



Технический комитет по стандартизации ТК164 «Искусственный интеллект»

Первое заседание. 6 августа 2019 г.

Актуальность

1. Зарождение ИИ (1940-1950)

- термин AI на собрании Д. Маккарти
- ИНС (У. Мак-Каллок, У. Питтс)
- перцептрон Розенблатта
- программы для игры в шахматы (К. Шеннон)
- тест Тьюринга
- программу для игры в шашки (А. Сэмюэл)

2. Развитие ИИ (1951-1973)

- язык LISP (Д. Маккарти)
- первые AI-программы (ELIZA, General Problem Solver, Logic Theorist, STUDENT, SAINT, SHRDLU, SNARC)
- машинный перевод (США)
- метод обратного распространения ошибки кибернетики в СССР, шахматная программа «Каисса»
- книга «Перцептроны (М. Мински, С. Пейперт), принципиальные ограничения метода сокращения финансирования работ в области ИИ со стороны DARPA

3. Первая зима ИИ (1974-1980)

Сдерживающие факторы

- отсутствие методов и алгоритмов, адекватных сложности поставленных задач
- исключение ИИ из числа национальных приоритетов, значительное сокращение финансирования

4. Возобновление работ (1981 - 1986)

- ЭВМ 5-го поколения (поддержка диалога, перевод языков, интерпретация изображений, построение причинно-следственных связей, Япония, \$850 млн)
- программа Alvey (интеллектуальные системы представления знаний, HMI, микроэлектроника, параллельные архитектуры, UK, £350 млн)
- экспертных систем, ИСПЗ, генетические алгоритмы;
- беспилотный автофургон от Mercedes
- методы и алгоритмы ИИ, в т.ч., метод обратного распространения ошибки для ИНС

5. Вторая зима (1987-1993)

Сдерживающие факторы

- ограниченные вычислительные возможности
- ограниченные возможности методов машинного обучения и анализа данных
- высокие требования к разработчикам, значительная трудоемкость создания технологий
- ограниченный объем данных для обучения и настройки систем, слабая обобщающая способность;
- значительное сокращение финансирования

6. Ренессанс ИИ (1994 г. - н.в.)

- шахматный СК DeepBlue обыграл Г. Каспарова, Google DeepMind's AlphaGo - чемпиона мира в Go, IBM Watson побеждает в Jeopardy
- Google, Yandex, Facebook;
- технологические конкурсы DARPA;
- общедоступные ИД большого объема (ImageNet)
- глобальная цифровизация: гетерогенные сенсоры непрерывно генерируют BigData
- появление GPU, высокопараллельных архитектур, общий рост производительности
- совершенствование математического аппарата: появление глубоких архитектур ИНС
- успехи в решении различных прикладных задач: машинный перевод, семантический поиск, вопросно-ответные системы, системы распознавания лиц, системы транскрибирования речи, системы беспилотного управления и др.

Сдерживающие факторы

- отсутствие доверия к системам ИИ;
- метрологические сложности;
- проблемы интероперабельности

Н.В.

Сдерживающие факторы, связанные с недостатками нормативно-технического регулирования ИИ

Отсутствие доверия к системам ИИ

Негарантированное качество работы, непрозрачная логика принятия решений, подверженность «состязательным» атакам

Ограничения на применение систем ИИ при решении ответственных задач

Метрологические сложности

Отсутствие стандартных процедур сравнения функциональных возможностей систем ИИ друг с другом и с возможностями квалифицированного человека-оператора

Проблемы оценки прикладного эффекта от применения систем ИИ

Проблемы интероперабельности

Отсутствие стандартных форматов представления данных, требований к формированию обучающих и контрольных выборок...

Проблемы эффективного встраивания систем ИИ в информационно-коммуникационную инфраструктуру

Международные организации по стандартизации



Resolution 12 – Establishment of JTC 1/SC 42, Artificial Intelligence

JTC 1 establishes a Systems Integration entity (see SD 24, Systems Integration Standardization Guidelines) in the form of a new Subcommittee 42 on Artificial Intelligence as follows:

Title: *Artificial Intelligence*

Scope:

Standardization in the area of Artificial Intelligence

1. *Serve as the focus and proponent for JTC 1's standardization program on Artificial Intelligence*
2. *Provide guidance to JTC 1, IEC, and ISO committees developing Artificial Intelligence applications*

While SC 42 will establish its own substructure at its first meeting, JTC 1 recommends that SC 42 includes the following topics:

- Foundational standards
- Computational methods
- Trustworthiness
- Societal concerns

The National Body of the United States will serve as Secretariat for SC 42. Wael Diab will serve as the Chair for SC 42.

The JTC 1 Secretariat is instructed to issue a call for membership in SC 42 and to work with ITTF to set up the committee.

JTC 1 notes that the National Body of China has expressed an interest in pursuing a Vice-chair position via the ISO Twinning process.

The National Body of China has offered to host the first meeting of JTC 1/SC 42.

JTC 1 notes that two NWIPs on the topic of artificial intelligence are currently out for JTC 1 letter ballot. If these NWIPs are approved, JTC 1 recommends that they be placed in JTC 1/SC 42.

