнительного анализа англоязычных публикаций, опубликованных за 2022 и 2023 год и находящихся в базе данных «Dimensions» были выявлены «слова-маркеры»:

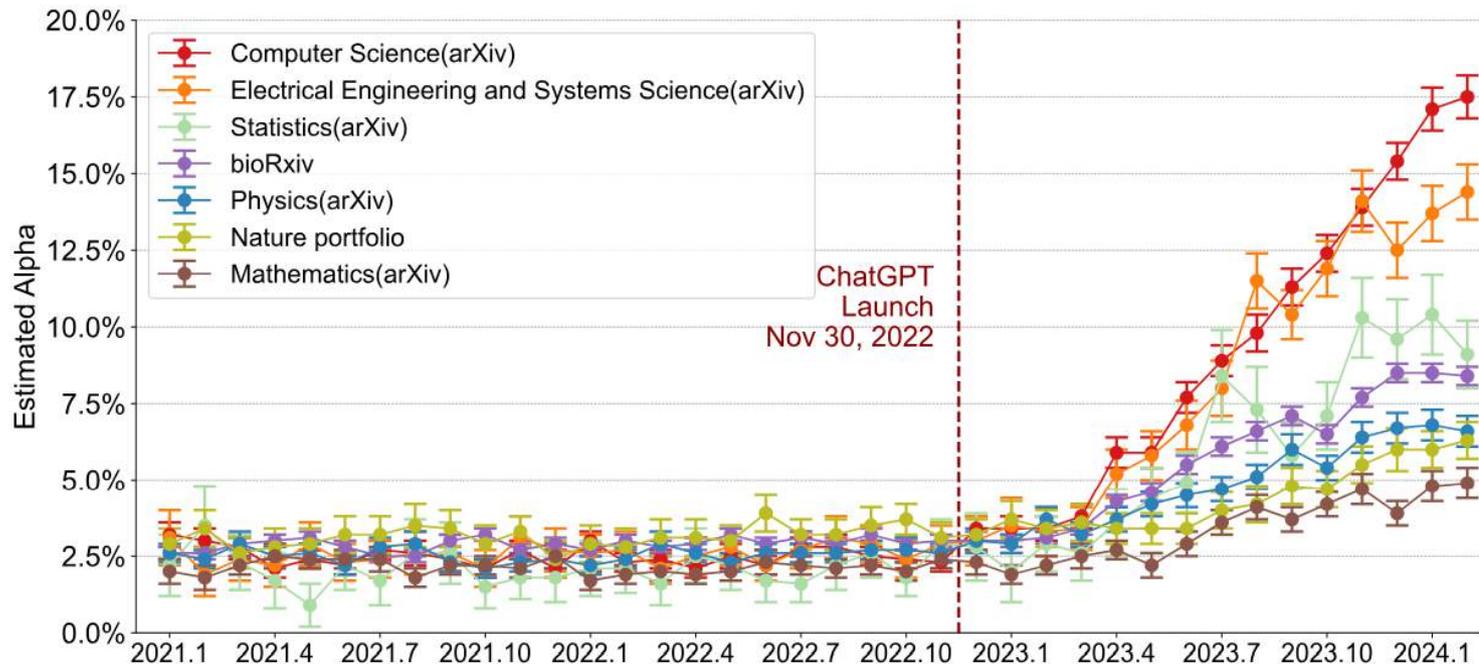
- «intricate»
- «meticulous»
- «meticulously»
- «commendable»

Рост использования составил **83,5%**.

