

МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

Департамент философии

РЕФЕРАТ ПО ИСТОРИИ НАУКИ

Проблемы этики искусственного интеллекта

Аспирант — Юсупов Идрис Фаргатович

Научный руководитель аспиранта _____ Бурцев М.А.

Преподаватель департамента философии — Храмов О.С.

Москва 2017

Содержание

Введение	3
Краткий экскурс в историю искусственного интеллекта	5
Этические проблемы связанные с ИИ	10
Направления в этике	10
Технологическая безработица	11
Проблема вагонетки	13
Этика больших данных	14
Этические нормы разработки и использования ИИ	16
Заключение	17
Список источников	18

Введение

Если взглянуть на человечество в исторической перспективе, то можно увидеть, что все время происходит рост качества жизни. Например, такие показатели как средняя продолжительность жизни, ВВП на душу населения, средний уровень образованности сильно выросли во всем мире за последние 200 лет [11]. Такое улучшение было бы невозможно без интеллектуальной способности человека. Именно это свойство дало ему возможность совершать научные открытия и применять научные знания для решения практических задач.

Человек - самое умное существо на планете. Но что если когда-нибудь появится разум превосходящий человека?

На текущий момент стремительно развивается наука и технология создания интеллектуальных машин и программ - т.н. искусственный интеллект (ИИ). В этой области работают ведущие исследовательские группы (Stanford, CMU, MIT, ...) и крупнейшие компании (Google, Facebook, Amazon, ...). Уже сейчас есть программы, которые решают некоторые задачи лучше человека. Например, в 1997 году компания IBM создала программу Deep Blue, которая обыграла гроссмейстера Каспарова. На сегодняшний день человеку не под силу обыграть компьютерную программу в шахматы. Программная реализация Deep Blue в основе своей состоит из алгоритмов и математических моделей, которые подходят только под шахматы. Другими словами перенести опыт, например на игру го, не получится.

В 2014 году Бостром в своей книге “Искусственный интеллект” [2, стр. 35] писал, что программы превосходящие человека в игре го появятся приблизительно через 10 лет. Но уже в 2016 году подразделение компании Google - DeepMind, представила компьютерную программу AlphaGo, которая выиграла со счетом 4-1 одного из сильнейших игроков в го. В отличие от Deep Blue, в AlphaGo были использованы методы машинного обучения, ключевая

особенность которых заключается в том, что они “обучаются” совершать заданные действия на основе заданных данных. Другими словами методы используемые в AlphaGo, можно использовать не только для игры в го, но и для других задач [3].

Методы машинного обучения сейчас применяются повсеместно - в поисковых системах для ранжирования поисковой выдачи пользователю; в разработке автономных автомобилей; в распознавании и генерации изображений; в диалоговых агентах и т.д. Все эти примеры относятся к т.н. слабому искусственному интеллекту - когда компьютерная программа может решать какую-то определенную (не очень сложную) проблему, которую может решить человек. Сильный искусственный интеллект - это уже мыслящая и осознающая себя машина, которая может решать любые задачи, которые под силу человеку. Также некоторые исследователи выделяют сверхразум - искусственный интеллект, который на порядки превосходит интеллектуальную способность человека [2].

В связи с развитием искусственного интеллекта, возникает множество этических проблем. Уже сейчас, в эру слабого ИИ, большое количество людей, которые занимаются низкоквалифицированным трудом, рискует остаться без работы.

Компании применяют огромные объемы данных для обучения ИИ. Это позволяет бизнесу достигать своих целей, получать большую прибыль, но при этом возникает целый ряд негативных последствий. При использовании больших данных встает вопрос о конфиденциальности пользовательской информации, возникает проблема черного ящика.

Еще одной важной этической проблемой является выбор слабым искусственным интеллектом действия в критической ситуации. Например, автономный автомобиль может оказаться в такой ситуации на дороге, где он должен выбрать - сбить одного или пятерых.