2017.10 (Владивосток, Россия) - решение о создании технического подкомитета SC42 “Artificial Intelligence” на заседании Объединённого ТК ISO/IEC JTC 1 “Information Technologies”

- 1) 2018.04 - Пекин, КНР
- 2) 2018.10 - Саннивэйл, США
- 3) 2019.04 - Дублин, Ирландия
- 4) 2019.10 - Токио (Япония)
- 5) 2020.04 - Париж (Франция)
- 6) 2020.10 - Монреаль (Канада)

- AHG 01 “Dissemination & outreach” (распространение и охват)
- AHG 02 “Liaison with SC 38” (связь с ПК 38 «Облачные и распределённые вычисления»)
- AHG 03 “Intelligent systems engineering” (инжиниринг интеллектуальных систем)
- WG 01 “Foundational standards” (основополагающие стандарты)
- WG 02 “Big Data” (большие данные)
- WG 03 “Trustworthiness” (доверенность систем ИИ)
- WG 04 “Use cases and applications” (практические примеры и приложения)
- WG 05 “Computational approaches and computational characteristics of AI systems” (вычислительные методы и характеристики систем ИИ)
- JWG 01 with SC 40 “Governance implications of AI” (последствия ИИ для сферы

Документы, разрабатываемые в ISO/IEC JTC1 SC42

#	Наименование проекта	WG 1	WG 2	WG 3	WG 4	WG 5	JWG 1
1	WD 22989 Artificial intelligence -- Concepts and terminology						
2	WD 23053 Framework for AI Systems Using Machine Learning (ML)						
3	FDIS 20546:2019 Big data -- Overview and vocabulary						
4	NP 24300 Process management framework for Big data analytics						
5	AWI TR 20547-1 Big data reference architecture - Part 1: Framework and application process						
6	TR 20547-2:2018 Big data reference architecture - Part 2: Use cases and requirements						
7	20547-3 Big data reference architecture -- Part 3: Reference architecture						
8	TR 20547-5:2018 Big data reference architecture -- Part 5: Standards roadmap						
9	NP TR 24027 AI -- Bias in AI systems and AI aided decision making						
10	NP TR 24028 AI -- Overview of trustworthiness in Artificial Intelligence						
11	TR 24029-1 Assessment of the robustness of neural networks -- Part 1: Overview						
12	NP TR 24368 Overview of ethical and societal concerns						
13	NP 23894 Artificial Intelligence -- Risk Management						
14	NP TR 24030 Artificial Intelligence (AI) -- Use cases						
15	NP TR 24372 Overview of computational approaches for AI systems						
16	NP 38507 Governance of IT -- Governance implications of the use of artificial intelligence by organizations						

Руководители рабочих групп SC42

Рабочая группа	Руководитель	Страна
WG 01 «Foundational standards»	Paul Cotton	Канада
WG 02 «Big Data»	Wo Chang	США
WG 03 «Trustworthiness»	David Filip	Ирландия
WG 04 «Use cases and applications»	Fumihiko Murayama	Япония
WG 05 «Computational approaches and computational characteristics of AI systems»	Tangli Liu	Китай
JWG 01 with SC 40 «Governance implications of AI»	Janna Lingenfelder Gyeong-Min Kim	Германия Южная Корея

A PLAN FOR FEDERAL ENGAGEMENT IN AI STANDARDS -DRAFT FOR PUBLIC REVIEW 2-JUL-2019

U.S. LEADERSHIP IN AI:
A PLAN FOR FEDERAL ENGAGEMENT IN
DEVELOPING TECHNICAL STANDARDS
AND RELATED TOOLS

DRAFT FOR PUBLIC COMMENT

PREPARED IN RESPONSE TO EXECUTIVE ORDER 13859
SUBMITTED ON AUGUST XX, 2019



- Межотраслевые (cross-sector) «горизонтальные» стандарты: обеспечение интероперабельности, безопасности, унификация форматов данных и т.д.
- Отраслевые (sector-specific) «вертикальные» стандарты: сравнение технологий ИИ друг с другом и с заданным критериальным порогом, включая человеческие возможности
- Стандарты разрабатываются в основном организациями частного сектора на основе принципа консенсуса
- Федеральные ведомства
 - определяют требования к стандартам
 - способствуют повышению уровня технической экспертизы стандартов
 - обеспечивают включение добровольных стандартов в нормативные акты
 - способствуют использованию стандартов при организации госзакупок

Приказ Росстандарта от 25 июля 2019 г. №1732 «О создании технического комитета по стандартизации «Искусственный интеллект»



МИНИСТЕРСТВО ПРОМЫШЛЕННОСТИ И ТОРГОВЛИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ АГЕНТСТВО
ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ
(Росстандарт)

П Р И К А З

25 июля 2019 г.

№ 1732

Москва

О создании технического комитета по стандартизации «Искусственный интеллект»

В целях реализации Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации», повышения эффективности работ по стандартизации на национальном, межгосударственном и международном уровнях и по согласованию с заинтересованными организациями приказываю:

1. Создать технический комитет по стандартизации «Искусственный интеллект» (далее – технический комитет) в соответствии с настоящим приказом.