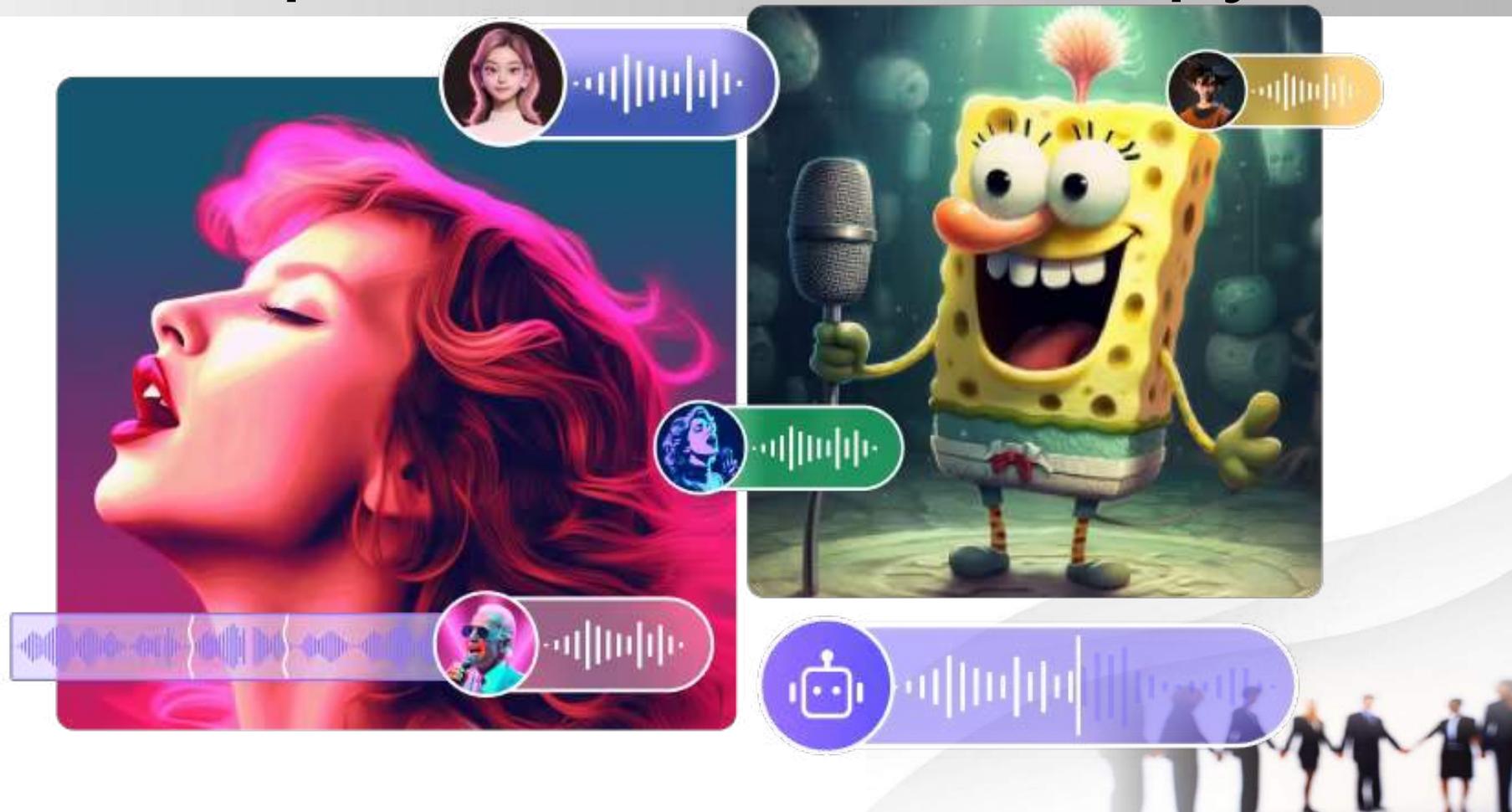


Генеративные ИИ и высшая школа

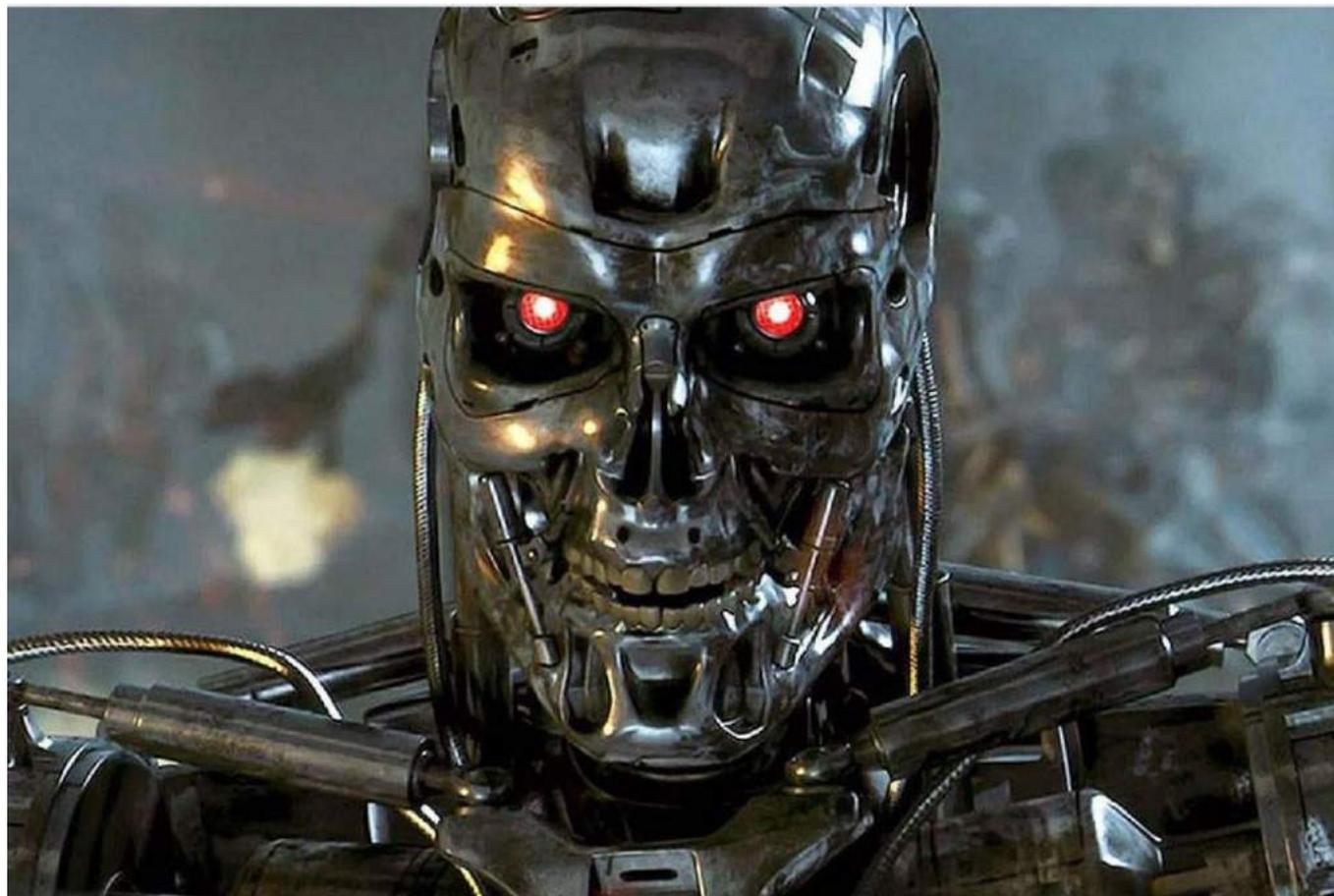
Анализ 950,965 публикаций из агрегатора «arXiv» за период с января 2020 г. по февраль 2024 г. выявило уверенный рост количества работ, включающих в себя использование БЯМ.



ИИ воспринимается как «игрушка»



...которая может иметь последствия



Влияние генеративного «фейк-контента» на воспринимаемую реальность



Marcelo Philip , Truth Tracker: Fake image of Pentagon explosion briefly sends jitters through stock market, 2023

ФБР определяет генерируемый нейросетями контент как «синтетический»



«Синтетическая реальность»