Область искусственного интеллекта очень быстро развивается, возможно скоро появится интеллект сопоставимый с человеческим или превосходящим его. Это ставит человечество в ситуацию в которой оно никогда не существовало. Какие риски несет эта ситуация, какие новые возможности открывает перед человечеством? Даже сейчас, в эру слабого ИИ, имеются различные этические проблемы, такие как технологическая безработица, этика больших данных и другие.

Анализу различных этических проблем связанных с ИИ посвящается данная работа.

Краткий экскурс в историю искусственного интеллекта

Говоря о проблемах этики ИИ нельзя не рассказать про историю ИИ, т.к. они неразрывно связаны.

Ада Лавлейс (1815-1852) - дочь английского поэта Байрона считается первым в мире программистом. Она изучала статью об аналитической машине ученого Чарльза Бэббиджа (1791-1871). Эта машина фактически является прообразом современных компьютеров. К заслугам Лавлейс относится то, что она описала схему вычисления чисел Бернулли на такой машине. Т.е. она придумала первый в мире алгоритм. Леди Лавлейс не сомневалась, что на такой машине можно будет производить не только математические вычисления, но и использовать её для создания музыки, картин и прочих видов искусства.

Леди Лавлейс не верила в то, что машина может мыслить. Она считала, что аналитическая машина может только точно выполнять инструкции и поэтому не может создать что-то новое, иметь свободу воли и сознание. Столетие спустя основоположник информатики Алан Тьюринг (1912-1954) назвал это утверждение “Возражение леди Лавлейс” [1, стр. 43].

Тьюринг в своей книге “Могут ли машины мыслить?” ответил на “Возражение леди Лавлейс” следующим образом: он выдвинул аргумент в

пользу программы, которая сама себя дописывает. Таким образом выходит, что программа - это не фиксированный набор инструкции заданный человеком. Он задает вопрос: “Не форма ли это обучения?”. Тьюринг говорит, что не нужно сразу писать программу, которая должна имитировать мышление взрослого человека. Надо написать программу, которая имитирует мышление ребенка с соответствующим процессом обучения, чтобы в итоге программа эволюционировала до взрослого человека [1, стр. 44].

Алан Тьюринг в 1950 году выдвинул одну из широко используемых сейчас проверок на то, что машина может мыслить - т.н. игра в имитацию или тест Тьюринга. Дано две сущности А и В, где А - это человек, В - человек или компьютерная программа. А ведет диалог с В с помощью компьютера, причем А не знает, что скрывается за В. По прошествии некоторого времени А должен дать ответ является ли В человеком или компьютерной программой [1, стр. 19].

В 1980 году американский философ Джон Сёрл (род. 1932) предложил мысленный эксперимент под названием “Китайская комната”, который является критикой теста Тьюринга и концепции сильного искусственного интеллекта. Представим себе комнату, где находится Джон Сёрл, знающий только английский язык. Человек вне комнаты - наблюдатель, передает через щель вопрос на китайском языке. Сёрл должен написать правильный ответ на вопрос и направить его в то же окно, причем ответ должен быть на китайском. В комнате имеются китайские иероглифы с помощью которых можно составить правильный ответ, если расположить их в правильном порядке. Джону доступна книга с инструкциями по манипуляции иероглифами, вида “возьми иероглиф Х и положи его рядом с иероглифом Y”. Причем в книге отсутствует информация о значениях иероглифов. Поэтому Джон Сёрл, для получения правильного ответа, просто следует инструкциям из книги подобно компьютеру и передает его в щель. Наблюдатель будет задавать осмысленный вопрос и

получать осмысленный ответ. При этом Сёрл, хоть и дает правильный ответ, не понимает ни вопроса, ни ответа [10].

Джон Сёрл своим экспериментом показывает, что даже если программа проходит тест Тьюринга - это еще не значит, что она способна мыслить, возможно, она только научилась расставлять буквы, чтобы дать правильный ответ.