2. Закрепить за техническим комитетом объекты стандартизации в соответствии с кодами ОКПД2 после подписания соответствующих соглашений о взаимодействии с техническими комитетами по стандартизации, имеющими смежную область деятельности:

26.30 – Оборудование коммуникационное (в области искусственного интеллекта);

30 – Средства транспортные и оборудование, прочие (в области искусственного интеллекта);

58.29 – Услуги по изданию прочего программного обеспечения (в области искусственного интеллекта);

62 – Продукты программные и услуги по разработке программного обеспечения; консультационные и аналогичные услуги в области информационных технологий (в области искусственного интеллекта);

63 – Услуги в области информационных технологий (в области искусственного интеллекта);

69.10.1 – Услуги юридические (в области искусственного интеллекта);

71.12.39 – Услуги в области геологических, геофизических и взаимосвязанных изыскательных работ прочие, не включенные в другие группировки (в области искусственного интеллекта);

- Ведение дел секретариата - АО «РВК»
- Председатель ТК - Гарбук С.В. (НИУ ВШЭ)
- Ответственный секретарь ТК - Лобунов А.А. (НИЦ «Курчатовский институт»)
- Участие с правом голосования в SC42 “Artificial Intelligence” ISO/IEC JTC1 “Information Technologies”
- Выполнение функций постоянно действующего национального органа в SC42
- Организация взаимодействия с ТК: 022 «Информационные технологии», 026 «Криптографическая защита информации», 098 «Биометрия и биомониторинг», 159 «Блокчейн», 194 «Кибер-физические системы», 355 «Технологии автоматической идентификации и сбора данных», 362 «Защита информации», 461 «ИКТ в образовании», 468 «Информатизация здоровья»

TK 164 как зеркальный орган SC42

SC42

- WG 01 “Foundational standards”

- WG 02 “Big Data”

- WG 03 “Trustworthiness”

- WG 05 “Computational approaches and computational characteristics of AI systems”

- WG 04 “Use cases and applications”

TK 164

(2019)
РГ 01

«Основополагающие стандарты»

- РГ 02 «Большие данные»

- РГ 03
«Качество систем ИИ»

- РГ 04 «Прикладные технологии ИИ»

- РГ 05 «Технологии ИИ в образовании»

TK 164

(2021)
ПК 01

- ПК 02

- ПК 03

- ПК 04 - ПК ...
 - ИИ в образовании
 - ИИ на транспорте
 - ИИ в медицине
 - ИИ в сфере безопасности

Основные задачи

Унификация и стандартизация терминологии

Обеспечение методологической преемственности в области методов и алгоритмов ИИ

Обеспечение интероперабельности систем ИИ

Повышение эффективности коллективных работ по созданию систем ИИ

Основные документы

WD 22989 Artificial intelligence -- Concepts and terminology 2021.03

WD 23053 Framework for AI Systems Using Machine Learning (ML)

2022.03

Технологии интеллектуальной обработки данных. Классификация

[НИИ сертификации]

2020.12

Информационные технологии. Искусственный интеллект. Дорожная карта стандартов

2020.12

Этические аспекты прикладного применения систем искусственного интеллекта (ТО)

2021.12

Основные задачи

Унификация и стандартизация требований к процедурам и средствам хранения БД, используемых при разработке, тестировании и эксплуатации систем ИИ

Снятие нормативных барьеров, связанных с обработкой персональных данных (гарантированная деперсонафикация, управление согласиями и т.п.)

Унификация метрик, определяющих качество наборов БД

Основные стандарты

FDIS 20546:2019 Big data -- Overview and vocabulary	2019.10
NP 24300 Process management framework for Big data analytics	
AWI TR 20547-1 Big data reference architecture - Part 1: Framework and application process	2020.05
TR 20547-2:2018 Big data reference architecture - Part 2: Use cases and requirements	2019.05
20547-3 Big data reference architecture -- Part 3: Reference architecture	2020.05
TR 20547-5:2018 Big data reference architecture -- Part 5: Standards roadmap	2019.12
Гармонизированные отечественные стандарты [МГУ им.М.В.Ломоносова, ИРИО]	TBD

Основные задачи

Стандартизация общих процедур подтверждения характеристик доверия и функциональных характеристик

Стандартизация процедур подтверждения безопасности функционирования системы ИИ (отсутствие неприемлемых угроз для пользователей, третьих лиц, окружающей среды)

Стандартизация процедур подтверждения предсказуемости поведения системы ИИ при определённых условиях эксплуатации (domain)

Основные стандарты

NP TR 24027 AI -- Bias in AI systems and AI aided decision making
2021.01

NP TR 24028 AI -- Overview of trustworthiness in Artificial Intelligence 2019.08

TR 24029-1 Assessment of the robustness of neural networks -- Part 1: Overview 2019.08

NP TR 24368 Overview of ethical and societal concerns
TBD

NP 23894 Artificial Intelligence -- Risk Management TBD

Технологии интеллектуальной обработки данных.