VS



ЛДОР

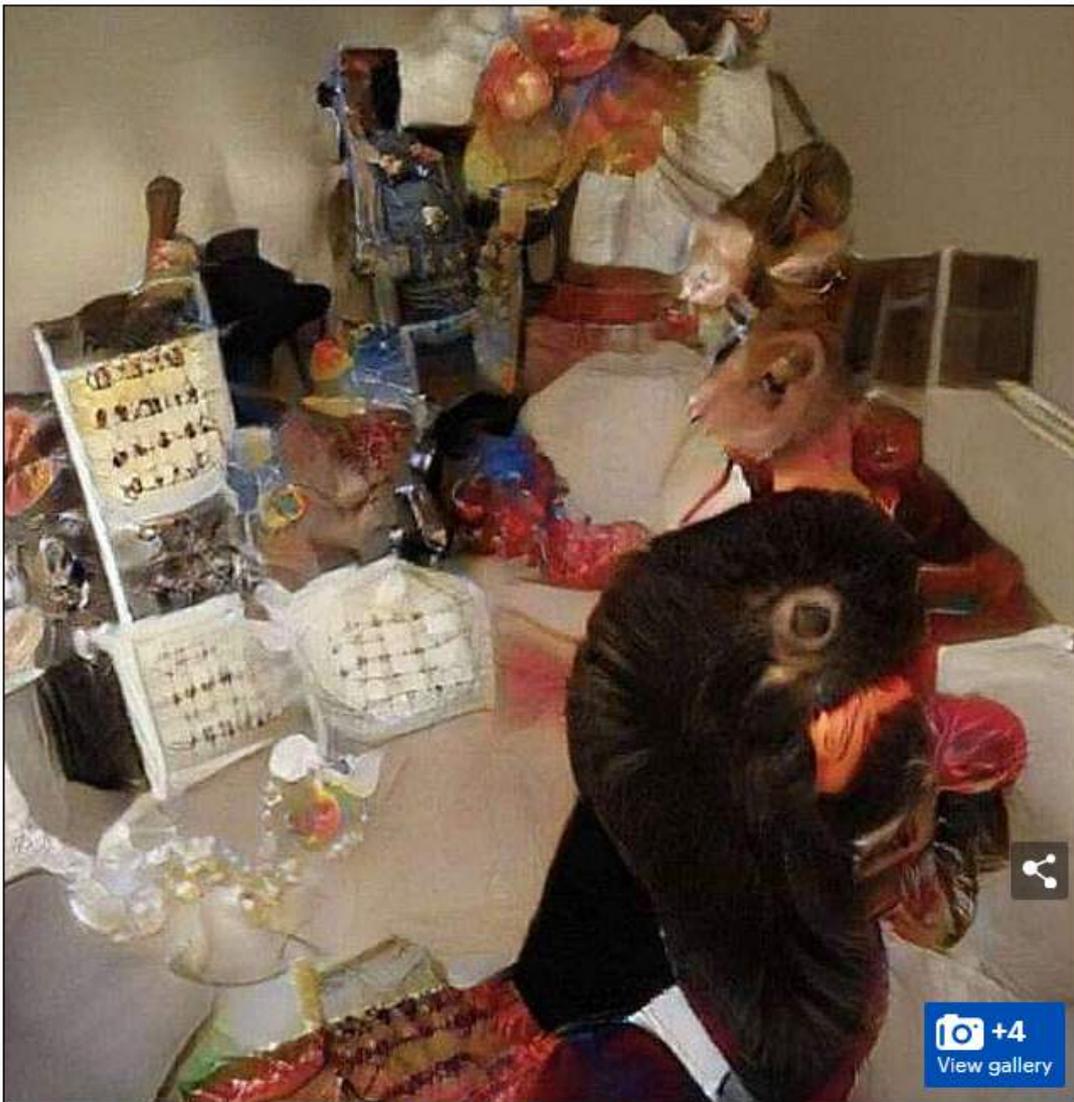
ПЕРВАЯ ЦИФРОВАЯ



ЛДОР

ПЕРВАЯ ЦИФРОВАЯ





The mysterious image was uploaded by @melipOne on Tuesday with the caption: 'Name one thing in this photo'

Искажение восприятия реальности с помощью ИИ

*Сгенерированное
изображение, на
котором невозможно
идентифицировать
объект.*



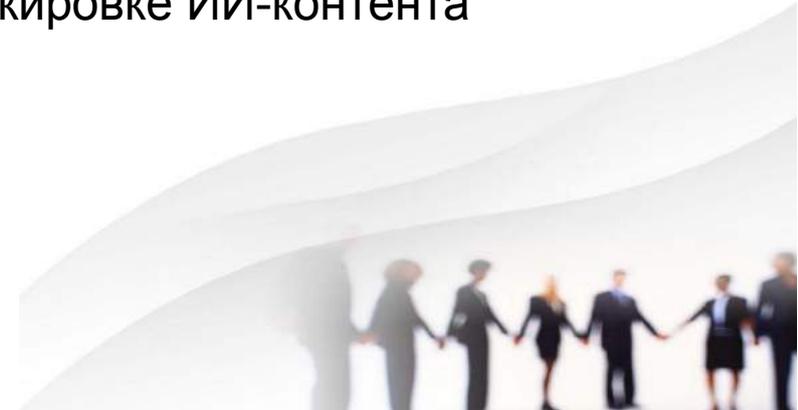
Регуляция сферы ИИ

США

- C2PA
- Маркирование ИИ при демонстрации полит. рекламы

Россия

- РТУ МИРЭА предложил в 2023 маркировать сгенерированный контент
- 2024 — Государственная дума начала работу над законопроектом о маркировке ИИ-контента