У этого эксперимента есть важный предшественник, придуманный Готфридом Лейбницем (1646-1716). Его мысленный эксперимент носит название “Мельница Лейбница”. Лейбниц предлагает представить механизм размером с мельницу, который симулирует сознание. Внутри такой мельницы будет множество взаимосвязанных элементов. В совокупности все эти элементы представляют собой сознание. Но понять, что же происходит в текущий момент времени внутри сознания, смотря на эти элементы, не получится, т.к. наблюдатель будет видеть шестеренки, рычаги и т.п. Этот мысленный эксперимент показывает, что одних физических элементов (или программных инструкций) недостаточно, чтобы понять как происходит мышление [10].

Вернемся в середину 20 века. Летом 1956 года на Дартмутской конференции была предложена новая область науки под названием “Искусственный интеллект” (ИИ). В этой конференции принимали участие такие ученые как Марвин Мински и Клод Шеннон. После нее началась “Золотая эра” ИИ (1956-1974). В это время был реализован первый диалоговый агент под названием ELIZA, общение с которой происходило через текстовый терминал. Многие ученые были уверены, что в течении 10-20 лет будет создан ИИ уровня человека.

Несмотря на полученные результаты, например, программа “Логик теоретик” (Logic Theorist) смогла доказать многие теоремы из книги Альфреда Норта Уайтхеда и Бертрانا Рассела “Начала математики”, ожидания скорого

появления ИИ сопоставимого с человеком не оправдались [14]. Это положило конец финансированию многих проектов и наступила “Зима ИИ”. Почти все реализованные проекты не были применены на практике из-за малой производительности компьютеров того времени. Например, производительность лучшего суперкомпьютера 1976 года Cray-1 составляла 80-130 MIPS (million instructions per second - число миллионов инструкций, выполняемых процессором за одну секунду). Для сравнения, современный процессор Intel Core i5 имеет 83000 MIPS. А программа использующая технологии компьютерного зрения, которая распознает номера автомобилей, может требовать от 10000 до 1000000 MIPS. Зима ИИ продолжалась до 1980 года.

С 1980 по 1987 год интерес к области ИИ вызвали экспертные системы - компьютерные программы, использующие логический вывод, базы знаний о конкретной предметной области и таким образом, частично заменяющие узкого специалиста. Примером экспертной системы может служить программа, которая по симптомам определяет каким заболеванием болен пациент. Многие компании начали использовать экспертные системы и получали на этом прибыль. Но с 1987 года началась еще одна “Зима ИИ”. Появление персональных компьютеров сделало невыгодным содержание экспертных систем на специализированных компьютерах. Финансирование исследовательских групп по искусственному интеллекту также урезали.

Начиная с 1993 область искусственного интеллекта стала развиваться с новой силой. Начали появляться поисковые движки - в 1993 был запущен первый интернет-поисковик W3Catalog. В 1995 вышли более известные Yahoo и AltaVista. В 1997 году компания IBM создала шахматную программу Deep Blue, которая обыграла чемпиона мира гроссмейстера Каспарова.

Можно сказать, что в это время формировались различные дисциплины, такие как компьютерная лингвистика, компьютерное зрение, автоматическое распознавание голоса, анализ данных и другие, которые в своей совокупности

представляют собой область искусственного интеллекта. Исследовательские группы публикуют и продолжают публиковать различные наборы данных на которых следует проверять качество работы методов из области ИИ. Например, в 1998 выпущен MNIST - набор данных содержащих изображения рукописных цифр. В 2003 году выпущен CoNLL 2003 - набор данных содержащих тексты с отмеченными на них именами собственными (персоны, организации, локации и другое).

У многих компаний и в Интернете накопился огромный массив данных за последние 27 лет. В 2010-х большую популярность набрала технология работы с большими данными (Big Data). В сочетании с методами на основе глубоких нейронных сетей (Deep Learning) компании и исследовательские группы, показывают на некоторых задачах результаты лучше, чем у человека.