Способы обеспечения доверия к системам с искусственным интеллектом [НИИ сертификации]
2020.12

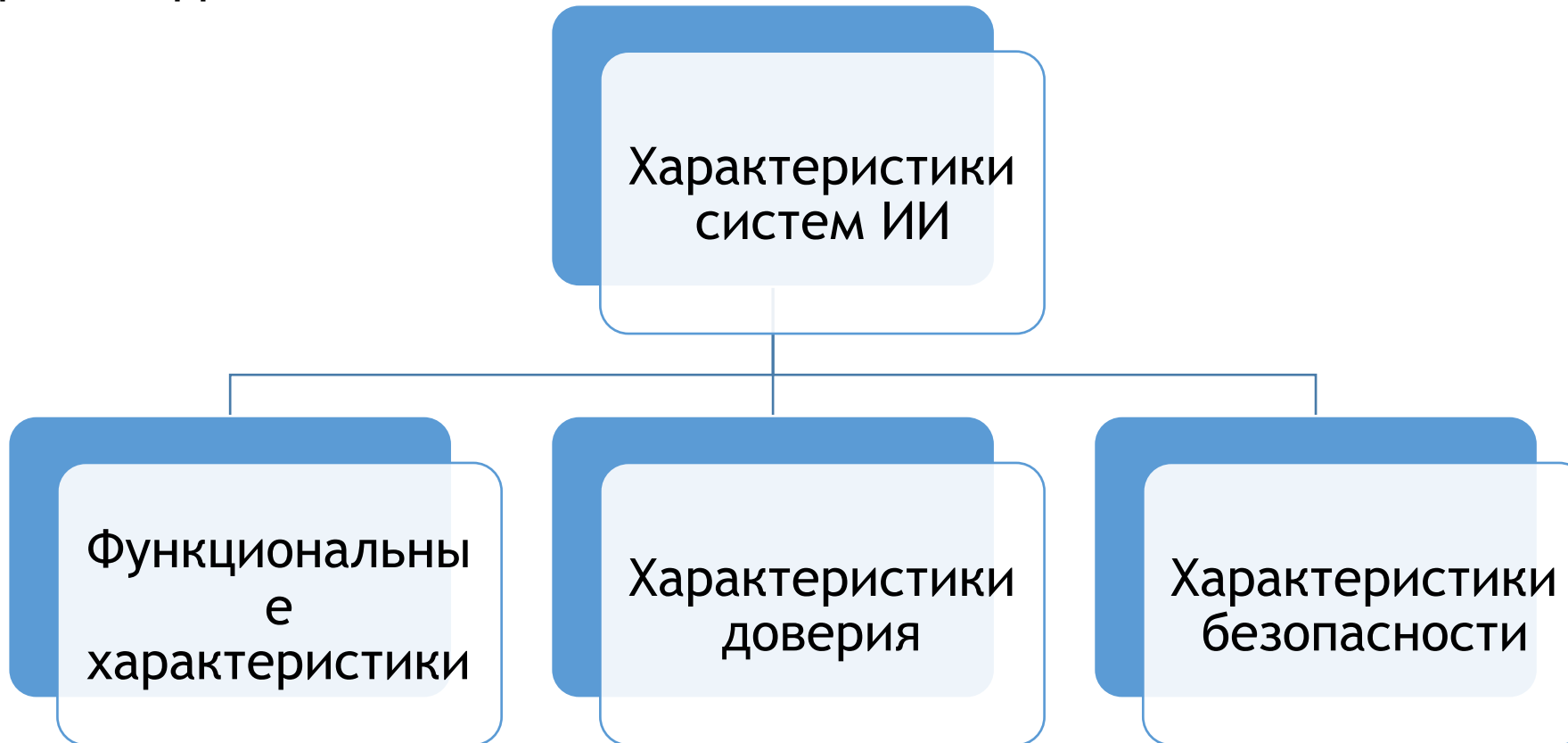
Стандарты в области оценки устойчивости нейросетевых систем ИИ к воздействию «состязательных» атак

TBD

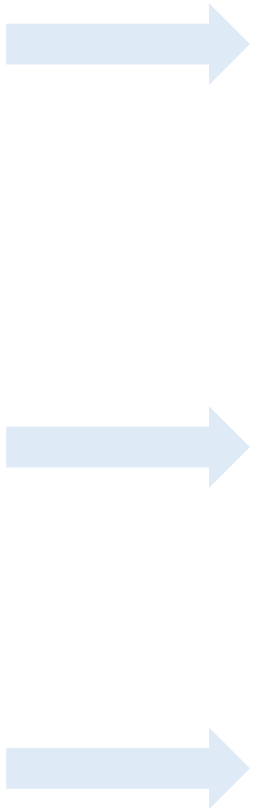
Стандарты, содержащие требования к искусственному интеллекту в защищенном исполнении

Понятие качества системы ИИ ISO 9000

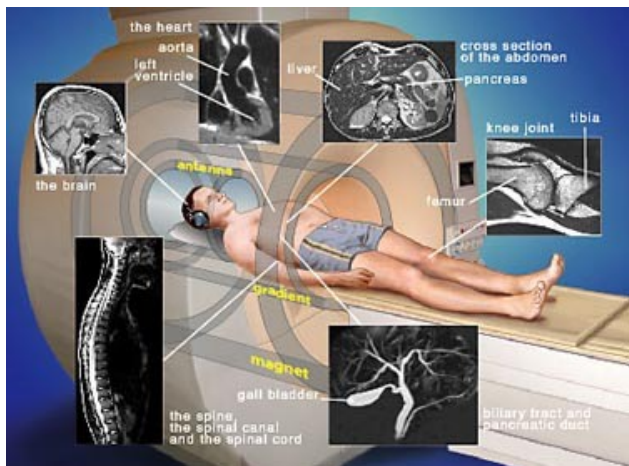
Качество – степень соответствия значимых характеристик системы ИИ требованиям, предъявляемым потребителем при решении конкретной прикладной задачи