ИИ и авторское право

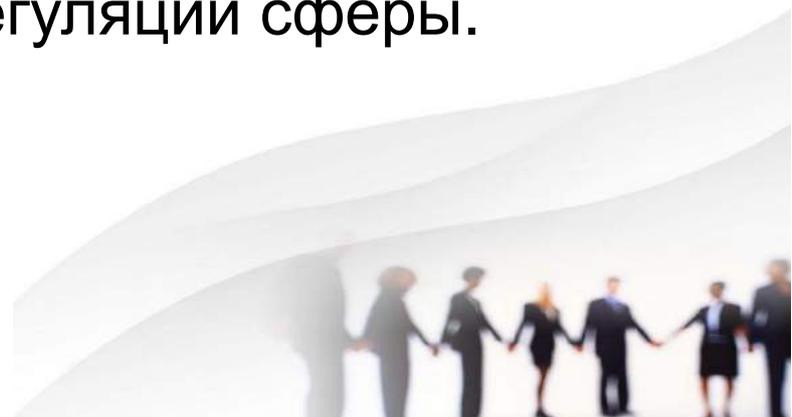


Джейсон Аллен (?) «Театр космической оперы»



Разработка стандарта

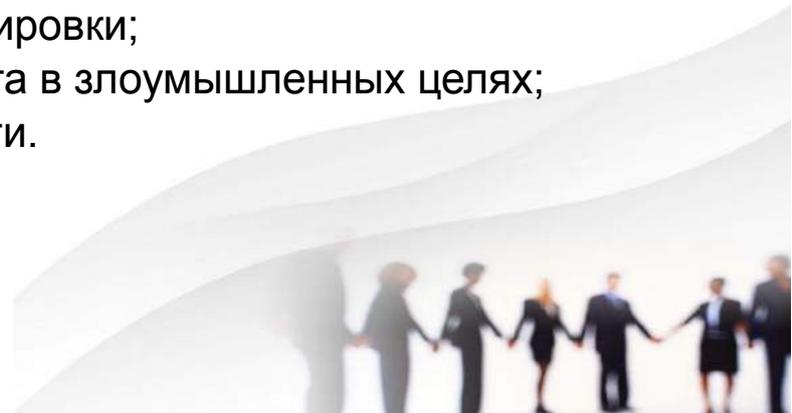
Организациям (как коммерческим, так и образовательным, например профильным организациям высшего образования), развивающим СПО необходимо принять участие в разработке таких стандартов, позволяющих проводить аудит программного кода со стороны сообщества СПО, что было бы актуально и для государства с целью обеспечения регуляции сферы.



Разработка стандарта

Выдвигаемое предположение заключается в том, что открытый стандарт маркировки сгенерированного контента должен в первую очередь:

- Обеспечить открытость процесса маркировки и критериев контента, подпадающего под маркировку;
- Защищать авторские права лиц, чьи произведения использовались при тренировке моделей;
- Позволить сообществу свободного программного обеспечения принимать участие в разработке и развитии открытого стандарта маркировки сгенерированного контента, а также проводить аудит программного кода средств маркировки;
- Гарантировать неприменимость маркированного контента в злоумышленных целях;
- Способствовать формированию доверия общественности.



Разработка стандарта

Роль государства заключается в обеспечении и регуляции «общих правил игры» для всех находящихся в его правовом поле субъектов. Открытость позволит обеспечить добросовестность использования ИИ в производстве аудиовизуального контента, в работе над научными публикациями и поможет обеспечить инфраструктуру общественной безопасности и доверительного отношения к развивающимся информационным технологиям.



Контакты



Эл. почта:
chaikinPP@yandex.ru



Вместо послесловия

- <https://youtu.be/JoAAuWNTB48?si=EEOzH9yVhHEmhniB>