На сегодняшний день благодаря развитию области ИИ имеются:

- голосовые помощники, такие как Apple Siri, Microsoft Cortana, Google Now, Amazon Echo;
- победа автоматической вопросно-ответной системы IBM Watson над человеком в игре Jeopardy (Своя игра);
- победа программы DeepMind AlphaGo над человеком в игре го,
- ...

Человек всю свою историю изобретает механизмы, которые упрощают ему жизнь. Слабый искусственный интеллект - один из них. Но он кардинально отличается от предыдущих механизмов - т.к. в отличие от них ИИ в силах автоматизировать некоторые аспекты интеллектуальной деятельности человека. Поэтому кроме очевидных вопросов с уменьшением числа низкоквалифицированных вакансий, возникает целый ряд других этических проблем.

Этические проблемы связанные с ИИ

Направления в этике

Как было показано в предыдущем разделе - некоторые аспекты интеллектуальной деятельности человека реализуются в компьютерных программах. Вполне вероятно, что однажды компьютеры смогут превзойти человека в интеллекте. Но как же быть с моралью - другой отличительной особенностью человека?

Мораль - это принятые в обществе представления о хорошем и плохом, правильном и неправильном, добре и зле, а также совокупность норм поведения, вытекающих из этих представлений. Учением о морали является этика. Рассмотрим некоторые направления в этике.

Этический эгоизм - направление в этике, в котором говорится, что человек должен поступать так, чтобы его личное счастье было максимальным. При этом позиция индивида всегда ставится выше позиции общества.

Эпикуреизм - учение в этике, которое утверждает, что человек должен максимизировать свое счастье в долгосрочной перспективе и минимизировать страдание. Эпикуреизм считается предшественником классического утилитаризма.

Английский философ-моралист Джереми Бентам (1748 - 1832) в работе "Введение в принципы морали и законодательства" (1789), положил начало утилитаризму, где сказал, что человек должен жить так, чтобы максимизировать счастье и минимизировать страдание для наибольшего количества людей. Другой английский философ Джон Стюарт Милль (1806 - 1873) развил эту идею, указав на то, что некоторые виды счастья являются более ценными. Он говорил, что лучше быть неудовлетворенным как Сократ, чем довольным как свинья.

В искусственном интеллекте, чем он более развит, необходим модуль, который проводит этический анализ предлагаемого решения. В [4] предлагается следующий подход для этического анализа решения:

1. Необходимо узнать релевантные факты, касающиеся вопроса.
2. Требуется оценить насколько вопрос и возможные решения вопроса соответствуют фундаментальным принципам этики различных направлений.
3. Необходимо оценить пользу того или иного решения вопроса.
4. Необходимо оценить насколько то или иное решение вопроса влияет на максимизацию добра и минимизацию зла.

Технологическая безработица

Технологический прогресс не стоит на месте. В связи с этим у некоторых групп людей есть страх потерять работу. Пожалуй самым известным примером общественного движения порожденного таким страхом являются луддиты - участники стихийных протестов первой половины XIX века в Англии. В то время в Англии началась промышленная революция. С точки зрения луддитов - внедрение ткацких станков являлось негативным явлением, т.к. они заменяли ткачей-ремесленников и таким образом приводили к безработице.

Действительно, технологические изменения вполне способны привести к краткосрочной безработице. Но в долгосрочной перспективе такой проблемы не наблюдалось, т.к. в связи с технологическим прогрессом возникают различные компенсирующие эффекты. Например, новые машины требуют операторов для их управления, рабочих для их создания; они уменьшают цену на конечный продукт и прочее.

Мнение, что технологический прогресс не приводит к безработице было доминирующим до конца первого десятилетия XXI века. Но исследования

показали, что в некоторых секторах экономики безработица растет, хотя там и наблюдается экономический рост. По прогнозам экономистов, в ближайшие десять лет 47% рабочих мест в США будет сокращено в связи с технологическим прогрессом [5].