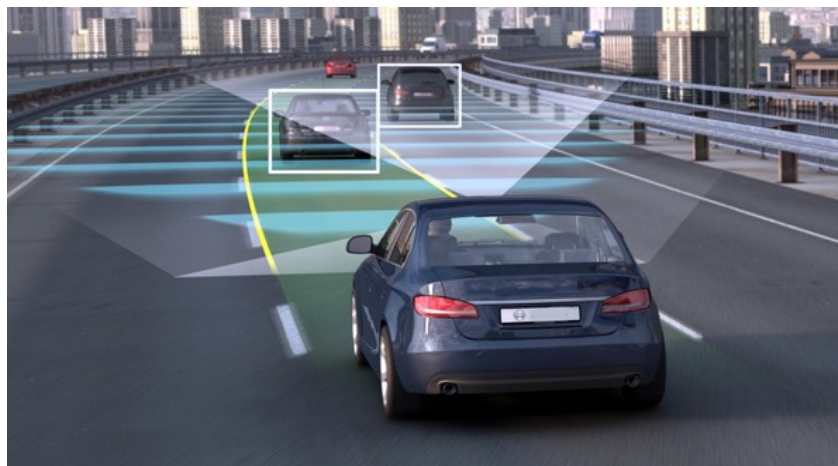
Доверие к системе (AIS Trustworthiness)

- 
- 1. Функциональность.** Возможности системы ИИ при решении конкретной прикладной задачи превышают заданный порог (например, не уступают возможностям квалифицированного человека-оператора)
 - 2. Безопасность.** Функционирование системы ИИ не приводит к появлению неприемлемых угроз для пользователей, третьих лиц, окружающей среды
 - 3. Предсказуемость.** Требования 1 и 2 с заданной вероятностью выполняются для всех возможных условий эксплуатации системы ИИ

Медицинская диагностика



Беспилотный транспорт

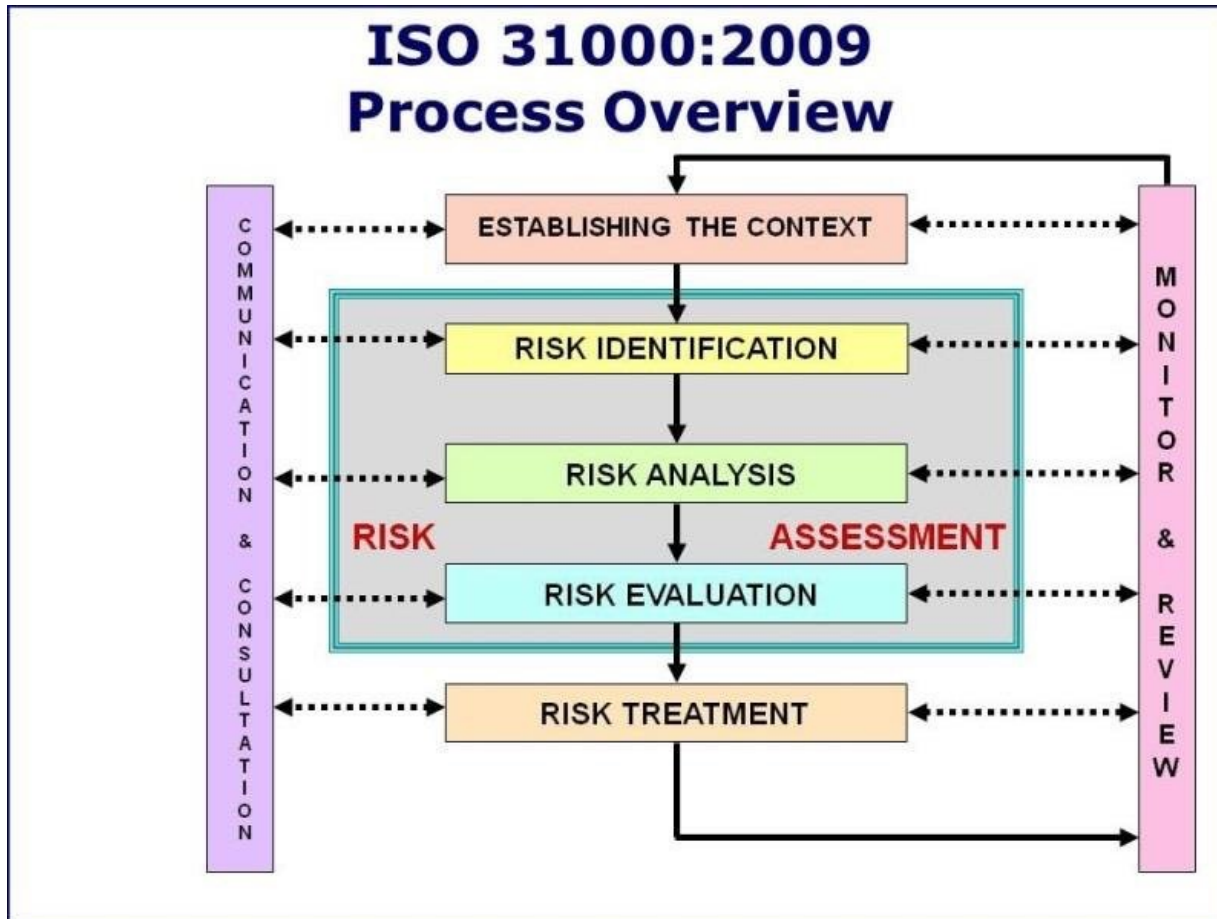


Системы безопасности



Некорректная работа систем ИИ сопряжена с риском для жизни и здоровья людей, значительным экономическим и экологическим ущербом