Главной причиной такой ситуации является стремительное развитие технологий искусственного интеллекта. Если раньше в основном автоматизировались действия различных физических механизмов, то сейчас автоматизируются интеллектуальные способности человека.

Экономисты в качестве решения проблемы технологической безработицы предлагают много разных вариантов. Несмотря на повышения уровня безработицы, происходит огромный экономический рост из-за развития технологий. Одним из вариантов решения проблемы является безусловный доход - это такой доход, которые получает любой гражданин страны, вне зависимости работает он или нет. Другим, уже существующим во многих странах мира вариантом, являются социальные выплаты.

С одной стороны, вместо работы многие люди смогут посвящать свое время тому, что им действительно важно. С другой стороны не приведет ли это к деградации населения, как в знаменитом эксперименте этолога и психолога Джона Кэлхуна (1917-1995) “Вселенная 25”? В эксперименте “Вселенная 25” для мышей были созданы утопические условия. Вначале в большой загон были помещены 4 пары мышей. Через некоторое время их количество разрослось до 2200, после которого популяция начала резко сокращаться. В конце концов вся популяция мышей погибла. В результате исследования Кэлхун пришел к выводу, что при заполнении всех социальных ролей и при превышении определенной популяции мышей возникают т.н. “молодые изгои”. Конфликт между ними и старшими особями приводит к крушению социальных связей. Общество приходит к такому состоянию, когда при малой смертности, примитивное бездеятельное поведение становится нормой. Кэлхун сказал что в

этом случае происходит “смерть в квадрате” - сначала погибает дух мышей, а затем и сами мыши [6].

Проблема вагонетки

В 1967 году английский философ Филиппа Фут (1920-2010) сформулировала мысленный эксперимент получивший название “Проблема вагонетки”:

“Тяжелая неуправляемая вагонетка несётся по рельсам. На пути её следования находятся пять человек, привязанные к рельсам сумасшедшим философом. К счастью, вы можете переключить стрелку — и тогда вагонетка поедет по другому, запасному пути. К несчастью, на запасном пути находится один человек, также привязанный к рельсам. Каковы ваши действия?” [7]

С точки зрения утилитаризма правильным решением будет переключить стрелку, чтобы пожертвовать одним человеком вместо пяти.

Все больше и больше компаний вкладываются в разработку автономных автомобилей. Известная марка автомобилей Tesla уже внедрила автопилот в свои автомобили. Согласно статистике компании Tesla, это позволило сократить количество аварий на 40 процентов [12]. Но все же, в мае 2016 произошла автокатастрофа, где погиб человек используя автопилот Tesla. В январе 2017 автопилот был признан невиновным, т.к. человек не следовал правилам пользования - он должен был следить за дорогой, держать руки на руле и быть готовым среагировать на опасную ситуацию.

Вопросом этики автопилота машины задались ученые из MIT и выпустили “Машину морали”¹ - среду для создания и проверки гипотетических ситуаций наподобие “Проблемы вагонетки”. Согласно их исследованиям, большинство людей согласны с тем, чтобы автопилот пожертвовал одним ради спасения пяти. В то же время, когда участники эксперимента ставили себя на место жертвы, они больше склонялись к автопилоту, который защитит их.

¹ <http://moralmachine.mit.edu/hl/ru>

Получается интересная дилемма: для себя человек купит автопилота-защитника, а на дорогах предпочтет автопилота следующего утилитаризму.

Как же поступить машине? Это открытый вопрос. Вполне возможно, что подобные системы должны будут получить лицензию, говорящую о том, что автопилот успешно прошел тесты по этике.

В случае аварийной ситуации с машиной-автопилотом выяснить кто виноват становится тривиальной задачей. Ведь методы ИИ реализованные в таких машинах принимают определенное решение используя большой набор сенсорных данных, применяя такое оборудование как видеокамеры, лидары, ультразвуковые датчики, радиолокационные станции, спутниковую систему навигации. Соответственно появляется возможность восстановить картину произошедшего.

Этика больших данных

Когда-то люди передавали информацию через рисунки в пещере. Потом были летописи и книги. Затем человечество овладело передачей аналоговых и электрических сигналов. Сейчас же создается более 2.5 квинтиллион байт данных в день. Это самые разные данные: gps сигналы, телефонные разговоры, посты в социальных сетях, фотографии, видео и т.д. Это и есть большие данные.

Большие данные - неотъемлемая часть существующих систем искусственного интеллекта. Автопилот для машин, поисковые системы, контекстная реклама - все они используют большие данные.

Использование больших данных позволяет найти корреляции, которые изначально не были заданы. Широко известна история с несовершеннолетней девушкой, которой прислали купоны для беременных. Решение о купонах было принято автоматической рекомендательной системой - программой, которая

предсказывает, какие объекты (купоны, фильмы, музыка и др.) будут интересны пользователю, имея определенную информацию о нем. В результате эти купоны увидели ее родители и сильно удивились узнав о положении дочери. Здесь возникает первая этическая проблема больших данных - многие использующие их методы принимают определенное решение используя огромное количество факторов, поэтому заранее узнать как именно поведет себя система очень трудозатратно.

Такая непредсказуемость работы системы с этической точки зрения, компенсируется экономической эффективностью. В конце концов цель компаний - это максимизация прибыли. Поисковые системы с контекстной рекламой отличный тому пример. Основная цель поисковых систем - выдать пользователю релевантную информацию по его поисковому запросу. Алгоритмы ранжирования поисковой выдачи используют большие данные и методы ИИ. Согласно эксперименту американского психолога Роберта Эпштейна (род. 1953), если в топе поисковой выдачи показывать положительную информацию о кандидате в президенты, то соотношение людей предпочитающих этого кандидата увеличится почти на 50%, причем 75% людей не замечало, что поисковая выдача была подставной. Этот эффект был назван “Манипуляционным эффектом поисковой системы” (Search Engine Manipulation Effect) [9].

Таким образом можно сказать, что сейчас в формировании общественного мнения по многим вопросам участвуют не только люди, которые пишут статьи или снимают фильмы, но и компьютерные программы, которые используют большие данные и методы ИИ. Хотя люди и являются авторами таких программ, но предсказание поведения такой системы очень трудоемкая задача.

Этические нормы разработки и использования ИИ

Если взглянуть на историю, то можно увидеть, что многие методы, на которых построены сегодняшние системы искусственного интеллекта, были известны уже к 80-м годам прошлого века. Ученые и футурологи того времени утверждали, что вот-вот появятся умные роботы и программы, которые кардинально изменят жизнь людей. Высокопоставленные политики и экономисты не обращали сильного внимания на такие утверждения. Сегодняшний день похож на те времена, с тем отличием, что политики и экономисты обратили свой пристальный взгляд на область ИИ. В их числе американское правительство, бывший президент США Барак Обама, множество докладов на Всемирном экономическом форуме 2017 в Давосе [15, 16].

Самой представительной конференцией в области проблем этики ИИ можно считать прошедшую в начале 2017 года в Асиломаре конференцию под названием “Полезный ИИ” (Beneficial AI). В ней выступали такие личности как Ник Бостром, Йошуа Бенжио, Дэвис Хассабис, Элон Маск и другие. По результатам конференции был предложен ряд норм, которыми следует руководствоваться при разработке систем искусственного интеллекта. Эти нормы называли Асиломарскими и они были поддержаны ведущими учеными и предпринимателями².

Один из посылов предложенных норм состоит в том, что должен быть способ понять почему система ИИ приняла то или иное решение. Другим посылом является то, что ИИ должен следовать общечеловеческим ценностям, приносить пользу, быть доступным всему человечеству; экономические блага, созданные ИИ, также должны принадлежать обществу в целом.