Обеспечение доверия к системам искусственного интеллекта на основе менеджмента рисков



Примеры рисков систем ИИ

Социальные риски

Экономические риски

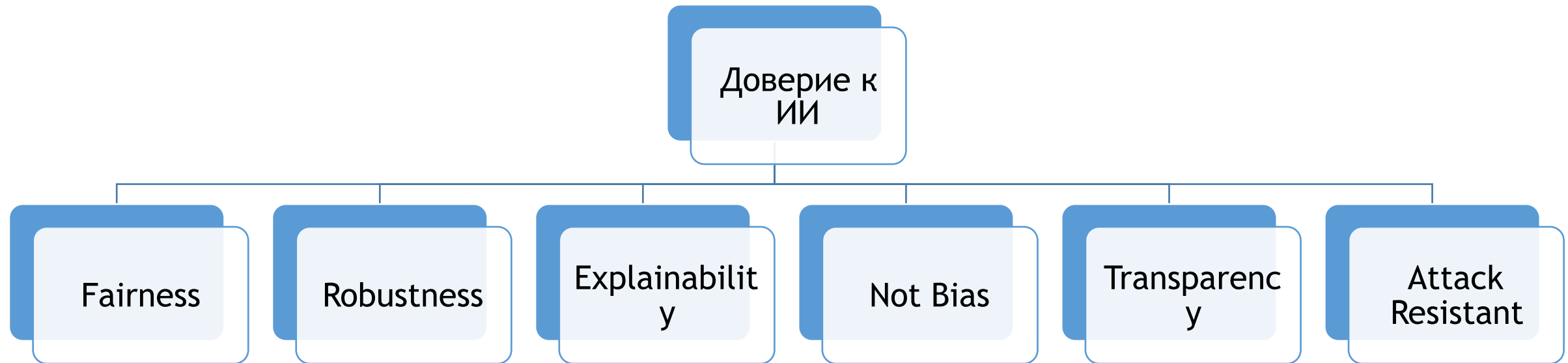
Этические риски

Риски, связанные с эксплуатацией

Риски безопасности

Риски утраты контроля

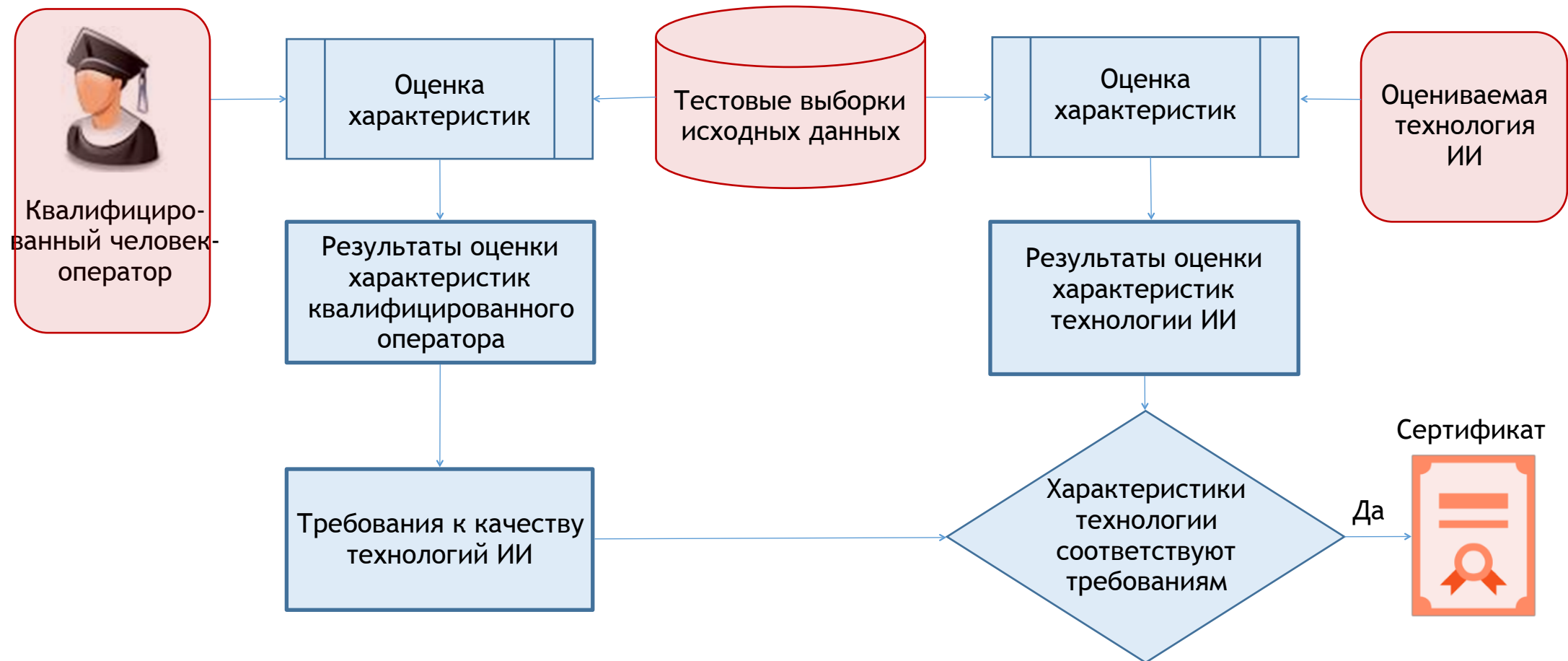
Обеспечение доверия преимущественно к процессу разраб РВК систем ИИ посредством прозрачности, объяснимости, контролируемости, робастности



IBM (2018, август)

Компаниями - разработчиками систем ИИ подготавливается и публикуется «Пользовательская декларация соответствия» (Supplier's Declaration of Conformity, DoC)

Обеспечения доверия к результату работы систем ИИ путем сравнения с функциональными возможностями человека



Основные задачи

Унификация характеристик качества систем ИИ, направленных на решение конкретных прикладных задач обработки данных

Стандартизация подходов к измерению функциональных характеристик прикладных систем ИИ

Основные стандарты

NP TR 24030 Artificial Intelligence (AI) -- Use cases

2020.01

Информационные технологии поддержки жизненного цикла продукции.

Интерактивные электронные технические руководства с применением технологий ИИ и дополненной реальности. Требования к электронной модели изделия [НИИ сертификации]

2020.12

Средства мониторинга поведения и прогнозирования намерений людей.

Термины и определения [Университет ИТМО, СПб политехнический университет]

2021.06

Воздушный транспорт. Аэропорты. Технические средства досмотра.

Методика определения показателей качества распознавания незаконных

вложений по тeneвым рентгеновским изображениям [ГосНИИАС, СПб политехнический университет]

2021.06

Стандарты в области измерения характеристик систем распознавания аэрокосмических изображений [ГосНИИАС, МФТИ]

TBD

Стандарты в области измерения характеристик систем распознавания радарных (тепловизионных) изображений [ИИТ]

Отрасль	Примеры задач
1. Медицина	Диагностика, поддержка принятия медицинских решений Преобразование медицинских документов в унифицированные цифровые форматы
2. Промышленность, строительство	Автоматический контроль ручных операций сборочного производства, строительных операций, выявление отклонений Неразрушающий контроль, контроль скрытых работ Техническое зрение, управление робототехническими комплексами
3. Транспорт	Беспилотное управление автомобильной техникой Выявление нарушений ПДД
4. Право	Анализ нормативных документов на непротиворечивость, полноту и безизбыточность Автоматизированная юридическая поддержка типовых ситуаций
5. Финансовый сектор	Персонализация отношений с клиентами Выявление недобросовестных заёмщиков в кредитных организациях
6. Управление природными ресурсами	Сравнение кадастровой карты со снимками местности, полученными с БПЛА или КА ДЗЗ
7. Безопасность	Идентификация и аутентификация Выявление признаков противоправных действий и намерений
8. Маркетинг, ритейл	Контроль качества и эффективности персонала, специфические задачи безопасности Автоматические call-центры

Человек-оператор в различных прикладных отраслях



1. Интеллектуальные АС (замена человека-оператора)

- Существование квалифицированного оператора с удовлетворительными функциональными возможностями при решении критериальных интеллектуальных задач (без учёта ограничений по времени обработки)
- Обработка данных преимущественно антропоморфной модальности

2. АС с элементами искусственного интеллекта (повышение эффективности обработки информации за счёт применения методов ИИ)

- Синтетические (не соответствующие модальности органов чувств человека) данные
- Большая размерность данных
- Принципиально слабые способности человека при обработке данных

Основные задачи

Стандартизация требований к учебным материалам с целью формирования персональных образовательных траекторий при помощи технологий ИИ

Стандартизация информации об образовательной активности обучающегося (цифровой след) и формирование стандартов по ее использованию, распространению и интерпретации с целью внедрения адаптивности и нелинейности образовательного процесса при помощи ИИ

Стандартизация требований к образовательным платформам и средствам проведения обучения с целью интеграции в них технологий ИИ

Основные стандарты и спецификации

TR 24372 Overview of computational approaches for AI systems

TR 20547-2:2018 Big data reference architecture - Part 2: Use cases and requirements

2019.05

Tin Can (xAPI)

#	Наименование программы	Головные организации	Программно-плановые документы
1	Национальная технологическая инициатива	Отраслевые ассоциации НТИ	ДК по нормативному регулированию «Нейронет», «Аэронет», «Автонет», «Маринет», «Технет»...
		Центры компетенции НТИ по направлению интеллектуальных технологий и больших данных	Программы работ ЦК в МФТИ, Университет ИТМО, МГУ им.М.В.Ломоносова
2	Национальная программа «Цифровая экономика Российской Федерации»	Сколково	План мероприятий федерального проекта «Нормативное регулирование цифровой среды»
		ПАО «Сбербанк»	План мероприятий федерального проекта «Искусственный интеллект»
3	Государственная программа вооружений	Росстандарт	План стандартизации оборонной продукции

Взаимодействие с другими техническими комитетами



#	Наименование ТК	Предмет взаимодействия
1	ТК 022 «Информационные технологии»	Интероперабельность ИС, большие данные, качество ИС
2	ТК 026 «Криптографическая защита информации»	Анализ устойчивости, надёжности функционирования ИС, в том числе, использующих нейросетевые алгоритмы. Применение нейросетевых алгоритмов для криптографической защиты данных
3	ТК 098 «Биометрия и биомониторинг»	Применение методов ИИ для решения задач биометрии и биомониторинга
4	ТК 159 «Блокчейн»	Применение методов ИИ для управления смарт-контрактами
5	ТК 194 «Кибер-физические системы»	Интеллектуальные системы управления киберфизическими системами
6	ТК 355 «Технологии автоматической идентификации и сбора данных»	Интеллектуальные системы идентификации и аутентификации людей
7	ТК 362 «Защита информации»	Деперсонафикация персональных данных, управление согласиями
8	ТК 461 «ИКТ в образовании»	Интеллектуальные образовательные системы
9	ТК 468 «Информатизация	Интеллектуальные системы медицинской диагностики,

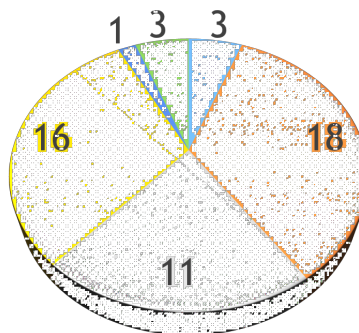
Спасибо за внимание!

Гарбук Сергей Владимирович
garbuk@list.ru



Национальное экспертное сообщество SC 42 “Artificial Intelligence”

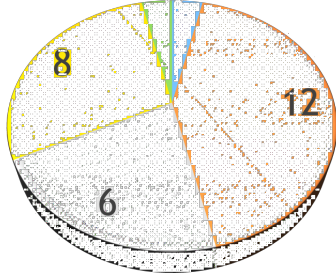
Состав экспертов для участия в проектах подкомитета SC42



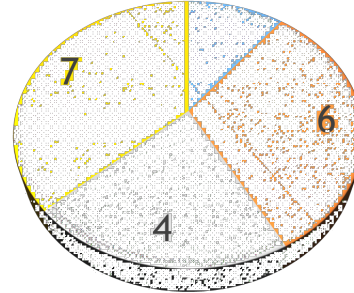
Всего 52 эксперта:

- АНО и общественные организации
- АО и иные коммерческие организации
- ВУЗы
- Научно-исследовательские, аналитические и проектные организации
- ФОИВ
- Независимые эксперты

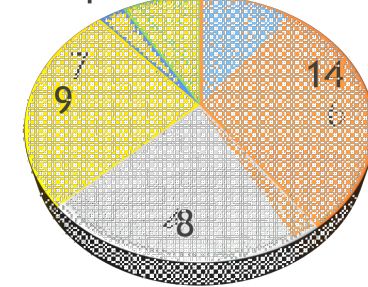
Распределение
участников
WG 1 SC 42



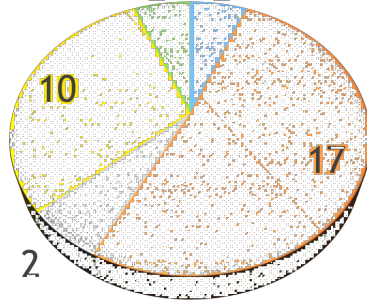
Распределение
участников
WG 2 SC 42



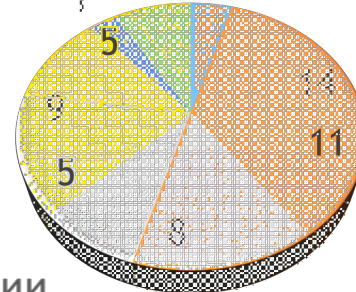
Распределение
участников
WG 3 SC 42



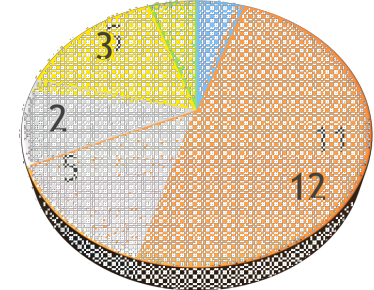
Распределение
участников
WG 4 SC 42



Распределение
участников
WG 5 SC 42



Распределение
участников
JWG 1 SC 42



- АНО и общественные организации
- АО и иные коммерческие организации
- ВУЗы
- Научно-исследовательские, аналитические и проектные организации
- ФОИВ
- Независимые эксперты

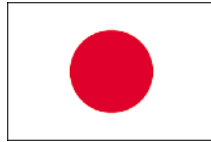


Заседание подкомитета SC 42 “Artificial Intelligence” в России

Заседания ISO/IEC JTC1/SC42



1) 2018.04 - Пекин (КНР)



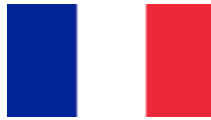
4) 2019.10 - Токио
(Япония)



7) 2021.04 - ???



2) 2018.10 - Саннивэйл
(США)



5) 2020.04 - Париж
(Франция)



3) 2019.04 - Дублин
(Ирландия)



6) 2020.10 -
Монреаль
(Канада)