Действительно, если придерживаться предложенных норм, то будет продвижение в решении “Проблемы вагонетки” и “Этики больших данных”,

² <https://futureoflife.org/ai-principles/>

т.к. появляется возможность перепрограммировать ИИ. Также должен быть снижен эффект технологической безработицы, т.к. блага, которые создает ИИ, должны принадлежать обществу, а не какой-то группе лиц.

Мое мнение по решению этических проблем ИИ совпадает с предложенными Асиломарскими нормами.

Заключение

В данной работе были рассмотрены некоторые из наиболее важных на сегодняшний день этических проблем связанных с искусственным интеллектом. Первой рассмотренной этической проблемой была технологическая безработица. Второй - проблема вагонетки, в контексте принятия решения программы автопилота автомобиля в аварийной ситуации. Последняя рассмотренная проблема касалась использования больших данных в системах ИИ. Также в данной работе была описана краткая история ИИ, т.к. этические проблемы не возникли бы без соответствующих исторических предпосылок и приведены различные направления из этики, которые могут быть учтены разработчиками систем ИИ. Кроме этого, было показано современное состояние дел с проблемами этики ИИ - сейчас они обсуждаются высокопоставленными политиками и экономистами, проводятся различные форумы и предлагаются нормы, которых следует придерживаться при разработке и использовании систем ИИ.

СПИСОК ИСТОЧНИКОВ

1. Тьюринг А. и др. Могут ли машины мыслить?. – ГИФМЛ; М.; 1960, 1960.
2. Бостром Н. Искусственный интеллект: Этапы. Угрозы. Стратегии. – "Манн, Иванов и Фербер", 2015.
3. Moyer C. How Google's AlphaGo beat a Go world champion //The Atlantic, March. – 2016. – Т. 28.
4. Edgar S. L. Morality and machines: Perspectives on computer ethics. – Jones & Bartlett Learning, 2003.
5. Frey C. B., Osborne M. A. The future of employment: how susceptible are jobs to computerisation? //Technological Forecasting and Social Change. – 2017. – Т. 114. – С. 254-280.
6. Calhoun J. B. Death squared: the explosive growth and demise of a mouse population. – 1973.
7. Foot P. The problem of abortion and the doctrine of double effect. – 1967.
8. Bonnefon J. F., Shariff A., Rahwan I. The social dilemma of autonomous vehicles //Science. – 2016. – Т. 352. – №. 6293. – С. 1573-1576.
9. Epstein R., Robertson R. E. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections //Proceedings of the National Academy of Sciences. – 2015. – Т. 112. – №. 33. – С. E4512-E4521.
10. Cole, David, "The Chinese Room Argument", The Stanford Encyclopedia of Philosophy (2015), Edward N. Zalta (ed.). [Электронный ресурс], URL: <https://plato.stanford.edu/archives/win2015/entries/chinese-room>
11. Max Roser – ‘A history of global living conditions in 5 charts’, (2017). [Электронный ресурс], URL: <https://ourworldindata.org/a-history-of-global-living-conditions-in-5-charts/>
12. Andrew J. Hawkins - ‘Tesla's crash rate dropped 40 percent after Autopilot was installed’, (2017). [Электронный ресурс], URL:

<http://www.theverge.com/2017/1/19/14326258/teslas-crash-rate-dropped-40-percent-after-autopilot-was-installed-feds-say>

13. “Ответ Алана Тьюринга на возражение леди Лавлейс”, (2015).

[Электронный ресурс], URL:

<https://polymus.ru/ru/museum/news/otvet-alana-tyuringa-na-vozrazhenie-ledi-lavleys/>

14. O'Regan G. Introduction to the History of Computing: A Computing History Primer. – Springer, 2016.

15. “Barack Obama on AI, Autonomous Cars, and the Future of Humanity”, (2016). [Электронный ресурс], URL:

<https://www.wired.com/2016/10/president-obama-mit-joi-ito-interview>

16. “The National AI Research and Development Strategic Plan”, (2016).

[Электронный ресурс], URL:

https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf